



ÉCOLE NATIONALE SUPÉRIEURE
D'INFORMATIQUE ET D'ANALYSE DES SYSTÈMES
- RABAT

PROJET BI

BBS Statistical DW : Industry Business Wing

Élèves :

Ayman FENKOUCH
Yahya MOUDRIK
Hiba ASGHAR
Youssef RIZKI

Enseignant :
Mme. BENHIBA

Table des matières

0.1	Data Warehouse Physical Design	2
0.1.1	Dimension Tables	2
0.1.2	Fact Table	2
0.2	ETL Conception	2
0.2.1	ETL Architecture	2
0.2.2	ETL Design for Dimension Tables	3
0.2.3	ETL Design for the Fact Table	3
0.3	ETL Development Using SSIS	3
0.3.1	Development of Dimension ETL Data Flows	3
0.3.2	Development of the Fact Table ETL Data Flow	3
0.3.3	ETL	4
0.4	User Application : Power BI Dashboards	4
0.4.1	Key Performance Indicators	4
0.4.2	Analytical Visualizations	4
0.4.3	Dashboard	5
0.5	Test Dataset	5
1	Conclusion	5

Introduction

Implementation and Realization

This chapter describes the technical realization of the Business Intelligence system developed for the Statistical Bureau of Bangladesh, focusing on the analysis of external trade data (Imports, Exports, and Trade Balance). The implementation covers the physical design of the Data Warehouse, the conception and development of ETL processes using SSIS, and the implementation of the user application using Power BI.

0.1 Data Warehouse Physical Design

The Data Warehouse was designed following a **star schema architecture**, which is well suited for analytical and reporting purposes. This architecture ensures simplicity, high query performance, and ease of use for analytical users.

0.1.1 Dimension Tables

The following dimension tables were implemented in Microsoft SQL Server :

- **DimDate** : provides the temporal analysis axis, including year information.
- **DimCountry** : represents trading partner countries.
- **DimAccount** : defines the type of trade account (Imports or Exports).
- **DimRoute** : indicates the transportation mode (Sea, Land, Air).
- **DimComposition** : describes the type of traded goods (Consumer goods, Capital goods, Materials, etc.).

Each dimension table uses a surrogate key as its primary key, referenced by the fact table to ensure referential integrity.

0.1.2 Fact Table

The central fact table, **FactTrade**, stores quantitative measures related to external trade activities. The main measures include :

- Export value
- Import value
- Trade balance

The fact table references all dimension tables through foreign keys, allowing multidimensional analysis of trade data.

0.2 ETL Conception

The ETL processes were designed to integrate data from heterogeneous sources into the Data Warehouse in a reliable and structured manner. The source data consists of **Excel files containing yearly survey data** on Bangladeshs external trade sector.

0.2.1 ETL Architecture

The ETL architecture is based on a clear separation between dimension loading processes and fact loading processes. Dimension tables are populated first in order to provide reference data required by the fact table.

0.2.2 ETL Design for Dimension Tables

For each dimension, a dedicated ETL data flow was designed. The design principles applied to dimension ETL processes are as follows :

- Extraction of data from Excel survey files.
 - Selection of descriptive attributes relevant to each dimension.
 - Identification and elimination of duplicate records.
 - Preparation of clean and consistent reference data for loading.
- This design ensures the uniqueness and consistency of dimension data.

0.2.3 ETL Design for the Fact Table

The ETL design for the fact table is based on the use of descriptive attributes present in the source data and their transformation into foreign keys.

The design includes :

- Mapping of source attributes to dimension tables.
- Definition of lookup relationships between fact data and dimensions.
- Preparation of quantitative measures for analytical use.

0.3 ETL Development Using SSIS

The ETL processes were implemented using **SQL Server Integration Services (SSIS)**, in accordance with the previously defined ETL design.

0.3.1 Development of Dimension ETL Data Flows

Each dimension ETL data flow was developed following the same implementation logic :

1. Connection to Excel files containing yearly survey data.
2. Selection of the columns required for the dimension.
3. Use of a *Sort* transformation with the *Remove Duplicates* option to ensure uniqueness.
4. Loading of distinct records into the corresponding dimension table.

This implementation guarantees clean and non-redundant dimension data.

0.3.2 Development of the Fact Table ETL Data Flow

The ETL data flow for the **FactTrade** table was developed to integrate transactional survey data into the Data Warehouse.

The implementation steps are as follows :

1. Extraction of trade data from Excel survey files.
2. Application of multiple *Lookup* transformations to retrieve surrogate keys from each dimension table.
3. Replacement of descriptive attributes with foreign keys.
4. Loading of the transformed records into the **FactTrade** table.

This approach ensures data consistency and referential integrity.

0.3.3 ETL

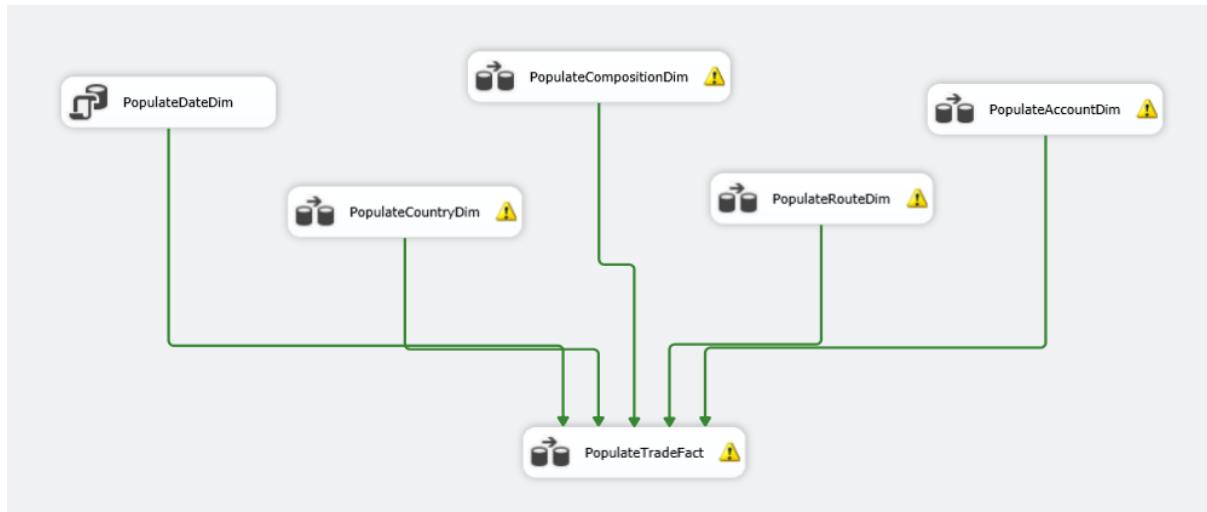


FIGURE 1 – Enter Caption

0.4 User Application : Power BI Dashboards

The user application was developed using **Microsoft Power BI** and provides an interactive interface for analyzing external trade data stored in the Data Warehouse.

0.4.1 Key Performance Indicators

The application displays several key performance indicators that summarize the state of Bangladeshs external trade :

- Total Imports
- Total Exports
- Trade Balance
- Comparison with the previous year

These indicators provide a quick overview for decision-makers and analysts.

0.4.2 Analytical Visualizations

The Power BI application includes interactive visualizations such as :

- Imports and exports by country.
- Trade values by goods composition type.
- Trade analysis by transportation route.
- Yearly evolution of imports, exports, and trade balance.

Dynamic filters (Year, Account Type, Country, Route, Composition) allow users to explore the data from multiple analytical perspectives.

0.4.3 Dashboard

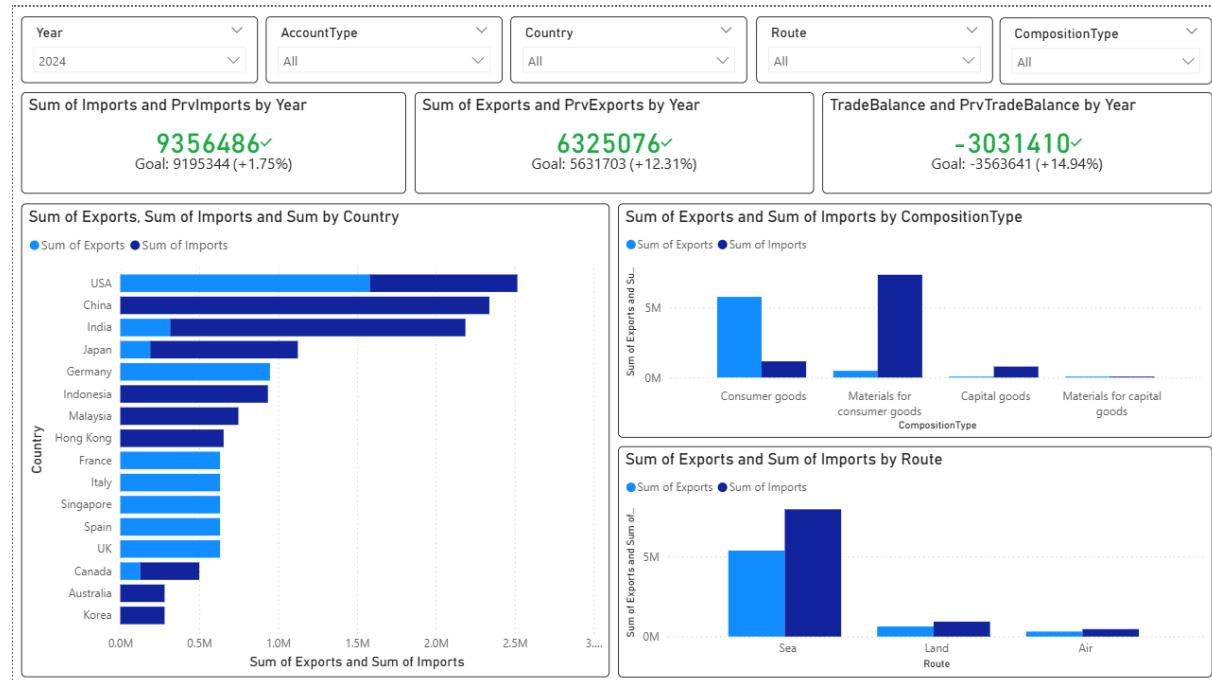


FIGURE 2 – Enter Caption

0.5 Test Dataset

A test dataset was created based on Excel survey files representing yearly external trade data for Bangladesh. This dataset was used to validate :

- The correctness of ETL development.
- The integrity of dimensional relationships.
- The accuracy and coherence of Power BI reports.

1 Conclusion