NOVA IMS Information Management School

# MGI

Master Degree Program in

**Information Management**

# Implementing an SQL Based ETL Platform for a Business Intelligence Solution

André Vieira da Silva

Internship Report

presented as partial requirement for obtaining the Master Degree Program in Information Management

**NOVA Information Management School**
**Instituto Superior de Estatística e Gestão de Informação**
Universidade Nova de Lisboa

**NOVA Information Management School**
**Instituto Superior de Estatística e Gestão de Informação**
Universidade Nova de Lisboa

# IMPLEMENTING AN SQL BASED ETL PLATFORM FOR A BUSINESS INTELLIGENCE SOLUTION

By

André Vieira da Silva

Project Work presented as partial requirement for obtaining the Master's degree in Information Management, with a specialization in Knowledge Management and Business Intelligence

**Supervisor:** Flávio L. Pinheiro

November 2022

# STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

André Vieira Silva

*Lisboa, 29 November 2022*

# ABSTRACT

The exponential growth and development of information technology in the last twenty years has compelled most industry segments to shift from focusing on core business to adopting digitally sophisticated and data-driven processes. Those who have followed its growth have benefited, but unfortunately, just a small percentage of them do. Having information systems that just hold a vast volume of data is no longer sufficient for businesses. To gain a competitive advantage, these businesses must make well-informed decisions. Every firm, regardless of industry, has access to a wealth of data that it can utilize to its advantage. This is where Business Intelligence comes in. Business intelligence enables these companies to make better use of their data by providing previously unusable data in an intelligible and interpretable format. This internship report aims to cover the development of the data warehousing and data analytics for HROps, a product owned by BI4ALL. HROps is being developed with the goal of facilitating, centralizing, and making people management processes in organizations more efficient. I will be working on a low-cost SQL based ETL Framework using T-SQL for developing standard ETL processes. I will also be working and creating Power BI dashboards and reports to gather useful information from the data collected.

# KEYWORDS

Data; Business Intelligence; Analytics; Data Visualization

# INDEX

# LIST OF FIGURES

# LIST OF TABLES

# 1. INTRODUCTION

"One of the most important assets of any organization is its information" (Kimball and Ross, 2013, p. 38). The data in our modern world grew in size and complexity as humans advanced both in culture as in technology. Data is what powers businesses, and without it organizations can't function. But having huge piles of datasets created challenges for organizations that would like to use this data for their advantage. This is commonly known as Big Data. The amount of data currently being collected and generated is expanding at an astounding rate. The fact that this growth rate is quickly outpacing our capacity to develop adequate systems to manage the data effectively and analyze it to extract important information for decision making, poses a significant challenge for organizations (Kaisler et al.). Right now, having information systems that just hold a vast volume of data is no longer sufficient for businesses. To gain a competitive advantage, these businesses must use it to make well-informed decisions.

Every firm, regardless of industry, has access to a wealth of data that it can utilize to its advantage. This is where Business Intelligence comes in. Due to its capacity to offer complicated and competitive information inputs for the decision process, business intelligence (BI) systems have gained a lot of attention from businesses in the recent period (Ain et al., 2019). "The term Business Intelligence (BI) refers to the technologies, applications, strategies, and practices used to collect, analyze, integrate, and present pertinent business information" (Efremov, 2019). BI allows better use of an organization's access to vast amounts of information by providing unusable data in an intelligible and interpretable format. In essence, BI allows you to combine data from numerous sources, evaluate it into an interpretable style, and then disseminate it to the appropriate stakeholders. This enables businesses to see the big picture and make informed decisions.

In order to get a better understanding of the fundamentals of a BI application, it is presented in Figure 1 the main components to consider according to the Kimball architecture: source systems, ETL process, data presentation area, and business intelligence applications (Kimball and Ross, 2013).

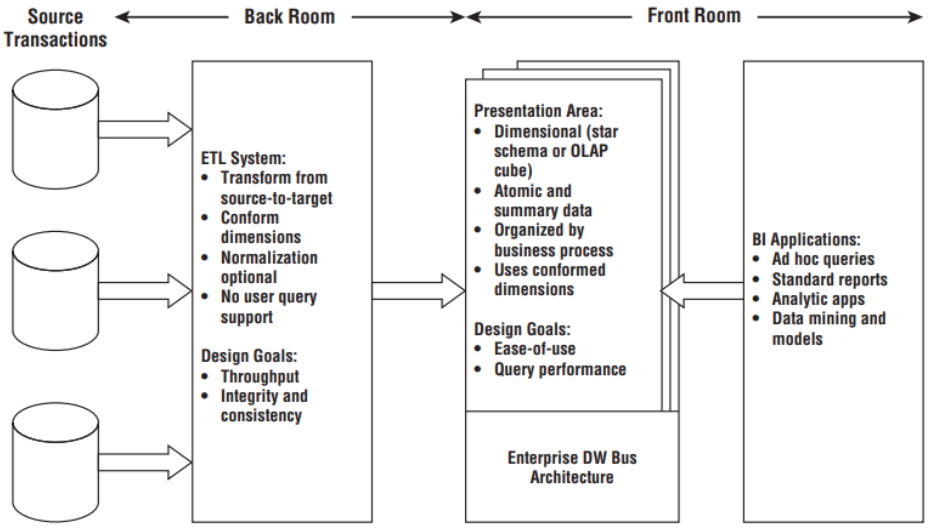


Figure 1.1 - Core elements of the Kimball BI architecture. Source: (Kimball and Ross, 2013, p.55)

Operational source systems work as data repositories that capture business's transactions. It is safe to assume that organizations have little control of the format and content of this data, and maintain little historical data. The source systems' biggest concerns are processing efficiency and availability (Kimball and Ross, 2013). Extraction, Transformation and Loading (ETL) consists of extracting data from different sources, manipulating it according to business needs, and loading the modified and clean data into a Data Warehouse (IBM Cloud Education, 2020). Because of the in-depth analytics data it provides, ETL function lies at the core of Business Intelligence systems. Data organization, storage, and direct querying by users, report writers, and other BI applications all take place in the presentation area. This is the layer that businesses see (Front Room), and it is now widely accepted that this data should be made available and presented in the form of dimensional modeling (Kimball and Ross, 2013). The final piece of the puzzle is the BI application, this "loosely refers to the range of capabilities provided to business users to leverage the presentation area for analytic decision making", according to (Kimball and Ross, 2013). In the case of this internship the final BI application will be in the form of data reports.

This report will focus on the activities performed in the implementation of an SQL based ETL Framework for a Human Resources (HR) product owned by BI4ALL called HROps. It is often said that human resources are the brain of every organization, because employees are the foundation of every business. HR used to be primarily concerned with hiring, firing, and yearly salary review. But in recent years, HR has been favorably reframed and has expanded in scope (Sands, 2021). Data collected by HR can be overwhelming, and its importance can be easily overlooked, so having a system in place that can collect data, organize it and successfully use it to an organization's benefit can be a game changer. The fact is that human resource data is more than just hours worked or vacations taken. If handled properly, this data has the potential to give valuable insights to issues including leadership development, productivity enhancement, and increased retention, among others (Erts, 2020). The goal of the HROps platform and this internship is to have this data cleaned and organized for later consumption. But clean and organized data isn't of much use if it can't help bringing valuable insights to an organization. This is where data visualization and reporting come in. The final goal of the project is to generate visualizations and gather important and useful information that can be presented and managed to/by the stakeholders.

### 1.1. ORGANIZATIONAL CONTEXT

The internship will be conducted at the company BI4ALL. BI4ALL is a Business Intelligence and Data Analytics Portuguese consulting company, based in Lisbon. Founded in 2004, BI4ALL has continued with the purpose of transforming data into insights, and its mission is to help organizations be more agile, flexible, and vigorous, anticipate the unpredictable and adapt quickly to market changes, thus being better prepared for the future.

Considered one of the Top 10 Big Data Analytics Consulting/Service Companies in Europe 2020, BI4ALL helps organizations in the Digital Transformation and Data Strategy process. Their main areas of competence and excellence focus on:

- Analytics
- Big Data
- Data Science

- Artificial Intelligence
- Data Visualizations
- Corporate Performance Management
- Software Engineering

The company has extensive expertise responding to a wide range of issues and business requirements from many industries and assists businesses in prospering and preparing for the challenges of the present and future. Some sectors include Government and Public Services, Food and Beverage, Automotive, Banking and Financial Services, Legal, Retail, and many more.

## 1.2. Project Context and Motivation

The motivation for this internship project was to help the Human Resources department make better use of the vast amount of information stored, make well-informed data driven decisions, having information regarding job opportunities, candidates, recruitment process, etc. cleaned and organized, by building a data model where they could easily have access to the information in order to produce dashboards and reports using tools like Power BI.

In order to achieve this, I will be working on a low-cost SQL based framework, based on Microsoft technology, for developing standard ETL processes in the universe of traditional BI projects, using only T-SQL Microsoft. This framework is based on a set of procedures and assumptions that allow accelerating the development of traditional BI projects. The idea is to reuse processes to perform different operations considering various parameterizations.

The platform is called HROps and is a product of intellectual property owned by BI4ALL developed with the goal of facilitating, centralizing, and making people management processes in organizations more efficient. Through this product, human resources management is carried out in a centralized and automated manner, tailored to the needs and strategy of each organization. Although the scope of the platform HROps will include all data regarding HR, for this internship the main area worked on and developed was the recruitment process. When I first started this project the HROps platform was in the early stages of development, the UI was built but because there was no dimensional model behind it the data was just stored in a database, without being able to use it in an efficient manner. In short, the goal was to translate the data from the HROps platform to dashboards and reports for the HR personnel to analyze and extract valuable insights. The objective was not only to implement a dimensional model and to generate dashboards, but to do it in a way that could be easily understood by non-technical HR personnel, so that they could manage the reports according to their needs.

# 2. LITERATURE REVIEW

## 2.1. BUSINESS INTELLIGENCE

Demand for Business Intelligence applications continues to grow as a way to reduce costs, improve service quality, and develop better decision-making processes. It is undeniable that in today's corporate environment, information has emerged as a primary source of competitive advantage. It cannot be disputed that information has become a source of major competitive advantage in today's business world (Foley and Guillemette, 2010).

### 2.1.1. Concepts

The first known use of the term Business Intelligence dates to 1868, by Richard M. Davens (Devens, 1868, p.210). However, it was only in 1958 that BI appeared in literature again, when Hans Peter Luhn used the term in an article. The article, titled "A Business Intelligence System", described BI as "an automatic system…developed to disseminate information to the various sections of any industrial, scientific, or government organization." (Heinze, 2014). Even though the term "business intelligence" has been around for a while, corporations have only recently started to invest more time and effort in researching the topic (Foley and Guillemette, 2010). Today, according to the Gartner IT Glossary ("Analytics and Business Intelligence (Abi)", 2019), BI is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance. Simply put BI is a management strategy used to create a more structured and effective approach to decision making (Nelson).

To make better decisions, organizations require better data and information. Professionals can use business intelligence to make wise and informed decisions about how their organization is run, thus making informed decisions and helping build a strong competitive advantage (Foley and Guillemette, 2010).

One real world example of this is the BI system implemented by Lotte.com. With 13 million customers, Lotte.com is the top online mall in Korea. Having a huge number of daily visitors on their website, the company wanted to know what made customers abandon their shopping cart. To help executives understand the challenge, customer experience analytics were implemented to perform targeted marketing, and comprehend consumer behavior. The knowledge obtained from the BI system helped increase customer loyalty and sales in just one year. This result was possible by finding the reasons behind shopping cart abandonment. (Morris, 2021)

### 2.1.2. State of the Art

Organizations consider technology, data, and analytics to be a transformative force. In order to enhance reporting and decision-making, many are integrating Business Intelligence systems (Rikhardsson and Yigitbasioglu, 2018)

A few ways that business intelligence can help companies make smarter, data-driven decisions include but are not limited to: smarter decision-making process, increased organizational efficiency, better

customer knowledge, reduce business costs, eliminate manual tasks, increased competitive advantage over business rivals, identifying market trends, trusted and governed data (Miller, n.d; Tableau, n.d; Datapine, n.d).

There are three types of analytics that businesses use to drive their decision making. The first one, descriptive analytics, is the process of collecting, cleaning, and presenting data to get immediate insights and getting data-driven answers to factual questions. Descriptive analytics is used to answer the question “Why did it happen?” It is characterized by techniques such as drill-down, data discovery, data mining, and correlations. Predictive analytics uses the findings of descriptive and diagnostic analytics to help understand likelihood and/or predict future events, such as providing revenue forecasts, using data science techniques. Finally, prescriptive analytics uses advanced tools and technologies, like machine learning, business rules and algorithms to tell us what should be done (Bekker, 2019; University of Bath, 2021).

According to (Chou, 2020) the top BI tool features based on the analysis of the Gartner Magic Quadrant for analytics and Business Intelligence platforms are: management, security and architecture of the BI platform, metadata management, analytics dashboards, interactive visual exploration, support mobile display, embed analysis content, embedded advanced analytics, self-service data preparation, publish and share analysis content, ease of use and visualization.


Figure 2.1 - Gartner Magic Quadrant for analytics and Business Intelligence Platforms 2022, Source: (Mitchell, 2020)

Overall, users' tastes for BI are evolving as their data sources become more varied. They favor interactive dashboards, self-service data exploration, and self-service creation.

Business intelligence used to be a secret weapon only accessible to big businesses with deep pockets. Fortunately, affordable data analysis tools made available by technology like cloud computing helps

small firms join the BI revolution (Anthony, 2019). Like all industries BI is becoming more accessible and evolving, according to (Anthony, 2019; Gilbert, 2022) here some BI trends for the upcoming years: storytelling, data governance, emergence of Voice-Activated assistants, defensive artificial intelligence on the rise, more businesses to use predictive analytics, data warehouse modernization, self-service BI interfaces, data security, data quality management.

### 2.1.3. Business Intelligence Environment

There are numerous BI architectures now in use, according to literature review (Ong et al., 2011; Wu et al., 2007; Liu, 2010). Different BI targets require different BI environments. However, we can conclude what a generic architecture of a BI system looks like, and in Figure X we can see it is composed of four major layers.

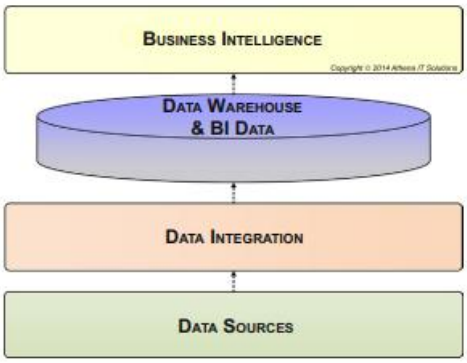


Figure 2.2 - BI generic architecture Source: (Sherman, 2015, p. 70)

**Data Sources**

A data source is anything that generates the digital information that business intelligence systems use. Data can be in the form of structured, unstructured or semi-structured data, and it comes from two types of sources. Data sources can be internal or external. Internal data sources refer to data that is collected and maintained within an organization (Customer relationship management system, financial software, Human Resources Management System, etc.), while external data sources refer to data that is generated outside of that organization (social media, public government data, other external data sources) (itransition, 2022).

**Data Integration**

Data integration is briefly described according to (Sherman, 2015) as “combining data from different sources and bringing it together to ultimately provide a unified view”. This layer is often called ETL, but actually ETL is one among other data integration methods. ETL stands for Extraction, Transformation and Loading. The first step in the ETL method is the extraction, it is the process of retrieval and gathering of data from the data sources, both internal and external. After the data is extracted there are a variety of potential transformations that can be done, for example: correcting misspellings, resolving domain conflicts, dealing with missing elements, or parsing into standard formats). Data transformation is the process of turning data into standard formats for reporting and analysis using a set of business rules. By doing this, the ETL process adds value to the data. Finally, the last step is to

load the transformed and cleaned data into a data warehouse form (Kimball and Ross, 2013; Watson, 2009; Sherman, 2015; Ong et al., 2011)

**Data Warehouse**

Data Warehouses are a type of management system intended to facilitate and assist business intelligence. They are used to store large amounts of historical data and the data is usually from multiple sources. A further and more detailed literature about Data Warehouses can be found in section 2.2.

**Business Intelligence**

The BI layer is where the receiving of business value begins. This layer enables data to be presented to business users in order to obtain and receive information, so they can use it to acquire valuable insights. This layer is the only visible layer of the system and is what the users see using tools and dashboards (Sherman, 2015; Watson, 2009). Here are a few examples of tools and applications that can be used to access the data Sherman, 2015; Watson, 2009):

- SQL queries
- OLAP reports
- Dashboards
- Data Visualization
- Excel
- Data Mining
- Mobile BI

## 2.2. DATA WAREHOUSE

### 2.2.1. Concept

Data warehouse (DW) is one of the most important components in BI architecture. They are used to store large amounts of data from multiple sources, both current and historical and this data can be in raw format as well as cleansed and filtered (MongoDB, n.d). According to (Inmon, 2005), a data warehouse should have the following properties:

- Subject-Oriented: a DW delivers information about a theme instead of organization's current operations (Inc, 2021; Tehreem Naeem, 2019)
- Integrated: a DW combines large amounts of data from multiple sources. In order to provide data quality, it is essential that the data is stored in a unified manner (Inmon, 2002; Tehreem Naeem, 2019)
- Nonvolatile: Data is permanent, old data is not deleted when new data is added, so historical data is preserved (Inc, 2021; Tehreem Naeem, 2019).
- Time-variant: Data in the data warehouse is always linked to a certain point in time (Inmon, 2002)

### 2.2.2. Models

When companies started to develop data warehouses as enterprise-wide data repositories two architectures for data warehousing arose (Watson, 2009), the Inmon approach, and the Kimball approach.

The approach proposed by Kimball focuses on a bottom-up approach. Data marts are first formed based on the business requirements according to the department's needs (Kimball and Ross, 2013). Using a data integration method such as ETL, the data is then loaded into a denormalized data model. The Kimball architecture uses the dimensional model such as star schemas or snowflakes to build a dimensional data warehouse (zentut, n.d). Facts and dimensions are the fundamental elements that define a star schema.

Fact tables are the core of the data warehouse. Facts are the measurable events that occur from a business process event. They contain quantitative information, and are almost always numeric. They also contain foreign keys for each of its associated dimensions, used to establish a link between the data in dimension and the fact table.

Dimensions, according to (Kimball and Ross, 2013), provide the “who, what, where, when, why, and how” context surrounding a business process event. They contain the descriptive attributes and the reference information about the facts (du Mortier, 2021, Kimball and Ross, 2013). Additionally, every dimension has a primary key field that is embedded as a foreign key in the associated fact tables.

According to (Kimball and Ross, 2013), this four key decisions should be taken into account when making a dimensional model: select the business process (this refers to operational activities performed by the organization), declare the grain (this means specify exactly what a fact table record contains, it means the level of detail associated with the measure), identify the dimensions (dimensions categorize and describe the facts recorded in the data warehouse), identify the facts (measurements that result from a business process event).

Bill Inmon advocated the enterprise data warehouse, and recommends following a top-down approach. This starts with identifying all the sources and main subject areas of the organization, and this is used to build a logical model for all primary entities. After that using the ETL process the physical model is constructed, following the normalized form. Inmon then proposes the creation of data marts for the analytic needs of each department or division. Thus, the DW acts as a single source of truth, and data is accessed through data marts for various reporting needs. (Inmon, 2002; zentut, n.d; Tehreem Naeem, 2019).

### 2.2.3. Data Warehouse vs Database

Table 1 - Data Warehouse vs Databases Source: (Tehreem Naeem, 2019; MongoDB, n.d)

| Data Warehouse | Database |
|---|---|
| Has an analytical workload<br>subject-oriented collection of data. | Has an operational and transactional workload<br>application-oriented collection of data |
| Serves as an information system that contains historical and commutative data from one or several sources. | Is a combination of related data |
| Uses Online Analytical Processing (OLAP) | Uses Online Transactional Processing (OLTP) |
| Tables and joins are denormalized | Tables and joins are normalized |

## 2.3. DATA VISUALIZATION

According to (Sadiku et al, 2016) data visualization is "the process of representing data in a graphical or pictorial way in a clear and effective manner". As a result of the increasing amounts of data available, it has become necessary to create effective visuals that display information in an easy, accessible and understandable format, this way making it easier for organizations to make sense of the data (Sadiku et al, 2016; Naidoo & Campbell, 2016).

### 2.3.1. Concepts and Benefits

"One of the greatest benefits of data visualization is the sheer quantity of information that can be rapidly interpreted if it is presented well" (Ware, 2004). The author explains several benefits of visualization, such as providing the ability to comprehend a huge amount of data; allows for the perception of unexpected properties that could go unnoticed; allows issues with the data itself to be quickly identified and frequently gives information about the data the methods used to gather it; facilitates understanding of both large-scale and small-scale features.

According to (Stoltzman, 2017), visualizations can be classified in three categories: good, bad and ugly.

Table 2 - Visualizations as classified by Stoltzman. Source: (Stoltzman, 2017)

| Good | Bad | Ugly |
|---|---|---|
| - Clearly illustrate a point<br>- Are tailored to the appropriate audience<br>- Are tailored to the presentation medium<br>- Are memorable to those who care about the material<br>- Make an impact which increases the understanding of the subject matter | - Are difficult to interpret<br>- Are unintentionally misleading<br>- Contain redundant and boring information | - Are almost impossible to interpret<br>- Are filled with completely worthless information<br>- Are intentionally created to mislead the audience<br>- Are inaccurate |

Depending on the kind of data you're working with and the final goal of the visualization, there is a wide variety of data visualization techniques, like line graphs, bar charts, scatter plots, pie charts, among many others (Miller, 2019, Sadiku et al, 2016). Although there is no strict rule on how to display the data, (Stoltzman, 2018) recommends focusing on what drives the visual representations (comparison, relationship, distribution, trend over time, and composition). Once the aim of the visual is known, it's much easier to reduce the number of charting possibilities. The author suggests using the following mappings:

- Trend: Column, Line
- Comparison: Area, Bar, Bullet, Column, Line, Scatter
- Relationship: Line, Scatter
- Distribution: Bar, Boxplot, Column
- Composition: Donut, Pie, Stacked Bar, Stacked Column

One way to display the data visualizations is through dashboards. They are used to consolidate different kinds of visualizations in an easy and understandable way (Tableau, 2022). There are three main types of dashboards: operational, analytical and strategic. Operational dashboards are used to see the ongoing operations, by monitoring and managing operations. Analytical dashboards help have a clear view of performance trends and potential issues, they help the company to progress at an executive level. Strategic dashboards are mainly focused on tracking goals and KPI's (Calzon, 2019; Costa, 2020)

### 2.3.2. Best Practices

According to several literature there are a lot of best practices when it comes to dashboard design. A few that are recurring according to several sources are (Calzon, 2019; Costa, 2020; Juice, 2009):

- Identify the audiences, and their interests and business purposes for the dashboards
- Define critical metrics and relevant KPIs
- Provide a dynamic and up-to-date view of the data
- Develop a user friendly and customizable dashboard interface

- Colors matter, use a color scheme
- Tell a story
- Present information in increasing levels of detail
- Use filters to allow flexibility and customization
- Use modularity to compact information

According to (Few, 2007), “Information cannot be placed just anywhere on the dashboard, nor can sections of the display be sized simply to fit the available space”. In order to have an effective dashboard the information placement and the space it occupies should be in relation to its importance, and a simple structure such as a grid is recommended.

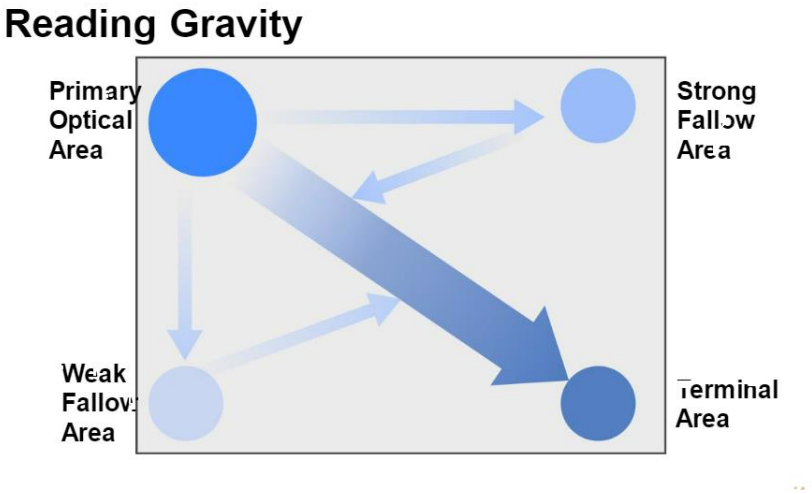


Figure 2.3 - Best Practices in Data Visualization, by Vihao Pham 2014

The use of interactive elements is also recommended as it allows to highlight key information and customization. Juice (2009) points to the importance of these features and divides them in two groups: basic (drill-downs, filters, comparisons, alerts, export/prints) and advanced (text-based-summary, starring/tagging, annotation, save/track changes, advanced visualizations).

# 3. DEVELOPMENT

## 3.1. PROJECT GOALS AND OBJECTIVES

From an internship point of view, the focus was on the process of recruiting and selecting candidates. HROps automates the management of recruitment processes, stakeholder applications, application review, and the selection process. The configuration of the existing phases in each hiring process is defined according to each organization, thus allowing a faithful image of what the existing processes are in organizations, with gains in efficiency, control, and information quality. The main objective is to create a dimensional model, so that all the data can be explored and valuable insights can be created and presented in PowerBI Platform.

## 3.2. METHODOLOGY

To make sure the goals of this project are accomplished I will have the opportunity to take advantage of my Business Intelligence knowledge and techniques acquired on my master's degree. With my existing knowledge I will be able and responsible to connect and extract data from an external data source, make the data flow through different conceptual Layers (ODS, Staging, DW) and create the dimensional model.

The Database of the HROps is called in our ecosystem as external data source and is the main source of data for the analytics ecosystem. To gather the data in the analytics database, which is the one I will be working on, I will be loading data from External Tables and store the data in the ODS Layer in the analytics ecosystem. After this integration, the next step will be to transform the data by applying the necessary business rules. This includes changing data types, excluding non-important columns, creating necessary columns, among others, this applies to dimensions as well as to factual tables. After this step is completed, the next step will be to load the data into the data warehouse.

For a better understanding of the ETL layered processing logic I will explain the existing layers and how I will be moving data between each of them.

It is necessary to create database objects (Tables) to materialize the data of each layer.

Existing Layers:

- **Virtualization**: External Tables for source systems
- **ODS**: Materialization of virtualization layer data into physical database tables
- **Staging**: Materialization of data in the Staging Area from business rules applied to ODS data (and potentially to other data sources)
- **DW**: Materialization of data in the Data Warehouse layer

Layers and Data Transformation Logic:

Data transformation logic presupposes the existence of a view to handle all defined business rules. In Figure 3.1 we can see what the process will look like:

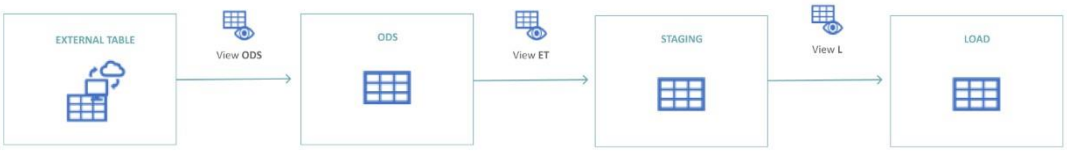


Figure 3.1 - Framework Overview

The business rules will be applied in each view layer, and this is also where the incremental loading logic and dealing with lookups will be applied.

To load and pass the data from one layer to the other we will be using Stored Procedures, which will help automate this process.

The procedures used encapsulate a set of generic actions that are typically performed in the same way. That is, the idea is to automate processes of incremental data loading, data updating and history refresh using only a set of input parameters, and the SPs will automatically adapt to the actions to be performed using dynamic SQL.

Although SP's were part of the project, I did not participate in the development of them.

The table below shows the Stored Procedures used and a brief description.

Table 2 - Stored Procedures used

| | |
|---|---|
| **ODS.LOAD_ODS_TABLE** | Loading existing data from a given source to a destination. The source and target schemas must have the same number of columns, with the same names and data types.<br><br>The target table will have 3 more audit fields: BI_INSERT_DATE, BI_ID_LOG, BI_ID_BATCH<br><br>Data Integration<br><br>- Full: Truncate-Insert<br><br>- Partial: Delete-Insert, considering a business-key |
| **ODS. LOAD_ODS_TABLE_HIST** | Loading existing data, typically in a view created by us where the Hash type1 and Hash Type 2 are already defined, as well as the business key.<br><br>The target table will have 5 more audit fields: COD_BEGIN_DATE, COD_END_DATE, BI_INSERT_DATE, BI_ID_LOG, BI_ID_BATCH |
| **ET. LOAD_STAGING_DIM_TABLE** | Starting from a source that is already a view with the necessary transformations, this procedure loads information from that view to the respective staging table |

| | |
|---|---|
| **ET. LOAD_STAGING_FACT_TABLE** | Starting from a source that is already a view with the necessary transformations, this procedure loads information from that view to the respective staging table |
| **L. LOAD_DW_DIM_TABLE** | Dimension loading based on a view with the necessary transformations parting from the staging table. This procedure responds to the needs of dimensions with and without history. |
| **L. LOAD_DW_FACT_TABLE** | Loading facts based on a view with the necessary transformations starting from the staging table. This procedure responds to the needs of loading factual tables in full format and partial loadings. |

In conclusion, HROps low cost ETL Framework will look like this:

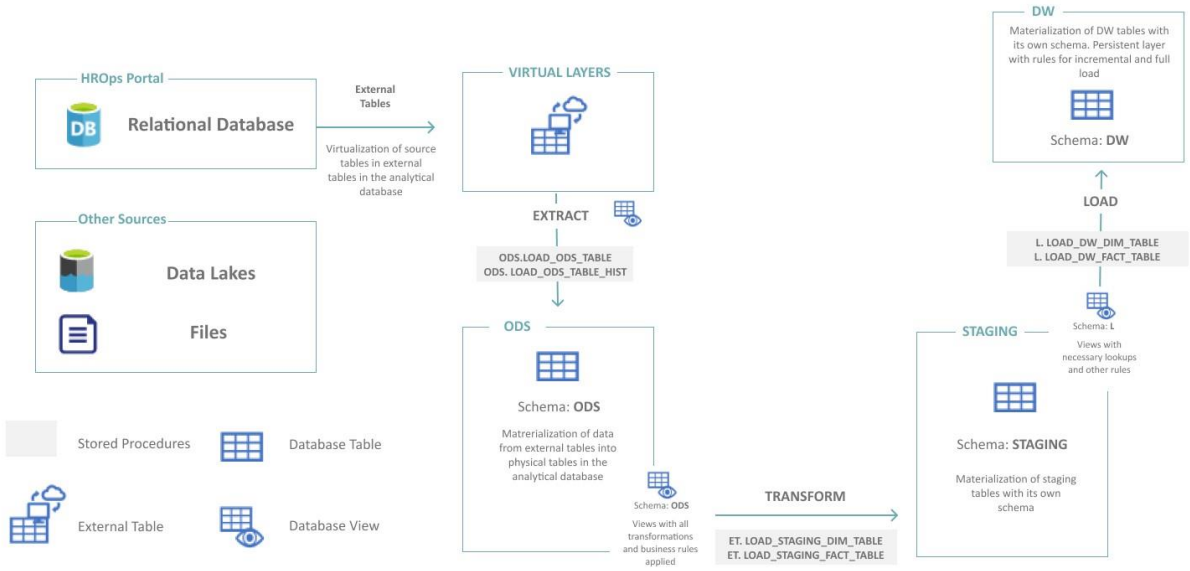


Figure 3.2 - Low Cost ETL Framework

In summary, our framework will have 4 main layers. The first step is to get the data from its original source into external tables (virtual layer), after that a standard ETL approach is followed with the ODS, STAGING and DW layers, where a view layer is used to apply the business rules, and stored procedures are used to load and pass the data between layers.

With all the necessary tables and relationships created, the ETL process is completed, and we now have a dimensional database. Finally, the last step will be to load this model to Power BI, so that we can create measures and compelling and useful visualizations to generate valuable insights.

**Maintenance and Control of ETL Processes**

In order to keep track of the processes executed on the server, whether they passed or failed, an Execution Log table is used. In table 4 it is shown what the execution log table looks like.

Table 3 - Example of Execution Log Table

| ID_LOG | ID_LOG_PARENT | ID_BATCH | PROCESS_NAME | START_TIME | END_TIME | STATUS | FAULIRE_MSG |
|---|---|---|---|---|---|---|---|
| 11613 | 11606 | 212 | STAGING_SA_DIM_CANDIDATE_ENTRY_TYPE | 2022-03-25 10:14:41.103 | 2022-03-25 10:14:41.170 | 1 | NULL |
| 11614 | 11606 | 212 | STAGING_SA_DIM_CANDIDATE | 2022-03-25 10:14:41.243 | 2022-03-25 10:14:42.420 | 1 | NULL |
| 11615 | 11606 | 212 | STAGING_SA_DIM_EMPLOYEE | 2022-03-25 10:14:42.430 | 2022-03-25 10:14:42.937 | 1 | NULL |
| 11616 | 11606 | 212 | STAGING_SA_DIM_HIRING_OFFER_CANDIDATE_STEP_STATUS | 2022-03-25 10:14:42.950 | 2022-03-25 10:14:42.983 | 1 | NULL |

This table allows to register the multiple process executions, monitor errors and times of execution.

Although I was not directly involved in the creation of the table, throughout the project I used it to keep track of processes and help me with debugging.

## 3.3. DEVELOPMENT ACTIVITIES

This chapter provides a process description of the different phases in the development of this project. The 1st phase starts with requirements gathering and construction of the dimensional model. The 2nd phase focuses on the data visualization part of the project, the elaboration of dashboards and creation of metrics.

### 3.3.1. Requirements gathering and construction of dimensional model

#### 3.3.1.1. Dimensional Model

The Scope of the Recruitment Module is to produce insights about the recruitment process. The stakeholders wanted to have a clear vision about the effort on the recruitment process and a blueprint about the candidates and their phases on the process. Everything from start to finish regarding the recruitment should be well organized and easy to access. In order to have a better understanding of the functional requirements of the system, below are a few business questions that the system should be able to answer:

- How many interviews were made last month?
- How many recruitment processes are active right now?
- How many recruitment processes are active right now by professional category?
- How many people were hired last month and for which job type?
- How many people rejected an offer?
- How has the number of hires changed over the year?
- How long does it take on average to hire someone based on seniority?
- What stage of the recruitment process takes the longest?
- What percentage of people, being selected after an HR interview, get to the financial proposal?
- In which process phase candidates failing the most?

The final dimensional model has sixteen dimensions and one fact table. A star schema approach was the preferred method for the design. Figure 7 shows the final model with the dimensions categorized by *Who*, *How*, *What* and *When* so their purpose can be better understood.

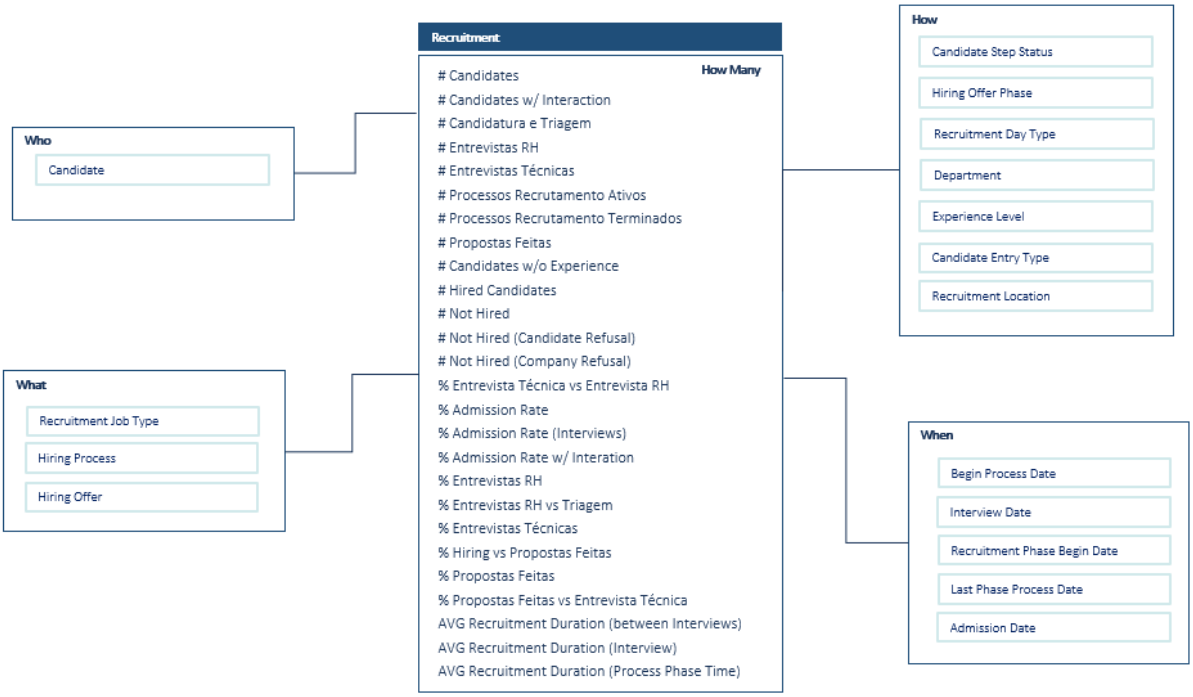


Figure 3.3 - Recruitment dimensional model

### 3.3.1.2. ETL Pipeline

The construction of the ETL framework was done with the help of the external tables that were already prepared in an SQL Server database (I was not involved in the creation and preparation of this tables). Since this was an internal project and had no specific timeline, both the requirements gathering and creation of the model were done simultaneously and intermittently. The team organized the tasks by having a daily meeting every morning, and giving feedback of what was done the previous day, and stating what they would be working on the present day. These meetings also served to lay out any problems encountered or coding errors that had to be talked to the team leader.

In order to complete the tasks appointed the team leader had already made the table mappings with the necessary fields for most the model dimensions. The tables were created in SQL Server using T-SQL and following the details provided by the team leader. Since the requirements gathering and construction of the tables were being done at the same time, sometimes after the creation of the table using the details given by the team leader adjustments still had to be made. Usually all the dimensions were created before moving on to the factual table. When the factual table was being created, sometimes the team had to go back and adjust the dimensions. For example, add an extra field, or remove a certain field, sometimes the table just wasn't necessary and ended up being deleted, and other times a new dimension had to be created. These modifications were talked about in the daily meetings already mentioned.

After the ODS layer was created (I was not involved in the development of this layer), is where the development of the model really started. The ODS layer, as mention previously, was only the materialization of the external tables, no transformations or data alterations were made.

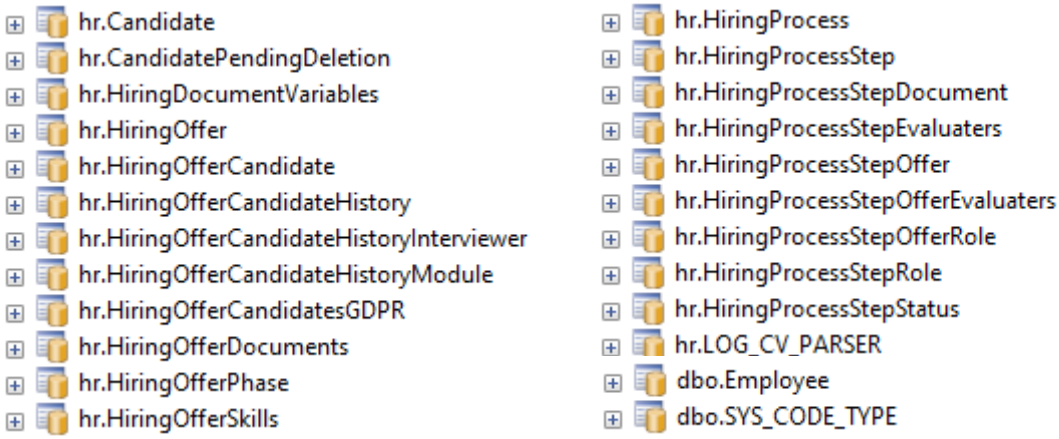


Figure 3.4 - External Tables

As shown in figure 6, besides the external tables, the framework is composed of views and tables that follow the standard ETL framework. The first view corresponds to the extract phase so its schema is E, this applies to the remaining views, the transformation views have the T schema, and the load views have the L schema. The extract and load views also have the "V_" prefix before the table name, and the transform views have the prefix "V_SA_", SA standing for Staging Area. For the actual tables, there are also three different schemas, ODS, STAGING and DW followed by the table names, with the exception of the staging table that have, again, the prefix "SA_".

In order to build the dimensions, the external tables had to be analyzed for the needed information. This was done with the help of the details already provided by the team leader. The mapping of the necessary fields for each view was done in the Transformation views. The views were generated using SQL code and would be used to pass the data to the Staging Area table.

After the view was created with the necessary fields, the staging table was created, this is the physical table that would save the data gathered by the view. This was done, as mentioned previously, using stored procedures. The SP also generated automatically three columns: BI_INSERT_DATE, BI_ID_LOG and BI_ID_BATCH.

Table 4 - Staging Area table for dimension JOB_TYPE

| COD_JOB_TYPE | DSC_JOB_TYPE | BI_INSERT_DATE | BI_ID_LOG | BI_ID_BATCH |
|---|---|---|---|---|
| 108001 | CEO | 20220331 | 12255 | 219 |
| 108002 | Executive Director | 20220331 | 12255 | 219 |
| 108003 | Manager | 20220331 | 12255 | 219 |
| 108004 | Partner | 20220331 | 12255 | 219 |
| 108005 | Senior Consultant | 20220331 | 12255 | 219 |

After the staging table is created, the process repeats but this time is for the Load views and Data Warehouse tables. For the load views, the only alteration is the insertion of a new column, a hash column, used to detect data changes.

After the final DW table is created, again, using an SP, the data from the view is loaded, and it automatically generates a new column, BI_UPDATE_DATE, in case data was altered, this column would store the date of the alteration.

Table 5 - Data Warehouse table for dimension JOB_TYPE

| SK_JOB_TYPE | COD_JOB_TYPE | DSC_JOB_TYPE | HASH_COLUMN_TYPE1 | BI_INSERT_DATE | BI_UPDATE_DATE | BI_ID_LOG | BI_ID_BATCH |
|---|---|---|---|---|---|---|---|
| 1 | 108001 | CEO | 0x57F0DF0E3D01F3F2B0AAF23F34E9555027766525CE779FBE5F0B044021D6FD15 | 20211124 | 20211124 | 1000 | 42 |
| 2 | 108002 | Executive Director | 0x844EC53558F277C4F9DA95CA311377773C25E1E1E1DD6F9C750A2E9A4FEBACF4 | 20211124 | NULL | 928 | 40 |
| 3 | 108003 | Manager | 0x8B2085F74DFA9C78A23B7D573C23D27D6D0B0E50C82A9B13138B193325BE3814 | 20211124 | NULL | 928 | 40 |
| 4 | 108004 | Partner | 0x29035E784F9305C0EB62AA33662B296306F6A77BEDACA7C67F2E3432DD43ACE2 | 20211124 | NULL | 928 | 40 |
| 5 | 108005 | Senior Consultant | 0xCFBB0A0BC8F3E3A7F3ADABE605F8608BFC10702922D3F92A3D6C2E512D11E98A | 20211124 | NULL | 928 | 40 |

Finally, there is also an admin table called PROCESS_ADMINISTRATION_MASTERS with the following fields:

- ID
- MASTER
- PROCESS_NAME
- SCRIPT
- FLAG_EXECUTION
- FLAG_ACTIVE

- MASTER_TYPE
- PARALLEL_EXECUTION
- EXECUTION_ORDER

This table is used to run the final stored procedure, SP_PROCESS_ORQUESTRATOR, that is used to run everything needed for the ETL process. The table has a MASTER column, used to identify which stage the process belongs to, ODS, STAGING or DW, the SCRIPT columns contains the code for the stored procedure needed to execute for that specific process, and the EXECUTION_ORDER is the order for which the scripts and tables need to be executed.

### 3.3.2. Power BI Dashboards Development

The second phase of this project consisted in using the dimensional model built to develop not only dashboards and visuals using Power BI, but also useful metrics. Two Power BI files were created, one was the analytics dataset, the other was the dashboards dataset. The information was passed from the SQL Server to the analytics dataset, and it was in this file that the model was imported and constructed. All the metrics and tests were done on this file. This file was then used to import the clean data into the dashboard dataset, and it was here that all the visuals were built. The option of having two files was chosen so we could have a development environment (the analytics dataset) and a clean file with just the visuals, because non-technical personal was meant to use the dashboards dataset, and because of their lack of technical knowledge it was thought it was a best practice to have the model and all the metrics in a separate file, so nothing could accidentally be ruined.
The metrics created were the following:

- # Active Recruitment Processes
- # Candidates
- # Candidates w/ Experience
- # Candidates w/ Interaction
- # Candidatura e Triagem
- # Candidates w/o Experience
- # Entrevistas RH
- # Entrevistas Técnicas
- # Processos Recrutamento Ativos
- # Terminated Recruitment Processes
- # Propostas Feitas
- # Candidates w/o Experience
- # Hired Candidates
- # Not Hired
- # Not Hired (Candidate Refusal)
- # Not Hired (Company Refusal)
- % Entrevista Técnica vs Entrevista RH
- % Admission Rate
- % Admission Rate (Interviews)
- % Admission Rate w/ Interation
- % Entrevistas RH
- % Entrevistas RH vs Triagem
- % Entrevistas Técnicas
- % Hiring vs Propostas Feitas
- % Propostas Feitas
- % Propostas Feitas vs Entrevista Técnica
- AVG Recruitment Duration (between Interviews)
- AVG Recruitment Duration (Interview)
- AVG Recruitment Duration (Process Phase Time)

The visualization was decided to be divided in three main areas: Interviews, Hiring and Actual state. There was a lot of information for each area, and it wouldn't make sense to mix visualizations between each one. This way allows for a better and clearer analysis and an easier access to the needed

information. It was also created a main dashboard with general information about the other dashboards, this was considered the home dashboard.

Each dashboard was created with multiple filters to allow for versatility, and also have general information that is the same for every dashboard (for example the number of candidates), but the main visuals are exclusive for each dashboard the area they are covering.

The *Interviews* dashboard has information regarding the interviews. It has information about the candidates and their experience, data about admissions and admission rates, as well as comparisons between different recruitment phases. The goal of this dashboard was to have the data regarding interviews summarized in one place. It allows stakeholders to see how many interviews are in progress or have been made, see how many interviews are done for each job position, a recruitment funnel of interview phases, and others.

The *Hiring* dashboard has information about the hiring process. Anything from active to terminated processes, to average process duration, to how many hired by job role, etc. The main focus was to understand how many people were being hired, to what type of job, how long it takes to hire, and what hiring phase takes the longest. The main focus was to have an overview of how time is spent in each recruitment phase and allow for its optimization as well as to see how hiring has evolved trough time periods.

The *Actual State* dashboard was meant to understand each candidate and hiring process actual state. Whether it was accepted, in review, dropped out, etc. It was also possible to understand which hiring processes were still active or not, and the average duration of them, among other insights. The goal of the dashboard was to have a clear picture of how things are right now.

With all four dashboards we believed a full analysis could be made to the recruitment model, and valuable insights could be extracted that accurately fulfilled the needs of the HR department. It is worth mentioning that the visuals created had the help of the feedback from the HR team, to whom we presented, like mentioned previously, the dashboards, and the team gave us feedback on what they think would be useful to them, not only for visuals but for metrics as well.

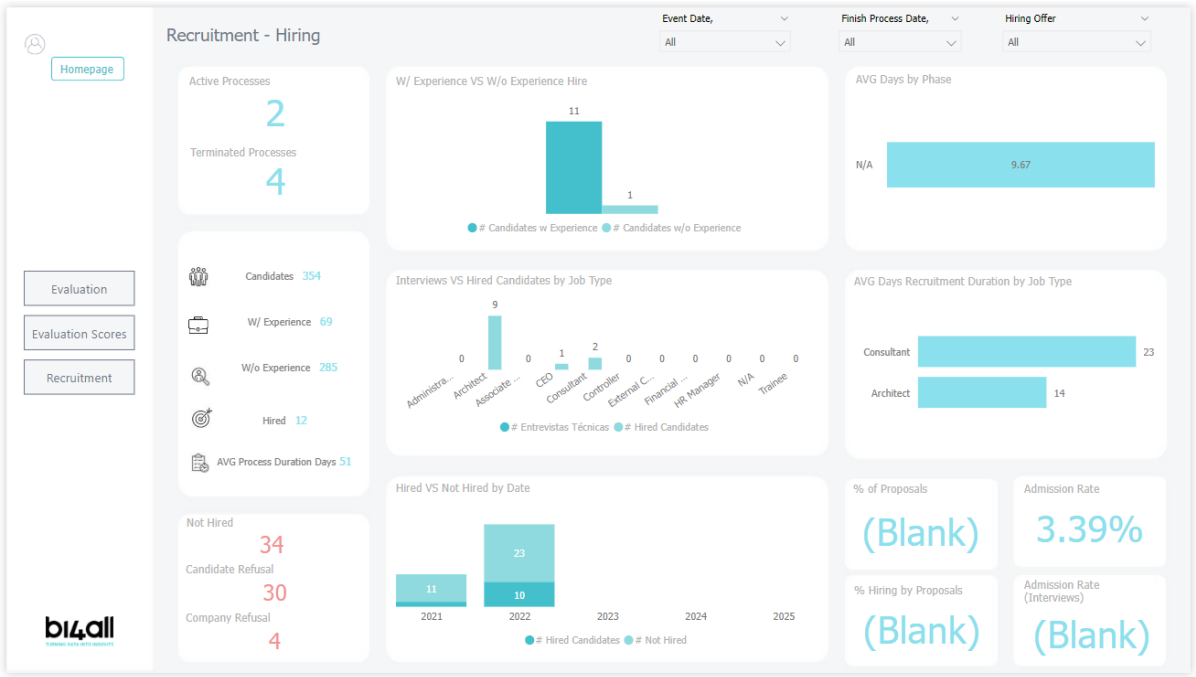


Figure 3.5 - Example of the final product, in this case the dashboard *Recruitment Hiring*

# 4. DISCUSSION

From the HR department point of view the dashboards are, at this point in time, not being used. In order to assess how impactful the project done in this internship was, a more qualitative approach is used. To do this, a few key subjects are discussed ahead.

Coming to this internship with no previous professional experience the impact on my knowledge and know-how was immense. It allowed me to develop as a professional and a person, developing both soft and hard skills. Although a lot of knowledge used in the project came from my previous education, the learning curve since the start of the internship was considerable.

In this modern era, organizations need to accelerate the digitalization of processes in order to meet customers' expectations (Markovitch, 2014). As customers become more digital and have increasingly more information, they seek a greater customer experience, thus obligating businesses to evolve technologically and culturally (Robledo, n.d.).

This project allowed to further implement process digitalization to the organization, by turning a process that was already being done to digital format, allowing for its optimization. By creating indicators and analytics associated to a digital transformation program of a core process of the company (recruitment), this project allowed to further implement process digitalization to the organization. By turning a process that was already being done to digital format, it allowed for its optimization

It is also worth mentioning that the project is in production and accessible to the HR department, but future work is implicit. The project followed an 80/20 approach, meaning that 80% of the functionalities were developed in 20% of the time. As the main functionalities of the system were developed, using our current knowledge of the data, the project requires continuous improvement, as the remaining functionalities require user feedback. This includes visualizations aligned with the specific needs of a team or department.

# 5. CONCLUSIONS

I have argued throughout this work how the exponential growth of data has made industries focus more on adopting digitally sophisticated data-driven processes. Having large amounts of data does not give a company a competitive advantage, not until this data can be used to give valuable insights, and that was exactly the main objective of this internship. The construction of this business intelligence application will allow the company to make better use of their data for the HR department.

The internship at BI4ALL has been a great experience, both professionally as well as personally. The internship's goals were directly aligned with the Master's program. Most of the tasks developed were related to ETL processes, data analytics and Business Intelligence. The majority of the knowledge that was learned during the Master's Program, both theoretical and practical, was applied. The internship has given me the ability to work on a real-world problem, and understand the complexity and architecture behind building an application, as well as a deeper understanding on the importance of data analytics and business intelligence in the modern world.

As the internship was divided in two different phases, it has given me the ability to work and improve on different fronts. The first one has allowed me to understand what goes through starting a project. All the requirements gathering, the trial and error building a dimensional model and the communication with third parties has allowed me to improve as a professional. Communication, organization and teamwork were the main soft skills developed, there were a good number of meetings throughout the project and a lot of tasks to be assigned, so these skills were essential to a good delivery. The technical skills developed during the first phase was mainly SQL, as the framework was built using it. The second phase has allowed me to expand my hard skills, by using PowerBI, and DAX language. In both phases communication and a will to learn were essential.

The hardest part of this project was the building of the dimensional model. Because the model dimensions and factual tables were built at the same time as the requirements gathering was made, there was a lot of trial and error, as well as going back and forth in order to see which solution was better, and even sometimes changes had to be made after talking to stakeholders. If I had to change one thing it would be to have a clear sketch of the requirements needed before starting the developments, as it would have saved time in the project, and made the developments easier. Another challenging aspect was the development in SQL, although I had experience with the language, the developments were more complex than what I had previously worked on, so a will to learn, to be self-taught, and to not be afraid of asking questions was crucial.

Although my part in this project was only to develop the application for the recruitment data, some future work may imply the same strategy for other areas, such as evaluation, training, certifications, among others. For this I strongly suggest to improve on the points mentioned previously.

Overall, it can be said that the project was successful, being delivered with all the requirements and meeting the deadlines. The internship has given a lot of practical experience, opportunities, and challenges in developing business intelligence applications.

## BIBLIOGRAPHICAL REFERENCES


Ain, N., Vaia, G., DeLone, W. H., & Waheed, M. (2019). Two decades of research on business intelligence system adoption, utilization and success – A systematic literature review. Decision Support Systems, 125, 113113. https://doi.org/10.1016/j.dss.2019.113113

Analytics And Business Intelligence (abi). (2019). Gartner. https://www.gartner.com/en/information-technology/glossary/business-intelligence-bi

Anthony, J. (2019, November 12). 11 Business Intelligence Trends for 2020/2021: Latest Predictions You Should Be Thinking About. Financesonline.com. https://financesonline.com/business-intelligence-trends/

Bekker, A. (2019, May 14). 4 types of data analytics to improve decision-making. Scnsoft.com. https://www.scnsoft.com/blog/4-types-of-data-analytics

Calzon, B. (2019, July 11). BI Blog | Data Visualization & Analytics Blog | datapine. BI Blog | Data Visualization & Analytics Blog | Datapine. https://www.datapine.com/blog/dashboard-design-principles-and-best-practices/

Chou, L. (2020, February 6). Top 10 Key Features of BI Tools in 2020. Medium. https://towardsdatascience.com/top-10-key-features-of-bi-tools-in-2020-fd02c1dd4c71

Costa, R. (2020, February 5). Dashboard Design: best practices and examples. Www.justinmind.com. https://www.justinmind.com/ui-design/dashboard-design-best-practices-ux

Datapine. (n.d.). Top 10 BI Tools - The Best BI Software Review List for 2020. Www.datapine.com. https://www.datapine.com/articles/best-bi-tools-software-review-list

Devens, R. M. (1868). Cyclopædia of Commercial and Business Anecdotes: Comprising Interesting Reminiscences and Facts, Remarkable Traits and Humors ... of Merchants, Traders, Bankers ... Etc. in All Ages and Countries ... In Google Books. D. Appleton and Company. https://books.google.pt/books?id=vqBDAAAAIAAJ&printsec=frontcover&hl=pt-PT&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

du Mortier, G. (2021, December 7). What Are Facts and Dimensions in a Data Warehouse? Vertabelo Data Modeler. https://vertabelo.com/blog/facts-dimensions-data-warehouse/

Efremov, O. (2019, November 17). Why Business Intelligence Is Important. Analytiks. https://analytiks.co/why-business-intelligence-is-important/

Erts, N. (2020, February 24). 3 Actionable Ways in Which the HR Data Management Process Can Be Changed for the Better - Sage HR Blog. Sage HR Blog | Easy to Implement HR Tips! https://blog.sage.hr/3-actionable-ways-in-which-the-hr-data-management-process-can-be-changed-for-the-better-human-resources/

Few, S. (2007). Pervasive Hurdles to Effective Dashboard Design. https://www.perceptualedge.com/articles/visual_business_intelligence/pervasive_hurdles_to_dd.pdf

Foley, É., & Guillemette, M. G. (2010). What is Business Intelligence? International Journal of Business Intelligence Research, 1(4), 1–28. https://doi.org/10.4018/jbir.2010100101

Gilbert, N. (2022, January 3). 8 Business Intelligence Software Trends for 2022/2023: Predictions You Should Be Thinking About - Financesonline.com. Financesonline.com. https://financesonline.com/business-intelligence-software-trends/

Heinze, J. (2014, September 26). History of Business Intelligence. BI Software Insight. http://www.bisoftwareinsight.com/history_of_business_intelligence/#topofpage

IBM Cloud Education. (2020, April 28). What is ETL (Extract, Transform, Load)? Www.ibm.com. https://www.ibm.com/cloud/learn/etl

Inc, F. (2021, June 17). 4 Key Characteristics of Data Warehouse. Faction Inc. https://www.factioninc.com/blog/4-characteristics-of-data-warehouse/

itransition. (2022, July 19). Business Intelligence Architecture: What It Is and How to Build One. Www.itransition.com. https://www.itransition.com/business-intelligence/architecture

Juice. (2009). A Guide to Creating Dashboards People Love to Use. https://static1.squarespace.com/static/52f42657e4b0b3416ff6b831/t/5310292ce4b08d35a87c9426/1393568044420/Guide_to_Dashboard_Design.pdf

Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013). Big Data: Issues and Challenges Moving Forward. 2013 46th Hawaii International Conference on System Sciences. https://doi.org/10.1109/hicss.2013.645

Kimball, R., & Ross, M. (2013). The data warehouse toolkit : the definitive guide to dimensional modeling. John Wiley & Sons, Inc.

Liu, L. (2010, August 1). Supply Chain Integration through Business Intelligence. IEEE Xplore. https://doi.org/10.1109/ICMSS.2010.5576813

Miller, C. (2017, June 21). What Is the Purpose of Business Intelligence in a Business? Financesonline.com. https://financesonline.com/purpose-business-intelligence-business

Miller, K. (2019, September 17). Data Visualization Techniques for All Professionals | HBS Online. Business Insights - Blog. https://online.hbs.edu/blog/post/data-visualization-techniques

Mitchell, C. (2022, March 30). Gartner Magic Quadrant for Analytics and Business Intelligence Platforms 2022. CX Today. https://www.cxtoday.com/data-analytics/gartner-magic-quadrant-for-analytics-and-business-intelligence-platforms-2022/

MongoDB. (n.d.). Databases vs. Data Warehouses vs. Data Lakes. MongoDB. https://www.mongodb.com/databases/data-lake-vs-data-warehouse-vs-database

Morris, A. (2021, April 16). 23 Examples of Business Intelligence. Oracle NetSuite. https://www.netsuite.com/portal/resource/articles/business-strategy/business-intelligence-examples.shtml

Naidoo, J., & Campbell, K. (2016, October 1). Extended abstract: Best practices for data visualization. IEEE Xplore. https://doi.org/10.1109/IPCC.2016.7740509

Nelson, G. (n.d.). Business Intelligence 2.0: Are we there yet? https://www.lexjansen.com/pharmasug/2010/MA/MA07.pdf

Ong, I., Siew, P., & Wong, S. (2011). A Five-Layered Business Intelligence Architecture. Communications of the IBIMA, 1–11. https://doi.org/10.5171/2011.695619

Rikhardsson, P., & Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. International Journal of Accounting Information Systems, 29(C), 37–58. https://ideas.repec.org/a/eee/ijoais/v29y2018icp37-58.html

Salesmate.io. (2019, March 29). 7 Ways business intelligence helps businesses grow - Salesmate. Salesmate. https://www.salesmate.io/blog/business-intelligence-helps-businesses-grow/

Sands, L. (2021, September 30). Why Is Human Resources (HR) Important? Www.breathehr.com. https://www.breathehr.com/en-gb/blog/topic/business-process/why-is-human-resources-important

Sherman, R. (2015). Business intelligence guidebook : from data integration to analytics. Elsevier.

Stoltzman, S. (2017, October 16). Data Visualization – Part 1. Open Data Science - Your News Source for AI, Machine Learning & More. https://opendatascience.com/data-visualization-part-1/

Stoltzman, S. (2018, February 11). Data Visualization – Part 3. Open Data Science - Your News Source for AI, Machine Learning & More. https://opendatascience.com/data-visualization-part-3/

Tableau. (n.d.). The Top 7 Benefits of Business Intelligence. Tableau Software. https://www.tableau.com/learn/articles/business-intelligence/enterprise-business-intelligence/benefits

Tableau. (2022). What is a dashboard? A complete overview. Tableau. https://www.tableau.com/learn/articles/dashboards/what-is

Tehreem Naeem. (2019, October 3). Data Warehouse Concepts: Kimball vs. Inmon Approach | Astera. Astera. https://www.astera.com/type/blog/data-warehouse-concepts/

University of Bath. (2021, April 22). Descriptive, predictive and prescriptive: three types of business analytics. The University of Bath. https://online.bath.ac.uk/content/descriptive-predictive-and-prescriptive-three-types-business-analytics

Watson, H. J. (2009). Tutorial: Business Intelligence – Past, Present, and Future. Communications of the Association for Information Systems, 25. https://doi.org/10.17705/1cais.02539

What is an operational data store? | Stitch resource. (2020). Stitch. https://www.stitchdata.com/resources/operational-data-store/

Wu, L., Barash, G., & Bartolini, C. (2007). A Service-oriented Architecture for Business Intelligence. IEEE International Conference on Service-Oriented Computing and Applications (SOCA '07). https://doi.org/10.1109/soca.2007.6

zentut. (n.d.). Kimball vs. Inmon in Data Warehouse Architecture. ZenTut. https://www.zentut.com/data-warehouse/kimball-and-inmon-data-warehouse-architectures/

Ware , Colin. (2004). Information Visualization: Perception for Design: Second Edition.

Sadiku, M. N. O. ., Shadare, A. E. ., Musa, S. M. ., & Akujuobi, C. M. . (2016). DATA VISUALIZATION. International Journal of Engineering Research and Advanced Technology (ijerat) (E-ISSN 2454-6135) DOI: 10.31695/IJERAT, 2(12), 11–16. Retrieved from https://ijerat.com/index.php/ijerat/article/view/191

Markovitch, S. (2014). Accelerating the digitization of business processes | McKinsey. Www.mckinsey.com. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/accelerating-the-digitization-of-business-processes

Robledo, P. (n.d.). Process Digitalization in Digital Transformation - Albatian. Albatian.com. https://albatian.com/en/blog-ingles/process-digitalization-in-digital-transformation/