NOVA IMS Information Management School

MGI

Master Degree Program in
Information Management

BUSINESS INTELLIGENCE IN A DATA DRIVEN BANK – POTENTIATING COMPLIANCE

José Guilherme Nóbua Coelho

Master Thesis

presented as partial requirement for obtaining the Master Degree 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

Business Intelligence in a data driven bank – Potentiating Compliance

by

José Guilherme Nóbua Coelho

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

Supervised by

Prof. Dr. André Barriguinha

July, 2024

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 acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

[Lisbon, 15-07-2024]

DEDICATION

To you, dad. You know what this means to us. I know you're watching Johnny Boy, you'd be proud.

To my mom, Helena, who is the biggest warrior on this earth. Thank you. Love commands life, we got this.

To Sara, who is the love of my life. Thank you for pushing and empowering me whenever.

To my grammas, I love you so much. Thank you for always asking how it was going and worrying.

To my grandparents, even though you are no longer here, I know you'd be proud.

To my friends. Thank you for always ensuring I was going full steam ahead, you all rock, legends.

ACKNOWLEDGEMENTS

To Prof. Dr. André Barriguinha, many thanks for always being available to hear my questions and giving so many great inputs which made it possible to finish this work.

ABSTRACT

Nowadays the whole business panorama is becoming more and more competitive, which means that they need to use all the possible competitive advantages that are available. Business Intelligence is a powerful tool that may be used to enhance companies' decision-making processes and help them make the best strategic decisions.

This master thesis addresses a research gap in the integration of business intelligence systems within the compliance departments of private banking companies. Compliance departments are often overwhelmed with vast amounts of data from various sources, making it challenging to monitor and manage compliance activities effectively. This specific Compliance department never had a framework that cleaned, stored and outputted data inherent to their tasks, operations and key metrics.

The compliance department is now able to make decisions based on accurate and timely information, which is crucial for an organization to be compliant regarding laws and regulations. The data mart was constructed using the necessary tables from the different raw sources, which was transformed and used to build all the dimensions and factual tables, using an ETL process (Extract; Transform; Load) to populate tables daily. A dashboard was built to provide utility,sustain the department's needs, and enhance the company's performance, more specifically of the compliance department with the impact BI (Business Intelligence) tools have on the compliance spectrum, leaving users more open to add value through other manners as they have key insights which can lead to optimized performance, AML risk being diminished and identifying possible problems within the compliance scope.

Ensuring data quality and consistency across all sources was a significant challenge. Inconsistencies in data formats, completeness, and accuracy could potentially compromise the reliability of the Data Mart (DM) and Data Warehouse (DW). Extensive testing activities were conducted in a quality (QA) environment to verify the correct creation of surrogate keys (SKs) in the fact tables and to ensure that no duplicates were present.

Secondly, the technical complexities associated with processes such as Extract, Transform, Load (ETL) in SQL Server Integration Services (SSIS), DM construction, and dashboard development posed significant challenges.

Future work could explore integrating machine learning applications with the DM. Leveraging machine learning tools can enhance decision-making processes, provide deeper insights, and offer various scenarios within the compliance scope. Implementing detection algorithms to identify unusual patterns and significant deviations would be an effective way to flag suspicious Anti-Money Laundering (AML) transfers for further investigation.

KEYWORDS

Business Analytics; Business Intelligence; Data Warehousing; Data Visualization; Compliance; Banking

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AML	Anti Money Laundering
BI	Business Intelligence
DFT	Data Flow Task
DM	Data Mart
DW	Data Warehouse
ETL	Extract Transform Load
IBM	International Business Machines Corporation
OLPT	Online Transaction Processing
PBI	Power BI
PEP	Politically Exposed Person
SQL	Structured Query Language
STG	Staging Area
SSIS	SQL Server Integration Services
SSMS	SQL Server Management Studio

1. INTRODUCTION

1.1. CONTEXT AND PROBLEM IDENTIFICATION

This chapter shall have the following structure, as per figure 1:

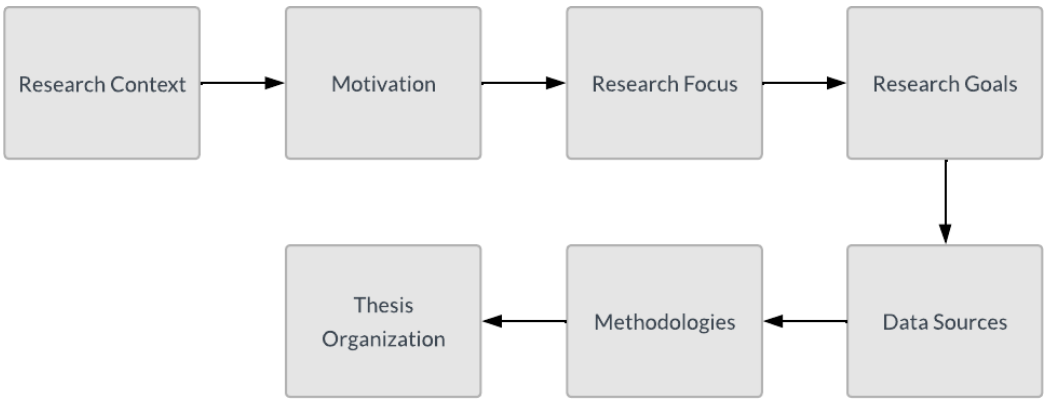

Figure 1 - Context and Problem Identification schema

The banking industry is highly competitive, with numerous institutions fighting for the loyalty and trust of customers. The need for financial institutions to stay ahead of the curve has never been more pressing, as financial institutions are constantly looking to gain a competitive advantage in the fast-paced, constantly changing environment of today's financial world. The issue of bank competition and compliance risk management is extremely important because it has significant impacts on both the banking sector and society, especially since it has not been explored adequately (Leo et al., 2019).

Similarly to other industries, having a robust framework supporting decision-making in the financial sector significantly impacts the efficiency with which financial services are produced, the quality of financial products offered, and the extent of innovation within the sector. Higher competition can lead to more efficient operations, better quality services and products, and increased innovation as financial institutions strive to differentiate themselves and meet customer needs effectively (Mirzaei & Moore, 2014). According to Tao and Tao (2008), business intelligence has become a critical tool in the banking industry for resolving numerous existing challenges, such as improving customer service, controlling financial risks, enhancing operational performance, and ensuring sustained profit growth. By leveraging ETL processing, data warehousing, online analytical processing, data mining, and data analysis tools, business intelligence enables large-scale enterprises to conduct comprehensive data processing, analysis, and display. This integrated application supports rational decision-making and helps achieve specific business objectives, providing quantitative decision support and enhancing overall performance. This triggers the need to develop a data mart, topped with a dashboard,

creating a data-driven environment that fuels informed decisions and efficiency. By centralizing information required by the compliance department and ensuring it is up-to-date, this development aims to enhance the overall efficiency and effectiveness of the department, improving the compliance process and reducing the risk of non-compliance.

The motivation for this research is driven by the need to address gaps in current Business Intelligence (BI) solutions used by banks in their compliance departments (Rahman, 2023). Although BI has gained widespread acceptance and is utilized by numerous leading organizations globally, there is a notable lack of research investigating the underlying factors that contribute to its successful implementation (El-Adaileh & Foster, 2019).Today, organizations harness data to drive informed decisions and optimize business operations, with top management relying on accurate information for effective decision-making (Kaula, 2015). However, existing BI solutions in the context of this development solely provide aggregated information that is not tailored to the specific needs of the compliance department, so, priorly to this research and development, the compliance department wasn't relying on accurate information for effective decision-making. This research seeks to fill this gap by developing a tailored BI infrastructure that offers granular, usable data for compliance teams, enabling them to identify trends, patterns, and insights that were previously unavailable. This will allow them to identify potential problems and take swift action to fix them before they escalate into a potential business crisis.

This research focuses on designing and implementing a BI infrastructure specifically tailored for a bank's compliance department. This involves analysing the needed sources, alongside with their ETL, designing and developing a data mart, alongside with its ETL, that centralizes all necessary compliance-related information and finally designing a dashboard that provides updated, accurate data for the compliance department.

The goals of this research include developing a data mart that centralizes all the information needed by the compliance department, ensuring that this information is up-to-date and reliable. The implementation of this solution is expected to deliver substantial value to the bank by improving the compliance process and reducing the risk of non-compliance, which can result in hefty fines and damage to the bank's reputation. According to Brar (2018), organizations optimize operations through data collection, understanding, and processing, enhancing informed decision-making for competitive advantage. By facilitating data-driven decision-making, this solution aims to help the bank achieve a competitive advantage through improved compliance processes.

The data sources for this research include various portals and systems used by the bank's compliance department. This encompasses data related to AML (Anti Money Laundering) alerts, account opening compliance analyses, and other regulatory requirements. Ensuring the consistency and accuracy of data from these diverse sources is a critical component of research. According to Panda (2012), banks often struggle to get the right data at the right time and place because data stored in repositories are of different dimensions from various

sources, leading to operational inefficiency. Developing a data mart with several key views inherent to the compliance department of a bank will enable data to be seen at the right time, in the right place, with assured standards and data quality.

The methodologies employed in this research involve extensive testing in a Quality Assurance (QA) environment to verify the correct creation of surrogate keys (SKs) in the fact tables and ensure no duplicates are present. Additionally, the study explores the integration of ETL (Extract, Transform, Load) processes in SQL Server Integration Services (SSIS), data mart construction, and dashboard development. Future work may include the application of machine learning tools to enhance decision-making and compliance monitoring, providing deeper insights and offering various scenarios within the compliance scope.

This thesis is organized into several chapters. The introduction includes the research context, motivation, research focus, goals, data sources, methodologies, and thesis organization. The literature review discusses existing BI solutions and their limitations. The methodology chapter details the design and implementation processes. The results chapter presents the outcomes of the data mart and dashboard development. The discussion chapter analyses the implications of the findings, and the conclusion chapter summarizes the research contributions and suggests future work.

1.2. OBJECTIVES

The primary objective of this research is to develop a solution that provides the compliance department of a bank with the necessary information to support its decision-making processes. The compliance department plays a crucial role in ensuring adherence to regulatory requirements and mitigating the risks associated with money laundering and other financial crimes (Rodrigues, 2019).

Specifically, the objectives of this research are:

- Design and Implement an ETL Process: Ensure the accuracy and completeness of data in the data mart and data warehouse by designing an efficient ETL process that extracts data from multiple sources.
- Develop a Data Mart: Create a centralized data mart that stores all relevant information regarding the compliance department's operations, ensuring fast querying and robust performance.
- Create a Comprehensive Dashboard: Design a dashboard that provides a comprehensive view of the compliance department's operations, enabling compliance members to quickly identify trends, patterns, and potential issues.
- Enhance Decision-Making & analytical analysis: Equip the compliance department with tools and insights that support informed decision-making, helping them meet regulatory obligations and mitigate financial crime risks.

By achieving these objectives, this research aims to deliver a clean and efficient business intelligence solution that empowers the compliance department to make informed decisions, monitor key client metrics and AML alerts, and manage compliance operations effectively in a centralized environment.

1.3. STUDY RELEVANCE

According to Pieket Weeserik & Spruit (2018), It's majorly important for the compliance department to have a concise and centralized repository of data which can guide them on making the best possible decisions, which will undoubtedly impact the business in a positive manner. Hence, it the data mart design includes all the needs of the compliance department and that the data is up-to-date, accurate and provides an holistic view of all the operations analysed by the compliance department, as well as information regarding AML alerts, that are triggered undercertain conditions being met in transactional/individual terms, adding also key metrics regarding the overall number of clients, politically exposed clients, the risk distribution of the banks' clients, pending compliance operations as well as status and key metrics inherent to the approval of operations by the compliance department.

According to Kimball and Ross (2002), data marts are a vital component of successful business intelligence solutions, offering a centralized repository for data that is easily accessible to decision-makers. Therefore, this research aims to design and implement a data mart aligned with the data warehouse (DW), generating views in SQL Server Management Studio (SSMS) that store relevant information regarding the compliance department's operations, including AML (Anti Money Laundering) alerts, client metrics, and historical data analysis. The data mart was designed for fast querying and solid performance under heavy loads to meet the needs of a busy compliance department.

Inmon (2005) highlights the necessity of a well-designed ETL (Extract, Transform, Load) process for an effective Business Intelligence solution, ensuring data accuracy and completeness in the data mart. This research aims to develop and implement an ETL process that extracts transactional data from various intranet portals and the bank's core financial system, providing the compliance department with a comprehensive view of their operations and the clients they monitor.

One of the most effective approaches to providing value and insights that impact the decision-making of a compliance department is through the development of an end-to-end tailored solution. This solution integrates data from multiple sources, contemplating all the portals and applications used by the compliance department. Integrating data from multiple sources is important to support comprehensive decision-making processes (Manolescu, 2020).By centralizing compliance-related data in a BI infrastructure, tailored to the specific needs of the compliance team, this solution can offer granularity and usability essential for more informed decision-making. (Pieket Weeserik & Spruit, 2018)

These systems analyse operational data and present insights through tables, graphs, charts, and other statistical formats. Such insights play a crucial role in determining the success or failure of an organization (Brar, 2018).To achieve this, it's crucial to follow best practices and models in its development, to understand what the best way is to generate value to the compliance department of a bank and impact their decision-making process. By addressing

the gap in providing comprehensive, centralized data and insights, this approach can significantly enhance the effectiveness of compliance management, reduce inefficiencies, mitigate risks, and safeguard the reputation of the bank (Rahman, 2023).

Nowadays the compliance department doesn't have a centralized repository of data regarding their operations, environment, working system and overall metrics, so the implementation of this project will propel the Compliance Department forward by providing them with the information they need in order to make well-informed decisions about their operations which will ultimately result in reduced risk and a positive impact on the business's overall growth. The project will also ensure that all the data will be easily accessible through an intuitive dashboard that will enable easy filtering of the data by the Compliance Team which will contribute positively for their day-to-day work as well as for their overall decision-making process.

Because the banking industry relies heavily on data, it's a natural fit for utilizing Business Intelligence (BI) capabilities. BI systems in banking excel at tackling complex real-world problems, making them great tools in this sector (Bany Mohammed et al., 2024).The dashboard will read the data mart that is going to be built, allowing a clean visualization and a form for the compliance department to analyse information in a friendly and intuitive manner. This development will provide value hence the compliance department will be able to monitor all its activities as well as key indicators, easily analyse the current situation and take immediate actions when needed, enabling the department to quickly identify areas of concern and take appropriate actions. Technology i.e., Business Intelligence, and the fact that technology propels advanced analytics techniques, contribute to providing valuable insights that will help to optimize its operations and better meet department and global goals.

2. LITERATURE REVIEW

According to Mositsa et al. (2023), having a comprehensive framework for data management in BI is essential for improved decision-making, so there is the urge to have theoretical insights regarding the whole business intelligence spectrum, to develop a great design for the data mart construction as well as on the planning to be done when it comes to the dashboard and how the information will be settled in the data mart, addressing the issues and conducting a solid research in order to define which will be the best methodologies and techniques applied for the development of the project.

Different approaches can impact the design of the data mart, what strategies to use regarding the ETL processes as well as graphical and visualization details that might impact the final output of the solution, as well as understanding what is the impact of BI tools and approaches on a compliance department of a bank, enhancing their decision-making and analysis efficiency (Bany Mohammad et al., 2022).

Akiotu (2022) states that failure to comply with regulatory sanctions can have severe consequences for banks, including financial losses and damage to reputation. Inadequate implementation of compliance strategies by bank managers can lead to reduced profitability, reputational harm, and even potential bank failure. This underscores the critical importance of effective compliance measures in safeguarding both financial viability and public trust in banking institutions (Akiotu, 2022). By using BI techniques there is the aim to improve the efficiency of the compliance department of a bank, therefore creating value by providing a better decision-making scenario and providing accurate and clear data that is suited for analytical purposes.

The best way to impact the compliance department of a bank regarding this development was to consider historical analysis (several indicators by timeframe), business performance analytics (operations and account opening processes analysed by the compliance department), employee performance measurement (pending operations and historical data of approvals),executive dashboards (final output solution), regulatory compliance and risk management (AML alerts insights to extract relevant information and clients by risk/Political exposed persons) (Brar,2018).

This chapter will approach key findings and methodologies developed by other authors that helped build concise research to facilitate the impacting process on decision making by applying BI tools in the compliance department of a bank. This shall be applied to build a successful data mart, improve the overall DW with new compliance related sources, develop the ETL and its logics for all the needed sources and dimensional model, design views in SSMS to feed a clean dashboard which will be set up and used in the compliance department of a bank to enhance the decision-making process by showing several key indicators.

2.1. COMPLIANCE, BANKING AND BUSINESS INTELLIGENCE

Business intelligence is deeply important for the compliance department of a determined bank as it provides considerable insights and several analyses inherent to customer behavior, transactions, determined compliance workflows as well as detecting patterns. Having this data at the disposal of a compliance department of a bank will facilitate the identification of potential risks, fraudulent activities, as well as key indicators incident on the whole compliance workflow activities, ensuring regulatory and financial compliance (Bany Mohammad et al., 2022). The goal is for BI to combine operational data with analytical tools to deliver competitive insights to planners and decisionmakers, improving the promptness and quality of the data that is used within the decision- making process (Panja & Paul, 2014). As previously mentioned, BI plays a huge role when it comes to providing advantages to organizations, enhancing operational flows and having key metrics at their disposal in real-time – this will largely impact a determined compliance department of a bank.

Using sophisticated software applications to collect and process data can help enhancing the

sanctions screening or AML processes to prevent misses that may end turning into sanctions violations (Böszörmenyi & Schweighofer, 2015). There are several BI tools and techniques that can be utilized by the compliance department to effectively monitor and analyse data, identify potential risks or compliance violations, and take proactive measures to prevent, more concretely fraud detection in terms of money laundering and terrorist financing. BI has the capability to provide organizations competitive advantages by transforming the decision-making processes, enabling decisions based on reliable data and insights, which may lead to improved efficiency, increased revenue and better customer satisfaction. Banks need to reassess many of their operational processes to comply efficiently with regulatory changes. This involves integrating the finance and risk segments of their businesses to strengthen performance goals while meeting new compliance requirements. By evaluating business performance on a risk-adjusted basis, banks can enhance their compliance strategies. Banks require advanced information analysis and reporting capabilities to manage risk, comply with regulations, and increase shareholder value (Panja & Paul, 2014).

With BI there is also the opportunity to monitor financial transactions in real time as well as detect patterns or suspicious activities that may indicate potential compliance violations, avoiding financial and reputational damage with associated penalties. Usually, BI solutions also tend to provide great data quality and accuracy, which is essential for a compliance department of a bank, in the sense that it provides reliable and trustworthy information for making informed decision, managing compliance workflow, have a macro analysis regarding clients of the bank and several others key indicators, providing a competitive advantage comparing to other banks improving operational and management efficiencies in the compliance department. Efficiency of internal operations will enhance profit margins of a bank, and BI plays a critical role in that allowing banks to monitor branch and employee information (Panja & Paul, 2014). In this development, the compliance department will have some key visions regarding to pending operations and other critical operational indicators that take place on the department's daily actions, thus providing a transparent and organized way to understand operational information enhancing the efficiency of the compliance department.

This will help a compliance department of a bank managing their daily operations and workflows, which will subsequently impact the overall success of the institution, hence BI systems are deeply important when it comes to improving data analysis and processing capabilities as well as leading to a more effective decision-making process by enhancing the identification and selection of relevant data (Caserio & Trucco, 2018).

There is the need to induce intuitive querying and reporting tools in order to gain competitive advantage as those facilitate access to crucial business information which is vital for strategic decision-making processes. BI is increasingly acknowledged as a strategic asset within various industries, particularly in the legal, financial and government sectors (Graves, 2005).

Nowadays there is a dynamic and very fiercely competitive environment in the banking sector, so it comes with naturality that this sector is adopting cutting-edge technologies to gain strategic insights, enhance operational efficiency, leverage the decision-making process. Integrating BI solutions within the scope of banking operations and processes enhances the extraction of valuable insights from vast volumes of data. To this extent banking and finance organizations are leveraging business intelligence technologies to increase profitability, mitigate risk and gain a competitive edge, once it enables banks to react to changing economic conditions in any economic scenario (Rouhani et al., 2016).

Using BI methods and technologies empowers banks to enhance their understanding of operations, clients, and prospects, paving the way for efficiency (Rahman, 2023). The research and methodologies conducted were based on the use of BI tools and methodologies to create value in the compliance department of a bank, enhancing their decision-making processes by boosting them with information that was previously unavailable, related to key client business indicators. Business Intelligence leverages technologies to convert data from diverse sources into actionable insights for business use, facilitating strategic decision-making, enabling the conduction of business operations with intelligent tools and techniques (Rahman, 2023). The methodology followed contemplated the integration of more than 50 sources inherent to information from the various portals the compliance department uses to approve operations, to manage client opening processes (screening) and AML alerts.

The work carried out by Rahman (2023), consisted on the examination of the effects of BI on bank operational efficiency using data from 27 branches of a commercial bank. He found that BI is positively significant to improve operational efficiency, and assure a competitive advantage.

BI has a deep important role in the banking sector by enabling experts and managers to make better, accurate, timely and relevant decisions being able to comply with the various regulatory dimensions of this sector (Bany Mohammad et al., 2022). There is a lot of complexity linked with the banking sector and their operations have a considerable transactional volume, BI implementations are a solution to manage their processes more effectively (Nithya & Kiruthika, 2021). Indeed, it was verified that there is a lot of complexity linked with the banking sector as the data was spread across different sources, not centralized in any platform, additionally the volume of data was considerable which lead the solution to contemplate special BI techniques to enhance the efficiency of the ETL processes, nevertheless all these techniques aimed to improve the decision-making and analytical ability of the compliance department.

According to a survey of 300 business-technology professionals, the adoption of BI tools is significantly driven by compliance and regulatory requirements in two out of five businesses surveyed. Organizations aiming to leverage reporting capabilities for compliance and risk mitigation are discovering that BI tools facilitate faster and easier achievement of these goals. Moreover, these organizations are realizing that BI tools also provide a competitive advantage

(Graves, 2005)

According to Mohammad et al. (2022), in today's dynamic environment, it's crucial for banks to prioritize the implementation of robust business intelligence systems to remain competitive in such a rapidly evolving technology, highlighting the value that this technology can bring specifically to the banking sector. BI has been defined as been defined as the ability of companies to think, plan, predict, solve problems appropriately, enable effective actions and help users creating and achieving business goals (AL-Okaily et al., 2021).

2.2. ETL

The ETL is a crucial aspect of data warehousing and BI, hence it involves the extraction of data from multiple sources, transforming it into a concise format, with consolidated data which is suitable to be analysed and loaded into a single data repository (Kimball & Ross, 2002). ETL, is the core and soul of BI, integrating and increasing the value of data according to the uniformed rules (Jun et al., 2009).

Typically, ETL processes involve cleansing data and implementing various transformations to ensure the data remains consistent, preserving its quality and integrity, as well as careful planning, design and testing to ensure data quality, consistency, and accuracy. ETL tools are a specialized category of software designed to address the challenges of data warehouse homogeneity, cleaning, and loading (Vassiliadis, P., Simitsis, A., & Skiadopoulos, S., 2002). Considering the findings of Rao & Kumar (2011), BI systems like ETL tools have transformed decision-making specifically in banking. ETL processes deal with the integration of data from diverse sources, a daunting task given the vast amount of data that exists in banking. This integration empowers users with valuable insights crucial for decision-making in a volatile industry like banking, where market dynamics constantly fluctuate.

Business intelligence tools such as the ones that perform the ETL process (in this case it was developed within SSIS), collaborate to produce an intelligent system that helps banking institutions in developing their current capabilities to implement and enhance their existing strategies, helping organizations achieving their goals (Alshehadeh & Al-khawaja, 2022).

As per Alshehadeh et al. (2023), which conducted an analysis regarding ETL processes in the banking sector, has concluded that the ETL process specifically has a vital role in the success of banking institutions, enabling faster and more accurate data and information exchange, leading to enhancing banking services more efficiently and cost-effectively. At the heart of efficient data management and subsequent BI processes, surges ETL (Extract, Transform and Load), aiming to facilitate the transfer of data from its source to a designated repository, representing a crucial phase in the implementation of an efficient and deep functioning DW. As present on figure 2, the ETL can be divided into five modules: data extraction, data validation, data cleaning, data conversion and data loading (Jun et al., 2009).

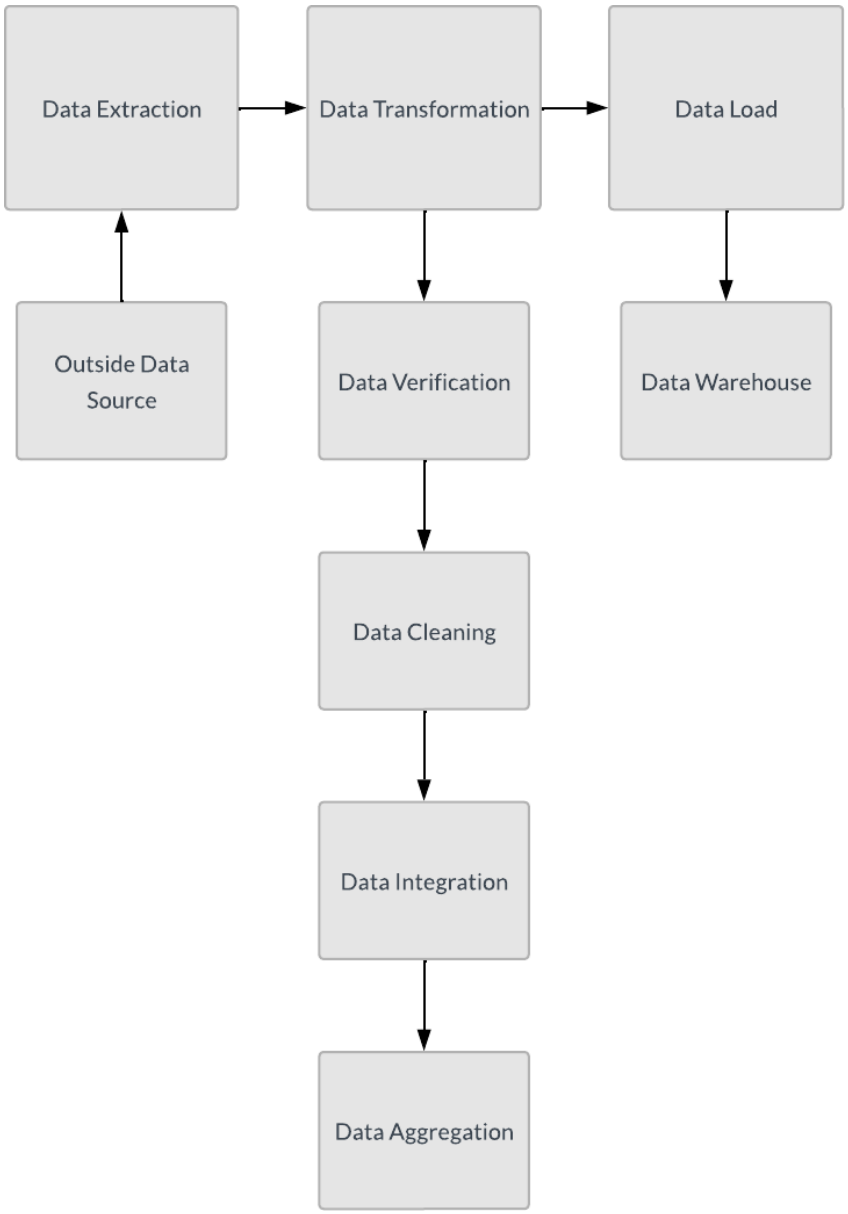

Figure 2 – ETL modules, adapted from Jun et al (2009).

Extract → This step is focused on extracting data from source systems, including several types of data, such as structured data from databases or unstructured data. The focus is to gather all the necessary information from the various data sources that are relevant to the project being developed. Generally, this stage involves substantial research and collaboration with various departments or external entities to identify the actual data sources. Extracting means reading and understanding the source data and copying the data needed for the data warehouse into the staging area for further manipulation (Kimball,1996).In order to identify all the necessary sources for the development of the project, there was the need to discuss and interact with the compliance department as well as with all the other IT teams responsible

for the population of the tables inherent to the visions that the compliance department uses on their daily basis, on which the indicators were be developed.

Transform → Very important part of the ETL process, in the sense that this step transformed the raw data, to data that is suitable for analytical purposes, to build an optimal solution. It is considered that there is a high probability of encountering source tables with column names that are non-standardized which couldn't be interpreted by users, so there was the need to transform the column names in order for it to be interpreted. In this step, data is cleaned, duplicates are dealt, and the normalization of tables happens. There is also the urge to configure data types correctly so that the information is properly stored in the repository. Kimball (1996) considered that this phase is the one that often demands the most attention and effort, as this phase demands considerable attention and effort because significant work occurs during the transform (T) step. This step involves combining data, addressing quality issues, identifying updated data, managing surrogate keys, constructing aggregates, and handling errors. The knowledge regarding advanced transformations will enable more complex logics to be applied on the ETL process, as a set of advanced, composite transformations is crucial, particularly in data warehousing tasks that involve combining simple transformations. Special attention must be given to tasks specific to data warehouses, such as handling slowly changing dimensions and resolving format mismatches (Vassiliadis et al., 2002)

Load → This is the final step of the ETL process, which mainly focus on loading the extracted and transformed data into the final data repository. It's on this stage that schemas are created, relationships between tables are defined and information is usually stored in dimensions and fact tables. Regardless of whether we're working with a series of flat files or a normalized data structure in the staging area, the final step of the ETL process is the loading of data. Loading in the data warehouse environment usually takes the form of presenting the quality-assured dimensional tables to the bulk loading facilities of each data mart (Kimball,1996). It is of extreme importance that quality and integrity of data is maintained during the loading part of ETL, as well as ensuring an optimized loading performance, which can consider developing partitions and indexes in order to enhance loading processes. To achieve the best loading performance, an analysis was developed with the goal of identifying which tables needed to be partitioned and indexed – as well as the development of an automatic routine job that executes daily, running all the packages inherent to the sources, by populating all the tables linked to the sources and afterwards executes other packages that insert data in the dimension and factual tables.

The ETL design and implementation followed established principles by developing it within a BI tool (SSIS + SSMS) , adhering to the extract, transform, and load (ETL) concepts commonly recognized in the field, as stated in the literature above. Despite this, alternative ETL approaches have emerged within the banking sector, such as the task-based model proposed by Li et al. (2009). This model employs a metadata-driven approach, combining coding and commercial tools to implement a task-centric framework at the core of the ETL process. Each task operates independently, offering flexibility in handling the ETL workload. However, while

this model proved effective for large commercial banks, it may not be suitable for the specific needs of this work, given the likely variations in data between large commercial banks and more specialized institutions like investment banks that work with niche clients. Therefore, the ETL approach developed continues to function smoothly within the niche banking environment this solution was deployed. Secondly, this approach would consider imputing wrong or NULL information on the ETL of the data mart developed on this particular situation, as it considers that the ETL process would be divided into several standalone parts and whenever a error occurs during some part is executing, instead of re-execute the whole ETL process ,there will be a re-execution of this part after the entire ETL process ends or recovering from problems. The ETL model that was developed follows a hierarchical approach, where data is sequentially loaded from sources to the data warehouse's source area (ODS). Subsequently, the STG phase processes data from the ODS to the staging (STG) tables, followed by loading dimensions and fact tables in a separate stage. However, the approach proposed by Li et al. (2009). presents a potential issue: if the ETL process encounters errors in a package from the sources, it may continue to run until completion, resulting in fact tables being loaded with incorrect or null information. This scenario occurs because the ETL process proceeds despite encountering errors in the source data, potentially impacting the accuracy of the loaded information in the fact tables or dimensions.

Implementing the full load strategy is quite simple: just truncate the data mart or table and reload the entire structure. This ensures a clean, fresh setup, eliminating tasks like key management and concerns about inconsistent data. Every time this strategy is employed, all data is uploaded from the source to the destination in SSMS. However, this simplicity comes with drawbacks when it comes to sources that have millions of records: updating a single data point in a repository with extensive longitudinal data is inefficient in terms of resource use. This inefficiency poses sustainability challenges and complicates risk management (Anywar et al., 2022). This challenge was faced during the implementation of this BI solution, so it involved the adaptation of the existing ETL processes related to some specific high memory consuming tables, by implementing a parameter-driven mechanism, aimed at selectively ingesting data based on predetermined criteria, such as specific date ranges or record identifiers. This approach has significantly enhanced the overall performance of the ETL process. Consequently, the need to load millions of extra rows on a daily basis was removed, leading to a marked improvement in operational efficiency, not needing to load millions of rows everyday but instead just load new records.

2.3. DATA WAREHOUSE

The data warehouse is a critical component in the whole enterprise architecture process because it serves as the central data repository and provides a wide variety of business users with a centralized source of high-quality business intelligence that supports decision-making processes across the organization. All this information can be analysed and leveraged to extract business advantages facing competitors. The data warehouse is considered a large, relational database that has been optimized for reporting and analysis purposes. It tends to be designed as scalable and flexible, being able to incorporate data from different sources (Kimball & Ross, 2002). The main goal considered regarding the development of this project was to build a DW/DM that accommodates data inherent to several sources i.e. working portals, websites and applications that the compliance department uses to analyse operations and clients. A data warehouse supports management decisions by providing a collection of subject-oriented, integrated, nonvolatile, and time-variant data (Inmon, 1996).

The data warehouse provides organizations with a centralized, consistent view of data, which enables the organization to make more pondered and accurate decisions, leading to hypothetical, optimal decision-making. These structures tend to adapt quite comfortably to changes, being scalable and perfectly adjustable when it comes to new business needs and data sources (Kimball & Ross, 2002). Data warehouses always need an explicit date dimension table. There are manydate attributes not supported by the SQL date function, including fiscal periods, seasons, holidays, and weekends. Rather than attempting to determine these non-standard calendar calculations in a query, we should look them up in a date dimension table (Kimball,1996). A date dimension shall be designed in order to be able to accommodate all the required dates and key indicators incident on specific festive dates. Designing a robust DW takes time and it'sa whole process, not something to be instantly made as it requires a step-by-step methodology, Kimball (1996) states that data warehouses are constructed incrementally rather than all at once. This step-by-step design and population process makes data warehouses evolutionary rather than revolutionary. Data warehouses are critical for modern organizations, as they provide a centralized repository that enables optimal decision-making.

Two approaches are mainly considered in the BI spectrum: Inmon (2005) and Kimball (1996). The Kimball approach emphasizes simplicity and ease of use, using a bottom-up approach, involving starting with a departmental data mart that focuses on end-users needs from the beginning of the process, while the Inmon approach emphasizes data integration and data quality, suggesting a top-down architecture that is designed to feed departmental databases, focusingon data-driven decision making. A data warehouse is designed to support decision-making by transforming data from various operational databases and other sources into new structures suited for business analysis (Vaisman & Zimányi, 2014).

According to Inmon (2005), a DW tends to acknowledge certain foundations:

Subject Oriented → Meaning it is organized around the main subjects or areas of interest for the business. By organizing data in this form, it allows the extraction of meaningful insights easily. In the specific banking scenario this can include customers, accounts, transactions, products, among others.

Integrated → It is deeply important to integrate all the data in a single repository of data, centralized. This ensures that all the data used in BI is consistent, accurate, and up-to-date providing a complete and unified view of the business. Within all the components of a DW, integration is the most important (Inmon,1996). Data undergoes multiple processes, highlighting the necessity to standardize and integrate all information. This involves transforming data to ensure seamless integration into the Data Warehouse (DW). As an illustration, consider standardizing measures, such as converting between centimeters and meters. It is essential to bring all measures onto a consistent scale, requiring the transformation of sources operating on different scales to the one most relevant for effective DW development.

Nonvolatile → In the sense that once the data is loaded in the data warehouse, it can't be changed. This is crucial to guarantee the accuracy and consistency of the information, providing a stable foundation for BI applications. It could make sense to delete some outdated data in the future, as nonvolatility in data warehouses means that they accumulate data from operational systems over extended periods, prohibiting data modification or removal. The only permissible operation is the purging of obsolete data no longer needed (Vaisman & Zimányi, 2014).

Time variant → It usually includes historical data as well as current data. This allows the business to analyse trends and patterns over time. Time variancy in a data warehouse signifies that each unit of data is accurate at a specific point in time (Inmon,1996). Historical data is considered to be valuable for business analysis and decision-making.

As stated by Kimball (1996), there are several requirements to ponder when building a data warehouse:

The data warehouse must make an organization's information easily accessible.

The data within the DW should be easily comprehensible to the business user, not only for the developer, as all the data should be intuitive and clearly labeled. If the information is easily accessible then probably there will be a data-driven culture that provides better decision making, as stated by Ross & Kimball (2002), understandability implies legibility; the contents of the data warehouse must be meaningfully labeled. Business users need the ability to separate and combine data in various ways, a process commonly known as slicing and dicing.

The data warehouse must present the organization's information consistently.

The data contained in the DW should be trustworthy and dependable, in the sense that it should be carefully collected from various sources within the organization. Consistency is key and it is in line with high-quality data, completeness and having all the necessary data to meet all the business needs in order to provide value. Users need to have confidence and feel that the solution is reliable, considering that data from different sources is being unified in a single repository, making it a reliable source of information for its users. Ross & Kimball (2002) considered that credibility is one of the key aspects of a DW as for a data warehouse to be credible, data must be meticulously collected from various sources within the organization, thoroughly cleansed, quality assured, and only released when it meets the necessary standards for user consumption.

The data warehouse must be adaptive and resilient to change.

Change is inherent to any business, therefore flexibility is a must for successful data management. The DW should be able to adapt to its business environment, data and technology. Meaning that existing data and applications won't be disrupted by amends on the existing DW. Adding new information to the DW as a business requirement may happen, so the DW must have a structure that doesn't get affected and is able to handle through all these changes and additions. Ross & Kimball (2002) have stated that the existing data and applications should remain unchanged and uninterrupted when the business community poses new questions or introduces new data into the warehouse.

The data warehouse must be a secure bastion that protects our information assets.

DW's store organizations' vital information, including sensitive and personal details regarding the whole parties involved. Having effective management and control of access to determined pieces of data is crucial to safeguard all the confidential data, making sure that data is properly secured and having proper access mechanisms and methodologies to access certain types of data. In contemporary data management, the introduction of data masking has emerged as a pivotal strategy for mitigating security risks associated with sensitive information in databases. This approach involves obscuring or disguising specific data elements to protect against unauthorized access. Moreover, the implementation of a group access mechanism offers an additional layer of security. For instance, by creating distinct user groups representing various departments within a bank,with members of departments all being under the same user group, organizations can precisely control access permissions.

The business community must accept the data warehouse if it is to be deemed successful.

The DW must be embraced by all business users as it serves business needs and it's critical in that sense. User acceptance is a major variable that indicates if the development is to be considered a failure. No matter how technically advanced and well-designed the DW, the true

acceptance and success is simply determined by whether the business users actively use it or not.

2.4. DATA MART

A data mart is a subset of a data warehouse that focuses on a specific set of business processes or functions or reporting requirements of a business organization. Typically, they are smaller than the DW and contain relevant data for a specific department, aiming to be highly focused and providing quick insights to a specific business function (Inmon, 2005). Regarding the development, a DM inherent to the compliance department of a bank shall be created, with information organizedwithin dimensions, taking into account all the inputs and key indicators required by the compliance department of a bank. All the sources were created within the DW layer which was later used to feed data to the DM. Aiming to provide efficient and meaningful analysis of specific business processes, data marts are built and shaped by specific requirements of a single department, as per Inmon (2005) data that is part of the data mart is fundamentally different from the data within the data warehouse, as it's denormalized,summarized and tailored to meet the specific requirements of individual departments, despite being related to data found in the operational level of the data warehouse. Data marts are a subset of a DW that supports the information needs of certain processes or departments (Utami et al., 2020).

Data marts provide an increase in speed and efficiency, in the sense that the critical information is highly focused so that users can access it quickly and in a manner that aims to provide better decision-making skills. A data mart should also be designed in a way that it can be easily transformed, being flexible and scalable, to promptly continue giving insightful business insights. On their most basic form, a data mart displays information derived from a single business process, mainly being related to different the different organization functions (Kimball & Ross, 2002).The development of the data mart has allowed the compliance department a better support for their management activities, this is one of the goals of the BI approach in banking (Brar, 2018). As per Jr et al. (2007) computer based systems that assist management activities and leverage functionality to summarize and analyse business information is called management support systems (MSS). This is exactly why there is the need to investigate if developing a BI solution containing a data mart may positively influence and generate more value to the decision-making process of the compliance department of a bank.

Operational business intelligence integrates analytical processes into operational business structures, enabling near real-time decision-making and collaboration. It enhances businesses to make informed decisions and act with efficiency regarding daily operations. Such efforts and developments will diminish fraud which is one of the main compliance targets (Brar,2018).In the end of the day the methodology that was designed is within the scope of operational BI in the banking scope as compliance will be able to have up to date information regarding operational data i.e. AML alerts which are automatically generated when certain conditions occur on a specific transaction and compliance approval and pending actions

inherent to clients transactions – that information is present on the dashboard allowing compliance department to closely monitor their operations and tasks. The data mart contains two fact tables, being both of them inherent to operational data (AML alerts and operations).

It is considered that data marts can mostly assume three types: dependent, independent and hybrid, being a combination of the first two. Dependent data marts play a vital role in an extensive data warehousing structure, extracting information directly from the centralized repository. They are considered subsets of a larger DW, integrated in the whole organizational data infrastructure, mainly referred to specific business units or departments with specific insights. Independent data marts are integrated as standalone entities, being developed taking into account the unique needs of a specific business function or department, thus design and maintained independently of a centralized DW. Hybrid approaches combine elements from both prior types of data marts. Regarding the development, an hybrid DM shall be built since some key indicators can be extracted from the DW but others, since they are very specific, shall be designed specifically inherently to the compliance department.

2.5. DASHBOARDS AND DATA VISUALIZATION

Dashboards and data visualization tools help enhance data and have a key role when it comes to providing business insights in an intuitive form, making data easily understandable and easy to interpret. It is considered difficult for users to visualize, explore and use the enormous amount of data being produced by organizations, hence having the ability to use dashboards and data visualization in business intelligence can greatly support decision-making processes, helping discovering patterns and comprehend information more easily (Sadiku et al., 2016).

Data visualization is an essential component of business intelligence systems as it has a wide range of application areas such as sales forecasting, performance analysis, and customer behavior analysis. Fraud detection and compliance thematic also greatly benefit from data visualization capabilities as they allow for quick identification of anomalies and patterns that may indicate fraudulent activity or non-compliance (Chen et al., 2008).

Data visualization plays a crucial role in the initial stages of fraud investigation by enabling investigators to detect patterns indicative of fraudulent activity (Sadiku et al., 2016).Data visualization may change the way individuals perceive information and evenimpact their lives, as data visualization transforms the way people interact with information and influences our daily lives (Aparicio & Costa, 2015).

It is to be noted that Cairo (2013), considers that visualization is a very powerful tool, as considering visualization mainly as a tool acknowledges that it is not solely an art form, but a form of functional art. This form of art achieves its aesthetic value not through the subjective, free expression of a painter or sculptor, but through the precise and deliberate craftsmanship of an engineer.

There are a few key aspects to consider when it comes to effective dashboard design, including the intended audience, the purpose of the development and key metrics included on the dashboard.

As per a study conducted by Allaymoun et al. (2023) inherent to five Islamic banks, generating great graphical results in banking may lead to simpler and greater financial performance analysis. Data can be presented in various manners, through summarized visions, aggregated visions, and exploring KPI's and BSC (Balance Score Cards). Design is crucial when it comes to Data Visualization, being deeply important, as it impacts users' understandability of the information, thus impacting decision-makers (Chen et al., 2008). It also can be outputted via different data visualization tools, in the case of this investigation the tool used was Power BI as it provides various personalized views of data and it ticks all the dots when it comes to the representations of the indicators that the compliance department of a bank wants to see, being also an end product that fits really well with the whole spectrum of applications used to develop the ETL, data mart and views that feed the dashboard

3. METHODOLOGY

To ensure the success of this project, it is imperative to follow a systematic approach that guarantees both the conceptualization and efficacy of the envisioned solution, outlined in figure 3:

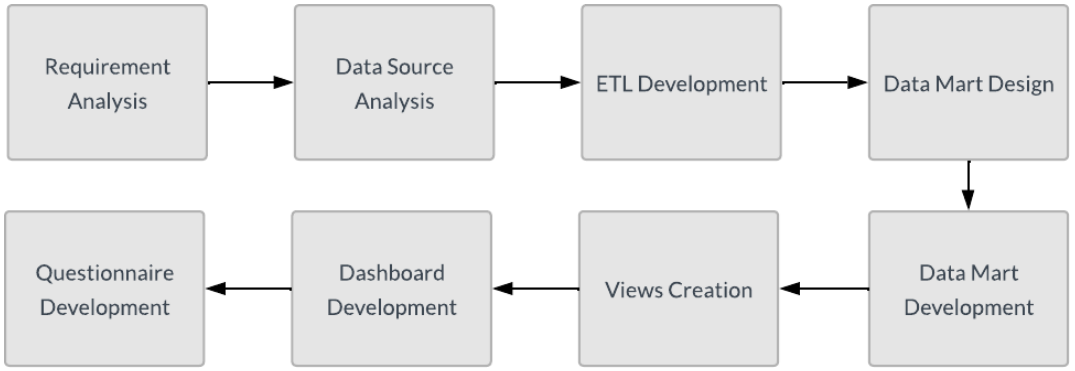

Figure 3 – Methodology

To ensure the success of this project, it is imperative to follow a systematic approach that guarantees both the conceptualization and efficacy of the envisioned solution. This development starts with an in-depth engagement with the compliance department, where there was a need to dive into their operational processes, check their core business needs, and pinpoint the precise data requirements essential for the envisioned final solution – this was made by several meetings during the development process, initially the meetings had the goal to define the business requirements and the indicators the compliance department needed and wanted to measure to enhance their decision-making process, as the methodology was applied the goal of the meetings shifted to a more deep understanding of the business indicators and doubts regarding the final visions compliance needed to enhance their decision-making processes.

The next step was to map out the data sources to fulfill the compliance department's information needs. This involves a comprehensive examination of existing data repositories, as well as a thorough exploration of any additional sources deemed necessary to augment the data landscape.

To provide a more complex understanding of the methodology phases, consider the following process descriptions:

Requirement analysis → This initial phase is crucial for understanding the specific needs and

objectives of the compliance department, engaging with key managers through several meetings to align all the steps of the process. Clear communication and documentation of requirements will lay the foundation for a successful solution.

Data Source Analysis → Identifying and assessing relevant data sources is essential for ensuring that the dashboard provides actionable insights. Evaluating both internal and external sources of data to determine their suitability for integration into the solution. This may include existing data within the Data Warehouse (DW) as well as data from external systems or sources that will need to be integrated within the DW.

ETL development → The Extract, Transform, Load (ETL) process is critical for integrating data from various sources into the DW. There will be a need to develop ETL processes for each identified data source, focusing on extracting relevant data, transforming it into a suitable format for analysis, and loading it into the DW. Utilize tools such as SQL Server Integration Services (SSIS) to create efficient and reliable ETL pipelines, ensuring data integrity and consistency throughout the process.

Data mart design → Designing a data mart tailored to the compliance department's needs is essential for providing meaningful insights. Define the dimensions and facts necessary to support analysis and reporting.

Data mart development → Implement the designed data mart within the DW, following the defined structure. Populate the data mart with relevant data using the ETL processes developed in the previous step. Ensure that the data mart is optimized for performance and scalability, allowing for efficient analysis of large datasets.

Views creation → Creating views is an essential step in this development. These views act as virtual representations of the underlying data, offering users simplified access to relevant information. Initially, the specific requirements and analytical perspectives necessary for decision-making will be identified. Based on these insights, views shall be designed for each vision that will be outputted in the final solution.

Dashboard development → Develop a visually appealing and user-friendly dashboard using Power BI. Integrate data from the data mart and data warehouse into the dashboard to provide real-time reporting and analysis capabilities. Design interactive elements and visualizations that allow users to explore data and gain insights easily. Automate the data loading process to ensure that the dashboard is regularly updated with the latest information, providing users with timely and accurate insights.

Questionnaire development → Create a user satisfaction questionnaire in order to obtain insights regarding the final output, with the goal of investigating if the implementation of a BI approach can enhance decision making within the compliance department of a bank.

Financial institutions are vulnerable to several risks that may threaten their financial safety and cause reputational damage, such as non-compliance with regulatory requirements and ethical standards.

To mitigate this risk, banks have established compliance departments that are responsible for ensuring that the organization complies with all applicable laws, regulations, and internal policies, helping to maintain reputation and trust with stakeholders, including customers, investors, and regulators.

The development will consist of an overview of the whole spectrum of operations that the compliance department of a bank analyses, as well as key metrics regarding clients. It's crucial for the compliance department to monitor and analyse certain operations, including the opening of new accounts. Additionally, there is a need to identify and prevent suspicious transactions that may be linked to money laundering, terrorist financing, or other illegal activities.

Figure 4 - Compliance key processes in requirement analysis

The **requirement analysis** phase considers several key concepts and metrics, according to figure 4, which can be described as follows:

OPE - Sent Operations. Consists of the whole spectrum of emitted operations by the bank's clients.

OPR - Received Operations. Consists of the whole spectrum of received operations by the bank's clients.

CB - Correspondent banking. Consist of all the SWIFT payment operations, to be considered MT103 (SWIFT payment message used specifically for cross-border and international wire transfers) and MT202 (SWIFT message for financial institution fund transfers between financial institutions). The compliance department has a working portal for all the correspondent banking operations.

TS - Transaction Screening. Consists of all the transactions that are inserted in a specific compliance working portal, where they are analyzed.

AML Alerts - Consists of alerts that are emitted and placed on a specific compliance working platform when certain parameters are reached in transactional terms (X EUR accumulated operations or operations above Y EUR).

PEP - Political Exposed Person.PEPs usually present a higher risk for money laundering activities or terrorist financing, so typically financial institutions must identify these clients and apply extra due diligence measures.

In terms of business insights that the department aims to achieve by having the dashboard developed, the following can be considered the requirements made by the compliance department for this project:

Last 7 days

- Number of OPE operations validated by the Compliance department
- Number of OPR operations validated by the Compliance department
- Number of CB operations validated by the Compliance department
- Number of TS operations validated by the Compliance department
- Number of AML alerts emitted split by each existing status
- New business relations (clients) validated by the Compliance department

Last 12 months

- Number of OPE operations validated by the Compliance department
- Number of OPR operations validated by the Compliance department
- Number of CB operations validated by the Compliance department
- Number of TS operations validated by the Compliance department
- Number of monthly aggregated AML alerts emitted split by each existing status
- New business relations (clients) validated by the Compliance department

Last 90 days

- Number of AML alerts generated split by each rule that triggers each alert

D0 (Today's numbers)

- Number of OPE operations pending validation from Compliance department
- Number of OPR pending operations validation from Compliance department
- Number of CB operations pending validation from Compliance department
- Number of TS operations with pending validation from Compliance department

- Total AML alerts having pending revision
- Total AML alerts in revision
- Total AML alerts in study
- New business relations (clients) with pending approval from Compliance department
- Total clients split by each risk level
- Total entities split by each risk level
- Total entities with PEP category
- Total entities with infected PEP category

4. PROJECT DEVELOPMENT

With the goal of providing a clearer and more concise vision about the project development, the following schema was developed, as follows on figure 5:

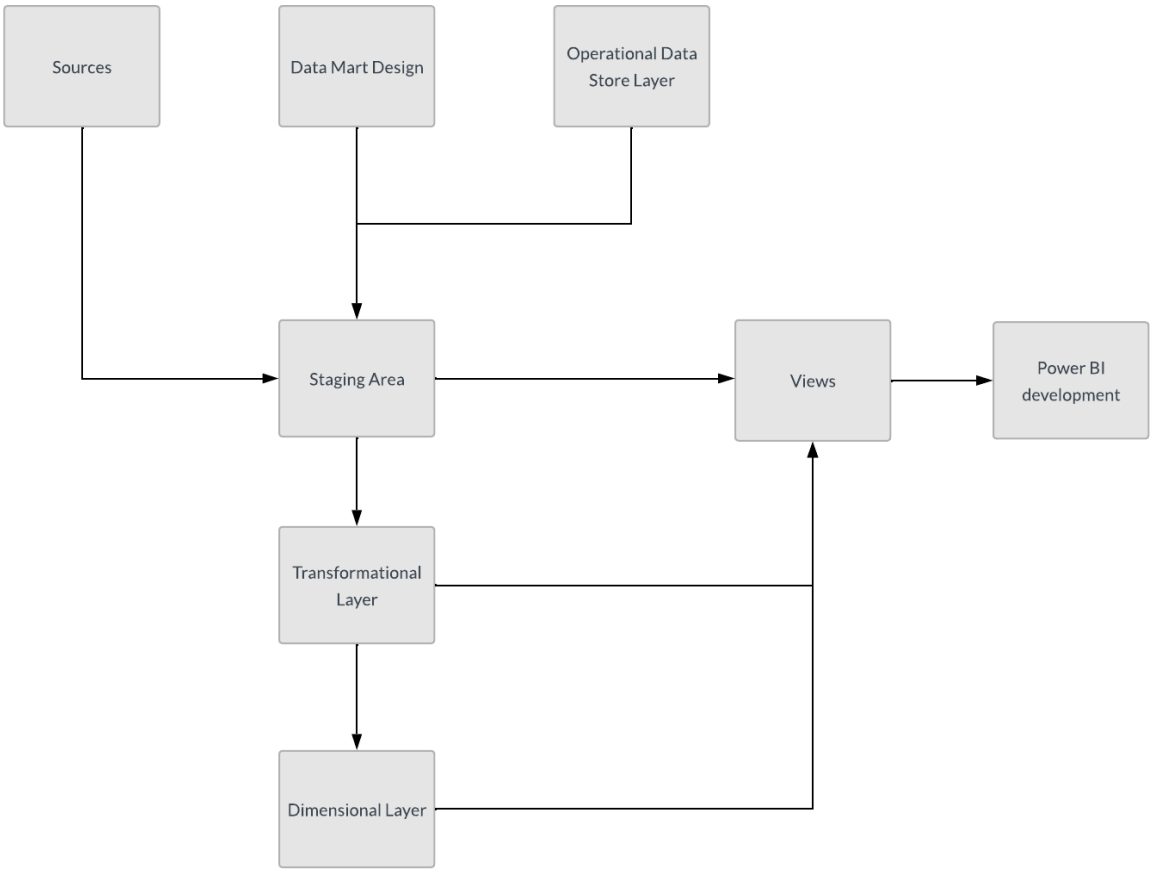

Figure 5 - Project development

4.1. SOURCES

The needed sources to carry with the development of the project, were determined through meetings with the compliance department and the internal development team, responsible for developing most of the compliance team workflows and data. The data required was available in different tables on a specific server. In this stage, all the sources were brought into the DW, more specifically into the ODS section, which will be covered later. These tables feed the DM dimensions and fact tables.

There is a quality environment, which is considered a testing environment that replicates the production environment, where all the users access the developments that are built. When the project is fully developed, users are granted access to a quality environment to test and analyse the development. If there is positive feedback inherent to the development made,

then there is the possibility to implement the development in the production environment, so that end users may use it daily.

Transactional systems, such as Online Transaction Processing (OLTP) databases, emerge as key sources of data. As elucidated by Oracle (n.d.), OLTP systems handle concurrent transactions in real time, encompassing activities like online banking, shopping, and order entry. Recognizing the significance of extracting data from such systems, the applied methodology prioritizes the extraction of relevant information from OLTP databases to populate the data warehouse and feed the data mart for the final dashboard. All the SSIS packages designed for the sources followed a standard approach that is implemented in the bank environment where the project was developed. Most packages inherent to the sources include a truncate statement on an expression and a DFT (data flow task) that contains 4 steps and is entirely built through variables on SSIS.

Source packages are built through a direct connection, meaning they are directly accessing the source server. The first step is a source task that is linked to the source table via direct connection, and thereis a selection of a table from the direction connection being used, afterwards, there is the addition of log columns, which are part of every table that is inserted in the repository. A total of 6 log columns are created and inserted into the DW table, plus all the columns with the same format as the source.

The next step was counting task to retrieve the total rows being inserted. The last task is inherent to the destination of the data that is being loaded, i.e., in this step, the DW table that will be loaded is selected, needing to map all the source inputs with the destination matching columns to guarantee that columns are correctly mapped and being correctly loaded.

Some tables, that were placed within the transformational layer represent sources and also exist on ODS, nevertheless, a TBL was created and inserted in the transformational layer.

TBL tables aim to optimize the loading process of the DW and was applied only in some specific tables, which consume a reasonable amount of CPU and memory, impacting the whole process. Instead of going to the source and retrieving all the present information, a different ETL process shall becreated linking the ODS table to the transformational layer table, having a delta number that is read every day, in the sense that only new information will be loaded to the ODS and inserted in the TBL instead of loading the whole source every day. Imagine populating a table that has 5 million rows every day; this process will only allow to load 5 million once, and every new day,only new data is loaded on top of those already existing 5 million rows.

The following schemas provide a more comprehensive view regarding the approach taken

regarding the different sources, according to figure 6:

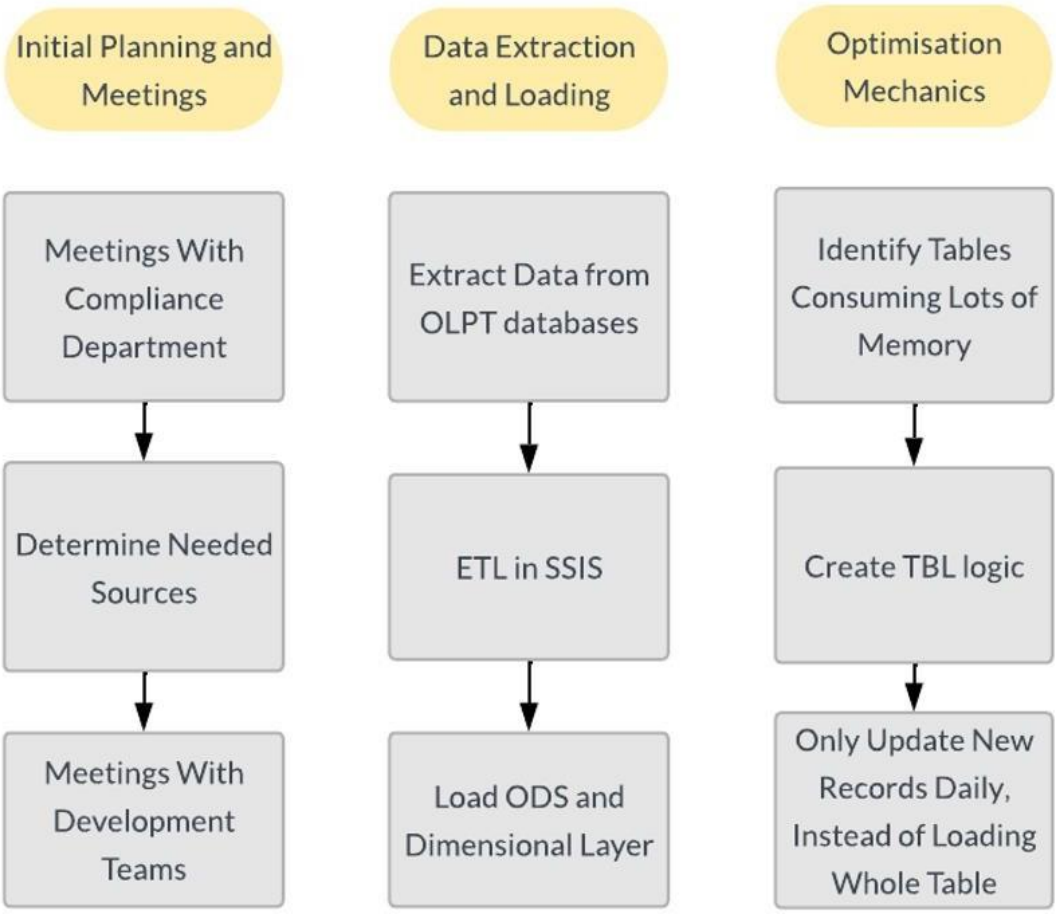

Figure 6 - Sources & mechanics

4.2. DATA MART DESIGN

Regarding DM design, several key decisions were made, especially since there are several types of actions, processes, and indicators requested that should be integrated into the dashboard. Thus, there is a need to organize data in a concise business format so that it provides solid business insights.

As stated in the literature review, the DM that was developed focuses on a specific set of business functions, more concretely on the compliance functions that are developed within a bank. All this information is present on the DW, which is feeding the DM with meaningful data that enhances the decision-making process of the compliance department ofa bank.

Following Inmon's (2005) approach, which states that the data warehouse as the foundational source of all departmental data, the data mart is conceptualized as an extension of this

infrastructure. The data mart is strategically positioned to leverage the comprehensive data stored within the data warehouse, ensuring consistency and coherence in the information ecosystem while also accommodating the unique demands of the compliance domain. This development has considered this approach, and it has indeed leveraged comprehensive data infrastructure and building compliance derived information.

The methodology adopted for this study is inherently aligned with Inmon's approach, emphasizing the creation of a data mart that is tightly aligned with the operational requirements and reporting needs of the compliance department. By consolidating relevant data within a concise business format, the data mart aims to deliver actionable insights that facilitate informed decision-making within the compliance function.

In line with the hybrid approach delineated in the literature, which combines elements from both the data warehouse and standalone data marts, the methodology for data mart design acknowledges the interplay between centralized data storage and department-specific analytics. The data mart architecture incorporates components sourced from the data warehouse, such as core operational data, while also integrating specialized elements tailored to the unique needs of compliance functions.

This methodology enables an increase in speed and efficiency, as all the compliance information is aggregated and summarized in a specific vision, providing faster and more accurate insights while still being flexible and scalable.

To achieve this goal, the DM was split between two main fact tables and their respective dimensions, mainly because the requirements mostly had indicators ranging between operations and AML alerts, which represent two different analytical workflows executed by the compliance department.

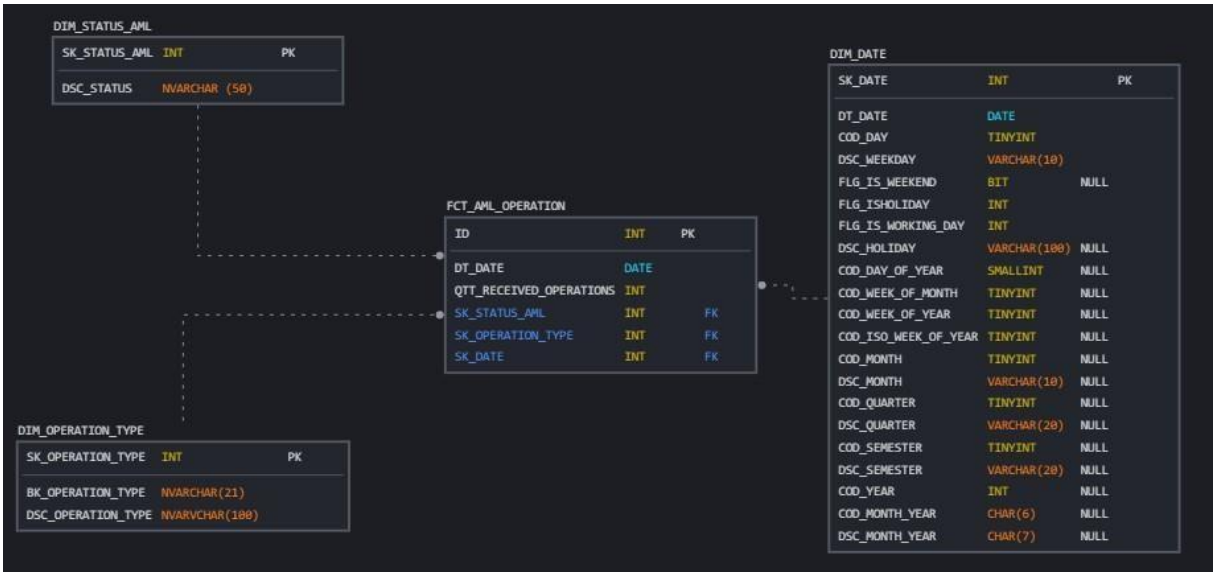

Figure 7 - Operations Fact Table

Firstly a fact table inherent to the operations analysed by the compliance department was developed, with the modelling that can be seen on figure 7, containing the quantity of operations that were analysed by the compliance department during a certain period. This fact table includes information inherent to all the operations analysed by the compliance department. This fact table contains one metric, quantity, as it is the only necessary metric to include in the fact table to produce all the final outputs that the compliance department wants to see on the dashboard.

This table makes it possible to measure the quantity of occurrences of a determined operation analysed by the compliance department for a specific status at a specific moment in time, contemplating information inherent to all the platforms a determined compliance department uses to analyse operations. This fact table has relationships with two dimensions, namely DIM_STATUS_AML, DIM_OPERATION_TYPE and DIM_DATE, because there is the need to identify the quantity of operations that have a certain status that are part of the whole range of operations analysed by the compliance department.

DIM_STATUS_AML has information inherent to AML operations possible status, combining multiple statuses from the whole spectrum of portals and applications a determined compliance department of a bank uses daily to approve several different types of operations from various portals and workflows.

DIM_OPERATION_TYPE contains information related to the broad range of operations scrutinized by a compliance department. Its purpose is to distinguish and categorize the various operations subject to analysis by the department through their various working portals.

DIM_DATE contains information regarding all the possible data dynamics in order to be able to present data within different timestamps and periods, considering weekdays, holidays, semesters, trimesters, and all the other timeframes that could make sense to include. Enabling temporal analysis is crucial to extracting valuable insights.

Figure 8 - Alerts Fact Table

The second fact table developed was the Alerts fact table which is related to all the AML alerts that are emitted on a certain platform (alerts are emitted because there are certain operational rules that trigger them) where the compliance department analyses the alerts. This fact table has information inherent to the whole spectrum of AML alerts that can be analysed by the compliance department.

These alerts are emitted on a certain platform, and they follow a specific workflow composed of certain activities that are managed by the compliance department. This fact table allows to measure and scrutinize the wholespectrum of alerts that are emitted. It can be seen on figure 8 that this fact table is linked to DIM_STATUS_AML_ALERTS and DIM_DATE as there is the need to measure the emitted alerts that were inserted on a certain status on a determined date.

DIM_DATE contains information regarding all the possible data dynamics in order to be able to present data within different timestamps and periods, considering weekdays, holidays, semesters, trimesters, and all the other timeframes that could make sense to include. Enabling temporal analysis is crucial to extracting valuable insights.

DIM_STATUS_AML_ALERTS contains information regarding the whole possible status inherent to AML alerts, including status from various working portals used by the compliance department of a bank to analyse and validate AML alerts.

4.3. OPERATIONAL DATA STORE LAYER

This layer is part of the DW solution and contemplates all the source tables. In order to correctly load the DM, the ETL for all the source tables must be developed within the DW. Instead of loading the dimensions directly through the servers from the data sources, this process develops the ETL on the ODS which then loads STG, DM and DW – which are correspondent to different layers of the Data Warehouse, as present on figure 9:

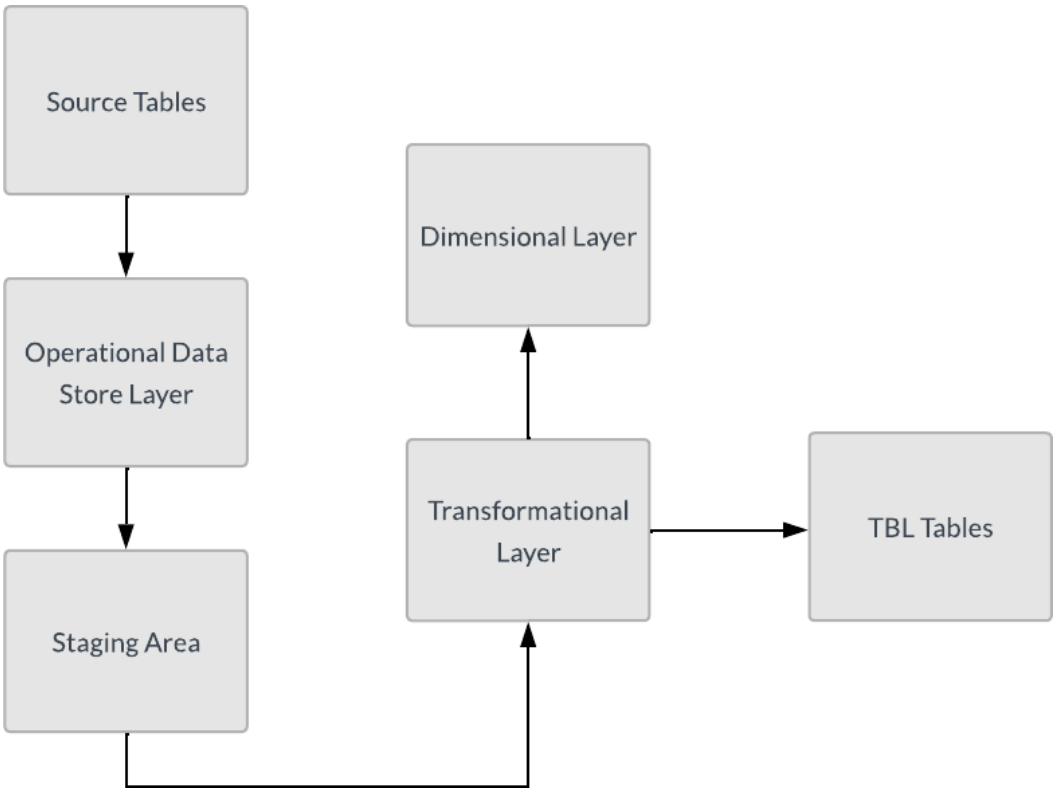

Figure 9 – ODS framework

In this section we will approach all the ETL processes that were built in SSIS, loading all the tables within the ODS environment, which in the end of the day refers to the source columns, all being present on the Data Warehouse solution The packages developed in SSIS and the tables being loaded for each ETL were also considered.

The packages created within ODS all follow the same approach and standardization is effective to promote the best results within the development of the project. Packages usually follow the below structure, being composed of a truncate expression, to truncate the destination table whenever the package runs and a DFT (Data flow task) which has the function of connecting to the table on the source, retrieve the data and load it on the ODS layer of the Data Warehouse.

Several tables within the DW layer serve as sources and are also present in ODS. However, a TBL table was specifically created and inserted into the transformational layer. The purpose of TBL tables is to optimize the loading process of the data warehouse and was only applied to selected tables that consume a significant amount of CPU and memory, thereby impacting the overall process.

Instead of retrieving all information directly from the source every day, a distinct Extract, Transform, Load (ETL) process was established to connect the ODS layer table to the DW layer table. This process involved reading a delta number daily, ensuring that only new information is loaded into the Operational Data Store (ODS) and inserted into the TBL within the DW. This approach avoids the need to load the entire dataset from the source each day.

Some specific packages contain a different logic because the source of the package has many rows, consuming lots of memory, to gain efficiency in the ETL process, this table follows a different and unique approach, in the sense that it loads the full information from the source just for its first execution, all the other executions only load new rows. This logic is made on SSIS by using variables with conditions on a source query variable and reading a control table that has the maximum ID or date present on another table. A table shall be created on the DW layer,which is a TBL, loading all the information from this table, which feeds the control table with the maximum ID, and the table in ODS only retrieves the new records.

Here the ETL of the sources takes place, where several transformations are made on the sources, transforming data into a concise format that can be loaded within each table, representing each source. The development made is in line with Vassiliadis, P., Simitsis, A., & Skiadopoulos, S., (2002), as all the transformations made guaranteed that data remained consistent and accurate, which is truly important to provide reliable data. This was only possible by using specific ETL tools, more concretely SSIS, which has all the tools to address the challenges that happen when building a DW or DM when it comes to data homogeneity, cleaning, and loading.

The development focuses on the three stages of the ETL process mentioned in the literature review: Initially, data is extracted from source systems, which requires plenty of research and collaboration with several other departments of the bank to identify the actual needed data sources. Then, transformations occur, changing raw data into data that is suitable for analytical purposes. During this development, several column names were modified so that users could interpret the meaning of each column. Lastly, all the tables are loaded with the transformed data and integrated within the DW.

In conclusion, the Operational Data Store (ODS) layer is an integral part of our Data Warehouse (DW) solution, encompassing all source tables. By developing the ETL processes within the DW using SSIS, we have ensured the systematic extraction, transformation, and loading of data into the ODS, STG, DM, and DW layers. The implementation of a delta load strategy and the creation of TBL tables have optimized the loading process, reducing CPU and memory consumption and ensuring that only new data is processed daily. This approach maintains the

integrity and performance of the DW. Transformations have been carefully applied to convert raw data into a concise, business-friendly format, enhancing the accuracy and usability of the data. As a result, the DW effectively supports the compliance department's functions, providing reliable and timely insights for informed decision-making.

4.4. STAGING AREA LAYER

This database represents the staging area layer, and as mentioned by Kimball (1996), in this stage all the source data was integrated into the staging area for further manipulation. This area plays a crucial role inherent to data integration and preparation as it serves as an intermediate layer where information from various sources is consolidated before being loaded into the DW. The staging area tables contain information from various sources, inherent to operational systems and applications used by the compliance department of a bank that have been consolidated, transformed, and integrated into this layer.

In order to build a meaningful DW, it was verified that Inmon's (1996) approach is a good model, as this was the model followed at this stage to enhance data quality and overall DW and DM performance. This development consisted of a subject-oriented and integrated foundation, as it is organized by the main areas of interest for the compliance department, i.e., AML alerts and operations analysed by the compliance department of a bank.

Several packages were created within SSIS to perform the ETL of several staging tables that, at a later stage, would be loaded into the data warehouse, integrating dimensions and fact tables. The development of this layer has proven to be very helpful, as it provides another stage where tracking and analysis can be made before loading the data into the data warehouse.

One of the primary challenges encountered pertained to the standardization of columns across disparate sources. The lack of uniformity in column naming and data formats posed a significant hurdle in the integration process. To address this challenge, all the needed columns from the various sources were standardized in the staging area, gaining business meaning, as per the literature review.

Developing SQL queries for ETL operations within SSIS to be executed, particularly those that had the role of updating, inserting, or replacing rows, proved to be exceedingly complex. These logics necessitated a deep understanding of the underlying data structures and businesslogic. By breaking down complex ETL logic into smaller, more manageable components, it was possible to enhance the maintainability and efficiency of the queries. Moreover, the implementation of error handling and logging mechanisms facilitated the identification and resolution of issues during query execution, ensuring the integrity and reliability of the staged data.

4.5. DIMENSIONAL LAYER

In this development, the data mart was located on the same layer as the broader data warehouse with the main facts and dimensions. This was made to unify the analytical environment and explore and analyse data across various levels of granularity. This promotes consistency in data interpretation and facilitates insight generation.

Co-locating the data mart with the primary data warehouse components may lead to performance improvements. By leveraging the same infrastructure and processing capabilities, queries using both systems can benefit from optimized resource utilization and query execution efficiency.

Also,co-locating the data mart with the main facts and dimensions within the rest of the data warehouse layer may offer benefits, like simplified data access, unified analytics, enhanced performance, reduced maintenance overhead, scalability, flexibility, and improved data governance. These advantages contribute to the effectiveness and efficiency of BI processes, as stated in the literature review.

In this layer, various packages on SSIS integrate data into the data warehouse, being the sources of those packages the tables inherent to the staging area, having each a dimension and fact table a corresponding staging table that serves as the source of the information being loaded on the final dimension and fact tables, within the data warehouse.

4.6. TRANSFORMATION LAYER

Throughout this project, certain tables emerged as pivotal components affecting the efficiency of the Extract, Transform, Load (ETL) process due to their involvement in handling substantial volumes of data on a daily basis. This involved the adaptation of the existing ETL processes to incorporate a parameter-driven mechanism, aimed at selectively ingesting data based on predetermined criteria, such as specific date ranges or record identifiers.

By embracing this approach, the ETL process was refined to dynamically adjust its data ingestion strategy, focusing on data surpassing predefined thresholds. This approach not only optimized resource utilization but also significantly enhanced the overall performance of the ETL process. Consequently, the need to load millions of extra rows on a daily basis was removed, leading to a marked improvement in operational efficiency and resource allocation.

To operationalize this strategy, a dedicated parameter table was developed in the ETL framework. This table is a repository for storing and managing the required parameter values governing the data ingestion process. Leveraging this parameter table, the ETL process from the ODS tables reads the maximum record inputted on this control table, which is present on the DW, after inserting data.

The integration of these logics into the SSIS (SQL Server Integration Services) framework helps to leverage robust data management tools to address complex challenges efficiently. Within SSIS, various components such as Data Flow Tasks (DFTs) and Execute SQL Tasks were used from the SSIS toolbox, each orchestrated through meticulously designed variables fed with queries tailored to execute the needed logic.

This approach lies in the strategic utilization of variables containing queries to implement the desired functionality. These queries are meticulously crafted to dynamically adapt the ETL process based on the predefined parameters, such as specific date ranges or record identifiers. Through SSIS's flexible architecture, these variables seamlessly integrate into the ETL pipeline, facilitating the execution of tailored data ingestion strategies.

The data sources fueling these ETL processes are drawn directly from the corresponding Operational Data Store (ODS) tables, which have been configured to implement the logic outlined in the earlier ODS chapter. Importantly, these processes are designed to retrieve only new records from the ODS tables, optimizing resource utilization and minimizing unnecessary data transfer.

The final step of these SSIS packages lies in updating the parameter table, a pivotal component that governs the subsequent day's data ingestion. By dynamically populating thisparameter table with the maximum date or ID from the TBL tables, the ETL process ensures that only records newer than the previously loaded value is imported from the ODS tables on subsequent runs.

4.7. VIEWS

This development has a framework that envisioned the creation of multiple views within SQL Server Management Studio (SSMS), tailored to meet the diverse analytical requirements of various compliance business indicator visions defined at the requirement analysis phase. These views serve as the sources for the later developed dashboards created with Power BI, which, according to (Gonçalves et al., 2023) can be used to create integrated performance dashboards that help in visualizing and analyzing business metrics effectively. The study highlights the importance of business intelligence tools in decision-making processes enabling data- driven insights and informed decision-making. According to Microsoft (n.d.), PBI is a unified and scalable platform for both self-service and enterprise BI, allowing for seamless connection and visualization of any type of data, providing detailed and intuitive analysis. Furthermore, Power BI facilitates the integration of visual elements directly into daily applications, enhancing the efficiency and accessibility of analyzed information

Instead of feeding the PBI directly with information originating from the DM or DW, a view for each vision of the dashboard was created, making the system more scalable, enabling debugging techniques and enabling security concerns.

Twelve distinct views were created within SSMS, each strategically aligned with specific business indicator visions identified by the compliance department of a bank. These views include curated datasets to address the analytical needs of diverse users, ranging from senior management to frontline compliance officers within PBI. Each view serves as a vehicle for analyzing specific visions of the bank's operations, ranging from transaction monitoring to customer due diligence or AML alerts. By mapping views to individual business indicator visions, the development framework facilitates granular analysis and focused reporting on key performance metrics and risk indicators.

The views designed were the following:

- VW_AML_ALERTS_RULES→ This view contains information regarding the diverse rules that trigger AML alerts on a compliance working portal. It counts the total number of each alert that are triggered everyday.
- VW_AML_ALERTS_STATES → This view contains information regarding the current status of the operations present on the AML compliance working portal.
- VW_AML_APPROVALS_MADE → This view contains information regarding the total approvalsmade by the compliance department of a bank in the AML working portal.
- VW_AML_APPROVED → This view contains information regarding the total transactions approved by the compliance department of a bank on all the operational working portals andwebsites used.
- VW_AML_CLIENT_RISK → This view contains information inherent to the risk

classification ofall the active clients of the bank.

- VW_AML_ENTITY_RISK → This view contains information inherent to the risk classification ofall the active entities of the bank.
- VW_AML_INFECTED_PEP_ENTITIES → This view contains information inherent to the entitiesthat are classified as political exposed persons because they have been infected by other entities which are originally considered PEP clients.
- VW_AML_NEW_ALERTS → This view contains information about all the new AML alerts that are emitted everyday.
- VW_AML_PENDING → This view contains information inherent to all the transactions that have pending compliance actions.
- VW_AML_PENDING_APPROVALS → This view contains information regarding the total clientopening processes that are pending compliance actions.
- VW_AML_PEP_ENTITIES → This view contains information corresponding to all the entities that are considered political exposed persons.

By implementing the strategy of mapping views to specific business indicator visions within the development, it enabled the enhancement of flexibility in the solution. This approach is consistent with the data warehouse (DW) concepts outlined by Kimball (1996). As articulated in the literature, a well-designed DW should be adaptive to its business environment, capable of accommodating changes and evolving requirements seamlessly. Moreover, it should serve as a secure bastion protecting the organization's information assets. This strategy enables these foundations specified on the literature review.

This solution incorporates tailored views aligned with distinct business indicator visions, embodies the adaptability and flexibility advocated in the literature. By modularizing the development and structuring the solution around individual views, the DW framework becomes inherently agile, capable of accommodating future adjustments or refinements to business indicator visions without necessitating a comprehensive and exhaustive redesign.

Furthermore, this approach aligns with the imperative of DW security highlighted in the literature. By restricting user access solely to the specific views corresponding to their analytical needs, the solution ensures the integrity and security of the DW environment. Access controls and permissions can be meticulously configured to safeguard sensitive data and prevent unauthorized access.

4.8. POWER BI DEVELOPMENT

Dashboards and data visualization tools play a key role in enhancing data comprehension and facilitating informed decision-making processes within organizations. As stated by Sadiku et al. (2016), the sheer volume of data generated by organizations poses a significant challenge for users in terms of visualization, exploration, and interpretation. Incorporating dashboards and data visualization capabilities into business intelligence frameworks can alleviate this challenge, enabling users to uncover patterns and recognize insights more intuitively. The development of a dashboard is also an important deliverable of this project, aimed at providing key insights for the compliance department of a designated bank.

Data visualization emerges as an indispensable component of business intelligence systems. As highlighted by Sadiku et al. (2016), data visualization serves as a proactive detection approach in fraud investigation, facilitating the identification of anomalies and patterns indicative of fraudulent activity. In this case, to exemplify, the dashboard is able to provide several insights related to the most alerts being emitted following to a certain rule which may trigger analysis inherent to that rule and all the operations generating alerts under that specific rule.

According to (Gonçalves et al., 2023), PBI is a business analytics and visualization tool designed for business decision-makers. It enables users to monitor business performance and interact with data to derive detailed business insights. PBI encompasses a suite of tools that facilitate data analysis, integration of data from various sources, and the creation of reports, graphs, visualizations, and dashboards. These features help in analyzing key performance indicators and metrics crucial for organizational growth and support the decision-making process.

By adopting PBI, organizations can foster a data-driven culture, make informed decisions, and gain a competitive edge in today's data-centric environment (Metre et al., 2024).

The development of the dashboard started with the creation of a user-friendly menu interface, aimed at providing a cutting-edge sensation for users while ensuring a minimalist and intuitive navigation experience. This initial step was pivotal in laying the groundwork for seamless interaction and efficient access to the dashboard's various visualization components.

A key aspect of the menu's development involved implementing interactive features to enhance user engagement and streamline navigation. Specifically, configurations were developed to enable selected visualizations to hover dynamically when users interacted with each menu item. This dynamic behaviour not only provided visual feedback but also embedded the dashboard with a sense of responsiveness and interactivity. Each menu item corresponding to each vision was developed with actionable functionality, allowing users to seamlessly transition to the desired vision with a single click.The menu can be observed in figure 10, as follows:

Figure 10 - Dashboard menu

The Operations vision was conceptualized to provide comprehensive insights into current compliance pending operations across various operation types, including Transaction Screening, Received Operations, Sent Operations, and Correspondent Banking. Additionally, the dashboard tracks the status of operations approved throughout the year, allowing users the flexibility to slice data as desired. To facilitate data exploration and analysis, a date filter and operation filter were incorporated, empowering users to segment their vision based on specific operation types and timeframes. Additionally, an option to go back to previous page was added on the top left corner of the dashboard.

A key decision in the development process was to enhance performance by incorporating forms as images into the dashboard, rather than directly loading and designing the forms within Power BI. This diminished the loading time in Power BI, resulting in improved performance and responsiveness. This vision reflects the core values of effective dashboard design outlined in the literature review. Specifically, the emphasis on user-centric design,

performance optimization, and interactive features resonates with the recommendations of Cairo (2013) regarding the functional artistry of visualization tools.

Throughout the development lifecycle, all the visions underwent iterative refinement guided by user feedback and usability testing. Several changes were made to enhance the clarity of insights presented, optimize performance, and improve overall user experience. This iterative approach ensured that the final dashboard met the requirements and expectations of dashboard users while delivering actionable insights effectively. The operations vision is presented in figure 11.

Figure 11 - Dashboard - Operations vision

As per Sadiku et al. (2016), dashboards serve as instrumental tools in facilitating data comprehension and supporting decision-making processes within organizations. In this context, the alerts vision of this dashboard was conceived to provide insights into currently pending compliance actions, particularly focusing on Anti-Money Laundering (AML) alerts from a particular software used by the compliance department of a bank. Firstly, this vision presents all currently pending compliance actions related to AML alerts, offering users visibility into critical compliance issues. Additionally, it includes a 90-day count of the rules triggering the launch of AML alerts during that period, enabling users to understand the frequency and distribution of alert-generating events over time.

Figure 12 - Dashboard - Alerts vision

The alerts visual representation, present on figure 12, allows compliance users to track the progression of alerts through different stages over time, providing key insights into patterns that may be due to some further investigation. In line with the operations vision, the alerts vision incorporates user-friendly features to enhance usability. Additionally, a slicer allows users to filter alerts based on their status, empowering users to focus their analysis on specific subsets of alerts according to their analytical objectives. This vision is designed to prioritize clarity, interactivity,and performance efficiency, aligning with the recommendations of Cairo (2013) regarding thefunctional artistry of visualization tools.

Figure 13 - Dashboard - Clients vision

The clients vision, displayed on figure 13, contains several key features aimed at enhancing analytical capabilities and enabling informed decision-making. Firstly, it presents a total count of politically exposed entities, providing users a clear view inherent to the existence of high-riskentities within the bank's client base. Additionally, it includes a count of entities classified as PEPs due to their association with original PEP entities, offering insights into the transmission and propagation of PEP-related risks.

Another aspect of this vision is the presentation of risk classification metrics by client and entity. This visualization allows users to identify and analyse the risk exposure associated with

individual clients and entities, facilitating risk mitigation strategies and compliance actions regarding those risks. Moreover, it also provides a clear view of risk distribution across the whole existing clients, enabling proactive risk management practices. It also includes insights inherent to pending compliance actions related to opening client accounts. By highlighting compliance actions associated with account opening processes, users can identify potential gaps or issues, thereby enhancing regulatory compliance and operational efficiency. A dynamic visualization representing the opening of accounts where compliance has participated in the account opening process over time further complements the clients vision. This visual representation, developed with slicers per date, enables users to track the involvement of compliance in account opening activities and identify temporal trends or patterns.

4.9. FURTHER REMARKS

In the context of developing a comprehensive data-driven strategy, there are several aspects that, while integral to the overall process, may not have the depth required for standalone chapters. To enhance the understanding of the development approach, some points were consolidated into this section. This ensures a cohesive presentation, enabling a clearer comprehension of the development processes carried through.

4.9.1. DATA DICTIONARY

All the tables created, in every layer except STG, contain metadata with their data dictionary, having on the extended properties a definition on the information present on the table as well as a description of each column that is part of the table.

The aim of this development is to enhance the understanding of the data present in the tables, facilitating users to identify which information is on each table and what is the meaning of each column. This is a factor that makes developers thrive and not lose time trying to understand the meaning of a determined column or table within the whole DW.

4.9.2. ORCHESTRATOR

The whole structure of this development is executed and managed through the orchestrator, which takes a key role when it comes to managing and integrating SSIS packages on SSMS and performing the ETL of all the needed packages daily. The Orchestrator is responsible for ensuring that all packages deployed to SSIS catalog are executed in the correct order with proper precedencies on SSMS, in the sense that a specific order has to be designed as source packages (ODS) for example can't be executed after the STG, because the source of the STG packages are the tables loaded on theODS layer.

All the packages developed in SSIS must be deployed into a catalog so that they can automatically run everyday on SSMS, through the packages to execute regarding the orchestrator.

There is a table,SSIS Package Configurations, which aims to parameterize all package execution data to ensure that packages are executed in the proper order and all data is correctly loaded into the databases tables.

4.9.3. SSIS DELTA CONFIGURATIONS

This is the table used within all the TBL logics → the delta value is updated for the maximum value of the selected column every time the TBL packages execute and load the TBL, as show in figure 14. In the next day, ODS packages read this table and have a condition to only integrate new records i.e where ID > Delta_Value.

ID	TABLE_SCHEMA	TABLE_NAME	COLUMN_NAME	DELTA_VALUE
11	MYATL_AML_EVO_PRT	CREDITNDEBIT	ID	20965031
12	MYATL_PORTALATLANTICO_PRT	WTHISTORY	ID	1222656
16	MYATL_PAYMENTSHUB_PRT	TRANSACTION	ID	7684938
25	MYATL_PORTALATLANTICO_PRT	HISTORYOPERATION	OPERATION_ID	61539

Figure 14- Delta Configurations example

4.9.4. TABLE MAX DATE

This table is used as a management table in order to fill the column “Value” with the maximum date loaded on each fact table, as shown on figure 15. The goal of this table is to enhance and optimize the DW performance by making the condition on any select statement through this table, to select only the current and actual records that were loaded more recently, instead of filtering through the FCT itself – this would consume much more memory and not optimize query performance as every query had to search within the FCTs whole data what is the maximum loaded date, by having this control table there is no need to do that, hence the select statement using this table is performed in under one second.

ID	TABLE_DATABASE	TABLE_SCHEMA	TABLE_NAME	COLUMN_NAME	VALUE	TABLE_UPDATE_DATE
3	BAE_DM	dbo	FCT_AML_OPERATION	DT_DATE	2024-02-16	2024-02-17 06:18:57.033
4	BAE_DM	dbo	FCT_AML_ALERTS	DT_DATE	2024-02-08	2024-02-13 06:30:36.597

Figure 15- Table max date

4.10. QUESTIONNAIRE

This section aims to gather insights from the end users of the entire BI system developed, which encompasses more than just the dashboard, nevertheless the compliance department, primarily analyzes the final output without having knowledge related to the technical details of ETL, DM, or other BI processes. The objective is to understand how the entire BI solution, including the data integration, transformation, and visualization processes, can positively impact decision-making and analytical approaches. The questions were crafted to assess the perception of the effectiveness, ease of use, and overall impact of the complete BI system on the compliance department, once all this has impact on the final output.

The compliance team is composed of 14 people, composed of junior analysts, senior analysts, and compliance managers. In figure 16, the percentage distribution pie chart of the roles within the compliance department:

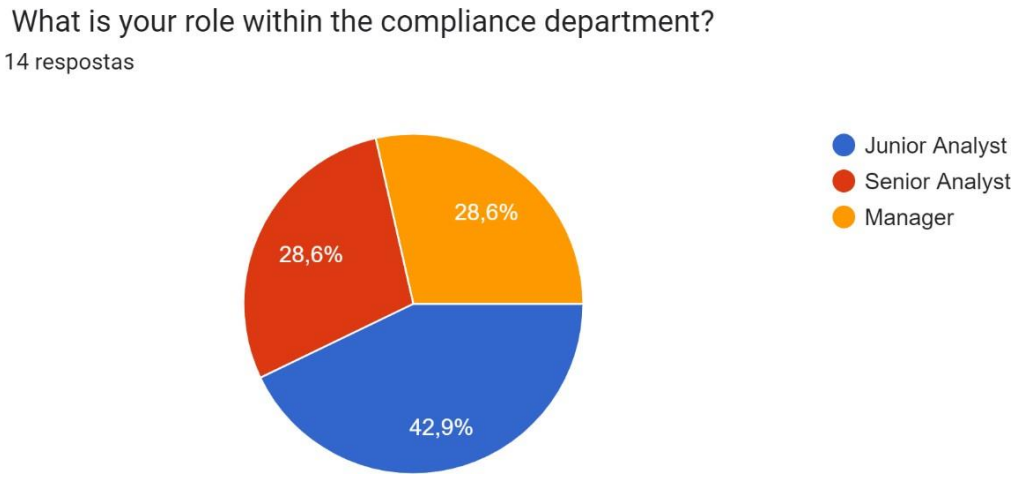

Figure 16 - Role in the Compliance Department

Of all the inquired, as present on figure 17, most are familiar with BI tools, being 1 not familiar and 5 somewhat familiar.

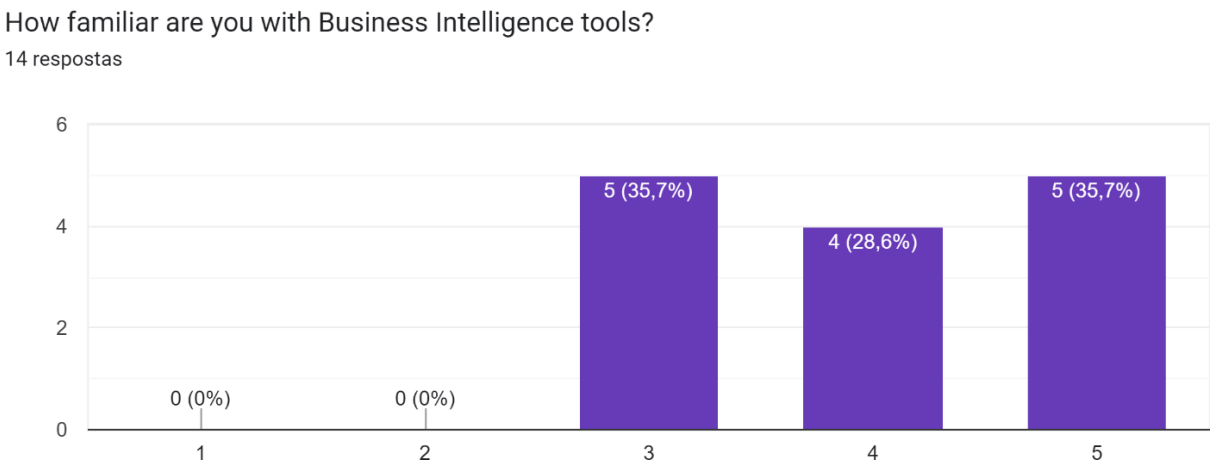

Figure 17 - Familiarity with BI tools

As seen on figure 18, most of the users tend to use the dashboard to monitor compliance metrics, while 28,6% tend to use it mainly to identify potential issues and to support decision-making processes.

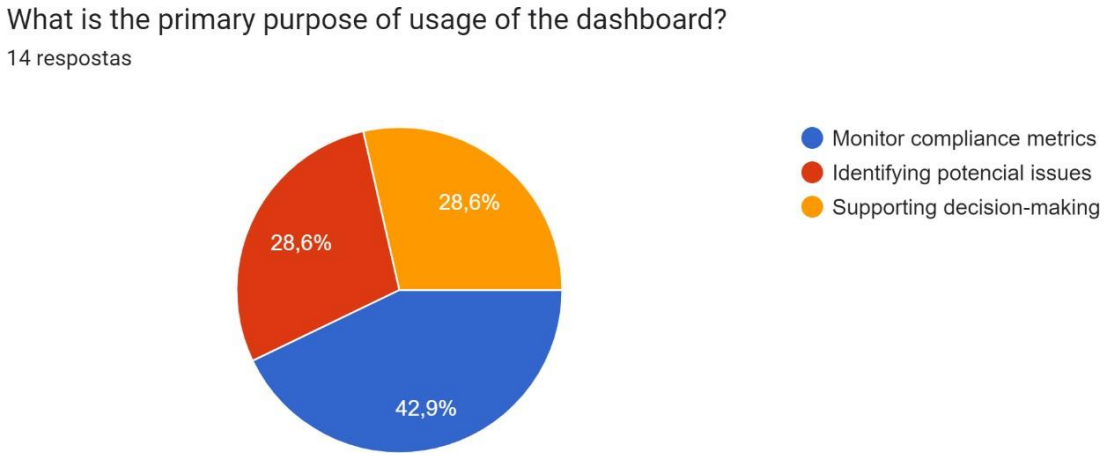

Figure 18 - Dashboard Usage

A key metric that was part of the survey was directly, in a scale of 1 to 5, where 1 stands for strongly disagree and 5 for strongly agree which can be seen on figure 19, what is the impact

of the dashboard on the decision-making process and more than 90% of the users consider that it impacts. This solution provides data visions and dynamic views that the compliance department of this bank never had in the past.

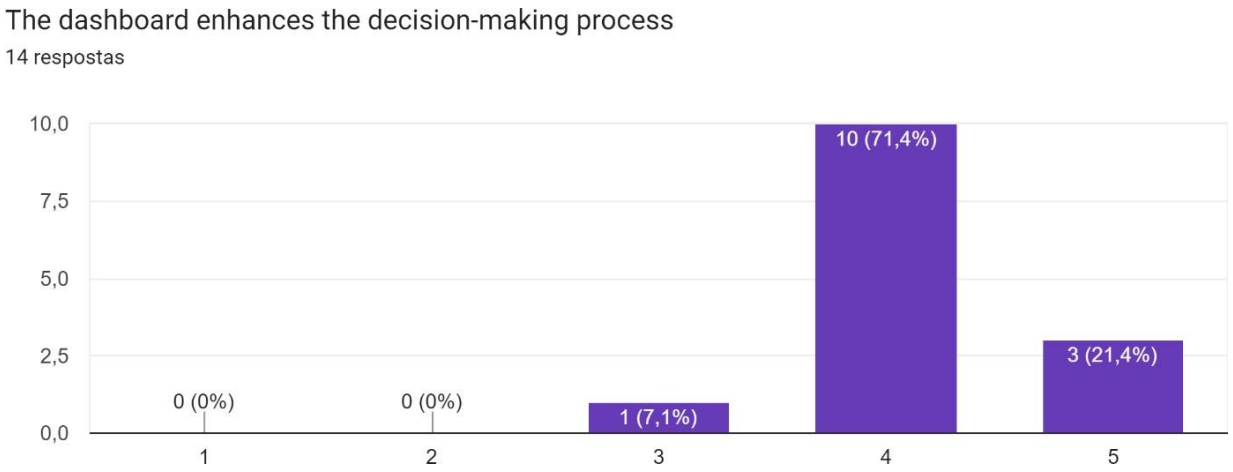

Figure 19 – Impact on decision making process

Another key metric that was considered was, as present on figure 20, again, a scale from 1-5 where 1 stands for strongly disagree and 5 for strongly agree, to helping identifying compliance issues through using the dashboard developed, and more than 90% of the users considered it

would impact on this end. By having the visions of the pending operations, the compliance department of a bank can easily identify where something on a specific process or workflow is not working that good.

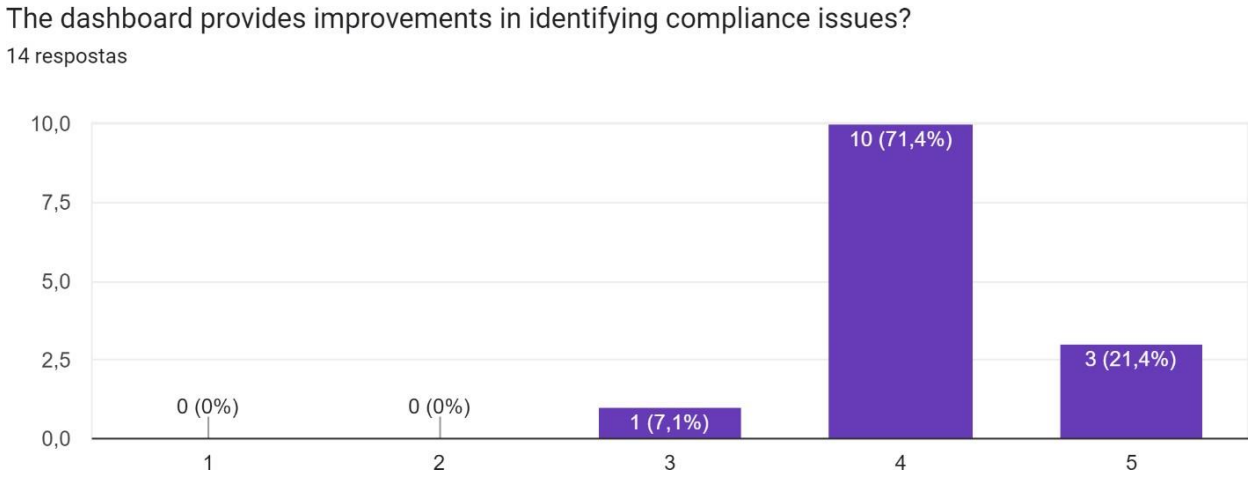

Figure 20 - Identification of compliance issues

On figure 21, most of the inquired considered that the dashboard helps to identify risks and impacts decision-making, increasing the efficiency of the compliance department.

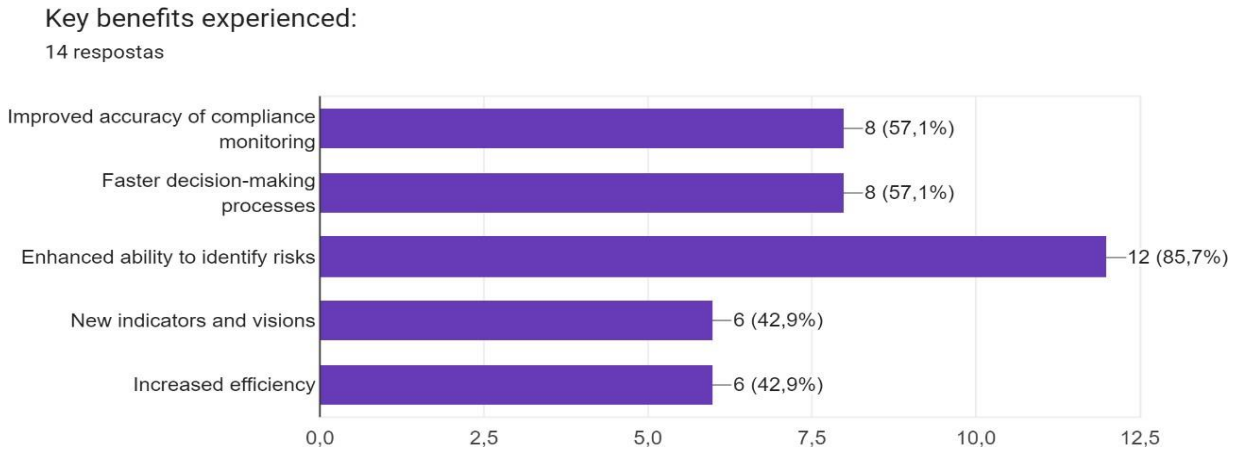

Figure 21 - Key Benefits

The questionnaire also featured some other questions, but these are the ones that are deeply linked to the need to investigate the impact of a BI application/solution within the decision-

making process and analytical approach taken by the compliance department of a bank.

Some other questions are related to user interface satisfaction, if the solution provides clear and accurate information, if the information is provided in a timely manner, if it helps identifying compliance issues, what is the overall satisfaction with the solution, the challenges experienced and suggestions for improvement.

5. CONCLUSIONS

In conclusion, the primary focus of this master thesis has been to address the critical need for enhanced decision-making capabilities and up-to-date analytics within the compliance department of a particular bank institution. By developing a comprehensive solution based on a functional BI framework including a data mart and dedicated views, a seamless and intuitive experience was provided to the compliance department, enabling them to navigate complex data landscapes with ease and efficiency. This involved the design and execution of various components, including the implementation of ETL processes using SSIS, the construction of a robust data mart architecture, and the development of metadata structures to facilitate data organization and accessibility.

The successful completion of the ETL processes allowed for the integration of data from several sources, ranging from client information to risk factors and AML alerts emitted by transactional activities. This data was then consolidated and stored within a data mart infrastructure. Additionally, the creation of views and metadata structures facilitated quick interpretation and analysis of compliance data.

One of the key highlights of this solution was the construction of a dynamic dashboard using Power BI. This dashboard served as the final piece of the solution, providing a comprehensive overview of client aspects, risk factors, AML alerts, and operational analyses. By leveraging the capabilities of Power BI, the solution developed provides a user-friendly interface that facilitated informed decision-making and proactive monitoring of compliance-related activities. Through interactive visualizations and customizable features, users can now gain valuable insights into their compliance operations, enabling them to identify bottlenecks, prioritize pending transactions, and respond promptly to emerging AML risks.

The developed solution offers tangible benefits to banking institutions seeking to enhance their compliance operations through technology-driven approaches, thereby mitigating risks and ensuring regulatory adherence.

6. LIMITATIONS AND FUTURE WORK

There were several challenges and adversities during the implementation of this design. Firstly, challenges related to data quality and consistency across all the sources was a challenge. Inconsistencies in data formats, completeness, and accuracy could potentially compromise the reliability of the DM and DW. The logics used to build the Data Mart logics involved several testing activities in QA (Quality) environment, which is a test environment, to indeed confirm and assure the correct SK's (surrogate keys) were being created in the fact tables, as well as assuring that no duplicates were found.

Secondly, technical complexities inherent in processes such as ETL (Extract, Transform, Load) in SSIS, data mart construction, and dashboard development have posed significant challenges during implementation.

Additionally, the daily activities of the bank and all the processes and applications used, which are constantly in flux, may present a limitation in the sense that it is quite complicated to keep up with the changes because managers may not have time to give the current business knowledge that teams need to develop solutions.

Future work might rely on exploring the integration of machine learning applications on top of the DM and leveraging machine learning tools will enhance decision-making processes as well as insights and different scenarios inherent to the compliance scope. Detection algorithms to identify unusual patterns and major deviations would be a great manner to identify suspicious AML transfers, flagging them for further investigation.

REFERENCES

Agrawal, R., Kadadi, A., Dai, X., & Andres, F. (2015). Challenges and opportunities with big data visualization. *Proceedings of the 7th International Conference on Management of Computational and Collective intElligence in Digital EcoSystems*, 169–173. https://doi.org/10.1145/2857218.2857256

Akiotu, H. (2022). Bank Management Compliance Strategies to Avoid Regulatory Sanctions. *Walden Dissertations and Doctoral Studies*. https://scholarworks.waldenu.edu/dissertations/13947

Allaymoun, M. H., Qaradh, S., Salman, M., & Hasan, M. (2023). Big Data Analysis and Data Visualization to Help Make a Decision—Islamic Banks Case Study. *International Conference on Business and Technology , ICBT 2021*, 54–63.

AL-Okaily, A., Ping, T., & Al-Okaily, M. (2021). Towards Business Intelligence Success Measurement in an Organization: A Conceptual Study. *Journal of System and Management Sciences*, *11*(2), 155–170.

Alshehadeh, A., Elrefae, G., Belarbi, A., Qasim, A., & Al-Khawaja, H. (2023). The impact of business intelligence tools on sustaining financial report quality in Jordanian commercial banks. *Uncertain Supply Chain Management*, *11*(4), 1667–1676.

Alshehadeh, A. R., & Al-khawaja, H. (2022). Financial Technology as a Basis for Financial Inclusion and its Impact on Profitability: Evidence from Commercial Banks. *International Journal of Advances in Soft Computing and Its Applications*, *14*, 126–138. https://doi.org/10.15849/IJASCA.220720.09

Analysis of the Introduction of Business Intelligence and Data Warehousing into Businesses in Latvia—[Scite report]. (n.d.). Retrieved 18 December 2023, from https://scite.ai/reports/analysis-of-the-introduction-of-vJkYmv5O?showReferences=true

Anywar, M., Schreiweis, B., & Ulrich, H. (2022). Strategies and Recommendation for Data Loading of FHIR-Based Data Marts with Focus on GDPR Compliance. In *Studies in health technology and informatics* (Vol. 298). https://doi.org/10.3233/SHTI220921

Aparicio, M., & Costa, C. J. (2015). Data visualization. *Communication Design Quarterly*, *3*(1), 7–11.

https://doi.org/10.1145/2721882.2721883

Ariyachandra, T., & Watson, H. (2010). Key organizational factors in data warehouse architecture selection. *Decision Support Systems*, *49*(2), 200–212. Scopus. https://doi.org/10.1016/j.dss.2010.02.006

Aziz, A., Saha, S., & Arifuzzaman, M. (2021). Analyzing Banking Data Using Business Intelligence: A Data Mining Approach. In M. S. Uddin & J. C. Bansal (Eds.), *Proceedings of International Joint Conference on Advances in Computational Intelligence* (pp. 245–256). Springer. https://doi.org/10.1007/978-981-16-0586-4_20

Baars, H., & Kemper, H.-G. (2008). Management support with structured and unstructured data—An integrated business intelligence framework. *Information Systems Management*, *25*(2), 132–148. Scopus. https://doi.org/10.1080/10580530801941058

Bany Mohammad, A., Al-Okaily, M., Al-Majali, M., & Masa'deh, R. (2022). Business Intelligence and Analytics (BIA) Usage in the Banking Industry Sector: An Application of the TOE Framework. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(4), 189. https://doi.org/10.3390/joitmc8040189

Bany Mohammed, A., Al-Okaily, M., Qasim, D., & Khalaf Al-Majali, M. (2024). Towards an understanding of business intelligence and analytics usage: Evidence from the banking industry. *International Journal of Information Management Data Insights*, *4*(1), 100215. https://doi.org/10.1016/j.jjimei.2024.100215

Basile, L. J., Carbonara, N., Pellegrino, R., & Panniello, U. (2023). Business intelligence in the healthcare industry: The utilization of a data-driven approach to support clinical decision making. *Technovation*, *120*, 102482. https://doi.org/10.1016/j.technovation.2022.102482

Böszörmenyi, J., & Schweighofer, E. (2015). A review of tools to comply with the Fourth EU anti-money laundering directive. *International Review of Law, Computers & Technology*, *29*(1), 63–77. https://doi.org/10.1080/13600869.2015.1016276

Brar, T. P. S. (2018). Business intelligence in banking: A study of bi technology implementation and challenges. *CGC International Journal of Contemporary Technology and Research*, *1*(1). https://www.cgcijctr.com/document/1.5.pdf

Cairo, A. (2013). *The Functional Art: An Introduction to Information Graphics and Visualization*. New Riders.

Caseiro, N., & Coelho, A. (2019). The influence of Business Intelligence capacity, network learning and innovativeness on startups performance. *Journal of Innovation &Amp; Knowledge*, *4*(3). https://doi.org/10.1016/j.jik.2018.03.009

Caserio, C., & Trucco, S. (2018). Business Intelligence Systems. In C. Caserio & S. Trucco (Eds.), *Enterprise Resource Planning and Business Intelligence Systems for Information Quality: An Empirical Analysis in the Italian Setting* (pp. 43–73). Springer International Publishing. https://doi.org/10.1007/978-3-319-77679-8_3

Chatzistefanou, D. (2023). *Data Warehousing in Business Intelligence and ETL Processes*.

Chaudhuri, S., & Dayal, U. (1997). An overview of data warehousing and OLAP technology. *ACM SIGMOD Record*, *26*(1), 65–74. https://doi.org/10.1145/248603.248616

Chen, C., Härdle, W., & Unwin, A. (2008). *Handbook of Data Visualization*. Springer. https://doi.org/10.1007/978-3-540-33037-0

Côrte-Real, N., Ruivo, P., & Oliveira, T. (2020). Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value? *Information &Amp; Management*, *57*(1). https://doi.org/10.1016/j.im.2019.01.003

Data Visualization. (n.d.). Springerprofessional.De. Retrieved 9 December 2023, from https://www.springerprofessional.de/en/data-visualization/7578166

Data Warehouse & Business Intelligence Architecture Guide. (n.d.). Retrieved 13 January 2024, from https://www.datapine.com/blog/data-warehousing-and-business-intelligence-architecture/

Data Warehouse Performance (world). (n.d.). Guide Books. https://doi.org/10.5555/288855

Dayal, U., Castellanos, M., Simitsis, A., & Wilkinson, K. (2009). *Data integration flows for Business Intelligence*. 1–11. Scopus. https://doi.org/10.1145/1516360.1516362

Design and Implementation of Efficient Decision Support System Using Data Mart Architecture | IEEE Conference Publication | IEEE Xplore. (n.d.). Retrieved 14 December 2023, from https://ieeexplore.ieee.org/abstract/document/9179313

Dilla, W. N., & Raschke, R. L. (2015). Data visualization for fraud detection: Practice implications and a call for future research. *International Journal of Accounting Information Systems*, *16*, 1–22. https://doi.org/10.1016/j.accinf.2015.01.001

Duggineni, S. (2022). Data Analytics in Modern Business Intelligence. *Journal of Marketing & Supply Chain Management*, 1–4. https://doi.org/10.47363/JMSCM/2023(2)114

Edwards, J., & Wolfe, S. (2004). The compliance function in banks. *Journal of Financial Regulation and Compliance*, *12*(3), 216–224. https://doi.org/10.1108/13581980410810795

Eidizadeh, R., Salehzadeh, R., & Esfahani, A. C. (2017). Analysing the role of business intelligence, knowledge sharing and organisational innovation on gaining competitive advantage. *Journal of Workplace Learning*, *29*(4). https://doi.org/10.1108/jwl-07-2016-0070

El-Adaileh, N. A., & Foster, S. (2019). Successful business intelligence implementation: A systematic literature review. *Journal of Work-Applied Management*, *11*(2), 121–132. https://doi.org/10.1108/JWAM-09-2019-0027

Gadda, K. R., & Dey, S. (2014, July 1). *Business Intelligence for Public Sector Banks in India: A Case study-Design, Development and Deployment. | Journal of Finance, Accounting & Management | EBSCOhost*. https://openurl.ebsco.com/contentitem/gcd:97594068?sid=ebsco:plink:crawler&id=ebsco:gcd:97594068

Gaol, F. L., Abdillah, L., & Matsuo, T. (2020). Adoption of Business Intelligence to Support Cost Accounting Based Financial Systems—Case Study of XYZ Company. *Open Engineering*, *11*(1). https://doi.org/10.1515/eng-2021-0002

Gardner, S. R. (1998). Building the data warehouse. *Communications of the ACM*, *41*(9), 52–60. https://doi.org/10.1145/285070.285080

Golfarelli, M., & Rizzi, S. (2009). A survey on temporal data warehousing. *International Journal of Data Warehousing and Mining*, *5*(1), 1–17. Scopus. https://doi.org/10.4018/jdwm.2009010101

Gonçalves, C. T., Gonçalves, M. J. A., & Campante, M. I. (2023a). Developing Integrated Performance Dashboards Visualisations Using Power BI as a Platform. *Information*, *14*(11), Article 11. https://doi.org/10.3390/info14110614

Gonçalves, C. T., Gonçalves, M. J. A., & Campante, M. I. (2023b). Developing Integrated Performance Dashboards Visualisations Using Power BI as a Platform. *Information*, *14*(11), Article 11. https://doi.org/10.3390/info14110614

Graves, R. (2005). *Business Intelligence Tools: The Smart Way to Achieve Compliance*. *15*(12). https://www.proquest.com/openview/4340cdd6104a126e0d3841afc630bc72/1?pq-origsite=gscholar&cbl=51938

Grossmann, W., & Rinderle-Ma, S. (2015). *Fundamentals of Business Intelligence*. Springer. https://doi.org/10.1007/978-3-662-46531-8

Ha, S. H., & Park, S. C. (1998). Application of data mining tools to hotel data mart on the Intranet for database marketing. *Expert Systems with Applications*, *15*(1), 1–31. https://doi.org/10.1016/S0957-4174(98)00008-6

Hamidinava, F., Ebrahimy, A., Samiee, R., & Didehkhani, H. (2021). A model of business intelligence on cloud for managing SMEs in COVID-19 pandemic *(Case: Iranian SMEs)*. *Kybernetes*, *52*(1). https://doi.org/10.1108/k-05-2021-0375

Hamoud, A. K., Hussien, H. N., Fadhil, A. A., & Ekal, Z. R. (2020). Improving Service Quality Using Consumers' Complaints Data Mart which Effect on Financial Customer Satisfaction. *Journal of Physics: Conference Series*, *1530*(1), 012060. https://doi.org/10.1088/1742-6596/1530/1/012060

Hamoud, A. K., Ulkareem, M. A., Hussain, H. N., Mohammed, Z. A., & Salih, G. M. (2020). Improve HR Decision-Making Based On Data Mart and OLAP. *Journal of Physics: Conference Series*, *1530*(1), 012058. https://doi.org/10.1088/1742-6596/1530/1/012058

Hariharan, N. K. (2018). DATA SOURCES FOR BUSINESS INTELLIGENCE. *International Journal of Innovations in Engineering Research and Technology*, *5*(11), 75–80.

Hendayun, M., Yulianto, E., Rusdi, J. F., Setiawan, A., & Ilman, B. (2021). Extract transform load process in banking reporting system. *MethodsX*, *8*, 101260. https://doi.org/10.1016/j.mex.2021.101260

Horani, O. M., Khatibi, A., AL-Soud, A. R., Tham, J., & Al-Adwan, A. S. (2023). Determining the Factors Influencing Business Analytics Adoption at Organizational Level: A Systematic Literature Review. *Big Data and Cognitive Computing*, *7*(3), Article 3. https://doi.org/10.3390/bdcc7030125

Hussein, M. F., Daud, P., Musa, O., Mohamad, N., & Ismail, N. L. (2023). Design and Implementation of Data Warehouse Solution at Kumpulan Wang Persaraan (KWAP). *Lecture Notes on Data Engineering and Communications Technologies*, *165*, 195–208. Scopus. https://doi.org/10.1007/978-981-99-0741-0_14

IBM Documentation. (2021, March 3). https://www.ibm.com/docs/en/atlas-policy-suite/6.0.3?topic=suite-data-source-definitions

Inmon, W. H. (2005). *Building the data warehouse* (3rd ed.). John Wiley & Sons, Inc.

Impact of Business Intelligence Adoption on performance of banks: A conceptual framework. (n.d.). Retrieved 1 May 2024, from http://ouci.dntb.gov.ua/en/works/9jxBMNn4/

Introducing Microsoft Power BI | Microsoft Press Store. (n.d.). Retrieved 9 December 2023, from https://www.microsoftpressstore.com/store/introducing-microsoft-power-bi-9781509302284

Iris, N., & Nagalingham, S. (2023). Implementation of Business Intelligence Solution for United Airlines. *International Journal of Advanced Computer Science and Applications*, *14*(1), 843–852. Scopus. https://doi.org/10.14569/IJACSA.2023.0140192

Jha, A. C., Jha, S. K., & Simha, J. B. (2023). Development of Analytical DataMart and Data Pipeline for

Recruitment Analytics. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, *471 LNICST*, 55–65. Scopus. https://doi.org/10.1007/978-3-031-35081-8_5

Jiang, L., Barone, D., Amyot, D., & Mylopoulos, J. (2011). Strategic Models for Business Intelligence. In M. Jeusfeld, L. Delcambre, & T.-W. Ling (Eds.), *Conceptual Modeling – ER 2011* (pp. 429–439). Springer. https://doi.org/10.1007/978-3-642-24606-7_33

Jourdan, Z., Rainer, R. K., & Marshall, T. E. (2008). Business intelligence: An analysis of the literature. *Information Systems Management*, *25*(2), 121–131. Scopus. https://doi.org/10.1080/10580530801941512

Jr, T., Jones, M., & Armstrong, C. (2007). The Dynamic Structure of Management Support Systems: Theory Development, Research Focus, and Direction. *MIS Quarterly*, *31*, 579–615. https://doi.org/10.2307/25148808

Jun, T., Kai, C., Yu, F., & Gang, T. (2009). The Research & Application of ETL Tool in Business Intelligence Project. *2009 International Forum on Information Technology and Applications*, *2*, 620–623. https://doi.org/10.1109/IFITA.2009.48

Kaula, R. (2015). Business Intelligence Rationalization: A Business Rules Approach. *Business and Management*. https://bearworks.missouristate.edu/articles-cob/30

Khorasani, S. A., & Chalmeta, R. (2022). Internet of Things Data Visualization for Business Intelligence. *Big Data*. https://doi.org/10.1089/big.2021.0200

Kimball, R., & Ross, M. (2002). *The data warehouse toolkit: The complete guide to dimensional modeling* (2nd ed). Wiley. http://www.books24x7.com/marc.asp?bookid=6694

Leo, M., Sharma, S., & Maddulety, K. (2019). Machine Learning in Banking Risk Management: A Literature Review. *Risks*, *7*(1), Article 1. https://doi.org/10.3390/risks7010029

Li, X., Deng, F., & Li, W. (2009). The Research and Application of an ETL Model Based on Task. *2009 First*

International Conference on Information Science and Engineering, 1006–1009. https://doi.org/10.1109/ICISE.2009.1257

Lin, Z., Zhu, M., Yin, W., & Dong, J. (2008). Banking Intelligence: Application of data warehouse in bank operations. *2008 IEEE International Conference on Service Operations and Logistics, and Informatics*, *1*, 143–146. https://doi.org/10.1109/SOLI.2008.4686380

Madlener, J. (2009). *The Implications of Integrating Governance, Risk and Compliance in Business Intelligence Systems on Corporate Performance Management*. https://repub.eur.nl/pub/15580/

Manolescu, I. (2020). Integrating (Very) Heterogeneous Data Sources: A Structured and an Unstructured Perspective. In J. Darmont, B. Novikov, & R. Wrembel (Eds.), *Advances in Databases and Information Systems* (pp. 15–20). Springer International Publishing. https://doi.org/10.1007/978-3-030-54832-2_3

March, S. T., & Hevner, A. R. (2007). Integrated decision support systems: A data warehousing perspective. *Decision Support Systems*, *43*(3), 1031–1043. Scopus. https://doi.org/10.1016/j.dss.2005.05.029

Martins, J., Mamede, H. S., & Correia, J. (2022). Risk compliance and master data management in banking – A novel BCBS 239 compliance action-plan proposal. *Heliyon*, *8*(6), e09627. https://doi.org/10.1016/j.heliyon.2022.e09627

Mathias, A., Kessler, K., & Bhatnagar, S. (2011). Using business intelligence for strategic advantage in REMS. *Journal of Medical Marketing*, *11*(1), 84–89. https://doi.org/10.1057/jmm.2010.40

Metre, K. V., Mathur, A., Dahake, R. P., Bhapkar, Y., Ghadge, J., Jain, P., & Gore, S. (2024). An Introduction to Power BI for Data Analysis. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(1s), Article 1s.

Mirzaei, A., & Moore, T. (2014). What are the driving forces of bank competition across different income groups of countries? *Journal of International Financial Markets, Institutions and Money*, *32*, 38–71. https://doi.org/10.1016/j.intfin.2014.05.003

Moody, D., & Kortink, M. (2000). *From enterprise models to dimensional models: A methodology for data*

warehouse and data mart design. 5.

Mositsa, R. J., Van der Poll, J. A., & Dongmo, C. (2023). Towards a Conceptual Framework for Data Management in Business Intelligence. *Information*, *14*(10), Article 10. https://doi.org/10.3390/info14100547

Muriithi, G. M., & Kotzé, J. E. (2013). A conceptual framework for delivering cost effective business intelligence solutions as a service. *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference*, 96–100. https://doi.org/10.1145/2513456.2513502

Negash, S. (2004). Business Intelligence. *Communications of the Association for Information Systems*, *13*(1). https://doi.org/10.17705/1CAIS.01315

Orlovskyi, D., & Kopp, A. (2021). *A business intelligence dashboard design approach to improve data analytics and decision making*. *2833*, 48–59. Scopus.

Ortiz, L., & Hallo, M. (2019). Analytical Data Mart for the Monitoring of University Accreditation Indicators. *2019 IEEE World Conference on Engineering Education (EDUNINE)*, 1–6. https://doi.org/10.1109/EDUNINE.2019.8875826

Owusu, A. (2017). Business intelligence systems and bank performance in Ghana: The balanced scorecard approach. *Cogent Business & Management*, *4*(1), 1364056. https://doi.org/10.1080/23311975.2017.1364056

Owusu, A., & Said, A. (2016). An integrated model for determining business intelligence systems adoption and post-adoption benefits in banking sector. *Journal of Administrative and Business Studies*, *2*, 84–100.

Panda, B. (2012, May 30). *Problems of Bank for Different Data source and its solution through Business Intelligence*. SlideShare. https://www.slideshare.net/pandabishnu/problems-of-bank-for-different-data-source-and-its-solution-through-business-intelligence

Panja, R., & Paul, S. (2014). A Review Report on the Evolution and Implementation of Business Intelligence Technique in the Banking Sector. *International Journal of Computer Science and Mobile Applications*, *2*, 109–114.

Pieket Weeserik, B., & Spruit, M. (2018a). Improving Operational Risk Management Using Business Performance Management Technologies. *Sustainability*, *10*(3), Article 3. https://doi.org/10.3390/su10030640

Pieket Weeserik, B., & Spruit, M. (2018b). Improving Operational Risk Management Using Business Performance Management Technologies. *Sustainability*, *10*(3), Article 3. https://doi.org/10.3390/su10030640

Ponelis, S. R. (Shana R. (2002). *Data marts as management information delivery mechanisms: Utilisation in manufacturing organisations with third party distribution* [Thesis, University of Pretoria]. https://repository.up.ac.za/handle/2263/27061

Post, F. H., Nielson, G., & Bonneau, G.-P. (2002). *Data Visualization: The State of the Art*. Springer Science & Business Media.

Power BI - Visualização de Dados | Microsoft Power Platform. (n.d.). Retrieved 13 July 2024, from https://www.microsoft.com/pt-pt/power-platform/products/power-bi

Pugna, I., Felicia, A., & Babeanu, D. (2009). The role of business intelligence in business performance management. *Annals of Faculty of Economics*, *4*, 1025–1029.

Rahman, Md. M. (2023). The Effect of Business Intelligence on Bank Operational Efficiency and Perceptions of Profitability. *FinTech*, *2*, 99–119. https://doi.org/10.3390/fintech2010008

Rai, A., Dubey, V., Chaturvedi, K. K., & Malhotra, P. K. (2008). Design and development of data mart for animal resources. *Computers and Electronics in Agriculture*, *64*(2), 111–119. https://doi.org/10.1016/j.compag.2008.04.009

Raman, V., & Hellerstein, J. M. (2001). *Potter's wheel: An interactive data cleaning system*. 381–390. Scopus.

Rao, G. K., & Kumar, R. (2011). *Framework to Integrate Business Intelligence and Knowledge Management in Banking Industry* (arXiv:1109.0614). arXiv. https://doi.org/10.48550/arXiv.1109.0614

Rodrigues, J. V. F. (2019). *Compliance nas instituições bancárias: Relação com o sistema de controlo interno e a auditoria interna* [masterThesis]. https://repositorio.ucp.pt/handle/10400.14/30249

Rouhani, S., Ashrafi, A., Zareravasan, A., & Afshari, S. (2016). The impact of business intelligence on decision support and organizational benefits. *Journal of Enterprise Information Management*, *29*. https://doi.org/10.1108/JEIM-12-2014-0126

Sadiku, M., Shadare, A., Musa, S., Akujuobi, C., & Perry, R. (2016). DATA VISUALIZATION. *International Journal of Engineering Research and Advanced Technology (IJERAT)*, *12*, 2454–6135.

Santos, M. Y., & Ramos, I. (2006). *Business Intelligence: Tecnologias da informação na gestão de conhecimento*. FCA - Editora de Informática, Lda. https://repositorium.sdum.uminho.pt/handle/1822/6198

Shao, C., Yang, Y., Juneja, S., & GSeetharam, T. (2022). IoT data visualization for business intelligence in corporate finance. *Information Processing and Management: An International Journal*, *59*(1). https://doi.org/10.1016/j.ipm.2021.102736

Shi, Y., & Lu, X. (2010). The Role of Business Intelligence in Business Performance Management. *2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering*, *4*, 184–186. https://doi.org/10.1109/ICIII.2010.522

Stancu, A.-M., & Cristescu, M. (2017). *SOLUTION FOR DATA ORGANIZATION IN THE BANKING DOMAIN*. *69*, 96–104.

Su, S. I., & Chiong, R. (2011). Business Intelligence. In *Encyclopedia of Knowledge Management, Second Edition* (pp. 72–80). IGI Global. https://doi.org/10.4018/978-1-59904-931-1.ch008

Sundjaja, A. (2013). IMPLEMENTATION OF BUSINESS INTELLIGENCE ON BANKING, RETAIL, AND EDUCATIONAL INDUSTRY. *CommIT (Communication and Information Technology) Journal*, *7*, 65.

https://doi.org/10.21512/commit.v7i2.586

Teaching Students to Focus on the Data in Data Visualization—Joanna Wolfe, 2015. (n.d.). Retrieved 18 December 2023, from https://journals.sagepub.com/doi/full/10.1177/1050651915573944

Technology, V. (n.d.). *What's the Difference Between Internal and External Data?* Retrieved 6 December 2023, from https://www.ventivtech.com/blog/whats-the-difference-between-internal-and-external-data

Trujillo, J., & Palomar, M. (1998). *An object oriented approach to multidimensional database conceptual Modeling (OOMD)*. *Part F129242*, 16–21. Scopus. https://doi.org/10.1145/294260.294266

Tryfona, N., Busborg, F., & Borch Christiansen, J. G. (1999). starER: A conceptual model for data warehouse design. *Proceedings of the 2nd ACM International Workshop on Data Warehousing and OLAP*, 3–8. https://doi.org/10.1145/319757.319776

Utami, A., Pratama, B. R., & Widianto, S. R. (2020). DATA MART DESIGN IN BKPP BANDUNG USING FROM ENTERPRISE MODELS TO DIMENSIONAL MODELS METHOD. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, *5*(2), Article 2. https://doi.org/10.33480/jitk.v5i2.1219

Vaisman, A., & Zimányi, E. (2014). *Data Warehouse Systems: Design and Implementation*. Springer. https://doi.org/10.1007/978-3-642-54655-6

Vassiliadis, P. (2009). A survey of extract-transform-load technology. *International Journal of Data Warehousing and Mining*, *5*(3), 1–27. Scopus. https://doi.org/10.4018/jdwm.2009070101

Vassiliadis, P., Simitsis, A., & Skiadopoulos, S. (2002). *Conceptual Modeling for ETL Processes*. https://doi.org/10.1145/583890.583893

Vuksic, V., Pejic Bach, M., & Popovič, A. (2013). Supporting performance management with business process management and business intelligence: A case analysis of integration and orchestration. *International Journal of Information Management*, *33*, 613–619. https://doi.org/10.1016/j.ijinfomgt.2013.03.008

Watson, H. (2001). Recent Developments in Data Warehousing. *Communications of the Association for*

Information Systems, *8*. https://doi.org/10.17705/1CAIS.00801

What are Data Source Types? (n.d.). Insightsoftware. Retrieved 1 December 2023, from https://insightsoftware.com/encyclopedia/data-source-types/

What is Data Masking? - Static and Dynamic Data Masking Explained - AWS. (n.d.). Amazon Web Services, Inc. Retrieved 14 December 2023, from https://aws.amazon.com/what-is/data-masking/

What Is OLTP? (n.d.). Retrieved 8 July 2024, from https://www.oracle.com/database/what-is-oltp/

Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly: Management Information Systems*, *25*(1), 17–41. Scopus. https://doi.org/10.2307/3250957

Xi, X., & Hongfeng, X. (2009). Developing a Framework for Business Intelligence Systems Integration Based on Ontology. *2009 International Conference on Networking and Digital Society*, *2*, 288–291. https://doi.org/10.1109/ICNDS.2009.151

Zasada, A., & Bui, T. (2018, August 18). *More Than Meets the Eye: A Case Study on the Role of IT Affordances in Supporting Compliance*.

Zhang, H., Ren, S., Li, X., Baharin, H., Alghamdi, A., & Alghamdi, O. A. (2023). Developing scalable management information system with big financial data using data mart and mining architecture. *Information Processing & Management*, *60*(3), 103326. https://doi.org/10.1016/j.ipm.2023.103326

Zhen, H., & Yee, M. (2023). Data Rather Than Germs on Your Fingertips: Leveraging Business Intelligence to Improve and Sustain Hand Hygiene Compliance. *American Journal of Infection Control*, *51*(7, Supplement), S13. https://doi.org/10.1016/j.ajic.2023.04.152

Zhou, Q., Tao, H., & Tao, W. (2008). Analysis of Business Intelligence and Its Derivative—Financial Intelligence. In *Proceedings of the International Symposium on Electronic Commerce and Security, ISECS 2008* (p. 1000). https://doi.org/10.1109/ISECS.2008.28

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa