Project: Sales Data Ingestion and Analytics Pipeline with Azure SQL Database





Submitted By: Narendran G

(2347132)

**Design Document**

**Proposed Solution Overview**

The project implements a sales data pipeline that ingests, processes, and visualizes data using SSMS Tool connecting Via Azure SQL Database and Power BI. The pipeline is designed to transform raw sales data from CSV format into meaningful insights for decision-making.

**Choice of Data Storage**

**Azure SQL Database** has been selected as the data storage solution for the following reasons:

* **Structured Data**: The sales dataset is structured, with clear relationships between tables (e.g., customer details, products, sales transactions). Azure SQL is optimized for handling relational data.
* **Performance and Scalability**: Azure SQL Database offers automated scaling and performance tuning, ensuring efficient querying for large datasets.
* **Security and Compliance**: Azure SQL Database provides advanced security features such as data encryption, threat detection, and role-based access control, ensuring the confidentiality and integrity of the data.

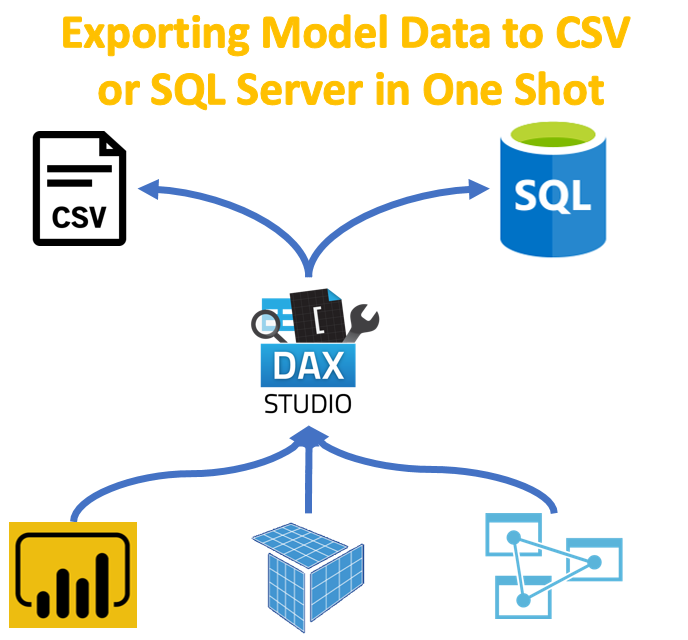
While **ADLS** is effective for storing unstructured or semi-structured data, it is less suitable for purely relational datasets like sales data, which benefit from SQL-like querying capabilities.

**Data Flow Architecture**

The pipeline comprises the following steps:

1. **Data Source**: CSV file containing raw sales data.
2. **Ingestion**: Data imported into Azure SQL using SSMS.
3. **Processing**: Structure and queries prepared in Azure SQL.
4. **Visualization**: Power BI dashboard connected to Azure SQL for creating interactive reports.

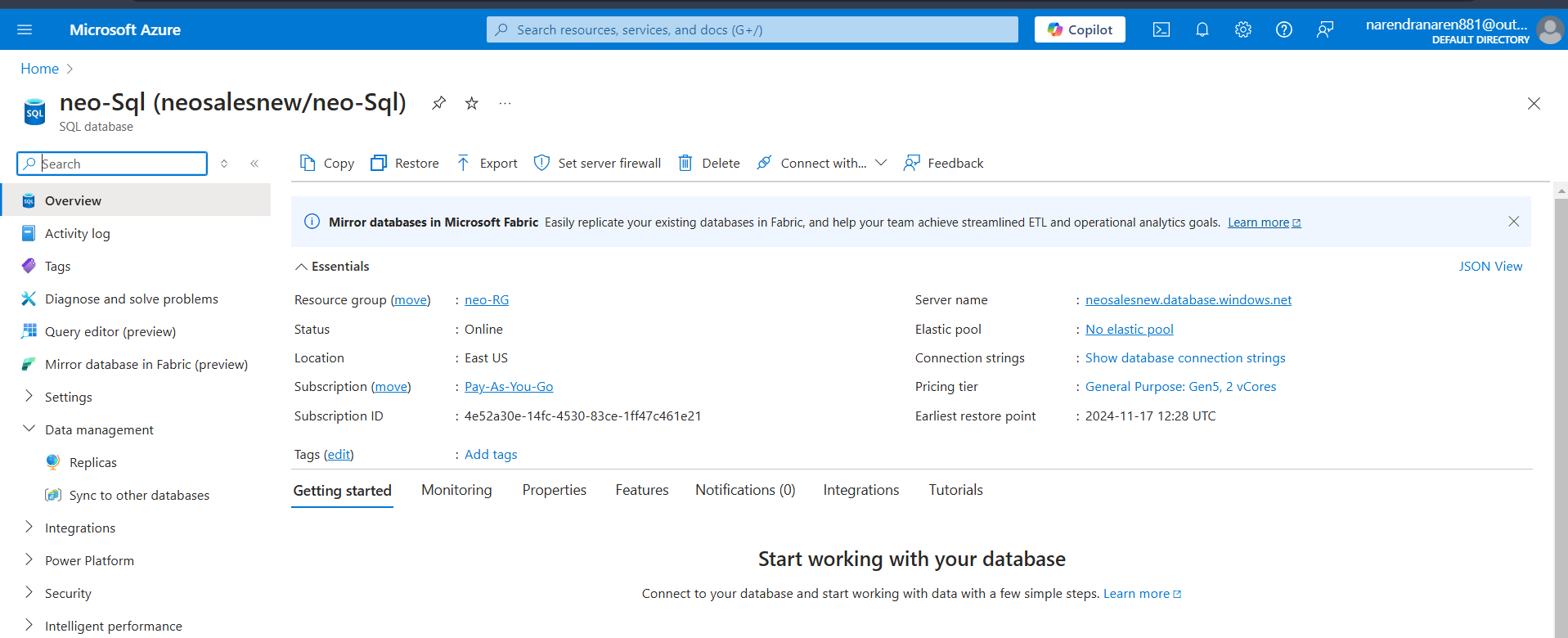
**Architecture Diagram**

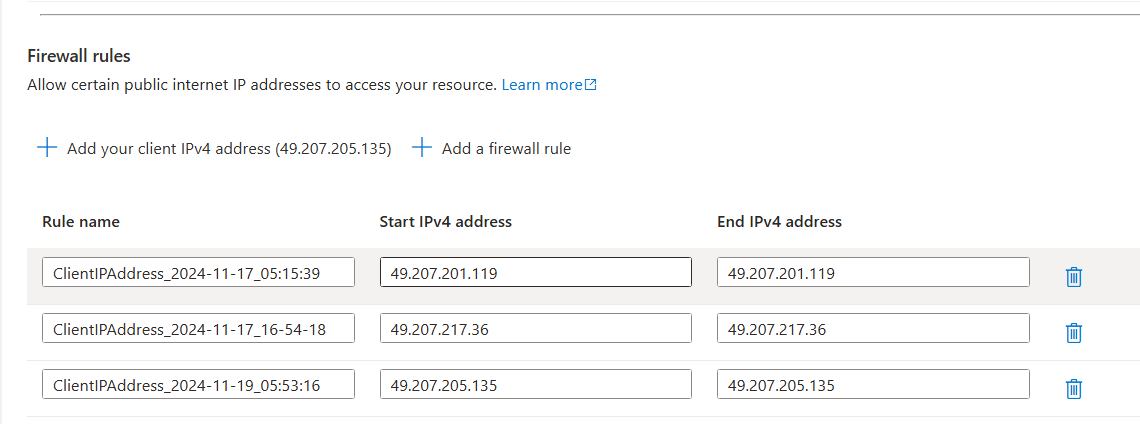


**Database Creation and Structure**

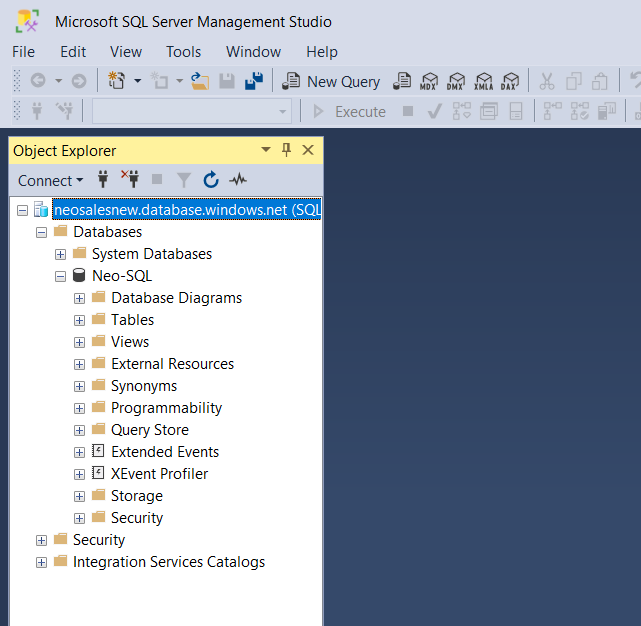
Using SSMS, a database named **NEO-SQL** was created to store and process sales data.

**Azure SQL Database:**

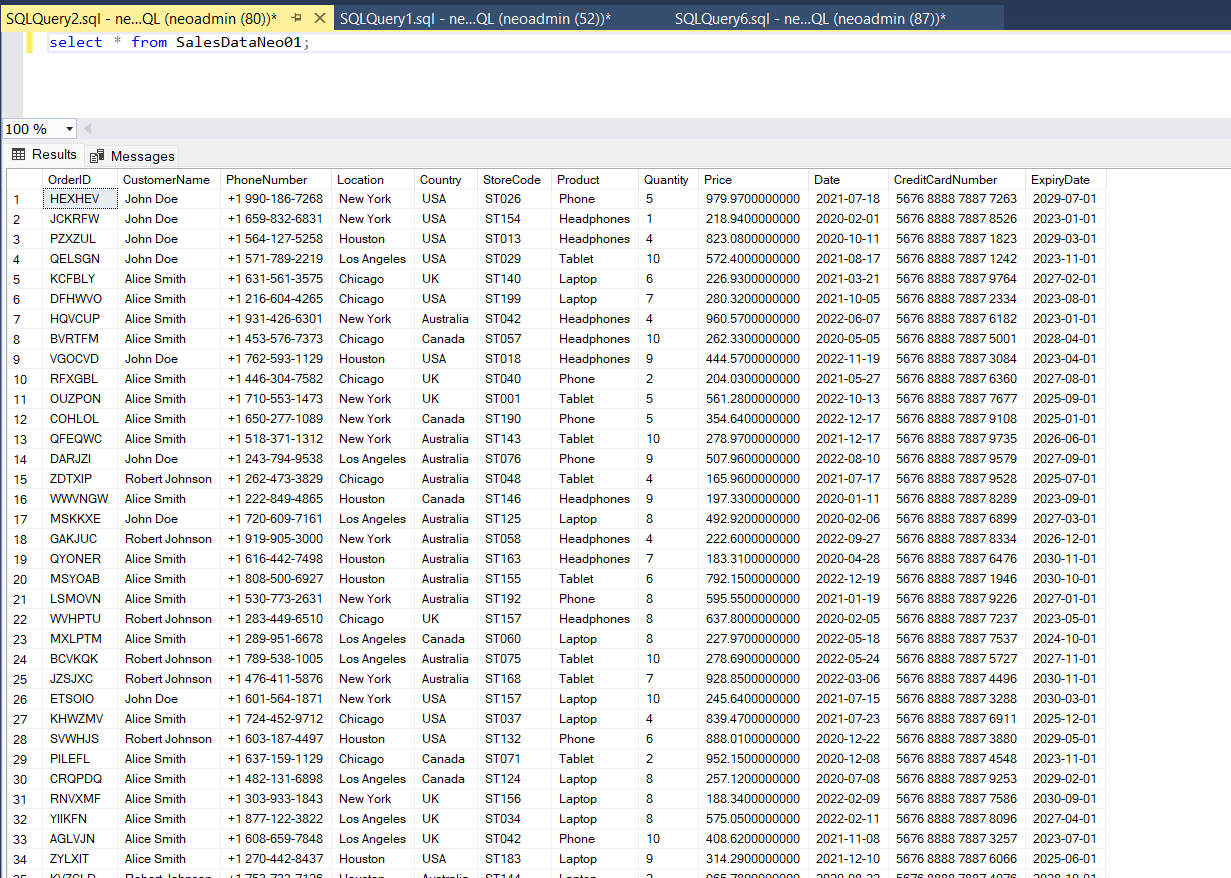


**Added Firewall (Allowed Access only for my Pc):**  


**SSMS Tools:**



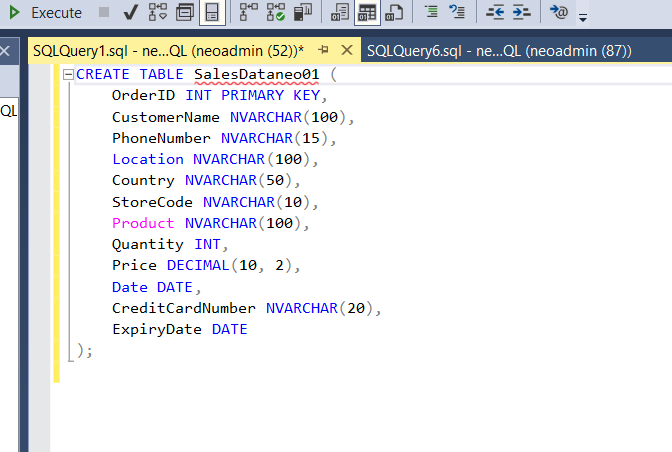
Imported Table As per Structure:

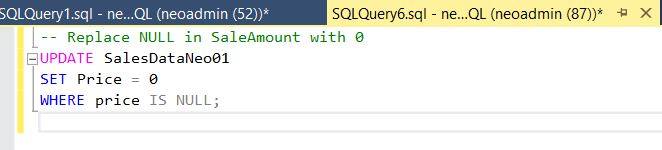


**Structure of the Data:**

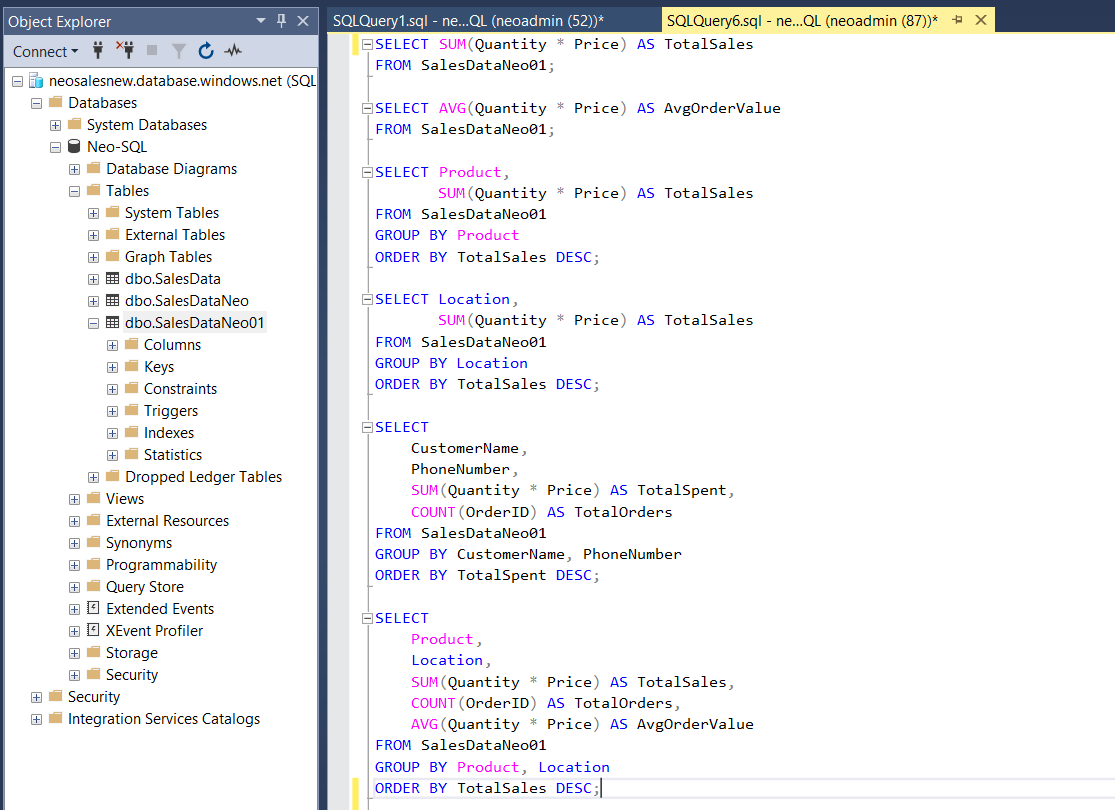
**Dataset Description**

* + OrderID (Integer): Unique identifier for orders.
  + CustomerName (String): Name of the customer.
  + PhoneNumber (String): Customer’s phone number.
  + Location (String): City or region of the customer.
  + Country (String): Customer’s country.
  + StoreCode (String): Code identifying the store.
  + Product (String): Name of the purchased product.
  + Quantity (Integer): Number of units sold.
  + Price (Float): Price per unit.
  + Date (Date): Date of the transaction.
  + CreditCardNumber (String): Masked credit card number.
  + ExpiryDate (Date): Expiry date of the card.
* Data Types: Specify SQL data types (e.g., INT, VARCHAR, FLOAT, DATE).

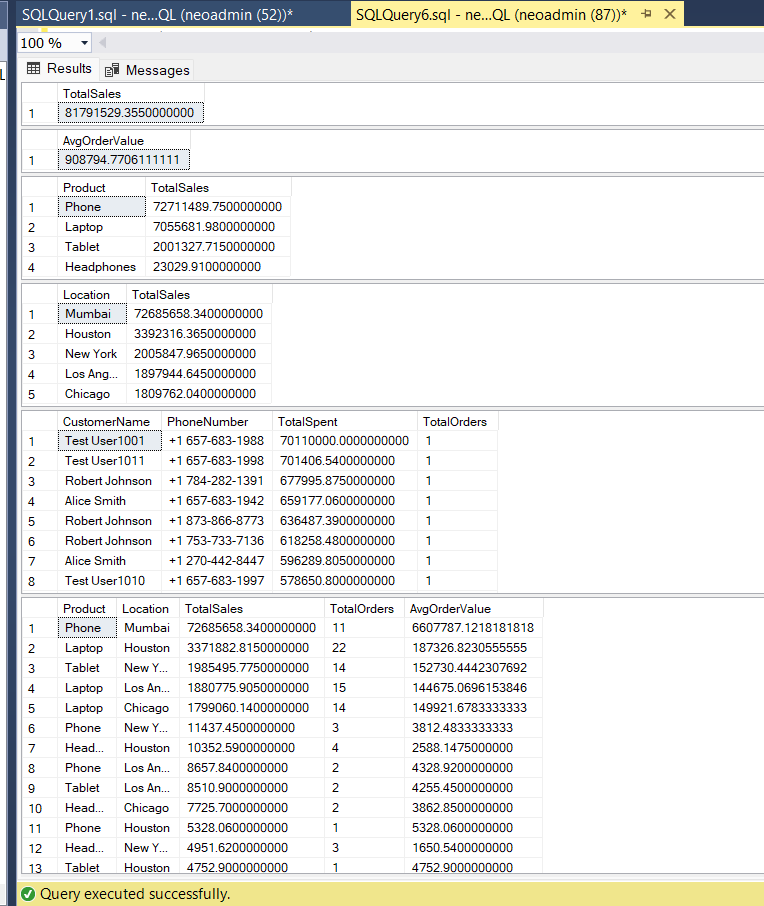




SQL Queries to perform the cleaning and further processing



Outputs:



**Power BI Overview for Sales Analysis**

Power BI is a powerful business analytics tool from Microsoft that enables users to visualize and share insights from their data. It allows for the creation of interactive dashboards, reports, and visualizations, making it easier to understand and present business data. In the context of sales analysis, Power BI is ideal for transforming raw transactional data into meaningful insights and visualizations.

**Steps:**

1.Open Power BI tool and Access data via SQL database

2.Transform data

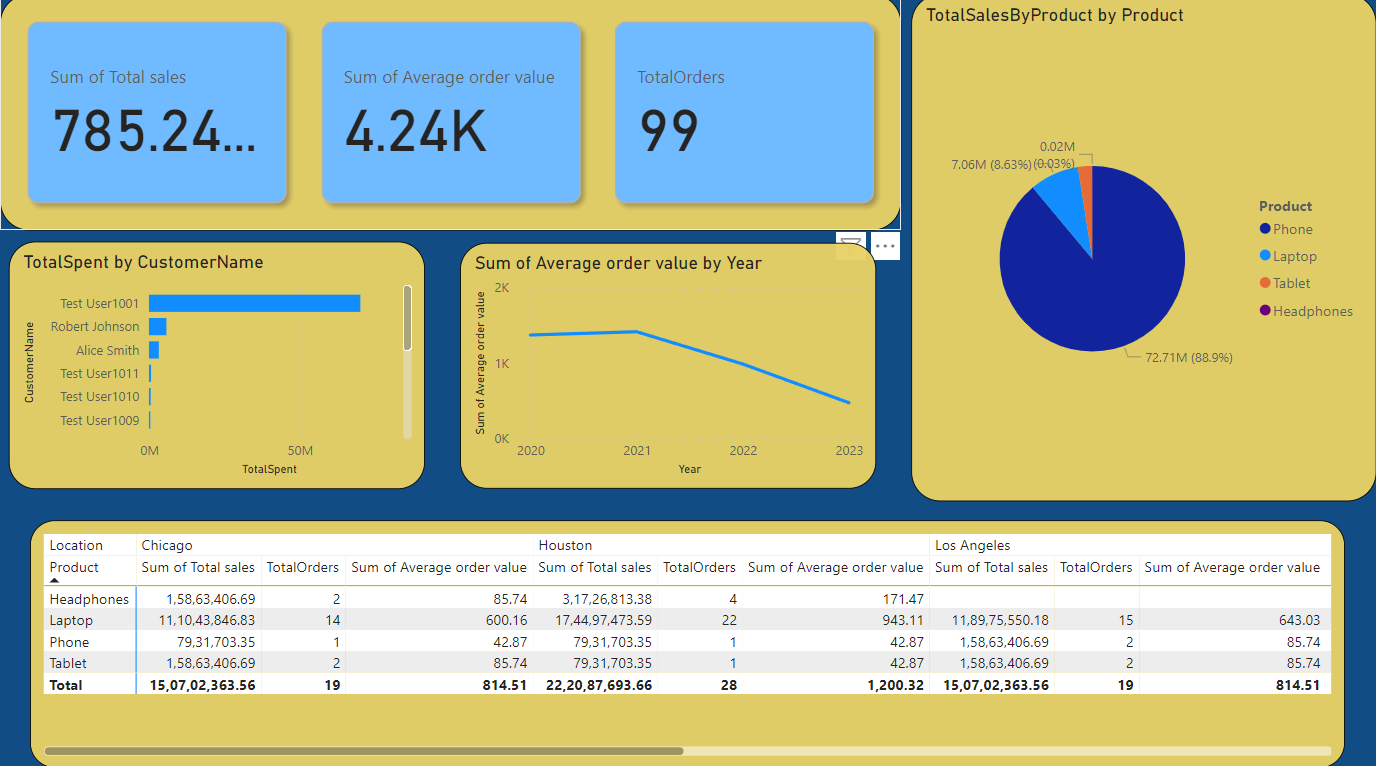
3.Create the Required measures for performing Task

4.Create the Graphs and Charts according to Requirements

5.Create a Dashboard

Dashboard:





**Thank YOU**