A black and white logo

Description automatically generated

**BI System Specifications Document**

**Written by: Noa Finkeltein**

1. **General**
   1. **Project Objective**

This project's objective is the creation of a comprehensive BI solution for the sales department of Western Digital, to support the company's growth in product sales. The project was designed according to the sales department's KPIs and is aimed at increasing the company's overall ROI.

Western Digital is an American computer drive manufacturer and data storage company. It designs, manufactures and sells data technology products, including data storage devices, data center systems and cloud storage services.

The Data Mart creation will be done using information derived from the Priority ERP operational database. The solution will include summarized data tables, focusing on product sales data, as well as data regarding the company's customers, sales personnel, products, and stores. In addition, the BI solution will include customized reports containing sales analysis, customer analysis, and an executive dashboard. These reports will be tailored to the sales department's needs and will contribute to Western Digital continued growth and success.

**1.2 Project Contents**

The project includes the building of a Data Mart which contains sales data. The data will be transferred through an ETL process from the PriorityERP operational database to the Data Mart – WesternDev.

1.2.1 The Data Mart will include 1 fact table and 6 MRR tables:

* FactSales: Data regarding all sales, including the id of the order, id of the customer, id of the employee, id of the product, products bought quantities, and prices. The Data loading process for this table will be incremental.
* DIM\_Customers: Data regarding the company’s customers.
* DIM\_Employees: Data regarding the company’s employees.
* DIM\_Stores: Data regarding the company’s stores.
* DIM\_Products: Data regarding the company’s products.
* DIM\_Categories: Data regarding the company’s categories.
* DIM\_SubCategories: Data regarding the company’s subcategories.

To provide a more comprehensive detail, attached is a link to Source To Target:

[Source To Target Link](https://1drv.ms/x/s!AtGhoISmIJ2sky8CsQj0sjp3v8CS?e=2HqGgE)

1.2.2 The reports will include data visualizations that will support the project’s objective in the following ways in power bi:

* Administration Report
* Company Report
* Customers Report

1. **Gnatt**

[**Gantt Link**](Gantt.html)

1. **Technical Specification**
   1. **Prerequisites**

|  |  |
| --- | --- |
| SQL Server | ERP system in the operational DB (PriorityERP) – tables, data (SQL files) |
| SSIS | ETL processes using SSIS in Visual Studio |
| Data refresh processes | Definition of JOBS in SSMS |
| Power BI | Creating reports and dashboards using Power B |

1. **Functional Specification**
   1. Creation of final Source to Target and ERD models.
      1. Source To Target

[Source To Target Link](https://1drv.ms/x/s!AtGhoISmIJ2sky8CsQj0sjp3v8CS?e=2HqGgE)

The total of all the tables will be used from the operational database.

* + 1. ERD Model

[ERD Link](https://drawsql.app/teams/noa-finkelstein/diagrams/final-project)

* 1. **ETL processes**

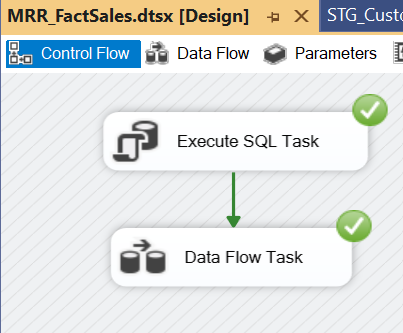
The ETL process was done in SSIS with packages.

All the packages include Execute SQL tasks in the control flow, and data flow processes.

Mirror tables are truncated, using a stored procedure, followed by loading data from the PriorityERP database. The MRR process was executed to perform a transfer of all data, across the tables in the project, as follows:

MRR\_FactSales:

For table MRR\_FactSales, data was extracted from tables Invoices and InvoiceItems in PriorityERP database.

****

A screenshot of a computer

Description automatically generated

DIM\_Employees:

For table DIM\_Employees, data was extracted from tables Employees and Invoices in PriorityERP database. Additionally, include additional data.

DIM\_Customers:

For table DIM\_Customers, data was extracted from tables Customers, Destcodes, States and Countries in PriorityERP database.

DIM\_Products:

For table DIM\_Products, data was extracted from tables Part, Family, and Familytypes in PriorityERP database.

DIM\_Stores:

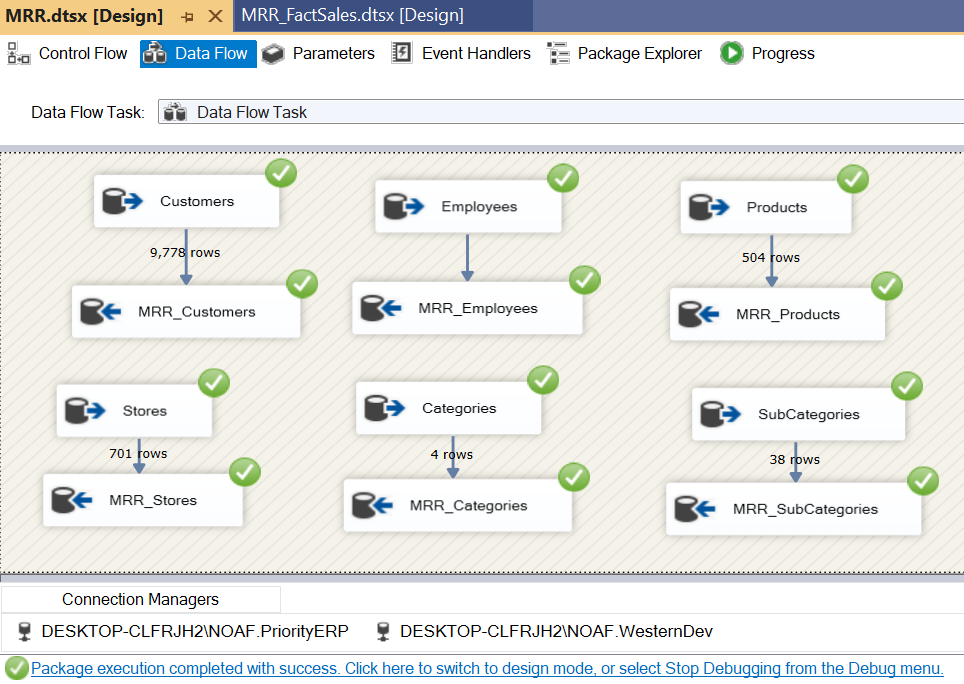
For table DIM\_Stores, data was extracted from tables Branches, Customers, Invoices, Destcodes and States in PriorityERP database.

DIM\_Categories:

For table DIM\_Categories, data was extracted from table Familytypes in PriorityERP database.

DIM\_SubCategories:

For table DIM\_SubCategories, data was extracted from table Family in PriorityERP database.

****

**A screenshot of a computer

Description automatically generated**

A screenshot of a computer

Description automatically generated

For the packages of STG processes, the following variables were defined:

* Table Name, type: string
* Row Count, type: int32
* Start Time, type: datetime
* End Time, type: datetime

All these variables were inserted into the DataTransfer table.

All the STG tables are truncated, and the mirror tables are joined and loaded using a data flow task, executed in an Execute SQL task.

A screenshot of a computer

Description automatically generatedSTG\_Employees:

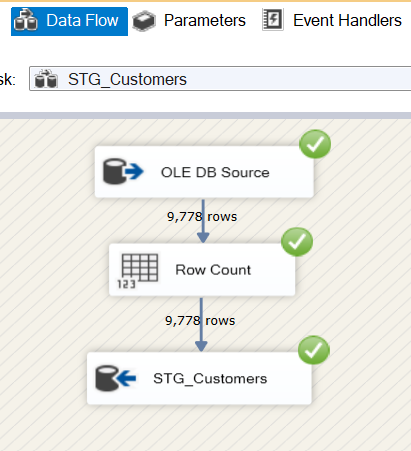
A screenshot of a computer

Description automatically generated

STG\_Customers:

**A screenshot of a computer

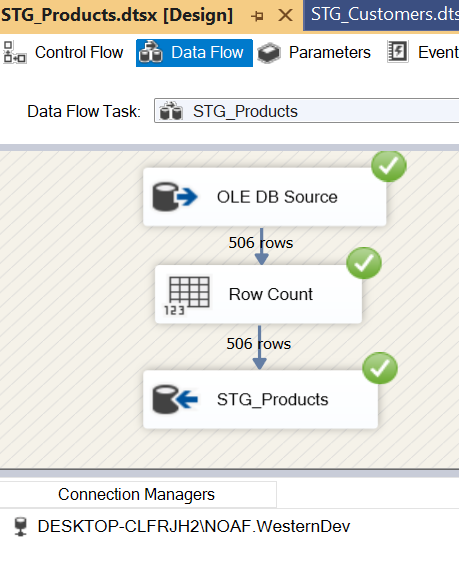
Description automatically generated**

****

STG\_Products:

**A diagram of a process

Description automatically generated**

****

STG\_Stores:

**A diagram of a process

Description automatically generated**

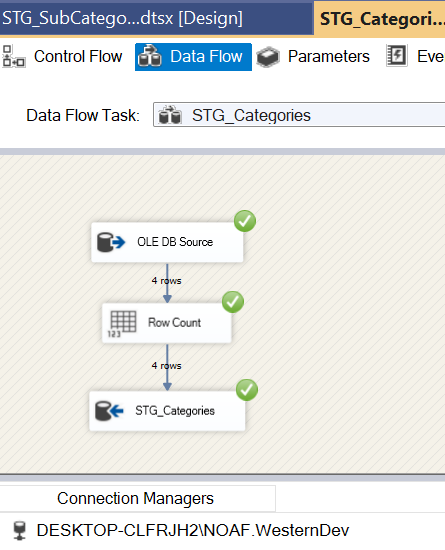
**A screenshot of a computer

Description automatically generated**

STG\_Categories:

**A screenshot of a computer

Description automatically generated**

****

**A screenshot of a computer

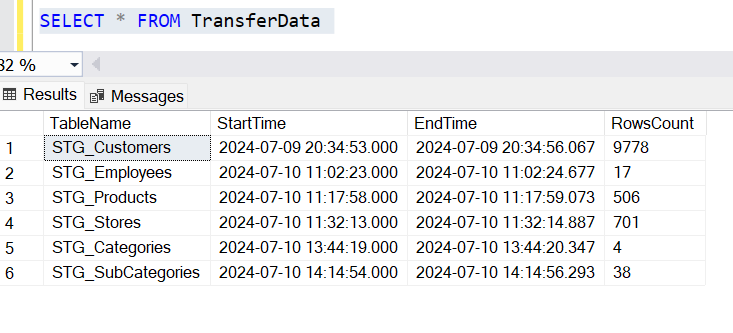
Description automatically generated**STG\_SubCategories:

**A screenshot of a computer

Description automatically generated**

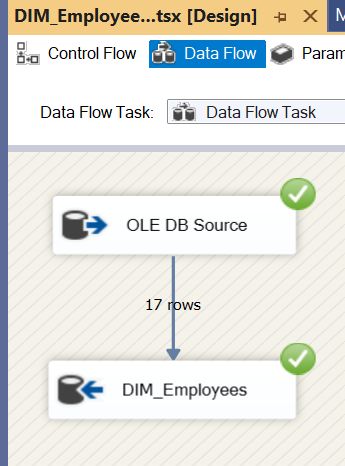
The following data is stored in TransferData Table, for each of the STG tables,

as follows:



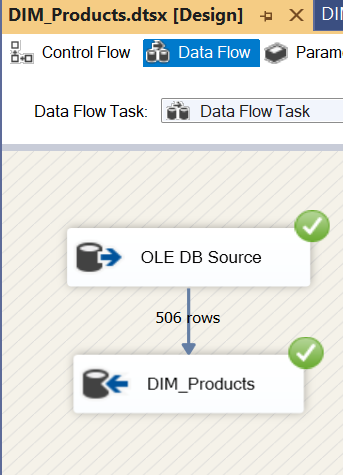
A screenshot of a computer

Description automatically generatedIn package FactSales, the column EmployyeID from table DIM\_Employee, the column CustomerID from table DIM\_Customers, the column ProductID from table DIM\_Products, were extracted at Merge Join, and then the operations were performed.

****DIM\_Employees:

A screenshot of a computer

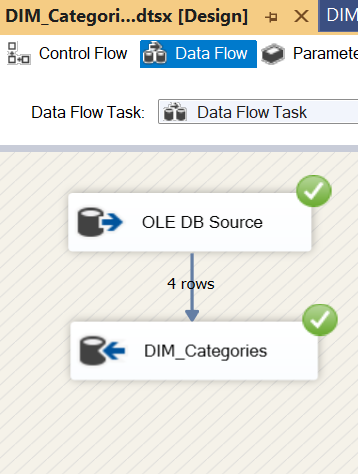
Description automatically generated DIM\_Customers:

 DIM\_Products:

A screenshot of a computer

Description automatically generated DIM\_Stores:

DIM\_Categories:



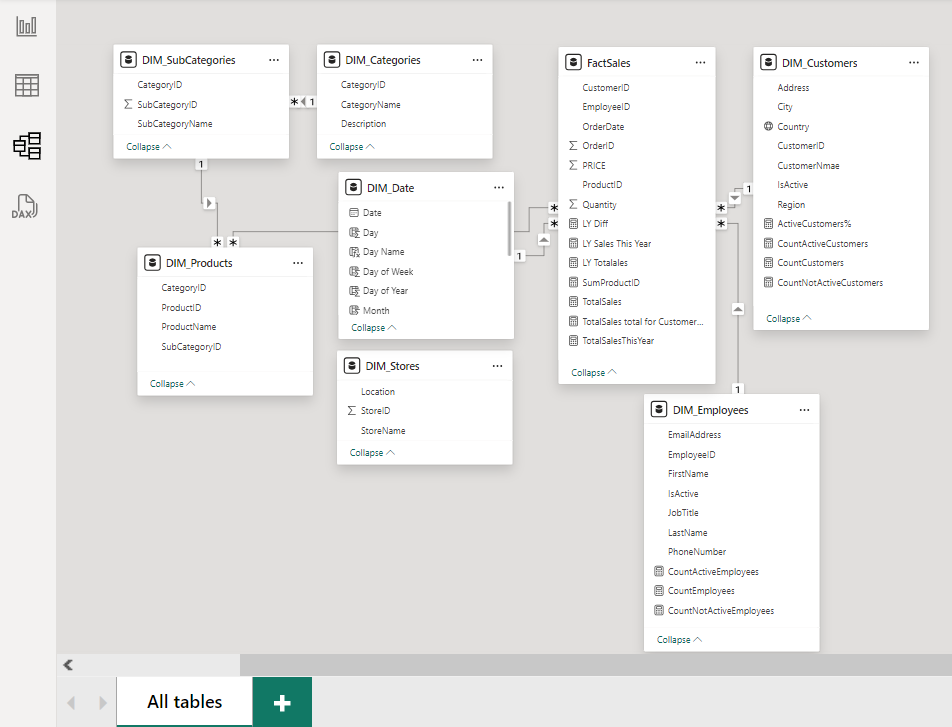
DIM\_SubCategories:

A screenshot of a computer

Description automatically generated

**4.3. Visualization in Power BI:**

The reports were created using Power BI Desktop and were published to Power BI Service. The model in the Power BI includes the Fact Sales table and the 6 Dimension tables. To these tables, a Dim Date table was added.



Measures for visuals in FactSales table:

TotalSales = SUMX(FactSales,FactSales[PRICE] \* FactSales[Quantity])

TotalSalesThisYear =

VAR MaxYear = YEAR(MAX(FactSales[OrderDate]))

RETURN (

CALCULATE([TotalSales],DIM\_Date[Year] = MaxYear))

TotalSales total for CustomerName =

CALCULATE([TotalSales], ALLSELECTED('DIM\_Customers'[CustomerNmae]))

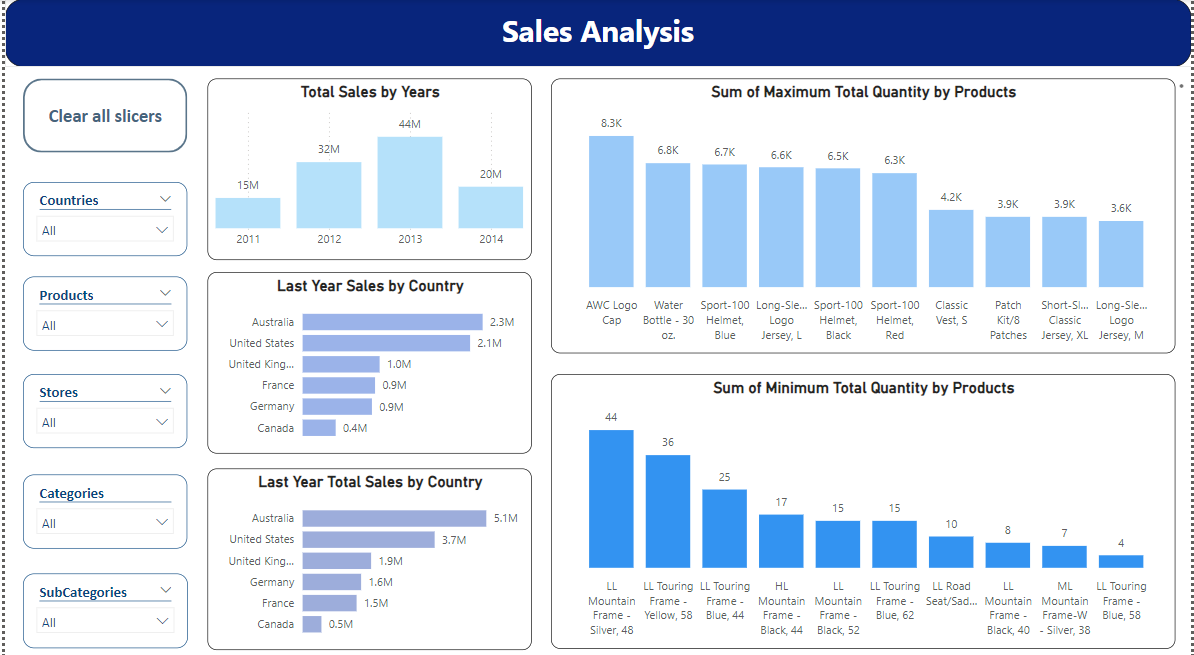
SumProductID = SUM(FactSales[ProductID])

LY Totalales = CALCULATE([TotalSales],SAMEPERIODLASTYEAR(DIM\_Date[Date]))

LY Sales This Year = CALCULATE([TotalSalesThisYear],SAMEPERIODLASTYEAR(DIM\_Date[Date]))

LY Diff = [TotalSales] - [LY Totalales]

Administration Report:



Company Report:

A screenshot of a company report

Description automatically generated

Customers Report:

