RETAIL INVENTORY MANAGEMENT AND FORECASTING

PROJECT

OVERVIEW

1

Our Team

2

About Project

3

Objectives

4

Project Steps

1 OUR TEAM

MOHAMED MAGDY

Database

ABDULLAH HATEM

ETL

MAHMOUD SAMIR

ETL

AHMED YASSER

ML on Azure

MANSOUR ELASHRY

Dashboard

2

ABOUT PROJECT

This project focuses on delivering a comprehensive analysis of business data to improve decision-making through a combination of database management, data transformation, machine learning, and data visualization.

3 OBJECTIVES

01 Efficient data storage and retrieval through well-structured SQL tables.

02 Streamlined data processing with ETL workflows using SSIS.

03 Predictive analytics powered by machine learning models in Azure.

04 Clear, actionable insights provided through an interactive Power BI dashboard.

4

PROJECT STEPS

STEP 1

Creating Tables with SQL

SQL was used to create structured tables that store customer, order, and product data, ensuring easy data management and querying.

STEP 2

ETL using SSIS

Data from various sources was extracted, transformed, and loaded (ETL) into a centralized data warehouse using SQL Server Integration Services (SSIS), ensuring the data is clean and ready for analysis.

STEP 3

Machine Learning with Python in Azure

Python scripts were implemented in Azure to build machine learning models, enabling predictive analysis and deeper insights into business trends.

STEP 4

Power BI Dashboard

Finally, a Power BI dashboard was created to visualize key metrics and performance indicators, making data easily accessible and understandable for stakeholders.

03 STEP 1

Database Creation

1) Imported the dataset as CSV file to SQL Server enviroment

2) Normalized the dataset into smaller tables based on relevant detailes

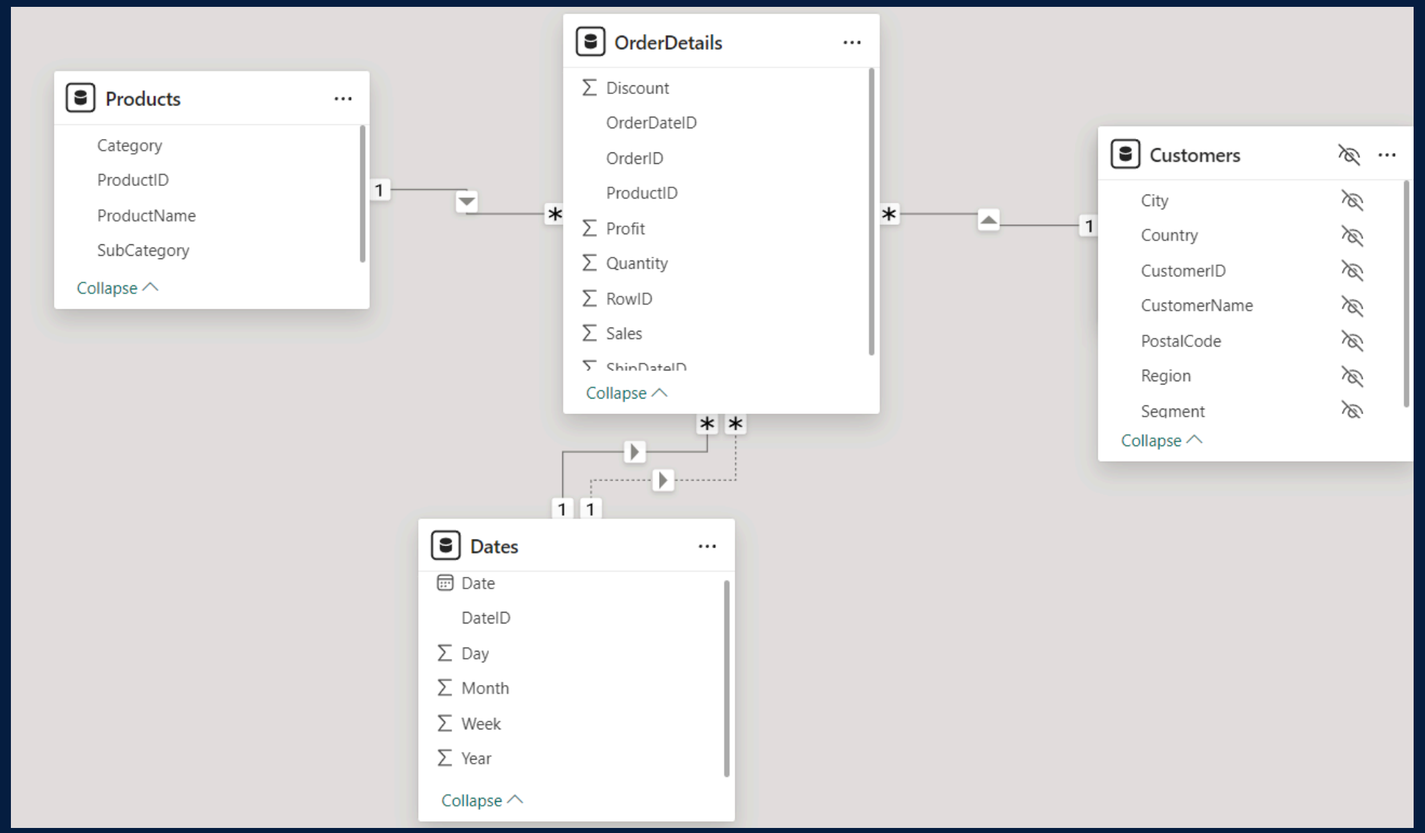
3) Used insert based on select statement to avoid the hassel of bulk insertion

```
CREATE TABLE Customers (
    CustomerID NVARCHAR(10) PRIMARY KEY,
    CustomerName NVARCHAR(255),
    Segment NVARCHAR(50),
    Country NVARCHAR(50),
    City NVARCHAR(100),
    Region NVARCHAR(50),
    PostalCode NVARCHAR(20)
);
```

```
);
INSERT INTO Customers (CustomerID, CustomerName, Segment, Country, City, Region, PostalCode)
SELECT
    Customer_ID,
    Customer_Name,
    Segment,
    Country,
    City,
    Region,
    ISNULL(Postal_Code, '00000') -- Ensure no NULLs in PostalCode
FROM superstore_cleaned;
```

03 STEP 1

Database Schema



03 STEP 1

Database Exploration

Exploration.sql -...(MAGDY\muham (52))

```sql
--What is the total sales amount for each product in the Pr
SELECT p.ProductName, SUM(od.Sales) AS TotalSales
FROM Products p
JOIN OrderDetails od ON p.ProductID = od.ProductID
GROUP BY p.ProductName;

--How many orders were placed by each customer?
SELECT c.CustomerName, COUNT(o.OrderID) AS NumberOfOrders
FROM Customers c
LEFT JOIN Orders o ON c.CustomerID = o.CustomerID
GROUP BY c.CustomerName;

--What are the top 5 products by total profit?
SELECT p.ProductName, SUM(od.Profit) AS TotalProfit
FROM Products p
JOIN OrderDetails od ON p.ProductID = od.ProductID
GROUP BY p.ProductName
ORDER BY TotalProfit DESC
OFFSET 0 ROWS FETCH NEXT 5 ROWS ONLY;
```

Results | Messages

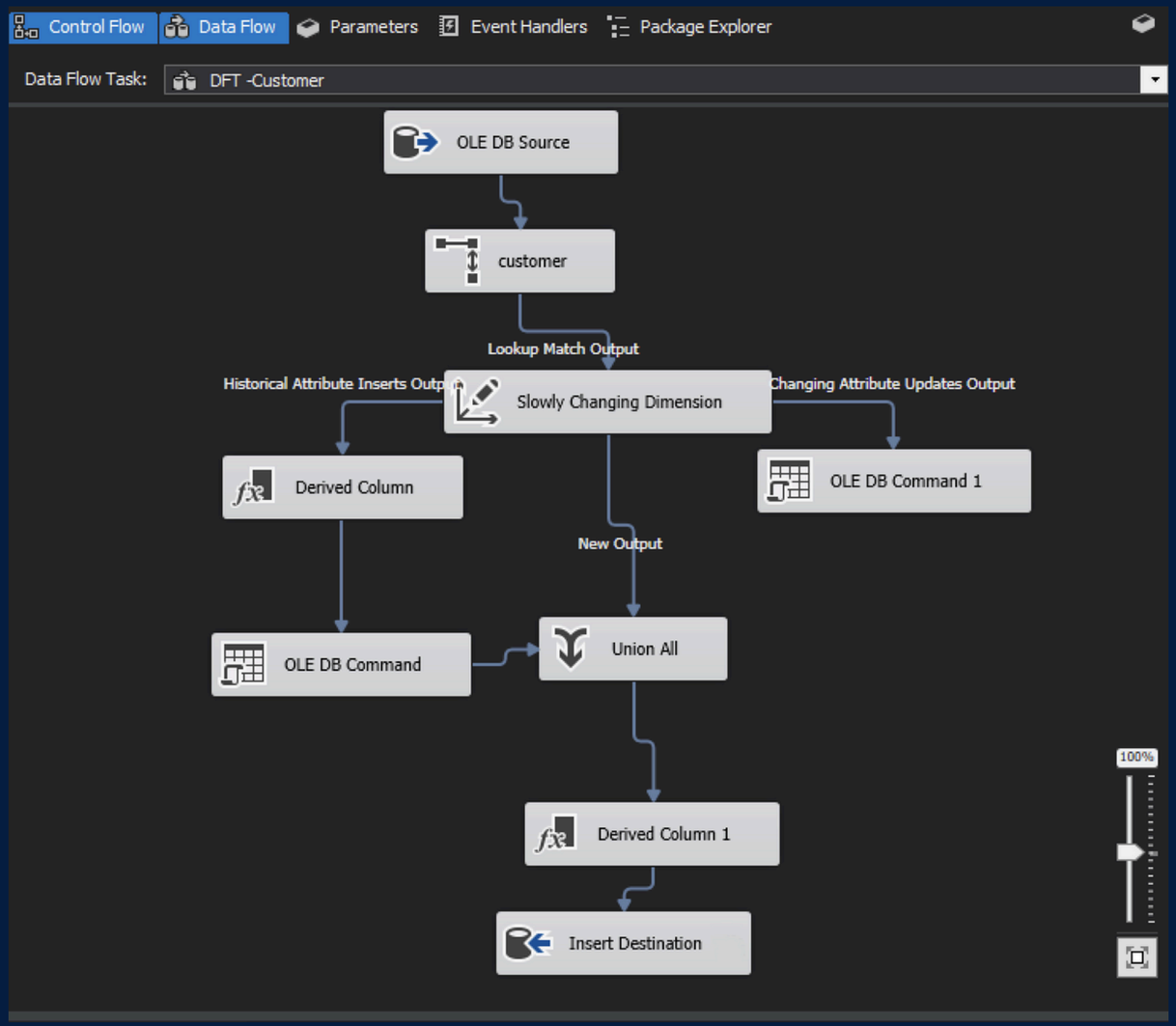
	ProductName	TotalProfit
1	Canon imageCLASS 2200 Advanced Copier	3919.99
2	Fellowes PB500 Electric Punch Plastic Comb Binding Mac...	2414.89
3	Ativa V4110MDD Micro-Cut Shredder	2400.97
4	GBC DocuBind TL300 Electric Binding System	2107.93
5	Canon imageCLASS MF7460 Monochrome Digital Laser ...	1995.99

Top 5 products by total profit

03 STEP 2

ETL

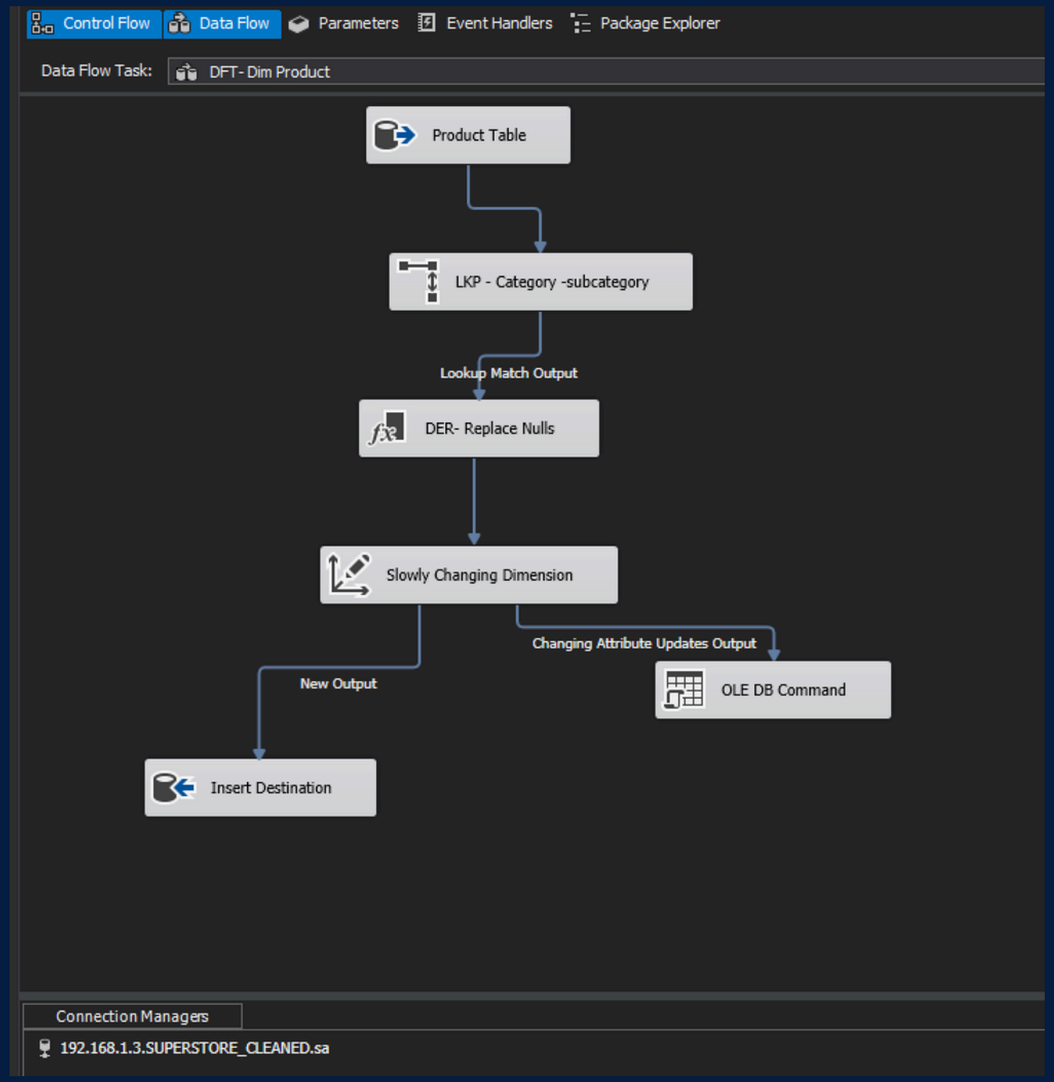
Dim custumer packge
Cleaning Data & Selecting
Data types



03 STEP 2

ETL

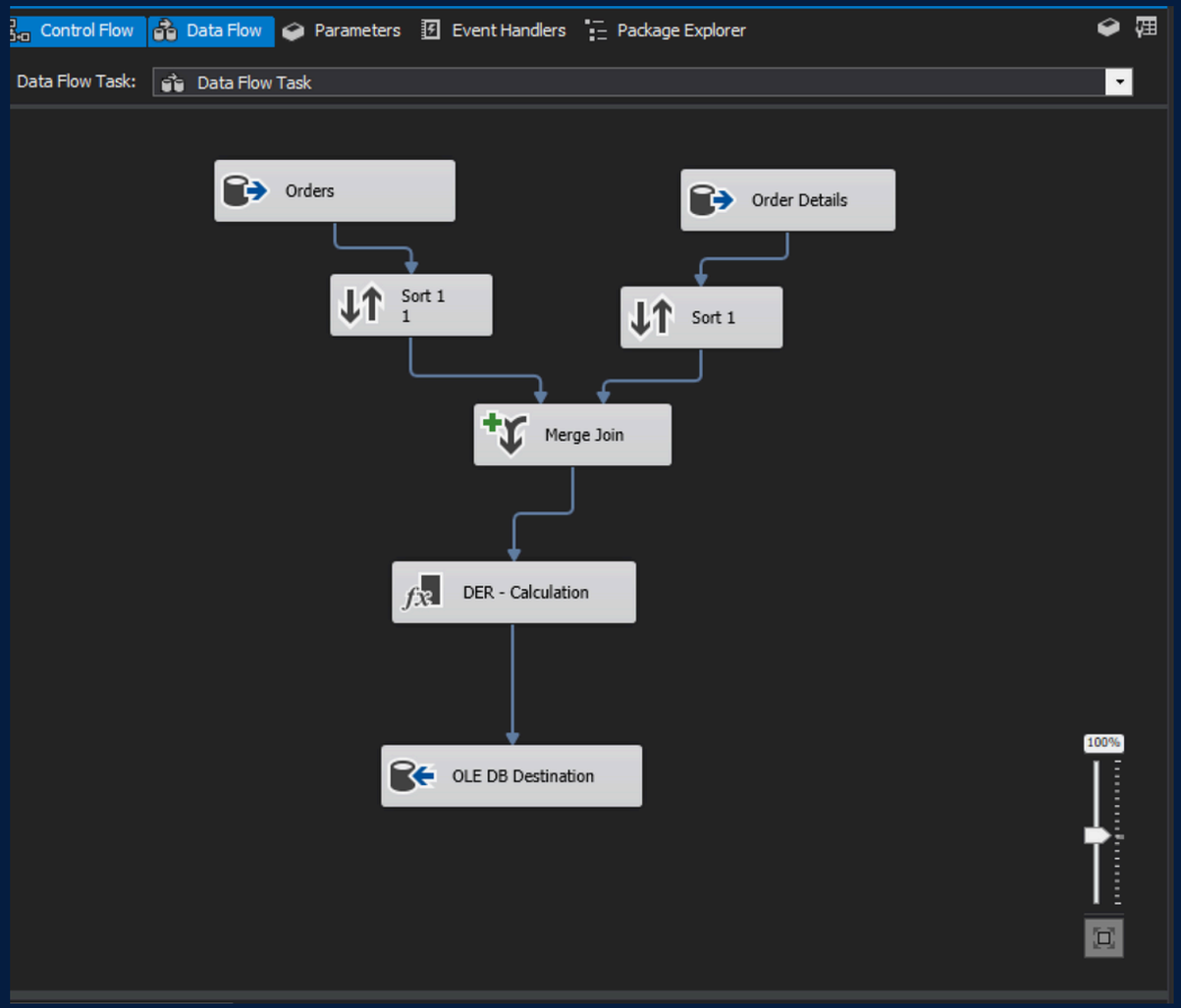
Dim product packge
Cleaning Data & Selecting Data types.



03 STEP 2

ETL

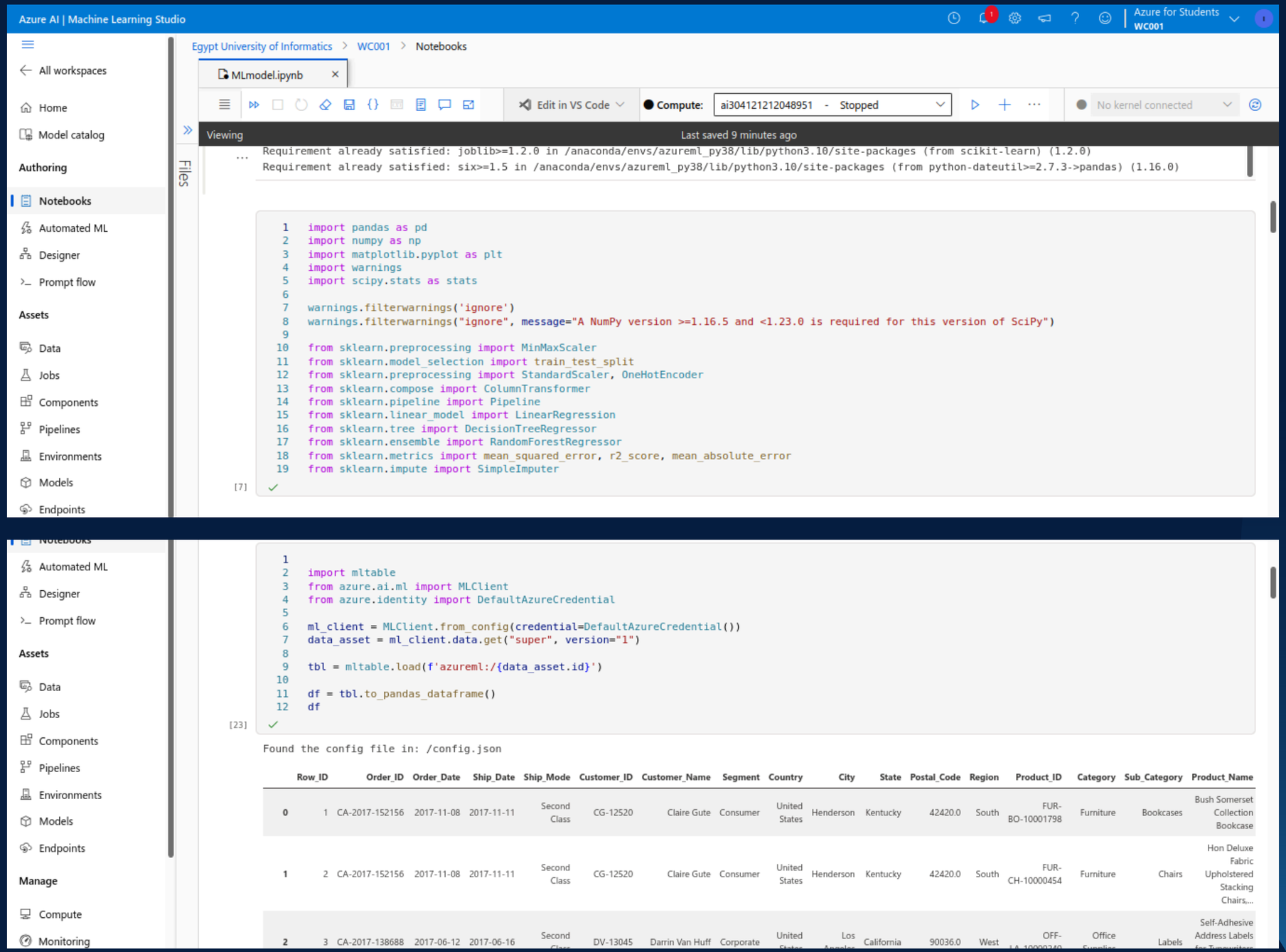
Fact sales package
Merge & Calculation
process



03 STEP 3

ML on Azure

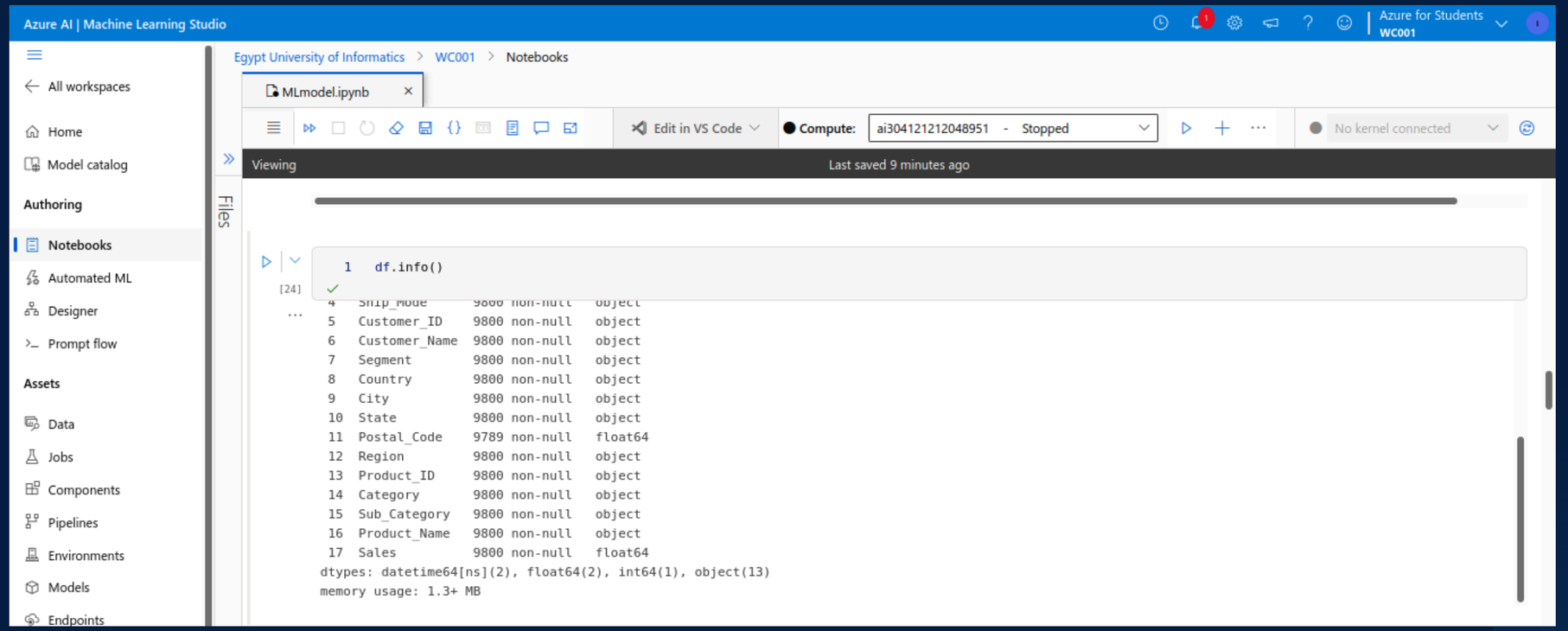
- import library
- show dataset

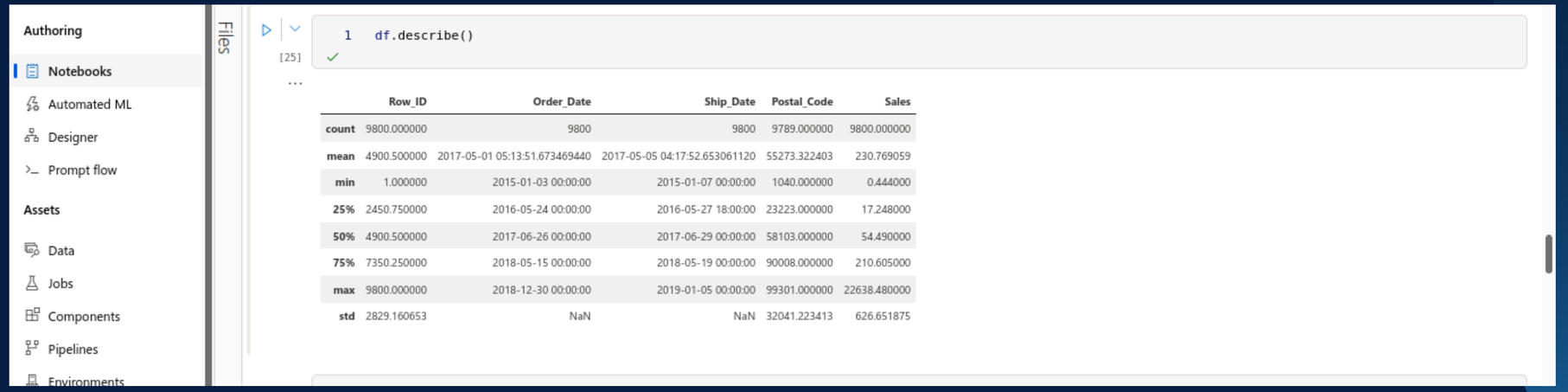


03 STEP 3

ML on Azure

- data info



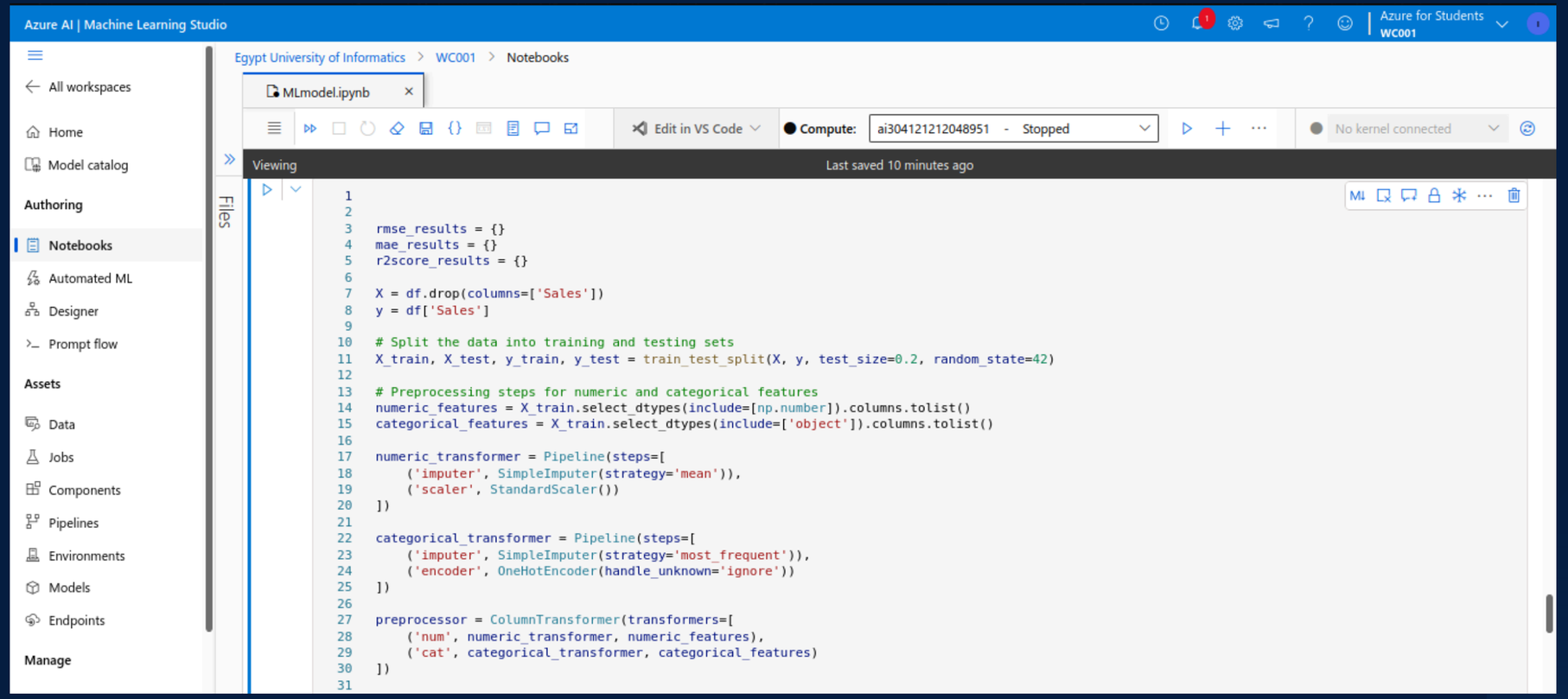


	Row_ID	Order_Date	Ship_Date	Postal_Code	Sales
count	9800.000000	9800	9800	9789.000000	9800.000000
mean	4900.500000	2017-05-01 05:13:51.673469440	2017-05-05 04:17:52.653061120	55273.322403	230.769059
min	1.000000	2015-01-03 00:00:00	2015-01-07 00:00:00	1040.000000	0.444000
25%	2450.750000	2016-05-24 00:00:00	2016-05-27 18:00:00	23223.000000	17.248000
50%	4900.500000	2017-06-26 00:00:00	2017-06-29 00:00:00	58103.000000	54.490000
75%	7350.250000	2018-05-15 00:00:00	2018-05-19 00:00:00	90008.000000	210.605000
max	9800.000000	2018-12-30 00:00:00	2019-01-05 00:00:00	99301.000000	22638.480000
std	2829.160653	NaN	NaN	32041.223413	626.651875

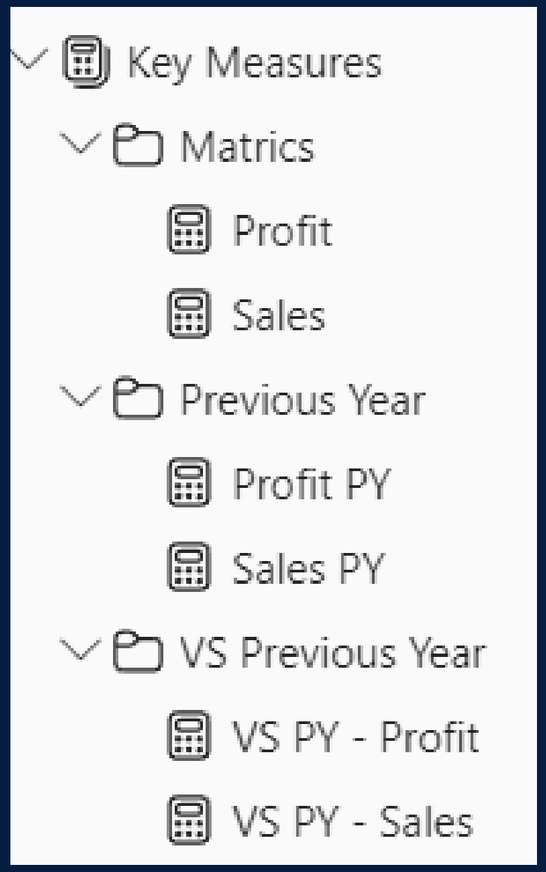
03 STEP 3

ML on Azure

- Data Preparation
- Train-Test Split
- Identifying Numeric and Categorical Features
- Preprocessing Pipelines
- Defining Models
- Training and Evaluation



03 STEP 4



The DAX Measures

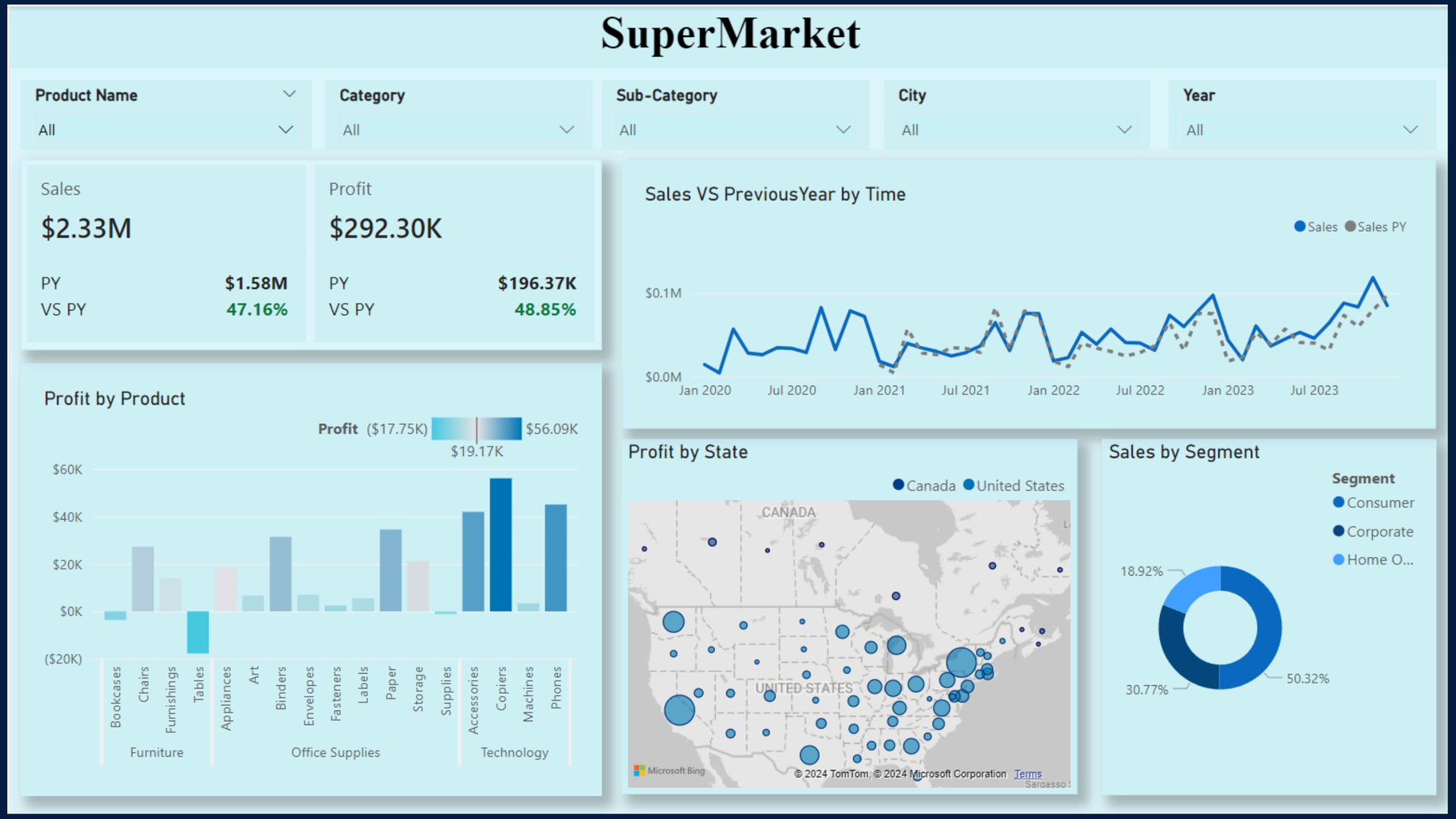
```
1  Sales = SUM(Orders[Sales])
```

```
1  Sales PY =
2  CALCULATE(
3      [Sales],
4      SAMEPERIODLASTYEAR('Date Table'[Date]))
```

```
1  VS PY - Sales =
2  DIVIDE( [Sales] - [Sales PY] , [Sales PY])
```

03 STEP 4

The Dashboard



03 STEP 4

The Key Points

Interactive Visualization

Allows for dynamic filtering by product, category, sub-category, city, and year.

Key Performance

Displays total sales, profit, and year-over-year comparisons.

Sales and Profit Trends

Graphical insights into sales over time, segmented by product and region.

Geographical Breakdown

Visual map of profit by state with a comparison between the U.S. and Canada.

Customer Segmentation

Insights into sales distribution across consumer, corporate, and home office segments.

THANK YOU