Department: Mathematics and Computer Science

End-of-study internship Report

Under the theme:

Implementation of a decision support system for optimizing the sales of agri-food products.

Realized by:

-Zakaria Najib

Project Supervisor:

-Mr. Mouhsine AIT BELLA

Academic Year 2023-2024

Dedication

Alhamdo liLLAH, who has guided us on this path and raised us in knowledge, without Him we would not have reached where we are today.

I dedicate this humble report to:
My dear parents for their financial and moral support. And, of course, no tribute could match the love they continue to shower on me, their sacrifice, and their selflessness to help me reach this level. I pray to God to grant them health and a long life.

To my sister and brothers for their kindness, valuable support, encouragement throughout our school years, and their love and affection. May this work be an expression of my esteem for you, and may God protect you.

To my dear friends and family. To all my dear teachers and mentors, for their help and support throughout my training period.
And to all those who have contributed directly or indirectly to bringing this project to fruition. I thank you immensely.

Acknowledgment

First and foremost, I extend my thanks, respect, and gratitude to all the individuals who contributed to the success of my work and assisted me during the writing of this report.

I extend my deepest appreciation to Mr. ESSALHI NABIL, Director of Information Systems Management in COPAG cooperative, who kindly welcomed me into this the cooperative. I would also like to express my gratitude to Mr. MOHSSINE AIT BELLA my internship supervisor, who greatly assisted me in adapting to the professional environment and provided high-level guidance and advice throughout my internship at COPAG.

We also extend our thanks to our other professors for their dedication and teaching throughout our academic journey. Their knowledge and advice have been invaluable for our professional and personal development.

We also want to thank our friends and family for their unwavering support, encouragement, and patience during this busy period. Their moral support has been a source of inspiration and motivation throughout this journey.

Each of these individuals played a crucial role in the realization of this project, and we are deeply grateful to them. Their contribution was essential to its success.

Summary

As part of marking the end of my second-year studying in DUT IDSD (Decision-Making Informatics and Data Sciences) at the School of Technology in Essaouira, a part of Cadi Ayyad University, I was required to complete an end-of-study internship at the cooperative Copag in Taroudant.

The key features of our project are : the development of a decision support system for optimizing the sales of agro-food products.

Several computer tools studied during the internship period are presented, along with their utilization in the context of supervising the IT infrastructure of the cooperative Copag in Taroudant.

Notably, I worked with two important tools: Power BI and SSIS. Power BI was used to create interactive and insightful dashboards, while SSIS (SQL Server Integration Services) enhanced our data analysis and reporting capabilities.

During my internship, I was involved in various aspects of the business intelligence (BI) project. Specifically, the global BI workflow architecture we followed had three important phases: Collect, Analyze, and Dashboard, based on OLAP (Online Analytical Processing) and Kimball architecture.

This document outlines the various stages undertaken for the completion of this project, beginning with the introduction, followed by the design of the information systems, and ultimately the implementation and deployment of the systeme.

The report also presents the decision-making informatics aspect I utilized during my internship. This part of my work involved the development and implementation of a prototype decision support system for optimizing the sales of agro-food products. Specifically, I employed Power BI and SSIS to enhance data analysis and reporting capabilities, ultimately improving the cooperative's sales strategies and operations.

Table of Contents

List of Figures

General Introduction

As part of my end-of-study internship at the Essaouira Higher School of Technology, I had the opportunity to carry out my final project at COPAG cooperative in Taroudant. I chose this cooperative because they provide an excellent data internship program, which aligns with my professional aspirations in the field of data science and business intelligence.

The objective of this internship is to put into practice the knowledge and various skills acquired during my academic training.

This report is the account of the internship I conducted at COPAG cooperative from 16/04/2024 to 01/06/2024 During this month and a half of work, I underwent training in data analysis and business intelligence tools such as Power BI and SSIS. This training facilitated the development and deployment of a decision support system for optimizing the sales of agro-food products.

The selection of this project aims to enhance the knowledge acquired during my academic journey and serves as a platform to amalgamate my theoretical insights. It provides an opportunity to gain practical experience in data analysis and business intelligence, refine my understanding of the corporate landscape, and conduct a comprehensive and profound analysis of the cooperative and its operating environment. Simultaneously, it equips me with valuable skills and knowledge that will serve as a solid foundation for my professional career.

This report is structured into four distinct chapters. The initial chapter focuses on introducing the cooperative and presenting the overarching context of the project. The second chapter delves into Concepts of business intelligence. In the third chapter, we detail the tools leveraged for the project's implementation. Lastly, the fourth chapter is dedicated to the implementation phase, accompanied by the presentation of interface screenshots.

We will conclude this report with a comprehensive summary that not only encapsulates our findings but also offers valuable insights for potential future enhancements, ensuring our work remains focused and tailored to a specific target audience.

Chapter 1: General Project Context

1.Introduction

Before delving into the specifics of how this project was executed, it's essential to begin by introducing the host organization, which refers to the cooperative where the internship was conducted. Furthermore, provides an overview of the project's framework and general context, which entails defining its boundaries, elucidating the methodology employed, outlining the schedule adhered to, elucidating the problem statement, and articulating the diverse set of objectives aimed to be accomplished.

2.Presentation of the host organization

2.1Definition of the Cooperative:

A cooperative is a type of enterprise based on the principle of cooperation. Its objective is to best serve the economic interests of its participants. It differs from a non-profit association, which is less focused on economic activities, and from a commercial company, which makes a distinction between its associates and its clients.

A cooperative is an autonomous association of individuals who voluntarily unite to meet their common economic, social, and cultural needs and aspirations through a jointly-owned and democratically-controlled enterprise.

The economic objective of cooperatives is not the pursuit of profit, but the satisfaction of the members' economic, social, and cultural aspirations and needs, such as reducing production costs or the purchase cost of certain products.

2.2History and Creation:

The cooperative COPAG was established during the general assembly on May 7, 1987, benefiting from the export liberalization policy initiated by the Moroccan state. At that time, 39 farmers from the Taroudant region felt the need and necessity to associate within a cooperative to take control of their agricultural products, from production to a more advanced stage of distribution.

Thus, the cooperative COPAG was born, representing one of the greatest and most successful stories in Morocco.

Since its creation, COPAG has experienced remarkable success, becoming a major player in Morocco's agricultural sector. Its cooperative model has enabled farmers in the Taroudant region to develop and market their agricultural products efficiently and profitably. Thanks to a strategic vision and strong commitment from its members, COPAG has established itself as a reference in the Moroccan agricultural industry, thereby contributing to the economic and social development of the region.

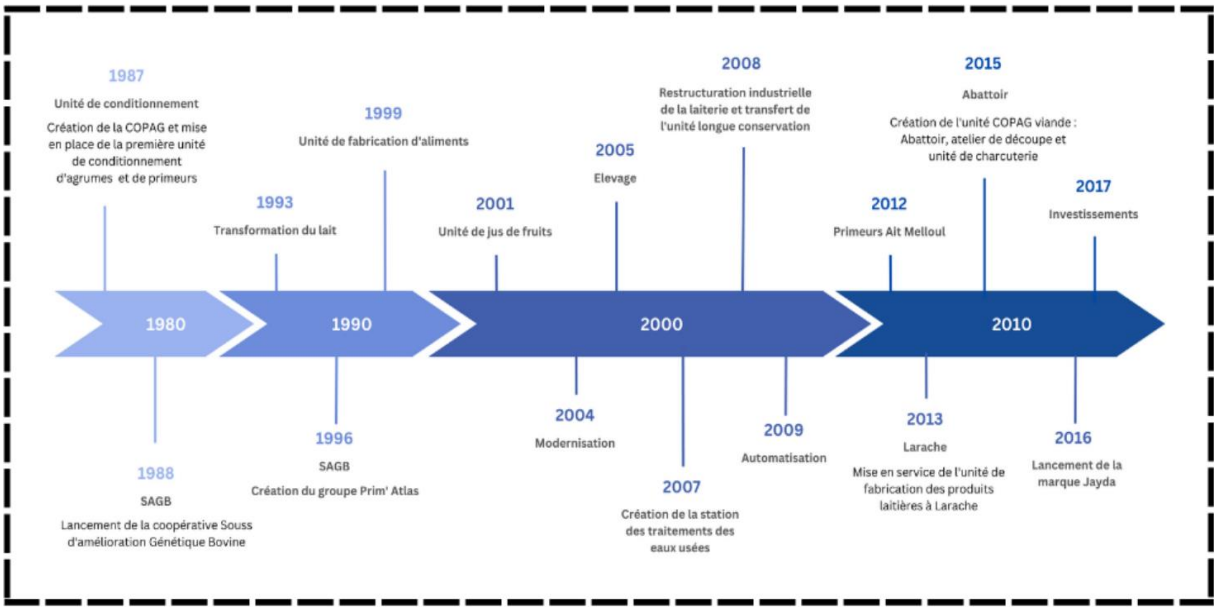


Figure 1 The history of COPAG.

2.3Cooperative COPAG

Is an economic organization composed of several coordinated services to form a well-organized network, aiming to ensure smooth operations and consequently contribute to economic development. It operates in the agricultural sector in the broadest sense: animal production (milk, meat), plant production (citrus fruits, early produce), agro-food, and more.

The development of COPAG is the result of a dual strategy: integrating the stages of the agricultural sector and maintaining an open-door policy to increase the number of members and grow the quantity of milk collected and processed.

Being a versatile cooperative, COPAG does not limit itself to the production, processing, and commercialization of milk and its derivatives. It also engages in the production and export of citrus fruits and early produce. The total exploited area reaches 11,100 hectares, distributed as follows: 4,000 hectares of citrus fruits, 1,100 hectares of early produce, and 6,000 hectares of forage crops. Regarding the bovine livestock, it consists of 80,000 heads, including 40,000 dairy cows.

2.4 Technical Data Sheet of the Cooperative

Company Name:	COPAG Agricultural Cooperative Taroudant
Legal Form:	Agricultural Cooperative
Creation Date:	May 7, 1987
Social Capital:	169,220,000 DH
Head Office:	Center AIT IAZZA circle Freija province of Taroudant
Name:	"COPAG-Taroudant" Cooperative of First Producers and Citrus Growers of Taroudant
Main Activity:	Agriculture
Number of Members:	179 farmer-breeders (67 cooperatives and 13,000 producers)
Workforce:	7,000 people
Registration Number at CNSS:	1283 085
Address:	Ait IAZZA Freija Taroudant B.P: 1001
Revenue:	3.23 Billion DH
Investments:	1.27 Billion DH

Figure 2 Technical data sheet

2.5Organization

COPAG operates in various fields of activity related to both plant and animal production, structured as follows:

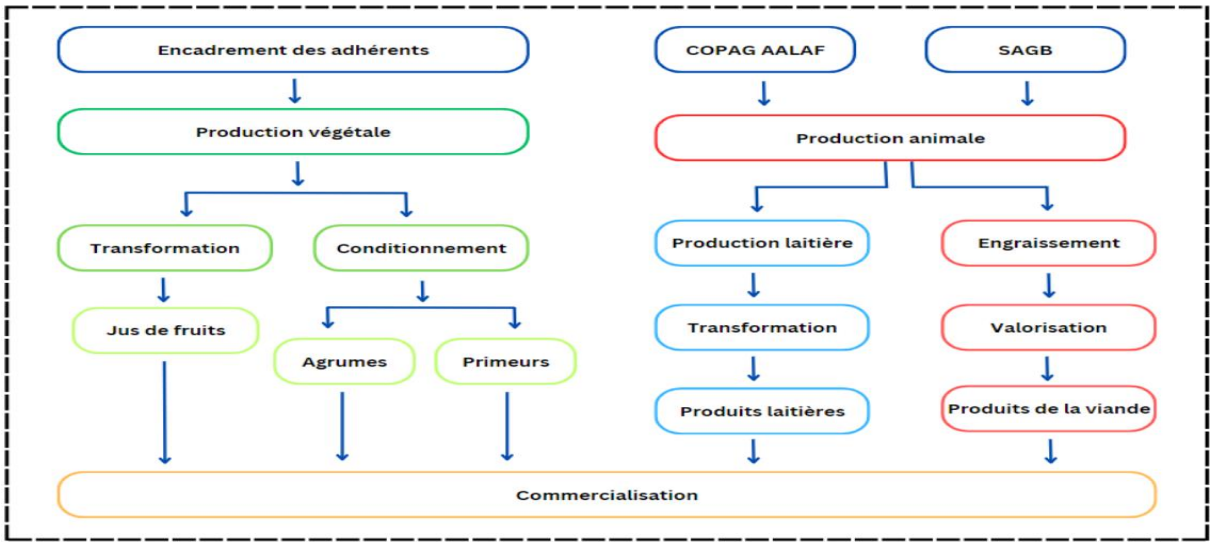


Figure 3 The organization of COPAG

2.6Organizational Chart

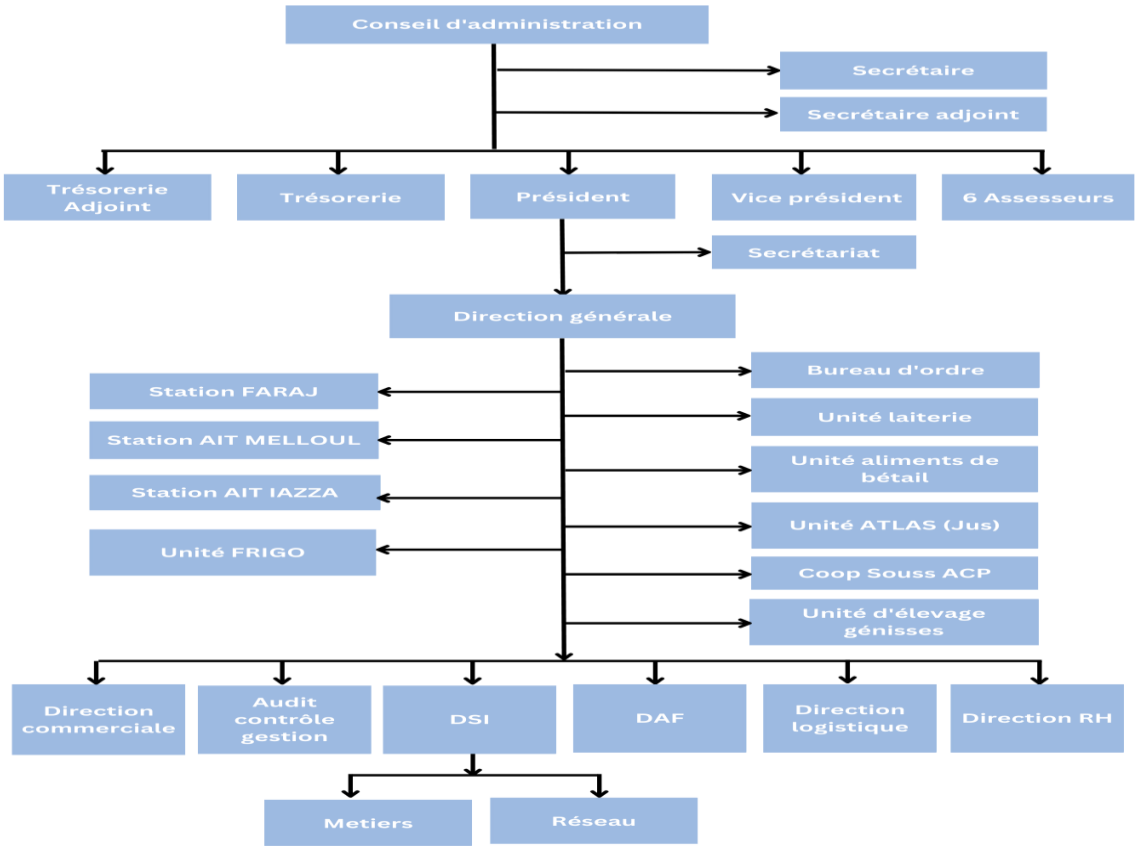


Figure 4 The organizational chart of COPAG

2.7Objectives of COPAG

All COPAG units aim to fully satisfy the needs of farmers, specifically:

- The reception, storage, preparation, packaging, and processing of plant and animal products from the farms of its members.
- The study, research, and implementation of means to improve the situation of producers, product quality, and the modernization of farms and the profession in general.
- The marketing of products from the farms of its members both for export and the domestic market, as well as products or by-products resulting from their processing or transformation.
- The joint purchase on the domestic or foreign market of fertilizers, treatment products, seeds, plants, veterinary products, livestock feed, cattle, sheep, poultry, agricultural equipment, breeding equipment, irrigation equipment, greenhouse equipment and supplies, packaging,

conservation and wrapping materials, spare parts, fuels, lubricants, and all supplies or tools necessary for the farms of its members and its main purpose.

- The evacuation, supply, and setup of all products used, harvested, bought, sold, or to be sold to meet the needs of the cooperative, transportation, shipping, receipt, customs formalities, and other operations.

- Grouping for all matters concerning the aforementioned goods and products.

3.Project Issues and Objectives

3.1Problem Statement

To design and implement a decision support system based on Business Intelligence (BI) that enables COPAG to effectively analyze, interpret, and use sales data of agro-food products to optimize commercial strategies, maximize revenues, and proactively and efficiently respond to changing market needs, several key steps and challenges must be addressed:

- **Data Volume Management**: Handling the massive influx of data from various sources without compromising accuracy and processing efficiency can be overwhelming.

- **Tool Integration**: Integrate and optimize various tools like SQL Server, Visual Studio, Power BI Desktop, and Power BI Services for smooth operation.

- **Data Integrity Risks**: Ensuring the accuracy and consistency of data in staging and fact tables before loading into the data warehouse involves rigorous validation processes that can be prone to errors.

- **ETL Process Optimization**: Optimize the Extract, Transform, Load (ETL) process to handle large data volumes efficiently and avoid performance bottlenecks.

- **User-Friendly Interface**: Designing a user-friendly interface that allows stakeholders to easily access, interpret, and utilize the data for strategic planning and operational adjustments.

By addressing these components, we can implement a decision support system that effectively optimizes the sales of agri-food products, maximizes revenue, and enhances the company's ability to meet market demands proactively and efficiently.

3.2Study of the Existing Situation

The selection of Business Intelligence (BI) tools is often seen as the first difficulty in a BI project. Indeed, choosing a tool that is poorly suited to the need can prove to be a real obstacle to the project's progress.

Figure 5 ETL tools

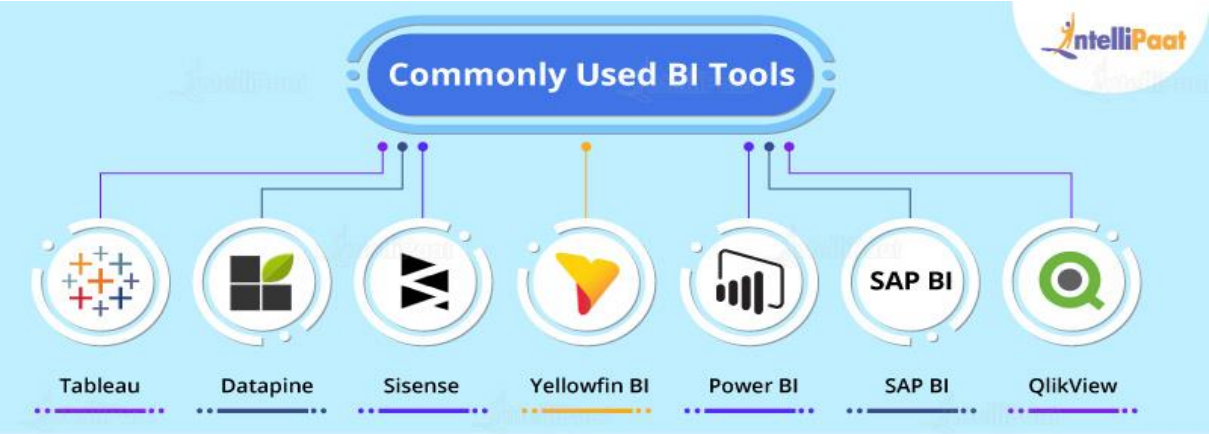


Figure 6 (BI) tools

For our ETL processes, we have chosen SQL Server Integration Services (SSIS) due to its robust data integration capabilities. SSIS provides a comprehensive platform for efficiently extracting, transforming, and loading data from various sources into a centralized data warehouse. This tool can handle complex data integration tasks, making it ideal for managing the diverse and high-volume data requirements of COPAG.

For our business intelligence and reporting needs, we have chosen Power BI due to its powerful capabilities for creating interactive and real-time dashboards and reports. Power BI allows us to visualize data dynamically, enabling users to explore data in depth and make informed decisions quickly. Its intuitive and user-friendly interface makes it accessible to a wide range of users, from technical analysts to business managers, facilitating widespread adoption across the organization.

3.3Objectives

The primary objective of this BI project is to design and implement a comprehensive decision support system for COPAG, which will enhance the company's ability to analyze, interpret, and utilize sales data related to agro-food products effectively. This system aims to optimize COPAG's commercial strategies, maximize revenues, and respond proactively and efficiently to the changing demands of the market.

- **Data Integration and Management**: Develop an efficient ETL process using SQL Server Integration Services (SSIS) to streamline data collection, transformation, and loading from multiple sources into a centralized data warehouse.

- **Advanced Analytics and Reporting:** Utilize Power BI to create interactive, real-time dashboards and reports, enabling dynamic data visualization and in-depth analysis.

- **User Accessibility and Experience:** Design a user-friendly interface in Power BI that is accessible to a wide range of stakeholders, from technical analysts to business managers.

- **Data Warehouse Design:** Implement a robust data warehouse architecture using Star and Snowflake Schemas to organize data efficiently for analytical processing**.**

By achieving these objectives, the project will enable COPAG to enhance its data-driven decision-making capabilities, optimize sales strategies, and respond more effectively to market trends and customer needs. This comprehensive BI system will support COPAG in achieving its business goals and sustaining competitive advantage in the agro-food industry.

4.Project Schedule and Timeline

4.1 Gantt Chart

Project planning aims to establish reasonable forecasts and scheduling of phases regarding the timing of engineering work and project management implementation. These forecasts are essential for effective project management. Among the project planning tools, we have chosen the Gantt chart, a tool used in scheduling and project management, which allows visualizing the timing of the various tasks that make up a project. It is a representation of a connected and directed graph that graphically shows the project's progress during this internship.

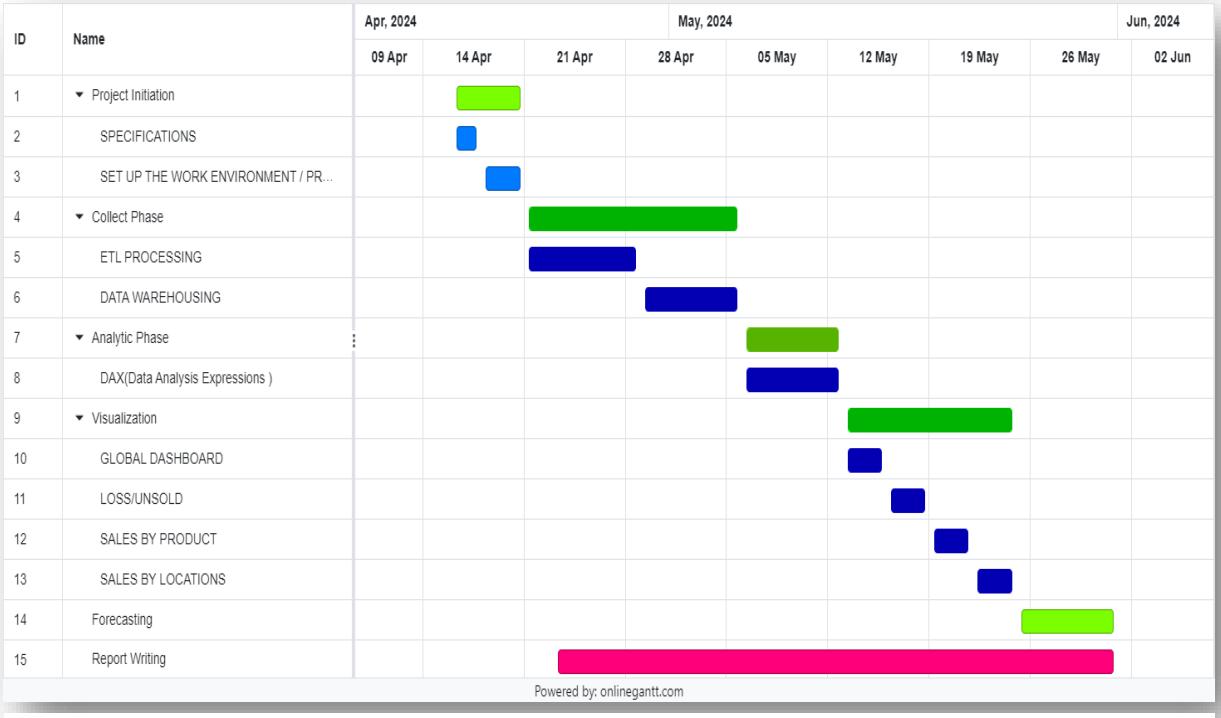


Figure 7 Gantt Chart

4.2Work Breakdown Structure (WBS)

The Work Breakdown Structure (WBS) is a hierarchical decomposition of the work required to achieve the project's objectives. Its purpose is to assist in organizing the project by defining the complete project scope. It is also used to guide risk management and identify necessary acquisitions.

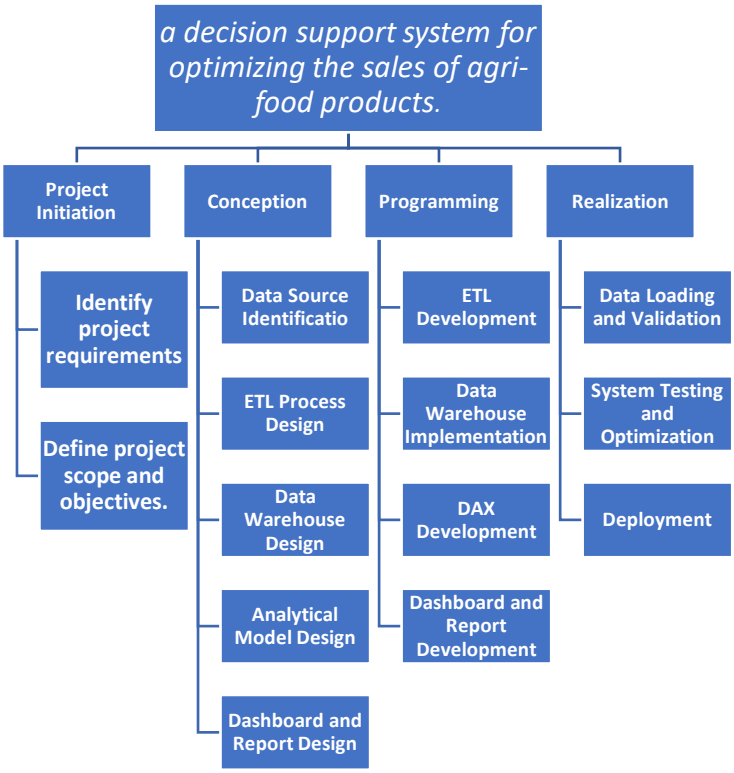


Figure 8 Work Breakdown Structure

Conclusion

The Global BI Workflow Architecture for COPAG is a comprehensive framework designed to enhance data-driven decision-making. By following a structured approach to data collection, analysis, and presentation, this architecture ensures that COPAG can efficiently transform raw data into valuable insights. The use of advanced tools like SSIS for ETL processes and Power BI for visualization ensures that the BI system is robust, scalable, and user-friendly, ultimately supporting the company's strategic objectives and improving overall business performance.

Chapter 2 : Concept of Business Intelligence

1.Introduction to Business Intelligence

Business Intelligence (BI) refers to the technologies, applications, and practices used for the collection, integration, analysis, and presentation of business information. The main purpose of BI is to support better business decision-making. Essentially, BI systems are data-driven decision support systems (DSS). BI encompasses various tools and methodologies that enable organizations to gather data from internal and external sources, prepare it for analysis, develop and run queries, and create reports, dashboards, and data visualizations to make the analytical results available to corporate decision-makers as well as operational workers.

The BI process can be diagrammed as follows:

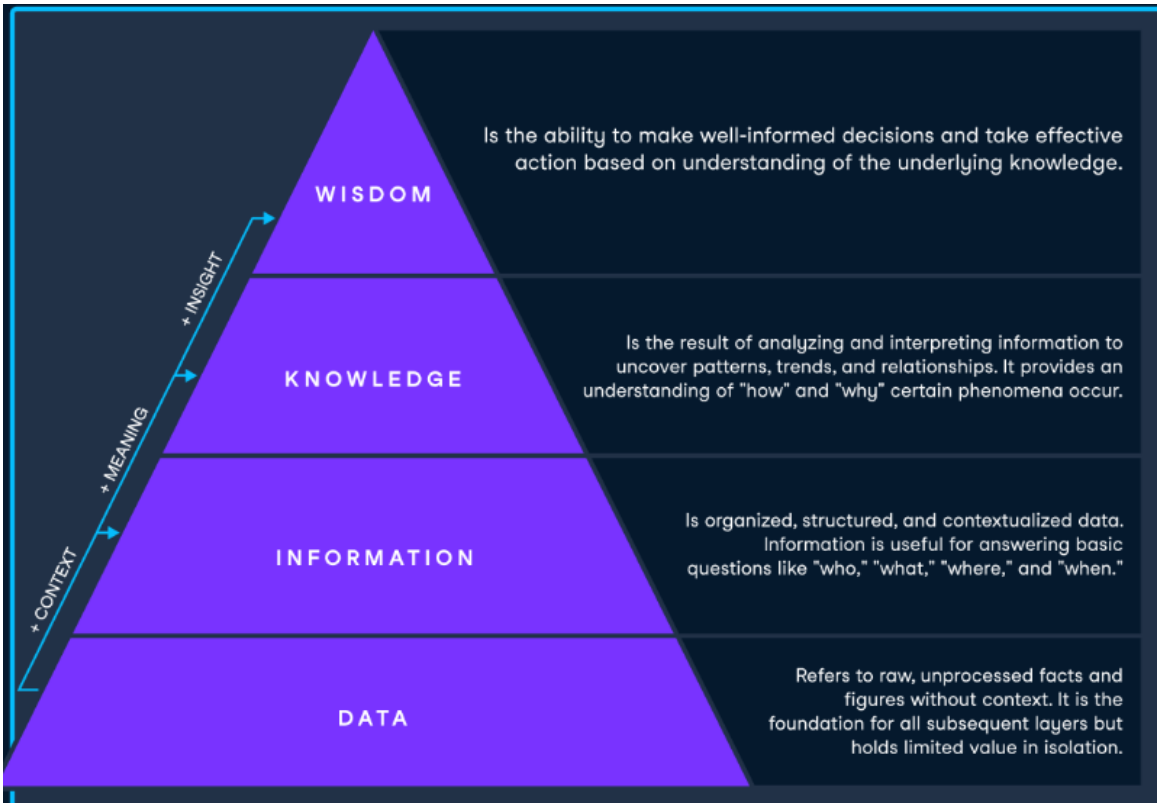


Figure 9 Knowledge Pyramid

The BI process can transform data into information and information into knowledge, thereby enabling the company to evolve, gain a competitive advantage, and build a knowledge base for Knowledge Management.

2. Opérationnel vs Décisionnel

To understand the essence of BI, it is essential to distinguish between two computing worlds: operational and decision-making. This distinction is key to integrating BI.

The Operational World:

The operational world consists of daily and repetitive tasks performed by company employees to keep the business running. Processing an order, issuing an invoice, and packaging a product are necessary tasks for a company's operations. These types of actions are referred to as operations.

Operational or transactional information systems (OLTP for OnLine Transactional Processing) are management or production systems that capture the company's activities in a more manageable and flexible IT environment. The characteristics of operational systems are:

- **Broad User Base**: These systems are intended for anyone involved in the company's daily activities. Decision-makers are excluded from this group as they operate at a higher level than daily management.

- **Extremely Fast** :They must be fast with reduced response times. Any downtime or blockage in OLTP systems leads to a halt in the respective activity.

- **Small Data Volume** :Management systems do not handle terabytes of data. These systems focus on current activities.

- **Transactional**: They operate using the transaction principle.

- **Read, Write, and Modify Data** :In an OLTP, information can be added, deleted if not useful for production, and modified.

The Decision-Making World:

It may seem obvious to define the decision-making world as the opposite of the operational world; however, the two worlds are complementary. While OLTP systems keep the company running, the decision-making world analyzes, predicts, and advises to better understand overall activities. The following characteristics are common to any decision-making product:

- **Small Number of Users**: Strategic decision support is intended for individuals typically at the top of the hierarchy.

- **General and Detailed Data**: Users' needs may vary regarding the level of detail required in the data.

- **Large Data Volumes**: Years of accumulated data generate gigabytes that must be managed by BI systems, from the smallest detail to the largest.

- **Non-Transactional**: BI does not follow a rigid process. Users must be able to start an analysis, backtrack, and begin another analysis.

- **Read-Only Data**: In the decision-making world, data loss should not occur. Data is never deleted; it is archived.

3. GLOBAL BI WORKFLOW ARCHITECTURE

The Global BI Workflow Architecture for COPAG is designed to streamline the process of collecting, analyzing, and presenting business data to support strategic decision-making. This architecture follows a structured approach based on OLAP (online analytical processing) and the Kimball methodology. The workflow is divided into three main phases: Collect, Analyze, and Dashboard.

The figure below illustrates all the phases :

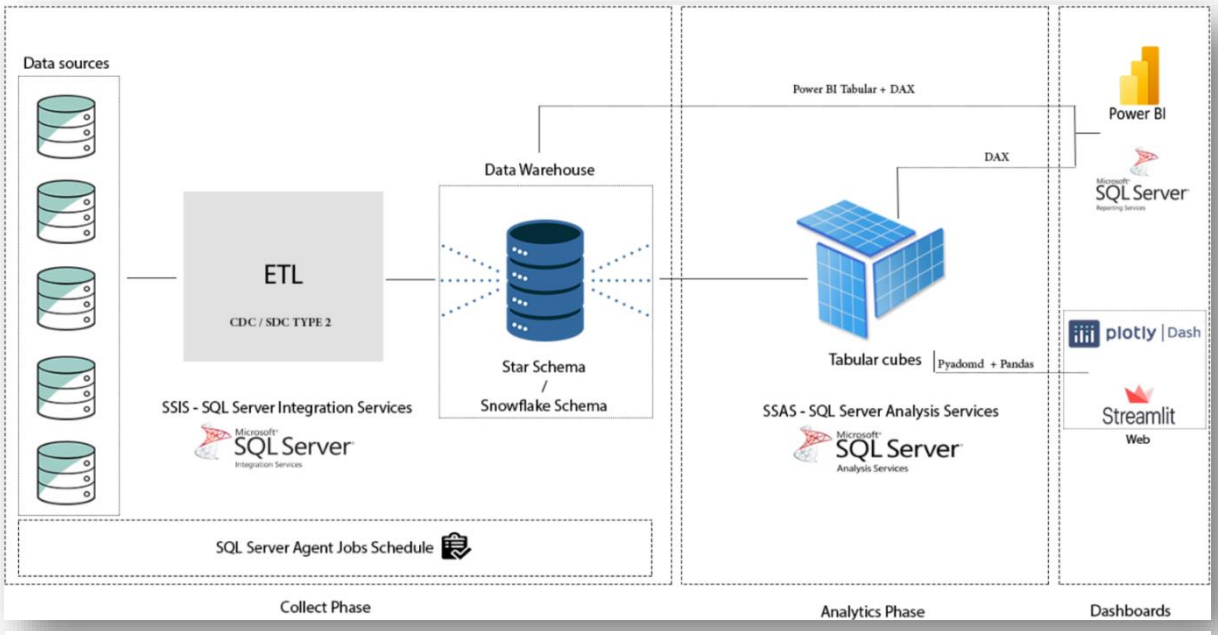


Figure 10 GLOBAL BI WORKFLOW ARCHITECTURE

3.1 Collect Phase:

The Collect Phase is the foundational stage where raw data from various sources is gathered and prepared for analysis. This phase includes several critical steps to ensure data accuracy and integrity.

ETL Steps:

- **Limit Incoming Data (CDC)**: Implement Change Data Capture (CDC) to extract and load only the new or changed records from the source systems. This reduces the load on the system and improves efficiency.
- **Generate History of Change Data (SCD Type 1/2)**: Use Slowly Changing Dimensions (SCD) Type 1 and Type 2 to preserve the history of changes in dimension tables, which is essential for historical analysis.
- **Transform & Process Dimension Tables**: Process dimension tables first to establish the context for the data. This involves cleaning, transforming, and loading dimension data.
- **Process Fact Tables**: After dimension tables, process the fact tables that contain transactional data. This step ensures that all relevant facts are correctly linked to the dimensions.
- **Load Data into Schemas**: Finally, load the processed data into the data warehouse using star or snowflake schemas. These schemas organize the data in a way that optimizes query performance and analysis.

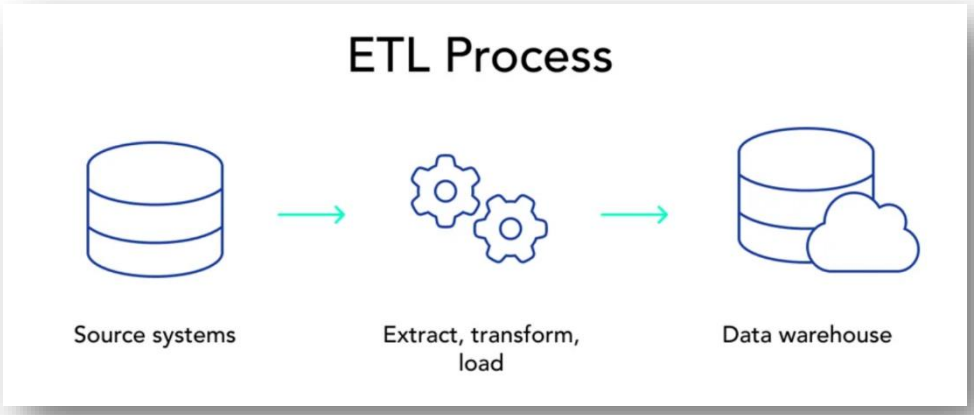


Figure 11 ETL Process

3.2Analyze Phase

The Analyze Phase involves processing the collected data to generate meaningful insights. This phase leverages advanced analytical tools and techniques to transform raw data into actionable intelligence.

Data Warehouse:

Star Schema: The Star Schema provides the simplest way of organizing data in a Data Warehouse. The center of the Star Schema can have one or more Fact Tables indexing a series

of Dimension Tables. Dimension Tables are used to describe Dimensions; they contain Dimension Keys, Values, and Attributes.

The goal of the Star Schema is to separate the numerical "FACT" data that relates to a business from the descriptive or "DIMENSIONAL" data.

Fact data can include information such as price, speed, weight, and quantities, that is, data presented in a numerical format. Dimensional Data can include uncountable things such as colors, geographical locations, salesperson names, employee names, etc., alongside the numerical information.

The Fact Data is organized inside the Fact Tables, while the Dimensional Data is inside the Dimension Tables. The Fact Tables form the integration points at the center of a Star in a Star Schema.

.

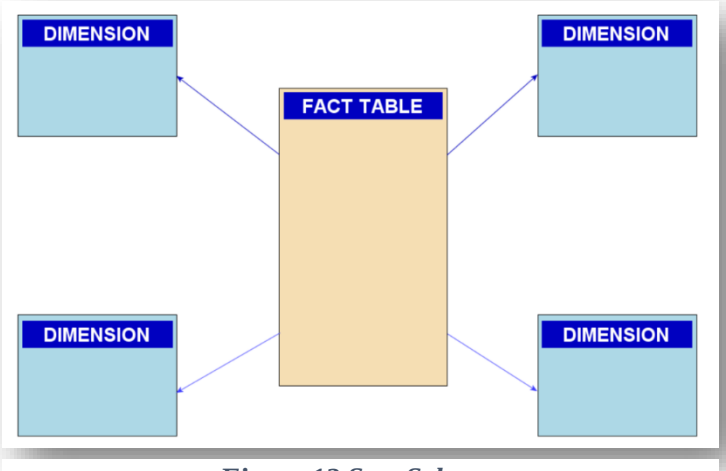


Figure 12 Star Schema

Snowflake Schema: A snowflake schema is very similar to the simple star schema above. The main difference is that snowflake schemas split dimensional tables into further dimensional tables (also called lookup tables).

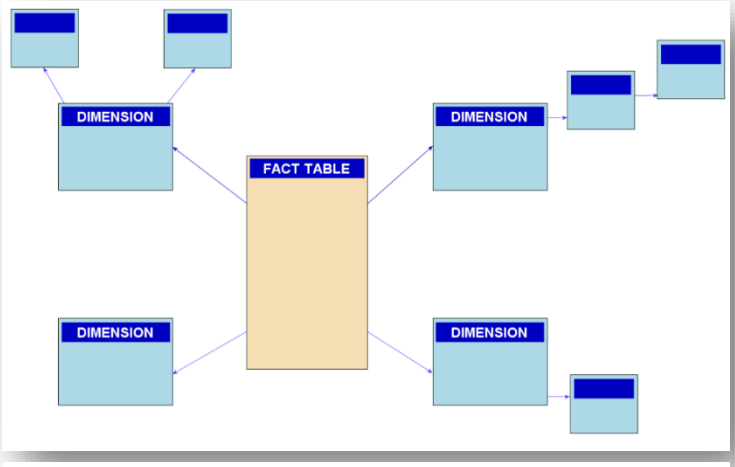


Figure 13 Snowflake Schema

Analytics Using SSAS:

- **Tabular Cubes**: Utilize SQL Server Analysis Services (SSAS) to create tabular cubes. These in-memory cubes use columnar storage, which enhances data compression and query speed. The tabular model is also easier to develop and manage compared to traditional OLAP cubes.

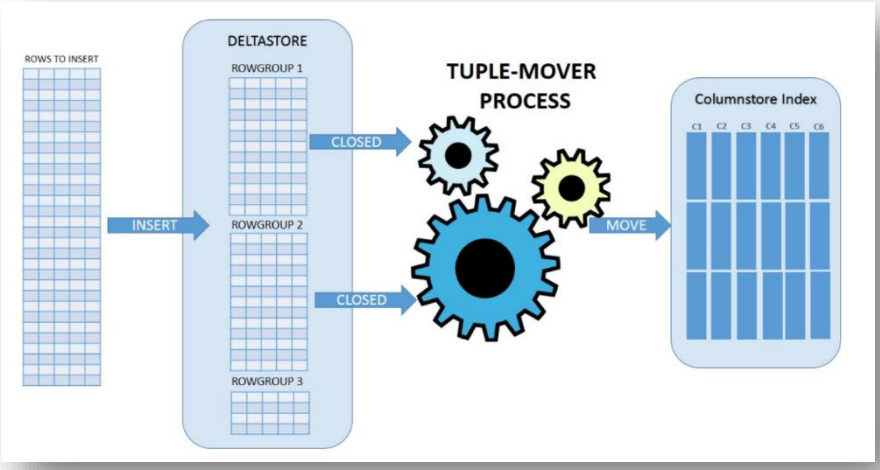


Figure 14 Tabular Cubes

3.3Dashboard Phase

The Dashboard Phase focuses on presenting the analyzed data through interactive and intuitive visualizations. This phase is crucial for enabling stakeholders to understand and act on the insights generated.

Power BI for Visualization:

- **Interactive Dashboards**: Use Power BI to create dynamic dashboards that provide real-time insights. These dashboards allow users to drill down into the data, explore different dimensions, and uncover trends and patterns.
- **Reports**: Generate detailed reports that summarize key metrics and performance indicators. These reports can be customized to meet the specific needs of different departments and stakeholders within COPAG.

User Training and Documentation:

- **Training Sessions**: Conduct training sessions for users to ensure they can effectively use the BI tools and understand the data presented in the dashboards and reports.
- **User Manuals and Guides**: Provide comprehensive documentation to help users navigate the BI system and leverage its full capabilities.

Conclusion

The Global BI Workflow Architecture for COPAG is a comprehensive framework designed to enhance data-driven decision-making. By following a structured approach to data collection, analysis, and presentation, this architecture ensures that COPAG can efficiently transform raw data into valuable insights. The use of advanced tools like SSIS for ETL processes and Power BI for visualization ensures that the BI system is robust, scalable, and user-friendly, ultimately supporting the company's strategic objectives and improving overall business performance.

Chapter 3: Tools Leveraged for the Project's Implementation

1 Introduction

The successful implementation of the COPAG Business Intelligence (BI) project relies on a set of robust tools that facilitate data integration, storage, analysis, and visualization. This chapter provides an overview of the primary tools leveraged for the project, detailing their functionalities and the reasons for their selection.

2.SQL Server Integration Services (SSIS)

2.1 Overview

SQL Server Integration Services (SSIS): is a powerful data integration tool provided by Microsoft. It is designed to solve complex business data integration problems by automating the process of extracting, transforming, and loading (ETL) data from various sources into a data warehouse.

2.2 Key Features

ETL Capabilities: SSIS offers comprehensive ETL functionalities, enabling efficient data extraction from multiple sources, data transformation, and loading into target systems.

- **Data Connectivity**: Supports a wide range of data sources including databases, flat files, Excel, and more.
- **Workflow Automation**: Allows automation of data workflows, reducing manual intervention and enhancing process efficiency.
- **Error Handling and Logging**: Provides robust error handling and logging features to ensure data integrity and facilitate troubleshooting.

2.3 Why SSIS?

- **Integration with SQL Server**: Seamless integration with Microsoft SQL Server ensures efficient data processing and management.
- **Scalability**: SSIS is designed to handle large volumes of data, making it suitable for the extensive data requirements of COPAG.
- **User-Friendly**: The drag-and-drop interface and pre-built components simplify the development of ETL processes.

2.4 SSIS in Action

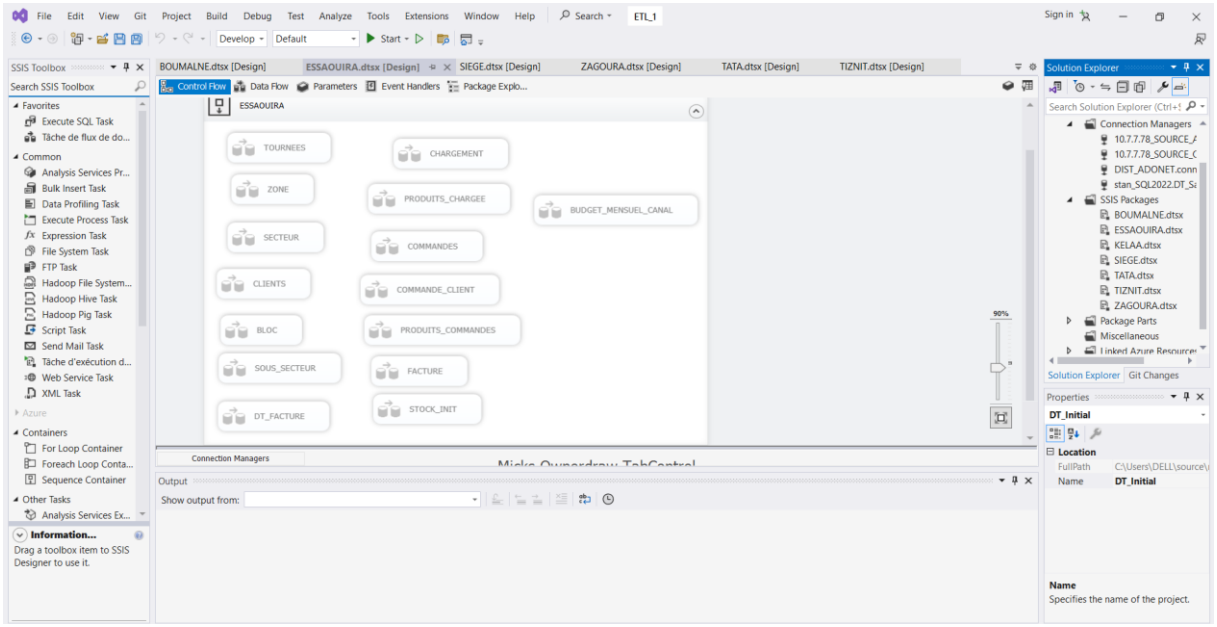


Figure 15 ETL package Essaouira

The figure illustrates the SQL Server Integration Services (SSIS) interface within Visual Studio, focusing on the design of the "ESSAOUIRA.dtsx" ETL package. This package orchestrates the extraction, transformation, and loading of data into a data lake. The main panel shows the Control Flow tab with a Sequence Container organizing tasks like loading data for "TOURNEES," "ZONE," "SECTEUR," and other entities.

The SSIS Toolbox on the left provides ETL components, facilitating easy drag-and-drop functionality to build workflows. The Properties panel on the right displays detailed information about the selected package or task, aiding in configuration.

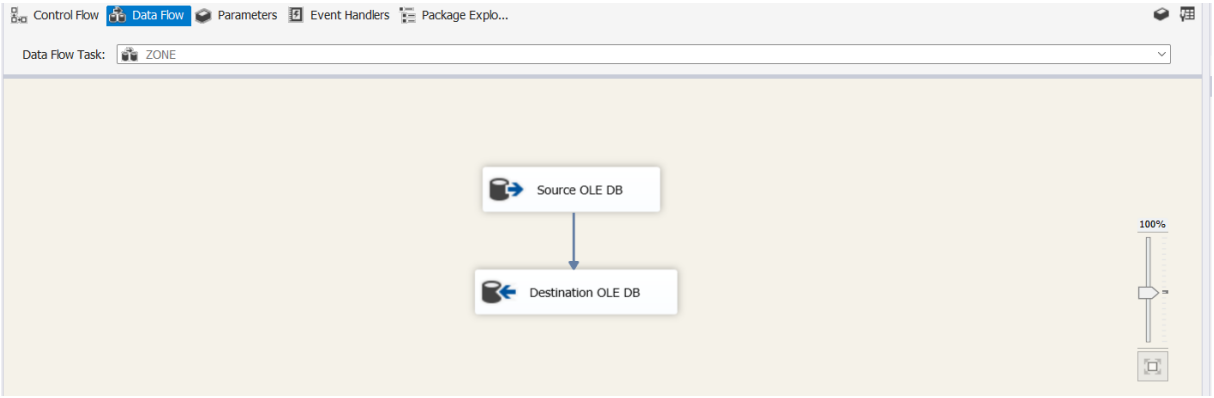


Figure 16 ETL package Essaouira

In this figure we focused on the "ESSAOUIRA.dtsx" ETL package. Specifically, it shows the Data Flow tab where data is transferred from a "Source OLE DB" to a "Destination OLE DB". This task, labeled "ZONE," is part of the ETL process designed to move and transform data into a data lake. The SSIS Toolbox on the left provides various ETL components, such as "Data Conversion" and "Aggregation," for building complex workflows. The Connection Managers panel at the bottom lists database connections, facilitating data extraction and loading. The Properties panel on the right displays detailed configurations for the selected task.

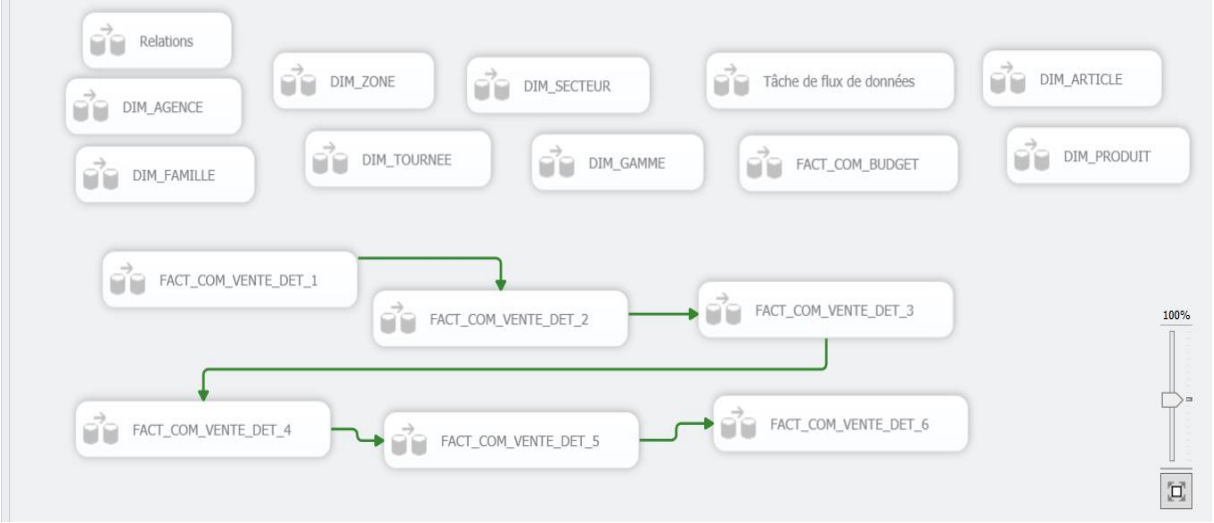


Figure 17 Table FACT_COM_VENTE_DET

This figure is highlighting the process of building a data warehouse . The Control Flow tab in the main panel includes tasks for loading and transforming data into the data warehouse, such as "DIM_AGENCE," "DIM_FAMILLE," "DIM_ZONE," and "DIM_SECTEUR" for dimension tables, and "FACT_COM_VENTE_DET_1" through "FACT_COM_VENTE_DET_6" for fact tables, with green arrows indicating the data flow between them.

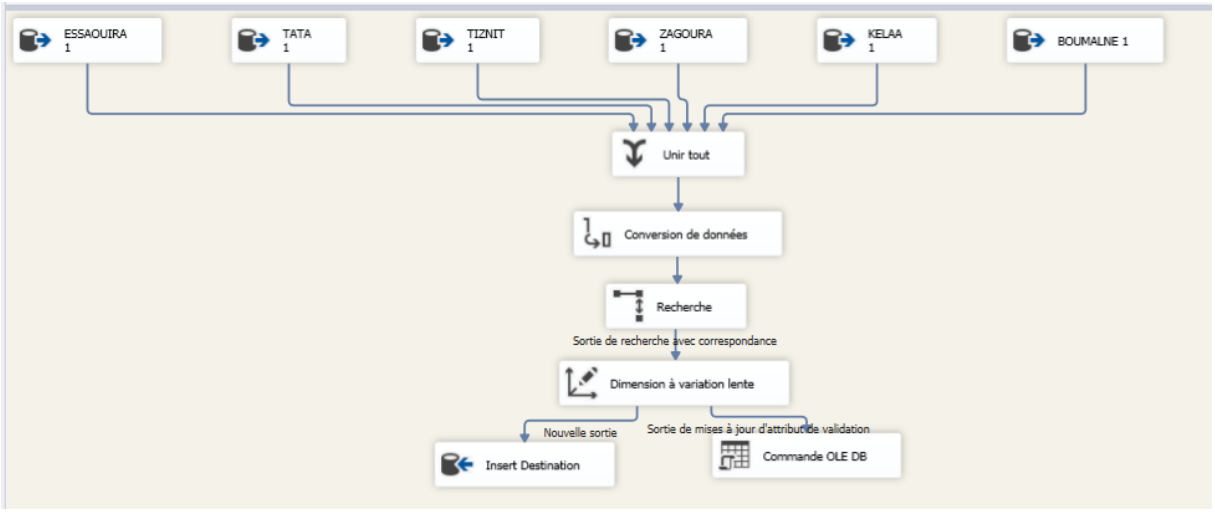


Figure 18 Table FACT_COM_VENTE_DET

focusing on the ETL package. The Data Flow tab shows the "DIM_TOURNEE" task, which consolidates and transforms data from multiple sources like ESSAOUIRA, TATA, TIZNIT, ZAGOURA, KELAA, and BOUMALNE.

The data flow includes merging sources with "Union All," performing "Data Conversion," validating with "Lookup," handling changes with "Slowly Changing Dimension," and finally inserting the data into the destination database.

3. Power BI

3.1 Overview

Power BI is a business analytics service by Microsoft that provides interactive visualizations and business intelligence capabilities with an interface simple enough for end users to create their own reports and dashboards.

3.2 Key Features

- **Interactive Dashboards**: Offers a wide range of visualization options to create interactive and real-time dashboards.
- **Data Connectivity**: Connects to a multitude of data sources, both on-premises and in the cloud.
- **Custom Visuals**: Allows the creation of custom visuals to meet specific business requirements.

- **DAX**: Utilizes Data Analysis Expressions (DAX) for creating sophisticated measures and calculations.

3.3 Why Power BI?

- **Ease of Use**: User-friendly interface that allows users at all technical levels to create and share reports and dashboards.
- **Real-Time Analytics**: Provides real-time data updates, enabling timely and informed decision-making.
- **Integration with Microsoft Ecosystem**: Seamless integration with other Microsoft products enhances data sharing and collaboration across the organization.

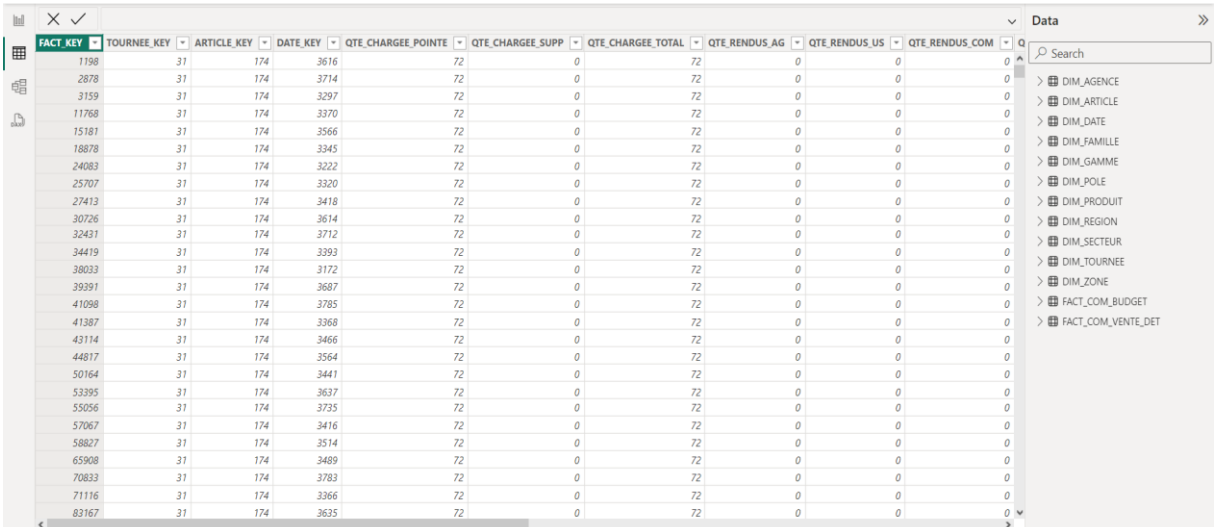


Figure 19 Power BI loaded data

This figure displays the Power BI interface, showcasing the data loaded from MSSQL to Power BI. The dataset includes columns such as "FACT_KEY," "TOURNEE_KEY," "ARTICLE_KEY," "DATE_KEY," and various quantity fields, part of the fact table "FACT_COM_VENTE_DET." The right panel lists different tables in the data model, including "DIM_AGENCE," "DIM_ARTICLE," "DIM_DATE," and "DIM_TOURNEE," essential for comprehensive data analysis.

3.4 Dax

DAX :stands for Data Analysis Expressions. It is a collection of functions, operators, and constants that can be used in a formula or expression to calculate and return one or more values. DAX helps create new information from data already in the model, which is essential for data analysis.

3.4. Key Concepts in DAX

- **Calculated Columns**: These are new columns added to a table in the data model. Calculated columns are computed using DAX formulas and can be used to create new data fields that do not exist in the original data source.
- **Measures**: Unlike calculated columns, measures are calculations used in data summarization and are dynamic based on the context of the data being analyzed. They are often used in aggregations such as sums, averages, and ratios.
- **Calculated Tables**: These are tables derived from DAX expressions. They are often used to create intermediate tables that are necessary for complex calculations.

3.5 DAX in Action

DAX is instrumental in transforming raw data into meaningful insights. For instance, in sales analysis, DAX can be used to calculate key performance indicators (KPIs) such as the following in the figure beside.

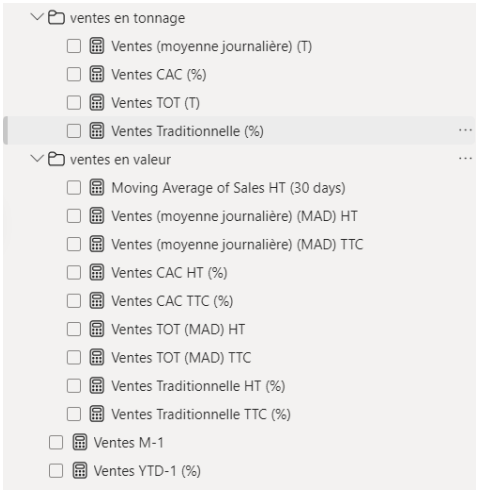


Figure 20 DAX measures

4. Microsoft SQL Server

4.1 Overview

Microsoft SQL Server is a relational database management system developed by Microsoft. It is used to store and retrieve data as requested by other software applications.

4.2Key Features

- **Data Storage**: Provides a reliable and scalable platform for storing large volumes of data.
- **Security**: Offers advanced security features to protect data integrity and confidentiality.
- **Performance**: High-performance database engine that supports fast data retrieval and processing.

4.3 Why Microsoft SQL Server?

- **Reliability**: Proven reliability and performance in handling large-scale data operations.
- **Scalability**: Scalable to meet the growing data needs of COPAG.
- **Integration**: Seamlessly integrates with SSIS, SSAS, and Power BI, providing a unified BI platform.

4.4 MSSM in Action

This figure displays the SQL Server Management Studio (SSMS) interface, highlighting the database structure on the server "10.7.7.78\ST_3". This server hosts the "OLTP_ITERN_COM_CP_T" database, which serves as a crucial source for extracting data into the data lake. The data from this server is meticulously integrated and processed to ensure that it is comprehensive, clean, and accurate.

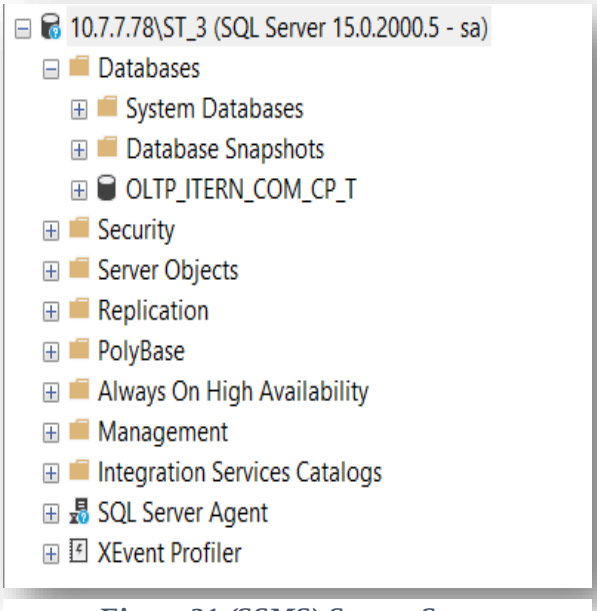


Figure 21 (SSMS) Source Server

The highlighted database, "WH_COPAG_COM", represents the data warehouse for the COPAG BI project. This data warehouse stores the integrated and processed data extracted from various sources, including the "DT_Sales_CP" database.

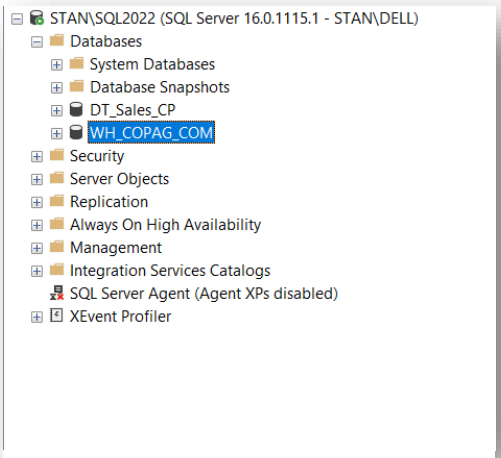


Figure 22 (SSMS) WearHouse

In our project, we used the Snowflake schema to structure the data warehouse. This normalized schema reduces data redundancy by breaking down dimension tables into smaller, related sub-tables. This approach facilitates data maintenance and enhances the efficiency of complex queries by minimizing the duplication of information. As a result, the Snowflake schema contributes to a more organized and efficient data warehouse.

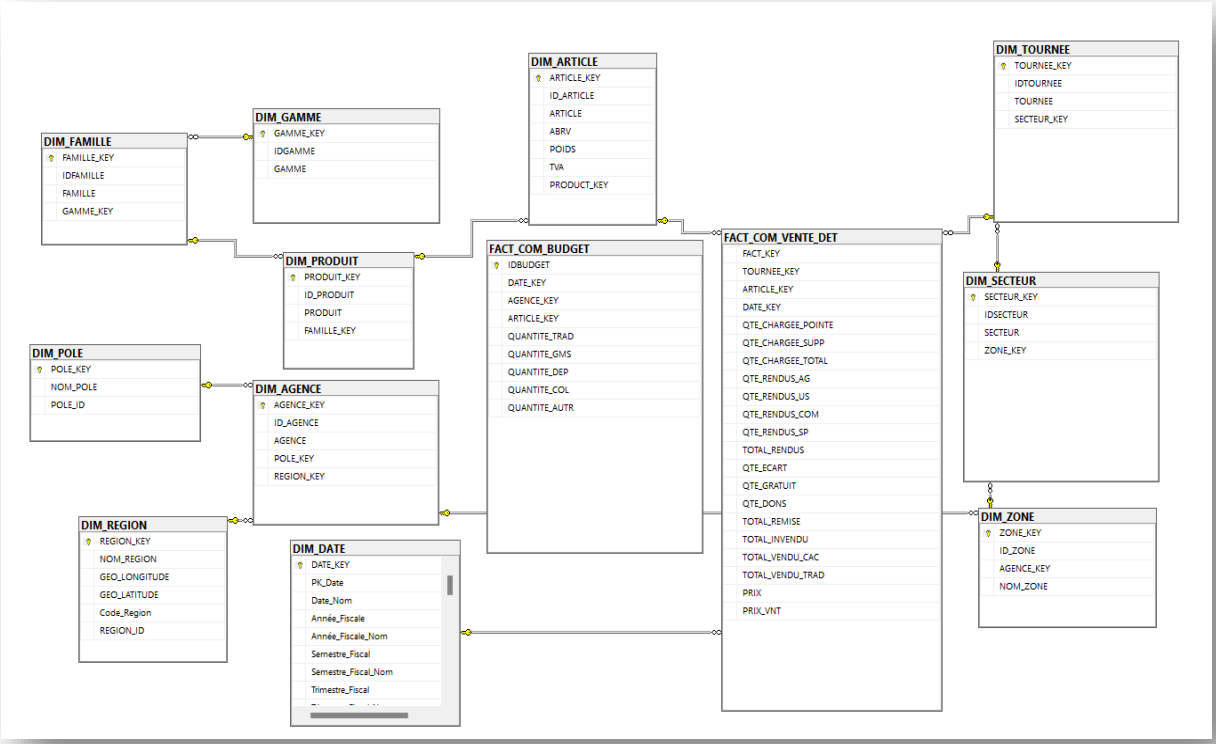


Figure 23 Data Wearhouse Structure

5. Python

5.1 Overview

Python is a versatile programming language widely used for data analysis, machine learning, and automation tasks . Python is utilized for data preprocessing, statistical analysis, and creating machine learning models, with a particular focus on forecasting.

5.2Key Features

- **Data Analysis Libraries**: Libraries such as Pandas, NumPy, and SciPy offer robust data manipulation and analysis capabilities.
- **Machine Learning**: Libraries like Scikit-Learn, TensorFlow, and PyTorch facilitate the development of machine learning models, including those used for forecasting.
- **Automation**: Python scripts can automate repetitive tasks, improving efficiency and consistency.
- **Integration**: Python can easily integrate with other tools and databases, enhancing its utility in the BI project.

5.3Why Python?

- **Versatility**: Python's extensive libraries and frameworks support a wide range of data analysis and machine learning tasks.
- **Ease of Use**: Python's simple syntax and readability make it accessible for both beginners and experienced programmers.
- **Community Support**: A large, active community contributes to continuous improvements and a wealth of resources.

6. Visual Studio

6.1 Overview

Visual Studio is an integrated development environment (IDE) from Microsoft. It is used for developing computer programs, websites, web apps, web services, and mobile apps.

6.2 Key Features

- **SSIS Development**: Supports the development of SSIS packages with a user-friendly interface and debugging tools.
- **Code Management**: Provides robust tools for code management and version control.

6.3 Why Visual Studio?

- **Development Efficiency**: Enhances development efficiency with integrated tools and features.
- **Debugging and Testing**: Advanced debugging and testing capabilities ensure the quality and reliability of BI solutions.
- **Integration with Microsoft BI Tools**: Seamlessly integrates with SSIS, SSAS, and SQL Server, providing a cohesive development environment.

Conclusion

The selection of these tools—SSIS, SSAS, Power BI, Microsoft SQL Server, Visual Studio, and Python—ensures a robust, scalable, and efficient BI system for COPAG. Each tool brings unique strengths to the project, collectively enhancing data integration, storage, analysis, and visualization capabilities. This integrated toolset supports COPAG's goal of optimizing sales strategies, improving decision-making, and responding proactively to market demands

CHAPTER 4: IMPLEMENTATION AND DEPLOYMENT

1 Introduction

After completing the requirements definition and the analysis and design phase, as well as specifying the technical development tools to be used, we will now embark on the implementation stage. This phase is a crucial part of the time dedicated to this project, involving the translation of functionalities defined during the design phases into . In this section, we will present the interfaces according to the types of users to give you an overview of the project.

2. Presentation of the Dashboards

2.1 Global Sales Analysis Interface:

Our interface includes sales dashboards, presenting global indicators such as the total pre-tax and after-tax sales, unsold items, losses, and total sales by region and range. A set of filters also allows for refining the data by region, month, and year, thus offering a detailed and customized analysis of business performance.

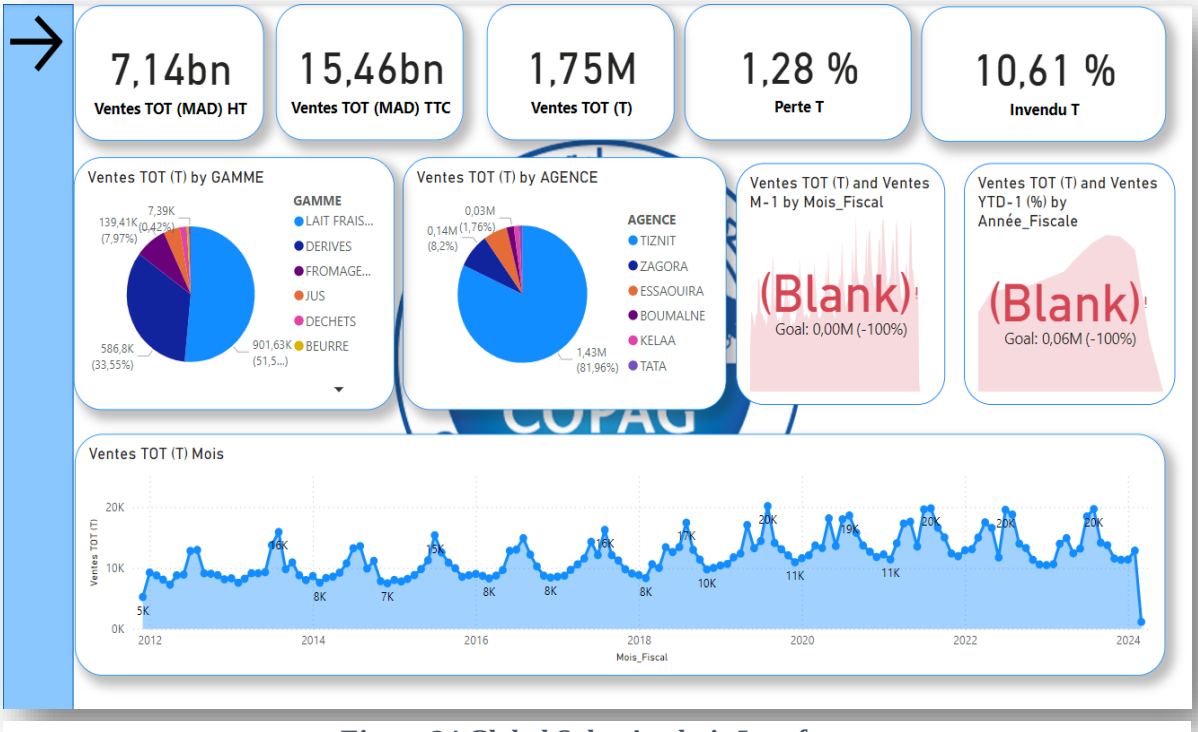


Figure 24 Global Sales Analysis Interface:

Our dashboard interface includes an advanced filter bar, allowing data to be filtered according to several criteria. Users can refine their analysis by selecting product ranges, regions, years, months, specific products, product families, individual items, and agencies. This feature provides maximum flexibility to explore the data from various perspectives and gain precise and relevant insights.

Figure 25 Filter Bar

This dashboard features maps that display total sales including and excluding tax, as well as average daily sales including and excluding tax. These indicators are presented by month and year, offering a detailed and chronological view of business performance. These visualizations allow for tracking sales trends and identifying periods of high or low activity.

Figure 26 Sales Values

This figure is showing total sales (in tons) over time, labeled "VENTES TOT (T) MOIS." The x-axis covers fiscal months from 2012 to 2024, while the y-axis represents total sales in tons. Key points include noticeable peaks in 2013, 2021, and 2023, indicating significant sales periods. This visualization helps identify trends, seasonal patterns, and fluctuations in sales, supporting data-driven decision-making.

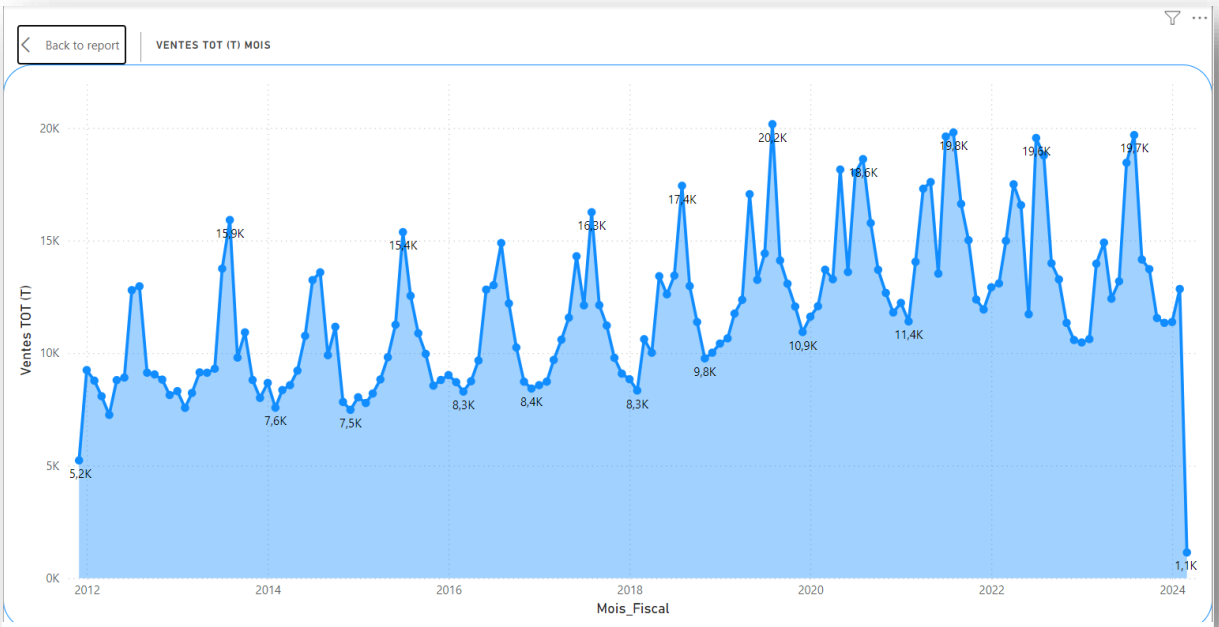


Figure 27 Sales Graph

This dashboard displays key performance indicators (KPIs) for comparing current annual sales with those of the previous year, as well as current monthly sales with those of the previous month. These KPIs provide users with a quick view of sales variations over time and enable them to identify trends of growth or decline.

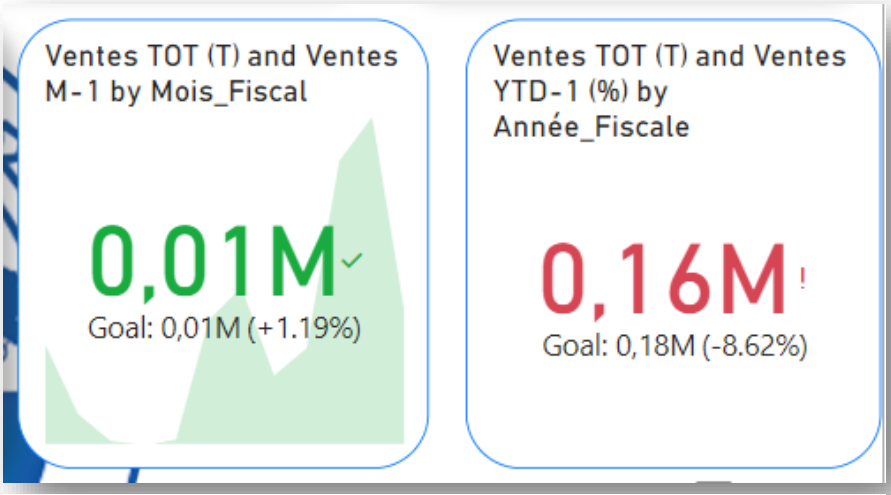


Figure 28 KPI's

2.2 Sales Performance Interface :

This interface displays sales details, including specific information on products, items, categories, families, and ranges. It allows users to examine sales performance in a granular manner, facilitating the identification of the highest performing products and the analysis of dynamics within each category and range.

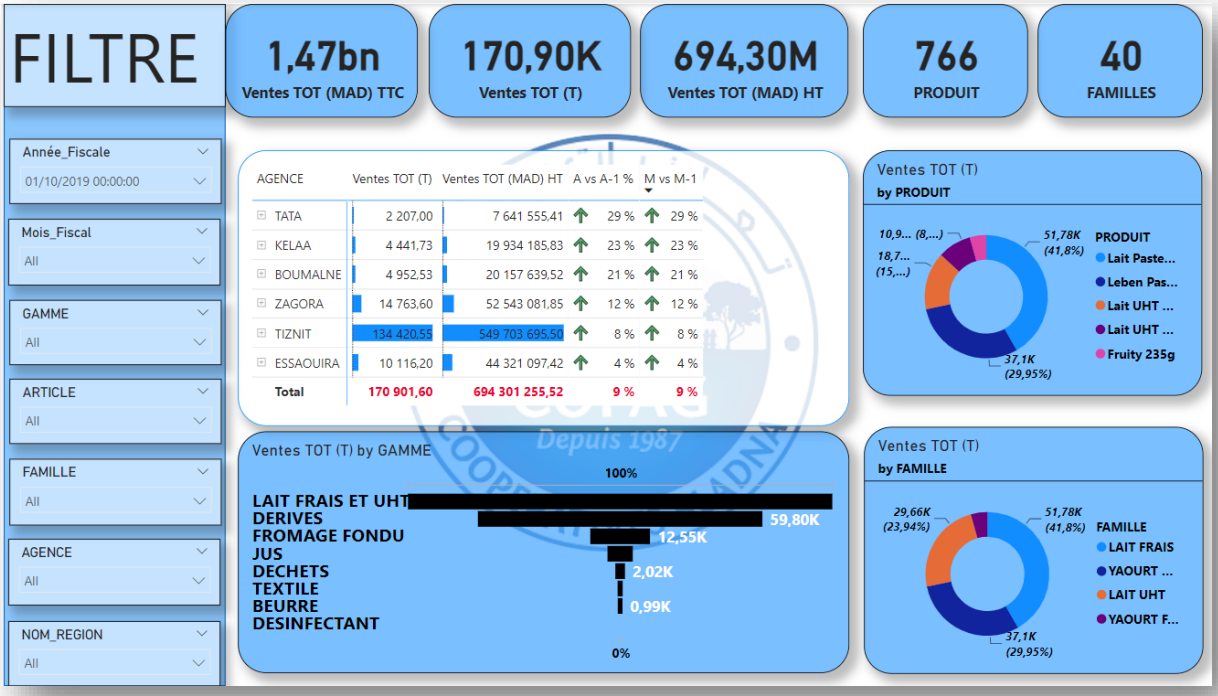


Figure 29 Sales Performance Interface

This dashboard features key performance indicators (KPIs) that display important metrics related to total sales and product information. It includes the total sales in Moroccan Dirhams, both with and without tax, as well as the total sales in tons. Additionally, the dashboard highlights the number of products and product families. These KPIs offer a quick overview of the sales performance and the diversity of product categories within the database, providing valuable insights for decision-making.

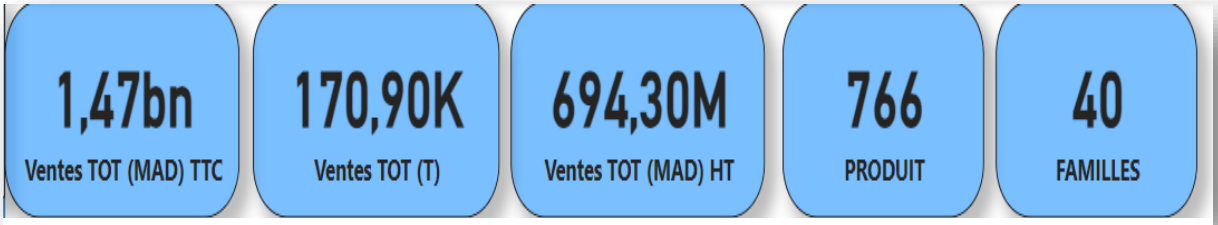


Figure 30 Sales Values

This dashboard includes a donut chart showing total sales by family. This visual representation makes it easy to understand the distribution of sales among different product

families, thus providing a clear and intuitive view of each family's respective contributions to total sales.

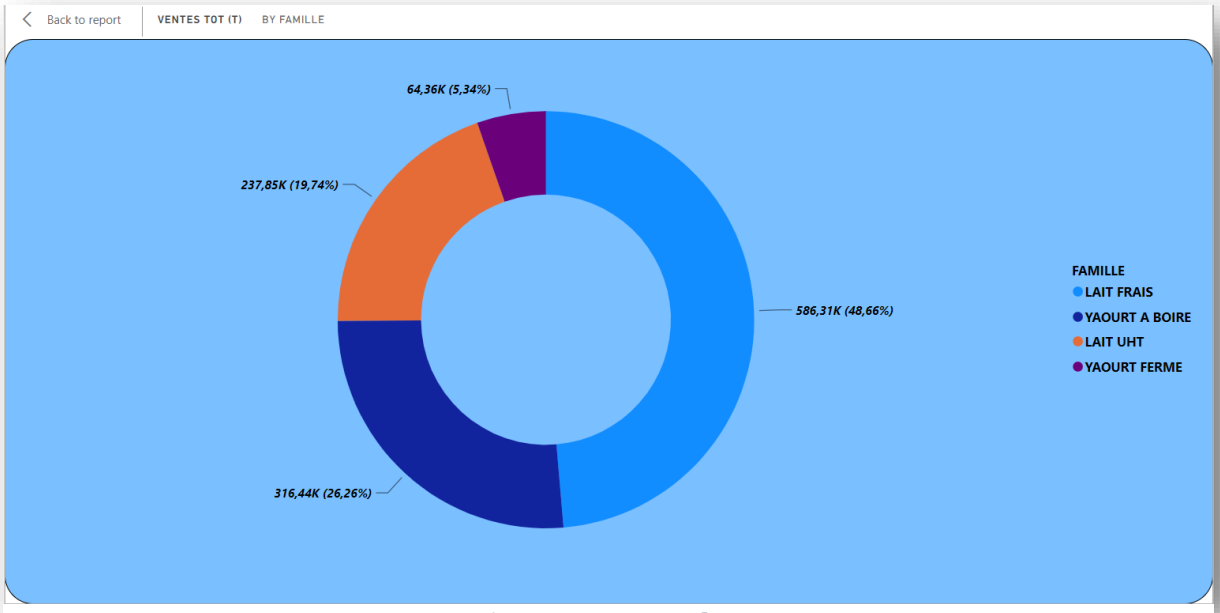


Figure 31 Donut Chart

This dashboard includes a horizontal bar chart showing total sales by product range. This visual representation allows for an easy understanding of the distribution of sales across different product ranges. "Fresh and UHT Milk" leads the sales, followed by "Derivatives" and "Melted Cheese". Other categories like "Juice", "Waste", "Butter", "Textiles", "Meat Derivatives", and "Disinfectants" complete the sales distribution. This chart provides a clear and intuitive view of each product range's contribution to total sales, facilitating performance analysis by product category.

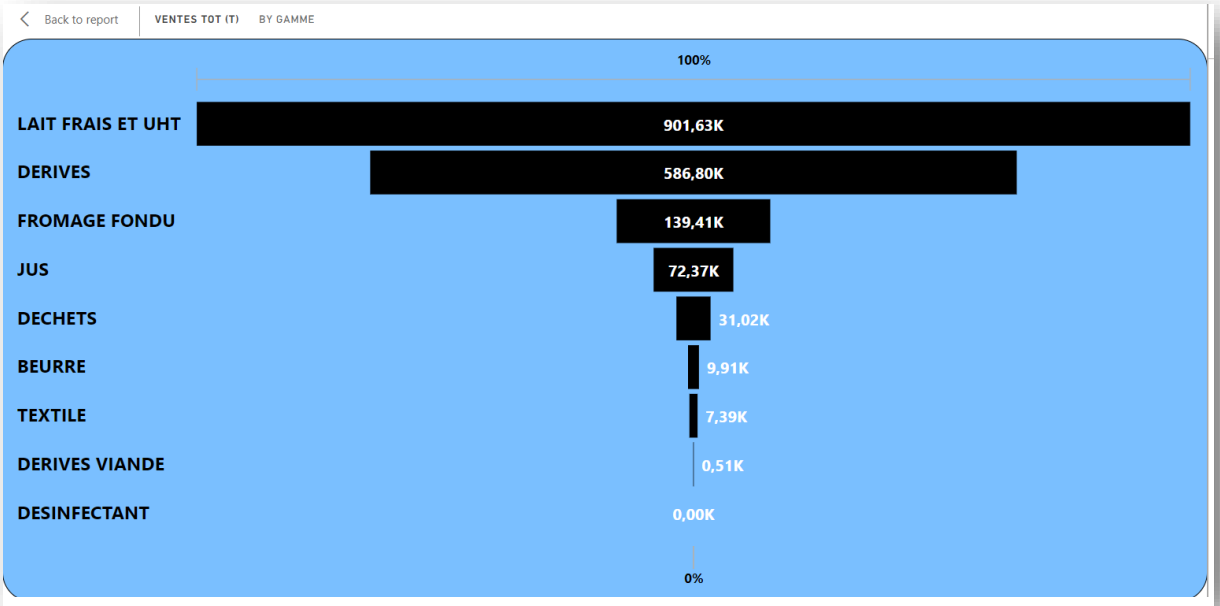


Figure 32 Horizontal Bar Chart

This dashboard includes a matrix table detailing the total sales by product category and location. Each product range is listed with total sales in tons and total sales in Moroccan Dirhams (excluding tax). Additionally, it displays the percentage change in sales compared to the same period last year and the previous month. This visualization allows users to examine the performance of each product and location in detail, while analyzing sales trends over time.

AGENCE	GAMME	Ventes TOT (T)	Ventes TOT (MAD) HT	A vs A-1 %	M vs M-1
KELAA	DECHETS	142,28	56 309,81	24 %	24 %
	JUS	488,92	152 998,03	16 %	16 %
	LAIT FRAIS ET UHT	2 951,91	23 982 068,22	7 %	7 %
	DERIVES VIANDE	5,50	2 882,67	4 %	4 %
	TEXTILE	59,76	58 504,44	4 %	4 %
	BEURRE	7,37	650 258,97	4 %	4 %
	FROMAGE FONDU	497,80	426 831,68	-3 %	-3 %
	DERIVES	1 033,35	700 752,37	-18 %	-18 %
	Total	**5 186,89**	**26 030 606,19**	**2 %**	**2 %**
BOUMALNE		**6 088,06**	**26 651 868,77**	**-3 %**	**-3 %**
ZAGORA		**15 587,85**	**61 168 404,81**	**-3 %**	**-3 %**
TATA		**2 446,28**	**7 733 707,21**	**-6 %**	**-6 %**
ESSAOUIRA		**10 223,81**	**50 209 804,62**	**-7 %**	**-7 %**
TIZNIT		**123 517,37**	**565 137 349,09**	**-11 %**	**-11 %**
Total		**163 050,25**	**736 931 740,68**	**-9 %**	**-9 %**

Figure 33 Matrix

2.3 Regional Performance Interface:

This dashboard includes a matrix table detailing the total sales by region and product range. It shows total sales in tons and Moroccan Dirhams (excluding tax) for each product range within each region. Additionally, it displays the percentage change in sales compared to the same period last year and the previous month. This visualization allows users to analyze sales performance by region and product range, offering a detailed and comparative view of how each area and product category is performing over time.

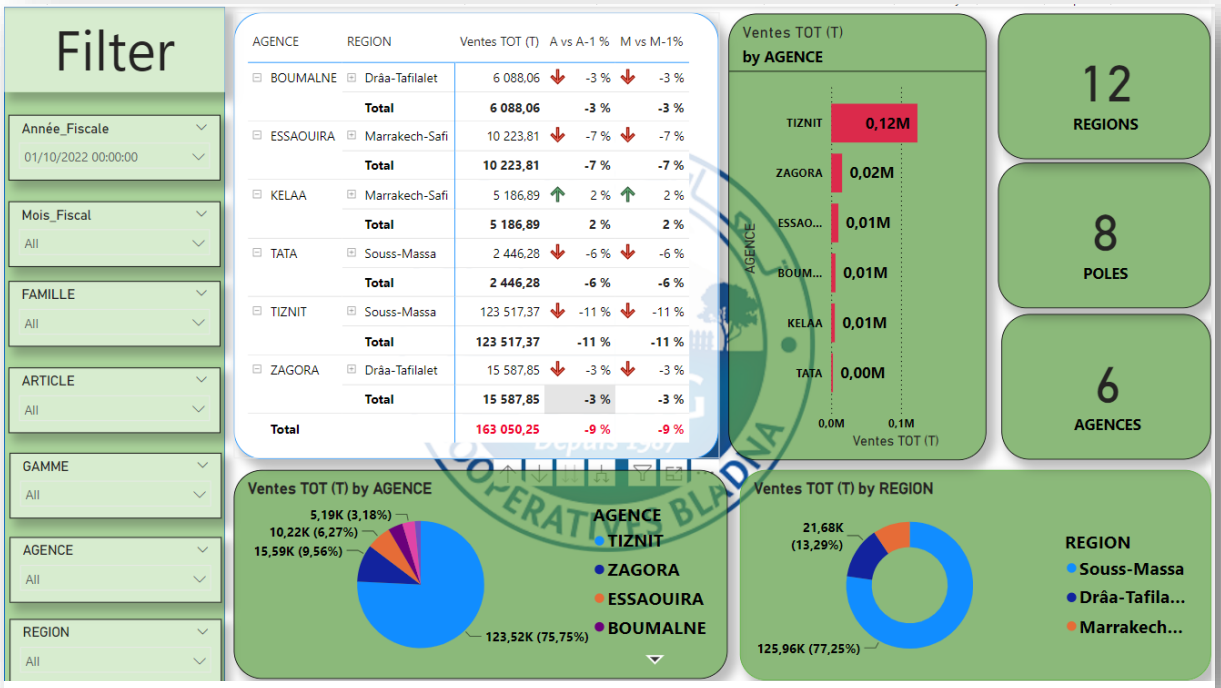
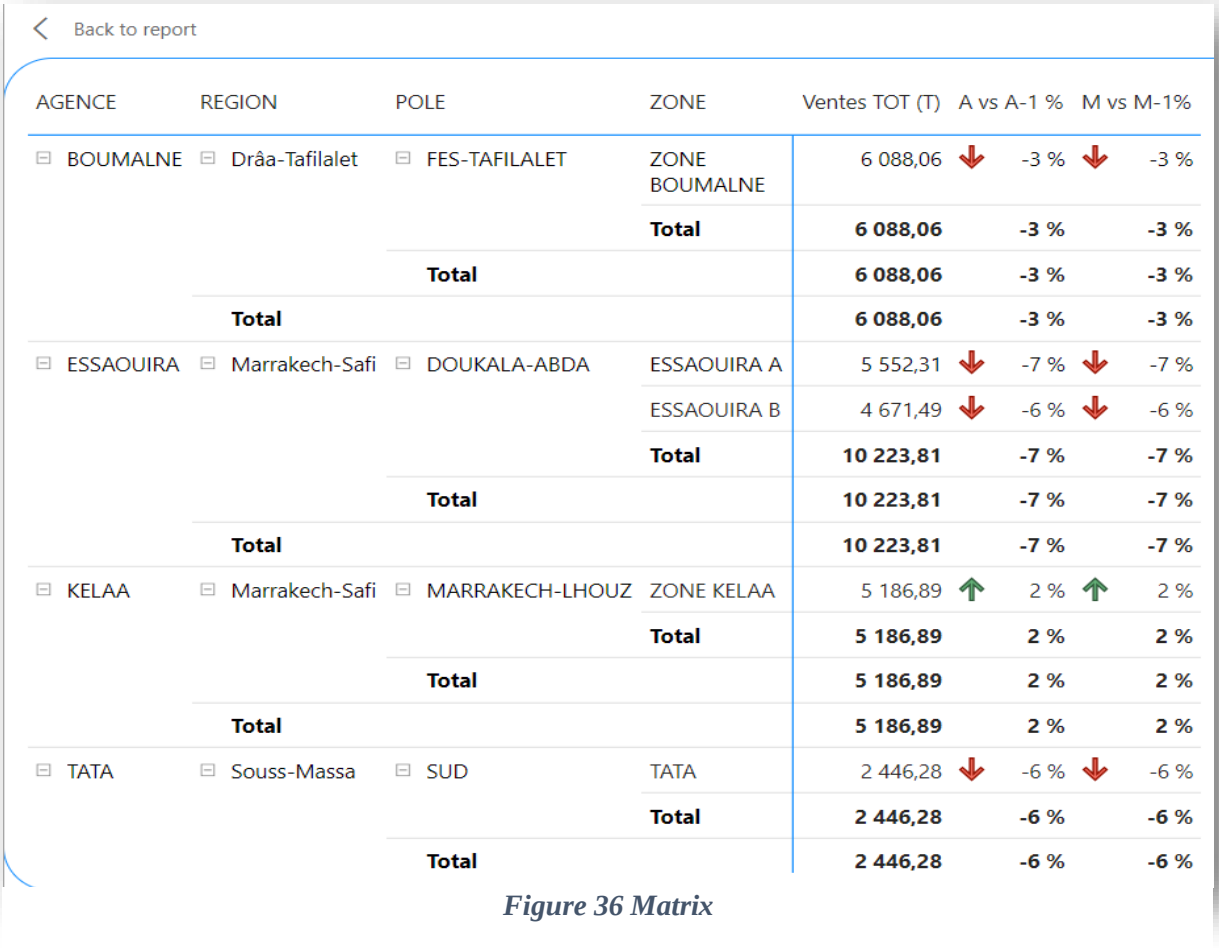


Figure 34 Regional Performance Interface

This dashboard section highlights key organizational metrics. It shows the total number of regions, poles, and agencies. Specifically, there are 12 regions, 8 poles, and 6 agencies. This visualization provides a quick overview of the organizational structure, facilitating an understanding of the distribution and reach of the organization's operations.

Figure 35 Organizational Metrics

This dashboard includes a matrix table detailing total sales by agency, region, pole, and zone. Each row provides the total sales in tons for different organizational units. It also shows the percentage change in sales compared to the same period last year (A vs A-1%) and the previous month (M vs M-1%). This comprehensive view allows users to analyze sales performance across various geographic and organizational divisions, facilitating a detailed understanding of how each unit is performing over time.



AGENCE	REGION	POLE	ZONE	Ventes TOT (T)	A vs A-1 %	M vs M-1%
BOUMALNE	Drâa-Tafilalet	FES-TAFILALET	ZONE BOUMALNE	6 088,06	-3 %	-3 %
			Total	**6 088,06**	**-3 %**	**-3 %**
		Total		**6 088,06**	**-3 %**	**-3 %**
	Total			**6 088,06**	**-3 %**	**-3 %**
ESSAOUIRA	Marrakech-Safi	DOUKALA-ABDA	ESSAOUIRA A	5 552,31	-7 %	-7 %
			ESSAOUIRA B	4 671,49	-6 %	-6 %
			Total	**10 223,81**	**-7 %**	**-7 %**
		Total		**10 223,81**	**-7 %**	**-7 %**
	Total			**10 223,81**	**-7 %**	**-7 %**
KELAA	Marrakech-Safi	MARRAKECH-LHOUZ	ZONE KELAA	5 186,89	2 %	2 %
			Total	**5 186,89**	**2 %**	**2 %**
		Total		**5 186,89**	**2 %**	**2 %**
	Total			**5 186,89**	**2 %**	**2 %**
TATA	Souss-Massa	SUD	TATA	2 446,28	-6 %	-6 %
			Total	**2 446,28**	**-6 %**	**-6 %**
		Total		**2 446,28**	**-6 %**	**-6 %**

Figure 36 Matrix

This dashboard includes a horizontal bar chart showing total sales in tons by agency. The chart provides a visual comparison of sales performance across different agencies. Tiznit leads with the highest sales, followed by Zagora, Essaouira, Boumalne, Kelaa, and Tata. This visualization allows users to quickly assess and compare the sales volumes of each agency, facilitating a clear understanding of regional sales distribution.

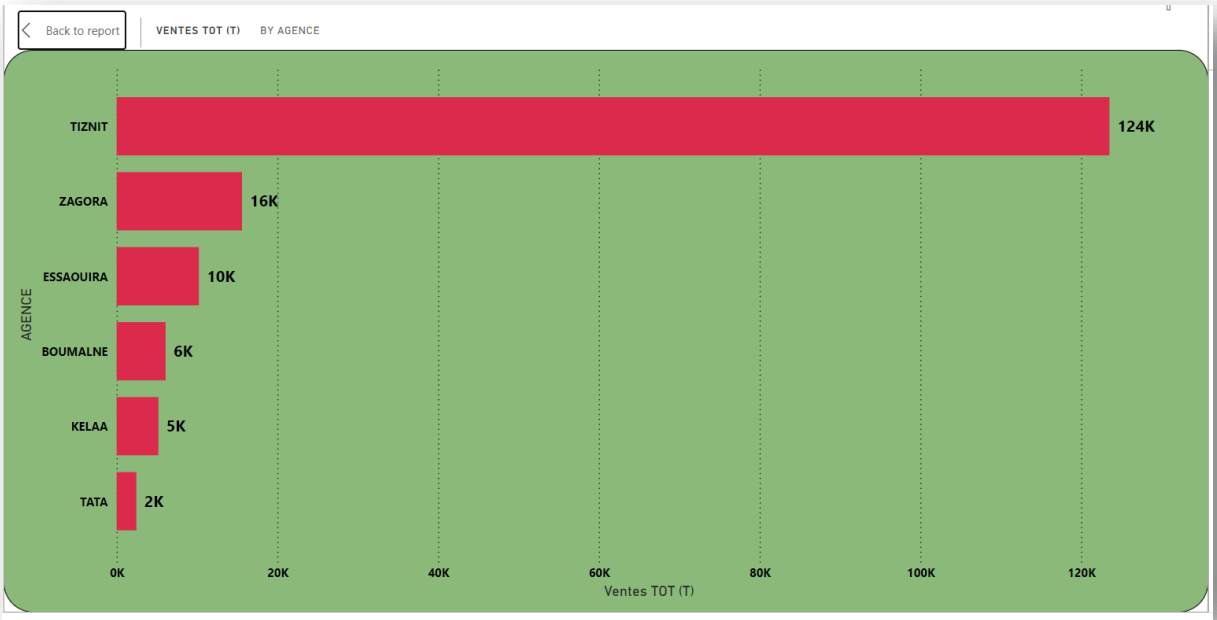


Figure 37 Horizontal Bar Chart

2.4 losses and unsold items

This dashboard presents a comprehensive view of sales and inventory metrics by agency. Key performance indicators (KPIs) at the top display overall metrics such as percentage of loss, unsold inventory, total sales excluding tax, total sales including tax, and total sales in tons. The matrix table in the center details unsold inventory and loss percentages by agency, showing month-over-month and year-over-year comparisons. Pie charts on the right visualize the distribution of unsold inventory and losses across different agencies. Filters on the left allow users to adjust the data view by fiscal year, month, product family, article, product range, agency, and region.

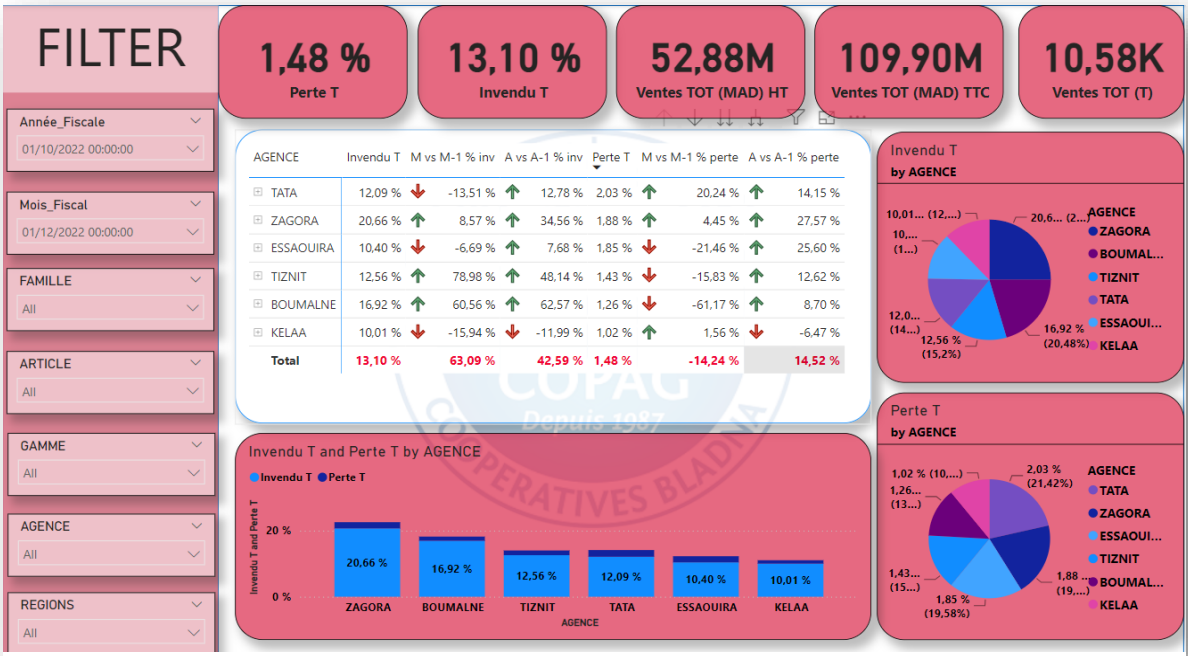


Figure 38 Losses and Unsold Items Interface

This dashboard features a pie chart displaying the unsold inventory percentage (Invendu T) by agency. Each segment of the pie chart represents an agency, with the size of each segment corresponding to the percentage of unsold inventory. This visual allows for an easy comparison of unsold inventory across different agencies. The chart shows that Zagora has the highest percentage of unsold inventory, followed by Boumalne, Tiznit, Tata, Essaouira, and Kelaa.

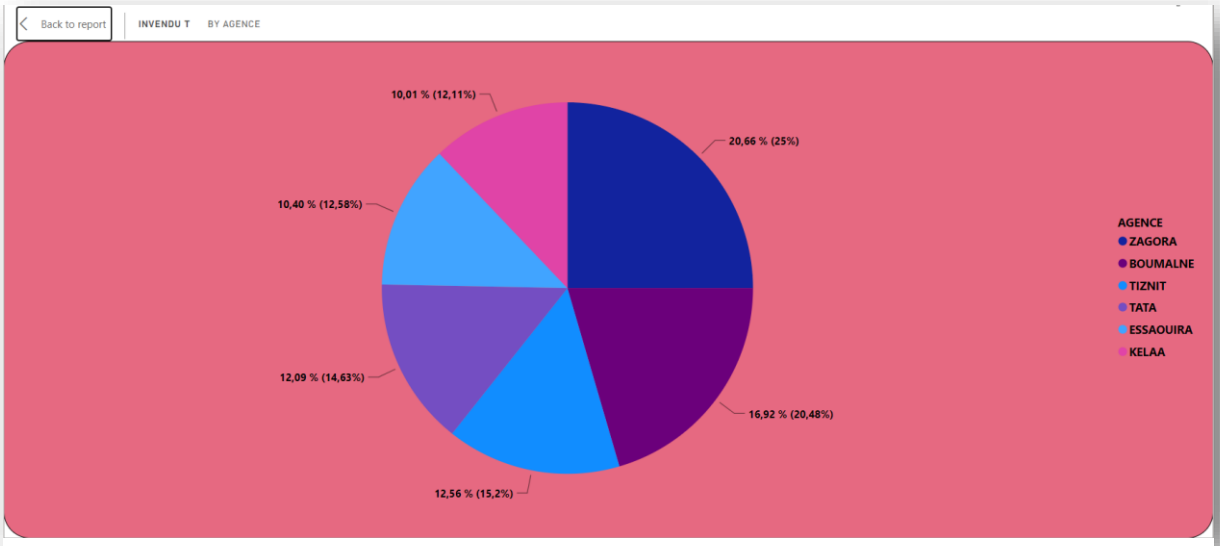


Figure 39 Pie Chart

This dashboard features a stacked bar chart comparing unsold inventory (Invendu T) and loss (Perte T) percentages by agency. Each bar represents an agency, with the blue section indicating the percentage of unsold inventory and the dark blue section representing the percentage of loss. The chart shows that Zagora has the highest unsold inventory percentage, followed by Boumalne and Tiznit. Tata, Essaouira, and Kelaa have lower percentages.

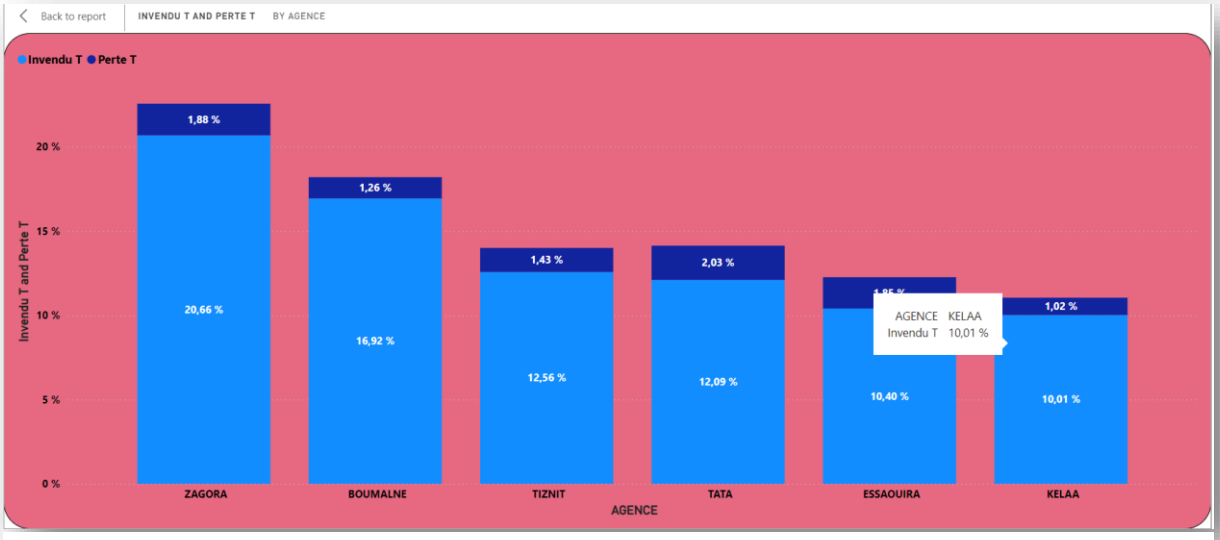
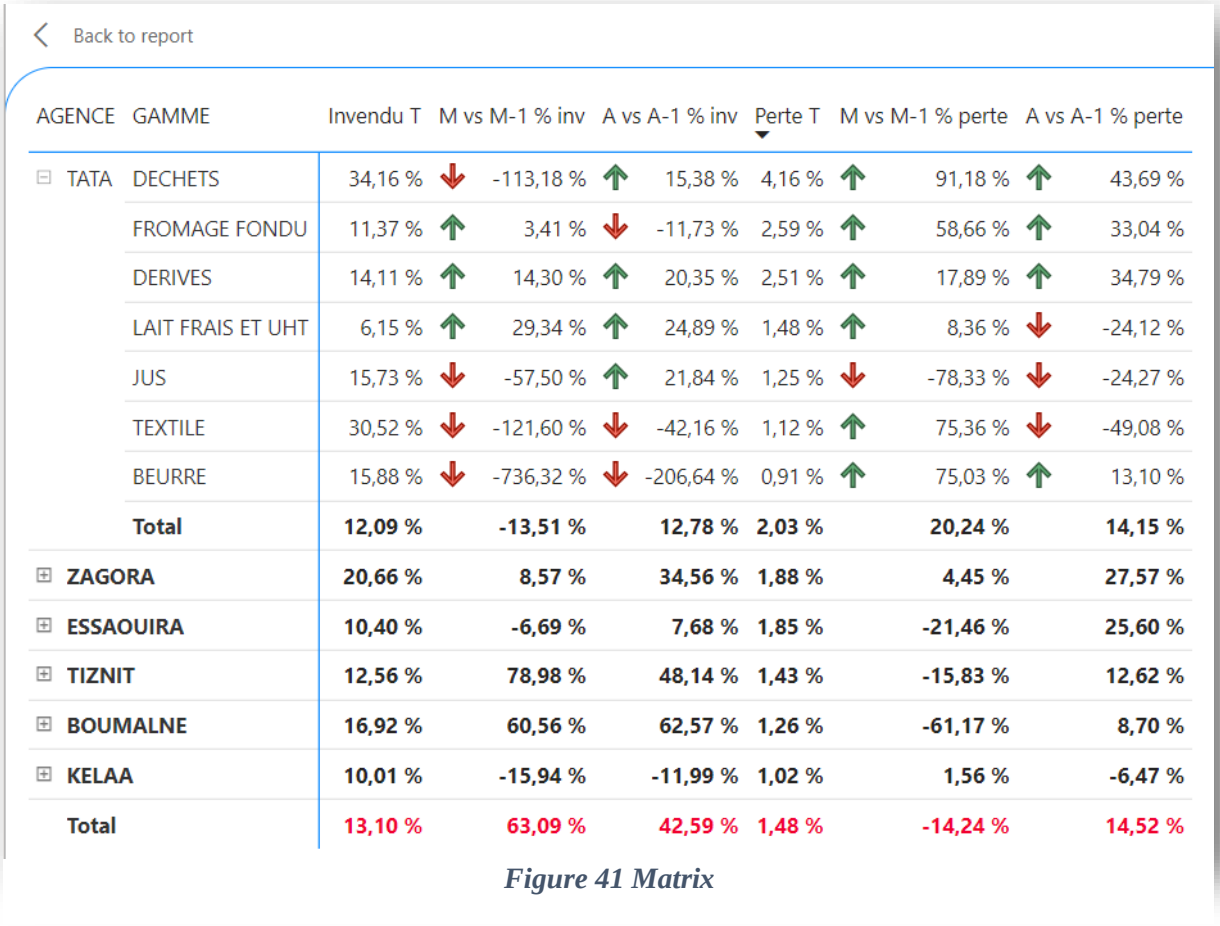


Figure 40 Stacked Bar Chart

This dashboard presents a comprehensive matrix table detailing the percentages of unsold inventory (Invendu T) and loss (Perte T) by agency and product range. The table includes comparisons of month-over-month (M vs M-1) and year-over-year (A vs A-1) changes for both unsold inventory and losses. Key metrics highlighted are the percentage of inventory that remains unsold and the percentage of inventory lost or wasted. This visualization enables a detailed analysis of the performance of each product range within different agencies, providing valuable insights into inventory management. It helps identify trends and areas requiring improvement, facilitating data-driven decisions to reduce unsold inventory and losses.



AGENCE	GAMME	Invendu T	M vs M-1 % inv	A vs A-1 % inv	Perte T	M vs M-1 % perte	A vs A-1 % perte
TATA	DECHETS	34,16 %	-113,18 %	15,38 %	4,16 %	91,18 %	43,69 %
	FROMAGE FONDU	11,37 %	3,41 %	-11,73 %	2,59 %	58,66 %	33,04 %
	DERIVES	14,11 %	14,30 %	20,35 %	2,51 %	17,89 %	34,79 %
	LAIT FRAIS ET UHT	6,15 %	29,34 %	24,89 %	1,48 %	8,36 %	-24,12 %
	JUS	15,73 %	-57,50 %	21,84 %	1,25 %	-78,33 %	-24,27 %
	TEXTILE	30,52 %	-121,60 %	-42,16 %	1,12 %	75,36 %	-49,08 %
	BEURRE	15,88 %	-736,32 %	-206,64 %	0,91 %	75,03 %	13,10 %
	Total	**12,09 %**	**-13,51 %**	**12,78 %**	**2,03 %**	**20,24 %**	**14,15 %**
ZAGORA		**20,66 %**	**8,57 %**	**34,56 %**	**1,88 %**	**4,45 %**	**27,57 %**
ESSAOUIRA		**10,40 %**	**-6,69 %**	**7,68 %**	**1,85 %**	**-21,46 %**	**25,60 %**
TIZNIT		**12,56 %**	**78,98 %**	**48,14 %**	**1,43 %**	**-15,83 %**	**12,62 %**
BOUMALNE		**16,92 %**	**60,56 %**	**62,57 %**	**1,26 %**	**-61,17 %**	**8,70 %**
KELAA		**10,01 %**	**-15,94 %**	**-11,99 %**	**1,02 %**	**1,56 %**	**-6,47 %**
Total		**13,10 %**	**63,09 %**	**42,59 %**	**1,48 %**	**-14,24 %**	**14,52 %**

Figure 41 Matrix

3 Forecasting:

Before forecasting with our data, we will use a set of models to evaluate which one offers the best performance. This step involves testing the performance of each model on our historical data to determine the most accurate and reliable for forecasting. For this, we will use only the data prior to October 1, 2021, for training. Then, we will test our models over the period from October 1, 2021, to October 1, 2022. By comparing the results obtained by each model during this testing phase, we can identify the most effective model. This model will

then be used to make forecasts for the period from October 1, 2022, to October 1, 2023, ensuring more robust and relevant projections.

3.1 Time series models:

Time series models are techniques used to analyze data collected at regular time intervals and to forecast future values. They account for trends, seasons, and cycles in the data, such as predicting monthly sales or daily temperatures.

ARIMA:

A statistical model used to analyze and forecast time series by capturing dependencies among past values and forecast errors.

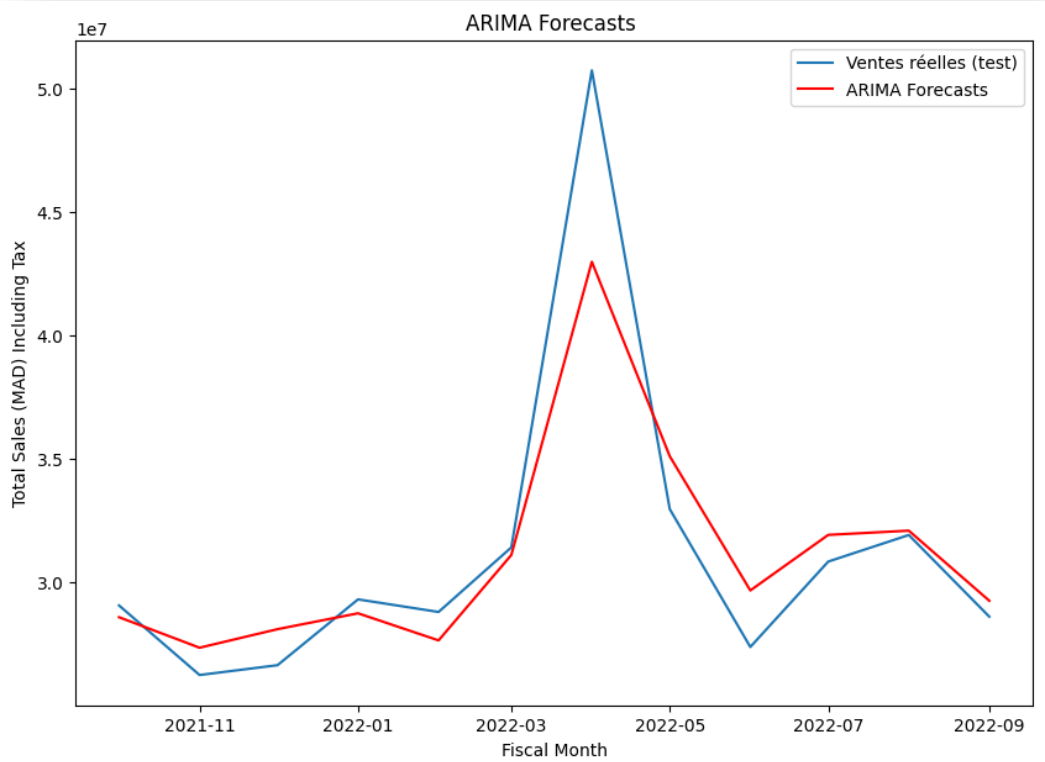


Figure 42 ARIMA Forecast test

In this graph displays ARIMA forecasts for total sales (in MAD including tax) alongside the actual sales for a test period. The x-axis represents the fiscal months from November 2021 to September 2022, while the y-axis indicates total sales in millions of Moroccan Dirhams. The blue line depicts the actual sales, while the red line shows the ARIMA forecasted sales. The comparison reveals how closely the ARIMA model's predictions align with the real sales data, with noticeable peaks and troughs indicating periods of high and low sales, respectively. This

visualization aids in evaluating the accuracy of the forecasting model and its effectiveness in predicting future sales trends.

SARIMAX:

SARIMAX is a statistical model for time series forecasting that includes seasonality and external factors. It extends ARIMA by adding seasonal components and exogenous variables.

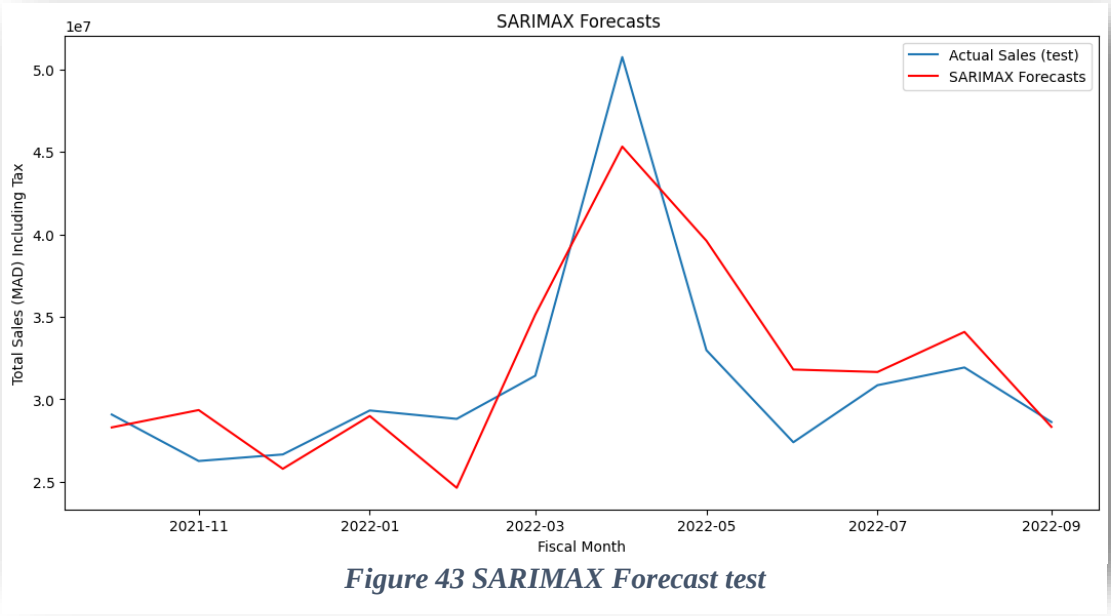


Figure 43 SARIMAX Forecast test

The SARIMAX model shows a similar overall trend and seasonality as the ARIMA model. However, it appears to align more closely with the actual sales during the peak and trough periods, particularly capturing the sharp increase and subsequent decrease around early 2022 more accurately.

Prophet:

A forecasting model developed by Facebook, designed to be flexible and easy to use. It is particularly effective for time series with seasonal trends and change points.

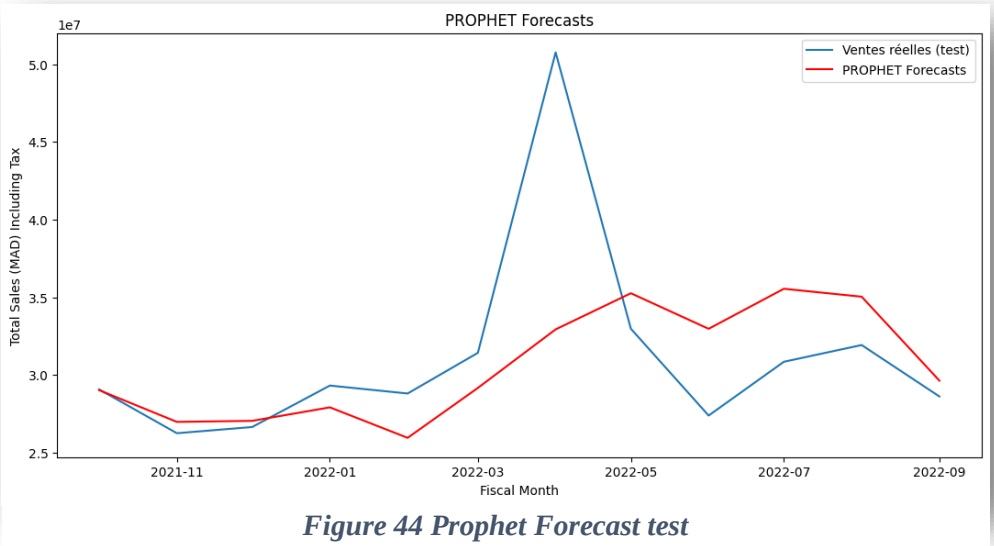


Figure 44 Prophet Forecast test

The Prophet model captures the general trend and seasonality. However, it underestimates the sharp increase in sales and slightly overestimates during some periods of decline.

Exponential Smoothing:

A forecasting technique that uses a weighted average of past values, where the weights decrease exponentially, giving more importance to recent observations to predict future values.

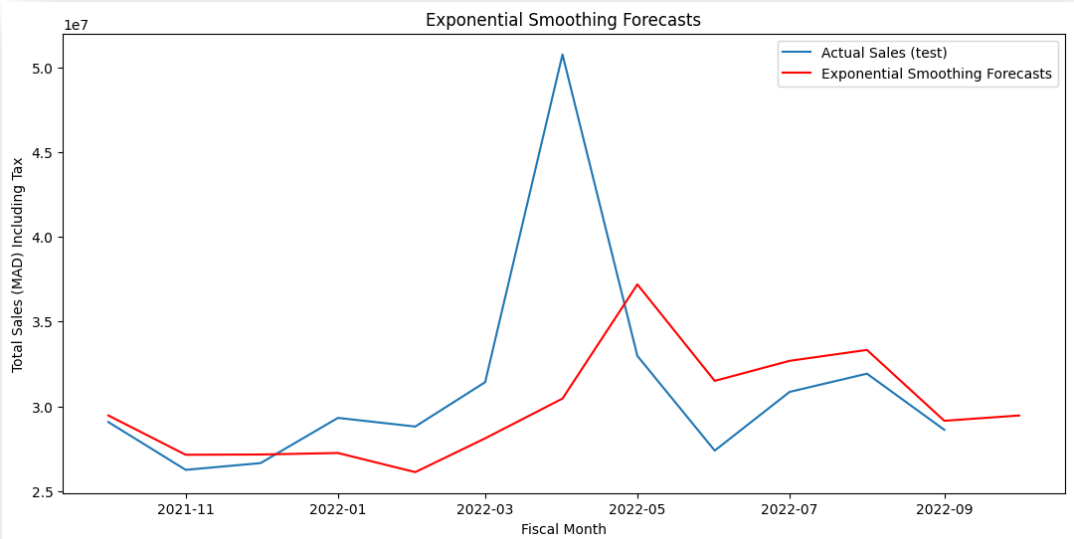


Figure 45 Exponential Smoothing Forecast test

Exponential Smoothing captures the general trend but significantly underestimates the peaks and overestimates the troughs. The model provides a smoother forecast compared to the other models

Model Comparison

Modele	RMSE
ARIMA	2531555.62
SARIMAX	3423079.88
PROPHET	5791308.9
Exponential Smoothing	6296748.18

Figure 46 Models Comparison

Based on RMSE values, the ARIMA model is the most accurate for forecasting total sales in this context. However, the choice of model can depend on specific forecasting needs:

- **ARIMA** is best for overall accuracy.
- **SARIMAX** is good for capturing sharp fluctuations.
- **PROPHET** provides a balanced, smoother forecast.
- **Exponential Smoothing** is less accurate but useful for stable short-term predictions.

Each model has its strengths, and the selection should be aligned with the specific requirements and characteristics of the sales data being analyzed.

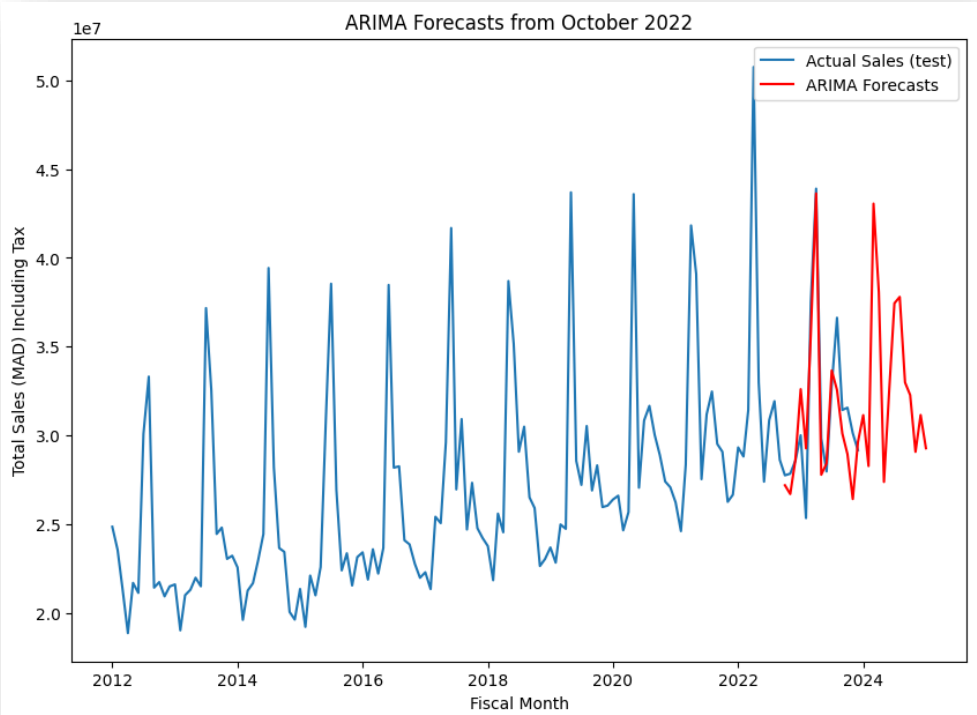


Figure 47 ARIMA Forecast

In this graph we display the ARIMA forecasts for total sales from October 2022 onwards. The blue line represents the actual sales data, while the red line indicates the predicted sales using the ARIMA model. The ARIMA model was chosen for forecasting because it demonstrated the highest accuracy compared to other models tested, as indicated by its lowest Root Mean Squared Error (RMSE) value. The model captures the overall trend and seasonal patterns in the sales data, with some deviations during periods of significant fluctuation. This forecast can help in planning and decision-making by providing insights into future sales performance based on historical data trends.

4. Conclusion

In this Chapter we detailed the implementation and deployment phase, showcasing the user interfaces and dashboards developed for the COPAG BI project. These tools provide comprehensive insights into sales performance, enabling detailed analysis and informed decision-making. The forecasting models evaluated during this phase, particularly the ARIMA model, enhance the ability to predict future sales trends. This implementation phase demonstrates the effective use of BI tools and techniques in optimizing sales strategies and supporting COPAG's strategic objectives.

GENERAL CONCLUSION

In conclusion, this internship has been an invaluable opportunity to explore and apply new technologies in the field of computer science, specifically in the realm of business intelligence and data analytics. The experience gained during this period allowed me to put into practice the knowledge acquired during my studies at the Higher School of Technology – Essaouira.

The internship provided a platform to enhance my skills in self-training and effective information research, as well as adapting to new work environments. Through the development and implementation of a decision support system for optimizing the sales of agro-food products at COPAG, I was able to gain hands-on experience with various BI tools such as SQL Server Integration Services (SSIS), Power BI, and Python. This project involved designing and deploying interactive dashboards and detailed reports that provide comprehensive insights into sales performance and trends.

The main objective was to create a robust BI system that integrates data from multiple sources, processes it efficiently, and presents it in a user-friendly manner to facilitate strategic decision-making. This was achieved by leveraging advanced data processing and visualization tools, ultimately supporting COPAG's goal of optimizing sales strategies and enhancing business operations.

Looking ahead, there are opportunities to further enhance the BI system by incorporating additional functionalities and improving data integration processes. The experience gained from this internship serves as a solid foundation for future projects and professional growth in the field of computer science and business intelligence.

Despite the challenges faced in analyzing requirements and adhering to time constraints, the majority of the project objectives were successfully met. This internship marks the beginning of my professional career in the field of computer science, serving as an initiation for future internships and projects that will further refine my skills and knowledge in this dynamic and rich domain.

WEBOGRAPHIE

DAX

- https://learn.microsoft.com/en-us/dax/dateadd-function-dax
- https://learn.microsoft.com/en-us/dax/average-function-dax
- https://learn.microsoft.com/en-us/dax/count-function-dax

POWER BI

- https://learn.microsoft.com/en-us/power-bi/visuals/desktop-matrix-visual
- https://learn.microsoft.com/en-us/power-bi/create-reports/power-bi-report-filter-types

SSIS

- https://learn.microsoft.com/en-us/sql/integration-services/data-flow/transformations/merge-join-transformation?view=sql-server-ver16
- https://learn.microsoft.com/en-us/sql/integration-services/data-flow/ole-db-source?view=sql-server-ver16
- https://learn.microsoft.com/en-us/sql/integration-services/data-flow/ole-db-destination?view=sql-server-ver16
- https://learn.microsoft.com/en-us/sql/integration-services/connection-manager/ado-connection-manager?view=sql-server-ver16
- https://learn.microsoft.com/en-us/sql/integration-services/data-flow/transformations/slowly-changing-dimension-transformation?view=sql-server-ver16