



Universidad de Valladolid

ESCUELA DE INGENIERÍA INFORMÁTICA (SG)

**Grado en Ingeniería Informática de Servicios y
Aplicaciones**

**BUSINESS INTELLIGENCE:
ANÁLISIS DE UNA CLÍNICA VETERINARIA**

*Alumnos: Miguel Ángel López Garzón
Leticia Aceves Tardón*

Tutor: Juan José Álvarez Sánchez

Leticia Aceves Tardón:

Desde estas líneas me gustaría expresar mi más sincero agradecimiento:

A mi familia que siempre me han apoyado en todo, ayudándome en los peores momentos.

En especial a mi abuelo, Ovidio, que finalmente no ha podido verme titulada y que siempre confió en mí, enseñándome que con esfuerzo, trabajo y constancia todo se consigue, y que en esta vida nadie regala nada.

A mi compañero Miguel Ángel López por soportar mis broncas.

Miguel Ángel López Garzón:

Con mi más sincero agradecimiento:

A mi familia y amigos, por aguantarme en todo momento.

Y, en especial, a mi compañera de proyecto, por su paciencia.

Resumen

Cada día nacen más y más empresas de diferentes sectores con un objetivo común: obtener beneficios. Para conseguir esta finalidad deben competir con otras empresas del sector y otros productos similares al suyo. Los análisis de Business Intelligence, permiten a una empresa o compañía, visualizar de una forma sencilla las ganancias o pérdidas que puede tener en un periodo determinado de tiempo.

Gestionar la información en las empresas es, una herramienta clave para poder sobrevivir en un mercado cambiante. Manejar esta información es fundamental para la toma de decisiones, el crecimiento y la gestión de una empresa. Business Intelligence acerca a la empresa a los sistemas de información que ayudan a la toma de decisiones. Toda empresa, da igual el tamaño, dispone de sistemas de información más o menos sofisticados y que es conveniente analizar y optimizar.

El presente proyecto es un acercamiento a los sistemas analíticos de Business Intelligence. La finalidad del mismo, es acercarnos al funcionamiento de los análisis BI y hacer de su complejidad algo sencillo de comprender.

Abstract

Everyday more and more companies rises from different sectors with a common goal: getting profit. To achieve this purpose they must compete with each other and with other products similar to theirs. Business Intelligence analysis, allow a business or company, to have a simple view of their gains and losses that may have in a certain period of time.

Managing information in companies is the key to survive in a changing marketplace. Operate with this information is critical for decision making, growth and management of a company. Business Intelligence approaches the company to the information systems that helps decision making. Every company, no matter what size, it has information systems which can be more or less sophisticated and it is convenient to analyze and optimize.

This project is an analytical approach to business intelligence systems. The purpose of it is to approach the performance of BI analysis and make something easy to understand its complexity.

Índice de Contenidos

CAPITULO 1: INTRODUCCIÓN.

1.1- Motivación.....	pag 14
1.2- Propósito General.....	pag 15
1.3- Objetivos y Alcance del Sistema.....	pag 15
1.3.1- Alcance.....	pag 15
1.3.2- Objetivos.....	pag 16
1.3.3- Reglas de Negocio.....	pag 17
1.4- Organización del Documento.....	pag 17

CAPITULO 2: ESTADO DEL ARTE

2.1- ¿Qué es Business Intelligence?.....	pag 19
2.2- Ventajas de BI.....	pag 20
2.3- Componentes de Business Intelligence.....	pag 21
2.4-Fases de BI.....	pag 22

CAPITULO 3: PLAN DE PROYECTO Y PRESUPUESTO.

3.1- Descripción de Nuestro Sistema.....	pag 27
3.1.1- Funcionalidad.....	pag 27
3.2- Metodología.....	pag 28
3.3- Herramientas Utilizadas.....	pag 30
3.4- Planificación Temporal.....	pag 34
3.5- Presupuesto.....	pag 36
3.5.1- Presupuesto Hardware y Software.....	pag 36
3.5.2- Presupuesto de Desarrollo.....	pag 37
3.5.3- Presupuesto Total.....	pag 38

CAPITULO 4: ANALISIS DEL PROYECTO

4.1- Características Principales.....	pag 40
4.2- Descripción de Actores del Sistema.....	pag 41
4.3- Requisitos de Información.....	pag 42
4.3.1- Granularidad.....	pag 43
4.3.2- Definición de Dimensiones.....	pag 45
4.3.3- Definición Tablas de Hechos.....	pag 46
4.3.4- Especificación de Dimensiones.....	pag 48
4.3.5- Especificación Tablas de Hechos.....	pag 53
4.4- Requisitos de Usuario.....	pag 54
4.4.1- Casos de Uso.....	pag 54
4.4.2- Diagramas de Casos de Uso.....	pag 55
4.4.3- Especificación de Casos de Uso.....	pag 57
4.5- Requisitos Funcionales.....	pag 65
4.5.1- Especificación de Requisitos Funcionales.....	pag 66
4.6- Requisitos No Funcionales.....	pag 73
4.6.1- Especificación de Requisitos No Funcionales.....	pag 75
4.7-Trazabilidad.....	pag 79

CAPITULO 5: DISEÑO

5.1- Arquitectura.....	pag 82
5.1.1- Arquitectura Lógica.....	pag 83
5.1.1- Arquitectura Física.....	pag 84
5.2- Diagrama de Clases.....	pag 85
5.3- Diagramas de Secuencia.....	pag 86
5.4- Diagramas de Estados.....	pag 90
5.5- Diseño de Interfaz de Usuario.....	pag 93

CAPITULO 6: IMPLEMENTACIÓN

6.1- Creación de la Estructura del Data Warehouse.....	pag 96
6.1.1- DDL: Estructura.....	pag 96
6.1.2- DML: Contenido de DIM_TIEMPO.....	pag 98
6.1.3- Usuarios y Permisos.....	pag 99
6.2-Proceso de Carga de Datos (ETLs).....	pag 101
6.2.1- Aspectos Previos.....	pag 101
6.2.2- Dimensiones.....	pag 108
6.2.3- Tablas de Hechos.....	pag 140
6.3- Proceso de Creación de Informes (Reporting).....	pag 151
6.3.1- Configuración de la Herramienta.....	pag 151
6.3.2- Creación del Proyecto: Microstrategy.....	pag 152
6.3.3- Informes.....	pag 165

CAPITULO 7: PRUEBAS Y VALIDACIONES

7.1- Validaciones a nivel de EtlS.....	pag 172
7.1.1- Dimensiones.....	pag 173
7.1.2- Tablas de Hechos.....	pag 189
7.1.3- Prueba de Relanzamiento.....	pag 196
7.2- Validaciones a nivel de Informes.....	pag 198
7.2.1- Informe de Consultas.....	pag 198
7.2.2- Informe de Variación.....	pag 201
7.2.3- Informe de Ventas.....	pag 205

CAPITULO 8: CONCLUSIONES

8.1- Conclusiones Personales.....	pag 207
-----------------------------------	---------

CAPITULO 9: BIBLIOGRAFIA.....pag 209

ANEXO I: CREACIÓN DEL MODELO OPERACIONAL..... pag 211

ANEXO II: CONTENIDO DEL CD-ROM.....pag 219

ANEXO III: GLOSARIO DE TERMINOS.....pag 220

Índice de Tablas

CAPITULO 3: PLAN DE PROYECTO Y PRESUPUESTO.

Tabla 001. Presupuesto; Hardware.....	pag 36
Tabla 002. Presupuesto; Software.....	pag 37
Tabla 003. Presupuesto; Horas.....	pag 37
Tabla 004. Presupuesto; Coste Desarrollo.....	pag 38
Tabla 005. Presupuesto; Total.....	pag 38

CAPITULO 4: ANALISIS DEL PROYECTO

Tabla 006. Actor-01; Administrador.....	pag 41
Tabla 007. Actor-02; Usuario.....	pag 41
Tabla 008. Dimensiones; DIM_TIEMPO.....	pag 48
Tabla 009. Dimensiones; DIM_EMPLEADO.....	pag 48
Tabla 010. Dimensiones; DIM_CARGO.....	pag 49
Tabla 011. Dimensiones; DIM_MASCOTA.....	pag 49
Tabla 012. Dimensiones; DIM_PROPIETARIO.....	pag 50
Tabla 013. Dimensiones; DIM_TIPOMASCOTA.....	pag 50
Tabla 014. Dimensiones; DIM_CONSULTA.....	pag 51
Tabla 015. Dimensiones; DIM_CLASECONSULTA.....	pag 51
Tabla 016. Dimensiones; DIM_PRODUCTO.....	pag 52
Tabla 017. Dimensiones; DIM_TIPOPUESTO.....	pag 52
Tabla 018. Tabla de Hechos; TH_CONSULTAS.....	pag 53
Tabla 019. Tabla de Hechos; TH_VENTAS.....	pag 53
Tabla 020. Caso de Uso (ETLs) 01; Autenticar Usuario.....	pag 57
Tabla 021. Caso de Uso (ETLs) 02; Consultar Base de Datos.....	pag 58
Tabla 022. Caso de Uso (Informes) 01; Autenticar Usuario.....	pag 59
Tabla 023. Caso de Uso (Informes) 02; Visualizar Informe.....	pag 60
Tabla 024. Caso de Uso (Informes) 03; Filtrar Información.....	pag 61
Tabla 025. Caso de Uso (Informes) 04; Seleccionar Sección del Informe.....	pag 62
Tabla 026. Caso de Uso (Informes) 05; Cerrar Informe.....	pag 64
Tabla 027. Requisito Funcional (ETL) 01; Ajustarse al Modelo.....	pag 66
Tabla 028. Requisito Funcional (ETL) 02; Integridad de los Datos.....	pag 66
Tabla 029. Requisito Funcional (ETL) 03; "No Informados".....	pag 66
Tabla 030. Requisito Funcional (ETL) 04; Manejo de Errores.....	pag 67
Tabla 031. Requisito Funcional (ETL) 05; Dimensión Temporal.....	pag 67
Tabla 032. Requisito Funcional (ETL) 06; Usuarios.....	pag 67
Tabla 033. Requisito Funcional (ETL) 07; Usuario Administrador.....	pag 67
Tabla 034. Requisito Funcional (ETL) 08; Usuario Cliente.....	pag 68
Tabla 035. Requisito Funcional (ETL) 09; Cuantías Numéricas.....	pag 68
Tabla 036. Requisito Funcional (ETL) 10; Formatos.....	pag 68
Tabla 037. Requisito Funcional (Informes) 01; Identificación de Usuario.....	pag 68
Tabla 038. Requisito Funcional (Informes) 02; Autenticación.....	pag 69
Tabla 039. Requisito Funcional (Informes) 03; Validación de Autenticación.....	pag 69
Tabla 040. Requisito Funcional (Informes) 04; Abrir Sesión.....	pag 69
Tabla 041. Requisito Funcional (Informes) 05; Permisos Usuario Cliente.....	pag 69
Tabla 042. Requisito Funcional (Informes) 06; Permisos Usuario Administrador.....	pag 70
Tabla 043. Requisito Funcional (Informes) 07; Menú.....	pag 70
Tabla 044. Requisito Funcional (Informes) 08; Abrir Informe.....	pag 70
Tabla 045. Requisito Funcional (Informes) 09; Acceso a Datos.....	pag 70
Tabla 046. Requisito Funcional (Informes) 10; Datos Adecuados.....	pag 71
Tabla 047. Requisito Funcional (Informes) 11; Filtros.....	pag 71
Tabla 048. Requisito Funcional (Informes) 12; Orden Cronológico.....	pag 71
Tabla 049. Requisito Funcional (Informes) 13; Formatos.....	pag 71
Tabla 050. Requisito Funcional (Informes) 14; Cuantías Numéricas.....	pag 72
Tabla 051. Requisito Funcional (Informes) 15; Información no Encontrada.....	pag 72
Tabla 052. Requisito Funcional (Informes) 16; Ampliar Sección.....	pag 72
Tabla 053. Requisito Funcional (Informes) 17; Cerrar Informe.....	pag 72
Tabla 054. Requisito No Funcional 01; Tiempo de Autenticación.....	pag 75
Tabla 055. Requisito No Funcional 02; Tiempo Abrir Menú.....	pag 75
Tabla 056. Requisito No Funcional 03; Tiempo Abrir Informe Usuario Experimentado.....	pag 75
Tabla 057. Requisito No Funcional 04; Tiempo Abrir Informe Usuario Nuevo.....	pag 75
Tabla 058. Requisito No Funcional 05; Tiempo de Validación.....	pag 76
Tabla 059. Requisito No Funcional 06; Tiempo de Apertura de la Página.....	pag 76
Tabla 060. Requisito No Funcional 07; Tiempo de Búsqueda.....	pag 76
Tabla 061. Requisito No Funcional 08; Tipos de Usuarios.....	pag 76
Tabla 062. Requisito No Funcional 09; Usuario Cliente.....	pag 77
Tabla 063. Requisito No Funcional 10; Usuario Administrador.....	pag 77
Tabla 064. Requisito No Funcional 11; Seguridad Física.....	pag 77
Tabla 065. Requisito No Funcional 12; Backups.....	pag 77
Tabla 066. Requisito No Funcional 13; Restaurar Backup.....	pag 78
Tabla 067. Requisito No Funcional 14; Disponibilidad.....	pag 78

Tabla 068. Requisito No Funcional 15; Mantenimiento.....	pag 78
Tabla 069. Trazabilidad 01; Requisitos Funcionales – Casos de Uso.....	pag 79
Tabla 070. Trazabilidad 02; Requisitos No Funcionales – Casos de Uso.....	pag 79
Tabla 071. Trazabilidad 03; Requisitos Funcionales – Requisitos No Funcionales.....	pag 80
Tabla 072. Trazabilidad 04; Requisitos y Casos de Uso –Objetivos y Características.....	pag 80

CAPITULO 5: DISEÑO

Tabla 073. Interfaz Usuario 01; Principal.....	pag 93
Tabla 074. Interfaz Usuario 02; Menú de Informes.....	pag 94
Tabla 075. Interfaz Usuario 03; Informes.....	pag 94

CAPITULO 6: IMPLEMENTACIÓN

Tabla 076. Usuarios y Permisos.....	pag 99
Tabla 077. Fuentes de Información Dimensiones; General.....	pag 108
Tabla 078. Destinos de Información Dimensiones; General.....	pag 109
Tabla 079. Mapeos de Origen/Destino; General.....	pag 113
Tabla 080. Fuentes de Información; DIM_TIEMPO.....	pag 113
Tabla 081. Destinos de Información; DIM_TIEMPO.....	pag 114
Tabla 082. Destinos de Información; Primary Key; DIM_TIEMPO.....	pag 114
Tabla 083. Mapeos de Origen/Destino; DIM_TIEMPO.....	pag 115
Tabla 084. Fuentes de Información; DIM_PROPIETARIO.....	pag 115
Tabla 085. Destinos de Información; DIM_PROPIETARIO.....	pag 116
Tabla 086. Destinos de Información; Primary Key; DIM_PROPIETARIO.....	pag 116
Tabla 087. Mapeos Origen/Destino; DIM_PROPIETARIO.....	pag 117
Tabla 088. Fuentes de Información; DIM_TIPOMASCOTA.....	pag 118
Tabla 089. Destinos de Información; DIM_TIPOMASCOTA.....	pag 118
Tabla 090. Destinos de Información; Primary Key; DIM_TIPOMASCOTA.....	pag 119
Tabla 091. Mapeos Origen/Destino; DIM_TIPOMASCOTA.....	pag 120
Tabla 092. Fuentes de Información; DIM_MASCOTA.....	pag 121
Tabla 093. Destinos de Información; DIM_MASCOTA.....	pag 122
Tabla 094. Destinos de Información; Primary Key; DIM_MASCOTA.....	pag 122
Tabla 095. Destinos de Información; Foreign Key; DIM_MASCOTA.....	pag 122
Tabla 096. Mapeos Origen/Destino; DIM_MASCOTA.....	pag 124
Tabla 097. Fuentes de Información; DIM_CARGO.....	pag 125
Tabla 098. Destinos de Información; DIM_CARGO.....	pag 125
Tabla 099. Destinos de Información; Primary Key; DIM_CARGO.....	pag 125
Tabla 100. Mapeos Origen/Destino; DIM_CARGO.....	pag 126
Tabla 101. Fuentes de Información; DIM_EMPLEADO.....	pag 127
Tabla 102. Destinos de Información; DIM_EMPLEADO.....	pag 127
Tabla 103. Destinos de Información; Primary Key; DIM_EMPLEADO.....	pag 128
Tabla 104. Destinos de Información; Foreign Key; DIM_EMPLEADO.....	pag 128
Tabla 105. Mapeos Origen/Destino; DIM_EMPLEADO.....	pag 129
Tabla 106. Fuentes de Información; DIM_CLASECONSULTA.....	pag 130
Tabla 107. Destinos de Información; DIM_CLASECONSULTA.....	pag 130
Tabla 108. Destinos de Información; Primary Key; DIM_CLASECONSULTA.....	pag 130
Tabla 109. Mapeos Origen/Destino; DIM_CLASECONSULTA.....	pag 131
Tabla 110. Fuentes de Información; DIM_CONSULTA.....	pag 132
Tabla 111. Destinos de Información; DIM_CONSULTA.....	pag 132
Tabla 112. Destinos de Información; Primary Key; DIM_CONSULTA.....	pag 133
Tabla 113. Destinos de Información; Foreign Key; DIM_CONSULTA.....	pag 133
Tabla 114. Mapeos Origen/Destino; DIM_CONSULTA.....	pag 134
Tabla 115. Fuentes de Información; DIM_TIPO_PRODUCTO.....	pag 135
Tabla 116. Destinos de Información; DIM_TIPO_PRODUCTO.....	pag 135
Tabla 117. Destinos de Información; Primary Key; DIM_TIPO_PRODUCTO.....	pag 135
Tabla 118. Mapeos Origen/Destino; DIM_TIPO_PRODUCTO.....	pag 136
Tabla 119. Fuentes de Información; DIM_PRODUCTO.....	pag 137
Tabla 120. Destinos de Información; DIM_PRODUCTO.....	pag 137
Tabla 121. Destinos de Información; Primary Key; DIM_PRODUCTO.....	pag 138
Tabla 122. Destinos de Información; Foreign Key; DIM_PRODUCTO.....	pag 138
Tabla 123. Mapeos Origen/Destino; DIM_PRODUCTO.....	pag 139
Tabla 124. Fuentes de Información Tablas de Hechos; General.....	pag 140
Tabla 125. Destinos de Información Tablas de Hechos; General.....	pag 140
Tabla 126. Mapeos Origen/Destino; Tablas de Hechos.....	pag 142
Tabla 127. Fuentes de Información; TH_CONSULTAS.....	pag 142
Tabla 128. Destinos de Información; TH_CONSULTAS.....	pag 143
Tabla 129. Destinos de Información; Primary Key; TH_CONSULTAS.....	pag 143
Tabla 130. Destinos de Información; Foreign Key; TH_CONSULTAS.....	pag 143
Tabla 131. Mapeos Origen/Destino; TH_CONSULTAS.....	pag 146
Tabla 132. Fuentes de Información; TH_VENTAS.....	pag 147
Tabla 133. Destinos de Información; TH_VENTAS.....	pag 147
Tabla 134. Destinos de Información; Primary Key; TH_VENTAS.....	pag 148
Tabla 135. Destinos de Información; Foreign Key; TH_VENTAS.....	pag 148
Tabla 136. Mapeos Origen/Destino; TH_VENTAS.....	pag 150
Tabla 137. Glosario.....	pag 220

Índice de Ilustraciones

CAPITULO 2: ESTADO DEL ARTE

Ilustración 001. Diagrama de Componentes de Business Intelligence.....	pag 21
Ilustración 002. Diagrama de Modelo en Estrella.....	pag 23
Ilustración 003. Diagrama de Modelo copo de Nieve.....	pag 24

CAPITULO 3: PLAN DE PROYECTO Y PRESUPUESTO.

Ilustración 004. Diagrama de Metodología Cascada.....	pag 28
Ilustración 005. Diagrama de Gantt.....	pag 34

CAPITULO 4: ANALISIS DEL PROYECTO

Ilustración 006. Diagrama de Árbol de Características.....	pag 40
Ilustración 007. Diagrama de Modelado en Copo de Nieve de Nuestro Sistema.....	pag 42
Ilustración 008. Diagrama de Casos de Uso (ETLs).....	pag 55
Ilustración 009. Diagrama de Casos de Uso (ETLs).....	pag 56
Ilustración 010. Caso de Uso (ETLs) 01; Autenticar Usuario.....	pag 58
Ilustración 011. Caso de Uso (ETLs) 02; Consultar Base de Datos.....	pag 59
Ilustración 012. Caso de Uso (Informes) 01; Autenticar Usuario.....	pag 60
Ilustración 013. Caso de Uso (Informes) 02; Visualizar Informe.....	pag 61
Ilustración 014. Caso de Uso (Informes) 03; Filtrar Información.....	pag 62
Ilustración 015. Caso de Uso (Informes) 04; Seleccionar Sección del Informe.....	pag 63
Ilustración 016. Caso de Uso (Informes) 05; Cerrar Informe.....	pag 64

CAPITULO 5: DISEÑO

Ilustración 017. Arquitectura Lógica.....	pag 83
Ilustración 018. Arquitectura Física.....	pag 84
Ilustración 019. Diagrama Clases.....	pag 85
Ilustración 020. Diagrama Secuencia (ETL) 01; Autenticar Usuario.....	pag 86
Ilustración 021. Diagrama Secuencia (ETL) 02; Consultar Base de Datos.....	pag 87
Ilustración 022. Diagrama Secuencia (Informes) 01; Autenticar Usuario.....	pag 87
Ilustración 023. Diagrama Secuencia (Informes) 02; Visualizar Informe.....	pag 88
Ilustración 024. Diagrama Secuencia (Informes) 03; Filtrar Informe.....	pag 88
Ilustración 025. Diagrama Secuencia (Informes) 04; Seleccionar Sección de Informe.....	pag 89
Ilustración 026. Diagrama Secuencia (Informes) 05; Cerrar Informe.....	pag 89
Ilustración 027. Diagrama Estados (ETL) 01; Autenticar Usuario.....	pag 90
Ilustración 028. Diagrama Estados (ETL) 02; Consultar Base de Datos.....	pag 91
Ilustración 029. Diagrama Estados (Informes) 01; Autenticar Usuario.....	pag 91
Ilustración 030. Diagrama Estados (Informes) 02; Visualizar Informe.....	pag 91
Ilustración 031. Diagrama Estados (Informes) 03; Filtrar Informe.....	pag 92
Ilustración 032. Diagrama Estados (Informes) 04; Seleccionar Sección de Informe.....	pag 92
Ilustración 033. Diagrama Estados (Informes) 05; Cerrar Informe.....	pag 92
Ilustración 034. Interfaz Usuario 01; Página Principal.....	pag 93
Ilustración 035. Interfaz Usuario 02; Menú de Informes.....	pag 93
Ilustración 036. Interfaz Usuario 03; Informe.....	pag 94

CAPITULO 6: IMPLEMENTACIÓN

Ilustración 037. Usuarios y Permisos; Configuración.....	pag 100
Ilustración 038. Nuevo Proyecto; Configuración.....	pag 102
Ilustración 039. ETLs; Configuración Origen de Datos: Administrador de Conexiones.....	pag 103
Ilustración 040. ETLs; Configuración Origen de datos: Columnas.....	pag 104
Ilustración 041. ETLs; Configuración Salida de Errores.....	pag 104
Ilustración 042. ETLs; Configuración Conversiones de Datos.....	pag 104
Ilustración 043. ETLs; Configuración Búsqueda: Conexión.....	pag 105
Ilustración 044. ETLs; Configuración Búsqueda: Columnas.....	pag 105
Ilustración 045. ETLs; Configuración Columna Derivada.....	pag 105
Ilustración 046. ETLs; Configuración Unión.....	pag 106
Ilustración 047. ETLs; Configuración Agregación.....	pag 106
Ilustración 048. ETLs; Configuración Destino: Administrador de Conexión.....	pag 106
Ilustración 049. ETLs; Configuración Destino: Asignaciones.....	pag 107
Ilustración 050. ETLs; Configuración Update.....	pag 107
Ilustración 051. ETLs; Flujo de Carga Completo.....	pag 110
Ilustración 052. ETLs; Configuración Quitar Claves.....	pag 110
Ilustración 053. ETLs; Configuración Quitar Claves; Consulta SQL.....	pag 111
Ilustración 054. ETLs; Configuración de Tareas del Flujo de Carga.....	pag 111
Ilustración 055. ETLs; Configuración Poner Claves.....	pag 111
Ilustración 056. ETLs; Configuración Poner Claves; Consulta SQL.....	pag 112
Ilustración 057. ETLs; Flujo de Control.....	pag 112
Ilustración 058. ETLs; Configuración "No Informado".....	pag 112

Ilustración 059.	Detalles del Proceso de Carga; DIM_PROPIETARIO.....	pag 116
Ilustración 060.	Detalles del Proceso de Carga; DIM_TIPOMASCOTA.....	pag 119
Ilustración 061.	Detalles del Proceso de Carga; DIM_MASCOTA.....	pag 123
Ilustración 062.	Detalles del Proceso de Carga; DIM_CARGO.....	pag 126
Ilustración 063.	Detalles del Proceso de Carga; DIM_EMPLEADO.....	pag 128
Ilustración 064.	Detalles del Proceso de Carga; DIM_CLASECONSULTA.....	pag 131
Ilustración 065.	Detalles del Proceso de Carga; DIM_CONSULTA.....	pag 133
Ilustración 066.	Detalles del Proceso de Carga; DIM_TIPO_PRODUCTO.....	pag 136
Ilustración 067.	Detalles del Proceso de Carga; DIM_PRODUCTO.....	pag 138
Ilustración 068.	Tablas de Hechos; Flujo de Control.....	pag 141
Ilustración 069.	Tablas de Hechos; Borrado de la Tabla de Hechos; General.....	pag 141
Ilustración 070.	Tablas de Hechos; Borrado de la Tabla de Hechos; Asignación de Parámetros.....	pag 141
Ilustración 071.	Tablas de Hechos; Borrado de la Tabla de Hechos; Creación de Variables.....	pag 142
Ilustración 072.	Detalles del Proceso de Carga; TH_CONSULTAS.....	pag 144
Ilustración 073.	Detalles del Proceso de Carga; TH_VENTAS.....	pag 148
Ilustración 074.	Microstrategy; Origen de Datos.....	pag 152
Ilustración 075.	Microstrategy; Nuevo Proyecto.....	pag 153
Ilustración 076.	Microstrategy; Selección de Tablas.....	pag 153
Ilustración 077.	Microstrategy; Tablas.....	pag 154
Ilustración 078.	Microstrategy; Usuarios y Permisos.....	pag 155
Ilustración 079.	Microstrategy; Login.....	pag 156
Ilustración 080.	Microstrategy; Hechos.....	pag 157
Ilustración 081.	Microstrategy; Hechos; Importe Consulta y Tipo Consulta.....	pag 157
Ilustración 082.	Microstrategy; Hechos; Precio Venta y Unidades.....	pag 157
Ilustración 083.	Microstrategy; Atributos.....	pag 158
Ilustración 084.	Microstrategy; Atributos; Configuración.....	pag 158
Ilustración 085.	Microstrategy; Jerarquías.....	pag 160
Ilustración 086.	Microstrategy; Transformaciones.....	pag 161
Ilustración 087.	Microstrategy; Transformaciones; Configuración.....	pag 161
Ilustración 088.	Microstrategy; Filtros; Mesano.....	pag 162
Ilustración 089.	Microstrategy; Filtros; Mesano; Configuración.....	pag 162
Ilustración 090.	Microstrategy; Filtros; Fecha – Recta.....	pag 162
Ilustración 091.	Microstrategy; Filtros; Fecha – Checkbox.....	pag 162
Ilustración 092.	Microstrategy; Filtros; Clase Consulta y Tipo Producto.....	pag 163
Ilustración 093.	Microstrategy; Filtros; Consulta y Producto.....	pag 163
Ilustración 094.	Microstrategy; Indicadores.....	pag 163
Ilustración 095.	Microstrategy; Indicadores; Configuración.....	pag 164
Ilustración 096.	Microstrategy; Informes.....	pag 165
Ilustración 097.	Microstrategy; Informes; Consultas.....	pag 165
Ilustración 098.	Microstrategy; Tablero; Consultas.....	pag 166
Ilustración 099.	Microstrategy; Tablero; Urgencias.....	pag 167
Ilustración 100.	Microstrategy; Tableros; Variación Anual.....	pag 168
Ilustración 101.	Microstrategy; Tablero; Variación Mensual.....	pag 168
Ilustración 102.	Microstrategy; Tablero; Variación Interanual.....	pag 169
Ilustración 103.	Microstrategy; Tablero; Ventas.....	pag 170

CAPITULO 7: PRUEBAS Y VALIDACIONES

Ilustración 104.	Flujo de Control; Flujo de Carga Completo.....	pag 172
Ilustración 105.	Flujo de Control; Inserción de "No Informados".....	pag 172
Ilustración 106.	Flujo de Carga; DIM_TIPOMASCOTA.....	pag 173
Ilustración 107.	Excel; Tipos de Mascotas.....	pag 174
Ilustración 108.	Data Warehouse; DIM_TIPOMASCOTA.....	pag 174
Ilustración 109.	Flujo de Carga; DIM_PROPIETARIO.....	pag 175
Ilustración 110.	Excel; Propietarios.....	pag 175
Ilustración 111.	Data Warehouse; DIM_PROPIETARIO; "No Informado".....	pag 176
Ilustración 112.	Data Warehouse; DIM_PROPIETARIO; Total.....	pag 176
Ilustración 113.	Flujo de Carga; DIM_MASCOTA.....	pag 177
Ilustración 114.	Excel; Mascotas.....	pag 177
Ilustración 115.	Data Warehouse; DIM_MASCOTA; "No Informado".....	pag 178
Ilustración 116.	Data Warehouse; DIM_PROPIETARIO; Total.....	pag 178
Ilustración 117.	Data Warehouse; DIM_MASCOTA; ID_PROPIETARIO.....	pag 178
Ilustración 118.	Data Warehouse; DIM_CLASEMASCOTA; ID_MASCOTA.....	pag 179
Ilustración 119.	Excel; Mascotas; Propietario.....	pag 179
Ilustración 120.	Data Warehouse; DIM_PROPIETARIO; ID_PROPIETARIO.....	pag 179
Ilustración 121.	Flujo de Carga; DIM_CLASECONSULTA.....	pag 179
Ilustración 122.	Excel; Consultas.....	pag 180
Ilustración 123.	Data Warehouse; DIM_CLASECONSULTA.....	pag 181
Ilustración 124.	Flujo de Carga; DIM_CONSULTA.....	pag 181
Ilustración 125.	Excel; Consultas.....	pag 182
Ilustración 126.	Data Warehouse; DIM_CONSULTA; "No Informados".....	pag 182
Ilustración 127.	Data Warehouse; DIM_CONSULTA; Total.....	pag 182
Ilustración 128.	Data Warehouse; DIM_CONSULTA; ID_CLASE.....	pag 182
Ilustración 129.	Excel; Consultas; Código Consulta.....	pag 183
Ilustración 130.	Data Warehouse; DIM_CLASECONSULTA; ID_CLASE.....	pag 183
Ilustración 131.	Flujo de Carga; DIM_CARGO.....	pag 183

Ilustración 132.	Excel; Empleados.....	pag 183
Ilustración 133.	Data Warehouse; DIM_CARGO.....	pag 184
Ilustración 134.	Flujo de Carga; DIM_EMPLEADO.....	pag 184
Ilustración 135.	Excel; Empleados.....	pag 184
Ilustración 136.	Data Warehouse; DIM_EMPLEADO.....	pag 185
Ilustración 137.	Data Warehouse; DIM_CARGO; ID_CARGO.....	pag 185
Ilustración 138.	Flujo de Carga; DIM_TIPO_PRODUCTO.....	pag 185
Ilustración 139.	Excel; Producto.....	pag 186
Ilustración 140.	Data Warehouse; DIM_TIPO_PRODUCTO; "No Informado".....	pag 186
Ilustración 141.	Data Warehouse; DIM_TIPO_PRODUCTO; Total.....	pag 186
Ilustración 142.	Flujo de Carga; DIM_PRODUCTO.....	pag 187
Ilustración 143.	Excel; Producto.....	pag 187
Ilustración 144.	Data Warehouse; DIM_PRODUCTO; "No Informado".....	pag 188
Ilustración 145.	Data Warehouse; DIM_PRODUCTO; Total.....	pag 188
Ilustración 146.	Data Warehouse; DIM_PRODUCTO; Codigo_Producto.....	pag 188
Ilustración 147.	Excel; Producto; Codigo_Producto.....	pag 188
Ilustración 148.	Data Warehouse; DIM_TIPO_PRODUCTO; Cod_TipoProducto.....	pag 188
Ilustración 149.	Flujo de Datos; Tablas de Hechos.....	pag 189
Ilustración 150.	Flujo de Carga; TH_CONSULTAS.....	pag 190
Ilustración 151.	Excel; Diario.....	pag 191
Ilustración 152.	Data Warehouse; TH_CONSULTAS; Total.....	pag 191
Ilustración 153.	Data Warehouse; TH_CONSULTAS; Importe.....	pag 191
Ilustración 154.	Excel; Diario; Importe.....	pag 192
Ilustración 155.	Excel; Diario; Mascota.....	pag 192
Ilustración 156.	Excel; Diario; Empleado.....	pag 192
Ilustración 157.	Excel; Diario; Consulta.....	pag 192
Ilustración 158.	Data Warehouse; DIM_MASCOTA; ID_MASCOTA.....	pag 192
Ilustración 159.	Data Warehouse; DIM_EMPLEADO; ID_EMPLEADO.....	pag 192
Ilustración 160.	Data Warehouse; DIM_CONSULTAS; ID_CONSULTA.....	pag 192
Ilustración 161.	Flujo de Carga; TH_VENTAS.....	pag 193
Ilustración 162.	Excel; Ventas.....	pag 194
Ilustración 163.	Data Warehouse; TH_VENTAS; Total.....	pag 194
Ilustración 164.	Data Warehouse; TH_VENTAS; Precio.....	pag 194
Ilustración 165.	Excel; Ventas; Precio Venta.....	pag 195
Ilustración 166.	Excel; Producto; Cod_producto.....	pag 195
Ilustración 167.	Excel; Diario; Empleado.....	pag 195
Ilustración 168.	Data Warehouse; DIM_PRODUCTO; ID_PRODUCTO.....	pag 195
Ilustración 169.	Data Warehouse; DIM_EMPLEADO; ID_EMPLEADO.....	pag 195
Ilustración 170.	Relanzamiento; Dimensiones.....	pag 196
Ilustración 171.	Relanzamiento; Tablas de Hechos.....	pag 197
Ilustración 172.	Relanzamiento; Tablas de Hechos; Salida de Errores.....	pag 197
Ilustración 173.	Informes; Consultas; Tabla General.....	pag 198
Ilustración 174.	Data Warehouse; TH_CONSULTAS; Validación Tabla General.....	pag 198
Ilustración 175.	Informes; Consultas; Validación Porcentaje Meses.....	pag 199
Ilustración 176.	Data Warehouse; TH_CONSULTAS; Validación Porcentaje Meses.....	pag 199
Ilustración 177.	Informes; Urgencias; Tabla.....	pag 200
Ilustración 178.	Data Warehouse; TH_CONSULTAS; Validación Urgencias.....	pag 200
Ilustración 179.	Informes; Variación Consultas; Tabla Mensual.....	pag 201
Ilustración 180.	Data Warehouse; TH_CONSULTAS; Validación Variación Mensual Consultas.....	pag 201
Ilustración 181.	Informes; Variación Ventas; Tabla Mensual.....	pag 202
Ilustración 182.	Data Warehouse; TH_VENTAS; Validación Variación Mensual Ventas.....	pag 202
Ilustración 183.	Informes; Variación; Anual Consultas.....	pag 203
Ilustración 184.	Informes; Variación; Interanual Consultas.....	pag 203
Ilustración 185.	Data Warehouse; TH_CONSULTAS; Validación Variación Anual e Interanual.....	pag 203
Ilustración 186.	Informes; Variación; Anual Ventas.....	pag 204
Ilustración 187.	Informes; Variación; Interanual Ventas.....	pag 204
Ilustración 188.	Data Warehouse; TH_VENTAS; Validación Variación Anual e Interanual.....	pag 204
Ilustración 189.	Informes; Ventas; Tabla General.....	pag 205
Ilustración 190.	Data Warehouse; TH_VENTAS; Validación Ventas.....	pag 205

ANEXO I: CREACIÓN DEL MODELO OPERACIONAL

Ilustración 191.	Fuente de Información; Mascotas.....	pag 213
Ilustración 192.	Fuente de Información; Consultas.....	pag 214
Ilustración 193.	Fuente de Información; Empleados.....	pag 215
Ilustración 194.	Fuente de Información; Productos.....	pag 215
Ilustración 195.	Fuente de Información; Diario.....	pag 216
Ilustración 196.	Fuente de Información; Aleatorio Fecha.....	pag 216
Ilustración 197.	Fuente de Información; Aleatorio Código Mascotas.....	pag 216
Ilustración 198.	Fuente de Información; Aleatorio Empleado.....	pag 216
Ilustración 199.	Fuente de Información; Aleatorio Código Consulta.....	pag 217
Ilustración 200.	Fuente de Información; Precio.....	pag 217
Ilustración 201.	Fuente de Información; Ventas.....	pag 217
Ilustración 202.	Fuente de Información; Aleatorio Ventas.....	pag 218

CAPITULO I: INTRODUCCIÓN

CAPITULO 1: INTRODUCCIÓN.

Este primer capítulo de la documentación de este proyecto fin de grado, define y expone los motivos por los que se decidió desarrollar este proyecto, con qué propósito se crea y los objetivos que debe cumplir.

En este primer acercamiento abarcaremos, en primer lugar, la motivación, que responde a la pregunta. ¿Por qué este proyecto?. En segundo lugar, el propósito general, donde se hace referencia al qué se va a hacer con este proyecto. Los objetivos y el alcance del sistema, que explicará hasta dónde se puede llegar, lo que podemos acoger en el proyecto. El entorno tecnológico, hace alusión al cómo se va a realizar y porqué se van a utilizar unas herramientas de trabajo y no otras. Y, por último, la organización del resto del documento.

1.1- Motivación

En los últimos años la tecnología ha evolucionado mucho y cada vez tiene más importancia en las empresas. Mediante las tecnologías de Business Intelligence, una compañía, sea grande o pequeña, puede sacar conclusiones sobre pérdidas y/o beneficios a lo largo de un periodo de tiempo sobre sus productos o servicios prestados.

La meta principal de toda organización es maximizar el número de ventas o beneficio sobre el servicio que ofrece. Esta meta se mantiene sin importar la naturaleza de la empresa. Por tanto, el hecho de conocer las pérdidas y ganancias de un producto o servicio, es para las organizaciones, fundamental y necesario: Si estas compañías poseen información sobre qué producto se vende más o cual es el servicio que mejor funciona en un periodo determinado de tiempo, podrán realizar acciones al respecto para garantizar el progreso de beneficios o menor número de pérdidas en su negocio.

En cuanto a la motivación de este proyecto en cuestión, es generada por la compañía Vector ITC group. Dicha organización propuso un curso de Business Intelligence para alumnos de la Universidad de Valladolid que estuvieran en los últimos años del Grado en Ingeniería Informática de Servicios y Aplicaciones o para personas en posesión de un título similar.

Tras realizar unas pruebas de acceso, comenzamos la realización de dicho curso, con una duración estimada de casi 4 meses (440 horas). Entre otras actividades, nos solicitaron la realización de un proyecto fin de beca. Este proyecto fin de beca consistiría en la realización, por parejas, de una simulación de datos y posterior análisis para la generación de informes, es decir, la creación de un proyecto entero de Business Intelligence, con información de una empresa simulada o real.

Una vez finalizado el curso de Business Intelligence, estábamos tan satisfechos con el trabajo realizado y sobre todo con los conocimientos adquiridos en este sector que nos pareció interesante realizar los cambios oportunos para presentar el proyecto realizado para la beca ofrecida por la compañía Vector ITC Group como Trabajo Fin de Grado.

La temática del proyecto era libre, por lo que cada grupo podía elegir la organización sobre la que realizar el proceso de BI. En nuestro caso, optamos por una clínica veterinaria. El motivo de esta elección es porque las clínicas veterinarias realizan dos tipos de actividades diferentes: por un lado la gestión de consultas y por otro la venta de productos para mascotas.

1.2- Propósito General

El propósito del proyecto es el de presentar un sistema de Business Intelligence (BI) con todas sus fases de forma sencilla y comprensible.

El proceso de BI se va a realizar sobre una clínica veterinaria. Realizando las operaciones oportunas sobre la información que ésta organización contiene en su base de datos, obtendremos una serie de informes estadísticos que permitirán a dicha empresa la toma de decisiones sobre sus servicios a clientes.

Los informes finales hacen que el cliente, en nuestro caso la compañía veterinaria, pueda ver de forma centralizada y ordenada los datos que considere oportunos y necesarios. Como toda compañía, la información más relevante son los beneficios obtenidos en un periodo de tiempo, ya sea de forma global o de una determinada consulta o venta de producto. Nos centraremos, por tanto, en dichos beneficios.

El sistema de informes proporcionará unos filtros de búsqueda para que el cliente pueda buscar los productos/ consultas vendidos/ realizadas en un periodo de tiempo determinado.

1.3- Alcance y Objetivos del Sistema

Se hace imprescindible, en todo proyecto, definir de forma clara los objetivos y el alcance del sistema para tener una visión conjunta de hasta dónde se quiere llegar con el proyecto y la envergadura del mismo.

1.3.1- Alcance

En el caso de los proyectos de Business Intelligence, el alcance viene determinado por los modelos de negocio de cada empresa y por los datos que ésta posea. También deberemos definir las funcionalidades u objetivos que incorporará el sistema.

Los factores críticos de éxito para definir el alcance son: Definir correctamente los requerimientos e identificar qué está dentro y qué está fuera del proyecto. Estos dos componentes son fundamentales para poder estimar correctamente los plazos y los recursos que se necesitan. En caso de que se produzcan cambios de requerimientos que afecten al alcance deberemos gestionar los cambios, lo que significa: Identificarlos, analizarlos, valorarlos, tomar una decisión y comunicarla.

En el alcance del proyecto es dependiente de la información proporcionada por el cliente y de las necesidades de éste. Esto es; el cliente proporciona los datos desde donde se extraen los informes, y por tanto, estos informes, son el reflejo de esta información. Por otro lado, el cliente tiene unas determinadas necesidades de información que, en el momento en que nos proporciona sus datos, nos hace saber. En un momento dado, estas necesidades pueden crecer a lo largo del proyecto, nosotros debemos tener en cuenta todos estos requisitos y la necesidad de futuros requerimientos sobre los datos proporcionados.

En conclusión, podemos decir que el alcance se limita en cada caso, a los datos proporcionados por el cliente y, a la necesidad de información del mismo.

1.3.2- Objetivos

Diferenciaremos, en este punto, dos tipos de objetivos: los objetivos del proyecto en sí y los objetivos requeridos por el cliente que solicita el sistema de BI.

Los proyectos de BI responden siempre a un objetivo básico: dotar a la empresa u organización de los medios necesarios para que pueda tomar las decisiones estratégicas y operacionales más adecuadas tomando como base el análisis de datos. La principal pretensión de todo proyecto de BI es que cualquier persona con responsabilidad o poder de decisión en una organización disponga de información adecuada, precisa, relevante y, en definitiva, convertida en un instrumento válido y útil como base, refuerzo o argumento en la toma de decisiones importantes para la organización.

Por otro lado, y en este proyecto en particular, hay que tener en cuenta los objetivos que la compañía contratante nos demanda. Estos son:

- Informe con ganancias de consultas realizadas a mascotas:
 - o Deberá contener filtros de tiempo, para poder visualizar las ganancias obtenidas en un periodo de tiempo determinado, este filtro se hará mensual y anual.
 - o Deberá contener filtros de tipo de consultas y consultas, para visualizar las consultas que se requieran.
 - o Contendrá gráficas explicativas para que los datos sean más comprensibles.
 - o Se deberán presentar variaciones anuales, mensuales e interanuales.

- Informe con ganancias de ventas realizadas a mascotas:
 - o Como en el caso anterior, deberá contener filtros temporales, para que se puedan seleccionar los meses y años.
 - o Deberá contener filtros sobre el tipo de producto y producto, para realizar comparativas.
 - o Contendrá gráficos lo suficientemente claros.

Una vez tenemos claro el objetivo de todo proyecto BI y los requerimientos, en términos generales, del cliente, podemos asumir algún otro objetivo:

- El usuario o cliente podrá acceder a los informes desde cualquier navegador.
- El usuario entrará con una contraseña específica a los informes.

Por tanto, los objetivos los podemos resumir como:

- ✓ **OBJ-01:** el proyecto debe realizar informes lo más precisos posibles para que la compañía contratante pueda tomar decisiones estratégicas.
- ✓ **OBJ-02:** los informes creados deberán ser lo suficientemente sencillos y entendibles.
- ✓ **OBJ-03:** los informes finales deberán contener filtros tanto temporales como otros que se crean convenientes.
- ✓ **OBJ-04:** los informes contendrán gráficos, que cambiarán con los filtros, para facilitar la comprensión de los datos.
- ✓ **OBJ-05:** los informes deberán ser visibles desde cualquier navegador.
- ✓ **OBJ-06:** los informes sólo serán visibles a usuarios autorizados mediante usuario y contraseña.

1.3.3- Reglas de Negocio

Una vez establecidos el alcance y los objetivos a conseguir, se procede a describir unas reglas básicas a tener en cuenta que definan su ámbito de aplicación.

- **RN-01:** La empresa cliente deberá proporcionar los datos suficientes para realizar el proceso de BI.
- **RN-02:** Los informes finales serán lo suficientemente claros, concisos y entendibles.
- **RN-03:** Todos los informes constarán de gráficos para facilitar la lectura de la información y la comparativa de datos.
- **RN-04:** El tiempo de carga de la información de los informes no debe superar los 3 segundos.

1.4- Organización del Documento.

La documentación ha sido estructurada de la siguiente forma:

- **Capítulo 1:** En el primer capítulo se ha realizado una presentación, definiendo los conceptos que se creen más importantes para comprender qué se va a hacer y porqué, así como el origen del proyecto. De este modo queda especificada claramente la idea del proyecto con sus objetivos establecidos, los cuales se irán desarrollando a lo largo de esta documentación.
- **Capítulo 2:** Seguidamente, en el segundo capítulo designado “estado del arte”, se especifica en profundidad lo que es un sistema de BI con sus beneficios y las herramientas que normalmente se utilizan.
- **Capítulo 3:** En tercer lugar, se realiza toda la planificación del proyecto, comenzando por los requisitos principales, seguido por estimaciones de tiempo, recursos, herramientas, costes, etc.
- **Capítulo 4:** A continuación, en el capítulo de análisis se realiza el estudio más exhaustivo de los requisitos y usuarios del sistema, con sus correspondientes tablas explicativas y todos los diagramas de planificación del proyecto (módulos, casos de uso, secuencia,...).
- **Capítulo 5:** En quinto lugar, en el capítulo dedicado al diseño, se realizara como su título indica, el diseño completo de la aplicación, desde su arquitectura, sus diagramas finales y el diseño de datos e interfaces.
- **Capítulo 6:** En sexto lugar, realizaremos el capítulo de pruebas o validaciones, en el cual se realizan las validaciones pertinentes para verificar que todo el proceso ha sido correcto y los datos no sean equivocados.
- **Capítulo 7:** Por último, en el octavo capítulo, las conclusiones personales acerca del desarrollo del proyecto.
- **Capítulo 8:** Finalizando la documentación, se enumerará la bibliografía utilizada para el desarrollo general tanto de la aplicación como de la documentación.
- **Anexos:** En estos puntos finales, desarrollamos la creación del operacional donde se detallará cómo se ha creado la fuente de información para nuestro proyecto, contenido del CD Rom adjunto y un glosario de términos.

CAPITULO II : ESTADO DEL ARTE

CAPITULO 2: ESTADO DEL ARTE

En este capítulo se puntualizará lo que conocemos como Business Intelligence y porqué se ha vuelto tan importante en el mundo empresarial de hoy en día.

Una vez que tenemos claro el motivo del actual proyecto, sus objetivos y alcance del mismo, vamos a definir de un modo global lo que es BI, sus fases y las ventajas que ofrece a las compañías que en un momento dado, contratan estos servicios de BI.

2.1- ¿Qué es *Business Intelligence*?

Actualmente, en el día a día de las operaciones de cualquier organización o empresa, generan datos como producto de todas las transacciones que se llevan a cabo. Típicamente, ésta información se almacena a través de sistemas transaccionales en bases de datos relacionales, o bien, hojas de cálculo tipo Excel, que cumplen la función de almacenaje y administración de éstos datos. Pero la idea es que estos dejen de ser simples datos para convertirse en información que enriquezca las decisiones de los ejecutivos.

Lo que se busca es encontrar información que no solamente conteste preguntas de lo que paso o lo que está sucediendo en el negocio sino que también a través de esta información se construyan modelos en los que se ayude a predecir futuros eventos. Esto es así puesto que las compañías siempre han querido explotar y maximizar el valor de su información con la idea principal de tener mayor ventaja sobre sus competidores.

Por otra parte a fin de incrementar ganancias, las compañías deben enfocarse en retener sus clientes así como buscar la manera de adquirir nuevos clientes. La adquisición de nuevos clientes siempre es un factor muy atractivo para ellas, aunque cabe mencionar que algo muy importante y que no se debe perder de vista es entender la rentabilidad de un cliente ya existente.

Un sistema de BI ofrece la posibilidad a la empresa de sacar ventaja a sus competidores, así como afianzar clientes y conseguir clientes nuevos.

Business Intelligence (Inteligencia de Negocios - Inteligencia Empresarial – BI), es un concepto que integra el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales, en información estructurada para su explotación directa mediante reporting (Informes) o para su análisis y conversión en conocimiento que de soporte a la toma de decisiones sobre el negocio.

El término BI también se refiere al uso de la tecnología para recolectar y usar efectivamente la información, a fin de mejorar la operación del negocio. Un sistema ideal de BI ofrece a los empleados, socios y altos ejecutivos de una compañía, acceso a la información clave que necesitan para realizar sus tareas del día con día, y principalmente para poder tomar decisiones basadas en datos correctos.

Podemos definir BI, por tanto, como el uso de los datos que una entidad tiene almacenados en sus sistemas, transfórmalos y adaptarlos para presentárselos de forma fácilmente interpretable de tal forma que se facilite la toma de decisiones.

2.2- Ventajas de BI

Uno de los objetivos básicos de los sistemas de información es que nos ayuden a la toma de decisiones. Cuando un responsable tiene que tomar una decisión pide o busca información, que le servirá para reducir la incertidumbre. Sin embargo, aunque todos la utilicen, no todos los responsables recogen la misma información: depende de muchos factores, como pueden ser su experiencia, formación, disponibilidad, etc. Del mismo modo, los responsables pueden necesitar recoger más o menos información dependiendo que su mayor o menor aversión al riesgo.

A partir de los datos que nos proporciona el sistema de Business Intelligence podemos descubrir conocimiento.

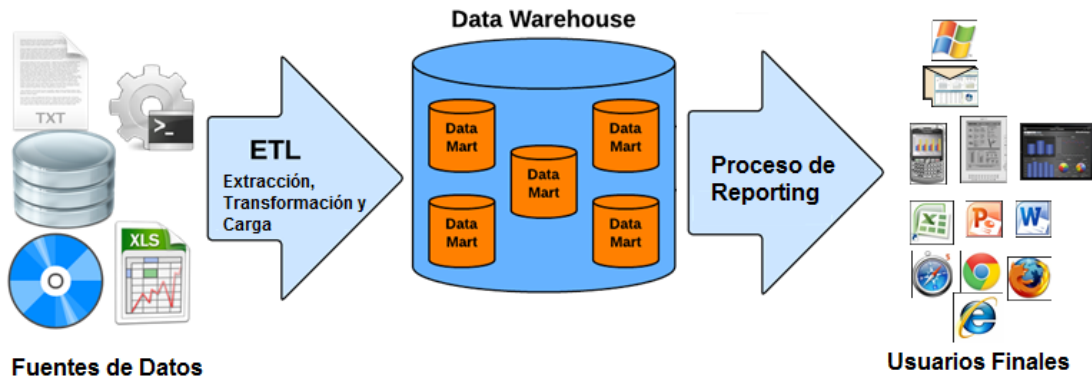
Business Intelligence nos servirá como ayuda para la toma de decisiones y, posteriormente, para descubrir cosas que hasta ahora desconocíamos. Los beneficios que se pueden obtener a través del uso de BI pueden ser de distintos tipos:

- **Beneficios tangibles:** reducción de costes, generación de ingresos, reducción de tiempos para las distintas actividades del negocio. En cuanto a generación de ingresos, que es lo que más interesa a las empresas, se tiene:
 - Mejorar la adquisición de clientes.
 - Reducir la tasa de abandono de clientes, incrementar su fidelidad, teniendo en cuenta cuál es su valor.
 - Incrementar los ingresos por crecimiento de las ventas.
 - Aumentar los resultados, consiguiendo que nuestros clientes actuales compren más productos o servicios.
 - Evitar las pérdidas producidas por las ventas de nuestros competidores.
 - Aumentar la rentabilidad por el acceso a información detallada de productos, clientes, etc.
- **Beneficios intangibles:** el hecho de que tengamos disponible la información para la toma de decisiones hará que más usuarios utilicen dicha información para tomar decisiones y mejorar la posición competitiva.
 - Optimizar la atención a los clientes.
 - Aumentar la satisfacción de los clientes.
 - Mejorar el acceso a los datos a través de consultas, análisis informes.
 - Información más actualizada.
 - Dotar a la información de mayor precisión.
- **Beneficios estratégicos:** Todos aquellos que nos facilitan la formulación de la estrategia, es decir, a qué clientes, mercados o con qué productos dirigirnos.
 - Mayor habilidad para analizar estrategias de precios.
 - Identificar y nutrir a aquellos clientes con mayor potencial.
 - Mejorar la toma de decisiones, realizándola de forma más rápida, informada y basada en hechos.
 - Mayor visibilidad de la gestión.

2.3- Componentes de Business Intelligence

En el siguiente gráfico vemos los distintos componentes de Business Intelligence que vamos a desarrollar a continuación:

Ilustración 001. Diagrama de Componentes de Business Intelligence



Los componentes son, por tanto:

- ✓ Fuentes de información, de las cuales partiremos para alimentar de información el Data Warehouse.
- ✓ Proceso ETL de extracción, transformación y carga de los datos en el Data Warehouse. Antes de almacenar los datos en un Data Warehouse, éstos deben ser transformados, limpiados, filtrados y redefinidos. Normalmente, la información que tenemos en los sistemas transaccionales no está preparada para la toma de decisiones.
- ✓ El propio Data Warehouse o almacén de datos, con el Metadata o Diccionario de datos. Se busca almacenar los datos de una forma que maximice su flexibilidad, facilidad de acceso y administración.
- ✓ Proceso de Reporting, proceso mediante el cual, la información contenida en el Data Warehouse se lleva a informes para que el usuario final pueda ver la información que necesite. La generación de informes requiere partir de la reflexión (operativa o estratégica) para saber qué información se necesita, quien la necesita y con qué periodicidad.
- ✓ Las herramientas de visualización, que nos permitirán el análisis y la navegación a través de los mismos.

Las principales herramientas de Business Intelligence son:

- ✓ **Generadoras de ETLs:** Utilizadas por desarrolladores profesionales para crear los procesos ETLs para pasar de la fuente de información a una información estructurada y cargada en un Data Warehouse.
- ✓ **Generadores de informes:** Utilizadas por desarrolladores profesionales para crear informes estándar para grupos, departamentos o la organización.
- ✓ **Herramientas de usuario final de consultas e informes:** Empleadas por usuarios finales para crear informes para ellos mismos o para otros, no requieren programación.

2.4-Fases de BI

Fase 1 – Problemática:

El proceso de BI empieza con los beneficiarios (Ejecutivos, Directivos, Líderes de Negocio etc.), y aquí se generan las preguntas que les va ayudar a ellos a alcanzar sus objetivos.

Esta fase inicial conlleva el hecho de que una determinada organización necesita dar respuesta a diferentes tipos de preguntas para poder sacar más rendimiento a sus productos o servicios. Estas preguntas suelen ser ¿Qué producto/servicio se vende más? ¿En qué época del año? ¿En qué año se ha sacado más beneficio o pérdida?....

Fase 2 – Recolección de Información:

Hay diversas fuentes de información dentro de una compañía. Cada una de estas fuentes crea, procesa y almacena diferente tipo de información. Este es un proceso continuo, y es importante entender que los datos de esas fuentes son simplemente información y no Inteligencia. Los datos en crudo frecuentemente son incompletos y confusos. La información se convierte en Inteligente mediante el procesado y el análisis.

El proceso de recolección de información es cuando las diferentes fuentes son analizadas para determinar los datos necesarios para encontrar las respuestas a las preguntas.

Fase 3 – Modelado:

Las dos primeras fases descritas anteriormente son propias de la empresa cliente que solicita el proceso de BI.

Una vez que tengamos la información necesaria para las preguntas que la compañía se plantea, se hace imprescindible estudiar los datos que tenemos y realizar el proceso de modelado.

El proceso de modelado consiste en recoger la información que facilita la empresa, examinarla y crear una estructura utilizable para el análisis. Dicha estructura es muy similar para todos los proyectos de BI.

Existen dos tipos de modelos básicos:

- **Modelo Estrella:** es el más sencillo en estructura. Modelo de datos que tiene una tabla de hechos (o *tabla fact*) central, que contiene los datos para el análisis, rodeada de las tablas de dimensiones incluida la dimensión de “Tiempo”.

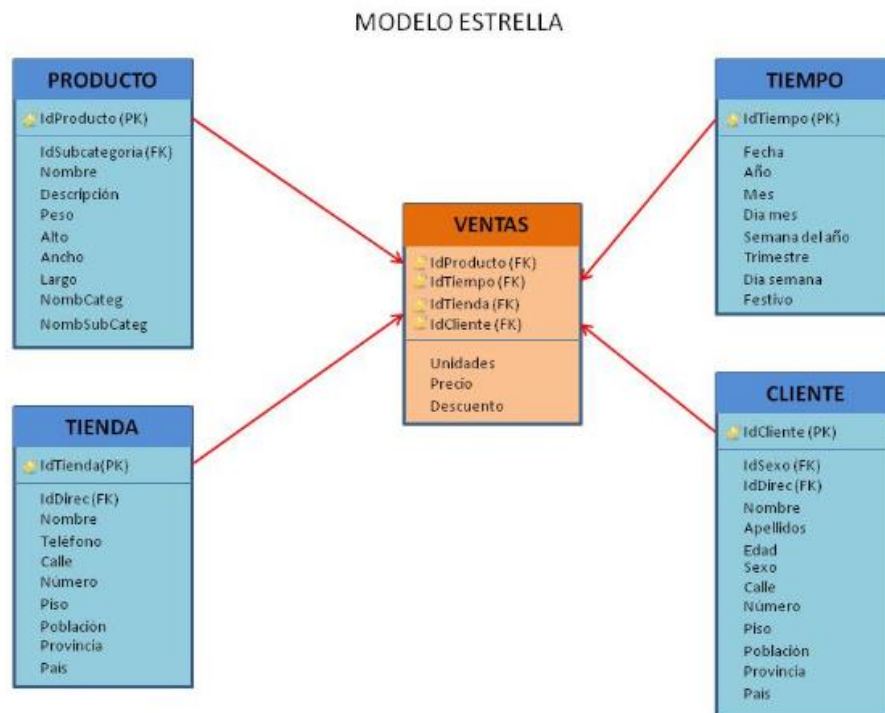
Lo característico de la arquitectura de estrella es que sólo existe una tabla de dimensiones para cada dimensión. Esto quiere decir que la única tabla que tiene relación con otra es la de hechos, lo que significa que toda la información relacionada con una dimensión debe estar en una sola tabla. Este aspecto, de tabla de hechos (o central) más grande rodeada de radios o tablas más pequeñas es lo que asemeja a una estrella, dándole nombre a este tipo de construcciones.

Las tablas de dimensiones tendrán siempre una clave primaria simple, mientras que en la tabla de hechos, la clave principal estará compuesta por las claves principales de las tablas dimensionales.

Los hechos contienen datos medibles, cuantitativos, relacionados a la transacción del negocio, y las dimensiones son atributos que describen los datos indicados en los hechos (una especie de meta-datos, o sea datos que describen otros datos)

Por tanto, en el esquema en estrella la tabla de hechos es la única tabla del esquema que tiene múltiples “joins” o uniones que la conectan con otras tablas (*foreign keys* hacia otras tablas). El resto de tablas del esquema (tablas de dimensión) únicamente hacen “join” con esta tabla de hechos. Las tablas de dimensión se encuentran además totalmente desnormalizadas, es decir, toda la información referente a una dimensión se almacena en la misma tabla.

Ilustración 002. Diagrama de Modelo en Estrella



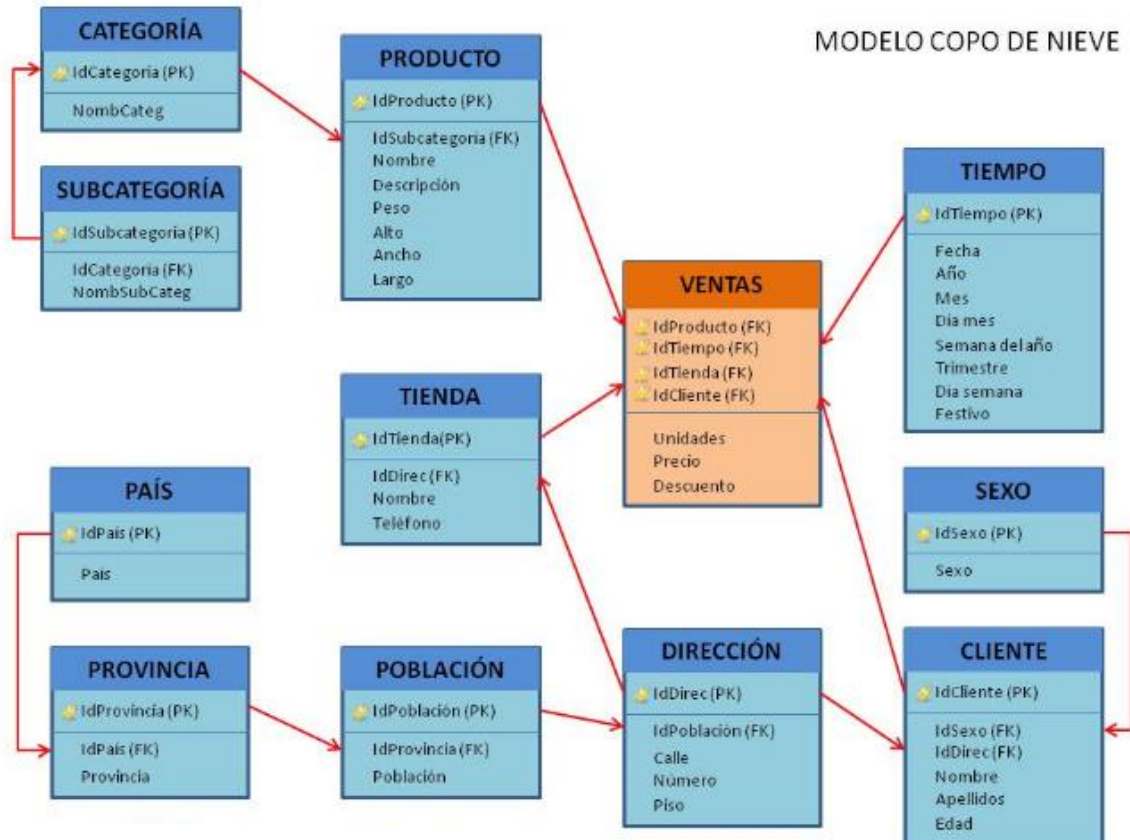
- **Modelo Copo de Nieve:** es un esquema de representación derivado del esquema en estrella. En este modelo la tabla de hechos deja de ser la única relacionada con otras tablas ya que existen otras tablas que se relacionan con las dimensiones y que no tienen relación directa con la tabla de hechos. Por esta razón aparecen nuevas “joins” o agrupaciones, gracias a que las dimensiones de análisis se representan ahora en tablas de dimensión normalizadas.

En la estructura dimensional normalizada, la tabla que representa el nivel base de la dimensión es la que se relaciona directamente con la tabla de hechos. La diferencia entre ambos esquemas (estrella y copo de nieve) reside entonces en la estructura de las tablas de dimensión. Para conseguir un esquema en copo de nieve se ha de tomar un esquema en estrella y conservar la tabla de hechos, centrándose únicamente en el modelado de las tablas de dimensión, que si bien en el esquema en estrella se encontraban totalmente desnormalizadas, ahora se dividen en subtablas tras un proceso de normalización.

Es posible distinguir dos tipos de esquemas en copo de nieve, un copo de nieve completo (en el que todas las tablas de dimensión en el esquema en estrella aparecen ahora normalizadas) o un copo de nieve parcial (sólo se lleva a cabo la normalización de algunas de ellas).

El modelo fue concebido para facilitar el mantenimiento de las dimensiones, sin embargo esto hace que se vinculen más tablas, haciendo la extracción de datos más difícil así como vuelve compleja la tarea de mantener el modelo.

Ilustración 003. Diagrama de Modelo copo de Nieve



Pros y contras de los dos esquemas:

- Modelo en estrella:
 - Este esquema es simple y rápido para ser usado en análisis multidimensional. Permite acceder tanto a datos agregados como de detalle.
 - El diseño de esquemas en estrella permite implementar la funcionalidad de una base de datos multidimensional utilizando una clásica base de datos relacional.
 - Las consultas a la base de datos no son complicadas, ya que las condiciones y los “joins” sólo involucran a la tabla de hechos y a las de dimensiones.
 - Son más simples de manejar que los modelos copo de nieve.
 - Es la opción con mejor rendimiento y velocidad pues permite indexar las dimensiones de forma individualizada son que repercute en el rendimiento de la base de datos en su conjunto

- Modelo copo de nieve:
 - El único argumento a favor de los esquemas en copo de nieve es que al estar normalizadas las tablas de dimensiones, se evita la redundancia de datos y con ello se ahorra espacio.
 - Se puede usar un esquema de copo de nieve en un Data Warehouse, aunque estos sean grandes y complejos, pero nunca en sistemas donde el tiempo de respuesta sea un factor crítico para los usuarios.

Fase 4 – **Procesamientos de Datos:**

Una vez que tengamos el modelado claro, formatearemos los datos en crudo hasta tener el modelo elegido. Esto puede ser posible, creando una nueva base de datos, agregar datos a bases de datos existente o consolidando información.

Esta fase es vista como Extracción, Transformación y Carga (proceso ETL) que ocurre en los ambientes de BI. Es necesario prestar especial cuidado a los datos que se manejan, puesto que, los datos entrantes en un proceso ETL y los salientes tienen que ser iguales, es decir, la consistencia e integridad de los datos debe respetarse en todo momento para que la información obtenida en las siguientes fases del proyecto sea correcta.

Fase 5- **Análisis y Producción:**

El grupo de análisis de negocios utiliza herramientas y técnicas para ordenar sobre los datos y crear “inteligencia”. Esta fase es también denominada de “Reporting”, el proceso de reporting consiste en recoger la información ya transformada (mediante ETLs) y realizar una serie de informes con estos datos. Los informes se deben ajustar a los requerimientos de la empresa cliente.

En esta fase se hace imprescindible una gran comunicación con la empresa cliente para establecer las vistas del informe y para validar que se van cumpliendo los objetivos que se persiguen con el proyecto.

Por esta comunicación fluida con el cliente, tal vez en esta fase, sean generados requerimientos adicionales, pues los analistas pueden que encuentren nuevas preguntas que necesiten ser contestadas. Estos nuevos requerimientos pueden que inviten a realizar algún cambio en el modelado y en las fases posteriores, lo que implica una modificación en la estimación temporal del proyecto, alargándola notablemente.

Fase 6- **Difusión:**

En esta fase de difusión, se entrega el producto final, es decir, los informes, a los diversos clientes que lo requieren. Esto básicamente implica el uso de herramientas BI para la publicación de dichos informes o la posibilidad de tener herramientas de fácil uso para que los mismos usuarios tengan la capacidad de revisar los datos de manera rápida y sencilla.

CAPITULO III : PLAN DE PROYECTO Y PRESUPUESTO

CAPITULO 3: PLAN DE PROYECTO Y PRESUPUESTO.

Una vez tenemos claro qué es lo que se quiere hacer en el presente proyecto y lo que es Business Intelligence, en este tercer capítulo, vamos a organizar temporalmente las fases de que consta el proyecto, la metodología que se va a utilizar y el presupuesto establecido.

La finalidad de un plan de proyecto es que un gestor de proyectos, o persona encargada, pueda usar para acompañar y supervisar el progreso del equipo de trabajo.

3.1- Descripción de Nuestro Sistema

Ahora que ya tenemos claro lo que es un sistema BI, sus características, sus ventajas y sus fases, vamos a describir el proyecto que se quiere realizar y cómo se va a plantear y a desarrollar.

Como ya se ha explicado anteriormente, hemos elegido una empresa ficticia, una clínica veterinaria denominada Vicherinario, como objeto de nuestro proyecto. Al ser una empresa de invención propia, los datos o información de esta compañía serán creados por nosotros con información extraída de internet principalmente (se adjuntará un documento con la explicación de cómo se ha creado la información de esta empresa).

En su caso, tenemos que Vicherinario es una empresa pequeña y demanda poder explotar sus datos para poder obtener información notable para la compañía.

Los orígenes de datos disponibles en la clínica están recogidos en un documento tipo hoja de cálculo. Para la realización del proceso BI se capturan los datos desde el documento, se analizan, se tratan y finalmente se insertan en una base de datos. Esta base de datos, o Data Warehouse como se ha definido anteriormente, servirá a posteriori, para visualizar los datos en informes.

Los orígenes de datos son denominados en BI como “Operacional”. Dependiendo de la empresa, estos datos del operacional son documentos o bases de datos. Cualquiera de las dos formas en que se nos presenten estos orígenes de datos son óptimas. En nuestro proyecto, al ser una clínica pequeña, se ha optado por un documento tipo hoja de cálculo.

3.1.1- Funcionalidad

En el proyecto que nos acontece, se ha creado una solución tecnológica en la cual, se genera una plataforma de reporting BI lo suficientemente eficiente para dar cobertura a las necesidades de información demandadas por la empresa ficticia Vicherinario dentro del contexto de Data Warehouse.

Por tanto, la funcionalidad que se persigue con el diseño de este sistema es que la compañía pueda tomar decisiones estratégicas de manera ágil permitiéndole así, tener una ventaja competitiva frente al resto de compañías de su mismo sector y mejorando, en su caso, el trato con el cliente. Del mismo modo, como se trata de una empresa joven, se hace necesario conocer si finalmente obtiene algún tipo de beneficio.

3.2- Metodología.

Todo proceso de creación de Software tiene un método o una metodología ligada a él. Un modelo de proceso, es una plantilla, patrón o marco que define el proceso a través del cual se crea software. Un modelo de proceso es, por tanto, una estructura aplicada al desarrollo de un producto de software. Hay varios modelos a seguir para el establecimiento de un proceso para el desarrollo de software, cada uno de los cuales describe un enfoque diferente para diferentes actividades que tienen lugar durante el proceso.

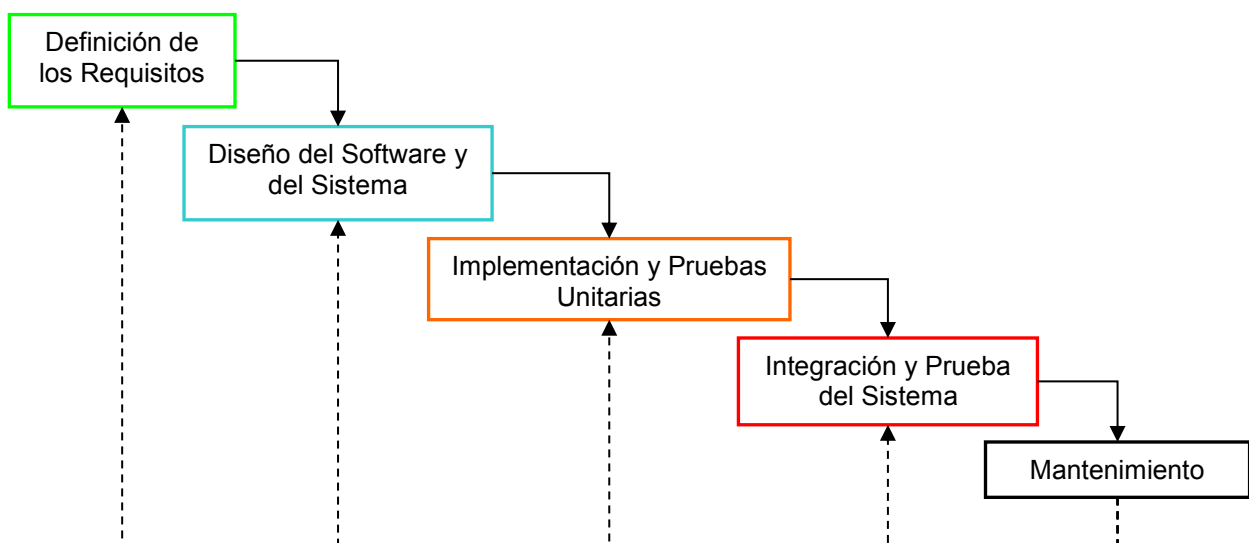
En el caso del proyecto que nos acontece, el método o metodología utilizado para la creación de software, es claramente el método denominado en cascada o lineal.

El modelo en cascada es el más antiguo de todos los modelos de Ingeniería de Software. Presenta una estructura secuencial, de ahí su nombre, formada por cinco fases o etapas diferenciadas:

El modelo en cascada consta de las siguientes fases:

- **Definición de los requisitos:** Los servicios, restricciones y objetivos son establecidos con los usuarios del sistema. Se busca hacer esta definición en detalle.
- **Diseño de software:** Se particiona el sistema en sistemas de software o hardware. Se establece la arquitectura total del sistema. Se identifican y describen las abstracciones y relaciones de los componentes del sistema.
- **Implementación y pruebas unitarias:** Construcción de los módulos y unidades de software. Se realizan pruebas de cada unidad.
- **Integración y pruebas del sistema:** Se integran todas las unidades. Se prueban en conjunto. Se entrega el conjunto probado al cliente.
- **Operación y mantenimiento:** Generalmente es la fase más larga. El sistema es puesto en marcha y se realiza la corrección de errores descubiertos. Se realizan mejoras de implementación. Se identifican nuevos requisitos.

Ilustración 004. Diagrama de Metodología Cascada



El desarrollo de las fases, como se ha mencionado antes, se produce de manera secuencial. Una vez se produce el análisis tanto del sistema como de los requisitos del software demandado por el cliente, (fases en las que la intervención del cliente es absolutamente necesaria), se procede a la fase de diseño de la arquitectura global de software. Un diseño elaborado de forma cuidadosa llevara a una rápida codificación. Tras haber traducido el programa a un lenguaje comprensible para el ordenador, se comprueban los elementos de forma individual y más tarde de manera homogénea (todos los sistemas a la vez). Una vez entregado el software al cliente, la fase de mantenimiento comprenderá las actualizaciones y las correcciones de errores que sean necesarias en el programa.

El modelo en cascada no permite retroceder, por lo que se hace estrictamente necesario que al final de cada fase el analista de sistemas o, en su caso, el programador, verifique y valide todo el trabajo realizado, ya que un error no detectado a tiempo podría perjudicar gravemente la fecha de entrega del software a nuestro cliente.

Este modelo tiene varios problemas, como pueden ser:

- Las iteraciones son costosas e implican rehacer trabajo debido a la producción y aprobación de documentos.
- Aunque son pocas iteraciones, es normal congelar parte del desarrollo y continuar con las siguientes fases.
- Los problemas se dejan para su posterior resolución, lo que lleva a que estos sean ignorados o corregidos de una forma poco elegante.
- Existe una alta probabilidad de que el software no cumpla con los requisitos del usuario por el largo tiempo de entrega del producto.
- Es inflexible a la hora de evolucionar para incorporar nuevos requisitos. Es difícil responder a cambios en los requisitos.

Por lo que este modelo sólo debe usarse si se entienden a plenitud los requisitos. En nuestro caso particular, en todos los proyectos de BI, el cliente o usuario final, establece los requisitos al inicio del proyecto, por lo que se tiene un conocimiento pleno de los requisitos que el sistema a de poseer. En caso de que el cliente tenga unos requerimientos diferentes o añada requisitos al final del proyecto, si el proyecto está bien construido, tan solo es necesario modificar los informes, es decir, las últimas fases del proyecto BI.

3.3- Herramientas Utilizadas

- **Microsoft Office: Excel, Word y Power Point**



Microsoft Office es una suite ofimática que abarca el mercado completo en Internet e interrelaciona aplicaciones de escritorio, servidores y servicios para los sistemas operativos Microsoft Windows y Mac OS X.

Microsoft Word es el procesador de texto. Word posee una posición dominante en el mercado de los procesadores de texto. Su formato propietario DOC es considerado un estándar, aunque en su versión Word 2007 utiliza un nuevo formato basado en XML llamado .DOCX, pero también tiene la capacidad de guardar y abrir documentos en el formato DOC.



Microsoft Excel es un programa de hoja o planilla de cálculo. Al igual que Microsoft Word, posee actualmente un mercado dominante. Está disponible para plataformas Windows y Macintosh.

Microsoft PowerPoint es un muy popular programa para desarrollar y desplegar presentaciones visuales en entornos Windows y Mac. Es usado para crear diapositivas multimediales, es decir, compuestas por texto, imágenes, sonido, animaciones y vídeos.



Utilizaremos Word en nuestro proyecto para redactar la toda la documentación. La utilización de Excel está ligada a la información que la empresa cliente nos facilita, en lugar de estar en una base de datos, en nuestro caso, la empresa cliente que contrata los servicios BI recopila toda su información en un documento hoja de cálculo tipo Excel. En cuanto a PowerPoint, se usará para la realización de las diapositivas de presentación de nuestro proyecto.

- **Star UML**



StarUML (Lenguaje Unificado de Modelado) es un lenguaje que nos permite modelar un sistema de software, es gráfico para visualizar, especificar, construir y documentar un sistema.

Esta herramienta ofrece la característica de describir un plano del sistema (modelo), donde podemos incluir o describir aspectos como: procesos de negocio o métodos, funciones del sistema entre otros. Ofrece la posibilidad de diagramar los requerimientos especificados de un sistema, starUML cuenta con varios tipos de diagramas, los cuales muestran diferentes aspectos a las entidades relacionadas.

Muy fácil de usar, debido a la simplicidad y rápida percepción de sus objetos, funciones y características, otra característica fundamental es que su código es compatible con lenguajes como C++ y Java.

El software heredó todas las características de la versión comercial y poco a poco ha ido mejorando sus características, entre las cuales se encuentran:

- Diagrama de casos de uso
- Diagrama de clase
- Diagrama de secuencia
- Diagrama de colaboración.
- Diagrama de estados
- Diagrama de actividad.
- Diagrama de componentes
- Diagrama de despliegue.
- Diagrama de composición estructural (UML 2.0)

- **Gliffy**

Gliffy es totalmente online y nos permitirá diseñar diagramas en línea de gran calidad en forma muy sencilla gracias a las herramientas de edición que ofrece. Si bien podemos utilizar esta aplicación sin necesidad de estar registrados, se requiere de una cuenta para guardar los diseños.



La principal característica de Gliffy es que no requiere de mayor experiencia para empezar a diseñar diagramas, tanto las herramientas que ofrece y la ventana de trabajo han sido diseñadas pensando en el usuario común y/o profesional.

Se utiliza para crear diagramas UML, planos, diagramas de Venn, diagramas de flujo y varios otros tipos de diagramas en línea. Los diagramas Gliffy pueden ser compartidos con y editados por los usuarios en tiempo real. Además es compatible con todos los navegadores web, como Google Chrome, Firefox, Safari e Internet Explorer 9 o superior.

- **Virtual Box**



Oracle VM VirtualBox es un software de virtualización para arquitecturas x86/amd64, creado originalmente por la empresa alemana innotek GmbH. Actualmente es desarrollado por Oracle Corporation como parte de su familia de productos de virtualización. Por medio de esta aplicación es posible instalar sistemas operativos adicionales, conocidos como «sistemas invitados», dentro de otro sistema operativo «anfitrión», cada uno con su propio ambiente virtual.

Resulta una aplicación extremadamente útil si estamos casi siempre trasteando con sistemas operativos nuevos o versiones Betas inestables. Un tema que tenemos que tener presente y que mucha gente me lo comenta es que la experiencia de uso del sistema no será igual a la de tener un sistema operativo instalado limpiamente en tu ordenador.

Entre los sistemas operativos soportados (en modo anfitrión) se encuentran GNU/Linux, Mac OS X, OS/2 Warp, Microsoft Windows, y Solaris/OpenSolaris, y dentro de ellos es posible virtualizar los sistemas operativos FreeBSD, GNU/Linux, OpenBSD, OS/2 Warp, Windows, Solaris, MS-DOS y muchos otros.

La aplicación fue inicialmente ofrecida bajo una licencia de software privativo, pero en enero de 2007, después de años de desarrollo, surgió VirtualBox OSE (Open Source Edition) bajo la licencia GPL 2. Actualmente existe la versión privativa Oracle VM VirtualBox, que es gratuita únicamente bajo uso personal o de evaluación, y está sujeta a la licencia de "Uso Personal y de Evaluación VirtualBox" (VirtualBox Personal Use and Evaluation License o PUEL) y la versión Open Source, VirtualBox OSE, que es software libre, sujeta a la licencia GPL.

Utilizaremos Virtual Box para todo nuestro proyecto. Todas las herramientas que utilizaremos en los diferentes procesos de BI serán instaladas, ejecutadas y utilizadas en nuestra máquina virtual.



Para trabajar en la máquina virtual utilizaremos el sistema operativo Windows Server 2012. Windows Server 2012 es la penúltima edición lanzada por Microsoft del sistema operativo Windows Server. Es la versión para servidores de Windows 8 y es el sucesor de Windows Server 2008 R2. El software está disponible para los consumidores desde el 4 de septiembre de 2012.

- **Microsoft Visual Studio**

Microsoft Visual Studio es un entorno de desarrollo integrado (IDE, por sus siglas en inglés) para sistemas operativos Windows. Soporta múltiples lenguajes de programación tales como C++, C#, Visual Basic .NET, F#, Java, Python, Ruby, PHP; al igual que entornos de desarrollo web como ASP.NET MVC, Django, etc., a lo cual sumarle las nuevas capacidades online bajo Windows Azure en forma del editor Monaco.



Visual Studio permite a los desarrolladores crear sitios y aplicaciones web, así como servicios web en cualquier entorno que soporte la plataforma .NET (a partir de la versión .NET 2002). Así se pueden crear aplicaciones que se comuniquen entre estaciones de trabajo, páginas web, dispositivos móviles, dispositivos embebidos, consolas, etc.

Visual Studio es, por tanto, un conjunto de herramientas de desarrollo de software y de otras tecnologías basado en componentes para crear aplicaciones eficaces de alto rendimiento.

Para nuestro caso en particular, agregaremos una extensión de Visual Studio para incluir una herramienta capaz de realizar ETLs e Informes. La extensión que debemos agregar a Visual Studio es la dedicada a Business Intelligence (SQL Server Data Tools - Business Intelligence (SSDT-BI)). Con esta ampliación, tenemos en el mismo IDE la capacidad de carga de información en el Data Warehouse y la capacidad de reporting, ambas imprescindibles para un proyecto de BI.

SQL Server Data Tools - Business Intelligence (SSDT-BI) para Visual Studio incluye herramientas de creación de proyectos de inteligencia empresarial y plantillas de proyectos para SQL Server Analysis Services, Reporting Services e Integration Services integrados en el shell de Visual Studio.

- **Sql Server**

Microsoft SQL Server es un sistema de manejo de bases de datos del modelo relacional (sistema de gestión de bases de datos relacionales - RDBMS), desarrollado por Microsoft para el entorno empresarial. SQL Server solo está disponible para sistemas operativos Windows de Microsoft.



El lenguaje de desarrollo utilizado (por línea de comandos o mediante la interfaz gráfica de Management Studio) es Transact-SQL (TSQL), una implementación del estándar ANSI del lenguaje SQL, utilizado para manipular y recuperar datos (DML), crear tablas y definir relaciones entre ellas (DDL). T-SQL es, por tanto, un conjunto de extensiones de programación de Sybase y Microsoft que añaden varias características a SQL estándar, incluyendo control de transacciones, excepción y manejo de errores, procesamiento fila, así como variables declaradas.

Características principales:

- Soporte de transacciones.
- Soporta procedimientos almacenados.
- Incluye también un entorno gráfico de administración, que permite el uso de comandos DDL y DML gráficamente.
- Permite trabajar en modo cliente-servidor, donde la información y datos se alojan en el servidor y los terminales o clientes de la red sólo acceden a la información.
- Además permite administrar información de otros servidores de datos.

- **Microstrategy**

MicroStrategy es una compañía que ofrece software de inteligencia de negocio y de informes para empresas. El software de MicroStrategy permite crear informes y análisis de datos almacenados en una Base de datos relacional y otras fuentes.



El software más reciente se llama MicroStrategy 9, Esta versión incluye importantes mejoras en una serie de áreas, proporcionando a los negocios la capacidad de construir sus propios cuadros de mando en cuestión de minutos. Asimismo, Visual Insigth es un producto que permite explorar los datos visualmente para descubrir ideas de negocio, analizar los datos importantes almacenados orientados a Big Data y mejorar la toma de decisiones.

Con la herramienta Microstrategy vamos a realizar el proceso de reporting de nuestro proyecto. Existen varias herramientas similares que podríamos utilizar: Qlick View, Microsoft Visual Studio. Pero nos resulta interesante la utilización de Microstrategy.

3.4- Planificación Temporal

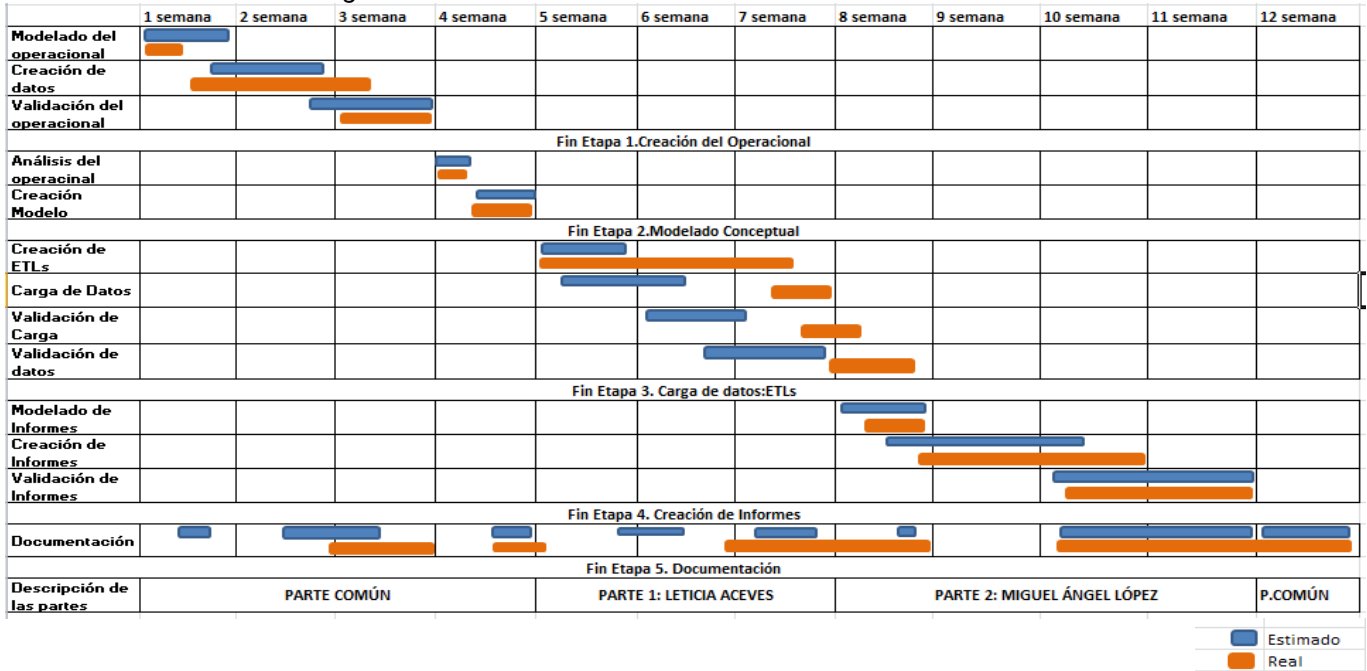
Definimos planificación temporal como la identificación de tareas, asignación de tiempos y recursos a dichas tareas y planificación de la secuencia de ejecución de forma que el tiempo de desarrollo del proyecto sea el mínimo posible.

La planificación temporal distribuye el esfuerzo estimado a lo largo de la duración prevista del proyecto. Los principios de la buena planificación temporal son:

- ✓ *Compartimentación*: descomposición del proyecto en un número manejable de actividades y tareas.
- ✓ *Interdependencia*: Se deben determinar las interdependencias de cada actividad o tarea compartimentada.
- ✓ *Asignación de tiempo*: A cada tarea que se vaya a programar se le deben asignar un cierto número de unidades de trabajo, una fecha de inicio y otra de finalización.
- ✓ *Validación del esfuerzo*: A medida que se realiza la asignación de tiempo, el gestor del proyecto se tiene que asegurar de que hay en plantilla el suficiente número de personas que se requiere en cada momento.
- ✓ *Responsabilidades definidas*: Cada tarea que se programe debe asignarse a un miembro específico del proyecto.
- ✓ *Resultados definidos*: El resultado de cada tarea, normalmente un producto, deberá estar definido. Los productos se combinan generalmente en entregas.
- ✓ *Hitos definidos*: Todas las tareas grupos de tareas deberían asociarse con algún hito del proyecto.

A partir de estas premisas, se ha elaborado un diagrama de Gantt de nuestro proyecto.

Ilustración 005. Diagrama de Gantt



Como se muestra en la figura, tenemos cinco etapas en el proyecto:

- La primera etapa del proyecto es la dedicada a la Creación del Operacional. Esta etapa es exclusiva de nuestro sistema, en los proyectos de BI no se tiene esta primera fase puesto que el operacional o la fuente de información viene dada por el cliente. En nuestro caso, es una empresa ficticia con datos creados por nosotros asemejándose lo más posible a datos reales.

A su vez, esta fase se encuentra dividida en tres, el Modelado del Operacional, donde se estudia cómo se va a estructurar la fuente de información. La creación de los datos de forma aleatoria y por último la validación de la información para que no existan errores ni incoherencias.

- La segunda etapa es la dedicada al Modelado Conceptual, esto es, a analizar los datos de los que disponemos y crear un modelado consistente (modelo en estrella o modelo en copo de nieve) frente a esos datos. Esta fase es muy importante puesto que un mal modelado puede hacer que el resto del proyecto, en un momento dado, no sea consistente y se tenga que volver a empezar con un modelado óptimo.
Esta etapa está dividida en dos partes, la primera es el análisis de información (en nuestro caso es corta porque el modelo operacional ha sido construido por nosotros) y una segunda parte de modelado conceptual en el que se elige entre la mejor opción (o el modelo en estrella o el copo de nieve) y se planifica su construcción.
- Una vez que tenemos el modelo conceptual, comienza la etapa de Carga de Datos en la base de datos.
Esta fase cuenta con cuatro partes: la primera es la creación en sí mismo de las ETLs para la carga de datos, seguidamente se realizan la carga de los datos en la base de datos y para finalizar esta fase, es importante validar que toda la información se ha cargado correctamente y que no existen inconsistencia en los datos.
- Cuando tenemos toda la información bien estructurada según el modelo conceptual en una base de datos o Data Warehouse, comienza la fase de Creación de Informes o Reporting. En esta fase se visualizan los datos contenidos en la base de datos de manera acorde con los requisitos del cliente.
En esta etapa, también tenemos divididas las fases, de tal modo que después de la creación de los informes como tal, es importante validar que los datos que se muestran en dichos informes coinciden con los datos contenidos en la base de datos.
- La última fase es la de Documentación, esta fase es común a todo proyecto, y dá comienzo en el mismo momento en que se inicia el proyecto y termina en las últimas semanas de la creación del mismo. La fase de documentación, por tanto, se debe llevar tener en cuenta durante todo el proceso de creación de software.

Es importante discernir qué persona se va a dedicar a qué función a lo largo de la creación del sistema. En nuestro caso, tenemos partes comunes a ambos desarrolladores, como puede ser la creación del operacional que servirá de base para el resto del proyecto, la creación del modelado la documentación final. Mientras que el proceso de carga de las ETLs será realizado por Leticia Aceves y el proceso de reporting o creación de informes, será responsabilidad de Miguel Ángel López.

En la ilustración, también encontramos diferenciados el tiempo estimado en un principio con lo que se ha tardado realmente en realizar todo el proceso de BI:

- En la creación de datos del operacional ha llevado una semana más de lo planeado, pero se ha compensado con el tiempo dedicado a la validación, que ha sido menor.
- La carga de datos se ha retrasado una semana, ya que el tiempo de la creación de ETLs ha sido más del doble de lo estimado previamente. En la última etapa de la validación de ETLs se ha comenzado con la creación de los informes, de esta forma el tiempo se ha podido ajustar a lo previsto.
- La creación de informes se ha ajustado bastante bien al tiempo previsto.
- Finalmente, en el caso de la documentación, la mayor parte se ha realizado en las últimas semanas.

3.5- Presupuesto

El Presupuesto constituye una de las herramientas fundamentales de las que dispone un profesional que se dedique al Desarrollo de Software. En muchos casos constituye nuestra primera toma de contacto y nuestra carta de presentación de cara al cliente. Un presupuesto mal redactado o con un precio incorrecto hará que nuestra relación se acabe antes de comenzar. También es fundamental el tiempo que tardemos en entregar el presupuesto, pues si se dilata excesivamente, el cliente tomará partida por otros que ya tenga en mano.

Estamos acostumbrados a utilizar el método del COCOMO para estimaciones presupuestarias, pero en nuestro caso particular, se hace complicado la utilización de este método.

Definimos el Modelo Constructivo de Costos (o COCOMO, por su acrónimo del inglés COConstructive COSt MOdel) como un modelo matemático de base empírica utilizado para estimación de costos de software. Pertenece a la categoría de modelos de subestimaciones basados en estimaciones matemáticas. Está orientado a la magnitud del producto final, midiendo el "tamaño" del proyecto, en líneas de código principalmente.

En la definición tenemos la respuesta de porqué es difícil realizar una estimación presupuestaria óptima utilizando COCOMO en nuestro sistema. COCOMO recurre a la magnitud del producto final midiendo las líneas de código utilizadas, pero en nuestro caso en particular, no tenemos líneas de código como tal, por lo tanto, una estimación con COCOMO no se aproximaría a la realidad.

Así pues, dividiremos el presupuesto en dos partes, la primera describe el coste hardware y el coste del software utilizado, y por último el coste del desarrollo o mano de obra. Para todas estas partes, además del tiempo de realización del proyecto, debemos tener en cuenta el tiempo de formación de los desarrolladores en el tema de Business Intelligence, este tiempo asciende a tres meses.

3.5.1- Presupuesto Hardware y Software

Para comenzar con el presupuesto, vamos a tener en cuenta el hardware y software necesario para la realización del proyecto.

- En cuanto a Hardware, se han utilizado:
 - Dos ordenadores personales de forma concurrente. Para el análisis, el desarrollo del sistema, la implantación y pruebas del mismo y la generación de la documentación.
 - Conexión a Internet para la obtención de información y la descarga de software.
 - Impresora para imprimir la documentación e información necesaria.

Tabla 001. Presupuesto; Hardware

HARDWARE	USO (%)	UNIDADES	COSTE (€)	TOTAL (€)
Ordenador Personal	20%	1	740	148
Ordenador Personal	34%	1	660	224,40
Conexión a Internet	(6 meses)	1	19,95/mes	119,7
Impresora	10%	1	55	5,5
			TOTAL:	497,60 €

- En cuanto a software, para el desarrollo de la aplicación se utilizarán las siguientes herramientas:
 - Windows 8
 - Oracle VM VirtualBox
 - Windows Server 2012
 - Microsoft Visual Studio
 - SQL Server 2012
 - Microstrategy
 - StarUML.
 - Gliffy
 - Microsoft Office 2007.

Tabla 002. Presupuesto; Software

SOFTWARE	USO (%)	UNIDADES	COSTE (€)	TOTAL (€)
Windows 8	12,5%	2	94,95€	23,74€
Oracle VM VirtualBox	100%	2	0	0
Windows Server 2012	100%	2	190€	380€
Microsoft Visual Studio	25%	1	400€	100€
SQL Server 2012	50%	2	320€	320€
Microstrategy	25%	1	1200€	300€
StartUML	1%	2	0	0
Gliffy	1%	2	0	0
Microsoft Office 2007	95%	2	50€	95€
TOTAL:				1.218,74€

3.5.2- Presupuesto de Desarrollo

Para el presupuesto invertido en desarrolladores, se ha tenido en cuenta que la estimación de los trabajos realizados por 2 personas durarán 3 meses, más el tiempo previo utilizado en su formación, que ha sido aproximadamente de otros 3 meses.

El tiempo dedicado a la formación ha sido de 7 horas al día durante 3 meses (12 semanas). Mientras que para el desarrollo del proyecto en sí, cada día se pretende trabajar 8 horas (solo días laborables, de lunes a viernes, festivos no incluidos) y que el número de días trabajados al mes son 22, calculamos el número de horas que va a trabajar una persona.

En la siguiente tabla se encuentra un desglose de las tareas necesarias para llevar a cabo el proyecto y la duración estimada de las mismas en horas de trabajo.

Tabla 003. Presupuesto; Horas

TAREA	DESARROLLADORES	DURACIÓN (HORAS)	TOTAL (HORAS)
Formación	2	420h	840h
Creación del Operacional	2	120h	240h
Modelado Conceptual	2	40h	80h
Carga de Datos ETIs	1	120h	120h
Creación de Informes	1	160h	160h
Documentación	2	40h	80h
TOTAL:			1.520h

***Nota:** a tener en cuenta que la documentación se realiza a lo largo de todo el proyecto, es decir, que parte de las horas dedicadas a cada tarea pertenecen a la documentación de esa tarea.

A este resultado de horas lo multiplicaremos por el sueldo por hora de un Titulado de Grado en Informática, se obtiene el coste total de la mano de obra:

Tabla 004. Presupuesto; Coste Desarrollo

	TIEMPO (2 personas)	COSTE	TOTAL (2 personas)
Ingeniero	1.520h	10€/h	15.200€

3.5.3- Presupuesto Total

La estimación del presupuesto total es la suma de las cuantías presupuestarias que se han estimado anteriormente.

Tabla 005. Presupuesto; Total

PRESUPUESTO	TOTAL (€)
Hardware	497,6€
Software	1.218,74€
Coste Desarrollo	15.200€
Total	16.916,34€

CAPITULO IV : ANÁLISIS DEL PROYECTO

CAPITULO 4: ANALISIS DEL PROYECTO

En este cuarto capítulo, se analizará el proyecto de forma más exhaustiva, explicando sus usuarios principales, y los principales requisitos que el proyecto deberá abordar una vez terminado. Probablemente la fase de análisis sea la más decisiva del proyecto, porque es en la que encontramos las necesidades del usuario final.

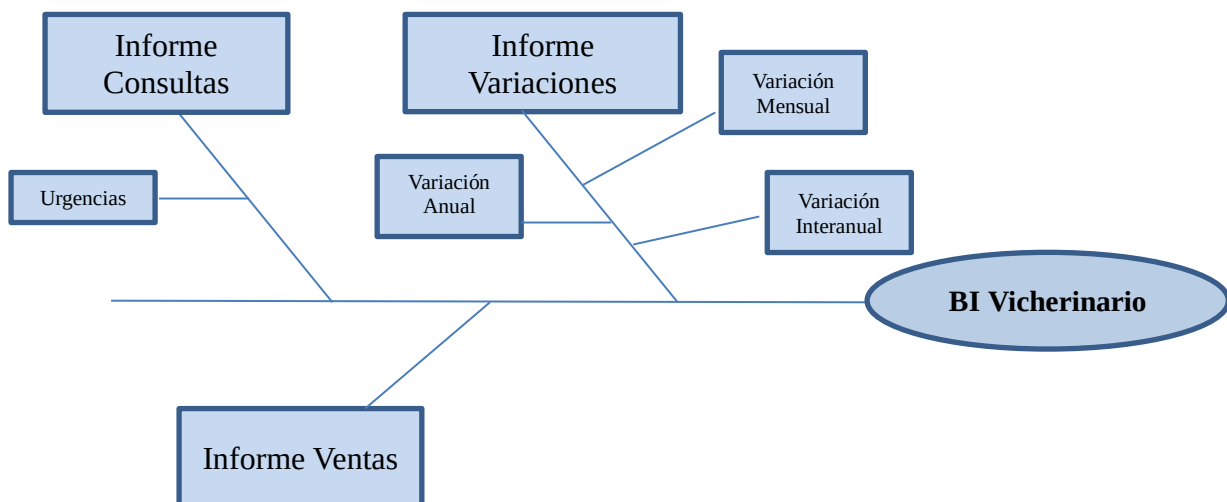
Debemos detallar bien todos y cada uno de los requisitos del sistema ya que este grado de detalle tiene como objetivo eliminar sorpresas durante la ejecución del proyecto y garantizar a ambas partes la claridad de los trabajos a realizar en la fase de ejecución.

4.1- Características Principales

En esta sección listaremos las características principales del producto. Estas características definirán nuestro proyecto de forma única.

El árbol de características es una representación visual de las características del producto agrupadas lógicamente.

Ilustración 006. Diagrama de Árbol de Características



- **C1:** El sistema de información deberá ser sencillo para el usuario, con una interfaz fácil y entendible.
- **C2:** El sistema tiene, como mínimo, que recoger todos los informes que el cliente nos ha demandado.
- **C3:** El sistema pedirá nombre de usuario y contraseña para acceder a la información
- **C4:** El usuario final o cliente, tan solo podrá visualizar los informes, sin, en ningún momento, poder hacer cambio alguno en ellos.
- **C5:** Se podrá acceder al sistema desde cualquier navegador, Tablet o móvil.
- **C6:** Los informes serán claros y entendibles, con los datos más importantes destacados de alguna forma: negrita, cursiva, otro color...
- **C7:** Cada informe deberá contener unos filtros que faciliten las búsquedas de información.
- **C8:** Los informes tendrán que contener al menos un gráfico lo suficientemente significativo.

4.2- Descripción de Actores del Sistema

En las siguientes tablas se describen los diferentes actores del sistema, estos son, los distintos tipos de personas o grupos de personas que van a interactuar con el sistema en algún momento.

En el caso de nuestro proyecto de BI, tan solo existen dos tipos de actores o usuarios del sistema. El administrador, que define a la persona o personas dedicadas a la realización del proyecto, este actor debe tener todos los permisos sobre el sistema. El administrador es el encargado de crear, modificar o borrar distintas partes del sistema a petición del cliente.

El usuario, es un actor que tan sólo va a interactuar con el sistema una vez esté terminado. Tan sólo tendrá permisos de lectura. El cliente final, en todo proyecto BI, recibe los informes creados y los visualiza.

Tabla 006. Actor-01; Administrador

ACT-01	Administrador
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Descripción	Este actor representa al Administrador del sistema
Comentarios	Este actor es el encargado de realizar todos los cambios oportunos en el sistema y en los informes, tales como: modificar informes, crear nuevos informes y borrar informes existentes.

Tabla 007. Actor-02; Usuario

ACT-02	Usuario
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Descripción	Este actor representa a un usuario del sistema
Comentarios	Este usuario final, tan sólo podrá visualizar los informes que se le presentan.

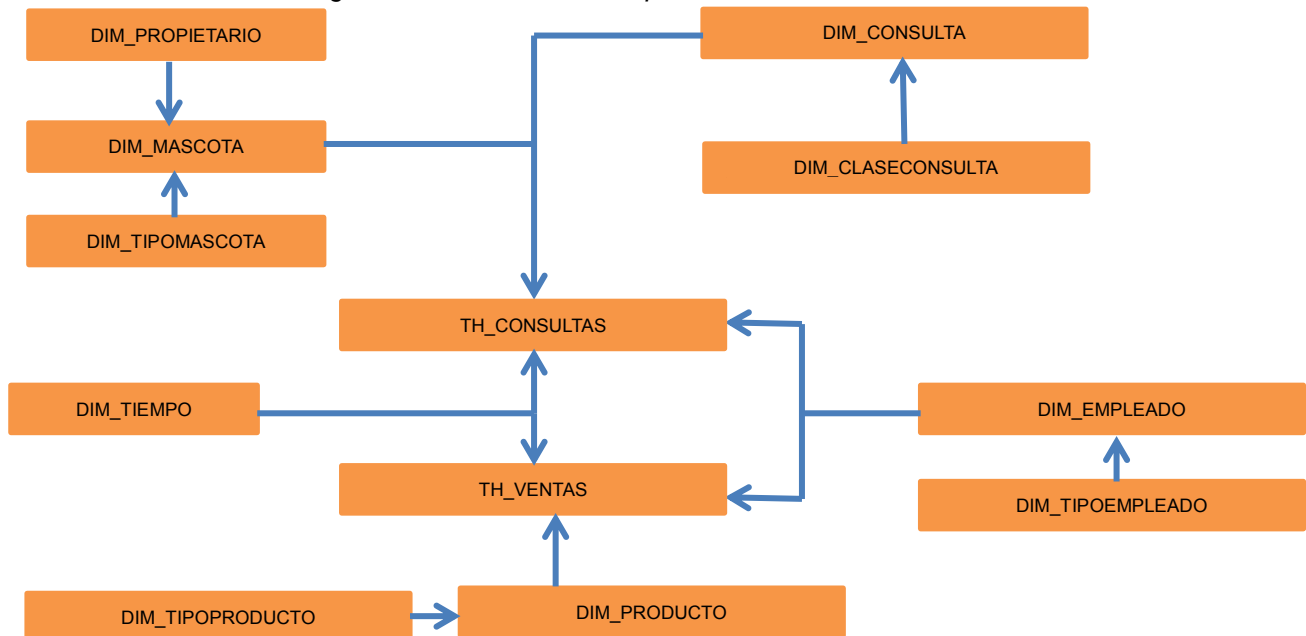
4.3- Requisitos de Información

Una vez que tenemos claro lo que vamos a hacer, las herramientas, las características que tiene que tener nuestro sistema y sobre todo, tenemos claro la información de la que disponemos y la finalidad de esta información. Debemos realizar la siguiente fase de un proyecto de BI: el modelado.

Para un correcto modelado debemos conocer con precisión las fuentes de información y tener claro lo que quiere el usuario final.

Como se ha explicado en las fases de un proyecto BI (en el capítulo II de este documento) existen dos modelados básicos, el modelo en estrella y el modelo en copo de nieve. Aunque cualquier de los dos modelos nos sirve para nuestro sistema BI de clínica veterinaria, nos vamos a quedar con el modelo en copo de nieve, puesto que este modelo permite jerarquías mientras que el modelo en estrella, no lo permite.

Ilustración 007. Diagrama de Modelado en Copo de Nieve de Nuestro Sistema



El modelo es sencillo, al ser una clínica veterinaria, vamos a tener dos operativas claramente diferenciadas dentro de la misma empresa. Por un lado tenemos las consultas realizadas a mascotas y por otro lado tenemos las ventas de productos para animales.

Si distinguimos dos actividades, tenemos dos tablas de hechos diferenciadas, una dedicada a consultas y otra a ventas.

Como en cualquier sistema de BI, tendremos una dimensión de tiempo, la dimensión temporal será común para las dos tablas de hechos, es lógico, pues en un mismo día puedes tener consultas y ventas.

En nuestro caso ambas tablas de hechos contienen otro dato que va a ser común, además del tiempo: el empleado. Esto es así, por qué no hay que olvidar que es una misma empresa y los empleados, cualquiera de ellos, puede realizar la venta de un producto, así como, un veterinario (Doctor), puede realizar consultas a las mascotas. Por tanto, un mismo empleado, puede realizar una consulta o una venta.

Si tenemos una tabla de hechos dedicada a las consultas, debemos saber a qué mascota se le realiza qué consulta, por qué empleado de la clínica se ha realizado esa consulta así como el importe que se ha cobrado. El importe a cobrar dependerá de si la consulta es de tipo urgente o normal, haciéndose un incremento del 20% en las consultas de tipo urgente.

Por lo tanto, a la tabla de hechos de consultas, irán relacionadas las dimensiones de: DIM_EMPLEADO, DIM_MASCOTA, DIM_CONSULTA y DIM_TIEMPO.

En cuanto a la tabla de hechos dedicada a las ventas, está claro que la clínica tiene que registrar el producto que se ha vendido, la cantidad de unidades de dicho producto, la persona que lo ha vendido y el importe de la compra.

A la tabla de hechos de ventas, se relacionarán por tanto: DIM_TIEMPO, DIM_PRODUCTO y DIM_EMPLEADO.

Las jerarquías de las dimensiones dependerán de la forma de trabajar de la empresa. En nuestro caso, cada dimensión tiene, al menos, una jerarquía:

- Las mascotas, que son clientes de la clínica, pertenecen a un solo propietario, pero el mismo propietario puede tener varias mascotas. En la dimensión de mascotas, tenemos otra jerarquía, y es que, la clínica con frecuencia imprime o visualiza listados con todos sus perros o gatos....es decir, de cada tipo de mascota, por lo que esto, lo tienen diferenciado y es otra jerarquía.
- En la empresa hay actualmente contratados cuatro especialistas, dos de ellos son administrativos y otros dos son veterinarios. Esta diferenciación es otra jerarquía.
- La compañía distingue entre varios tipos de consultas de forma genérica y dentro de estos tipos de consultas están las consultas en sí mismas. Un ejemplo de ello son las pruebas de laboratorio, que pueden ser: análisis de sangre, orina....
- En cuanto a los productos, diferenciamos, por ejemplo, los que son de tipo alimenticio y dentro de ellos: piensos para gatos, perros, hámster....por lo que, DIM_TIPOPRODUCTO y DIM_PRODUCTO son también una jerarquía.

4.3.1- Granularidad

La granularidad representa el nivel de detalle al que se desea almacenar la información sobre el negocio que se esté analizando. Mientras mayor sea el nivel de detalle de los datos, se tendrán mayores posibilidades de análisis, ya que los mismos podrán ser resumidos. Es decir, los datos que posean granularidad fina (nivel de detalle) podrán ser resumidos hasta obtener una granularidad media o gruesa.

Por tanto, una característica importante que define a una tabla, ya sea de dimensiones o de hechos, es el nivel de granularidad de los datos que en ella se almacenan. La granularidad de una tabla representa el nivel más atómico por el cual se definen los datos. Por ejemplo, no es lo mismo contar el tiempo por horas (grano fino) que por semanas (grano medio) que por meses (grado grueso); o en el caso de los productos, se puede considerar cada variante de un mismo artículo como un producto (por ejemplo, en una empresa textil, cada talla y color de pantalón podría ser un producto) o agrupar todos los artículos de una misma familia considerándolos como un único producto (por ejemplo, el producto pantalón genérico).

Como se puede observar, la granularidad afecta a la cardinalidad (cantidad de registros de una tabla), tanto de las dimensiones como de la tabla de hechos, a mayor granularidad (grano más fino) mayor será el número de registros final de la tabla de hechos.

Cuando la granularidad es mayor, es frecuente que se desee disponer de subtotaes parciales, es decir, si tenemos una tabla de hechos con las ventas por días, podría interesar disponer de los totales semanales o mensuales, estos datos se pueden calcular haciendo sumas parciales, pero es frecuente añadir a la tabla de hechos registros donde se almacenan dichos cálculos para no tener que repetirlos cada vez que se requieran y mejorar así el rendimiento de la aplicación. En este caso se dispondrá en la misma tabla de hechos de datos de grano fino y de grano más grueso aumentando aún más la cardinalidad de la tabla.

Una vez que se ha entendido el concepto de Granularidad, vamos a enumerar las distintas granularidades que en nuestro proyecto se abarcan por tabla.

Granularidad de las dimensiones:

- La granularidad de la dimensión de tiempo es ID_MESANO
- La granularidad de las dimensiones dedicadas a las mascotas es el ID_MASCOTA
- La granularidad en la tabla de propietarios es ID_PROPIETARIO
- La granularidad en la tabla de tipo de mascotas es ID_TIPOMASCOTA
- La granularidad de las dimensiones de empleado es el ID_EMPLEADO
- La granularidad en la tabla de cargo es ID_CARGO
- La granularidad de las dimensiones de consulta es ID_CONSULTA
- La granularidad en la tabla de clase consulta es ID_CLASECONSULTA
- La granularidad de las dimensiones de producto es ID_PRODUCTO
- La granularidad en la tabla de tipo producto es ID_TIPOPRODUCTO

Es normal que la granularidad de las dimensiones se corresponda con los ids o identificadores, ya que estos identificadores deben ser únicos por campo, por lo que, serán los que indicarán el grano en las dimensiones.

Granularidad de la tabla de hechos:

-La granularidad para la tabla de hechos dedicada a las consultas será: ID_MESANO.

-Por otro lado, la granularidad para la tabla de hechos dedicada a los productos vendidos, será igualmente de: ID_MESANO.

Lo lógico sería establecer una granularidad a nivel de día, puesto que la empresa contratante sobre la que se va a realizar el proceso BI, realiza consultas y ventas diarias, pero el cliente requiere variaciones a nivel mensual y anual, por lo que la granularidad se establecerá a nivel mensual. Para pasar de un grano diario a uno mensual se realizan agrupaciones en las ETLs en el proceso de carga de datos.

Se debe diferenciar el grano de las PK o primary key de las bases de datos. El campo ID_MESANO no es suficiente para establecer como PK de las tablas de hechos, debemos tener en cuenta otros atributos para evitar la pérdida de información y redundancia en datos. Por ejemplo, es más probable que se haga la misma venta dos veces en el mismo mes (por lo que se sumarian las cantidades), que se haga la misma venta dos veces en el mismo mes por el mismo empleado y del mismo producto.

4.3.2- Definición Dimensiones

Denominamos dimensiones a aquellos datos que nos permiten filtrar, agrupar o seccionar la información. Representan los aspectos de interés, mediante los cuales los usuarios podrán filtrar y manipular la información almacenada en la tabla de hechos. Las tablas de dimensiones definen como están los datos organizados lógicamente y proveen el medio para analizar el contexto del negocio. Contienen datos cualitativos

Las Dimensiones (también llamada entidad, perspectiva, característica, dato maestro, etc.) son los grupos de datos que permiten identificar quién, cuándo o dónde se genera una operación o transacción del negocio. Por ejemplo: clientes, proveedores, tiempo y lugar son dimensiones típicas.

La agrupación de dimensiones según una relación de dependencia lógica, se le denomina jerarquías.

Los datos dentro de estas tablas, que proveen información del negocio o que describen alguna de sus características, son llamados datos de referencia.

Más detalladamente, cada tabla de dimensión podrá contener los siguientes campos:

- Clave principal o identificador único.
- Clave foráneas.
- Datos de referencia primarios: datos que identifican la dimensión. Por ejemplo: nombre del cliente.
- Datos de referencia secundarios: datos que complementan la descripción de la dimensión. Por ejemplo: e-mail del cliente, fax del cliente, etc.

• Dimensión de Tiempo

En un DW, la creación y el mantenimiento de una tabla de dimensión Tiempo es obligatoria, y la definición de granularidad y estructuración de la misma depende de la dinámica del negocio que se esté analizando, en nuestro caso, a nivel mensual. Toda la información dentro del depósito, posee su propio sello de tiempo que determina la ocurrencia de un hecho específico.

Es importante tener en cuenta que la dimensión tiempo no es sola una secuencia cronológica representada de forma numérica, sino que mantiene niveles jerárquicos especiales que inciden notablemente en las actividades de la organización. Esto se debe a que los usuarios podrán por ejemplo analizar las ventas realizadas teniendo en cuenta el día de la semana en que se produjeron, quincena, mes, trimestre, semestre, año, estación, etc.

Existen muchas maneras de diseñar esta tabla, y en adición a ello no es una tarea sencilla de llevar a cabo. Por estas razones se considera una buena práctica evaluar con cuidado la temporalidad de los datos, la forma en que trabaja la organización, los resultados que se esperan obtener del almacén de datos relacionados con una unidad de tiempo y la flexibilidad que se desea obtener de dicha tabla.

4.3.3- Definición Tablas de Hechos

Una tabla de hechos es una representación de un proceso de negocio. A nivel de diseño es una tabla que permite guardar dos tipos de atributos diferenciados:

- Medidas del proceso / actividad / flujo de trabajo / evento que se pretende modelizar.
- Claves foráneas hacia registros en una tabla de dimensión (o en otras palabras, como ya sabemos, hacia una vista de negocio).

Cuando estamos construyendo nuestro Data Warehouse tenemos que diseñar la tabla central que es la que guardará los hechos. A diferencia de un sistema transaccional donde en una tabla tenemos el total de la factura, en otra el total de la orden de compra, en otro el tipo de cambio (y así sucesivamente) en un Data Warehouse los hechos (las cosas que sucedieron) están en una única tabla.

Normalmente las cosas que están en la tabla de hechos tienen afinidad entre sí. De esta forma tendremos una tabla de Hechos de Ventas, una de inventario, una de Recursos Humanos, una de producción, etc.

En las bases de datos, y más concretamente en un Data Warehouse, una tabla de hechos (o tabla fact) es la tabla central de un esquema dimensional (en estrella o en copo de nieve) y contiene los valores de las medidas de negocio o dicho de otra forma los indicadores de negocio. Cada medida se toma mediante la intersección de las dimensiones que la definen, dichas dimensiones estarán reflejadas en sus correspondientes tablas de dimensiones que rodearán la tabla de hechos y estarán relacionadas con ella.

Como ya se ha explicado anteriormente, como el cliente, la clínica veterinaria, tiene dos actividades diferenciadas, dos tipos de abonos diferentes, se han reflejado dos tablas de hechos con hechos distintos (TH_VENTAS y TH_CONSULTAS).

- **Indicadores**

Los Indicadores o Métricas (también llamados KPIs, Valores, etc.) son los valores numéricos generados en una o varias operaciones o transacciones de negocio. Por ejemplo: total ventas, total costos, total pagos. Ayudan a responder las preguntas referidas a cantidades o importes.

Los indicadores más útiles para incluir en una tabla de hechos son los aditivos, es decir, aquellas medidas que pueden ser sumadas como por ejemplo la cantidad de producto vendido, los costes de producción o el dinero obtenido por las ventas; son medidas numéricas que pueden calcularse con la suma de varias cantidades de la tabla.

En este apartado se describen los indicadores que van a formar parte de las tablas de hechos de nuestro proyecto. Estos indicadores son independientes de las dimensiones, en el sentido de que no están contenidos en ninguna de estas, pues reflejan el trabajo diario de la entidad.

En la tabla de hechos de consultas, tenemos los indicadores:

- *Tipo de Consulta*: este indicador hace referencia a si una consulta es de tipo urgente o normal. En el caso de que la consulta fuera de tipo urgente, el importe total de la consulta (otro indicador) se incrementará en un 20%. La urgencia está asociada a una mascota en concreto, en un momento puntual en el tiempo y a un empleado. Este indicador se identificará con el siguiente nombre: Tipo_consulta.

- *Importe Consulta*: como su propio nombre indica, hace referencia al importe total que la Entidad ha cobrado por una consulta determinada en un determinado momento y a una mascota en concreto. Este indicador se identificará con la siguiente nomenclatura *Importe_Consulta*.

En la tabla de hechos de ventas, tenemos los indicadores:

- *Unidades Vendidas*: hace referencia a la cantidad de productos que se venden en un momento determinado. Este indicador se identificará como *Unidades_Vendidas*.
- *Precio Venta*: refiere al saldo en euros asociado a una venta en un día determinado y de un producto y una cantidad de este producto, realizada por un empleado en cuestión. Se define como *Precio_Venta*.

4.3.4- Especificación Dimensiones

El modelo analítico deberá contemplar las siguientes capacidades de análisis con las que se le permitirá al usuario una navegación en las dimensiones que conforman cada informe.

Las dimensiones deben tener relación con otras dimensiones o bien con la tabla de hechos. De esta forma, desde una dimensión se puede llegar a una tabla de hechos y desde la tabla de hechos a las dimensiones de mayor y menor granularidad con las que tenga algún tipo de dependencia.

Tabla 008. Dimensiones; DIM_TIEMPO

Nombre	DIM_TIEMPO
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	Creada por la empresa de BI
Granularidad	ID_MESANO
Jerarquías	Ninguna
Relaciones	Con las tablas de hechos: TH_VENTAS y TH_CONSULTAS
Campos	Id_mesano, ano,mes, desc_mes, mes_ant, ano_ant, mesant_mismoano, mes_anoant
Descripción	Identifica el eje de análisis temporal.
Comentarios	La granularidad se hará a nivel de mes para poder realizar las variaciones de forma eficiente. Si se realizara a nivel diario se perdería información ya que no todos los días se pasa consulta o se venden productos.

Tabla 009. Dimensiones; DIM_EMPLEADO

Nombre	DIM_EMPLEADO
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	Documento Hoja de cálculo de Vicherinario; la hoja denominada "Empleados"
Granularidad	ID_EMPLEADO
Jerarquías	Hijo de DIM_CARGO
Relaciones	Con las tablas de hechos: TH_VENTAS y TH_CONSULTAS Además de con la DIM_CARGO como su hija
Campos	Id_empleado, cod_empleado, nombre, apellidos, teléfono, id_cargo
Descripción	Esta dimensión identifica los distintos empleados que tiene la empresa Vicherinario.
Comentarios	Aunque en el documento ambas dimensiones (cargo y empleado) aparecen en la misma hoja de cálculo, mediante un proceso ETL las diferenciaremos como dos dimensiones dispares.

Tabla 010. Dimensiones; DIM_CARGO

Nombre	DIM_CARGO
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	Documento Hoja de cálculo de Vicherinario; la hoja denominada "Empleados"
Granularidad	ID_CARGO
Jerarquías	Padre de DIM_EMPLEADO
Relaciones	Con la DIM_EMPLEADO como su padre
Campos	Id_cargo, cod_cargo, cargo
Descripción	Es jerárquicamente dependiente de la DIM_EMPLEADO, definiendo los diferentes tipos de cargos que se tiene en la empresa. Cada empleado tiene su cargo asignado: administrativo o doctor.
Comentarios	Aunque en el documento ambas dimensiones (cargo y empleado) aparecen en la misma hoja de cálculo, mediante un proceso ETL las diferenciaremos como dos dimensiones dispares.

Tabla 011. Dimensiones; DIM_MASCOTA

Nombre	DIM_MASCOTA
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	Documento Hoja de cálculo de Vicherinario; la hoja denominada "Mascotas"
Granularidad	ID_MASCOTA
Jerarquías	Hija de DIM_PROPIETARIO y DIM_TIPOMASCOTA
Relaciones	Con la DIM_PROPIETARIO y DIM_TIPOMASCOTA como su hija y con la Tabla de Hechos TH_CONSULTAS
Campos	Id_mascota, cod_mascota, nombre, raza, sexo, peso, fecha_nacimiento, fecha_defunción, id_tipomascota, id_propietario
Descripción	Esta dimensión identifica a todos los animales que son clientes habituales de la clínica veterinaria.
Comentarios	Aunque en el documento facilitado por la empresa las tres dimensiones: mascotas, propietarios y tipo mascota, aparecen en la misma hoja de cálculo, mediante un proceso ETL las diferenciaremos como tres dimensiones dispares.

Tabla 012. Dimensiones; DIM_PROPIETARIO

Nombre	DIM_PROPIETARIO
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	Documento Hoja de cálculo de Vicherinario; la hoja denominada "Mascotas"
Granularidad	ID_PROPIETARIO
Jerarquías	Padre de DIM_MASCOTA
Relaciones	Con la DIM_MASCOTA como su hija.
Campos	Id_propietario, cod_propietario, nombre, dirección, teléfono
Descripción	Esta dimensión, dependiente jerárquicamente de la anterior, identifica los distintos propietarios de los animales que la clínica atiende diariamente.
Comentarios	Aunque en el documento facilitado por la empresa las tres dimensiones: mascotas, propietarios y tipo mascota, aparecen en la misma hoja de cálculo, mediante un proceso ETL las diferenciaremos como tres dimensiones dispares.

Tabla 013. Dimensiones; DIM_TIPOMASCOTA

Nombre	DIM_TIPOMASCOTA
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	Documento Hoja de cálculo de Vicherinario; la hoja denominada "Mascotas"
Granularidad	ID_TIPOMASCOTA
Jerarquías	Padre de DIM_MASCOTA
Relaciones	Con la DIM_MASCOTA como su hija.
Campos	Id_tipomascota, cod_tipomascota, tipomascota
Descripción	En esta dimensión, dependiente jerárquicamente de DIM_MASCOTA, se identifica las distintas agrupaciones de mascotas a los que la entidad realiza sus consultas.
Comentarios	Aunque en el documento facilitado por la empresa las tres dimensiones: mascotas, propietarios y tipo mascota, aparecen en la misma hoja de cálculo, mediante un proceso ETL las diferenciaremos como tres dimensiones dispares.

Tabla 014. Dimensiones; DIM_CONSULTA

Nombre	DIM_CONSULTA
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	Fichero Excel proporcionado por la entidad. La hoja de cálculo a la que refiere es "Consultas".
Granularidad	ID_CONSULTA
Jerarquías	Hija de DIM_TIPOCONSULTA
Relaciones	Con la DIM_TIPOCONSULTA como su hija. Y la tabla de hechos TH_CONSULTAS
Campos	Id_consulta, cod_consulta, descripción, precio, id_clase
Descripción	Dimensión dedicada a definir las consultas o tratamientos que realiza la clínica veterinaria.
Comentarios	Aunque en el documento facilitado por la empresa las dos dimensiones: consultas y tipo consulta, aparecen en la misma hoja de cálculo como campos distintos, mediante un proceso ETL las diferenciaremos como dos dimensiones dispares.

Tabla 015. Dimensiones; DIM_CLASECONSULTA

Nombre	DIM_CLASECONSULTA
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	Fichero Excel proporcionado por la entidad. La hoja de cálculo a la que refiere es "Consultas".
Granularidad	ID_CLASECONSULTA
Jerarquías	Padre de DIM_CONSULTA
Relaciones	Con la DIM_CONSULTA como su hija.
Campos	Id_clase, cod_clase, claseconsulta.
Descripción	Esta dimensión identifica los tipos de consultas existentes en la entidad, esto es una dependencia jerárquica del anterior.
Comentarios	Aunque en el documento facilitado por la empresa las dos dimensiones: consultas y tipo consulta, aparecen en la misma hoja de cálculo como campos distintos, mediante un proceso ETL las diferenciaremos como dos dimensiones dispares.

Tabla 016. Dimensiones; DIM_PRODUCTO

Nombre	DIM_PRODUCTO
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	El documento Excel de la empresa y la hoja de cálculo nombrada como "Productos".
Granularidad	ID_PRODUCTO
Jerarquías	Hijo de DIM_TIOPRODUCTO
Relaciones	Con la DIM_TIOPRODUCTO como su hija.
Campos	Id_producto, cod_producto, nombre, descripción, precio, id_tipoproducto
Descripción	Esta dimensión constituye la enunciación de los distintos productos que la Entidad tiene a la venta.
Comentarios	Aunque en el documento facilitado por la empresa las dos dimensiones: producto y tipo producto, aparecen en la misma hoja de cálculo como campos distintos, mediante un proceso ETL las diferenciaremos como dos dimensiones dispares.

Tabla 017. Dimensiones; DIM_TIOPRODUCTO

Nombre	DIM_TIOPRODUCTO
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	El documento Excel de la empresa y la hoja de cálculo nombrada como "Productos".
Granularidad	ID_TIOPRODUCTO
Jerarquías	Padre de DIM_PRODUCTO
Relaciones	Con la DIM_PRODUCTO como su hija.
Campos	Id_tipoproducto, cod_tipoproducto, tipoproducto
Descripción	Dependiente jerárquicamente de la anterior, de mayor granularidad, identifica los diferentes tipos de productos de la empresa.
Comentarios	Aunque en el documento facilitado por la empresa las dos dimensiones: producto y tipo producto, aparecen en la misma hoja de cálculo como campos distintos, mediante un proceso ETL las diferenciaremos como dos dimensiones dispares.

4.3.5- Especificación Tablas de Hechos

Tabla 018. Tabla de Hechos; TH_CONSULTAS

Nombre	TH_CONSULTAS
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	El documento Excel de la empresa y la hoja de cálculo nombrada como "Diario".
Granularidad	ID_MESANO
Relaciones	DIM_TIEMPO, DIM_MASCOTA, DIM_EMPLEADO e DIM_CONSULTA
Indicadores	Tipo Consulta e Importe Consulta
Campos	ID_MESANO, ID_MASCOTA, ID_EMPLEADO, ID_CONSULTA, Tipo_consulta, importe_consulta
Descripción	Esta tabla recoge toda la información concerniente a las consultas realizadas a las mascotas durante los 2 últimos años.
Comentarios	El importe de la consulta se incrementa en un 20% si se trata de una consulta de tipo urgente.

Tabla 019. Tabla de Hechos; TH_VENTAS

Nombre	TH_VENTAS
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Fuente de Información	El documento Excel de la empresa y la hoja de cálculo nombrada como "Ventas".
Granularidad	ID_MESANO
Relaciones	DIM_TIEMPO, DIM_EMPLEADO e DIM_PRODUCTO
Indicadores	Unidades Vendidas y Precio Venta
Campos	ID_MESANO, ID_EMPLEADO e ID_PRODUCTO, Unidades, Precio_venta
Descripción	Esta tabla recoge toda la información concerniente a las ventas de productos realizadas en la clínica durante los 2 últimos años.
Comentarios	El importe de la venta dependerá de las unidades vendidas y del precio del producto.

4.4- Requisitos de Usuario

Los requisitos de usuario son declaraciones en lenguaje natural y en diagramas de los servicios que se espera que el sistema provea y de las restricciones bajo las cuales debe operar.

Describen los requerimientos funcionales y no funcionales de tal forma que sean comprensibles por los usuarios del sistema que no posean un conocimiento técnico detallado. Únicamente especifican el comportamiento externo del sistema y evitan, tanto como sea posible, las características de diseño del sistema. Deben redactarse utilizando el lenguaje natural, representaciones y diagramas intuitivos sencillos.

En nuestro sistema tenemos dos partes a desarrollar por dos ingenieros distintos, esto va a causar que existan dos tipos diferenciados de requerimientos dependiendo del usuario. Por un lado tenemos la creación del Data Warehouse que va a ser realizado por un desarrollador para que otro desarrollador. Y por otro lado, este último desarrollador va a realizar los informes finales para el cliente.

Teniendo esto en cuenta, tendremos dos tipos de casos de uso y dos tipos de requisitos funcionales, dependiendo del producto del que estemos hablando y de su usuario final.

4.4.1- Casos de Uso

Un caso de uso es una descripción de los pasos o las actividades que deberán realizarse para llevar a cabo algún proceso. Los personajes o entidades que participarán en un caso de uso se denominan actores.

En el contexto de ingeniería del software, un caso de uso es una secuencia de interacciones que se desarrollarán entre un sistema y sus actores en respuesta a un evento que inicia un actor principal sobre el propio sistema.

Al haber dos fases diferentes en el sistema y trabajar en ella distintos desarrolladores, tenemos que diferenciar los casos de uso que requiere el desarrollador para el cual se realiza la primera parte y los casos de uso que requiere el usuario final, que corresponde con el cliente. Por tanto, diferenciaremos entre casos de uso de ETLs y casos de uso de Informes.

- **Casos de Uso de ETLs**

En este caso, la finalización del proceso dará como resultado una base de datos denominada Data Warehouse que posteriormente será utilizada, por un compañero desarrollador, para la creación de informes que el usuario cliente (usuario final) ha requerido del sistema.

Como lo que se va a crear es una base de datos, los casos de uso se reducen a los siguientes:

- **CU(E)-01:** Autenticar Usuario.
- **CU(E)-02:** Consultar Base de Datos.

- **Casos de Uso de Informes**

Para continuar enumeraremos los distintos casos de uso que se han encontrado en la parte final del sistema, estos hacen referencia al usuario cliente:

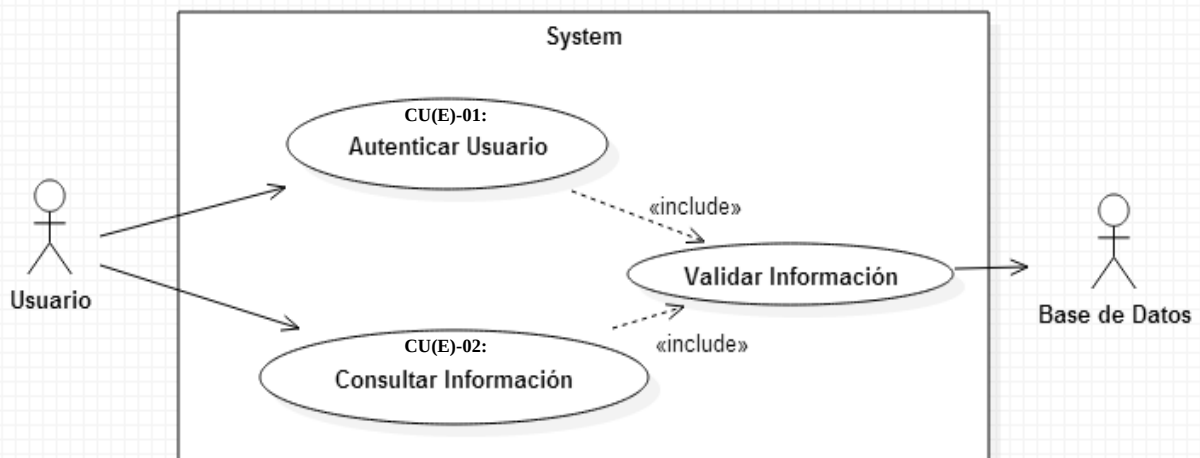
- **CU(I)-01:** Autenticar Usuario
- **CU(I)-02:** Visualizar Informe
- **CU(I)-03:** Filtrar Información
- **CU(I)-04:** Seleccionar Sección de Informe
- **CU(I)-05:** Cerrar Informe

El caso de uso 01 de Autenticar Usuario es común a ETLs e Informes. Esto es, las restricciones de usuario y creación de tipos de usuarios del sistema se crean en primer lugar en la base de datos y se recogen en segundo lugar para los informes, en ambos lugares, los usuarios son los mismos y las autenticaciones (nombre de usuario y contraseña) son iguales.

4.4.2- Diagramas de Casos de Uso

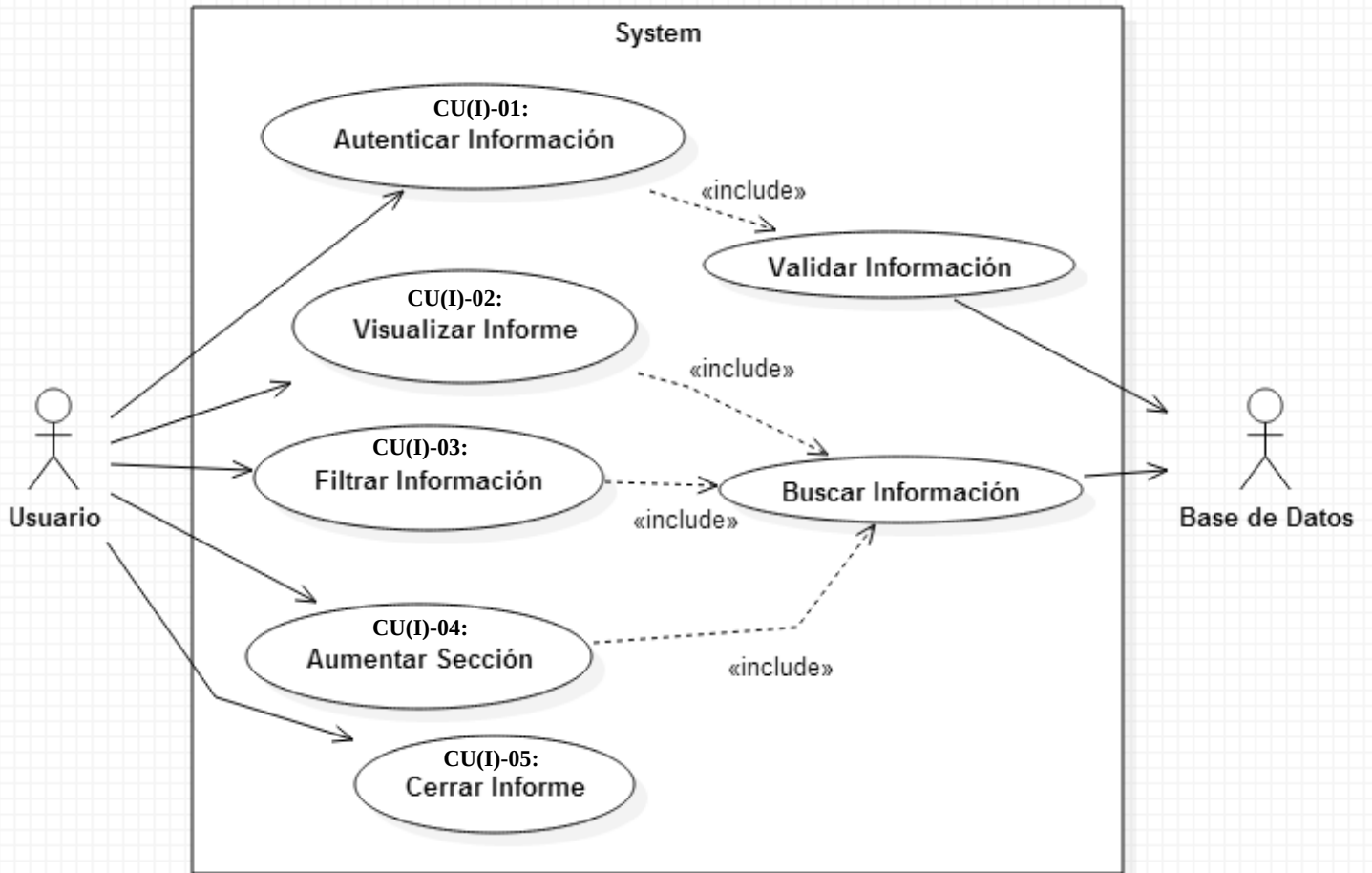
- **Diagrama de Casos de Uso de ETLs**

Ilustración 008. Diagrama de Casos de Uso (ETLs)



- **Diagrama de Casos de Uso de Informes**

Ilustración 009. Diagrama de Casos de Uso (Informes)



4.4.3- Especificación de Casos de Uso

Los diagramas de casos de uso sirven para especificar la comunicación y el comportamiento de un sistema mediante su interacción con los usuarios y/u otros sistemas. O lo que es igual, un diagrama que muestra la relación entre los actores y los casos de uso en un sistema. Una relación es una conexión entre los elementos del modelo, por ejemplo la especialización y la generalización son relaciones. Los diagramas de casos de uso se utilizan para ilustrar los requerimientos del sistema al mostrar cómo reacciona a eventos que se producen en su ámbito o en él mismo.

- **Especificación de Casos de Uso de ETLs**

Tabla 020. Caso de Uso (ETLs) 01; Autenticar Usuario

CU(E)-01	Autenticar Usuario	
Versión	1.0 (25-Junio-2016)	
Autores	Miguel Ángel López y Leticia Aceves	
Actor principal	Usuario o Administrador	
Actor secundario	-	
Descripción	El sistema deberá permitir la entrada de un usuario al sistema identificándose con nombre de usuario y contraseña para poder acceder a los informes.	
Precondición	PRE-1: El usuario debe estar registrado en el sistema previamente PRE-2: Base de datos disponible	
Secuencia normal	Paso	Acción
	p1	El usuario entra en la página de inicio de la aplicación
	p2	El usuario hace login en el sistema con sus datos de acceso (nombre de usuario y contraseña)
	p3	El sistema comprueba que el usuario está registrado.
	p4	El sistema abre la sesión del usuario.
Secuencia Alternativa	Paso	Acción
	p3	El sistema volverá al paso 2
Postcondiciones	POS-1: Abrir sesión de usuario	
Excepciones	Paso	Acción
	-	-
Frecuencia	5 /día	
Prioridad	Alta	
Dependencias	<ul style="list-style-type: none"> - RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05, RF(I)-06 - RNF-01, RNF-08, RNF-09, RNF-10 - OBJ-06 - C3 	
Supuestos	Suponemos que los usuarios están previamente creados en el sistema.	
Comentarios	La base de datos final, deberá permitir autenticarse a un usuario antes de poder utilizarla. Ya sea usuario administrador(todos los permisos) o cliente(solo permisos de lectura)	

Ilustración 010. Caso de Uso (ETLs) 01; Autenticar Usuario

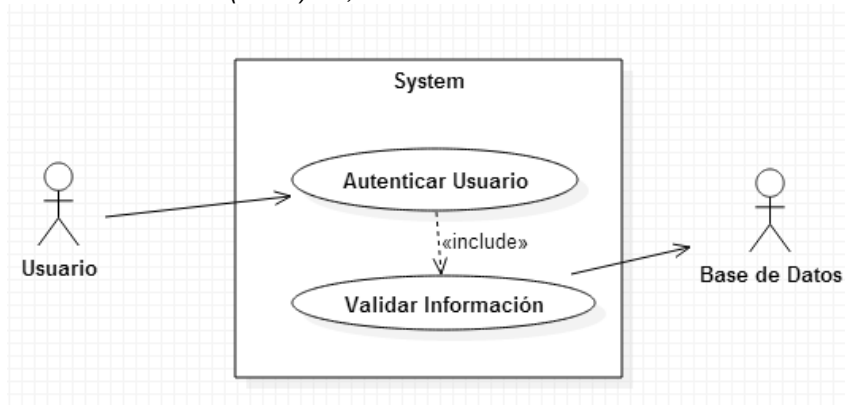
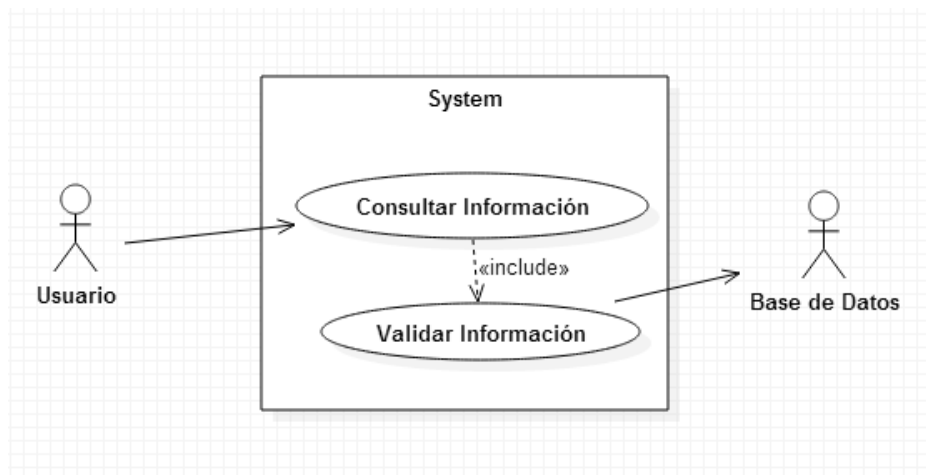


Tabla 021. Caso de Uso (ETLs) 02; Consultar Base de Datos

CU(E)-02	Consultar Base de Datos	
Versión	1.0 (25-Junio-2016)	
Autores	Miguel Ángel López y Leticia Aceves	
Actor principal	Usuario o Administrador	
Actor secundario	-	
Descripción	El sistema deberá permitir la consulta de información en la base de datos mediante consultas SQL	
Precondición	PRE-1: El usuario debe estar registrado en el sistema previamente PRE-2: El usuario debe autenticarse en el sistema PRE-3: Base de datos disponible	
Secuencia normal	Paso	Acción
	p1	El usuario accede a la base de datos
	p2	El sistema abre la información concerniente a dicha base de datos
	p3	El usuario realiza la consulta que necesita
	p4	El sistema accede a la información correspondiente a dicha consulta.
Secuencia Alternativa	Paso	Acción
	p3	Si la consulta está mal planteada: el sistema volverá al paso 2
Postcondiciones	-	
Excepciones	Paso	Acción
	-	-
Frecuencia	5 /día	
Prioridad	Alta	
Dependencias	<ul style="list-style-type: none"> - RF(E)-01, RF(E)-02, RF(E)-03, RF(E)-05, RF(E)-09, RF(E)-10, RF(I)-08, RF(I)-09, RF(I)-11, RF(I)-15 - RNF-03, RNF-04, RNF-07, RNF-12, RNF-14, RNF-15 - RN-04 	
Supuestos	Suponemos que la base de datos está creada previamente	
Comentarios	Tanto el administrador como el usuario cliente pueden acceder a la base de datos para realizar consultas. Los informes también realizan consultas a la base de datos.	

Ilustración 011. Caso de Uso (ETLs) 02; Consultar Base de Datos



• **Especificación de Casos de Uso de Informes**

Tabla 022. Caso de Uso (Informes) 01; Autenticar Usuario

CU(I)-01	Autenticar usuario	
Versión	1.0 (25-Junio-2016)	
Autores	Miguel Ángel López y Leticia Aceves	
Actor principal	Usuario o Administrador	
Actor secundario	-	
Descripción	El sistema deberá permitir la entrada de un usuario al sistema identificándose con nombre de usuario y contraseña para poder acceder a los informes.	
Precondición	PRE-1: El usuario debe estar registrado en el sistema previamente PRE-2: Base de datos disponible	
Secuencia normal	Paso	Acción
	p1	El usuario entra en la página de inicio de la aplicación
	p2	El usuario hace login en el sistema con sus datos de acceso (nombre de usuario y contraseña)
	p3	El sistema comprueba que el usuario está registrado.
	p4	El sistema abre la sesión del usuario.
Secuencia Alternativa	Paso	Acción
	p3	El sistema volverá al paso 2
Postcondiciones	POS-1: Abrir sesión de usuario	
Excepciones	Paso	Acción
	-	-
Frecuencia	5 /día	
Prioridad	Alta	

Dependencias	<ul style="list-style-type: none"> - RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05, RF(I)-06 - RNF-01, RNF-05, RNF-08, RNF-09, RNF-10 - OBJ-06 - C3
Supuestos	-
Comentarios	No todos los usuarios tienen acceso a la aplicación. Tan sólo el usuario cliente (con permisos de lectura) y el administrador del sistema (con todos los permisos) tienen acceso a los informes y su contenido.

Ilustración 012. Caso de Uso (Informes) 01; Autenticar Usuario

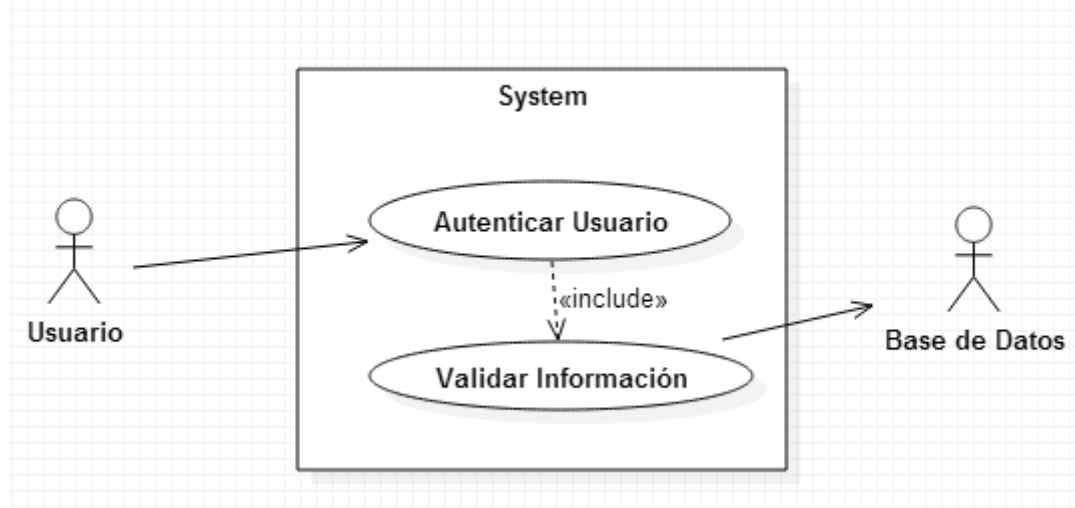


Tabla 023. Caso de Uso (Informes) 02; Visualizar Informe

CU(I)-02	Visualizar Informe	
Versión	1.0 (25-Junio-2016)	
Autores	Miguel Ángel López y Leticia Aceves	
Actor principal	Usuario o Administrador	
Actor secundario	-	
Descripción	El sistema deberá permitir la visualización de la información relativa a cada informe.	
Precondición	PRE-1: El usuario está registrado PRE-2: Abrir la sesión de usuario PRE-3: Base de datos disponible	
Secuencia normal	Paso	Acción
	p1	El usuario entra en el menú de informes de la aplicación
	p2	El usuario hace doble click en el informe que quiere visualizar.
	p3	El sistema abre el informe seleccionado.
Secuencia Alternativa	Paso	Acción
	-	-
Postcondiciones	POS-1: Informe abierto	
Excepciones	Paso	Acción
	-	-

	-	-
Frecuencia	10 /día	
Prioridad	Alta	
Dependencias	<ul style="list-style-type: none"> - RF(E)-02, RF(E)-05, RF(I)-08, RF(I)-09, RF(I)-10, RF(I)-12, RF(I)-13, RF(I)-14, RF(I)-17 - RNF-03, RNF-04, RNF-14, RNF-15 - OBJ-01, OBJ-02, OBJ-03, OBJ-04, OBJ-05 - RN-02, RN-03, RN-04 - C1, C2, C4, C5, C6, C7, C8 	
Supuestos	-	
Comentarios	Una vez que entras en el sistema, ya seas administrador o cliente, se accederá a un menú donde se encuentran todos los informes creados.	

Ilustración 013. Caso de Uso (Informes) 02; Visualizar Informe

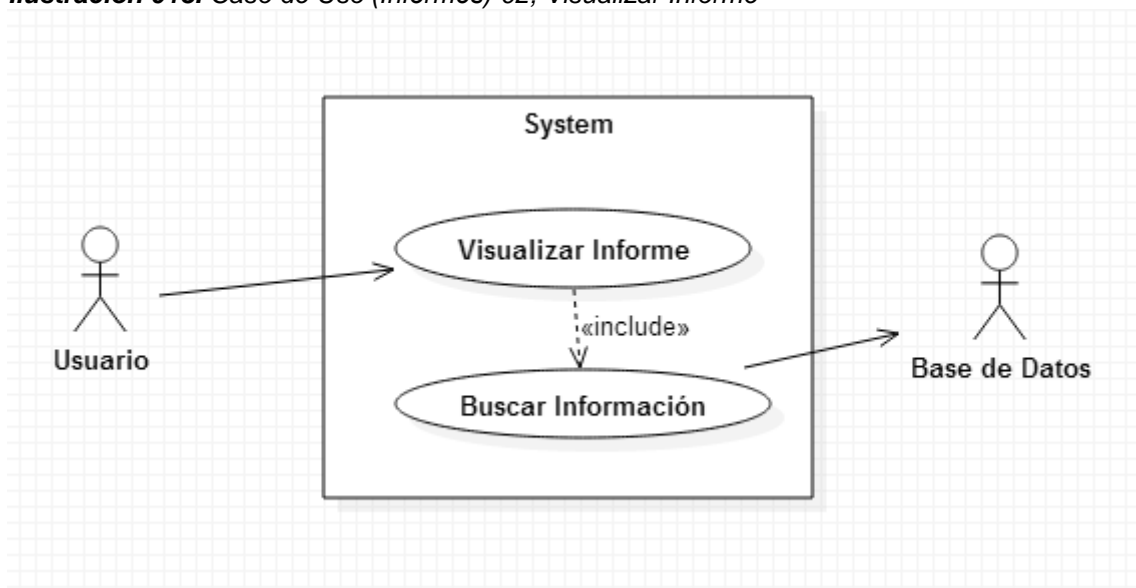


Tabla 024. Caso de Uso (Informes) 03; Filtrar Información

CU(I)-03	Filtrar Información	
Versión	1.0 (25-Junio-2016)	
Autores	Miguel Ángel López y Leticia Aceves	
Actor principal	Usuario o Administrador	
Actor secundario	-	
Descripción	El sistema deberá permitir realizar filtros en la información de un informe y posteriormente la visualización de los datos con estos filtros.	
Precondición	PRE-1: El usuario está registrado PRE-2: Abrir la sesión de usuario PRE-3: Base de datos disponible PRE-4: Visualizar Informe	
Secuencia normal	Paso	Acción
	p1	El usuario escoge, entre los filtros de la parte superior del informe, los datos que necesita.
	p2	El sistema realiza una nueva búsqueda de información conforme a los filtros aplicados.

	p3	El sistema muestra el informe con los nuevos datos.
Secuencia Alternativa	Paso	Acción
	-	-
Postcondiciones	POS-1: Filtros aplicados	
Excepciones	Paso	Acción
	p3	Si no existen datos para los filtros seleccionados, el informe aparecerá vacío.
Frecuencia	10 /día	
Prioridad	Media	
Dependencias	<ul style="list-style-type: none"> - RF(E)-05, RF(I)-09, RF(I)-10, RF(I)-11, RF(I)-12, RF(I)-13, RF(I)-14, RF(I)-15 - RNF-03, RNF-04, RNF-07, RNF-14, RNF-15 - OBJ-01, OBJ-02, OBJ-03, OBJ-04, OBJ-05 - RN-02, RN-03, RN-04 - C1, C2, C4, C5, C6, C7, C8 	
Supuestos	Suponemos que el filtro contiene información, si no es así y no existe información para los filtros aplicados por el usuario, el informe aparecerá vacío.	
Comentarios	En la parte superior de cada informe se encontrarán los filtros necesarios para poder visualizar la información que en un momento dado se necesite.	

Ilustración 014. Caso de Uso (Informes) 03; Filtrar Información

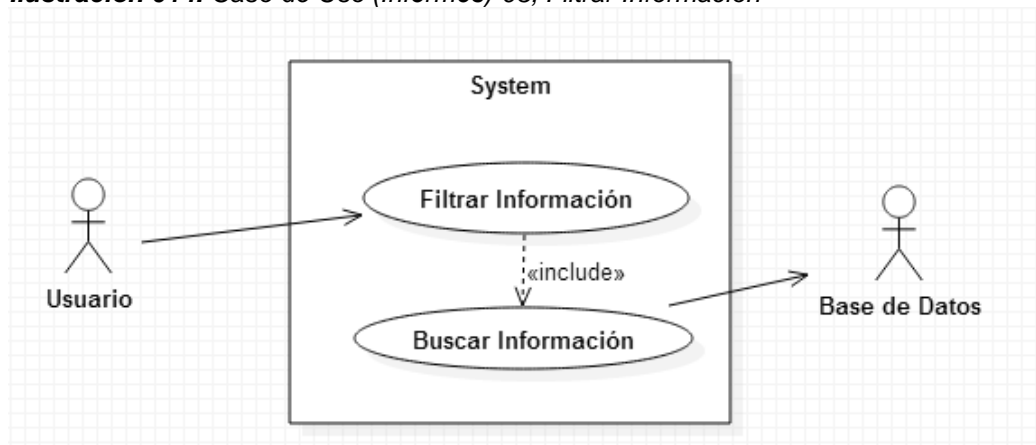


Tabla 025. Caso de Uso (Informes) 04; Aumentar Sección del Informe

CU(I)-04	Seleccionar Sección del Informe
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Actor principal	Usuario o Administrador
Actor secundario	-
Descripción	El sistema deberá permitir visualizar, de manera más detallada, una determinada sección dentro de un informe.
Precondición	PRE-1: El usuario está registrado PRE-2: Abrir la sesión de usuario PRE-3: Base de datos disponible PRE-4: Visualizar Informe

Secuencia normal	Paso	Acción
	p1	El usuario escoge una determinada sección del informe.
	p2	El usuario pincha en la parte superior derecha de dicha sección dentro del informe.
	p3	El sistema realiza una nueva búsqueda de información conforme a la sección escogida.
	p4	El sistema muestra a pantalla completa la sección deseada.
Secuencia Alternativa	Paso	Acción
	-	-
Postcondiciones	POS-1: Visualización a pantalla completa de sección escogida.	
Excepciones	Paso	Acción
	-	-
Frecuencia	10 /día	
Prioridad	Baja	
Dependencias	<ul style="list-style-type: none"> - RF(I)-09, RF(I)-10, RF(I)-11, RF(I)-12, RF(I)-13, RF(I)-14, RF(I)-15, RF(I)-16, RF(I)-17 - RNF-03, RNF-07, RNF-14, RNF-15 - OBJ-01, OBJ-02, OBJ-03, OBJ-04, OBJ-05 - RN-02, RN-03, RN-04 - C1, C4, C5, C6, C7, C8 	
Supuestos	-	
Comentarios	Todos los informes estarán divididos en diferentes secciones. En un mismo informe se encontrarán, al menos, gráficos e información numérica. Cualquiera de estas secciones individuales se deberán poder ampliar para visualizar plenamente su contenido.	

Ilustración 015. Caso de Uso (Informes) 04; Aumentar Sección del Informe

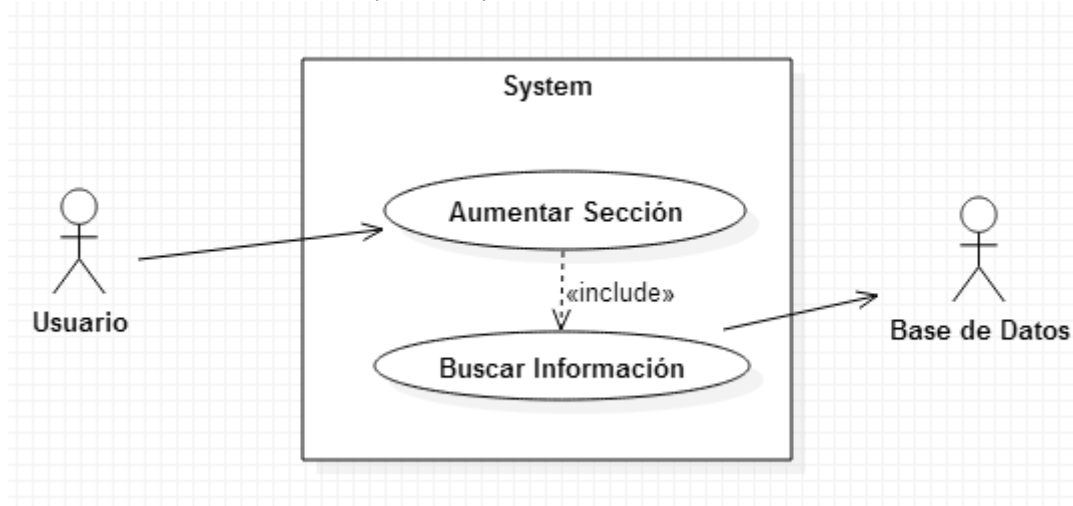
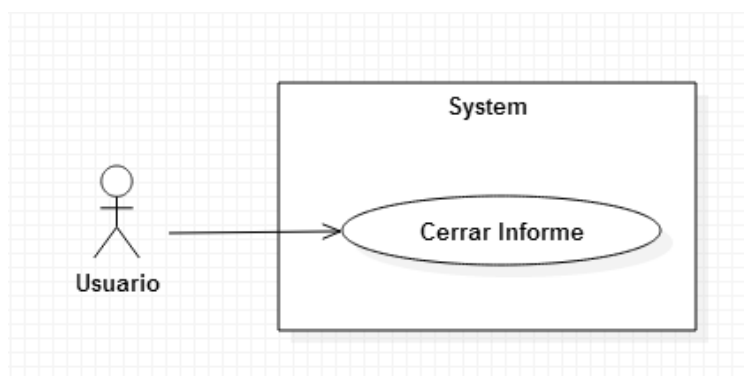


Tabla 026. Caso de Uso (Informes) 05; Cerrar Informe

CU(I)-05	Cerrar Informe	
Versión	1.0 (25-Junio-2016)	
Autores	Miguel Ángel López y Leticia Aceves	
Actor principal	Usuario o Administrador	
Actor secundario	-	
Descripción	El sistema deberá permitir cerrar un informe.	
Precondición	PRE-1: El usuario está registrado PRE-2: Abrir la sesión de usuario PRE-3: Base de datos disponible PRE-4: Visualizar Informe	
Secuencia normal	Paso	Acción
	p1	El usuario pincha en la parte superior derecha del informe
	p2	El sistema cierra el informe
	p3	El sistema muestra el menú de informes.
Secuencia Alternativa	Paso	Acción
	-	-
Postcondiciones	POS-1: Visualización del menú de informes.	
Excepciones	Paso	Acción
	-	-
Frecuencia	10 /día	
Prioridad	Baja	
Dependencias	<ul style="list-style-type: none"> - RF(I)-08, RF(I)-16, RF(I)-17 - RNF-03 - OBJ-05 - C1, C5 	
Supuestos	Suponemos que tenemos un informe abierto	
Comentarios	Todos los informes, al igual que abrir, se pueden cerrar. Cuando se cierra un informe, el sistema vuelve al menú de informes.	

Ilustración 016. Caso de Uso (Informes) 05; Cerrar Informe



4.5- Requisitos Funcionales

Los requisitos funcionales describen qué tendrán que implementar los programadores para satisfacer los requisitos de usuario. Especifican el comportamiento del producto software bajo unas condiciones especificadas.

En nuestro proyecto, existen dos clases de requisitos funcionales, ya que tenemos dos partes diferenciadas en el proyecto. La primera parte corresponde a la creación de la Base de Datos de donde se exportarán los datos a los informes finales, esta parte será desarrollada por programadores diferentes a los encargados de los informes. Para este proceso de carga (realización de ETLs), se tienen los siguientes requisitos funcionales:

- **RF(E)-01:** El sistema tendrá que ajustarse al modelo establecido.
- **RF(E)-02:** El sistema asegurará la integridad de los datos en todo momento, siendo los datos iniciales los mismos que de salida.
- **RF(E)-03:** El sistema deberá manejar los datos referentes a casos sin información, estos son, los “No Informados”.
- **RF(E)-04:** El sistema deberá manejar los errores o excepciones que puedan acontecer en el proceso de carga.
- **RF(E)-05:** El sistema deberá contener una dimensión temporal planteada de forma mensual.
- **RF(E)-06:** El sistema deberá contemplar dos tipos de usuarios: administrador y cliente
- **RF(E)-07:** El usuario administrador debe tener todos los permisos, para poder manejar adecuadamente la base de datos.
- **RF(E)-08:** El usuario cliente tan solo tendrá permisos de lectura sobre la base de datos.
- **RF(E)-09:** Las cuantías de tipo numérico deberán tener al menos dos decimales.
- **RF(E)-10:** El sistema deberá mantener en todo momento, el tipo de formato de los datos de entrada y salida.

Los anteriores requisitos son adecuados a un sistema cuyo uso va a estar ligado a la posterior creación de informes por un desarrollador distinto. Aunque estos requisitos de alguna forma están relacionados con el usuario final, su principal objetivo es facilitar la siguiente fase de creación de informes a otros desarrolladores.

A continuación se definen los requisitos funcionales que definen el comportamiento y los servicios que debe proveer el sistema en su etapa final. Estos requisitos son los que el usuario cliente necesita:

- **RF(I)-01:** El sistema mostrará la página de identificación de usuario.
- **RF(I)-02:** El sistema deberá permitir la entrada de un usuario al sistema identificándose con nombre de usuario y contraseña.
- **RF(I)-03:** El sistema deberá validar los datos de inicio de sesión del usuario
- **RF(I)-04:** El sistema podrá abrir sesión de usuario, con sus privilegios de usuario.
- **RF(I)-05:** El sistema sólo permitirá al usuario cliente leer informes, pues este actor solo tiene permisos de lectura.
- **RF(I)-06:** El sistema permitirá al administrador realizar cambios en los informes. El administrador tiene todos los permisos sobre el programa: puede modificar, crear y borrar informes.
- **RF(I)-07:** El sistema mostrará un menú con los diferentes informes creados.
- **RF(I)-08:** El sistema abrirá el informe que seleccione el usuario.
- **RF(I)-09:** El sistema podrá acceder a los datos correspondientes de la base de datos.
- **RF(I)-10:** El sistema mostrará los datos adecuados a cada informe.
- **RF(I)-11:** El sistema será capaz de mostrar la información según los filtros que el

usuario le proporcione.

- **RF(I)-12:** El sistema mostrará los datos de los informes por orden cronológico.
- **RF(I)-13:** El sistema mostrará los datos de los informes de forma clara, concisa y con totales en todos los casos concernientes a cuantías numéricas.
- **RF(I)-14:** Las cantidades de tipo numérico deberán tener al menos dos decimales.
- **RF(I)-15:** Si los datos introducidos en los filtros no contienen información en la base de datos, el informe aparecerá vacío.
- **RF(I)-16:** El sistema permitirá al usuario ampliar una sección de un informe a pantalla completa y volver a minimizar esta sección.
- **RF(I)-17:** El sistema permitirá cerrar un informe y volver al menú de informes.

4.5.1- Especificación Requisitos Funcionales

- **Especificación Requisitos Funcionales de ETLs**

Tabla 027. Requisito Funcional (ETL) 01; Ajustarse al Modelo

RF(E)-01	El sistema tendrá que ajustarse al modelo establecido
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	- CU(E)-02 - RF(I)-14 - RN-01
Descripción	La carga de datos en la base de datos del Data Warehouse debe finalizar con todas las dimensiones y tablas de hechos conteniendo la información correspondiente, de tal manera, que se ajuste por completo al modelo de datos realizado en la fase de modelado.
Importancia	Alta
Prioridad	Alta
Comentarios	El desarrollador de ETLs debe concordar la carga de datos del modelo operacional que se le ofrece al modelo Data Warehouse que finalmente se le pide. En todo momento se deberá respetar la integridad de los datos.

Tabla 028. Requisito Funcional (ETL) 02; Integridad de los Datos

RF(E)-02	El sistema asegurará la integridad de los datos
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-02 -RF(E)-03, RF(E)-09, RF(E)-10,RF(I)-09, RF(I)-10, RF(I)-11, RF(I)-14, RF(I)-15 -RNF-12, RNF-13, RNF-14, RNF-15 -RN-01
Descripción	El sistema asegurará la integridad de los datos en todo momento, siendo los datos iniciales los mismos que de salida.
Importancia	Alta
Prioridad	Alta
Comentarios	La carga de datos establecida por el desarrollador de ETLs debe asegurar la integridad de la información. Realizando el proceso de carga convenientemente.

Tabla 029. Requisito Funcional (ETL) 03; "No Informados"

RF(E)-03	Tratamiento de los "No Informados"
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02 -RF(E)-02, , RF(I)-15
Descripción	El sistema deberá manejar los datos referentes a casos sin información, estos son, los "No Informados".
Importancia	Alta
Prioridad	Normal
Comentarios	En aquellos casos en que algún dato esté incompleto o falte, se asumirá esta falta de información de tal manera que en el Data Warehouse aparecerá la ausencia como un dato "No Informado"

Tabla 030. Requisito Funcional (ETL) 04; Manejo de Errores

RF(E)-04	El sistema deberá manejar los errores
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-RNF-12, RNF-13
Descripción	El sistema deberá manejar los errores o excepciones que puedan acontecer en el proceso de carga.
Importancia	Normal
Prioridad	Normal
Comentarios	Si el flujo de carga es correcto, el manejo de errores se hace innecesario, aun así, es conveniente realizar una administración de excepciones para, en su caso, poder detectar incidencias en el proceso de ETL.

Tabla 031. Requisito Funcional (ETL) 05; Dimensión Temporal

RF(E)-05	El sistema tendrá una dimensión temporal mensual
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-02, CU(I)-03 -RF(I)-10, RF(I)-11, RF(I)-12,RF(I)-15 -OBJ-03, OBJ-04 -RN-03 -C7, C8
Descripción	El sistema deberá contener una dimensión temporal planteada de forma mensual.
Importancia	Alta
Prioridad	Alta
Comentarios	La dimensión temporal se hará mensual por petición del cliente, puesto que de forma diaria, a la hora de visualizar los datos en los informes, se harían muy extensos

Tabla 032. Requisito Funcional (ETL) 06; Usuarios

RF(E)-06	El sistema tendrá dos tipos de Usuarios
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-04,RF(I)-05,RF(I)-06 -RNF-01,RNF-02, RNF-05, RNF-08, RNF-09, RNF-10 -OBJ-06 -C3, C4
Descripción	El sistema deberá contemplar dos tipos de usuarios: administrador y cliente
Importancia	Alta
Prioridad	Alta
Comentarios	-

Tabla 033. Requisito Funcional (ETL) 07; Usuario Administrador

RF(E)-07	Usuario Administrador con todos los Permisos
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(I)-01,RF(I)-02, RF(I)-03, RF(I)-04,RF(I)-06 -RNF-01,RNF-02, RNF-05,RNF-08, RNF-09, RNF-10 -OBJ-06 -C3, C4
Descripción	El usuario administrador debe tener todos los permisos, para poder manejar adecuadamente la base de datos.
Importancia	Alta
Prioridad	Alta
Comentarios	El Administrador tendrá todos los permisos tanto en la base de datos como en los informes, esto es así por si hubiera cambios en los requisitos de usuario que conlleven modificaciones en el sistema, el administrador podrá realizar estos cambios sin problemas.

Tabla 034. Requisito Funcional (ETL) 08; Usuario Cliente

RF(E)-08	Usuario Cliente con Permisos de Lectura
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05 -RNF-01, RNF-02, RNF-05, RNF-08, RNF-09, RNF-10 -OBJ-06 -C3, C4
Descripción	El usuario cliente tan solo tendrá permisos de lectura sobre la base de datos.
Importancia	Alta
Prioridad	Alta
Comentarios	El usuario cliente sólo tendrá permisos de lectura sobre la base de datos y sobre los informes posteriores puesto que, no es necesario ningún tipo de permiso más. El cliente tan solo visualiza los informes, no realiza modificación alguna en ellos.

Tabla 035. Requisito Funcional (ETL) 09; Cuantías Numéricas

RF(E)-09	Cuantías Numéricas con dos Decimales
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02 -RF(E)-02, RF(E)-10, RF(I)-09, RF(I)-10, , RF(I)-13, RF(I)-14 -OBJ-01
Descripción	Las cuantías de tipo numérico deberán tener al menos dos decimales.
Importancia	Alta
Prioridad	Normal
Comentarios	Esto debe ser así ya que si truncamos las cuantías numéricas y no consideramos los decimales, puede perderse mucha información de precios y cantidades.

Tabla 036. Requisito Funcional (ETL) 10; Formatos

RF(E)-10	Mismo Formato de Entrada y Salida
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02 -RF(E)-02, RF(E)-09, RF(I)-08, RF(I)-09, RF(I)-10, RF(I)-12, , RF(I)-13, RF(I)-14 -OBJ-01
Descripción	El sistema deberá mantener en todo momento, el tipo de formato de los datos de entrada y salida.
Importancia	Normal
Prioridad	Normal
Comentarios	Para evitar errores en la inserción y tratamiento de datos en la base de datos.

• Especificación Requisitos Funcionales de Informes

Tabla 037. Requisito Funcional (Informes) 01; Identificación de Usuario

RF(I)-01	Identificación de Usuario
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05, RF(I)-06 -RNF-01, RNF-02, RNF-06, RNF-08, RNF-10 -OBJ-06 -C3, C4
Descripción	El sistema mostrará la página de identificación de usuario.
Importancia	Alta
Prioridad	Alta
Comentarios	Al iniciar la aplicación, lo primero que se debe hacer es autenticarse. El sistema deberá mostrar una página de inicio para que el usuario ingrese su contraseña.

Tabla 038. Requisito Funcional (Informes) 02; Autenticación

RF(I)-02	Autenticación
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-03, RF(I)-04, RF(I)-05, RF(I)-06 -RNF-01, RNF-02, RNF-08, RNF-10 -OBJ-06 -C3, C4
Descripción	El sistema deberá permitir la entrada de un usuario al sistema identificándose con nombre de usuario y contraseña.
Importancia	Alta
Prioridad	Alta
Comentarios	En la página de inicio al sistema, el usuario deberá ingresar sus credenciales para poder acceder a los informes con los permisos adecuados.

Tabla 039. Requisito Funcional (Informes) 03; Validación de Autenticación

RF(I)-03	Validación de datos de Autenticación de Usuario
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-04, RF(I)-05, RF(I)-06 -RNF-01, RNF-02, RNF-05, RNF-08, RNF-10 -OBJ-06 -C3, C4
Descripción	El sistema deberá validar los datos de inicio de sesión del usuario
Importancia	Alta
Prioridad	Alta
Comentarios	El sistema comprobará los credenciales de cada usuario, permitiendo entrar en el sistema con los privilegios de usuario que corresponden

Tabla 040. Requisito Funcional (Informes) 04; Abrir Sesión

RF(I)-04	Abrir Sesión
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-05, RF(I)-06, RF(I)-07 - RNF-01, RNF-02, RNF-05, RNF-08, RNF-10 -OBJ-06 -C2, C4, C5
Descripción	El sistema podrá abrir sesión de usuario, con sus privilegios de usuario.
Importancia	Alta
Prioridad	Alta
Comentarios	El sistema abrirá la sesión al usuario que corresponda con los privilegios que se le hayan otorgado.

Tabla 041. Requisito Funcional (Informes) 05; Permisos Usuario Cliente

RF(I)-05	Permisos Usuario Cliente
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05 -RNF-01, RNF-02, RNF-05, RNF-08, RNF-09, RNF-10 -OBJ-06 -C3, C4
Descripción	El sistema sólo permitirá al usuario leer informes, pues este actor solo tiene permisos de lectura.
Importancia	Alta
Prioridad	Alta
Comentarios	-

Tabla 042. Requisito Funcional (Informes) 06; Permisos Usuario Administrador

RF(I)-06	Permisos Usuario Administrador
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-06 -RNF-01, RNF-02, RNF-05, RNF-08, RNF-09, RNF-10 -OBJ-06 -C3, C4
Descripción	El sistema permitirá al administrador realizar cambios en los informes. El administrador tiene todos los permisos sobre el programa: puede modificar, crear y borrar informes.
Importancia	Alta
Prioridad	Alta
Comentarios	-

Tabla 043. Requisito Funcional (Informes) 07; Menú

RF(I)-07	Menú
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-RF(I)-04 -RNF-02, RNF-06, RNF-14, RNF-15 -OBJ-05, OBJ-06 -RN-04 -C1, C2, C5
Descripción	El sistema mostrará un menú con los diferentes informes creados.
Importancia	Alta
Prioridad	Alta
Comentarios	El sistema, nada más abrir sesión de usuario, mostrará un menú con los diferentes informes de los que se compone., pudiendo así seleccionar el que se quiera.

Tabla 044. Requisito Funcional (Informes) 08; Abrir Informe

RF(I)-08	Abrir Informe
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-02, CU(I)-05 -RF(E)-10, RF(I)-09, RF(I)-10, RF(I)-11, RF(I)-12, RF(I)-13, , RF(I)-17 -RNF-03, RNF-04, RNF-06, RNF-14, RNF-15 -OBJ-01, OBJ-02, OBJ-3, OBJ-04, OBJ-05, OBJ-06 -RN-02, RN-03, RN-04 -C1, C4, C5, C6, C7, C8
Descripción	El sistema abrirá el informe que seleccione el usuario.
Importancia	Alta
Prioridad	Alta
Comentarios	Al hacer doble click sobre un informe en el menú de informes, este se abrirá, permitiendo la visualización de los datos al usuario.

Tabla 045. Requisito Funcional (Informes) 09; Acceso a Datos

RF(I)-09	Acceso a Datos
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-02, CU(I)-03, CU(I)-04 -RF(E)-02, RF(E)-09, RF(E)-10, RF(I)-08, RF(I)-10, RF(I)-11, , RF(I)-15 -RNF-03, RNF-04, RNF-05, RNF-06, RNF-07, RNF-14
Descripción	El sistema podrá acceder a los datos correspondientes de la base de datos.
Importancia	Alta
Prioridad	Alta
Comentarios	Para que el sistema pueda abrir adecuadamente los informes y mostrar la información de forma apropiada, el sistema deberá tener acceso a estos datos en la base de datos, en este caso, al Data Warehouse.

Tabla 046. Requisito Funcional (Informes) 10; Datos Adecuados

RF(I)-10	Datos Adecuados
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(I)-02, CU(I)-03, CU(I)-04 -RF(E)-02, RF(E)-05, RF(E)-09, RF(E)-10, RF(I)-08, RF(I)-09, RF(I)-12, RF(I)-13, RF(I)-14, , RF(I)-15 - RNF-03, RNF-04,RNF-05, RNF-07 -RN-04
Descripción	El sistema mostrará los datos adecuados a cada informe.
Importancia	Alta
Prioridad	Alta
Comentarios	Los datos que se muestran en cada informe deben ser los ajustados a dicho informe, sin errores de lecturas.

Tabla 047. Requisito Funcional (Informes) 11; Filtros

RF(I)-11	Filtros
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-03, CU(I)-04 -RF(E)-02, RF(E)-05, RF(I)-08, RF(I)-09, RF(I)-12, RF(I)-15 - RNF-05, RNF-06, RNF-07 -OBJ-03 -RN-04 -C1,C6,C7
Descripción	El sistema será capaz de mostrar la información según los filtros que el usuario le proporcione.
Importancia	Alta
Prioridad	Normal
Comentarios	Cuando el usuario filtra la información que compone un informe este debe adecuarse a dicho filtro, mostrando los datos óptimos en cada caso.

Tabla 048. Requisito Funcional (Informes) 12; Orden Cronológico

RF(I)-12	Datos en Orden Cronológico
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(I)-02, CU(I)-03, CU(I)-04 -RF(E)-05, RF(E)-10, RF(I)-08, RF(I)-10, RF(I)-11, RF(I)-13, RF(I)-15 -OBJ-01
Descripción	El sistema mostrará los datos de los informes por orden cronológico.
Importancia	Normal
Prioridad	Normal
Comentarios	El orden cronológico será ascendente

Tabla 049. Requisito Funcional (Informes) 13; Formatos

RF(I)-13	Formato de Datos
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(I)-02, CU(I)-03, CU(I)-04 - RF(E)-09, RF(I)-08, RF(E)-10, RF(I)-10, RF(I)-12, , RF(I)-14, , RF(I)-15 -OBJ-01, OBJ-02 -RN-02 -C1, C6
Descripción	El sistema mostrará los datos de los informes de forma clara, concisa y con totales en todos los casos concernientes a cuantías numéricas.
Importancia	Normal
Prioridad	Normal
Comentarios	Los formatos deben ser los adecuados en cada circunstancia, mostrando totales en los casos necesarios.

Tabla 050. Requisito Funcional (Informes) 14; Cuantías Numéricas

RF(I)-14	Cuantías Numéricas con Decimales
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(I)-02, CU(I)-03, CU(I)-04 -RF(E)-01, RF(E)-02, RF(E)-09, RF(E)-10, RF(I)-10, RF(I)-13 -OBJ-01
Descripción	Las cantidades de tipo numérico deberán tener al menos dos decimales.
Importancia	Alta
Prioridad	Normal
Comentarios	Las cuantías numéricas deben ir con al menos dos decimales para evitar la pérdida de cantidades que pueden llegar a ser cuantiosas.

Tabla 051. Requisito Funcional (Informes) 15; Información no Encontrada

RF(I)-15	Información No Encontrada
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-03, CU(I)-04 -RF(E)-02, RF(E)-03, RF(E)-05, RF(I)-09, RF(I)-10, RF(I)-11, RF(I)-12, RF(I)-13 -RNF-03, RNF-04, RNF-05, RNF-06, RNF-07 -OBJ-01, OBJ-03, OBJ-05 -RN-02, RN-04 -C7
Descripción	Si los datos introducidos en los filtros no contienen información en la base de datos, el informe aparecerá vacío.
Importancia	Normal
Prioridad	Normal
Comentarios	-

Tabla 052. Requisito Funcional (Informes) 16; Ampliar Sección

RF(I)-16	Ampliar Sección
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(I)-04, CU(I)-05
Descripción	El sistema permitirá al usuario ampliar una sección de un informe a pantalla completa y volver a minimizar esta sección.
Importancia	Normal
Prioridad	Normal
Comentarios	Los informes contienen diferentes secciones que el usuario, en un momento dado, quiere ampliar. El sistema ofrece la posibilidad de aumentar las secciones de un informe para ver sus datos de forma más detallada.

Tabla 053. Requisito Funcional (Informes) 17; Cerrar Informe

RF(I)-17	Cerrar Informe
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	- CU(I)-02, CU(I)-04, CU(I)-05 -RF(I)-08 -RNF-03 -RN-04
Descripción	El sistema permitirá cerrar un informe y volver al menú de informes.
Importancia	Normal
Prioridad	Normal
Comentarios	-

4.6- Requisitos No Funcionales.

Como requisito no funcional entenderemos la descripción de una propiedad que debe tener el sistema o una restricción que debe respetar. Si los requisitos funcionales especifican que hace el sistema, los no funcionales indican cómo de bien lo hace. Varios tipos de requisitos no funcionales:

- Los requisitos no funcionales describen características importantes del sistema: disponibilidad, usabilidad, seguridad, rendimiento, etc. (atributos de calidad). Son cuantitativos y verificables.
 - Otros requisitos no funcionales describen el entorno en el que opera el sistema: plataforma, portabilidad, compatibilidad, etc.
 - Requisitos de localización: idioma, moneda, leyes, terminología...
 - Otros describen las interfaces externas entre el sistema y el mundo exterior.
 - Conexión con otros sistemas software, hardware, interfaces de comunicación...
 - Finalmente, también se consideran como no funcionales las restricciones impuestas sobre el diseño o implementación del producto.
- **Requisitos No Funcionales: Usabilidad**

En este punto, se especificarán los requisitos respecto a las características que harán que el software que parecen ser "fácil de usar."

Usabilidad abarca:

- La facilidad de uso
- Facilidad de aprendizaje
- Evitar los errores
- La recuperación
- La eficiencia de las interacciones
- Accesibilidad

Obviamente en el sistema se quiere que todas y cada una de estas características de usabilidad se tengan en cuenta por diseñadores y desarrolladores.

Para que el sistema se utilice adecuadamente se tendrán en cuenta los siguientes factores:

- **RNF-01:** Un usuario podrá autenticarse en el sistema en un tiempo máximo de 30 segundos.
- **RNF-02:** El sistema abrirá el menú de informes en un tiempo máximo de 30 segundos.
- **RNF-03:** Un usuario experimentado realizará una solicitud de visualización de informe en un tiempo promedio de 30 segundos y máximo de 1 minuto, en el 95 % de las veces.
- **RNF-04:** Un usuario nuevo realizará una solicitud de visualización de informe en un tiempo promedio de 2 minutos y máximo de 3 minutos, en el 95 % de las veces.

- **Requisitos No Funcionales: Rendimiento**

Estos requisitos determinan el rendimiento esperado para las diferentes funciones o características del sistema. El rendimiento mide la respuesta del sistema ante las solicitudes y acciones de los usuarios. El rendimiento (atributo externo), está muy ligado a la eficiencia (atributo interno).

- **RNF-05:** El sistema no tardará más de 2 segundos en realizar cualquier tipo de validación de datos.
- **RNF-06:** Las páginas web se mostrarán completamente en un tiempo promedio de 3 segundos, para una conexión de 30Mbps.
- **RNF-07:** El sistema no tardará más de 5 segundos en realizar cualquier tipo de búsqueda.

- **Requisitos No Funcionales: Seguridad**

En este punto, se especifican requisitos relativos a los problemas de seguridad o privacidad que restringen el acceso o uso del producto. Estos podrían referirse a la integridad física, datos o software de seguridad. Requisitos de seguridad a menudo se originan en las reglas de negocio, por lo que se identifican las políticas o regulaciones a las que el producto debe ajustarse de seguridad o privacidad.

- **RNF-08:** En el sistema solo habrá dos tipos de usuarios el administrador y el usuario cliente
- **RNF-09:** El usuario cliente solo tendrá permisos de lectura
- **RNF-10:** El administrador tendrá todos los privilegios en el sistema.

- **Requisitos No Funcionales: Integridad y protección**

Se deben especificar los requisitos que tienen que ver con una posible pérdida, daño o perjuicio que pudiera ocasionar por el uso del producto. Definir las acciones potencialmente peligrosas que deben prevenirse. Identificar cualquier seguridad certificaciones, políticas o regulaciones a las que el producto debe cumplir.

En definitiva, estos requisitos especifican aspectos relacionados con la posible pérdida o daños que pueden producirse con el uso del sistema, así como las salvaguardas y las acciones a tomar en caso de que sucedan situaciones de riesgo.

- **RNF-11:** Garantizar la seguridad física de los ordenadores y sistemas de almacenamiento
- **RNF-12:** Realizar, al menos, un backup de los datos de la base de datos.
- **RNF-13:** Si existiera algún error en los datos de la base de datos, se deberá restaurar los datos del backup.

- **Requisitos No Funcionales: Disponibilidad**

La disponibilidad ofrece una medida del tiempo en el que los servicios del sistema están operativos:

- La disponibilidad requiere de periodos de mantenimiento programados.
- Está relacionada directamente con la fiabilidad y la modificabilidad del sistema.
 - **RNF-14:** El sistema deberá estar disponible las 24 horas del día
 - **RNF-15:** El tiempo de mantenimiento, si lo hubiera, será el menor posible, sin afectar a todas las funcionalidades del sistema

4.6.1- Especificación de Requisitos No Funcionales

Tabla 054. Requisito No Funcional 01; Tiempo de Autenticación

RNF-01	Tiempo de Autenticación
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01,RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05, RF(I)-06 -RNF-08 -C3
Descripción	Un usuario podrá autenticarse en el sistema en un tiempo máximo de 30 segundos.
Importancia	Alto
Prioridad	Alto
Comentarios	Es importante que un usuario no tenga que esperar mucho para que la página se cargue. El tiempo de autenticación de un usuario debe ser el menor posible.

Tabla 055. Requisito No Funcional 02; Tiempo Abrir Menú

RNF-02	Tiempo Abrir Menú
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	- RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-04,RF(I)-05, RF(I)-06, RF(I)-07 -RNF-06, RNF-09, RNF-10 -C1, C2
Descripción	El sistema abrirá el menú de informes en un tiempo máximo de 30 segundos.
Importancia	Alto
Prioridad	Alto
Comentarios	El tiempo para que el sistema abra el menú de informes debe ser el menor posible.

Tabla 056. Requisito No Funcional 03; Tiempo Abrir Informe Usuario Experimentado

RNF-03	Tiempo Abrir Informe Usuario Experimentado
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-02, CU(I)-03, CU(I)-04 -RF(I)-08, RF(I)-09, RF(I)-10, RF(I)-15, RF(I)-17 -C1, C5
Descripción	Un usuario experimentado realizará una solicitud de visualización de informe en un tiempo promedio de 30 segundos y máximo de 1 minuto, en el 95 % de las veces.
Importancia	Alto
Prioridad	Alto
Comentarios	A la hora de abrir los informes, el tiempo de respuesta del sistema deberá ser el menor posible. Los informes tienen que realizar las consultas de forma óptima a la base de datos para que la respuesta de esta, sea temprana.

Tabla 057. Requisito No Funcional 04; Tiempo Abrir Informe Usuario Nuevo

RNF-04	Tiempo Abrir Informe Usuario Nuevo
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-02, CU(I)-03 -RF(I)-08, RF(I)-09, RF(I)-10, RF(I)-15 -C1, C5
Descripción	Un usuario nuevo realizará una solicitud de visualización de informe en un tiempo promedio de 2 minutos y máximo de 3 minutos, en el 95 % de las veces.
Importancia	Alto
Prioridad	Alto
Comentarios	La organización del menú de informes debe ser sencilla, de tal forma que un usuario que nunca ha entrado en el sistema, encuentre el informe que necesita en el menor tiempo posible.

Tabla 058. Requisito No Funcional 05; Tiempo de Validación

RNF-05	Tiempo de Validación
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01, CU(I)-03 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05, RF(I)-06, RF(I)-09, RF(I)-11, RF(I)-15 -RNF-08, RNF-09, RNF-10 -OBJ-03 -RN-04 -C3, C7
Descripción	El sistema no tardará más de 2 segundos en realizar cualquier tipo de validación de datos.
Importancia	Alto
Prioridad	Alto
Comentarios	El tiempo de validación de información en cuanto a filtros o autenticación de un usuario en el sistema, deberá ser el menor posible, realizando la búsqueda en la base de datos con la mayor brevedad viable.

Tabla 059. Requisito No Funcional 06; Tiempo de Apertura de la Página

RNF-06	Tiempo de Apertura de la Página
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-RF(I)-01, RF(I)-07, RF(I)-08, RF(I)-09, RF(I)-11, RF(I)-15 -RNF-02 -OBJ-05 -RN-04
Descripción	Las páginas web se mostrarán completamente en un tiempo promedio de 3 segundos.
Importancia	Alto
Prioridad	Alto
Comentarios	El tiempo de recarga de la página tiene que ser corto.

Tabla 060. Requisito No Funcional 07; Tiempo de Búsqueda

RNF-07	Tiempo de Búsqueda
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02, CU(I)-03, CU(I)-04 -RF(I)-09, RF(I)-10, RF(I)-11, RF(I)-15 -OBJ-03 -RN-04 -C7
Descripción	El sistema no tardará más de 5 segundos en realizar cualquier tipo de búsqueda.
Importancia	Alto
Prioridad	Alto
Comentarios	El tiempo de búsqueda de información en cuanto a filtros o apertura de la página, deberá ser el menor posible, realizando la búsqueda en la base de datos con la mayor brevedad viable.

Tabla 061. Requisito No Funcional 08; Tipos de Usuarios

RNF-08	Tipos de Usuarios
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05, RF(I)-06 -RNF-01, RNF-05, RNF-09, RNF-10 -OBJ-06 -C3, C4
Descripción	En el sistema solo habrá dos tipos de usuarios el administrador y el usuario cliente
Importancia	Alto
Prioridad	Alto
Comentarios	-

Tabla 062. Requisito No Funcional 09; Usuario Cliente

RNF-09	Usuario Cliente
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-05, RF(I)-06 - RNF-02, RNF-05, RNF-08, RNF-10 -OBJ-06 -C3, C4
Descripción	El usuario cliente solo tendrá permisos de lectura
Importancia	Alto
Prioridad	Alto
Comentarios	El Usuario Cliente tan solo tendrá permisos de lectura en el sistema puesto que no necesita ninguno más. El cliente sólo va a visualizar los informes y realizar los filtros oportunos para poder ver la información que crea conveniente en cada caso.

Tabla 063. Requisito No Funcional 10; Usuario Administrador

RNF-10	Usuario Administrador
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-01, CU(I)-01 -RF(E)-06, RF(E)-07, RF(E)-08, RF(I)-01, RF(I)-02, RF(I)-03, RF(I)-04, RF(I)-05, RF(I)-06, RF(I)-10 - RNF-02, RNF-05, RNF-08, RNF-09 -OBJ-06 -C3, C4
Descripción	El administrador tendrá todos los privilegios en el sistema.
Importancia	Alto
Prioridad	Alto
Comentarios	El usuario administrador tendrá todos los privilegios en el sistema por si, en un momento dado, es necesario realizar algún tipo de cambio en el: crear, modificar o borrar algún informe.

Tabla 064. Requisito No Funcional 11; Seguridad Física

RNF-11	Seguridad Física
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-
Descripción	Garantizar la seguridad física de los ordenadores y sistemas de almacenamiento
Importancia	Alto
Prioridad	Normal
Comentarios	-

Tabla 065. Requisito No Funcional 12; Backups

RNF-12	Backups
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-CU(E)-02 -RF(E)-02, RF(E)-04 -RNF-13, RNF-14, RNF-15
Descripción	Realizar, al menos, un backup de los datos de la base de datos.
Importancia	Alto
Prioridad	Normal
Comentarios	Por seguridad, se hace necesario realizar siempre una copia de seguridad de toda la información de la base de datos así como de la metodología utilizada.

Tabla 066. Requisito No Funcional 13; Restaurar Backup

RNF-13	Restaurar Backup
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	-RF(E)-02, RF(E)-04 -RNF-12, RNF-14, RNF-15
Descripción	Si existiera algún error en los datos de la base de datos, se deberá restaurar los datos del backup.
Importancia	Normal
Prioridad	Normal
Comentarios	-

Tabla 067. Requisito No Funcional 14; Disponibilidad

RNF-14	Disponibilidad
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	- CU(E)-02, CU(I)-02, CU(I)-03, CU(I)-04 -RF(E)-02, RF(I)-07, RF(I)-08, RF(I)-09 -RNF-12, RNF-13, RNF-15
Descripción	El sistema deberá estar disponible las 24 horas del día
Importancia	Normal
Prioridad	Alto
Comentarios	Los informes, una vez realizados, estarán dispuestos siempre y en todo momento para el usuario cliente.

Tabla 068. Requisito No Funcional 15; Mantenimiento

RNF-15	Mantenimiento
Versión	1.0 (25-Junio-2016)
Autores	Miguel Ángel López y Leticia Aceves
Dependencias	- CU(E)-02, CU(I)-02, CU(I)-03, CU(I)-04 -RF(E)-02, RF(I)-07, RF(I)-08 -RNF-12, RNF-13, RNF-14
Descripción	El tiempo de mantenimiento, si lo hubiera, será el menor posible, sin afectar a todas las funcionalidades del sistema.
Importancia	Normal
Prioridad	Normal
Comentarios	El mantenimiento o modificación de algún informe se realizará de forma transparente al cliente, mientras este dispone de la anterior versión del sistema.

4.7- Trazabilidad

Esta sección debe contener una matriz *objetivo-requisito*, de forma que para cada objetivo se pueda conocer con qué requisitos está asociado. El formato de la matriz de rastreabilidad es el siguiente:

Tabla 069. Trazabilidad 01; Requisitos Funcionales – Casos de Uso

Requisitos Funcionales	CU(E)-01	CU(E)-02		CU(I)-01	CU(I)-02	CU(I)-03	CU(I)-04	CU(I)-05
RF(E)-01		X						
RF(E)-02		X		X	X			
RF(E)-03		X						
RF(E)-04								
RF(E)-05		X			X	X		
RF(E)-06	X			X				
RF(E)-07	X			X				
RF(E)-08	X			X				
RF(E)-09		X						
RF(E)-10		X						
RF(I)-01	X			X				
RF(I)-02	X			X				
RF(I)-03	X			X				
RF(I)-04	X			X				
RF(I)-05	X			X				
RF(I)-06	X			X				
RF(I)-07								
RF(I)-08		X			X			X
RF(I)-09		X			X	X	X	
RF(I)-10					X	X	X	
RF(I)-11		X				X	X	
RF(I)-12		X			X	X	X	
RF(I)-13					X	X	X	
RF(I)-14					X	X	X	
RF(I)-15		X				X	X	
RF(I)-16							X	X
RF(I)-17					X		X	X

Tabla 070. Trazabilidad 02; Requisitos No Funcionales – Casos de Uso

Requisitos No Funcionales	CU(E)-01	CU(E)-02		CU(I)-01	CU(I)-02	CU(I)-03	CU(I)-04	CU(I)-05
RNF-01	X			X				
RNF-02								
RNF-03		X			X	X	X	X
RNF-04		X			X	X		
RNF-05	X			X		X		
RNF-06								
RNF-07		X				X	X	
RNF-08	X			X				
RNF-09	X			X				
RNF-10	X			X				
RNF-11								
RNF-12		X						
RNF-13								
RNF-14		X			X	X	X	
RNF-15		X			X	X	X	

Tabla 071. Trazabilidad 03; Requisitos Funcionales – Requisitos No Funcionales

Requisitos Funcionales	RNF-01	RNF-02	RNF-03	RNF-04	RNF-05	RNF-06	RNF-07	RNF-08	RNF-09	RNF-10	RNF-11	RNF-12	RNF-13	RNF-14	RNF-15
RF(E)-01												X		X	X
RF(E)-02												X	X		
RF(E)-03															
RF(E)-04												X	X		
RF(E)-05															
RF(E)-06	X	X			X			X	X	X					
RF(E)-07	X	X			X			X	X	X					
RF(E)-08	X	X			X			X	X	X					
RF(E)-09															
RF(E)-10															
RF(I)-01	X	X				X		X		X					
RF(I)-02	X	X			X			X		X					
RF(I)-03	X	X			X			X		X					
RF(I)-04	X	X			X			X		X					
RF(I)-05	X	X			X			X	X	X					
RF(I)-06	X	X			X			X	X	X					
RF(I)-07		X				X								X	X
RF(I)-08			X	X		X								X	X
RF(I)-09			X	X	X	X	X							X	
RF(I)-10			X	X	X		X								
RF(I)-11					X	X	X								
RF(I)-12															
RF(I)-13															
RF(I)-14															
RF(I)-15			X	X	X	X	X								
RF(I)-16															
RF(I)-17			X												

Tabla 072. Trazabilidad 04; Requisitos y Casos de Uso –Objetivos y Características

Requisitos Funcionales	OBJ-01	OBJ-02	OBJ-03	OBJ-04	OBJ-05	OBJ-06		C1	C2	C3	C4	C5	C6	C7	C8
RF(E)-01															
RF(E)-02															
RF(E)-03															
RF(E)-04															
RF(E)-05			X	X										X	X
RF(E)-06						X				X	X				
RF(E)-07						X				X	X				
RF(E)-08						X				X	X				
RF(E)-09	X														
RF(E)-10	X														
RF(I)-01						X				X	X				
RF(I)-02						X				X	X				
RF(I)-03						X				X	X				
RF(I)-04						X			X		X	X			
RF(I)-05						X				X	X				
RF(I)-06						X				X	X				
RF(I)-07								X	X			X			
RF(I)-08	X	X	X	X	X	X		X			X	X	X	X	X
RF(I)-09															
RF(I)-10															
RF(I)-11			X					X					X	X	
RF(I)-12	X														
RF(I)-13	X	X						X							X
RF(I)-14	X														
RF(I)-15	X		X		X									X	
RF(I)-16															
RF(I)-17															
CU(E)-01						X				X					
CU(E)-02															
CU(I)-01						X				X					
CU(I)-02	X	X	X	X	X			X	X		X	X	X	X	X
CU(I)-03	X	X	X	X	X			X	X		X	X	X	X	X
CU(I)-04	X	X	X	X	X			X			X	X	X	X	X
CU(I)-05					X			X				X			

CAPITULO V: DISEÑO

CAPITULO 5: DISEÑO

El diseño de Software juega un papel importante en el desarrollo de software lo cual permite al ingeniero producir varios modelos del sistema o producto de que se va a construir. El diseño es el primer paso de la fase de desarrollo de cualquier producto o sistema de ingeniería.

Una vez que se han establecido los requisitos del software, el diseño es la primera de tres actividades técnicas: diseño, codificación y prueba. Cada actividad transforma la información de forma que al final se obtiene un software validado.

El diseño es técnicamente la parte central de la ingeniería del software. Durante el diseño se desarrollan, revisan y se documentan los refinamientos progresivos de las estructuras de datos, de la estructura del programa y de los detalles procedimentales. El diseño da como resultado representaciones cuya calidad puede ser evaluada.

5.1- Arquitectura

El concepto de arquitectura de software se refiere a la estructuración del sistema que, idealmente, se crea en etapas tempranas del desarrollo. Esta estructuración representa un diseño de alto nivel del sistema que tiene dos propósitos primarios: satisfacer los atributos de calidad (desempeño, seguridad, modificabilidad), y servir como guía en el desarrollo.

La arquitectura de software es un conjunto de patrones que proporcionan un marco de referencia necesario para guiar la construcción de un software, permitiendo a los programadores, analistas y todo el conjunto de desarrolladores del software compartir una misma línea de trabajo y cubrir todos los objetivos y restricciones de la aplicación.

Arquitectura es la organización fundamental de un sistema descrita en:

- ✓ Sus componentes.
- ✓ Relación entre ellos y con el ambiente.
- ✓ Principios que guían su diseño y evolución.

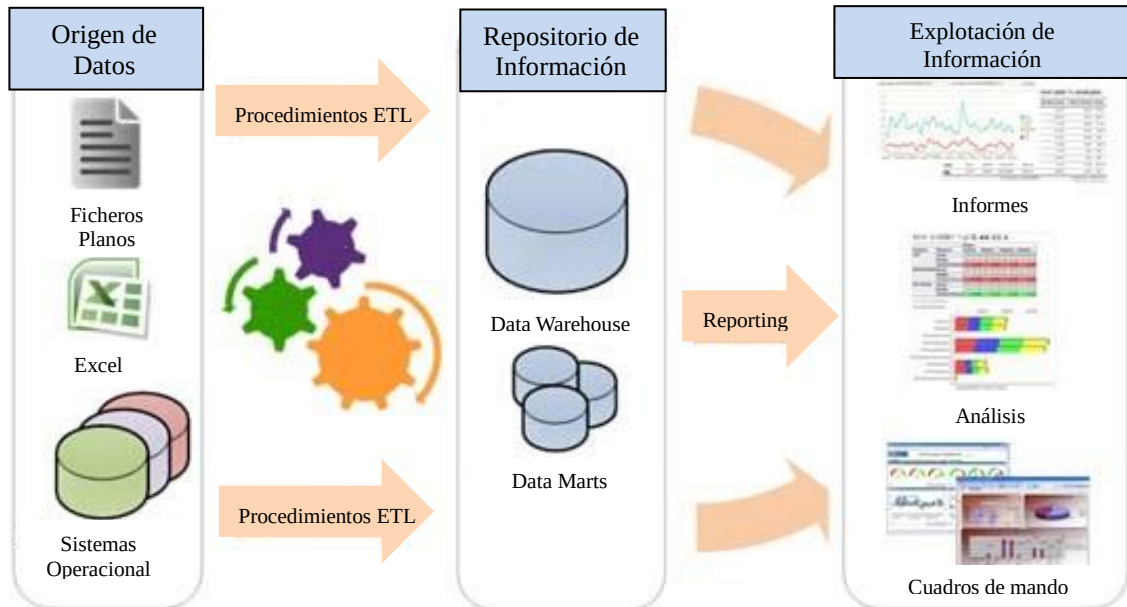
Objetivos de la arquitectura:

- ✓ Alinear la tecnología a las necesidades del negocio
- ✓ Dar soporte al resto de áreas operativas de forma eficiente
 - Permitiendo el máximo grado de agilidad empresarial
 - Implantando nuevos modelos de negocio fácilmente

5.1.1- Arquitectura Lógica

La arquitectura lógica expresa cuales son los componente lógicos (subsistemas, o macrofunciones) que participan en nuestra solución y la relación entre ellos.

Ilustración 017. Arquitectura Lógica



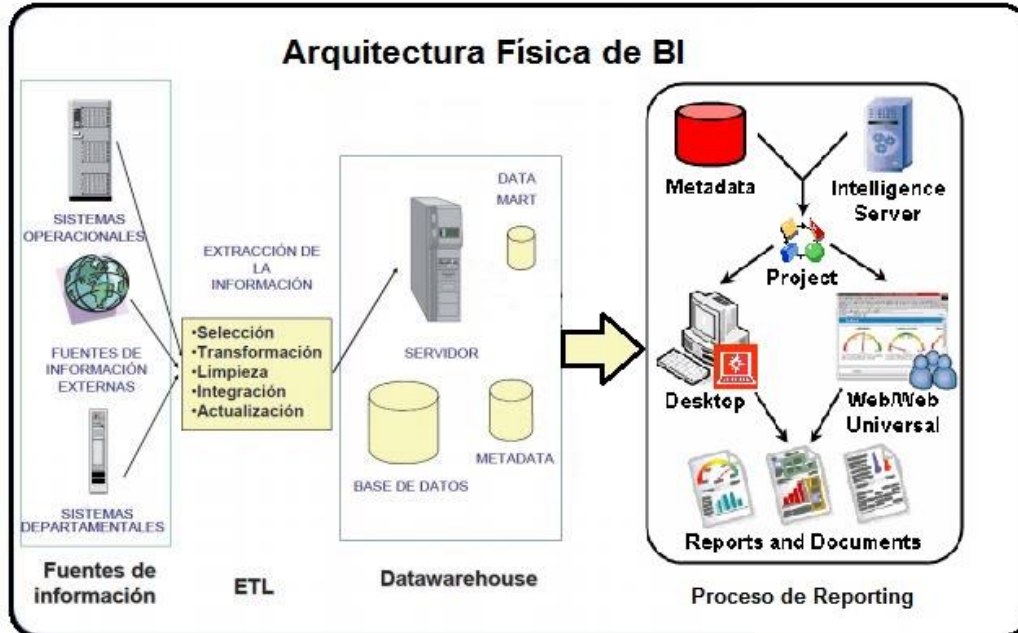
En el capítulo dos de este documento hablábamos de las diferentes fases de un proyecto de BI. Teniendo estas etapas en cuenta y viendo el diagrama anterior podemos distinguir los siguientes componentes principales:

- ✓ **Origen de datos:** Las distintas fuentes de información origen sobre la que se empieza a montar el sistema BI. Estas pueden ser múltiples con distintos formatos. En el caso de nuestro proyecto, el origen de datos es un documento tipo hoja de cálculo (Excel) que la empresa cliente utiliza para organizar la información que usa a diario en su negocio.
- ✓ **ETL:** Sobre las fuentes de información origen se montan los procesos ETL (Extracción, Transformación y Carga) que recogen la información de las fuentes de datos origen, realizan las transformaciones oportunas y carga la información en un nuevo repositorio de información (Data Warehouse) adaptado para poder realizar sobre él la exploración de la información.
- ✓ **Exploración de la información (Reporting):** Conjunto de herramientas que permiten recuperar la información del Data Warehouse adaptada a las necesidades que el cliente requiera.

5.1.2- Arquitectura Física

La arquitectura física expresa cuáles son los componentes físicos (Cliente, servidor, servidor web, BD...) que participan en nuestra solución, así como la relación entre ellos.

Ilustración 018. Arquitectura Física



Una vez que se ha descrito la arquitectura lógica, vamos a describir la arquitectura física, los elementos físicos que se van a utilizar en nuestro proyecto de BI:

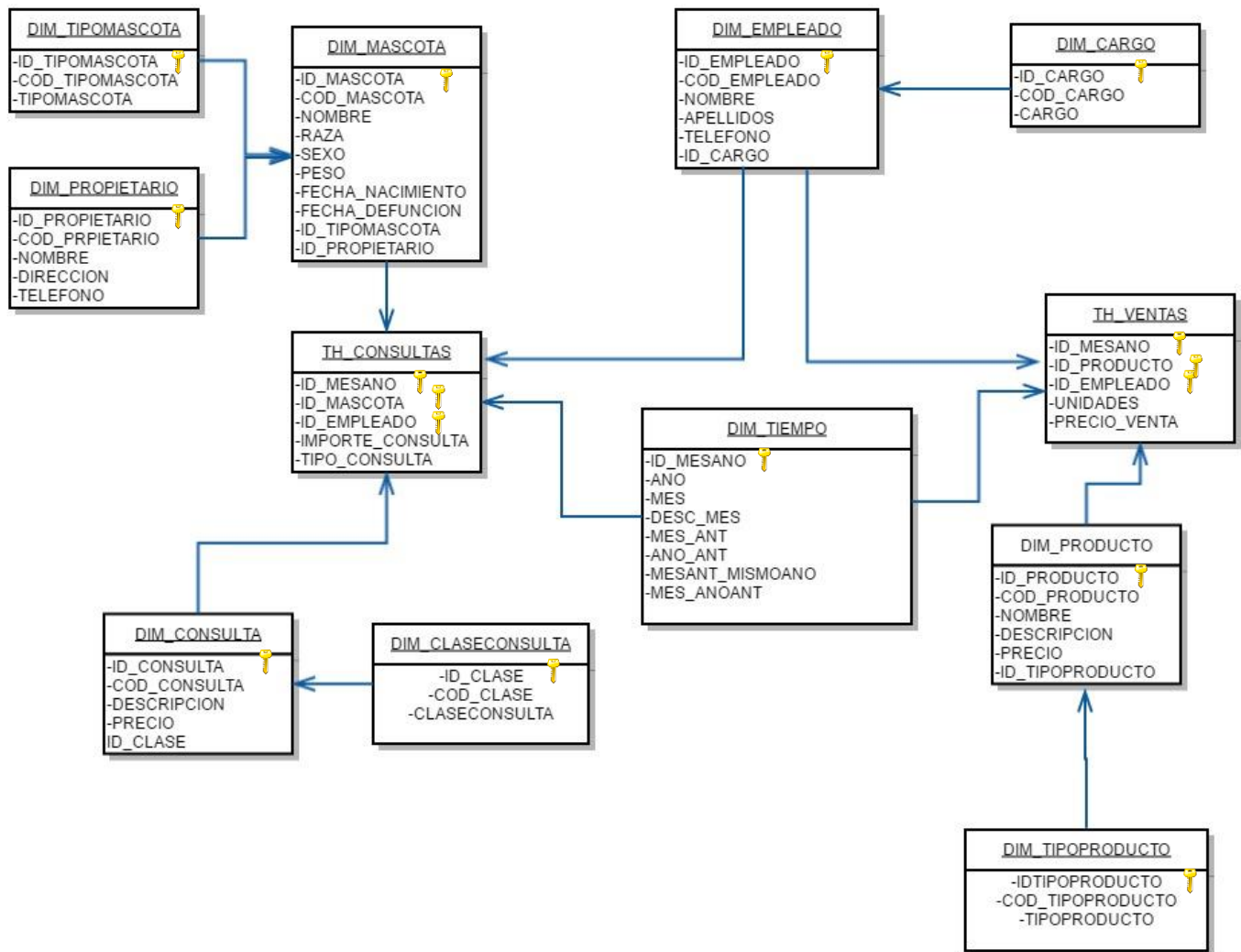
- **Documentos de Origen (Base de datos o documentos):** El cliente nos debe de facilitar una base de datos o un documento en donde se contenga la información concerniente al negocio propio de la empresa cliente. Esto es absolutamente necesario para el comienzo del proyecto.
- **Data Warehouse DW (Almacén de datos o Base de Datos):** Una de las definiciones más famosas sobre DW, es la de William Harvey Inmon, quien define: “Un Data Warehouse es una colección de datos orientada al negocio, integrada, variante en el tiempo y no volátil para el soporte del proceso de toma de decisiones de la gerencia”. Esto es, una base de datos que contiene la información que el cliente nos proporciona pero tratada y modelada adecuadamente (mediante ETLs) para su posterior visualización en informes.
- **Data marts:** Son subconjuntos de datos de un Data Warehouse para áreas específicas. Su función principal es dividir la información a analizar por áreas de análisis, simplificando el almacén de datos global y ayudando a la toma de decisiones de un área específica del negocio.
- **Metadata:** repositorio que almacena la definición de los objetos de reporting: la información sobre el data warehouse. Son datos que definen los datos del Data Warehouse.
- **Intelligence Server:** servidor analítico optimizado para el reporting empresarial.
- **Desktop:** aplicación en entorno windows que proporciona un completo abanico de funciones analíticas diseñadas para facilitar el desarrollo de informes.
- **Web and Web Universal:** interfaz de usuario altamente interactivo para la ejecución de informes y análisis.

5.2- Diagrama de Clases

Un diagrama de clases es un tipo de diagrama de estructura estática que describe la distribución de un sistema mostrando las clases del sistema, sus atributos, operaciones (o métodos), y las relaciones entre los objetos.

A continuación vemos el diagrama de clases correspondiente con nuestro sistema:

Ilustración 019. Diagrama Clases



Ya se ha explicado el modelado del sistema en puntos anteriores. Este diagrama de clases, no es más que una representación gráfica de todo lo que hemos visto hasta el momento.

En el gráfico se ven todas y cada una de las tablas creadas para el Data Warehouse con cada uno de sus campos. Se especifica con una llave amarilla al lado, los campos que actuarán como clave primaria en la base de datos.

Para poder navegar por las jerarquias sin problema, el id de las tablas con mayor granularidad pasa a ser campo de sus tablas hijas con menor grano. Permitiendo así, poder visualizar la información de cualquier tabla relacionada.

5.3- Diagramas de Secuencia

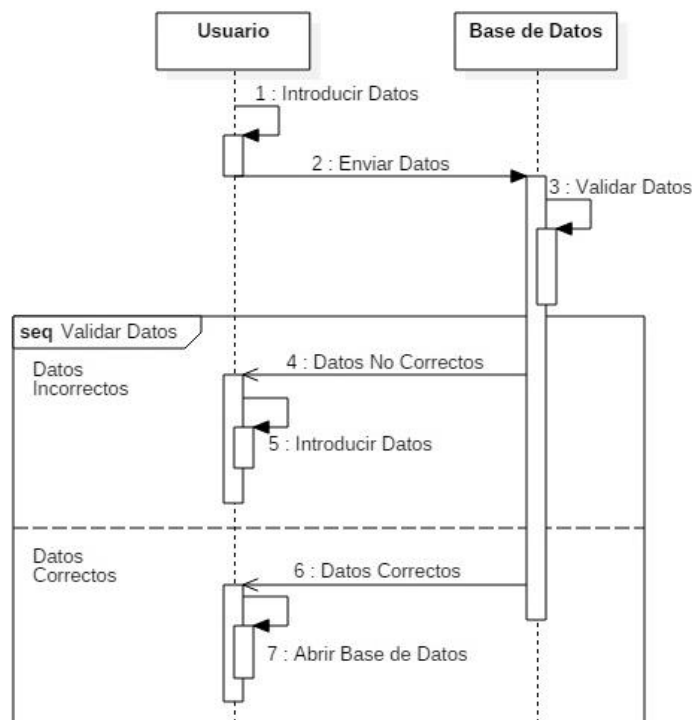
Un diagrama de secuencia muestra la interacción de un conjunto de objetos en una aplicación a través del tiempo. El diagrama de secuencia contiene detalles de implementación del escenario, incluyendo los objetos y clases que se usan para implementar el escenario y mensajes intercambiados entre los objetos.

El diagrama de secuencia es uno de los diagramas más efectivos para modelar interacción entre objetos en un sistema. Este diagrama también es llamado diagrama de interacción; muestra las interacciones entre un conjunto de objetos ya sean clases o actores, ordenadas según el tiempo en que tienen el lugar.

Como en nuestro sistema tenemos dos partes diferenciadas, con sus casos de uso correspondientes, debemos discernir también dos clases de diagramas de secuencias: los diagramas de secuencias dedicados a el proceso de carga de datos o ETLs, cuyo usuario final es otro desarrollador del sistema, y los diagramas de secuencia dedicados a los informes, cuyo usuario final es el cliente.

- **Diagramas de secuencia ETLs**

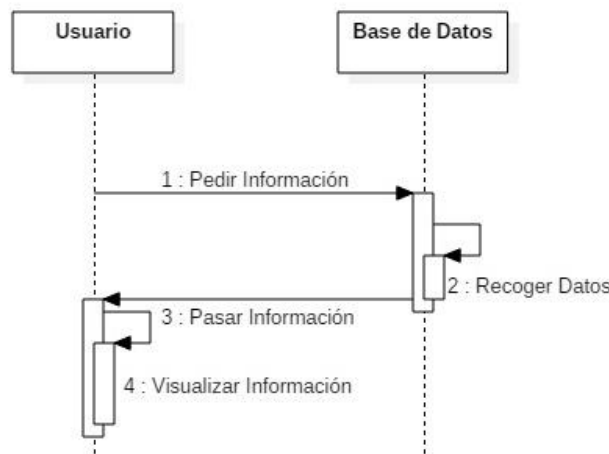
Ilustración 020. Diagrama Secuencia (ETL) 01; Autenticar Usuario



Para autenticar un usuario en la base de datos es sencillo, primero accedes al sistema gestor de base de datos, en nuestro caso SQL Server, y para poder realizar consultas o modificaciones en la base de datos de nuestro sistema nos pide un usuario y contraseña.

Así comienza la secuencia de acciones, el usuario debe autenticarse con sus credenciales, el sistema gestor envía estos credenciales a la base de datos de usuarios del sistema para comprobar que son correctos. Si los datos son correctos, el usuario puede consultar la base de datos del sistema. Si por el contrario, los credenciales son incorrectos el sistema gestor de base de datos pedirá que ingrese de nuevo el usuario y contraseña.

Ilustración 021. Diagrama Secuencia (ETL) 02; Consultar Base de Datos

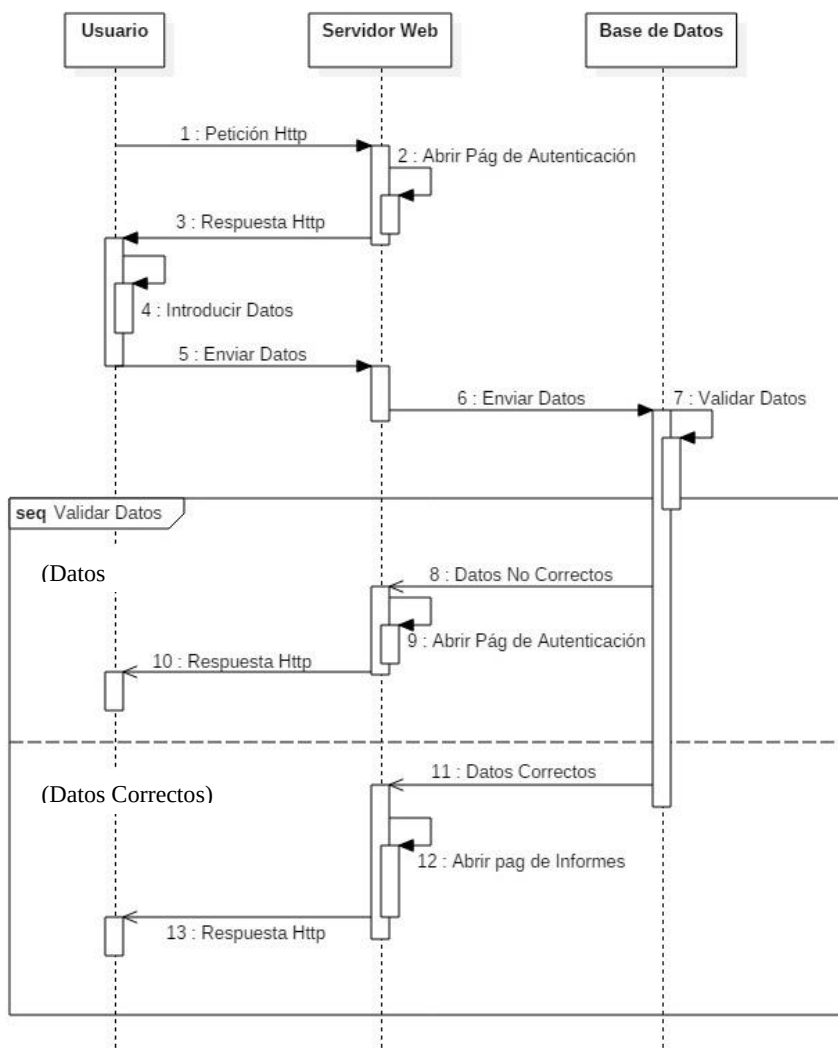


En diagrama de secuencias anterior muestra el procedimiento para poder consultar la base de datos del sistema. Este es muy sencillo de entender, tan solo consta de cuatro pasos.

El usuario accede a la base de datos con el sistema gestor de base de datos (SQL Server), mediante sentencias SQL, la base de datos procesa la petición realizada y muestra un mensaje con la ejecución de la petición.

- **Diagramas de secuencia Informes**

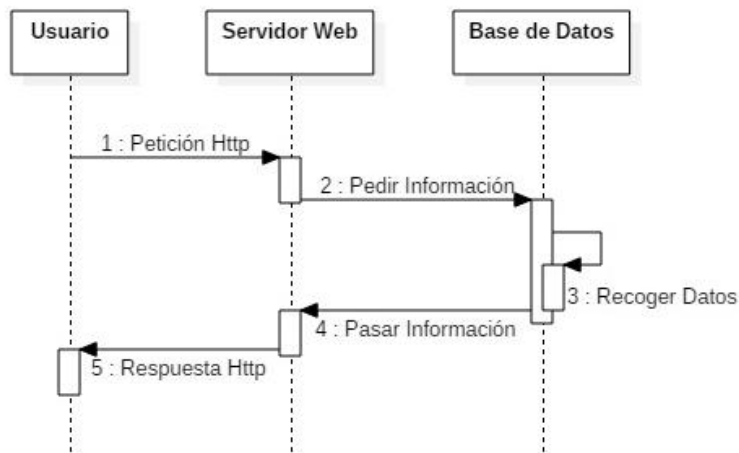
Ilustración 022. Diagrama Secuencia (Informes) 01; Autenticar Usuario



En la anterior imagen podemos ver el funcionamiento del sistema al autenticar un usuario en la interfaz principal. Al acceder al sistema, se muestra la página de autenticación, donde el usuario debe introducir su nombre de usuario y su contraseña.

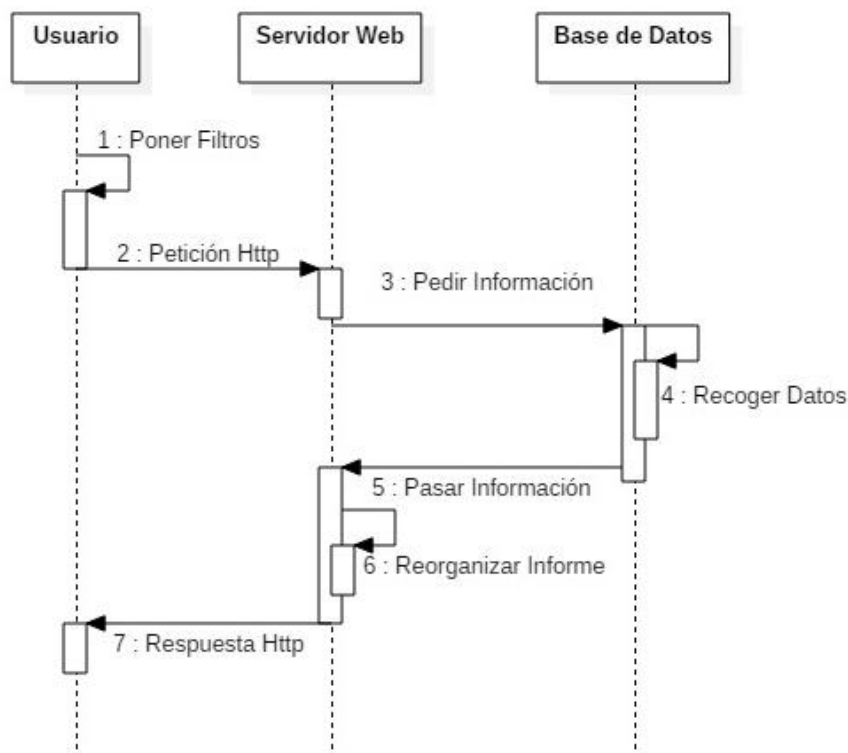
Una vez que el usuario ha introducido sus credenciales, la base de datos debe validar esta información. Si los datos son correctos, se abrirá el Menú de informes. Si por el contrario, la información proporcionada por el usuario es errónea, se volverán a pedir estos datos abriendo de nuevo la página de autenticación.

Ilustración 023. Diagrama Secuencia (Informes) 02; Visualizar Informe



Para la visualización de informes, la secuencia es más sencilla. El usuario, una vez dentro de la aplicación, visualiza el menú de informes, tan solo debe seleccionar uno de estos y después hacer doble click en el, de este modo, el correspondiente informe se abre y muestra su información correspondiente.

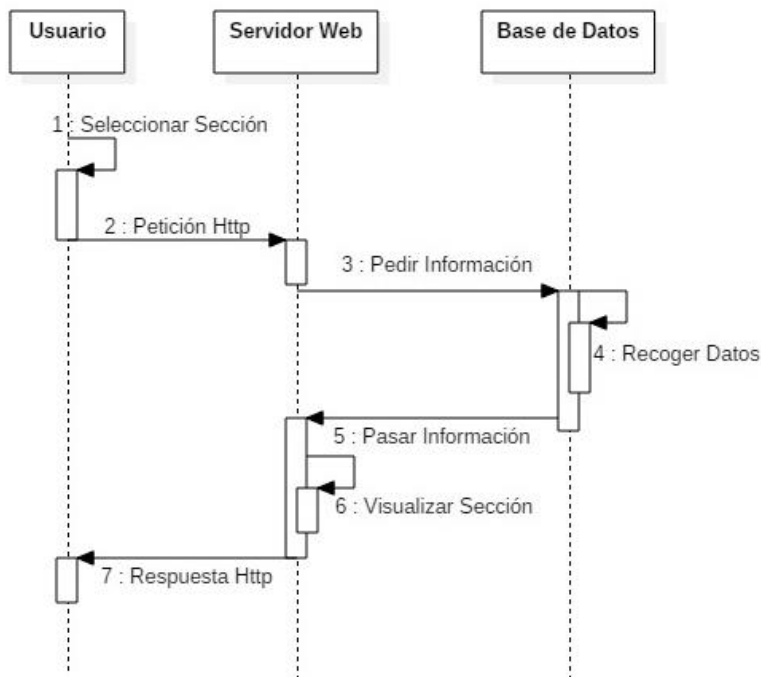
Ilustración 024. Diagrama Secuencia (Informes) 03; Filtrar Informe



Para poner filtros a un informe, este debe estar abierto previamente. En la parte superior es donde el usuario tendrá los filtros correspondientes a dicho informe, tan solo deberá seleccionar los datos que mejor le parezcan.

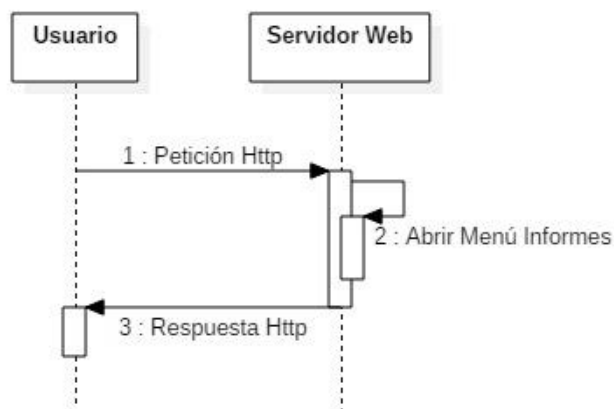
Una vez se cambie un filtro, el sistema buscará los datos correspondientes a dicha información en la base de datos y se visualizará el informe con los cambios oportunos. Deberán modificarse todas las secciones del informe respecto de los filtros seleccionados.

Ilustración 025. Diagrama Secuencia (Informes) 04; Aumentar Sección de Informe



Este caso se asemeja mucho al anterior. El usuario, en un momento determinado puede querer visualizar la información de una determinada sección, ya sea de un gráfico o de los datos en crudo. Para realizar esta acción no tiene más que pinchar en la parte superior derecha del cuadro de dicha sección y esta ocupará toda la pantalla, reorganizando los datos.

Ilustración 026. Diagrama Secuencia (Informes) 05; Cerrar Informe



Para cerrar un informe tan solo hay que pinchar en la parte superior derecha del mismo. Al cerrar un informe, el sistema volverá a mostrar el menú de informes.

5.4- Diagramas de Estados

Los diagramas de estado son un método conocido para explicar el comportamiento de un sistema. Explican todos los estados posibles en los que puede ingresar un objeto particular y la manera en que modifica el estado del objeto, como resultado de los eventos que llegan a él.

Un diagrama de estados es un diagrama utilizado para determinar cada una de las rutas o caminos que puede tomar un movimiento de información luego de ejecutarse cada proceso. Permite identificar bajo qué pruebas se ejecuta cada uno de los procesos y en qué momento podrían tener una variación. El diagrama de estados consiente visualizar de una forma ordenada la ejecución de cada uno de los procesos.

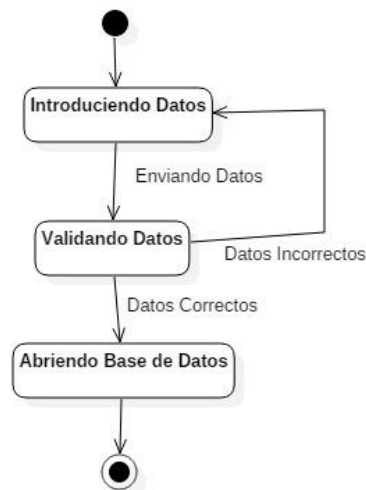
Un estado es una situación durante la vida de un objeto, de forma que cuando dicha situación se satisface se lleva a cabo alguna acción o se espera por un evento. El estado de un objeto se puede caracterizar por el valor de una o varias de las características de su clase, además, el estado de un objeto también se puede caracterizar por la existencia de un enlace con otro objeto.

El diagrama de estados y transiciones incluye todos los mensajes que un objeto puede enviar o recibir. En un diagrama de estados, un escenario simboliza un camino dentro del diagrama. Dado que generalmente el espacio entre dos envíos de mensajes representa un estado, se pueden utilizar los diagramas de secuencia para buscar los diferentes estados de un objeto.

Del mismo modo que en los diagramas de secuencia, hay dos tipos de diagramas de estados, los dedicados al proceso de carga de datos o ETLs y los dedicados a informes.

- **Diagramas de Estados de ETLs**

Ilustración 027. Diagrama Estados (ETL) 01; Autenticar Usuario

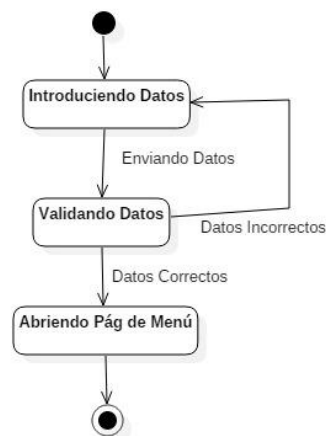


En el diagrama anterior, la máquina pasa de un estado inicial a que el usuario introduzca sus credenciales. Los datos de usuario y contraseña son validados por el sistema. Si son correctos, el sistema abre la base de datos, permitiendo hacer consultas. Si por el contrario los datos no son correctos, el sistema volverá a pedir al usuario que introduzca sus credenciales.

Ilustración 028. Diagrama Estados (ETL) 02; Consultar Base de Datos

En este caso pasamos de un estado inicial a realizar una consulta determinada. Una vez se ha escrito la consulta, el sistema gestor de base de datos busca la información correspondiente a la sentencia SQL escrita por el usuario y muestra los datos en pantalla.

• Diagramas de Estados de Informes

Ilustración 029. Diagrama Estados (Informes) 01; Autenticar Usuario

Para autenticar usuario en el sistema, es muy similar al caso de la base de datos. Tan solo tenemos una diferenciación, una vez los datos se han validado, el sistema abre la página dedicada a mostrar el menú con los distintos informes. Si las credenciales introducidos por el usuario no son correctos, el sistema vuelve a mostrar la página de autenticarse.

Ilustración 030. Diagrama Estados (Informes) 02; Visualizar Informe

Para visualizar un informe, el usuario debe hacer doble click sobre el informe que quiera visualizar en el menú de informes y este automáticamente se abrirá.

Ilustración 031. Diagrama Estados (Informes) 03; Filtrar Informe



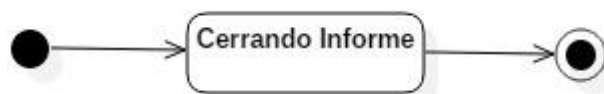
Una vez que el informe se está visualizando, el usuario puede modificar los filtros de la parte superior del mismo para ajustar la información que se muestra a lo que se quiere ver. Una vez que se utilizan los filtros el sistema realiza la búsqueda necesaria a la base de datos y muestra la información oportuna en el informe.

Ilustración 032. Diagrama Estados (Informes) 04; Aumentar Sección de Informe



Cuando el usuario necesita ampliar una sección del informe, pinchará en la parte superior derecha de esta sección. La información se amplía ocupando toda la pantalla.

Ilustración 033. Diagrama Estados (Informes) 05; Cerrar Informe



Para cerrar un informe basta con pinchar en la parte superior derecha y el sistema volverá al menú de informes.

5.5- Diseño de Interfaz de Usuario

En cuanto a interfaces de usuario deben ser sencillas y manejables, fácilmente comprensible para todo tipo de personas. En el caso del sistema que nos acontece, tenemos tres interfaces básicas: principal, menú de informes e informes. A continuación explicaremos cada una de ellas en detalle.

Como interfaz de la página principal, tenemos el siguiente prototipo:

Ilustración 034. Interfaz Usuario 01; Página Principal

El prototipo muestra un formulario de inicio de sesión con los siguientes elementos:

- Etiqueta 'Usuario:' seguida de un campo de entrada de texto rectangular.
- Etiqueta 'Contraseña:' seguida de un campo de entrada de texto rectangular.
- Un botón rectangular centrado con el texto 'Iniciar sesión'.

Tabla 073. Interfaz Usuario 01; Principal

I-01	Principal
Descripción	Es la interfaz genérica de la aplicación, se abre en cuanto accedemos al sistema. Es viable a todos los usuarios (cliente y administrador).
Activación	Se activa al acceder a la aplicación
Eventos	Cambio de pantalla al inicio de sesión de un usuario registrado.

Una vez ingresados nuestros credenciales, el sistema nos llevará a un menú de informes:

Ilustración 035. Interfaz Usuario 02; Menú de Informes

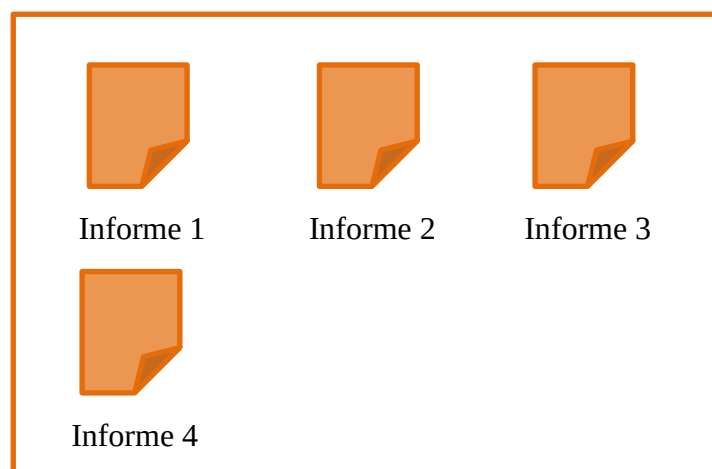
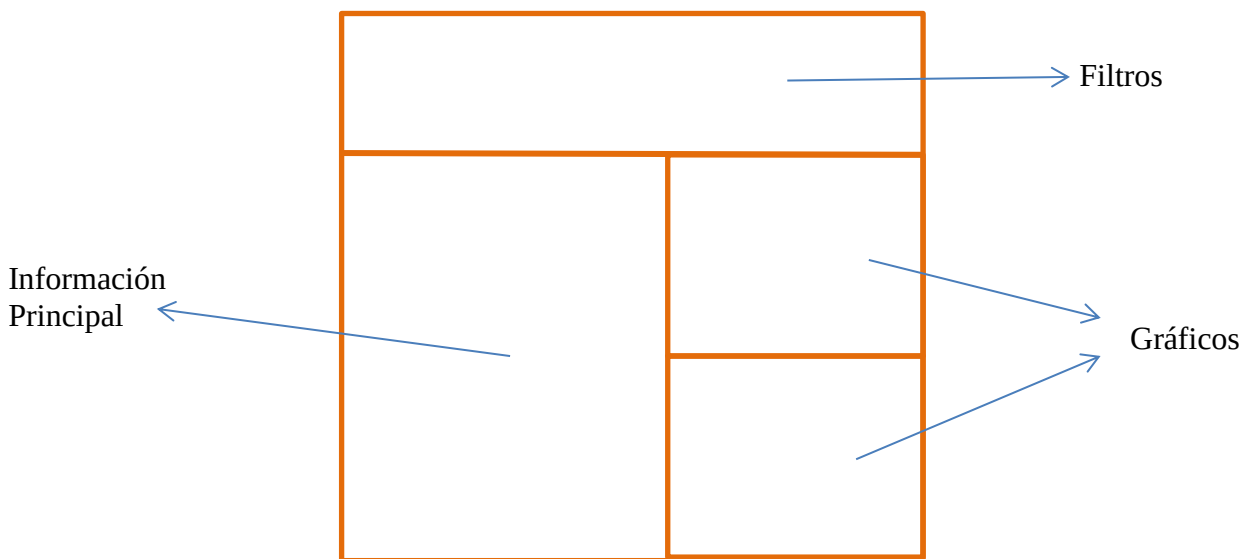


Tabla 074. Interfaz Usuario 02; Menú de Informes

I-02	Menú de Informes
Descripción	Esta interfaz se activa cuando un usuario introduce su nombre de usuario y contraseña en la aplicación (en el interfaz principal). El sistema comprueba que los datos introducidos por el usuario son correctos y abre esta nueva interfaz, en la cual, el usuario puede acceder a los distintos informes.
Activación	Se activa cuando un usuario inicia su sesión en el sistema
Eventos	Cambio de interfaz al acceder a algún informe en concreto, en este caso se activa la interfaz I-03.

Al seleccionar uno de estos informes, se abrirá y tendrá este aspecto:

Ilustración 036. Interfaz Usuario 03; Informe**Tabla 075. Interfaz Usuario 03; Informes**

I-03	Informes
Descripción	Esta interfaz se activa cuando un usuario abre un informe (en el interfaz Menú de Informes). El sistema abre esta nueva interfaz, en la cual, el usuario puede visualizar la información que contienen los distintos informes.
Activación	Se activa cuando un usuario hace doble click sobre un informe en el menú de informes.
Eventos	Cambio de interfaz al cerrar el informe, que se activa nuevamente el interfaz I-02.

CAPITULO VI: IMPLEMENTACIÓN

CAPITULO 6: IMPLEMENTACIÓN

La implementación es la instalación de una aplicación informática, realización o la ejecución de un plan, idea, modelo científico, diseño, especificación, estándar, algoritmo o política. Es decir, es la realización de una especificación técnica o algoritmos como un programa, componente software, u otro sistema de cómputo.

En este capítulo vamos a entrar en la realización del sistema entero, cómo se ha creado paso a paso. Se explicará exhaustivamente todo el desarrollo del proyecto. Posiblemente sea el capítulo más importante (sobre todo para desarrolladores), puesto que define desde el principio y hasta el final el proceso que se ha ido siguiendo para la creación del sistema de BI, se explican las herramientas usadas, en los momentos en que se utilizan y para qué se han utilizado.

Como en el proyecto hay dos partes diferenciadas, vamos a explicar ambas partes por separado. En la primera parte, se definirán las herramientas del proceso de carga de información en el Data Warehouse. Mientras que en la segunda parte se especificará el funcionamiento de la herramienta de Microstrategy, y cómo se ha utilizado para la creación de los informes finales.

6.1- Creación de la Estructura del Data Warehouse

Antes de comenzar con el proceso de carga de información en la base de datos, tenemos que tener las tablas construidas de la misma. El proceso de carga o ETL tan sólo recoge los datos del operacional, los trata y los vuelca en un destino, pero este destino tiene que estar ya creado y estructurado según el modelo que se ha elegido previamente.

Toda la información contenida en las tablas del Data Warehouse, ya sean dimensiones o tablas de hechos, se obtienen directamente del operacional u origen de datos. Toda la información menos los registros de la dimensión de tiempo. La dimensión de tiempo es creada para el proceso de BI. En el operacional no tiene cabida dedicar una hoja de cálculo o una tabla exclusivamente a fechas, pero en el Data Warehouse es importantísimo tener una tabla dedicada únicamente a las fechas, puesto que sobre estos datos se van a sustentar los informes.

6.1.1- DDL: Estructura

El lenguaje de definición de datos o DDL es un sencillo lenguaje artificial para definir y describir los objetos de una base de datos, su estructura, relaciones y restricciones. En el caso del Data Warehouse que nos acontece, el DDL de la base de datos será el siguiente:

```
CREATE DATABASE DWH_VETERINARIA;
USE DWH_VETERINARIA;
```

```
CREATE TABLE [dbo].[DIM_TIEMPO](
    [ID_MESANO] [int] NOT NULL,
    [AÑO] [int] NOT NULL,
    [MES] [int] NOT NULL,
    [DESC_MES] [varchar](15) NOT NULL,
    [MES_ANT] [int] NOT NULL,
    [MESANT_AÑO] [int] NOT NULL,
    [AÑO_ANT] [int] NOT NULL,
    [MES_AÑOANT] [int] NOT NULL,
    PRIMARY KEY([ID_MESANO])
)
```

```
CREATE TABLE [dbo].[DIM_PROPIETARIO](
    [ID_PROPIETARIO] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_PROPIETARIO] [int] NOT NULL,
    [NOMBRE] [varchar](100) NOT NULL,
    [DIRECCION] [varchar](150) NOT NULL,
```

```

        [TELEFONO] [int] NOT NULL
    )
CREATE TABLE [dbo].[DIM_TIPOMASCOTA](
    [ID_TIPOMASCOTA] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_TIPOMASCOTA] [int] NOT NULL,
    [TIPO_MASCOTA] [varchar](50) NOT NULL
)
CREATE TABLE [dbo].[DIM_MASCOTA](
    [ID_MASCOTA] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_MASCOTA] [int] NOT NULL,
    [NOMBRE] [varchar](50) NOT NULL,
    [RAZA] [varchar](50) NOT NULL,
    [SEXO] [char](1) NOT NULL,
    [PESO] [DECIMAL](5,2),
    [FECHA_NACIMIENTO] [DATE] NOT NULL,
    [FECHA_DEFUNCION] [DATE],
    [ID_TIPOMASCOTA] [int] NOT NULL,
    [ID_PROPIETARIO] [int] NOT NULL,
    FOREIGN KEY([ID_TIPOMASCOTA]) REFERENCES [dbo].[DIM_TIPOMASCOTA] ([ID_TIPOMASCOTA]),
    FOREIGN KEY([ID_PROPIETARIO]) REFERENCES [dbo].[DIM_PROPIETARIO] ([ID_PROPIETARIO])
)

CREATE TABLE [dbo].[DIM_CLASECONSULTA](
    [ID_CLASE] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_CLASE] [int] NOT NULL,
    [CLASE_CONSULTA] [varchar](50) NOT NULL
)

CREATE TABLE [dbo].[DIM_CONSULTA](
    [ID_CONSULTA] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_CONSULTA] [int] NOT NULL,
    [DESCRIPCION] [varchar](50) NOT NULL,
    [PRECIO] [DECIMAL](6,2) NOT NULL,
    [ID_CLASE] [int] FOREIGN KEY([ID_CLASE])
    REFERENCES [dbo].[DIM_CLASECONSULTA] ([ID_CLASE]) NOT NULL
)

CREATE TABLE [dbo].[DIM_TIPO_PRODUCTO](
    [ID_TIPOPRODUCTO] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_TIPOPRODUCTO] [int] NOT NULL,
    [TIPO_PRODUCTO] [varchar](50) NOT NULL
)

CREATE TABLE [dbo].[DIM_PRODUCTO](
    [ID_PRODUCTO] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_PRODUCTO] [int] NOT NULL,
    [NOMBRE] [varchar](60) NOT NULL,
    [DESCRIPCION] [varchar](255) NOT NULL,
    [PRECIO] [DECIMAL](6,2),
    [ID_TIPOPRODUCTO] [int] FOREIGN KEY([ID_TIPOPRODUCTO])
    REFERENCES [dbo].[DIM_TIPO_PRODUCTO] ([ID_TIPOPRODUCTO]) NOT NULL
)

CREATE TABLE [dbo].[DIM_CARGO](
    [ID_CARGO] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_CARGO] [int] NOT NULL,
    [CARGO] [varchar](50) NOT NULL
)

CREATE TABLE [dbo].[DIM_EMPLEADO](
    [ID_EMPLEADO] [int] IDENTITY(1,1) PRIMARY KEY NOT NULL,
    [COD_EMPLEADO] [int] NOT NULL,
    [NOMBRE] [varchar](50) NOT NULL,
    [APELLIDOS] [varchar](50) NOT NULL,
    [TELEFONO] [int] NOT NULL,
    [ID_CARGO] [int] FOREIGN KEY([ID_CARGO])
    REFERENCES [dbo].[DIM_CARGO] ([ID_CARGO]) NOT NULL
)

CREATE TABLE [dbo].[TH_CONSULTAS](
    [ID_MESANO] [int] NOT NULL,
    [ID_MASCOTA] [int] NOT NULL,
    [ID_EMPLEADO] [int] NOT NULL,
    [ID_CONSULTA] [int] NOT NULL,
    [TIPO_CONSULTA] [int] NOT NULL,
    [IMPORTE_CONSULTA] [decimal](7,2) NOT NULL,

```

```

PRIMARY KEY([ID_MASCOTA],[ID_EMPLEADO],[ID_CONSULTA],[ID_MESANO]),
FOREIGN KEY([ID_MESANO]) REFERENCES [dbo].[DIM_TIEMPO] ([ID_MESANO]),
FOREIGN KEY([ID_MASCOTA]) REFERENCES [dbo].[DIM_MASCOTA] ([ID_MASCOTA]),
FOREIGN KEY([ID_EMPLEADO]) REFERENCES [dbo].[DIM_EMPLEADO] ([ID_EMPLEADO]),
FOREIGN KEY([ID_CONSULTA]) REFERENCES [dbo].[DIM_CONSULTA] ([ID_CONSULTA])
)
)

CREATE TABLE [dbo].[TH_VENTAS](
    [ID_MESANO] [int] NOT NULL,
    [ID_PRODUCTO] [int] NOT NULL,
    [ID_EMPLEADO] [int] NOT NULL,
    [UNIDADES] [int] NOT NULL,
    [PRECIO_VENTA] [decimal](7,2) NOT NULL,
    PRIMARY KEY([ID_MESANO],[ID_PRODUCTO],[ID_EMPLEADO]),
    FOREIGN KEY([ID_MESANO]) REFERENCES [dbo].[DIM_TIEMPO] ([ID_MESANO]),
    FOREIGN KEY([ID_PRODUCTO]) REFERENCES [dbo].[DIM_PRODUCTO] ([ID_PRODUCTO]),
    FOREIGN KEY([ID_EMPLEADO]) REFERENCES [dbo].[DIM_EMPLEADO] ([ID_EMPLEADO])
)
)

```

Obviamente la creación de tablas debe ser ordenada debido a las referencias cruzadas entre ellas:

-Primero las dimensiones de mayor grano: DIM_TIEMPO, DIM_PROPIETARIO, DIM_TIPOMASCOTA, DIM_CLASECONSULTA, DIM_TIPO_PRODUCTO y DIM_CARGO. El orden de creación de estas dimensiones no es importante.

-Posteriormente las de menos grano, dependientes de las anteriores: DIM_CONSULTA, DIM_MASCOTA, DIM_CONSULTA, DIM_PRODUCTO y DIM_EMPLEADO.

-Finalmente las tablas de hechos: TH_VENTAS y TH_CONSULTAS, pues hacen referencia a las anteriores dimensiones.

Los diferentes atributos de cada tabla son los propios al operacional dado. Toda esta estructura de la base de datos, la hemos explicado con anterioridad en el punto dedicado al diagrama de clases.

6.1.2- DML: Contenido de DIM_TIEMPO

El lenguaje manipulación de datos o DML es un lenguaje artificial de cierta complejidad que permite el manejo y procesamiento del contenido de la base de datos. En nuestro proyecto utilizaremos el DML tan sólo en la tabla dedicada a la dimensión de tiempo, puesto que, la dimensión temporal es la única que no podemos cargar de información mediante procesos ETL, ya que no existen datos en el operacional para ello.

El DML con la información de la dimensión temporal es el siguiente:

```

INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,
MES_ANOANT) VALUES (201401,2014,01,'ENERO',12,201312,2013,201301)
INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,
MES_ANOANT) VALUES (201402,2014,02,'FEBRERO',01,201401,2013,201302)
INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,
MES_ANOANT) VALUES (201403,2014,03,'MARZO',02,201402,2013,201303)
INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,
MES_ANOANT) VALUES (201404,2014,04,'ABRIL',03,201403,2013,201304)
INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,
MES_ANOANT) VALUES (201405,2014,05,'MAYO',04,201404,2013,201305)
INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,
MES_ANOANT) VALUES (201406,2014,06,'JUNIO',05,201405,2013,201306)
INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,
MES_ANOANT) VALUES (201407,2014,07,'JULIO',06,201406,2013,201307)
INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,
MES_ANOANT) VALUES (201408,2014,08,'AGOSTO',07,201407,2013,201308)
INSERT INTO DIM_TIEMPO (ID_MESANO, ANO, MES, DESC_MES, MES_ANT, MESANT_ANO, ANO_ANT,

```

```

MES_AÑOANT) VALUES (201409,2014,09,'SEPTIEMBRE',08,201408,2013,201309)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201410,2014,10,'OCTUBRE',09,201409,2013,201310)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201411,2014,11,'NOVIEMBRE',10,201410,2013,201311)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201412,2014,12,'DICIEMBRE',11,201411,2013,201312)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201501,2015,01,'ENERO',12,201412,2014,201401)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201502,2015,02,'FEBRERO',01,201501,2014,201402)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201503,2015,03,'MARZO',02,201502,2014,201403)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201504,2015,04,'ABRIL',03,201503,2014,201404)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201505,2015,05,'MAYO',04,201504,2014,201405)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201506,2015,06,'JUNIO',05,201505,2014,201406)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201507,2015,07,'JULIO',06,201506,2014,201407)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201508,2015,08,'AGOSTO',07,201507,2014,201408)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201509,2015,09,'SEPTIEMBRE',08,201508,2014,201409)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201510,2015,10,'OCTUBRE',09,201509,2014,201410)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201511,2015,11,'NOVIEMBRE',10,201510,2014,201411)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201512,2015,12,'DICIEMBRE',11,201511,2014,201412)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201601,2016,01,'ENERO',12,201512,2015,201501)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201602,2016,02,'FEBRERO',01,201601,2015,201502)
INSERT INTO DIM_TIEMPO (ID_MESANO, AÑO, MES, DESC_MES, MES_ANT, MESANT_AÑO, AÑO_ANT,
MES_AÑOANT) VALUES (201603,2016,03,'MARZO',02,201602,2015,201503)

```

Todos estos registros corresponden a dos años, desde enero del 2014 a marzo de 2016 exactamente, y recogen todos los meses de estos años.

6.1.3- Usuarios y Permisos

En el proyecto que nos acontece, existen dos tipos de usuarios como ya hemos explicado anteriormente. Por un lado el administrador, con todos los privilegios sobre la base de datos, por tanto, en todo momento, el administrador, podrá leer, escribir y modificar la información existente en el Data Warehouse.

Por otro lado, tenemos el usuario cliente, el cual, no podrá más que leer en la base de datos la información.

Tabla 076. Usuarios y Permisos; General

BASE DE DATOS	PERMISOS	USUARIO
DWH_VETERINARIA	R,W,D,U	Administrador
DWH_VETERINARIA	R	Cliente

Para crear los usuarios en SQL Server es sencillo:

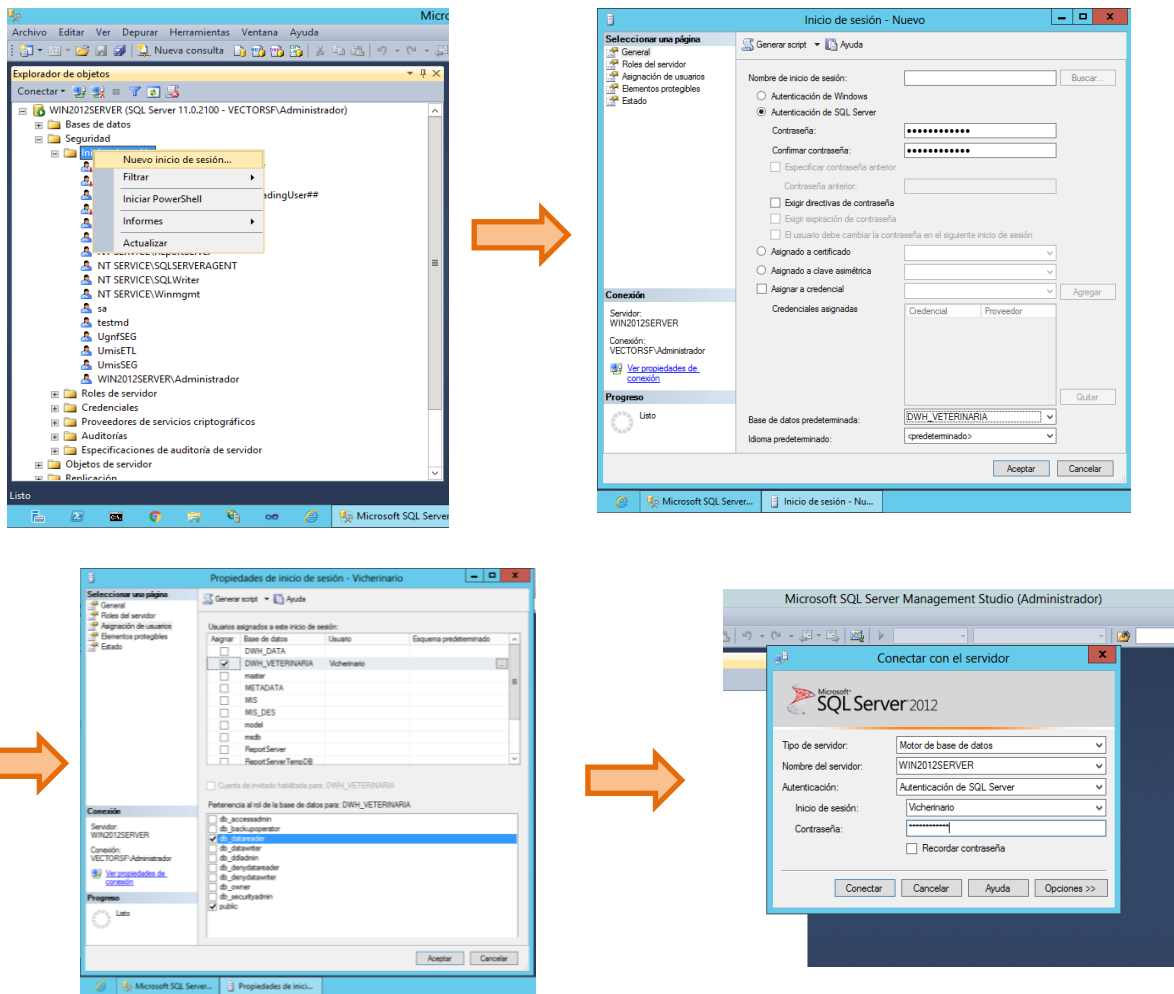
- 1- Abrimos SQL server en la pantalla principal hacemos clic seguridad y a continuación clic derecho en inicios de sesión, nuevo inicio de sesión.
- 2- En el cuadro nuevo inicio de sesión situamos apartado general, damos nombre de inicio “Vicherinario”, marcamos autenticación de SQL Server, establecemos

contraseña, en este caso “Vicherinario” y confirmamos la contraseña, desmarcamos exigir directivas contraseña y abajo en base de datos predeterminada seleccionamos la nuestra DWH_VETERINARIA

- 3- Después, En el apartado Asignación de usuarios, marcar nuestra BBDD y abajo ponemos en los permisos el de db_datareader

De esta forma, la próxima vez que iniciemos SQL server podremos iniciar con este nuevo usuario.

Ilustración 037. Usuarios y Permisos; Configuración



6.2- Proceso de Carga de Datos (ETLs)

6.2.1-Aspectos Previos

- **Detalles del Proceso**

- **Lógica de Carga**

En este apartado se explicará los procesos de ejecución de las ETLs comunes para todas las dimensiones.

Para el proyecto, la lógica de carga de las ETLs a seguir es:

- En primer lugar las dimensiones con menor granularidad: DIM_TIPOMASCOTA, DIM_PROPIETARIO, DIM_CARGO, DIM_TIPOCONSULTA y DIM_TIPOPRODUCTO, sin ningún orden en particular.
- Seguidamente, las dimensiones con más grano: DIM_MASCOTA, DIM_EMPLEADO, DIM_CONSULTA, DIM_PRODUCTO y las dimensiones temporales: DIM_TIEMPOV y DIM_TIEMPO.
- Para finalizar, se cargarán las tablas de hechos: TH_CONSULTAS y TH_VENTAS, en cualquier orden.

Esta ordenación debe mantenerse en cualquier caso para que no den errores las claves foráneas al referir a una tabla sin datos existentes.

- **Estrategia de Carga**

A la hora de realizar una descripción de alto nivel acerca de los procesos ETL que van a componer el proyecto, se tienen bien claras y definidas dos tipos estrategias de carga dependiendo del tipo de tabla que vaya a cargar el proceso ETL. Estas estrategias de carga son:

- Estrategia de carga para las dimensiones: la estrategia de carga acordada para realizar la carga de las dimensiones será INSERT/UPDATE. Siendo el caso de que si un registro no existe se inserta uno nuevo para el mismo y en el caso de que si exista lo que se hace es actualizarlo siempre.
- Estrategia de carga para las tablas de hechos: la estrategia de carga acordada para realizar la carga de las tablas de hechos será INSERT/DELETE. Siendo el caso de que únicamente se insertaran registros en las tablas de hechos, nunca se realizaran UPDATES.

Las cargas se realizarán una sola vez, puesto que los informes del cliente nos los ha encomendado para un determinado momento. Pero, si es necesario, se pueden cargar las ETLs con frecuencia. Para poder tener resultados óptimos, y puesto que la empresa no es muy extensa, ni recibe gran número de datos, la frecuencia de las cargas de los procesos ETL recomendable podrá ser:

- Dimensiones que se cargan una vez.
 - DIM_TIEMPO.
- Tablas que se cargaran Mensualmente.

- DIM_MASCOTA.
- DIM_PROPIETARIO.
- DIM_TIPOMASCOTA.
- DIM_CONSULTA.
- DIM_CLASECONSULTA.
- DIM_EMPLEADO.
- DIM_CARGO.
- DIM_PRODUCTO.
- DIM_TIPOPRODUCTO.

Cargas de las tablas de hechos mensualmente:

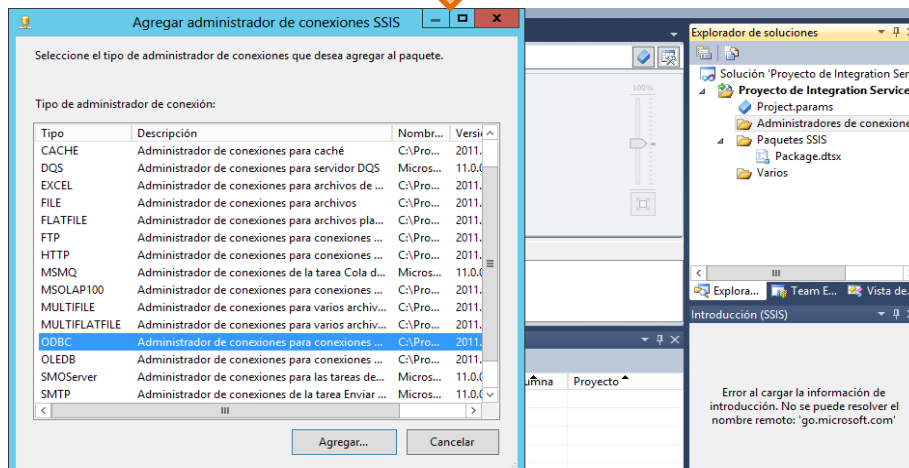
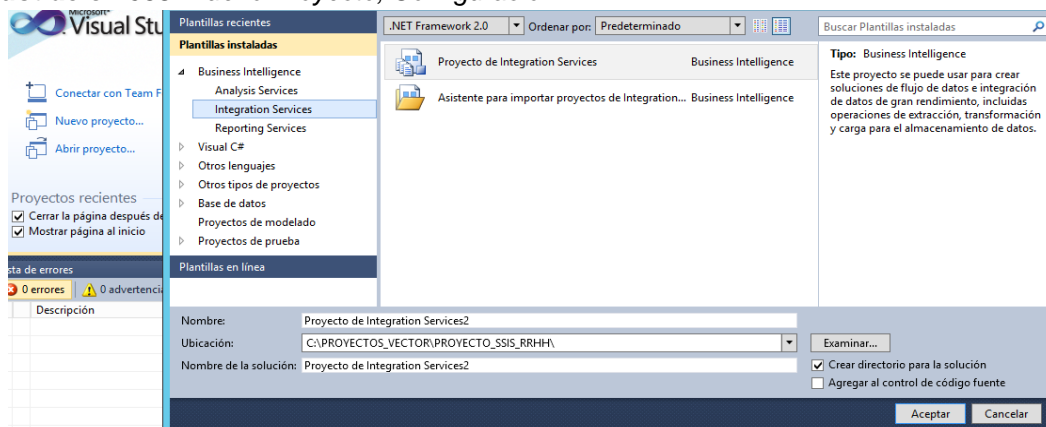
- TH_CONSULTAS.
- TH_VENTAS.

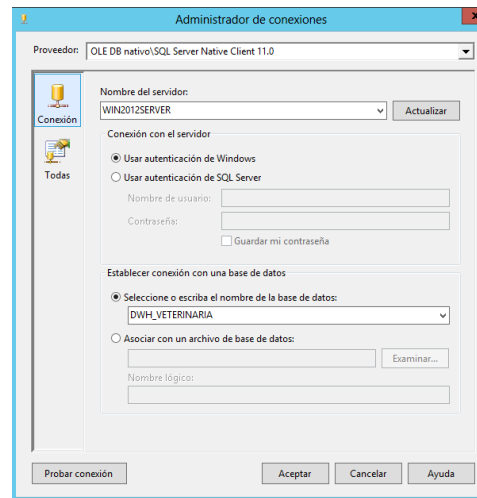
• Crear un Nuevo Proyecto

Para comenzar a utilizar la herramienta de Visual Studio, al lanzarla hay que crear un nuevo proyecto. El proyecto será de tipo Business Intelligence e Integration Services. Le damos nombre y generamos nuevo proyecto.

Continuaremos generando un “Administrador de conexiones” (En la parte derecha de la herramienta). En nuestro caso es de tipo OLEDB, generamos uno nuevo con la configuración apropiada, tal y como se muestra en la figura. Ya tenemos nuestra conexión establecida.

Ilustración 038. Nuevo Proyecto; Configuración





Una vez que tenemos la conexión, podemos empezar a trabajar con la herramienta. Debemos ir creando paquetes (En la parte derecha “Paquetes SSIS”), uno por cada ETL que queramos generar.

• Configuración de Algunas Herramientas

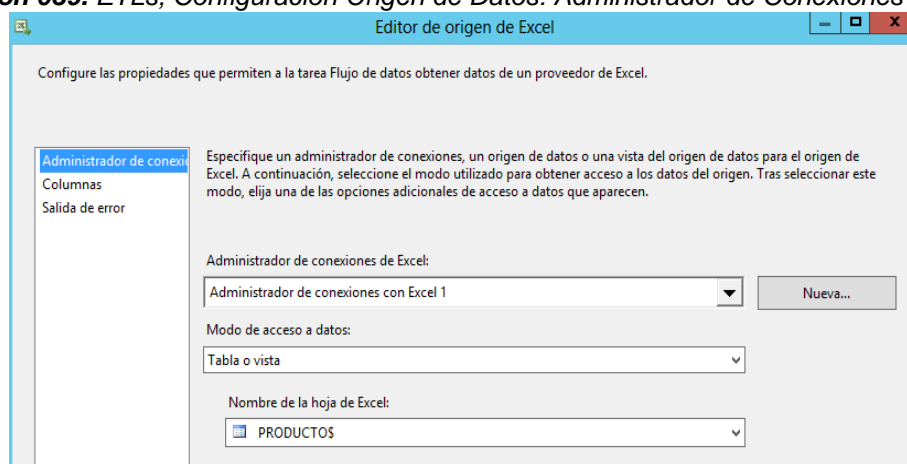
Orígenes de datos: Los orígenes de datos recogen toda la información que va a ser tratada en la ETL.

En este proyecto, para todas las ETLs el origen de datos es el mismo, un archivo Excel. Por ello utilizaremos la herramienta de origen de datos tipo Excel, y lo configuramos para que importe la información del documento en cuestión.

Dependiendo de la ETL, lo configuramos para que los datos sean de una hoja de cálculo u otra.

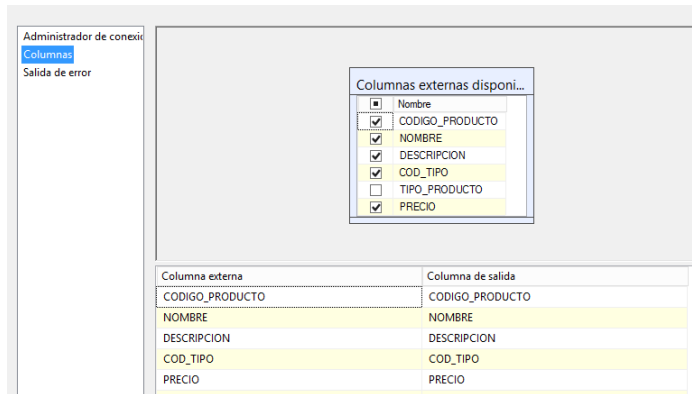


Ilustración 039. ETLs; Configuración Origen de Datos: Administrador de Conexiones



Una vez configurada la entrada de datos, aparecerán las columnas que contenga la hoja de cálculo escogida. Aquí, simplemente elegimos las columnas que nos sean necesarias en la ETL. Las columnas de entrada, pueden no ser las mismas que las de salida.

Ilustración 040. ETLs; Configuración Origen de datos: Columnas



Salida de Errores : esta herramienta no nos es estrictamente necesaria, pero nos es útil si en un momento dado perdemos algún registro por el camino. La salida de errores nos permite almacenar los errores que se puedan dar a lo largo de la ETL en un fichero plano, de esta forma, al abrir este fichero, se visualizarán los registros con errores. En nuestro proyecto, pondremos salidas de errores en los orígenes de datos y en los destinos principalmente.

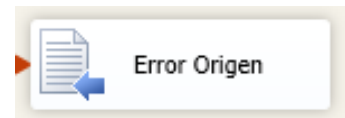
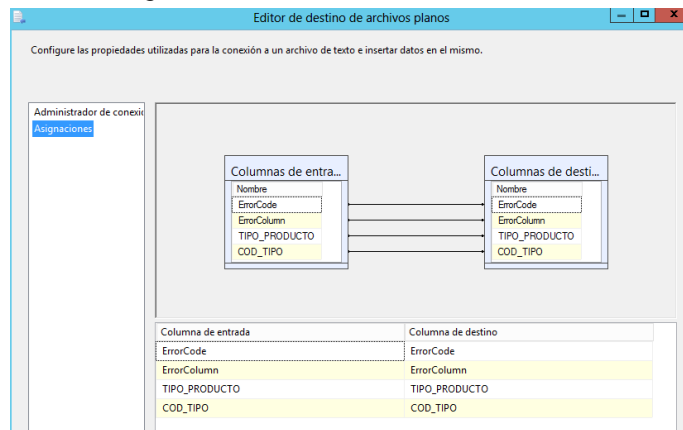


Ilustración 041. ETLs; Configuración Salida de Errores



Conversiones de datos: con esta herramienta convertimos los tipos de datos que vienen del documento excel al tipo que necesitemos de salida, para que concuerden los tipos y no haya fallos en la carga de datos. Elegimos los datos de los que queremos cambiar el tipo, y solo tenemos que elegir un formato nuevo entre todas las posibilidades que te ofrece.

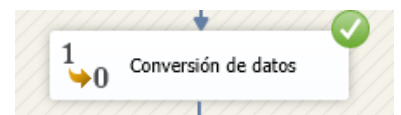
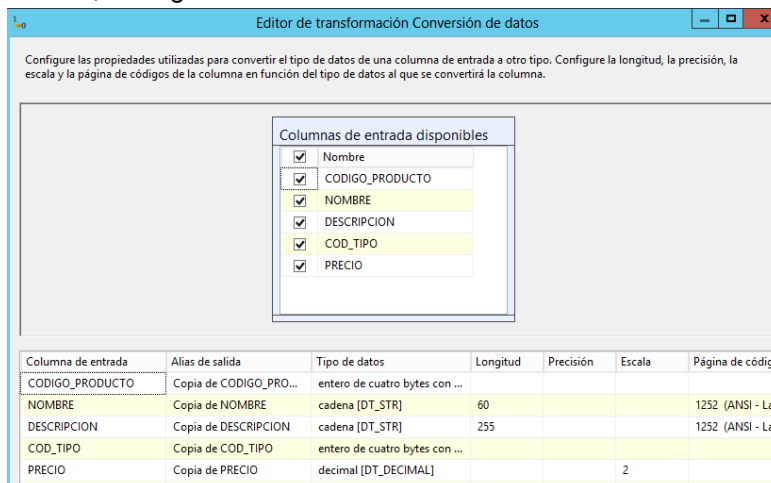


Ilustración 042. ETLs; Configuración Conversiones de Datos



Búsquedas: La búsqueda establece un tipo de conexión con la base de datos en la que se van a cargar los datos. Una vez instaurada esta conexión, establece una comparación de campos, entre los datos que le llegan desde el origen y los de destino, creando, en su caso, un identificador único para el registro.

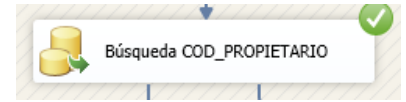
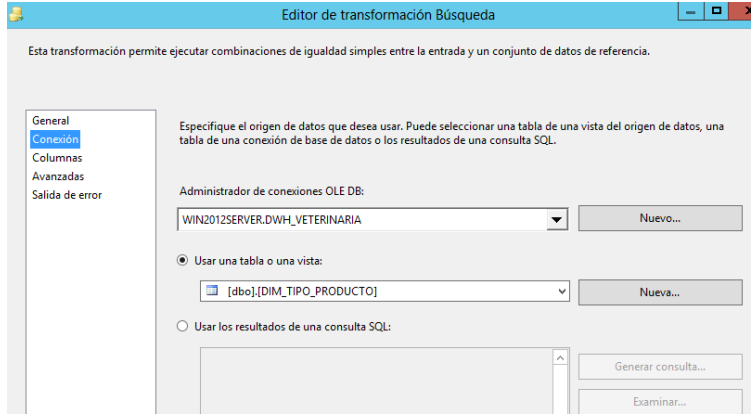
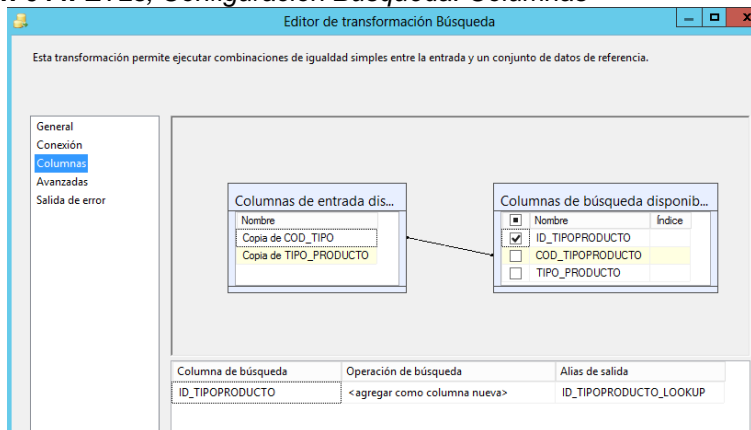


Ilustración 043. ETLs; Configuración Búsqueda: Conexión



En el caso del ejemplo, comparación de códigos y generación de identificadores:

Ilustración 044. ETLs; Configuración Búsqueda: Columnas



Columnas derivadas: Con esta herramienta, podemos introducir un registro nuevo con un valor establecido. En el proyecto, lo utilizaremos para generar los datos no informados.

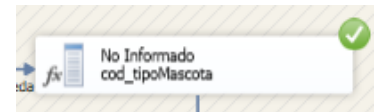
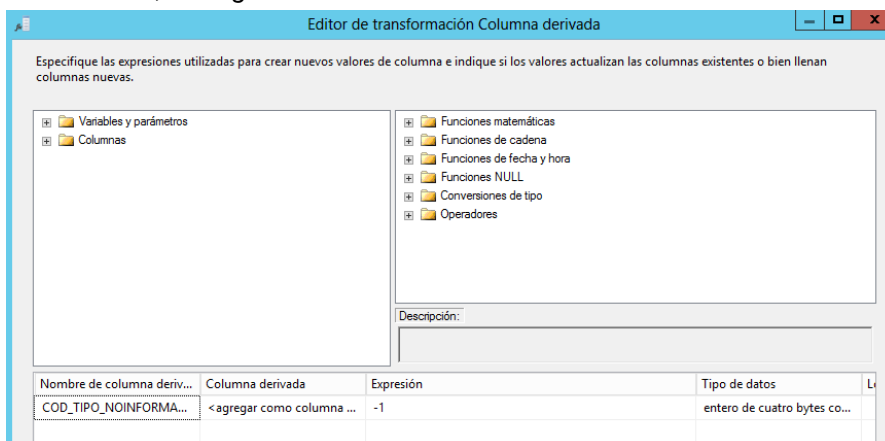


Ilustración 045. ETLs; Configuración Columna Derivada



Uniones: esta herramienta nos sirve para unir varios datos. En el caso de la intrusión de un dato no informado, es necesario que se inserte con el resto de registros, por ello lo debemos unir al flujo de la ETL.

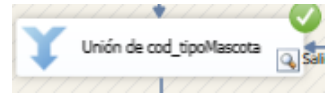
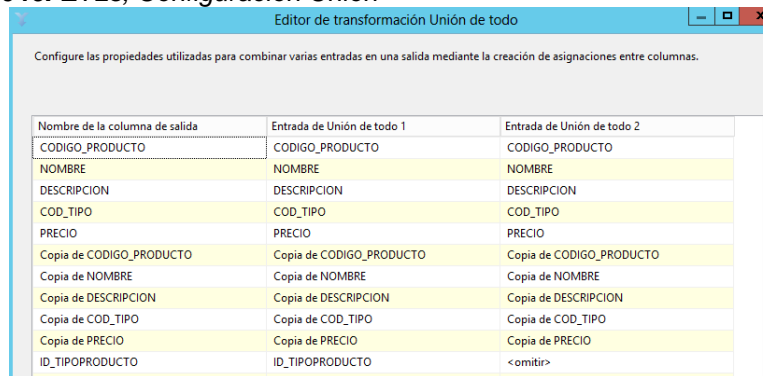


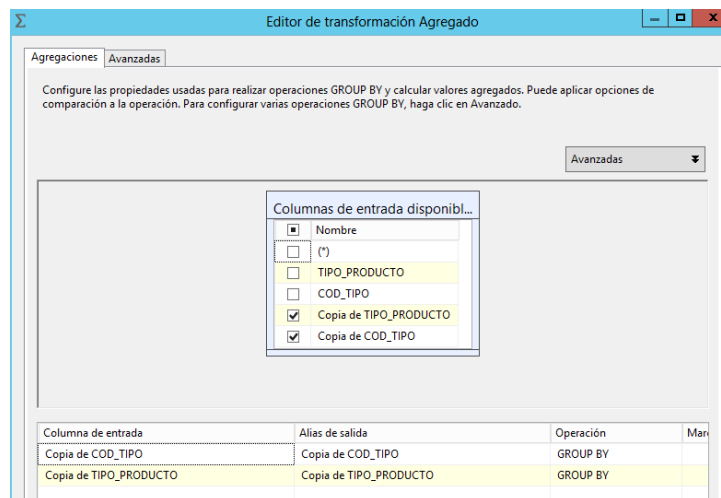
Ilustración 046. ETLs; Configuración Unión



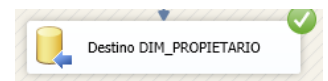
Agregaciones: esta herramienta es capaz de realizar agrupaciones de datos iguales y suma de importes. Es muy útil en caso de registros duplicados, como en el caso de los tipos de productos, clases de consultas o cargo de los empleados, ya que elimina estos duplicados.



Ilustración 047. ETLs; Configuración Agregación

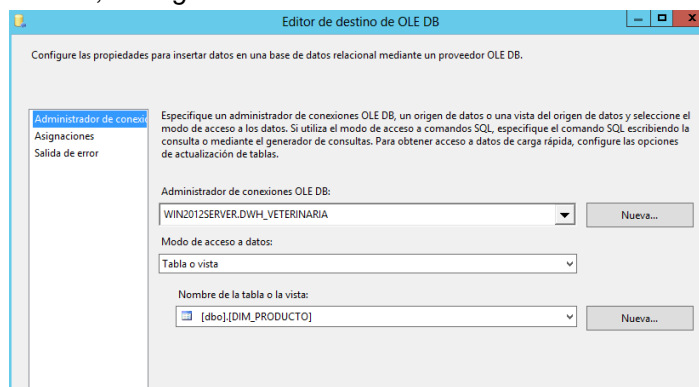


Destinos: Para finalizar con la ETL se deben cargar los datos en la tabla que proceda en cada caso. Para ello esta herramienta nos permite configurar una salida para los registros.



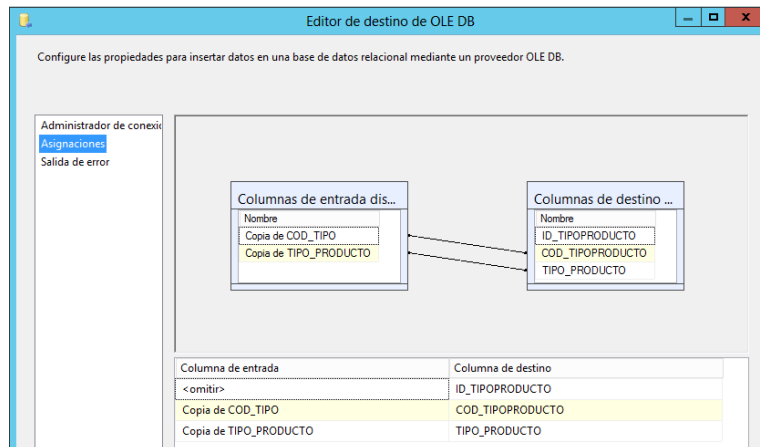
La configuración de esta herramienta es muy similar a la de Origen.

Ilustración 048. ETLs; Configuración Destino: Administrador de Conexión



Se deben asignar cada columna con su destino en la tabla de la base de datos:

Ilustración 049. ETLs; Configuración Destino: Asignaciones



Update: Esta herramienta es importante, si una tabla está cargada con registros, actualiza los datos existentes con los nuevos. Las dimensiones tienen estrategia de carga de Insert/update, por ello es necesario que en las dimensiones aparezcan las dos herramientas: la de destino y la de actualización.

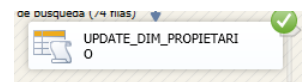
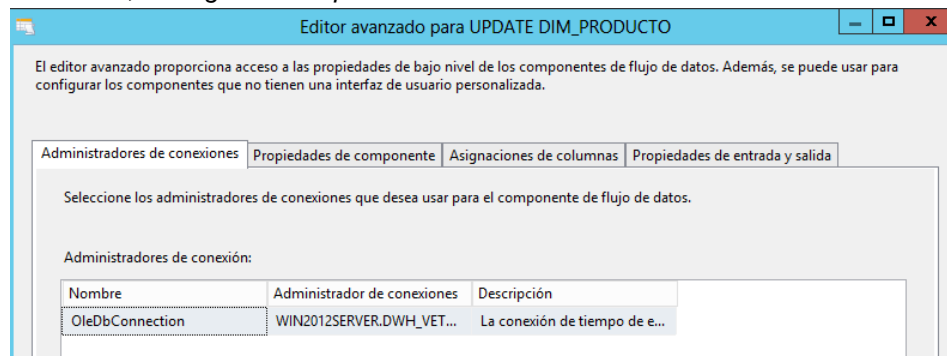


Ilustración 050. ETLs; Configuración Update



- **Modelo de Datos: Volumetría de Carga**

La volumetría total de la carga de todas las ETLs es de unos 6.500 registros en total, de ellos, unos 300 son de las dimensiones y el resto corresponden a las tablas de hechos. Esta es la cantidad de registros que finalmente tendrá que soportar nuestra base de datos. Pudiéndose incrementar notablemente con el paso del tiempo, ya que la clínica veterinaria tiene consultas y realiza ventas de productos a diario.

6.2.2- Dimensiones

Como se explicará en este apartado, el proceso de implementación y más concretamente la carga de información del operacional en el Data Warehouse, debe tener un orden en concreto. Se debe comenzar con las Dimensiones para poder garantizar que las claves foráneas funcionarán correctamente al realizar las referencias a otras tablas.

Así pues, comenzaremos con la información genérica a todas las dimensiones y las herramientas comunes que se utilizarán para la creación de las ETLs. Posteriormente analizaremos una a una todas las dimensiones. Se definirán los orígenes de información y su destino en cada caso.

• Información General para Todas las Dimensiones

En este punto se explicarán el desarrollo común para todas las dimensiones de forma genérica, detallando una a una las herramientas utilizadas en la creación de las ETLs y su configuración.

○ Fuentes de Información

Tabla 077. Fuentes de Información Dimensiones; General

Nombre del Archivo	Dim_sin_97	
Tipo Archivo	Documento Hoja de Cálculo	
Descripción	Colección de todos los datos que tiene la clínica veterinaria.	
Estructura		
Hoja de cálculo	Columnas	Descripción
MASCOTAS	Cod_mascota, Nombre, Raza, Sexo, Peso, cod_propietario, Propietario, Dirección, Teléfono, fecha_nacimiento, fecha_defuncion, cod_tipo y tipo mascota	Contiene información referente a las mascotas que son clientes de la clínica veterinaria. Tiene también información de los propietarios de cada mascota, así como el tipo de mascota del que se trata.
EMPLEADO	Cod_empleado, nombre, apellidos, dirección, cod_cargo y cargo	Todos los datos necesarios para identificar cada empleado y el puesto que tiene en la empresa.
CONSULTA	Cod_Consulta, Descripción, Precio, Cod_Clase y Clase_Consulta	Define cualquier consulta que se puede llevar a cabo en la clínica, junto con su tipo (más genérico) y su descripción
PRODUCTO	Cod_Producto, Nombre, Descripción, Cod_Tipo, Tipo_Producto y Precio	En esta hoja se tienen todos los productos que se venden en la clínica.

○ **Destinos de Información**

El destino de toda la información concerniente a la clínica veterinaria, es una base de datos denominada DWH_VETERINARIA. La estructura de esta base de datos ha sido creada desde el inicio.

Tabla 078. Destinos de Información Dimensiones; General

Base de Datos	DWH_VETERINARIA	
Tipo	Base de Datos	
Descripción	Base de datos con todas las tablas necesarias para estructurar adecuadamente las dimensiones de la clínica veterinaria y para un manejo eficaz de sus datos.	
Estructura		
Tablas	Columnas	Descripción
DIM_TIEMPO	Id_Mesano, Ano, Mes, Desc_Mes, Mes_Ant, Ano_Ant, Mesant_Mismoano y Mes_Anoant	Dimensión que nos sirve para establecer periodos temporales. Realizada y cargada completamente desde 0.
DIM_TIPOMASCOTA	Id_TipoMascota, Cod_tipoMascota y TipoMascota	Identifica a los diferentes tipos de mascotas. Por ejemplo: aves, perros, gatos...
DIM_PROPIETARIO	Id_propietario, Cod_propietario, Nombre, Dirección y Telefono	Recoge los datos concernientes a los propietarios de las mascotas.
DIM_MASCOTA	Id_mascota, Cod_mascota, Nombre, Raza, Sexo, Peso, Id_propietario, Fecha_nac, Fecha_defuncion, ,Id_TipoMascota	Almacena la información referida a las mascotas. Las mascotas tienen un tipo y un propietario.
DIM_CLASECONSULTA	Id_ClaseConsulta, Cod_claseConsulta y Descripción	Descripción global del tipo de las consultas que se pueden llevar a cabo en la clínica veterinaria.
DIM_CONSULTA	Id_Consulta, Cod_consulta, Descripción, Precio, Id_TipoConsulta	Definición de cada consulta, cada una tiene referencia a su tipo.
DIM_CARGO	Id_cargo, Cod_Cargo, Cargo	Designación del cargo de cada empleado de la clínica.
DIM_EMPLEADO	Id_Empleado, Cod_Empleado, Nombre, Apellidos, Telefono, Id_cargo	Identifica a cada empleado y se le asigna a cada uno el cargo que tengan dentro de la empresa.
DIM_TIPO_PRODUCTO	Id_TipoProducto, Cod_TipoProducto, Descripción	En esta dimensión se enumeran los tipos de productos de forma genérica
DIM_PRODUCTO	Id_Producto, Cod_Producto, Nombre, Descripción, Precio, Id_TipoProducto	Se especifican los todos los productos y al tipo de producto al que pertenecen

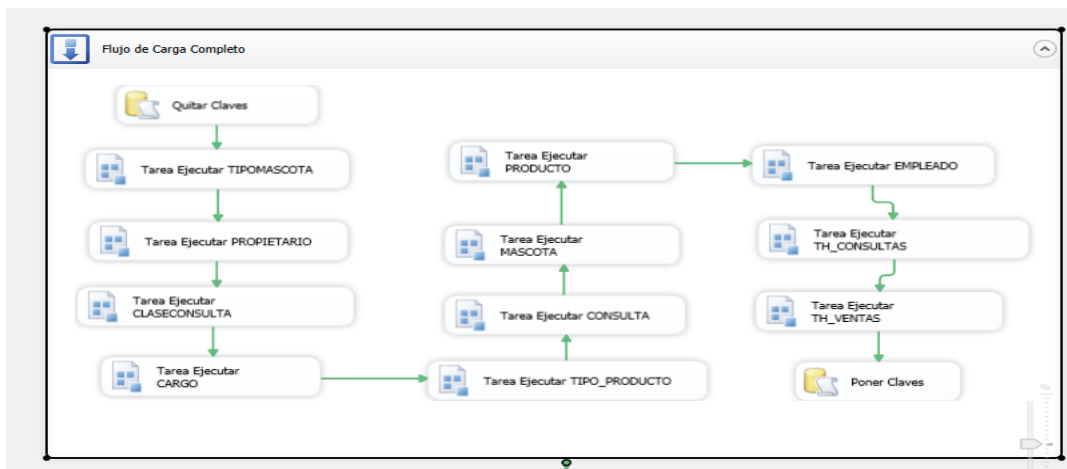
▪ *Flujo de Carga Completo*

Se realizará una ETL global de forma que, desde un solo flujo de control, se ejecuten las demás ETLs, cargando todas las dimensiones. El flujo de control debe llevar un orden determinado, comenzando con las dimensiones que contengan algún tipo de jerarquía respecto de otra y, dentro de éstas, por las que tengan mayor grano.

De no hacerse de este modo, se crearían errores a la hora de cargar las diferentes tablas del Data Warehouse, debido a las claves foráneas de la base de datos. Una columna no puede hacer referencia a otra que esté vacía.

Comenzamos creando un contenedor, y dentro de él, tantas tareas de “ejecutar paquete” como ETLs tengamos:

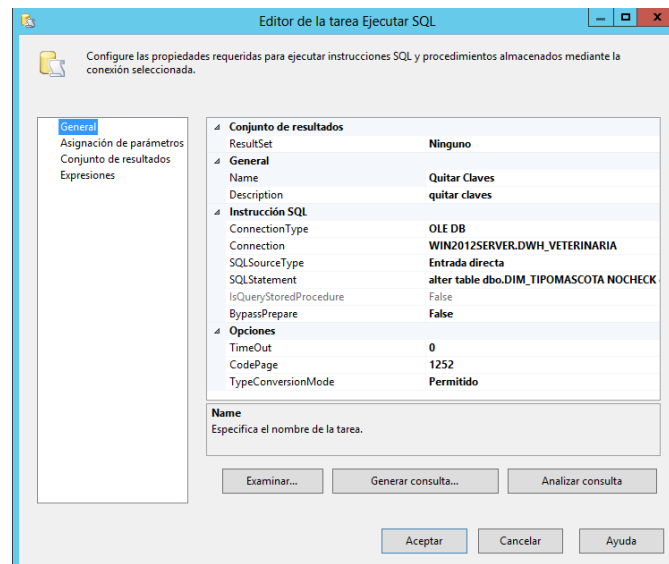
Ilustración 051. ETLs; Flujo de Carga Completo



Como se vé en la ilustración, antes de ejecutar las ETLs, hemos considerado quitar, mediante una sentencia sql, las claves foráneas y las primary keys para que no supongan problema en los datos a la hora de introducir información en la base de datos. En principio, las referencias cruzadas no debieran suponer un problema si la carga de datos se realiza con el orden establecido. Posteriormente en cada tarea, configuramos para que ejecute la ETL correspondiente y finalmente activaremos las claves de referencia nuevamente.

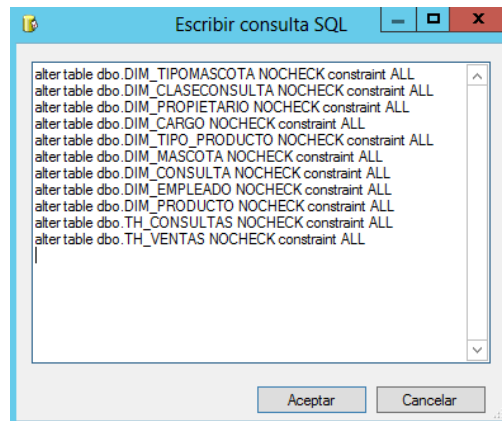
Primero quitamos las restricciones de las claves:

Ilustración 052. ETLs; Configuración Quitar Claves



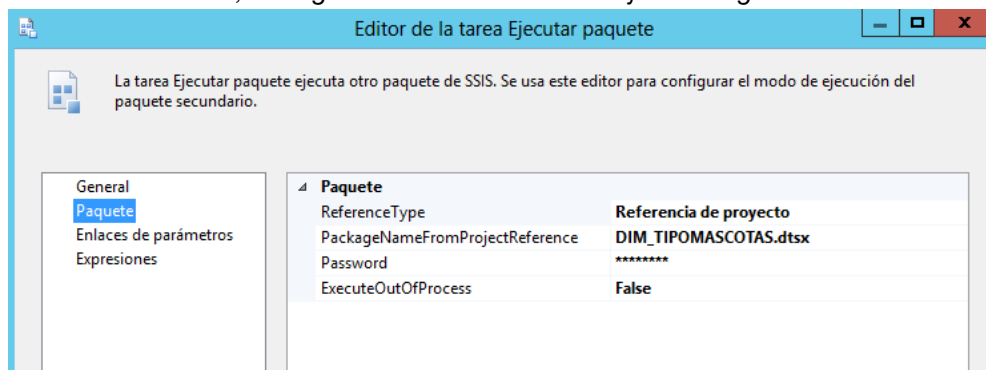
La sentencia utilizada para quitar las restricciones de claves es la siguiente:

Ilustración 053. ETLs; Configuración Quitar Claves; Consulta SQL



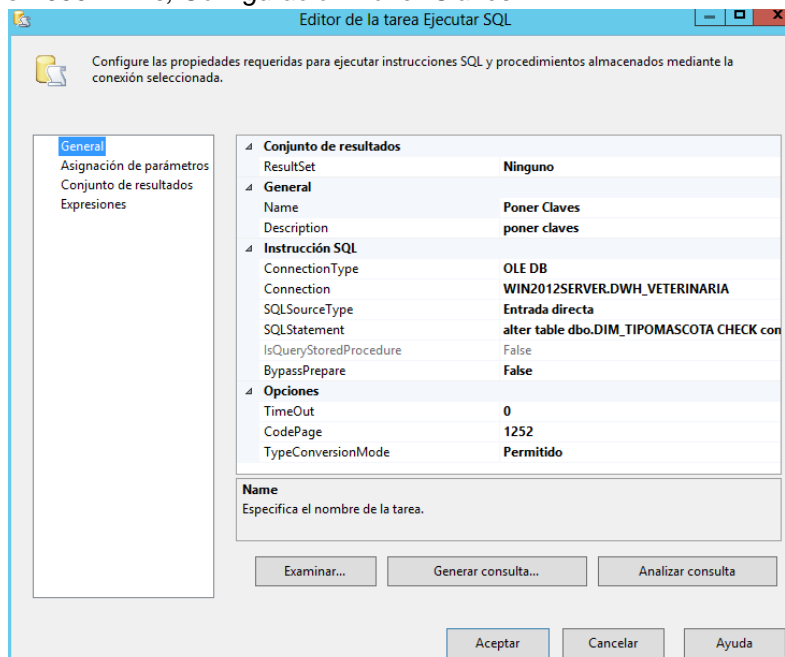
Después de quitar las claves de todas las tablas, ejecutamos todas las dimensiones en orden. Por lo que configuramos las tareas del flujo de carga como se ve en la figura.

Ilustración 054. ETLs; Configuración de Tareas del Flujo de Carga



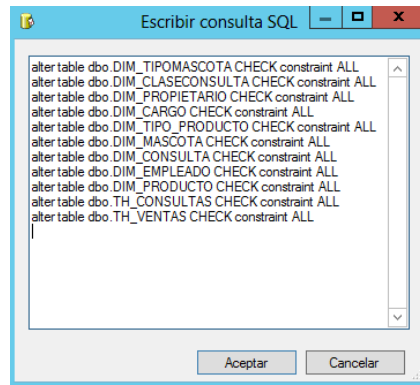
Por último, volveremos a activar las restricciones de las claves de referencia en la base de datos.

Ilustración 055. ETLs; Configuración Poner Claves



La sentencia utilizada para activar nuevamente las restricciones de claves es la siguiente:

Ilustración 056. ETLs; Configuración Poner Claves; Consulta SQL



Una vez tengamos todas las ETLs creadas, y el flujo de carga bien configurado, si ejecutamos este flujo de carga global, automáticamente ejecutamos también todas las ETLs a las que reseñamos desde aquí.

▪ **Tratamiento de los “No Informados”**

Antes de la carga de las dimensiones referentes a tipos o clases, es decir, las dimensiones que tengan alguna jerarquía, es necesario realizar una inserción: la inserción de los no informados.

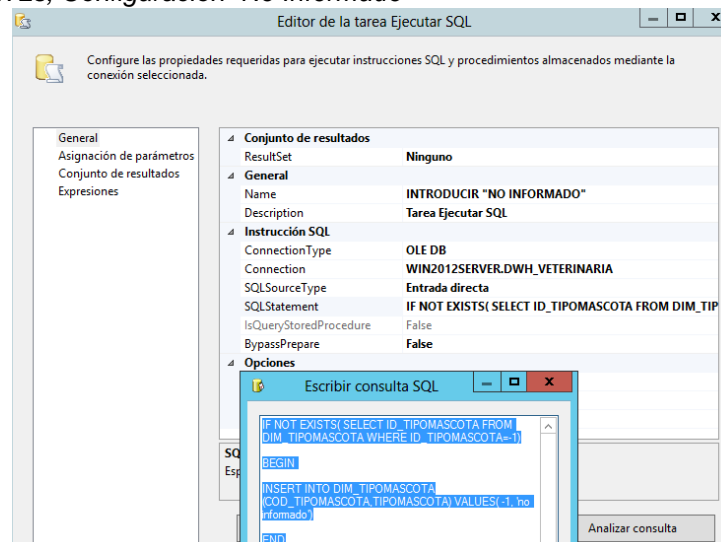
El caso “no informado” se utilizará en las ocasiones en las cuales falte cierta información en algún campo de la tabla. Como puede ser el caso de un teléfono o una dirección.

Ilustración 057. ETLs; Flujo de Control



Para la implantación del caso “no informado” lo configuramos de la siguiente manera:

Ilustración 058. ETLs; Configuración “No Informado”



- **Mapeos Origen/Destino**

Tabla 079. Mapeos de Origen/Destino; General

ORIGEN (Excel)	DESTINOS (DataWereHouse)
TIEMPO	-DWH_VETERINARIA.dbo.DIM_TIEMPO
MASCOTAS	-DWH_VETERINARIA.dbo.DIM_PROPIETARIO -DWH_VETERINARIA.dbo.DIM_TIPOMASCOTA -DWH_VETERINARIA.dbo.DIM_MASCOTA
EMPLEADO	-DWH_VETERINARIA.dbo.DIM_CARGO -DWH_VETERINARIA.dbo.DIM_EMPLEADO
CONSULTA	-DWH_VETERINARIA.dbo.DIM_CLASECONSULTA -DWH_VETERINARIA.dbo.DIM_CONSULTA
PRODUCTO	-DWH_VETERINARIA.dbo.DIM_TIPO_PRODUCTO -DWH_VETERINARIA.dbo.DIM_PRODUCTO

- **DIM_TIEMPO**

- **Fuentes de Información**

Esta dimensión difiere de las demás en que la clínica veterinaria no nos la facilita, pues no les es necesario para su funcionamiento diario.

La dimensión de tiempo la crearemos para el buen funcionamiento de nuestro modelo. Es necesaria para la fase de reporting en la que se visualizan los datos.

Por lo tanto, la fuente de información no está proveída por la empresa, sino que la crearemos nosotros.

Tabla 080. Fuentes de Información; DIM_TIEMPO

Nombre del Archivo	Dim_sin_97.xls	
Nombre	DIM_TIEMPO	
Tipo	Hoja de cálculo: "DIM_TIEMPO"	
Descripción	Colección de las distintas fechas que se pueden abarcar en las tablas de hechos.	
Estructura		
Columna	Formato	Descripción
ID_MESANO	Numérico	Representa el mes y año de una fecha concreta
ANO	Numérico	Representa el año
MES	Numérico	Constituye el mes
DESCRIPCION_MES	Alfanumérico	Describe el mes en letra.
MES_ANTERIOR	Numérico	Número del mes anterior
ANO_ANTERIOR	Numérico	Año anterior
MESANTERIORMISMOAÑO	Numérico	Fecha con el mes anterior y el año actual
MESYANOANTERIOR	Numérico	Fecha con el mismo mes del año anterior

- **Destino/s de Información**

Tabla 081. Destinos de Información; DIM_TIEMPO

Base de Datos	DWH_VETERINARIA			
Nombre	DIM_TIEMPO			
Tipo	Tabla			
Descripción	Colección de datos temporales.			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_MESANO	Numérico	6	PK	Identificador único que define el mes y el año
ANO	Numérico	4		Descripción numérica del año
MES	Numérico	2		Número de mes
DESC_MES	Alfanumérico	15		Definición del mes del año
MES_ANT	Numérico	2		Número de mes anterior
ANO_ANT	Numérico	4		Declaración del año anterior
MESANT_MISMOANO	Numérico	6		Identifica el mes anterior del mismo año
MISMOMES_ANOANT	Numérico	6		Define el mismo mes pero del año anterior

Tabla 082. Destinos de Información; Primary Key; DIM_TIEMPO

Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_TIEMPO	PK	(ID_MESANO)

- **Detalles del Proceso**
 - **Lógica de Carga**

En cuanto a la lógica de carga de esta dimensión temporal es importante reseñar que, como la empresa no utiliza estos datos, no nos los puede proporcionar. Esta dimensión es creada por y para el proyecto. No es necesario realizar la carga mediante ETL, la cargamos directamente en la base de datos mediante el comando “Insert into”.

- **Mapeos Origen/Destino**

Tabla 083. Mapeos de Origen/Destino; DIM_TIEMPO

ORIGEN (Excel)	DESTINO (DataWereHouse)
DIM_TIEMPO.ID_MESANO	DWH_VETERINARIA.dbo.DIM_TIEMPO.ID_MESANO
DIM_TIEMPO.ANO	DWH_VETERINARIA.dbo.DIM_TIEMPO.ANO
DIM_TIEMPO.MES	DWH_VETERINARIA.dbo.DIM_TIEMPO.MES
DIM_TIEMPO.DESCRIPCION_MES	DWH_VETERINARIA.dbo.DIM_TIEMPO.DESC_MES
DIM_TIEMPO.MES_ANTERIOR	DWH_VETERINARIA.dbo.DIM_TIEMPO.MES_ANT
DIM_TIEMPO.ANO_ANTERIOR	DWH_VETERINARIA.dbo.DIM_TIEMPO.ANO_ANT
DIM_TIEMPO.MESANTERIORMISMOAÑO	DWH_VETERINARIA.dbo.DIM_TIEMPO.MESANT_MISMOANO
DIM_TIEMPO.MESYANOANTERIOR	DWH_VETERINARIA.dbo.DIM_TIEMPO.MES_ANOANT

- **Modelo de Datos: Volumetría de Carga**

Como la clínica veterinaria tiene histórico de dos años atrás, es decir de unos 27 meses. Los datos que debemos contemplar en la dimensión de tiempo creada deben ser los correspondientes a estos 27 meses. Los registros, por tanto, ascienden a 821.

- **DIM_PROPIETARIO**

- **Fuentes de Información**

Como origen de información en la DIM_PROPIETARIO, utilizaremos la hoja de cálculo de “Mascotas”. Como la empresa no es muy experta en bases de datos, la información concerniente a los propietarios de las mascotas, la tiene incluida en la misma hoja de cálculo donde se encuentra la información de las mascotas.

Tabla 084. Fuentes de Información; DIM_PROPIETARIO

Nombre del Archivo	Dim_sin_97.xls	
Nombre	MASCOTAS	
Tipo	Hoja de cálculo: “Mascotas”	
Descripción	Colección de los distintos propietarios de las mascotas de la clínica veterinaria.	
Estructura		
Columna	Formato	Descripción
COD_PROPIETARIO	N Numérico	Código identificativo, impuesto por la empresa a cada propietario de mascota.
PROPIETARIO	A Alfanumérico	Nombre y apellidos del propietario de una mascota.
DIRECCIÓN	A Alfanumérico	Calle, número y ciudad donde reside el propietario de una mascota.
TELÉFONO	N Numérico	Número de teléfono de contacto del propietario.

○ **Destino/s de Información**

Tabla 085. Destinos de Información; DIM_PROPIETARIO

Base de Datos	DWH_VETERINARIA			
Nombre	DIM_PROPIETARIO			
Tipo	Tabla			
Descripción	Agrupación de los distintos propietarios de las mascotas atendidas en la clínica.			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_PROPIETARIO	Numérico	6	PK	Identificador autoincremental y único por propietario de mascota.
COD_PROPIETARIO	Numérico	6		Código de propietario establecido por la empresa contratante.
NOMBRE	Alfanumérico	100		Nombre y apellidos del propietario de la mascota.
DIRECCIÓN	Alfanumérico	150		Dirección completa del propietario de la mascota
TELÉFONO	Numérico	9		Número de teléfono del propietario de la mascota.

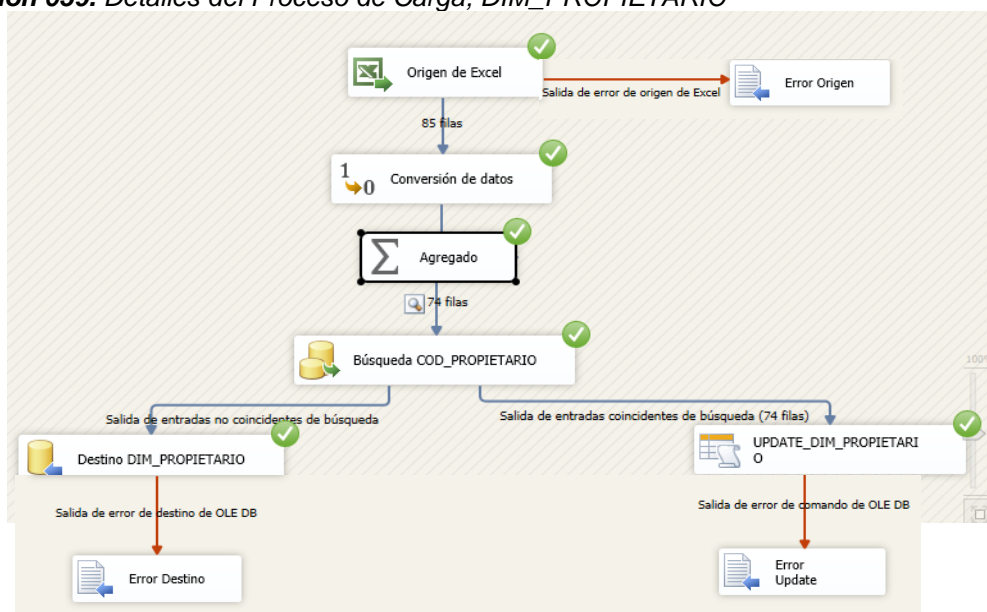
Tabla 086. Destinos de Información; Primary Key; DIM_PROPIETARIO

Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_PROPIETARIO	PK	(ID_PROPIETARIO)

○ **Detalles del Proceso**
 ▪ **Lógica de Carga**

La ETL creada para cargar la DIM_PROPIETARIO es la siguiente:

Ilustración 059. Detalles del Proceso de Carga; DIM_PROPIETARIO



1. *Origen de Excel*: Configuramos para que el origen de datos de nuestra ETL sea el documento Excel facilitado por la empresa y, en este caso, la hoja de cálculo de “Mascotas”.
2. *Conversión de datos*: Es necesario para convertir los tipos de datos y que no dé errores a la hora de insertar en la base de datos.
3. *Agregado*: Agrupa por cod_propietario, de esta forma, no habrá duplicados en la DIM_PROPIETARIO
4. *Búsqueda COD_PROPIETARIO*: De esta forma podemos crear los identificadores autonumericos para la base de datos.
5. *Destino DIM_PROPIETARIO*: Carga la DIM_PROPIETARIO de la base de datos de salida cuando mientras está permanece vacía.
6. *UPDATE_DIM_PROPIETARIO*: Actualiza la tabla DIM_PROPIETARIO si esta contiene información previa.

Como se vé en la imagen, en un principio tenemos un total de 85 filas, pero tras el agregado, nos quedan 74, esto es porque agrupa los propietarios repetidos en una sola fila, así pues, tendremos un id único por propietario.

▪ **Mapeos Origen/Destino**

Tabla 087. Mapeos Origen/Destino; DIM_PROPIETARIO

ORIGEN (Excel)	DESTINO (DataWereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_PROPIETARIO.ID_PROPIETARIO
MASCOTAS.COD_PROPIETARIO	DWH_VETERINARIA.dbo.DIM_PROPIETARIO.COD_PROPIETARIO
MASCOTAS.NOMBRE	DWH_VETERINARIA.dbo.DIM_PROPIETARIO.NOMBRE
MASCOTAS.DIRECCIÓN	DWH_VETERINARIA.dbo.DIM_PROPIETARIO.DIRECCIÓN
MASCOTAS.TELÉFONO	DWH_VETERINARIA.dbo.DIM_PROPIETARIO.TELÉFONO

○ **Modelo de Datos: Volumetría de Carga**

La volumetría de carga de esta dimensión son 74 registros de los 85 existentes en la información de origen. Esta reducción es debido a que en el caso de los propietarios, la clínica veterinaria, tiene sus datos en conjunto con el de las mascotas, esto es, como hay 85 mascotas diferentes, tienen 85 propietarios.

Esta situación no es óptima, ya que en algunos casos, un propietario puede tener varias mascotas. En este escenario, los datos de propietarios se duplican.

Nosotros quitamos estos duplicados mediante el proceso ETL.

- **DIM_TIPOMASCOTA**
 - **Fuentes de Información**

La clínica veterinaria, define el tipo de mascotas dentro de la hoja Excel de las Mascotas. Por tanto, la fuente de información para esta dimensión se encuentra en la hoja de cálculo denominada “Mascotas”

Tabla 088. Fuentes de Información; DIM_TIPOMASCOTA

Nombre del Archivo	Dim_sin_97.xls	
Nombre	MASCOTAS	
Tipo	Hoja de cálculo: “Mascotas”	
Descripción	Determina los diferentes tipos de mascotas existentes en la clínica.	
Estructura		
Columna	Formato	Descripción
COD_TIPOMASCOTA	Numérico	Código establecido por la empresa para diferencial los tipos de mascotas existentes.
TIPOMASCOTA	Carácter	Descripción de cada tipo de mascota: ave, perro, gato...

- **Destino/s de Información**

Tabla 089. Destinos de Información; DIM_TIPOMASCOTA

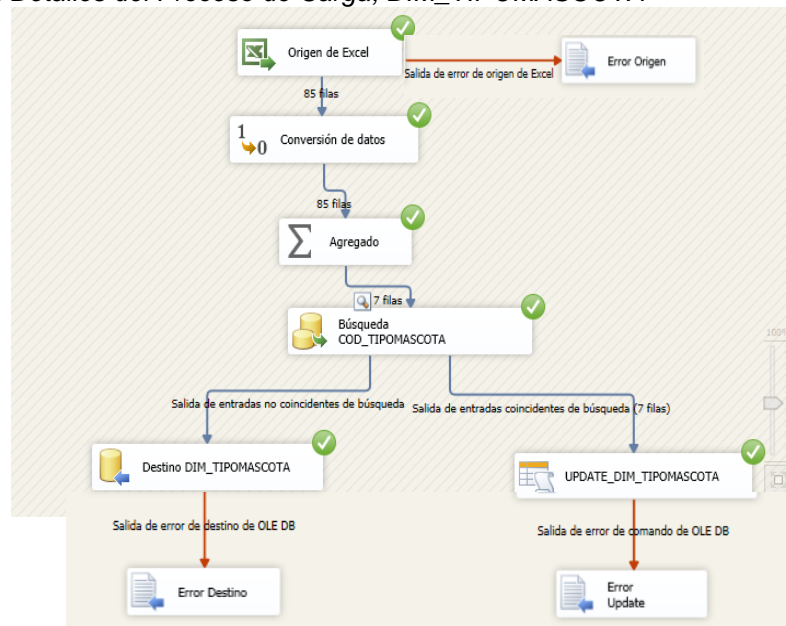
Base de Datos	DWH_VETERINARIA			
Nombre	DIM_TIPOMASCOTA			
Tipo	Tabla			
Descripción	Diferencia a cada mascota por su tipo, es un nivel más genérico.			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_TIPOMASCOTA	Numérico	6	PK	Identificador único y autoincremental establecido para diferenciar los tipos de mascotas
COD_TIPOMASCOTA	Numérico	6		Código importado de la documentación de la empresa, utilizado para diferenciar los tipos de mascotas
TIPOMASCOTA	Alfanumérico	50		Descripción del tipo de mascota.

Tabla 090. Destinos de Información; Primary Key; DIM_TIPOMASCOTA

Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_TIPOMASCOTA	PK	(ID_TIPOMASCOTA)

- **Detalles del Proceso**
 - **Lógica de Carga**

La ETL construida para cargar la DIM_TIPOMASCOTA es la siguiente:

Ilustración 060. Detalles del Proceso de Carga; DIM_TIPOMASCOTA

1. *Origen de Excel*: Configura que el origen de información para coger los datos de la hoja de cálculo denominada “Mascotas”
2. *Conversión de datos*: Necesario para que los tipos de datos no den problemas a la hora de cargarlos en la base de datos.
3. *Agregado*: Agrupa la información para que no existan duplicados y el id sea único para los diferentes tipos de mascotas.
4. *Búsqueda COD_TIPOMASCOTA*: Introduce un id único por fila correspondiente a cada tipo de mascota.
5. *Destino DIM_TIPOMASCOTA*: Si la dimensión a cargar está vacía, la llena con los datos oportunos, según su configuración.
6. *UPDATE_DIM_TIPOMASCOTA*: Si la DIM_TIPOMASCOTA de la base de datos de salida, ya contiene datos, los actualiza.

- **Mapeos Origen/Destino**

Tabla 091. Mapeos Origen/Destino; DIM_TIPOMASCOTA

ORIGEN (Excel)	DESTINO (Data WereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_TIPOMASCOTA.ID_TIPOMASCOTA
MASCOTAS.COD_TIPOMASCOTA	DWH_VETERINARIA.dbo.DIM_TIPOMASCOTA.COD_TIPOMASCOTA
MASCOTAS.TIPOMASCOTA	DWH_VETERINARIA.dbo.DIM_TIPOMASCOTA.TIPOMASCOTA

- **Modelo de Datos: Volumetría de Carga**

En el caso del tipo de mascota, se recogen los datos de la hoja de cálculo de “Mascotas”, en donde existen 85 registros que corresponden a 85 mascotas diferentes. Pero el tipo de las mascotas, como es lógico, se repite. Por lo que, de estos 85 tipos de mascotas, en realidad, al quitar los duplicados mediante la ETL, se nos quedan en 7 tipos de mascotas diferentes.

- **DIM_MASCOTA**
 - **Fuentes de Información**

Todos los datos referentes a las mascotas que realizan consultas en la clínica veterinaria la encontramos en la hoja del cálculo del Excel denominada “Mascotas”. En esta hoja de cálculo también encontramos la información de los propietarios de los animales y del tipo de animal al que pertenece.

En la DIM_MASCOTA, sólo cargaremos la información correspondiente a cada mascota.

Tabla 092. Fuentes de Información; DIM_MASCOTA

Nombre del Archivo	Dim_sin_97.xls	
Nombre	MASCOTAS	
Tipo	Hoja de cálculo: “Mascotas”	
Descripción	Conjunto de todas las mascotas clientes de la clínica veterinaria.	
Estructura		
Columna	Formato	Descripción
COD_MASCOTA	Númérico	Código único por mascota.
NOMBRE	Alfanumérico	Nombre de la mascota
RAZA	Alfanumérico	Raza de cada mascota.
SEXO	Carácter (1)	(H) hembras, (M) machos.
PESO	Decimal	Cantidad en kilos. Peso del animal
FECHA_NACIMIENTO	Fecha	Fecha del nacimiento del animal
FECHA_DEFUNCION	Fecha	En su caso, fecha de defunción de la mascota
CODIGO_PROPIETARIO	Númérico	Refiere al propietario de la mascota.
CODIGO_TIPOMASCOTA	Númérico	Refiere al tipo de mascota de que se trata.

○ **Destino/s de Información**

Tabla 093. Destinos de Información; DIM_MASCOTA

Base de Datos	DWH_VETERINARIA			
Nombre	DIM_MASCOTA			
Tipo	Tabla			
Descripción	Diferencia a cada mascota.			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_MASCOTA	Numérico	6	PK	Identificador único y autoincremental que identifica de manera única a cada mascota
COD_MASCOTA	Numérico	6		Código identificador de la mascota. Utilizado por la empresa.
NOMBRE	Alfanumérico	50		Nombre de la mascota
RAZA	Alfanumérico	50		Raza del animal.
SEXO	Carácter	1		(H) hembras, (M) machos
PESO	Decimal	5,2		Peso en kilos del animal.
FECHA_NACIMIENTO	Fecha	10		Fecha de nacimiento del animal
FECHA_DEFUNCION	Fecha	10		En su caso, fecha de fallecimiento de la mascota.
ID_TIPOMASCOTA	Numérico	6	FK	Hace referencia al tipo de mascota que estamos tratando. Desde aquí podemos navegar a la DIM_TIPO_MASCOTA
ID_PROPIETARIO	Numérico	6	FK	Refiere al propietario de la mascota. Este identificador permite navegar hasta la

Tabla 094. Destinos de Información; Primary Key; DIM_MASCOTA

Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_MASCOTA	PK	(ID_MASCOTA)

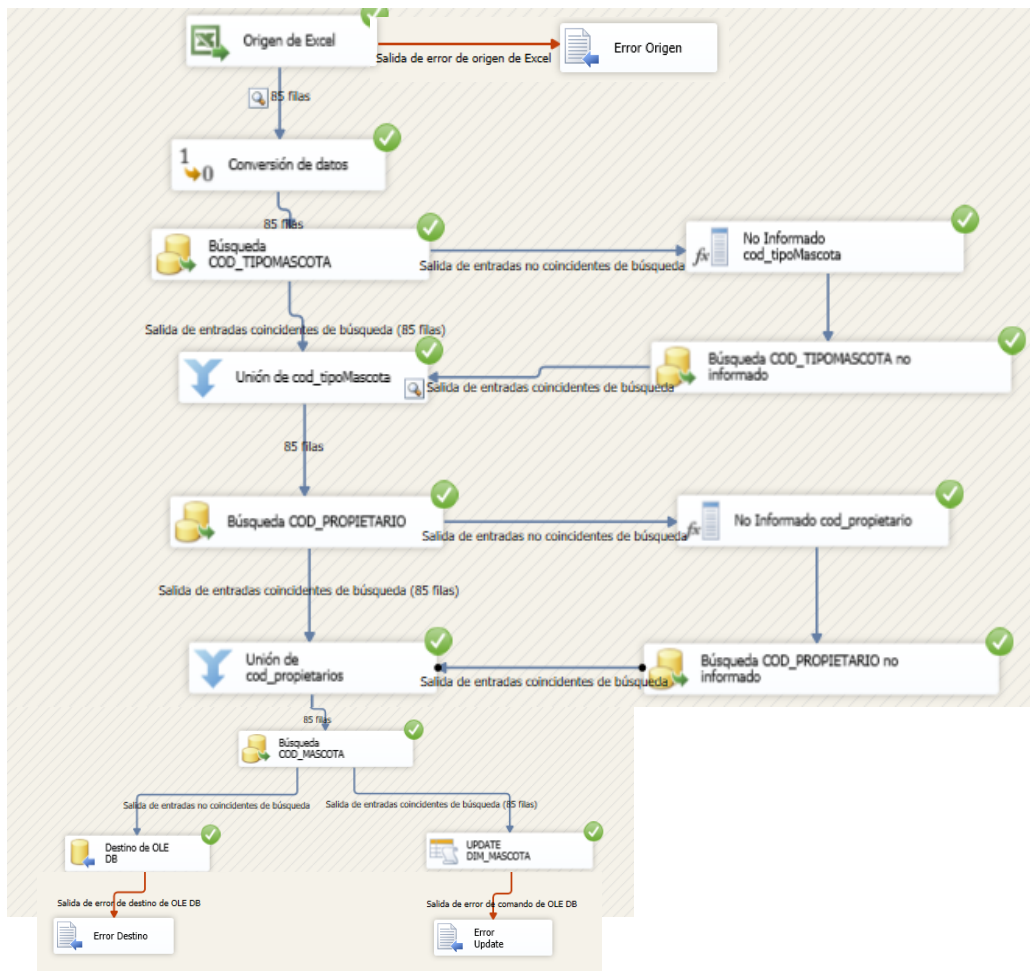
Tabla 095. Destinos de Información; Foreign Key; DIM_MASCOTA

Foreign Key		
Foreign Key	Tipo	Miembros
FK_DIM_TIPOMASCOTA	FK	(ID_TIPOMASCOTA)
FK_DIM_PROPIETARIO	FK	(ID_PROPIETARIO)

- **Detalles del Proceso**
 - **Lógica de Carga**

La ETL creada para cargar la DIM_MASCOTA es la siguiente:

Ilustración 061. Detalles del Proceso de Carga; DIM_MASCOTA



1. *Origen de Excel*: configuración para establecer el origen de datos en el documento Excel y más concretamente de la hoja de cálculo denominada “Mascotas”, que, es de donde extraemos la información referente a las mascotas.
2. *Conversión de datos*: para que los tipos de datos no causen conflicto con los que se insertarán en la tabla de dimensión establecida en la base de datos, se deben realizar conversiones en los tipos de datos.
3. *Búsqueda COD_TIPOMASCOTA*: establece los ids de los tipos de mascotas, de esta forma podemos referenciar a una mascota con su tipo.
4. *No Informado COD_TIPOMASCOTA*: si en algún caso, no existe un cod_tipomascota asignado a una mascota, la columna derivada establece un “no informado” en el campo vacío.
5. *Búsqueda COD_TIPOMASCOTA no informado*: establece un id para el caso del no informado establecido con anterioridad.

6. *Unión de COD_TIPOMASCOTA*: une todos los datos originales con el nuevo dato, si existiera, denominado “No Informado”.
7. *Búsqueda COD_PROPIETARIO*: inserta los identificadores de los propietarios para poder unir la dimensión actual de Mascotas con la de Propietarios.
8. *No Informado COD_PROPIETARIO*: establece un nuevo dato en el caso de que se encuentre alguna mascota sin id del propietario. Este nuevo dato es el “No informado”.
9. *Búsqueda COD_PROPIETARIO no informado*: asigna un id nuevo al nuevo registro “No informado”.
10. *Unión de COD_PROPIETARIO*: para unir el propietario nuevo establecido como “No Informado” con los datos que ya se tenían.
11. *Búsqueda COD_MASCOTA*: establece un identificador único para cada una de las mascotas.
12. *Destino*: configura los datos de salida, para poder cargar adecuadamente la DIM_MASCOTA.
13. *Update DIM_MASCOTA*: en el caso de que la dimensión esté cargada de datos, esta segunda opción de salida actualiza los datos.

▪ **Mapeos Origen/Destino**

Tabla 096. Mapeos Origen/Destino; DIM_MASCOTA

ORIGEN (Varios)	DESTINO (Data WereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_MASCOTA.ID_MASCOTA
MASCOTAS.COD_MASCOTA	DWH_VETERINARIA.dbo.DIM_MASCOTA.COD_MASCOTA
MASCOTAS.NOMBRE	DWH_VETERINARIA.dbo.DIM_MASCOTA.NOMBRE
MASCOTAS.RAZA	DWH_VETERINARIA.dbo.DIM_MASCOTA.RAZA
MASCOTAS.SEXO	DWH_VETERINARIA.dbo.DIM_MASCOTA.SEXO
MASCOTAS.PESO	DWH_VETERINARIA.dbo.DIM_MASCOTA.PESO
MASCOTAS.FECHA_NACIMIENTO	DWH_VETERINARIA.dbo.DIM_MASCOTA.FECHA_NACIMIENTO
MASCOTAS.FECHA_DEFUNCION	DWH_VETERINARIA.dbo.DIM_MASCOTA.FECHA_DEFUNCION
DWH_VETERINARIA.dbo.DIM_TIPOMASCOTA.ID_TIPOMASCOTA	DWH_VETERINARIA.dbo.DIM_MASCOTA.ID_TIPOMASCOTA
DWH_VETERINARIA.dbo.DIM_PROPIETARIO.ID_TIPOMASCOTA	DWH_VETERINARIA.dbo.DIM_MASCOTA.ID_PROPIETARIO

○ **Modelo de Datos: Volumetría de Carga**

Esta dimensión corresponde con todos los registros existentes en la hoja de cálculo de “Mascota”, en dicho documento hay un total de 85 registros de animales diferentes. Por lo tanto debieran recogerse 85 registros en la DIM_MASCOTA.

- **DIM_CARGO**

- **Fuentes de Información**

Los datos concernientes a esta dimensión, los encontramos en la hoja de cálculo de “Empleados” del documento Excel facilitado por la empresa. Esta información la extraemos y diferenciamos de los datos referentes a empleados como tal.

Tabla 097. Fuentes de Información; DIM_CARGO

Nombre del Archivo	Dim_sin_97.xls	
Nombre	EMPLEADOS	
Tipo	Hoja de cálculo: “Empleados”	
Descripción	Contiene todos los tipos de cargos que pueden tener los empleados de la clínica veterinaria.	
Estructura		
Columna	Formato	Descripción
CODIGO_CARGO	Numérico	Código establecido por la clínica para distinguir los cargos que puede poseer un empleado.
CARGO	Alfanumérico	Descripción del tipo de cargo.

NOTA: cada empleado sólo puede tener un cargo en la empresa.

- **Destino/s de Información**

Tabla 098. Destinos de Información; DIM_CARGO

Base de Datos	DWH_VETERINARIA			
Nombre	DIM_CARGO			
Tipo	Tabla			
Descripción	Define los cargos que puede tener un empleado de la clínica			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_CARGO	Numérico	6	PK	Identificador único y autoincremental que permite distinguir cada cargo de un empleado.
COD_CARGO	Numérico	6		Código para diferenciar el tipo de cargo de un empleado.
CARGO	Alfanumérico	50		Descripción del tipo de Cargo.

Tabla 099. Destinos de Información; Primary Key; DIM_CARGO

Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_CARGO	PK	(ID_CARGO)

- **Detalles del Proceso**
 - **Lógica de Carga**

La carga de los datos correspondiente a la DIM_CARGO se hará con la ETL:

Ilustración 062. Detalles del Proceso de Carga; DIM_CARGO



1. *Origen de Excel*: configurado para que extraiga la información del Excel de la hoja llamada “Empleados”. Las columnas a coger son las que tienen que ver con el cargo que disfrutaban los empleados en la empresa.
2. *Conversión de datos*: necesario para cambiar el tipo de datos y que no tengan conflictos con los de la DIM_MASCOTA.
3. *Agregado*: agrupa los datos iguales para que no aparezcan duplicados en la dimensión.
4. *Búsqueda COD_CARGO*: concede identificadores únicos a cada cargo. De esta forma se completan los datos de todas las columnas.
5. *Destino*: dispuesto para unir los datos de la ETL a la DIM_CARGO, para que ninguna columna quede sin asignar.
6. *Update DIM_CARGO*: en caso de que la DIM_CARGO ya contenga datos, los actualiza.

- **Mapeos Origen/Destino**

Tabla 100. Mapeos Origen/Destino; DIM_CARGO

ORIGEN (Excel)	DESTINO (Data WereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_CARGO.ID_CARGO
EMPLEADO.CODIGO_CARGO	DWH_VETERINARIA.dbo.DIM_CARGO.COD_CARGO
EMPLEADO.CARGO	DWH_VETERINARIA.dbo.DIM_CARGO.CARGO

- **Modelo de Datos: Volumetría de Carga**

Esta dimensión es cargada desde el documento “Empleados”, en esta hoja de cálculo hay 4 trabajadores, pero solo hay dos tipos de cargos: administrativo y doctor. En este caso, la ETL debe obviar los datos duplicados. Por tanto, la volumetría de carga será de 2 registros.

- **DIM_EMPLEADO**
 - **Fuentes de Información**

La fuente de información de la DIM_EMPLEADO, la tenemos en la hoja de cálculo designada "Empleado". Extraeremos, tan sólo, los datos referidos a los empleados.

Tabla 101. Fuentes de Información; DIM_EMPLEADO

Nombre del Archivo	Dim_sin_97.xls	
Nombre	EMPLEADO	
Tipo	Hoja de cálculo: "Empleado"	
Descripción	Diferencia todos los empleados que trabajan en la clínica veterinaria.	
Estructura		
Columna	Formato	Descripción
CODIGO_EMPLEADO	N Numérico	Código establecido por la empresa para discernir a los empleados
NOMBRE	N Numérico	Nombre del empleado
APELLIDOS	N Alfanumérico	Apellidos del empleado de la empresa
TELEFONO	N Numérico	Número de contacto del empleado
CODIGO_CARGO	N Numérico	Refiere al cargo que ocupa un empleado en la empresa

- **Destino/s de Información**

Tabla 102. Destinos de Información; DIM_EMPLEADO

Base de Datos	DWH_VETERINARIA			
Nombre	DIM_EMPLEADO			
Tipo	Tabla			
Descripción	Describe los datos de cada empleado.			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_EMPLEADO	N Numérico	6	PK	Identificador autoincremental y único por empleado.
COD_EMPLEADO	N Numérico	6		Código numérico, establecido por la empresa para identificar de forma única a cada trabajador.
NOMBRE	N Alfanumérico	50		Nombre del trabajador
APELLIDOS	N Alfanumérico	50		Apellidos del trabajador
TELEFONO	N Numérico	9		Teléfono de contacto del empleado
ID_CARGO	N Numérico	6	FK	Refiere a la DIM_CARGO para poder navegar y conocer el cargo del trabajador

Tabla 103. Destinos de Información; Primary Key; DIM_EMPLEADO

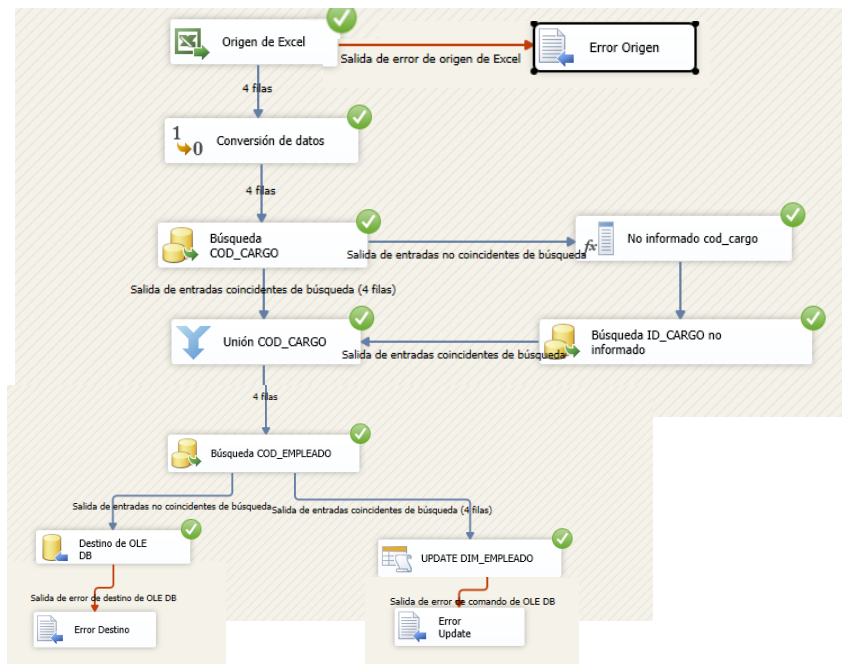
Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_EMPLEADO	PK	(ID_EMPLEADO)

Tabla 104. Destinos de Información; Foreign Key; DIM_EMPLEADO

Foreign Key		
Foreign Key	Tipo	Miembros
FK_DIM_CARGO	FK	(ID_CARGO)

- **Detalles del Proceso**
 - **Lógica de Carga**

La ETL para cargar los datos de la DIM_EMPLEADO tiene este esquema:

Ilustración 063. Detalles del Proceso de Carga; DIM_EMPLEADO

1. *Origen de Excel*: establece el origen en el Excel que nos ha facilitado la empresa. Con más exactitud en la hoja de cálculo denominada “Empleados”.
2. *Conversión de datos*: como en los casos anteriores, es necesario para que el tipo de datos que se cargan en la dimensión sean los mismos que en la estructura de la tabla en la base de datos.
3. *Búsqueda COD_CARGO*: establece un identificador para el cargo que tiene cada empleado, este código refiere a la DIM_CARGO, anteriormente cargada.
4. *No informado COD_CARGO*: añade un campo “no informado” en aquellos casos en los que no haya información de cod_cargo.

5. *Búsqueda COD_CARGO no informado*: agrega un identificador al campo añadido de “No informado”.
6. *Unión COD_CARGO*: junta los datos de la tabla con el dato nuevo de “No informado” en aquellos casos en los que sea necesario.
7. *Búsqueda COD_EMPLEADO*: instaure un identificador para cada empleado.
8. *Destino*: se configura a la tabla donde deben cargar los datos: DIM_EMPLEADO. Y se establece que dato pertenece a cada columna.
9. *Update DIM_EMPLEADO*: si la DIM_EMPLEADO está ya cargada, esta segunda opción, actualiza los datos ya cargados.

▪ Mapeos Origen/Destino

Tabla 105. Mapeos Origen/Destino; DIM_EMPLEADO

ORIGEN (Varios)	DESTINO (DataWereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_EMPLEADO.ID_EMPLEADO
EMPLEADO.CODIGO_EMPLEADO	DWH_VETERINARIA.dbo.DIM_EMPLEADO.COD_EMPLEADO
EMPLEADO.NOMBRE	DWH_VETERINARIA.dbo.DIM_EMPLEADO.NOMBRE
EMPLEADO.APELLIDOS	DWH_VETERINARIA.dbo.DIM_EMPLEADO.APELLIDOS
EMPLEADO.TELEFONO	DWH_VETERINARIA.dbo.DIM_EMPLEADO.TELEFONO
DWH_VETERINARIA.dbo.DIM_CARGO.ID_CARGO	DWH_VETERINARIA.dbo.DIM_EMPLEADO.ID_CARGO

○ *Modelo de Datos: Volumetría de Carga*

La volumetría de carga para la DIM_EMPLEADO es de 4. La empresa tiene 4 trabajadores. Esta información es cargada desde la ETL que coge la información de la hoja de cálculo de “Empleado” del documento Excel suministrado por la empresa.

- **DIM_CLASECONSULTA**
 - **Fuentes de Información**

La dimensión de clase consulta establece una relación jerárquica con la dimensión de consultas. Podemos ver esta relación en el hecho de que la fuente de información es la misma: la hoja de cálculo de “Consultas”.

Tabla 106. Fuentes de Información; DIM_CLASECONSULTA

Nombre del Archivo	Dim_sin_97.xls	
Nombre	CONSULTAS	
Tipo	Hoja de cálculo: “Consultas”	
Descripción	Colección genérica de los distintos tipos de consultas que ofrece la clínica veterinaria a sus mascotas.	
Estructura		
Columna	Formato	Descripción
CODIGO_CLASE	Numérico	Código establecido por la empresa para discernir las clases de consultas que pueden darse.
CLASE_CONSULTA	Alfanumérico	Descripción del tipo de clase de consulta

- **Destino/s de Información**

Tabla 107. Destinos de Información; DIM_CLASECONSULTA

Base de Datos	DWH_VETERINARIA			
Nombre	DIM_CLASECONSULTA			
Tipo	Tabla			
Descripción	Enumera de forma genérica las clases de consultas que se realizan en la clínica.			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_CLASE	Numérico	6	PK	Identificador único y autoincremental para diferenciar las clases de consultas.
COD_CLASE	Numérico	6		Código generado por la empresa para identificar cada clase de consulta
CLASE_CONSULTA	Alfanumérico	50		Definición de la clase de consulta.

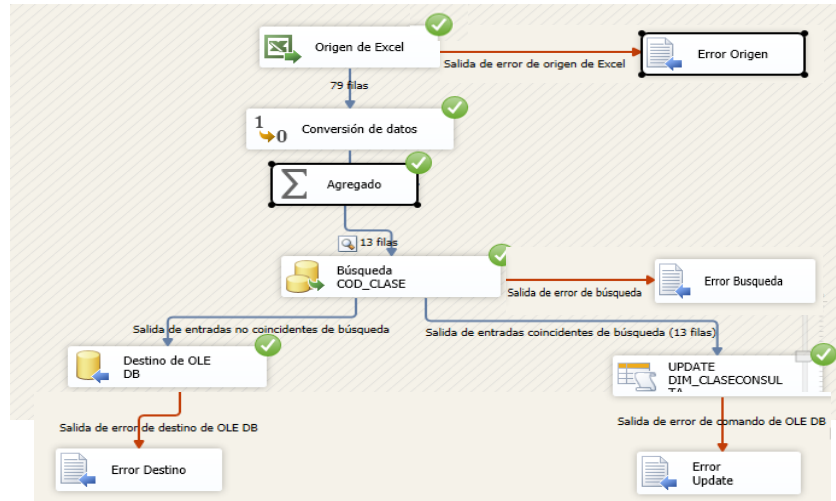
Tabla 108. Destinos de Información; Primary Key; DIM_CLASECONSULTA

Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_CLASECONSULTA	PK	(ID_CLASE)

- **Detalles del Proceso**
 - **Lógica de Carga**

Para cargar la DIM_CLASECONSULTA utilizaremos la ETL:

Ilustración 064. Detalles del Proceso de Carga; DIM_CLASECONSULTA



1. *Origen de Excel*: refiere a la configuración de entrada de datos. Este cuadro de herramientas coge la información del documento Excel, más concretamente de la hoja de cálculo denominada “Consulta”.
2. *Conversión de Datos*: es necesario convertir los tipos de datos para que no den errores a la hora de cargarlos en nuestra dimensión DIM_CLASECONSULTA.
3. *Agregado*: realiza las agrupaciones correspondientes para que la información no contenga duplicados.
4. *Búsqueda COD_CLASE*: asignamos un identificador único por cada tipo de consulta diferente para que queden diferenciados.
5. *Destino*: en caso de que la dimensión esté vacía la llena con los datos que recoge, según corresponda.
6. *Update DIM_CLASECONSULTA*: en caso de que la dimensión que queremos cargar esté llena de información, la actualiza.

- **Mapeos Origen/Destino**

Tabla 109. Mapeos Origen/Destino; DIM_CLASECONSULTA

ORIGEN (Excel)	DESTINO (Data WereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_CLASECONSULTA.ID_CLASE
CONSULTA.CODIGO_CLASE	DWH_VETERINARIA.dbo.DIM_CLASECONSULTA.COD_CLASE
CONSULTA.CLASE_CONSULTA	DWH_VETERINARIA.dbo.DIM_CLASECONSULTA.CLASE_CONSULTA

- **Modelo de Datos: Volumetría de Carga**

El tipo de cada consulta viene recogido por la misma hoja de cálculo que contiene la información concerniente a las consultas. Esto causa duplicados en los datos referentes al tipo de la consulta. Estos duplicados son tratados en la ETL y finalmente son impedidos.

Como consecuencia de la obviedad de datos duplicados, finalmente suman un total de 13 registros distintos.

- **DIM_CONSULTA**

- **Fuentes de Información**

La fuente de información de la DIM_CONSULTA, está en la hoja de cálculo de “Consultas”. Aquí encontramos todos los datos que la empresa considera importantes acerca de las consultas que realiza a diario.

Tabla 110. Fuentes de Información; DIM_CONSULTA

Nombre del Archivo	Dim_sin_97.xls	
Nombre	CONSULTAS	
Tipo	Hoja de cálculo: “Consultas”	
Descripción	Contiene todas las consultas que la clínica ofrece a sus clientes o mascotas.	
Estructura		
Columna	Formato	Descripción
CODIGO_CONSULTA	Numérico	Código numérico que genera la empresa para distinguir cada consulta que se realiza de forma única.
DESCRIPCIÓN	Alfanumérico	Definición de la consulta
PRECIO	Decimal	Importe en €, precio de la consulta.
CODIGO_CLASE	Numérico	Hace referencia a la clase de la consulta.

- **Destino/s de Información**

Tabla 111. Destinos de Información; DIM_CONSULTA

Base de Datos	DWH_VETERINARIA
Nombre	DIM_CONSULTA
Tipo	Tabla
Descripción	Enumera cada una de las consultas ofrecidas a las mascotas de la clínica.

Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_CONSULTA	Numérico	6	PK	Identificador único y autoincremental para unificar cada consulta
COD_CONSULTA	Numérico	6		Número formado por la empresa para distinguir cada consulta.
DESCRIPCIÓN	Alfanumérico	50		Definición de la consulta
PRECIO	Decimal	6,2		Importe en € que indica el precio de una consulta en concreto.
ID_CLASE	Numérico	6	FK	Referencia a la DIM_CLASECONSULTA, para identificar a que clase pertenece cada consulta

Tabla 112. Destinos de Información; Primary Key; DIM_CONSULTA

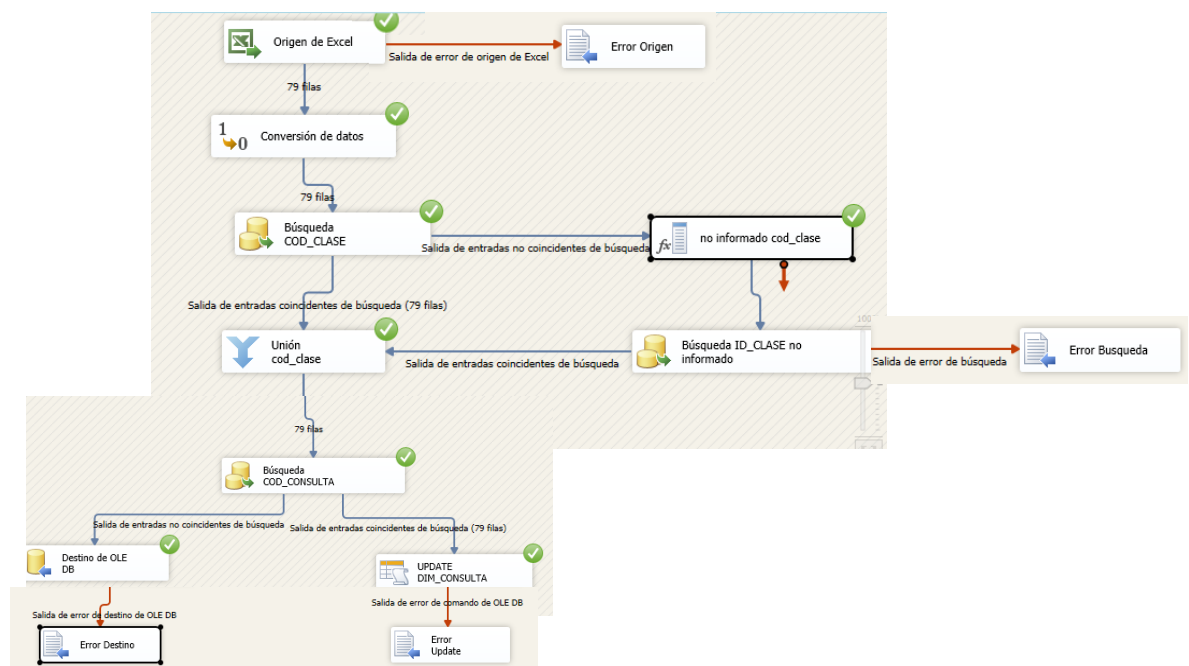
Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_CONSULTA	PK	(ID_CONSULTA)

Tabla 113. Destinos de Información; Foreign Key; DIM_CONSULTA

Foreign Key		
Foreign Key	Tipo	Miembros
FK_DIM_CLASECONSULTA	FK	(ID_CLASE)

- **Detalles del Proceso**
 - **Lógica de Carga**

Ilustración 065. Detalles del Proceso de Carga; DIM_CONSULTA



1. *Origen de Excel*: como en el caso de las anteriores, se cargan los datos de un documento Excel, más concretamente de la hoja de cálculo de “Consultas”.
2. *Conversión de datos*: una vez más, es necesario un cambio en el tipo de datos para poder cargarlos sin errores en el destino.
3. *Búsqueda COD_CLASE*: asigna el cod_claseconsulta que corresponda a cada consulta, pueden repetirse, ya que una clase de consulta tiene varias consultas asignadas.
4. *No informado cod_clase*: en caso de que no exista un cod_clase asignado a una consulta, lo describe con un -1, para referirse a un dato “no informado”.
5. *Búsqueda COD_CLASE no informado*: asigna un identificador único por cada consulta.
6. *Destino*: si la dimensión está vacía, la carga con los datos de la ETL.
7. *Update DIM_CONSULTA*: en caso de que la dimensión ya contenga datos, los actualiza con los nuevos datos.

- **Mapeos Origen/Destino**

Tabla 114. Mapeos Origen/Destino; DIM_CONSULTA

ORIGEN (Varios)	DESTINO (Data WereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_CONSULTA.ID_CONSULTA
CONSULTA.CODIGO_CONSULTA	DWH_VETERINARIA.dbo.DIM_CONSULTA.COD_CONSULTA
CONSULTA.DESCRIPCIÓN	DWH_VETERINARIA.dbo.DIM_CONSULTA.DESCRIPCIÓN
CONSULTA.PRECIO	DWH_VETERINARIA.dbo.DIM_CONSULTA.PRECIO
DWH_VETERINARIA.dbo.DIM_CLAS ESCONSULTA.ID_CLASE	DWH_VETERINARIA.dbo.DIM_CONSULTA.ID_CLASE

- **Modelo de Datos: Volumetría de Carga**

La clínica veterinaria tiene un total de 79 consultas diferentes. Por tanto en la ETL deben pasar los 79 registros para cargar la dimensión, sin excluir ninguna de las consultas. La cantidad de consultas la podemos verificar en la hoja de cálculo de “Consultas”, viendo la cantidad de filas que contiene, excluyendo la fila de los encabezados. Y en la propia ETL nos indica cuantos registros se cargan.

- **DIM_TIPO_PRODUCTO**
 - **Fuentes de Información**

La DIM_TIPO_PRODUCTO la cargaremos con los datos que encontramos en la hoja de cálculo designada como “Producto”. La empresa tiene esta hoja de cálculo dedicada a los productos que se tienen a la venta.

Tabla 115. Fuentes de Información; DIM_TIPO_PRODUCTO

Nombre del Archivo	Dim_sin_97.xls	
Nombre	PRODUCTO	
Tipo	Hoja de cálculo: “Producto”	
Descripción	Agrupa en tipos de productos todos los productos que se venden en la clínica veterinaria.	
Estructura		
Columna	Formato	Descripción
CODIGO_TIPO	Numérico	Código establecido por la empresa como único por tipo de producto.
TIPO_PRODUCTO	Alfanumérico	Descripción del tipo de producto.

- **Destino/s de Información**

Tabla 116. Destinos de Información; DIM_TIPO_PRODUCTO

Base de Datos	DWH_VETERINARIA			
Nombre	DIM_TIPO_PRODUCTO			
Tipo	Tabla			
Descripción	Agrupa los productos por su tipo más genérico.			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_TIPOPDUCTO	Numérico	6	PK	Identificador autoincremental y único para cada tipo de producto
COD_TIPOPDUCTO	Numérico	6		Código realizado por la empresa para identificar de forma única al tipo de producto.
TIPO_PRODUCTO	Alfanumérico	50		Definición del tipo de producto

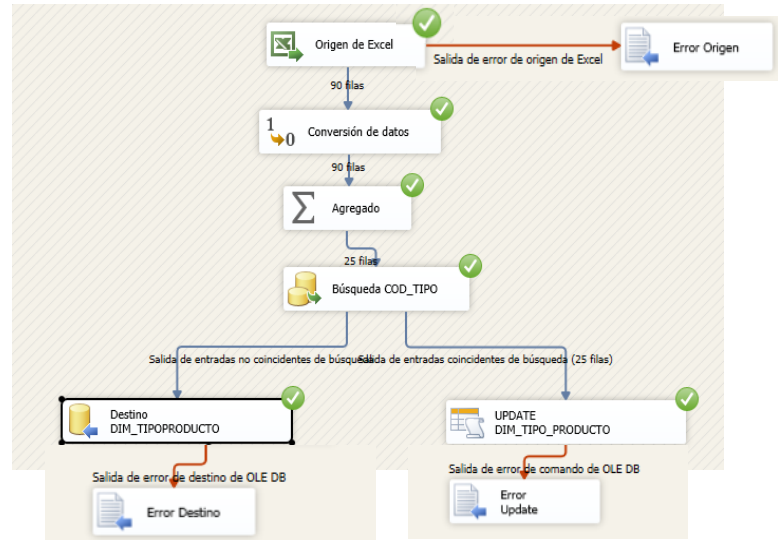
Tabla 117. Destinos de Información; Primary Key; DIM_TIPO_PRODUCTO

Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_TIPO_PRODUCTO	PK	(ID_TIPOPDUCTO)

- **Detalles del Proceso**
 - **Lógica de Carga**

ETL encargada de cargar los datos en la DIM_TIPO_PRODUCTO:

Ilustración 066. Detalles del Proceso de Carga; DIM_TIPO_PRODUCTO



1. *Origen de Excel*: establece un origen de datos. Lo configuramos para que nos adquiera la información pertinente. En este caso la correspondiente al tipo de producto
2. *Conversión de datos*: con esta herramienta conseguimos cambiar el tipo de datos de cada columna para que se corresponda perfectamente con el tipo de datos de las columnas de la dimensión a cargar.
3. *Agregado*: agrupa los datos iguales del conjunto de datos que le llegan, de esta manera, no habrá duplicados en los tipos de productos.
4. *Búsqueda COD_TIPO*: introduce un código identificador único por cada tipo de producto.
5. *Destino*: en caso de que la DIM_TIPOPRODUCTO esté vacía la carga con los datos que le llega. Asignando a cada columna los datos correspondientes que le llegan de la ETL.
6. *Update DIM_TIPO_PRODUCTO*: por esta rama irá, si la DIM_TIPOPRODUCTO está ya cargada de información. Se encarga de actualizar los datos ya existentes.

- **Mapeos Origen/Destino**

Tabla 118. Mapeos Origen/Destino; DIM_TIPO_PRODUCTO

ORIGEN (Excel)	DESTINO (Data WereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_TIPO_PRODUCTO.ID_TIPOP DUCTO
PRODUCTO.CODIGO_TIPO	DWH_VETERINARIA.dbo.DIM_TIPO_PRODUCTO.COD_TIPOP RODUCTO
PRODUCTO.TIPO_PRODUCTO	DWH_VETERINARIA.dbo.DIM_TIPO_PRODUCTO.TIPO_PROD UCTO

- **Modelo de Datos: Volumetría de Carga**

Los 90 productos diferentes se agrupan en tipos. Mediante ETLs se quitan los duplicados que se generan en el Excel. De esta manera nos quedan 25 tipos de productos diferentes a la venta en la clínica veterinaria y cargados en la DIM_TIPO_PRODUCTO.

- **DIM_PRODUCTO**

- **Fuentes de Información**

La carga de la dimensión de producto se realiza en función de los datos que se encuentran almacenados en la hoja de cálculo de “Producto”, donde se enumeran y describen los productos que la clínica veterinaria tiene a la venta para sus clientes.

Tabla 119. Fuentes de Información; DIM_PRODUCTO

Nombre del Archivo	Dim_sin_97.xls		
Nombre	PRODUCTO		
Tipo	Hoja de cálculo: “Producto”		
Descripción	Se enumeran todos los productos que se tienen a la venta en la clínica veterinaria.		
Estructura			
Columna	Formato	Descripción	
CODIGO_PRODUCTO	Numérico	Código generado por la empresa para discernir los productos a la venta.	
NOMBRE	Alfanumérico	Nombre del producto	
DESCRIPCION	Alfanumérico	Definición del producto.	
COD_TIPO	Numérico	Código que identifica al tipo de producto al que pertenece	
PRECIO	Decimal	Importe en €.	

- **Destino/s de Información**

Tabla 120. Destinos de Información; DIM_PRODUCTO

Base de Datos	DWH_VETERINARIA			
Nombre	DIM_PRODUCTO			
Tipo	Tabla			
Descripción	Describe todos y cada uno de los productos que están a la venta en la clínica veterinaria			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_PRODUCTO	Numérico	6	PK	Identificador autoincremental y único para diferenciar cada producto en venta
COD_PRODUCTO	Numérico	6		Código establecido previamente por la empresa

NOMBRE	Alfanumérico	60		Nombre del producto
DESCRIPCION	Alfanumérico	255		Definición del producto
PRECIO	Decimal	6,2		Importe en € del precio del producto.
ID_TIOPRODUCTO	Numérico	6	FK	Refiere a la DIM_TIPO_PRODUCTO, para saber de qué tipo es el producto

Tabla 121. Destinos de Información; Primary Key; DIM_PRODUCTO

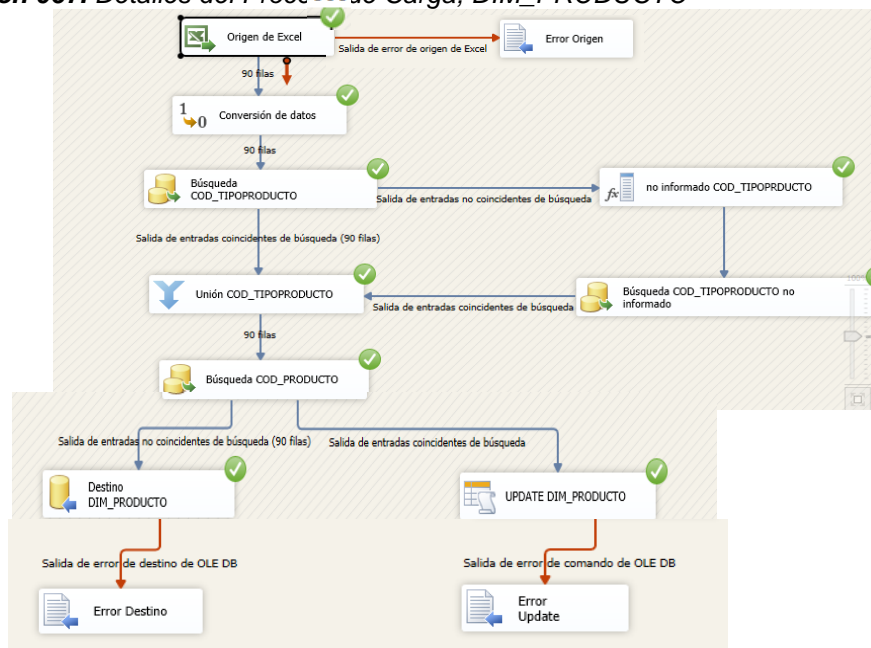
Primary Key		
Primary Key	Tipo	Miembros
PK_DIM_PRODUCTO	PK	(ID_PRODUCTO)

Tabla 122. Destinos de Información; Foreign Key; DIM_PRODUCTO

Foreign Key		
Foreign Key	Tipo	Miembros
FK_DIM_TIPO_PRODUCTO	FK	(ID_TIOPRODUCTO)

- **Detalles del Proceso**
 - **Lógica de Carga**

Ilustración 067. Detalles del Proceso de Carga; DIM_PRODUCTO



1. *Origen de Datos:* recoge los datos del documento Excel, de la hoja de cálculo denominada “Productos”.
2. *Conversión de Datos:* fundamental para evitar errores al insertar datos en el destino, puesto que el tipo de datos debe ser el mismo.

3. *Búsqueda COD_TIOPRODUCTO*: introduce identificadores de los tipos de productos a los que pertenece cada producto.
4. *No informado COD_TIOPRODUCTO*: si en el paso anterior encontrara algún valor nulo, lo cambiaría por un tipo producto “no informado”.
5. *Búsqueda COD_TIOPRODCUTO no informado*: inserta un id nuevo al campo no informado si hiciera falta.
6. *Unión COD_TIOPRODUCTO*: mediante este objeto unimos el campo no informado con el resto de campos.
7. *Búsqueda COD_PRODUCTO*: inserta códigos únicos para los productos, para poder diferenciarlos y hacer referencias.
8. *Destino DIM_PRODUCTO*: carga la DIM_PRODUCTO con los datos que le vienen de la ETL. Se configura qué dato se inserta en qué columna.
9. *UPDATE DIM_PRODUCTO*: si la DIM_PRODUCTO está cargada, se actualizan los datos con los nuevos, en su caso.

- **Mapeos Origen/Destino**

Tabla 123. Mapeos Origen/Destino; DIM_PRODUCTO

ORIGEN (Varios)	DESTINO (Data WereHouse)
AUTONUMERICO	DWH_VETERINARIA.dbo.DIM_PRODUCTO.ID_PRODUCTO
PRODUCTO.CODIGO_PRODUCTO	DWH_VETERINARIA.dbo.DIM_PRODUCTO.COD_PRODUCTO
PRODUCTO.NOMBRE	DWH_VETERINARIA.dbo.DIM_PRODUCTO.NOMBRE
PRODUCTO.DESCRIPCION	DWH_VETERINARIA.dbo.DIM_PRODUCTO.DESCRIPCION
DWH_VETERINARIA.dbo.DIM_TIPO_PRODUCTO.ID_TIOPRODUCTO	DWH_VETERINARIA.dbo.DIM_PRODUCTO.ID_TIOPRODUCTO
PRODUCTO.PRECIO	DWH_VETERINARIA.dbo.DIM_PRODUCTO.PRECIO

- **Modelo de Datos: Volumetría de Carga**

Los datos que se van a cargar en la DIM_PRODUCTO están extraídos de la hoja de cálculo de “Productos”, que contiene 90 filas correspondientes a 90 productos diferentes puestos a la venta en la clínica veterinaria. Por lo tanto, los registros a cargar en la dimensión es de 90.

6.2.3- Tablas de Hechos

Después de cargar todas las dimensiones, lo último serán las tablas de hechos, puesto que contienen información referente a las dimensiones principalmente, además de los indicadores (importes).

Al igual que hicimos con las dimensiones, vamos a definir primero las partes genéricas y en común de las dos tablas de hechos y finalmente explicaremos ambas en detalle.

- **Información General para las Tablas de Hechos**

- **Fuentes de Información**

Tabla 124. Fuentes de Información Tablas de Hechos; General

Nombre del Archivo	Dim_sin_97.xls	
Tipo Archivo	Documento Hoja de Cálculo	
Descripción	Colección de todos los datos que tiene la clínica veterinaria.	
Estructura		
Hoja de cálculo	Columnas	Descripción
DIARIO	Fecha, Cod_Mascota, Cod_Empleado, Cod_Consulta, Tipo_Consulta e Importe	Recoge los movimientos de la clínica respecto a las consultas. Todas las consultas que se realizan cada día.
VENTAS	Fecha, Cod_Producto, Cod_Empleado, Unidades, y Precio_Venta	Se describen todas las ventas de productos que se realizan en la clínica cada día.

- **Destino/s de Información**

El destino de toda la información concerniente a la clínica veterinaria, es una base de datos denominada DWH_VETERINARIA. La estructura de esta base de datos ha sido creada desde el inicio.

Tabla 125. Destinos de Información Tablas de Hechos; General

Base de Datos	DWH_VETERINARIA	
Tipo	Base de Datos	
Descripción	Base de datos con todas las tablas necesarias para estructurar adecuadamente las dimensiones de la clínica veterinaria y para un manejo eficaz de sus datos.	
Estructura		
Tablas	Columnas	Descripción
TH_CONSULTAS	Id_Mesano, Id_Consulta, Id_Empleado, Id_mascota, Tipo_consulta, Importe	Tabla de hechos que recoge las consultas realizadas en la clínica diariamente
TH_VENTAS	Id_Mesano, Id_Empleado, Id_Producto, Unidades y Precio	En esta tabla de hechos se apuntan todas las ventas diarias de productos de la empresa.

- **Detalles del Proceso**
 - **Borrado de Tabla**

Antes de la carga de las tablas de hechos es necesario realizar un borrado de datos, esto es debido a que la estrategia de carga de las tablas de hechos es DELETE/INSERT, no se pueden actualizar. El borrado tan sólo se realiza de unos cuantos registros, esto es así puesto que si se borrarán todos los registros, en una tabla de hechos mucho más grande, sería poco eficiente y tardaría mucho tiempo en realizar este proceso.

En los proyectos de BI, por tanto, se suele realizar un borrado de la tabla de hechos seleccionando los registros de los dos últimos años en adelante, de esta forma es más rápido.

A diferencia de las dimensiones, las tablas de hechos no contienen registros “no informados”, esto es lógico, la información que se recoge en las tablas de hechos son las ventas y consultas, estas se dan en una fecha determinada, por lo que este campo nunca será “no informado”. Por el contrario, sí se puede dar que a uno de los registros le falte algún tipo de dato, esto se soluciona con un “no informado” en ese campo.

Y por último cargaremos la información en la base de datos mediante el flujo de datos.

Ilustración 068. Tablas de Hechos; Flujo de Control



Borrado de la tabla de hechos:

Ilustración 069. Tablas de Hechos; Borrado de la Tabla de Hechos; General

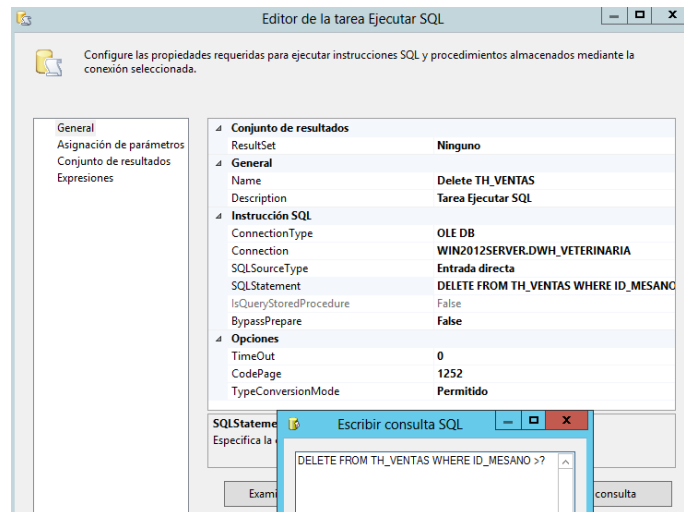


Ilustración 070. Tablas de Hechos; Borrado de la Tabla de Hechos; Asignación de Parámetros

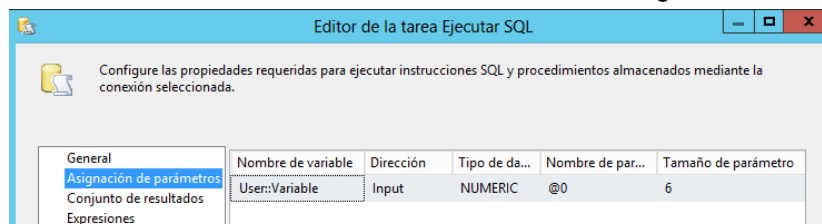
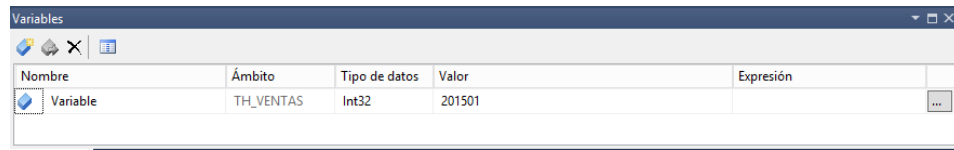


Ilustración 071. Tablas de Hechos; Borrado de la Tabla de Hechos; Creación de Variables

- **Mapeos Origen/Destino**

Tabla 126. Mapeos Origen/Destino; Tablas de Hechos

ORIGEN (Excel)	DESTINOS (Data WereHouse)
DIARIO	DWH_VETERINARIA.dbo.TH_CONSULTAS
VENTAS	DWH_VETERINARIA.dbo.TH_VENTAS

- **TH_CONSULTAS**
 - **Fuentes de Información**

Como origen de información de la TH_CONSULTAS, utilizaremos la hoja de cálculo de “Diario”. En esta hoja de cálculo se encuentra el registro diario realizado por la clínica de las diferentes consultas a los diferentes animales.

Tabla 127. Fuentes de Información; TH_CONSULTAS

Nombre del Archivo	Dim_sin_97.xls	
Nombre	DIARIO	
Tipo	Hoja de cálculo: “Diario”	
Descripción	Registro de todas las consultas realizadas a diario por la clínica.	
Estructura		
Columna	Formato	Descripción
FECHA	Fecha	Fecha en que se realiza la consulta
COD_MASCOTA	Numérico	Identificador de la mascota a quien se realiza la consulta
COD_EMPLEADO	Numérico	Empleado que realiza la consulta. (atendiendo a que las consultas solo las pueden realizar doctores)
COD_CONSULTA	Numérico	Identificador del tipo de consulta que se le realiza a un animal
TIPO_CONSULTA	Numérico	Refiere a si la consulta es de tipo urgente (1) o normal (0)
IMPORTE	Decimal	Precio en € de la consulta. Se incrementa en un 20% si la consulta es de tipo urgente.

○ **Destino/s de Información**

Tabla 128. Destinos de Información; TH_CONSULTAS

Base de Datos	DWH_VETERINARIA			
Nombre	TH_CONSULTAS			
Tipo	Tabla			
Descripción	Definición de las consultas realizadas por la clínica.			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_MESANO	Numérico	6	PK/ FK	Identifica el mes y el año en el que se realizan las consultas
ID_MASCOTA	Numérico	6	PK/ FK	Identificador de la mascota a quien se realiza la consulta
ID_EMPLEADO	Numérico	6	PK/ FK	Empleado que realiza la consulta
ID_CONSULTA	Numérico	6	PK/ FK	Refiere a la consulta que se realiza
TIPO_CONSULTA	Numérico	1		Si es urgente (1) y (0) si es normal
IMPORTE_CONSULTA	Decimal	7,2		Precio en € de la consulta

Tabla 129. Destinos de Información; Primary Key; TH_CONSULTAS

Primary Key		
Primary Key	Tipo	Miembros
PK_TH_CONSULTAS	PK	(ID_MESANO, ID_MASCOTA, ID_EMPLEADO, ID_CONSULTA)

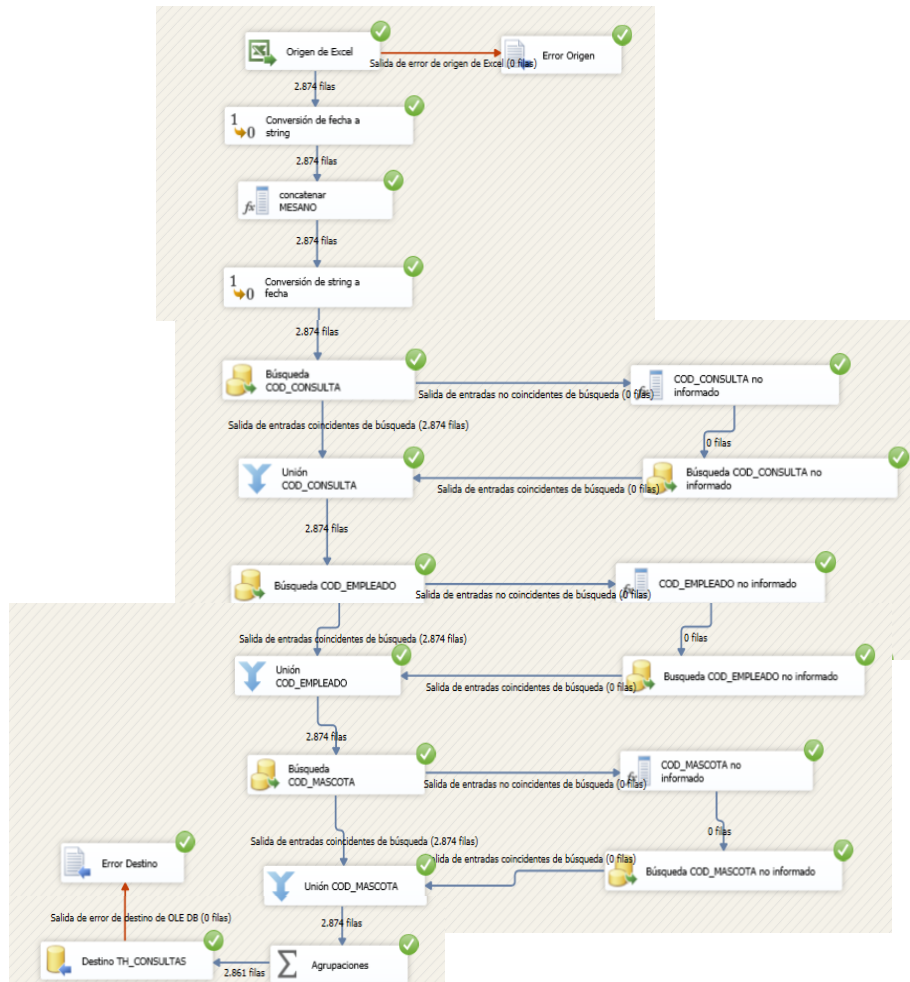
Tabla 130. Destinos de Información; Foreign Key; TH_CONSULTAS

Foreign Key		
Foreign Key	Tipo	Miembros
FK_DIM_TIEMPO	FK	(ID_MESANO)
FK_DIM_MASCOTA	FK	(ID_MASCOTA)
FK_DIM_EMPLEADO	FK	(ID_EMPLEADO)
FK_DIM_CONSULTA	FK	(ID_CONSULTA)

- **Detalles del Proceso**
 - **Lógica de Carga**

La ETL creada para cargar la TH_CONSULTAS es la siguiente:

Ilustración 072. Detalles del Proceso de Carga; TH_CONSULTAS



1. *Origen de Excel*: Configuramos para que el origen de datos de nuestra ETL sea el documento Excel facilitado por la empresa y, en este caso, la hoja de cálculo de “Diario”.
2. *Conversión de fecha a string*: Es necesario para convertir las fechas a tipo de datos string, para poder quitar posteriormente los separadores, y que no dé errores a la hora de insertar en la base de datos.
3. *Concatenar MESANO*: Para poder quitar los separadores del string de la fecha es necesario coger tan sólo los caracteres que precisamos y concatenarlos.
4. *Conversión de string a fecha*: una vez que hemos quitado los separadores del string, podemos volver a convertir el tipo de datos a fecha.
5. *Búsqueda COD_CONSULTA*: Comparamos los códigos de las consultas para evitar duplicados y ponemos los identificadores que correspondan.

6. *COD_CONSULTA no informado*: en caso de un código nulo, se añadirá un “no informado”
7. *Búsqueda COD_CONSULTA no informado*: se implanta un identificador para el caso del “no informado”
8. *Unión COD_CONSULTA*: se une el “no informado” con los demás códigos de consultas.
9. *Búsqueda COD_EMPLEADO*: como hemos hecho en el caso anterior, se buscan los códigos de los empleados correspondientes y se cambia este código por su identificador.
10. *COD_EMPLEADO no informado*: en caso de códigos de empleados a nulos, insertamos un campo “no informado”.
11. *Búsqueda COD_EMPLEADO no informado*: asignamos el identificador al caso “no informado”.
12. *Unión COD_EMPLEADO*: unimos el caso “no informado” con el resto de registros.
13. *Búsqueda COD_MASCOTA*: Realizamos la misma operación con los códigos de mascotas: buscamos los códigos de las mascotas que afectan y se modifica este código con su identificador.
14. *COD_MASCOTA no informado*: se crea un caso “no informado” para los códigos de mascotas que resulten nulos
15. *Búsqueda COD_MASCOTA no informado*: estipulamos un identificador al campo “no informado”.
16. *Unión COD_MASCOTA*: unimos el caso “no informado” con el resto de registros
17. *Agrupaciones*: Agrupa por los campos de ID_MESANO, ID_CONSULTA, ID_EMPLEADO, ID_MASCOTA, de esta forma, no habrá duplicados en la TH_CONSULTAS. Y al hacer estas agrupaciones, debemos sumar los importes para no perder información.
18. *Destino TH_CONSULTAS*: Carga la TH_CONSULTAS de la base de datos de salida con los datos correspondientes.

Como se aprecia en la imagen, en un principio tenemos un total de 2.874 filas, pero tras el agregado, nos quedan 2.861, esto es porque agrupa los registros repetidos en una sola fila y suma los importes de dichos duplicados.

- **Mapeos Origen/Destino**

Tabla 131. Mapeos Origen/Destino; TH_CONSULTAS

ORIGEN (Excel o Base de Datos)	DESTINO (Data WereHouse)
DWH_VETERINARIA.dbo.DIM_TIEMPO.ID_MESANO	DWH_VETERINARIA.dbo.TH_CONSULTAS.ID_MESANO
DWH_VETERINARIA.dbo.DIM_MASCOTA.ID_MASCOTA	DWH_VETERINARIA.dbo.TH_CONSULTAS.ID_MASCOTA
DWH_VETERINARIA.dbo.DIM_EMPLEADO.ID_EMPLEADO	DWH_VETERINARIA.dbo.TH_CONSULTAS.ID_EMPLEADO
DWH_VETERINARIA.dbo.DIM_CONSULTA.ID_CONSULTA	DWH_VETERINARIA.dbo.TH_CONSULTAS.ID_CONSULTA
DIARIO.TIPO_CONSULTA	DWH_VETERINARIA.dbo.TH_CONSULTAS.TIPO_CONSULTA
DIARIO.IMPORTE	DWH_VETERINARIA.dbo.TH_CONSULTAS.IMPORTE_CONSULTA

- **Modelo de Datos: Volumetría de Carga**

La volumetría de carga de esta tabla es 2.861 registros de los 2.874 existentes en la información de origen. Esta reducción es debido a la agrupación de datos que se hace en la ETL.

La cantidad de registros de esta tabla puede aumentar mucho puesto que se realizan consultas a diario en la clínica.

- **TH_VENTAS**
 - **Fuentes de Información**

El origen de información en la TH_VENTAS, se utilizará la hoja de cálculo de “Ventas”. En donde se recoge la información referente a las ventas diarias de productos de la clínica.

Tabla 132. Fuentes de Información; TH_VENTAS

Nombre del Archivo	Dim_sin_97.xls	
Nombre	VENTAS	
Tipo	Hoja de cálculo: “Ventas”	
Descripción	Datos de ventas de productos.	
Estructura		
Columna	Formato	Descripción
FECHA	Fecha	Fecha en la que se vende un producto
COD_PRODUCTO	Numérico	Código identificador del producto vendido
COD_EMPLEADO	Numérico	Empleado que realiza la venta
UNIDADES	Numérico	Unidades vendidas de un producto
PRECIO_VENTA	Decimal	Importe en € de la compra

- **Destino/s de Información**

Tabla 133. Destinos de Información; TH_VENTAS

Base de Datos	DWH_VETERINARIA			
Nombre	TH_VENTAS			
Tipo	Tabla			
Descripción	Ventas de productos de la clínica			
Estructura				
Columna	Formato	Long.	PK / FK	Descripción
ID_MESANO	Numérico	6	PK / FK	Mes y año en que se vende un producto
ID_PRODUCTO	Numérico	6	PK / FK	Identificador del producto vendido
ID_EMPLEADO	Numérico	6	PK / FK	Identificador del empleado que vende un producto
UNIDADES	Numérico	6		Cantidad numérica de los productos que se venden
PRECIO_VENTA	Decimal	7,2		Importe en € de la venta

Tabla 134. Destinos de Información; Primary Key; TH_VENTAS

Primary Key		
Primary Key	Tipo	Miembros
PK_TH_VENTAS	PK	(ID_MESANO, ID_PRODUCTO, ID_EMPLEADO)

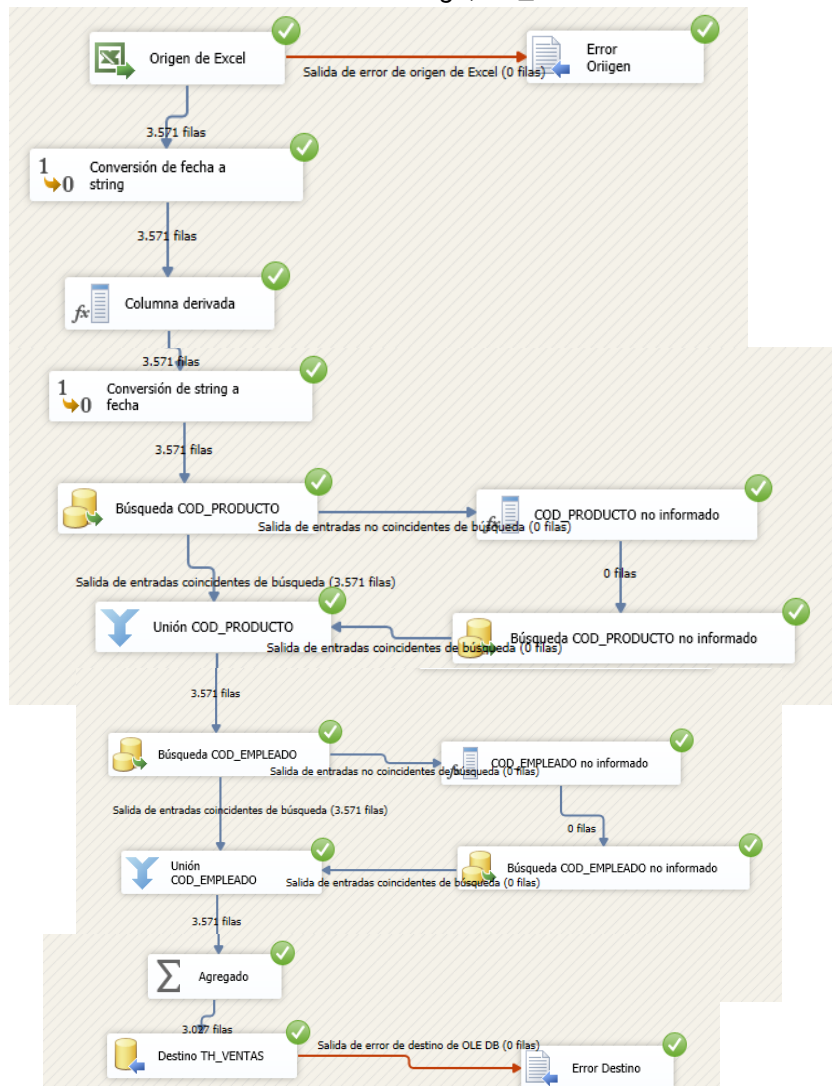
Tabla 135. Destinos de Información; Foreign Key; TH_VENTAS

Foreign Key		
Foreign Key	Tipo	Miembros
FK_DIM_TIEMPO	FK	(ID_MESANO)
FK_DIM_PRODUCTO	FK	(ID_PRODUCTO)
FK_DIM_EMPLEADO	FK	(ID_EMPLEADO)

- **Detalles del Proceso**
 - **Lógica de Carga**

La ETL creada para cargar la TH_VENTAS es la siguiente:

Ilustración 073. Detalles del Proceso de Carga; TH_VENTAS



1. *Origen de Excel*: Configuramos para que el origen de datos de nuestra ETL sea el documento Excel facilitado por la empresa y, en este caso, la hoja de cálculo de “Ventas”.
2. *Conversión de fecha a string*: Es necesario para convertir las fechas a tipo cadena de caracteres, y quitar los separadores, los cuales darán errores al cargarlos en base de datos
3. *Concatenar MESANO*: Recoge, del string de fecha, los datos importantes y los concatenan.
4. *Conversión de string a fecha*: una vez quitados los separadores, volvemos a convertir la cadena de caracteres a string.
5. *Búsqueda COD_PRODUCTO*: realiza una búsqueda de los códigos de los productos relacionando los que sean iguales y asigna los identificadores correspondientes.
6. *COD_PRODUCTO no informado*: en el caso de que existan códigos de productos a nulo los cambia a un registro “no informado”.
7. *Búsqueda COD_PRODUCTO no informado*: insertar un identificador para el campo de “no informado”.
8. *Unión COD_PRODUCTO*: dedicado a unir los campos no informados con el resto de campos.
9. *Búsqueda COD_EMPLEADO*: asigna identificadores a los códigos de los empleados según corresponda
10. *COD_EMPLEADO no informado*: en el caso de un código de empleado a nulo, lo definimos como “no informado”
11. *Búsqueda COD_EMPLEADO no informado*: asigna un identificador al registro “no informado”.
12. *Unión COD_EMPLEADO*: une el registro “no informado” con los demás registros.
13. *Agrupaciones*: Agrupa por identificadores y suma en el caso del precio o importe.
14. *Destino TH_VENTAS*: Carga la TH_VENTAS de la base de datos de salida.

- **Mapeos Origen/Destino**

Tabla 136. Mapeos Origen/Destino; TH_VENTAS

ORIGEN (Excel o base de datos)	DESTINO (Data WereHouse)
DWH_VETERINARIA.dbo.DIM_TIEMPO.ID_MESANO	DWH_VETERINARIA.dbo.TH_VENTAS.ID_MESANO
DWH_VETERINARIA.dbo.DIM_PRODUCTO.ID_PRODUCTO	DWH_VETERINARIA.dbo.TH_VENTAS.ID_PRODUCTO
DWH_VETERINARIA.dbo.DIM_EMPLEADO.ID_EMPLEADO	DWH_VETERINARIA.dbo.TH_VENTAS.ID_EMPLEADO
VENTAS.UNIDADES	DWH_VETERINARIA.dbo.TH_VENTAS.UNIDADES
VENTAS.PRECIO	DWH_VETERINARIA.dbo.TH_VENTAS.PRECIO_VENTA

- **Modelo de Datos: Volumetría de Carga**

Como se aprecia en la imagen de la lógica de carga, la volumetría de carga es de 3.571 registros, desde el inicio hasta el fin de la ETL. Aunque, al realizar agrupaciones de datos duplicados, finalmente se quedan en 3.027 registros almacenados.

Como podemos ver, no hay datos nulos, ya que en ningún caso el flujo de datos va por las ramas dedicadas a los “no informados”.

6.3- Proceso de Creación de Informes (Reporting)

Una vez que se ha terminado con el proceso de carga de información en la base de datos, este Data Warehouse resultante va a ser el utilizado en el proceso de reportig para realizar los informes que finalmente se entregarán al usuario final o cliente. En este punto vamos a conocer cómo se realiza en proceso de reporting paso a paso.

6.3.1- Configuración de la Herramienta

Para realizar los informes se pueden usar diferentes herramientas. En nuestro proyecto hemos optado por utilizar Microstrategy.

- **Parámetros de Configuración**

Para este proyecto sólo será necesaria una instancia de base de datos: DWH_VETERINARIA, que conectará con el Data Mart de consolidación

El número de filas permitidas por report es de 32000, para este proyecto concreto no habría problemas a priori, ya que dispone de un número mucho menor de datos y las previsiones de crecimiento no son tan elevadas, si se diera el caso en el futuro, se evaluaría.

- **Modelo de Datos: Tablas de Proyecto**

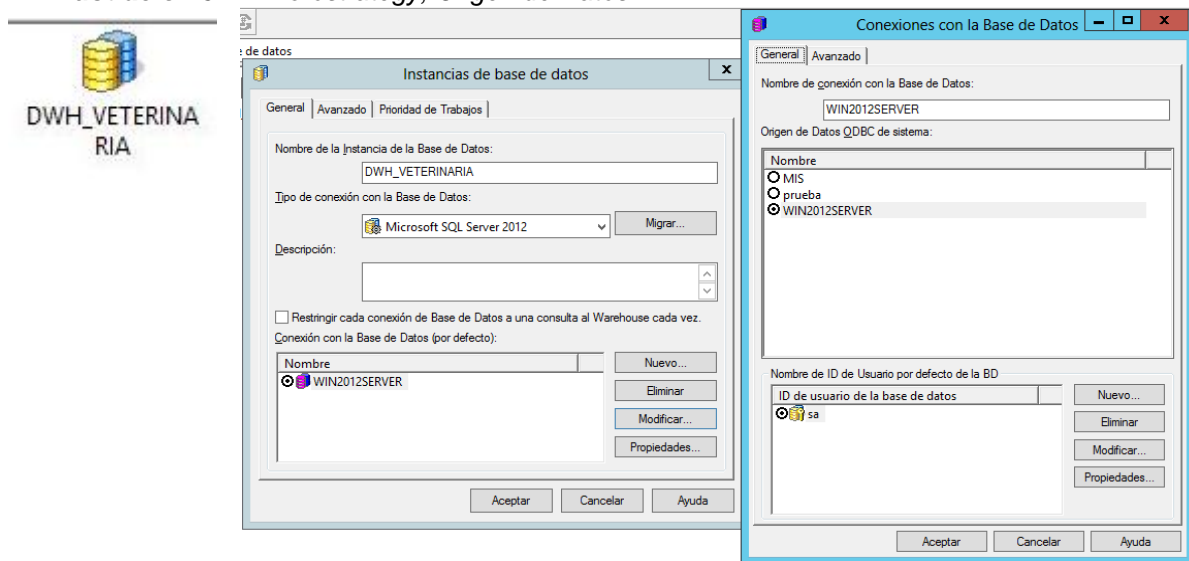
- ✓ *DIM_CARGO*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_CARGO, se empleara como tabla para el atributo Cargo
- ✓ *DIM_CLASECONSULTA*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_CLASECONSULTA, se empleara como tabla para el atributo Clase
- ✓ *DIM_CONSULTA*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_CONSULTA, se empleara como tabla lookup para el atributo Consulta
- ✓ *DIM_EMPLEADO*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_EMPLEADO, se empleara como tabla lookup para el atributo empleado
- ✓ *DIM_MASCOTA*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_MASCOTA, se empleara como tabla para el atributo Mascota
- ✓ *DIM_PRODUCTO*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_PRODUCTO, se empleara como tabla lookup para el atributo Producto
- ✓ *DIM_PROPIETARIO*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_PROPIETARIO, se empleara como tabla para el atributo Propietario

- ✓ *DIM_TIEMPO*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_TIEMPO, se empleara como tabla lookup para el atributo Mesano
- ✓ *DIM_TIPOPUESTO*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es DIM_TIPOPUESTO, se empleara como tabla para el atributo Tipopuesto
- ✓ *TH_CONSULTAS*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es TH_CONSULTAS, se empleara como tabla de hechos
- ✓ *TH_VENTAS*: esta tabla se encuentra en la base de datos DWH_VETERINARIA y su esquema es TH_VENTAS, se empleara como tabla de hechos

6.3.2- Creación del Proyecto: Microstrategy

Para comenzar con Microstrategy, debemos configurar un origen de datos para nuestros informes, este origen de datos se corresponde con el Data Warehouse que se ha realizado en el proceso de carga anterior.

Ilustración 074. Microstrategy; Origen de Datos



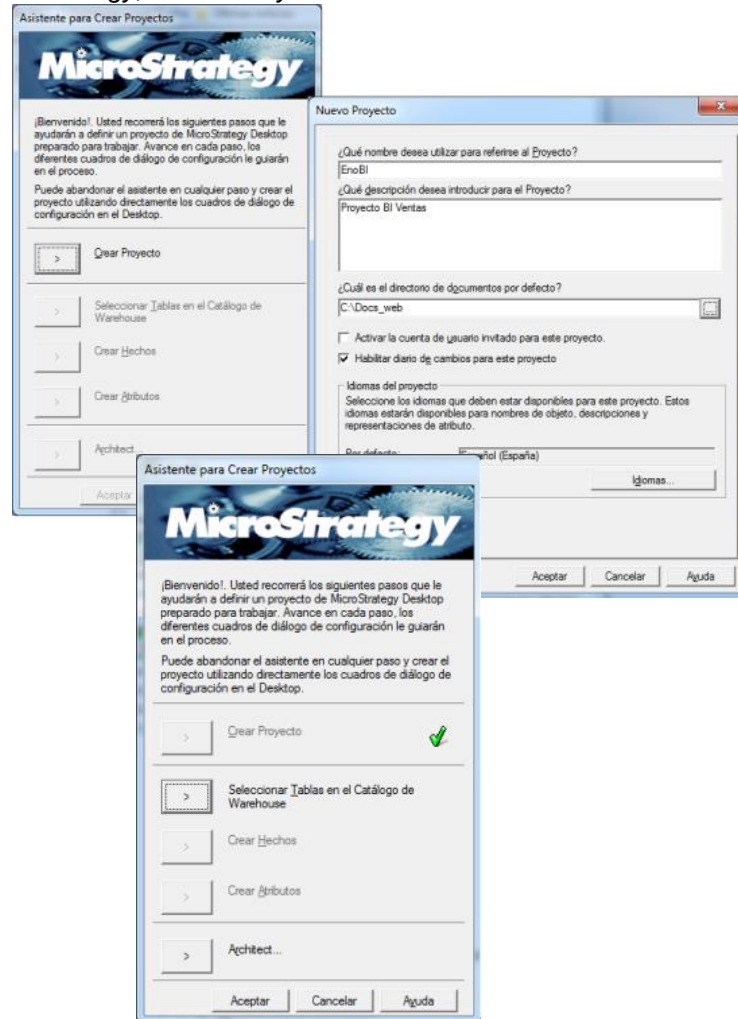
Un origen de proyecto es la ubicación centralizada de los diferentes proyectos de reporting. Es un contenedor que luego nos permitirá definir dentro de él los proyectos que tengamos en nuestra infraestructura de BI. Se pueden definir varios orígenes de datos. Al crear un origen de proyectos, le asociamos un tipo de origen y le indicamos el tipo de validación de usuarios que se va a realizar para el (usuario de windows, usuario de Microstrategy, usuario LDAP, usuario de base de datos, etc).

En nuestro caso debemos tener en cuenta los dos usuarios que se van a utilizar para el sistema; el usuario administrador con todos los privilegios y el usuario cliente con opciones sólo de lectura.

Tras la creación del origen de datos podemos crear un nuevo proyecto asociado a este origen y asumir todas o algunas de las tablas o elementos que tenemos en este origen de datos. Asociaremos las que necesitemos para nuestros informes, en el caso de nuestro sistema, como

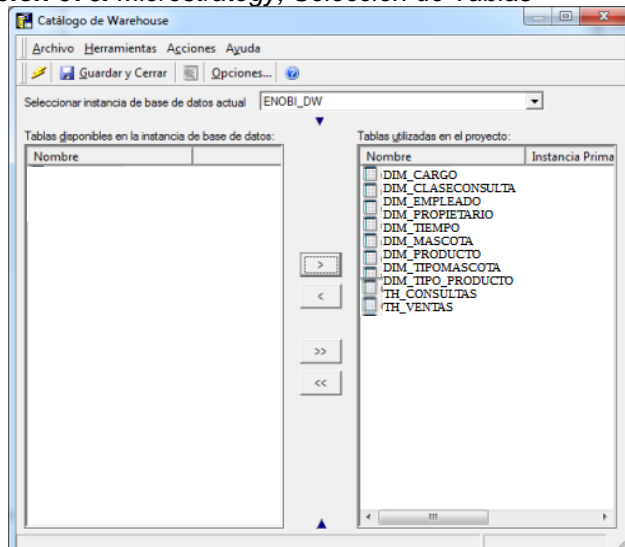
está preparado así, cogeremos todas las tablas con toda su información. Para realizar un nuevo proyecto, usaremos el asistente.

Ilustración 075. Microstrategy; Nuevo Proyecto



Una vez hemos creado el proyecto y seleccionado las tablas, la creación de hechos y de atributos las haremos a mano directamente en el programa, por lo que aceptaremos los cambios y cerraremos el asistente

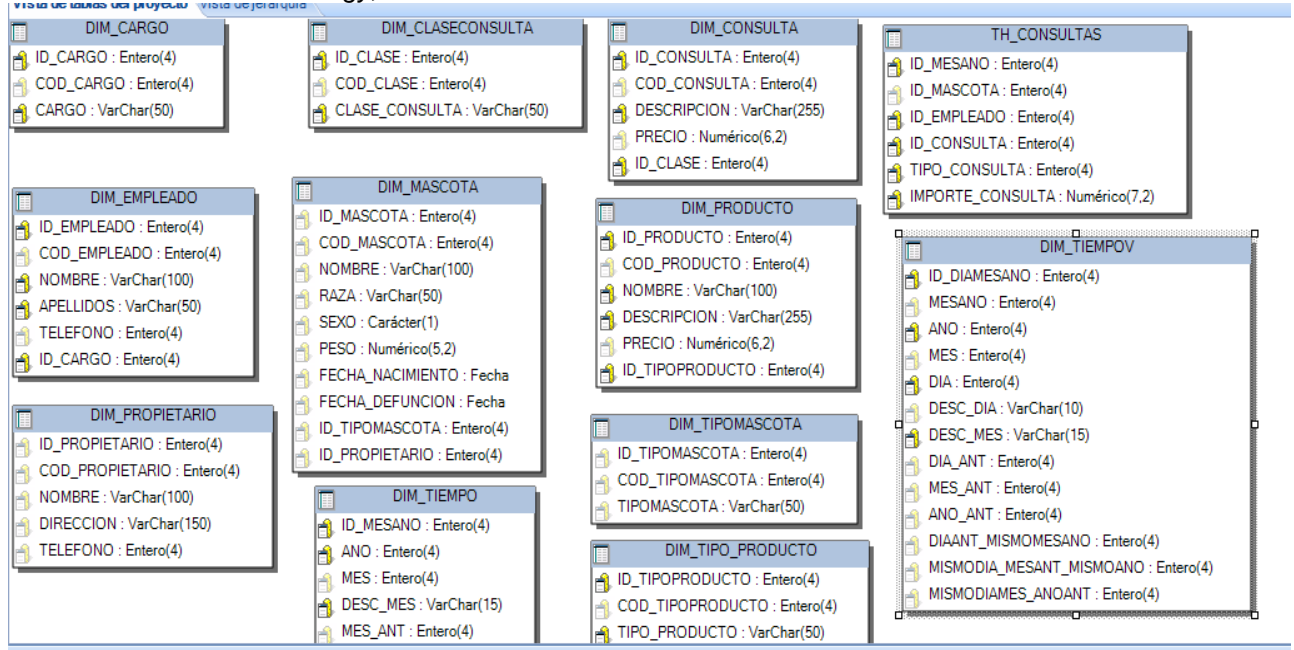
Ilustración 076. Microstrategy; Selección de Tablas



Microstrategy nos permite trabajar con la misma tabla varias veces a través de los “alias” de tabla. Puede ser útil cuando la misma dimensión física se utiliza de forma lógica en varios lugares (y no es necesario tener que tener una tabla física para cada una de las dimensiones). Igualmente, también nos permite trabajar con vistas. Tras la creación del proyecto y con las tablas en Microstrategy y ya podemos trabajar con la herramienta.

En el siguiente diagrama podemos ver como se han cargado todas las tablas en el proyecto.

Ilustración 077. Microstrategy; Tablas



• Usuarios y Permisos

Para Microstrategy también hay que configurar los usuarios y sus permisos, al igual que hemos hecho con SQL Server. En este caso es un poco más complicado.

- 1- En la pantalla principal, desplegamos el apartado administración de nuestro origen de proyecto Vicherinario y hacemos clic derecho en administración de usuarios y a continuación en nuevo grupo.
- 2- Creamos grupo cliente, damos nombre y aceptamos
- 3- Una vez el grupo está creado, hacemos doble clic para seleccionarlo y abrirlo, deberá aparecer en blanco ya que lo acabamos de crear y no tiene usuarios, hacemos clic derecho, nuevo, y creamos un usuario
- 4- Establecemos el nombre de este nuevo usuario Vicherinario, y confirmamos contraseña “Vicherinario”, el resto de opciones se dejan por defecto por si en un futuro se necesita cambiar la contraseña
- 5- Haciendo clic derecho en el apartado administración de usuarios, el cual antes hemos mencionado, hacemos clic en propiedades
- 6- Una vez ahí nos situaremos en el apartado seguridad y pulsaremos el botón agregar
- 7- En el siguiente cuadro buscamos nuestro usuario Vicherinario dentro del grupo Cliente, activando el check de mostrar, teniendo seleccionada la opción de permiso “Ver”, podemos dar a aceptar, tras esto podremos observar como el usuario se ha agregado a la lista
- 8- A continuación debemos seleccionar qué permisos queremos darle a este usuario, al ser el cliente le daremos permisos solo de lectura para los informes web

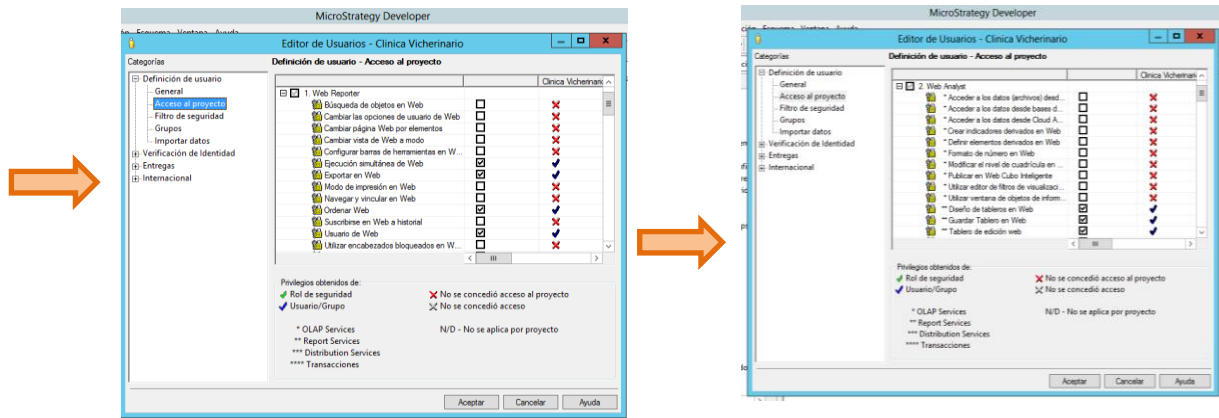
Ilustración 078. Microstrategy; Usuarios y Permisos

The illustration shows the following steps in MicroStrategy Developer:

- Step 1:** In the 'Administrador de Usuarios' folder, a right-click context menu is shown with 'Nuevo' > 'Grupo' selected.
- Step 2:** The 'Definición de usuario - General' dialog is shown for the user 'Clínica Vichermano'. Fields include 'Nombre completo' and 'Descripción'.
- Step 3:** The 'Propiedades - Seguridad' dialog for the 'Administrador de Usuarios' role is shown. It contains a table of permissions:

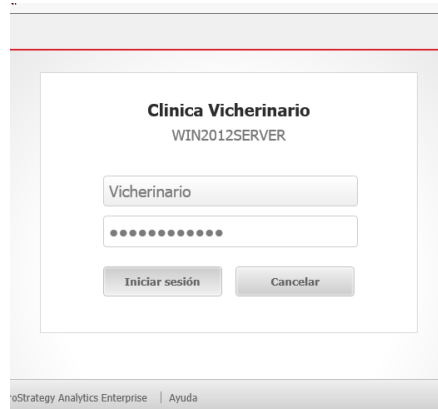
Usuario	Objeto	Hijos
Administrador	Control total	Por defecto
Everyone	Personalizado	Personalizado
Public / Guest	Personalizado	Personalizado
System Administrators	Personalizado	Control total
System Monitors	Por defecto	Personalizado
User Administrators	Personalizado	Por defecto

- Step 4:** The 'Seguridad' tab in the 'Propiedades - Seguridad' dialog is active, showing a list of users. 'Vichermano (Clínica Vichermano)' is selected.
- Step 5:** The 'Propiedades - Seguridad' dialog is shown with the 'Cliente' user selected in the 'Usuarios y grupos' list.



Tras haber creado el usuario y otorgado todos los permisos necesarios, debemos recordar que al terminar la realización de nuestros informes, deberemos mover todos los documentos que queramos que el cliente tenga acceso, de la carpeta de mis informes a la de informes compartidos dentro del navegador. Para ello colocamos una copia de los tableros en la carpeta de informes compartidos y con clic derecho seleccionamos “compartir”

Ilustración 079. Microstrategy; Login



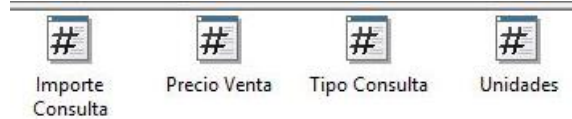
Por defecto, Microstrategy viene con el usuario administrador creado, en estos momentos ya hemos creado el usuario cliente. Cuando queramos acceder a los informes desde el navegador o en el propio sistema, nos pedirá siempre los credenciales para tener acceso con unos permisos o con otros.

- **Hechos**

De las tablas seleccionadas en el paso anterior, indicaremos que campos son los que corresponden a los hechos. Estos campos nos servirán de base para la creación de las métricas, que serán las que utilizaremos en los informes, documentos y/o tableros. Estas métricas, partiendo de las base de los hechos, podrán incluir operaciones, cálculos de uno o más campos, así como el uso de funciones complejas (Microstrategy incluye un gran número de funciones para realizar cálculos complejos sobre los datos, incluyendo funciones estadísticas). Esto nos permitirá disponer de valores que se calculan y que realmente no están guardados en la base de datos.

Los hechos que se definen y emplean en nuestro proyecto son todos dedicados a importes y son los siguientes:

Ilustración 080. Microstrategy; Hechos

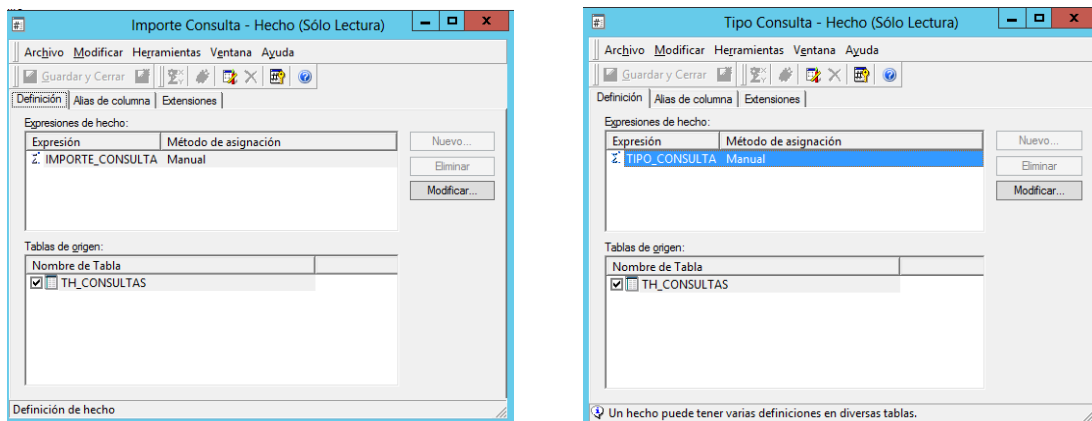


- *Importe consulta:* importe en moneda local (euros). La tabla y columna que lo componen son TH_CONSULTAS.IMPORTE_CONSULTA
- *Precio venta:* importe de las ventas en moneda local (euros). La tabla y columna que lo componen son TH_VENTAS.PRECIO_VENTA
- *Tipo consulta:* muestra si la consulta ha sido urgente. La tabla y columna que lo componen son TH_CONSULTAS.IMPORTE_CONSULTA
- *Unidades:* muestra el número de unidades que se vendieron de un artículo. La tabla y columna que lo componen son TH_VENTAS.IMPORTE_CONSULTA

Estos hechos se consolidan en el nivel de detalle más bajo.

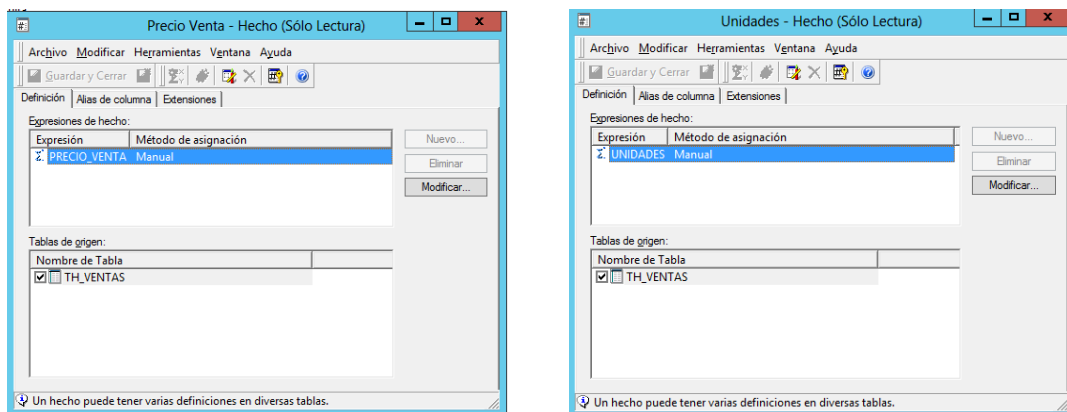
- “Importe consulta” y “tipo consulta” son indicadores pertenecientes a la TH_Conultas y están relacionados con las consultas y las mascotas.

Ilustración 081. Microstrategy; Hechos; Importe Consulta y Tipo Consulta



- “precio venta” y “unidades” son indicadores pertenecientes a la TH_Ventas y se relacionan con los productos.

Ilustración 082. Microstrategy; Hechos; Precio Venta y Unidades



• Atributos

En este paso, seleccionaremos los campos que corresponden a los atributos y realizaremos la configuración básica de ellos, como descripciones, ordenación, asignación de descripciones a códigos (lookups), así como la configuración de la estructura jerárquica de los diferentes componentes que forman una dimensión (a través de las relaciones padres e hijos).

Los atributos, en Microstrategy, se corresponden con los del Data Warehouse, esto es, tendremos en cuenta la información que queremos mostrar en el informe y crearemos un atributo por cada campo de la base de datos que queremos mostrar.

Un atributo, en resumen, son los datos que le dan sentido a los hechos, es decir sus descripciones, se suelen llamar atributos o dimensiones.

La creación de atributos es un poco más compleja y lleva asociados varios pasos. En primer lugar, de todos los campos de las tablas, seleccionaremos cuales son los que corresponden a nuestros atributos. En el caso de que un atributo lleve asociado un campo identificador y un campo descripción, solo seleccionaremos el campo identificador (pues después se establecerá la relación entre campos de código y campos de descripción o de lookup).

A continuación, para cada uno de los atributos, indicaremos su campo de lookup. Cuando un campo no dispone de este, indicaremos “Utilizar ID como descripción”.

Como último paso en la creación de los atributos, para cada uno de ellos indicaremos que atributos son hijos de él (están después en la jerarquía de la dimensión) o cuales son padre (están arriba en el árbol).

Los atributos que tenemos en nuestro proyecto son los siguientes:

Ilustración 083. Microstrategy; Atributos

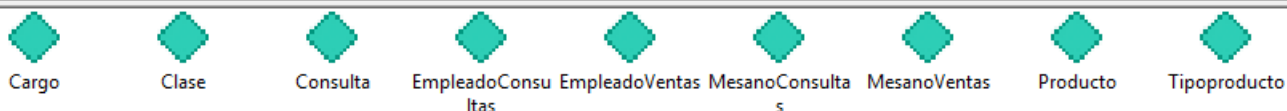
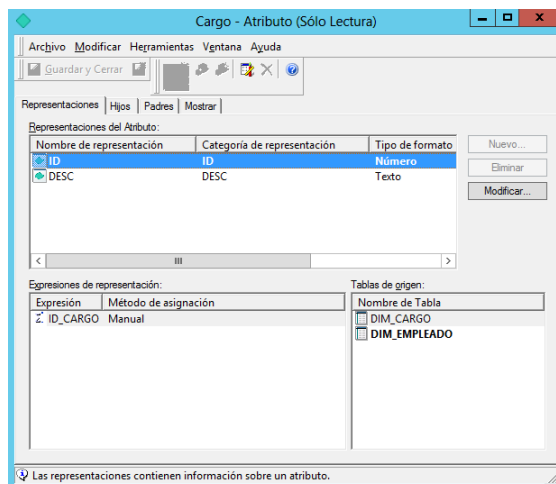


Ilustración 084. Microstrategy; Atributos; Configuración



Cargo: se corresponde con el cargo de los empleados

ID → DIM_CARGO.ID_CARGO

DESC → DIM_CARGO.CARGO

Clase: se corresponde con la clase de consulta que se ha realizado

ID → DIM_CLASECONSULTA.ID_CLASE

DESC → DIM_CLASECONSULTA.CLASE_CONSULTA

Consulta: se corresponde con los datos de la consulta realizada

ID → DIM_CONSULTA.ID_CONSULTA

DESC → DIM_CONSULTA.DESCRIPCION

EmpleadoConsultas y EmpleadoVentas: se corresponde con los datos de empleados que realizan consultas y ventas

ID → DIM_EMPLEADO.ID_EMPLEADO

DESC → DIM_EMPLEADO.NOMBRE

MesanoConsultas y MesanoVentas: se corresponde con el mes y el año de la fecha de la consulta

ID → DIM_TIEMPO.ID_MESANO

DESC → DIM_TIEMPO.DESC_MES

Nueva representación → DIM_TIEMPO.ANO

Producto: se corresponde con los datos de los productos vendidos

ID → DIM_PRODUCTO.ID_PRODUCTO

DESC → DIM_PRODUCTO.NOMBRE

Nueva representación → DIM_PRODUCTO.DESCRIPCION

Tipoproducto: Muestra el tipo de producto vendido

ID → DIM_TIPOPRODUCTO.ID_TIPOPRODUCTO

DESC → DIM_TIPOPRODUCTO.TIPO_PRODUCTO

La explicación de porqué empleado y Mesano aparecen en dos ocasiones es sencilla, ambas dimensiones (DIM_EMPLEADO y DIM_TIEMPO), son comunes a las tablas de hechos. A la hora de la realización de los atributos para los informes es necesario configurar a la tabla de hechos que apuntan, y como tablas de hechos hay dos, es necesario que figuren dos atributos de cada, uno por cada tabla de hechos.

• Jerarquías

Las jerarquías nos permiten organizar y navegar por los atributos. En Microstrategy, disponemos de dos tipos de jerarquías:

System hierarchy: Es la jerarquía del sistema, la que incluye las relaciones definidas entre los atributos del proyecto. Esta jerarquía no es necesario definirla, sino que se genera automáticamente en la herramienta Desktop al crear el proyecto (fue la que construimos utilizando el Architect).

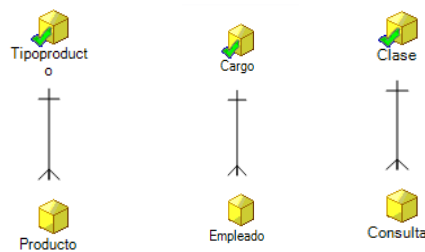
Es la jerarquía por defecto, y contiene todos los atributos del proyecto, formando parte de la definición del esquema. Al terminar la creación del proyecto, es la única jerarquía que existe. La jerarquía de sistema no se edita, pero se actualiza cada vez que añadimos o quitamos atributos hijo o padre utilizando el editor de atributos o el architect, o añadimos nuevos atributos al proyecto.

User hierarchy: las jerarquías de usuario son grupos de atributos y las relaciones entre ellos, juntados de manera que tenga significado en el contexto de análisis. No necesariamente tendrá que seguir el modelo de datos lógico. Además, a diferencia de la jerarquía de sistema, nos permite definir limitaciones al número de valores visualizados, establecer filtros sobre los valores a visualizar o definir un punto de entrada, que será el atributo donde nos posicionaremos al utilizar la jerarquía (por ejemplo, al navegar o al realizar filtrado).

Las jerarquías de usuario son las que tendremos disponibles cuando naveguemos por los datos en los informes y documentos de Microstrategy. Por tanto, son un elemento indispensable en la configuración del sistema, pues son el punto de partida para sacarle todo el partido a la navegación dimensional.

Las jerarquías de sistema las tenemos definidas, por tanto, una vez hemos creado los atributos con su herencia, y son:

Ilustración 085. Microstrategy; Jerarquías



Tipoproducto: Esta jerarquía define la relación entre el tipo de producto y el producto vendido. Un producto pertenece a un tipo de producto.

Cargo: Jerarquía que define la relación entre el cargo que desempeña y el empleado. Un empleado se corresponde con un cargo.

Clase: Jerarquía que define la relación entre la consulta y la clase que ocupa. Una consulta se corresponde con una clase.

En cuanto a las jerarquías de usuario, nos tenemos que fijar en los informes finales que se van a realizar. Es muy normal que se fije en primer lugar en todos los informes una jerarquía temporal. Es lógico que la información que representan los informes esté organizada de forma temporal para que el usuario pueda visualizar de forma clara en qué año/mes/día ha ganado más o menos, que se vende más en qué época del año, si la empresa va mejorando o empeorando con el paso del tiempo...

Después del tiempo, el orden de los atributos que se suele llevar es desde el más genérico al más específico. Esto es, por ejemplo, en el caso de las consultas en nuestra clínica veterinaria, la clase de la consulta y luego la consulta, primero tendremos las pruebas de laboratorio y posteriormente, de forma más específica el tipo de análisis: de sangre, orina....

• Transformaciones

Las transformaciones son objetos del esquema que podemos definir utilizando los atributos. Se utilizan para ejecutar análisis de series de tiempo como periodo a fecha, periodo actual contra periodo anterior o análisis de periodos cambiantes. Las transformaciones las definimos sobre los atributos. Esto nos configura un valor variable, que dependerá del valor del atributo con el que está relacionado, y que luego utilizaremos al crear los indicadores.

Ilustración 086. Microstrategy; Transformaciones

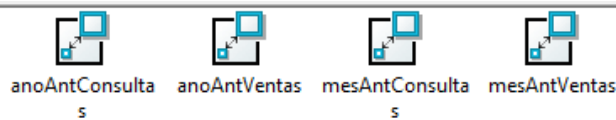
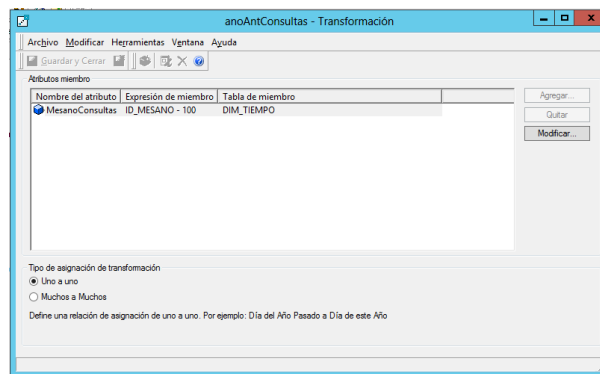


Ilustración 087. Microstrategy; Transformaciones; Configuración



- AnoAnt (Ejercicio a Ejercicio anterior): Devuelve el año anterior al periodo seleccionado. La tabla a la que accede es DIM_TIEMPO. El atributo que se utiliza es MESANO.
- MesAnt (Mes a mismo Mes Ejercicio anterior): Devuelve el mismo mes pero del año anterior al periodo seleccionado. La tabla a la que accede es DIM_TIEMPO. El atributo que se utiliza es MESANO.

• Filtros

Los filtros nos permiten establecer restricciones de valores a los diferentes elementos de nuestro sistema. Por ejemplo, podremos establecer un filtro sobre el atributo año y seleccionar un valor en concreto (por ejemplo 2009), y al ejecutar un informe con ese filtro, solo se recuperan los valores de ese año. Podríamos igualmente haber definido ese filtro de forma dinámica, de forma que al ejecutar el informe, nos aparezca una lista de años y seleccionemos el que vamos a analizar (será un filtro de selección dinámica).

○ *Filtros a Nivel de Herramienta*

Tan solo tenemos dos filtros a nivel de herramienta, y ambos son iguales, uno para el informe de consultas y otro para el de ventas. Los hemos utilizado para visualizar sólo los meses del año 2016 en la variación interanual. Estos filtros son estáticos, es decir, el usuario no puede interactuar con los filtros, los filtros son lo que son y no se pueden modificar por el usuario cuando el informe se abre.

Ilustración 088. Microstrategy; Filtros; Mesano

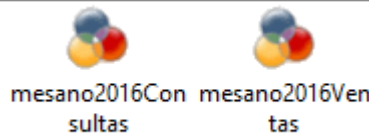
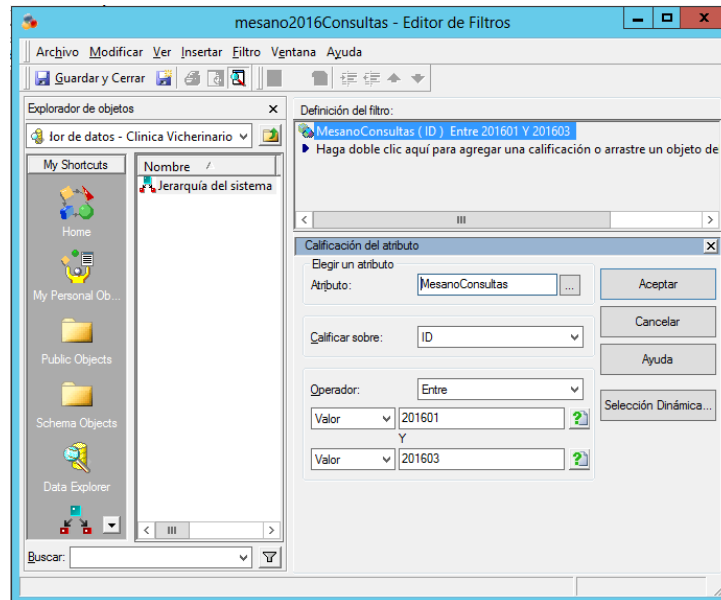


Ilustración 089. Microstrategy; Filtros; Mesano; Configuración



Filtro MESANO2016: Nos permite realizar filtrado entre los meses de 2016.

○ **Filtros a Nivel de Informe:**

Este tipo de filtros son dinámicos, es decir, el usuario puede interactuar con ellos y elegir los datos que mejor le parezcan. Los filtros a nivel de informe, como se indica, aparecen en la parte superior del propio informe.

En nuestros informes van a aparecer los siguientes filtros a nivel de informe:

Tiempo: Nos permite realizar filtrados por meses y años, para el se emplea el atributo mesano.

Ilustración 090. Microstrategy; Filtros; Fecha - Recta

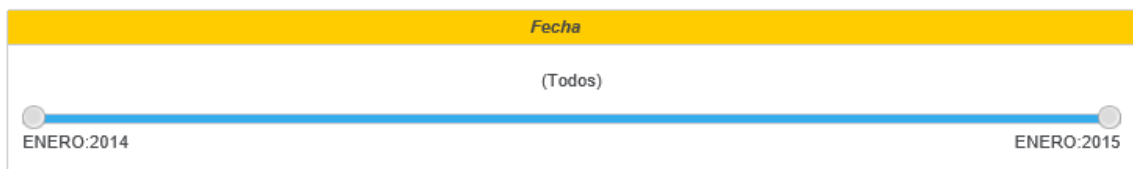
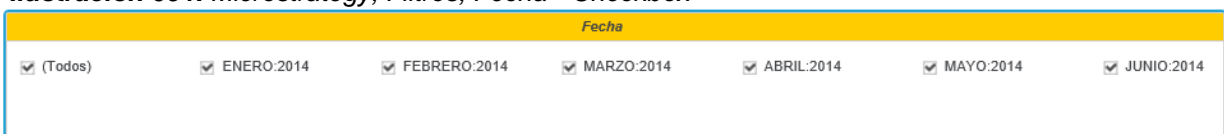
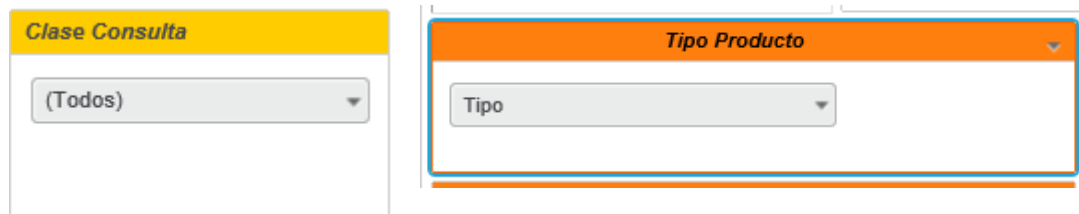


Ilustración 091. Microstrategy; Filtros; Fecha - Checkbox



Clase Consulta o Tipo Producto: Nos permiten realizar filtrados por clase de consulta o por tipo de producto, empleando el atributo correspondiente en cada caso, clase consulta y tipo producto respectivamente.

Ilustración 092. Microstrategy; Filtros; Clase Consulta y Tipo Producto



Consulta y Producto: Permite realizar filtrados por consultas y productos de forma más específica al anterior, empleando el atributo consultas y producto respectivamente.

Ilustración 093. Microstrategy; Filtros; Consulta y Producto



• Indicadores

Los indicadores son objetos de MicroStrategy que representan medidas de negocio e indicadores de rendimiento clave. Desde un punto de vista práctico, los indicadores son los cálculos realizados en los datos almacenados en la base de datos y cuyos resultados se muestran en un informe. Son parecidos a las fórmulas de un programa de hoja de cálculo.

No sería una exageración afirmar que el centro de prácticamente todos los informes son sus indicadores. La mayoría de las decisiones que se toman acerca del resto objetos que se deben incluir en un informe dependen de los indicadores que se utilicen en el informe.

En particular, los indicadores definen los cálculos analíticos que se realizan con los datos almacenados en el origen de datos. Un indicador se compone de hechos del origen de datos y de operaciones matemáticas que se van a realizar en tales hechos para, de este modo, poder efectuar un análisis de negocio con sentido en los resultados.

Ilustración 094. Microstrategy; Indicadores

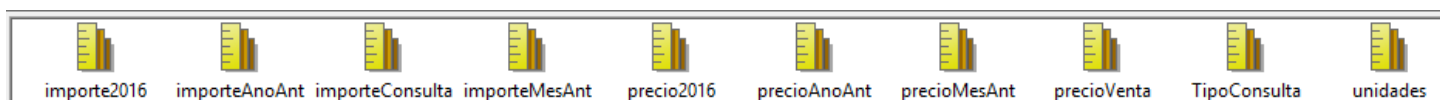
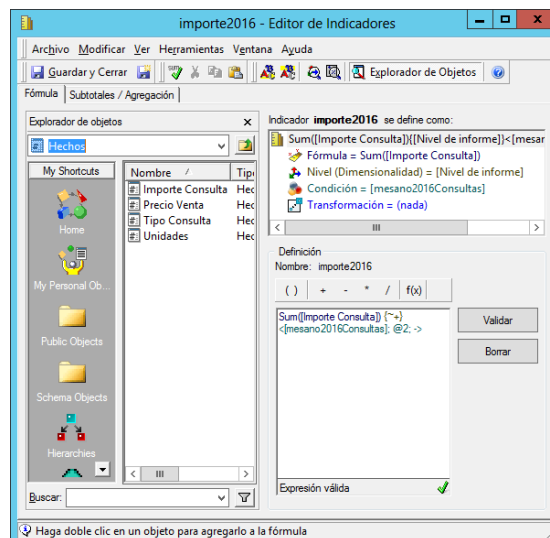


Ilustración 095. Microstrategy; Indicadores; Configuración

- **Importe2016:** Recogerá el sumatorio de los importes por las consultas realizadas en euros en 2016.
Formula: Sum(importe_consulta, filtro MESANO2016CONSULTAS)
- **ImporteAnoAnt:** Recogerá el sumatorio de los importes por las consultas realizadas en euros en el ejercicio anterior.
Formula: Sum(importe_consulta, transformación añoAntConsulta)
- **ImporteConsulta:** Recogerá el sumatorio de los importes por las consultas realizadas.
Formula: Sum(importe_consulta)
- **ImporteMesAnt:** Recogerá el sumatorio de los importes por las consultas realizadas en euros en el mes anterior del ejercicio actual.
Formula: Sum(importe_consulta, transformación mesAntConsulta)
- **Precio2016:** Recogerá el sumatorio de los importes por las ventas realizadas en euros en 2016 .
Formula: Sum(importe_consulta, filtro MESANO2016VENTAS)
- **PrecioAnoAnt:** Recogerá el sumatorio de los importes por las ventas realizadas en euros en el ejercicio anterior.
Formula: Sum(precio_venta, transformación añoAntVentas)
- **ImporteMesAnt:** Recogerá el sumatorio de los importes por las consultas realizadas en euros en el mes anterior del ejercicio actual.
Formula: Sum(precio_venta, transformación mesAntVentas)
- **PrecioVenta:** Recogerá el sumatorio de los importes por las ventas realizadas por productos.
Formula: Sum(precio_venta)
- **TipoConsulta:** muestra la suma de los tipos de consulta realizadas.
Formula: Sum(tipo_consulta)
- **Unidades:** muestra la suma de unidades vendidas por producto.
Formula: Sum(unidades)

6.3.3- Informes

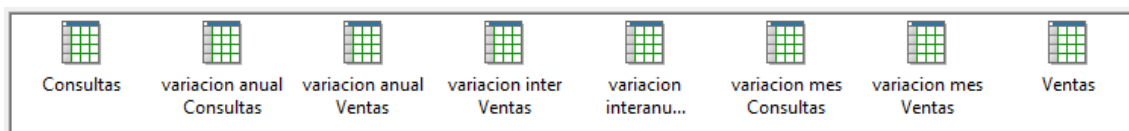
Como ya hemos visto, Business Intelligence (BI) es una amplia categoría de soluciones de software informático que permite a una empresa u organización, obtener información de sus operaciones más críticas a través de aplicaciones de generación de informes y herramientas de análisis.

Las aplicaciones de BI pueden incluir diversos componentes, como informes tabulares, hojas de cálculo, gráficas... Un informe es una vista de varias perspectivas de los datos, con visualizaciones que representan diferentes resultados e información de esos datos.

Una vez se ha terminado el proceso de carga y una vez creados todos los elementos necesarios para la instauración de informes, vamos a proceder a la creación de los mismos.

Según lo requisitos del cliente que se han explicado anteriormente, vamos a tener un mínimo de informes:

Ilustración 096. Microstrategy; Informes



Estos informes están creados a partir de los atributos y los indicadores que hemos preparado previamente, y contienen la información que el cliente nos ha pedido.

Ilustración 097. Microstrategy; Informes; Consultas

Nombre		Filtro de visualización				
		Indicadores	TipoConsulta	importeConsulta	porcentaje del total por meses	
MesanoConsultas	Clase	Consulta				
		Fractura diafisaria clavos	1	1.100,00 €	3,47%	
		Fractura diafisaria fijador	2	1.368,00 €	4,31%	
		Fractura diafisaria placa	0	800,00 €	2,52%	
		Fractura mandibular	1	516,00 €	1,63%	
		Ligamentos cruzados	1	1.952,00 €	6,15%	
		Osteocondritis disecante	0	1.220,00 €	3,84%	
		Proceso anconeo	2	1.308,00 €	4,12%	
		Proceso coronoides	0	1.220,00 €	3,84%	
		Trocleooplastia/cresta tibial	1	1.188,00 €	3,74%	
		Higroma de codo (penrose)	2	660,00 €	2,08%	
		Total	10	11.332,00 €	35,71%	
		Hernia diafragmática	1	1.100,00 €	3,47%	

• Creación de Informes: Microstrategy

Una vez que hemos creado todos los componentes necesarios, vamos a crear los informes. La herramienta Microstrategy ofrece múltiples posibilidades a la hora de “pintar” informes. La realización final de informes la haremos directamente en nuestro navegador, en la opción web de la herramienta, desde la URL <http://localhost/microstrategy/asp/Main.asp> podemos ver todos los proyectos que tenemos de Microstrategy. Tan sólo debemos entrar en el proyecto que nos acontece y comenzar con el proceso de creación de informes finales.

Después de varias pruebas en la realización de informes, elegimos hacer los informes mediante los denominados “Tableros”.

Esta decisión no fue tomada a la ligera. Los tableros ofrecen multitud de posibilidades en cuanto a gráficos que los informes dentro de la herramienta no brindan.

Los tableros son documentos que tratan de concentrar en un único pantallazo (aunque pueda tener varias secciones o pestañas), toda la información. Con una presentación de calidad, gráfica e intuitiva para permitir el análisis de la información de una forma fácil.

Además, los tableros nos ofrecen una vista final del informe. Lo que finalmente se le brinda al cliente es la dirección web, pero con los credenciales que utilizará el cliente, tan sólo tendrá la opción de visualización de informes (además de poder filtrar los datos).

Además, con los tableros, tenemos opción de introducir subinformes en el mismo informe utilizando las pestañas de la parte inferior.

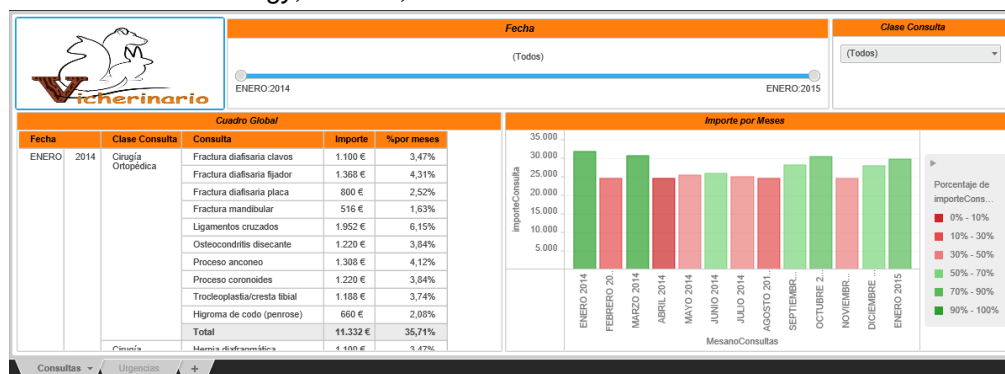
o Informe Global de Consultas:

En este informe tendremos los datos relacionados con las consultas realizadas por la clínica en los últimos dos años. Para tener la información más ordenada, se realizarán dos informes dentro del mismo, uno dedicado a todas las consultas y otro dedicado tan sólo a las urgencias.

Las urgencias son importantes. La clínica veterinaria asume consultas urgentes, las cuales tienen un incremento en el coste de un 20%. Consideramos significativo tener un informe tan sólo de las urgencias, puesto que el cliente quiere tener claro de dónde vienen sus beneficios.

-Consultas: Este informe facilita al análisis de volumen de negocio para las consultas de la clínica. Refleja el grosor de consultas y los beneficios de las mismas en un tiempo determinado. El informe es dinámico, permitiendo filtrar por tiempo y tipo de consultas. El análisis temporal para este informe será mensual.

Ilustración 098. Microstrategy; Tablero; Consultas



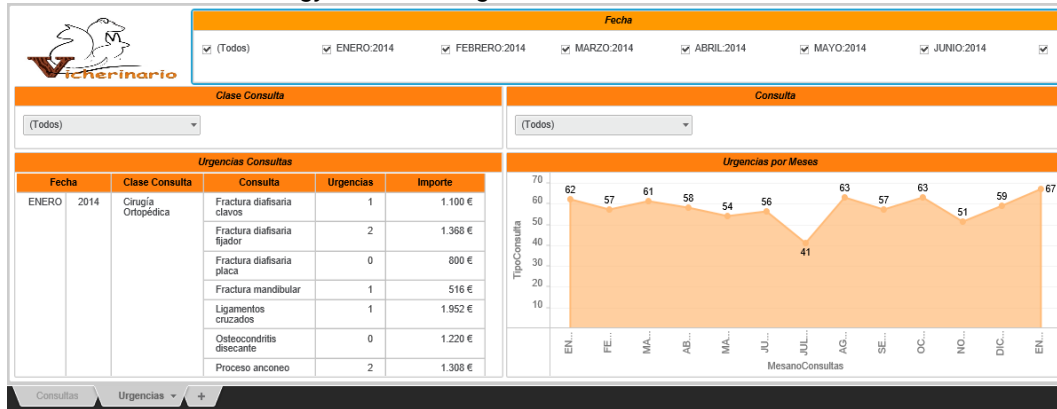
El cuadro global nos muestra los datos de forma numérica, con importes y porcentajes respecto de los totales mensuales.

Como todos los informes, este contiene gráficos dinámicos, cuyos datos se modifican cuando el usuario cambia los filtros y selecciona los datos que necesita.

En el gráfico de la derecha se muestran las ganancias por meses, siendo los verdes buenos resultados y los rojos peores resultados.

-**Urgencias:** Refleja el número de consultas atendidas y el importe de beneficios en un tiempo establecido y por tipo de consultas. Es un informe dinámico en el que se puede filtrar por tiempo, por clase de consulta y consultas. El análisis temporal es, igualmente, por meses.

Ilustración 099. Microstrategy; Tablero; Urgencias



En este caso, tenemos representadas las sumas de urgencias que se han tratado en el mes de cada consulta y las consultas normales.

En el gráfico se representan las urgencias totales de cada mes en el rango de tiempo establecido, para poder discernir el mes en que se atienden más urgencias. Mediante los filtros podemos seleccionar las clases de consultas y las consultas para poder visualizar qué consulta dentro de una clase tiene más urgencias.

o Informe Variaciones:

En este segundo informe tendremos las diferentes comparativas de beneficios en un tiempo determinado respecto del mismo tiempo en un año o dos antes. Este informe ha sido pedido expresamente por el cliente para verificar la constancia o mejora del negocio en un periodo de tiempo.

Como en el caso anterior, y como el cliente así lo requiere, tenemos dentro de este informe varios subinformes:

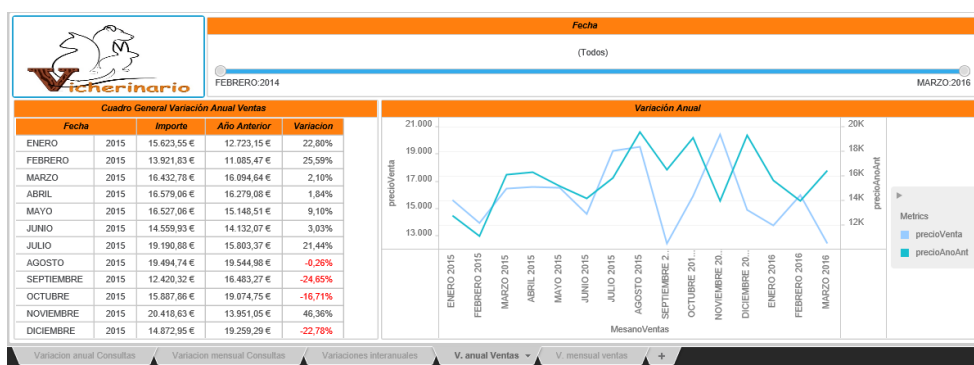
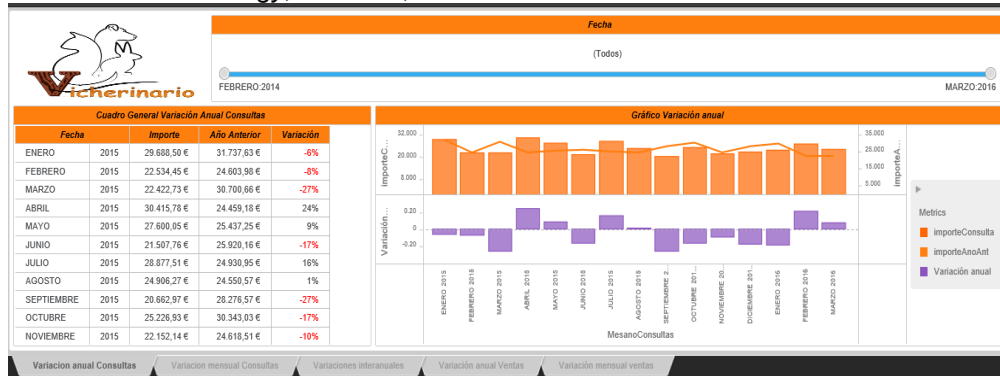
-**Variación Anual:** estos subinformes reflejan las ganancias o pérdidas de la empresa en cuanto a importes, tanto para ventas como para consultas, en un periodo de tiempo determinado y lo compara con el mismo periodo de tiempo del año anterior (variación anual). Realiza comparativas anuales de un mes con el mismo mes del año anterior. Son informes de tipo dinámico, pudiendo filtrar de forma temporal siendo el análisis temporal mensual.

En los datos, podemos ver, en la última columna del cuadro general, la variación calculada del mes del año actual respecto del mismo mes del año anterior. Si la variación es negativa aparece en color rojo.

Las gráficas muestran primeramente los importes del año actual en las barras y en línea roja vemos las ganancias del mismo periodo de tiempo del año anterior y en las barras inferiores se muestra la variación en sí, en el caso de las consultas.

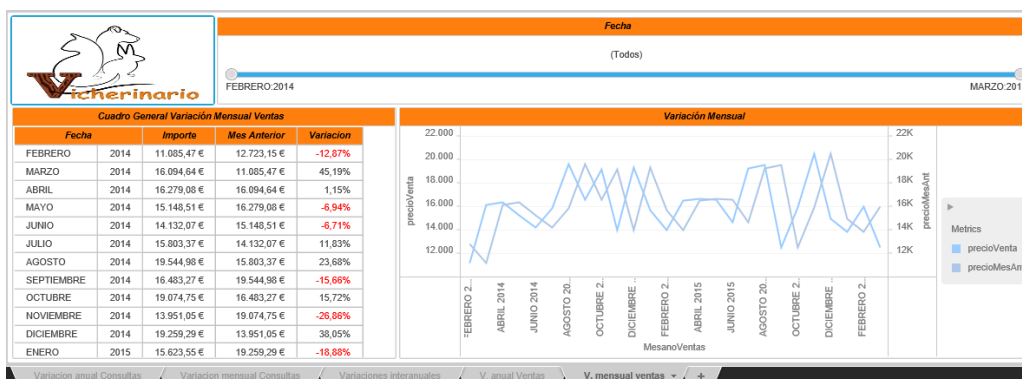
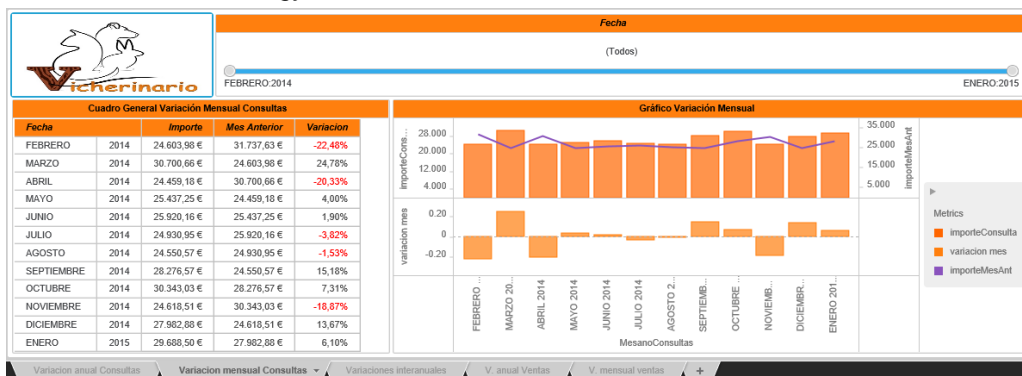
En caso de ventas de producto el grafico es lineal, siendo el azul claro las ganancias del año más actual y el azul oscuro las ganancias del mismo mes del año anterior.

Ilustración 100. Microstrategy; Tableros; Variación Anual



- Variación Mensual: estos subinformes, uno para ventas y otro para consultas, reflejan el crecimiento o decrecimiento neto del volumen de negocio mes a mes, así como la variación mensual producida. El análisis temporal para este informe será mensual. Y los filtros dinámicos son temporales.

Ilustración 101. Microstrategy; Tablero; Variación Mensual



La diferencia con los informes anteriores, es que en el mensual se reflejan mes a mes desde el primer mes comparable con el mes anterior del mismo año (excepto en enero que lo compara con diciembre del mes anterior).

En cuanto a los gráficos, son iguales: en barras tenemos los importes del mes actual, en línea los importes del mes anterior y en el gráfico de abajo, la variación, en el caso de consultas, y en cuanto a ventas, gráfico lineal en el que el azul claro refleja los beneficios del mes actual y en azul oscuro el mes anterior.

-Variación interanual: este subinforme refleja el crecimiento o decrecimiento neto del volumen de negocio del periodo de tiempo que va desde el inicio del año hasta el mes actual, comparado con el mismo periodo de tiempo del año pasado. En este caso, en la parte de la izquierda del informe, tenemos la variación interanual de consultas y en la parte derecha, el de ventas. El análisis temporal para este subinforme será mensual. Es un informe de tipo estático puesto que no hay necesidad de filtro alguno.

Ilustración 102. Microstrategy; Tablero; Variación Interanual



En los informes interanuales tendremos menos información, por ello podemos tener las variaciones interanuales de consultas y ventas en el mismo subinforme. Esto es porque en el caso en que nos acontece, el cliente nos ha facilitado información hasta el mes de marzo del año actual, por lo que se recogerán los datos de tres meses del año actual con respecto a los mismos tres meses del año anterior.

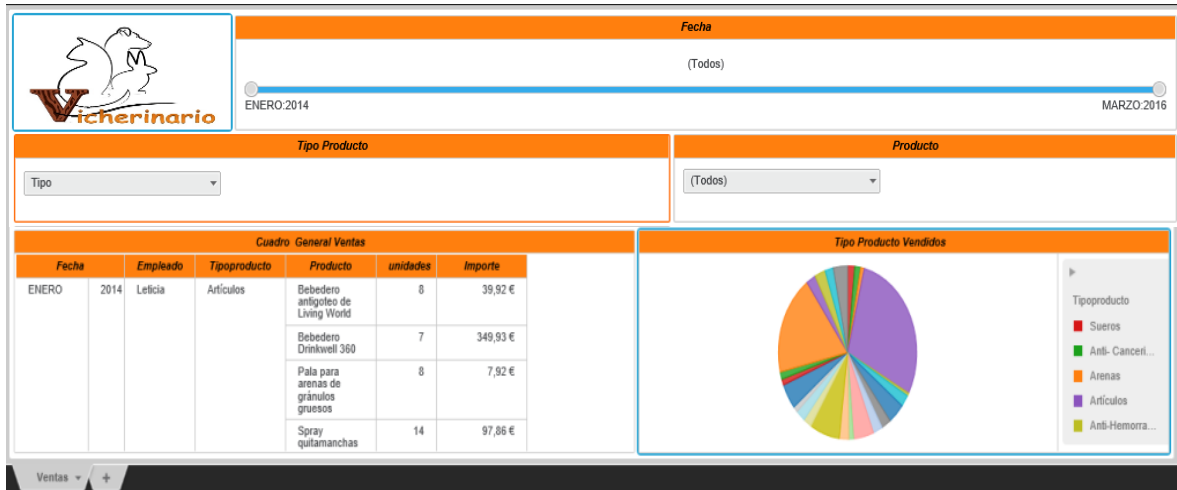
Los gráficos son iguales a los casos de los subinformes anteriores, las barras identifican los meses más recientes y la línea los mismos meses del año anterior.

o Informe Ventas:

Para finalizar realizaremos un informe dedicado a ventas, para poder visualizar cual es el beneficio de las ventas de productos en la clínica veterinaria.

-Informe Ventas: este informe permite analizar el volumen de negocio para las ventas de productos de la clínica. Refleja el volumen de ventas y los beneficios de estas ventas en un periodo de tiempo determinado. El análisis temporal es mensual, es dinámico, permitiendo filtrar por tiempo.

Ilustración 103. Microstrategy; Tablero; Ventas



Este informe, como se puede apreciar en la imagen, es muy similar al primero (General de consultas). Ofrece datos sobre las ventas de productos en la clínica, viendo las unidades vendidas de cada producto y el coste total de la venta.

Gráficamente podemos ver el producto que más se vende en un periodo de tiempo determinado.

CAPITULO VII: PRUEBAS Y VALIDACIONES

CAPITULO 7: PRUEBAS Y VALIDACIONES

Las pruebas de validación en la ingeniería de software son el proceso de revisión que verifica que el sistema de software producido cumple con las especificaciones y que logra su cometido. La validación es el proceso de comprobar que lo que se ha especificado es lo que el usuario realmente quería.

Se debe asegurar que el software cumple las expectativas del cliente. Es importante llevar a cabo la validación de los requerimientos del sistema de forma inicial, intermedia y final. Es fácil cometer errores y omisiones durante la fase de análisis de requerimientos del sistema y, en tales casos, el software final no cumplirá las expectativas de los clientes.

En este capítulo comprobaremos que la implementación ha sido correcta y que todo funciona correctamente. Como tenemos dos partes diferenciadas, debemos validar ambas partes para que finalmente los datos de los informes sean los correctos.

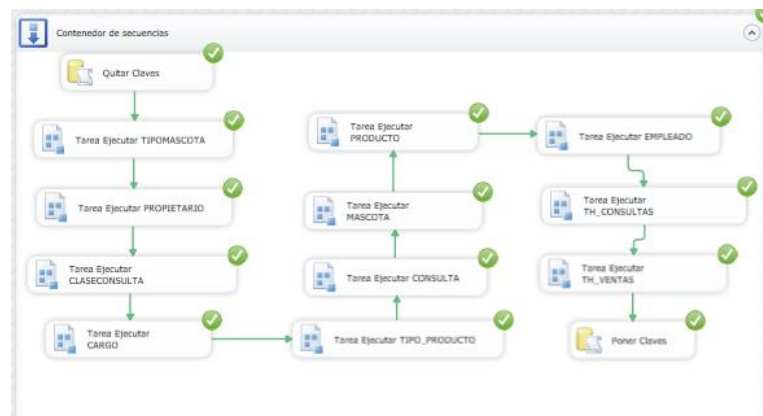
Por un lado validaremos y comprobaremos que el proceso de carga de información al Data Warehouse ha sido correcto, que los datos tienen consistencia e integridad, y lo más importante, que la información de entrada es igual que la de salida. Cada dato debe ir donde corresponde en la nueva base de datos.

Por otro lado, debemos comprobar que lo que se ha “pintado” en los informes finales se corresponde con la información del Data Warehouse. Estas comprobaciones se deben hacer mediante consultas sql a la base de datos, puesto que la información contenida en informes son totales de importes y porcentajes de estos importes.

7.1- Validaciones a nivel de Carga de Datos (Etl)

Como ya hemos visto, para lanzar todas las ETLs, se ha creado un flujo de carga con todas las ejecuciones. En principio este flujo de carga ha resultado correcto, por lo que deducimos que las ETLs han realizado su función, aun así, vamos a realizar una serie de comprobaciones para cerciorarnos de que la información de entrada (operacional) es la misma que de salida (Data Warehouse) y que el nuevo esquema se ha cargado de forma correcta.

Ilustración 104. Flujo de Control; Flujo de Carga Completo

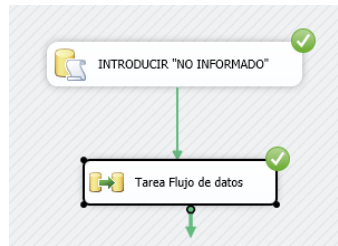


7.1.1- Dimensiones

Comenzaremos a comprobar por las dimensiones. Todas ellas tienen un flujo de control realizado del mismo modo para introducir los “No informados” como registros en la carga de datos. La carga de los “No Informados” se realiza con anterioridad al resto de registros, por lo que deberán aparecer en la base de datos en primer lugar.

En todas las dimensiones aparece como correctamente cargados y creados. Los deberemos comprobar.

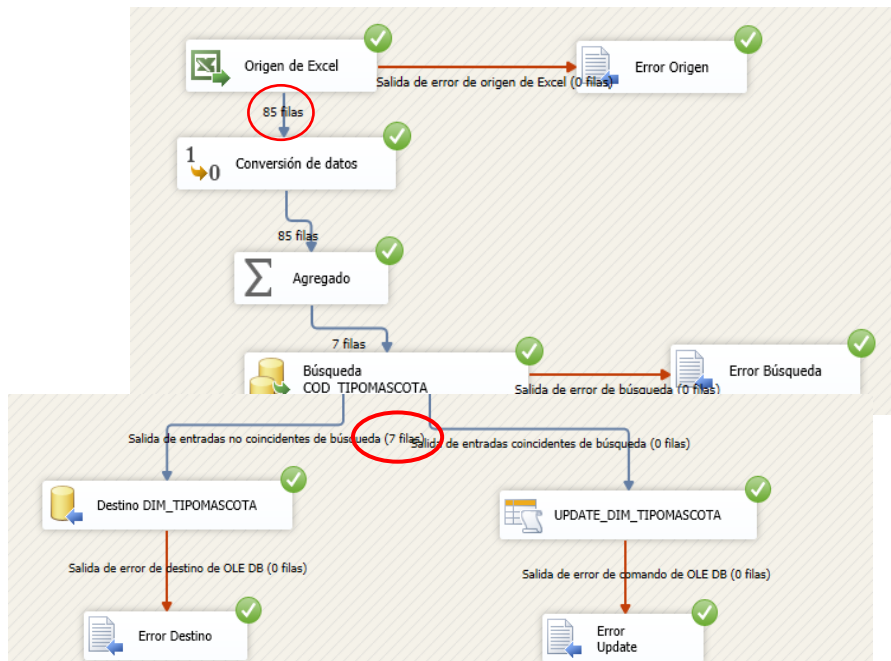
Ilustración 105. Flujo de Control; Inserción de “No Informados”



- **DIM_TIPOMASCOTA**
 - **Validación Flujo de Datos**

Tras la comprobación del buen comportamiento del flujo control, comprobamos que el flujo de carga ha ido bien. En principio no ha habido ningún error en la carga:

Ilustración 106. Flujo de Carga; DIM_TIPOMASCOTA



Si comprobamos el origen de información como en la foto, filtrando por cod_tipo en el Excel y ordenamos, comprobamos cuantos tipos de mascotas hay en realidad y cuantos se deberían cargar. Tenemos suerte porque el cod_tipo es autoincremental y podemos saber con facilidad todos los tipos existentes:

Ilustración 107. Excel; Tipos de Mascotas

HA_DEFUN	COD_TIP	TIPO
		1 perro
		1 perro
		1 perro
		1 perro
		1 perro
		1 perro
		1 perro
		1 perro
20/01/2014		1 perro
		1 perro
		1 perro

	3 ave
	3 ave
	3 ave
	4 conejo
	4 conejo
	4 conejo
	4 conejo
	4 conejo
	4 conejo
	4 conejo
016	5 hamster
	5 hamster
	5 hamster
	5 hamster
	5 hamster
	6 hurón
	6 hurón
	7 cerdo

Como vemos en el origen de datos, hay 7 tipos de mascotas diferentes. En la imagen de la ETL, podemos ver que se cargan un total de 7 registros en la base de datos. En principio, todo ha ido correcto.

○ **Validación Carga de Datos**

Veamos ahora que la carga de datos ha sido correcta para la DIM_TIPOMASCOTA. Como hemos visto, en el origen de los datos, tenemos un total de 7 tipos diferentes, y en la ETL, terminan 7 registros. Debemos comprobar cuantos registros se han cargado en la realidad.

Ilustración 108. Data Warehouse; DIM_TIPOMASCOTA

```

SQLQuery1.sql - WI...Administrador (52) X
/***** Script para el comando SelectTopNRows de SS
SELECT TOP 1000 [ID_TIPOMASCOTA]
, [COD_TIPOMASCOTA]
, [TIPOMASCOTA]
FROM [DWH VETERINARIA].[dbo].[DIM TIPOMASCOTA]
    
```

ID_TIPOMASCOTA	COD_TIPOMASCOTA	TIPOMASCOTA
1	-1	no informado
2	7	cerdo
3	3	ave
4	6	hurón
5	1	perro
6	4	conejo
7	2	gato
8	5	hamster

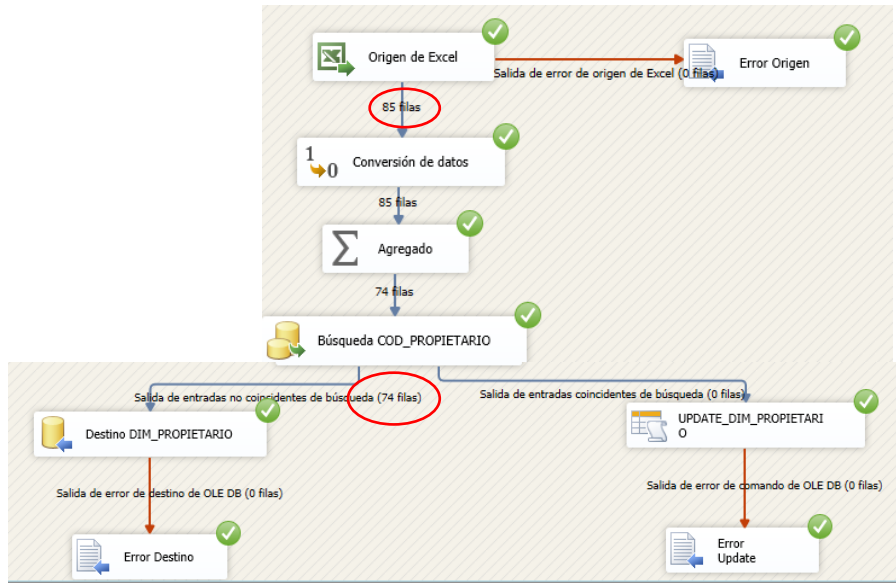
Como se presume en la imagen, el dato “no informado” aparece en primer lugar y el resto de registros posteriormente. Como los identificadores son incrementales y el número de diferentes registros debe ser 7 más el registro “no informado” deben aparecer 8 datos, y por tanto, debe haber hasta 8 identificadores.

Podemos decir finalmente que la carga de datos ha sido la correcta.

- **DIM_PROPIETARIO**
 - **Validación Flujo de Datos**

Analizaremos ahora a los propietarios, estos son más complicados de analizar porque hay más número de propietarios de animales que tipos de mascotas.

Ilustración 109. Flujo de Carga; DIM_PROPIETARIO



El flujo de carga de datos, aunque comienza con 85 registros, finaliza con 74 filas cargadas, esto es 74 registros diferentes, puesto que con el “agregado” agrupamos los registros iguales. Comprobamos cuantos registros diferentes existen en el origen:

Ilustración 110. Excel; Propietarios

CODIGO_PROPIETARIO	PROPIETARIO	Dirección	Teléfono
1	Aceves Olmos, Félix	C/ Salida ronda, 9, 1º, 40008, Segovia	998556622
2	Guijarro Sanz, Oscar	C/ Caño Grande, 25, 4A, 40102, Segovia	975826481
3	Martín Rubio, Gema	C/ Carretera La Granja, 25, 5B, 40100, Segovia	948522117
3	Martín Rubio, Gema	C/ Carretera La Granja, 25, 5B, 40100, Segovia	948522117
4	García Ferrado, Santiago	C/ Los Rosales, 15, 3B, 40100, Segovia	958648572
4	García Ferrado, Santiago	C/ Los Rosales, 15, 3B, 40100, Segovia	958648572
5	Domingo Morgado, Fátima	C/ Tomás Peña, 17, bajo, 40100, Segovia	958647125
6	López Escolar, Miguel Angel	C/ Polar, 31, 1A, 40003, Segovia	918545214
7	Mateo Matesanz, Ernesto	C/ Claveles. 10.3C. 40100. Segovia	925871156
87	Cebrian Pineda, Irene	C/ Laiga, 27, bajo A, 40102, Segovia	914736130
68	Zamora Agüado, Francisco	C/ Florón, 24, bajo A, 40008, Segovia	945728145
69	Parra Serna, Javier	C/ San Fernando, 36, 3 derecha, 40101, Segovia	945871145
70	Hereder de la Fuente, Aleje	C/ Eutiquia Martín, 13, bajo A, 40005, Segovia	998765548
71	Pérez Gallego, Pablo	C/ Los Marqueses, 21, 5B, 40102, Segovia	954712465
72	Escolar Llanos, Mateo	C/ Melocotón, 31, bajo D, 40008, Segovia	914758466
73	Castaño Cuesta, Eva	Avda De las Americas, 19, 4A, 40101, Segovia	953328471
74	Garzón Sancho, Alberto	D/ Miguel de la Fuente, 37, 1A, 40003, Segovia	935419873

Como vemos en las imágenes, en el origen, los códigos de los propietarios también es autoincremental y los duplicados tienen el mismo código, esto nos facilita el trabajo a la hora de las comprobaciones. Tenemos por tanto, un total de 74 propietarios diferentes y en le ETL se cargan 74 registros, a primera vista, todo parece correcto.

○ Validación Carga de Datos

Comprobamos ahora en la base de datos que se han cargado los 75 registros más el dato “no informado”, es decir, la carga ha tenido que ser de un total de 75 registros en el Data Warehouse.

Ilustración 111. Data Warehouse; DIM_PROPIETARIO; “No Informado”

```

/***** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP 1000 [ID_PROPIETARIO]
, [COD_PROPIETARIO]
, [NOMBRE]
, [DIRECCION]
, [TELEFONO]
FROM [DWH_VETERINARIA].[dbo].[DIM_PROPIETARIO]
  
```

ID_PROPIETARIO	COD_PROPIETARIO	NOMBRE	DIRECCION	TELEFONO
1	-1	no informado	no informado	0
2	46	Roldán Gómez, Mario	Cerrotera La Grana, 2, 6A, 40003, Segovia	924756148
3	39	Tardón Aranz, Leticia	Avda de las Flores, 14, 1C, 40101, Segovia	991554872
4	49	Pelaez Galicia, Abel	C/ Los Molinos, 6, bajo, 40101, Segovia	984576581
5	65	Camión Diez, Inés	C/ Falsa, 12, 3A, 40008, Segovia	997581145
6	66	Aragón Peña, Lucía	Avda Constitución, 12, 3D, 40005, Segovia	914756845
7	45	Miguel Calle, Roberto	C/ Leopoldo el Grande, 24, bajo, 40008, Segovia	915758847
8	20	Vega Serrano, Felipe	C/ Escopeta, 3, 3B, 40003, Segovia	971548475
9	72	Escobar Llanos, Mateo	C/ Melocotón, 31, bajo D, 40008, Segovia	914758466

El dato “no informado” aparece en primer lugar y cargado. Veamos cuantos registros existen:

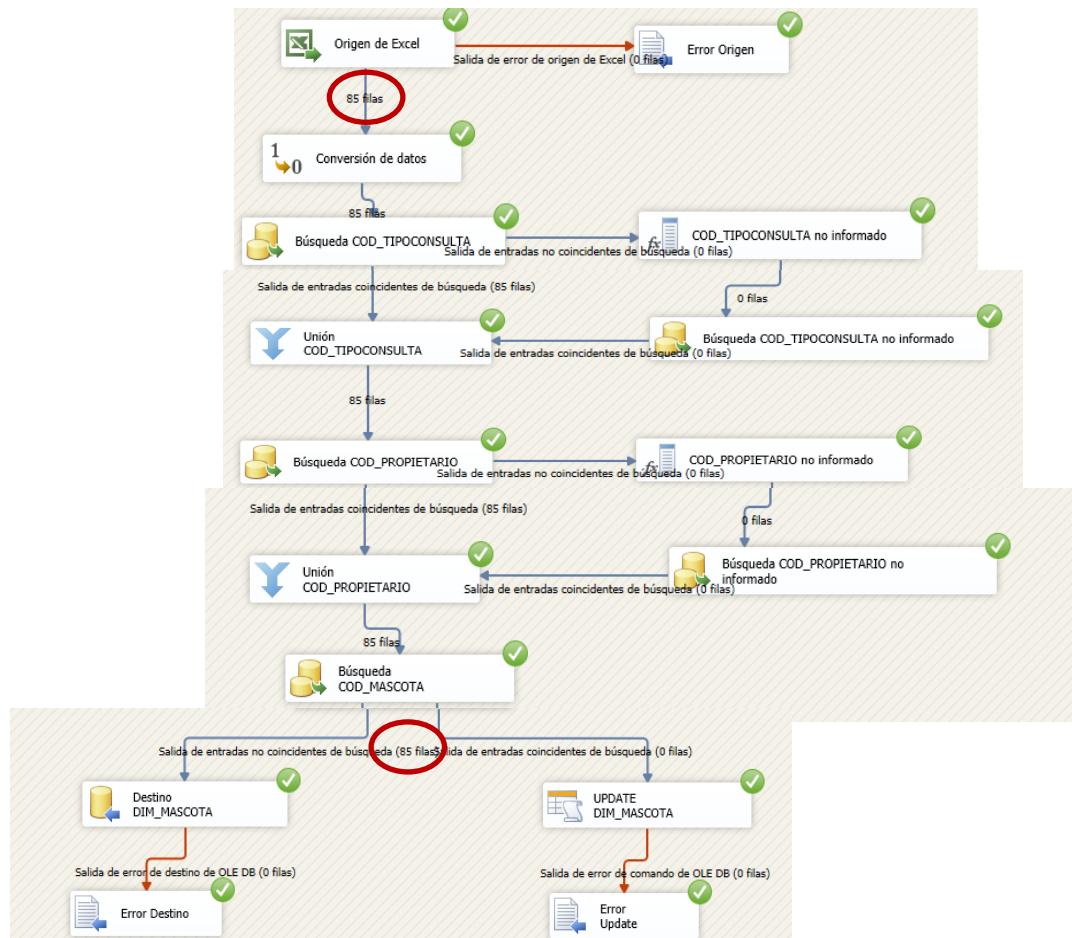
Ilustración 112. Data Warehouse; DIM_PROPIETARIO; Total

ID_PROPIETARIO	COD_PROPIETARIO	NOMBRE	DIRECCION	TELEFONO
63	16	Muñoz Yusta, Héctor	C/ El Pastelero, 6, 2C, 40005, Segovia	974845561
64	9	Escribano García, Luis	C/ Ovidio Aceves, 23, bajo C, 40008, Segovia	915768248
65	67	Cebrian Pinela, Irene	C/ Larga, 27, bajo A, 40102, Segovia	914758156
66	53	Diez Gil, Ricardo	C/ Sonsoto, 37, 4B, 40005, Segovia	945876591
67	25	Sema Aguilar, Hugo	C/ Santa Teresa, 21, 3B, 40003, Segovia	914578456
68	5	Domingo Morgado, Fátima	C/ Tomás Peña, 17, bajo, 40100, Segovia	958647125
69	60	Ramos Hernández, Andrés	C/ las Monjas, 37, 4B, 40101, Segovia	947555481
70	12	Rubio De Dios, Inmacula...	C/ Las Ranas, 32, 2A, 40008, Segovia	971486584
71	21	Plaza Heredero, Miguel	C/ Ave del Paraíso, 21, 4C, 40101, Segovia	956887154
72	50	Pinilla Rujas, María	C/ Vía Unida, 29, 3C, 40003, Segovia	984568715
73	40	Santos Ruiz, Eduardo	C/ Gran Vía, 10, 1D, 40100, Segovia	948114495
74	38	González Santos, Adriana	C/ La Paloma, 28, 5A, 40102, Segovia	982546587
75	32	Barbado Vega, Rubén	Plaza León, 14, 4A, 40101, Segovia	925415687

La carga de datos ha sido correcta: aparecen 75 registros incluido el “No Informado”.

- DIM_MASCOTA
 - Validación Flujo de Datos

Ilustración 113. Flujo de Carga; DIM_MASCOTA



El origen de datos recoge un total de 85 filas, esto es que existen 85 mascotas diferentes. La ETL termina con la carga de 85 registros. Comprobamos el origen si hay en realidad 85 animales.

Ilustración 114. Excel; Mascotas

	A	B	C	D	E	CC
	CODIGO_M	NOMBRE	RAZA	SEX	PESO	ET
1	ASCOTA					
2	101	Thor	Labrador	M	28,50	
3	102	Lola	Pit Bull	H	27,40	
4	103	Carla	Mestizo	H	8,50	
5	104	Santi	Setter Inglés	M	21,31	
6	105	Linka	Husky	H	24,42	
7	106	Nala	Chow Chow	H	23,10	
...
77	405	Blanquita	Holandes	H	1,55	
78	406	Socrates	Toy	M	1,24	
79	501	Sugar	Ruso	M	0,35	
80	502	Chestnut	Chino	M	0,32	
81	503	Speedy Gonzales	Ruso	M	0,41	
82	504	Perez	Ruso	M	0,30	
83	505	Quesito	Campbell	M	0,29	
84	601	Tyson	Hurón	M	2,50	
85	602	Duque	Hurón	M	1,75	
86	701	Congo	vietnamita	M	42,63	
87						

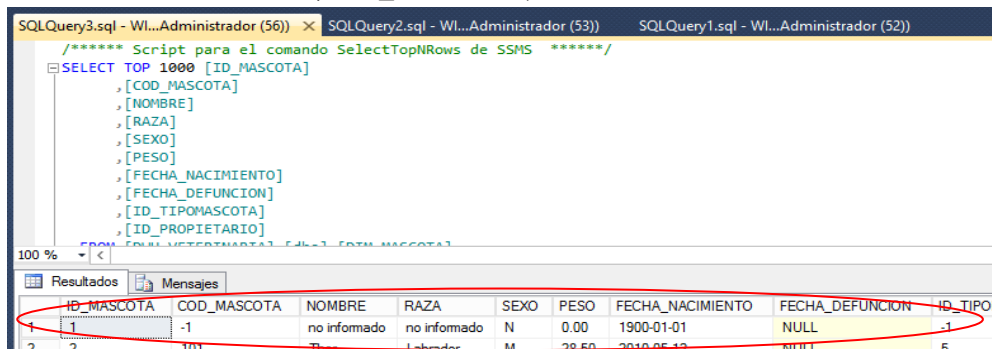
En el caso de las mascotas, tienen unos códigos que no son autoincrementales, pero esto no es problema, en la parte izquierda, el Excel nos dice cuantas filas hay ocupadas. Como podemos ver, hay un total de 86 filas ocupadas en el Excel, pero la primera fila está dedicada a los encabezados, por lo que en realidad existen 85 registros que se deben cargar.

En principio, a simple vista, no existen errores de la carga.

○ **Validación Carga de Datos**

Vamos ahora con la base de datos, y comprobaremos si la carga se ha realizado correctamente:

Ilustración 115. Data Warehouse; DIM_MASCOTA; "No Informado"



La carga del “no informado” ha sido correcta.

Ilustración 116. Data Warehouse; DIM_PROPIETARIO; Total

	ID_MASCOTA	COD_MASCOTA	NOMBRE	RAZA	SEXO	PESO	FECHA_NACIMIENTO	FECHA_DEFUNCION	ID_TIPO
78	78	309	Romina	Cacatua	H	0.95	2008-04-23	NULL	3
79	79	310	Mika	Periquito	H	0.40	2008-10-26	NULL	3
80	80	120	Rex	Pastor Ale...	M	30.45	2002-06-08	NULL	5
81	81	505	Quesito	Campbell	M	0.29	2016-02-14	NULL	8
82	82	303	Fido	Periquito	M	0.30	2010-07-10	NULL	3
83	83	123	Cora	Pomerania	H	2.43	2009-11-15	NULL	5
84	84	122	Gala	Pastor Ale...	H	26.35	2005-04-23	NULL	5
85	85	220	Homer	Mestizo	M	2.89	2015-03-17	NULL	7
86	86	502	Chestnut	Chino	M	0.32	2015-06-30	NULL	8

Hay un total de 86 registros diferentes en la DIM_MASCOTAS, esto es, 85 mascotas distintas, más el datos no informado, son 86 registros distintos. La carga de información ha sido la correcta.

Esta no es la única comprobación que debemos realizar, en esta dimensión tenemos el identificador del propietario y del tipo de mascota a la que hace referencia. Vamos a comprobar que esta información se ha cargado correctamente también.

Ilustración 117. Data Warehouse; DIM_MASCOTA; ID_PROPIETARIO

	COD_MASCOTA	NOMBRE	RAZA	SEXO	PESO	FECHA_NACIMIENTO	FECHA_DEFUNCION	ID_TIPO	ID_PROPIETARIO
1	-1	no informado	no informado	N	0.00	1900-01-01	NULL	-1	-1
2	101	Thor	Labrador	M	28.50	2010-05-12	NULL	5	54
3	115	Snoopy	Caniche	M	9.47	2008-01-27	NULL	5	44
4	308	Reina	Canario	H	0.15	2013-04-27	NULL	3	44
5	315	R...	C...	H	7.18	2013-04-27	NULL	7	25

Vamos a comprobar a Thor, que tiene como tipo mascota el valor 5 y el id_propietario es el 54.

Si nos vamos a la dimensión del tipo de mascota, vemos que el valor 5 corresponde con un perro, y Thor es un perro, por lo que los datos son correctos para este identificador.

Ilustración 118. Data Warehouse; DIM_CLASEMASCOTA; ID_MASCOTA

2	2	7	cerdo
3	3	3	ave
4	4	6	hurón
5	5	1	perro
6	6	4	conejo

Origen:

Ilustración 119. Excel; Mascotas; Propietario

MASCOTA	NOMBRE	RAZA	SEXO	PESO	ETARIO	PROPIETARIO	DIR	TELEFONO	NTO	CIEN	U	TIPO
101	Thor	Labrador	M	28,50		1 Aceves Olmos, Félix	C/ Salic	998556622		12/05/10		1 perro
102	Lola	Pit Bull	H	27,40		2 Guijarro Sanz, Oscar	C/ Cañic	975826481		04/10/11		1 perro

Veamos el propietario, que como comprobamos en el origen, se llama Félix.

Ilustración 120. Data Warehouse; DIM_PROPIETARIO; ID_PROPIETARIO

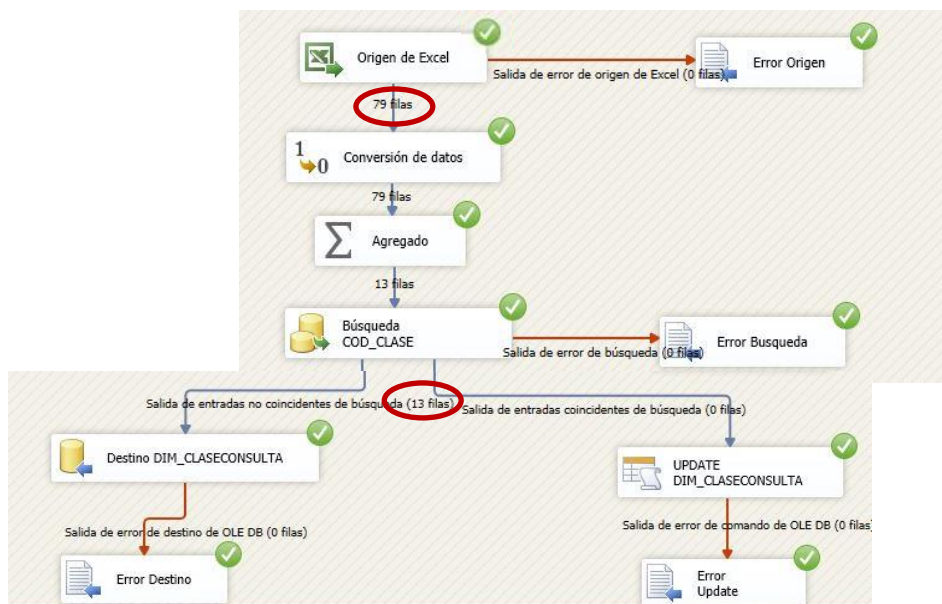
49	49	37	Herrero Serrano, M.ª Pilar	C/ Vía Unida, 13, 1C, 40003, Segovia	961548725
50	50	73	Castañó Cuesta, Eva	Avda De las Americas, 19, 4A, 40101, Segovia	953328471
51	51	54	Carreras Fernández, Álv...	Travesía Encantadora, 17, 3A, 40005, Segovia	925894781
52	52	70	Herederro de la Fuente, ...	C/ Eutiquia Martín, 13, bajo A, 40005, Segovia	998765548
53	53	2	Guijano Sanz, Oscar	C/ Caño Grande, 25, 4A, 40102, Segovia	975826481
54	54	1	Aceves Olmos, Félix	C/ Salida ronda, 9, 1º, 40008, Segovia	998556622
55	55	52	Yuste Rojas, Diana	C/ El Caño, 26,3A, 40101, Segovia	945876185
56	56	19	Sanz Riaño, Irene	C/ Afrodita, 27, 1C, 40102, Segovia	921547884
57	57	71	Pérez Gallego, Pablo	C/ Los Marqueses, 21, 5B, 40102, Segovia	954712465

Efectivamente los datos son correctos, los indicadores que hacen referencia a otras dimensiones, también se han cargado correctamente.

- **DIM_CLASECONSULTA**
 - **Validación Flujo de Datos**

La carga del registro “no informado” no da, a priori, ningún tipo de error. Y la carga del flujo de datos, tampoco parece que de errores:

Ilustración 121. Flujo de Carga; DIM_CLASECONSULTA



De las 79 filas que se cargan en el origen de datos, finalmente llegan a su destino un total de 13. Esto indica que había datos repetidos, cosa normal, porque la clase consulta refiere a tipos de consultas.

Veamos el origen de datos:

Ilustración 122. Excel; Consultas

	CODIGO_CONSULTA	DESCRIPCION	PRECIO	COD_CLASE	CLASE_CONSULTA
1					
2	1409	Tumor Mamario	232,00	14	Cirugía General
3	1410	Mastectomía	300,00	14	Cirugía General
4	1411	Uretrostomía	375,00	14	Cirugía General
5	1702	Ovariohisterectomía	300,00	17	Cirugía Reproductiva
6	1703	Cesárea	400,00	17	Cirugía Reproductiva
77	1608	Proceso coronoides	610,00	16	Cirugía Ortopédica
78	1609	Trocleooplastia/cresta tibial	540,00	16	Cirugía Ortopédica
79	1610	Higroma de codo (penrose)	275,00	16	Cirugía Ortopédica
80	1701	Esterilización	300,00	17	Cirugía Reproductiva
81					

Como vemos en las imágenes, hay un total de 80 filas de consultas, menos el encabezado, son 79 registros, son los que entran en el origen de datos, en principio, todo correcto.

Se cargan 13 filas pero, aparentemente, hay 17 clases de consultas diferentes. Esto es porque, aunque lo parezca, el código designado a las clases de consulta no es autoincremental.

La nomenclatura utilizada por la clínica es diferente dependiendo del tipo de consultas:

- Las Consultas Generales (cod_clase 01) están numeradas del 0101 al 0108
- Los Diagnósticos (Cod_clase 02) están numerados de 0201 al 0206
- Las Pruebas de Laboratorio (Cod_clase 03) están numerados de 0301 al 0306
- Los Tratamientos Preventivos (Cod_clase 04) están numerados de 0401 al 0408
- Los Procedimiento clínico (Cod_clase 05) están numerados de 0501 al 0512
- Las cirugías (Cod_clase 1XXX) están numerados según el tipo que sean:
 - o Cirugía Cardiovascular (Cod_Clase 10) y cod_consulta 1001
 - o Cirugía de oído (Cod_Clase 11) y Cod_Consulta del 1101 al 1103
 - o Cirugía Dental (Cod_Clase 12) y Cod_Consulta 1201
 - o Cirugía Digestiva (Cod_Clase 13) y Cod_Consulta del 1301 al 1304
 - o Cirugía General (Cod_Clase 14) y Cod_Consulta del 1401 al 1411
 - o Cirugía Ocular (Cod_Clase 15) y Cod_Consulta del 1501 al 1505
 - o Cirugía Ortopédica (Cod_Clase 16) y Cod_Consulta del 1601 al 1610
 - o Cirugía Reproductiva (Cod_Clase 17) y Cod_Consulta del 1701 al 1704

Si hacemos recuento, en realidad existen solo 13 clases de consultas diferentes, que son las que realmente llegan al destino en la ETL.

○ **Validación Carga de Datos**

Verificamos la carga de información en la base de datos:

Ilustración 123. Data Warehouse; DIM_CLASECONSULTA

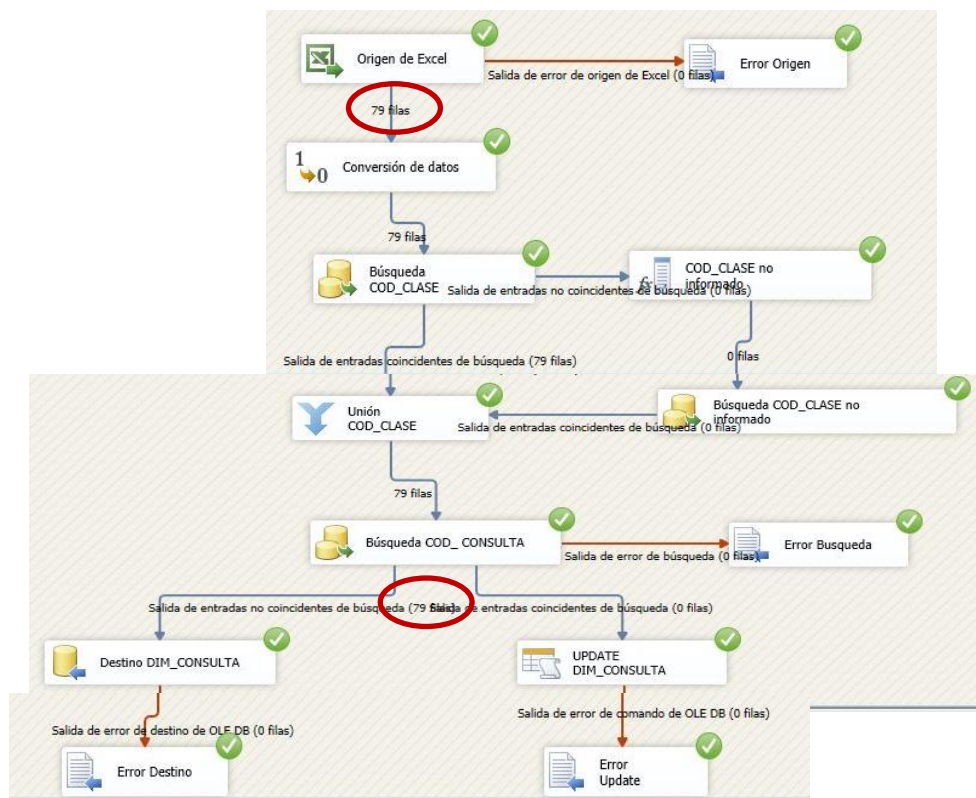
ID_CLASE	COD_CLASE	CLASE_CONSULTA
1	-1	no informado
2	16	Cirugia Ortopedica
3	14	Cirugia General
4	1	Consultas
5	13	Cirugia Digestiva
6	3	Pruebas de Laboratorio
7	2	Diagnóstico
8	15	Cirugia Ocular
9	10	Cirugia Cardiovascular
10	5	Procedimiento Clínico
11	11	Cirugia del Oído
12	12	Cirugia Dental
13	4	Tratamientos Prevent...
14	17	Cirugia Reproductiva

Como vemos en el destino, se cargan 14 registros, de los cuales, el primero corresponde al “no informado”. En principio, todo es correcto.

● **DIM_CONSULTA**

○ **Validación Flujo de Datos**

Ilustración 124. Flujo de Carga; DIM_CONSULTA



Vemos que desde el origen pasan 79 registros y hasta el final de la ETL llegan los 79 registros. Originalmente, deberían aparecer 79 consultas diferentes:

Ilustración 125. Excel; Consultas

	CODIGO_CONSULTA	DESCRIPCION	PRECIO	COD_CLASE	CLASE_CONSULTA
1					
2	1409	Tumor Mamario	232,00	14	Cirugía General
3	1410	Mastectomía	300,00	14	Cirugía General
4	1411	Uretróstomía	375,00	14	Cirugía General
5	1702	Ovariohisterectomía	300,00	17	Cirugía Reproductiva
6	1703	Cesárea	400,00	17	Cirugía Reproductiva
77	1608	Proceso coronoides	610,00	16	Cirugía Ortopédica
78	1609	Trocleooplastia/cresta tibial	540,00	16	Cirugía Ortopédica
79	1610	Higroma de codo (penrose)	275,00	16	Cirugía Ortopédica
80	1701	Esterilización	300,00	17	Cirugía Reproductiva
81					

Existen, como origen de datos, 80 filas, pero si restamos el encabezado, hay un total de 79 consultas diferentes.

○ **Validación Carga de Datos**

Nos cercioramos que el resultado de la carga de información ha sido satisfactoria:

Ilustración 126. Data Warehouse; DIM_CONSULTA; “No Informados”

```

/***** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP 1000 [ID_CONSULTA]
, [COD_CONSULTA]
, [DESCRIPCION]
, [PRECIO]
, [ID_CLASE]
FROM [DWH_VETERINARIA1].[dbo].[DIM_CONSULTA]
    
```

ID_CONSULTA	COD_CONSULTA	DESCRIPCION	PRECIO	ID_CLASE
1	-1	no informado	0.00	-1
2	1001	Cirugía Cardiovascular	630.00	9
3	1101	Espigas auriculares	92.00	11
4	1102	Otohematoma	185.00	11
5	1103	Omiectomía	171.00	11

La carga del “no informado” es correcta y el registro aparece en primer lugar.

Ilustración 127. Data Warehouse; DIM_CONSULTA; Total

73	73	511	Sondaje	160.00	10
74	74	1702	Ovariohisterectomía	300.00	14
75	75	1703	Cesárea	400.00	14
76	76	512	Citología	33.00	10
77	77	1704	Criptóquido	300.00	14
78	78	1409	Tumor Mamario	232.00	3
79	79	1410	Mastectomía	300.00	3
80	80	1411	Uretróstomía	375.00	3

Como se ha cargado el “no informado” con el identificador 1, si sumamos 79 registros más, hacen un total de 80 registros. Los datos parecen correctos.

Vamos a realizar una última comprobación:

Ilustración 128. Data Warehouse; DIM_CONSULTA; ID_CLASE

	ID_CONSULTA	COD_CONSULTA	DESCRIPCION	PRECIO	ID_CLASE
50	50	302	Análisis Orina	32.37	6
51	51	303	Análisis Coprológico	38.36	6
52	52	304	Cultivo y antibiograma	32.00	6

Escogemos una consulta al azar, el cod_consulta 303 corresponde con el “Análisis Coprológico”, tiene un precio de 38,36 y pertenece al id_clase 6

Ilustración 129. Excel; Consultas; Código_Conсульта

	A	B	C	D	E
	CODIGO_CONSULTA	DESCRIPCION	PRECIO	COD_CLASE	CLASE_CONSULTA
1					
24	0303	Análisis Coprológico	38,36	03	Pruebas de Laboratorio
81					

En el origen parece todo correcto, veamos la clase consulta:

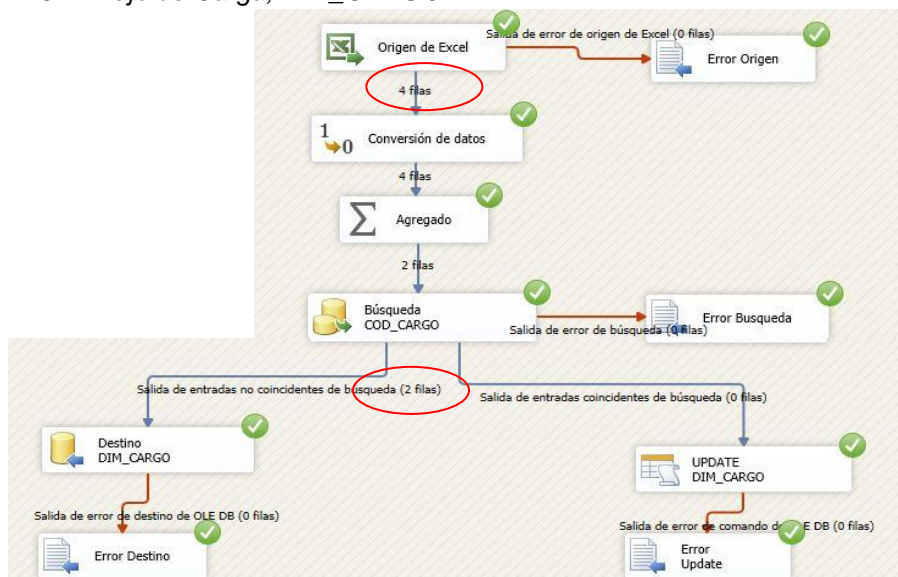
Ilustración 130. Data Warehouse; DIM_CLASECONSULTA; ID_CLASE

	ID_CLASE	COD_CLASE	CLASE_CONSULTA
1	1	-1	no informado
2	2	16	Cirugía Ortopédica
3	3	14	Cirugía General
4	4	1	Consultas
5	5	13	Cirugía Digestiva
6	6	3	Pruebas de Laboratorio

El identificador 6 corresponde con “Pruebas laboratorio”, por lo que los datos son correctos y las referencias también.

- **DIM_CARGO**
 - **Validación Flujo de Datos**

Ilustración 131. Flujo de Carga; DIM_CARGO



Del origen de datos salen 4 filas que son las que realmente tenemos en la hoja de cálculo. Pero al destino solo llegan 2 filas, esto es porque tan solo hay dos cargos diferentes:

Ilustración 132. Excel; Empleados

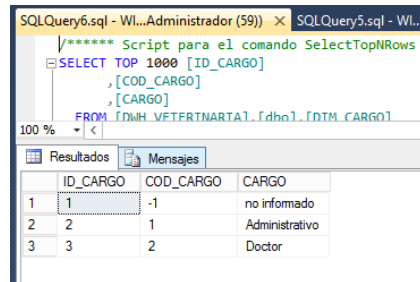
	A	B	C	D	E	F	G
	CODIGO_EMPLEAD	NOMBRE	APELLIDOS	TELÉFONO	COD_CARG	CARGO	
1							
2	1	Leticia	Aceves Martín	999887711	1	Administrativo	
3	2	Miguel Ángel	Garzón López	999333544	2	Doctor	
4	3	Iván	Sastre García	999228877	1	Administrativo	
5	4	Susana	Alonso Herrero	999115533	2	Doctor	

El proceso de validación en caso de esta dimensión es muy sencillo porque tan solo hay dos datos. Efectivamente, solo hay 2 tipos de cargos: administrativo o doctor.

○ **Validación Carga de Datos**

Vamos a ver el destino de la ETL:

Ilustración 133. Data Warehouse; DIM_CARGO



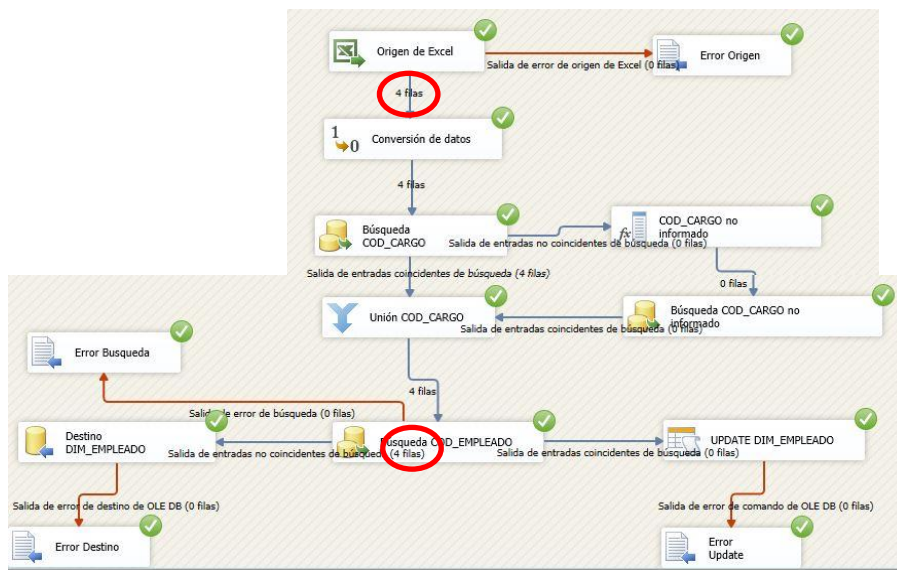
La carga de datos en la base de datos es correcta: tenemos el “no informado”, el doctor y el administrativo.

- **DIM_EMPLEADO**
 - **Validación Flujo de Datos**

La DIM_EMPEADO, la hemos validado, prácticamente, a la vez que la DIM_CARGO:

EL “no informado” se carga correctamente.

Ilustración 134. Flujo de Carga; DIM_EMPLEADO



De las 4 filas que entran en el proceso ETL, se cargan 4 filas, y como hemos comprobado antes en el origen, efectivamente hay 4 trabajadores diferentes:

Ilustración 135. Excel; Empleados

	A	B	C	D	E	F
1	CODIGO_EMPLEADO	NOMBRE	APELLIDOS	TELÉFONO	COD_CARGO	CARGO
2	1	Leticia	Aceves Martín	999887711	1	Administrativo
3	2	Miguel Ángel	Garzón López	999333544	2	Doctor
4	3	Iván	Sastre García	999228877	1	Administrativo
5	4	Susana	Alonso Herrero	999115533	2	Doctor

○ **Validación Carga de Datos**

Verificamos ahora la base de datos:

Ilustración 136. Data Warehouse; DIM_EMPLEADO

```

/***** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP 1000 [ID_EMPLEADO]
, [COD_EMPLEADO]
, [NOMBRE]
, [APELLIDOS]
, [TELEFONO]
, [ID_CARGO]
FROM [DWH_VETERINARIA].[dbo].[DIM_EMPLEADO]
    
```

ID_EMPLEADO	COD_EMPLEADO	NOMBRE	APELLIDOS	TELEFONO	ID_CARGO
1	-1	no informado	no informado	0	-1
2	1	Leticia	Aceves Martín	999887711	2
3	2	Miguel Ángel	Garzón López	999333544	3
4	3	Iván	Sastre García	999228877	2
5	4	Susana	Alonso Herrero	999115533	3

El registro “no informado” aparece en primer lugar y posteriormente los 4 empleados de la clínica, veamos que los datos se corresponden y que Iván es administrativo:

Ilustración 137. Data Warehouse; DIM_CARGO; ID_CARGO

```

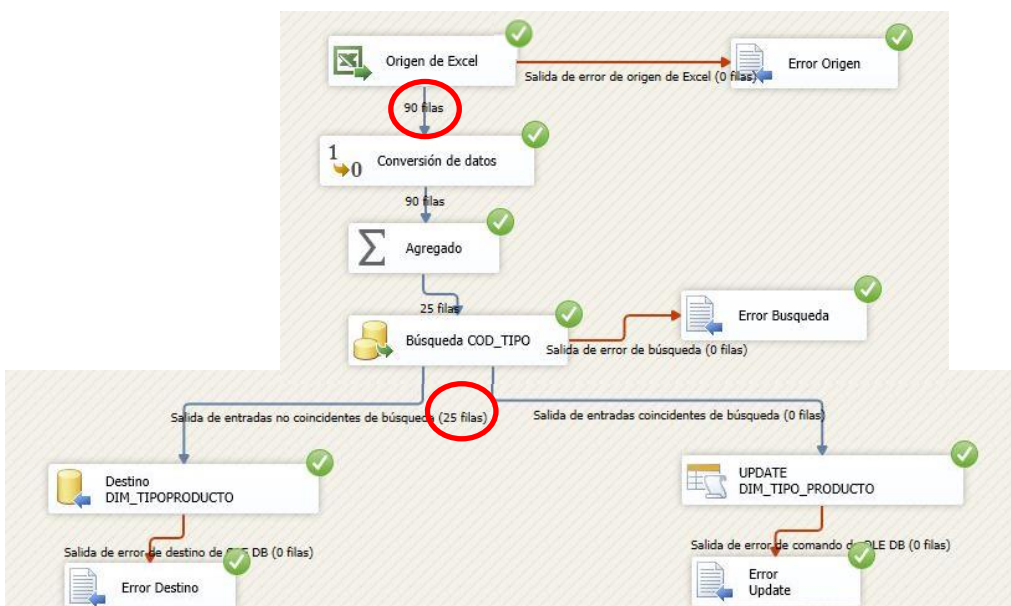
/***** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP 1000 [ID_CARGO]
, [COD_CARGO]
, [CARGO]
FROM [DWH_VETERINARIA].[dbo].[DIM_CARGO]
    
```

ID_CARGO	COD_CARGO	CARGO
1	-1	no informado
2	1	Administrativo
3	2	Doctor

Iván tiene el id_cargo = 2 y efectivamente, es administrativo. Los datos cargados son correctos.

- **DIM_TIPO_PRODUCTO**
 - **Validación Flujo de Datos**

Ilustración 138. Flujo de Carga; DIM_TIPO_PRODUCTO



Entran 90 filas desde el origen de datos y después del agregado, quedan 25 filas. Demostramos que existen en el origen de datos 25 tipos de productos diferentes:

Ilustración 139. Excel; Producto

	A	B	C	D	E	F
	CODIGO_PRODUC	NOMBRE	DESCRIPCION	COD_TIPO	TIPO_PRODUCTO	PRECIO
1						
2	1	Blocan 100 ml	Via oral, directamente en la boca del animal o añadido a la comida o al agua. Uso en casos de estreñimiento crónico o intermitente. (Perros y gatos)	1	Digestivo	4,93
3	2	Daforte Suspensión 20 ml	Compacta las heces, reduciendo la severidad y duración de la diarrea. (Perros)	1	Digestivo	9,18
			Inmunización activa para prevenir infección frente al virus del moquillo, adenovirus tipo 2, parvovirus, virus de parainfluenza			
90	15	CALMEX PERRO 12 CAPSULAS	veterinario, estancias en residencias caninas, etc. (Perros)		24 Tranquilizantes	10,89
			Para preparar al ave para desarrollar su máxima potencia y resistencia a la fatiga. Para mejorar la calidad del canto, la apariencia y el comportamiento del ave en concursos y exposiciones. (Aves)			
91	38	RED CELL AVIAN 500 ML			25 Vitaminas	20,48
92						
93						

Si ordenamos el Cod_tipo, vemos que están numerados de forma autoincremental y que el último tipo de producto tiene el código 25. Por lo tanto, sí existen 25 tipos de productos diferentes en el origen de información.

○ **Validación Carga de Datos**

Verificamos la carga de información en el destino:

Ilustración 140. Data Warehouse; DIM_TIPO_PRODUCTO; “No Informado”

```

SQLQuery8.sql - WI...Administrador (61) x SQLQuery7.sql - WI...Administrador (55)
/***** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP 1000 [ID_TIPOPRODUCTO]
, [COD_TIPOPRODUCTO]
, [TIPO_PRODUCTO]
FROM [DWH_VETERINARIA].[dbo].[DIM_TIPO_PRODUCTO]
    
```

ID_TIPOPRODUCTO	COD_TIPOPRODUCTO	TIPO_PRODUCTO
1	-1	no informado
2	23	Sueros
3	4	Antibioticos
4	10	Arenas
5	11	Articulos
6	5	Anti-Hemorragicos
7	12	Cardiología

El dato “no informado” aparece en primer lugar, y hay 26 registros incluyendo este, por tanto la carga de datos ha sido satisfactoria.

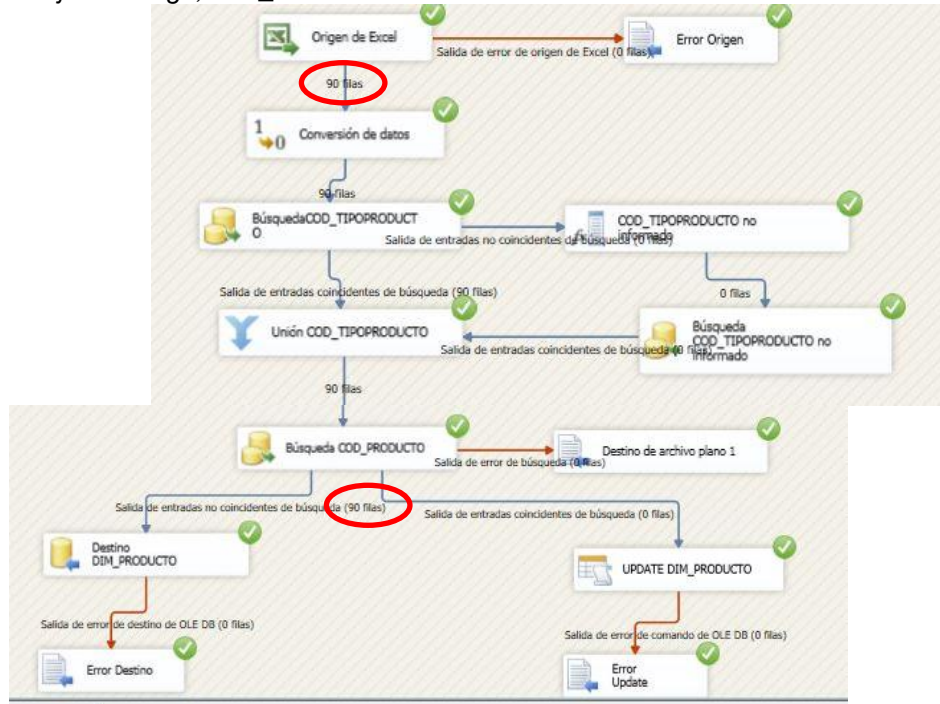
Ilustración 141. Data Warehouse; DIM_TIPO_PRODUCTO; Total

20	20	25	Vitaminas
21	21	14	Desinfectantes
22	22	19	Pienso
23	23	17	Oticos
24	24	6	Anti-Infeciosos
25	25	13	Condoprotectores
26	26	15	Hepáticos

Consulta ejecutada correctamente. WIN2012SERVER (11.0 RTM)

- **DIM_PRODUCTO**
 - **Validación Flujo de Datos**

Ilustración 142. Flujo de Carga; DIM_PRODUCTO



De 90 filas que recoge el origen de datos, se cargan las 90 en el destino, esto quiere decir que no hay información duplicada. Comprobamos:

Ilustración 143. Excel; Producto

CODIGO_PRODUC	NOMBRE	DESCRIPCION	COD_TIPO	TIPO_PRODUCTO	PRECIO
1	TO				
2	1 Blocan 100 ml	Vía oral, directamente en la boca del animal o añadido a la comida o al agua. Uso en casos de estreñimiento crónico o intermitente. (Perros y gatos)	1	Digestivo	4,93
3	2 Daforte Suspensión 20 ml	Compacta las heces, reduciendo la severidad y duración de la diarrea. (Perros)	1	Digestivo	9,18
4	3 CANIGEN 7 10 Dosis	Inmunización activa para prevenir infección frente al virus del moquillo, adenovirus tipo 2, parvovirus, virus de parainfluenza canina y las leptospirosis caninas. (Perros)	2	Vacuna	53,15
5	4 DURAMUNE cachorros DP + C 1 4 DOSIS	reducir la enfermedad causada por el virus del moquillo canino y el parvovirus canino; para reducir la infección intestinal causada por el coronavirus canino. (Perros)	2	Vacuna	6,49
6	5 CANLEISH 1 Dosis	Inmunización activa de perros negativos a leishmania a partir de los 6 meses. (Perros)	2	Vacuna	31,99
89	88 Cama osito para mascotas	Mullida cama con imitación de lana que invita a relajarse. (Perros y Gatos)	11	Artículos	19,99
90	89 Cama Basic para mascotas	Cama confortable de pelo artificial y exterior resistente de material textil, con aplicación en forma de huella, combina muy bien con cualquier decoración y es lavable a máquina. (Perros y Gatos)	11	Artículos	12,99
91	90 Cama Technicolor	Cama resistente con funda de poliéster que repele la suciedad y el agua, con borde acolchado. (Perros y Gatos)	11	Artículos	44,99

Como vemos en la imagen, efectivamente, hay 90 productos diferentes a la venta en la clínica.

○ **Validación Carga de Datos**

Ahora comprobaremos que la carga de datos en la base de datos se ha dado correctamente y sin inconsistencias.

Ilustración 144. Data Warehouse; DIM_PRODUCTO; “No Informado”

```

SELECT TOP 1000 [ID_PRODUCTO]
, [COD_PRODUCTO]
, [NOMBRE]
, [DESCRIPCION]
, [PRECIO]
, [ID_TIPOPRODUCTO]
FROM [DMH_VETERINARIA].[dbo].[DIM_PRODUCTO]
    
```

ID_PRODUCTO	COD_PRODUCTO	NOMBRE	DESCRIPCION	PRECIO	ID_TIP
1	-1	no informado	no informado	0.00	-1
2	1	Blocao 100 ml	Via oral, directamente en la boca del animal o aña...	4.93	14
3	2	Daforte Suspensión 20 ml	Compacta las heces, reduciendo la severidad y du...	9.18	14
4	3	CANIGEN 7 10 Dosis	Inmunización activa para prevenir infección frente ...	53.15	19

El registro “no informado” es correcto y se ha cargado sin errores.

Ilustración 145. Data Warehouse; DIM_PRODUCTO; Total

84	84	83	Bañera para pájaros Hagen Vision	Perfecta para jaulas con pocas puertas o puertas ...	3.99	5
85	85	84	JR Birds Barra de bar	Palo de madera con sano mijo en rama y una delici...	6.99	5
86	86	85	Comedero Savic Crock con fijación de ...	Se sujeta a la jaula con una fijación por tornillo. Po...	5.99	5
87	87	86	Cama Vainilla y Chocolate para mascotas	Muy cómoda y calentita, de tela aterciopelada en ...	12.99	5
88	88	87	Cama Lille para mascotas	Cama original en forma de gato, con la cara borda...	11.99	5
89	89	88	Cama osito para mascotas	Mullida cama con imitación de lana que invita a rel...	19.99	5
90	90	89	Cama Basic para mascotas	Cama confortable de pelo artificial y exterior resiste	12.99	5
91	91	90	Cama Technicolor	Cama resistente con funda de poliéster que repele ...	44.99	5

Hay un total de 91 registros, esto es, los 90 que teníamos en el origen de datos, más el “no informado”. La carga se ha hecho aparentemente bien.

Comprobamos que los tipos se corresponden. Cogemos el producto con código 80, por ejemplo, cuyo precio es 2,49 y su id_tipo es 5:

Ilustración 146. Data Warehouse; DIM_PRODUCTO; Codigo_Producto

ID_PRODUCTO	COD_PRODUCTO	NOMBRE	DESCRIPCION	PRECIO	ID_TIP
80	80	Pala para arenas de gránulos gruesos	La manera sencilla y práctica de limpiar la bandeja ...	0.99	5
81	81	Set de juguetes de gato, con pelotas y ...	Un set de juguetes extragrande para gatos juguetoo...	2.49	5
82	82	Spray quitamanchas de orina de Trixie	Elimina las manchas viejas y persistentes así como...	6.99	5
83	83	Jaula Finca pequeña	Para aves como canarios, periquitos y pinzones ce...	34.99	5
84	84	Bañera para pájaros Hagen Vision	Perfecta para jaulas con pocas puertas o puertas ...	3.99	5
85	85	JR Birds Barra de bar	Palo de madera con sano mijo en rama y una delici...	6.99	5
86	86	Comedero Savic Crock con fijación de ...	Se sujeta a la jaula con una fijación por tornillo. Po...	5.99	5
87	87	Cama Vainilla y Chocolate para mascotas	Muy cómoda y calentita, de tela aterciopelada en ...	12.99	5
88	88	Cama Lille para mascotas	Cama original en forma de gato, con la cara borda...	11.99	5

Ilustración 147. Excel; Producto; Codigo_Producto

CODIGO_PRODUCTO	NOMBRE	DESCRIPCION	COD_TIPO	TIPO_PRODUCTO	PRECIO
81	Set de juguetes de gato, con 80 pelotas y ratones	Un set de juguetes extragrande para gatos juguetones: 12 divertidos juguetes con campanas, cascabeles o suave peluche. (Gatos)	11	Artículos	2,49

Con el origen se corresponde perfectamente, y es un artículo. Confrontamos:

Ilustración 148. Data Warehouse; DIM_TIPO_PRODUCTO; Cod_TipoProducto

ID_TIPOPRODUCTO	COD_TIPOPRODUCTO	TIPO_PRODUCTO
1	-1	no informado
2	23	Sueros
3	4	Antibioticos
4	10	Arenas
5	11	Artículos
6	5	Anti-Hemorragicos

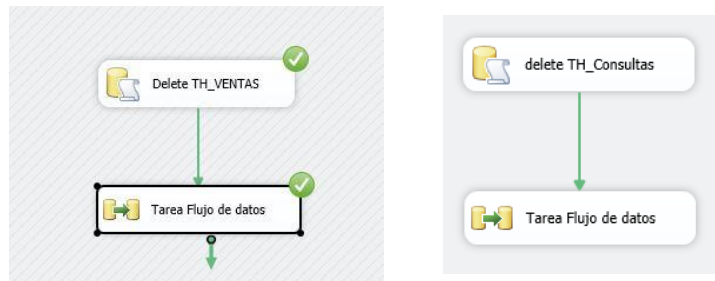
7.1.2- Tablas de Hechos

Para las tablas de hechos, igual que para las dimensiones, tenemos un flujo de carga previo al flujo de la ETL en sí.

Las tablas de hechos no tienen actualizaciones, los datos no se actualizan, esto es lógico puesto que los registros de ventas o consultas realizadas son las que son y las que se han producido y no van a cambiar. Eso sí, en un momento dado, posiblemente se haya que introducir una consulta o venta en una fecha pasada (no es lo avitual), por ello, antes de la ejecución de la ETL se realiza un proceso de borrado, este borrado no será completo debido a que las tablas de hechos pueden contener muchos registros y el borrado de la tabla entera puede tardar más de la cuenta. El borrado se realizará desde una fecha determinada hasta la última fecha en la que se encuentren registros.

De esta forma se asegura la integridad de los datos y la rapidez en la carga de los mismos.

Ilustración 149. Flujo de Datos; Tablas de Hechos

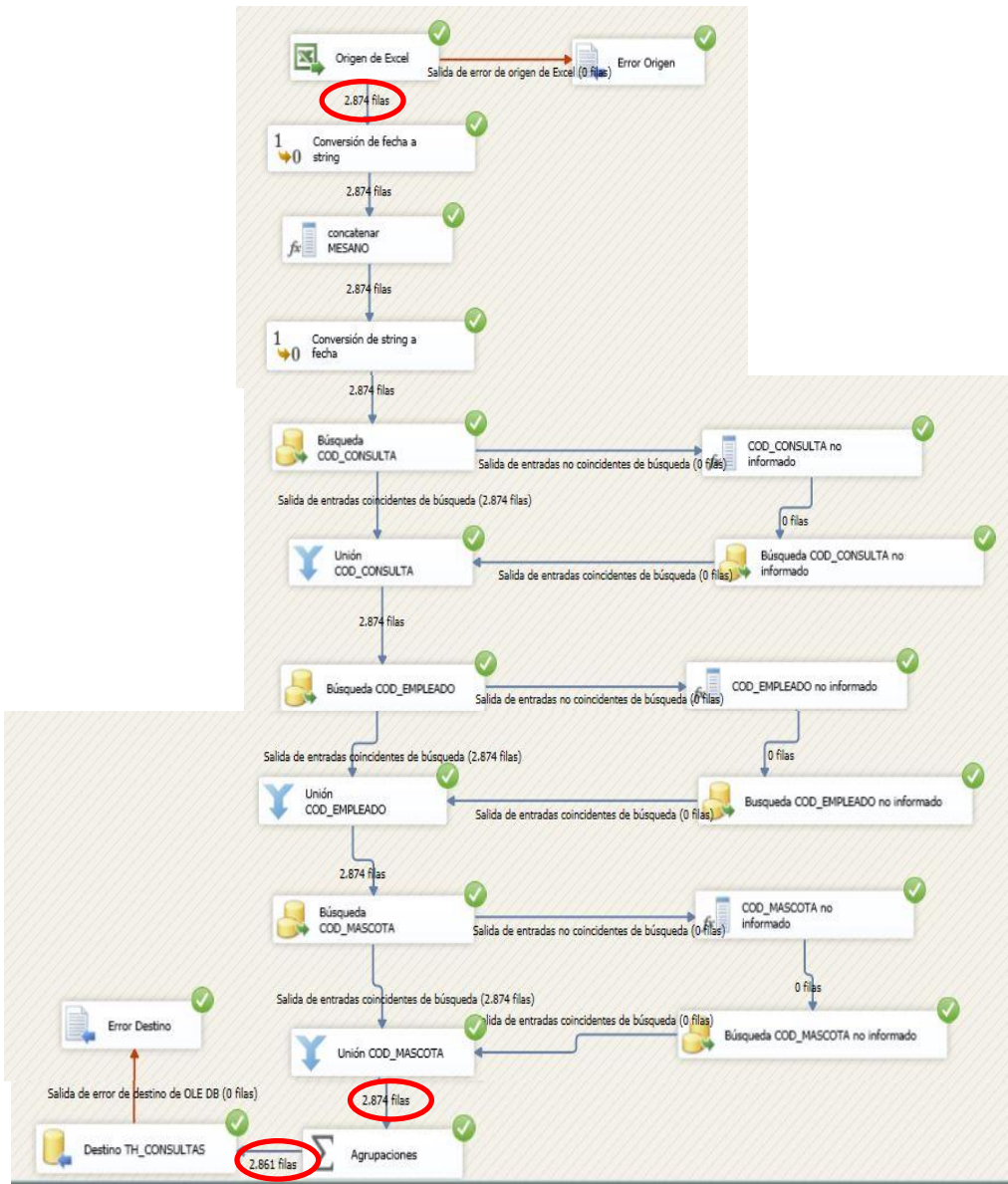


Como vemos en la ilustración, no cargamos el registro “no informado”, esto es lógico, puesto que las tablas de hechos son tablas con datos finales. De estos registros, si faltara algún tipo de dato, este sí aparecería como “No Informado” sin problemas porque en las dimensiones sí que aparece este registro.

- **TH_CONSULTAS**

- **Validación Flujo de Datos**

Ilustración 150. Flujo de Carga; TH_CONSULTAS



Desde el origen de datos, se cogen 2.874 filas, estos registros llegan hasta la agrupación sin ningún problema. Una vez llega al agregado, alguno de los registros se pierde. Esto es así debido a la dimensión temporal, en lugar de recoger datos diarios, los vamos a agrupar en datos mensuales.

En el agregado se agrupan los días y se suman aquellos importes cuyos identificadores sean iguales, es decir, se suman los importes de aquellas consultas que se hayan hecho 2 veces al mismo animal en el mismo mes. Esto hace que se agrupe un registro con otro.

Por lo tanto, finalmente se cargan en la base de datos un total de 2.861 filas o registros.

Veamos los registros que aparecen en el origen:

Ilustración 151. Excel; Diario

	A	B	C	D	E	F
1	FECHA	CODIGO_MASCOTA	CODIGO_EMPLEADO	CODIGO_CONSULTA	TIPO_CONSULTA	IMPORTE
2	01/01/14	406	2	1103	1	205,20
3	01/01/14	210	2	0504	0	28,90
4	01/01/14	217	4	0303	1	46,03
5	01/01/14	309	4	0103	1	30,00
6	01/01/14	118	4	1404	0	395,00
7	03/01/14	110	2	1302	0	420,00
8	03/01/14	130	2	1604	1	516,00
...
2869	29/03/16	103	2	0201	0	72,32
2870	29/03/16	127	4	1603	1	960,00
2871	30/03/16	129	2	1602	1	684,00
2872	30/03/16	504	2	0501	0	135,00
2873	30/03/16	112	4	0202	0	40,95
2874	30/03/16	602	2	0105	1	108,00
2875	31/03/16	209	2	1403	1	234,00

En el documento hay 2.875 datos, si no contamos el encabezado son los 2.874 que recoge la ETL y llegan hasta el agregado. En principio, todo correcto.

○ **Validación Carga de Datos**

Una vez comprobados los datos originales del proceso de carga y que todo parece ser correcto, veamos los datos que realmente están cargados en la base de datos:

Ilustración 152. Data Warehouse; TH_CONSULTAS; Total

```

/***** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP 4000 [ID_MESANO]
,[ID_MASCOTA]
,[ID_EMPLEADO]
,[ID_CONSULTA]
,[TIPO_CONSULTA]
,[IMPORTE_CONSULTA]
FROM [DW_VETERINARIA].[dbo].[TH_CONSULTAS]
    
```

	ID_MESANO	ID_MASCOTA	ID_EMPLEADO	ID_CONSULTA	TIPO_CONSULTA	IMPORTE_CONSULTA
2849	201407	86	3	71	1	78.00
2850	201407	86	5	4	0	185.00
2851	201507	86	5	8	1	504.00
2852	201407	86	5	10	1	546.00
2853	201401	86	5	22	1	212.40
2854	201506	86	5	22	1	212.40
2855	201402	86	5	24	1	600.00
2856	201601	86	5	43	0	72.32
2857	201407	86	5	51	1	46.03
2858	201603	86	5	52	0	32.00
2859	201412	86	5	57	0	31.00
2860	201403	86	5	68	0	85.00
2861	201504	86	5	77	1	360.00

Haremos alguna prueba más:

Ilustración 153. Data Warehouse; TH_CONSULTAS; Importe

	ID_MESANO	ID_MASCOTA	ID_EMPLEADO	ID_CONSULTA	TIPO_CONSULTA	IMPORTE_CONSULTA
583	201502	85	3	28	0	610.00
584	201407	85	3	34	0	300.00
585	201409	85	3	76	0	33.00
586	201505	85	5	25	0	570.00
587	201511	85	5	44	1	49.14
588	201602	85	5	47	1	354.00

Vamos a comparar el dato con importe 570, correspondiente a la mascota 85, y la consulta 25 que es “no urgente” y se realiza en el mes 5 de 2015:

Ilustración 154. Excel; Diario; Importe

1811	06/05/15	701	4	0406	0	30,00
1812	06/05/15	121	2	1701	1	360,00
1813	06/05/15	307	4	1405	1	240,00
1814	07/05/15	220	4	1602	0	570,00
1815	07/05/15	219	2	0402	1	37,20
1816	07/05/15	307	4	1609	0	540,00
1817	07/05/15	601	2	1103	1	205,20
1818	08/05/15	505	4	0507	0	43,00

En el origen se corresponde con la mascota 220, el empleado 4 y la consulta 1602. De momento el importe, las fechas y el tipo de consulta se corresponden.

Ilustración 155. Excel; Diario; Mascota

CODIGO_M	ASCOTA	NOMBRE	RAZA	SEX	PESO	ETAR
57	220	Homer	Mestizo	M	2,89	
87						

Ilustración 156. Excel; Diario; Empleado

CODIGO_EMPLEAD	NOMBRE	APELLIDOS	TELÉFONO	COD_CARG	CARGO
1	Leticia	Aceves Martín	999887711	1	Administrativo
2	Miguel Ángel	Garzón López	999333544	2	Doctor
3	Iván	Sastre García	999228877	1	Administrativo
4	Susana	Alonso Herrero	999115533	2	Doctor
5					

Ilustración 157. Excel; Diario; Consulta

CODIGO_CONSULTA	DESCRIPCION	PRECIO	COD_CLASE	CLASE_CONSULTA
71	1602 Fractura diafisaria fijador	570,00	16	Cirugía Ortopédica
81				

Destino:

Ilustración 158. Data Warehouse; DIM_MASCOTA; ID_MASCOTA

ID_MASCOTA	COD_MASCOTA	NOMBRE	RAZA	SEXO	PESO	F
80	80	120	Rex	Pastor Ale...	M	30.45
81	81	505	Quesito	Campbell	M	0.29
82	82	303	Fido	Periquito	M	0.30
83	83	123	Cora	Pomerania	H	2.43
84	84	122	Gala	Pastor Ale...	H	26.35
85	85	220	Homer	Mestizo	M	2.89
86	86	502	Chestnut	Chino	M	0.32

Se corresponde, veamos con el empleado:

Ilustración 159. Data Warehouse; DIM_EMPLEADO; ID_EMPLEADO

ID_EMPLEADO	COD_EMPLEADO	NOMBRE	APELLIDOS	TELEFONO	ID_CARGO
1	1	no informado	no informado	0	-1
2	2	Leticia	Aceves Martín	999887711	2
3	3	Miguel Ángel	Garzón López	999333544	3
4	4	Iván	Sastre García	999228877	2
5	5	Susana	Alonso Herrero	999115533	3

Corresponde también, ahora veamos la consulta y el precio:

Ilustración 160. Data Warehouse; DIM_CONSULTAS; ID_CONSULTA

ID_CONSULTA	COD_CONSULTA	DESCRIPCION	PRECIO	ID_CLASE
23	23	1505	Tumor del Párpado	200.00
24	24	1601	Fractura diafisaria clavos	500.00
25	25	1602	Fractura diafisaria fijador	570.00
26	26	1603	Fractura diafisaria placa	800.00
27	27	1604	Fractura mandibular	430.00
28	28	1605	Ligamentos cruzados	610.00

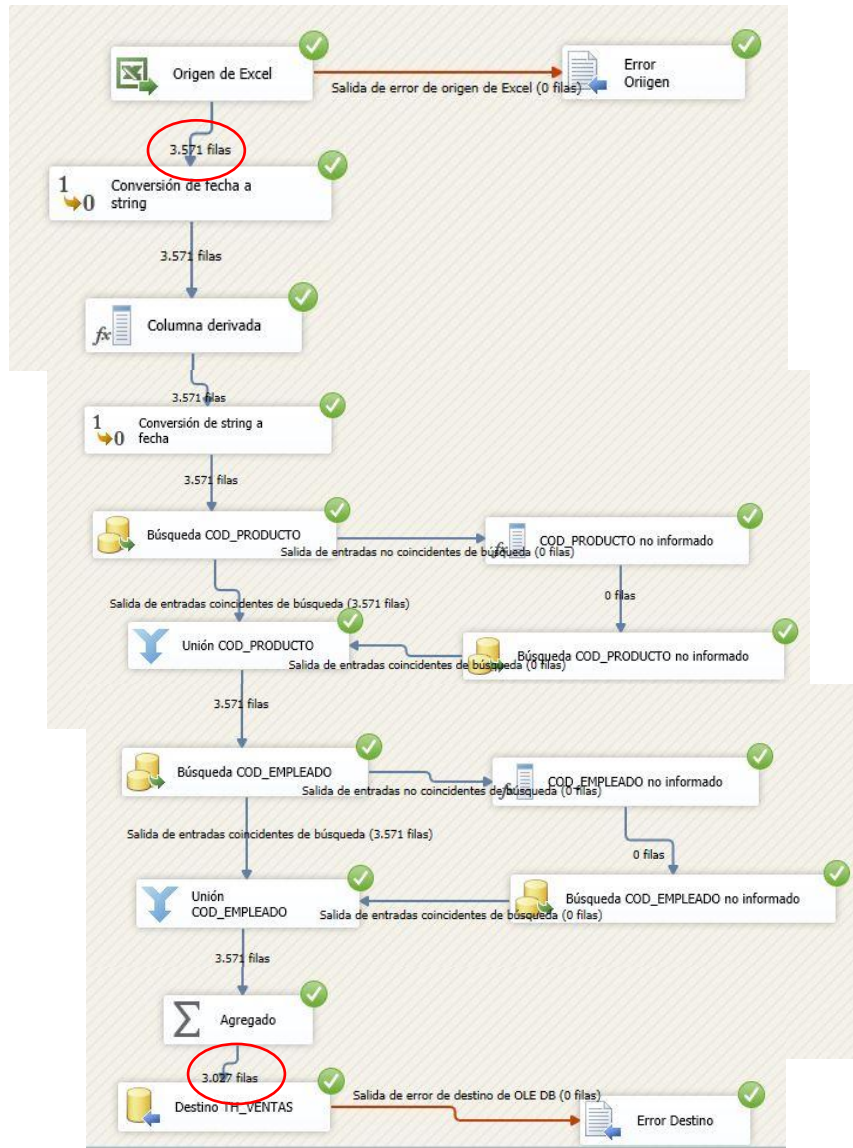
Corresponde perfectamente con todos los datos.

- **TH_VENTAS**

- **Validación Flujo de Datos**

Para finalizar, vamos a comprobar la carga correcta de la información de la tabla de hechos de ventas. Como ha sucedido en el resto de los casos, el flujo de carga parece correcto.

Ilustración 161. Flujo de Carga; TH_VENTAS



Desde el origen al destino se transfieren 3.571 filas. La agrupación de registros, al igual que en el caso anterior, ha sumado las ventas pertenecientes al mismo empleado en el mismo mes y del mismo producto aunque fueran en días diferentes, puesto que la agrupación se realiza por meses y finalmente, en el Data Warehouse, se insertan un total de 3.027 registros.

Verificamos en el origen de datos:

Ilustración 162. Excel; Ventas

	A	B	C	D	E	F
1	FECHA	CODIGO_PR ODUCTO	CODIGO_EM PLEADO	UNIDADE S	PRECIO_UNIT ARIO	PRECIO_VENTA
2	01/01/2014	28	1	3	27,72	83,16
3	01/01/2014	40	2	6	77,71	466,26
4	01/01/2014	13	1	8	24,86	198,88
5	01/01/2014	59	2	7	2,29	16,03
6	02/01/2014	42	2	9	43,99	395,91
7	02/01/2014	62	4	3	11,99	35,97
3567	30/03/2016	60	1	6	5,49	32,94
3568	30/03/2016	90	4	3	44,99	134,97
3569	31/03/2016	56	4	1	23,99	23,99
3570	31/03/2016	37	2	3	4,70	14,10
3571	31/03/2016	6	3	3	6,31	18,93
3572	31/03/2016	34	3	8	17,00	136,00

En el origen, hay un total de 3.572 registros diferentes, si descontamos la fila dedicada a la cabecera, hacen 3.571, que son exactamente los que se cargan en el origen, por lo tanto, a priori, la carga es correcta.

○ **Validación Carga de Datos**

Manifestamos el destino, es decir, la base de datos.

Ilustración 163. Data Warehouse; TH_VENTAS; Total

```

/***** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP 4000 [ID_MESANO]
,[ID_PRODUCTO]
,[ID_EMPLEADO]
,[UNIDADES]
,[PRECIO_VENTA]
FROM [DWH_VETERINARIA].[dbo].[TH_VENTAS]
    
```

	ID_MESANO	ID_PRODUCTO	ID_EMPLEADO	UNIDADES	PRECIO_VENTA
3014	201603	83	5	8	279.92
3015	201603	84	4	11	43.89
3016	201603	85	3	8	55.92
3017	201603	85	4	16	111.84
3018	201603	86	3	5	29.95
3019	201603	87	2	6	77.94
3020	201603	88	2	10	119.90
3021	201603	88	5	5	59.95
3022	201603	89	3	10	199.90
3023	201603	89	4	2	39.98
3024	201603	90	2	6	77.94
3025	201603	90	4	1	12.99
3026	201603	90	5	10	129.90
3027	201603	91	5	3	134.97

Tenemos cargados un total de 3.027 registros, que son los que finalizan la ETL con éxito.

Veamos si los datos coinciden con el origen:

Ilustración 164. Data Warehouse; TH_VENTAS; Precio

	ID_DIAMESANO	ID_PRODUCTO	ID_EMPLEADO	UNIDADES	PRECIO_VENTA
273	20150123	23	5	8	60.88
274	20150123	44	4	4	199.96
275	20150124	27	2	1	27.27
276	20150124	87	5	5	64.95

El 24 – 01 – 2015 se realizó una venta con importe de 27,27 de un producto, vendido por el empleado número 2 y el producto con id 27.

Ilustración 165. Excel; Ventas; Precio Venta

	A	B	C	D	E	F
1	FECHA	CODIGO_PR	CODIGO_EM	UNIDADE	PRECIO_UNIT	PRECIO_VENTA
218	24/01/2015	26	1	1	27,27	27,27

En el origen corresponde con el producto 26 y el empleado 1.

Ilustración 166. Excel; Producto; Cod_producto

	A	B	C	D	E	F
1	CODIGO_PRODUC	NOMBRE	DESCRIPCION	COD_TIPO	TIPO_PRODUCTO	PRECIO
27	26	HEPATOSIL desde 10 Kg 30 Comprimidos	Suplemento nutricional terapéutico especialmente formulado para ayudar a la función hepática en perros y gatos con problemas hepáticos. (Perros y Gatos)	15	Hepáticos	27,27

Ilustración 167. Excel; Diario; Empleado

	A	B	C	D	E	F
1	CODIGO_EMPLEADO	NOMBRE	APELLIDOS	TELÉFONO	COD_CARGO	CARGO
2	1	Leticia	Aceves Martín	999887711	1	Administrativo
3	2	Miguel Ángel	Garzón López	999333544	2	Doctor
4	3	Iván	Sastre García	999228877	1	Administrativo
5	4	Susana	Alonso Herrero	999115533	2	Doctor

Veamos el destino si se corresponde:

Ilustración 168. Data Warehouse; DIM_PRODUCTO; ID_PRODUCTO

	ID_PRODUCTO	COD_PRODUCTO	NOMBRE	DESCRIPCION	PRECIO	ID_TI
24	24	23	KENOFIX AEROSOL 250 ML	una película resistente que impide la entrada de s...	12.47	24
25	25	24	DENAMARIN BLISTER 90 MG 30 CO...	Recupera la funcionalidad del hígado, siendo esp...	28.30	26
26	26	25	HEPATOSIL Perro-Gato hasta 10 Kg	Suplemento nutricional terapéutico especialmente	21.19	26
27	27	26	HEPATOSIL desde 10 Kg 30 Comprimi...	Suplemento nutricional terapéutico especialmente ...	27.27	26
28	28	27	FELIWAY SPRAY 60 ml	Ayuda a parar o prevenir el marcaje urinario del gat...	22.07	16

Ilustración 169. Data Warehouse; DIM_EMPLEADO; ID_EMPLEADO

	ID_EMPLEADO	COD_EMPLEADO	NOMBRE	APELLIDOS	TELEFONO	ID_CARGO
1	1	-1	no informado	no informado	0	-1
2	2	1	Leticia	Aceves Martín	999887711	2
3	3	2	Miguel Ángel	Garzón López	999333544	3
4	4	3	Iván	Sastre García	999228877	2
5	5	4	Susana	Alonso Herrero	999115533	3

Corresponde todo a la perfección.

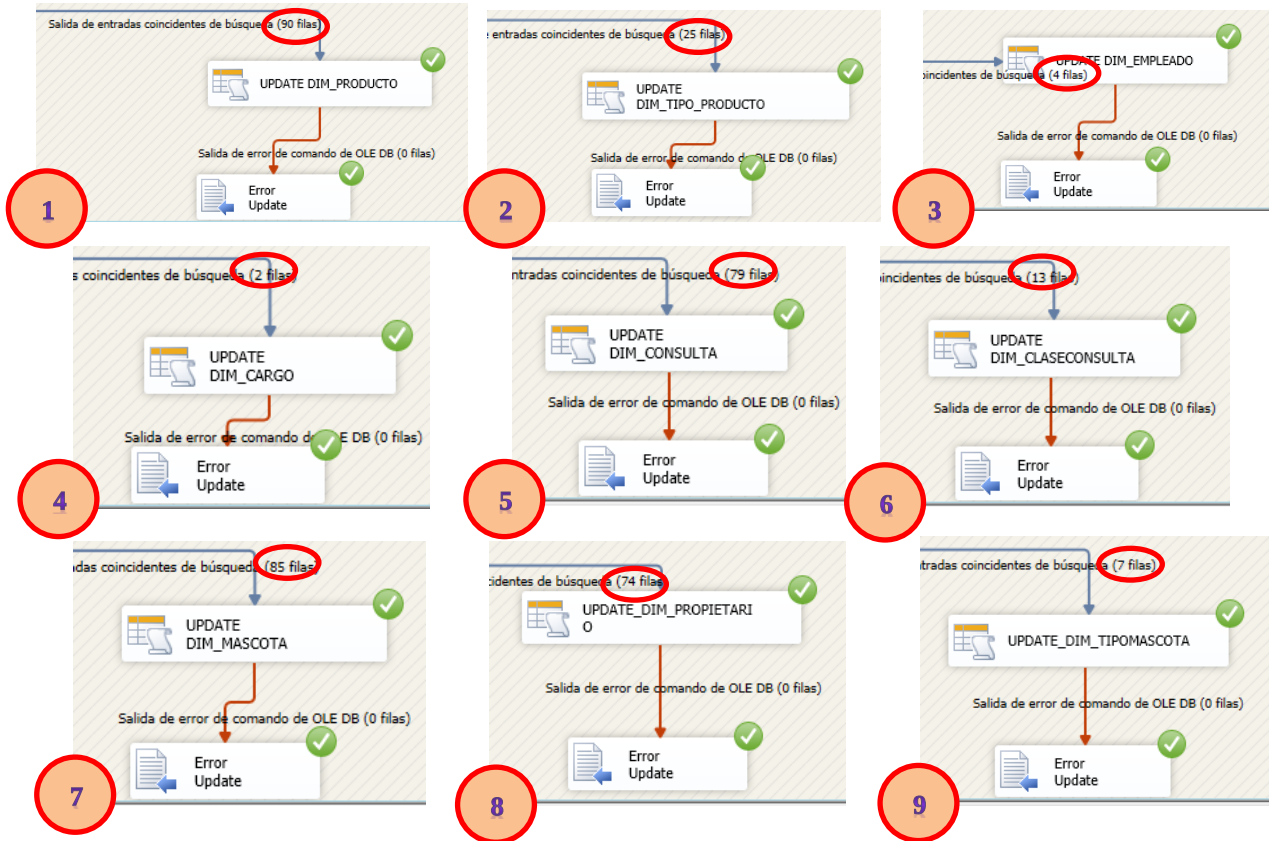
7.1.3- Prueba de Relanzamiento

Una de las características de las ETLs, es que son relanzables, esto es, que una vez hechas y cargados todos los datos en el Data Warehouse, las mismas ETLs se pueden volver a lanzar y si los datos son los mismos, no producir cambio alguno en el resultado de la información en la base de datos. En el caso de que los datos sean distintos, obviamente los actualizará en el caso de las dimensiones o los insertará en el caso de las tablas de hechos.

De este modo, volveremos a lanzar los flujos de carga para comprobar la salida de datos en todos y cada uno de los casos es la correcta y no se producen errores ni en el proceso de carga ni, por supuesto, en la base de datos.

Cuando volvemos a lanzar las ETLs, en principio el flujo de carga completo parece correcto. Visualizaremos todas las ETLs para comprobar que las salidas de los datos se realizan por los updates en el caso de las dimensiones:

Ilustración 170. Relanzamiento; Dimensiones



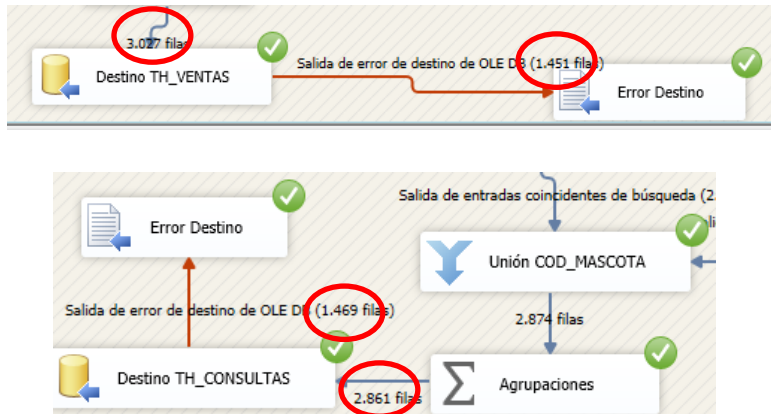
Como vemos en las imágenes, todas las ETLs se han vuelto a lanzar correctamente. Todas realizan el update de todos los registros:

- 4- DIM_PRODUCTO: un total de 90 filas cargadas y en el nuevo lanzamiento un total de 90 registros actualizados.
- 5- DIM_TIPO_PRODUCTO: 25 productos diferentes cargados en el primer lanzamiento y 25 registros actualizados.
- 6- DIM_EMPLEADO: 4 empleados contratados y 4 empleados actualizados
- 7- DIM_CARGO: 2 registros cargados en primer lugar y 2 en el relanzamiento
- 8- DIM_CONSULTA: 79 datos en la base de datos y 79 actualizados

- 9- DIM_CLASECONSULTA: 13 filas cargadas en un primer momento y 13 actualizadas
- 10- DIM_MASCOTA: 85 mascotas cargadas en la base de datos y 85 actualizadas
- 11- DIM_PROPIETARIO: 74 propietarios diferentes teníamos y 74 han sido actualizados.
- 12- DIM_TIPOMASCOTA: 7 tipos de mascotas actualizados y 7 tipos de mascotas había.

En el caso de las tablas de hechos:

Ilustración 171. Relanzamiento; Tablas de Hechos



En el caso de las tablas de hechos es un poco diferente. Parte de los registros se cargan con total normalidad, pero otra parte, la ETL la expulsa a la salida de errores. Esto es debido al delete previo a ejecución del flujo de carga, solamente se eliminan algunos de los registros, los que pertenecen a los últimos años, estos registros eliminados son los que se cargan correctamente en la tabla de la base de datos. Los otros registros, como no tenemos salida de update para poder sobre-escribirlos, el sistema les descarta al fichero de errores.

Ilustración 172. Relanzamiento; Tablas de Hechos; Salida de Errores

errorDestino.txt: Bloc de notas									
Archivo	Edición	Formato	Ver	Ayuda					
201411	65	5	19	1	71	24	-1071607683	0	
201410	7	3	17	1	456	-1071607683	0		
201410	7	3	58	1	456	-1071607683	0		
201409	71	3	65	0	65	-1071607683	0		
201410	7	3	81	1	456	-1071607683	0		
201409	71	3	81	1	78	-1071607683	0		
201409	71	5	15	0	65	-1071607683	0		
201409	71	5	59	0	65	-1071607683	0		
201408	13	3	80	1	234	-1071607683	0		

Como vemos, la salida de errores contiene los datos referentes al primer año 2014.

Finalmente, comprobaremos nuevamente de una en una todas las tablas de la base de datos para acreditar que todas tienen los datos correctos y el relanzamiento se ha realizado con éxito.

7.2- Validaciones a nivel de Informes

En este punto se definirán las pruebas realizadas para comprobar que el proceso de reporting de datos suministrador por el Data Warehouse se realizó correctamente y que los datos del informe son correctos y válidos.

Para ello evaluaremos todos los informes y comprobaremos mediante sentencias sql que los datos se corresponden perfectamente con los que se encuentran en el Data Warehouse.

7.2.1- Informe de Consultas

- Consultas

Para comenzar a validar los informes, empezaremos con el dedicado a todas las consultas realizadas en la clínica en los últimos 2 años.

El gráfico coge los datos de la tabla, por lo que necesitamos validar la tabla de datos. Como la información es de tipo numérico las comparaciones se hacen más sencillas.

Ilustración 173. Informes; Consultas; Tabla General

Fecha	Año	Clase	Consulta	Importe	%por meses			
ENERO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1.100 €	3,47%			
			Fractura diafisaria fijador	1.368 €	4,31%			
			Fractura diafisaria placa	800 €	2,52%			
			Fractura mandibular	516 €	1,63%			
			Ligamentos cruzados	1.952 €	6,15%			
			Osteocondritis disecante	1.220 €	3,84%			
			Proceso anconeo	1.308 €	4,12%			
			Proceso coronoideas	1.220 €	3,84%			
			Trocleooplastia/cresta tibial	1.188 €	3,74%			
			Higroma de codo (penrose)	660 €	2,08%			
			Total	11.332 €	35,71%			
			Cirugía General	2014	Cirugía General	Hernia diafragmática	1.100 €	3,47%
						Hernia umbilical	624 €	1,97%
Hernia perineal	2.133 €	6,72%						
Tumor/quiste cutáneo	440 €	1,39%						
Laparatomía exploratoria	638 €	2,01%						

Comparamos los datos directamente con los de la base de datos mediante secuencias SQL.

Ilustración 174. Data Warehouse; TH_CONSULTAS; Validación Tabla General

```

--VALIDACION GENERAL DE DATOS
DECLARE @idConsulta as INT = 24;
SELECT
    T.DESC_MES DESC_MES,
    T.ANO ANO,
    CC.CLASE_CONSULTA CLASE_CONSULTA,
    C.DESCRIPCION DESCRIPCION,
    sum((TH.IMPORTE_CONSULTA) IMPORTE_CONSULTA
FROM TH_CONSULTAS TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T ON (TH.ID_MESANO = T.ID_MESANO)
INNER join DIM_CLASECONSULTA CC ON (C.ID_CLASE = CC.ID_CLASE])
where C.ID_CONSULTA= @idConsulta
group by TH.ID_MESANO, T.DESC_MES, T.ANO, CC.CLASE_CONSULTA, C.DESCRIPCION
ORDER BY TH.ID_MESANO
    
```

DESCR_MES	ANO	CLASE_CONSULTA	DESCRIPCION	IMPORTE_CONSULTA
ENERO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1100,00
FEBRERO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1000,00
MAYO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	600,00
JUNIO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	500,00
JULIO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1100,00
AGOSTO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1100,00
SEPTIEMBRE	2014	Cirugía Ortopédica	Fractura diafisaria clavos	500,00

Consulta ejecutada correctamente. WIN2012SERVER (11.0 RTM) | VECTORSF\Administrador... | DWH_VETERINARIA | 00:00:00 | 20 filas

En enero del 2014 se realizaron cirugías ortopédicas, en concreto nos fijaremos en la fractura disfasia de clavos, por importe de 1.100€. Como vemos en ambas imágenes, los datos coinciden a la perfección, tanto en las cuantías como en fechas.

Del mismo modo se han comprobado otras cantidades dando un resultado exacto en ambos casos tanto en el informe como en la base de datos, por lo que en principio, los datos del informe son correctos.

Vamos a validar los porcentajes:

Ilustración 175. Informes; Consultas; Validación Porcentaje Meses

Cuadro General Consultas							
Fecha		Clase	Consulta	importe	%por meses		
ENERO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1.100 €	3,47%		
			Fractura diafisaria fijador	1.368 €	4,31%		
			Fractura diafisaria placa	800 €	2,52%		
			Fractura mandibular	516 €	1,63%		
			Ligamentos cruzados	1.952 €	6,15%		
			Osteocondritis disecante	1.220 €	3,84%		
			Proceso anconeo	1.308 €	4,12%		
			Proceso coronoides	1.220 €	3,84%		
			Trocleooplastia/cresta tibial	1.188 €	3,74%		
			Higroma de codo (penrose)	660 €	2,08%		
			Total			44.222 €	25,74%

Cogemos el caso de una consulta de fractura disfasaria placa, con un importe de 800€. Veamos mediante una sentencia sql si los datos se corresponden con los del Data Warehouse.

Ilustración 176. Data Warehouse; TH_CONSULTAS; Validación Porcentaje Meses

```

-- VALIDACION PORCENTAJE CONSULTAS POR MESES
DECLARE @MESANO INT =201401;
DECLARE @idConsulta2 as INT = 26;
SELECT
    T.DESC_MES DESC_MES,
    T.ANO ANO,
    CC.CLASE_CONSULTA CLASE_CONSULTA,
    C.DESCRIPCION DESCRIPCION,
    TH.TIPO_CONSULTA TIPO_CONSULTA,
    TH.IMPORTE_CONSULTA IMPORTE_CONSULTA,
    ((TH.IMPORTE_CONSULTA)/(SELECT SUM(IMPORTE_CONSULTA) FROM TH_CONSULTAS TH
        INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
        INNER join DIM_TIEMPO T ON (TH.ID_MESANO = T.ID_MESANO)
        INNER join DIM_CLASECONSULTA CC ON (C.ID_CLASE = CC.ID_CLASE)
        WHERE TH.ID_MESANO=@MESANO)*100) PORCENTAJE_MESES
FROM TH_CONSULTAS TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T ON (TH.ID_MESANO = T.ID_MESANO)
INNER join DIM_CLASECONSULTA CC ON (C.ID_CLASE = CC.ID_CLASE)
    
```

DESC_MES	ANO	CLASE_CONSULTA	DESCRIPCION	TIPO_CONSULTA	IMPORTE_CONSULTA	FORNENTAJE_MESES
ENERO	2014	Cirugía Ortopédica	Fractura diafisaria placa	0	800.00	2.52066710715329641695993053

Consulta ejecutada correctamente. WIN2012SERVER (11.0 RTM) | VECTORSF\Administrador... | DWH_VETERINARIA | 00:00:00 | 1 filas

Los datos coinciden, por lo que el informe, en estos datos, también es correcto.

- **Urgencias**

Este informe está dedicado el número de urgencias en un periodo de tiempo que se realizan en la clínica veterinaria. Confirmaremos la realidad de estos datos con los que tenemos de base.

Como en el caso anterior, realizaremos las comparaciones de datos con la tabla, que es la que recoge datos numéricos:

Ilustración 177. Informes; Urgencias; Tabla

Cuadro General Urgencias				
Fecha	Clase	Consulta	Urgencias	importeConsulta
ENERO	Cirugía Ortopédica	Fractura diafisaria clavos	1	1.100 €
		Fractura diafisaria fijador	2	1.368 €
		Fractura diafisaria placa	0	800 €
		Fractura mandibular	1	516 €
		Ligamentos cruzados	1	1.952 €
		Osteocondritis disecante	0	1.220 €
		Proceso anconeo	2	1.308 €

La tabla es muy parecida a la del informe anterior y por tanto, la validación es similar. Por ejemplo, validaremos la fractura diafisaria de clavos.

Ilustración 178. Data Warehouse; TH_CONSULTAS; Validación Urgencias

```

-- VALIDA URGENCIAS TOTALES
DECLARE @idConsultaU as INT = 24;
SELECT
    T.DESC_MES DESC_MES,
    T.ANO ANO,
    CC.CLASE_CONSULTA CLASE_CONSULTA,
    C.DESCRIPCION DESCRIPCION,
    sum(TH.TIPO_CONSULTA) TIPO_CONSULTA,
    sum(TH.IMPORTE_CONSULTA) IMPORTE_CONSULTA
FROM TH_CONSULTAS TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T ON (TH.ID_MESANO = T.ID_MESANO)
INNER join DIM_CLASECONSULTA CC ON (C.ID_CLASE = CC.ID_CLASE)
where c.ID_CONSULTA= @idConsultaU
    
```

DESC_MES	ANO	CLASE_CONSULTA	DESCRIPCION	TIPO_CONSULTA	IMPORTE_CONSULTA
ENERO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	0	500.00
ENERO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1	600.00
FEBRERO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	3	1800.00
MAYO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1	600.00
JUNIO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	0	500.00
JULIO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	0	500.00
JULIO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1	600.00
AGOSTO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	0	500.00
AGOSTO	2014	Cirugía Ortopédica	Fractura diafisaria clavos	1	600.00

Consulta ejecutada correctamente. | WIN2012SERVER (11.0 RTM) | VECTORSP\Administrador... | DWH_VETERINARIA | 00:00:00 | 30 filas

Como vemos en la sentencia sql, tenemos 2 consultas de fractura diafisaria de clavos en enero de 2014, una de tipo normal y otra de tipo urgente. En el informe, tan solo se cuentan las urgencias, las consultas normales no suman en tipo pero sí se suman en cantidad o importe, por lo que en el informe aparece una cantidad de 1.100€ que es exactamente lo que da como resultado de la suma de 500€ de la consulta normal con los 600€ de la consulta tipo urgente. Por lo tanto, confirmado, los datos están validados y correctos.

7.2.2- Informe Variación

- Variación Mensual
 - Consultas

Se han realizado informes de variaciones, comenzamos con la variación mensual de las consultas realizadas por la clínica veterinaria en un periodo de tiempo.

Como en los casos anteriores, nos fijaremos en la tabla de datos numéricos:

Ilustración 179. Informes; Variación Consultas; Tabla Mensual

Cuadro General Variación Mensual Consultas				
Fecha		Importe	Mes Anterior	variación
NOVIEMBRE	2014	24.618,51 €	30.343,03 €	-18,87%
NOVIEMBRE	2014	27.982,88 €	24.618,51 €	13,87%
ENERO	2015	29.688,50 €	27.982,88 €	6,10%
FEBRERO	2015	22.534,45 €	29.688,50 €	-24,10%
MARZO	2015	22.422,73 €	22.534,45 €	-0,50%
ABRIL	2015	30.415,78 €	22.422,73 €	35,65%
MAYO	2015	27.600,05 €	30.415,78 €	-9,26%
JUNIO	2015	21.507,76 €	27.600,05 €	-22,07%
JULIO	2015	28.877,51 €	21.507,76 €	34,27%
AGOSTO	2015	24.906,27 €	28.877,51 €	-13,75%
SEPTIEMBRE	2015	20.662,97 €	24.906,27 €	-17,04%

Para realizar este tipo de validaciones, filtraremos la consulta a la base de datos mediante una variable.

Confirmaremos la variación mensual para el mes de enero de 2015, por lo que la variable toma el valor 201501. Para realizar la confirmación de otro periodo temporal, es suficiente con cambiar el valor de la variable.

Ilustración 180. Data Warehouse; TH_CONSULTAS; Validación Variación Mensual Consultas

```
--VALIDA VARIACION MENSUAL CONSULTAS
DECLARE @MESANO INT =201501;
SELECT TH.[ID_MESANO]
,SUM(TH.[IMPORTE_CONSULTA]) SUMA
,(SELECT SUM(TH.[IMPORTE_CONSULTA])
FROM [TH_CONSULTAS] TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MESANT_AÑO)
INNER join DIM_CLASECONSULTA CC ON (C.ID_CLASE = CC.ID_CLASE)
WHERE @MESANO=T2.ID_MESANO
) IMPOR_ANT
,((SUM(TH.[IMPORTE_CONSULTA])-(SELECT SUM(TH.[IMPORTE_CONSULTA])
FROM [TH_CONSULTAS] TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MESANT_AÑO)
INNER join DIM_CLASECONSULTA CC ON (C.ID_CLASE = CC.ID_CLASE)
WHERE @MESANO=T2.ID_MESANO))
/(SELECT SUM(TH.[IMPORTE_CONSULTA])
FROM [TH_CONSULTAS] TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MESANT_AÑO))
) VARIACION
```

ID_MESANO	SUMA	IMPOR_ANT	VARIACION
1	29688.50	27982.88	6.09200

Como vemos los datos se corresponden con los de la tabla, el importe es 29.688,50€, el importe del mes anterior es 27.982,88€ y la variación es de 6,09%.

- Ventas

Realizaremos el mismo proceso con la tabla de las variaciones mensuales del caso de las ventas de productos.

Ilustración 181. Informes; Variación Ventas; Tabla Mensual

Cuadro General Variación Mensual Ventas				
Fecha		Importe	Mes Anterior	Variación
SEPTIEMBRE	2014	16.483,27 €	19.544,98 €	-15,66%
OCTUBRE	2014	19.074,75 €	16.483,27 €	15,72%
NOVIEMBRE	2014	13.951,05 €	19.074,75 €	-26,86%
DICIEMBRE	2014	19.259,29 €	13.951,05 €	36,06%
ENERO	2015	15.623,55 €	19.259,29 €	-18,88%
FEBRERO	2015	13.921,83 €	15.623,55 €	-10,95%
MARZO	2015	16.432,78 €	13.921,83 €	18,04%
ABRIL	2015	16.579,06 €	16.432,78 €	0,89%
MAYO	2015	16.527,06 €	16.579,06 €	-0,31%
JUNIO	2015	14.559,93 €	16.527,06 €	-11,90%
JULIO	2015	19.190,88 €	14.559,93 €	31,81%

Ilustración 182. Data Warehouse; TH_VENTAS; Validación Variación Mensual Ventas

```
--VALIDA VARIACION MENSUAL VENTAS
DECLARE @MESANOV INT -201501;
SELECT TH.[ID_MESANO]
, SUM([TH].[PRECIO_VENTA]) SUMA
, (SELECT SUM([TH].[PRECIO_VENTA])
FROM [TH_VENTAS] TH
INNER join DIM_PRODUCTO P ON (TH.ID_PRODUCTO = P.ID_PRODUCTO)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MESANT_ANO)
INNER join DIM_TIPO_PRODUCTO TP ON (P.ID_TIPOPRODUCTO = TP.ID_TIPOPRODUCTO)
WHERE @MESANOV=T2.ID_MESANO) IMPOR_ANT
, ((SUM([TH].[PRECIO_VENTA]) - (SELECT SUM([TH].[PRECIO_VENTA])
FROM [TH_VENTAS] TH
INNER join DIM_PRODUCTO P ON (TH.ID_PRODUCTO = P.ID_PRODUCTO)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MESANT_ANO)
INNER join DIM_TIPO_PRODUCTO TP ON (P.ID_TIPOPRODUCTO = TP.ID_TIPOPRODUCTO)
WHERE @MESANOV=T2.ID_MESANO))
/ (SELECT SUM([TH].[PRECIO_VENTA])
FROM [TH_VENTAS] TH
INNER join DIM_PRODUCTO P ON (TH.ID_PRODUCTO = P.ID_PRODUCTO)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MESANT_ANO)
INNER join DIM_TIPO_PRODUCTO TP ON (P.ID_TIPOPRODUCTO = TP.ID_TIPOPRODUCTO))
) VARIACION
FROM [TH_VENTAS] TH
```

ID_MESANO	SUMA	IMPOR_ANT	VARIACION
201501	15623.55	19259.29	-18.877800

El importe del mes que nos interesa es de 15.623,55€ en ambos casos y el importe total de beneficios del mes anterior es 19.259,29€ con una variación de -18,87%, es una variación negativa porque el mes anterior se obtuvieron más beneficios. Todos los datos coinciden, la validación ha sido correcta.

- **Variación Anual e Interanual**
 - **Consultas**

La variación anual e interanual la podemos validar ambas juntas ya que las consultas sql es igual en ambos casos: los datos refieren al importe de un mes con el mismo mes del año anterior. En el caso del anual son todos los meses del año, en el caso del interanual tan solo contamos con los tres primeros meses del año.

Ilustración 183. Informes; Variación; Anual Consultas

Cuadro General Variación Anual Consultas				
Fecha	Año	Importe	Año Anterior	Variación
JUNIO	2015	21.507,76 €	25.920,16 €	-17%
JULIO	2015	28.877,51 €	24.930,95 €	16%
AGOSTO	2015	24.906,27 €	24.550,57 €	1%
SEPTIEMBRE	2015	20.662,97 €	28.276,57 €	-27%
OCTUBRE	2015	25.226,93 €	30.343,03 €	-17%
NOVIEMBRE	2015	22.152,14 €	24.618,51 €	-10%
DICIEMBRE	2015	22.966,77 €	27.982,88 €	-18%
ENERO	2016	23.938,58 €	29.688,50 €	-19%
FEBRERO	2016	27.373,97 €	22.534,45 €	21,48%
MARZO	2016	24.213,68 €	22.422,73 €	8%
Total		374.488,09 €	398.207,05 €	-6%

Como se ve en el cuadro del informe, en marzo del 2016 se alcanzó un importe de 24.213,68€ y en el mismo mes del año anterior fueron 22.422,73€, esto quiere decir que en el año 2016 se ha tenido una ganancia de un 8%.

Ahora vamos a ver la tabla del informe interanual:

Ilustración 184. Informes; Variación; Interanual Consultas

Cuadro General Variación Interanual Consultas				
Fecha	Año	Importe 2016	Año Anterior	variacion
ENERO	2016	23.939	29.689	-19,37%
FEBRERO	2016	27.374	22.534	21,48%
MARZO	2016	24.214	22.423	7,99%

Para la validación, como en el caso anterior, se ha utilizado una variable, donde podemos incluir el mes y año que queremos validar, en este caso, el valor que tomará la variable es de 201603.

Ilustración 185. Data Warehouse; TH_CONSULTAS; Validación Variación Anual e Interanual

```
-- VALIDACION VARIACION ANUAL E INTERANUAL DE CONSULTAS
DECLARE @MESANO2 INT =201603;
SELECT TOP 1000 TH.[ID_MESANO]
,SUM((TH.[IMPORTE_CONSULTA])) SUMA
,(SELECT SUM((TH.[IMPORTE_CONSULTA]))
FROM [TH_CONSULTAS] TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MES_ANOMANT)
INNER join DIM_CLASECONSULTA CC ON (C.ID_CLASE = CC.ID_CLASE)
WHERE @MESANO2=T2.ID_MESANO
) IMPOR_ANT
,((SUM((TH.[IMPORTE_CONSULTA]))-(SELECT SUM((TH.[IMPORTE_CONSULTA]))
FROM [TH_CONSULTAS] TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MES_ANOMANT)
INNER join DIM_CLASECONSULTA CC ON (C.ID_CLASE = CC.ID_CLASE)
WHERE @MESANO2=T2.ID_MESANO))
/(SELECT SUM((TH.[IMPORTE_CONSULTA]))
FROM [TH_CONSULTAS] TH
INNER join DIM_CONSULTA C ON (TH.ID_CONSULTA = C.ID_CONSULTA)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MES_ANOMANT)

```

ID_MESANO	SUMA	IMPOR_ANT	VARIACION
201603	24213.68	22422.73	7.987200

Como vemos, la validación de la variación anual e interanual en el caso de las consultas se ha realizado con éxito.

○ Ventas

Ahora vamos a realizar las mismas comprobaciones con las ventas en el caso de las variaciones anuales e interanuales.

Ilustración 186. Informes; Variación; Anual Ventas

Cuadro General Variación Anual Ventas				
Fecha		Importe	Año Anterior	variación
MAYO	2015	16.527,06 €	15.148,51 €	9,10%
JUNIO	2015	14.559,93 €	14.132,07 €	3,03%
JULIO	2015	19.190,88 €	15.803,37 €	21,44%
AGOSTO	2015	19.494,74 €	19.544,98 €	-0,26%
SEPTIEMBRE	2015	12.420,32 €	16.483,27 €	-24,65%
OCTUBRE	2015	15.887,86 €	19.074,75 €	-16,71%
NOVIEMBRE	2015	20.418,63 €	13.951,05 €	46,36%
DICIEMBRE	2015	14.872,95 €	19.259,29 €	-22,78%
ENERO	2016	13.766,85 €	15.623,55 €	-11,88%
FEBRERO	2016	15.956,96 €	13.921,83 €	14,38%
MARZO	2016	12.426,94 €	16.432,78 €	-24,38%

El importe de marzo del 2016 es de 12.426,94€ y del mismo mes del año anterior 16.432,78€ con una variación negativa del 24,38%.

Ilustración 187. Informes; Variación; Interanual Ventas

Cuadro General Variación Interanual Ventas				
Fecha		Importe 2016	Año Anterior	variación
ENERO	2016	13.766,85 €	15.623,55 €	-11,88%
FEBRERO	2016	15.956,96 €	13.921,83 €	14,38%
MARZO	2016	12.426,94 €	16.432,78 €	-24,38%

Vamos a comprobar los importes en la base de datos.

Ilustración 188. Data Warehouse; TH_VENTAS; Validación Variación Anual e Interanual

```

-- VALIDACION VARIACION ANUAL E INTERANUAL DE VENTAS
DECLARE @MESANOV2 INT =201603;
SELECT TOP 1000 TH.[ID_MESANO]
,SUM(TH.[PRECIO_VENTA]) SUMA
,(SELECT SUM(TH.[PRECIO_VENTA])
FROM [TH_VENTAS] TH
INNER join DIM_PRODUCTO P ON (TH.ID_PRODUCTO = P.ID_PRODUCTO)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MES_AÑOANT)
INNER join DIM_TIPO_PRODUCTO TP ON (P.ID_TIPOPRODUCTO = TP.ID_TIPOPRODUCTO)
WHERE @MESANOV2=T2.ID_MESANO
) IMPOR_ANT
,((SUM(TH.[PRECIO_VENTA])-(SELECT SUM(TH.[PRECIO_VENTA])
FROM [TH_VENTAS] TH
INNER join DIM_PRODUCTO P ON (TH.ID_PRODUCTO = P.ID_PRODUCTO)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MES_AÑOANT)
INNER join DIM_TIPO_PRODUCTO TP ON (P.ID_TIPOPRODUCTO = TP.ID_TIPOPRODUCTO)
WHERE @MESANOV2=T2.ID_MESANO))
/(SELECT SUM(TH.[PRECIO_VENTA])
FROM [TH_VENTAS] TH
INNER join DIM_PRODUCTO P ON (TH.ID_PRODUCTO = P.ID_PRODUCTO)
INNER join DIM_TIEMPO T2 ON (TH.ID_MESANO = T2.MES_AÑOANT)

```

ID_MESANO	SUMA	IMPOR_ANT	VARIACION
201603	12426.94	16432.78	-24.377100

Como vemos en el resultado de la sentencia sql, la suma se corresponde con el importe del mes de marzo del 2016. El importe del mismo mes del año anterior también coincide y la variación es la misma en todos los casos.

7.2.3- Informe Ventas

- Ventas

El último informe está dedicado a las ventas de productos realizadas a diario. En el informe vemos el producto que se ha vendido, cuántas unidades se han vendido y qué trabajador ha realizado la venta en un mes determinado.

Ilustración 189. Informes; Ventas; Tabla General

Cuadro General Ventas						
Fecha	Empleado	Tipoproducto	Producto	unidades	Importe	
ENERO	2014	Leticia	Artículos	Bebedero antigoteo de Living World	8	39,92 €
				Bebedero Drinkwell 360	7	349,93 €
				Pala para arenas de gránulos gruesos	8	7,92 €
				Spray quitamanchas	14	97,86 €

En enero del 2014, Leticia vendió 8 unidades de Bebederos con un importe total de 39,92€. Como podemos ver, concuerda a la perfección con los datos de la base de datos.

Ilustración 190. Data Warehouse; TH_VENTAS; Validación Ventas

```
-- VALIDACION DEL INFORME DE VENTAS
SELECT TH.ID_MESANO
      ,E.NOMBRE
      ,TP.TIPO_PRODUCTO
      ,P.NOMBRE
      ,TH.UNIDADES
      ,SUM(TH.PRECIO_VENTA) PRECIO
FROM [TH_VENTAS] TH
INNER join DIM_PRODUCTO P ON (TH.ID_PRODUCTO = P.ID_PRODUCTO)
INNER join DIM_TIEMPO T ON (TH.ID_MESANO = T.ID_MESANO)
INNER join DIM_TIPO_PRODUCTO TP ON (TP.ID_TIPOPRODUCTO = P.ID_TIPOPRODUCTO)
INNER join DIM_EMPLEADO E ON (TH.ID_EMPLEADO = E.ID_EMPLEADO)
WHERE E.ID_EMPLEADO=2 AND P.ID_PRODUCTO=75
GROUP BY TH.ID_MESANO
```

ID_MESANO	NOMBRE	TIPO_PRODUCTO	NOMBRE	UNIDADES	PRECIO
201401	Leticia	Artículos	Bebedero antigoteo de Living World	8	39.92
201402	Leticia	Artículos	Bebedero antigoteo de Living World	1	4.99
201408	Leticia	Artículos	Bebedero antigoteo de Living World	2	9.98
201410	Leticia	Artículos	Bebedero antigoteo de Living World	4	19.96
201411	Leticia	Artículos	Bebedero antigoteo de Living World	8	39.92
201412	Leticia	Artículos	Bebedero antigoteo de Living World	8	39.92
201504	Leticia	Artículos	Bebedero antigoteo de Living World	8	39.92
201508	Leticia	Artículos	Bebedero antigoteo de Living World	4	19.96
201510	Leticia	Artículos	Bebedero antigoteo de Living World	1	4.99

Consulta ejecutada correctamente. WIN2012SERVER (11.0 RTM) | VECTORSF\Administrador... | DWH_VETERINARIA | 00:00:00 | 11 filas

La validación de este informe es correcta, como vemos en el resultado de la base de datos, en enero del 2014 Leticia, efectivamente, vendió 8 unidades de bebedero con un importe total de 39,92€.

CAPITULO VIII: CONCLUSIONES

CAPITULO 8: CONCLUSIONES

Ya finalizando con nuestro proyecto, no nos podemos olvidar de las conclusiones. En este capítulo vamos a valorar nuestra experiencia en Business Intelligence, tanto para bien, como para mal.

8.1- Conclusiones Personales

De forma genérica, nos ha gustado mucho tener la oportunidad de realizar este proyecto. No solo por el proyecto en sí, sino por el proceso de aprendizaje. No conocíamos nada del mundo de Business Intelligence, ahora que lo estamos comenzando a conocer, podemos decir que nos encantaría seguir trabajando en ello.

La realización de este proyecto nos ha dado la oportunidad de conocer el verdadero manejo de bases de datos que tienen las empresas, introduciéndonos en el mundo empresarial de forma real. En el momento de la realización del proyecto para la empresa Vector ITC Group, nos facilitaron modelos de documentación reales propios de proyectos BI para clientes. En esta memoria hemos tenido que transformar toda esa documentación y modificarla casi por completo.

El mayor inconveniente que se nos ha planteado a la hora de la realización del proyecto es el trabajo con máquinas virtuales y licencias de programas como Microstrategy. Ambas cosas nos las facilitaron en la empresa Vector ITC Group, pero, evidentemente las licencias no son permanentes. En varias ocasiones hemos tenido que volver sobre nuestros pasos y realizar nuevas importaciones de máquinas virtuales para poder seguir trabajando con las mismas licencias.

Las máquinas virtuales consumen mucha energía de nuestros ordenadores y tardan mucho tiempo en encenderse, esto ha ralentizado nuestro trabajo notablemente.

Además, la información sobre BI que podemos encontrar en libros o en la web, todavía es pequeña, siendo casi imposible encontrar manuales de cada uno de los programas o herramientas utilizados, y, los manuales encontrados, en muchas ocasiones son incompletos.

CAPITULO IX: BIBLIOGRAFÍA

CAPITULO 9: BIBLIOGRAFIA

CAPITULO I

<http://www.captio.net/blog/%C2%BFen-que-consiste-un-proyecto-de-business-intelligence-bi>
<https://basesdatoscms.files.wordpress.com/2012/10/resumen-businessintelligence.pdf>

CAPITULO II

<http://www.gestiopolis.com/procesos-inteligencia-negocios/>
<http://www.contunegocio.es/gestion/business-intelligence-estrategias-o-herramientas/>
<http://www.davinci-ti.es/introduccion-al-business-intelligence-pasos-previos/>
http://185.5.126.23/innowiki/index.php/Business_Intelligence
https://es.wikipedia.org/wiki/Esquema_en_estrella
<http://mundodb.es/diseño-data-warehouse-hechos-y-dimensiones>
<http://www.businessintelligence.info/serie-dwh/jerarquias-business-intelligence.html>
<http://r-pita.blogspot.com.es/2011/09/modelo-estrella-o-copo-de-nieve.html>
<http://josepcurto.com/2007/11/19/diseño-de-un-data-warehouse-estrella-y-copo-de-nieve/>

CAPITULO III

https://es.wikipedia.org/wiki/Proceso_para_el_desarrollo_de_software (Metodología)
https://es.wikipedia.org/wiki/Microsoft_Office (Microsoft Office)
<https://es.wikipedia.org/wiki/VirtualBox> (Virtual Box)
https://es.wikipedia.org/wiki/Microsoft_Visual_Studio (Visual Studio)
https://es.wikipedia.org/wiki/Microsoft_SQL_Server (SQL Server)
<https://es.wikipedia.org/wiki/MicroStrategy> (Microstrategy)
<https://es.scribd.com/doc/57058768> (Planificación Temporal)
<http://blog.nubelo.com/como-redactar-presupuesto-para-desarrollo-software/> (Presupuesto)
<https://es.wikipedia.org/wiki/COCOMO> (COCOMO)

CAPITULO IV

<https://eduardorene.wordpress.com/2012/08/17/> (Requisitos de Usuario)
https://es.wikipedia.org/wiki/Caso_de_uso (Casos de Uso)

CAPITULO V

<http://indalog.ual.es/mtorres/LP/FundamentosDiseno.pdf> (Diseño)
http://www.ecured.cu/Arquitectura_de_software (Arquitectura Software)
<https://sg.com.mx/revista/27/arquitectura-software#.V1BmsvmLTIU> (Arquitectura Software)
https://es.wikipedia.org/wiki/Diagrama_de_secuencia (Diagrama de Secuencia)
<https://wikiuml.wikispaces.com/Diagrama+de+Secuencia> (Diagrama de Secuencia)
<https://wikiuml.wikispaces.com/Diagrama+de+Estados> (Diagrama de Estados)

CAPITULO VI

<https://es.wikipedia.org/wiki/Implementaci%C3%B3n> (Implementación)
<http://es.calameo.com/books/002299301667571c7ab05> (Granularidad)
https://es.wikipedia.org/wiki/Tabla_de_hechos (Granularidad)
<http://www.dataprix.com/datawarehouse-manager> (Dimensiones y Tablas de hechos)
<http://www.informationbuilders.es/business-intelligence> (Informes)
<https://powerbi.microsoft.com/es-es/documentation/powerbi-service-reports/> (Informes)

CAPITULO VI

https://es.wikipedia.org/wiki/Pruebas_de_validaci%C3%B3n (Validación)

LIBROS

-“Introducción a Business Intelligence” por Josep Curto Díaz, Editorial OUC, S.L., 2010. ISBN 8497888863, 9788497888868
 -“SQL Server 2014. Implementación de una solución de Business Intelligence” por Thomas Gauchet. ISBN: 9782746095236
 -“Business Intelligence. Técnicas, Herramientas y Aplicaciones” por María Pérez Marqués. ISBN: 9788494305528

ANEXOS

ANEXO I: CREACIÓN DEL MODELO OPERACIONAL

Ya se ha definido con anterioridad en este documento lo que es un modelo operacional o fuente de información: Todas las empresas, ya sean grandes empresas o PYMES, tienen un sistema de almacenamiento de información respecto de sus beneficios/ pérdidas y gastos. A este almacenamiento de información es lo que denominamos fuentes de información o modelo operacional.

Las fuentes de información es, por tanto, la base de todo el sistema BI. Un modelo operacional bien estructurado, facilita notablemente el trabajo a la hora de realizar el análisis de la información.

El modelo operacional puede estar en distintas formas de almacenamiento. La forma más habitual es encontrar estos datos en un documento tipo documento hoja de cálculo, normalmente en el caso de las pequeñas empresas o en una base de datos, en el caso de las grandes empresas.

En este punto se expondrá cómo se ha creado nuestro modelo operacional, definiendo la creación de cada dato. Si bien, el operacional habitualmente es suministrado por una empresa real, por lo que en circunstancias normales no se tiene que crear. En el caso de este proyecto, se ha creado unos datos ficticios de forma aleatoria.

Teniendo todo esto en cuenta, y queriendo realizar un modelo operacional con datos ficticios de la pequeña empresa inexistente Vicherinario, optamos por el almacenamiento en documento hoja de cálculo. Aunque estos datos se han creado por y para el proyecto, se ha intentado asemejar lo máximo posible a una empresa real.

Estructura Organizacional del Documento

En primer lugar se definirá cómo se ha estructurado el documento hoja de cálculo para guardar los datos oportunos y por qué se ha hecho así. Se define la información que contendrá cada hoja de cálculo y cómo trabaja la clínica veterinaria en su día a día, de qué manera vamos a almacenar la información en el documento hoja de cálculo para que se asemeje a una clínica veterinaria real.

- **Organización**

En cuanto a la organización del documento tipo hoja de cálculo (Excel) que se va a crear, tenemos que diferenciar, primeramente, que la clínica veterinaria va a efectuar dos funciones:

- La realización de consultas a mascotas
- La venta de productos de mascotas

Estas dos actividades se tratan de forma distinta en las dos tablas de hechos. Teniendo así, una tabla de hechos para cada actividad: TH_CONSULTAS y TH_VENTAS.

Asumiendo que es la misma empresa (Vicherinario) quien realiza las dos actividades, no es de extrañar que las dos tablas de hechos tengan datos comunes. Por ejemplo, el código de empleado que realiza la venta o que hace la consulta. Por otro lado, es obvio que también existen datos que solo conciernen a una tabla de hechos en cuestión. Esta situación la vemos

reflejada en el diagrama de diseño lógico de la página 28 de este documento.

Para estudiar las consultas, como si fuera una clínica veterinaria real, vamos a tomar en cuenta datos referentes a: mascotas, propietarios, tipos de mascotas, empleado y cargo que posee dentro de la empresa, consulta y tipo de consulta.

Por otro lado, para las ventas, tendremos en cuenta: empleado, cargo del empleado, el tipo de producto y el producto.

- **Estructura de cada Hoja**

Una vez aclarados los datos que necesitaremos, pensamos en una clínica veterinaria y en cómo tendría estructurada esta información dentro del documento hoja de cálculo.

Necesariamente, los datos deben estar diferenciados en distintas hojas. Se debe dedicar una hoja de cálculo para registrar la información diaria concerniente a cada actividad: una para las consultas y otra para las ventas de productos.

Toda clínica veterinaria, debe tener también, un registro de las mascotas que atiende. Todos estos animales tienen una dirección y unos datos del propietario que es necesario conocer y almacenar.

Cualquier empresa lleva un registro de empleados. Se tendrá una hoja de cálculo dedicada a los empleados y el puesto de trabajo que realiza en la empresa.

A diario, se realizan distintos tipos de consultas, cada una con un precio estipulado. Este precio y las consultas hay que registrarlas, para mantener una lista de precios. Igual pasaría con los productos, es importante llevar un registro de los productos que están en venta y de su precio.

Con todo esto, nos salen un total de 6 hojas de cálculo, plenamente diferenciadas:

- Diario (Consultas realizadas al día)
- Ventas
- Empleados
- Mascotas
- Productos
- Consultas

Creación de la Información

Una vez entendida la información que queremos reflejar en la documentación y como organizaremos estos datos en el documento, vamos a evaluar hoja por hoja los datos que se deben contemplar en cada lugar. Los datos de cada hoja de cálculo no serán más que los necesarios para realizar nuestro proyecto de BI.

• Hoja de cálculo “Mascotas”

En esta hoja de cálculo, la clínica veterinaria va a tener toda la información concerniente a las mascotas que atiende.

La información que se ha creído imprescindible para una veterinaria respecto de sus mascotas es: Nombre, Raza, Sexo, Peso, la fecha de nacimiento del animal, en el caso de defunción también se registra, así como el tipo de mascota del que se trata.

Otros datos a tener en cuenta que se consideran importantes concernientes a las mascotas son: Nombre del Propietario, Dirección, Teléfono de contacto. Para poder mantener un contacto con los propietarios de los animales.

Ilustración 191. Fuente de Información; Mascotas

	A	B	C	D	E	F	G
	CODIGO_M					CODIGO_PROPI	
1	ASCOTA	NOMBRE	RAZA	SEX	PESO	ETARIO	PROPIETARIO
2	101	Thor	Labrador	M	28,50	1	Aceves Olmos, Félix
3	115	Snoopy	Caniche	M	9,47	15	Alonso Serna, Victoria
4	308	Reina	Canario	H	0,15	15	Alonso Serna, Victoria
5	215	Bola	Stames	H	7,10	47	Álvarez Antona, Samuel
6	113	Marea	Galgo	H	24,70	13	Antón Escribano, José María
7	405	Blanquita	Holandés	H	1,55	66	Aragón Peña, Lucía
8	111	Nora	Mestizo	H	12,00	11	Arranz Gómez, Raquel
9	207	Ron	Mestizo	M	2,60	11	Arranz Gómez, Raquel
10	136	Lucero	Galgo	M	26,00	36	Arribas Galindo, Natalia
11	132	Betty	Yorkshire Terrier	H	4,20	32	Barbado Vega, Rubén
12	206	Terry	Mestizo	M	4,10	41	Bartolomé Cañas, Iván

	G	H	I	J	K	L	M
	PROPIETARIO	Dirección	Teléfono	NTO	FECHA_NACIMIE	FECHA_DEFUN	COD_TIP
	Aceves Olmos, Félix	C/ Salida ronda, 9, 1º, 40008, Segovia	998556622		12/05/10		1 perro
	Alonso Serna, Victoria	C/ Santa Isabel, 23, 6A, 40101, Segovia	911547884		27/01/08		1 perro
	Alonso Serna, Victoria	C/ Santa Isabel, 23, 6A, 40101, Segovia	911547884		27/04/13		3 ave
	Álvarez Antona, Samuel	C/ San Vicente, 11, 1B, 40102, Segovia	941254745		27/04/13		2 gato
	Antón Escribano, José María	C/ Rueda, 17, 3A, 40008, Segovia	912547862		29/08/14		1 perro
	Aragón Peña, Lucía	Avda Constitución, 12, 3D, 40005, Segovia	914756845		07/03/15		4 conejo
	Arranz Gómez, Raquel	C/ Obispo Francisco, 14, 1C, 40003, Segovia	964715481		11/08/08		1 perro
	Arranz Gómez, Raquel	C/ Obispo Francisco, 14, 1C, 40003, Segovia	964715481		16/04/10		2 gato
	Arribas Galindo, Natalia	C/ Los Tejedores, 6, 3A, 40102, Segovia	962551472		16/10/04		1 perro
	Barbado Vega, Rubén	Plaza León, 14, 4A, 40101, Segovia	925415687		26/08/13		1 perro
	Bartolomé Cañas, Iván	C/ Los Castillos, 11, 2C, 40101, Segovia	914547133		28/10/15		2 gato

Toda esta información referente a las mascotas y sus dueños se ha generado individualmente y teniendo en cuenta la raza de las mascotas se ha estimado su peso. En el caso de la edad, se han puesto fechas de nacimiento posibles en la realidad, valorando el tipo de mascota.

Hay que destacar la forma de numeración que tienen las mascotas en cuanto a su código. El código de las mascotas no es aleatorio no casual:

- Los perros (Cod_tipo = 1) tienen el código de mascota 1XX
- Los gatos (Cod_tipo = 2) tienen el código de mascota 2XX
- Las aves (Cod_tipo = 3) tienen el código de mascota 3XX
- Los conejos (Cod_tipo = 4) tienen el código de mascota 4XX
- Los hamsters (Cod_tipo = 5) tienen el código de mascota 5XX
- Los hurones (Cod_tipo = 6) tienen el código de mascota 6XX
- Los cerdos (Cod_tipo = 7) tienen el código de mascota 7XX

- **Hoja de cálculo “Consultas”**

El caso de las consultas es parecido al anterior. Se ha recogido la información que se cree importante sobre las consultas que se realizan en la clínica.

Esta información es: descripción de la consulta y precio, así como un tipo colectivo para diferenciar de manera más genérica a las consultas.

Ilustración 192. Fuente de Información; Consultas

	A	B	C	D	E
1	CODIGO_CONSULTA	DESCRIPCION	PRECIO	COD_CLASE	CLASE_CONSULTA
2	1001	Cirugía Cardiovascular	690,00	10	Cirugía Cardiovascular
3	1101	Espigas auriculares	92,00	11	Cirugía del Oído
4	1102	Otohematoma	185,00	11	Cirugía del Oído
5	1103	Orquiectomía	171,00	11	Cirugía del Oído
6	1201	Colmillos Deciduos	160,00	12	Cirugía Dental
7	1301	Enterotomía	380,00	13	Cirugía Digestiva
8	1302	Enterectomía	420,00	13	Cirugía Digestiva
9	1303	Gastrotomía	390,00	13	Cirugía Digestiva
10	1304	Gastropexia	455,00	13	Cirugía Digestiva
11	1401	Hernia diafragmática	500,00	14	Cirugía General
12	1402	Hernia inguinal	400,00	14	Cirugía General
13	1403	Hernia umbilical	195,00	14	Cirugía General
14	1404	Hernia perineal	395,00	14	Cirugía General
15	1405	Tumor/quiste cutáneo	200,00	14	Cirugía General
16	1406	Laparatomía exploratoria	290,00	14	Cirugía General
17	1407	Esplenectomía	365,00	14	Cirugía General
18	1408	Cistotomía	350,00	14	Cirugía General
19	1409	Tumor Mamario	232,00	14	Cirugía General
20	1410	Mastectomía	300,00	14	Cirugía General
21	1411	Uretrastomía	375,00	14	Cirugía General

Las consultas y los precios son reales de clínicas veterinarias. Están conseguidas de internet de diversas páginas. Los precios son de tipo decimal.

En cuanto a los códigos de las consultas, no están asignados de forma aleatoria, siguen un patón:

- Las Consultas Generales (cod_clase 01) están numeradas del 0101 al 0108
- Los Diagnósticos (Cod_clase 02) están numerados de 0201 al 0206
- Las Pruebas de Laboratorio (Cod_clase 03) están numerados de 0301 al 0306
- Los Tratamientos Preventivos (Cod_clase 04) están numerados de 0401 al 0408
- Los Procedimiento clínico (Cod_clase 05) están numerados de 0501 al 0512
- Las cirugías (Cod_clase 1XXX) están numerados según el tipo que sean:
 - o Cirugía Cardiovascular (Cod_Clase 10) y cod_consulta 1001
 - o Cirugía de oído (Cod_Clase 11) y Cod_Consulta del 1101 al 1103
 - o Cirugía Dental (Cod_Clase 12) y Cod_Consulta 1201
 - o Cirugía Digestiva (Cod_Clase 13) y Cod_Consulta del 1301 al 1304
 - o Cirugía General (Cod_Clase 14) y Cod_Consulta del 1401 al 1411
 - o Cirugía Ocular (Cod_Clase 15) y Cod_Consulta del 1501 al 1505
 - o Cirugía Ortopédica (Cod_Clase 16) y Cod_Consulta del 1601 al 1610
 - o Cirugía Reproductiva (Cod_Clase 17) y Cod_Consulta del 1701 al 1704

- **Hoja de cálculo “Empleados”**

Los empleados de la empresa son cuatro, y cada uno con su cargo correspondiente. Los datos que se reúnen de ellos son: nombre, apellidos, teléfono y el cargo.

Ilustración 193. Fuente de Información; Empleados

	A	B	C	D	E	F
1	CODIGO_EMPLEADO	NOMBRE	APELLIDOS	TELÉFONO	COD_CARGO	CARGO
2	1	Leticia	Aceves Martín	999887711	1	Administrativo
3	2	Miguel Ángel	Garzón López	999333544	2	Doctor
4	3	Iván	Sastre García	999228877	1	Administrativo
5	4	Susana	Alonso Herrero	999115533	2	Doctor

Los datos de los empleados se han creado de forma manual uno a uno.

- **Hoja de cálculo “Productos”**

Los productos como las consultas, tienen pocos campos, en este caso, es lógico tener una columna dedicada a la descripción de lo que hace el producto a la venta, puesto que el nombre es poco aclarativo.

La información a tener en cuenta es: nombre del producto, su descripción y su precio. Además, como en el resto de casos, tenemos una columna para el tipo de producto.

Ilustración 194. Fuente de Información; Productos

	A	B	C	D	E	F
1	CODIGO_PRODUC	NOMBRE	DESCRIPCION	COD_TIPO	TIPO_PRODUCTO	PRECIO
2	1	Blocan 100 ml	Via oral, directamente en la boca del animal o añadido a la comida o al agua. Uso en casos de estreñimiento crónico o intermitente. (Perros y gatos)	1	Digestivo	4,93
3	2	Daforte Suspensión 20 ml	Compacta las heces, reduciendo la severidad y duración de la diarrea. (Perros)	1	Digestivo	9,18
4	3	CANIGEN 7 10 Dosis	Inmunización activa para prevenir infección frente al virus del moquillo, adenovirus tipo 2, parvovirus, virus de parainfluenza canina y las leptospirosis caninas. (Perros)	2	Vacuna	53,15
5	4	DURAMUNE cachorros DP + C 1 DOSIS	reducir la enfermedad causada por el virus del moquillo canino y el parvovirus canino; para reducir la infección intestinal causada por el coronavirus canino. (Perros)	2	Vacuna	6,49
6	5	CANILEISH 1 Dosis	Inmunización activa de perros negativos a leishmania a partir de los 6 meses. (Perros)	2	Vacuna	31,99
7	6	EURICAN PNEUMODOG 1 Dosis	Inmunización activa de perros contra las afecciones respiratorias causadas por Bordetella bronchiseptica y parainfluenza de tipo 2. (Perros)	2	Vacuna	6,31
8	7	CAZITEL PLUS 104 COMPRIMIDOS	Este medicamento contiene antihelmínticos activos contra nematodos y tenias gastrointestinales. (Perros)	9	Anti-parásitos	121,95
9	8	CESTEM PERROS GRANDES 48	Praziquantel actúa contra los cestodos, provocando la contracción y parálisis de los parásitos. (Perros)	9	Anti-parásitos	168,77

Los datos de esta tabla están generados uno a uno, con información extraída de internet de productos de clínicas veterinarias reales.

- **Hoja de cálculo “Diario”**

La hoja de cálculo denominada “Diario” servirá para alimentar una de las tablas de hechos.

La información que se recoge es la de las consultas que se hacen a diario en la clínica veterinaria, por lo que los campos que se deben recoger aquí, son: la fecha en la que se hace la consulta, código de la mascota a quien se realiza la consulta, código del empleado que realiza la consulta, código que identifica la consulta que se realiza, el tipo de la consulta: si es urgente aparecerá un 1 y el importe de la consulta se incrementará en un 20%, y finalmente el precio total de la consulta.

Ilustración 195. Fuente de Información; Diario

	A	B	C	D	E	F
1	FECHA	CODIGO_MASCOTA	CODIGO_EMPLEADO	CODIGO_CONSULTA	TIPO_CONSULTA	IMPORTE
2	01/01/14	406	2	1103	1	205,20
3	01/01/14	210	2	0504	0	28,90
4	01/01/14	217	4	0303	1	46,03
5	01/01/14	309	4	0103	1	30,00
6	01/01/14	118	4	1404	0	395,00
7	03/01/14	110	2	1302	0	420,00
8	03/01/14	130	2	1604	1	516,00
9	03/01/14	125	2	0302	0	32,37
10	03/01/14	113	2	1101	1	110,40
11	04/01/14	503	2	1101	1	110,40

La información recogida en esta tabla, debe pertenecer a los códigos de los datos de las hojas de cálculo que corresponda.

Teniendo en cuenta que la frecuencia de consultas es diaria, vamos a generar los datos de forma aleatoria, de este modo es sencilla la creación de, al menos 500 datos para esta tabla.

- ✓ La *fecha* se ha generado de la siguiente forma:

Ilustración 196. Fuente de Información; Aleatorio Fecha

Fuente	Alineación
<i>f_x</i>	=ALEATORIO.ENTRE(41640;42460)

Este dato, se formatea para que sea de tipo fecha. Debe ser un aleatorio en ese rango porque necesitamos datos entre el 01-01-2014 hasta 31-03-2016

- ✓ Para el *código de las mascotas* atendidas:

Ilustración 197. Fuente de Información; Aleatorio Código Mascotas

Fuente	Alineación	Número
<i>f_x</i>	=INDICE(MASCOTAS!\$A\$2:\$A\$86;ALEATORIO.ENTRE(1;CONTARA(MASCOTAS!\$A\$2:\$A\$86)))	

Devuelve los códigos de mascotas de forma aleatoria.

- ✓ El *código del empleado* se ha generado de la siguiente forma:

Ilustración 198. Fuente de Información; Aleatorio Empleado

Fuente	Alineación
<i>f_x</i>	=SI(ALEATORIO.ENTRE(2;4)=3;4;2)

Devuelve un aleatorio entre dos valores, pero, hay que tener en cuenta, como se ha mencionado anteriormente, que tan solo los doctores podrán pasar consultas. Siendo de esta forma, el aleatorio generado, no puede ser el número 3 (los doctores tienen los números 2 y 4), si esto ocurriera, se pondría un 4 en lugar de un 3.

- ✓ El *código de la consulta*, lo generaremos igual que el código de mascota, atraparemos la columna que corresponda y le creamos los números de forma aleatoria:

Ilustración 199. Fuente de Información; Aleatorio Código Consulta

Biblioteca de funciones	Nombres definidos
<i>f_x</i>	=INDICE(CONSULTA!\$A\$2:\$A\$80;ALEATORIO.ENTRE(1;CONTARA(CONSULTA!\$A\$2:\$A\$80)))

- ✓ El *tipo de la consulta* es un aleatorio entre 0 y 1: “=ALEATORIO.ENTRE(1;4)”
- ✓ El *precio*, tiene más complejidad:

Ilustración 200. Fuente de Información; Precio

Biblioteca de funciones	Nombres definidos	Biblioteca de funciones	Nombres definidos
<i>f_x</i>	=SI(E2=1;(BUSCARV(\$D\$2:\$D\$630;CONSULTA!\$A\$2:\$C\$80;3;0)*0,2)+BUSCARV(\$D\$2:\$D\$630;CONSULTA!\$A\$2:\$C\$80;3;0);BUSCARV(\$D\$2:\$D\$630;CONSULTA!\$A\$2:\$C\$80;3;0))		

Se comprueba si el tipo de consulta es urgente para aumentar el importe de la consulta un 20%. Para aumentar el precio primero hay que comparar el código de la consulta generado de forma aleatoria con el código de consulta que corresponda y adquirir su precio.

Si el tipo de consulta es normal, simplemente comparamos los códigos de consultas y cogemos el precio correspondiente.

Esta columna será de tipo decimal.

- **Hoja de cálculo “Ventas”**

Esta última hoja de cálculo está dedicada a recoger la información de las ventas que se realizan en la clínica veterinaria.

La información, como en el caso del diario, está creada aleatoriamente.

Los datos son generados de forma diaria, puesto que las ventas de productos se realizan a diario. Los datos que se deben tomar en cuenta son: la fecha de la venta, el código del producto que se vende, el trabajador que realiza dicha venta, las unidades vendidas del producto y el importe de los productos vendidos (unidades por precio).

Ilustración 201. Fuente de Información; Ventas

	A	B	C	D	E	F
1	FECHA	CODIGO_PR ODUCTO	CODIGO_EM PLEADO	UNIDADE S	PRECIO_UNIT ARIO	PRECIO_VENTA
2	20/09/2015	38	2	6	20,48	122,88
3	26/10/2015	38	3	5	20,48	102,40
4	19/10/2014	36	3	5	11,87	59,35
5	11/11/2015	72	4	4	2,99	11,96
6	03/02/2014	26	2	8	27,27	218,16
7	04/01/2014	37	4	10	4,70	47,00
8	15/03/2016	33	3	10	11,04	110,40
9	25/08/2015	30	2	5	31,21	156,05
10	08/10/2014	32	3	5	36,79	183,95
11	04/02/2014	41	3	2	4,68	9,36
12	16/02/2014	15	2	2	10,89	21,78
13	03/06/2014	58	2	1	1,49	1,49

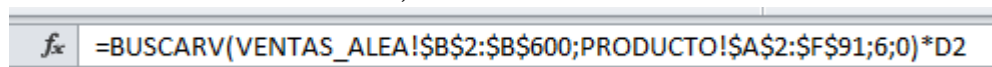
La información recogida en esta tabla, como en la anterior, debe pertenecer a los códigos de los datos de las hojas de cálculo que corresponda.

Teniendo en cuenta que la frecuencia de ventas es diaria, y vamos a generar los datos de forma aleatoria, se crearan al menos 500 datos para esta tabla. Con los 500 registros

anteriores y estos 500, son un total de 1000 registros.

- ✓ Las *fechas*, están creadas de forma aleatoria y pasadas a formato fecha, como en el caso de la hoja de cálculo de diario.
- ✓ El *código del producto* está creado de forma aleatoria entre el rango de valores de códigos de productos existentes, que es de 1 al 90: “=ALEATORIO.ENTRE(1;90)”
- ✓ Para el *código del empleado*, es un aleatorio entre 1 y 4, ya que cualquier trabajador de la clínica puede realizar una venta: “=ALEATORIO.ENTRE(1;4)”
- ✓ Para las *unidades* de productos vendidos, se creará un aleatorio entre 1 y 10: “=ALEATORIO.ENTRE(1;10)”
- ✓ Finalmente, para el *importe*, debe corresponder con el precio del producto por el número de unidades de este que se venden:

Ilustración 202. Fuente de Información; Aleatorio Ventas



`=BUSCARV(VENTAS_ALEA!B2:B600;PRODUCTO!A2:F91;6;0)*D2`

Se realiza una comparación entre el código del producto aleatorio de esta tabla con el código de producto de la hoja de cálculo “Productos”, y se coge el precio que corresponda a dicho producto. Posteriormente se multiplica el precio por el número de unidades vendidas.

Otras Consideraciones a Tener en Cuenta

- **Hojas de cálculo de Diario y Ventas:**

Como la información de las hojas de cálculo está generada de forma aleatoria, cada vez que se abre y cierra el documento Excel se vuelven a generar datos diferentes a los anteriores. Esto hace obligado la copia de la información creada en un momento dado, en otra hoja de cálculo de manera estática, es decir, sin formulas.

Una vez realizada esta copia, nos debemos asegurar que no existen en la información en cuanto a los códigos.

Comprobados los duplicados, vamos a realizar otra comprobación: que no existan consultas inapropiadas, esto es, por ejemplo, que no se haya realizado un tratamiento para el parto o cesáreas a aves o a machos. Que no existan incoherencias de datos.

- **Para finalizar**

Para poder cargar los datos mediante ETL, se han tenido algunos problemas de compatibilidad de formatos del Excel. Es necesario guardar todos los datos en un documento Excel sin formulas ni formato alguno, solo datos, y además, guardar el documento como Excel del tipo 1997.

Los únicos datos que se pueden guardar con formato en celda, serán los correspondientes a importes y fechas. En los importes, es obligatorio que sea decimal para no perder información. En las fechas, para no perder el formato.

ANEXO II: CONTENIDO DEL CD-ROM

En este anexo, vamos a explicar el contenido del CD-Rom adjunto al documento actual. El CD-Rom deberá contener todo lo referente al proyecto y su desarrollo.

- **Documentación:** en el que se encuentra este documento en formato PDF.
- **Operacional:** donde encontramos el fichero tipo Excel de la creación del modelo operacional y el fichero Excel, si formatos, desde el que se extrae la información para realización de carga de datos en el Data Warehouse.
- **ETLs:** carpeta donde se ubica el proyecto de Visual Studio con todas las ETLs.
- **Bases de Datos:** encontramos el script de la estructura de la base de datos donde se cargará la información del operacional y el script con los “Insert into” de la dimensión temporal. Igualmente estarán, la base de datos cargada a partir de las ETLs, y la denominada “Metadata” con el proyecto de Microstrategy.
- **Validaciones Informes:** Directorio llamado Código, donde se encuentra todo el código fuente de la aplicación.

ANEXO III: GLOSARIO

Tabla 137. Glosario

Término	Definición
Proceso Software	Conjunto estructurado de actividades y resultados asociados requeridos para desarrollar un sistema de software
BI o Business intelligence	Business intelligence, es el conjunto de estrategias y aspectos relevantes enfocados a la administración y creación de conocimiento sobre el medio, a través del análisis de los datos existentes en una organización o empresa.
ETL	Proceso de carga de datos en un proyecto de BI, por sus siglas, extration, transformation y Load, es el proceso de extraer y cargar los datos de un operacional en un data mart
Data mart	Repositorio de datos, que contiene el conjunto de datos de un operacional una vez aplicados los procesos de etl
Data warehouse	Repositorio de datos, similar al data mart, pero que puede ser mas amplio en concepto, es decir puede englobar varios data mart (de otros departamentos, por ejemplo)
BBDD operacional	Es la base de datos proporcionada por una empresa cliente para la conformación de un producto de BI.
BBDD informacional	Es un sinónimo en la definición de data warehouse o data mart, es decir, es la BBDD que contiene los datos del proyecto de BI, ya transformados
Dimensión o DIM	Es cada una de las tablas que conforman un data mart o data warehouse
Tabla de hechos o TH	Tabla de bbdd que contiene los IDs de los maestros de todas las dimensiones de un data mart, esto permite tenerla enlazadas para la creación de informes
Informe	Representación de los datos del data warehouse según las necesidades y peticiones del cliente
Requisito	Los requisitos son una especificación de lo que debería ser implementado. Los requisitos son descripciones de cómo el sistema debe comportarse o son propiedades/atributos del sistema.
Requisito Funcional o RF	Descripción del comportamiento que debe mostrar el sistema bajo condiciones determinadas
Requisito no Funcional	Descripción de una propiedad que debe tener el sistema o una restricción que debe respetar
Requisito de Negocio	Objetivo de alto nivel de la organización que solicita el software
Requisito de Usuario	Objetivo o tarea que una clase de usuario podrá realizar con el sistema
Interfaz Usuario	Medio que permite a una persona comunicarse con la máquina
Usuario	Persona que utiliza la plataforma.
Administrador	Persona o usuario del sistema con todos los privilegios sobre el mismo.
Caso de Uso o CU	Descripción de los pasos o las actividades que deberán realizarse para llevar a cabo algún proceso.
Precondición	Condición previa que se tiene que satisfacer
Postcondición	Condición posterior que se tiene que satisfacer.
Granularidad	Especificidad a la que se define un nivel de detalle en una tabla, es decir, si hablamos de una jerarquía la granularidad empieza por la parte más alta, siendo la granularidad mínima el nivel más bajo.
Cardinalidad	Indica el número de elementos de un conjunto, sea esta cantidad finita o infinita. En una tabla de base de datos, la cardinalidad refiere al número de registros contenidos en la misma.