**Somatic’s Sales Data Mart**

תמונה שמכילה גופן, גרפיקה, טקסט, צילום מסך

התיאור נוצר באופן אוטומטי

Date: 02/07/2024

Version: 4.00

Written by: Yanai Bardosh

Content

1. General

1.1 Purpose of the Project

1.2 Project Content

2. Preparation of a Work Plan (Gantt)

3. Technical Characterization

3.1 Prerequisites - Complete list of systems including access methods

3.2 Solution Architecture(HLD)

4 .Functional Characterization

4. 1 Creation of final Source to Target and ERD models

4. 2 Detailed Description of All ETL Processes

4. 3 Details of Indices, Tables, and Filters for the Reports in PBI

1.General

1.1 Purpose of the Project

The objective of this project is to develop a comprehensive BI solution for Somatix's sales department to support the growth of SafeBeing™ sales. This project is tailored to meet the KPIs of Somatix's sales department and aims to enhance the company's overall ROI.

Somatix is an AI-powered digital health company founded on the belief that all patients deserve equitable, affordable, high-quality healthcare. Their technology is designed to detect early behavioral and emotional indicators of declining health and provide actionable clinical insights. Using its unique and patented gesture detection technology, Somatix’s digital health platform enables providers to remotely and passively track for risk factors of adverse events, including poor medical compliance, falls, and dehydration. This novel, yet highly intuitive approach simplifies provider workflows and supports connected care through real-time remote patient monitoring and telemedicine capabilities.

The Data Mart will be created using data from the PriorityERP database, Somatix's operational database. This solution will include summarized data tables focusing on company sales and customer data. Additionally, the BI solution will feature customized reports on sales analysis, customer analysis, and an executive dashboard. Tailored to the needs of the sales department, these reports will contribute to the growth of Somatix’s SafeBeing™ sales.

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 – SomatixSalesDM.

1.2.1. The Data Mart will include 1 fact table and 4 dimension tables, and 1 history table:

• **Fact\_Sales** – Data regarding all sales, including the id of the order, products bought, quantities, and prices.

• **Dim\_Person\_Customers** – Data regarding the company’s Person customers.

• **Dim\_ Location\_Customers** – Data regarding the company’s Location customers(Hospitals, Research Centers and rehab Centers).

• **Dim\_Sales\_Person** – Data regarding the company’s Salespersons.

• **Dim\_Products** - Data regarding the company’s products.

• **Dim\_Customers\_History** – Historic data regarding the company’s Person Customers.

The tables will be updated daily at 00:00:00 using an automated process configured in the SQL Server Management Studio.

1.2.2. The project will include a Transfer Table. This table will allow us to monitor data movement by showing the number of rows transferred and the time taken for each package in the ETL process.

1.2.3. The reports will include data visualizations in Power BI that will support the project’s objective in the following ways:

• **Sales Analysis:**

The Sales Analysis Dashboard will offer a detailed overview of the company's sales activities, providing insights into overall revenue generation, performance across different products, and sales trends over time. It will highlight the contributions of individual salespersons. The dashboard will also include Location Customers sales distribution, helping to identify strong and weak location's type. This view will support the sales team in understanding sales dynamics and optimizing their strategies to boost performance.

• **Customer Analysis:**

The Customer Analysis Dashboard will focus on understanding the customer base and their behaviors. It will present demographic data, purchasing patterns, and customer segmentation to reveal who the customers are and how they interact with the company's products. The dashboard will also track customer engagement levels, identifying active and inactive customers, and providing insights into customer acquisition trends. This analysis will help the sales team tailor their approaches to meet customer needs better and enhance customer satisfaction and loyalty.

• **Executive Dashboard:**

The Executive Dashboard will provide a high-level summary of the company's overall performance, focusing on strategic insights and trends. It will offer a snapshot of sales performance, key growth indicators, and ROI metrics, allowing executives to quickly assess the company's performance. This dashboard will support strategic planning and drive the company's growth and success.

2.Gantt

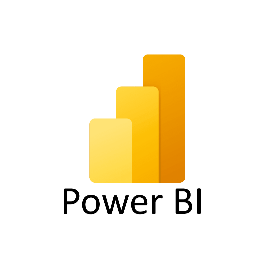
I used ClickUp Platform to create my Gant, if you don't have an account please use my account: Email - [Yanaib1215@Gmail.com](mailto:Yanaib1215@Gmail.com), Password - Tsv1860m. The Gantt is in the following [**Link**](https://app.clickup.com/9018377767/v/g/8crk1h7-398)

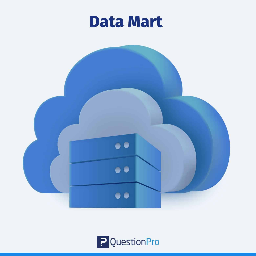
3.Technical Characterization

3.1. Prerequisites

|  |  |
| --- | --- |
| SQL SERVER | ERP system in the operational DB (PriorityERP) |
| SSIS | ETL Process Using SSIS In Visual Studio |
| Data refresh processes | Definition of JOBS in SSMS |
| POWER BI | Creating repots and dashboard using Power BI |

3.2. Solution Architecture

3.2.1. High Level Design:



Visualization

Sales Data Mart

SSIS

Operational DB

4.Functional Characterization

4.1. Creation of final Source to Target and ERD models.

4.1.1. [**Source To Target link**](S2T%20Mapping.xlsx)

A total of 11 tables will be used from the operational database.

4.1.2. [**ERD link**](https://drawsql.app/teams/bi-60/diagrams/somatic-s-sales-data-mart)

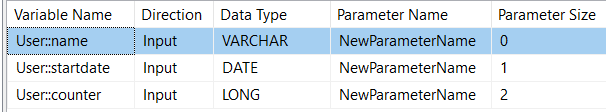
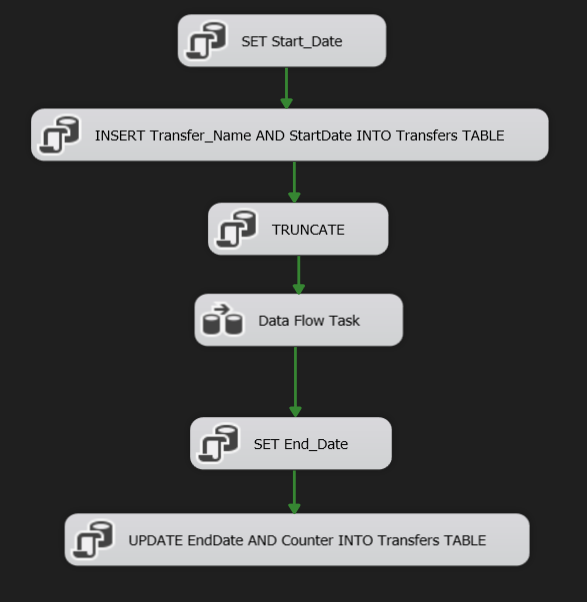
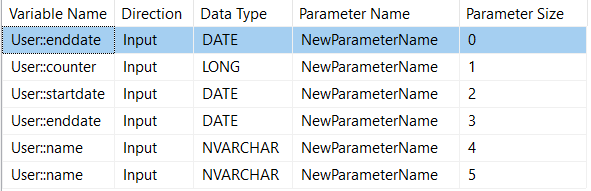
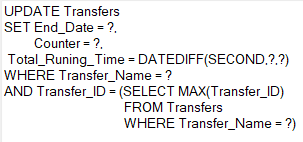
4.2. ETL processes

**• Transfer Table**

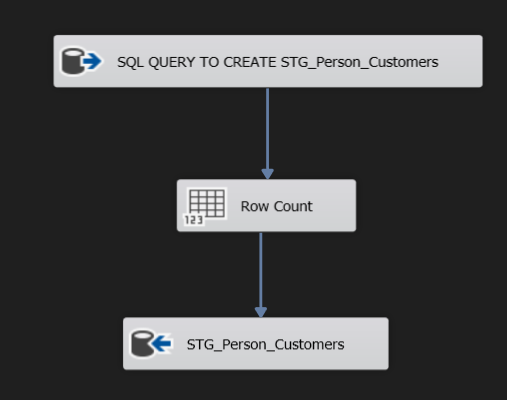
To monitor the ETL process, I created a Transfer table that documents each data insert, including which table was updated, how many rows were inserted, the start and end times, and the run duration of the package. The tasks and transformations responsible for the updates are included in all the packages.

Example from STG\_Person\_Customers package:

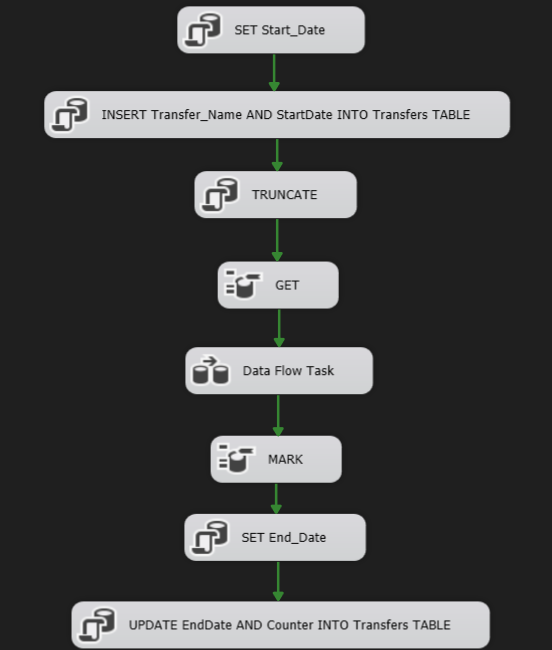
In the control flow user variable startdate is set in the first task with GETDATE() as startdate, an insert statement is executed in the second task, inserting the values of the user variables: name, startdate, counter (which is updated in the data flow), user variable enddate is set in the fifth task with GETDATE() as enddate. and an update statement is executed in the last task.



תמונה שמכילה טקסט, צילום מסך, גופן, מספר

התיאור נוצר באופן אוטומטיIn the data flow the user variable counter is updated using a Row Count transformation.

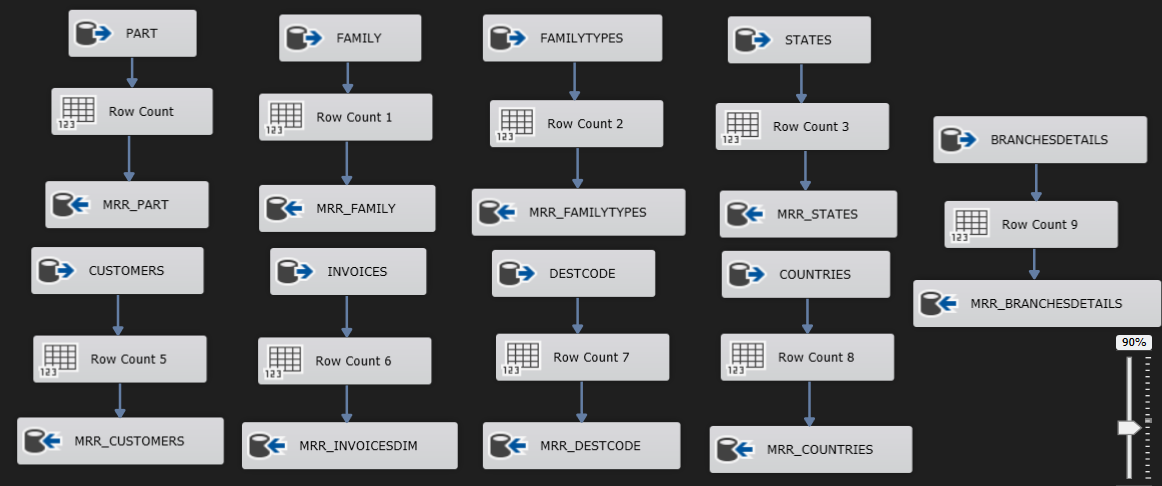
**• MRR\_DIM\_Tabels Package**

This package is responsible for loading data from PriorityERP tables to all mirror tables relevant for the dim tables (10 tables in total). All the tables in this package are truncated using a stored procedure.



For Change Data Capture on BRANCHES Table

In the Data Flow Task:

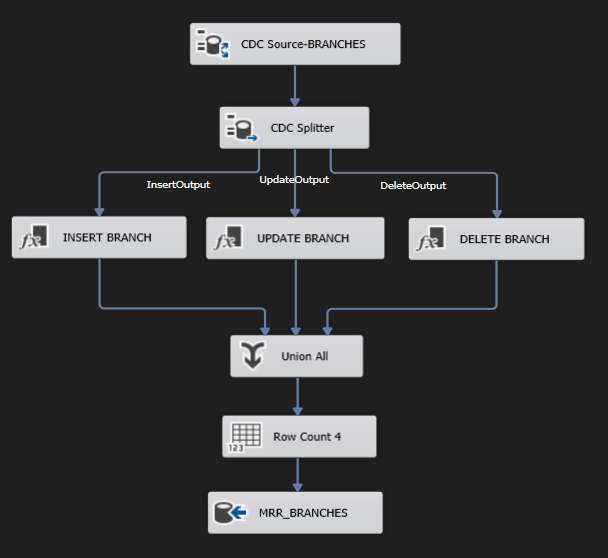
Regular MRR

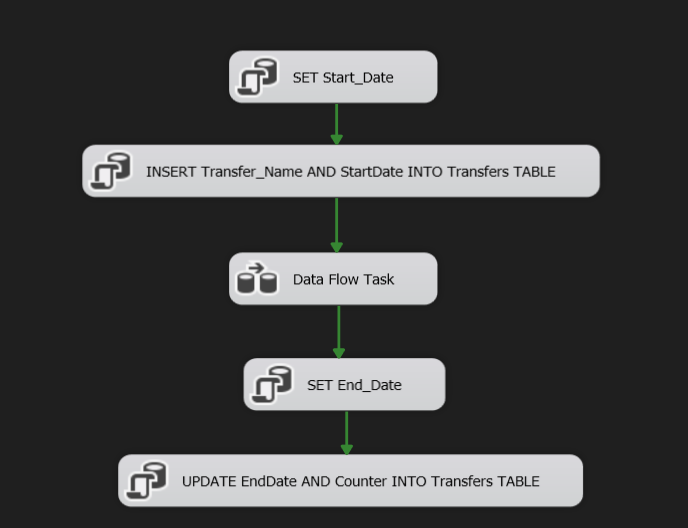
Change Data Capture on BRANCHES Table

Change data capture (CDC) refers to the process of identifying and capturing changes made to data in BRANCHES table and then delivering those changes in the ETL proccess.

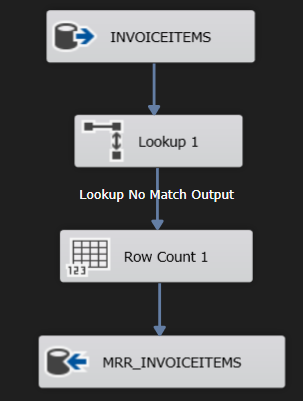
תמונה שמכילה טקסט, צילום מסך, גופן, קו

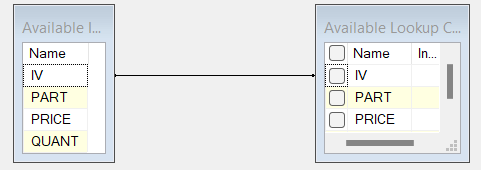
התיאור נוצר באופן אוטומטיתמונה שמכילה טקסט, צילום מסך, גופן, קו

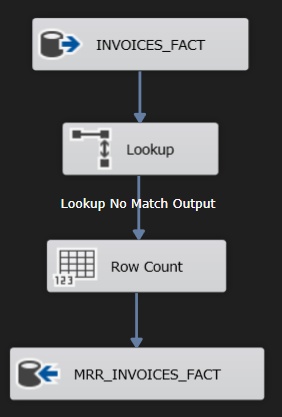
התיאור נוצר באופן אוטומטי

**• MRR\_FACT\_Tabels Package**

In the Data Flow Task:

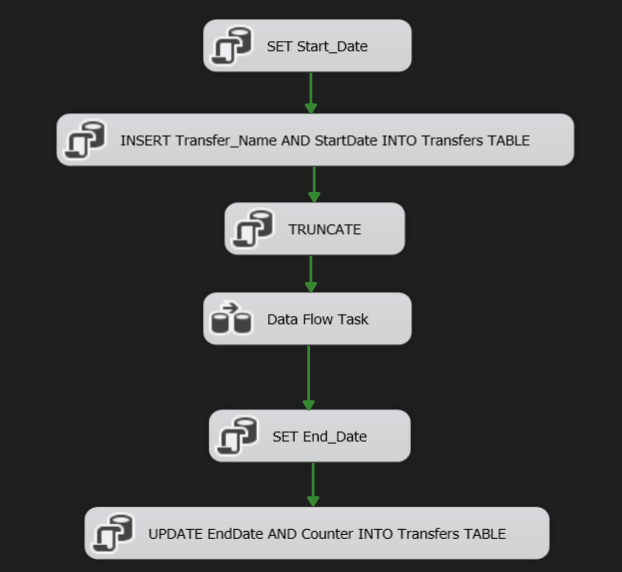
Loading data into MRR\_INVOICES\_FACT and MRR\_INVOICEITEMS is performed incrementally using the Lookup transformation, which compares records to existing entries in MRR\_INVOICES\_FACT and MRR\_INVOICEITEMS. The tables are not truncated during this process.





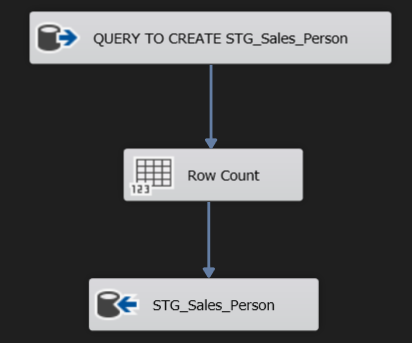
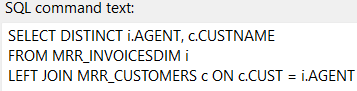
תמונה שמכילה טקסט, צילום מסך, תרשים, קו

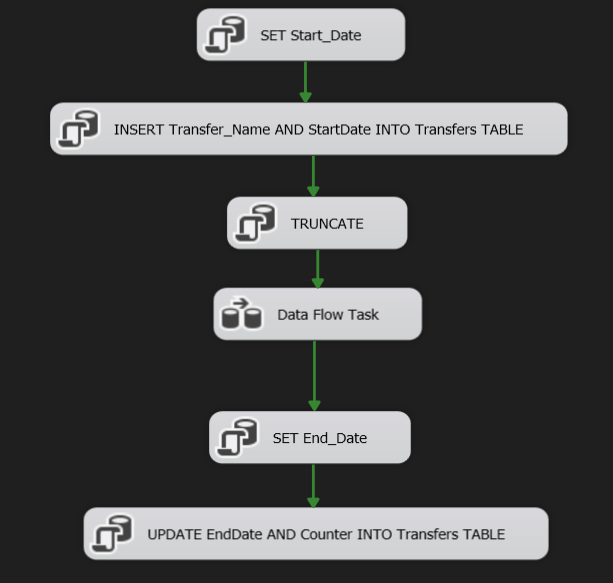
התיאור נוצר באופן אוטומטי

**• STG\_Sales\_Person Package**

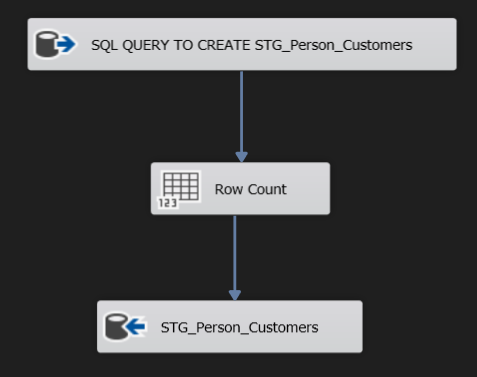


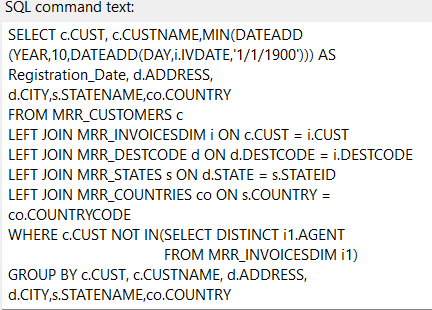
In the Data Flow Task:

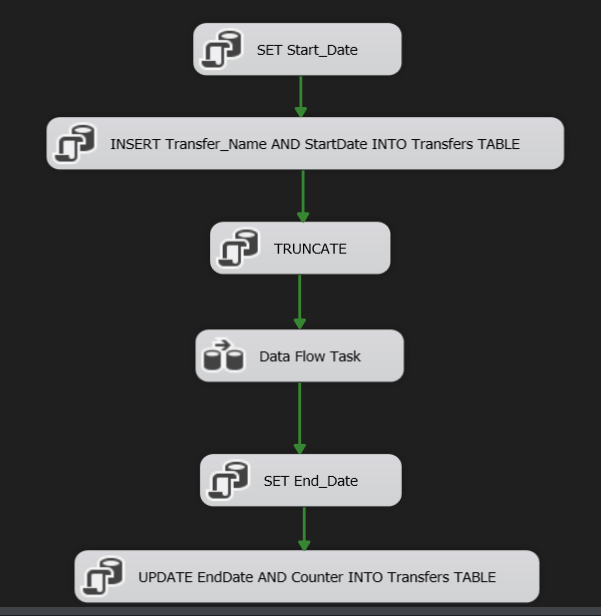


**• STG\_Person\_Customers Package**

****

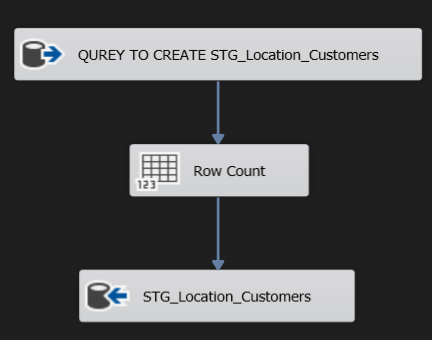
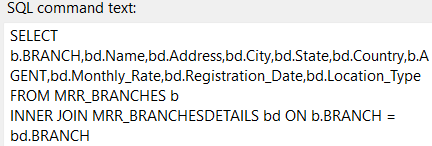
****In the Data Flow Task:

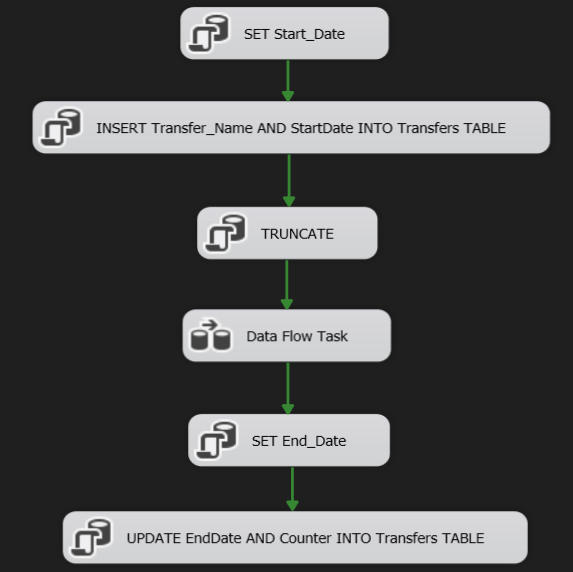
****

**• STG\_Location\_Customers Package**

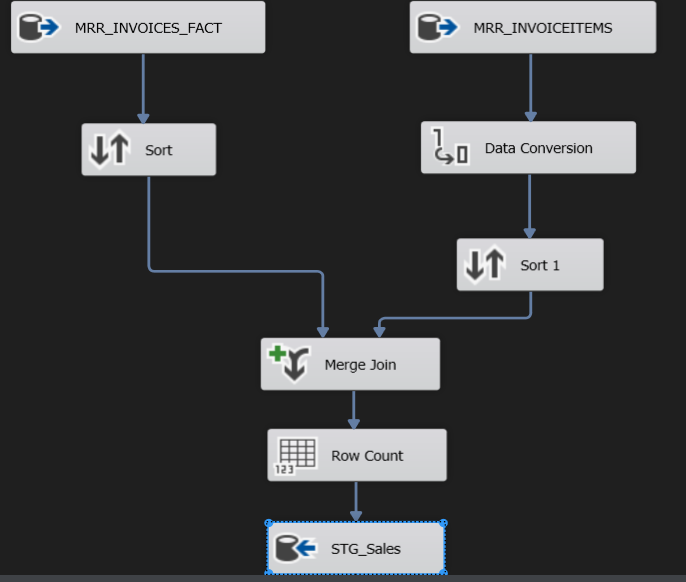
****

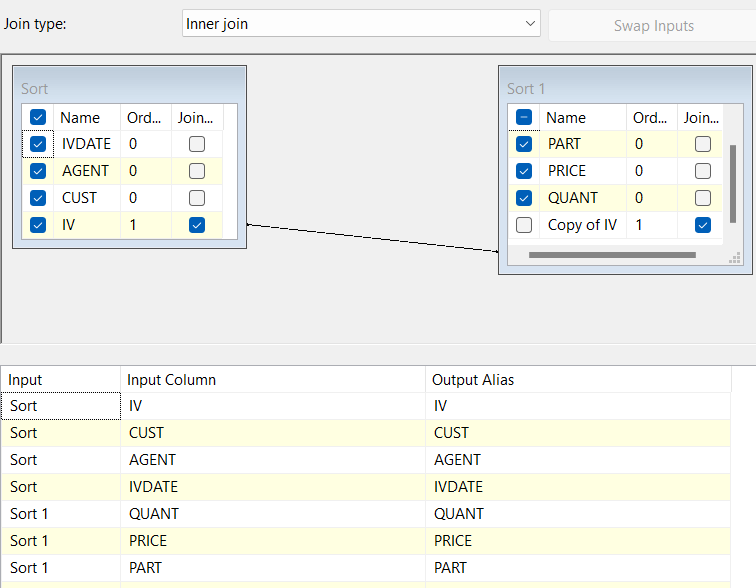
In the Data Flow Task:

****

**• STG\_Sales Package**

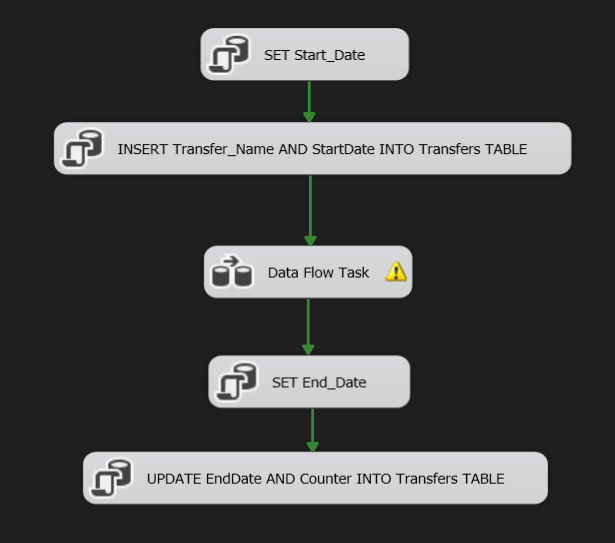
****

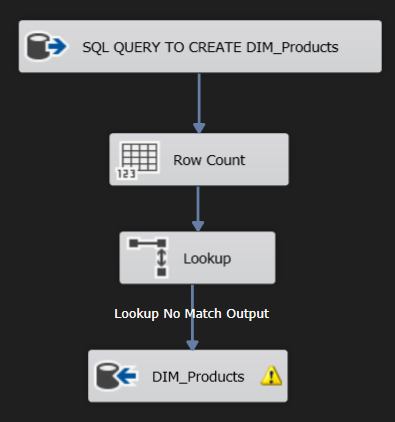
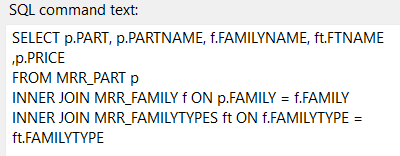
****In the Data Flow Task:

****

**• DM\_Products Package**

Loading data into DIM\_Products is performed incrementally using the Lookup transformation, which compares records to existing entries in DIM\_Products. The table is not truncated during this process.

****

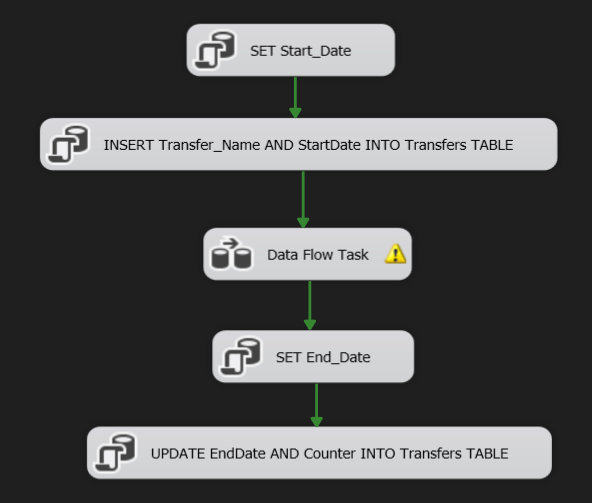
****In the Data Flow Task:

**תמונה שמכילה טקסט, צילום מסך, גופן, מספר

התיאור נוצר באופן אוטומטי**

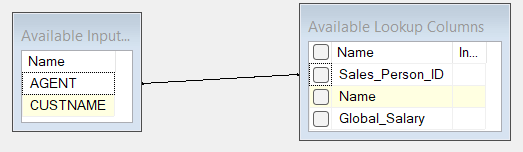
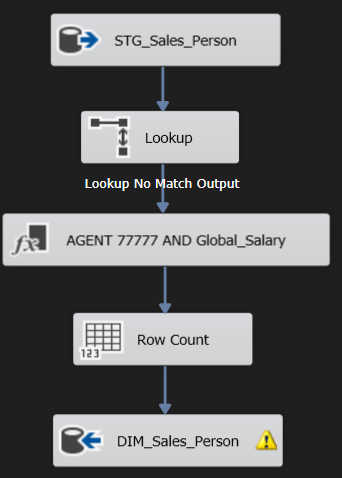
**• DM\_****Sales\_Person Package**

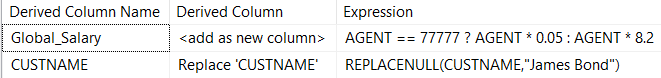
Loading data into DIM\_Sales\_Person is performed incrementally using the Lookup transformation, which compares records to existing entries in DIM\_Sales\_Person. The table is not truncated during this process.

****

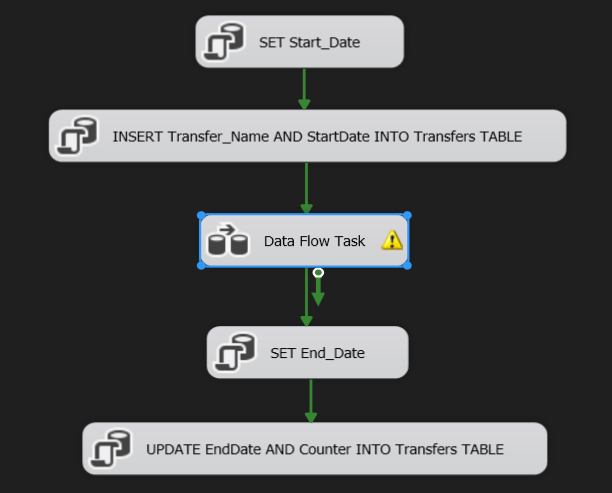
In the Data Flow Task:

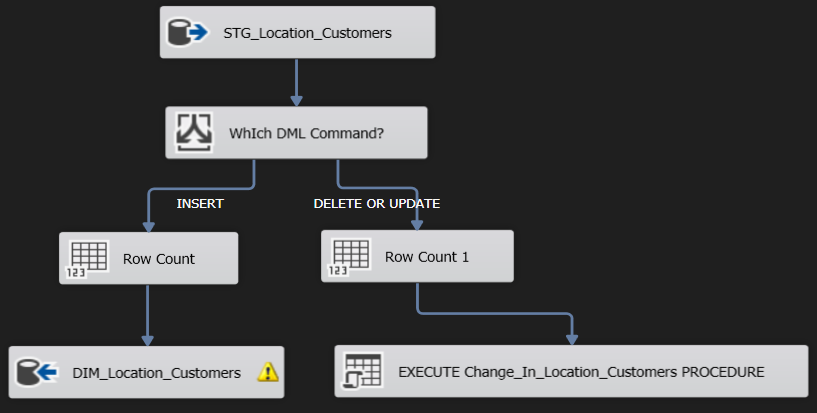
I used the Derived Column Transform to calculate the global salary for each worker and to assign a name to agent 77777.





**• DM\_Location\_Customers Package**

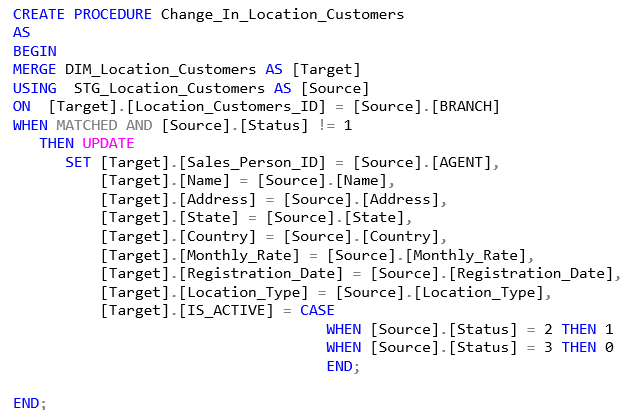
Loading data into DIM\_Location\_Customers is performed using CDC with the Split Condition transformation. When Status = 1, a regular insert is performed; otherwise, it is an update or delete. I used a stored procedure to execute this process. The table is not truncated during this operation.

תמונה שמכילה טקסט, גופן, צילום מסך, קו

התיאור נוצר באופן אוטומטיIn the Data Flow Task:

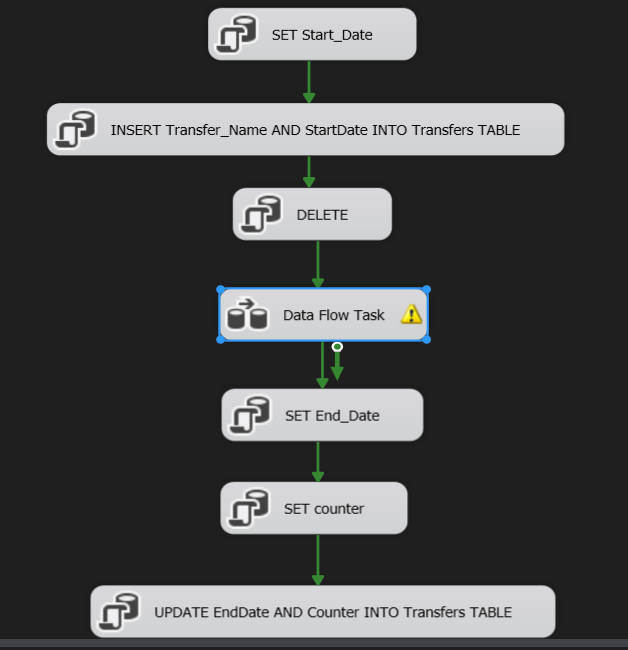




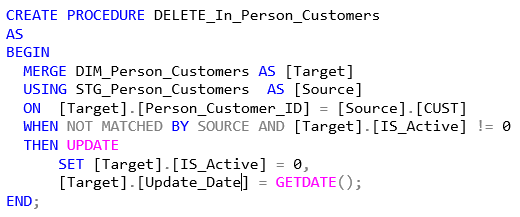


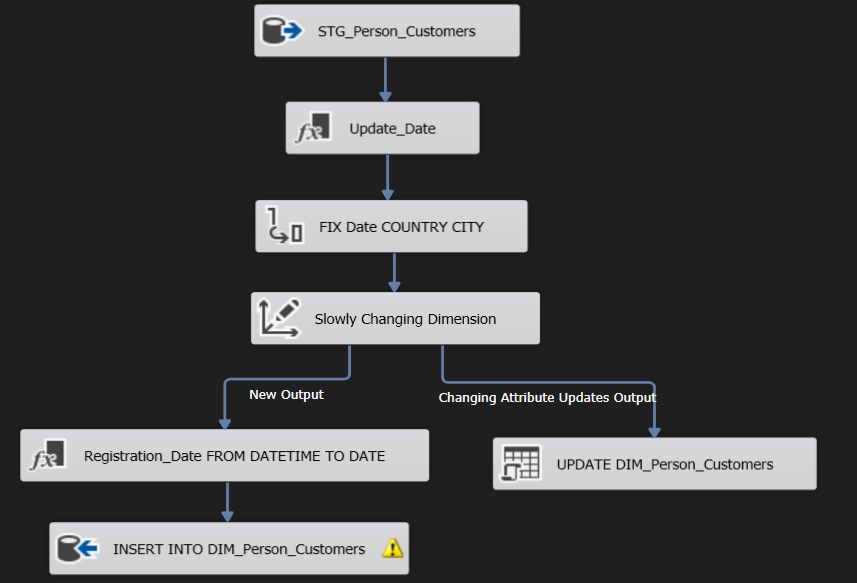
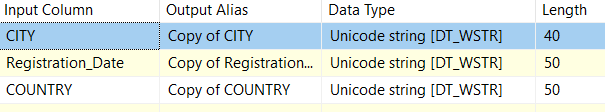
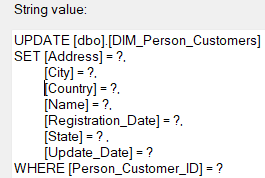
**• DM\_Person\_Customers Package**

Loading data into DIM\_Person\_Customers is performed using an incremental load with the Slowly Changing Dimension transformation (change type: Changing Attribute). To deleted entries, I created a stored procedure that updates the IS\_Active column in DIM\_Person\_Customers to 0.



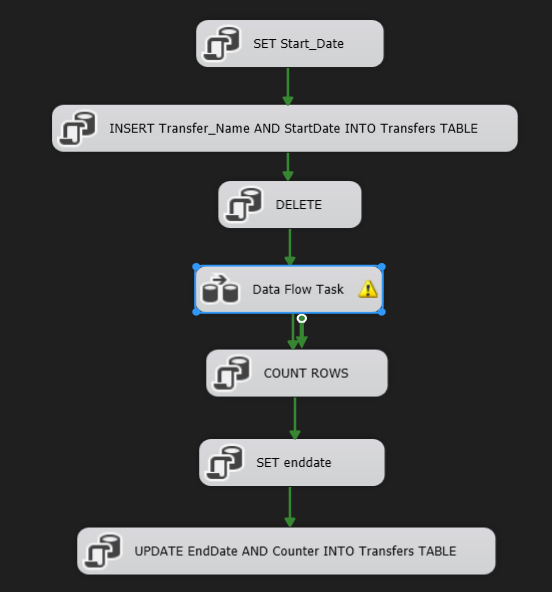


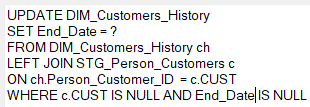


In the Data Flow Task, I created the Update\_Date column using the Derived Column transformation and converted the Registration\_Date, City, and Country column types using the Data Conversion transformation. In addition, Update statement created to update the new values in DIM\_Person\_Customers

**• DM\_Customers\_History Package**

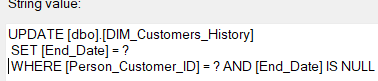
Loading data into DIM\_Customers\_History is performed using an incremental load with the Slowly Changing Dimension transformation (change type: Historical Attribute). To deleted entries, I created a update statement that updates the End\_Date column in DIM\_Customers\_History to current datetime.

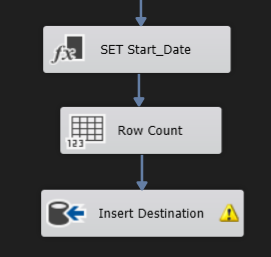
****



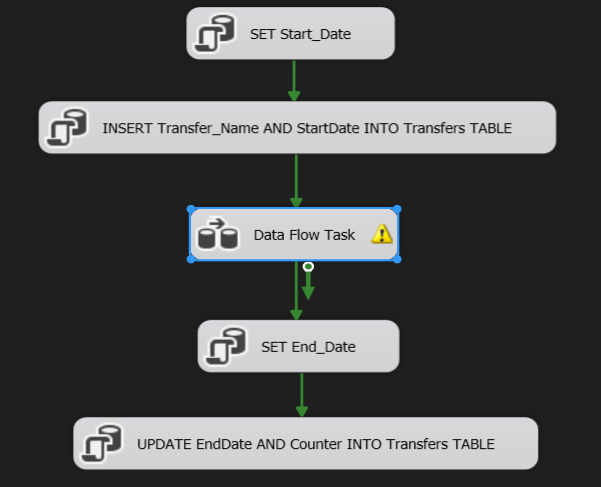
תמונה שמכילה טקסט, צילום מסך, גופן, קו

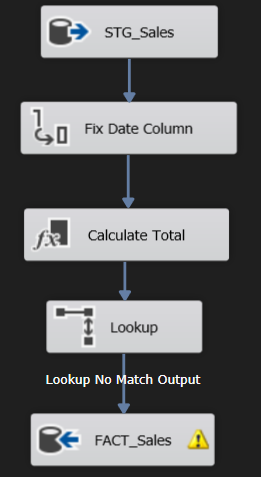
התיאור נוצר באופן אוטומטיIn the Data Flow Task, I created the End\_Date and Start\_Date columns using the Derived Column transformation and converted the Registration\_Date, City, and Country column types using the Data Conversion transformation. In addition, Update statement created to set the End\_Date column in the old values.

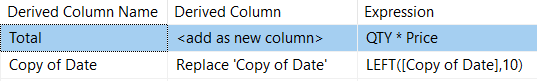
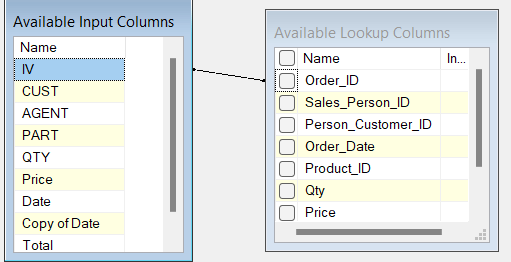




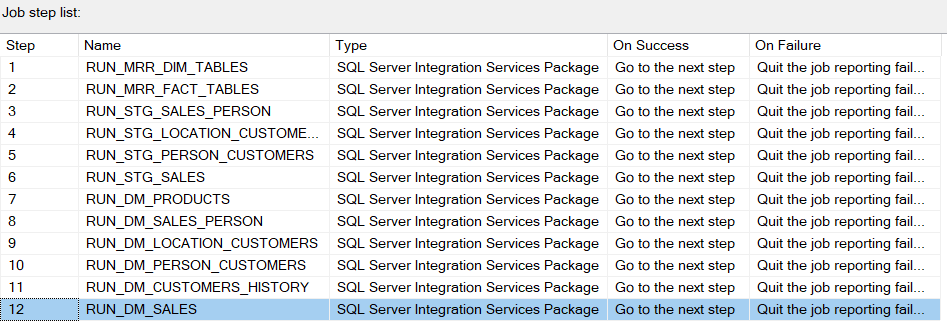
**• DM\_Sales Package**

****Loading data into FACT\_Sales is performed incrementally using the Lookup transformation, which compares records to existing entries in FACT\_Sales. The table is not truncated during this process.I calculated the Total column using the Derived Column transformation and converted the Date column type using the Data Conversion transformation**.**

****In the Data Flow Task:

****

**• Automatic Processing**

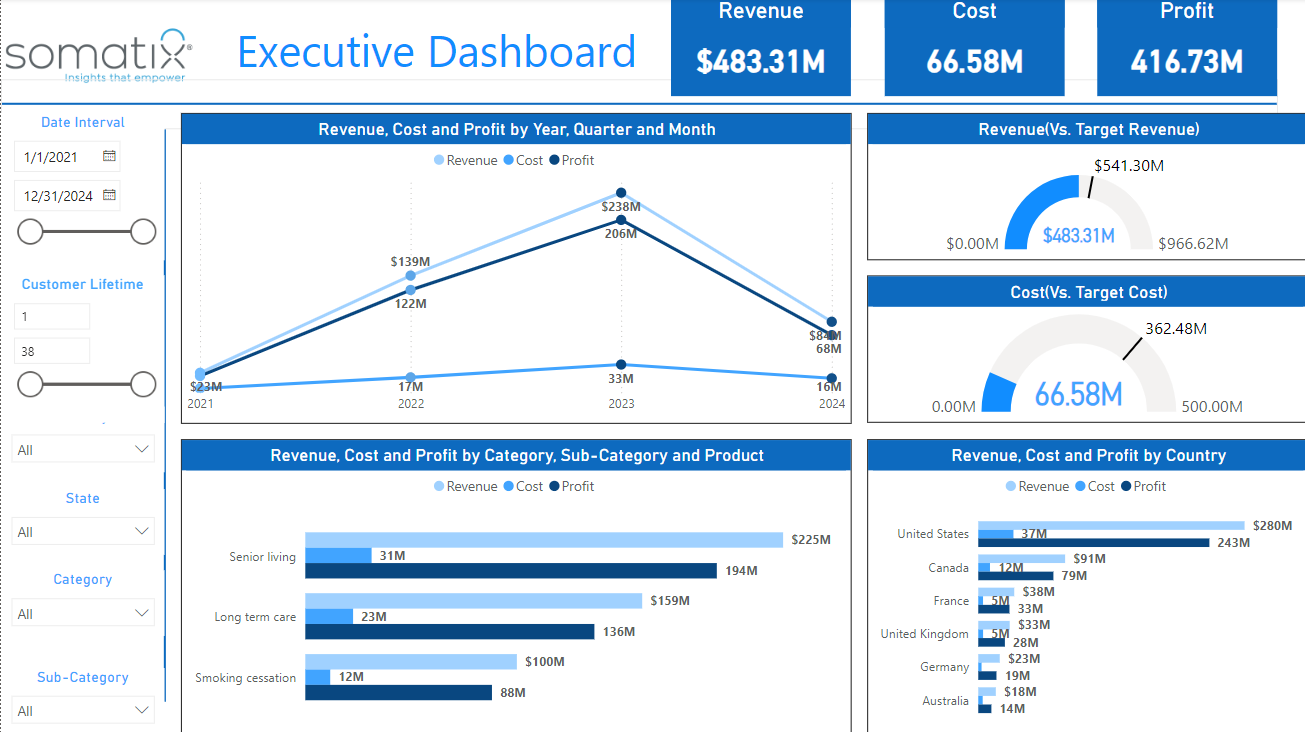
The data is automatically refreshed daily at 00:00:00 using SQL Agent jobs, the first step executed is the RUN\_MRR\_DIM\_TABLES, the second step executed is EUN\_MRR\_FACT\_TABLES and so on.

4.3. Visualization in Power BI:

**• Reports**

The project includes 3 reports: Executive Dashboard, Sales Analysis, Customer Analysis.

**Executive Dashboard:**

This dashboard provides a comprehensive overview of the company's financial performance, highlighting key metrics and performance against targets.

Key Metrics

* Revenue
* Cost
* Profit

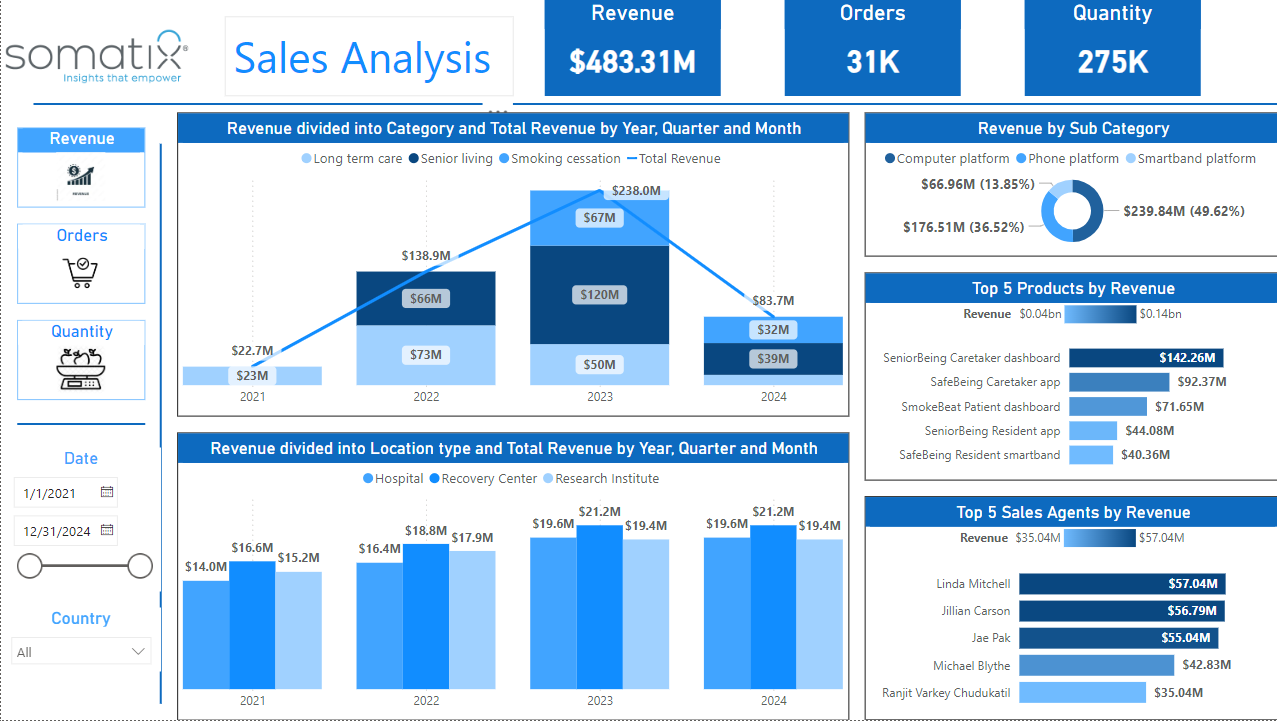
Visuals

* Revenue, Cost, and Profit by Year, Quarter, and Month
* Revenue (Vs. Target Revenue)
* Cost (Vs. Target Cost)
* Revenue, Cost, and Profit by Category, Sub-Category, and Product
* Revenue, Cost, and Profit by Country

Filters

* Date Interval: Allows selection of data between 1/1/2021 and 12/31/2024.
* Customer Lifetime: Range from 1 to 38.
* State
* Category
* Sub-Category

**Sales Analysis:**

This dashboard provides an in-depth analysis of the company’s sales performance, broken down by various categories, products, locations, and sales agents. It highlights key revenue figures and top-performing segments.

Same for al States(Revenue State, Orders State, Quantity State)

Key Metrics

* Revenue
* Orders
* Quantity

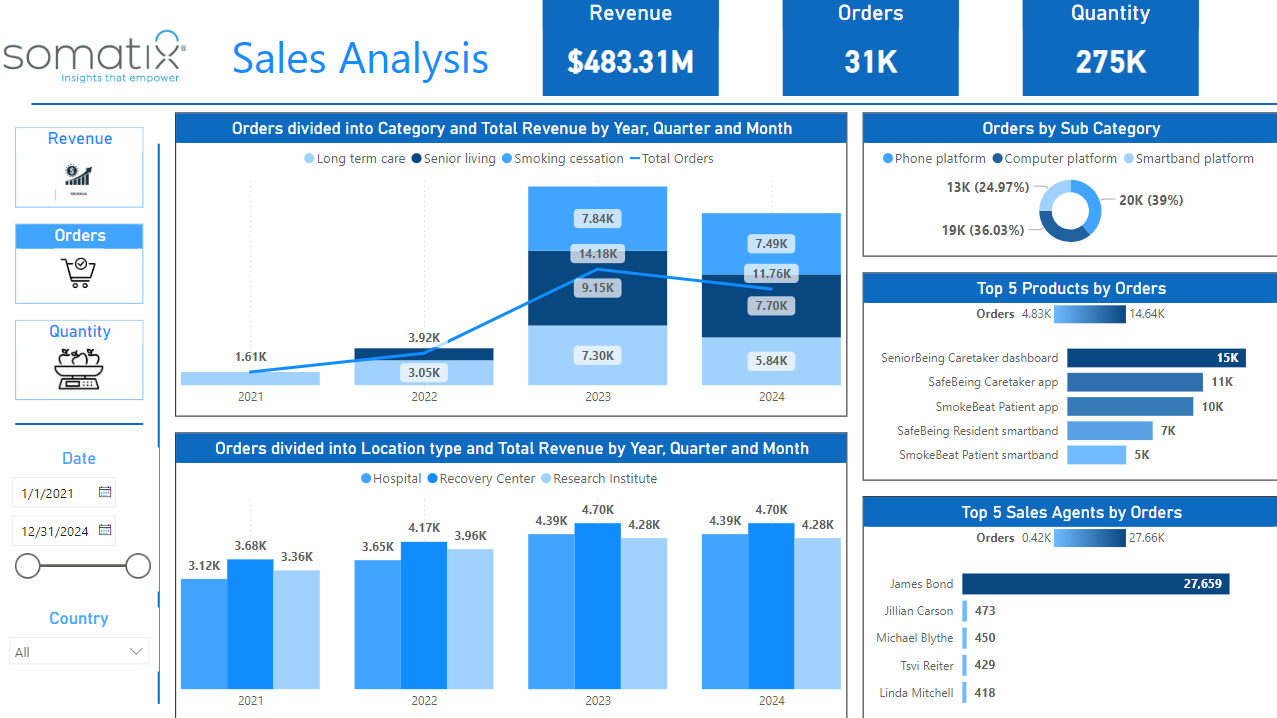
Filters

* Date: From 1/1/2021 to 12/31/2024
* Country

Revenue State

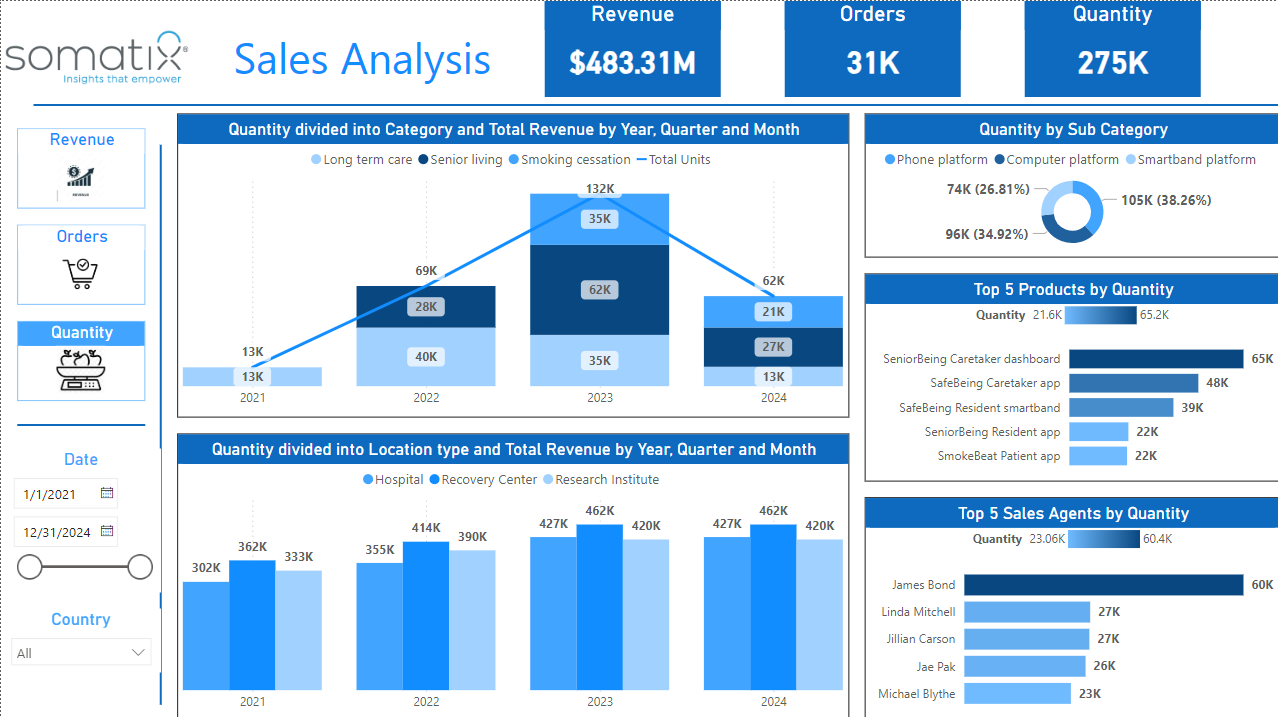
Visuals

* Revenue divided into Category and Total Revenue by Year, Quarter and Month
* Revenue by Sub Category
* Top 5 Products by Revenue
* Revenue divided into Location type and Total Revenue by Year, Quarter and Month
* Top 5 Sales Agents by Revenue

Orders State

Visuals

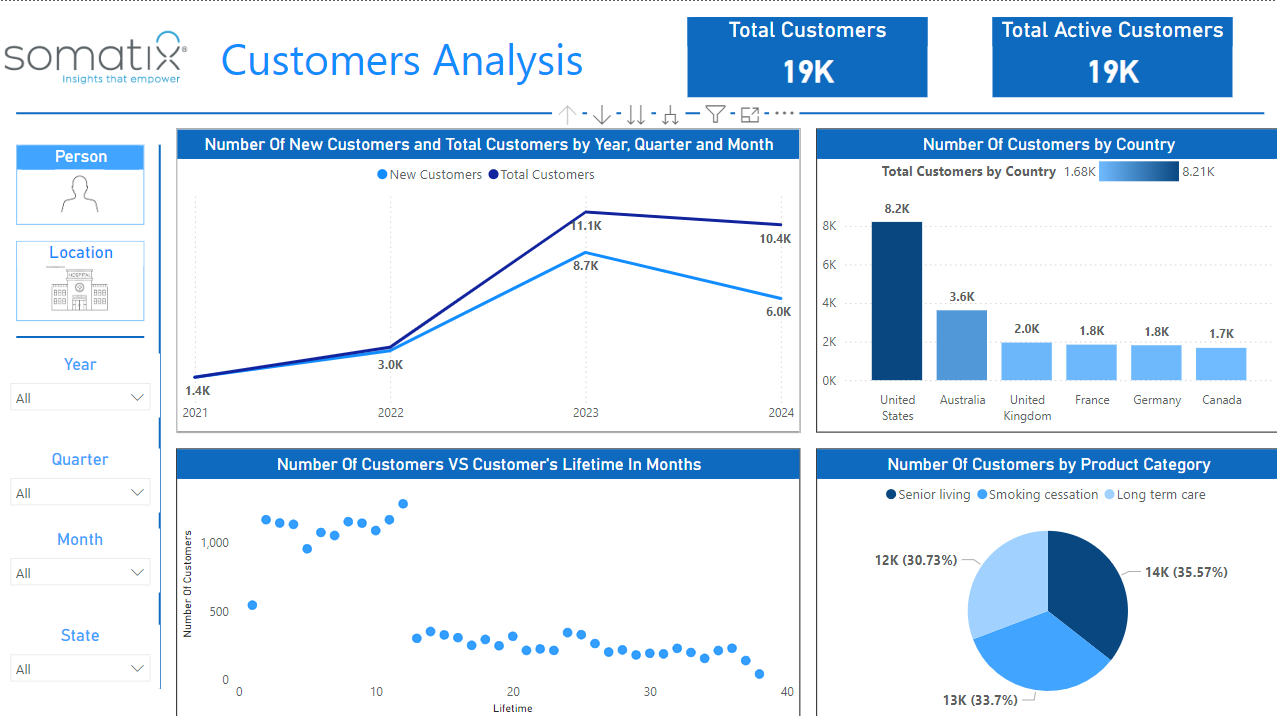
* Orders divided into Category and Total Revenue by Year, Quarter and Month
* Orders by Sub Category
* Top 5 Products by Orders
* Orders divided into Location type and Total Revenue by Year, Quarter and Month
* Top 5 Sales Agents by Orders

Quantity State

Visuals

* Quantity divided into Category and Total Revenue by Year, Quarter and Month
* Quantity by Sub Category
* Top 5 Products by Quantity
* Quantity divided into Location type and Total Revenue by Year, Quarter and Month
* Top 5 Sales Agents by Quantity

**Customers Analysis:**

This dashboard provides insights into the customer base, including the number of new and total customers over time, customer distribution by country, and product category preferences. It also includes an analysis of customer lifetime.

Same for al States(Person State, Location State)

Filters

* Year
* Quarter
* Month
* State

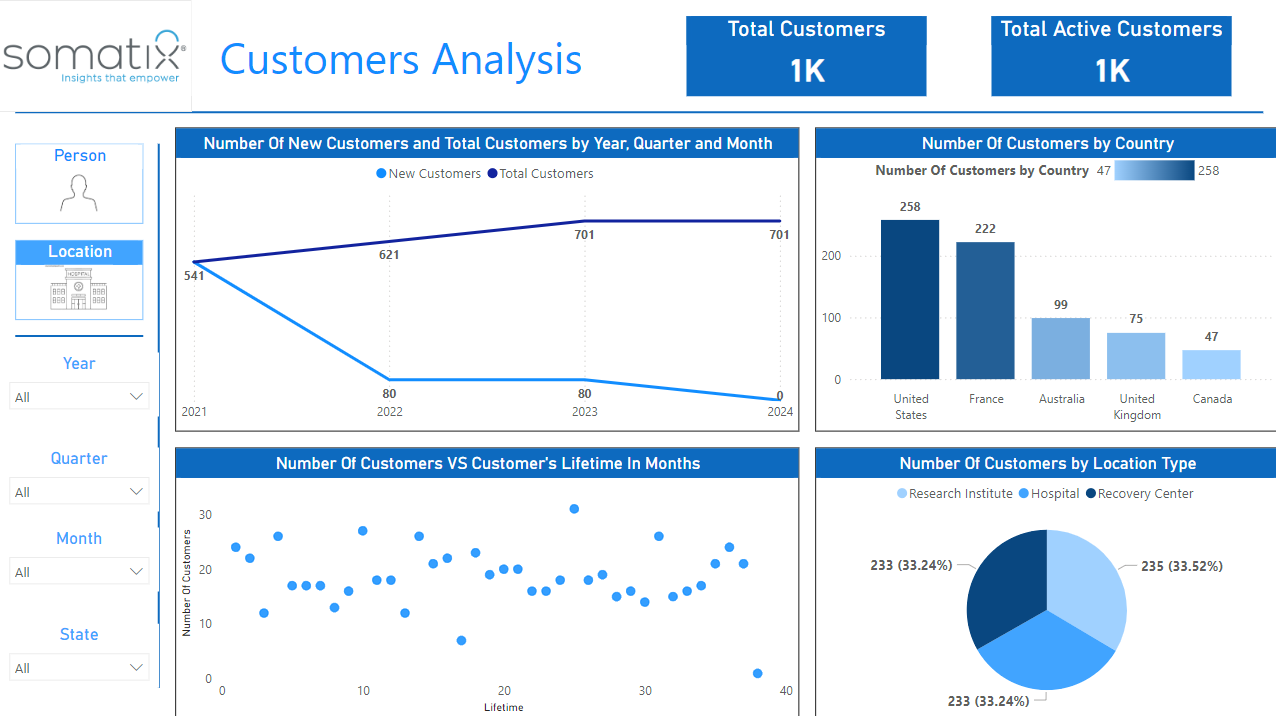
Person State

Key Metrics

* Total Customers
* Total Active Customers

Visuals

* Number of New Customers and Total Customers by Year, Quarter, and Month
* Number of Customers by Country
* Number of Customers vs. Customer's Lifetime in Months
* Number of Customers by Product Category



Location State

Key Metrics

* Total Customers
* Total Active Customers

Visuals

* Number of New Customers and Total Customers by Year, Quarter, and Month
* Number of Customers by Country
* Number of Customers vs. Customer's Lifetime in Months
* Number of Customers by Location Type