

UNIVERSIDAD DE EL SALVADOR

FACULTAD DE INGENIERIA Y ARQUITECTURA

ESCUELA DE INGENIERIA DE SISTEMAS INFORMATICOS



**PROPUESTA DE FORMACIÓN ACADÉMICA Y
PROFESIONAL PARA EL PERFIL DE
ESPECIALIZACIÓN EN EL ÁREA DE INTELIGENCIA
DE NEGOCIOS Y ANALÍTICA DE DATOS PARA LA
CARRERA DE PREGRADO DE INGENIERÍA DE
SISTEMAS INFORMÁTICOS DE LA UNIVERSIDAD DE
EL SALVADOR**

PRESENTADO POR:

MÓNICA GUADALUPE FLORES BONILLA

MARÍA JOSÉ FLORES RAMOS

MELVIN IVÁN REYES ORTIZ

JUAN DIEGO ZAVALETA MEJÍA

PARA OPTAR AL TÍTULO DE:

INGENIERO DE SISTEMAS INFORMATICOS

CIUDAD UNIVERSITARIA, JULIO DE 2018

UNIVERSIDAD DE EL SALVADOR

RECTOR:

MSc. ROGER ARMANDO ARIAS ALVARADO

SECRETARIO GENERAL:

MSc. CRISTOBAL HERNAN RIOS BENITEZ

FACULTAD DE INGENIERIA Y ARQUITECTURA

DECANO:

ING. FRANCISCO ANTONIO ALARCON SANDOVAL

SECRETARIO:

ING. JULIO ALBERTO PORTILLO

ESCUELA DE INGENIERIA DE SISTEMAS INFORMATICOS

DIRECTOR:

ING. JOSE MARIA SANCHEZ CORNEJO

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERIA Y ARQUITECTURA
ESCUELA DE INGENIERIA DE SISTEMAS INFORMATICOS

Trabajo de Graduación previo a la opción al Grado de:
INGENIERO DE SISTEMAS INFORMATICOS

Título:

**PROPUESTA DE FORMACIÓN ACADÉMICA Y
PROFESIONAL PARA EL PERFIL DE
ESPECIALIZACIÓN EN EL ÁREA DE INTELIGENCIA
DE NEGOCIOS Y ANALÍTICA DE DATOS PARA LA
CARRERA DE PREGRADO DE INGENIERÍA DE
SISTEMAS INFORMÁTICOS DE LA UNIVERSIDAD DE
EL SALVADOR**

Presentado por:

**MÓNICA GUADALUPE FLORES BONILLA
MARÍA JOSÉ FLORES RAMOS
MELVIN IVÁN REYES ORTIZ
JUAN DIEGO ZAVALETA MEJÍA**

Trabajo de Graduación Aprobado por:

Docente Asesor:
ING. ELMER ARTURO CARBALLO RUIZ MSc

SAN SALVADOR, JULIO 2018

Trabajo de Graduación Aprobado por:

Docente Asesor:

ING. ELMER ARTURO CARBALLO RUIZ MSc

AGRADECIMIENTOS

Quiero agradecer a Dios todo poderoso y a María Auxiliadora por haberme iluminado a lo largo de toda la carrera y permitirme ahora poder culminar este trabajo de graduación gracias por siempre darme la fortaleza necesaria para afrontar las adversidades sin rendirme en ningún momento; Gracias Diosito por poner excelentes personas que me ayudaron a crecer a lo largo de este camino y que sin pensarlo se convirtieron en grandes amigos.

A mi madre Norma Bonilla por siempre apoyarme, consentirme y darme sabios consejos; gracias mamá por siempre animarme a seguir adelante y cumplir todas mis metas, comprenderme a lo largo de todos estos años y a tenerme paciencia cuando me sentía abatida , no tengo las suficientes palabras para agradecerte todo lo que has hecho por mí , mamá este logro es para ti.

A mi padre José Flores por siempre confiar en que su tigrita podía con todo , gracias por estar siempre pendiente y apoyarme en todo momento.

A mis segundas Juana López mamá Patty Bonilla , Aminta Orellana y Susy Lovos gracias por su apoyo incondicional, su preocupación por mí y por sus consejos en el momento indicado; un especial gracias mi prima Aida gracias por sacarme sonrisas.

A mi novio William por escucharme , apoyarme y comprenderme , gracia por tus consejos y ayudarme cuando sentía que era demasiado.

Un especial agradecimiento a mi amiga incondicional y hermana de otra mamá María José Flores que desde el inicio de la carrera siempre me apoyo, gracias por tus consejos y paciencia a lo largo de estos años, gracias por todas las risas y los obstáculos que superamos juntas. Gracias a mis amigos y compañeros de tesis Iván Reyes y Juan Diego Zavaleta por su esfuerzo y comprensión durante la carrera y sobre todo en el trabajo de graduación.

A mis otros hermanitos Edgar y Eduardo por todos los trabajos, las desveladas y momentos de alegrías que pasamos juntos en la carrera, gracias por siempre echarme porras y apoyarme a seguir perseverando.

A mis amigas Jackeline y Jessica por su apoyo incondicional, por escucharme cuando más lo necesitaba gracias por estar siempre pendientes y confiar en mí.

A mis amigos de la carrera Andrés, Ricardo, Douglas, Nelson y a todos los demás compañeros que compartimos en esta aventura llamada Ingeniería

A nuestro asesor el Ing, Elmer Arturo Carballo por haber apoyado y confiado en este proyecto su guía fue muy importante para desarrollarlo, gracias por hacernos dar lo mejor de nosotros. Un agradecimiento a todas aquellas profesionales que nos apoyaron en responder las diferentes encuestas que conforman este trabajo, su apoyo fue muy valioso.

Mónica Guadalupe Flores Bonilla

AGRADECIMIENTOS

Agradezco a Dios Todopoderoso y a la Santísima Virgen María por la oportunidad que recibí para educarme y por darme la sabiduría necesaria para completar mis estudios universitarios, por las muchísimas bendiciones que pusieron en mi camino para poder terminar este ciclo de mi vida, por su amor al permitirme conocer excelentes personas en la universidad que ahora puedo llamar amigos, pero sobre todo por darme una madre buena, fuerte y trabajadora, quien es la mayor impulsora de este logro.

Dedico este trabajo a mi madre, Olimpia Concepción Ramos, pues gracias a su esfuerzo pude tener una buena educación y una excelente calidad de vida. Agradezco su amor incondicional demostrado de tantas maneras, agradezco su sabiduría a través de sus consejos y correcciones, así como su apoyo durante este proceso cuando me veía abatida y estresada. Mamá, siéntete orgullosa, porque tus sacrificios dieron ya sus frutos, lograste entre muchas otras cosas, convertir a tus dos hijas en profesionales.

Agradezco a mi hermana mayor Valeria, porque siempre se ha esforzado por darme un buen ejemplo, por siempre cuidarme a pesar que ya estoy grande, por apoyarme y ser comprensiva y por ser incondicional en las buenas y en las malas. Agradezco los buenos momentos que me haces pasar hermana, porque en ellos me das tu gran amor y tu alegría.

Agradezco a mi abuelita Marina, porque me cuidó y atendió en mis años de estudio en el colegio y en la universidad, porque siempre confió en la capacidad de sus dos nietas y ha esperado pacientemente que llegue el momento de mi graduación. Gracias a Romy, a mis tías, a mis padrinos y demás familia por su amor y su apoyo.

Gracias a mi equipo de trabajo, porque a pesar de los momentos de desacuerdo y de alta tensión logramos trabajar y culminar este proyecto. A Mónica, que ha sido mi amiga incondicional desde el primer ciclo y la mejor compañera que pude tener en toda la carrera, a Juan Diego e Iván por su esfuerzo en este proyecto.

Gracias a mi amiga Kimberly, por siempre escucharme, hacerme reír y por confiar en mí, por siempre darme ánimos y cariño, pero sobre todo por haberme ayudado dos veces en este proceso para que mi equipo pudiera realizar sus últimas dos defensas. Comparto mi felicidad con mis compañeros y fieles amigos Eduardo y Edgardo, ya que iniciamos juntos esta carrera, estudiamos y nos divertimos muchísimo en los cinco años de la carrera, gracias niños por su apoyo en los momentos malos y por tanto cariño hacia mi persona.

Agradezco a mis demás compañeros de la carrera por haber compartido conmigo su conocimiento y alegría en estos años de estudio y a las nuevas personas que Dios ha puesto en mi camino que se unen a mi felicidad.

Agradecimiento especial a las personas que permitieron el desarrollo de este proyecto, invirtiendo su tiempo en responder las encuestas.

María José Flores Ramos

AGRADECIMIENTOS

A mi Padre Celestial que me ha acompañado siempre, en las buenas y en las malas, en los éxitos y en los fracasos y porque nunca me ha abandonado. Porque le ha dado sentido a mi vida y me ha guiado en sus sabios caminos. “Porque yo sé muy bien los planes que tengo para ustedes —afirma el Señor—, planes de bienestar y no de calamidad, a fin de darles un futuro y una esperanza”. Jer. 29:11.

A mi madre por haberme apoyado en mis estudios y en mi vida, por su sacrificio sin haber esperado nada a cambio. Te agradezco por todo lo que has hecho y por haberme enseñado tanto.

A mi padre por sus consejos y por su deseo de verme prosperar y llegar a ser un profesional de bien. Muchas gracias por tu apoyo, tu dirección y tu ejemplo.

A mi hermano quien ha compartido conmigo durante toda mi vida y en quien he encontrado apoyo incondicional y una gran amistad.

A mis hermanos del Ministerio Manos de Dios que me tuvieron presente en sus oraciones. Quienes semana a semana muestran ejemplo de su desinteresada labor con el prójimo ayudando al necesitado y al desprotegido.

A Walter Gavarrete por haberme apoyado en todo el proceso de mi trabajo de graduación. Y por darme la oportunidad para desarrollarme como profesional.

A mis compañeros Mónica, María José y Juan Diego, por su colaboración y compromiso para con este proyecto, les estoy eternamente agradecido.

A niña Cony por habernos recibido en su hogar y habernos tratado como familia. Le agradezco por su generosidad y por haber abierto las puertas de su casa.

A las personas que colaboraron con los instrumentos de sondeo utilizados en esta investigación, gracias por el tiempo prestado y por aportar al desarrollo de esta propuesta que busca el beneficio de la educación salvadoreña.

Melvin Iván Reyes Ortiz

AGRADECIMIENTOS

Primeramente, a Dios todo poderoso, por bendecirme, darme vida y fortaleza para seguir adelante ante cualquier obstáculo presentado en esta etapa de la vida.

A mi madre, Edith América Mejía, que siempre está a mi lado, que con mucho esfuerzo, me ayudo a emprender este largo camino, gracias por su apoyo, una persona que siempre admiró y admiraré, que me enseñó el significado de seguir adelante pase lo que pase; una persona fuerte que siempre iniciaba sus días a las cuatro de la mañana para conseguir el dinero para que pudiera seguir adelante con mis estudios; que cuando yo no creía en mí mismo, ella siempre lo hizo por mí; este triunfo es fruto de ella, te amo mucho madrecita.

A mis hermanas Karla Irene y Clara Lizeth, por cada uno de sus consejos, gracias hermanitas por siempre estar conmigo, las quiero mucho y estoy bendecido con ustedes.

A mi abuelita Marta Mejía, mis tías, mis primos, sobrinos en especial Victor Manuel y Florence Camila, padre y toda mi familia por siempre creer en mí, y su apoyo incondicional, los quiero mucho.

A mis compañeros de equipo: María José, Mónica Guadalupe y Melvin Iván, que más que compañeros, son mis amigos; los cuales les tengo un gran aprecio, y a pesar de todas las diferencias debidas a los obstáculos presentados en este proyecto, jamás me arrepentiré de haberlos escogidos como mi equipo de trabajo, los quiero mucho.

A cada uno de mis amigos y amigas: Zaira Figueroa, Fernando Bolaños, Samuel Flores, Nelson Rubio, Rhina Tejada, Roger Nieto, Christian Guillen, Mónica Solorzano, Gabriela Esmeralda, Nelson Alberto, Débora Guardado, Yamileth Alas, Elvis Ticas, Mario Vides, Sarahi Torrento, Luis Morales, Beatriz Rubio, Alan Quijada, Kevin Núñez, Henry Gutiérrez, Ivania Rivas, Victoria Carranza, Cesar Martínez, Miguel Zelaya, Claudia Guerrero, Enrique Bermudez, Marcela López, Samuel Zepeda, William García, Erick Domínguez, Will Bonilla, Douglas Figueroa, Milton Aragón, Andrés Rodríguez, Milton Orellana, Ricardo Rivas entre muchos más; que fueron mis compañeros y que se convirtieron en mis amigos, que siempre me hacen reír, y me ayudan con cada uno de sus consejos, los quiero mucho y siempre serán como una familia para mí.

A mi asesor Ing. Elmer Carballo por todo su apoyo y paciencia, gracias por guiarnos en esta etapa de la carrera y vida, en lo cual demostró ser un gran profesional y persona. A los docentes de la Escuela de Ingeniería de Sistema Informáticos, en especial a las partes involucradas en este proyecto Ing. José María Sánchez, Ing. Cesar Gonzales, Ing. Balmore Ortiz e Ing. Yesenia Vigil.

A los diferentes profesionales que compartieron de su conocimiento con sus respuestas, comentarios y observaciones en las diferentes encuestas y sondeos realizados en este proyecto.

Muchas gracias a todos, bendiciones por siempre...

Juan Diego Zavaleta Mejía

CONTENIDO

CONTENIDO.....	v
ÍNDICE DE FIGURAS	viii
ÍNDICE DE TABLAS	viii
ÍNDICE DE GRÁFICOS	viii
INTRODUCCIÓN	ix
OBJETIVOS.....	xi
Objetivo General.....	xi
Objetivos Específicos	xi
CAPÍTULO I: EL PROBLEMA DE INVESTIGACIÓN.....	1
1.1. Situación problemática.....	2
1.2. Planteamiento del problema	4
1.2.1. Variables de estudio.....	4
1.3. Situación deseada	4
1.4. Preguntas de investigación.....	5
1.5. Alcances y Limitaciones.....	5
1.5.1. Limitaciones de la investigación	5
1.5.2. Alcances.....	5
1.6. Justificación del estudio	6
1.6.1. Importancia	7
1.6.2. Resultados Esperados	10
1.7 Definición tipo de investigación.....	10
CAPÍTULO II: MARCO TEÓRICO.....	14
2.1. Antecedentes del estudio.....	15
2.2. Bases teóricas	20
2.2.1. Inteligencia de Negocios	20
a) Definición.....	20
b) Componentes de la Inteligencia de Negocios	21
c) Modelado de Data Warehouse	23
d) Elementos de un Data Warehouse	26
e) Proceso Analítico en Línea (OLAP)	27
f) Extracción, Transformación y Carga (ETL).....	30

g) Diseño de reportes / Visualizaciones	32
2.2.2. Analítica de Datos	35
2.3. Marco Conceptual.....	46
CAPÍTULO III: ESTUDIO DE LA SITUACIÓN ACTUAL DE LA INTELIGENCIA DE NEGOCIOS Y ANALÍTICA DE DATOS EN EL SALVADOR.....	51
3.1 Metodología para la recolección de datos.....	52
3.1.1. Definición de la población y muestra	52
3.1.2. Diseño del instrumento.....	53
3.2 Resultados del estudio.....	57
CAPÍTULO IV: CONSTRUCCIÓN DEL PERFIL ACADÉMICO PARA EL ESTUDIANTE DE ESPECIALIZACIÓN DE INTELIGENCIA DE NEGOCIOS Y ANALITICA DE DATOS	59
4.1 Metodología para el diseño curricular	60
4.1.1. Investigación de metodologías de diseño curricular.....	60
4.1.1.1. ACM (Association for Computing Machinery).....	60
4.1.1.2. ACAI- LA (Adopción de enfoques de calidad, accesibilidad e innovación en la educación superior de Latinoamérica).....	64
4.1.1.3. Tuning.....	66
4.1.2. Comparación de metodologías estudiadas.....	70
4.1.3. Selección de metodología curricular para el proyecto.....	75
4.2. Herramientas para recolección de datos.....	75
4.2.1 Encuesta a profesionales en el área de Inteligencia de Negocios y Analítica de datos.....	76
4.2.2 Sílabos de instituciones educativas.....	77
4.2.3 Bolsas de trabajo nacional e internacional.....	79
4.3. Perfil académico para el estudiante de especialización de Inteligencia de Negocios y Analítica de datos.....	80
4.3.1. Objetivo general de la especialización.....	80
4.3.2. Competencia general de la Especialización	80
4.3.3. Sub-competencias de la Especialización.....	80
4.4. Modelo Profesional.....	84
4.3.5. Descripción del Perfil del profesional en Inteligencia de Negocios y Analítica de datos.....	85
4.3.6. Roles del Perfiles académicos de la especialización	85

CAPÍTULO V: PROPUESTA DE ESPECIALIZACIÓN DE INTELIGENCIA DE NEGOCIOS Y ANALITICA DE DATOS	87
5.1. Valorización de los Programas de Estudio propuestos	88
5.2. Material Didáctico de la especialización.....	92
5.3. Sesiones.....	93
5.4. Programas de estudio.....	94
CAPITULO VI: PROPUESTA DE CAPACITACION DE DOCENTES EISI.....	130
6.1 Propuesta de Capacitación para el área de Fundamentos de Negocio.....	131
6.2 Perfil Profesional Externo.....	148
CONCLUSIONES	150
RECOMENDACIONES	151
GLOSARIO	152
ANEXOS	158
Anexo 1: Flujo de Trabajo del Proyecto	159
Anexo 2: Lluvia de Ideas para el análisis y determinación de la situación problemática.....	160
Anexo 3: Análisis de Resultado de Encuesta “Investigación de la Situación Actual de la inteligencia de negocios y analítica de datos”	161
Anexo 4: Cálculo de muestra para poblaciones infinitas o desconocidas.....	161
Anexo 5: Diseño de la Encuesta de Sondeo del Perfil de Inteligencia de Negocios y Analítica de Datos	162
Anexo 6: Análisis de Resultados de la Encuesta de Sondeo del Perfil de Inteligencia de Negocios y Analítica de Datos	162
Anexo 7: Matrices de resumen de Sílabos de Universidades.....	162
Anexo 8: Análisis de Bolsas de Trabajo.....	162
Anexo 9: Análisis de Resultados de Encuesta de Valorización	162
Anexo 10: Diagnóstico de las competencias en Inteligencia de Negocios y Analítica de datos de docentes EISI.....	163
BIBLIOGRAFÍA	192

ÍNDICE DE FIGURAS

Figura 1 Diagrama de árbol.....	3
Figura 2 Enfoque de Inmon.	24
Figura 3 Enfoque de Kimball.	25
Figura 4 Tipos de gráficos.	34
Figura 5 Tipos de Analítica de Big Data.	45
Figura 6 Estructura del modelo curricular IS 2010.....	62
Figura 7 Roles del perfil profesional	85

ÍNDICE DE TABLAS

Tabla 1 Comparación de Metodologías	71
Tabla 2 Análisis preguntas cerradas	76
Tabla 3 Análisis preguntas abiertas.....	76
Tabla 4 Consolidación de categorías	77
Tabla 5 Matriz resumen.....	78
Tabla 6 Áreas de conocimiento comunes por institución	78
Tabla 7 Comparativo de instituciones.....	78
Tabla 8 Competencias por área de conocimiento.....	79
Tabla 9 Plan de capacitación	132
Tabla 10 Monto total de las capacitaciones.....	147

ÍNDICE DE GRÁFICOS

Gráfico 1 ¿Por medio de qué fuentes conoce de Inteligencia de Negocios y Analítica de datos?.....	57
Gráfico 2 ¿Considera que en su empresa o en su gerencia/departamento es importante que los datos se analicen con mayor énfasis?.....	58
Gráfico 3 ¿Considera que la información que actualmente posee es suficiente para tomar decisiones?	58
Gráfico 4 Instrumento de valorización para Fundamento de Negocio.....	89
Gráfico 5 Instrumento de valorización para Data Warehouse.....	90
Gráfico 6 Instrumento de valorización para Inteligencia de Negocios.....	91
Gráfico 7 Instrumento de valorización para Gestión de Big Data.....	91
Gráfico 8 Instrumento de valorización para Análisis Predictivo de Datos.....	92

INTRODUCCIÓN

La Escuela de Ingeniería de Sistemas Informáticos de la Universidad de El Salvador ofrece el Plan de Estudios 1998 para la carrera de Ingeniería de Sistemas Informáticos; Sin embargo, el Comité Curricular de la Escuela de Ingeniería de Sistemas Informáticos está trabajando en una iniciativa de actualización de currícula, por lo que surge el trabajo “PROPUESTA DE FORMACIÓN ACADÉMICA Y PROFESIONAL PARA EL PERFIL DE ESPECIALIZACIÓN EN EL ÁREA DE INTELIGENCIA DE NEGOCIOS Y ANALÍTICA DE DATOS PARA LA CARRERA DE PREGRADO DE INGENIERÍA DE SISTEMAS INFORMÁTICOS DE LA UNIVERSIDAD DE EL SALVADOR” como un apoyo a la iniciativa en construcción.

El actual plan de estudios no contempla las áreas de Inteligencia de Negocios y Analítica de Datos, las cuales son una necesidad real en las organizaciones y empresas en El Salvador y a nivel centroamericano. Por tanto, se planteó dar solución a la situación a través de una investigación bibliográfica y por medio de encuestas, cuyos resultados permitieran la creación de una propuesta de especialización de pregrado.

La propuesta que se presenta, plantea un perfil académico y profesional basado en competencias y sub competencias, las cuales se alcanzan adquiriendo un conjunto de conocimientos representados en seis roles: Analista de Negocios, Desarrollador de Data Warehouse, Analista de Inteligencia de Negocios, Arquitecto de Big Data, Analista de Big Data y Minero de datos.

Los resultados de la investigación realizada se materializan en cinco programas de estudio que representan las asignaturas de la propuesta de especialización, guiones teóricos y presentaciones de clases, guías de laboratorio/discusiones y sesiones de clase. Estas últimas proponen una guía a los docentes para impartir el contenido de cada asignatura.

Dicho esto, se presenta la estructura del documento, el cual consta de siete capítulos.

En el Capítulo I denominado El Problema de Investigación se describe la situación problemática y se plantean las preguntas de investigación y los objetivos a alcanzar. Se detalla además los alcances y limitaciones de la investigación realizada y la importancia de la misma.

En el Capítulo II se describe los antecedentes de la Inteligencia de Negocios y la Analítica de datos y se presenta una base teórica de los temas en estudio, abarcando los temas de Data Warehouse, Big Data y Minería de datos.

En el Capítulo III se define el tipo de investigación que caracteriza a este proyecto.

En el Capítulo IV se detalla la investigación realizada para determinar la situación actual de la Inteligencia de Negocios y Analítica de datos en El Salvador. Este

instrumento es de gran importancia para el proyecto pues con él se justifica que las áreas en mención se requieren dentro del perfil profesional de un Ingeniero Informático.

En el Capítulo V se presenta en primera instancia el estudio que se hizo sobre metodologías de diseño curricular que guiaron a la determinación de un perfil en base a competencias. Además, en este capítulo se detallan todos los instrumentos que sirvieron de insumo a la determinación de conocimientos y habilidades, presentando finalmente el perfil académico y profesional de la propuesta de especialización.

En el Capítulo VI se detallan los resultados de la investigación (material didáctico, sesiones de clases para docentes y los programas de estudio de cada asignatura). Se presenta además el instrumento que permitió retroalimentar los programas de estudio.

En el Capítulo VII se presenta una propuesta de plan de capacitación de docentes en base a los resultados obtenidos en el último instrumento realizado con el afán de proponer las áreas de conocimiento que los docentes pueden reforzar para estar en sintonía con el contenido propuesto en la especialización.

Se espera que con el producto final de este proyecto se aporte un insumo tangible que agregue más valor a la ya iniciada tarea de actualización del plan de estudio de la carrera de Ingeniería de Sistemas Informáticos; así como también se marque un punto de referencia para la creación de otras especializaciones en la carrera que fortalezcan el perfil profesional de los estudiantes.

OBJETIVOS

Objetivo General

- Construir una propuesta de especialización de pregrado en el área de Inteligencia de Negocios y Analítica de Datos para la carrera de Ingeniería de Sistemas Informáticos de la Facultad de Ingeniería y Arquitectura de la Universidad de El Salvador a través de la determinación de los conocimientos y las habilidades técnicas que forman parte del perfil académico y profesional del área a nivel centroamericano.

Objetivos Específicos

- Definir el perfil académico y profesional de Inteligencia de Negocios y Analítica de Datos que determine los conocimientos y habilidades de la propuesta de especialización.
- Diseñar el programa de estudio de cada una de las asignaturas que conforman la Especialización de Inteligencia de Negocios y Analítica de Datos.
- Elaborar el material didáctico de enseñanza y aprendizaje: guiones de clase, guías de laboratorio, guías de discusión, presentaciones y sesiones de clase para facilitar el desarrollo de las competencias formativas.
- Hacer una propuesta de capacitación para los docentes del Departamento de Desarrollo de Sistemas de la Escuela de Ingeniería de Sistemas Informáticos que garantice el dominio avanzado de los temas que conforman la especialización.

CAPÍTULO I: EL PROBLEMA DE INVESTIGACIÓN

1.1. Situación problemática

La Inteligencia de Negocios y Analítica de datos se ha posicionado en las empresas públicas y privadas de El Salvador como una necesidad emergente. Aproximadamente desde hace más de una década se inició la creación de unidades de Inteligencia de Negocios en las empresas salvadoreñas y por ende la oferta de plazas. En el año 2013 se empezaron a impartir cursos de especialización y posgrados en el área por parte de instituciones educativas superiores y organizaciones. Estos hechos hacen que los estudiantes de sistemas informáticos interesados en el tema requieran como parte de su aprendizaje, recibir los conocimientos, las habilidades técnicas y destrezas en el manejo de herramientas de Inteligencia de Negocios y Analítica de Datos, lo cual al nivel de pregrado no se satisface en la mayoría de las universidades del país.

La Escuela de Ingeniería de Sistemas Informáticos de la Facultad de Ingeniería y Arquitectura de La Universidad de El Salvador, es conocedora de dicha situación y requiere generar soluciones a las necesidades actuales y futuras para satisfacer las nuevas tendencias en formación profesional. Proporcionar esta solución a los estudiantes de la carrera de Ingeniería de Sistemas Informáticos, implica dos acontecimientos; el primero está orientado a la inclusión en la malla curricular de los dos temas en cuestión; y el segundo está relacionado con el dominio avanzado de los temas a un nivel conceptual y práctico por parte de los docentes de la Escuela, que garanticen una enseñanza de calidad.

Para identificar la situación problemática de la investigación se utilizó la técnica denominada lluvia de ideas (Ver anexo 2), en donde los investigadores colocaron sus criterios en base a la experiencia en el tema, a la búsqueda de fuentes secundarias y a los resultados obtenidos de un estudio de sondeo realizado, titulado “Investigación sobre la situación actual de la Inteligencia de Negocios y Analítica de datos” (Ver anexo 3). Posteriormente se analizaron y consolidaron las ideas en un “árbol del problema”¹ con el cual se describe el problema de estudio, así como sus causas y efectos. Dicha técnica se asemeja en su representación a un árbol; en la parte de en medio, simulando el tronco, se encuentra el problema; en la parte inferior, simulando las raíces, se han colocado las causas y en la parte superior simulando las hojas se han colocado los efectos.

¹ Betancourt, D. (2016). Cómo hacer un árbol de problemas. mayo 26, 2017, de Ingenio Empresa Sitio web: <https://ingenioempresa.com/arbol-de-problemas/>

Diagrama de árbol para identificar la Situación Problemática

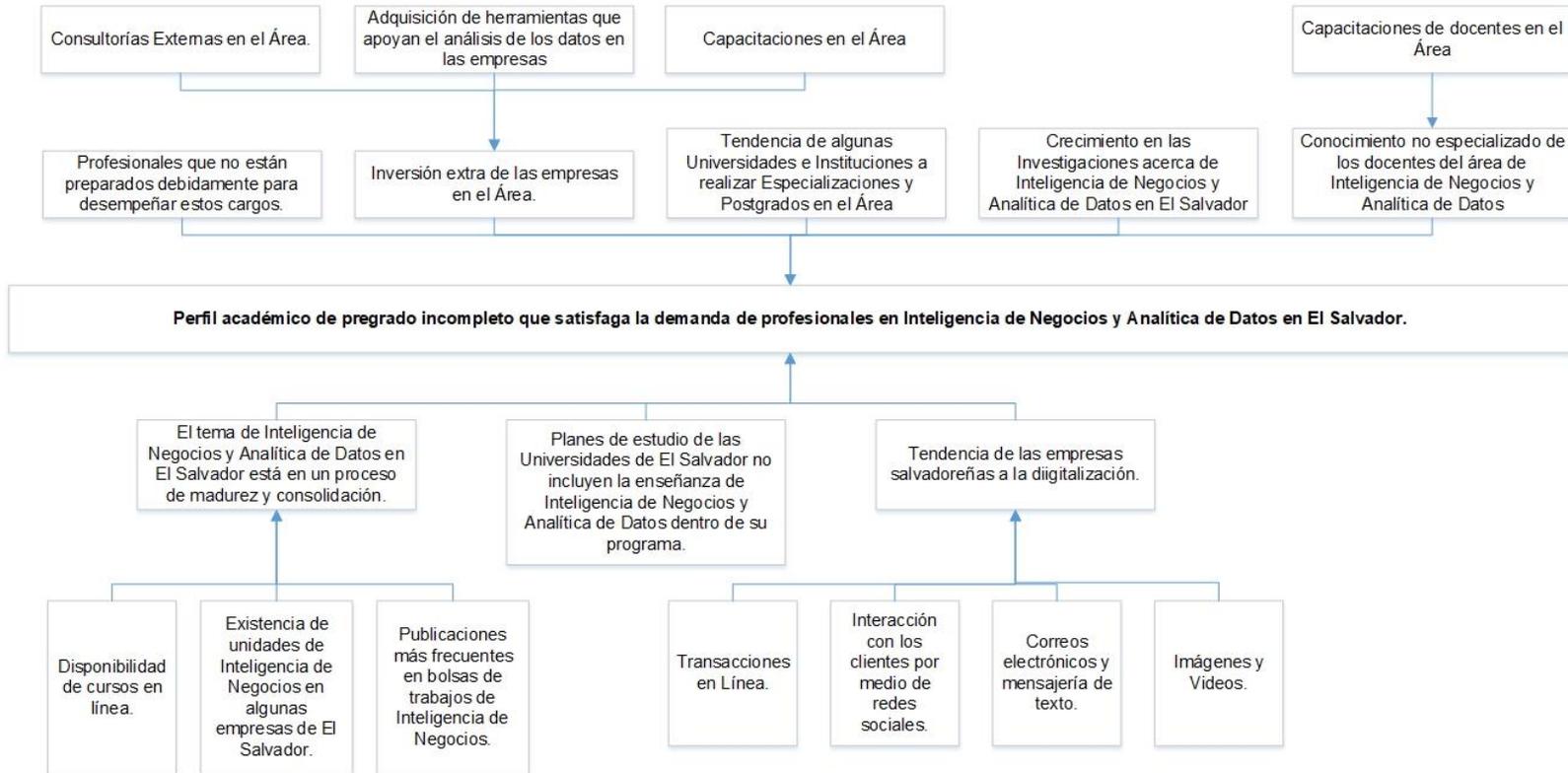


Figura 1 Diagrama de árbol

1.2. Planteamiento del problema

Tomando como base la situación problemática, surge la siguiente interrogante: ¿Cuáles son los conocimientos, las habilidades técnicas y las herramientas que deben formar parte del perfil académico y profesional en el área de Inteligencia de Negocios y Analítica de Datos para poder formar en esta área a los estudiantes de pregrado de la carrera de Ingeniería de Sistemas Informáticos de la Universidad de El Salvador?

1.2.1. Variables de estudio

- La necesidad de profesionales especializados en Inteligencia de Negocios y Analítica de Datos en El Salvador.
- La oferta de cursos de especialización y posgrados de parte de instituciones educativas superiores y otras organizaciones.
- Los conocimientos que forman parte del perfil de especialización de Inteligencia de Negocios y Analítica de Datos.
- Las habilidades técnicas y herramientas de Inteligencia de Negocios y Analítica de Datos fundamentales para un perfil académico y profesional en el área.
- El grado de dominio que tienen los docentes del Departamento de Desarrollo de Sistemas en las áreas de Inteligencia de Negocio y Analítica de Datos.

1.3. Situación deseada

Un perfil académico y profesional de Inteligencia de Negocios donde se definan los conocimientos que se requieren para desarrollarse en esta área.

Un perfil académico y profesional de Analítica de datos en donde se definan los conocimientos que se requieren para desarrollarse en esta área.

Una currícula actualizada con las áreas de Inteligencia de Negocios y Analítica de datos para la carrera de Ingeniería de Sistemas Informáticos de la Universidad de El Salvador.

Docentes del Departamento de Desarrollo de Sistemas con conocimientos especializados en el área de Inteligencia de Negocios y Analítica de Datos.

1.4. Preguntas de investigación

- ¿Qué conocimientos debe tener un profesional que se dedique a la Inteligencia de Negocios y Analítica de Datos?
- ¿Cuáles son los requerimientos técnicos más demandados en el área laboral de Inteligencia de Negocios y Analítica de Datos?
- ¿Cuáles son las técnicas y herramientas que ayudan a los profesionales a aplicar Inteligencia de Negocios y Analítica de Datos?
- ¿De qué manera se formará al estudiante para que se desempeñe en esas áreas?
- ¿Qué conocimientos requieren ser reforzados en el docente del Departamento de Desarrollo de Sistemas para que tenga un dominio avanzado en estas áreas?

1.5. Alcances y Limitaciones

1.5.1. Limitaciones de la investigación

- Las investigaciones de campo se limitan solo a empresas e instituciones tanto públicas como privadas del área metropolitana.
- La selección de la(s) herramienta(s) propuestas para el desarrollo del curso académico y profesional estará condicionado por el acceso gratuito o bien por licencias estudiantiles.

1.5.2. Alcances

- Identificación de las necesidades a nivel centroamericano sobre el perfil de formación en el área de Inteligencia de Negocios y Analítica de Datos.
- Investigación sobre el cuerpo de conocimiento necesario para un perfil de un especialista en Inteligencia de Negocios y Analítica de Datos de acuerdo con las necesidades de la región centroamericana.
- Propuesta de conocimientos, habilidades y valores que el estudiante en el área de Inteligencia de Negocios y Analítica de Datos que deberá poseer en su formación profesional.

- Propuesta de asignaturas para el curso de especialización en Inteligencia de Negocios y Analítica de Datos.
- Propuesta de Capacitación en las áreas de Inteligencia de Negocios y Analítica de Datos para los docentes del Departamento de Desarrollo de Sistemas de la Escuela de Ingeniería de Sistemas Informáticos.

1.6. Justificación del estudio

La Universidad de El Salvador, crítica de la realidad y con capacidad de proponer soluciones a los problemas nacionales se visualiza ser un transformador y líder de la Educación Superior. Dado que la situación problemática identificada tiene sus alcances en el área de la informática le corresponde a la Escuela de Ingeniería de Sistemas Informáticos, de la Facultad de Ingeniería y Arquitectura investigar, innovar y desarrollar la educación de tal manera que los profesionales que forman, tengan las capacidades que satisfagan las necesidades presentes del sector laboral salvadoreño.

A partir del año 2000 la Inteligencia de Negocios y la Analítica de Datos vinieron a impactar a las empresas del país, reconociéndose la necesidad de profesionales especializados en el área. Desde el año 2013 hasta la fecha en que se está realizando este estudio, existen universidades e instituciones que ofrecen especializaciones y postgrados en la materia:

- La Universidad Tecnológica ofrece la pre especialización de graduación “Desarrollo de soluciones en Inteligencia de Negocios para entornos empresariales” para sus estudiantes de pregrado.
- La Universidad Francisco Gavidia posee una especialización de graduación en Administración y Desarrollo de Base de datos e Inteligencia de Negocios.
- Fepade ofreció un postgrado internacional en Inteligencia de Negocios en el año 2015.

El contenido que se ofrece en los cursos de la especialización de graduación mencionados no se ha actualizado desde su creación. Además, están pensados para llevarse a cabo en un periodo de seis a ocho meses lo que limita la profundidad del aprendizaje y el desarrollo de las competencias que se demandan. Estos dos hechos, hacen que surja la oportunidad de crear una propuesta de formación académica y profesional para el perfil de especialización en el área de Inteligencia de Negocios y Analítica de Datos para la carrera de pregrado de esta alma mater.

Se realizó un estudio para conocer la situación actual de la Inteligencia de Negocios y Analítica de Datos visto desde la perspectiva de los profesionales que consumen la información, es decir profesionales que ejercen cargos de jefaturas, gerencias y altos mandos en empresas públicas y privadas. Este instrumento se detalla en el capítulo IV

De los resultados obtenidos (Ver anexo 3) se analiza lo siguiente:

1. Se evaluó el conocimiento de las personas que consideraron estar familiarizadas con el tema de estudio, En promedio el 59.53% de las personas identificó correctamente el concepto de Inteligencia de Negocios y el de Analítica de Datos, sin embargo, para el segundo caso, sólo el 4.76% separa a las personas que si saben de las que no.

Una proporción tan pequeña demanda interés y esfuerzos por especializar a los profesionales en las áreas de su competencia, para garantizar que sean expertos en lo que se desenvuelven y no se conviertan en ingenieros genéricos de conocimiento que tratan saber de todo.

2. Las 24 personas encuestadas consideran de manera unánime que los datos que se generan en sus empresas deben analizarse con un mayor énfasis; lo que se traduce en un reconocimiento de la necesidad de aplicar Inteligencia de Negocios y Analítica de Datos en los procesos de sus negocios.
3. El 52% de las personas que participaron en la encuesta conocen la existencia de la Inteligencia de Negocios porque en su empresa existe un área como tal y un 33% sabe que por lo menos en una empresa del país existe esta área. Este 85% en conjunto representa el posicionamiento de la Inteligencia de Negocios en las empresas salvadoreñas y por ende el crecimiento de la demanda de profesionales que las trabajen.
4. Existe una disconformidad del 54% por parte de los directivos con respecto a la información que actualmente poseen para tomar decisiones, reconociendo que con la Inteligencia de Negocios y Analítica de Datos la toma de decisiones es más oportuna y precisa. La realización de proyecciones, el descubrimiento de nuevos datos y el conocimiento de comportamientos futuros de los clientes fueron los beneficios de la Inteligencia de Negocios y Analítica de datos que los directivos identificaron con más recurrencia.

Por tanto, existe un pronunciamiento de Inteligencia de Negocios y Analítica de Datos en las empresas salvadoreñas y por ende de profesionales dedicados al área, es preciso entonces, que la Universidad de El Salvador cumpliendo en su misión educativa, se enfoque en darle solución a las necesidades actuales desde el área de su competencia: la educación.

1.6.1. Importancia

La propuesta de solución planteada es la creación de una especialización de pregrado en la carrera de Ingeniería de Sistemas Informáticos; la cual está orientada a que el estudiante adquiera las competencias definidas en el perfil de la especialización.

Esto impacta a tres diferentes sectores: los estudiantes y futuros ingenieros de sistemas informáticos, la Universidad de El Salvador y la Sociedad Salvadoreña.

1. Estudiantes y futuros ingenieros de Sistemas Informáticos

Dado que las áreas de estudio tienden al crecimiento a nivel centroamericano y nacional, se espera que la necesidad de profesionales crezca. El Foro Económico Mundial (The World Economic Forum) en colaboración con Business Insider publicó en mayo del 2016 “Los 8 trabajos que las compañías demandarán para el 2020 ²”. En éste, los analistas de datos serán una de las profesiones más demandadas dado a la necesidad de darle sentido a la mezcla de datos no estructurados y estructurados que está generando la tecnología. Por tanto, esto beneficia a los futuros profesionales, al tener mejores oportunidades laborales y mayores posibilidades de selectividad gracias a los conocimientos técnicos específicos y las habilidades que adquieran.

El Índice Global de Habilidades 2015 - 2016, elaborado por Hays en colaboración con Oxford Economics, a petición de Forbes México, expresa la demanda de profesionales en las industrias de alta especialización: tecnológica, telecomunicaciones, energética, farmacéutica y manufactura. Gerente en Inteligencia de Negocios, en el sector TI es una de las diez tendencias laborales del 2016 publicadas por Forbes México ³. Según éste, la tendencia indica que cada vez más compañías alrededor del mundo necesitan profesionales con un perfil de Inteligencia de Negocios.

2. Universidad de El Salvador

La Universidad de El Salvador ha innovado recientemente en la manera de educar, inaugurando en el año 2016 la nueva modalidad de educación en línea, con la cual se pretende la inclusión de los jóvenes salvadoreños a la educación superior. Los desafíos por seguir transformado la educación y mantenerse en un panorama protagónico de entre todas las universidades del país se encuentran latentes.

La Escuela de Ingeniería de Sistemas Informáticos de la Facultad de Ingeniería y Arquitectura de la Universidad de El Salvador está trabajando en una propuesta de actualización del plan de estudios desde hace ocho años, considerando al área de Inteligencia de Negocios y Analítica de Datos desde hace tres.

La creación de un perfil de especialización en esta área es conveniente al servir de apoyo a la ya comenzada labor de la escuela, que actualmente se encuentra en la fase de elaboración del contenido temático de las asignaturas.

² Thompson, C.. (2016). 8 jobs every company will be hiring for by 2020. abril 09,2017, de World Economic Forum Sitio web: <https://www.weforum.org/agenda/2016/01/8-jobs-every-company-will-be-hiring-for-by-2020/>

³ Forbes Staff. (2016). Las 10 profesiones más demandadas en 2016. Febrero,14,2016, de Forbes México Sitio web: <http://www.forbes.com.mx/las-10-profesiones-mas-demandadas-en-2016/#gs.fzevR8o>

El aporte del curso de especialización propuesto al esfuerzo de los docentes de la escuela servirá de ejemplo para otras escuelas de las diferentes carreras de la universidad, para que se preocupen por investigar la situación actual del campo de su competencia y se contraste con los conocimientos que actualmente se están impartiendo en sus programas de educación.

Los resultados de una investigación con fines educativos y propositivos pueden incluir el fortalecimiento del sistema educativo universitario actual (al incluir nuevos contenidos a la currícula y reconocer puntos de mejora en los conocimientos de los docentes), la promoción de políticas y estrategias de estado en pro de la educación superior que involucren actores gubernamentales y/o no gubernamentales, apertura de nuevas alianzas educativas que se reflejen en beneficios para los estudiantes como becas, trabajos en el exterior o intercambios entre algunos.

3. Sociedad Salvadoreña

En un contexto del crecimiento de la economía, y según el Programa de las Naciones Unidas para el Desarrollo (PNUD en El Salvador), la generación de trabajo decente sigue siendo un gran desafío. De cada cien personas que forman parte de la Población Económicamente Activa (PEA), 7 están desempleadas y 44 subempleadas (con ingresos inferiores al salario mínimo vigente en el sector económico en el que laboran). Únicamente la quinta parte de la fuerza laboral cuenta con trabajo decente y solamente el 28% de la PEA es cotizante para recibir una pensión.⁴

En septiembre de 2015, los Estados Miembros de la ONU aprobaron la Agenda 2030 para el Desarrollo Sostenible, que incluye un conjunto de 17 objetivos⁵ que deberán ser alcanzados por sus países miembros (El Salvador es uno de ellos). Dos de ellos están enfocados a la educación de calidad y al trabajo decente y crecimiento económico:

“... Para 2030, aumentar sustancialmente el número de jóvenes y adultos que tienen las competencias necesarias, en particular técnicas y profesionales, para acceder al empleo, el trabajo decente y el emprendimiento” ...

“... Lograr niveles más elevados de productividad económica mediante la diversificación, la modernización tecnológica y la innovación, entre otras cosas centrandose la atención en sectores de mayor valor añadido y uso intensivo de mano de obra” ...

⁴ PNUD. (2017). El Salvador en breve. abril 09,2017, de Programas de las Naciones Unidas para el Desarrollo Sitio web: http://www.sv.undp.org/content/el_salvador/es/home/countryinfo.html

⁵ PNUD. (2017). Agenda de desarrollo post 2015. abril 09,2017, de Programas de las Naciones Unidas para el Desarrollo Sitio web; http://www.sv.undp.org/content/el_salvador/es/home/post-2015/sdg-overview.html

Dicho esto, se espera que al tener profesionales con competencias que satisfagan las demandas laborales, se está contribuyendo al desarrollo económico del país al mover la actividad económica de otros sectores y aumentar el ingreso salarial en las familias salvadoreñas.

1.6.2. Resultados Esperados

- Perfil Académico de Pregrado que satisfaga la demanda profesional del Área de Inteligencia de Negocios.
- Perfil Académico de Pregrado que satisfaga la demanda profesional del Área de Analítica de Datos.
- Curso Académico de Inteligencia de Negocios y Analítica de Datos a partir de los perfiles investigados.
- Propuesta de Capacitación de temas de Inteligencia de Negocios y Analítica de Datos para los docentes del departamento de Desarrollo de Sistemas de la Escuela de Ingeniería de Sistemas Informáticos.

1.7 Definición tipo de investigación

El tipo de investigación se puede definir por medio de distintos criterios; uno de ellos es por el objetivo de la investigación es decir cuál es la finalidad que se busca al llevar a cabo ésta; la cual puede ser investigación básica o investigación aplicada.

INVESTIGACIÓN BÁSICA

También recibe el nombre de investigación pura, teórica, dogmática y fundamental. Se caracteriza porque parte de un marco teórico y permanece en él; la finalidad radica en formular nuevas teorías o modificar las existentes, en incrementar los conocimientos científicos o filosóficos, pero sin contrastarlos con ningún aspecto práctico.

INVESTIGACIÓN APLICADA

El proyecto o investigación busca la aplicación, el uso y generar posibles beneficios de los conocimientos resultantes de la misma, es decir, busca conocer para actuar, le interesa la aplicación sobre la realidad antes que el mero desarrollo de teorías generales, está dirigida hacia un fin u objetivo práctico determinado.⁶

⁶ Enciclopedia de clasificaciones. (2016). Tipos de investigación. Marzo, 10, 2017, de Enciclopedia de clasificaciones Sitio web: <http://www.tiposde.org/general/484-tipos-de-investigacion/>

Al analizar estos conceptos, se identifica que, debido a la finalidad de la investigación de formar un curso de especialización de pregrado a partir de la investigación de los conocimientos, habilidades y técnicas presentes en un profesional de Inteligencia de Negocios y Analítica de Big Data, se establece que la investigación es **aplicada**.

Otro aspecto importante de la clasificación de la investigación es según el enfoque de investigación que se realizó; este puede ser cuantitativo, cualitativo o mixto.

Los métodos mixtos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos y cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada (meta inferencias) y lograr un mayor entendimiento del fenómeno bajo estudio (Hernández Sampieri y Mendoza, 2008).

La definición anterior demuestra que, para ésta investigación, el hecho de necesitar datos tanto cualitativos como cuantitativos hizo que se seleccionara el método mixto; esto se fundamenta con las siguientes preguntas que ayudaron a elegir el diseño mixto:

1. ¿Qué clase de datos tienen prioridad los cuantitativos, los cualitativos o ambos por igual?

Cualitativos.

El tema de Inteligencia de Negocios y Analítica de Big Data se encuentra en una etapa de madurez a nivel internacional, y en nuestro ambiente, en proceso de crecimiento y posicionamiento en el área laboral, por tanto, prevalece toda la información disponible del avance de la Inteligencia de Negocios y Analítica de Big Data, que es resultado de estudios realizados por expertos en el área. No siendo menos importante, la experiencia actual que ha tenido el tema en nuestro ambiente. Por tanto, los datos cuantitativos resultantes de la investigación asisten en la interpretación de lo cualitativo.

2. ¿Qué resulta más apropiado para el estudio en particular: recolectar los datos cuantitativos y cualitativos de manera simultánea (concurrente) o secuencial (un tipo de datos primero y luego el otro)?

Concurrente

Se aplican ambos métodos (CUAN y CUAL) de manera simultánea, los datos cuantitativos y cualitativos se recolectan y analizan más o menos en el mismo tiempo. Desde luego, los datos cualitativos requieren de mayor tiempo para su obtención y análisis.

Los diseños concurrentes implican cuatro condiciones (Onwuegbuzie y Johnson, 2008):

- i. Se recaban en paralelo y de forma separada datos cuantitativos y cualitativos.
- ii. Ni el análisis de los datos cuantitativos ni el análisis de los datos cualitativos se construye sobre la base del otro análisis.
- iii. Los resultados de ambos tipos de análisis no son consolidados en la fase de interpretación de los datos de cada método, sino hasta que ambos conjuntos de datos han sido recolectados y analizados de manera separada se lleva a cabo la consolidación.
- iv. Después de la recolección e interpretación de los datos de los componentes CUAN y CUAL, se efectúa una o varias “meta inferencias” que integran las inferencias y conclusiones de los datos y resultados cuantitativos y cualitativos realizadas de manera independiente.

Se propone el **Diseño anidado o incrustado concurrente de modelo dominante (DIAC)**.

El diseño anidado concurrente colecta simultáneamente datos cuantitativos y cualitativos teniendo en consideración que un método predominante guía el proyecto (cualitativo). El método que posee menor prioridad es anidado o incrustado dentro del que se considera central. Tal incrustación puede significar que el método secundario responda a diferentes preguntas de investigación respecto al método primario.

Los datos recolectados por ambos métodos son comparados y/o mezclados en la fase de análisis.

Este diseño suele proporcionar una visión más amplia del fenómeno estudiado que si usáramos un solo método.

3. ¿Cuál es el propósito central de la integración de los datos cuantitativos y cualitativos? Por ejemplo: triangulación, complementación, exploración o explicación.

Complementación

Obtener una visión más comprensible sobre el planteamiento si se emplean ambos métodos.

4. ¿En qué parte del proceso, fase o nivel es más conveniente que se inicie y desarrolle la estrategia mixta? Por ejemplo: desde y/o durante el planteamiento del problema, en el diseño de investigación, recolección de los datos, análisis de los datos, interpretación de resultados o elaboración del reporte de resultados.

Desde el diseño de la investigación, ya que desde esa etapa se aprovechó las bondades del enfoque mixto pues se hizo uso de instrumentos cuantitativos (encuestas), y cualitativos (revisión de sílabos y bolsas de trabajo) para obtener datos que en la etapa de análisis ayudaron a profundizar en la construcción del perfil de conocimientos, habilidades y técnicas al proporcionar una visión más amplia del fenómeno estudiado.⁷

⁷ Hernández Sampieri, R., Fernández Collado, C. and Baptista Lucio, P. (2010). *Metodología de la investigación*. 1st ed. México: McGraw-Hill.p572

CAPÍTULO II: MARCO TEÓRICO

2.1. Antecedentes del estudio

La Inteligencia de Negocios y Analítica de Datos y el campo intrínseco a estas dos: Big Data, han ido ganando importancia y protagonismo en el área de los negocios debido a las ventajas que se obtienen de los datos y del análisis de los mismos, lo que ha hecho que las personas generen especial interés en la Inteligencia de Negocios y Analítica de datos, de ahora en adelante IN&A.

IN&A es generalmente entendida como las técnicas, tecnologías, sistemas, prácticas, metodologías y aplicaciones que analizan datos críticos de las empresas en función del mejor entendimiento del negocio y del mercado y la toma de decisiones en tiempo real.⁸

Las tecnologías y aplicaciones de IN&A que adoptaron las empresas entre la década de los 90's y el año 2000 se conoce como IN&A 1.0. En sus inicios, los datos de las industrias eran en su mayoría estructurados, los cual se obtenían de sistemas heredados y se almacenaban en bases de datos relacionales (RDB).

Se considera que la gestión de datos y data warehousing son la base de esta primera versión de IN&A:

- El diseño de data marts y las herramientas de extracción, transformación y carga (ETL) son esenciales para convertir e integrar datos,
- Consultas a la base de datos, el procesamiento analítico en línea (OLAP) y las herramientas de reportería basados en gráficos simples pero intuitivos utilizados para la exploración de datos,
- La gestión del rendimiento empresarial (Business Performance Management o BPM) a través del uso de cuadros de mandos (dashboards) y tableros(scorecards) para ayudar a visualizar el desempeño por medio de métricas.

Adicional a estas tres funciones de la IN&A 1.0 se suman los análisis estadísticos desarrollados en los años 70 y las técnicas de minería de datos desarrolladas en los 80's, a saber: análisis de asociación, segmentación de datos, análisis de clasificación y regresión y detección de anomalías. La mayoría de estas técnicas de procesamiento y análisis de datos han sido incorporados a las herramientas de Inteligencia de Negocios incluyendo Microsoft, IBM, Oracle y SAP (Sallam et al. 2011).

⁸ Chen,H., Chiang, R. & Storey, V.. (2012). Business Intelligence and Analytics: From Big Data to Big Data. En Special Issue: Business Intelligence Research (1188). Estados Unidos: Special Issue Associate Editors.

A partir del año 2000, el internet y la web empezaron a ofrecer una manera diferente de recolección datos y búsquedas analíticas:

- Los motores de búsqueda como Google y Yahoo, así como la tendencia del e-commerce como Amazon e eBay, permitieron a las empresas presentar sus negocios en línea e interactuar con sus clientes directamente.
- Los datos que se recogen a través de las cookies y logs de los servidores se convirtieron en una mina de oro para el entendimiento de las necesidades de los consumidores.

La inteligencia en la web, el análisis web y todo el contenido de los usuarios generados a través de la web 2.0 inauguró la nueva era de la IN&A 2.0, centrada en el contenido web desestructurado. Las fuentes de datos de la web 2.0, por mencionar algunos son: foros, grupos en línea, blogs, redes sociales, sitios multimedia y juegos sociales.

Con el análisis web se logra determinar aspectos como el diseño de sitios web, la optimización de ubicar un producto, analizar las transacciones de un consumidor, analizar la estructura del mercado y hacer recomendaciones de productos.

En el 2011 The Economist⁹ publicó un artículo donde se reportaba que los teléfonos móviles y las tabletas (más de 480 millones de unidades) sobrepasaban el número de laptops y PCs (cerca de 380 millones de unidades). El estudio comparó la tendencia conocida y la estimada del crecimiento de los dispositivos móviles desde el año 2005-2013, proyectando que para el 2020 la cantidad de dispositivos móviles conectados alcanzarían los 10 billones.

Los dispositivos móviles además de otros dispositivos con sensores y con conexión a internet, dispositivos con RFID(Radio Frequency Identification) y códigos de barra, son los que dan inicio la versión 3.0 de IN&A y es aquí donde el IN&A y Big Data se encuentran.

El término de Big Data se utilizó por primera vez para referirse a volúmenes de datos en aumento a mediados de los años noventa. En 2001, Doug Laney, entonces analista de la consultora Meta Group Inc., amplió la noción de grandes datos para incluir también aumentos en la variedad de datos generados por las organizaciones y la velocidad a la que se creaban y actualizaban esos datos. Estos tres factores -volumen, velocidad y variedad- se conocieron como las 3Vs del Big Data, un concepto que Gartner popularizó después de adquirir Meta Group y contratar a Laney en 2005.

⁹ Special Report on Personal Technology. (2011). Beyond PC. abril 10,2017, de The Economist Sitio web: <http://www.economist.com/node/21531109>

Posteriormente algunos autores como Mark van Rijmenam han añadido más factores para describir de una mejor manera Big Data.¹⁰

La Analítica de Datos es el proceso de examinar volúmenes crecientes y variados conjuntos de datos para descubrir patrones, correlaciones, tendencias de mercado y preferencias para ayudar a las organizaciones en la toma de decisiones basadas en información de valor.¹¹

El análisis de Big Data incluye datos estructurados, además de otras formas de datos que a menudo se dejan sin explotar por los programas convencionales de IN&A. Esto incluye una mezcla de datos semi estructurados y no estructurados, por ejemplo: datos de clics de Internet, registros de servidores web, contenido de redes sociales, texto de correos electrónicos de clientes y respuestas de encuestas, registros de llamadas de teléfonos móviles y datos de máquinas capturados por sensores conectados (Internet de las cosas).

Los tipos de datos no estructurados y semi estructurados no encajan bien en los almacenes de datos tradicionales que se basan en bases de datos relacionales orientadas a conjuntos de datos estructurados. Además, estos pueden no ser capaces de manejar las demandas de procesamiento que plantean los conjuntos de Big Data que necesitan actualizarse con frecuencia, o incluso continuamente. Como resultado, muchas organizaciones que recopilan, procesan y analizan Big Data recurren a Hadoop y sus herramientas complementarias, como YARN, MapReduce, Spark, HBase, Hive, Kafka y Pig, así como bases de datos NoSQL.

Antecedentes de Inteligencia de Negocios y Analítica de Datos en El Salvador.

Educación

- El plan de estudios 2009 de la carrera de Ingeniería en Ciencias de la Computación de la Universidad Don Bosco contempla la materia “Inteligencia de Negocios” como una opción electiva técnica.¹²
- En el 2010, estudiantes de la Universidad de El Salvador realizaron el primer trabajo de graduación sobre Inteligencia de Negocios en la Escuela de Ingeniería de Sistemas Informáticos denominado “Desarrollo de un Data Warehouse para el proceso de Denuncias de la Defensoría del Consumidor”.

¹⁰ Why The 3V's Are Not Sufficient To Describe Big Data. van Rijmenam, M. (2018). junio 1,2018, de Datafloq Sitio web: <https://datafloq.com/read/3vs-sufficient-describe-big-data/166>

¹¹ Rouse, M., (2017). Big Data Analytics. abril 10,2017, de Tech Target Sitio web: <http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics>

¹² Universidad Don Bosco. (2014). Ingeniería en Ciencias de la Computación Plan 2009. abril 11,2017, de Universidad Don Bosco Sitio web: http://www.udb.edu.sv/udb/archivo/pensum/ing_compu.pdf

- En el 2012, estudiantes de la Universidad Matías Delgado llevaron a cabo el seminario de trabajo de investigación “Uso de la Inteligencia de Negocios en las PYMES en El Salvador”.
- La Universidad Tecnológica ofrece desde el año 2013 la pre especialización de graduación “Desarrollo de soluciones en Inteligencia de Negocios para entornos empresariales” para sus estudiantes de pregrado.¹³
- El 26 de julio de 2014 la Escuela Europea de Negocios impartió un seminario gratuito en San Salvador abordando el tema “Las nuevas tecnologías en los modelos de negocio (BIG DATA)”.¹⁴
- Fepade ofreció un postgrado internacional en Inteligencia de Negocios en el año 2015.¹⁵
- La Universidad Francisco Gavidia posee desde el 2015 una especialización de graduación en Administración y Desarrollo de Base de datos e Inteligencia de Negocios.¹⁶
- En el 2016, estudiantes de la Universidad de El Salvador llevaron a cabo el trabajo de graduación “Business Intelligence como soporte de las decisiones estratégicas, tácticas y operacionales de las organizaciones”.¹⁷
- En el 2016, la Universidad Tecnológica publicó en su revista Entorno los resultados de su investigación “Big Data, análisis de datos en la nube” con el objetivo de conocer sobre el auge de Big Data, y cómo se podrían utilizar algunas herramientas para el procesamiento, análisis y visualización de los datos.¹⁸

¹³ Universidad Tecnológica de El Salvador. (2013). Programación Preespecialidad. Marzo 14, 2017, de Universidad Tecnológica de El Salvador Sitio web: <http://graduadoutec.edu.sv/banner/Programacion-Preespecialidad-Prueba01-2013.pdf>

¹⁴ Escuela Europea de Negocios. (2014). Las nuevas tecnologías en los modelos de negocio (BIG DATA). abril 11, 2017, de Escuela Europea de Negocios Sitio web: <http://www.escolaeuropea.eu/SV/san-salvador-1/las-nuevas-tecnologias-en-los-modelos-de-negocio-big-data/evento/es>

¹⁵ Isdae Fepade. (2015). Postgrado Internacional en Business Intelligence - Septiembre 2015. Marzo 14, 2017, de Fepade Sitio web:

<https://www.iseade.edu.sv/index.php/postgrados-especializados/30-postgrados-especializados/historico-de-postgrados/259-postgrado-internacional-en-business-intelligence-septiembre-2015>

¹⁶ Universidad Francisco Gavidia. (2015). Curso de Especialización Administración y Desarrollo de Base de datos e Inteligencia de Negocios. Marzo 14, 2017, de Universidad Francisco Gavidia Sitio web:

<http://registro.ufg.edu.sv/InformacionEgresados/Especializacion.aspx>

¹⁷ Argueta Torres, Claudia Ivonne y Maldonado, Cristina Abigail (2016) Business Intelligence como soporte de las decisiones estratégicas, tácticas y operacionales de las organizaciones. Maestría thesis, Universidad de El Salvador.

¹⁸ Rosa, V. & Rivera, J.. (2016, junio 01). Big Data, análisis de datos en la nube. Revista entorno, Universidad Tecnológica de El Salvador, 61, 24.

- El 29 de marzo de 2017 la Universidad Matías Delgado abordó el tema de Big Data en la actualidad y los Millennials en el área financiera, como parte de su evento Global Money Week 2017.¹⁹
- La Universidad Don Bosco ha ofrecido a estudiantes y profesionales, por medio de su Centro de Tecnología de la información y las telecomunicaciones CTIC las siguientes conferencias:
 - Conferencia: "Analítica de Datos", 30 de agosto de 2016.²⁰
 - Taller Internacional de Big Data:" Procesamiento de Big Data a gran escala usando Spark", 22 y 23 de septiembre de 2016. ²¹
 - Conferencia Internacional Big Data,14 de marzo de 2017. ²²
 - Taller Internacional: Analítica y Gestión de Big Data, 16 y 17 de marzo de 2017.²³
- La Escuela Superior de Economía y Negocios ESEN ofreció a partir de mayo 2017 culminando en el año 2018 los siguientes programas:
 - Postgrado en Analítica para Ciencia de Datos.
 - Certificado superior en Minería de Texto y Analítica de Datos.
 - Postgrado en Estadística para Ciencias Sociales.

¹⁹Reyes, G., (2017). Segundo día del Global Money Week 2017. abril 11,2017, de Universidad Dr. José Matías Delgado Sitio web: <http://www.ujmd.edu.sv/contenidoslideshow/29-noticias/1354-segundo-dia-del-global-money-week-2017>

²⁰ CTIC-UDB. (2016, agosto, 18).La Universidad Don Bosco, a través del CTIC (Centro de Tecnología de la Información y las Comunicaciones) le invitan a la Conferencia: "Analítica de Big Data". Recuperado de: <https://www.facebook.com/cticudb/photos/a.104339129684803.4866.104214889697227/1004593346326039/?type=3&theater>

²¹CTIC-UDB. (2016, septiembre, 1). La Universidad Don Bosco a través del CTIC le invita a su 3er Taller Internacional de Big Data:" Procesamiento de Big Data a gran escala usando Spark" impartido por el especialista Magister Willis Polanco de República Dominicana. Recuperado de: <https://www.facebook.com/cticudb/photos/a.104339129684803.4866.104214889697227/1016800408438666/?type=3&theater>

²² CTIC-UDB. (2017b, marzo, 3).Quieres saber más acerca de BIG DATA? Te invitamos a participar de nuestra Conferencia Internacional, entrada gratis, entérate de toda la información necesaria, danos like y comparte esta publicación. Recuperado de: <https://www.facebook.com/cticudb/photos/a.104339129684803.4866.104214889697227/1188136064638432/?type=3&theater>

²³ CTIC-UDB. (2017a, marzo, 3).CTIC le invita a participar en el Taller Internacional: "Analítica y Gestión de Big Data" impartido por el especialista Magister Willis Polanco de República Dominicana. Recuperado de: <https://www.facebook.com/cticudb/photos/a.104339129684803.4866.104214889697227/1188022041316501/?type=3&theater>

2.2. Bases teóricas

2.2.1. Inteligencia de Negocios

a) Definición

1. Inteligencia de Negocios puede ser definida como un conjunto de modelos matemáticos y metodologías de análisis que explotan los datos disponibles para generar información y el conocimiento útil para procesos de toma de decisiones complejos.²⁴
2. Proceso de tomar grandes cantidades de datos, analizarlos y presentar un conjunto de reportes de alto nivel que condensan la esencia de esos datos en las bases de las actividades de los negocios, permitiendo a la gerencia tomar decisiones fundamentales de negocio diarias.²⁵
3. Conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización.²⁶

Áreas o categorías en las que se aplica Inteligencia de Negocios

1. AQL - Lógica de consulta asociativa.
2. Scorecarding.
3. Gestión del rendimiento empresarial y Medición del desempeño.
4. Planificación empresarial.
5. Reingeniería de procesos de negocio.
6. Análisis competitivo.
7. Gestión de relaciones con los clientes (CRM) y Marketing.
8. Sistemas de Apoyo a la Decisión (DSS) y Predicción.
9. Almacenes de documentos y administración de documentos.
10. Sistemas de gestión empresarial.

²⁴ Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. Reino Unido: Wiley.

²⁵ Rajan, J.. (2005). *Business Intelligence: Concepts, Components, Techniques And Benefits*. Journal of Theoretical and Applied Information Technology, Vol. 9, pp 60-70.

²⁶ Curto, J. & Conesa, J. . (2010). Capítulo I: Introducción al Business Intelligence. En *Introducción al Business Intelligence* (p. 18). Barcelona: UOC.

11. Sistemas de Información Ejecutiva (EIS).
12. Finanzas y Presupuesto.
13. Recursos humanos.
14. Gestión del Conocimiento.
15. Mapeo, visualización de información y Tableros (Dashboards).
16. Sistemas de Información Gerencial (MIS).
17. Sistemas de Información Geográfica (SIG).
18. Inteligencia de Negocios en tiempo real.
19. Estadísticas y análisis de datos técnicos.
20. Gestión de la Cadena de Suministro / Gestión de la cadena de Demandas.
21. Inteligencia de sistemas.
22. Análisis de tendencia.
23. Consulta del usuario / usuario final y generación de informes.
24. Personalización Web y Minería Web.
25. Extracción de textos.

b) Componentes de la Inteligencia de Negocios

- **Data Warehouse:** es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización –independientemente de cómo se vayan a utilizar posteriormente por los consumidores o usuarios–, con las propiedades siguientes: estable, coherente, fiable y con información histórica. Al abarcar un ámbito global de la organización y con un amplio alcance histórico, el volumen de datos puede ser muy grande (centenas de terabytes). Las bases de datos relacionales son el soporte técnico más comúnmente usado para almacenar las estructuras de estos datos y sus grandes volúmenes.

Resumiendo, presenta las siguientes características:

- Temático: los datos están vinculados al negocio y organizados por las diferentes funciones que lleva a cabo el mismo.
- Integrado: incluye datos de múltiples orígenes y presenta consistencia de datos.
- Variable en el tiempo: se realizan fotos de los datos basadas en fechas o hechos.
- No volátil: sólo de lectura para los usuarios finales.

- **Data Warehousing:** es el proceso de extraer y filtrar datos de las operaciones comunes de la organización, procedentes de los distintos sistemas de información operacionales y/o sistemas externos, para transformarlos, integrarlos y almacenarlos en un almacén de datos con el fin de acceder a ellos para dar soporte en el proceso de toma de decisiones de una organización.
- **Data Mart:** es un subconjunto de los datos del Data Warehouse cuyo objetivo es responder a un determinado análisis, función o necesidad, con una población de usuarios específica. Al igual que en un Data Warehouse, los datos están estructurados en modelos de estrella o copo de nieve. Un data mart puede ser dependiente o independiente de un Data Warehouse. Está pensado para cubrir las necesidades de un determinado departamento dentro de la organización.
- **Operational Data Store:** es un tipo de almacén de datos que proporciona sólo los últimos valores de los datos y no su historial; además, generalmente admite un pequeño desfase o retraso sobre los datos operacionales.
- **Staging Area:** es el sistema que permanece entre las fuentes de datos y el Data Warehouse con el objetivo de:
 - Facilitar la extracción de datos desde fuentes de origen con una gran variedad y complejidad grande.
 - Mejorar la calidad de datos.
 - Ser usado como caché de datos operacionales con el que posteriormente se realiza el proceso de data warehousing.
 - Para acceder en detalle a información no contenida en el Data Warehouse.
- **Procesos ETL:** tecnología de integración de datos basada en la consolidación de datos que se usa tradicionalmente para alimentar Data Warehouse, data mart, staging área y Almacén Operacional de datos (ODS). Usualmente se combina con otras técnicas de consolidación de datos.
- **Metadatos:** Son datos acerca de los datos, para un Data Warehouse los metadatos almacenan la ruta del mapa del almacén de datos, funcionan como un directorio para localizar los objetos y el contenido que conforma el Data Warehouse.

c) Modelado de Data Warehouse

Es el proceso para la creación de un Data Warehouse, existen varios enfoques para implementarlos:

Enfoque de Inmon

Este enfoque es fundado por Bill Inmon en los años 90 para satisfacer los requisitos de las empresas y permitirles desarrollar sistemas para la toma de decisiones. Permite almacenar todos los eventos de una empresa; este enfoque se basa en los diagramas entidad-relación de los sistemas transaccionales. Los datos de la empresa se cargan sin conocer a priori los requisitos del usuario; esta es la razón por la que se llama enfoque basado en datos.

La arquitectura del almacén de datos Inmon (Ver Figura 2) incluye todos los sistemas de información de la empresa en lugar de considerar sólo fragmentos de información.

Inmon divide el entorno de las bases de datos de la empresa en cuatro niveles, incluyendo: operacional, atómico (Data Warehouse), departamentales (data marts) y niveles individuales.

Los tres últimos niveles se refieren al Data Warehouse mientras el primer nivel (operacional) apoya las operaciones diarias y contiene los datos transaccionales de la empresa. Estos datos se transforman y cargan en el atómico (Data Warehouse) utilizando ETL (Extract, Transform and Load).

De acuerdo con los requerimientos de los departamentos de la empresa, los datos del nivel departamental están orientados a un tema, siempre consistentes porque provienen de un mismo Data Warehouse. Inmon considera que una Data Warehouse y un data mart están físicamente separados.

El almacén de datos tiene su propia existencia física y se orienta al almacenamiento, trazabilidad y escalabilidad en respuesta a nuevos requerimientos. Además, los data marts tienen su propia existencia física y ofrecen restitución, estructura orientada al rendimiento en respuesta a los requerimientos del usuario.

Los usuarios crean el último nivel cuando analizan y explotan los datos cargados en los data marts (OLAP Analysis, Informes, Dashboards, etc.).

Inmon propone una arquitectura denominada Corporate Information Factory (o Enterprise Data Warehouse), donde existe un Data Warehouse corporativo y unos data marts (o incluso cubos OLAP) dependientes del mismo. El acceso a datos se realiza a los data marts o al ODS en caso de existir, pero nunca al propio Data Warehouse. Puede existir en el caso de ser necesaria una staging área.

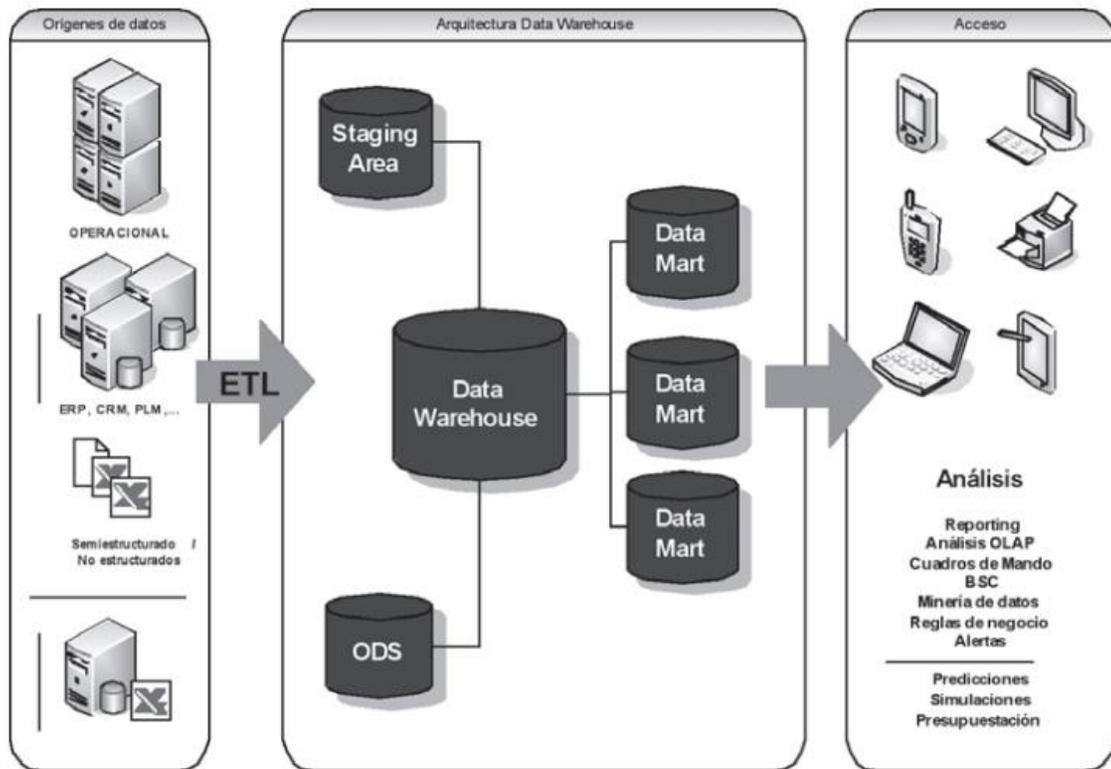


Figura 2 Enfoque de Inmon. Corporate Information Factory. Introducción al business Intelligence (p.40), por Josep Curto Díaz y Jordi Conesa Caralt, 2010, Barcelona: UOC, Copyright [2010] por Josep Curto Díaz y Jordi Conesa i Caral. Reimpresión autorizada

Enfoque de Kimball

Kimball creó su enfoque en los años 90 ofreciendo una nueva arquitectura, una nueva visión y un modelo innovador de Data Warehouse.

Este enfoque se basa en el concepto de modelado dimensional. Se opone al principio de aislamiento de los usuarios finales enunciado por Inmon. De hecho, en esta propuesta se involucra a los usuarios finales en las primeras etapas del proyecto, por eso se le denomina enfoque basado en los requerimientos del usuario. Se presenta una visión diferente de los Data Warehouse pues considera que estos pueden ser visto como un conjunto de data marts de datos consistentes y basados en dimensiones compartidas.

Kimball propone una arquitectura denominada Enterprise Bus Architecture (o Data Warehouse Virtual/Federado) la cual permite un desarrollo incremental del Data Warehouse y de Inteligencia de Negocios (DW / BI). Descompone el proceso de planificación DW / BI en partes manejables, centrándose en los procesos de negocio de la organización, junto con las dimensiones conformadas asociadas.

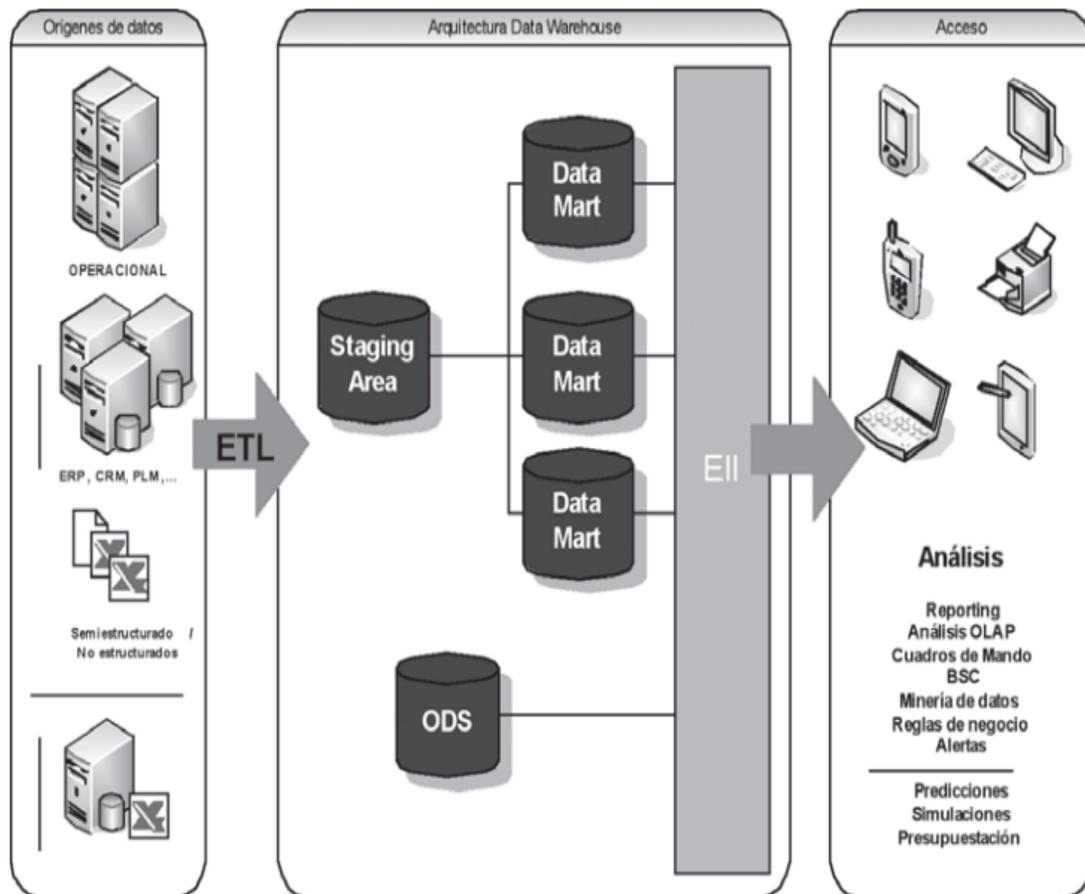


Figura 3 Enfoque de Kimball. *Enterprise bus. Introducción al business Intelligence* (p.39), por Curto, J. & Conesa, J., 2010, Barcelona: UOC, Copyright [2010] por Josep Curto Díaz y Jordi Conesa i Caral. Reimpresión autorizada

Esquemas

Sirve para estructurar los datos en el Data Warehouse; este proceso se realiza mediante un esquema (en inglés, Database Schema) el cual describe la estructura de una Base de datos, en un lenguaje formal soportado por un Sistema Administrador de Base de datos (DBMS). En una Base de datos Relacional, el esquema define sus tablas, sus campos en cada tabla y las relaciones entre cada campo y cada tabla. En la práctica el término esquema de la base de datos se refiere al diseño físico de la base de datos.

Existen tres tipos de esquemas:

- **Esquema en estrella:** consiste en estructurar la información en procesos, vistas y métricas recordando a una estrella (por ello el nombre). A nivel de diseño, consiste en una tabla de hechos (fact table) en el centro para el hecho objeto de análisis y una o varias tablas de dimensión por cada punto de vista de análisis que participa de la descripción de ese hecho.

En la tabla de hecho encontramos los atributos destinados a medir (cuantificar): sus métricas. La tabla de hechos sólo presenta uniones con dimensiones.

- **Esquema en copo de nieve:** es un esquema de representación derivado del esquema en estrella, en el que las tablas de dimensión se normalizan en múltiples tablas. Por esta razón, la tabla de hechos deja de ser la única tabla del esquema que se relaciona con otras tablas, y aparecen nuevas uniones. Es posible distinguir dos tipos de esquemas en copo de nieve:
 - Completo: en el que todas las tablas de dimensión en el esquema en estrella aparecen ahora normalizadas.
 - Parcial: sólo se lleva a cabo la normalización de algunas de ellas.

Una tabla de hecho 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 modelar.
- Claves foráneas hacia registros en una tabla de dimensión (o hacia una vista de negocio).

Las dimensiones recogen los puntos de análisis de un hecho. Por ejemplo, una venta se puede analizar en función del día de venta, producto, cliente, vendedor o canal de venta, entre otros.

- **Esquema Constelación:** este modelo está compuesto por una serie de esquemas en estrella, está formado por una tabla de hechos principal y por una o más tablas de hechos auxiliares, las cuales pueden ser sumalizaciones de la principal. Dichas tablas yacen en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones.

No es necesario que las diferentes tablas de hechos compartan las mismas tablas de dimensiones, ya que, las tablas de hechos auxiliares pueden vincularse con solo algunas de las tablas de dimensiones asignadas a la tabla de hechos principal, y también pueden hacerlo con nuevas tablas de dimensiones.

d) Elementos de un Data Warehouse

La idea principal es que la información sea presentada desnormalizada para optimizar las consultas. Para ello se debe identificar, en el núcleo de organización, los procesos de negocio, las vistas para el proceso de negocio y las medidas cuantificables asociadas a los mismos. De esta manera se habla de:

- **Tabla de hecho:** es la representación en el Data Warehouse de los procesos de negocio de la organización. Por ejemplo, una venta puede identificarse como un proceso de negocio de manera que es factible, si corresponde en la organización, considerar la tabla de hecho ventas.
- **Dimensión:** es la representación en el Data Warehouse de una vista para un cierto proceso de negocio.
- **Métrica (Medida):** son los indicadores de negocio de un proceso del mismo. Aquellos conceptos cuantificables que permiten medir el proceso de negocio. Por ejemplo, en una venta se tiene el importe de la misma.

e) Proceso Analítico en Línea (OLAP)

El concepto de Inteligencia de Negocios engloba múltiples componentes. Uno de los más importantes es el concepto OLAP que es una tecnología que permite un análisis multidimensional²⁷ a través de tablas matriciales o pivotantes.

Se entiende por OLAP, al método ágil y flexible para organizar datos, especialmente metadatos, sobre un objeto o jerarquía de objetos como en un sistema u organización multidimensional, y cuyo objetivo es recuperar y manipular datos y combinaciones de los mismos a través de consultas o incluso informes.

OLAP forma parte de lo que se conoce como sistemas analíticos, que permiten responder preguntas como: ¿Por qué paso? Estos sistemas pueden encontrarse tanto integrados en suites de Inteligencia de Negocios o ser simplemente una aplicación independiente.

Una herramienta OLAP está formada por un motor y un visor. El motor es, en realidad, justo el concepto que acabamos de definir. El visor OLAP es una interfaz que permite consultar, manipular, reordenar y filtrar datos existentes en una estructura OLAP mediante una interfaz gráfica de usuario que dispone funciones de consulta MDX (Expresiones Multidimensionales) y otras. Las estructuras OLAP permiten realizar consultas que serían sumamente complejas mediante SQL.

²⁷ En 1962, se introduce el análisis multidimensional en el libro de Ken Iverson A Programming Language.

Tipos de OLAP

Existen diferentes tipos de OLAP, que principalmente difieren en cómo se guardan los datos:

- **MOLAP (OLAP Multidimensional):** es la forma clásica de OLAP y frecuentemente es referida con dicho acrónimo. MOLAP utiliza estructuras de bases de datos generalmente optimizadas para la recuperación de los mismos. Es lo que se conoce como bases de datos multidimensionales (o, más coloquialmente, cubos). En definitiva, se crea un fichero que contiene todas las posibles consultas pre calculadas. A diferencia de las bases de datos relacionales, éstas formas de almacenaje están optimizadas para la velocidad de cálculo. También se optimizan a menudo para la recuperación a lo largo de patrones jerárquicos de acceso. Las dimensiones de cada cubo son típicamente atributos tales como período, localización, producto o código de la cuenta. La forma en la que cada dimensión será agregada se define por adelantado.
- **ROLAP (OLAP Relacional):** trabaja directamente con las bases de datos relacionales, que almacenan los datos base y las tablas dimensionales como tablas relacionales mientras se crean nuevas tablas para guardar la información agregada.
- **HOLAP (OLAP Híbrido):** no hay acuerdo claro en la industria en cuanto a qué constituye el OLAP híbrido, exceptuando el hecho de que es una base de datos en la que los datos se dividen en almacenaje relacional y multidimensional. Por ejemplo, para algunos vendedores, HOLAP consiste en utilizar las tablas relacionales para guardar las cantidades más grandes de datos detallados, y utiliza el almacenaje multidimensional para algunos aspectos de cantidades más pequeñas de datos menos detallados o agregados.
- **DOLAP (OLAP de Escritorio):** es un caso particular de OLAP ya que está orientado a equipos de escritorio. Consiste en obtener la información necesaria desde la base de datos relacional y guardarla en el escritorio. Las consultas y los análisis son realizados contra los datos guardados en el escritorio.
- **In-memory OLAP (OLAP en memoria):** es un enfoque por el que muchos nuevos fabricantes están optando. Consiste en que la estructura dimensional se genera sólo a nivel de memoria y se guarda el dato original en algún formato que potencia su despliegue de esta forma (por ejemplo, comprimido o mediante una base de datos de lógica asociativa).

Reglas para evaluar soluciones OLAP de E. F. Codd

- **Vista conceptual multidimensional:** se trabaja a partir de métricas de negocio y sus dimensiones, el negocio debe verse por naturaleza de forma multidimensional.
- **Transparencia:** el sistema OLAP debe formar parte de un sistema abierto que soporta fuentes de datos de varios orígenes (como se conoce actualmente arquitectura orientada a servicios).
- **Accesibilidad:** se debe presentar el servicio OLAP al usuario con un único esquema lógico de datos (lo que, en definitiva, indica que debe presentarse respecto una capa de abstracción directa con el modelo de negocio).
- **Rendimiento de informes consistente:** el rendimiento de los informes no debería degradarse cuando el número de dimensiones del modelo se incrementa.
- **Arquitectura cliente/servidor:** basado en sistemas modulares y abiertos que permitan la interacción y la colaboración.
- **Dimensionalidad genérica:** capacidad de crear todo tipo de dimensiones y con funcionalidades aplicables de una dimensión a otra.
- **Tratamiento sistemático de valores nulos:** la manipulación de datos en los sistemas OLAP debe poder diferenciar valores vacíos de valores nulos y además poder ignorar las celdas sin datos.
- **Operaciones cruzadas entre dimensiones sin restricciones:** todas las dimensiones son creadas igual y las operaciones entre dimensiones no deben restringir las relaciones entre celdas.
- **Manipulación de datos intuitiva:** dado que los usuarios a los que se destinan este tipo de sistemas son frecuentemente analistas y altos ejecutivos, la interacción debe considerarse desde el prisma de la máxima usabilidad de los usuarios.
- **Elaboración de reportes flexibles:** los usuarios deben ser capaces de manipular los resultados que se ajusten a sus necesidades conformando informes. Además, los cambios en el modelo de datos deben reflejarse automáticamente en esos informes.

- **Niveles de dimensiones y de agregación ilimitados:** no deben existir restricciones para construir cubos OLAP con dimensiones y con niveles de agregación ilimitados.

f) Extracción, Transformación y Carga (ETL)

Los procesos de ETL están íntimamente ligados con la creación de Data Warehouse. Para que los datos contenidos en un Data Warehouse sean precisos y tengan relevancia, tienen que pasar por un proceso que permita ser transformado de acuerdo a como la estructura donde se almacenarán lo requieran. El objetivo principal de los procesos ETL es facilitar la migración y la transformación de los datos.

Los procesos ETL se pueden definir como la tecnología de integración de datos basada en la consolidación de datos que se usa tradicionalmente para alimentar un Data Warehouse, data mart, staging área y ODS. Usualmente se combina con otras técnicas de consolidación de datos.

Los procesos ETL son una colección de aplicaciones o programas que aportan valor añadido a los datos. Esto va más allá de solo depurar los datos que serán almacenados en un Data Warehouse. Los sistemas ETL realizan varias tareas:

- Remueven datos incorrectos o corrigen datos faltantes.
- Proveen de indicadores de confiabilidad en los datos.
- Capturan el flujo de los datos transaccionales de manera segura.
- Ajustan los datos provenientes de múltiples fuentes para ser usados de manera conjunta.
- Reestructuran los datos para ajustarse a las herramientas usadas por el usuario final.

Técnicas de modelado

La construcción del ETL puede dividirse en tres subprocesos o componentes: componente de extracción, componente de transformación y componente de carga.

La tarea de diseñar procesos ETL involucra: (1) analizar las fuentes de datos existentes para encontrar la semántica oculta en ellas y (2) diseñar el flujo de trabajo que extraiga los datos desde las fuentes, repare sus inconsistencias, los transforme en un formato deseado, y, finalmente, los inserte en la Data Warehouse. Con el propósito de facilitar la tarea de modelar los procesos ETL, se han desarrollado diferentes técnicas, dos categorías que sobresalen son las siguientes.

Enfoque orientado al modelado de procesos.

Con este enfoque existen dos tipos de modelado: el conceptual y el lógico. Para el modelado conceptual se plantea identificar cuáles son las fuentes, los atributos y las transformaciones que se necesitan realizar.

Para el modelado lógico se plantea considerar además otros elementos como el conjunto de valores que el atributo admite, cuales son los parámetros necesarios para que las funciones se realicen de manera correcta.

Enfoque orientado a objetos.

En este enfoque se plantea usar diagramas de UML. Al emplear la adopción del diagrama de clases de UML, se logra cubrir las especificaciones de conceptos como clases, atributos y métodos. Además de ello se logra acortar la curva de aprendizaje al ser una buena práctica del desarrollo de software, al asimilar una nueva tecnología.

Técnicas de integración de datos

Las herramientas de ETL han sido la opción usual para alimentar el Data Warehouse. La funcionalidad básica de estas herramientas consiste básicamente en la extracción, transformación y carga de datos.

La integración de datos se conoce como al conjunto de aplicaciones, productos, técnicas y tecnologías que permiten tener una visión única y consistente de los datos del negocio. Existen diferentes tecnologías basadas en estas técnicas.

- **ETL de generación de código:** constan de un entorno gráfico en donde se diseñan y se especifican los datos de origen, sus transformaciones y sus entornos de destino. El resultado es un programa de tercera generación, usualmente basado en Cobol que permite realizar las transformaciones de datos.
- **ETL basados en motor:** permite crear flujos de trabajo en tiempo de ejecución definidos mediante herramientas gráficas. El entorno gráfico permite hacer un mapeo de los entornos de datos de origen y destino, las transformaciones de datos necesarias, el flujo de procesos y los procesos por lotes necesarios. Los motores generalmente usados para realizar estos procesos se pueden componer por diversos adaptadores como ODBC, JDBC, JNDI, etc.
- **ETL integrado en la base de datos:** algunos fabricantes incluyen capacidades de ETL dentro del motor de la base de datos. En general, presentan menos funcionalidades y complejidad, y son solución menos completa que los ETL comerciales basados en motor o generación de código.

g) Diseño de reportes / Visualizaciones

La intención de los reportes es que el usuario del negocio pueda recibir información pertinente que pueda contribuir a la toma de decisiones. Existen diferentes tipos de informe en función de la interacción ofrecida al usuario.

- **Estáticos:** tienen un formato preestablecido inamovible.
- **Paramétricos:** presentan parámetros de entrada y permiten múltiples consultas.
- **Ad-hoc:** son creados por el usuario final a partir de la capa de metadatos que permite usar el lenguaje de negocio propio.

Elementos de un informe

Es necesario destacar cuáles son los componentes que tiene un informe, puede estar conformado por:

- **Texto:** que describe el estado del proceso de negocio o proporciona las descripciones necesarias para entender el resto de elementos del informe.
- **Tablas:** este elemento tiene forma de matriz y permite presentar una gran cantidad de información.
- **Gráficos:** este elemento persigue el objetivo de mostrar información con un alto impacto visual que sirva para obtener información agregada o resumida con mucha más rapidez que a través de tablas.
- **Mapas:** este elemento permite mostrar información geolocalizada.
- **Métricas:** que permiten conocer cuantitativamente el estado de un proceso de negocio.
- **Alertas visuales y automáticas:** consiste en avisos del cambio de estado de información que pueden estar formadas por elementos gráficos como fechas o colores, resultados y que deben estar automatizadas en función de reglas de negocio encapsuladas en el cuadro de mando.

Tipos de métricas

Los informes poseen métricas que son propias del negocio, existen diferentes tipos de medida que se clasifican según el tipo de información que se requiere.

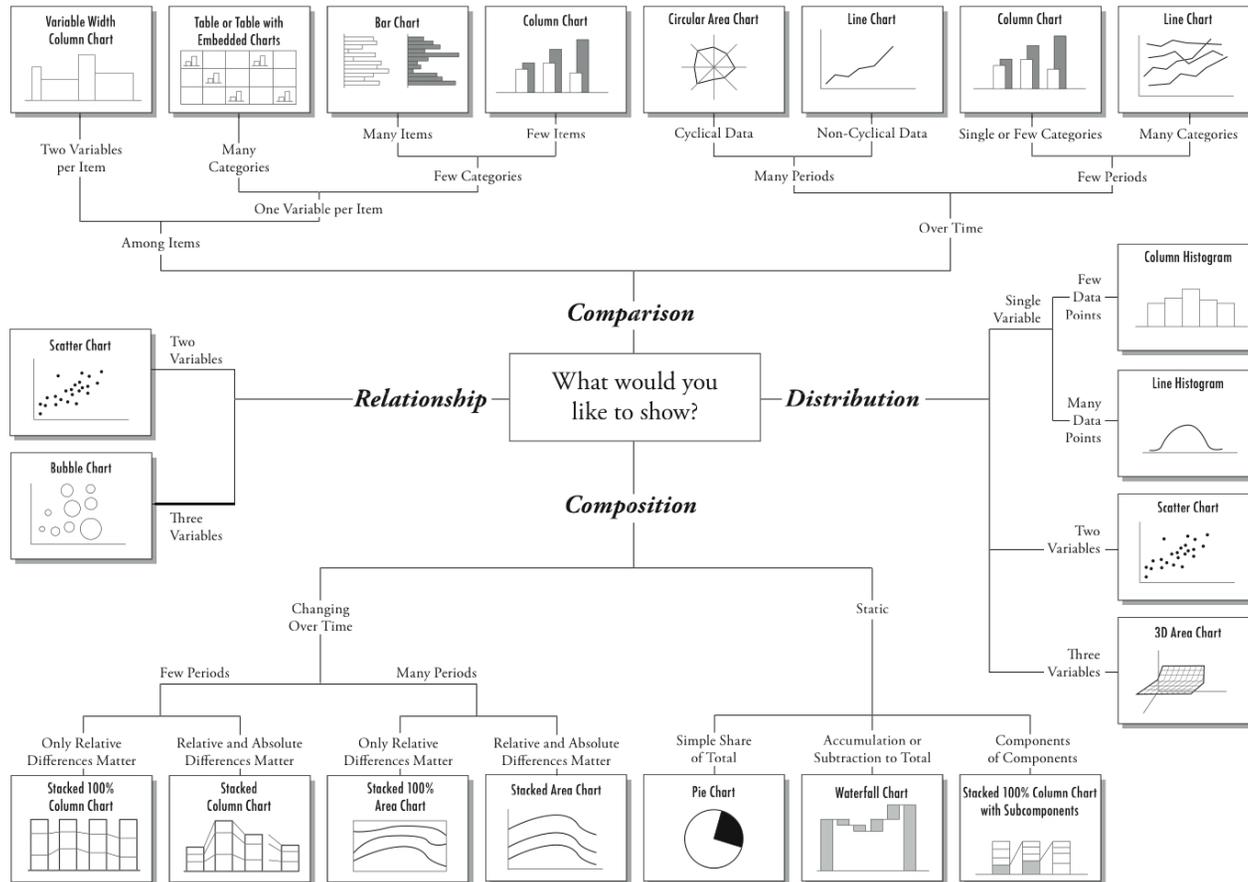
- **Métricas:** valores que recogen el proceso de una actividad o los resultados de la misma. Estas medidas proceden del resultado de la actividad de negocio.
 - **Métricas de realización de actividad (leading):** miden la realización de una actividad. Por ejemplo, la participación de una persona en un evento.
 - **Métricas de resultado de una actividad (lagging):** recogen los resultados de una actividad. Por ejemplo, la cantidad de puntos de un jugador en un partido.

- **Indicadores clave:** son valores correspondientes que hay que alcanzar, y que suponen el grado de asunción de los objetivos. Estas medidas proporcionan información sobre el rendimiento de una actividad o sobre la consecución de una meta.
 - **Indicadores claves de rendimiento (KPI):** se definen valores que explican en qué rango óptimo de rendimiento se deberían situar para alcanzar objetivos. Son métricas del proceso.
 - **Indicadores de meta (KGI):** definen mediciones para informar a la dirección general si un proceso TIC ha alcanzado sus requisitos de negocio, y se expresan por lo general en términos de criterios de información.

Tipos de gráficos

Existen diferentes tipos de gráficos que ayudan a una mejor visualización y dependiendo de ciertos criterios es mejor utilizar un gráfico por sobre otro, el siguiente diagrama reconoce los criterios necesarios para la elección de un gráfico adecuado de acuerdo a la información que se desea presentar.

Chart Suggestions—A Thought-Starter



www.ExtremePresentation.com
 © 2009 A. Abela — a.v.abela@gmail.com

Figura 4 Tipos de gráficos. Chart Suggestions - A Thought-Starter. En "Extreme presentation tools", por A. Abela, 2009, <http://extremepresentation.com>. Copyright [2009]. The Extreme Presentation Method. Reimpresión autorizada.

2.2.2. Analítica de Datos

Hasta este punto ya se ha hablado sobre Inteligencia de Negocios, pero para poder abordar el tema de Analítica de Datos y su enfoque en Big Data, es necesario abordar la diferencia entre estos tres tópicos.

La Analítica de Datos es el medio para llegar a la Inteligencia de Negocios y Big Data. Para hacer Inteligencia de Negocios nos basamos en el análisis de los datos, para hacer Analítica de Big Data nos basamos igualmente del mismo proceso. Entonces, el Análisis o la Analítica de Datos es un término general para cualquier tipo de procesamiento, que implica la recopilación, integración y preparación de los datos (perfilado y limpieza de datos). Además, si el análisis es más avanzado se habla del desarrollo, prueba y revisión de modelos analíticos para asegurar que se produzcan resultados precisos.

La Analítica de Datos tiene diferentes metodologías que corresponden a diferentes niveles de aplicación. A groso modo, las metodologías de análisis de datos incluyen el *análisis exploratorio de datos (EDA)*, cuyo objetivo es encontrar patrones y relaciones en los datos y *el análisis de datos confirmatorios (CDA)*, que aplica técnicas estadísticas para determinar si las hipótesis sobre un conjunto de datos son verdaderas o falsas.

El análisis de datos también puede separarse en *análisis de datos cuantitativos* y *análisis de datos cualitativos*. La primera implica el análisis de datos numéricos con variables cuantificables que pueden ser comparadas o medidas estadísticamente. El enfoque cualitativo es más interpretativo - se centra en la comprensión del contenido de datos no numéricos como texto, imágenes, audio y video, incluyendo frases comunes, temas y puntos de vista.

Los tipos de análisis de datos más avanzados incluyen la *minería de datos*, que implica clasificar grandes conjuntos de datos para identificar tendencias, patrones y relaciones; *analítica predictiva*, que busca predecir el comportamiento del cliente, fallas de equipos y otros eventos futuros; y *machine learning*, una técnica de inteligencia artificial que utiliza algoritmos automatizados para manipular conjuntos de datos más rápidamente de lo que puede hacerse a través del modelado analítico convencional.

Análisis Exploratorio de Datos

Es un modo de análisis de los datos que utiliza el resumen numérico y visual para explorar datos en busca de patrones no anticipados. En general, el análisis exploratorio de datos se caracteriza por el uso de herramientas o técnicas con mucha carga visual o gráfica, con énfasis en revelar información vital sobre la data examinada. El arsenal correspondiente está compuesto, entre otros, por instrumentos como:

- Diagrama de caja y bigotes (Box and whisker): un resumen visual de la distribución (comportamiento) de una variable que provee detalles acerca de si uno o ambos extremos de la distribución contienen valores inusualmente grandes o pequeños.
- Diagrama de tallo y hojas (Stem and leaf): es una muestra visual de la distribución de una variable. Se asemeja a un histograma “y” se usa mucho para mostrar tanto el orden de rangos como la forma de un conjunto de datos en forma simultánea. (Anderson y otros, 1999: 40). Cada caso es representado por uno o más dígitos colocados a la derecha de una línea vertical y en la fila correspondiente al primer dígito del valor observado.
- Diagrama de dispersión (Scatter plot): gráfico que muestra la relación entre variables. Es útil para examinar la dirección, fuerza y forma de la relación.

Análisis Confirmatorio de Datos

Es un modo de análisis de datos que utiliza estadísticos numéricos de resúmenes generados a partir del empleo de un modelo, definido a priori, para confirmar o no una hipótesis. Se caracteriza por el empleo de indicadores como la media, la varianza y los coeficientes de correlación y regresión, así como las pruebas de hipótesis.

Así mismo, puede decirse que este es el modelo de análisis de datos cuya rutina es más sencilla de enseñar y de estandarizar mediante el uso de computadores, como se puede constatar por el desarrollo de programas informáticos como el SPSS, SAS, SPAD y STATGRAPHICS, entre otros.

Análisis Cuantitativo (Análisis descriptivo)

Análisis Univariante

Los métodos de análisis univariable se utilizan para estudiar el comportamiento de las variables de forma individual.

- **Distribuciones de frecuencias:** Permiten obtener una primera aproximación de la tendencia de los resultados, ya que indican el número de individuos que tanto en valores absolutos como en valores porcentuales han mencionado cada uno de los códigos posibles (respuestas) que puede tomar una variable determinada.
- **Medidas de tendencia central:** Los tipos estadísticos que miden la tendencia central permiten apreciar cuál es el comportamiento medio de cada variable. Los tres más utilizados son la moda, la mediana y la media.

- **Medidas de dispersión:** Permiten analizar la variabilidad de la distribución, es decir, determinar si las respuestas que han dado las personas entrevistadas se han concentrado sólo en unos cuantos valores o si, por el contrario, han sido muy variadas. La dispersión se mide respecto del comportamiento medio de la variable, por lo que la elección de la medida de dispersión que hay que utilizar también depende de la escala en que esté medida la variable que se analiza. La medida que permite evaluar la dispersión de las respuestas respecto de la media cuando la escala de medida es cuantitativa es la varianza. Otras medidas de dispersión que permiten completar la información suministrada por la varianza son las siguientes:
 - El coeficiente de simetría (skewness): indica el grado de simetría o asimetría de la distribución.
 - El coeficiente de apuntamiento (curtosis): valora si las respuestas están concentradas en pocos valores o están repartidas.
- **Inferencia estadística:** Una vez obtenidos los resultados es importante validarlos, es decir, ver si hay diferencias entre los resultados obtenidos en la investigación y unos valores determinados conocidos a priori o teóricos; en caso de que las haya, hay que comprobar si estas diferencias son estadísticamente significativas o si, por el contrario, se deben al azar. El proceso que hay que seguir para validar los resultados se denomina test de inferencia estadística.

Análisis Multivariante

Según Rosana Ferrero, científica de datos y profesora del Máster en Estadística Aplicada con R Software, en su artículo “Análisis Multivariado para Big Data” clasifica las técnicas de análisis multivariante de la siguiente manera:

1. Por grupos
 - a. Encontrar grupos (Clúster).
 - b. Comparar grupos (análisis de varianza multivariado MANOVA, el test de permutación multirespuesta MRPP, el análisis de similitudes de grupos ANOSIM, y el test de mantel MANTEL).
 - c. Discriminar entre grupos (análisis discriminante lineal LDA, árbol de regresión o clasificación CART/MRT).
2. Gradientes
 - a. Ordenación no restringida (análisis de componentes principales PCA, análisis de correspondencia CA, análisis de correspondencias sin tendencia DCA, escalamiento multidimensional métrico MDS y no métrico NMDS).
 - b. Ordenación restringida (análisis de redundancia RDA, análisis de correspondencia canónico CCA).

Análisis cualitativo (Análisis predictivo)

Hoy en día, se utiliza cada vez más software informático para análisis de datos cualitativos (QDA). Un software QDA se utiliza en muchos campos académicos, como la sociología, la psicología, la ciencia política, la medicina y la ciencia educativa, y es también una herramienta popular para las empresas y los investigadores de mercado.

Un software QDA sirve para proporcionar información sobre conjuntos de datos cualitativos sin sugerir interpretaciones. Basándose en un análisis de contenido, se puede sacar conclusiones sobre el objeto que se está estudiando (por ejemplo, datos de entrevistas). Las herramientas de software para el análisis cualitativo de datos y textos permiten una fácil clasificación, estructuración y análisis de grandes cantidades de texto u otros datos y facilitan la gestión de las interpretaciones y evaluaciones resultantes.

Existe una tendencia hacia la inclusión y el análisis de los archivos de imágenes, así como de los materiales de audio y video. Por lo tanto, un software de análisis de datos cualitativo de última generación apoya idealmente el análisis de textos, así como de archivos multimedia (imágenes, videos, grabaciones de audio) y otros tipos de documentos.

Ejemplo de herramientas que apoyan al análisis cualitativo son: QDA Miner, Atlas.ti, HyperRESEARCH, MAXQDA, NVivo, entre otras.

Aclarado este punto, corresponde tratar la diferencia entre la inteligencia de negocios y Big Data. La primera ayuda a encontrar respuestas a preguntas conocidas (puesto que define qué información es útil y relevante para la toma de decisiones en base a los datos existentes) mientras que Big Data ayuda a encontrar las cuestiones que no se sabe encontrar (se requiere de la integración de los datos de las diferentes fuentes para sacar conclusiones y aprovechar oportunidades futuras).

Análisis predictivo

La analítica predictiva es una forma de análisis avanzado que utiliza datos nuevos e históricos para predecir la actividad futura, el comportamiento y las tendencias. Implica la aplicación de técnicas de análisis estadístico, consultas analíticas y algoritmos automáticos de machine learning sobre conjuntos de datos para crear modelos predictivos que sitúen un valor numérico o puntuación en la probabilidad de que ocurra un suceso particular.

Una amplia gama de herramientas y técnicas se utiliza en los modelos y análisis predictivos. IBM, Microsoft, el SAS Institute y muchos otros proveedores de software ofrecen herramientas analíticas predictivas, incluyendo software de machine learning y tecnologías relacionadas que soportan aplicaciones de aprendizaje profundo.

Además, el software de código abierto juega un papel importante en el mercado de análisis predictivo. El lenguaje de código abierto R se utiliza comúnmente en aplicaciones de análisis predictivo, al igual que los lenguajes de programación Python y Scala.

Data mining

La minería de datos o data mining es el proceso de clasificar grandes conjuntos de datos para identificar patrones y establecer relaciones para resolver problemas a través del análisis de datos. Las herramientas de minería de datos permiten a las empresas predecir tendencias futuras.

En la minería de datos, las reglas de asociación se crean analizando datos para patrones frecuentes si/entonces (if/then), usando los criterios de soporte y confianza para localizar las relaciones más importantes dentro de los datos. El soporte es la frecuencia con la que aparecen los elementos en la base de datos, mientras que la confianza es el número de veces que las sentencias if/then son exactas.

Otros parámetros de minería de datos incluyen Análisis de secuencia o trayectoria (Sequence or Path Analysis), Clasificación y Agrupamiento (clustering).

- Los parámetros de Análisis de secuencia o trayectoria buscan patrones en los que un evento conduce a otro evento posterior. Una secuencia es una lista ordenada de conjuntos de elementos y es un tipo común de estructura de datos que se encuentra en muchas bases de datos.
- Un parámetro de clasificación busca nuevos patrones y puede resultar en un cambio en la forma en que se organizan los datos. Los algoritmos de clasificación predicen variables basadas en otros factores dentro de la base de datos.
- Los parámetros de agrupación encuentran y documentan visualmente grupos de hechos previamente desconocidos. Clustering agrupa un conjunto de objetos y los agrega basados en lo similares que son entre sí.

Big Data

En general, casi todas las definiciones están de acuerdo en la esencia de Big Data: crecimiento exponencial de la creación de grandes volúmenes de datos y la necesidad de su captura, almacenamiento y análisis para conseguir el mayor beneficio para organizaciones y empresas.

La empresa multinacional de auditoría Deloitte lo define como: “El término que se aplica a conjuntos de datos cuyo volumen supera la capacidad de las herramientas informáticas de uso común, para capturar, gestionar y procesar datos en un lapso de tiempo razonable. Los volúmenes.”

Una de las definiciones de la consultora Gartner es: “Big Data son los grandes conjuntos de datos que tiene tres características principales: volumen (cantidad), velocidad (velocidad de creación y utilización) y variedad (tipos de fuentes de datos no estructurados, tales como la interacción social, video, audio)”.

Características de Big Data

Big Data abarca diversas tecnologías. Los datos de entrada a los sistemas de Big Data pueden proceder de redes sociales, logs, registros de servidores web, sensores de flujos de tráfico, imágenes de satélites, flujos de audio y de radio, transacciones bancarias, datos de mercados financieros. ¿Todos estos datos son realmente los mismos?

Uno de los mayores desafíos ha sido identificar las características de dichos datos. Diferentes investigadores han tenido puntos de vistas distintos hacia estas características; como algunos mencionan las 3Vs [*Volumen, Velocidad y Variedad*] de datos (por ejemplo Shah, Rabhi, & Ray, 2015).

Otros informaron 4Vs [*Volumen, Velocidad, Variedad y Variabilidad*] de datos (por ejemplo Liao, Yin, Huang, & Sheng, 2014) y 6Vs [*Volumen, Velocidad, Veracidad, Variedad, Variabilidad y Valor*] de datos (Gandomi y Haider, 2015).

Según una Crítica de la Revista de Investigación Empresarial²⁸ se identificaron 7 Vs, siete características de los datos que son: Volumen, Variedad, Velocidad, Variabilidad, Veracidad, Visualización y Valor:

1. Volumen:

Son datos a gran escala y volumen, lo cual conlleva a un gran desafío para la recuperación, procesamiento, integración e inferencia de los mismos. Podemos estar hablando de terabytes, petabytes o zettabytes generados en un momento. Facebook genera diariamente más de 500 terabytes de datos y Wal-Mart recolecta cada hora más de 2.5 petabytes como resultado de las transacciones con sus clientes.

2. Variedad:

Se refiere a las diferentes formas y cualidades de los datos, estos no siguen un formato específico, debido a que son capturados de diversas formas y de múltiples fuentes.

Por ejemplo: mensajes (texto, correo electrónico, tweets, blogs), contenido transaccional (registros web, transacciones comerciales), datos científicos (datos procedentes de experimentos, del área de la salud, bases de datos biológicas) y datos web (imágenes publicadas en medios de comunicación y lecturas de sensores).

3. Veracidad:

IBM presentó esta característica de Big Data, asociándola a la falta de confianza inherente en muchas fuentes de datos estructurados, así como no estructurados.

²⁸ Uthayasankar, S., Muhammad, M., Zahir, I., & Vishanth, W.. (2016, agosto 10). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 263–286, p.7.

Akerkar (2014) y Zicari (2014) refieren la veracidad a cómo hacer frente a los sesgos, dudas, imprecisiones, ambigüedad y evidencia extraviada en los datos.

4. Velocidad:

Esto se aplica principalmente a los conjuntos de datos que se generan a través de grandes redes complejas, incluidos los datos generados por la proliferación de dispositivos digitales. Los datos procedentes de dispositivos móviles y que fluyen a través de aplicaciones móviles o el uso de tarjetas de tiendas (por ejemplo, la tarjeta de Sainsbury para recoger puntos de néctar) genera información que se puede utilizar al tiempo en que se produce, como para ofertar productos personalizados según comportamientos, patrones de compra o la ubicación espacial. (Gandomi & Haider, 2015).

5. Variabilidad:

La variabilidad no debe confundirse con la variedad, ya que variabilidad hace referencia a los diferentes significados que puede tomar un dato en el tiempo y no a los distintos tipos de datos de los que describe la variedad.

Variabilidad también está relacionado con el análisis de sentimientos (sentiment analysis). Por ejemplo, en tweets una misma palabra puede tener un significado totalmente diferente. Para realizar un análisis de sentimiento adecuado, los defensores afirman que los algoritmos necesitan ser capaces de entender el contexto y ser capaces de descifrar el significado exacto de una palabra en ese contexto (Zhang, Hu et al., 2015).

6. Visualización:

Consiste en representar información clave y el conocimiento de una manera más intuitiva y efectiva mediante el uso de diferentes formatos visuales tal como en una representación pictórica o gráfica (Taheri, Zomaya, Siegel, & Tari, 2014).

7. Valor:

Extraer conocimiento de la cantidad de datos estructurados y no estructurados para los usuarios finales. Los investigadores de Big Data consideran el valor de los datos como una característica esencial, ya que en algún lugar dentro de estos existe información valiosa (datos de alto valor), aun si la mayoría de los datos pudiera parecer insignificantes (Zaslavsky, Perera, & Georgakopoulos, 2012).

Tipos de datos

Los Big Data son diferentes de las fuentes de datos tradicionales que almacenan datos estructurados en las bases de datos relacionales. Es frecuente dividir las categorías de datos en dos grandes tipos: estructurados (datos tradicionales) y no estructurados (datos Big Data).

Sin embargo, las múltiples fuentes que están generando datos propician a la creación de nuevas categorías dentro de los tipos de datos no estructurados: datos semiestructurados y datos no estructurados propiamente dichos.

- **Datos Estructurados:**
La mayoría de las fuentes de datos tradicionales son datos estructurados. En estas fuentes, los datos vienen en un formato bien definido y conforman las bases de datos relacionales, las hojas de cálculo y los archivos, fundamentalmente. Los datos estructurados se componen de piezas de información que se conocen de antemano, es decir, vienen en un formato específico, y se producen en un orden específico, facilitando el trabajo con dichos datos. Formatos típicos son: fecha de nacimiento (DD, MM, AA); documento nacional de identidad o pasaporte (por ejemplo, 8 dígitos y una letra); número de la cuenta corriente en un banco (20 dígitos), entre otros.
- **Datos Semiestructurados:**
Los datos semiestructurados tienen un flujo lógico y un formato que puede ser definido, pero no es fácil su comprensión para el usuario. Son datos que no tienen formatos fijos, pero contienen etiquetas y otros marcadores que permiten separar los elementos.
La lectura de datos semiestructurados requiere el uso de reglas complejas que determinan cómo proceder después de la lectura de cada pieza de información. Ejemplos típicos de datos semiestructurados son: el texto de etiquetas de lenguajes XML y HTML y los registros de web logs de las conexiones a Internet. Un web log se compone de diferentes piezas de información, cada una de las cuales sirve para un propósito específico.
- **Datos No Estructurados:**
Los datos no estructurados son datos sin tipos predefinidos. Se almacenan como “documentos” u “objetos” sin estructura uniforme, y se tiene poco o ningún control sobre ellos. Por ejemplo, las imágenes se clasifican por su resolución en píxeles y no tienen campos fijos. Ejemplos típicos de datos no estructurados son: audio, video, fotografías, documentos impresos, cartas, hojas electrónicas, imágenes digitales, formularios especiales, mensajes de correo electrónico y texto, mensajes instantáneos SMS, artículos, libros, mensajes de mensajería instantánea tipo WhatsApp, Une, Joyn, Viber, Une, Wechat, Spotbros. Al menos, el 80% de la información de las organizaciones no reside en las bases de datos relacionales o archivos de datos, sino que se encuentran esparcidos a lo largo y ancho de la organización.

Plataformas

Sin duda, los datos más difíciles de dominar por los analistas son los datos no estructurados, pero su continuo crecimiento ha provocado el nacimiento de frameworks y plataformas para su manipulación como es el caso de Hadoop, Spark y NoSQL.

1. **Hadoop:**
Es un framework de software que soporta aplicaciones distribuidas bajo una licencia libre. Permite a las aplicaciones trabajar con miles de nodos y petabytes de datos.

Hadoop se inspiró en los documentos Google para MapReduce y Google File System (GFS). Hadoop es un proyecto de alto nivel Apache que está siendo construido y usado por una comunidad global de contribuyentes, mediante el lenguaje de programación Java. Yahoo ha sido el mayor contribuyente al proyecto, y usa Hadoop extensivamente en su negocio.

Al hablar de Hadoop es inevitable hablar del tema de MapReduce, que es considerado el corazón de este, al ser su entorno de programación.

MapReduce es un modelo de programación para dar soporte a la computación paralela sobre grandes colecciones de datos en grupos de computadoras. Su nombre está inspirado en los nombres de dos importantes métodos, macros o funciones en programación funcional: Map y Reduce. Su desarrollo fue liderado inicialmente por Yahoo y actualmente lo realiza el proyecto Apache. Se han escrito implementaciones de bibliotecas de MapReduce en diversos lenguajes de programación como C++, Java y Python.

2. Spark:

Apache Spark es un motor de procesamiento rápido y en memoria que permite ejecutar eficientemente grandes cargas de datos que requieran procesamiento continuo de datos (data streaming), machine learning o procesamiento SQL.

Spark puede integrarse con Hadoop y añadir nuevas funcionalidades, puede funcionar con muchos otros productos de Big Data como: CassandraDB, Google Big Query, almacenamiento de Amazon S3, Elastic Search, etc. O puede correr sobre cluster independiente.

3. Sistema de bases de datos NoSQL:

También llamadas No Solo SQL, son un enfoque hacia la gestión de datos y el diseño de base de datos que es útil para grandes conjuntos de datos distribuidos. Abarca una amplia gama de tecnologías y arquitecturas, busca resolver los problemas de escalabilidad y rendimiento de Big Data que las bases de datos relacionales no fueron diseñadas para abordar. Es especialmente útil cuando una empresa necesita acceder y analizar grandes cantidades de datos no estructurados o datos que se almacenan de forma remota en varios servidores virtuales en la nube.

Métodos de análisis de Big Data

Dado que Big Data puede mejorar la toma de decisiones e incrementar el rendimiento en las organizaciones, es necesario recurrir a métodos analíticos usados para extraer sentido a los datos.

La Universidad pública de London, Brunel University, clasifica los métodos analíticos para Big Data en función de la finalidad que se busca obtener del análisis, sea este: conocer la información que se tiene, tener mejor visión, tomar decisiones y tomar acciones.²⁹

1. Analítica descriptiva:

Hace un análisis de los datos e información para definir el estado actual de la situación del negocio en la que los acontecimientos y los patrones ya son evidentes, produciendo reportes estándares, reportes ad hoc y alertas (Joseph & Johnson, 2013).

2. Analítica inquisitiva:

Consiste en probar datos para aprobar o rechazar propuestas de negocios. Por ejemplo, análisis drill downs, análisis estadísticos, análisis de factores (Bihani & Patil, 2014).

3. Analítica predictiva:

Se ocupan de la predicción y el modelo estadístico para determinar las posibilidades futuras (Waller & Fawcett, 2013).

4. Analítica prescriptiva:

Se trata de pruebas aleatorias y de optimización para evaluar cómo los negocios mejoran sus niveles de servicio mientras disminuyen sus gastos (Joseph & Johnson, 2013).

5. Analítica preventiva:

Consiste en tener la capacidad de tomar acciones preventivas sobre eventos que puedan influir negativamente en el desempeño de las organizaciones. Por ejemplo: identificar posibles riesgos y recomendar estrategias de mitigación adelantadas en el tiempo (Szongott, Henne y von Voigt, 2012).

²⁹ Sivarajah, U., Kamal, M., Irani, Z. & Weerakkody, V.. (2016, Agosto 10). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 286. 2017, abril 14, De n.d Base de datos.

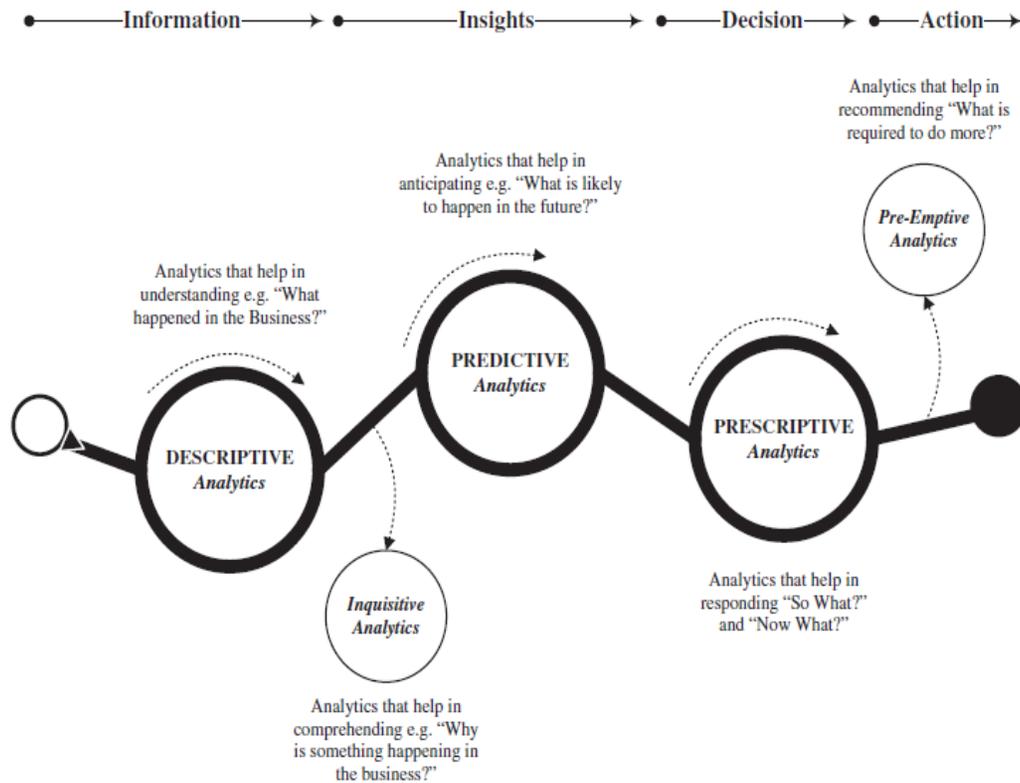


Figura 5 Tipos de Análítica de Big Data. methods. En *Critical analysis of Big Data challenges and analytical methods* (p.266). por *Journal of Business Research*, 2016

¿Quién usa Big Data en el mundo?

IBM publicó en el 2016 un artículo titulado "How 10 industries are using Big Data to win big" donde habla de las 10 industrias que están aplicando Big Data en sus actividades, estas son:

1. Telecomunicaciones.
2. Manufactura.
3. Comercio (retail).
4. Fitness.
5. Seguros.
6. Bancos.
7. Finanzas.
8. Gobiernos.
9. Salud.
10. Aerolíneas.

2.3. Marco Conceptual

Se presentan a continuación un conjunto de conceptos relacionados con el diseño y construcción del perfil profesional cuyo propósito es establecer el contexto de aplicación.

- **Análisis de Datos:**
Es la ciencia que examina datos en bruto con el propósito de sacar conclusiones sobre la información. El análisis de datos es usado en varias industrias para permitir que las compañías y las organizaciones tomen mejores decisiones empresariales y también es usado en las ciencias para verificar o reprobando modelos o teorías existentes. El análisis de datos se distingue de la extracción de datos por su alcance, su propósito y su enfoque sobre el análisis. Los extractores de datos clasifican inmensos conjuntos de datos usando software sofisticado para identificar patrones no descubiertos y establecer relaciones escondidas. El análisis de datos se centra en la inferencia, el proceso de derivar una conclusión basándose solamente en lo que conoce el investigador.³⁰
- **Análisis de Trayectoria o Path Analysis:**
El análisis de trayectoria es una forma de análisis estadístico de regresión múltiple utilizado para evaluar modelos causales examinando las relaciones entre una variable dependiente y dos o más variables independientes. Usando este método uno puede estimar tanto la magnitud como la importancia de las conexiones causales entre las variables.³¹
- **Big Data:**
Activos de información de gran volumen, alta velocidad y / o de alta variedad que exigen formas innovadoras y rentables de procesamiento de información que permiten un mejor conocimiento, toma de decisiones y automatización de procesos.³²
- **Clustering:**
Es un algoritmo de agrupamiento en donde datos con las mismas características son juntados para su posterior análisis.

³⁰ Rouse, M.. (2012). *Análisis de datos*. abril 14, 2017, de Data Profiling Sitio web: <http://searchdatacenter.techtarget.com/es/definicion/Analisis-de-Datos>

³¹ Crossman, A..(2017). *Understanding Path Analysis*. abril 15,2017, de Thought Co. Sitio web: <https://www.thoughtco.com/path-analysis-3026444>

³² Big Data.(n.d).Gartner IT Glossary.Recuperado de <http://www.gartner.com/it-glossary/big-data/>

- **Conocimientos:**
UNICEF (Unicef,2013), define el conocimiento como un estado o condición de comprensión sintetizado en conceptos amplios y aplicables. Por tanto, son todos esos temas agrupados en áreas conceptuales que un profesional posee³³.
- **Conocimientos de interés:**
Se ha definido para efectos de este proyecto, para referirse a aquellos conocimientos que el profesional considera importantes aprender dado que aún no los maneja y son requeridos para su crecimiento y desempeño profesional.
- **Cursos:**
Fuente de conocimiento para la profundización en entendimiento de un tema.
- **Data cleansing:**
También conocido como data scrubbing (borrado de datos), es el proceso de modificar o eliminar de una base de datos aquellos datos incorrectos, incompletos, formateados incorrectamente o duplicados.³⁴
- **Data mart:**
Es un subconjunto de los datos del Data Warehouse cuyo objetivo es responder a un determinado análisis, función o necesidad, con una población de usuarios específica.
- **Data profiling:**
También conocida como data archeology (arqueología de los datos). Es el análisis estadístico y la evaluación de los valores dentro de un conjunto de datos para determinar la coherencia, singularidad y lógica de los mismos.³⁵
- **Data source:**
Fuente de datos se define como los diferentes medios de los cuales se pueden obtener datos estas pueden ser bases de datos operacionales, datos externos o Información de los datos ya existentes en un Data Warehouse; así como también bases de datos relacionales o cualquier otra estructura de datos que soporté la línea de aplicaciones de negocio.

³³ UNICEF. (2013). Definition of Terms. septiembre 12,2017, de UNICEF Sitio web: https://www.unicef.org/lifeskills/index_7308.html

³⁴ Rouse, M..(2010). *data scrubbing (data cleansing)*. abril 14,2017, de TechTarget Sitio web: <http://searchdatamanagement.techtarget.com/definition/data-scrubbing>

³⁵ Rouse, M.. (2014). *Data Profiling*. abril 14,2017, de TechTarget Sitio web: <http://searchdatamanagement.techtarget.com/definition/data-profiling>

- **Data streaming:**
Proceso en el que los grandes flujos de datos se procesan en tiempo real con el objetivo de extraer conocimiento y tendencias útiles. Un flujo continuo de datos no estructurados se envía para su análisis en memoria antes de ser almacenarla en disco. La velocidad es lo más importante en el data streaming, el valor de los datos, si no se procesa rápidamente, disminuye con el tiempo.³⁶
- **Datawarehouse:**
Repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización, con las propiedades siguientes: estable, coherente, fiable y con información histórica.
- **ETL:**
Tecnología de integración de datos basada en la consolidación de datos que se usa tradicionalmente para alimentar Data Warehouse, data mart, staging área y ODS. Usualmente se combina con otras técnicas de consolidación de datos.
- **Habilidades:**
El glosario Cedefop de la Comisión Europea (Cedefop, 2008) define habilidad como la capacidad de realizar tareas y solucionar problemas.
Por tanto, es la capacidad de aplicar conocimientos, sean éstos producto de la experiencia o la teoría, a una situación específica se considerará una habilidad³⁷.
- **Herramientas:**
Son las aplicaciones, programas o estrategias que ayudan en el desempeño de las funciones del área.
- **Inteligencia de negocios:**
Conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización.
- **Metadatos:**
Datos estructurados y codificados que describen características de instancias; aportan informaciones para ayudar a identificar, descubrir, valorar y administrar las instancias descritas.

³⁶ Big data streaming.(n.d).Techopedia.Recuperado de <https://www.techopedia.com/definition/31752/big-data-streaming>

³⁷ Instituto de Tecnologías Educativas,. (2010). Habilidades y competencias del siglo XXI para los aprendices del nuevo milenio en los países de la OCDE. septiembre 12,2017, de OCDE Sitio web: http://guayama.inter.edu/wordpress/?wpfb_dl=140

- **ODS:**
Es un tipo de almacén de datos que proporciona sólo los últimos valores de los datos y no su historial; además, generalmente admite un pequeño desfase o retraso sobre los datos operacionales.
- **OLAP:**
Proceso analítico en línea, método ágil y flexible para organizar datos, especialmente metadatos, sobre un objeto o jerarquía de objetos como en un sistema u organización multidimensional, y cuyo objetivo es recuperar y manipular datos y combinaciones de los mismos a través de consultas o incluso informes.
- **Perfil Profesional:**
El perfil profesional es un conjunto de rasgos identificadores de competencias en un sujeto que recién recibe su título o grado que corresponde al profesional básico.
- **Reportes ad hoc:**
Reportes que son creados por usuarios finales en tiempo real, con los que se pretende dar respuesta a un objetivo o una pregunta específica. Se utilizan herramientas o asistentes que ofrecen la posibilidad a los usuarios de elegir las funciones de los informes y establecer el alcance, la variabilidad y otros aspectos. También pueden proporcionar una presentación visual personalizada con gráficos.³⁸
- **RFID (Radio Frequency Identification):**
Identificación por radiofrecuencia es una tecnología que incorpora el uso de acoplamiento electromagnético o electrostático en la porción de radiofrecuencia (RF) del espectro electromagnético para identificar de forma única un objeto, animal o persona. RFID está entrando en uso creciente en la industria como una alternativa al código de barras. La ventaja de la RFID es que no requiere el contacto directo.³⁹
- **Sentiment Analysis:**
Uso de métodos estadísticas, la programación neuro-lingüística, o machine learning para extraer, identificar o caracterizar el contenido de sentimiento de una unidad de texto.⁴⁰

³⁸ Ad hoc reporting. (n.d).Techopedia.Recuperado de <https://www.techopedia.com/definition/30294/ad-hoc-reporting>

³⁹ Rouse, M.. (2007). *RFID (radio frequency identification)*. mayo 4,2017, de TechTarget Sitio web: <http://internetofthingsagenda.techtarget.com/definition/RFID-radio-frequency-identification>

⁴⁰ n.d. (n.d). *Introduction to Sentiment Analysis*. abril 15,2017, de European Masters Program in Language and Communication Technologies (LCT) i Sitio web: <https://lct-master.org/files/MullenSentimentCourseSlides.pdf>

- **Sistemas transaccionales:**
Es un tipo de sistema diseñado para soportar las operaciones del negocio, generalmente es usado para almacenar, recolectar y modificar cualquier tipo de información.
- **Valores:**
Toda aquella manera de actuar o interactuar, considerada de importancia, y enfocados para este caso, a un ambiente laboral⁴¹.

⁴¹ Townsville Community Service Inc. (2008). The ethics, morals, values distinction. septiembre 12,2017, de Townsville Community Service Inc Sitio web: http://www.tcls.org.au/01_cms/details.asp?ID=39

CAPÍTULO III: ESTUDIO DE LA SITUACIÓN ACTUAL DE LA INTELIGENCIA DE NEGOCIOS Y ANALÍTICA DE DATOS EN EL SALVADOR

3.1 Metodología para la recolección de datos

3.1.1. Definición de la población y muestra

Técnica: Encuesta

Nombre:

Investigación de la Situación Actual de la inteligencia de negocios y analítica de datos (ver Anexo 3)

Unidad de análisis:

Profesionales o no que se desempeñan en cargos relacionados a Inteligencia de Negocios o análisis de datos.

Población:

Profesionales o no que actualmente desempeñen los siguientes cargos: desarrolladores de Inteligencia de Negocios, analistas de Inteligencia de Negocios, consultores de Inteligencia de Negocios, coordinadores de proyecto de Inteligencia de Negocios, analistas de datos y analista de Big Data en empresas de cualquier naturaleza y tamaño.

Lugar: Área metropolitana

Tiempo: año 2017

Muestra:

Para la determinación de la muestra se tomaron las siguientes consideraciones:

1. El Instituto del Seguro Social en su Anuario Estadístico 2016 totaliza los trabajadores cotizantes según actividad económica y sector público y privado, sin embargo no es posible conocer la cantidad de personas que se dediquen al área en estudio.⁴²
2. El MINED en su informe Resultados de la Información Estadística de Instituciones de Educación Superior 2015 publicó la cantidad de graduados de todas las universidades para todas las carreras, sin embargo no es posible determinar de estos datos la cantidad de profesionales que se dediquen al área de estudio.⁴³

⁴² Instituto Salvadoreño del Seguro Social. (2017, abril 01). *Estadísticas del ISSS 2016*. Anuario Estadístico, n.d, 96. 2017, abril 19, De n.d Base de datos.

⁴³ Ministerio de Educación República de El Salvador. (2016, diciembre 01). *Resultados de la Información Estadística de Instituciones de Educación Superior 2015*. Portal de Transparencia, n.d, 166. 2017, abril 19, De n.d Base de datos.

3. El cálculo probabilístico de muestras para poblaciones desconocidas o infinitas (Ver anexo 4) indica que la muestra es de 385 personas.

Trabajar con muestras grandes o que no puedan ser alcanzadas afectan la confiabilidad de los resultados obtenidos. Pentti Routio, docente de la Universidad de Arte y Diseño en Helsinki dice lo siguiente:

“Sucede a menudo que algunos casos en la muestra resultan infructuosos porque no pueden ser alcanzados, o las medidas fallan, o los entrevistados rechazan cooperar etc. El método normal es entonces sobredimensionar la muestra levemente, y después se olvida simplemente los casos que fallan.”⁴⁴

Ante estas consideraciones, y teniendo en mente el objetivo del instrumento, la unidad de análisis y la calidad de los resultados, se determina que la muestra a utilizar para este instrumento se considera intencional, dirigido o por conveniencia.⁴⁵ En esta técnica de muestreo, el investigador selecciona a la muestra en función del conocimiento que tiene de ella, es decir en función de su representatividad. Por tanto, los investigadores encuestarán a las personas que sepan que cumplen con las características de la población definida, es decir, que estén desempeñando un cargo de los especificados anteriormente.

3.1.2. Diseño del instrumento

Estimado encuestado

Somos estudiantes de la Universidad de El Salvador optando al título de Ingeniero de Sistemas Informáticos, estamos trabajando en una investigación que servirá para elaborar el trabajo de graduación “Propuesta de formación profesional y académica para el perfil de especialización en el área de Inteligencia de Negocios y Analítica de datos para la carrera de pregrado de Ingeniería de Sistemas informáticos de la Universidad de El Salvador”

Esta encuesta está destinada a profesionales o estudiantes de ingenierías o licenciaturas de sistemas informáticos, ciencias de la computación o carreras afines que tenga experiencia en el desarrollo de actividades en Inteligencia de Negocios, análisis de datos y/o Big Data. Solicitamos de su valioso tiempo para contestar unas preguntas.

Las opiniones de todos los encuestados serán de mucho aporte para los resultados de esta investigación, pero nunca se comunicarán datos individuales.

Objetivo: Conocer la experiencia de las personas que ejercen actividades relacionadas con la Inteligencia de Negocios, Analítica de datos y/o Big Data.

⁴⁴ Routio, P.. (2007). *Arteología, la ciencia de productos y profesiones: Muestreo*. mayo 1,2017, de Universidad de Arte y Diseño Sitio web: <http://www.uiah.fi/projects/metodi/252.htm>

⁴⁵ Gidalberto Bonilla. (1993). *Cómo hacer una tesis de investigación con técnicas estadísticas*. El Salvador: UCA Editores.

Indicaciones: Conteste las siguientes interrogantes según el criterio que usted crea conveniente. Si alguna respuesta no se acomoda a su criterio, elija la que más se asemeje.

Preguntas generales

Objetivo: Conocer el área de cada encuestado, para que a partir de esta segmentar las preguntas siguientes:

1. ¿En qué área se desempeña actualmente?

Inteligencia de Negocios

Análisis de datos

Big Data

Objetivo: Obtener una lista de las habilidades para realizar el perfil académico.

2. En base a su experiencia, mencione ¿cuáles son las habilidades que un profesional en su área debe poseer?

Objetivo: Obtener una lista de los valores profesionales para realizar el perfil académico

3. En base a su experiencia, menciona que valores profesionales debe poseer un profesional de su área

Analítica de datos

Objetivo: Medir el porcentaje de participación de los diferentes tipos de análisis, para inferir las necesidades actuales y futuras en el área laboral en cuanto a análisis de datos se trate.

4. ¿Qué tipo de análisis realiza en su cargo?

- a. Diagnóstico: Un vistazo al desempeño del pasado para determinar qué sucedió y por qué.

- b. Descriptivo: Lo que está sucediendo ahora basado en los datos entrantes.

- c. Prescriptivo: Este tipo de análisis revela qué acciones deben tomarse.

- d. Predictivo: Un análisis de los posibles escenarios de lo que podría suceder.

Objetivo: Obtener una lista de conocimientos técnicos para realizar el perfil académico.

5. En base a su experiencia, mencione que conocimientos técnicos o base teórica ha requerido para llevar a cabo análisis de datos

Objetivo: Conocer qué herramientas tiene mayor uso, para considerarla en el desarrollo del curso académico.

6. ¿Qué herramientas utiliza/ha utilizado para el análisis?
- a) SAS
 - b) IBM
 - c) KNIME
 - d) RapidMiner
 - e) Otro

Objetivo: Conocer los cursos que los encuestados han recibido, para su posterior indagación y obtención de temas de Big Data a considerar en la creación del curso académico y profesional.

7. En cuanto al tema de Big Data, ¿ha recibido algún curso o capacitación?

Si la Respuesta es **Sí**, responder a lo siguiente:

7.1 Modalidad: ¿Presencial o en línea?

7.2 ¿Cuál fue la duración (horas, meses, etc.)?

7.3 Menciona que cursos recibió y quién lo impartió

Si la respuesta es **No**, responder a lo siguiente:

7.4 ¿En qué temas relacionadas con Analítica de datos y Big Data considera debería recibir curso, capacitación o consultoría profesional en función de sus metas profesionales o los objetivos de la compañía a la que labora?

Objetivo: Obtener los conocimientos técnicos relacionados con Big Data para su incorporación en el curso académico a desarrollar.

8. En base a su experiencia o conocimiento teórico de Big Data, ¿Qué temas considera un analista de Big Data debe manejar?

Objetivo: Identificar los temas complejos relacionados al análisis de datos y Big Data para evaluar y determinar su incorporación al curso a crear en función de los alcances del estudio.

9. En base a su experiencia, ¿Qué temas relacionados a la Analítica de datos y Big Data considera que requieren de mayor estudio y profundidad debido a su alto nivel de complejidad?

Objetivo: Determinar necesidades futuras en el área de análisis de datos y Big Data para su incorporación en el curso.

10. ¿Cuál considera son las nuevas tendencias en el área de análisis de datos y Big Data?

Inteligencia de Negocios

Objetivo: Obtener una lista de conocimientos técnicos para realizar el perfil académico.

11. En base a su experiencia, mencione que conocimientos técnicos ha requerido para llevar a cabo Inteligencia de Negocios

Objetivo: Identificar qué herramientas tiene mayor uso, para considerarla en el desarrollo el curso académico.

12. ¿Qué herramientas utiliza/ha utilizado para realización de reportes?

- a. Tableau
- b. Qlik
- c. Microsoft Power BI
- d. Pentaho
- e. Microstrategy
- f. Otro

Objetivo: Identificar los cursos que los encuestados han recibido, para su posterior indagación y obtención de temas de Inteligencia de Negocios a considerar en la creación del curso académico y profesional.

13. En cuanto al tema de Inteligencia de Negocios, ¿ha recibido algún curso o capacitación?

Si la Respuesta es **Sí**, responder a lo siguiente:

13.1 Modalidad: ¿Presencial o en línea?

13.2 ¿Cuál fue la duración (horas, meses, etc.)?

13.3 Menciona que cursos recibió y quién lo impartió

Si la respuesta es **No**, responder a lo siguiente:

13.4 ¿En qué temas relacionados con Inteligencia de Negocios considera debería recibir curso, capacitación o consultoría profesional en función de sus metas profesionales o los objetivos de la compañía?

Objetivo: Determinar necesidades futuras en el área de Inteligencia de Negocios para su incorporación en el curso.

14. ¿Cuál considera son las nuevas tendencias en el área de Inteligencia de Negocios?

3.2 Resultados del estudio

De los resultados del estudio (Ver Anexo 3) se concluyó lo siguiente:

1. El 52% de las personas que participaron en la encuesta conocen la existencia de la inteligencia de negocios porque en su empresa existe un área y el 33% sabe que por lo menos en una empresa del país existe esta área.

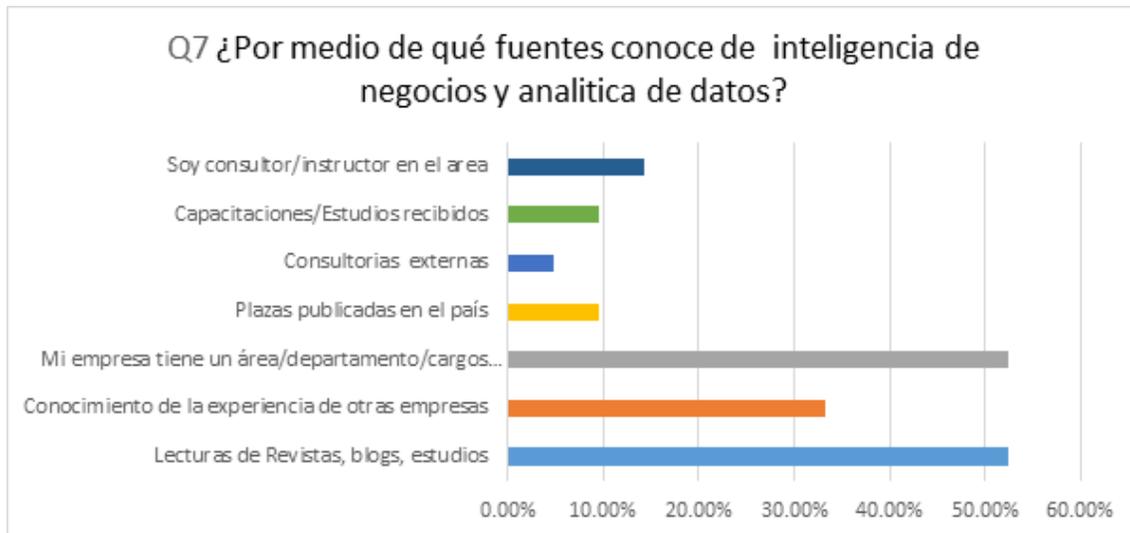


Gráfico 1 ¿Por medio de qué fuentes conoce de Inteligencia de Negocios y Analítica de datos?

2. Las 24 personas encuestadas consideran que los datos que se generan en sus empresas deben analizarse con un mayor énfasis.

Q10 ¿Considera que en su empresa o en su gerencia/departamento es importante que los datos se analicen con mayor énfasis?

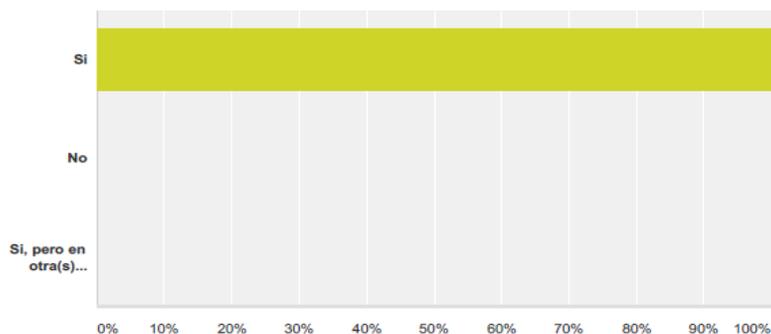


Gráfico 2 ¿Considera que en su empresa o en su gerencia/departamento es importante que los datos se analicen con mayor énfasis?

3. El 54% de los directivos se considera disconforme con respecto a la información que actualmente poseen para tomar decisiones.

Q13 ¿Considera que la información que actualmente posee es suficiente para tomar decisiones?

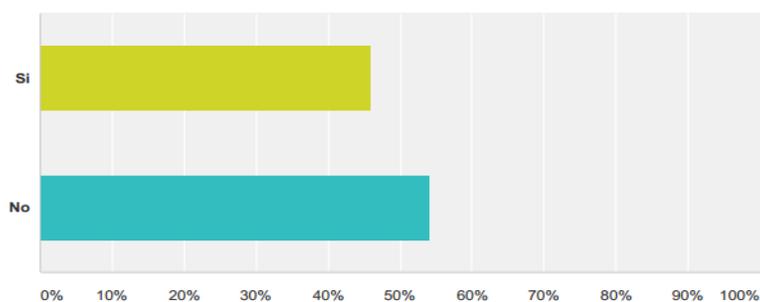


Gráfico 3 ¿Considera que la información que actualmente posee es suficiente para tomar decisiones?

**CAPÍTULO IV:
CONSTRUCCIÓN DEL
PERFIL ACADÉMICO
PARA EL ESTUDIANTE
DE ESPECIALIZACIÓN
DE INTELIGENCIA DE
NEGOCIOS Y ANALITICA
DE DATOS**

4.1 Metodología para el diseño curricular

4.1.1. Investigación de metodologías de diseño curricular

El primer paso para la construcción del perfil académico fue el diseño de la currícula. Se investigaron tres metodologías de diseño: ACM (Association for Computing Machinery) ya que es uno de los referentes en cuanto a recomendaciones de currícula y tiene el apoyo de la IEEE Computer Society, además contiene planes de estudio muy completos que van acorde a los campos relacionados a ciencias de computación y ciencias de información; Tuning Latinoamérica y ACAI (Adopción de enfoques de calidad, accesibilidad e innovación en la educación superior de Latinoamérica) son metodologías que están adaptadas al contexto regional, es decir fueron creadas adaptándose al contexto latinoamericano.

4.1.1.1. ACM (Association for Computing Machinery)

ACM proporciona recomendaciones para planes de estudios para cinco sub-disciplinas de la informática⁴⁶: Ciencias de la Computación, Ingeniería Informática, Sistemas de Información, Tecnología de la Información, e Ingeniería de Software. Por lo tanto, fue necesario comparar el enfoque de cada disciplina para definir cuál de entre todas las recomendaciones curriculares de ACM debería ser contrastada con ACAI y Tuning Latinoamérica.

Se presenta a continuación cada una de las sub disciplinas de ACM:

1. El enfoque de las Ciencias de la Computación en ACM desarrolla un estilo de pensamiento y de resolución de problemas que emerge de las experiencias obtenidas mediante el estudio del campo y la práctica profesional. Conecta y extrae de muchas disciplinas, incluyendo matemáticas, ingeniería eléctrica, psicología, estadísticas, artes y lingüística.
2. El enfoque de la Ingeniería Informática combina la ingeniería informática y la ingeniería eléctrica, incorporando la ciencia y la tecnología del diseño, la construcción, implementación y mantenimiento de componentes de software y hardware de sistemas informáticos modernos y equipo controlado por ordenador.
3. El enfoque de Sistemas de Información es el uso eficaz y eficiente de la información y las tecnologías de comunicaciones para lograr una ventaja competitiva y la excelencia en el servicio en las organizaciones.

⁴⁶ ACM . (2017). Curricula Recommendations. julio 30,2017, de ACM Sitio web: <http://www.acm.org/education/curricula-recommendations>

Además de apoyar a los procesos de gestión en todos los niveles: operacional, táctico y estratégico para la identificación, análisis y toma de decisiones.

4. El enfoque de la Tecnología de la Información es abogar por los usuarios y satisfacer sus necesidades desde el contexto organizacional y social a través de la selección, creación, aplicación, integración y administración de tecnologías de la información.
5. El enfoque de la Ingeniería de Software es la producción de software de manera sistemática, controlada y eficiente para una amplia gama de aplicaciones y contextos y con uso de diversas herramientas (tanto conceptuales como físicas), aplicando las mejores prácticas para desplegarlos.

Por tanto, se consideró apropiado utilizar el enfoque de *Sistemas de Información de ACM* por estar orientado a la gestión del negocio y el proceso de toma de decisiones y mantener en consideración en insumo matemático y estadístico del enfoque de Ciencias de la Computación.

ACM (Association for Computing Machinery) IS 2010 Curriculum Guidelines for Undergraduate Degree Programs in Information Systems⁴⁷

El informe IS 2010 es la última versión del modelo curricular para niveles de pregrado en Sistemas de Información (IS) que comenzó a principios de los años setenta; proporciona orientación sobre el contenido básico para programas de IS, sugiriendo un conjunto de cursos básicos comunes a los planes de estudio de esta disciplina; así como la consideración de cursos electivos dentro de un programa de estudios. Además, considera que las unidades académicas deben mantener sus programas académicos en función de las necesidades a nivel nacional, regional y mundial.

Esta revisión tiene cuatro grandes características:

1. El currículo va más allá de las escuelas de negocios y de gestión. Las versiones anteriores del currículo de IS se habían enfocado a una escuela de negocios norteamericana; este modelo, sin embargo, está guiado por la creencia de que, aunque los negocios continúen siendo probablemente el dominio principal de los Sistemas de Información, la disciplina es propicia para que se aplique a diversos dominios.

⁴⁷ Topi H., Valacich J., Wright R., Kaiser K., Nunamaker J.F., Sipior J. & Vreede G.J. 2010. IS 2010 Curriculum Guidelines for Undergraduate Degree Programs in Information Systems. En IS 2010 Curriculum Guidelines (1, 88) n/a: n/a.

2. Las expectativas de los resultados del plan de estudios han sido cuidadosamente reevaluadas y articuladas primero en forma de las capacidades de IS de alto nivel (cursos básicos o core courses) y también en tres categorías de conocimiento y habilidades: conocimientos y habilidades específicas de IS, Conocimiento y habilidades fundamentales y fundamentos del dominio.
3. El plan de estudios está estructurado, de manera que separa entre cursos básicos y cursos electivos. Los siete cursos básicos especifican las unidades de conocimientos requeridas y los temas que deben ser cubiertos en cada programa de Sistemas de Información. El plan de estudios modelo incluye ejemplos de cursos electivos, que amplían la cobertura proporcionada por el curso básico dentro de un área de conocimiento específica o introducen nuevas áreas de conocimiento en el currículo. Los cursos electivos son elementos esenciales de la trayectoria de la carrera de IS. La Figura 7 muestra los cursos básicos y electivos de muestra aplicados a una serie de carrera sugeridas por IS 2010.
4. El diseño de este plan de estudios permite su adopción en una variedad de contextos de sistemas educativos.

Structure of the IS Model Curriculum: Information Systems specific courses																		
Career Track:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	
Core IS Courses:																		A = Application Developer B = Business Analyst
Foundations of IS	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	C = Business Process Analyst
Enterprise Architecture	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	D = Database Administrator
IS Strategy, Management and Acquisition	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	E = Database Analyst
Data and Information Management:	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	F = e-Business Manager
Systems Analysis & Design	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	G = ERP Specialist
IT Infrastructure	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	H = Information Auditing and Compliance Specialist
IT Project Management	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	I = IT Architect J = IT Asset Manager
Elective IS Courses:																		K = IT Consultant
Application Development	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	L = IT Operations Manager
Business Process Management		●	●															M = IT Security and Risk Manager
Collaborative Computing																		N = Network Administrator
Data Mining / Business Intelligence		●		●	●	○	○	○	○	○	○	○	○	○	○	○	○	O = Project Manager
Enterprise Systems		●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	P = User Interface Designer
Human-Computer Interaction	●																●	Q = Web Content Manager
Information Search and Retrieval		○		○	●												●	
IT Audit and Controls	○		●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
IT Security and Risk Management	○			○	○	○	○	○	○	○	○	○	○	○	○	○	○	
Knowledge Management		●		○		○				○								
Social Informatics																		

Key:
● = Significant Coverage
○ = Some Coverage
Blank Cell = Not Required

Figura 6 Estructura del modelo curricular IS 2010

Cursos básicos o core courses

IS 2010 especifica siete resultados que un graduado de IS debe tener, estos son conocidos como las capacidades de alto nivel de SI o los cursos básicos, los cuales son:

1. Fundamentos de los Sistemas de Información.
2. Gestión de datos y de información.
3. Arquitectura de la empresa.
4. Infraestructura de TI.
5. Gestión de Proyectos IS.
6. Análisis y diseño de sistemas.
7. Estrategia, Gestión y Adquisición de IS.

Conocimientos y habilidades de los cursos básicos

Estas capacidades de alto nivel representan conocimientos y habilidades categorizados de la siguiente manera:

1. Conocimiento y habilidades específicos de IS (IS Specific)

IS Specific se refiere al conocimiento y las habilidades particulares de IS

- a. Identificar y diseñar oportunidades de mejora para organizaciones.
- b. Análisis de los compromisos.
- c. Diseñar e implementar soluciones de sistemas de información.
- d. Gestión de operaciones de tecnología de la información en curso.

2. Conocimientos básicos y habilidades Fundamentales (IS Foundation)

IS Foundation es un conjunto de conocimientos y habilidades analíticas y de comunicación que permite al profesional en IS competir en el mundo real.

- a. Liderazgo y colaboración.
- b. Comunicación.
- c. Negociación.
- d. El pensamiento analítico y crítico, incluyendo la creatividad y el análisis ético.
- e. Fundamentos Matemáticos.

3. Principios del dominio (Domain Fundamentals)

Domain Fundamentals es cualquier conocimiento o habilidad que le permite al profesional en IS hacer frente a cualquier dominio específico. El dominio específico puede ser algún tipo de organización (gobiernos, sin fines de lucro, etc), industrias (bancos, manufactura) o alguna función de negocios (recursos humanos, finanzas, ventas y distribución, etc.).

- a. Modelos generales de un dominio.
- b. Especializaciones clave dentro de un dominio.
- c. Evaluación del desempeño dentro de un dominio.

Muestra de cursos electivos de ACM IS 2010

- Desarrollo de aplicaciones.
- Gestión de Procesos de Negocio.
- Sistemas empresariales.
- Introducción a la interacción hombre-computadora.
- Auditoría y controles de TI.
- IS Innovación y Nuevas Tecnologías.
- Seguridad de TI y Gestión de Riesgos.

Por tanto, ACM IS 10 trata a través de su estructura, cubrir los dos roles centrales de un sistema de información: 1. La gestión de datos e información y 2. El diseño y análisis de sistemas.

4.1.1.2. ACAI- LA (Adopción de enfoques de calidad, accesibilidad e innovación en la educación superior de Latinoamérica)

El reto que hoy supone la internacionalización de la Educación Superior ha obligado a las universidades a romper los mecanismos de gestión tradicionales y a establecer iniciativas educativas abiertas y en red, que garanticen el contexto actual, la calidad y la pertinencia de sus programas académicos.

El objetivo del proyecto ACAI-LA “Adopción de enfoques de calidad, accesibilidad e innovación en la educación superior de Latinoamérica” es contribuir a la modernización de la educación superior virtual, asegurando su calidad, innovando en metodologías pedagógicas, garantizando la equidad en el acceso a la universidad de la población más vulnerable, fomentando el desarrollo de cualificaciones para la inserción laboral de egresados y actualizando los recursos de las universidades de América Latina.

El proyecto ACAI-LA abarca un conjunto de organizaciones con culturas, prácticas y estructuras organizativas distintas, muestra de la diversidad en la región América Latina. Se busca cubrir toda el área de América Latina para mejorar el impacto, diseminación y explotación de resultados.

Los socios participantes en el proyecto son:

- Universidad de Alcalá (España).
- Universidad Americana (Nicaragua).
- Universidad Nacional Autónoma de Nicaragua, León (Nicaragua).
- Universidad del Magdalena (Colombia).
- Católica del Norte Fundación Universitaria (Colombia).
- Universidad Panamericana (Guatemala).
- Universidad Galileo (Guatemala).
- Universidad Nacional del Litoral (Argentina).
- Universidad Nacional de Córdoba (Argentina).

- Helsinki Metropolia University of Applied Sciences (Finlandia).
- Universidad Telemática Internacional UNINETTUNO (Italia).

Objetivos

Los objetivos específicos de la iniciativa ACAI-LA son:

1. Cooperación para la mejora la educación superior, abierta y accesible, asegurando la calidad de los programas e incorporando la innovación educativa como estrategia de fortalecimiento del sistema universitario.
2. Modernización de los recursos tecnológicos para la implementación de programas en red, certificados, accesibles, multilingües y de contenidos abiertos, que soporten acciones de movilidad virtual de egresados y académicos.
3. Creación o actualización de servicios en red, que apoyen las prácticas docentes de innovación pedagógica, garanticen el acceso y permanencia en la universidad de población desfavorecida (con discapacidad sensorial, dificultades económicas o discriminadas por razones de género o etnia), y den soporte al desarrollo de cualificaciones profesionales mejorando la empleabilidad de los egresados estableciendo vínculos con la empresa.
4. Fomento de la cooperación entre países de la Unión Europea y América Latina, estableciendo una red transregional de investigación, innovación científica y tecnológica en las áreas de ACAI-LA, aportando recomendaciones a los responsables de reformas en la educación superior y promoviendo la convergencia entre ambas regiones.

Esta metodología se basa en cuatro fundamentos:

- **Fundamento 1:** Debido a la diversidad de procesos existentes para la creación de cursos virtuales, la elección del marco de referencia propuesto por la norma ISO/IEC 19796 (ISO, 2005) permite trabajar sobre una base replicable que sea fácilmente adaptable por otras instituciones educativas.

La norma ISO/IEC 19796 (ISO, 2005) identifica siete procesos relevantes dentro del ciclo de vida de los sistemas de información y comunicación utilizados para el aprendizaje, educación y formación. Los siete procesos son: (1) Análisis de necesidades; (2) Análisis del marco; (3) Concepción y Diseño; (4) Desarrollo y Producción; (5) Implementación; (6) Aprendizaje; (7) Evaluación y Optimización. Cada uno de los procesos contiene una serie de subprocesos para un total de 38 subprocesos.

- **Fundamento 2:** Las tres acciones para desarrollar prácticas de aprendizaje accesible propuestas por Seale (2007), deben ser incorporadas en los diferentes componentes de un proyecto virtual accesible.

Estas acciones se describen a continuación:

- a. Formación para desarrollar contenido multimedia y objetos de aprendizaje accesibles.
 - b. Implementar soluciones holísticas para desarrollar e-Learning y experiencias semi-presenciales accesibles.
 - c. Fomentar la creación de programas institucionales para crear conciencia y abordar los problemas de accesibilidad.
- **Fundamento 3:** Es importante contemplar la formación del personal involucrado en los diferentes procesos de un proyecto educativo virtual. La formación debe estar fundamentada en la base de generación de conciencia entre los participantes sobre la importancia de crear contenidos y actividades accesibles, en especial para los estudiantes con discapacidad.
 - **Fundamento 4:** Es indispensable incorporar métricas basadas en indicadores de accesibilidad que permitan utilizar la metodología propuesta como una referencia para comparación con modelos de calidad. Esto permitirá realizar auditorías sobre accesibilidad y generar modelos de madurez.

Tomando como acción de partida el fundamento 1, la norma ISO/IEC 19796 ha sido seleccionada como base para la propuesta inicial. A partir de un trabajo consensuado se han definido siete componentes interrelacionados hacia una metodología para crear cursos virtuales accesibles. Los componentes definidos son: procesos, actividades, tareas, productos, técnicas, métricas y participantes.

Los procesos representan las etapas principales en el ciclo de vida de un proyecto educativo virtual accesible. La secuencia de los procesos e interacción entre los mismos puede variar según el proyecto educativo, algunos ejemplos de secuencias son: ciclo de vida iterativo por ediciones, ciclo de vida iterativo con análisis y diseño compartido, ciclo de vida con evaluación y optimización por proceso. Los procesos definidos para la metodología propuesta coinciden con los siete procesos propuestos en la norma ISO/IEC 19796.

4.1.1.3. Tuning

El proyecto Tuning – América Latina surge en un contexto de intensa reflexión sobre educación superior, tanto a nivel regional como internacional. Hasta finales de 2004, Tuning había sido una experiencia exclusiva de Europa, un logro de más de 175 universidades europeas, que desde el año 2011 llevan adelante un intenso trabajo en pos de la creación del Espacio Europeo de Educación Superior como respuesta al desafío planteado por la Declaración de Bolonia⁴⁸.

48 Desarrollo cronológico. agosto 20, 2017, de Espacio Europeo Educación Superior Sitio web: <http://www.eees.es/es/eees-desarrollo-cronologico>

En términos teóricos, el proyecto Tuning América Latina remite implícitamente a un marco reflexivo-crítico, producto de una multi-referencialidad, tanto pedagógica como disciplinaria, para compatibilizar sus líneas de acción. El proyecto no puede enfocarse como una “receta”, sino como una metodología que procede de una perspectiva, cuya finalidad es incorporar los diferentes aspectos de la diversidad de los países que en él intervienen e interactúan.

Es un trabajo conjunto que busca y construye lenguajes y mecanismos para la comprensión recíproca de los sistemas de enseñanza superior, que faciliten los procesos de reconocimiento de carácter transnacional y transregional. Ha sido concebido como un espacio de reflexión de actores comprometidos con la educación superior, que, a través de la búsqueda de consensos, contribuye para avanzar en el desarrollo de titulaciones fácilmente comparables y comprensibles, de forma articulada, en América Latina.

¿Qué implica Tuning – América Latina?

El proyecto busca iniciar un debate cuya meta es identificar e intercambiar información y mejorar la colaboración entre las instituciones de educación superior, para el desarrollo de la calidad, efectividad y transparencia.

La protección de la rica diversidad de la educación superior latinoamericana es fundamental en el proyecto, así como la autonomía universitaria. El debate se nutre teniendo en cuenta las particularidades de las diferentes instituciones. Este punto es un pilar básico del proyecto. Uno de sus propósitos centrales es contribuir al desarrollo de titulaciones fácilmente comparables y comprensibles “desde adentro”, en base a los objetivos que la titulación se marque, desde los perfiles buscados para los egresados, ofreciendo elementos que posibiliten ampliar la articulación entre los sistemas de educación superior de los países de América Latina.

Mediante la búsqueda de perspectivas que pudiesen facilitar la movilidad de los poseedores de títulos universitarios y profesionales en América Latina y quizás también en Europa, el proyecto tiene como meta impulsar consensos a escala regional sobre la forma de entender los títulos, desde el punto de vista de las competencias que los poseedores de dichos títulos serían capaces de alcanzar.

De esta forma, el inicio del proyecto está dado por la búsqueda de puntos comunes de referencia, centrados en las competencias.

La elección de usar puntos de referencia comunes y no definiciones de asignaturas (cursos, materias) muestra un claro posicionamiento, ya que, si se quiere propiciar la movilidad profesional y académica entre nuestros países, la formación universitaria en cada uno tiene que presentar cierto nivel de consenso, con respecto a puntos de referencia acordados conjuntamente, y reconocidos dentro de cada una de las áreas y disciplinas específicas. Además, el uso de puntos de referencia respeta la diversidad, la libertad y la autonomía.

Siguiendo la metodología propia, Tuning – América Latina tiene cuatro grandes líneas de trabajo:

1. Competencias (genéricas y específicas de las áreas temáticas).
2. Enfoques de enseñanza, aprendizaje y evaluación de estas competencias.
3. Créditos académicos.
4. Calidad de los programas.

En lo concerniente a la primera línea, se trata de identificar competencias compartidas, que puedan generarse en cualquier titulación y que son consideradas importantes por ciertos grupos sociales. Hay ciertas competencias, como la capacidad de aprender y actualizarse permanentemente, la capacidad de abstracción, análisis y síntesis, etc., que son comunes a todas o casi todas las titulaciones.

En una sociedad en transformación, donde las demandas se están reformulando constantemente, estas competencias genéricas se vuelven muy importantes.

En esta primera línea, se analiza, además de las competencias genéricas, aquellas que se relacionan con las áreas temáticas. Éstas últimas están vinculadas con una disciplina y son las que confieren identidad y consistencia a un programa específico.

Las competencias específicas difieren de disciplina a disciplina. Para Tuning, es necesario desarrollar programas más transparentes y comparables a nivel latinoamericano, de forma de asegurar resultados del aprendizaje y competencias equivalentes para cada titulación.

La definición de estas competencias es responsabilidad de los académicos, en consulta con otros grupos interesados en el tema. Al definir competencias y resultados del aprendizaje, se desarrollan puntos de referencia consensuados, que sientan bases para la garantía de la calidad y contribuyen con los procesos de evaluación nacional e internacional.

Con la segunda línea, se propone preparar una serie de materiales que permitan visualizar cuáles serán los métodos de enseñanza, aprendizaje y evaluación más eficaces para el logro de los resultados del aprendizaje y las competencias identificadas. Esto implica desarrollar una combinación novedosa de enfoques de enseñanza y aprendizaje, para estimular —o permitir que se desarrollen— las competencias que se diseñan en el perfil.

Los cambios de enfoques, objetivos de enseñanza y aprendizaje implican también modificaciones en los métodos y criterios de evaluación, en función no sólo de los contenidos, sino también de habilidades, destrezas y valores. Cada estudiante debe experimentar una variedad de enfoques y tener acceso a diferentes contextos de aprendizaje, cualquiera sea su área de estudio.

Por supuesto, la transparencia y la comparabilidad de los métodos y criterios para evaluar el logro de las competencias son esenciales, si queremos contribuir a la mejora de la calidad.

Si la primera línea del proyecto busca la definición de las competencias genéricas y específicas, la segunda busca el modo más adecuado de aprenderlas, enseñarlas y evaluarlas.

En la tercera línea, se inicia una reflexión sobre el impacto y la relación de este sistema de competencias con el trabajo del estudiante, y su conexión con el tiempo resultante medido en créditos académicos. Se busca subrayar también su relación con los enfoques de enseñanza, aprendizaje y evaluación.

El proceso de Tuning requiere una definición clara de los conceptos asociados con los créditos, y de las metas, objetivos y resultados del aprendizaje. Por todo ello, es necesario lograr una mayor claridad y un mejor conocimiento, respecto a los siguientes temas: el papel de los créditos, la asignación de los créditos a los cursos, el diseño global del programa de estudio, el cálculo de los créditos en base al trabajo del estudiante, la relación entre el trabajo del estudiante, los métodos de enseñanza y los resultados del aprendizaje.

Esta línea no estaba incluida inicialmente en el proyecto, pero en las visitas preparatorias que se realizaron a los Ministerios de Educación, Consejos de Educación Superior y Conferencias de Rectores de América Latina, algunos países consideraron este tema como un aspecto emergente y relevante para ser integrado en las reflexiones del proyecto, ya que, en su opinión, facilita la cooperación académica.

Finalmente, la cuarta línea del proyecto destaca que la calidad es parte integrante del diseño del currículo basado en competencias, lo que resulta fundamental para articular las tres líneas expuestas anteriormente.

Si un grupo de académicos desea elaborar un programa de estudios, o redefinirlo, necesita un conjunto de elementos para contribuir a la calidad de estos. La confianza mutua entre las instituciones de educación superior y el reconocimiento de las titulaciones que éstas expidan debe tener como soporte básico una metodología común y contrastada de evaluación de la calidad.

La movilidad y el reconocimiento de estudios requieren un clima de confianza y de transparencia y también una correspondencia comprobada entre los elementos fundamentales de la formación, en los distintos sistemas de educación superior.

Un trabajo serio y articulado de las cuatro líneas propuestas redundará en el fomento de la transparencia de los perfiles profesionales y académicos de las titulaciones y los programas de estudio y favorecerá un énfasis cada vez mayor en los resultados.

La idea que los estudiantes adquieran más competencias determinadas afecta positivamente la claridad en la definición de los objetivos fijados para un programa específico.

Esto se logrará añadiendo indicadores que puedan ser medidos de forma precisa, mientras se establece que esos objetivos tienen que ser dinámicos y más acordes con las necesidades de la sociedad y el empleo.

Estos cambios podrán llevar a una transformación en el enfoque de las actividades educativas, puesto que favorecerá la participación sistemática del estudiante, individualmente o en grupo, en la preparación de trabajos pertinentes, presentaciones, etc.

El interés en el desarrollo de competencias en los programas concuerda con un enfoque de la educación centrado primordialmente en el estudiante y en su capacidad de aprender, exigiendo más protagonismo y cuotas más altas de compromiso puesto que es el estudiante quien debe desarrollar las competencias.

Asimismo, se facilita la innovación a través de la elaboración de nuevos materiales de enseñanza, que favorece a estudiantes y profesores; así como, al proceso de enseñanza, aprendizaje y evaluación, en su conjunto.

4.1.2. Comparación de metodologías estudiadas

Se construyó un cuadro de comparación de sus principales elementos, que son:

- Alcance.
- Campo de Estudio.
- Definición / Objetivo.
- Estructura.
- Metodología.

	ACM IS 2010	TUNING	ACAI
Alcance	Regional, Nacional y Mundial	Regional, Nacional y Mundial	América Latina y la Unión Europea (11 Universidades)
Campo	Sistemas de Información	Informática	En General
Definición /Objetivo	Ayudar a los programas en Sistemas de Información a producir graduados a nivel de pregrado competentes y aptos para adquirir responsabilidades laborales o de estudios superiores en el área de sistemas de información.	Construir titulaciones compatibles, comparables, relevantes para la sociedad y con niveles de calidad y excelencia, preservando la valiosa diversidad que viene de las tradiciones de cada uno de los países.	Contribuir a la modernización la Educación Superior virtual, innovando en metodologías pedagógicas, garantizando la equidad en el acceso a la universidad de la población, fomentando el desarrollo de cualificaciones para la inserción laboral de egresados y actualizando los recursos de las universidades de América Latina.

Tabla 1 Comparación de Metodologías

	ACM IS 2010	TUNING	ACAI
Estructura	<p>Siete capacidades de alto nivel para IS y tres categorías</p> <p>Categorías:</p> <ol style="list-style-type: none"> 1. Conocimientos y habilidades específicas de IS, 2. Conocimiento y habilidades fundamentales 3. Fundamentos del dominio. <p>Capacidad de alto nivel:</p> <ul style="list-style-type: none"> • Fundamentos de los Sistemas de Información • Gestión de datos y de información • Arquitectura de la empresa • Infraestructura de TI • Gestión de Proyectos IS • Análisis y diseño de sistemas • Estrategia, Gestión y Adquisición de IS 	<p>Cuatro líneas de trabajo:</p> <ol style="list-style-type: none"> 1) Competencias (genéricas y específicas de las áreas temáticas); 2) Enfoques de enseñanza, aprendizaje y evaluación de estas competencias; 3) Créditos académicos; 4) Calidad de los programas <p>Tres categorías del meta perfil:</p> <ol style="list-style-type: none"> 1) Ejercicio de la Profesión 2) Responsabilidad Social 3) Aspectos Disciplinarios 	<p>Se basa en 4 fundamentos: La metodología toma de referencia la norma ISO/IEC 19796 (ISO, 2005) para la creación de la metodología de cursos virtuales. Incluye también las prácticas de aprendizaje accesible propuestas por Seale (2007), contempla la formación del personal involucrado, Incorporará métricas para una posible auditoría con modelos de calidad.</p>

Tabla 1 Comparación de Metodologías

	ACM IS 2010	TUNING	ACAI
Metodología	No existe, sin embargo, propone un conjunto de áreas de aplicación, así como ejemplos de cursos núcleo y electivos.	<p>1.Revisión de las condiciones básicas</p> <p>2. Definición del perfil conducente al título.</p> <p>3. Descripción de los objetivos del programa y de los resultados de aprendizaje (en términos de conocimientos, comprensión, habilidades y capacidades) que deberán cumplirse.</p> <p>4. Identificación de las competencias genéricas y específicas de cada área que deberían alcanzarse en el programa.</p> <p>5. Traducción al plan de estudios: contenido (temas que habrán de cubrirse) y estructura (módulos y créditos).</p>	Norma ISO/IEC 19796 (ISO, 2005) identifica siete procesos relevantes dentro del ciclo de vida de los sistemas de información y comunicación utilizados para el aprendizaje, educación y formación. Los siete procesos son: (1) Análisis de necesidades; (2) Análisis del marco; (3) Concepción y Diseño; (4) Desarrollo y Producción; (5) Implementación; (6) Aprendizaje; (7) Evaluación y Optimización. Cada uno de los procesos contiene una serie de subprocesos para un total de 38 subprocesos.

Tabla 1 Comparación de Metodologías

	ACM IS 2010	TUNING	ACAI
Metodología		<p>6. Definición de los enfoques didácticos y de aprendizaje (tipos de métodos, técnicas y formatos) y de los métodos de evaluación (en caso necesario, desarrollando el material didáctico)</p> <p>7. Desarrollo de un sistema de evaluación concebido para incrementar de modo constante su calidad.</p>	

Tabla 1 Comparación de Metodologías

4.1.3. Selección de metodología curricular para el proyecto

Para la determinación del modelo curricular se ha seguido las directrices que indica el modelo Tuning, de las cuales se seleccionaron cuatro de las cinco propuestas por la metodología y se han resumido en los siguientes puntos.

- 1) Competencias genéricas y sub competencias.
- 2) Función de ECTS (European Credit Transfer System) como un sistema de acumulación.
- 3) Enfoques de aprendizaje, didácticos y evaluación.

Determinación de las competencias genéricas y sub competencias

Como primer punto para el diseño del programa académico se identificaron cuáles eran las competencias genéricas y las disciplinares o específicas, ello con el objetivo de elaborar el perfil del área de inteligencia de negocios y analítica de datos.

Sistema de acumulación de créditos académicos

La metodología Tuning utiliza el ECTS (European Credit Transfer System) para representar la carga de trabajo y los resultados de aprendizaje de un determinado curso; para este punto se tomará el Sistema de Unidades Valorativas que tiene la Universidad de El Salvador, a partir de la cantidad de cursos que conformará la especialización para poder cubrir con las competencias.

4.2. Herramientas para recolección de datos.

Para la identificación de las competencias se utilizaron tres herramientas para recopilar datos e información.

- 1) Encuesta dirigida a profesionales con experiencia en los campos afines.
- 2) Análisis de bolsas de trabajo nacional e internacional.
- 3) Análisis de cursos de inteligencia de negocios y analítica de datos de instituciones de enseñanza nacionales, regionales e internacionales.

En conjunto estas tres herramientas dieron los insumos para construir tales competencias.

4.2.1 Encuesta a profesionales en el área de Inteligencia de Negocios y Analítica de datos.

Se realizó una encuesta llamada “Sondeo del Perfil de Inteligencia de Negocios y Analítica de Datos”, la cual fue dirigida a profesionales que tuvieran experiencia en las áreas de inteligencia de negocios y/o analítica de datos, con el objetivo de obtener más insumos para la construcción del perfil.

La encuesta consistió en una serie de preguntas orientadas a profundizar sobre los conocimientos, habilidades y valores que un profesional debe tener, así como indagar acerca de las nuevas tendencias en el campo y las herramientas que se utilizan. Estuvo compuesta de preguntas tanto abiertas como cerradas, asegurando de esta manera que el profesional tuviera la libertad de dar a conocer su opinión sobre los temas y de no predisponer su pensamiento a una serie de respuestas (Ver Anexo 5). Las áreas de las encuestas fueron:

- Inteligencia de Negocios.
- Analítica de Datos.
- Minería de Datos.
- Machine Learning.

Preguntas Cerradas

Para el análisis de las preguntas cerradas se elaboró un cuadro resumen por cada pregunta donde se colocaron las opciones de la pregunta y se contabilizó su frecuencia como lo muestra el siguiente modelo.

Nombre de la pregunta	
Opciones	Frecuencia
...	...

Tabla 2 Análisis preguntas cerradas

Preguntas Abiertas

Para el análisis de las preguntas abiertas, de cada una de las respuestas proporcionadas por los encuestados se clasificó las palabras comunes como categorías, cada categoría fue puesta en una matriz y se contabilizaron las frecuencias.

Nombre de la pregunta	
Categoría	Frecuencia
...	...

Tabla 3 Análisis preguntas abiertas

Una vez clasificadas cada pregunta de la encuesta se procedió a escoger las opciones que fueron más representativas y colocarlas en una tabla con la siguiente estructura:

Área	
Elemento	
Conocimientos	...
Habilidades	...
Valores	...
Conocimientos de interés	...
Herramientas	...
Cursos	...

Tabla 4 Consolidación de categorías

El análisis de los resultados de este instrumento puede ser consultado en el Anexo 6.

4.2.2 Sílabos de instituciones educativas.

Con el objetivo de obtener más información para la construcción del perfil, se extrajeron áreas de conocimiento y competencias de cursos educativos impartidos por instituciones a nivel nacional, regional (latinoamericano) e internacional. Para la elección de las universidades a nivel internacional y regional se utilizó el sitio de Top Universities⁴⁹ en donde se clasifica a las universidades bajo los siguientes criterios: reputación académica, reputación de los empleadores, investigaciones realizadas por las instituciones y el h-índice, que es una métrica que mide la productividad y el impacto de un trabajo científico realizado en la institución. Se tomaron cuatro universidades con las mejores puntuaciones de cada región específicamente en el área de ciencias de la computación y sistemas de información y que tuvieran carreras o cursos afines a las áreas de Inteligencia de Negocios y Analítica de Datos. Para la elección de instituciones a nivel nacional se buscaron instituciones educativas que ofrecieran cursos relacionados al área.

Internacional

- Massachusetts Institute of Technology(MIT).
- University of Cambridge.
- Nanyang Technological University of Singapore (NTU).
- Carnegie Mellon University.

⁴⁹ Top Universities. (2017). agosto 20, 2017, de Top Universities Sitio web: <https://www.topuniversities.com>

Regional (Latinoamérica)

- Pontificia Universidad Católica de Chile (UC).
- Universidad Nacional Autónoma de México.
- Universidad Federal de Rio de Janeiro.
- Universidad de Buenos Aires.

Nacional

- Escuela Superior de Negocios (ESEN).
- Universidad Francisco Gavidia.
- Iseade Fepade.

Toda la información proveniente de los programas de estudios se resumió en una matriz para una mejor organización y resumen de los datos. (Ver Anexo 7).

País	Universidad o Institución Educativa	Carrera	Perfil de Conocimientos	Observaciones
...
...

Tabla 5 Matriz resumen

Posteriormente se analizaron los temas de cada curso, identificando los cursos que eran similares en contenido, colocando una X a la institución que impartiera el área. Como resultado se elaboró una matriz por cada región. (Ver Anexo 7).

Área	Institución 1	Institución 2	...
...		X	...
...	X		...

Tabla 6 Áreas de conocimiento comunes por institución

Una vez obtenidas las áreas que eran comunes para cada institución por región, se hizo un comparativo colocando cada área obtenida de la matriz y se colocaron todas las instituciones para contrastar y cruzar toda la información, se tomaron las frecuencias para obtener cuales eran las áreas más comunes de entre todas las regiones. (Ver Anexo 7).

Tema	Institución 1	Institución 2	...	Frecuencia
....	X	
...		X

Tabla 7 Comparativo de instituciones

4.2.3 Bolsas de trabajo nacional e internacional.

Para la recopilación de datos provenientes de las ofertas de trabajo se hizo uso de buscadores en línea de empleos de El Salvador, a saber: Computrabajo, Transdoc y Tecoloco en un periodo de quince días del 19 de junio de al 2 de julio de 2017. Se analizaron las ofertas de empleo relacionadas a las áreas de inteligencia de negocios y analítica de datos.

Asimismo, se tomaron ofertas provenientes de los mejores buscadores de empleo a nivel internacional tomando como referencia los sitios Forbes, Digital Trends y The Balance . De estos sitios se tomaron los buscadores: Indeed, Monster y LinkedIn en el mismo rango de fecha de las ofertas de empleo nacionales. Para la elección de las ofertas a nivel internacional, se utilizó el filtro proporcionado por el buscador para obtener las ofertas más relevantes, garantizando de esta manera que se considerara las mejores evaluadas de entre todas ofertas publicadas.

En cada oferta, se analizaron las funciones del puesto y los requisitos es decir los conocimientos y habilidades solicitados para el puesto. Para más detalles de las empresas y las solicitudes analizadas (Ver Anexo 8).

Para la elaboración de las competencias se tomó como insumo los resultados provenientes de las encuestas, las áreas de conocimiento obtenidas de los sílabos de las instituciones educativas y la información proporcionada por las ofertas de empleo. Se identificó para cada área las subcompetencias, conocimientos, habilidades y valores que tuvieran estrecha relación y con ello se formuló una competencia que sintetizara estos cuatro elementos colocándose la información en una tabla con la siguiente estructura, la cual obedece al formato utilizado por el Comité de Diseño Curricular en el proyecto de actualización de currícula de la EISI.

Competencia del área de conocimiento			
Sub competencia	Conocimiento	Habilidades	Valores

Tabla 8 Competencias por área de conocimiento

4.3. Perfil académico para el estudiante de especialización de Inteligencia de Negocios y Analítica de datos.

Con base a los análisis realizados se identificaron los conocimientos, habilidades y valores requeridos en un profesional en el área de Inteligencia de Negocios y Analítica de datos.

4.3.1. Objetivo general de la especialización

Formar estudiantes con conocimientos especializados en el área de Inteligencia de negocios y analítica de datos con la capacidad de proponer soluciones y mejoras estratégicas de los procesos del negocio de las organizaciones a nivel centroamericano.

4.3.2. Competencia general de la Especialización

Aplicar conocimientos matemáticos, estadísticos, informáticos y de negocios para crear soluciones de inteligencia de negocios y analítica de datos que ayuden a mejorar el desempeño y el crecimiento de las organizaciones proporcionando información confiable para una mejor toma de decisiones.

4.3.3. Sub-competencias de la Especialización

Competencia para el área de Inteligencia de Negocios Aplicar conocimientos de negocio y técnicos en el desarrollo de proyectos de inteligencia de negocios para potenciar el crecimiento y la eficiencia de las organizaciones.
Sub competencia Conceptualizar las actividades, procesos y elementos de negocio para proponer soluciones orientadas al cumplimiento de los objetivos organizacionales y estrategias del negocio.

Conocimiento	Habilidades	Valores
<ul style="list-style-type: none"> • Herramientas de diagnóstico • Gestión de Proyectos • Fundamentos de Contabilidad • Fundamentos de Finanzas • Fundamentos de gestión de inventarios • Modelado de Procesos de Negocios BPM • Metodología para la creación de indicadores de desempeño • Fundamentos del proceso general de ventas 	<ul style="list-style-type: none"> • Abstracción de conceptos y procesos • Crear y aplicar indicadores de desempeño • Coordinar proyectos y equipos • Comunicación oral y escrita • Negociación con usuarios • Trabajo en equipo 	<ul style="list-style-type: none"> • Ética • Liderazgo • Proactivo
<p>Sub competencia</p> <p>Construir soluciones óptimas para la integración de datos estructurados a través de la aplicación de metodologías y técnicas para la obtención de información lista para analizar.</p>		
Conocimiento	Habilidades	Valores
<ul style="list-style-type: none"> • Modelado dimensional y multidimensional • Lenguaje Procedimental de Base de Datos avanzado • Ciclo de vida de desarrollo de un Data Warehouse • ETL 	<ul style="list-style-type: none"> • Análisis de Requerimientos • Abstracción de requerimientos • Optimización de consultas • Optimización en la creación de código • Aplicación de técnicas de limpieza de datos. • Análisis de fuentes de información 	<ul style="list-style-type: none"> • Compromiso • Confidencialidad • Colaboración • Calidad • Integridad

<p>Sub competencia Diseñar reportes y visualizaciones para la interpretación de los datos provenientes de las áreas de negocio por medio de la aplicación de fundamentos estadísticos y métricas personalizadas como un insumo a la toma de decisiones.</p>		
Conocimiento	Habilidades	Valores
<ul style="list-style-type: none"> ● Estadística descriptiva ● Lenguaje Procedimental de Base de Datos avanzado ● Conocimientos de sistemas empresariales ● Indicadores de desempeño ● Técnica de construcción de visualizaciones 	<ul style="list-style-type: none"> ● Criterio de selección del tipo de visualización que mejor se adecue al análisis ● Interpretación lógica y numérica de los datos ● Organización y simplificación de los datos ● Optimización de consultas ● Crear métricas para procesos de negocios. 	<ul style="list-style-type: none"> ● Confiable ● Proactivo ● Creativo ● Previsor ● Excelencia ● Detallista ● Integro

<p>Competencia para el área de Analítica de datos Aplicar conocimientos técnicos que permitan realizar análisis descriptivos y predictivo a los datos para interpretarlos y diseñar estrategias que aumenten la competitividad de una organización.</p>
<p>Sub competencia Construir soluciones de analítica de Big Data para el descubrimiento de información valiosa a través del uso de tecnologías de procesamiento, almacenamiento, acceso y gestión de Big Data que sirvan como insumo a la toma de decisiones.</p>

Conocimiento	Habilidades	Valores
<ul style="list-style-type: none"> ● Fundamentos de lógica matemática ● Fundamentos de programación. ● Base de datos semiestructurados y no estructurados. ● Técnicas de visualización. ● Técnicas y entornos de procesamiento para Big Data. ● Lenguajes de programación orientados a la analítica. ● Arquitecturas de Big Data. 	<ul style="list-style-type: none"> ● Aplicar conocimiento estratégico para identificar oportunidades de mejora. ● Aplicar técnicas estadísticas para el descubrimiento de patrones ● Organizar y aplicar ciclo de vida de desarrollo de proyectos aplicados a analítica de Big Data ● Manejo de herramientas de Big Data. ● Creación de algoritmos. ● Trabajo en equipo 	<ul style="list-style-type: none"> ● Proactivo ● Detallista ● Liderazgo ● Integridad
<p>Sub competencia Identificar los fundamentos y técnicas estadísticas requeridas en minería de datos para su aplicación en modelos matemáticos.</p>		
Conocimientos	Habilidades	Valores
<ul style="list-style-type: none"> ● Análisis exploratorio de datos ● Tipos de análisis de datos ● Técnicas Estadísticas para el análisis de datos 	<ul style="list-style-type: none"> ● Agilidad en cálculos matemáticos ● Manejo de herramientas de análisis de datos y hojas de cálculo. ● Comunicación oral y escrita ● Trabajo en equipo 	<ul style="list-style-type: none"> ● Creativo ● Honesto ● Detallista ● Integro

Sub competencia Construir modelos de minería de datos para detectar patrones y tendencias existentes en los mismos con relaciones complejas mediante la aplicación de análisis matemáticos, estadísticos y modelos predictivos.		
Conocimientos	Habilidades	Valores
<ul style="list-style-type: none"> ● Análisis matemático ● Fundamentos de programación ● Lenguajes de programación orientados a la analítica ● Técnicas de limpieza de datos ● Estructuras de datos ● Análisis de agrupamiento ● Análisis de sentimiento ● Análisis de Varianza (ANOVA) ● Modelos predictivos ● Estándar de proyectos de minería de datos ● Visualización de datos 	<ul style="list-style-type: none"> ● Aplicar algoritmos de agrupamientos ● Aplicar modelos estadísticos ● Comunicación oral y escrita ● Diseño y creación de modelos de minería de datos ● Descubrimiento de patrones en los datos ● Creación de algoritmos ● Trabajo en equipo ● Heurística para la resolución de problemas 	<ul style="list-style-type: none"> ● Detallista ● Integro ● Confiable

4.4. Modelo Profesional

El profesional especializado en inteligencia de negocios y analítica de datos debe ser capaz de ofrecer soluciones informáticas basadas en el análisis de datos con el objetivo de proveer información útil para la toma de decisiones.

Debe ser capaz de emplear técnicas y metodologías que permitan el descubrimiento de información pertinente a fin de que las organizaciones mejoren su eficiencia y la calidad de sus procesos.

El profesional debe poder aplicar conocimientos estadísticos y matemáticos para realizar análisis de diagnóstico, descriptivos y predictivos que permitan tratar con los datos de una organización y de esta forma crear estrategias para ganar competitividad en el mercado.

Debe poder dar a conocer los hallazgos a través de visualizaciones, gráficos, reportes ejecutivos para los usuarios estratégicos y tácticos a fin de comunicar información de manera resumida, pero sin perder esencia. El desarrollo de sus actividades será realizado con responsabilidad, ética y calidad.

4.3.5. Descripción del Perfil del profesional en Inteligencia de Negocios y Analítica de datos

La especialización en el área de Inteligencia de negocios y analítica de datos aporta a la sociedad profesionales competentes que se desenvuelven satisfactoriamente en el área de la informática gracias a sus conocimientos en integración de datos, conceptos propios de negocio, formas de presentar información, análisis de Big Data y modelos predictivos; sustentando el accionar de dichos profesionales por medio de ética, integridad, confiabilidad, colaboración y pro actividad.

Propone soluciones óptimas para mejorar o solucionar los retos presentes en una empresa aplicando las metodologías y técnicas para encontrar información valiosa dentro de los datos; para ello se aplicará conocimientos fundamentales de negocios para realizar indicadores de desempeño especializados y a la medida, con el fin de contribuir a los objetivos estratégicos del negocio. Tiene la habilidad de preparar los datos en las estructuras adecuadas para su análisis; dependiendo de la complejidad de estos; además aplicará modelos matemáticos o algoritmos que buscarán patrones en los datos.

Se formarán profesionales integrales con conciencia en que la labor que realizan conlleva responsabilidad y confidencialidad al manejar uno de los insumos más importantes de las organizaciones.

4.3.6. Roles del Perfiles académicos de la especialización

Los roles que se determinaron para el desarrollo de esta especialización son:



Figura 7 Roles del perfil profesional

- **Analista de Negocios:**

Identifica las necesidades reales de la organización y propone soluciones e iniciativas de mejora en base al conocimiento que posee de los diferentes procesos del negocio.

- **Desarrollador de Data Warehouse:**

Se encarga de diseñar, construir y mantener el modelo multidimensional de acuerdo a los requerimientos solicitados por el analista de negocios, así mismo implementa la integración de los datos por medio de los procesos de ETL.

- **Analista de Inteligencia de Negocios:**

Se encarga de la recopilación de datos históricos y presentes; diseña indicadores para realizar análisis del negocio con la finalidad de facilitar la toma de decisiones tácticas y estratégicas más adecuadas para lograr los objetivos de la organización

- **Arquitecto de Big Data:**

Diseña la infraestructura de Big Data dado el conocimiento que posee en los componentes de plataformas y en las diferentes tecnologías relacionadas al área.

- **Analista de Big Data:**

Explora, integra y analiza fuentes de información estructurada y no estructurada para descubrir nuevas oportunidades de negocio al aplicar conocimientos de lenguajes de programación orientados al análisis.

- **Minero de Datos:**

Aplica técnicas para identificar patrones de comportamiento de los datos, con la finalidad de predecir escenarios que incrementen la competitividad en los negocios.

**CAPÍTULO V:
PROPUESTA DE
ESPECIALIZACIÓN DE
INTELIGENCIA DE
NEGOCIOS Y ANALITICA
DE DATOS**

5.1. Valorización de los Programas de Estudio propuestos

Con el instrumento “Encuesta de Valorización” se logró una retroalimentación de los componentes del programa para cada materia. Los participantes fueron profesionales expertos que laboran en puestos como: Jefe de gestión de la Información (DWH Senior), Analista de Data Warehouse, Data Scientist, Business Decision Support, Analista de BI, Jefe de Data Warehouse.

De los resultados obtenidos (Ver Anexo 9) se concluyen lo siguiente:

1. Objetivos

En general, las propuestas obtenidas de objetivos hacían referencia a adquirir conocimientos y habilidades que la especialización ya contempla en otras áreas de estudio predecesoras. Esto hace evidente que una materia es prerrequisito de otra, y que al cumplir una implícitamente se estará cumpliendo los objetivos anteriores.

2. Contenido del programa

Uno de los elementos principales de los programas son los contenidos que lo conforman. Como se pueden apreciar en las gráficas de cada programa, dichos contenidos han tenido una buena aceptación por parte de los encuestados, teniendo la categoría “Muy de acuerdo” un promedio ponderado del 70.4% para todos los programas en general.

Además de preguntar individualmente por cada tema que forma el programa, se quería conocer algún contenido o tema propuesto que los encuestados consideraban se debía añadir al programa.

En la mayoría de los casos los contenidos propuestos por los expertos estaban contemplados en otros programas, esto es un buen indicador porque señala una buena interrelación entre los programas y sus contenidos y que estos se cubren entre sí.

Otra de las observaciones con respecto al contenido de la materia fue agregar características que el estudiante debería lograr al cursar la materia. Por ejemplo: “Establecer comunicaciones con un equipo multidisciplinario para la implementación del proyecto de BI” para el programa de Business Intelligence.

Esto hace referencia a una competencia o habilidad que debe lograr el estudiante; dicha competencia o habilidad está tomada en cuenta en los perfiles, por lo tanto, es un hecho que confirma que se ha construido los programas teniendo en cuenta las habilidades y las competencias fomentadas en la primera etapa de este proyecto.

Se obtuvo observaciones de contenido, en las que el evaluador consideraba que el tema ya formaba parte de una materia del actual programa de estudios.

Para estos casos: Diseño de plantillas y SQL Avanzado, se decidió mantener la propuesta en la especialización puesto que el porcentaje de los evaluadores que estuvieron de acuerdo fue mayor y porque se tiene respaldo de las bibliografías (libros y bolsas de trabajo) que estas dos son áreas de conocimientos contemplados en inteligencia de negocios y Data Warehouse. Sin embargo, en aras de mantener la transparencia en los resultados se expresa en este análisis.

3. Metodología de Evaluación:

Se han tomado sugerencias con respecto a las metodologías de evaluación en los programas. Es importante contar con distintas formas de medir el conocimiento que el alumno ha aprendido a lo largo de los cursos, por ello se ha procedido a tomar en cuenta y mejorar el aspecto evaluativo de programas como BI al proponer implementar una solución de inteligencia de negocios por medio de una herramienta de visualización seleccionada por el estudiante que sea distinta a la impartida en la materia.

4. Bibliografía

Se obtuvieron dos propuestas de bibliografía para el área de Minería de datos, las cuales se han agregado en la parte de bibliografía complementaria de este programa. Esta bibliografía propuesta se sugerirá a los alumnos como fuente de consulta para la realización del caso de estudio (actividad de evaluación). De igual manera se agrega bibliografía para el área de negocios.

No se obtuvo observación de la bibliografía de los demás programas, por lo que no se ha realizado ninguna acción sobre los mismos.

Gráficos de los Resultados según Programa de Estudio

Programa Fundamentos de Negocios

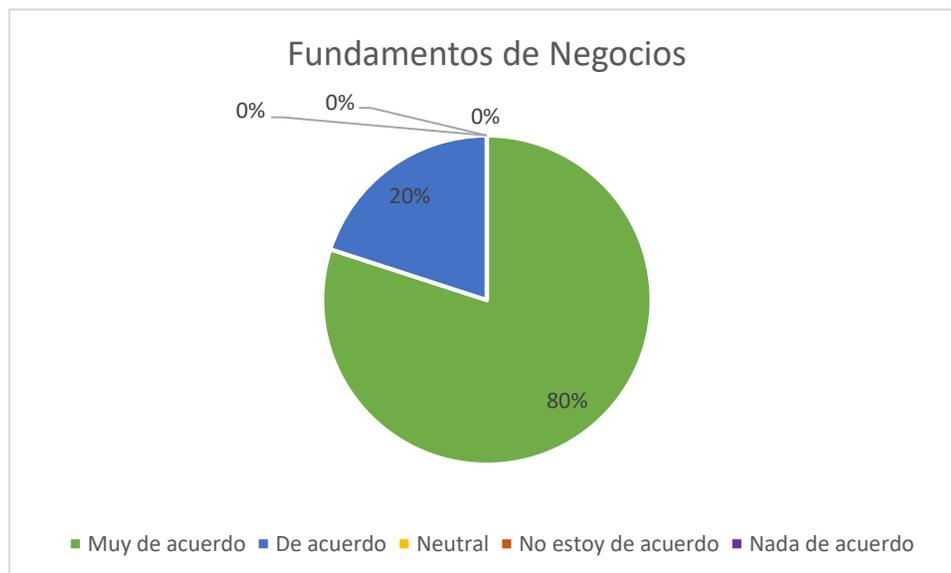


Gráfico 4 Instrumento de valorización para Fundamento de Negocio

Los contenidos, objetivos, metodología de evaluación y la bibliografía utilizada en el programa de Fundamentos de negocios se consideran aceptables desde el punto de vista de los encuestados al obtener un 80% en la mejor categoría dentro de la escala de evaluación (Muy de acuerdo).

Programa de Data Warehouse

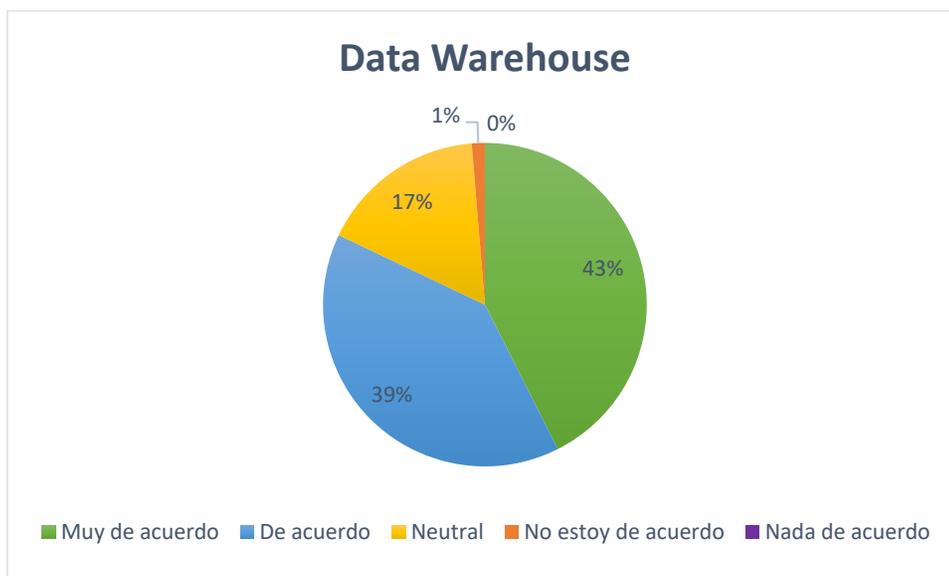


Gráfico 5 Instrumento de valorización para Data Warehouse

Los contenidos, objetivos, metodología de evaluación y la bibliografía utilizada en el programa de Data Warehouse se consideran aceptables desde el punto de vista de los encuestados al obtener un 82% en las dos mejores categorías dentro de la escala de evaluación (Muy de acuerdo y De acuerdo) y al obtener únicamente el 1% de los encuestados dentro de la escala de No estoy de acuerdo.

Programa de Inteligencia de Negocios

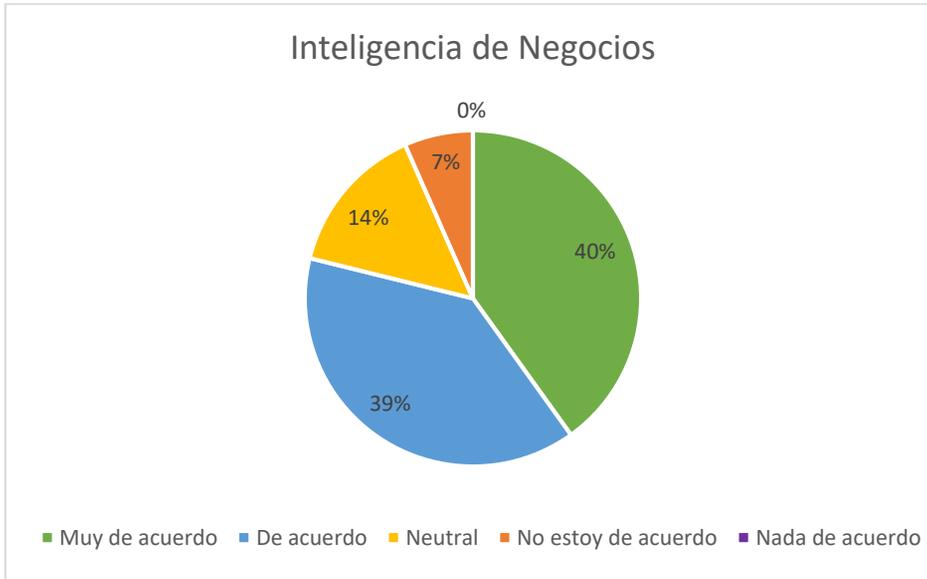


Gráfico 6 Instrumento de valoración para Inteligencia de Negocios

Los contenidos, objetivos, metodología de evaluación y la bibliografía utilizada en el programa de Inteligencia de Negocios se consideran aceptables desde el punto de vista de los encuestados al obtener un 79% en las dos mejores categorías dentro de la escala de evaluación (Muy de acuerdo y De acuerdo) y al obtener únicamente un 7% de los encuestados dentro de la escala de No estoy de acuerdo.

Programa de Gestión de Big Data

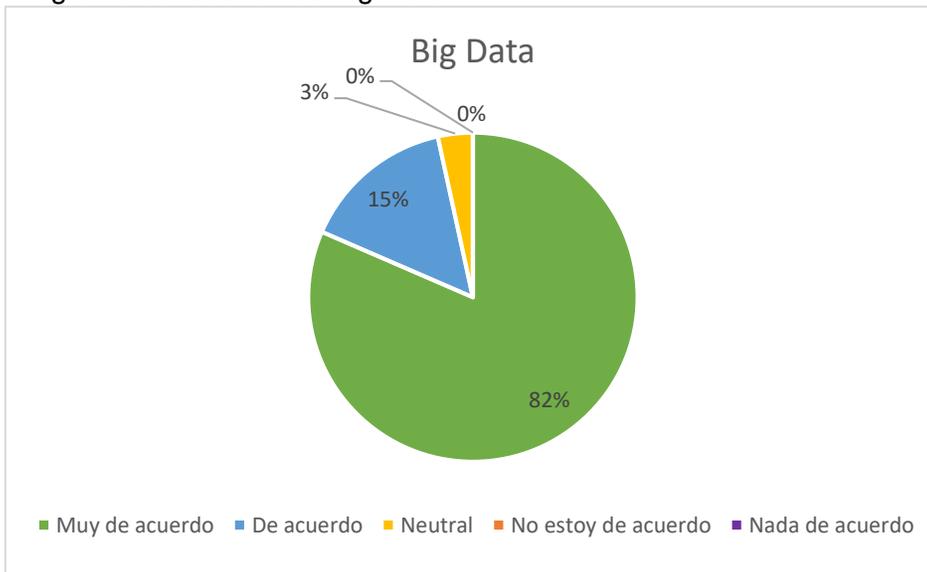


Gráfico 7 Instrumento de valoración para Gestión de Big Data

Los contenidos, objetivos, metodología de evaluación y la bibliografía utilizada en el programa de Gestión de Big Data se consideran aceptables desde el punto de vista de los encuestados al obtener un 82% en la mejor categoría dentro de la escala de evaluación (Muy de acuerdo).

Programa de Análisis Estadístico y Predictivo de Datos

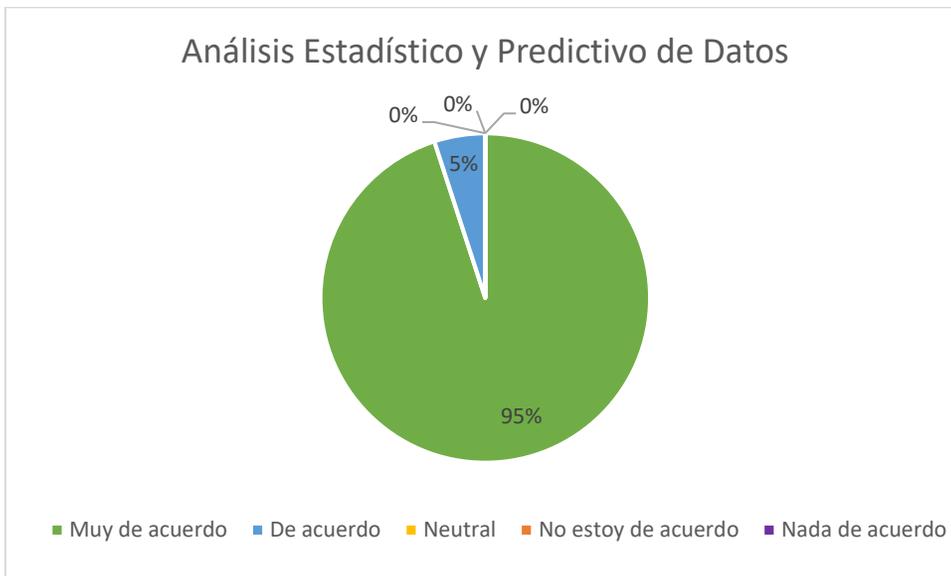


Gráfico 8 Instrumento de valoración para Análisis Predictivo de Datos

Los contenidos, objetivos, metodología de evaluación y la bibliografía utilizada en el programa de Análisis Estadístico y Predictivo de Datos se consideran aceptables desde el punto de vista de los encuestados al obtener un 95% en la mejor categoría dentro de la escala de evaluación (Muy de acuerdo).

5.2. Material Didáctico de la especialización

Se han desarrollado para cada programa:

1. **Guiones de clases:** desarrollo de la teoría de cada contenido del programa.
2. **Guías de laboratorio:** desarrollo de guías de ejemplos o ejercicios propuestos para los contenidos del programa que aplique un desarrollo práctico.

3. **Guías de discusiones:** desarrollo de guías de discusión en donde se presenten casos de estudio o guiones de lectura para aquellos contenidos del programa que aplique.
4. **Guías de instalación/configuración:** se proporciona las guías de instalación y configuración de las herramientas utilizadas en las guías de laboratorio.
5. **Presentaciones:** desarrollo de presentaciones con la información de los guiones de clases.

5.3. Sesiones

El contenido de los guiones de clase y las guías de laboratorio han sido organizados en sesiones, se compone de las siguientes partes.

Información general de la materia

- Nombre de la asignatura: Se asigna el nombre de la materia.
- Numero de sesión: Se especifica el número de la sesión.
- Catedrático: El docente catedrático de la materia.
- Fecha de sesión: Fecha de la sesión.
- Horario: Horario de la sesión.
- Lugar: Lugar donde se lleva a cabo la sesión.
- No. Participantes: El número de alumnos asignados a la clase.

Escenario inicial del curso

- Escenario inicial del curso: Contiene las actividades que dan introducción a la materia, como presentación del programa, evaluaciones, etc.
- Tiempo: Duración del escenario inicial.
- Recursos: Recursos necesarios para la introducción al curso.
- Responsable: Docente asignado.

Información de la sesión

- Objetivos de aprendizaje: Es el fin que se quiere alcanzar con el contenido propuesto.
- Contenidos y subcontenidos: Son los temas que se impartirán en la sesión.
- Estrategias metodológicas: Son las estrategias que se pretenden emplear para impartir los contenidos.
- Tiempo: La duración de la etapa.
- Recursos: Los recursos a utilizar.

- Evaluación: Para asegurar la comprensión del estudiante se especifica una forma de evaluación.

Insumos de sesiones

- Equipos y medios audiovisuales: Se especifican los materiales audiovisuales a ocupar durante el curso.

5.4. Programas de estudio

Los componentes que conforman un programa de estudio son:

I. Generalidades de la Asignatura

Se presenta en un cuadro resumen con la siguiente información: Nombre de la Asignatura, Código, Prerrequisito, Número de horas, Unidades valorativas, Ciclo, Número correlativo, Horas teóricas semanales, Horas prácticas semanales, Duración del ciclo (semanas), Duración horas clase (minutos) y Plan de estudios.

II. Fundamentación de la Asignatura

Explicación sobre la importancia y el propósito de la asignatura en la carrera de Ingeniería de Sistemas Informáticos, los prerrequisitos de la asignatura, y el perfil académico a alcanzar.

III. Objetivos

Los objetivos a alcanzar con cada una de las unidades propuestas en el programa.

IV. Contenidos

Tabla con el detalle de los contenidos de la asignatura separados por unidades.

V. Metodología

Explicación de las actividades académicas y sujetas a evaluación que se desarrollaran durante la duración de la asignatura.

VI. Evaluación

Detalle de los porcentajes para cada una de las actividades sujetas a evaluación.

VII. Bibliografía

Listado del material didáctico de la asignatura.

Se presentan a continuación el detalle de los programas para cada área de conocimiento propuesta.

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

PROGRAMA DE ASIGNATURA

I GENERALIDADES DE LA ASIGNATURA

Asignatura	Fundamentos de Negocios
Código	DS-FDN115
Prerrequisito	Herramientas de diagnóstico, Gestión de proyectos, Fundamentos de contabilidad y finanzas, Fundamentos de gestión de inventarios
Número de horas	64
Unidades valorativas	4
Ciclo	5
Número correlativo	TBD
Horas teóricas semanales	4
Horas prácticas semanales	2
Duración del ciclo (semanas)	16
Duración horas clase (minutos)	50
Plan de estudios	2018

II FUNDAMENTACIÓN DE LA ASIGNATURA

La asignatura de Fundamentos de Negocios está ubicada en el área de Desarrollo de Sistemas en el cuerpo del conocimiento de una especialización de Inteligencia de Negocios y Analítica de Datos, su objeto de estudio lo constituyen los conceptos de los sistemas de información, las estructuras de modelos de negocios, los tipos de sistemas empresariales y las herramientas de gestión estratégica.

La asignatura de Fundamentos de Negocios pretende desarrollar en el estudiante de Ingeniería de Sistemas Informáticos las habilidades y conocimientos para identificar las actividades y procesos que se realizan dentro de las organizaciones. Además, identificar las características y arquitectura de los principales tipos de sistemas informáticos usados en las organizaciones.

El curso es de suma importancia para el estudiante de Ingeniería de Sistemas Informáticos porque lo capacita en el uso de herramientas para el diseño y modelamiento de procesos de negocio, enlazados a las estrategias y objetivos en las organizaciones.

La asignatura de Fundamentos de Negocios tiene como prerequisites conocimientos en áreas de contabilidad, finanzas, ventas, mercadeo y recursos humanos que son las más comunes y elementales en las organizaciones. Además, requiere del manejo de conceptos básicos de administración.

III OBJETIVOS

1. Explicar las características y funciones que desempeñan los sistemas de información en los negocios para conocer los beneficios que aportan a las organizaciones.
2. Identificar los elementos esenciales y comunes en la estructura de los modelos de negocios para comprender la importancia de los mismos en el desempeño de las organizaciones.
3. Identificar los tipos de sistemas empresariales fundamentales presentes en las organizaciones para ofrecer soluciones estratégicas basadas en la información que estos gestionan.
4. Aplicar herramientas de gestión estratégica para proponer mejoras en procesos y estrategias de negocio.
5. Explicar la gestión de los datos/información de una organización a través del uso del programa de Gobierno de datos, con el fin de garantizar disponibilidad, integridad, usabilidad y seguridad de los mismos.

IV CONTENIDO

Unidad	Contenido
1.Sistemas de información y los negocios	1.1 El rol de los sistemas de información en los negocios actuales <ul style="list-style-type: none"> 1.1.1 La manera en que los sistemas de información están transformando los negocios 1.1.2 La empresa digital emergente 1.1.3 Objetivos estratégicos de negocios de los sistemas de información 1.2 Procesos de negocios y sistemas de información <ul style="list-style-type: none"> 1.2.1 Procesos de negocio 1.2.2 Cómo mejora la tecnología de información los procesos de negocios: eficiencia y transformación 1.3 Tipos de sistemas de información 1.4 Los sistemas desde una perspectiva funcional <ul style="list-style-type: none"> 1.4.1 Sistemas de ventas y mercadeo

Unidad	Contenido
1. Sistemas de información y los negocios	1.4.2 Sistemas de manufactura y producción 1.4.3 Sistemas financieros y contables 1.4.4 Sistemas de recursos humanos 1.5 Tipos de sistemas de información empresarial desde un punto de los usuarios 1.5.1 Sistema de procesamiento de transacciones 1.5.2 Sistemas de información gerencial y sistemas de apoyo a la toma de decisiones 1.5.3 Sistemas de apoyo a ejecutivos
2. Modelos de Negocio	2.1 Los Nueve Bloques de todo modelo de negocio 2.1.1 Segmentos de Mercado 2.1.2 Propuestas de valor 2.1.3 Canales 2.1.4 Relaciones con los clientes 2.1.5 Fuentes de ingresos 2.1.6 Recursos clave 2.1.7 Actividades clave 2.1.8 Asociaciones clave 2.1.9 Estructura de costes
3. Sistemas Empresariales	3.1 Sistemas de Planificación de Recursos Empresariales (ERP) ⁵⁰ 3.1.1 Definición 3.1.2 Características 3.1.3 Beneficios 3.2 Sistemas de Gestión de Relaciones con los Clientes (CRM) ⁵¹ 3.2.1 Definición 3.2.2 Características 3.2.3 Beneficios 3.3 Sistemas de Gestión de la Cadena de Suministro (SCM) ⁵² 3.3.1 Definición 3.3.2 Características 3.3.3 Beneficios 3.4 Sistemas de Soporte a las Decisiones (DSS) ⁵³ 3.4.1 Definición 3.4.2 Características 3.4.3 Beneficios

⁵⁰ ERP: Enterprise Resource Planning

⁵¹ CRM: Customer Relationship Management

⁵² SCM: Supply Chain Management

⁵³ DSS: Decision Support System

Unidad	Contenido
3. Sistemas Empresariales	3.5 Sistemas de Comercio Electrónico (E-Commerce) <ul style="list-style-type: none"> 3.5.1 Definición 3.5.2 El Crecimiento 3.5.3 Características Únicas 3.5.4 Mercados Digitales 3.5.5 Productos Digitales 3.5.6 Tipos de Comercio Electrónico 3.5.7 Modelos de Negocio en Internet 3.5.8 Modelos de Ingresos
4. Herramientas de gestión estratégica	4.1 Gestión de Procesos de Negocio (BPM) ⁵⁴ <ul style="list-style-type: none"> 4.1.1 Generalidades 4.1.2 Ciclo de vida de la Gestión de Procesos de Negocio 4.1.3 Notación de Gestión de Procesos de Negocio⁵⁵ 4.2 Evaluación Estratégica <ul style="list-style-type: none"> 4.2.1 Plan Estratégico 4.2.2 Beneficios de la Evaluación Estratégica 4.2.3 Situaciones que se apoyan de la Evaluación Estratégica 4.2.4 Mapa Estratégico: Las perspectivas de la Evaluación Estratégica 4.2.5 Elementos de la Evaluación Estratégica
5 Gobierno de datos	5.1 Introducción a Gobierno de datos. <ul style="list-style-type: none"> 5.1.1. Gestión de Datos 5.1.2. Gestión de Información Empresarial 5.1.3. Arquitectura de los datos (Información) 5.2. Conceptos relacionados a Gobierno de datos <ul style="list-style-type: none"> 5.2.1. Dominio de la Gestión de Datos 5.2.2. Calidad de los Datos 5.2.3. Inteligencia de Negocios 5.2.4. Principios 5.2.5. Políticas 5.2.6. Gestión de los activos de la información 5.3. Programa de Gobierno de datos <ul style="list-style-type: none"> 5.3.1 Alcances de un programa de Gobierno de datos <ul style="list-style-type: none"> 5.3.1.1 Modelo del Negocio 5.3.1.2 Contenido 5.3.1.3 Federación

⁵⁴ BPM: Business Process Management

⁵⁵ BPMN: Business Process Management Notation

Unidad	Contenido
5 Gobierno de datos	5.3.2 Elementos de un programa de Gobierno de datos <ul style="list-style-type: none"> 5.3.2.1 Organización 5.3.2.2 Principios 5.3.2.3 Políticas 5.3.2.4 Funciones 5.3.2.5 Métricas 5.3.2.6 Tecnologías y herramientas 5.4 El caso de negocio de un Gobierno de datos <ul style="list-style-type: none"> 5.4.1 Caso de negocio 5.4.2 Objetivos de un caso de negocio para un gobierno de datos <ul style="list-style-type: none"> 5.4.3 Contenido para un caso de negocio <ul style="list-style-type: none"> 5.4.3.1 Visión 5.4.3.2 Programas de riesgo 5.4.3.3 Alineación con el negocio 5.4.3.4 Costos de la calidad de datos 5.4.3.5 Costos de oportunidad 5.4.3.6 Obstáculos, Impacto y Cambios 5.4.3.7 Presentación del caso 5.5 Despliegue de Gobierno de datos <ul style="list-style-type: none"> 5.5.1 Alcance e iniciación 5.5.2 Evaluar 5.5.3 Visión 5.5.4 Alineación y valor del negocio 5.5.5 Diseño funcional 5.5.6 Diseño de un marco de trabajo 5.5.7 Hoja de ruta 5.5.8 Despliegue y sostenibilidad.

V METODOLOGÍA

- Se trabajará por medio de clases expositivas, examen parcial y laboratorios, un trabajo de investigación y desarrollo de casos prácticos en discusiones.
- Desarrollo de un proyecto en dos etapas.

VI EVALUACIÓN

A continuación, se determina las evaluaciones que se deben llevar a cabo y el peso correspondiente.

Evaluación	Peso específico en calificación final
Proyecto de Desarrollo	
Etapa 1	15%
Etapa 2	25%
Discusiones	30%
Examen Parcial	20%
Tarea de Investigación	10%

Detalle de las evaluaciones propuestas en el programa de estudios

Evaluación	Descripción
Proyecto de Desarrollo	<p>Realizar un trabajo de investigación aplicada en una empresa u organización sea de carácter pública o privada que consiste en dos etapas.</p> <p>Etapa I La primera etapa del proyecto comprende en la exploración de la situación actual de la empresa. El objetivo de esta etapa es elaborar un informe donde el estudiante pueda identificar la siguiente información:</p> <ul style="list-style-type: none"> ● Sistemas de información y empresariales presentes en la organización y que ayudan en los procesos que posee. ● Procesos que realizan dentro de la empresa. ● Bloques de modelo de negocio <p>Etapa II La segunda etapa del proyecto consiste en la elaboración de un documento con recomendaciones orientadas a mejorar la situación actual descrita en la primera etapa. Realizando sugerencias en base a los temas vistos en clase.</p>
Discusiones	Las discusiones presentan casos de estudio basados en situaciones reales para que el estudiante después de analizar dichos casos pueda responder a las interrogantes formuladas y desarrolle pensamiento crítico de las temáticas evaluadas.

Evaluación	Descripción
Examen parcial	Medir el conocimiento adquirido por el estudiante en el tema de Gobierno de datos, a través de una evaluación teórica de la última unidad de la materia, evaluando conceptos, los componentes de un programa de Gobierno de datos y el marco de trabajo propuesto por el Instituto de Gobernanza de datos.
Tarea de investigación	<p>Por equipos de trabajo, investigar acerca de temas de vanguardia relacionados con tendencias de Inteligencia de negocio y analítica de datos con el fin de que los estudiantes se familiaricen con dichas temáticas.</p> <p>Cada equipo de trabajo contará con un tema diferente el cual deberá investigar y sintetizar en un informe corto.</p> <p>Propuesta de Temas</p> <ul style="list-style-type: none"> • Tendencias de Inteligencia de Datos y los negocios • Data Lakes • Metodología Kanban • Tendencias de Big Data para el próximo año • Data mining • Internet de las cosas

VII BIBLIOGRAFÍA

- O'Brien, J & Marakas, G. (2006). Sistemas de Información gerencial. México, D.F.: McGraw-Hill.
- Laudon, K. & Laudon, J. (2008). Sistemas de información gerencial. México: Pearson Educación.
- Osterwalder, A., & Pigneur, Y. (2015). Generación de modelos de negocio. Barcelona: Deusto.
- Laudon, K., Laudon, J., & Vidal Romero Elizondo, A. (2012). Sistemas de información gerencial. México: Pearson Educación.
- Robert S. Kaplan. (2010). Conceptual Foundations of the Balanced Scorecard. Harvard Business School.

- Dumas, M., La Rosa, M., Mendling J & Reijers H.. (2013). Fundamentals of Business Process Management. Heidelberg: Springer.
- OMG. (2011). Business Process Model and Notation.
- Hitt, M., Ireland & R., Hoskisson R.. (2009). Strategic Management. Mason Ohio: South-Western.
- Soares, S.. (2014). Data Governance Tools. Boise, Idaho: MC PRes.
- Medina, M., Rivera, M. & Velarde S. . (2013). El Balanced Scorecard, Una Herramienta Para la Planeación Estratégica. octubre 8, 2017, de Itson Sitio web: http://www.itson.mx/publicaciones/pacioli/Documents/no66/17a-el_bsc_una_herramienta_para_la_planeacion_estrategicax.pdf
- Oracle. (2017). Enterprise Information Management: Best Practices in Data Governance. octubre 8, 2017, de Oracle Sitio web: https://iapp.org/media/pdf/knowledge_center/oea-best-practices-data-gov-400760.pdf
- Ladley, J.. (2012). Data Governance How to Design, Deploy and Sustain an Effective Data Governance Program. USA: Elsevier Inc.
- The DGI Data Governance Framework - The Data Governance Institute. (n.d). The Data Governance Institute. Noviembre 12, 2017, de http://www.datagovernance.com/wp-content/uploads/2014/11/dgi_framework.pdf
- Sampagar, V.. (noviembre 18, 2017). Data Warehouse Fact Constellation Schema and Design. noviembre 19,2017, de DWgeek Sitio web: <http://dwgeek.com/data-warehouse-fact-constellation-schema-design.html/>
- Rouse, M.. (2017). Definition Caso de negocio. noviembre 13,2017, de TechTarget Sitio web: <http://searchdatacenter.techtarget.com/es/definicion/Caso-de-negocio>

Bibliografía Complementaria

- Ritson, N. (2017). Strategic Managment (2da ed.). Bookboon.

Elaborado por	Juan Diego Zavaleta Mejía María José Flores Ramos Melvin Iván Reyes Ortiz Mónica Guadalupe Flores Bonilla	Fecha	02/03/2018
Revisado por	Elmer Arturo Carballo Ruiz MSc	Fecha	12/04/2018

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

PROGRAMA DE ASIGNATURA

I GENERALIDADES DE LA ASIGNATURA

Asignatura	Data Warehouse
Código	DS-FDD115
Prerrequisito	Bases de Datos, Fundamentos de Negocio
Número de horas	64
Unidades valorativas	4
Ciclo	6
Número correlativo	TBD
Horas teóricas semanales	4
Horas prácticas semanales	2
Duración del ciclo (semanas)	16
Duración horas clase (minutos)	50 minutos
Plan de estudios	2018

II FUNDAMENTACIÓN DE LA ASIGNATURA

La asignatura de Data Warehouse está ubicada en el área de Desarrollo de Sistemas en el cuerpo del conocimiento de una especialización de Inteligencia de Negocios y Analítica de Datos, su objeto de estudio lo constituyen los conceptos de modelado dimensional y multidimensional, limpieza de datos, diseño y construcción de procesos ETL y administración de metadatos.

La Asignatura de Data Warehouse tiene como propósito preparar al estudiante de Ingeniería de Sistemas Informáticos en los conocimientos y habilidades necesarias para diseñar y construir estructuras de datos dimensionales y multidimensionales, de acuerdo a las mejores prácticas que satisfagan las necesidades de análisis de información de un ambiente de negocios dado; junto con técnicas que permitan poblar de datos la estructura creada por medio de la gestión de sistemas ETL.

Adicionalmente se hará énfasis en la gestión de la metadata en las diferentes etapas del diseño de una Data Warehouse.

Para el perfil de un estudiante de Ingeniería de Sistemas Informáticos especializado en Inteligencia de negocios es importante que adquiera los conocimientos de diseño y construcción de Data Warehouse y cubos OLAP con el que garantice análisis de información efectivos para toma de decisiones ágiles.

Un estudiante de Ingeniería de Sistemas Informáticos debe poder diseñar una estructura de datos dimensional o multidimensional acorde a las necesidades de almacenamiento y generación de datos. Con el fin de permitir llegar al nivel de detalle requerido para la toma de decisiones; esto se logra con el diseño y construcción de un Data Warehouse o de un cubo OLAP que será capaz de soportar todos los posibles análisis de un área de negocio o del negocio en general; para mantener la información del Data Warehouse en sintonía con las transacciones que la originan, el estudiante deberá ser capaz de crear procesos ETL, que garanticen la consolidación de los datos que provienen de distintas fuentes en un solo lugar.

La asignatura tiene como prerrequisito los conocimientos de Bases de datos y Fundamentos de Negocios; pues es necesario conocer previamente de modelado entidad relación, algebra relacional, integridad referencial, lenguaje SQL, Normalización y aspectos de gestión de bases de datos, para aplicarlos en el modelamiento dimensional, multidimensional y procesos ETL.

III OBJETIVOS

1. Explicar las características, componentes y funciones principales para el diseño y modelado de un Data Warehouse.
2. Implementar el modelo multidimensional que mejor se adapte según los requerimientos dados.
3. Identificar el proceso, técnicas y medidas de limpieza, y conformación de datos que se requieren para asegurar la calidad de los mismos dentro de un Data Warehouse en la etapa previa a su carga.
4. Aplicar técnicas de carga de datos históricos o incrementales para las diferentes estructuras de un Data Warehouse.
5. Conceptualizar el término de metadata para aplicarlo a lo largo de las etapas de desarrollo de una Data Warehouse.

IV CONTENIDO

Unidad	Contenido
1. Fundamentos de Data Warehousing	1.1 Enfoques de Data Warehouse 1.2 Generalidades del Data Warehouse 1.3 Objetivos de un Data Warehouse 1.4 Arquitectura de una Data Warehouse 1.5 Data Warehouse y Data mart 1.6 Tabla de Dimensiones; Dimensiones Conformadas 1.7 Tablas de hechos: Agregados, Semi-Agregados y no Agregados, Llaves surrogadas 1.8 Técnica de Modelado <ul style="list-style-type: none"> 1.8.1 Esquema de Estrella 1.8.2 Copo de nieve 1.8.3 Constelación
2. OLAP	2.1 Concepto de OLAP 2.2 Reglas para Evaluar Soluciones OLAP 2.3 Características de los sistemas OLAP 2.4 Elementos 2.5 Funciones principales <ul style="list-style-type: none"> 2.5.1 Análisis dimensional y Multidimensional 2.5.2 Drill Down y Roll -up 2.5.3 Slice and Dice 2.6 Beneficios 2.7 Modelos OLAP <ul style="list-style-type: none"> 2.7.1 MOLAP 2.7.2 ROLAP 2.7.3 HOLAP 2.7.4 DOLAP 2.7.5 In Memory OLAP
3. Extracción y Transformación de datos	3.1 Generalidades 3.2 Data profiling (Perfilado de datos) 3.3 Mapa de datos lógico 3.4 Staging área (Datos temporales) <ul style="list-style-type: none"> 3.4.1 Diseño del staging area 3.4.2 Elementos del staging area 3.5 Estructuras de datos en un sistema ETL 3.6 Extracción de datos 3.7 SQL Avanzado <ul style="list-style-type: none"> 3.7.1 Funciones Analíticas 3.7.2 Agrupamiento de datos 3.7.3 Consultas Jerárquicas 3.7.4 Vistas Materializadas

Unidad	Contenido
4. Limpieza de datos	<p>4.1 Definición de Objetivos</p> <ul style="list-style-type: none"> 4.1.1 Roles para la definición de objetivos 4.1.2 Factores en competencia 4.1.3 Formulación de una política 4.1.4 Dimensión de Auditoría <p>4.2: Entregables de limpieza</p> <ul style="list-style-type: none"> 4.2.1 Resultados de perfilamiento de datos 4.2.2 Tabla de eventos de error 4.2.3 Dimensión de Auditoría <p>4.3: Pantallas (Screen) y sus medidas</p> <ul style="list-style-type: none"> 4.3.1 Fase de detección de anomalías <ul style="list-style-type: none"> 4.3.1.1 Data sampling 4.3.1.2 Tipos de aplicación 4.3.2 Diseño de las pantallas que conducen a la medición <ul style="list-style-type: none"> 4.3.2.1 Flujo de proceso general 4.3.2.2 Pantallas
5. Carga de datos	<p>5.1 Dimensiones de carga lenta</p> <ul style="list-style-type: none"> 5.1.1 Tipo 1 Sobrescribir 5.1.2 Tipo 2 Agregar nueva fila 5.1.3 Tipo 3 Agregar nuevo atributo 5.1.4 Tipo 4 Agregar Mini-Dimensión 5.1.5 Tipo 5 Agregar Mini-Dimensión y Outtrigger de Tipo 5.1.6 Tipo 6 Agregar atributos tipo 1 a la dimensión tipo 2 5.1.7 Tipo 7 Dimensiones dobles tipo 1 y tipo 2 <p>5.2 Manejo de índices</p> <ul style="list-style-type: none"> 5.2.1 Tipos de índices <p>5.3 Manejo de particiones</p> <ul style="list-style-type: none"> 5.3.1 Ejemplos de Particionado de tablas y creación de índices en Oracle <ul style="list-style-type: none"> 5.3.1.1 Creando Tablas Particionadas por Rango e Índices Globales 5.3.1.2 Creando índice global particionado por rango 5.3.1.3 Creando tablas particionadas por intervalo 5.3.1.4 Creando tablas particionadas por hash e índices globales 5.3.1.5 Creando tablas particionadas por lista 5.3.1.6 Creando tablas particionadas por

Unidad	Contenido
5. Carga de datos	<p style="text-align: center;">referencia</p> <p>5.4 Cargas incrementales</p> <p style="padding-left: 20px;">5.4.1 Cargas incrementales en tablas de dimensiones</p> <p style="padding-left: 20px;">5.4.2 Cargas incrementales en tablas de hechos</p> <p>5.5 Carga histórica de datos</p> <p style="padding-left: 20px;">5.5.1 Poblado de dimensiones con datos históricos</p> <p style="padding-left: 40px;">5.5.1.1 Poblar dimensiones de tipo 1</p> <p style="padding-left: 40px;">5.5.1.2 Poblar dimensiones de tipo 2</p> <p style="padding-left: 40px;">5.5.1.3 Poblar dimensiones de fechas y otras dimensiones estáticas</p> <p style="padding-left: 20px;">5.5.2 Poblado de tabla de hechos con datos históricos</p>
6. Gestión de sistemas ETL	<p>6.1 Programación y Soporte</p> <p style="padding-left: 20px;">6.1.1 Confiabilidad, disponibilidad, análisis de manejabilidad para ETL</p> <p style="padding-left: 20px;">6.1.2 Herramientas de Programación</p> <p style="padding-left: 40px;">6.1.2.1 Funcionalidad requerida de un planificador de ETL</p>
7. Metadata	<p>7.1 Definición de Metadata</p> <p style="padding-left: 20px;">7.1.1 ¿Qué es Metadata?</p> <p style="padding-left: 20px;">7.1.2 Sistema de Origen de los Metadatos</p> <p style="padding-left: 20px;">7.1.3 Front Room Metadata</p> <p>7.2 Metada del Negocio</p> <p style="padding-left: 20px;">7.2.1 Definición de Negocio</p> <p style="padding-left: 20px;">7.2.2 Información del Sistema Fuente</p> <p style="padding-left: 20px;">7.2.3 Diccionario de Datos del Almacén de Datos</p> <p style="padding-left: 20px;">7.2.4 Mapas de Datos Lógicos</p> <p>7.3 Metadata Técnica</p> <p style="padding-left: 20px;">7.3.1 Inventario del Sistema</p> <p style="padding-left: 20px;">7.3.2 Modelos de Datos</p> <p style="padding-left: 20px;">7.3.3 Definición de Datos</p> <p style="padding-left: 20px;">7.3.4 Reglas del Negocio</p> <p>7.4 Metadata Generada por ETL</p> <p style="padding-left: 20px;">7.4.1 ETL Job Metadata</p> <p style="padding-left: 20px;">7.4.2 Job</p> <p style="padding-left: 20px;">7.4.3 Transformación de Metadata</p> <p style="padding-left: 40px;">7.4.3.1 Nomenclatura de la Transformación</p> <p style="padding-left: 20px;">7.4.4 Lotes de Metadata</p> <p style="padding-left: 40px;">7.4.4.1 Load Schedule</p> <p>7.5 Proceso de ejecución</p> <p style="padding-left: 20px;">7.5.1 Ejecutar Resultados</p> <p style="padding-left: 20px;">7.5.2 Manejo de Excepciones</p> <p style="padding-left: 20px;">7.5.3 Lotes Programados (Batches Scheduled)</p>

V METODOLOGÍA

- Se trabajará por medio de clases expositivas, exámenes parciales, trabajos de investigación y desarrollo de casos prácticos en discusiones.
- Desarrollo de un proyecto

VI EVALUACIÓN

A continuación, se determina las evaluaciones que se deben llevar a cabo y el peso correspondiente.

Evaluación	Peso específico en calificación final
Examen Parcial I	10%
Examen Parcial II	20%
Proyecto de Desarrollo Etapa 1	15%
Etapa 2	25%
Laboratorios	30%

Detalle de las evaluaciones propuestas en el programa de estudios

Evaluación	Descripción
Examen Parcial I	Medir el conocimiento adquirido por el estudiante en las primeras cuatro unidades de la materia que corresponden a Fundamentos de Data warehousing, OLAP, ETL y Limpieza y Conformación de datos, por medio de una evaluación teórica y práctica. La parte teórica abarcaría las dos primeras unidades, y la parte práctica abordar la evaluación de perfilamientos de datos, técnicas de limpieza y las dimensiones conformadas.
Examen Parcial II	Medir el conocimiento adquirido por el estudiante de las unidades restantes de la materia que corresponden a Carga de datos, Gestión de sistemas ETL y Metadata, por medio de una evaluación teórica y práctica. La parte práctica evaluará la capacidad de creación de dimensiones SCD y la carga de datos histórica e incremental. Las unidades seis y siete serán evaluadas de manera teórica.

Evaluación	Descripción
Proyecto de desarrollo	<p>Etapa I Dado un caso de estudio en la primera etapa el alumno deberá identificar los requerimientos para realizar un modelamiento dimensional que aporte una solución de integración de datos (Data Warehouse) al caso propuesto el cual se verá reflejado en documento y en el desarrollo de uno de los datamarts.</p> <p>Etapa II Con base a la primera etapa el alumno deberá desarrollar un ETL que satisfaga la integración de datos necesaria para el datamart desarrollado.</p>
Laboratorios	Realización de los laboratorios propuestos para la unidad y entrega de los mismos para la revisión por el docente, en el periodo que se establezca.

VII BIBLIOGRAFÍA

- Díaz, J., & Caralt, J. (2011). Introducción al Business Intelligence. España: Editorial UOC.
- Inmon, W. (2002). Building the Data Warehouse; Third Edition (3rd ed.). New York: John Wiley.
- Kimball, R., & Ross, M. (2002). The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling; Second Edition (2nd ed.). New York: Wiley.
- Yessad, L., & Labiod, A. (2016). Comparative Study of Data Warehouses Modeling Approaches: Inmon, Kimball and Data Vault. International Conference On System Reliability And Science, 16, 95-99.
- Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., & Becker, B. (2011). The Data Warehouse Lifecycle Toolkit. Hoboken: John Wiley & Sons.
- Ponniah, P. (2001). Data Warehousing Fundamentals. Hoboken: Wiley-Interscience

- Caserta, J., & Kimball, R. (2004). The Data Warehouse ETL Toolkit: Practical Techniques for Extracting; Cleaning; Conforming; and Delivering Data. New York: Wiley.
- Oracle. (2010). Creating Partitions. diciembre 5,2017, de Oracle Sitio web: https://docs.oracle.com/cd/E18283_01/server.112/e16541/part_admin001.htm
- Microsoft. (2016). Indices. diciembre 1,2017, de Microsoft Sitio web: <https://docs.microsoft.com/es-es/sql/relational-databases/indexes/indexes>

Elaborado por	Juan Diego Zavaleta Mejía María José Flores Ramos Melvin Iván Reyes Ortiz Mónica Guadalupe Flores Bonilla	Fecha	02/03/2018
Revisado por	Elmer Arturo Carballo Ruiz MSc	Fecha	12/04/2018

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

PROGRAMA DE ASIGNATURA

I GENERALIDADES DE LA ASIGNATURA

Asignatura	Inteligencia de negocios
Código	DS-IN115
Prerrequisito	Probabilidad y estadística, Bases de Datos, Data Warehouse
Número de horas	64
Unidades valorativas	4
Ciclo	7
Número correlativo	TBD
Horas teóricas semanales	4
Horas prácticas semanales	2
Duración del ciclo (semanas)	16
Duración horas clase (minutos)	50
Plan de estudios	2018

II FUNDAMENTACIÓN DE LA ASIGNATURA

La asignatura de Inteligencia de Negocios está ubicada en el área de Desarrollo de Sistemas en el cuerpo del conocimiento de una especialización de Inteligencia de Negocios y Analítica de Datos, su objeto de estudio lo constituyen los conceptos esenciales de inteligencia de negocios, sus componentes, arquitectura y cómo a partir de los datos, que a través de la combinación de tecnología, herramientas y procesos se transforman en información almacenada y en conocimiento dirigido a un plan o estrategia organizacional.

La asignatura de Inteligencia de Negocios tiene como propósito esencial preparar al estudiante para desarrollar soluciones de nivel estratégico valiéndose de las diferentes fuentes de datos históricos y actuales que posee una organización a fin de convertirlos en información de utilidad para toma de decisiones y le permita optimizar la utilización de los recursos, monitorear el cumplimiento de los objetivos de la empresa para así obtener mejores resultados.

El curso tiene gran importancia para el estudiante de Ingeniería de Sistemas Informáticos porque lo prepara para hacer uso de los procesos de toma de decisiones, métricas e indicadores de negocio, repositorios de datos y herramientas de Inteligencia de negocios para conformar, diseñar y crear soluciones como reportes e informes que ayuden y apoyen a las áreas estratégicas y tácticas.

La asignatura tiene como prerrequisitos conocimientos de fundamentos de Negocios y Data Warehouse puesto que ambos conocimientos son aplicados de manera conjunta: la parte de negocios que define el tipo de información que se necesita y la parte de Data Warehouse que define el repositorio de los cuales se obtendrán los datos.

III OBJETIVOS

1. Identificar el papel que tiene la inteligencia de negocios en las organizaciones a través del conocimiento de su historia y proceso evolutivo.
2. Explicar los componentes de un sistema de inteligencia de negocios, su arquitectura, y la relación con el análisis de datos.
3. Conceptualizar los sistemas de soporte a las decisiones como un apoyo al desarrollo de aplicaciones de inteligencia de negocios.
4. Construir métricas de desempeño a través del empleo de metodologías.
5. Diseñar plantillas de informes de usuarios finales como un paso previo a la utilización de una herramienta de acceso y visualización de datos.
6. Usar herramientas de visualización de datos para presentar información resumida que sirva para el análisis de datos.

IV CONTENIDO

Unidad	Contenido
1. Generalidades de Inteligencia de Negocios	1.1 ¿Qué es Inteligencia de Negocios? 1.2 ¿Qué comprende Inteligencia de Negocios? 1.3 Historia y antecedentes.

Unidad	Contenido
1. Generalidades de Inteligencia de Negocios	1.4 Beneficios de Inteligencia de Negocios
2. Componentes de Inteligencia de Negocios	2.1 Decisiones efectivas y oportunas 2.2 Datos, información y conocimiento 2.3 El papel de los modelos matemáticos 2.4 Arquitectura de Inteligencia de Negocios <ul style="list-style-type: none"> 2.4.1 Ciclo de vida 2.4.2 Elementos facilitadores de proyectos 2.4.3 Desarrollo de un sistema Inteligencia de Negocios 2.5 La ética e Inteligencia de Negocios
3. Sistemas de soporte a las decisiones	3.1 Representación del proceso para resolver problemas 3.2 Tipos de decisión 3.3 Evolución de los sistemas de información 3.4 Definición de sistemas de soporte a las decisiones
4. Indicadores de desempeño	4.1 Generalidades 4.2 Metodología para el establecimiento de indicadores
5. Aplicaciones de Inteligencia de Negocios	5.1 Qué es una plantilla de aplicación de usuario final 5.2 Rol de la aplicación del usuario final 5.3 Modelos de aplicación: planificación, presupuesto, previsión y qué pasaría si (What if) 5.4 Especificación de aplicaciones <ul style="list-style-type: none"> 5.4.1 Determinar el conjunto de plantillas inicial 5.4.2 Diseño del enfoque de navegación de plantillas 5.4.3 Determinar estándares de plantilla 5.4.4 Especificación detallada de la plantilla 5.5 Desarrollo de aplicaciones de usuario final 5.6 Mantenimiento de plantilla de aplicación 5.7 Papeles claves

Unidad	Contenido
6. Inteligencia de Negocios con una herramienta de visualización	6.1 Conceptos clave de la herramienta de visualización

V METODOLOGÍA

- Se trabajará por medio de clases expositivas, exámenes parciales y laboratorios, un trabajo de investigación y desarrollo de una tarea de aplicación.

VI EVALUACIÓN

A continuación, se determina las evaluaciones que se deben llevar a cabo y el peso correspondiente.

Evaluación	Peso específico en calificación final
Tarea de Investigación	5%
Laboratorios	10%
Examen Parcial 1	10%
Examen Parcial 2	10%
Examen Parcial Práctico	25%
Tarea de Investigación/Aplicación	
Etapa 1 Investigativa	15%
Etapa 2 Aplicación	25%

Detalle de las evaluaciones propuestas en el programa de estudios

Evaluación	Descripción
Tarea de investigación	El estudiante deberá investigar sobre las tendencias en el área de inteligencia de negocios y deberá presentar un reporte con los resultados de la investigación.
Laboratorios	Realización de los laboratorios propuestos para la unidad y entrega de los mismos para la revisión por el docente, en el periodo que se establezca.

Evaluación	Descripción
Examen Parcial I y II	Medir el conocimiento adquirido por los estudiantes, el examen parcial I cubrirá los temas de las unidades 1 y 2; El segundo parcial cubrirá los temas del resto de unidades.
Examen Parcial Practico	Medir el conocimiento adquirido por el estudiante para desarrollar soluciones de inteligencia de negocios a través de la herramienta de visualización, generando reportes y las visualizaciones de los datos.
Tarea de Investigación/Aplicación	<p data-bbox="740 611 834 644">Etapa I</p> <p data-bbox="740 678 1382 940">El estudiante deberá investigar tres herramientas de inteligencia de negocios diferentes a la utilizada en los laboratorios de la materia y deberá presentar un reporte con los resultados de la investigación: generalidades de la herramienta, ventajas y desventajas, un ejemplo de uso de la herramienta y un cuadro comparativo de las herramientas investigadas.</p> <p data-bbox="740 947 1382 1094">Además; deberá seleccionar cual herramienta usará en la segunda etapa de la investigación anexando en el reporte una justificación del porque utilizará esa herramienta.</p> <p data-bbox="740 1140 842 1173">Etapa II</p> <p data-bbox="740 1207 1382 1507">El estudiante deberá plantear una situación ficticia o bien algún caso real en el cual se requiere de una solución de inteligencia de negocios. Deberá presentar la propuesta de solución: análisis de la situación, métricas a crear y reportes a generar (plantillas o cuadros de mando). Finalmente deberá presentar la solución creada a través de un cuadro de mando en la herramienta seleccionada en la etapa 1.</p>

VII BIBLIOGRAFÍA

- Vercellis, C.. (2009). Business Intelligence Aplicaciones. En Business Intelligence : Data Mining and Optimization for Decision Making(417). United Kingdom: John Wiley & Sons Ltd.

- Kimball, R., Ross, M., Thornthwaite, W., Mundy, J. and Becker, B. (2011). The Data Warehouse Lifecycle Toolkit. Hoboken: John Wiley & Sons.
- Paramenter, D. (2007). Key Performance Indicators. New Jersey: John Wiley & Sons, Inc
- Murray, D. (2016). Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software, 2n. John Wiley & Sons.
- Lamia, Y., & Aissa, L.. (2016, diciembre 31). Comparative Study of Data Warehouses Modeling Approaches. International Conference on System Reliability and Science, 1, 5.
- Rasmey, H., & Raghul, M.. (2016, diciembre 31). Literature Review of Business Intelligence. School of Business and Engineering Halmstad University, Sweden, 1, 10.
- Sinnexus. (2016). ¿Qué es Business Intelligence?. noviembre 30, 2017, de Sinnexus Sitio web: http://www.sinnexus.com/business_intelligence/
- Time Manager. (julio 31, 2017). Breve Historia de Business Intelligence. noviembre 30, 2017, de Time Manager Sitio web: <https://www.timemanagerweb.com/2017/01/31/breve-historia-del-business-intelligence/>
- Demilia, Brian; Peded, Michael; Jorgensen, Kenneth; and Subramanian, Ramesh (2012) "The Ethics of BI with Private and Public Entities," Communications of the IIMA: Vol. 12: Iss. 2, Article 2. Available at: <http://scholarworks.lib.csusb.edu/ciima/vol12/iss2/2>

Elaborado por	Juan Diego Zavaleta Mejía María José Flores Ramos Melvin Iván Reyes Ortiz Mónica Guadalupe Flores Bonilla	Fecha	02/03/2018
Revisado por	Elmer Arturo Carballo Ruiz MSc	Fecha	12/04/2018

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

PROGRAMA DE ASIGNATURA

I GENERALIDADES DE LA ASIGNATURA

Asignatura	Gestión de Big Data
Código	DS-GBD115
Prerrequisito	Inteligencia de Negocios
Número de horas	64
Unidades valorativas	6
Ciclo	8
Número correlativo	TBD
Horas teóricas semanales	4
Horas prácticas semanales	2
Duración del ciclo (semanas)	16
Duración horas clase (minutos)	50
Plan de estudios	2018

II FUNDAMENTACIÓN DE LA ASIGNATURA

La asignatura Gestión de Big Data ubicada en el en el área de Desarrollo de Sistemas como parte del cuerpo de conocimiento de una especialización de Inteligencia de Negocios y Analítica de Datos, pretende proveer los fundamentos enfocados al almacenamiento, procesamiento y análisis de grandes colecciones de datos no relacionados, generados recurrentemente por fuentes dispares e incluyendo las técnicas que aprovechan recursos computacionales.

Es importante que el estudiante se especialice en esta área de la informática para que a través de las capacidades adquiridas para gestionar y analizar grandes volúmenes de datos pueda extraer valor de todas las fuentes disponibles en una empresa y así impactar positivamente.

El estudiante especializado en Big Data deberá cumplir con dos perfiles importantes dentro del área: ser arquitecto y analista. Tendrá la capacidad de definir una estructura de Big Data en función de las necesidades y los recursos disponibles en una empresa.

Así mismo, podrá manejar los datos de diversas estructuras en tiempo real y obtener de ellos información a través del uso de tecnologías de procesamiento de grandes volúmenes de datos y de herramientas de visualizaciones.

El estudiante deberá poseer conocimientos de negocio, los conocimientos técnicos para el tratamiento de datos dentro un Data Warehouse y las habilidades para la presentación de los datos por medio de visualizaciones que apoyen al negocio. Además, la asignatura de Gestión de Big Data tiene como prerrequisitos el conocimiento de por lo menos algún lenguaje de programación orientado a objetos; así como una sólida base en SQL, modelado y diseño de bases de datos relacionales. Se recomienda que el estudiante esté familiarizado con el manejo de sentencias por consola de sistemas distribuidos libres para agilizar el aprendizaje de los contenidos propuestos.

III OBJETIVOS

1. Conceptualizar los fundamentos de Big Data que permita al estudiante el conocimiento de la existencia y de la importancia de Big Data en las empresas.
2. Analizar las diferentes estructuras de almacenamiento en memoria y en disco para que el estudiante decida cuál utilizar en función de las características de los datos presentes en una organización.
3. Examinar los tipos de procesamiento de Big Data para facilitar la comprensión de las tecnologías (frameworks de Big data que serán abordadas en la unidad).
4. Gestionar datos de diferente estructura a través de algunos componentes de los ecosistemas disponibles para Big Data.
5. Usar técnicas y herramientas de visualización de Big Data para presentar gráficamente los datos y que sean de provecho en su interpretación.

IV CONTENIDO

Unidad	Contenido
1. Fundamentos de Big Data	1.1 Características de Big Data 1.1.1 Definición de las V's de Big Data 1.1.2 Tipos de datos

	<ul style="list-style-type: none"> 1.1.2.1 Datos estructurados 1.1.2.2 Datos semi estructurados
Unidad	Contenido
1. Fundamentos de Big Data	<ul style="list-style-type: none"> 1.1.2.3 Datos no estructurados 1.1.2.4 Metadatos 1.2. Big Data, Open Data y Open Government <ul style="list-style-type: none"> 1.2.1 Definición de Open Data y Open Government 1.2.2 Principios de Open Data 1.2.3 Desafíos de Open Data 1.2.4 Interrelación entre big data, open data y open government
	<ul style="list-style-type: none"> 1.3 Data Lakes <ul style="list-style-type: none"> 1.3.1 Data lake y Data Warehouse 1.3.2 Impacto de Data lakes en una empresa 1.3.3 Pasos para el diseño de un Data Lake 1.3.4 Beneficios para el negocio de contar con un lago de datos
2. Almacenamiento y Procesamiento de Big Data	<ul style="list-style-type: none"> 2.1 Dispositivos de Almacenamiento en Disco <ul style="list-style-type: none"> 2.1.1 Sistemas de Archivos Distribuidos 2.1.2 Bases de Datos RDBMS 2.1.3 NoSQL 2.1.4 NewSQL 2.2 Dispositivos de Almacenamiento en Memoria <ul style="list-style-type: none"> 2.2.1 In-Memory Data Grid (IMDG) 2.2.2 In-Memory Database (IMDB) 2.3 Procesamiento de Big Data <ul style="list-style-type: none"> 2.3.1 Procesamiento en paralelo 2.3.2 Procesamiento Distribuido 2.3.3 Workloads <ul style="list-style-type: none"> 2.3.3.1 Modo Batch <ul style="list-style-type: none"> 2.3.3.1.1 MapReduce 2.3.3.2 Modo Tiempo Real 2.3.4 Clúster
3. Arquitecturas y Frameworks de Big Data	<ul style="list-style-type: none"> 3.1 OldSQL, NewSQL y NoSQL 3.2 Arquitectura Scale Out (Escalamiento horizontal) <ul style="list-style-type: none"> 3.2.1 Almacenamiento Key-Value (Clave-valor) 3.2.2 Tablas hash distribuidas 3.2.3 Almacenamiento de Documentos 3.2.4 Almacenamiento de familia de columnas 3.2.5 Almacenamiento de Grafo 3.3 Ecosistemas de Big Data: Arquitectura, componentes y funcionamiento. 3.4 Cloud Computing <ul style="list-style-type: none"> 3.4.1 Características de Cloud Computing

	3.4.2 Capas de cloud computing 3.4.2.1 IaaS: Infrastructure as a Service
Unidad	Contenido
3. Arquitecturas y Frameworks de Big Data	3.4.2.2 PaaS: Platform as a Service 3.4.2.3 SaaS: Software as a Service 3.4.3 Tecnologías de Habilitación 3.4.3.1 Virtualización 3.4.3.2 Servicio web y arquitectura orientada a servicios 3.4.3.3 Flujo de servicio y flujos de trabajo 3.4.3.4 Web 2.0 y Mashup
4. Visualización de Big Data	4.1 Análisis Visual 4.1.1 Mapas de calor 4.1.2 Series de tiempo 4.1.3 Diagrama de nodos 4.1.4 Mapeo espacial de datos 4.2 Herramientas de visualización de Big Data

V METODOLOGÍA

- Se trabajará por medio de clases expositivas, exámenes parciales, trabajos de investigación y desarrollo de laboratorios.
- Desarrollo de un proyecto

VI EVALUACIÓN

A continuación, se determina las evaluaciones que se deben llevar a cabo y el peso correspondiente.

Evaluación	Peso específico en calificación final
Laboratorios	30%
Examen Parcial 1	10%
Examen Parcial 2	15%
Examen Parcial Práctico	20%
Tarea de Investigación / Aplicación	25%

Detalle de las evaluaciones propuestas en el programa de estudios

Evaluación	Descripción
Laboratorios	Realización de los laboratorios propuestos para la unidad y entrega de los mismos para la revisión por el docente, en el periodo que se establezca.
Evaluación	Descripción
Examen Parcial 1 y 2	Medir el conocimiento adquirido por los estudiantes, el examen parcial uno cubrirá los temas de las unidades uno y dos; el segundo parcial cubrirá los temas del resto de unidades
Examen Parcial Practico	Medir el conocimiento adquirido por los estudiantes en el procesamiento de big data.
Tarea de Investigación/ Aplicación	<p>Etapa I</p> <p>Los alumnos deberán investigar y seleccionar una fuente de datos semi estructurada, la cual deberán explorar y preparar para poder alimentar un ecosistema big data; así como automatizar dichos procesos con una herramienta con el cual realizarán análisis posteriormente.</p> <p>Etapa II</p> <p>Los estudiantes deberán analizar y transformar los datos aplicando herramientas diferentes a las estudiadas en la materia, así como realizar una presentación con visualizaciones de los datos analizados.</p>

VII BIBLIOGRAFÍA

- Erl,T.,Khattak,W., & Buhler, P.. (2015). Big Data Fundamentals Concepts, Drivers & Techniques. Estados Unidos: Prentice Hall.
- Mohanty, S., Jagadeesh, M. and Srivatsa, H. (2013). Big data imperatives. New York: Apress..
- Science International (2015): Open Data in a Big Data World. Paris: International Council for Science (ICSU), International Social Science Council (ISSC), The World Academy of Sciences (TWAS), InterAcademy Partnership (IAP).
- Open Knowledge International. (2017). Open Data Handbook. noviembre7,2017, de OpenGovernmentData.org Sitio web: <http://opendatahandbook.org/guide/en/datos.gob.es>.

- (2017). Datos abiertos, big data y gobierno abierto: diferentes tipos de datos. noviembre 7,2017, de datos.gob.es Sitio web: <http://datos.gob.es/es/noticia/datos-abiertos-big-data-y-gobierno-abierto-diferentes-tipos-de-datos>
- PowerData. (2017, noviembre 13). Data Lake. Superando las limitaciones del Data Warehouse. PowerData, 1, 16.
- Handbook of cloud computing. (2014). [Place of publication not identified]: Springer.
- Karau, H. (2015). Learning Spark. Beijing ; Köln[u.a.]: O'Reilly.
- Aven, J. (2017). Hadoop in 24 hours, sams teach yourself. Indianapolis, Indiana: Sams.
- Apache Software Foundation . (2017). HDFS Architecture. noviembre 1,2017, de Apache Software Foundation Sitio web: <https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/HdfsDesign.html>
- Apache Software Foundation . (2017). YARN Architecture. noviembre 1,2017, de Apache Software Foundation Sitio web: <https://hadoop.apache.org/docs/current/hadoop-yarn/hadoop-yarn-site/YARN.html>
- Apache Software Foundation . (2017). Apache Hive. noviembre 10,2017, de Apache Software Foundation Sitio web: <https://cwiki.apache.org/confluence/display/Hive/Home#Home-UserDocumentation>
- Holmes, M.. (2013). Simple hive 'cheat sheet' for sql users. noviembre 2017, de Hortonworks Sitio web: <https://es.hortonworks.com/blog/hive-cheat-sheet-for-sql-users/>

Elaborado por	Juan Diego Zavaleta Mejía María José Flores Ramos Melvin Iván Reyes Ortiz Mónica Guadalupe Flores Bonilla	Fecha	02/03/2018
Revisado por	Elmer Arturo Carballo Ruiz MSc	Fecha	12/04/2018

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

PROGRAMA DE ASIGNATURA

I GENERALIDADES DE LA ASIGNATURA

Asignatura	Análisis Estadístico y Predictivo de Datos
Código	DS-FMD115
Prerrequisito	Gestión de Big Data
Número de horas	64
Unidades valorativas	4
Ciclo	9
Número correlativo	TBD
Horas teóricas semanales	4
Horas prácticas semanales	2
Duración del ciclo (semanas)	16
Duración horas clase (minutos)	50
Plan de estudios	2018

II FUNDAMENTACIÓN DE LA ASIGNATURA

La asignatura Análisis Estadístico y Predictivo de Datos, ubicada en el en el área de Desarrollo de Sistemas como parte del cuerpo de conocimiento de una especialización de Inteligencia de Negocios y Analítica de Datos, pretende desarrollar las habilidades y conocimientos de estudio de comportamientos de datos por medios de modelos matemáticos para expresar relaciones entre variables, parámetros, entidades; medidas de análisis para su exploración y predicciones de hechos; y el uso de un lenguaje estadístico siendo una herramienta que facilita y optimiza el tiempo para la representación y análisis de los datos.

Con esta asignatura el estudiante deberá ser capaz de analizar datos utilizando métodos o técnicas matemáticas, para interpretarlos, encontrar patrones que ayuden a pronosticar para una mejor toma de decisiones estratégicas y tácticas con precisión, integridad y confiabilidad.

La asignatura de Análisis Estadístico y Predictivo de Datos tiene como prerrequisitos los conocimientos de probabilidad y estadística, programación, métodos probabilísticos y análisis numérico.

Es de mucha importancia que el estudiante tenga la capacidad necesaria de lógica matemática, que conozca el manejo de arreglos y tenga la habilidad de modelar y representar por medio de variables matemáticas hechos o sucesos. Además, deberá cumplir con los roles de un analista de negocios, desarrollador de Data Warehouse, analista de inteligencia de negocios, arquitecto y analista de big data.

III OBJETIVOS

1. Identificar las clases de modelos matemáticos que mejor se adapten y representen los datos para su análisis e interpretación.
2. Utilizar lenguajes de programación orientado al análisis estadístico para dar un tratamiento adecuado a los datos.
3. Conceptualizar las metodologías de minería de datos para que el estudiante seleccione la más adecuada según el tipo de análisis a realizar.
4. Aplicar técnicas de minería a datos estructurados y no estructurados para identificar patrones.

IV CONTENIDO

Unidad	Contenido
1. Modelos Matemáticos para la toma de decisiones	1.1 Modelos Estocásticos 1.2 Modelos Determinísticos 1.3 Clases de Modelos 1.3.1 Modelos Predictivos 1.3.2 Modelos Descriptivos 1.3.3 Modelos de Diagnósticos 1.3.4 Modelos de Reconocimiento de Patrones 1.3.5 Modelos de Optimización 1.3.6 Modelos de Gestión de Proyectos 1.3.7 Modelos de Análisis de Riesgos 1.3.8 Modelos de Líneas de Espera
2. Lenguajes de Análisis Estadístico	2.1 Introducción 2.2 Objetos y Tipos de Datos 2.3 Subconjuntos de Datos 2.4 Lectura y Escritura de Datos 2.5 Funciones

Unidad	Contenido
2. Lenguajes de Análisis Estadístico	2.6 Creación de Gráficas
3. Exploración de Datos	3.1 Análisis Bivariado 3.1.1 Medidas de Correlación 3.2 Análisis Multivariado 3.2.1 Medidas de Correlación
4. Introducción a la Minería de Datos	4.1 Definición de Minería de Datos 4.2 Proceso de Minería de Datos 4.3 Metodología de Análisis por Minería de Datos 4.3.1 Caracterización y Discriminación 4.3.2 Clasificación 4.3.3 Regresión 4.3.4 Análisis de Series Temporales 4.3.5 Reglas de Asociación 4.3.6 Agrupación 4.3.7 Descripción y Visualización 4.3.8 Árboles de Decisión

V METODOLOGÍA

- La Se trabajará por medio de clases expositivas, exámenes parciales, tarea de investigación y desarrollo de laboratorios.
- Desarrollo de un proyecto

VI EVALUACIÓN

A continuación, se determina las evaluaciones que se deben llevar a cabo y el peso correspondiente.

Evaluación	Peso específico en calificación final
Tarea de Investigación	10%
Desarrollo de un Caso de Estudio Etapa 1	15%
Desarrollo de un Caso de Estudio Etapa 2	25%
Laboratorios	30%
Examen Parcial	20%

Detalle de las Evaluaciones propuestas en el programa de estudio

Evaluación	Descripción
Tarea de Investigación	Realizar un paper sobre herramientas de minería de datos y su aplicación en las técnicas de análisis. El docente deberá proponer las herramientas y se rifará entre los alumnos.
Desarrollo de Caso de Estudio	<p>Etapa I</p> <p>Presentar un caso de estudio real (o ficticio en su defecto) con requerimiento. En esta etapa deberá presentarse todo el ecosistema del problema: rubro de la empresa, antecedentes y descripción del problema, sistemas empresariales y otras fuentes de datos de la empresa, procesos de negocios involucrados en el problema. Además, en esta etapa deberá plantearse la solución de minería de datos y los beneficios que obtendrá el negocio al implementar la solución.</p> <p>Etapa II</p> <p>Se presentará la solución al problema planteado en la etapa uno. El estudiante deberá aplicar la técnica de minería que permita la solución. Deberá presentar las guías de instalación de la herramienta y la descripción de las funciones de la herramienta y/o pasos con que se llegó a la solución.</p>
Laboratorios	Desarrollar cada una de las temáticas por medio de guía prácticas, las cuales serán revisadas por el docente al final de cada una de las sesiones.
Examen Parcial	Medir el conocimiento obtenido de todo lo aprendido en minería de datos.

VII BIBLIOGRAFÍA

- Vercellis, C. (2009). Business Intelligence Data Mining and Optimization for Decision Making. Italia: Wiley.
- Uso de técnicas estadísticas para el análisis de datos (Coursera)

- Recolección y exploración de datos (Coursera)
- Introducción a Data Science: Programación Estadística con R (Coursera)
- Han, J., & Kamber, M.. (2012). Data Mining Concepts and Techniques. USA: Morgan Kaufmann.
- Regression, Data Mining, Text Mining, Forecasting using R (Udemy)
- Learn Data Mining - Clustering Segmentation Using R, Tableau (Udemy)
- Witten, I., (2012). Practical Data Mining. enero 21,2018, de Mexican International Conference and Artificial Intelligence Sitio web:
<https://www.micai.org/2012/tutorials/Weka%20tutorials%20Spanish.pdf>
- Neller, T., (2016). Clustering Iris Data with Weka. enero 6,2018, de EAAI-2018: The Eighth Symposium on Educational Advances in Artificial Intelligence Sitio web:
http://modelai.gettysburg.edu/2016/kmeans/assets/iris/Clustering_Iris_Data_with_Weka.pdf
- Villena, R.,(2014). Minería de datos. enero 5,2018, de Universidad Carlos III de Madrid Sitio web:
<http://www.it.uc3m.es/jvillena/irc/descarga.htm?url=practicass/03-04/18.mem.pdf>
- UCI KDD datasets. (n.d.). [arff]
<https://astuteinternet.dl.sourceforge.net/project/weka/datasets/UCI%20and%20StatLib/uci-20070111.tar.gz>, uci-20070111.
- Orallo, J. & Ramirez, C.. (2006). Práctica de Minería de Datos Introducción al Weka. enero 7,2018, de Universitat Politècnica de València Sitio web:
<http://users.dsic.upv.es/~jorallo/docent/doctorat/weka.pdf>

Bibliografía Complementaria

- Wickham, H., & Grolemund, G. (2017). R for data science. Sebastopol: O'Reilly.
- Lantz, B. (2015). Machine learning with R. Birmingham, UK: Packt Publishing.

- Bishop, C. (2007). Pattern recognition and machine learning. New York, NY: Springer.

Elaborado por	Juan Diego Zavaleta Mejía María José Flores Ramos Melvin Iván Reyes Ortiz Mónica Guadalupe Flores Bonilla	Fecha	02/03/2018
Revisado por	Elmer Arturo Carballo Ruiz MSc	Fecha	12/04/2018

CAPITULO VI: PROPUESTA DE CAPACITACION DE DOCENTES EISI

Se presenta a continuación una propuesta de capacitación a docentes del Departamento de Desarrollo de Sistemas de la Escuela de Ingeniería de Sistemas Informáticos, producto de la investigación realizada a través del instrumento “Diagnostico de las competencias en Inteligencia de Negocios y Analítica de Datos de docentes EISI” (Ver Anexo 10).

Esta propuesta ha sido construida alineándose al tercer horizonte del Plan Estratégico de la Facultad de Ingeniería y Arquitectura para los años 2015-2019:

“Un número importante de profesores contará con estudios de posgrado y atienden programas de capacitación continua de metodologías de enseñanza aprendizaje y el uso de recursos virtuales para aportarle un valor agregado a su actividad docente. Los docentes utilizarán las habilidades desarrolladas para realizar investigaciones en las diferentes áreas de la ingeniería y de la arquitectura, generando procesos de innovación con valor tecnológico y científico en concordancia con las necesidades de desarrollo y competitividad del país”.⁵⁶

Cumpliendo con lo anterior, se ha considerado en esta propuesta lo siguiente:

1. Que el número de docentes capacitados corresponda a la mayoría simple de los docentes encuestados (6 docentes) en las temáticas con un grado de dominio menor a cuatro, según la escala utilizada en el instrumento, donde 1 correspondía sin dominio y 5 dominio avanzado.

2. Que los cursos propuestos sean de grados avanzados, entendiéndose éstos como diplomados especializados y/o postgrados en las áreas de estudio.

6.1 Propuesta de Capacitación para el área de Fundamentos de Negocio

Objetivo:

Que los docentes de la Escuela de Ingeniería de Sistemas Informáticas del Departamento de Desarrollo de Sistemas refuercen sus conocimientos en las temáticas que componen la especialización de Inteligencia de Negocios y Analítica de Datos.

⁵⁶ Universidad de El Salvador, Facultad de Ingeniería y Arquitectura. (n.d). Plan Estratégico Período 2015-2019 (n.d). Recuperado de Unidad de Acceso a la Información Pública de Universidad de El Salvador.

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Fundamentos de Negocios						
Modelos de Negocio	Diplomado de especialización en Innovación Estratégica Empresarial Enlace: http://postgrados.udb.edu.sv/dinnovacion.php	Universidad Don Bosco	80 horas (4 módulos de 20 horas cada uno)	Presencial	\$ 250.00 /módulo	\$ 6,000.00
	Curso Customer Lifetime Value ¿Cómo crear un modelo de negocio basado en el Potencial de la Compañía?	Universidad Centroamericana José Simeón Cañas	16 horas	Presencial	\$ 200.00	\$ 1,200.00
	¿Cómo construir mi modelo de negocio?	Universidad de los Andes	5 semanas	Virtual	\$ 39.00	\$ 234.00
Sistemas de Gestión de Relaciones con los Clientes (CRM)	Customer Relationship Management (CRM) y la Planificación Estratégica	Universidad Centroamericana José Simeón Cañas	16 horas	Presencial	\$ 200.00	\$ 1,200.00

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Sistemas de Gestión de Relaciones con los Clientes (CRM)	Diplomado en CRM (Gerencia De Relaciones Con Los Clientes) Enlace: https://bit.ly/2xxTt9Q	Politécnico de Colombia	120 horas	Virtual	No disponible	No disponible
Sistemas de Soporte a las Decisiones (DSS)	Business Intelligence Concepts, Tools, and Applications Enlace: https://bit.ly/2LbVGKd	University of Colorado	5 semanas	Virtual	\$ 49.00	\$ 294.00
Sistemas de Gestión de la Cadena de Suministro (SCM)	Programa Gestión de la Logística y Cadena de Suministro	Alta Dirección S.A. de C.V. - Insaforp	130 horas	Virtual	No disponible	No disponible
	Seminario en Supply Chain Management Enlace: http://www.circulotec.com/web/guest/seminario-en-supply-chain	Tecnológico de Monterrey	15 horas	Virtual	\$ 210.00	\$ 1,260.00
Sistemas de Comercio Electrónico (E-Commerce)	E-commerce: Venta Online y Comercio Electrónico Enlace: https://bit.ly/2JkbIEd	Universidad Tecnológica Nacional. Facultad Regional Buenos Aires	20 horas	Virtual	\$ 136.00	\$ 816.00

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Gestión de Procesos de Negocio (BPM)	Business Process Management Professional Certification Enlace: http://www.bpminstitute.org/certificates/business-process-management	BPMInstitute	No disponible	Virtual	\$ 3,995.00	\$ 23,970.00
	Certified Business Process Professional (CBPP®) Enlace: https://abpmp.site-ym.com/?page=CBPP_App_Process	ABPMP International	No disponible	Virtual	\$ 725.00	\$ 4,350.00
	Omg Certified Expert In Bpm™ 2 (OCEB™ 2) Enlace: http://www.omg.org/oceb-2/	Object Management Group	No disponible	Virtual	No disponible	No disponible

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Gobernanza de datos	Diplomado en Gobierno Abierto y Gobierno Enlace: https://bit.ly/2kF0Pzd	Universidad Tecnológica Nacional. Facultad Regional Buenos Aires	90 horas	Virtual	\$ 294.00	\$ 1,764.00
	Certified in the Governance of Enterprise IT (CGEIT) Enlace: https://bit.ly/1cRn8IF https://bit.ly/2LcbBrM	ISACA	3 meses	Virtual	\$ 415.00	\$ 2,490.00
	CIMP Data Governance Package Enlace: https://ecm.elearningcurve.com/Online_Data_Governance_Certification_package-cimp-dg.htm http://ecm.elearningcurve.com/v/vspfiles/files/pdf/ELC_DG_Catalog.pdf	ELearning Curve	1 mes	Virtual	\$ 1,795.00	\$ 10,770.00

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Data warehouse						
Arquitectura de una Data Warehouse Esquema de Estrella Esquema de Copo de nieve Esquema de Constelación	Implementing a Data Warehouse with Microsoft® SQL Server® 2014 Enlace: https://www.microsoft.com/en-us/learning/course.aspx?cid=20463	Corporación Orbital	40 horas	Presencial	\$ 550.00	\$ 3,300.00
Data profiling (Perfilado de datos) Mapa de datos lógico Staging área (Datos temporales) Estructuras de datos en un sistema ETL	Curso Superior en Desarrollo de BBDD para Big Data y Gestión de un Datawarehouse Enlace: https://www.emagister.com/curso-superior-desarrollo-bbdd-para-big-data-gestion-datawarehouse-cursos-3395629.htm	Instituto Europeo de Estudios Empresariales	6 meses	Virtual	\$ 509.00	\$ 3,054.00

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Extracción de datos Técnicas de Limpieza de datos Dimensiones de carga lenta Carga Incremental Carga histórica de datos Gestión de sistemas ETL Metadata Tipos de Metadata	SQL Server 2008 R2 Integration Services: Alimentado el Datawarehouse Enlace: https://www.emagister.com/como-configurar-aplicacion-sql-server-2008-online-cursos-2850247.htm	Inter Grupo	100 horas	Virtual	No Disponible	No Disponible

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
OLAP	Inteligencia de Negocios con SQL Server	Universidad Don Bosco	16 horas	Presencial	\$ 350.00	\$ 2,100.00
	Análisis OLAP sobre SQL Server 2008 R2 Enlace: https://www.emagister.com/curso-online-analisis-multidimensional-cubos-olap-cursos-2829181.htm	Inter Grupo	140 horas	Virtual	No Disponible	No Disponible
Reglas para Evaluar Soluciones OLAP Características de los sistemas OLAP						
Modelos OLAP	Máster en Business Intelligence Enlace: https://www.emagister.com/master-online-inteligencia-empresarial-para-toma-decisiones-cursos-2738296.htm	Business School	12 meses	Virtual	\$ 7,885.00	\$ 47,310.00

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Inteligencia de Negocios						
Generalidades de Inteligencia de Negocios: Qué es, Historia, antecedentes y beneficios	Posgrado Internacional en Business Intelligence Enlace: https://www.iseade.edu.sv/index.php/postgrados-especializados/postgrados-impartidos/postgrado-internacional-en-business-intelligence-septiembre-2015	Iseade Fepade	6 meses	Presencial	\$ 1,988.00	\$ 11,928.00
El papel de los modelos matemáticos en Inteligencia de Negocios						
Ciclo de vida en proyectos de inteligencia de negocios						
Desarrollo de un sistema Inteligencia de Negocios						
La ética e Inteligencia de Negocios						

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
<p>Tipos de decisiones</p> <p>Plantilla de aplicación de usuario final</p> <p>Uso de herramienta de visualización en Inteligencia de Negocios</p>	<p>CBIP Preparation and Certification Bundle for Business Professionals</p> <p>Enlace:</p> <p>https://online-learning.tdwi.org/collections/cbip-preparation-and-certification-bundle-for-business-professionals</p>	TDWI Online Training	No disponible	Virtual	\$ 1,730.00	\$ 10,380.00
<p>Indicadores de desempeño</p> <p>Metodologías para el establecimiento de indicadores</p> <p>Sistemas de Planificación de Recursos Empresariales (ERP)</p>	<p>Posgrado en Innovación de la Gestión Empresarial</p> <p>Enlace:</p> <p>http://www.utec.edu.sv/maestrias/utec/postgrados/innovacion-gestion-empresarial</p>	UTEC	8 meses	Presencial	\$ 920.00	\$ 7,360.00

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Big Data						
Fundamentos de Big Data	Maestría en Data Management e Innovación Tecnológica Enlace: https://www.obs-edu.com/int/maestria-en-data-management-e-innovacion-tecnologica?c=I90436M3044&qclid=CjwKCAjwzoDXBRBbEiwAGZRIeD1Daizxfv54fzZPATmbrNSbTVINFm9C1ErkzLtie28FfyxMHu2ehoCmWgQAvD_BwE	Universidad de Barcelona	10 meses	Virtual	\$ 7,914.02	\$ 47,484.12
	Diplomado en Big Data Enlace: http://www.cuautitlan.unam.mx/descargas/educosn/diplomado_big_data.pdf	Universidad Nacional Autónoma de México	250 horas	Presencial	\$ 2,019.62	\$ 12,117.72

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Fundamentos de Big Data	Curso de Big Data Introducción a la Ciencia de Datos Enlace: https://www.educacionit.com/curso-de-big-data	educacionIT	6 semanas	Virtual	\$ 328.00	\$ 1,968.00
Almacenamiento y Procesamiento de Big Data	Diplomado en Big Data Enlace: http://www.universia.cl/estudios/uc/diplomado-big-data/st/258696	Pontificia Universidad Católica de Chile	144 horas	Presencial	\$ 2,900.00	\$ 17,400.00
	Maestría en Ciencia de los Datos y Procesamiento de Datos Masivos (Big Data) Enlace: https://bit.ly/2snMII3	Universidad Cuauhtémoc	20 meses	Presencial	No Disponible	No Disponible

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Almacenamiento y Procesamiento de Big Data	Ingeniería de Datos y Big Data Diploma de Posgrado Enlace: http://estudios.uoc.edu/es/masters-posgrados-especializaciones/diploma-posgrado/informatica-multimedia-telecomunicacion/ingenieria-datos-big-data/programa-academico	Universidad Abierta de Cataluña	6 meses	Presencial	No disponible	No Disponible
Arquitecturas y Frameworks de Big Data	Maestría en Ciencia de los Datos y Procesamiento de Datos Masivos (Big Data) Enlace: https://bit.ly/2snMII3	Universidad Rey Juan Carlos	12 meses	Virtual	No Disponible	No Disponible

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Arquitecturas y Frameworks de Big Data	Diplomado en Big Data: Arquitectura, Implantación y Aplicación Enlace: http://www.universia.net.mx/estudios/itesm/diplomado-big-data-arquitectura-implantacion-aplicacion/st/248356	Instituto Tecnológico y de Estudios Superiores de Monterrey	96 horas	Presencial	\$ 4,060.00	\$ 24,360.00
	Diplomado en Big Data Enlace: http://diplomadosonline.com/docs/BROCHURE_BD.pdf	Universidad Central de Venezuela	4 meses	Virtual	No Disponible	No Disponible
Visualización de Big Data	Máster Universitario en Análisis, visualización e inteligencia de datos Enlace: https://bit.ly/2J5Rcni	UNIR (La Universidad en Internet)	12 meses	Virtual	\$ 9,457.11	\$ 56,742.66

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Visualización de Big Data	Gestión y Visualización de Grandes Volúmenes de Datos Enlace: https://www.fing.edu.uy/cpap/cursos/gestion-y-visualizacion-de-grandes-volumenes-de-datos	Universidad de la República de Uruguay	4 semanas	Presencial	No Disponible	No Disponible
	Business Intelligence / Big Data, Data Visualization Enlace: https://panamericanlatam.com/certificacion/business-intelligence-big-data-data-visualization/	Florida International University	1 semana	Presencial	No Disponible	No Disponible

Tabla 9 Plan de capacitación

Contenido	Nombre del curso	Institución	Duración	Modalidad	Costo del curso(USD)	Costo Total(USD)
Proceso y Metodología de Minería de Datos	Master en Data Analytics & Cognitive Intelligence Enlace: https://www.emagister.com/master-data-analytics-cognitive-intelligence-cursos-3405619.htm	MSMK - Madrid School of Marketing En Madrid	8 meses	Virtual	\$ 17,573.58	\$ 105,441.48
	Master en Big Data y Business Analytics Enlace: https://bit.ly/2mDEL9y	Business School	Entre 6 y 24 meses	Virtual	No Disponible	No Disponible
	Data Analytics – MSc Enlace: https://bit.ly/2LOSuoY	University of Glasgow	36 meses	Virtual	No Disponible	No Disponible

Tabla 9 Plan de capacitación

Asignatura	Total Capacitación (en dólares americanos)
Fundamentos de Negocio	\$ 34,448.00
Data Warehouse	\$ 5,400.00
Inteligencia de Negocios	\$ 19,288.00
Gestión de Big Data	\$ 145,986.78
Análisis Estadístico y Predictivo de Datos	\$ 105,441.48

Tabla 10 Monto total de las capacitaciones

6.2 Perfil Profesional Externo

Se presenta además dos perfiles externos de profesionales con las competencias requeridas en las áreas de Inteligencia de Negocios y Analítica de Datos que permiten impartir la especialización propuesta. Los años de experiencia especificados en cada perfil profesional, tiene su base en los años solicitados en el mercado laboral al momento de realizar el presente trabajo.

Perfil de un Especialista en Inteligencia de Negocios:

Instrucción: Ingeniería o Licenciatura En Sistemas o afín

Postgrado o maestría (deseable) en Negocios, Economía, Matemáticas, Contabilidad o Finanzas.

Experiencia:

- Experiencia de 3 años mínimo en análisis de negocio, levantamiento de requerimientos, procesos, diseño especificaciones funcionales y técnicas.
- Experiencia de 2 años mínimo en mapeo de Procesos y procedimientos
- Experiencia de 3 años mínimo en análisis, diseño y construcción de Bodegas de Datos. Generación de procesos ETL en modelos tipo Data Warehouse.
- Experiencia de 2 años mínimo en administración y mantenimiento de bases de datos
- Experiencia demostrable con el manejo de alguna herramienta de inteligencia empresarial, reportes/visualización.

Conocimientos:

- Diseño e interpretación de diagramas de flujo.
- Diseño de modelos funcionales de las aplicaciones que soportan procesos de negocios.
- Desarrollo de bases de datos, incluido el diseño de objetos y la implementación de tablas, vistas, funciones y procedimientos almacenados
- Diseño e implementación de procesos ETL complejos diseñados para importar y normalizar datos de múltiples sistemas fuente.
- Gobernanza de datos (deseable)
- Inglés nivel intermedio.
- Certificaciones relacionadas con la especialidad.

Habilidades:

- Capacidad de gestionar proyectos.
- Capacidad de aprender y adaptarse rápidamente a nuevas herramientas y tecnologías.
- Excelentes habilidades de comunicación verbal/escrita y presentación.
- Enseñanza (deseable).

Perfil de un Especialista en Analítica de datos

Instrucción: Ingeniería o Licenciatura en Sistemas Informáticos.

Postgrado o maestría en Big Data y Analítica de datos o en Data Science (Deseable)

Experiencia:

- Experiencia de 5 años mínimo en tecnologías de Big data.
- Experiencia de 5 años mínima en procesamiento de datos no estructurados.
- Experiencia de 3 años mínimo en la aplicación de modelos estadísticos en conjuntos de datos.
- Experiencia de 3 años mínimo en análisis de datos orientado al descubrimiento de patrones.
- Experiencia de 3 años mínimo en Lenguajes de programación orientados al análisis de datos.

Conocimientos

- Gestión de Base de datos semi estructurados y no estructurados.
- Técnicas y entornos de procesamiento para Big Data.
- Arquitecturas de Big Data.
- Lenguajes de programación orientados a la analítica.
- Herramientas de visualización.
- Herramientas Analíticas.
- Modelos de agrupamiento de datos.
- Modelos predictivos de datos.
- Análisis de sentimiento en los datos.
- Creación de modelos para minería de datos.
- Inglés Intermedio.

Habilidades

- Capacidad de aprender y adaptarse rápidamente a nuevas herramientas y tecnologías.
- Capacidad de comunicación clara.
- Excelentes habilidades de comunicación verbal/escrita y presentación.
- Enseñanza (deseable).

CONCLUSIONES

Con el desarrollo de este proyecto se concluye lo siguiente:

1. Los resultados de este estudio de investigación son una propuesta de especialización para la carrera de Ingeniería de Sistemas Informáticos de la Universidad en dos macro áreas de la informática: inteligencia de negocios y analítica de datos, las cuales se demostraron ser parte de la demanda laboral en aumento a nivel centroamericano. Es importante tener conciencia, que las necesidades en conocimientos informáticos especializados son evolutivos y rápidos; y que este proyecto representa un punto de partida para especializaciones futuras más complejas dentro de la informática que están pronunciándose.
2. Las cinco áreas de conocimiento de esta especialización, a saber: Negocios, Data Warehouse, Inteligencia de Negocios, Big data y Minería de datos guardan relación con la descripción de la carrera de Ingeniería de Sistemas Informáticos la cual dicta que abarca conocimientos de Ciencias de Computación, Matemática, Ingeniería de Software y Gestión Empresarial. Así mismo, las habilidades que se pretenden desarrollar en esta especialización están encaminadas a solución de problemas y la toma de decisiones. Estas actividades, según lo propone la Facultad de Ingeniería y Arquitectura, son parte importante de la carrera de Ingeniería de Sistemas. Con esto se corrobora que la especialización propuesta se apega a las bases del actual programa, pero agrega valor al actualizar su contenido.
3. La naturaleza investigativa de este proyecto hizo necesario el estudio de mercado de diferentes escenarios (bolsas de trabajo, programas de estudio de universidades locales e internacionales y encuestas); asegurando con ello, un perfil académico que cumpla con su función de especializar al estudiante para que se convierta en un profesional competitivo en el entorno laboral. Involucrar en esta investigación a profesionales que ya se encuentran laborando en las áreas de estudio, fue de suma importancia, pues validó que el propósito de la especialización es alcanzable y valorizó los contenidos de los programas de estudio. Esto último, agrega más objetividad a los resultados presentados.
4. Con la propuesta de capacitación que se presenta en este documento se demuestra las fortalezas y los puntos de mejora de los docentes en las áreas de la especialización de este proyecto. Es importante que se tome esta iniciativa como un aspecto a considerar pues está alineado con las líneas estratégicas del Plan de la Facultad de Ingeniería y Arquitectura 2015-2019 y está encaminado al enriquecimiento del perfil profesional de los docentes de carrera de Ingeniería de Sistemas Informáticos, pues beneficia a la Universidad de El Salvador al tener personal preparado que garantice la calidad de la enseñanza y beneficia a los estudiantes de la carrera al ser ellos el punto esencial de esta labor.

RECOMENDACIONES

Se sugiere a la Escuela de Ingeniería de Sistemas Informáticos:

1. Construir una propuesta de maestría en el área de Inteligencia de Negocios y Analítica de datos para la carrera de Ingeniería de Sistemas Informáticos de la Universidad de El Salvador.
2. Desarrollar proyectos de investigación y/o aplicación en los temas de inteligencia artificial y aprendizaje automático, al ser áreas en tendencia de la informática. El desarrollo de proyectos de esta magnitud beneficiaría no solo a la educación, sino que también a la cultura y a la economía del país.
3. Implementar nuevas metodologías de evaluación de conocimientos y habilidades en los estudiantes. Se propone para los programas presentados, y para las materias de la carrera que apliquen las siguientes:
 - a. Completar un curso en línea gratuito.
 - b. Realizar talleres prácticos en convenio con universidades o empresas públicas o privadas nacionales o extranjeras.
4. Se sugiere que al momento de implementar esta propuesta de especialización se comparen los contenidos de ésta contra los de la nueva curricular para la óptima integración de temas.
5. Se sugiere realizar estudios de diagnóstico de competencias a los docentes de la EISI para otras áreas de la información con el propósito de implementar y/o reforzar programas de capacitación.

GLOSARIO

A

Ambari: Apache Ambari es un software que proporciona, gestiona y monitorea clusters de Apache Hadoop por medio de una interfaz web intuitiva por medio de una robusta API REST que es sumamente útil para automatizar operaciones en los clústeres.

Árboles de decisión: Es un algoritmo que induce reglas de manera natural y que puede ser usada en clasificación de datos y predicción.

Almacén operacional de datos (ODS): Durante el proceso ETL es lugar donde se almacenan y mapean los datos extraídos de los sistemas operacionales.

C

Cassandra: Es una base de datos no SQL distribuida masivamente escalable de código abierto, que ofrece disponibilidad continua capaz de escalar linealmente.

Cookie: conocida también como una cookie HTTP, cookies web o cookie del navegador, son archivos de tamaño pequeño que algunos sitios web guardan o almacenan en el navegador web de un usuario, mientras el usuario está navegando o navegó por un sitio web.

Coefficiente de correlación: Es una medida de regresión que pretende cuantificar el grado de variación conjunta entre dos variables.

Competencia: capacidad de movilizar, articular y colocar en acción, valores, conocimientos y habilidades necesarios para el desempeño eficiente y eficaz de actividades requeridas por la naturaleza del trabajo.

Conocimientos: estado o condición de comprensión sintetizado en conceptos amplios y aplicables.

Currículo: conjunto de competencias básicas, objetivos, contenidos, criterios metodológicos y de evaluación que el estudiante deberá lograr cuando se encuentren en un determinado nivel educativo.

Curso: Fuente de conocimiento para la profundización en entendimiento de un tema.

D

Diagrama de dispersión: Es una herramienta gráfica que ayuda a identificar la posible relación entre dos variables.

Diagrama de Pareto: Constituye un sencillo y gráfico método de análisis que permite discriminar entre las causas más importantes de un problema (los pocos y vitales) y las que lo son menos (los muchos y triviales).

H

Habilidades: capacidad de aplicar conocimientos, sean éstos producto de la experiencia o la teoría, a una situación específica.

Habilidades Técnicas: Capacidad de aplicar métodos, procedimientos y técnicas específicas en un campo especializado.

Hadoop: es un framework de programación basado en Java de código abierto que soporta el procesamiento y almacenamiento de conjuntos de datos extremadamente grandes en un entorno de computación distribuida. Los servicios de Hadoop proporcionan almacenamiento de datos, procesamiento de datos, acceso a datos, gobierno de datos, seguridad y operaciones.

I

IBM: Es un portafolio de herramientas creado por IBM provee un ecosistema de datos, analítica y capacidades cognitivas. Además, cuenta con suites para gestión de datos, Big data, business analytics y data science.

Intervalos de confianza: Un intervalo de confianza es un rango de valores, derivado de los estadísticos de la muestra, que posiblemente incluya el valor de un parámetro de población desconocido.

Internet de las cosas IoT: es una red de objetos físicos –vehículos, máquinas, electrodomésticos y más– que utiliza sensores e interfaces de programación de aplicaciones (API) para conectarse e intercambiar datos por internet.

IS: Information System (Sistemas de información), conjunto de componentes interrelacionados que permiten capturar, procesar, almacenar y distribuir la información para apoyar la toma de decisiones y el control en una institución.

K

KNIME: (Konstanz Information Miner): es una plataforma de código abierto, de fácil uso y comprensible para integración de datos, procesamiento, análisis, y exploración. Ofrece a los usuarios la capacidad de crear de forma visual flujos o tuberías de datos, ejecutar selectivamente algunos o todos los pasos de análisis, y luego estudiar los resultados, modelos y vistas interactivas. Está escrito en Java y está basado en Eclipse.

M

Machine Learning: (Aprendizaje Automático) es una rama de la Inteligencia Artificial que se dedica al estudio de los agentes/programas que aprenden o evolucionan basados en su experiencia, para realizar una tarea determinada cada vez mejor. El objetivo principal de todo proceso de aprendizaje es utilizar la evidencia conocida para poder crear una hipótesis y poder dar una respuesta a nuevas situaciones no conocidas.

Malla curricular: Es un instrumento que contiene la estructura del diseño de un determinado curso, de forma articulada e integrada, permitiendo una visión de conjunto sobre la estructura general de un área.

MathWorks; es el desarrollador líder de software de computación matemática.

Ingenieros y científicos de todo el mundo confían en sus productos (MATLAB y Simulink) para acelerar el ritmo de descubrimiento, innovación y desarrollo. MATLAB, el lenguaje de la computación técnica, es un entorno de programación para el desarrollo de algoritmos, análisis de datos, visualización y cálculo numérico. Simulink es un entorno gráfico para simulación y diseño basado en modelos para sistemas multidominio dinámicos y embebidos.

Minitab: es un programa de computadora diseñado para ejecutar funciones estadísticas básicas y avanzadas. Combina lo amigable del uso de Microsoft Excel con la capacidad de ejecución de análisis estadísticos.

L

Lógica Difusa: Es una lógica multivaluada que permite representar matemáticamente la incertidumbre y la vaguedad, proporcionando herramientas formales para su tratamiento.

P

Perfil Profesional: conjunto de rasgos identificadores de competencias en un sujeto que recién recibe su título o grado que corresponde al profesional básico.

R

Random Forest: Extensión de árboles de decisión. Con la diferencia que se construyen a partir de una selección aleatoria de muestras y una selección aleatoria de un conjunto de atributos.

RapidMiner: antes llamado YALE, es un ambiente de experimentos en aprendizaje automático y minería de datos que se utiliza para tareas de minería de datos tanto en investigación como en el mundo real. Permite a los experimentos componerse de un gran número de operadores anidables arbitrariamente, que se detallan en archivos XML y se hacen con la interfaz gráfica de usuario de RapidMiner. RapidMiner ofrece más de 500 operadores para todos los principales procedimientos de máquina de aprendizaje, y también combina esquemas de aprendizaje y evaluadores de atributos del entorno de aprendizaje Weka. Está disponible como una herramienta stand-alone para el análisis de datos y como motor para minería de datos que puede integrarse en tus propios productos.

Regresión Lineal Simple: Es un modelo matemático usado para aproximar la relación de dependencia entre una variable dependiente Y , las variables independientes X_i y un término aleatorio ϵ .

Regresión Lineal Múltiple: Es un modelo matemático como la regresión lineal simple, es utilizada para conocer la relación entre dos o más variables.

Regresión Logística: Resulta útil para los casos en los que se desea predecir la presencia o ausencia de una característica o resultado según los valores de un conjunto de predictores. Es similar a un modelo de regresión lineal, pero está adaptado para modelos en los que la variable dependiente es dicotómica.

Requerimiento: es la condición o capacidad que debe tener un sistema, producto, servicio o componente para satisfacer un contrato, estándar, especificación, u otros documentos formalmente establecido.

Rstudio: es un entorno de desarrollo integrado (IDE) para R (lenguaje de programación), incluye una consola, editor de sintaxis que apoya la ejecución de código, así como herramientas para el trazado, la depuración y la gestión del espacio de trabajo.

Rol: papel que un individuo o cosa determinada juega en un determinado contexto. Tiene que ver con la función que se ejerce, con el sentido de las actividades que se desarrollan

S

SAS: Conocida anteriormente como Statistical Analysis System, es una suite de herramientas desarrollada por Sas Institute para analítica avanzada, análisis multivariado, inteligencia de negocios, gestión de datos, y análisis predictivo.

Silabo: Es una herramienta de planificación de un curso que organiza los contenidos y el trabajo que se realizará para lograr el aprendizaje que se propone.

Sistemas Heredados: Sistema de información con misión crítica dentro de una organización que se resiste significativamente a cambios y modificaciones.

Slowly changing dimensión (Dimensiones de carga lenta): Dimensiones de un Data Warehouse que cambian lentamente en el transcurso del tiempo en vez de cambiar de manera periódica o regular.

SVM: Es un algoritmo supervisado de machine learning usado principalmente en problemas de regresión y clasificación. En este algoritmo se grafica cada item como un punto n-dimensional (donde n se refiere al número de atributos que se tienen) con el valor de cada atributo siendo el valor una coordenada particular.

T

Tableu: Es una herramienta de análisis profundo, que permite realizar descubrimientos, de los dos activos más importantes: gente y datos (tanto big data como datos de menor volumen). Admite análisis a medida por parte de prácticamente de cualquier usuario. De este modo otorga a todo el personal la capacidad para ver y comprender mejor sus datos. Además, permite a sus analistas de negocios publicar KPI a toda la empresa en una plataforma de análisis centralizada y fácil de usar.

Teoría de Bayes: El teorema de Bayes parte de una situación en la que es posible conocer las probabilidades de que ocurran una serie de sucesos A_i . A esta se añade un suceso B cuya ocurrencia proporciona cierta información, porque las probabilidades de ocurrencia de B son distintas según el suceso A_i que haya ocurrido.

V

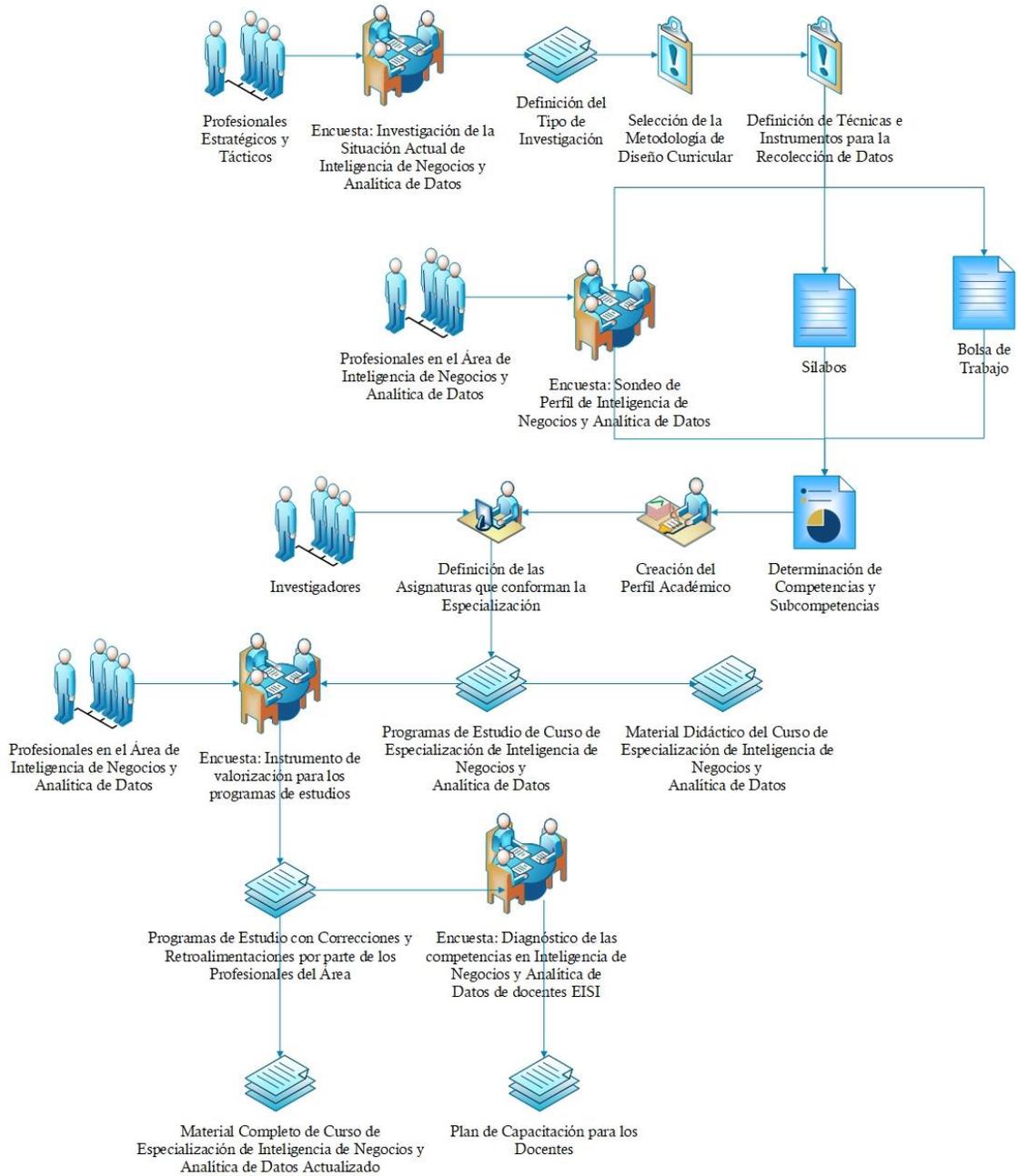
Valores: es toda aquella manera de actuar o interactuar, considerada de importancia, e enfocados para este caso, a un ambiente laboral.

W

Weka: (Waikato Environment for Knowledge Analysis, en español «entorno para análisis del conocimiento de la Universidad de Waikato») es una plataforma de software para el aprendizaje automático y la minería de datos escrito en Java y desarrollado en la Universidad de Waikato. Weka es software libre distribuido bajo la licencia GNU-GPL.

ANEXOS

Anexo 1: Flujo de Trabajo del Proyecto



Anexo 2: Lluvia de Ideas para el análisis y determinación de la situación problemática

- El plan de estudio de la carrera de Ingeniería de Sistemas Informáticos no se ha actualizado desde 1998.
- Existen otras tendencias para un ingeniero informático, el análisis y el desarrollo de sistemas puede ser realizado por técnicos, ahora se requiere que un ingeniero tenga un enfoque a los datos.
- El internet posibilita el acceso a la información, por este medio, es que se conocen las tendencias y los estudios que se están realizando en la carrera de Ingeniería de Sistemas Informáticos. Las bolsas de trabajo nacionales publican con más frecuencia plazas relacionadas con el Data Warehousing, elaboración de reportes, analistas.
- Existen universidades e instituciones que están innovando la enseñanza en ofrecer cursos de especialización y postgrados en inteligencia de negocios.
- La generación de información por medio de sistemas transaccionales, mensajes de texto, redes sociales, transacciones en línea, correos se está incrementando y ahora es necesario no solo almacenar los datos, sino más bien, darles un tratamiento y generar conocimiento a partir de ellos.
- Hay empresas que ya contemplan unidades destinadas a la inteligencia de negocios.
- El término Big Data se ha popularizado en El Salvador, la Escuela Superior de Economía y Negocios tiene a disposición a partir del 27 de mayo 2016 un Postgrado en Analítica para Ciencia de Datos.
- En el país existen empresas que necesitan aplicar tecnologías que permitan analizar de mejor manera los datos que manejan.
- Cubrir conocimientos que el estudiante o profesional no cumplen en el campo laboral.
- Identificar los conocimientos, las habilidades y técnicas actuales, que se requieren en un profesional especializado en inteligencia de negocios y analítica de datos en El Salvador.

- Existen sistemas transaccionales robustos que generan gran cantidad de datos de las diferentes operaciones de las grandes y medianas organizaciones; estos datos al ser detallados son necesarios que pasen por un análisis con el cual se genere una formación con la cual se puedan tomar decisiones, ese proceso de análisis se realiza por medio de la inteligencia de negocios y analítica de datos.

Anexo 3: Análisis de Resultado de Encuesta “Investigación de la Situación Actual de la inteligencia de negocios y analítica de datos”

(Ver en CD en la carpeta de Anexos)

Anexo 4: Cálculo de muestra para poblaciones infinitas o desconocidas

La fórmula para calcular el tamaño de muestra cuando se desconoce el tamaño de la población es la siguiente:

$$n_{opt.} = \frac{Z^2 \times p \times q}{d^2}$$

Donde

Z = nivel de confianza,

p = probabilidad de éxito, o proporción esperada

q = probabilidad de fracaso

d = precisión (error máximo admisible en términos de proporción)

Consideraciones:

1. Cuando se desconoce la probabilidad de éxito esperada, se tiene que utilizar el criterio conservador ($p = q = 0.5$), lo cual maximiza el tamaño de muestra.
2. Para el caso del nivel de confiabilidad de la muestra calculada (Z) existen algunos valores estandarizados en función del grado de confiabilidad asumido.
 Para un: 99 % ----- z = 2,58 (Empleado con frecuencia)
 95 % ----- z = 1,96 (El más empleado)
 90 % ----- z = 1,64

Por tanto, se establece se asumirá un grado de confianza del 95%.

3. Toda expresión que se calcula contiene un error de cálculo debido a las aproximaciones decimales que surgen en la división por decimales, error en la selección de la muestra, entre otras, por lo que este error se puede asumir entre un 1 hasta un 10%; es decir, que se asume en valores de probabilidad correspondiente entre un 0.01 hasta un 0.1.

Por tanto, se establece un error máximo admisible del 5%

$$n_{opt} = \frac{1.96^2 \times 0.5 \times 0.5}{0.05^2} = 385$$

Anexo 5: Diseño de la Encuesta de Sondeo del Perfil de Inteligencia de Negocios y Analítica de Datos

(Ver en CD en la carpeta de Anexos)

Anexo 6: Análisis de Resultados de la Encuesta de Sondeo del Perfil de Inteligencia de Negocios y Analítica de Datos

(Ver en CD en la carpeta de Anexos)

Anexo 7: Matrices de resumen de Sílabos de Universidades

(Ver en CD en la carpeta de Anexos)

Anexo 8: Análisis de Bolsas de Trabajo

(Ver en CD en la carpeta de Anexos)

Anexo 9: Análisis de Resultados de Encuesta de Valorización

(Ver en CD en la carpeta de Anexos)

Anexo 10: Diagnóstico de las competencias en Inteligencia de Negocios y Analítica de datos de docentes EISI

1. Mencione las últimas cinco asignaturas que ha impartido en la Universidad de El Salvador en los últimos tres años en orden cronológico (empezando de la más reciente a la más antigua).

Asignatura	Frecuencia
Introducción a la Informática	7
Programación I	5
Bases de Datos	2
Diseño de Sistemas I	2
Manejo de Software	2
Programación III	2
Programación para Dispositivos Móviles	2
Sistemas de Información Gerencial	2
Administración de Proyectos Informáticos	1
Análisis Numérico	1
Comercio Electrónico	1
Consultoría Profesional	1
Diseño de Sistemas II	1
Estructura de datos	1
Herramientas de Productividad	1
Implementación de Base de Datos	1
Ingeniería de Software	1
Ingeniería del software	1
Métodos Probabilísticos	1
Microprogramación	1
Programación II	1
Sistemas Operativos	1
Técnicas de Simulación	1
Tecnologías Orientadas a Objetos	1
Trabajo de Graduación	1

2. Mencione los cursos o capacitaciones que ha impartido en otras instituciones/organizaciones en los últimos tres años en orden cronológico (empezando de la más reciente a la más antigua).

Nota: Para esta pregunta se ha colocado cada respuesta en una categoría para efectos de presentación de resultados

Curso	Categoría
Taller de Enseñanza de Matlab.	Análisis de datos
Taller de Enseñanza de Scilab.	
Curso de Administración de Base de Datos	Bases de Datos
Curso de Especialización en base de datos	
Aplicaciones descentralizadas y su implementación en blockchain	Desarrollo de Aplicaciones/Sistemas
Desarrollo Web Symfony2 integrando Sonata Project	
Desarrollo Web Symfony2 y Sonata Project	
Desarrollo ágil de apis con ruby on rails 5	
Desarrollo Web con Symfony2	
Implementación de JasperReport Server en un entorno libre GNU Linux	
Implementación de Arquitecturas Basadas en Servicios - SOA	Diseño de Aplicaciones/Sistemas
SCRUMSCRUM	Gestión de Proyectos
Docencia en ambientes virtuales de enseñanza	Otros
Seguridad en Sistemas Operativos, base de Datos y aplicaciones	Seguridad
Protección y Respaldo de datos	
Personalización de seguridad y modificación de core de PENTHO 7	

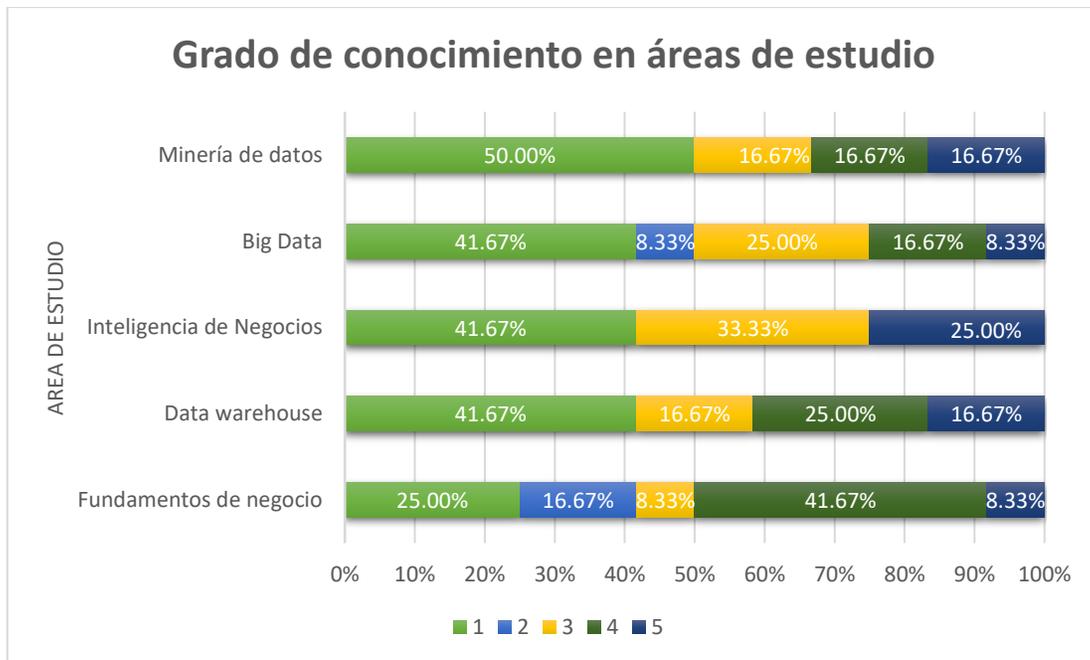
3. Mencione los cursos, capacitaciones o estudios que ha completado o está recibiendo.

Nota: Para esta pregunta se ha colocado cada respuesta en una categoría para efectos de presentación de resultados

Curso	Categoría
Curso de Big Data	Análisis de datos
Curso de Cloud Computing	
Curso de Analítica de Datos	
Curso de Machine Learning	
Curso Básico e Intermedio en el Software Estadístico R	
Blockchain for Business - An Introduction to Hyperledger Technologies	Desarrollo de Aplicaciones/Sistemas
Curso de Programación Orientada a Objetos	
Diplomado en Tecnologías Móviles	
Desarrollo de servicios en la nube con HTML5, Javascript y node.js	
Introducción a Ruby on Rails	

Scrum	Gestión de Proyectos	
Curso de Scrum		
Administración de Proyectos de PMI		
Seminario sobre Metodología de Desarrollo ágil Scrum.		
ITIL	Gobierno/Auditoría de Sistemas	
Gestión y Auditoría de la seguridad de la información		
COBIT		
Auditoría de Sistemas de Información Basados en COBIT y otros Marcos de Referencia		
Cobit 4	Otros	
Diplomado en la Educación a Distancia		
La nueva economía digital en América Latina y el Caribe		
Curso de Fortalecimiento Pedagógico para la Enseñanza Superior		
Diplomado en herramientas e-learning.		
Diplomado Formación Pedagógica para la Enseñanza Superior		
Diplomado en Formación Pedagógica		
Capacitación sobre escritura de Documentos científicos		
Curso de Introducción a los Sistemas de Información Geográficos		
Formación de Competencias Educativas de Insaforp		
Estudios de ingeniería de sistemas Informáticos		
Introducción a la Investigación Científica		
Cisco Certified Network Associate		Redes y Comunicaciones
Curso de Administración Básica de Servidores Debian GNU/Linux		
vo/lp		
Cybersecurity Fundamentals	Seguridad	
Criptología Matemática		
Gestión de Riesgos y Controles IT		
Curso de Seguridad		

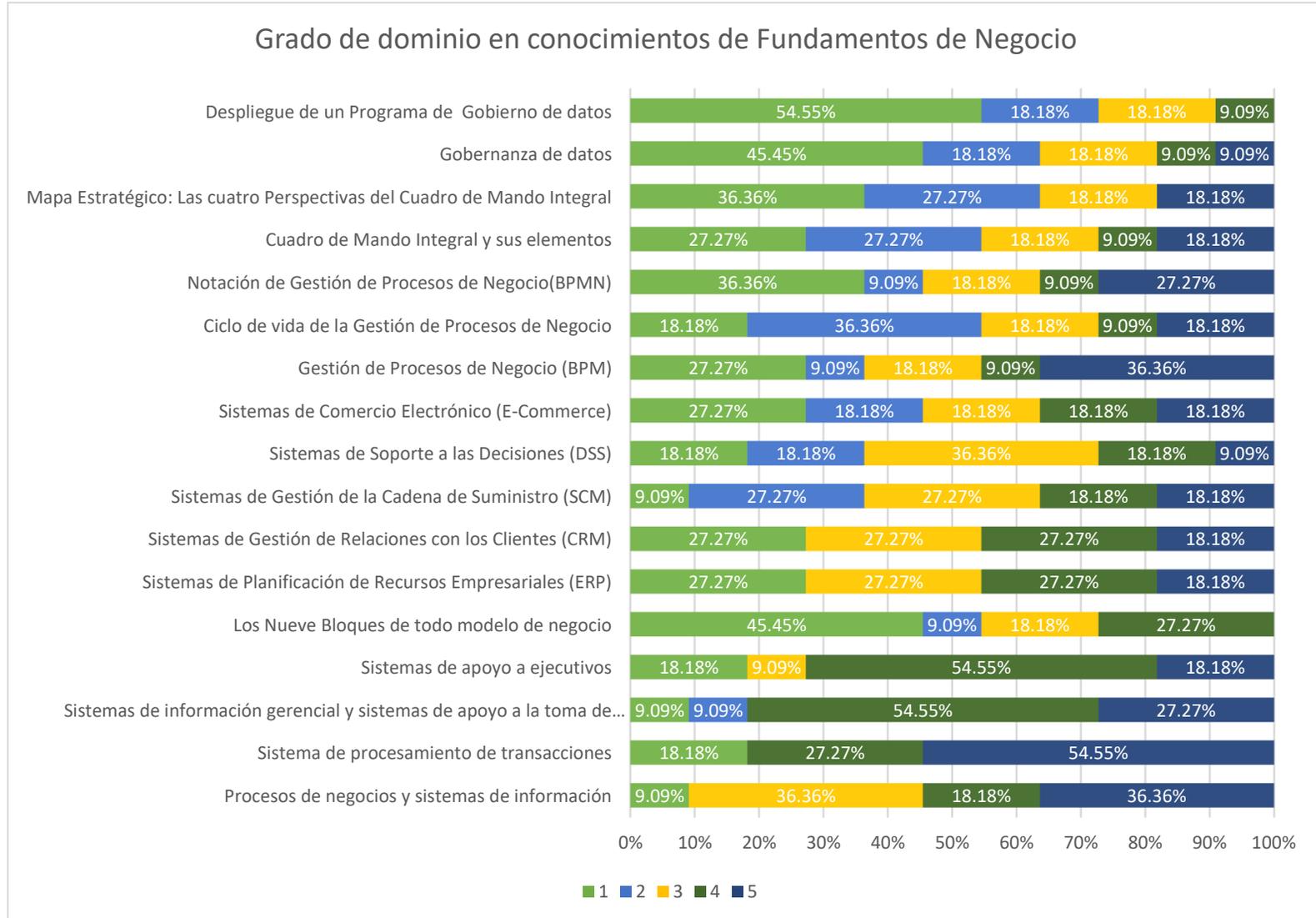
4. Para cada área listada indique el grado de conocimiento que posee; siendo 1 sin dominio del área y 5 dominio avanzado



Total respuestas: 12

Área de estudio	Grado de Dominio				
	1	2	3	4	5
Fundamentos de Negocio	25,00%	16,67%	8,33%	41,67%	8,33%
Data Warehouse	41,67%	0,00%	16,67%	25,00%	16,67%
Inteligencia de Negocios	41,67%	0,00%	33,33%	0,00%	25,00%
Big Data	41,67%	8,33%	25,00%	16,67%	8,33%
Minería de datos	50,00%	0,00%	16,67%	16,67%	16,67%

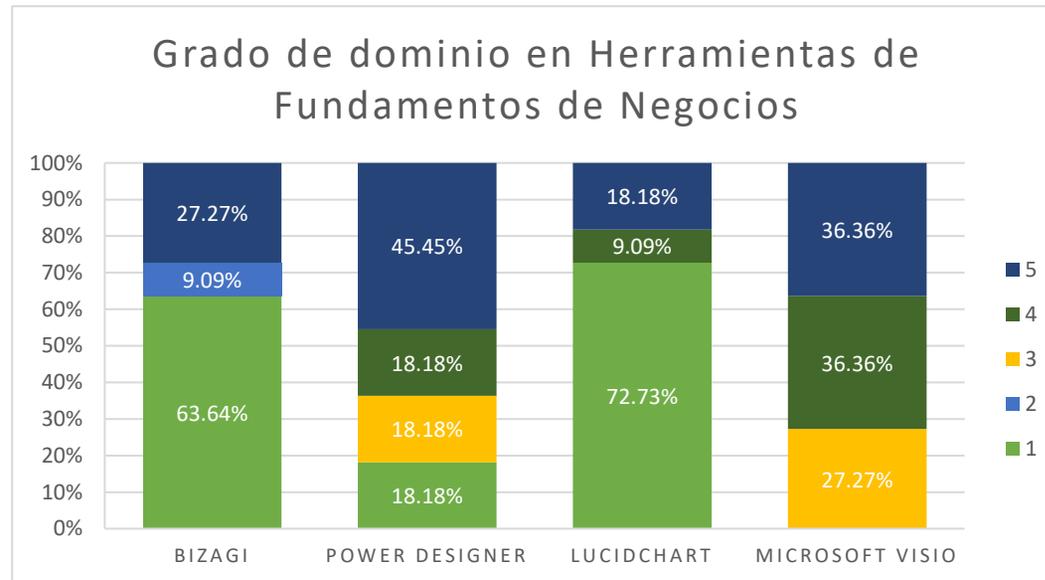
5. Seleccione la opción que más se ajuste a su criterio; siendo 1 sin dominio del área y 5 dominio avanzado



Contenidos de Fundamentos de Negocio	Grado de Dominio					Plan Capacitación 1	Plan Capacitación 2
	1	2	3	4	5		
Procesos de negocios y sistemas de información	9,09%	0,00%	36,36%	18,18%	36,36%	No	No
Sistema de procesamiento de transacciones	18,18%	0,00%	0,00%	27,27%	54,55%	No	No
Sistemas de información gerencial y sistemas de apoyo a la toma de decisiones	9,09%	9,09%	0,00%	54,55%	27,27%	No	No
Sistemas de apoyo a ejecutivos	18,18%	0,00%	9,09%	54,55%	18,18%	No	No
Los Nueve Bloques de todo modelo de negocio	45,45%	9,09%	18,18%	27,27%	0,00%	Si	No
Sistemas de Planificación de Recursos Empresariales (ERP)	27,27%	0,00%	27,27%	27,27%	18,18%	Si	No
Sistemas de Gestión de Relaciones con los Clientes (CRM)	27,27%	0,00%	27,27%	27,27%	18,18%	Si	No
Sistemas de Gestión de la Cadena de Suministro (SCM)	9,09%	27,27%	27,27%	18,18%	18,18%	Si	No
Sistemas de Soporte a las Decisiones (DSS)	18,18%	18,18%	36,36%	18,18%	9,09%	Si	No
Sistemas de Comercio Electrónico (E-Commerce)	27,27%	18,18%	18,18%	18,18%	18,18%	Si	No
Gestión de Procesos de Negocio (BPM)	27,27%	9,09%	18,18%	9,09%	36,36%	Si	No
Ciclo de vida de la Gestión de Procesos de Negocio	18,18%	36,36%	18,18%	9,09%	18,18%	Si	No
Notación de Gestión de Procesos de Negocio(BPMN)	36,36%	9,09%	18,18%	9,09%	27,27%	Si	No
Cuadro de Mando Integral y sus elementos	27,27%	27,27%	18,18%	9,09%	18,18%	Si	No
Mapa Estratégico: Las cuatro Perspectivas del Cuadro de Mando Integral	36,36%	27,27%	18,18%	0,00%	18,18%	Si	No
Gobernanza de datos	45,45%	18,18%	18,18%	9,09%	9,09%	Si	No
Despliegue de un Programa de Gobierno de datos	54,55%	18,18%	18,18%	9,09%	0,00%	Si	Si

Total respuestas: 11

6. Para cada herramienta listada indique el grado de dominio que posee; siendo 1 sin dominio de la herramienta y 5 dominio avanzado

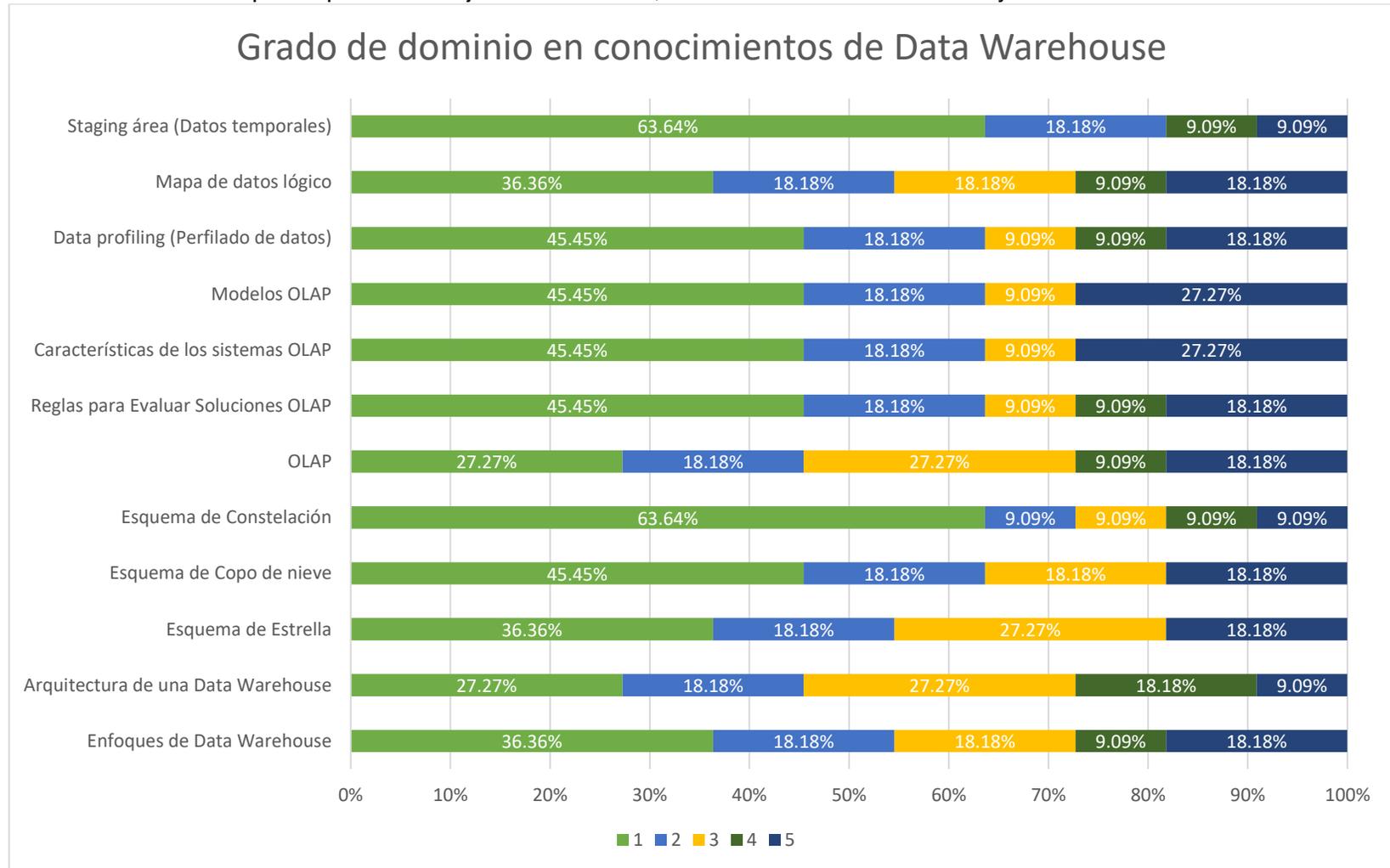


Total respuestas: 11

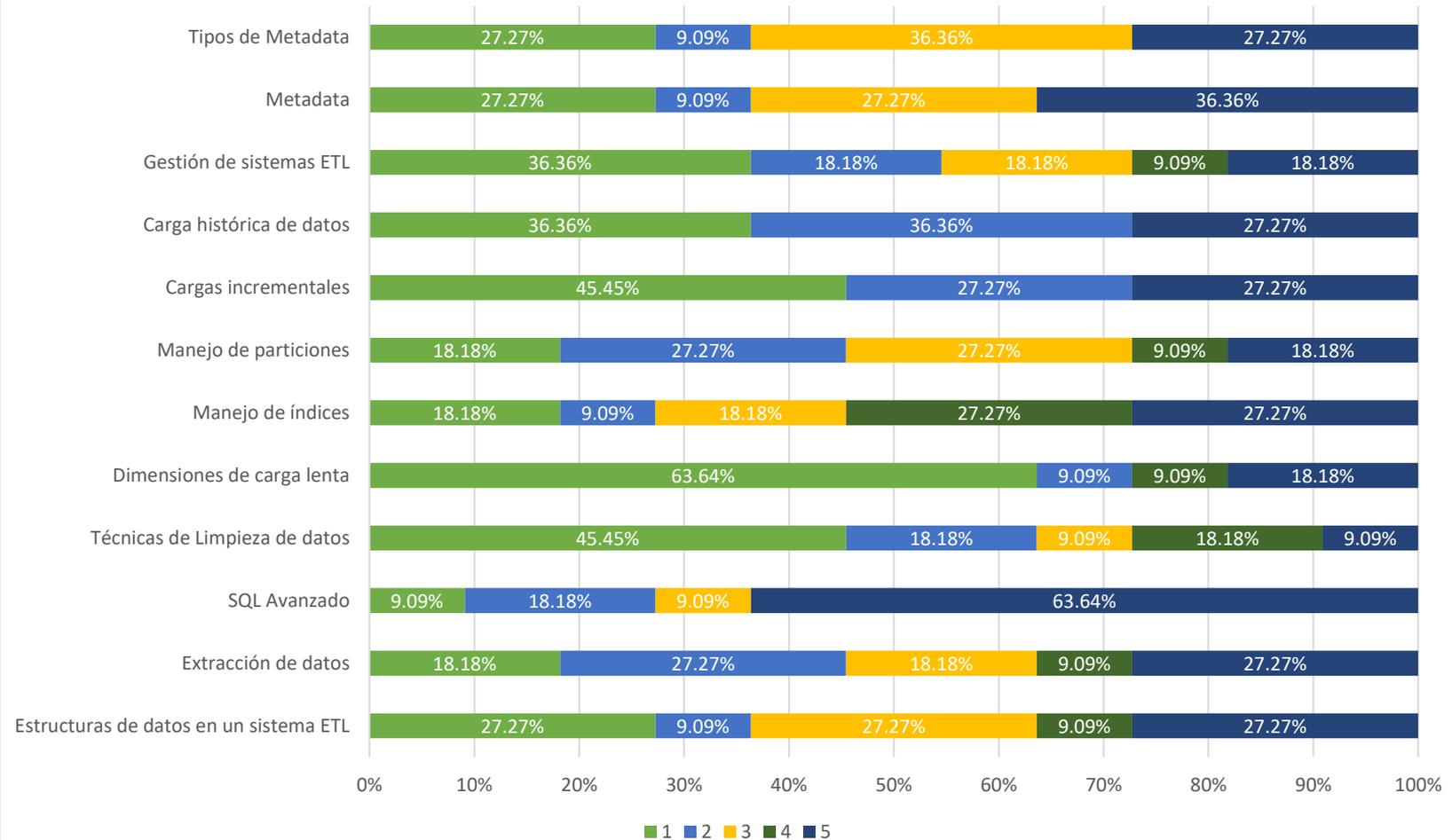
Los encuestados agregaron las siguientes herramientas: **Eclipse, NetBeans, TOAD, BonitaSoft, yEd Graph Editor, DIA**. No especificaron grado de dominio.

Herramientas de Fundamentos de Negocio	Grado de Dominio					Plan Capacitación 1	Plan Capacitación 2
	1	2	3	4	5		
Bizagi	63,64%	9,09%	0,00%	0,00%	27,27%	Si	No
Power Designer	18,18%	0,00%	18,18%	18,18%	45,45%	No	No
Lucidchart	72,73%	0,00%	0,00%	9,09%	18,18%	Si	No
Microsoft Visio	0,00%	0,00%	27,27%	36,36%	36,36%	No	No

7. Seleccione la opción que más se ajuste a su criterio; siendo 1 sin dominio del área y 5 dominio avanzado



Grado de dominio en conocimientos de Data Warehouse

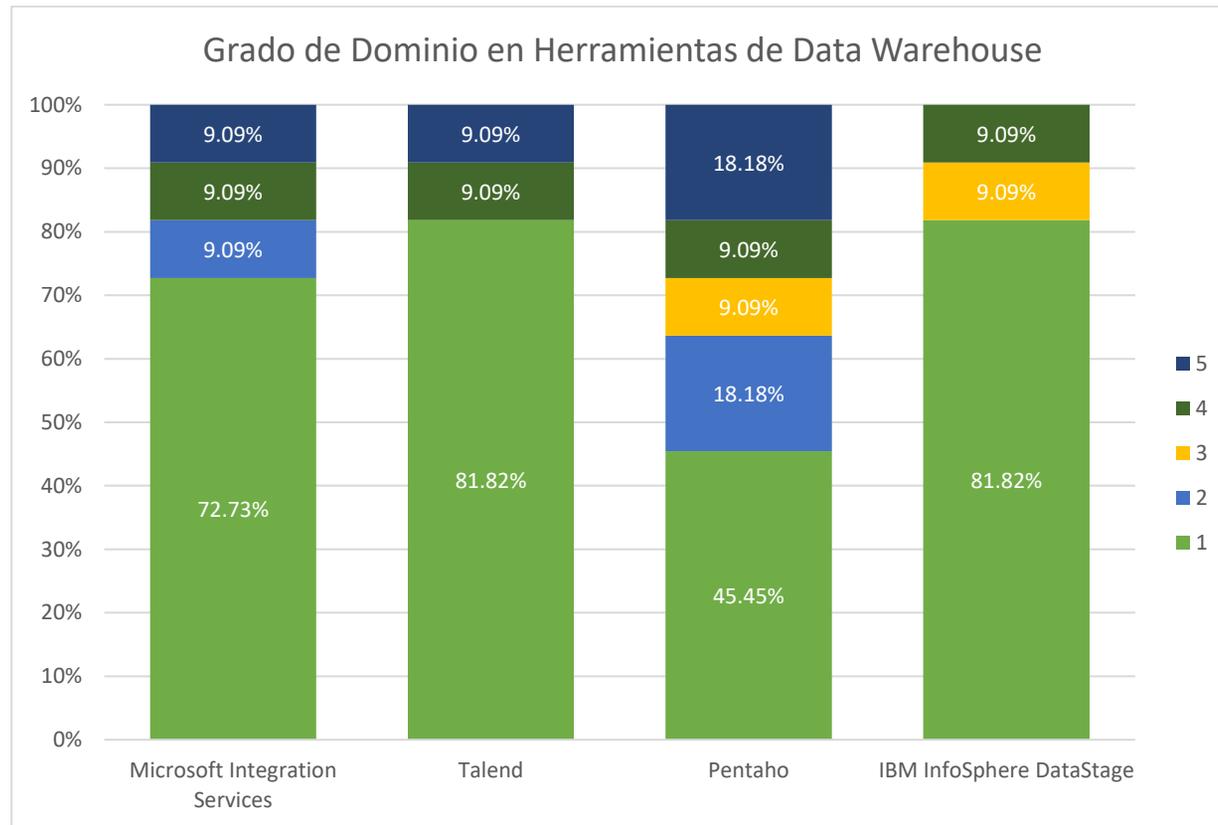


Contenidos de Data Warehouse	Grado de Dominio					Plan de capacitación 1	Plan de capacitación 2
	1	2	3	4	5		
Enfoques de Data Warehouse	36,36%	18,18%	18,18%	9,09%	18,18%	Si	No
Arquitectura de una Data Warehouse	27,27%	18,18%	27,27%	18,18%	9,09%	Si	No
Esquema de Estrella	36,36%	18,18%	27,27%	0,00%	18,18%	Si	No
Esquema de Copo de nieve	45,45%	18,18%	18,18%	0,00%	18,18%	Si	No
Esquema de Constelación	63,64%	9,09%	9,09%	9,09%	9,09%	Si	No
OLAP	27,27%	18,18%	27,27%	9,09%	18,18%	Si	No
Reglas para Evaluar Soluciones OLAP	45,45%	18,18%	9,09%	9,09%	18,18%	Si	No
Características de los sistemas OLAP	45,45%	18,18%	9,09%	0,00%	27,27%	Si	No
Modelos OLAP	45,45%	18,18%	9,09%	0,00%	27,27%	Si	No
Data profiling (Perfilado de datos)	45,45%	18,18%	9,09%	9,09%	18,18%	Si	No
Mapa de datos lógico	36,36%	18,18%	18,18%	9,09%	18,18%	Si	No
Staging área (Datos temporales)	63,64%	18,18%	0,00%	9,09%	9,09%	Si	No
Estructuras de datos en un sistema ETL	27,27%	9,09%	27,27%	9,09%	27,27%	Si	No
Extracción de datos	18,18%	27,27%	18,18%	9,09%	27,27%	Si	No
SQL Avanzado	9,09%	18,18%	9,09%	0,00%	63,64%	No	No
Técnicas de Limpieza de datos	45,45%	18,18%	9,09%	18,18%	9,09%	Si	No
Dimensiones de carga lenta	63,64%	9,09%	0,00%	9,09%	18,18%	Si	No
Manejo de índices	18,18%	9,09%	18,18%	27,27%	27,27%	No	No
Manejo de particiones	18,18%	27,27%	27,27%	9,09%	18,18%	Si	No
Cargas incrementales	45,45%	27,27%	0,00%	0,00%	27,27%	Si	No
Carga histórica de datos	36,36%	36,36%	0,00%	0,00%	27,27%	Si	No

Contenidos de Data Warehouse	Grado de Dominio					Plan de capacitación 1	Plan de capacitación 2
	1	2	3	4	5		
Gestión de sistemas ETL	36,36%	18,18%	18,18%	9,09%	18,18%	Si	No
Metadata	27,27%	9,09%	27,27%	0,00%	36,36%	Si	No
Tipos de Metadata	27,27%	9,09%	36,36%	0,00%	27,27%	Si	No

Total respuestas: 11

8. Para cada herramienta listada indique el grado de dominio que posee; siendo 1 sin dominio de la herramienta y 5 dominio avanzado

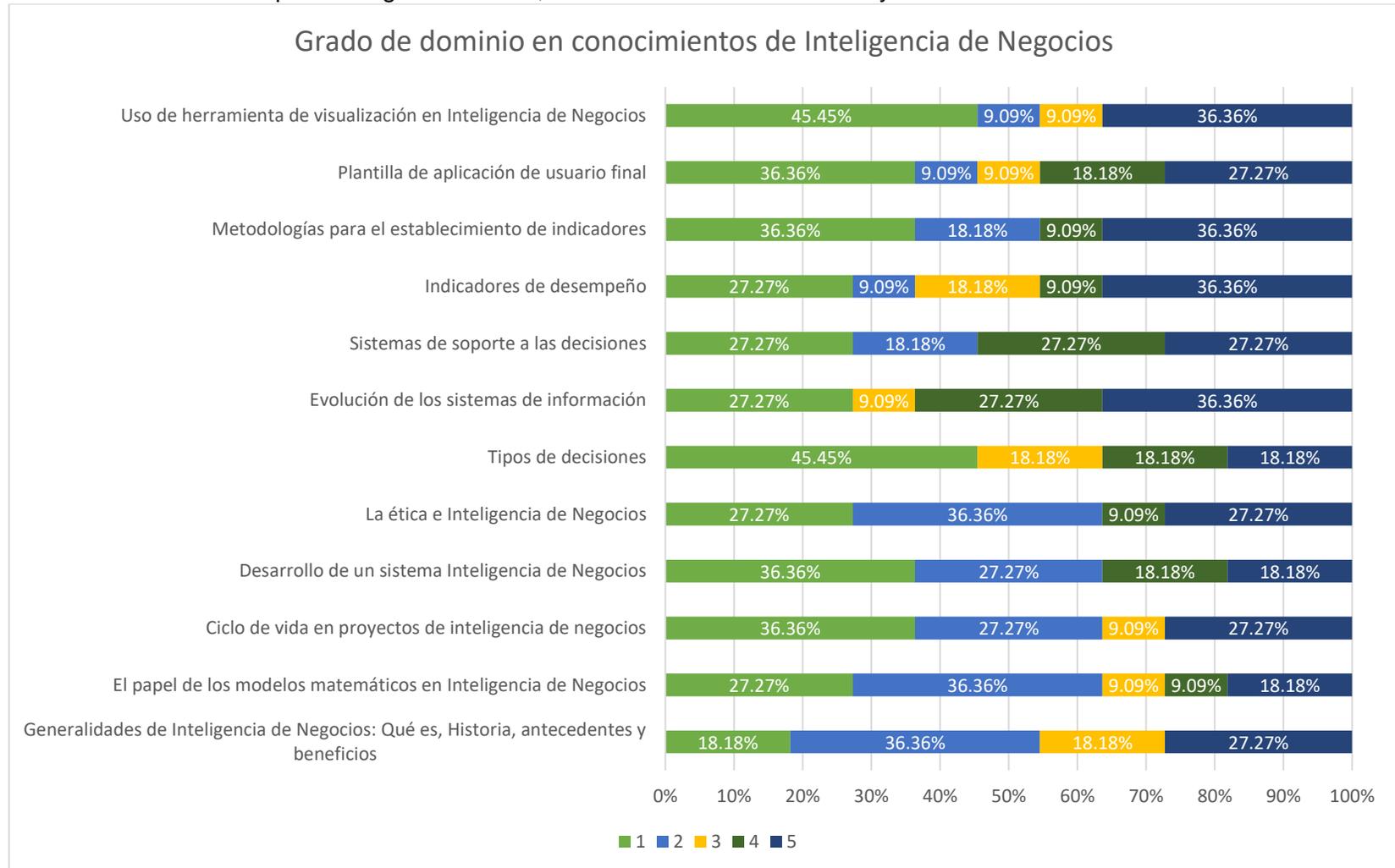


	Grado de Dominio					Plan de capacitación 1	Plan de capacitación 2
	1	2	3	4	5		
Herramientas de Data Warehouse							
Microsoft Integration Services	72,73%	9,09%	0,00%	9,09%	9,09%	Si	No
Talend	81,82%	0,00%	0,00%	9,09%	9,09%	Si	No
Pentaho	45,45%	18,18%	9,09%	9,09%	18,18%	Si	No
IBM InfoSphere DataStage	81,82%	0,00%	9,09%	9,09%	0,00%	Si	Si

Total respuestas: 11

Un encuestado agregó la siguiente herramienta: **JasperServer**. No se especificó grado de dominio.

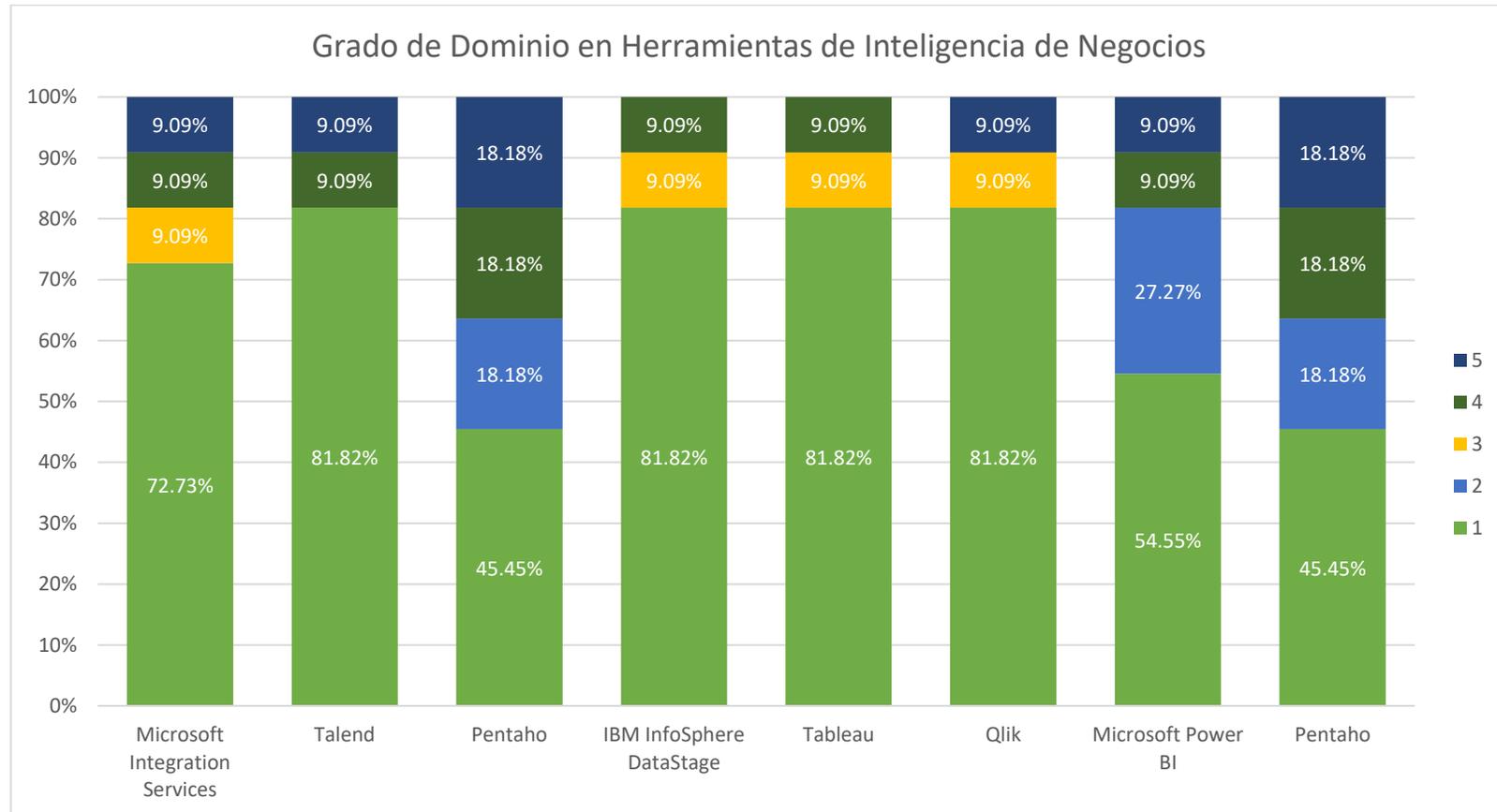
9. Seleccione la respuesta según su criterio; siendo 1 sin dominio del área y 5 dominio avanzado



Contenidos de Inteligencia de Negocios	Grado de Dominio					Plan de capacitación 1	Plan de capacitación 2
	1	2	3	4	5		
Generalidades de Inteligencia de Negocios: Qué es, Historia, antecedentes y beneficios	18,18%	36,36%	18,18%	0,00%	27,27%	Si	No
El papel de los modelos matemáticos en Inteligencia de Negocios	27,27%	36,36%	9,09%	9,09%	18,18%	Si	No
Ciclo de vida en proyectos de inteligencia de negocios	36,36%	27,27%	9,09%	0,00%	27,27%	Si	Si
Desarrollo de un sistema Inteligencia de Negocios	36,36%	27,27%	0,00%	18,18%	18,18%	Si	Si
La ética e Inteligencia de Negocios	27,27%	36,36%	0,00%	9,09%	27,27%	Si	Si
Tipos de decisiones	45,45%	0,00%	18,18%	18,18%	18,18%	No	No
Evolución de los sistemas de información	27,27%	0,00%	9,09%	27,27%	36,36%	No	No
Sistemas de soporte a las decisiones	27,27%	18,18%	0,00%	27,27%	27,27%	Si	No
Indicadores de desempeño	27,27%	9,09%	18,18%	9,09%	36,36%	Si	No
Metodologías para el establecimiento de indicadores	36,36%	18,18%	0,00%	9,09%	36,36%	Si	No
Plantilla de aplicación de usuario final	36,36%	9,09%	9,09%	18,18%	27,27%	Si	No
Uso de herramienta de visualización en Inteligencia de Negocios	45,45%	9,09%	9,09%	0,00%	36,36%	Si	Si

Total respuestas: 11

10. Para cada herramienta listada indique el grado de dominio que posee; siendo 1 sin dominio de la herramienta y 5 dominio avanzado

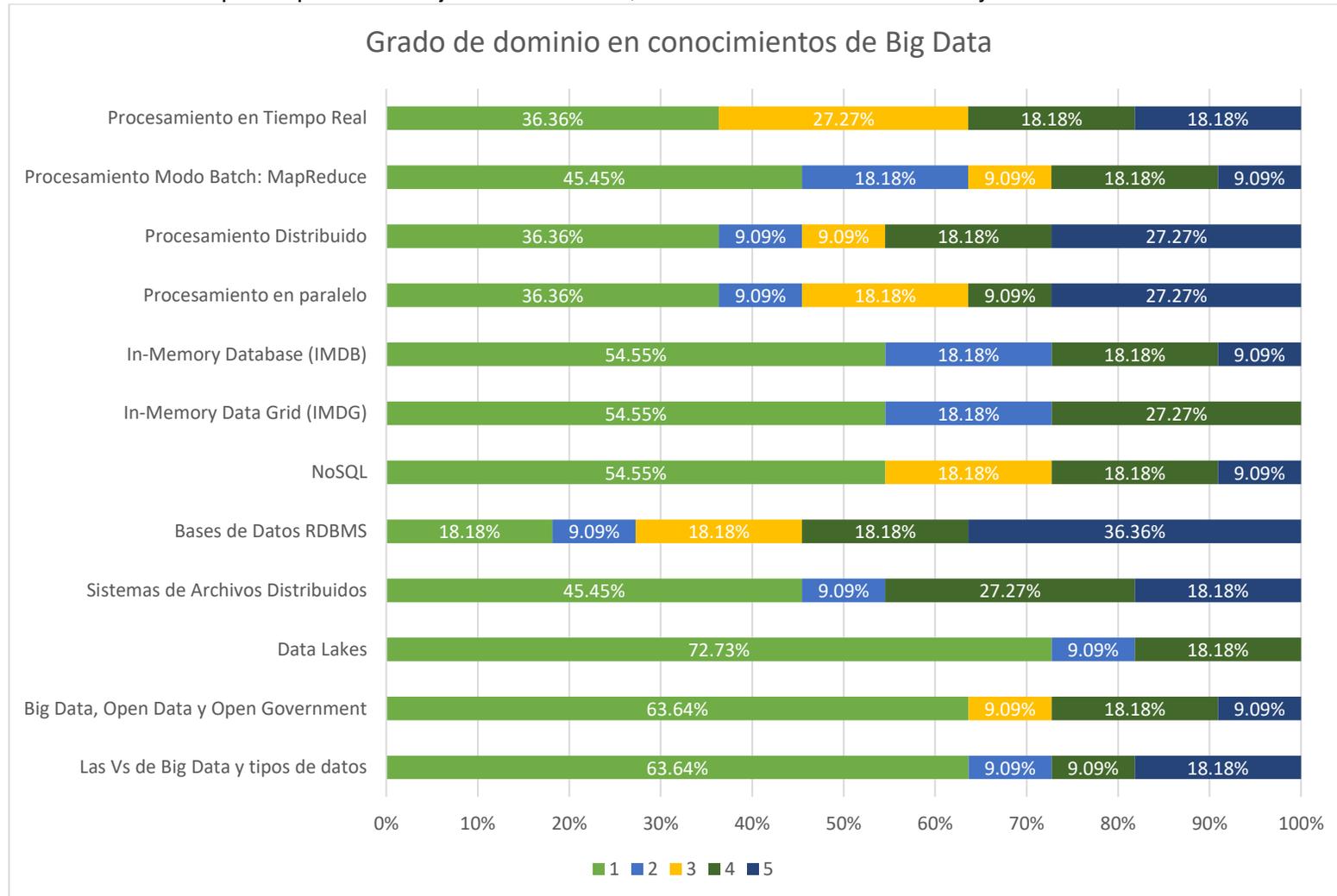


Herramientas de Inteligencia de Negocios	Grado de Dominio					Plan de Capacitación 1	Plan de Capacitación 2
	1	2	3	4	5		
Microsoft Integration Services	72,73%	0,00%	9,09%	9,09%	9,09%	Si	No
Talend	81,82%	0,00%	0,00%	9,09%	9,09%	Si	No
Pentaho	45,45%	18,18%	0,00%	18,18%	18,18%	Si	No
IBM InfoSphere DataStage	81,82%	0,00%	9,09%	9,09%	0,00%	Si	Si
Tableau	81,82%	0,00%	9,09%	9,09%	0,00%	Si	Si
Qlik	81,82%	0,00%	9,09%	0,00%	9,09%	Si	Si
Microsoft Power BI	54,55%	27,27%	0,00%	9,09%	9,09%	Si	Si
Pentaho	45,45%	18,18%	0,00%	18,18%	18,18%	Si	Si
Microstrategy	72,73%	9,09%	0,00%	18,18%	0,00%	Si	Si

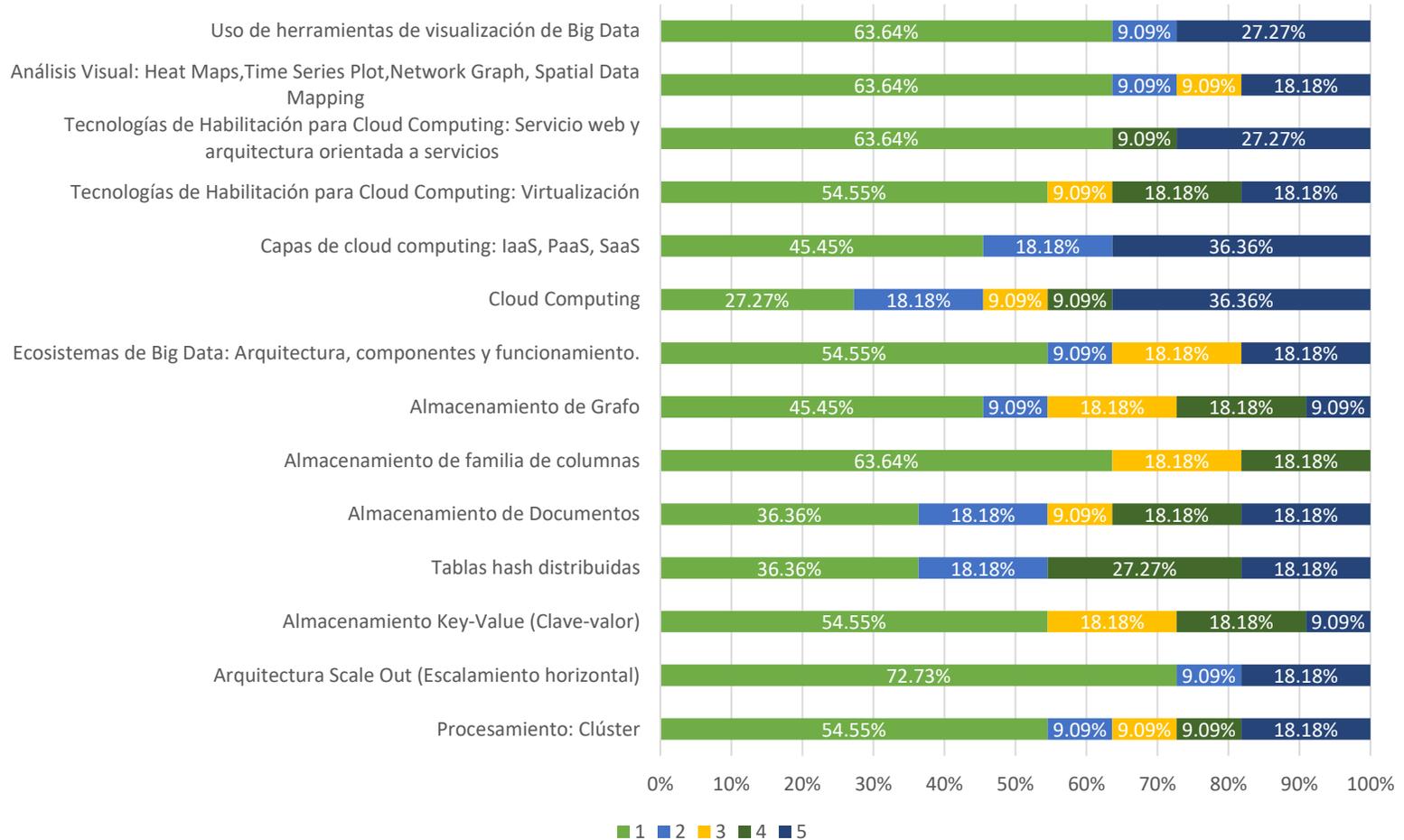
Total respuestas: 11

Un encuestado agregó la siguiente herramienta: **JasperServer**. No se especificó grado de dominio.

11. Seleccione la opción que más se ajuste a su criterio; siendo 1 sin dominio del área y 5 dominio avanzado



Grado de dominio en conocimientos de Big Data

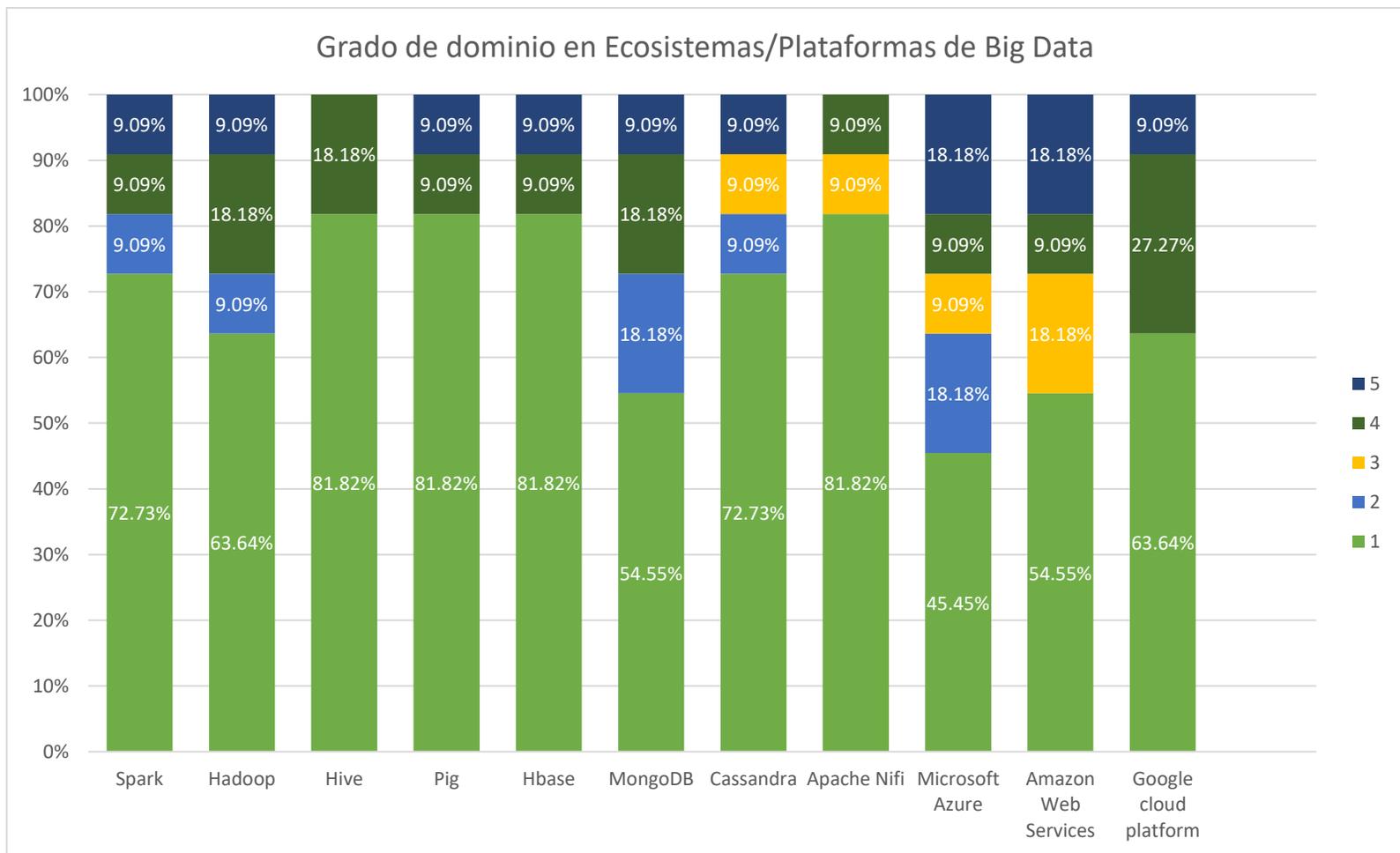


Conocimientos de Big Data	Grado de Dominio					Plan de Capacitación 1	Plan de Capacitación 2
	1	2	3	4	5		
Las Vs de Big Data y tipos de datos	63,64%	9,09%	0,00%	9,09%	18,18%	Sí	Sí
Big Data, Open Data y Open Government	63,64%	0,00%	9,09%	18,18%	9,09%	Sí	Sí
Data Lakes	72,73%	9,09%	0,00%	18,18%	0,00%	Sí	Sí
Sistemas de Archivos Distribuidos	45,45%	9,09%	0,00%	27,27%	18,18%	Sí	Sí
Bases de Datos RDBMS	18,18%	9,09%	18,18%	18,18%	36,36%	No	Sí
NoSQL	54,55%	0,00%	18,18%	18,18%	9,09%	Sí	Sí
In-Memory Data Grid (IMDG)	54,55%	18,18%	0,00%	27,27%	0,00%	Sí	Sí
In-Memory Database (IMDB)	54,55%	18,18%	0,00%	18,18%	9,09%	Sí	Sí
Procesamiento en paralelo	36,36%	9,09%	18,18%	9,09%	27,27%	Sí	Sí
Procesamiento Distribuido	36,36%	9,09%	9,09%	18,18%	27,27%	Sí	Sí
Procesamiento Modo Batch: MapReduce	45,45%	18,18%	9,09%	18,18%	9,09%	Sí	Sí
Procesamiento en Tiempo Real	36,36%	0,00%	27,27%	18,18%	18,18%	Sí	Sí
Procesamiento: Clúster	54,55%	9,09%	9,09%	9,09%	18,18%	Sí	Sí
Arquitectura Scale Out (Escalamiento horizontal)	72,73%	9,09%	0,00%	0,00%	18,18%	Sí	Sí
Almacenamiento Key-Value (Clave-valor)	54,55%	0,00%	18,18%	18,18%	9,09%	Sí	Sí
Tablas hash distribuidas	36,36%	18,18%	0,00%	27,27%	18,18%	Sí	Sí
Almacenamiento de Documentos	36,36%	18,18%	9,09%	18,18%	18,18%	Sí	Sí
Almacenamiento de familia de columnas	63,64%	0,00%	18,18%	18,18%	0,00%	Sí	Sí
Almacenamiento de Grafo	45,45%	9,09%	18,18%	18,18%	9,09%	Sí	Sí
Ecosistemas de Big Data: Arquitectura, componentes y funcionamiento.	54,55%	9,09%	18,18%	0,00%	18,18%	Sí	Sí
Cloud Computing	27,27%	18,18%	9,09%	9,09%	36,36%	Sí	Sí

Conocimientos de Big Data	Grado de Dominio					Plan de Capacitación 1	Plan de Capacitación 2
	1	2	3	4	5		
Capas de cloud computing: IaaS, PaaS, SaaS	45,45%	18,18%	0,00%	0,00%	36,36%	Sí	Sí
Tecnologías de Habilitación para Cloud Computing: Virtualización	54,55%	0,00%	9,09%	18,18%	18,18%	Sí	Sí
Tecnologías de Habilitación para Cloud Computing: Servicio web y arquitectura orientada a servicios	63,64%	0,00%	0,00%	9,09%	27,27%	Sí	Sí
Análisis Visual: Heat Maps, Time Series Plot, Network Graph, Spatial Data Mapping	63,64%	9,09%	9,09%	0,00%	18,18%	Sí	Sí
Uso de herramientas de visualización de Big Data	63,64%	9,09%	0,00%	0,00%	27,27%	Sí	Sí

Total respuestas: 11

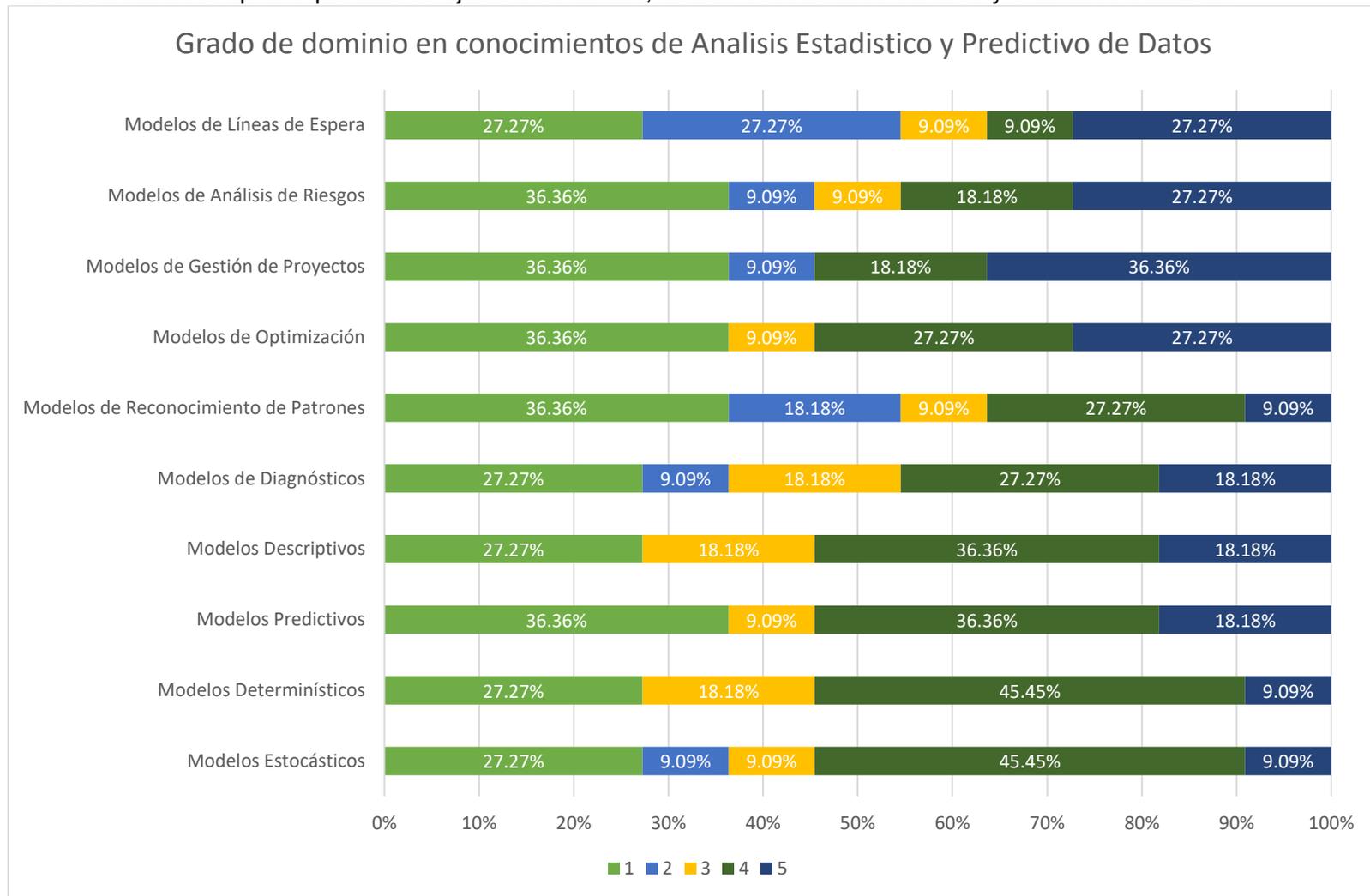
12. Para cada ecosistema o plataforma listada indique el grado de dominio que posee; siendo 1 sin dominio de la herramienta y 5 dominio avanzado



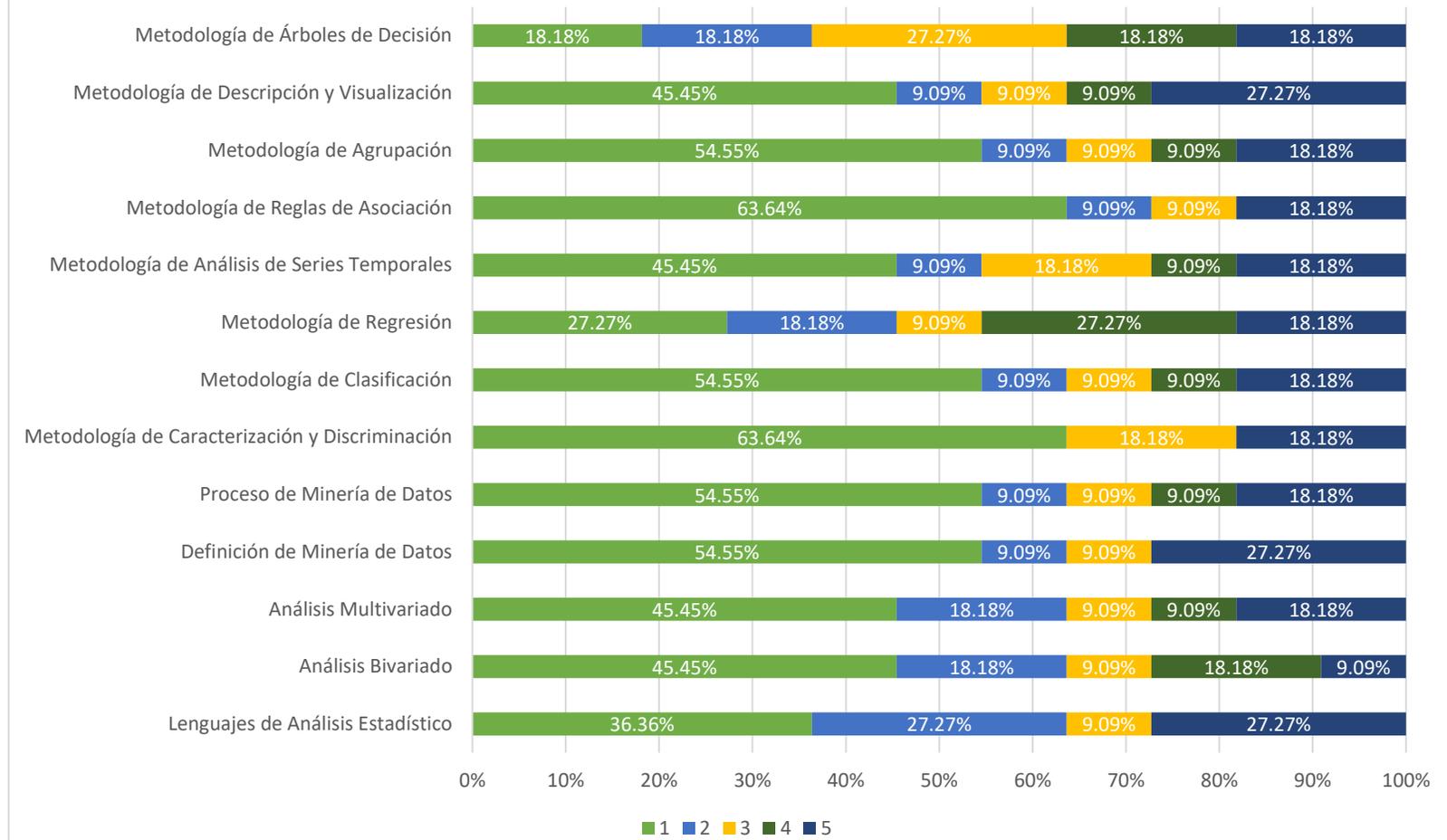
Ecosistemas/Plataformas de Big Data	Grado de Dominio					Plan de Capacitación 1	Plan de Capacitación 2
	1	2	3	4	5		
Spark	72,73%	9,09%	0,00%	9,09%	9,09%	Sí	Sí
Hadoop	63,64%	9,09%	0,00%	18,18%	9,09%	Sí	Sí
Hive	81,82%	0,00%	0,00%	18,18%	0,00%	Sí	Sí
Pig	81,82%	0,00%	0,00%	9,09%	9,09%	Sí	Sí
Hbase	81,82%	0,00%	0,00%	9,09%	9,09%	Sí	Sí
MongoDB	54,55%	18,18%	0,00%	18,18%	9,09%	Sí	Sí
Cassandra	72,73%	9,09%	9,09%	0,00%	9,09%	Sí	Sí
Apache Nifi	81,82%	0,00%	9,09%	9,09%	0,00%	Sí	Sí
Microsoft Azure	45,45%	18,18%	9,09%	9,09%	18,18%	Sí	Sí
Amazon Web Services	54,55%	0,00%	18,18%	9,09%	18,18%	Sí	Sí
Google cloud platform	63,64%	0,00%	0,00%	27,27%	9,09%	Sí	Sí

Total respuestas: 11

13. Seleccione la opción que más se ajuste a su criterio; siendo 1 sin dominio del área y 5 dominio avanzado



Grado de dominio en conocimientos de Analisis Estadístico y Predictivo de Datos

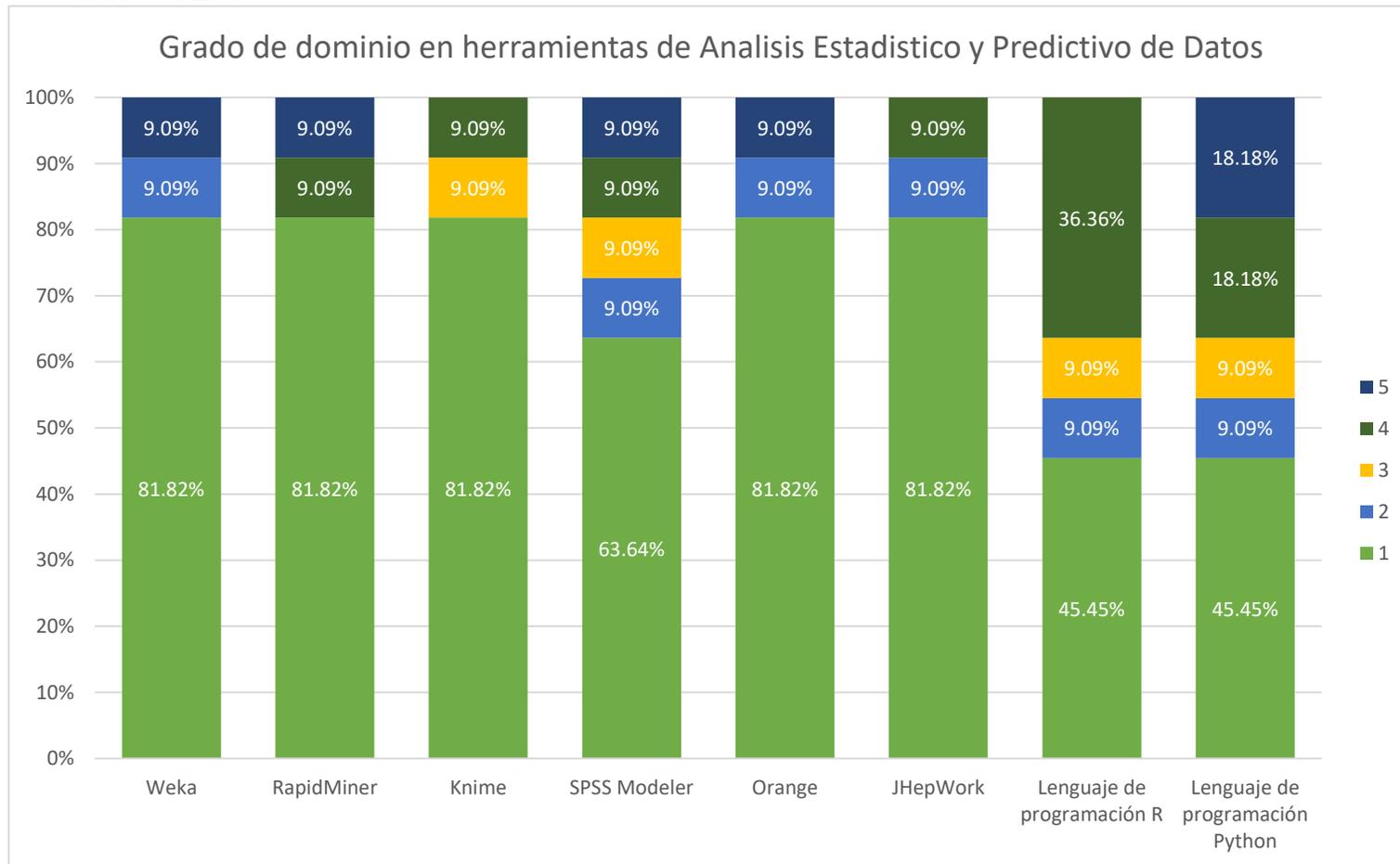


Conocimientos en Análisis Estadístico y Predictivo de Datos	Grado de Dominio					Plan de capacitación 2	Pan de capacitación 1
	1	2	3	4	5		
Modelos Estocásticos	27,27%	9,09%	9,09%	45,45%	9,09%	No	No
Modelos Determinísticos	27,27%	0,00%	18,18%	45,45%	9,09%	No	No
Modelos Predictivos	36,36%	0,00%	9,09%	36,36%	18,18%	No	No
Modelos Descriptivos	27,27%	0,00%	18,18%	36,36%	18,18%	No	No
Modelos de Diagnósticos	27,27%	9,09%	18,18%	27,27%	18,18%	Si	No
Modelos de Reconocimiento de Patrones	36,36%	18,18%	9,09%	27,27%	9,09%	Si	No
Modelos de Optimización	36,36%	0,00%	9,09%	27,27%	27,27%	No	No
Modelos de Gestión de Proyectos	36,36%	9,09%	0,00%	18,18%	36,36%	No	No
Modelos de Análisis de Riesgos	36,36%	9,09%	9,09%	18,18%	27,27%	Si	No
Modelos de Líneas de Espera	27,27%	27,27%	9,09%	9,09%	27,27%	Si	No
Lenguajes de Análisis Estadístico	36,36%	27,27%	9,09%	0,00%	27,27%	Si	No
Análisis Bivariado	45,45%	18,18%	9,09%	18,18%	9,09%	Si	No
Análisis Multivariado	45,45%	18,18%	9,09%	9,09%	18,18%	Si	No
Definición de Minería de Datos	54,55%	9,09%	9,09%	0,00%	27,27%	Si	No
Proceso de Minería de Datos	54,55%	9,09%	9,09%	9,09%	18,18%	Si	No
Metodología de Caracterización y Discriminación	63,64%	0,00%	18,18%	0,00%	18,18%	Si	No
Metodología de Clasificación	54,55%	9,09%	9,09%	9,09%	18,18%	Si	No
Metodología de Regresión	27,27%	18,18%	9,09%	27,27%	18,18%	Si	No
Metodología de Análisis de Series Temporales	45,45%	9,09%	18,18%	9,09%	18,18%	Si	No
Metodología de Reglas de Asociación	63,64%	9,09%	9,09%	0,00%	18,18%	Si	No
Metodología de Agrupación	54,55%	9,09%	9,09%	9,09%	18,18%	Si	No

Conocimientos en Análisis Estadístico y Predictivo de Datos	Grado de Dominio					Plan de capacitación 2	Pan de capacitación 1
	1	2	3	4	5		
Metodología de Descripción y Visualización	45,45%	9,09%	9,09%	9,09%	27,27%	Si	No
Metodología de Árboles de Decisión	18,18%	18,18%	27,27%	18,18%	18,18%	Si	No

Total respuestas: 11

14. Para cada herramienta listada indique el grado de dominio que posee; siendo 1 sin dominio de la herramienta y 5 dominio avanzado



Herramientas de Minería de Datos	Grado de Dominio					Plan de capacitación 1	Plan de capacitación 2
	1	2	3	4	5		
Weka	81,82%	9,09%	0,00%	0,00%	9,09%	Si	Si
RapidMiner	81,82%	0,00%	0,00%	9,09%	9,09%	Si	No
Knime	81,82%	0,00%	9,09%	9,09%	0,00%	Si	Si
SPSS Modeler	63,64%	9,09%	9,09%	9,09%	9,09%	Si	No
Orange	81,82%	9,09%	0,00%	0,00%	9,09%	Si	Si
JHepWork	81,82%	9,09%	0,00%	9,09%	0,00%	Si	Si
Lenguaje de programación R	45,45%	9,09%	9,09%	36,36%	0,00%	Si	No
Lenguaje de programación Python	45,45%	9,09%	9,09%	18,18%	18,18%	Si	No

Total respuestas: 11

15. Siéntase libre de especificar los contenidos de las asignaturas contempladas que no se mencionaron y que a su criterio pueden agregarse

1 - Calidad de los datos 2 - Técnicas indirectas de minería de datos, que garanticen la calidad de los datos
3- Un elemento de tendencias de las áreas de Analítica de Datos y Big Data... o solo de Datos.... como: Las bases de datos con procesamiento basado en GPUs, chatbots, etc. Cada año podría variar esa temática.
4- En todas las materias que he tenido la oportunidad de evaluar el desempeño de los estudiantes es evidente el bajo o nulo nivel de análisis de los estudiantes, todos los contenidos de este instrumento requieren que la persona que posea estos conocimientos sea analítico. Ese requisito no es facilitado en el aprendizaje cuando solo se tiene acceso a herramientas privadas o de pago, limitan a los usuarios finales en ese sentido.

BIBLIOGRAFÍA

- ACM (2017). Curricula Recommendations. julio 30,2017, de ACM Sitio web: <http://www.acm.org/education/curricula-recommendations>
- Topi H., Valacich J., Wright R., Kaiser K., Nunamaker J.F., Sipior J. & Vreede G.J. 2010. IS 2010 Curriculum Guidelines for Undergraduate Degree Programs in Information Systems. En IS 2010 Curriculum Guidelines (1, 88) .
- Organización Internacional del Trabajo. (Septiembre 04, 2012). Definiciones de competencia en las instituciones dedicadas a la formación y desarrollo de los recursos humanos. Agosto 3,2017, de OIT Sitio web: <http://www.oitcinterfor.org/p%C3%A1gina-libro/definiciones-competencia-instituciones-dedicadas-formaci%C3%B3n-desarrollo-recursos-humanos>
- Alfonso, R. F (abril 2008). Competencias Educativas. Agosto 03, 2017, de Instituto de Educación Media Superior de la ciudad de México Sitio web: http://academicos.iems.edu.mx/cired/docs/tg/macroacademiaquimica/Competencias%20educativas.aprendizajegenuino_FeitoAlonso.pdf
- Argudin, M. L.. (2008). Enfoques educativos. Agosto 3, 2017. Sitio web: <http://hadoc.azc.uam.mx/menu/menu.htm>
- Amado-Salvatierra, H.R.; Hilera González, J.R. y Otón Tortosa, S. (en prensa). Formalización de un marco metodológico para la implementación de un proyecto educativo virtual accesible (ACAI). Educación XX1.
- O'Brien, J & Marakas, G. (2006). Sistemas de Información gerencial. México, D.F.: McGraw-Hill.
- Laudon, K. & Laudon, J. (2008). Sistemas de información gerencial. México: Pearson Educación.
- OMG. (2011). Business Process Model and Notation.
- Robert S. Kaplan. (2010). Conceptual Foundations of the Balanced Scorecard. Harvard Business School.
- JMACOE. (2017). 5 de los mejores software de minería de datos de Código Libre y Abierto. Septiembre 12, 2017, de El rincón de JMACOE Sitio web: http://blog.jmacoe.com/gestion_ti/base_de_datos/5-mejores-software-mineria-datos-codigo-libre-abierto/
- LinkedIn. (2017). Mathworks. Septiembre 12,2017, de LinkedIn Sitio web: <https://www.linkedin.com/company/1194036/>
- Rouse M.. (2016). Definition Hadoop. Septiembre 12,2017, de TechTarget Sitio web: <http://searchcloudcomputing.techtarget.com/definition/Hadoop>
- Aiteco. (2015). Diagrama de Dispersión. Septiembre 14, 2017, de Aiteco Sitio web: <https://www.aiteco.com/diagrama-de-dispersion/>
- Conomipedia. (2017). Coeficiente de correlación lineal. Septiembre 14, 2017, de Conomipedia Sitio web: <http://economipedia.com/definiciones/coeficiente-de-correlacion-lineal.html>

- Wikipedia. (2017). Regresión lineal múltiple. Septiembre 14, 2017, de Wikipedia Sitio web: https://es.wikipedia.org/wiki/Regresi%C3%B3n_lineal#Regresi.C3.B3n_lineal_m.C3.BA1ltiple
- IBM. (2017). Regresión Logística. Septiembre 14, 2017, de IBM Sitio web: https://www.ibm.com/support/knowledgecenter/es/SSLVMB_22.0.0/com.ibm.sps.s.statistics.help/spss/regression/idh_lreg.htm
- Minitab. (2017) ¿Qué es un intervalo de confianza? Septiembre 14, 2017, de Minitab Sitio web: <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/basics/what-is-a-confidence-interval/>
- Salinas, J. (2005). Teorema de Bayes. Septiembre 14,2017, de Universidad de Granada Sitio web: <http://www.ugr.es/~jsalinas/bayes.htm>
- Wikipedia. (2017). Algoritmo de agrupamiento. Septiembre 14, 2017, de Wikipedia Sitio web: https://es.wikipedia.org/wiki/Algoritmo_de_agrupamiento
- Gollapudi, S. (2016). Practical Machine Learning. Birmingham B3 2PB, UK: Packt Publishing Ltd.
- Aiteco. (2015). Diagrama de Pareto. Septiembre 14, 2017, de Aiteco Sitio web: <https://www.aiteco.com/diagrama-de-pareto/>
- Data Warehouse4u. (2008). What are Slowly Changing Dimensions? Septiembre 14, 2017, de Data Warehouse4u Sitio web: <https://www.aiteco.com/diagrama-de-pareto/>
- González, C. (2017). Lógica Difusa. Septiembre 14, 2017, de Escuela Superior de Informática Sitio web: http://www.esi.uclm.es/www/cglez/downloads/docencia/2011_Softcomputing/LogicaDifusa.pdf
- Sunil, R. (2017). Understanding Support Vector Machine algorithm. Septiembre 14, 2017, de Analytics Vidhya Sitio web: <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>
- Wikipedia. (2017). SAS (software). Septiembre 14, 2017, de Wikipedia Sitio web: [https://en.wikipedia.org/wiki/SAS_\(software\)](https://en.wikipedia.org/wiki/SAS_(software))
- IBM. (2017). IBM Analytics. Septiembre 14, 2017, de IBM Sitio web: <https://www.ibm.com/analytics/us/en/>
- Wikipedia. (2017). Minitab. Septiembre 15, 2017, de Wikipedia Sitio web: <https://es.wikipedia.org/wiki/Minitab>
- Wikipedia. (2017). RStudio. Septiembre 15, 2017, de Wikipedia Sitio web: <https://es.wikipedia.org/wiki/RStudio>
- Tableau. (2017). Inteligencia de negocios para su gente. Septiembre 15, 2017, de Tableau Sitio web: <https://www.tableau.com/es-es/resource/business-intelligence>
- Wikipedia. (2017). Weka (aprendizaje automático). Septiembre 15, 2017, de Wikipedia Sitio web: [https://es.wikipedia.org/wiki/Weka_\(aprendizaje_autom%C3%A1tico\)](https://es.wikipedia.org/wiki/Weka_(aprendizaje_autom%C3%A1tico))

- Sistemas de información | Gerencie.com. (2017). Gerencie.com. Recuperado 13 Septiembre 2017, Sitio web: <https://www.gerencie.com/sistemas-de-informacion.html>
- Apache Ambari - Hortonworks. (2017). Hortonworks. Recuperado 13 Septiembre 2017, Sitio Web https://es.hortonworks.com/apache/ambari/#section_1
- Zaforas, M., & Zaforas, M. (2017). Cassandra, la dama de las bases de datos NoSQL - Paradigma. Paradigma. Recuperado Septiembre 13, 2017, Sitio Web <https://www.paradigmadigital.com/dev/cassandra-la-dama-de-las-bases-de-datos-nosql/>
- IBM Knowledge Center. (2018). Ibm.com. Recuperado Marzo 25, 2018, a partir de https://www.ibm.com/support/knowledgecenter/en/SSRL5J_1.1.0/com.ibm.rational.raer.help.doc/topics/c_ods_fact
- Sistemas heredados Caro Gutiérrez, M., Bocca, J., & Campos, D. (2002). MIGRACION DE SISTEMAS HEREDADOS: UNA METODOLOGÍA DE APOYO BASADA EN EL USO DE HERRAMIENTAS DE KDD (KNOWLEDGE DISCOVERY IN DATABASES). Revista Ingeniería de Sistemas, 16(1). Recuperado a partir de [http://migracion.de.sistemas.heredados:una.metodologia.de.apoyo.basada.en.el.uso.de.herramientas.de.kdd.\(knowledge.discovery.in.databases\)](http://migracion.de.sistemas.heredados:una.metodologia.de.apoyo.basada.en.el.uso.de.herramientas.de.kdd.(knowledge.discovery.in.databases))
- Cookies (2018). ¿Qué es Cookie? - Su Definición, Concepto y Significado. ConceptoDefinicion.de. Recuperado Marzo 25, 2018, a partir de <http://conceptoDefinicion.de/cookie/>
- ¿Qué es internet de las cosas? | Tecnología de IoT | SAP. (2018). SAP. Recuperado Marzo 25, 2018, a partir de <https://www.sap.com/latinamerica/trends/internet-of-things.html>
- Logs. Universo de la Comunicación. Ur.mx. Recuperado 25 March 2018, a partir de <http://www.ur.mx/Default.aspx?tabid=3804&language=en-US>
- Rose, K. (2013). A Guide to the Project Management Body of Knowledge (PMBOK® Guide)-Fifth Edition. Project Management Journal, 44(3), e1-e1. doi:10.1002/pmj.21345
- Habilidades técnicas. (2018). Cca.org.mx. Recuperado 26 Marzo 2018, a partir de http://www.cca.org.mx/cca/cursos/administracion/artra/habad/habadm/habtec_had.htm
- Universidad de El Salvador, Facultad de Ingeniería y Arquitectura. (n.d). Plan Estratégico Período 2015-2019 (n.d). Recuperado de Unidad de Acceso a la Información Pública de Universidad de El Salvador.
- Definición de Curricular. (2018). Definición ABC. Recuperado Abril 29, 2018, a partir de <https://www.definicionabc.com/general/curricular.php>
- Psicomentarios - ¿Qué es el sílabo?. (2018). Psicomentarios. Recuperado Abril 29, 2018, de <http://blogs.up.edu.pe/psicomentarios/que-es-el-silabo/>
- Malla Curricular | Ernesto Yturralde Worldwide Inc.. (2018). Mallacurricular.com. Recuperado Abril 29, 2018, de <http://www.mallacurricular.com/>.