MICROSOFT POWER BI TRACK

GRADUCATION PROJECT

GROUP [6]

MINISTRIY OF COMMUNICATION AND INFORMATION TECHNOLOGY

DEPI-ROUND 2

# **DEDICATION & ACKNOWLEDGEMENT**

With gratitude to Allah Almighty, we, Group 6 - Data Analysis Pioneers, extend our sincere appreciation to the Egyptian Ministry of Communications for adopting the Digital Egypt Pioneers initiative in alignment with the Egyptian government’s Vision 2030.

We also express our deep gratitude to Dr. Mohamed Atiya for his tremendous efforts throughout this round, providing comprehensive explanations and daily follow-ups.

Additionally, we extend our thanks and appreciation to CLS, the organizing and sponsoring company, for their dedication and excellent organization, which significantly contributed to the success of this free scholarship.

**Data Analysis Pioneers team**

# **ABSTRACT**

Within the framework of Egypt Vision 2030, and in alignment with the continuous directives of the President of Egypt emphasizing the necessity of investing in human development, the Ministry of Communications and Information Technology (MCIT) has launched the Digital Egypt Pioneers initiative. This initiative reflects the ministry’s commitment to advancing the ICT sector and developing highly skilled professionals.

The Vision is to achieve leadership in the ICT sector by building a skilled and innovative workforce capable of driving comprehensive digital transformation both locally and regionally.

The initiative is implemented in collaboration with leading global technology companies, as well as local and regional enterprises specializing in ICT skills development. These companies contribute by organizing practical workshops in various technological fields, enriching the educational process, and offering training programs to enhance students’ technical expertise.

The initiative's success partners include some of the world's leading technology companies, such as Google, Huawei, IBM, Linux, Microsoft, Oracle, SAP, and many other global industry leaders contributing to the initiative's objectives.

# **LIST OF TABLES**

[Table 1 Microsoft Power BI Track Overview 6](#_Toc196216116)

[Table 2 Data Analysis Project Charter 10](#_Toc196216117)

[Table 3 Dataset Column's Description 11](#_Toc196216118)

[Table 4 Breakdown of the visual elements presented in the sales dashboard 19](#_Toc196216119)

[Table 5 Breakdown of the visual elements presented in the customers dashboard 21](#_Toc196216120)

[Table 6 Breakdown of the visual elements presented in the product dashboard 23](#_Toc196216121)

[Table 7 Breakdown of the visual elements presented in the orders dashboard 25](#_Toc196216122)

# **LIST OF FIGURES**

[Figure 1 Documentation of WhatsApp group creation 7](#_Toc196216422)

[Figure 2 Project Selection Criteria 9](#_Toc196216423)

[Figure 3 Sample Dataset Screenshot 12](#_Toc196216424)

[Figure 4 Data Cleaning process 14](#_Toc196216425)

[Figure 5 Creating Customers, Geolocation and Product tables 15](#_Toc196216426)

[Figure 6 Create dates table using DAX 15](#_Toc196216427)

[Figure 7 Using Model View to create Star Schema 16](#_Toc196216428)

[Figure 8 Creating Measures 16](#_Toc196216429)

[Figure 9 Structure of dashboard 17](#_Toc196216430)

[Figure 10 Sales Dashboard 18](#_Toc196216431)

[Figure 11 Customer Dashboard 20](#_Toc196216432)

[Figure 12 Product Dashboard 22](#_Toc196216433)

[Figure 13 Orders Dashboard 24](#_Toc196216434)

Table of Contents

[**DEDICATION & ACKNOWLEDGEMENT** 1](#_Toc196216816)

[**ABSTRACT** 2](#_Toc196216817)

[**LIST OF TABLES** 3](#_Toc196216818)

[**LIST OF FIGURES** 4](#_Toc196216819)

[**I.** **Chapter (1): Introduction** 6](#_Toc196216820)

[**Track Overview** 6](#_Toc196216821)

[**Project Scope and Objective** 6](#_Toc196216822)

[**Agreed Policy** 7](#_Toc196216823)

[**Project Selection Criteria** 8](#_Toc196216824)

[**Project Charter** 9](#_Toc196216825)

[**1. General Project Information** 9](#_Toc196216826)

[**2. Project Team** 10](#_Toc196216827)

[**3. Stakeholders** 10](#_Toc196216828)

[**4. Project Scope Statement** 10](#_Toc196216829)

[**5. Communication Strategy** 10](#_Toc196216830)

[**II.** **Chapter (2): Selected Project Overview** 11](#_Toc196216831)

[**Introduction** 11](#_Toc196216832)

[**Agreed Metrics & KPIs** 12](#_Toc196216833)

[**Data Cleaning and preparation** 13](#_Toc196216834)

[**III.** **Chapter (3): Creating Dashboard** 17](#_Toc196216835)

[**Dashboard Structure** 17](#_Toc196216836)

[**Sales Dashboard** 17](#_Toc196216837)

[**Customers Dashboard** 20](#_Toc196216838)

[**Product Dashboard** 22](#_Toc196216839)

[**Orders Dashboard** 23](#_Toc196216840)

[**IV.** **Chapter (4): Conclusion** 26](#_Toc196216841)

# **Chapter (1): Introduction**

The team will discuss the following topics: Overview, Scope & Objectives, Group Policy for Project Selection Criteria, and the Project Charter.

## **Track Overview**

|  |  |
| --- | --- |
| Inquiry | Details |
| Round Number | 2 |
| Track Name | Microsoft Power BI Engineer |
| Instructor’s Name | Dr. Mohamed Atiya |
| Group Code | Microsoft Power BI Engineer - CLS ONL2\_DAT2\_G1 |
| Sponsor | Ministry of Communication and Information Technology |
| Provider | CLS Learning Solutions |
| Graduation Project Group Name | Data Analysis Pioneers team |
| Numbers of Team Members | 5 |

Table 1 Microsoft Power BI Track Overview

## **Project Scope and Objective**

The project scope includes the implementation and execution of data analysis for the selected project using Microsoft Power BI. This involves:

* Utilizing Power Query to clean and prepare the data.
* Applying various charts to support the overall concept.
* Using DAX formulas to calculate key indicators based on the data.

And, the project Objective include some of agreed criteria based on the project scope. This involves:

* Encouraging collective thinking and brainstorming to increase data presentation proposals by 25% compared to the initial project discussion.
* Enhancing communication and collaboration skills among team members.

## **Agreed Policy**

On November 1, 2024, the team created a WhatsApp group to collaborate on the project. One team member was chosen as the leader and project manager. The following rules were agreed upon to achieve the goals of the Data Analysis Pioneers team:

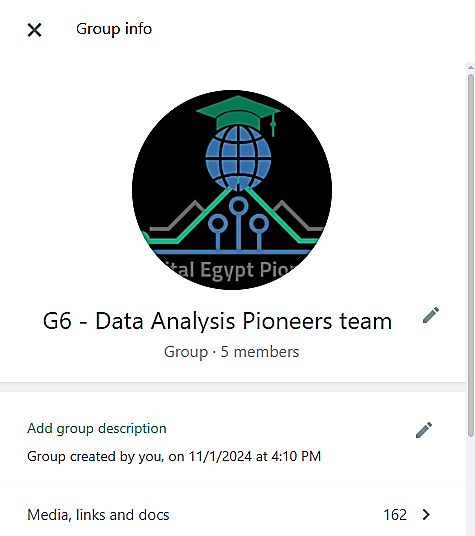
* Each team member selected a project idea and presented it in an online meeting to the rest of the team.
* The team chose the best idea to apply and implement for their graduation project.
* The team agreed that each member would complete their assigned tasks individually, then discuss their work together to reach the best results, make improvements, and finalize the project.

Figure 1 Documentation of WhatsApp group creation

## **Project Selection Criteria**

Among the six proposed projects provided by the Egypt Digital Pioneers Initiative Management as potential projects and databases for the graduation project, which included databases related to:

* **Human Resources (HR)**
* **Manufacturing Downtime**
* **MTA Daily Ridership**
* **UK Train Rides**
* **Superstore Sales Dataset**
* **Supply Chain Data**

Each team member individually selected a project and then presented their idea in a team meeting. Subsequently, the team collectively chose the best project to work on based on a set of criteria, these criteria was agreed and included:

* **Ease of understanding the data**
* **Data quality**
* **Richness of the data in terms of extractable insights**
* **Feasibility of utilizing the data in real-world scenarios**
* **Capability to develop advanced analytical models**

A relative weighting system ranging from 1 to 10 was used to compare all the proposed projects. Based on these relative weights, the team selected the most suitable project to work on.



Figure 2 Project Selection Criteria

After selecting the project, the team documented it in a foundational project charter to ensure an institutional approach that guarantees effective implementation and monitoring.

## **Project Charter**

|  |  |  |  |
| --- | --- | --- | --- |
| **1. General Project Information** | | | |
| **Project Name:** | | **Superstore Sales Dataset Analysis** | |
| **Executive Sponsors:** | | * **Ministry of Communication and Information Technology.** * **DEPI** * **CLS Learning Solutions** | |
| **Department Sponsor:** | | **Dr. Mohamed Ateya** | |
| **Impact of project:** | | * **Increasing the team's ability to use and analyze data using Power BI.** * **Enhancing the team's competitiveness in the job market by obtaining the scholarship certificate.** | |
| **2. Project Team** | | | |
|  | **Name** | | **Phone Number** |
| **Project Manager:** | Yousef Mohamed Elsayed | | +2 01026041908 |
| **Team Members:** | Alaa Hesham | | +20 10 64150642 |
| Marwa Taha | | +20 10 20499971 |
| Zeniab Elsadek | | +20 15 54305525 |
| Ahmed Fahmy | | +2 01063422021 |
| **3. Stakeholders** | | | |
| **Stakeholders includes the following :**   * **Ministry of Communication and Information Technology.** * **DEPI.** * **CLS Learning Solutions.** * **Data Analysis Pioneers team.** | | | |
| **4. Project Scope Statement** | | | | |
| **In the project approved by the Data Analysis Pioneers team, it was observed that data is recorded in its raw form, making it difficult to utilize effectively. Therefore, the project aims to transform this data into meaningful information, ensuring at least 90% data utilization.**  **The project scope includes:**   * **Designing an interactive report using Power BI to present information clearly and effectively.** * **Engaging with stakeholders to understand their requirements and ensure the report meets their needs.** * **Utilizing learned tools and concepts to implement the project efficiently.** | | | | |
| **Objectives** | | | | |
| **In the project approved by the Data Analysis Pioneers team, it was observed that data is recorded in its raw form, making it difficult to utilize effectively. Therefore, the project aims to transform this data into meaningful information, ensuring at least 90% data utilization.**  **The project scope includes:**   * **Designing an interactive report using Power BI to present information clearly and effectively.** * **Engaging with stakeholders to understand their requirements and ensure the report meets their needs.** * **Utilizing learned tools and concepts to implement the project efficiently.** | | | | |
| **Deliverables** | | | | |
| * **Utilizing Power BI tools, such as Power Query, to process, organize, and clean data.** * **Designing an interactive report using Power BI.** | | | | |
| **Scope** | | | | |
| **This project includes all stakeholder requirements, as well as the outcomes of brainstorming sessions among team members. It does not include any considerations outside DEPI.** | | | | |
| **5. Communication Strategy** | | | | |
| * **Weekly Lectures** * **Whats App Groups** * **Online Meetings & Cell Phones.** | | | | |

Table 2 Data Analysis Project Charter

# **Chapter (2): Selected Project Overview**

## **Introduction**

The Selected dataset, titled "Superstore Sales Dataset", contains transactional sales data from a retail store. The dataset is a Comma-Separated Values [CSV] file containing 9,800 rows and 18 columns.

It includes various details related to orders, shipping, customers, products, and sales performance. This dataset can be useful for analyzing business trends, customer behavior, and logistics efficiency.

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| Row ID | int64 | Unique row identifier |
| Order ID | object | Unique identifier for each order |
| Order Date | object | Date when the order was placed |
| Ship Date | object | Date when the order was shipped |
| Ship Mode | object | Shipping type (e.g., Standard, Express) |
| Customer ID | object | Unique customer identifier |
| Customer Name | object | Name of the customer |
| Segment | object | Customer segment (e.g., Consumer, Corporate) |
| Country | object | Country where the order was placed |
| City | object | City of the customer |
| State | object | State of the customer |
| Postal Code | int64 | Postal code of the shipping address |
| Region | object | Regional classification (e.g., West, South) |
| Product ID | object | Unique identifier for each product |
| Category | object | Product category (e.g., Furniture, Office Supplies) |
| Sub-Category | object | Product sub-category |
| Product Name | object | Name of the product |
| Sales | float64 | Sales revenue for the product |

Table 3 Dataset Column's Description

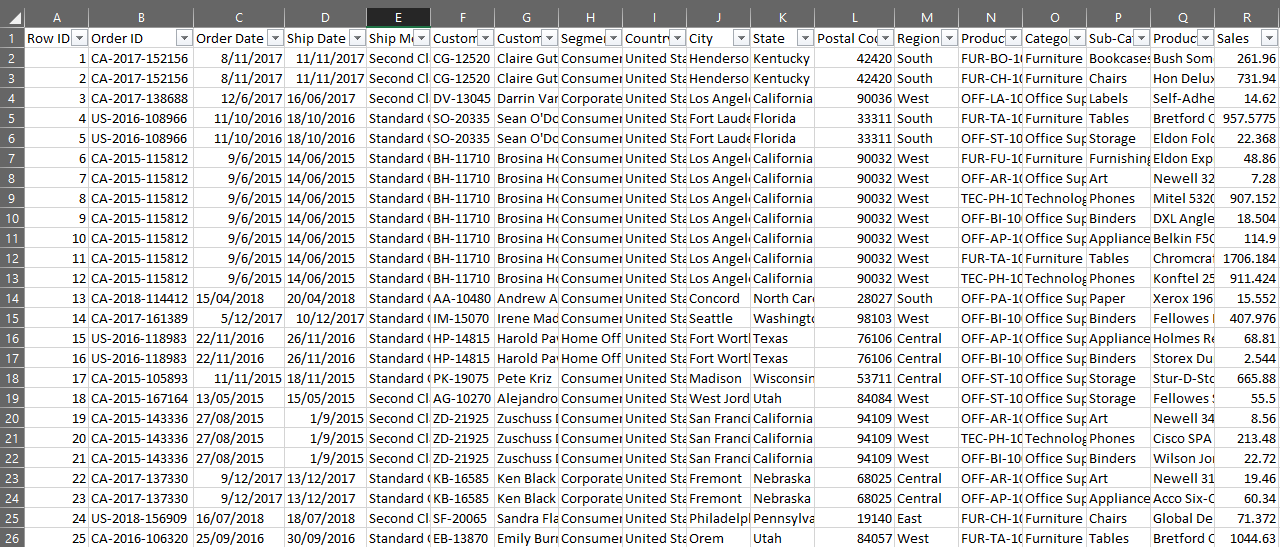


Figure 3 Sample Dataset Screenshot

## **Agreed Metrics & KPIs**

Through brainstorming and collaboration among team members, the team has settled on several key indicators and questions that are expected to be valuable to stakeholders. This Include

* Total Sales (Sum of Sales)
* Average Order Value
* Total Sales by Category
* Total Sales by Region
* Total Number of Orders
* Sales Trend Over Time
* Sales by Customer
* Sales Distribution by Segment (Consumer/Corporate)
* Average Shipping Time
* Orders by Ship Mode
* Orders by Sub-Category
* Top 5 Products by Sales
* Sales Distribution by Product

## **Data Cleaning and preparation**

Data cleaning is a crucial step in preparing data for analysis, and Power Query in Excel or Power BI is an effective tool for performing these tasks.

In this process, the data was imported from the source, the first row was promoted to headers, and data types were defined.

Duplicates and unnecessary columns such as "Country" were removed, and missing values and city names were standardized.

Order, customer, and product IDs were split using delimiters, converted to numbers, and the resulting temporary columns were deleted.

The cleaned columns were then renamed accordingly.

Afterwards, queries were merged to combine data from multiple sources, expanded to show detailed product information, and finally, irrelevant columns were removed to produce a well-structured, clean dataset ready for analysis.

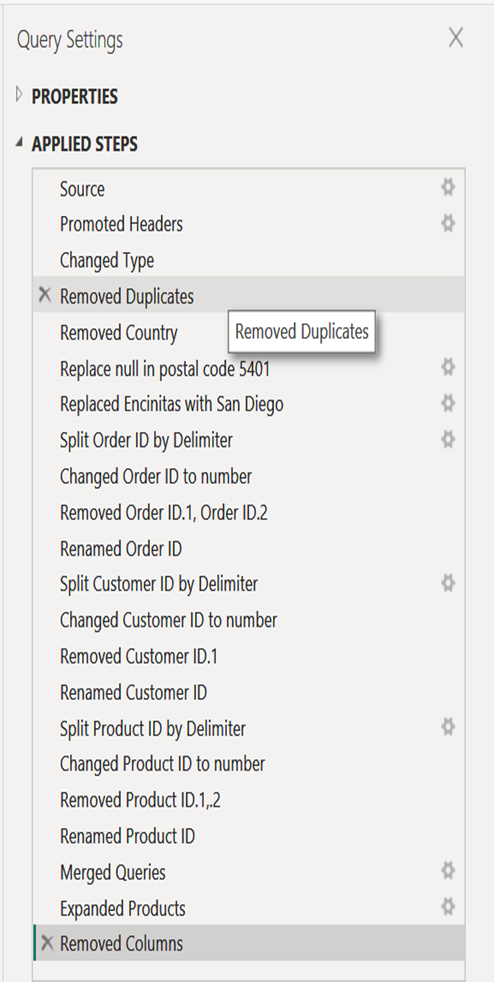


Figure 4 Data Cleaning process

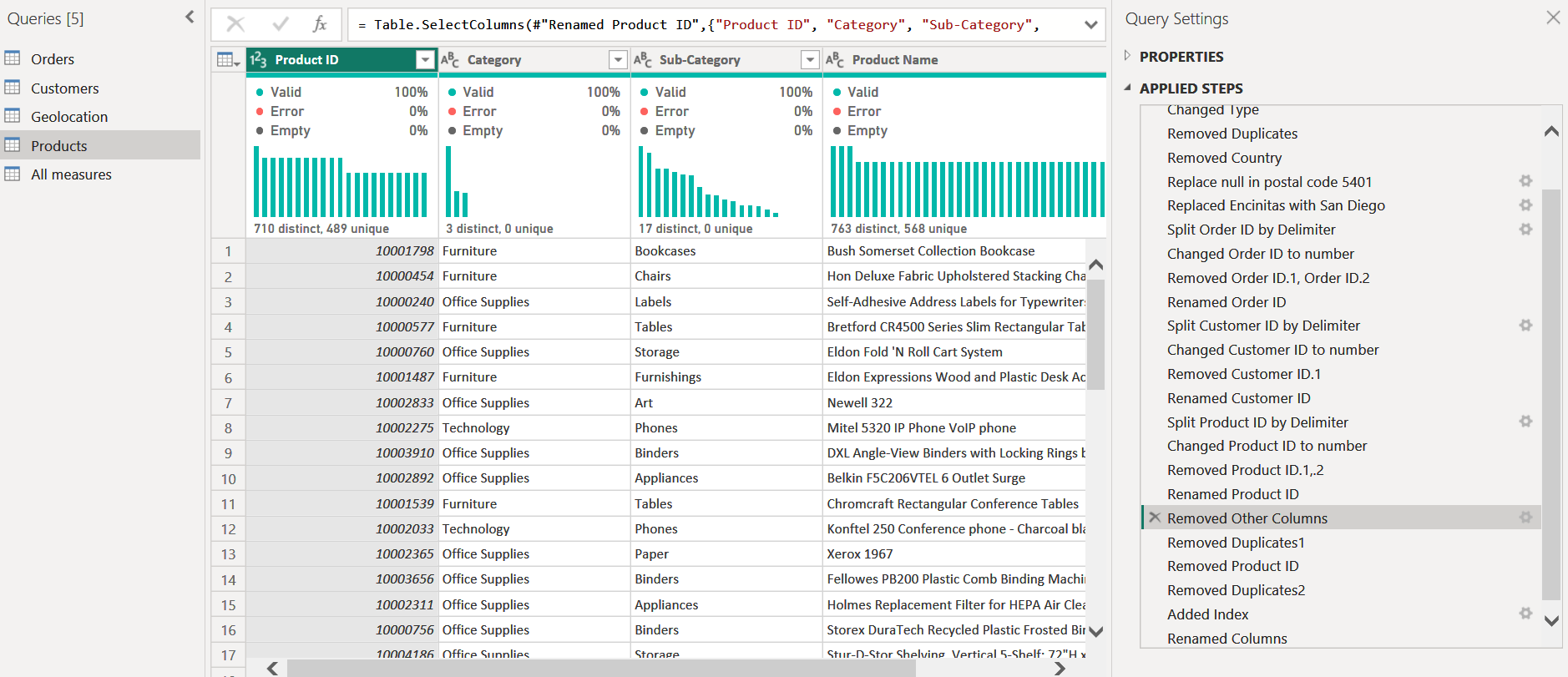
After completing the data cleaning process in Power Query, the Orders table was duplicated three times to create separate sub-tables for easier analysis and organization. In the first step, the Orders table was duplicated and renamed to Customers to extract customer-related information. Then it was duplicated again and renamed to Geolocation to work with geographic location data. Finally, the table was duplicated and renamed to Products to isolate and analyze product details. These steps help build a flexible and well-structured data model, making it easier to establish relationships between different tables using common keys.

Figure 5 Creating Customers, Geolocation and Product tables

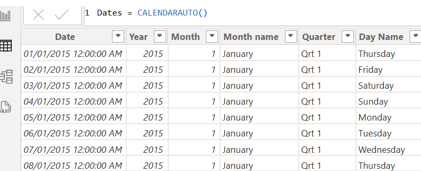
Then, we used DAX to create dates table

Figure 6 Create dates table using DAX

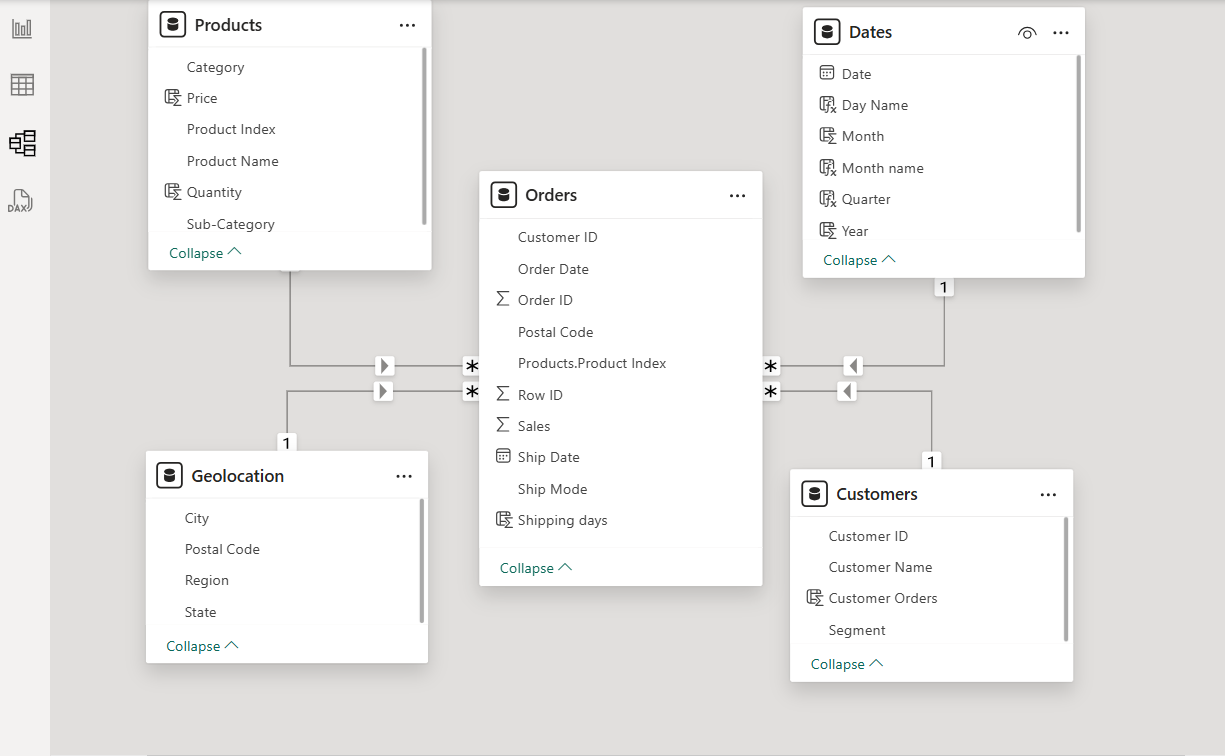
Also, we used the Model View to create relationship (Star schema)

Figure 7 Using Model View to create Star Schema

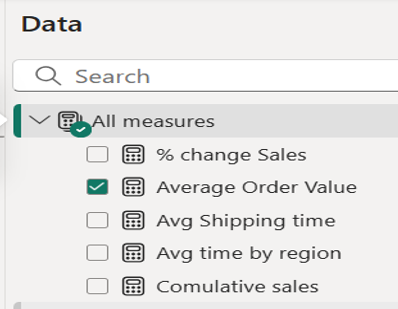
And then, we create a separated section for calculates

Figure 8 Creating Measures

# **Chapter (3): Creating Dashboard**

## **Dashboard Structure**

The team divided the visual presentation into four sections, distributed among the team members. The presentation included a visual report for Sales, another for Customers, one for Products, and finally a visual report for Orders.

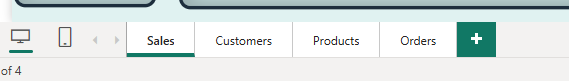


Figure 9 Structure of dashboard

### **Sales Dashboard**

Sales dashboard allowing users to understand sales performance from multiple perspectives (temporal, categorical, geographical, and logistical).

It also supports data-driven decision-making through interactive and analytical visualization.

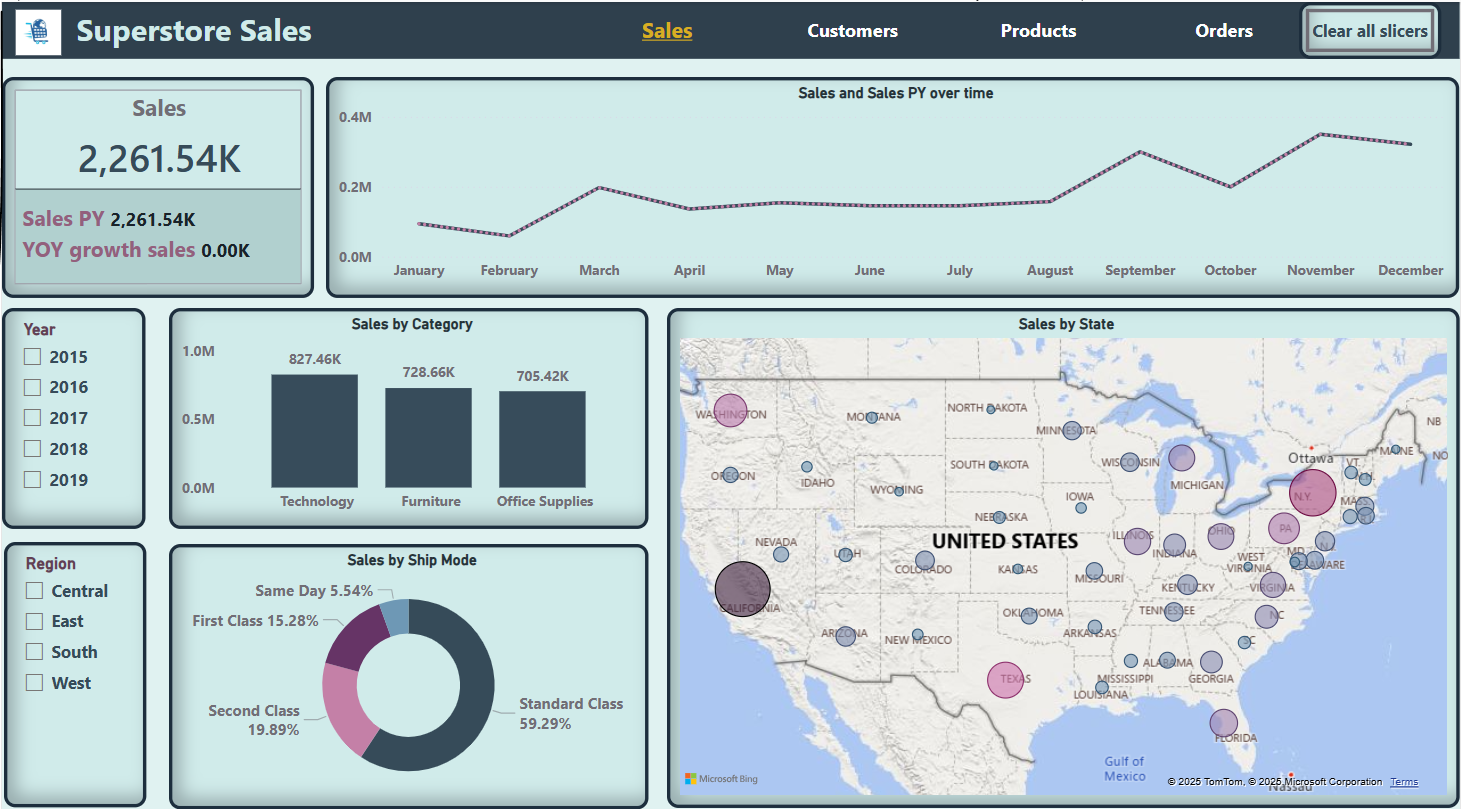


Figure 10 Sales Dashboard

|  |  |  |  |
| --- | --- | --- | --- |
| Indicator / Element | Description | Value | Chart Type / Element |
| **Sales** | Total current sales value | 2,261.54K | KPI Card |
| **Sales PY (Previous Year Sales)** | Sales value for the same period last year | 2,261.54K | KPI within the card |
| **YOY Growth Sales** | Year-over-year sales growth rate (no growth in this case) | 0.00K | KPI within the card |
| **Sales and Sales PY Over Time** | Comparison of monthly sales this year vs. last year | - | Line Chart |
| **Sales by Category** | Sales distribution across three categories: | Technology: 827.46K Furniture: 728.66K Office Supplies: 705.42K | Bar Chart |
| **Sales by State** | Geographical distribution of sales by state across the U.S. | - | Map with Bubble Markers |
| **Sales by Ship Mode** | Percentage of sales by shipping method | Standard: 59.29% Second: 19.89% First: 15.28% Same Day: 5.54% | Donut Chart |
| **Year Slicer** | Filter to select the year from 2015 to 2019 | 2015–2019 | Slicer (Interactive Filter) |
| **Region Slicer** | Filter to select region (Central – East – South – West) | Central, East, South, West | Slicer (Interactive Filter) |

Table 4 Breakdown of the visual elements presented in the sales dashboard

The table provides a detailed breakdown of the visual elements presented in the sales dashboard. Each indicator is described along with its corresponding value and the type of chart or visual component used to represent it.

The dashboard components cover various analytical perspectives, including overall performance (KPIs), trends over time, category-based analysis, geographical insights, shipping logistics, and interactive filters for year and region.

### **Customers Dashboard**

Customers dashboard enabling users to analyze customer distribution, behavior, and contribution to sales across different segments, time periods, and regions. It facilitates strategic insights through interactive filters and visualizations, helping identify key customer groups and top performers.

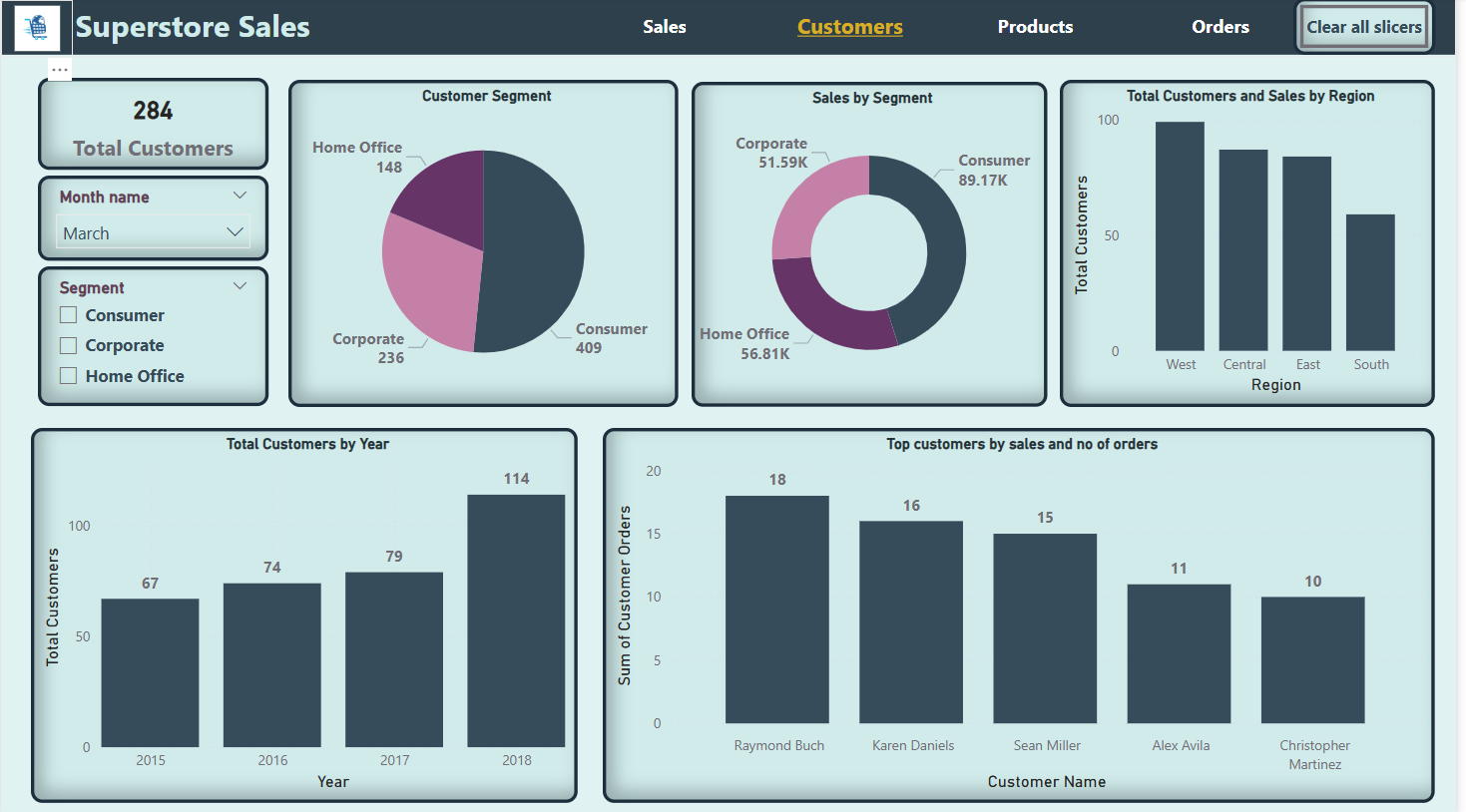


Figure 11 Customer Dashboard

|  |  |  |  |
| --- | --- | --- | --- |
| Indicator / Element | Description | Value | Chart Type / Element |
| Total Customers | Total number of customers in the selected time period | 284 | KPI Card |
| Month Filter | Drop-down filter to select a specific month | March | Slicer (Interactive Filter) |
| Segment Filter | Interactive filter to select customer segments | Consumer, Corporate, Home Office | Slicer (Interactive Filter) |
| Customer Segment | Distribution of customers across three segments | Consumer: 409 Corporate: 236 Home Office: 148 | Pie Chart |
| Sales by Segment | Total sales value per segment | Consumer: 89.17K Home Office: 56.81K Corporate: 51.59K | Donut Chart |
| Total Customers and Sales by Region | Geographical distribution of customers across regions | West, Central, East, South (with bar heights indicating customer count) | Bar Chart |
| Total Customers by Year | Annual customer count from 2015 to 2018 | 2015: 67 2016: 74 2017: 79 2018: 114 | Bar Chart |
| Top Customers by Sales and Number of Orders | Leading customers by number of orders placed | Raymond Buch: 18 Karen Daniels: 16 Sean Miller: 15 Alex Avila: 11 Christopher Martinez: 10 | Bar Chart |

Table 5 Breakdown of the visual elements presented in the customers dashboard

### **Product Dashboard**

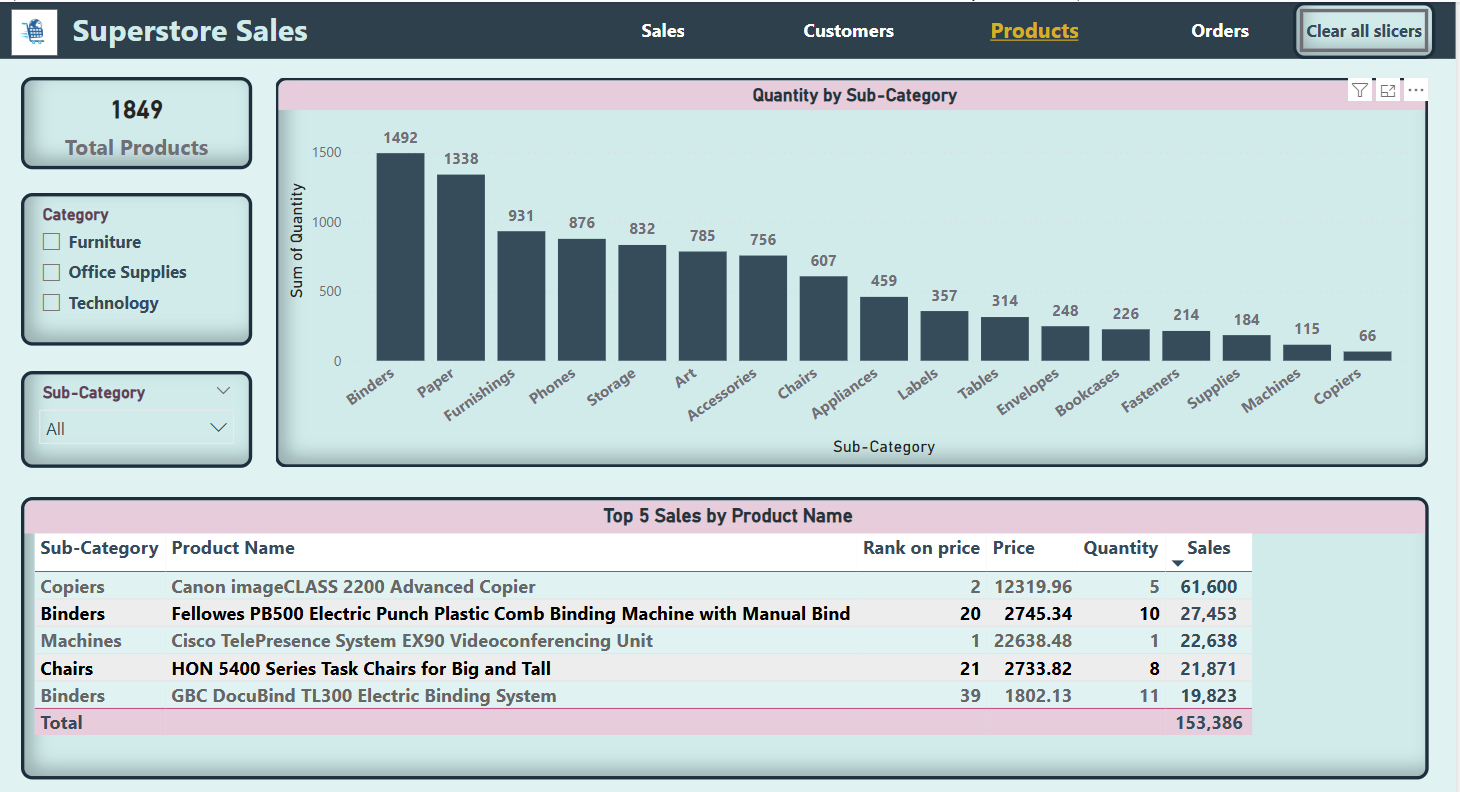
dashboard providing a comprehensive view of the superstore's product portfolio performance across various sub-categories. It highlights total product count, quantity by sub-category, and top-selling products, enabling users to identify popular items and sales trends.

Figure 12 Product Dashboard

|  |  |  |  |
| --- | --- | --- | --- |
| Indicator / Element | *Description* | *Value* | *Chart Type / Element* |
| Total Products | Total number of unique products in the dataset | 1849 | KPI Card |
| Category Filter | Interactive filter to select product categories (Furniture, Office Supplies, Technology) | Furniture, Office Supplies, Technology | Slicer |
| Sub-Category Filter | Interactive filter to select product sub-categories within the chosen category | All (with a dropdown list of sub-categories) | Slicer |
| Quantity by Sub-Category | Total quantity sold for each product sub-category | Binders: 1492, Paper: 1338, ... Copiers: 66 | Bar Chart |
| Top 5 Sales by Product Name | Top 5 products based on total sales value | (See table in the image for details) | Table |

Table 6 Breakdown of the visual elements presented in the product dashboard

### **Orders Dashboard**

dashboard providing a comprehensive overview of order-related metrics, including total orders, average shipping time, and average order value.

It allows users to analyze order patterns by ship mode, customer segment, and product sub-category, as well as track shipping time trends over different years and categories, facilitating insights into operational efficiency and sales distribution.



Figure 13 Orders Dashboard

|  |  |  |  |
| --- | --- | --- | --- |
| Indicator / Element | Description | Value | Chart Type / Element |
| Total Orders | Total number of orders placed | 9800 | KPI Card |
| Avg Shipping time | Average time taken for order shipment | 3.96 | KPI Card |
| Average Order Value | Average value of all orders | 459.48 | KPI Card |
| Year Filter | Drop-down filter to select a specific year for order analysis | All | Slicer |
| Orders by Ship Mode | Distribution of orders across different shipping modes | Same Day: 5.49%, First Class: 15.32%, ... | Pie Chart |
| Average Order Value by Segment | Average order value for each customer segment | Home Office: 475.37, Corporate: 461.77, ... | Bar Chart |
| Orders by Sub-Category | Number of orders containing items from each product sub-category | Binders: 1.5K, Paper: 1.3K, ... Copiers: 0.1K | Horizontal Bar Chart |
| Avg Shipping time by Year and Category | Trend of average shipping time over the years, broken down by product category | (Visual representation of lines for each category) | Line Chart |

Table 7 Breakdown of the visual elements presented in the orders dashboard

# **Chapter (4): Conclusion**

In conclusion, this project represents a significant achievement that aligns with Egypt's Vision 2030 for the development of the information and communication technology sector. The Digital Egypt Pioneers initiative has provided the Data Analysis Pioneers team with the opportunity to gain practical experience in using Microsoft Power BI to analyze sales data.

Through this project, we were able to:

* Clean and process data using Power Query.
* Design interactive dashboards that provide valuable insights into sales, customers, products, and orders.
* Utilize charts and Key Performance Indicators (KPIs) to facilitate data understanding and decision-making.

The team's ability to transform raw data into actionable information highlights the importance of investing in the development of data analysis skills. We hope that this project has contributed to the initiative's objectives and will have a positive impact on the participants' ability to compete in the job market.