Presented by :

Practical Business Intelligence Project: MACROBUS

Business Intelligence Project

**ADAMA FALL**

*Subject* :  **MacroBus** is a company that sells vehicles across various markets and aims to measure the performance of its sales representatives over time. Following a discussion with the CEO of **MacroBus**, he expresses the need to enhance the company's performance and effectively manage its operations. Given the current context, **MacroBus**'s priority is to have a decision support system in place to efficiently manage the thousands of orders it receives.

[INTRODUCTION 3](#_Toc143704055)

[A. Design and Implementation 4](#_Toc143704056)

[1) Propose Model : Snowflake Multidimentional Model. 4](#_Toc143704057)

[2) Implementation of this model using SSIS (Considering DimDate and the question) 5](#_Toc143704058)

[a. The MAIN package 5](#_Toc143704059)

[b. The ODS package 6](#_Toc143704060)

[i. Sequence Container 6](#_Toc143704062)

[c. The DIM Package of the DIM Task 8](#_Toc143704063)

[i. Sequence container 1 9](#_Toc143704064)

[ii. Sequence container 2 12](#_Toc143704065)

[d. Task FAIT and package FactCommande 14](#_Toc143704066)

[3) Proposed script for loading the fact table through a stored procedure 15](#_Toc143704067)

[4) Among other dimension tables, create the DimCommercial, DimSupplier, DimContact, DimGeography, and DimManager dimensions. 15](#_Toc143704068)

[5) Proposal for a default record in each dimension table. 16](#_Toc143704069)

[B. Functional Validation 18](#_Toc143704070)

[1) Number and total amount of orders per territory. (Territory) 18](#_Toc143704071)

[2) Top 2 sales representatives who receive the most orders. 19](#_Toc143704072)

[3) The top three most ordered products during the second quarter of the year 2003. 20](#_Toc143704073)

[4) Total amount of orders per subsidiary and per product category during the second half of the year 2003. 21](#_Toc143704074)

[C. Creating visualizations in response to the previous questions (SAP Predictive Analysis) 22](#_Toc143704075)

[1. Number and total amount of orders per territory (Territory) 22](#_Toc143704076)

[2. Top 2 sales representatives who receive the most orders. 22](#_Toc143704077)

[3. The top three most ordered products during the second quarter of the year 2003. 23](#_Toc143704078)

[4. Total amount of orders per subsidiary and per product category during the second half of 2003. 23](#_Toc143704079)

[D. Summary dashboard for all four previous questions. 24](#_Toc143704080)

[E. Some execution screenshots of our SSIS packages 25](#_Toc143704081)

[1. ODS 25](#_Toc143704082)

[2. DIM 26](#_Toc143704083)

[3. FactCommande 29](#_Toc143704084)

[4. MAIN 29](#_Toc143704085)

[F. Final Dashboard (Power BI) 30](#_Toc143704086)

[CONCLUSION 31](#_Toc143704088)

INTRODUCTION

As part of the Business Intelligence final training project, we are required to implement a dashboard to assist the business leaders of the company **MACROBUS** in their decision-making process.

Our work will be focused on four (04) main areas, namely: designing and implementing a snowflake model tailored to our subject, ensuring the functional validation of our Data Warehouse, creating visualizations to address requirements and reporting goals, finally developing a dashboard using these visualizations.

1. Design and Implementation

In this section, we will design our multidimensional model and proceed with its implementation.

Snowflake Multidimensional Model.

Initially, for the snowflake model, the dimensions DimDate, DimProduct, DimEmployee, DimCustomer, and DimOffice are the only dimensions linked to the FactCommande fact table. The ProductLine dimension is connected to the DimProduct dimension.

Hence, the following schema:

©©

* **Adding new dimensions**
* Dimensions DimManager and DimCommercial

Managers represent a category of employees who oversee other employees.

Sales Representatives are employees who do not supervise any other employees.

Therefore, the Managers and Sales Representatives allow us to fully cover the DimEmployee dimension. As a result, the dimensions Dim Commercial and Dim Manager will be linked to the DimEmployee dimension.

* Dimension DimContact

The DimCustomer dimension contains attributes: Contact FirstName and Contact LastName. These pieces of information can be grouped into a new dimension, resulting in the DimContact dimension.

* The dimension DimSupplier

The DimProduct dimension contains information about the Product Vendor. Hence, the new dimension DimSupplier will be created, containing information about the suppliers for each product. This dimension will be connected to the DimProduct dimension.

* The dimension DimGeographie

The DimCustomer and DimOffice dimensions both contain location information. Therefore, a new dimension called DimGeography will be created to hold location information for both DimCustomer and DimOffice. Technically, a new field "Type" will be inserted into DimGeography to facilitate the loading of DimCustomer and DimOffice dimensions:

©©

1. Implementation of this model using SSIS (Considering DimDate and the question)

The main packages created for the project are:

1. The MAIN package

It allows for sequencing the execution of our main packages, namely the ODS package, the DIM package, and the FACT package.

1. The ODS package

This package ensures the sequencing of tasks involved in loading the ODS (Operational Data Store) area of our data warehouse from the production MACROBUS\_PROD database's base tables. spTruncateODS

This represents a task for executing an SQL query, which in this case runs the stored procedure that clears all the contents of the tables in our ODS schema. This stored procedure is named dbo.spTruncateODS.

**Sequence Container**

It enables the parallel execution of tasks that facilitate loading our ODS schema from the MACROBUS\_PROD database tables. These tasks are contained within the following package execution tasks: PCK\_Product, PCK\_ProductLine, PCK\_Customer, PCK\_Employee, PCK\_Office, PCK\_OrderDetails, PCK\_Orders.

* **PCK\_Product**

Loads the ods.Product table from the dbo.Product table in MACROBUS\_PROD.

* **PCK\_ProductLine**

Loads the ods.ProductLine table from the dbo.Product table in MACROBUS\_PROD.

* **PCK\_Customer**

Loads the ods.Customer table from the dbo.Customer table in MACROBUS\_PROD.

* **PCK\_Employee**

Loads the ods.Employee table from the dbo.Employee table in MACROBUS\_PROD.

* **PCK\_Office**

Load the ods.Office table from the dbo.Office table in MACROBUS\_PROD.

* **PCK\_OrderDetails**

Loads the ods.OrderDetails table from the dbo.OrderDetails table in MACROBUS\_PROD.

* **PCK\_Orders**

Loads the ods.Orders table from the dbo.Orders table in MACROBUS\_PROD.

1. The DIM Package of the DIM Task

This package orchestrates the execution of package tasks that contribute to loading all dimensions of the DWH schema in our Data Warehouse.

To execute the package execution task PCK\_DimEmployee, the Sequence Container 2 needs to be executed beforehand. And to execute the tasks within Sequence Container 1...

1. Sequence container 1

* **Dimension DimDate**

It allows for the execution of the stored procedure that loads the DimDate dimension.

* **Task PCK\_DimContact**

Enables the execution of the package that loads the DimContact dimension into our Data Warehouse.

k

The Contact source points to the ods.Customer table, and we select only the columns of interest: Contact FirstName, Contact LastName, phone, and CustomerNumber.

* **Task PCK\_DimFournisseur**

Allows the execution of the package that loads the DimSupplier dimension into our Data Warehouse.

The **Supplier** source is pointed to the ods.Product table, and we select the columns: ProductCode and ProductVendor.

* **Task PCK\_DimGeography**

Allows the execution of the package that loads the DimGeography dimension into our Data Warehouse.

As mentioned in question 1, we are using the ods.Customer and ods.Office tables here to populate the Geography source. Hence, the new column "Type" will have a value of 'office' for a record originating from the ods.office table, and a value of 'customer' for a record originating from the ods.Customer table.

**Task PCK\_DimProductLine**

Allows the execution of the package that loads the DimProductLine dimension into our Data Warehouse.

The ProductLine source is pointed to the ods.Product table, and we select the columns: ProductLine and ProductDescription.

**Task PCK\_DimManager**

Allows the execution of the package that loads the DimManager dimension into our Data Warehouse.

The Manager source contains all employees from ods. Employee who supervise other employees. We obtain them through the following query:

Then we create a calculated column named **ManagerName** that will contain the manager's last name and first name.

**Task PCK\_DimCommercial**

Allows the execution of the package that loads the DimCommercial dimension into our Data Warehouse.

The Commercial source contains all employees from ods.Employee who do not supervise any employees. We obtain them through the following query:

Then we create a calculated column named **CommercialName** that will contain the commercial's last name and first name.

**Sequence container 2**

After the successful execution of all tasks within Sequence Container 1, we can now proceed to the tasks within Sequence Container 2.

**Task PCK\_DimOffice**

Allows the execution of the package that loads the DimOffice dimension into our Data Warehouse.

The **Office** source is pointed to the ods.Office table. Then, we create a calculated column named **Type** and set it to 'office'. Afterwards (using a lookup object), we join the OfficeCode and Type columns with the codeGeography and Type columns of the DimGeography dimension to retrieve its **Key\_DimGeography** key.

**Task PCK\_DimProduct**

Allows the execution of the package that loads the DimProduct dimension into our Data Warehouse.

The Product source is pointed to the ods.Product table, and we retrieve the information that will be useful to us: ProductCode, ProductName, ProductLine, and Quantity in Stock. Then, using a lookup object, we retrieve the attribute **Key\_DimFournisseur** of the dimension DimSupplier using the ProductCode attribute, which is shared with our source. Then, we retrieve the **Key\_DimProductLine** attribute from the DimProductLine dimension using the ProductLine attribute, which is also shared with our source.

**Task PCK\_DimCustomer**

Allows the execution of the package that loads the DimCustomer dimension into our Data Warehouse.

The Customer source is pointed to the **ods.Customer** table. Then, we create a calculated column named Type and set it to **'customer'**. After that, we create a lookup object where we join the CustomerNumber and Type columns with the codeGeography and Type columns of the DimGeography dimension to retrieve its Key\_DimGeography key. Subsequently, we create another lookup object where we join the CustomerNumber column from our source with the CustomerNumber column of the DimContact dimension to retrieve the Key\_DimContact key from the latter.

**Task PCK\_DimEmployee**

Allows the execution of the package that loads the DimEmployee dimension into our Data Warehouse

.

As mentioned in question 1), our employee source is populated from the DimCommercial and DimManager dimensions, as indicated by the following query:

1. Task and package FactCommande

This task enables the execution of the FactCommande package, which is responsible for loading the fact table in our Data Warehouse.

Our fact table is constructed from 5 regular dimensions (DimCustomer, DimProduct, DimDate, DimEmployee, and DimOffice) and a degenerate dimension (OrderDetails) since all attributes of the latter are found in the fact table.

We begin by retrieving the content of the ods.OrderDetails table. Then, using a lookup object, we join it with the ods.Orders table based on the common attribute OrderNumber to obtain the CustomerNumber. This CustomerNumber is subsequently used (through another lookup object) to retrieve the Key\_DimCustomer key through a join on their common attribute CustomerNumber.

Next, through a lookup object, we retrieve the Key\_DimProduct key from the DimProduct dimension using a join on the ProductCode attribute, which is shared with our degenerate dimension (ods.OrderDetails).

Further, using a lookup object, we retrieve the Key\_DimDate key from the DimDate dimension through a join between ods.Orders.OrderDate and dwh.DimDate.Date.

Following this, a calculated column is created to generate a new column called Amount, which results from the product of quantityOrdered and Unitprice.

Subsequently, through a lookup object and a join between dwh.DimCustomer. SalesRepEmployeeNumber and dwh.DimEmployee.EmployeeNumber, we retrieve the Key\_DimEmployee key.

Finally, using another lookup object and a join between dwh.DimEmployee.OfficeCode and dwh.DimOffice.OfficeCode, we retrieve the Key\_DimOffice key from the DimOffice dimension.

1. Proposed script for loading the fact table through a stored procedure
2. Among other dimension tables, create the DimCommercial, DimSupplier, DimContact, DimGeography, and DimManager dimensions.

**Please refer to questions 1) and 2).**

1. Proposal for a default record in each dimension table.

FUNCTIONAL VALIDATION

Answers to the questions asked in the production database and in the Data Warehouse (DWH).

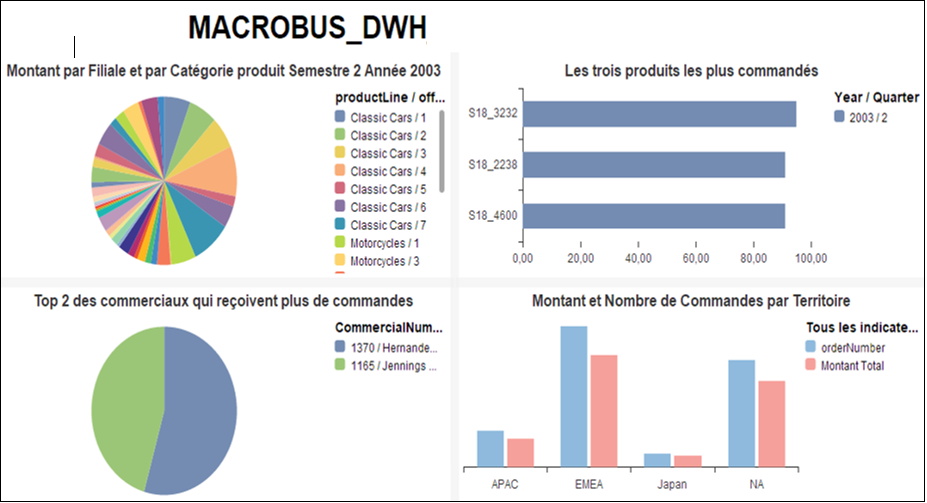
1. Number and total amount of orders per territory. (Territory)

1. Top 2 sales representatives who receive the most orders.
2. The top three most ordered products during the second quarter of the year 2003.
3. Determine the combined order value for each subsidiary and product
4. (SAP Predictive Analysis): Visualizations

using the tool ***SAP Predictive Analysis.***

1. Number and total amount of orders per territory (Territory)

1. Top 2 sales representatives who receive the most orders.
2. The top three most ordered products during the second quarter of the year 2003.
3. Total amount of orders per subsidiary and per product category during the second half of 2003.
4. Summary dashboard for all four previous questions.



1. screenshots of our SSIS packages
2. ODS

* SRC\_ Employee
* SRC\_Office
* SRC\_OrderDetails
* SRC\_Oders
* SRC\_Product
* SRC\_ProductLine

1. DIM

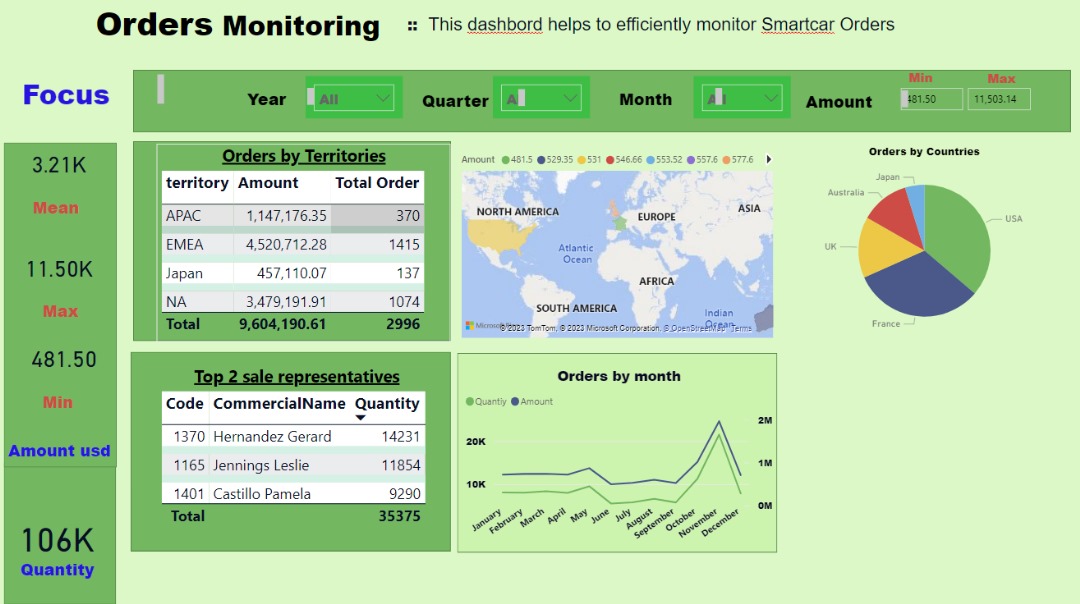
* DimCommercial
* DimContact
* DimCustomer
* DimEmployee
* DimFournisseur
* DimGeography
* DimManager
* DimOffice
* DimProduct
* DimProductLine

1. FactCommande

1. MAIN

1. Final Dashboard (Power BI)

We then realize the Dashboard below :



CONCLUSION:

This project has provided us with a valuable opportunity to comprehensively revisit all the subjects covered during our training. Our expedition commenced by conceptualizing the “MACROBUS\_DWH” data warehouse schema and subsequently extended to crafting a decision-support Dashboard. This endeavor entailed not only its full – fledged. implementation but also meticulous functional validation. Nonetheless, it is worth highlighting that the “data warehouse schema design” phase emerged as the most challenging juncture due to its pivotal significance in molding and influencing all ensuing facets of our MACROBUS project.

CONCLUSION:

This project has provided us with a valuable opportunity to comprehensively revisit all the subjects covered during our training. Our expedition commenced by conceptualizing the “MACROBUS\_DWH” data warehouse schema and subsequently extended to crafting a decision-support Dashboard. This endeavor entailed not only its full – fledged. implementation but also meticulous functional validation. Nonetheless, it is worth highlighting that the “data warehouse schema design” phase emerged as the most challenging juncture due to its pivotal significance in molding and influencing all ensuing facets of our MACROBUS project.