FMCG Company Sales and Performance Analysis

Track: Microsoft Power BI Engineer

Track Code: GHR_DAT2_M1e

Technical Provider: ITC

Team Members:

- Alaa Osama Rashad
- Mohamed Elsayed Ali
- Eman Ahmed Mekawy
- Mohamed Shawada
- Mohand Oraiby

Table of Contents

1. Introduction

The FMCG (Fast-Moving Consumer Goods) sector is highly competitive, requiring businesses to continuously monitor sales performance, employee efficiency, product popularity, and outlet visits. Our analysis focused on an FMCG company's sales performance for the first quarter of 2020, with particular attention on identifying trends, evaluating employee contributions, and recognizing opportunities for operational improvements.

This report outlines our findings, showcasing key insights into sales, top-performing products, outlet visits, and return trends. Our goal was to provide the company with actionable recommendations to optimize its operations and prepare for the upcoming quarter.

2. Project Objectives

The main objectives of this project were:

- To analyze the company’s sales performance during the first quarter.
- To evaluate employee performance against set targets.
- To determine which outlets required the most support.
- To identify sales trends across product categories and regions.
- To highlight key areas for improvement based on return rates and other factors.

3. Data Collection and Analysis Process

Step 1: Data Collection

We gathered the data, which included the following tables:

- Sales
- Visits
- Warehouse
- Reps List
- Targets (April 2021)
- Outlets
- products

Step 2: Data understanding:

Here's an explanation of each table and the associated columns in it.

1. Sales Table: This table contains detailed sales transaction data.

- **Date**: The date of the sale.
- **Sub_Db_Name**: The name of the warehouse.
- **Username**: The username of the sales rep making the sale.
- **Name_Of_The_User**: The name of the sales rep.
- **Outlet_Id**: The ID of the outlet where the sale was made.
- **PRODUCT_CODE**: The product code of the item sold.
- **Product Name**: The name of the product sold.
- **Quantity**: The quantity of the product sold.
- **Price_Per_Piece**: The price per unit of the product.
- **Total Price**: The total price of the transaction.

Purpose: This table captures details of all sales transactions, including the product sold, quantity, and total sale amount.

2. Visits Table

This table logs visits made by sales reps to various stores or outlets.

- **DB Name**: The name of the warehouse associated with the visit.
- **Sales Rep ID**: The unique identifier of the sales rep.
- **Sales Rep Name**: The name of the sales rep.
- **Date**: The date of the visit.
- **Classification**: The classification of the store ("تجزئة – جملة")
- **Store Code**: A unique identifier for the store visited.
- **Visit Starting Time**: The time when the visit began.
- **Visit Ending Time**: The time when the visit ended.
- **Sale Amount**: The total sales amount during the visit (if any).

Purpose: This table records the visits made by sales reps to outlets, including timing and sales data.

3. Warehouse Table

This table holds details about different warehouses.

- **Warehouse Name**: The name of the warehouse.
- **Code**: A unique identifier or code for the warehouse.
- **Region**: The geographical region where the warehouse is located.

Purpose: This table stores information about the distribution centers or warehouses used for storing and shipping goods.

4. Reps List Table

This table contains information about the sales representatives.

- **ID**: A unique identifier for each sales representative.
- **Name**: The name of the sales rep.
- **Username**: A unique code used by the sales rep.
- **Role**: The role or designation of the sales rep.
- **ZONE**: The area where the sales rep operates.
- **Distributor**: The warehouse associated with the sales rep.

Purpose: This table lists salespeople and their relevant information, including their operational areas and the distributor they work for.

5. Targets April 2021 Table

This table holds the target and actual sales data for April 2021.

- **ID**: A unique identifier for each record.
- **Username**: The username of the sales rep.
- **Name**: The name of the sales rep.
- **ZONE**: The zone in which the rep is responsible.
- **Distributor**: The distributor associated with the sales rep.
- **Target**: The sales target for the sales rep for the month.
- **AC**: The actual sales achieved.

Purpose: This table tracks each sales rep's target and their actual sales performance for April 2021.

6. Outlets Table

This table contains details about retail outlets where the products are sold.

- **OutletId**: A unique identifier for each outlet.
- **Outlet Name**: The name of the retail outlet.
- **Outlet Class**: Classification of the outlet ("تجزئه- بقالة").
- **Outlet Type**: Type of the retail outlet.
- **Employee_Code**: A unique code for the employee responsible for this outlet.
- **Warehouse Code**: Code of the warehouse servicing this outlet.

Purpose: This table stores details about various retail outlets and the salespeople responsible for them.

7. Products Table

This table holds information about the products being sold.

- **Product ID**: A unique identifier for each product.
- **Product Code**: A unique code used to identify the product.
- **Product Name**: The name of the product.
- **Category**: The category to which the product belongs.
- **Subcategory**: The subcategory of the product.
- **Price**: The price per unit of the product.

Purpose: This table lists the available products, their categorization, and their price details.

Step 3: Business Questions

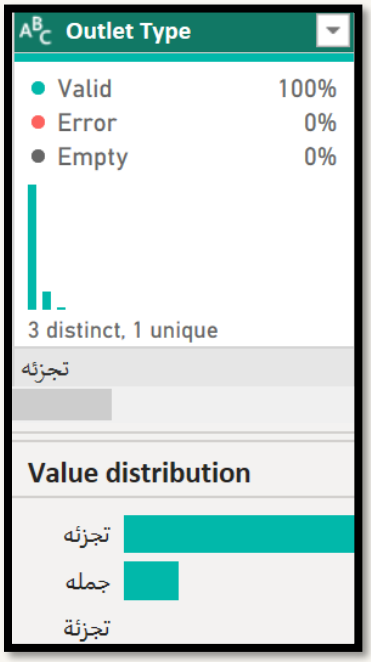
- Which outlet has the most visits?
- Which category generated the highest sales?
- Which employee is not meeting their target?
- What is the top-selling day of the week?
- What is the busiest day in terms of visits?
- What is the highest-selling month?
- Which region contributes the most to sales?
- Which month did the top-selling category perform best?
- Which product had the highest returns?
- Which outlet type performed best in terms of visits?
- Who is the top-performing employee in terms of target achievement?
- Who generated the highest sales?
- Which product generated the highest revenue?
- Which product had the highest quantity?

Step 4: Exploratory Data Analysis (EDA)

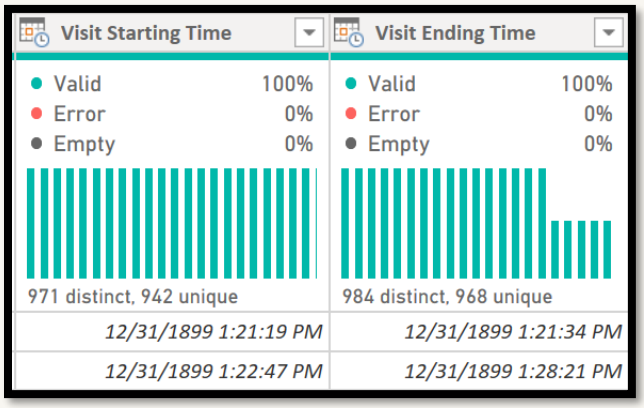
Before cleaning the data, we conducted an initial analysis (EDA) to uncover inconsistencies.

The following issues were identified:

- **Negative Quantities:** Discovered negative sales quantities indicating potential product returns.
- **Missing Employee Information:** Some employee names were missing.
- **Outlet Type Inconsistencies:** Outlet types were written inconsistently (e.g., “Retail” was written both as "تجزئة" and "تجزئه").

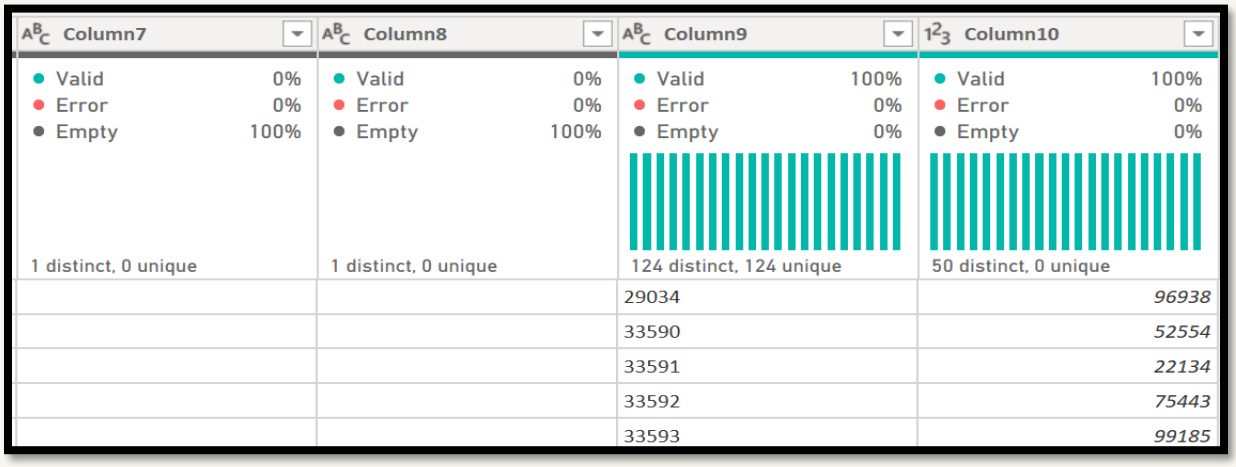


- **Data Type Mismatches:** The visit start time and visit end time column had inconsistent data types.



Visit Starting Time	Visit Ending Time
Valid 100%	Valid 100%
Error 0%	Error 0%
Empty 0%	Empty 0%
971 distinct, 942 unique	984 distinct, 968 unique
12/31/1899 1:21:19 PM	12/31/1899 1:21:34 PM
12/31/1899 1:22:47 PM	12/31/1899 1:28:21 PM

- **Null Values:** Detected missing values, especially in the product table.
- **Unnecessary Columns:** Identified unnecessary columns that weren't useful for analysis in product table.



Column7	Column8	Column9	Column10
Valid 0%	Valid 0%	Valid 100%	Valid 100%
Error 0%	Error 0%	Error 0%	Error 0%
Empty 100%	Empty 100%	Empty 0%	Empty 0%
1 distinct, 0 unique	1 distinct, 0 unique	124 distinct, 124 unique	50 distinct, 0 unique
		29034	96938
		33590	52554
		33591	22134
		33592	75443
		33593	99185

- **Outlets without Proper Names:** Some outlet names were written without proper names (e.g., “,”)

Step 5: Data Cleaning

After identifying these issues in the EDA phase, we proceeded to clean the data:

- **Negative Quantities:** Treated negative sales values as returns.
- **Employee Information:** Collaborated with HR to fill in missing employee codes.
- **Outlet Types:** Standardized the inconsistent outlet types to ensure uniformity.
- **Timestamp Formatting:** Corrected and standardized the data types for the visit start time and the visit end time column.
- **Handling Null Values:** Imputed or removed null values to ensure accuracy.
- **Column Pruning:** Removed redundant columns to maintain focus on relevant data.

Step 6: Data Modeling

Rather than following a traditional star or snowflake schema, we focused on how the **business process** works and structured our data model accordingly.

We began by understanding the flow of operations:

- The **company's warehouses** stock a variety of products.
- **Sales representatives** are assigned targets and are responsible for collecting products from these warehouses.
- They then deliver the products during their **visits to outlets** across different regions.

Based on this business process, we structured our data to reflect these relationships:

- We linked **warehouse stock** with the **sales representatives**, and their visits were connected to the **outlets**.
- This approach allowed us to model the data in a way that closely mirrors the real-world processes of the business, ensuring better insights and accurate analysis.

Step 7: Data Analysis

We conducted a detailed analysis focusing on:

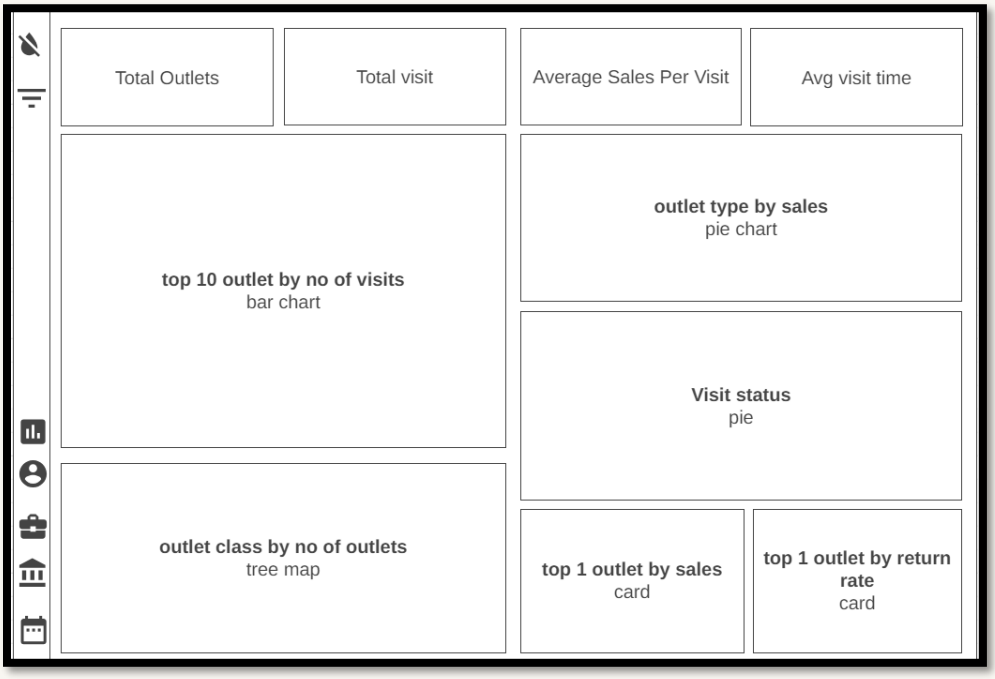
- **Top-performing employees** and their target achievements.
- **Outlet performance**, identifying the most visited and highest-selling locations.
- **Product analysis**, focusing on the best-selling and most returned items.
- **Regional sales**, identifying high-revenue areas.

Step 8: Data Visualization

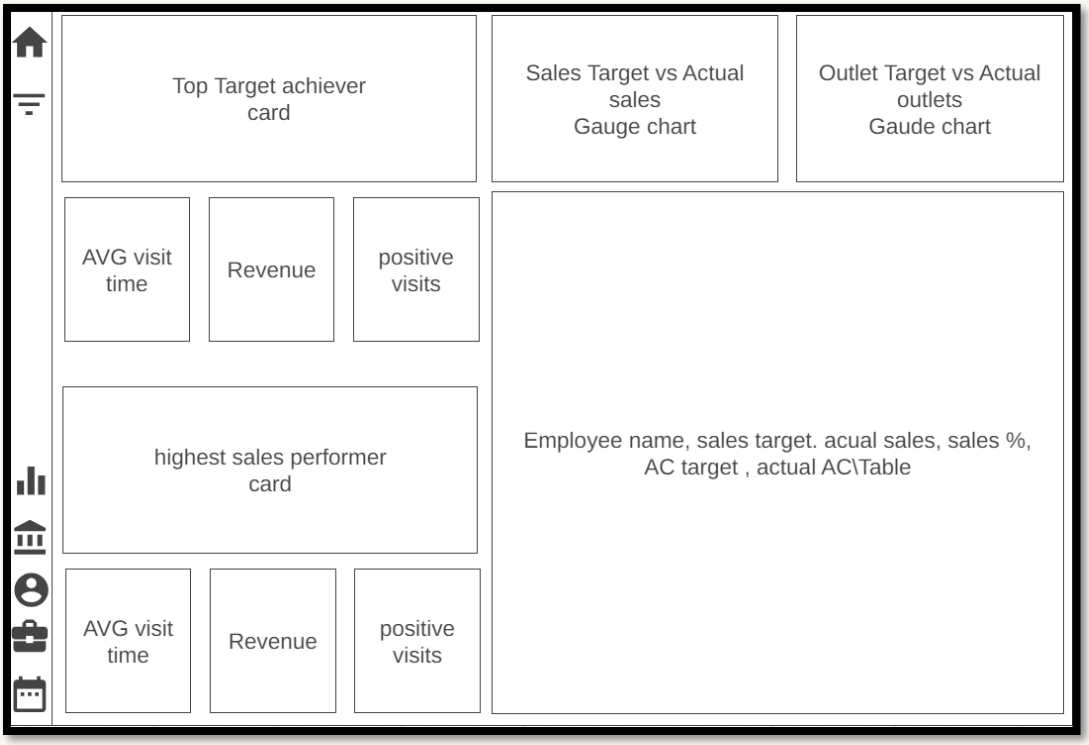
✓ Dashboard Design

- Overview Page

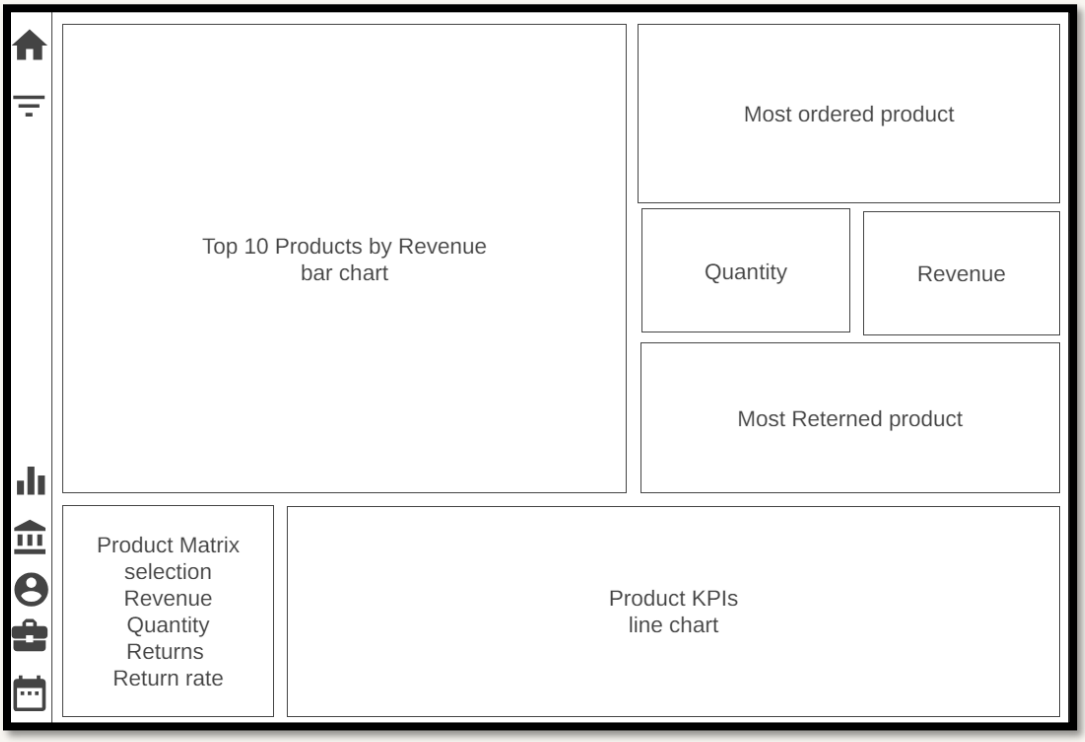
- Outlets Page



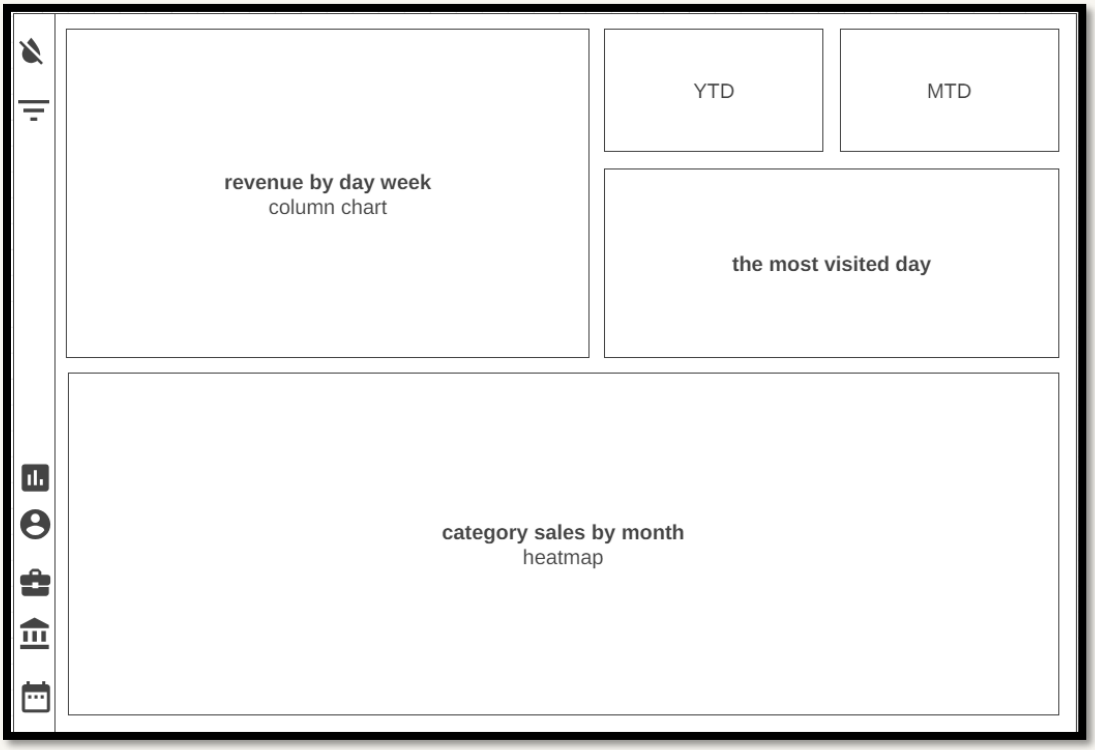
- Sales Person Page



- Product Page

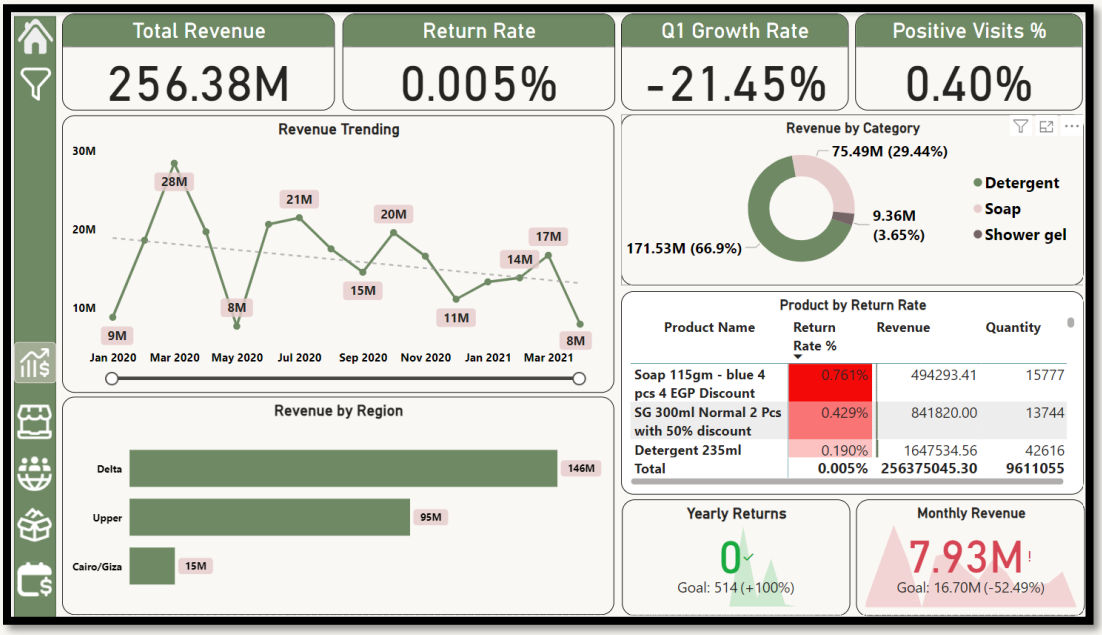


- Date Page

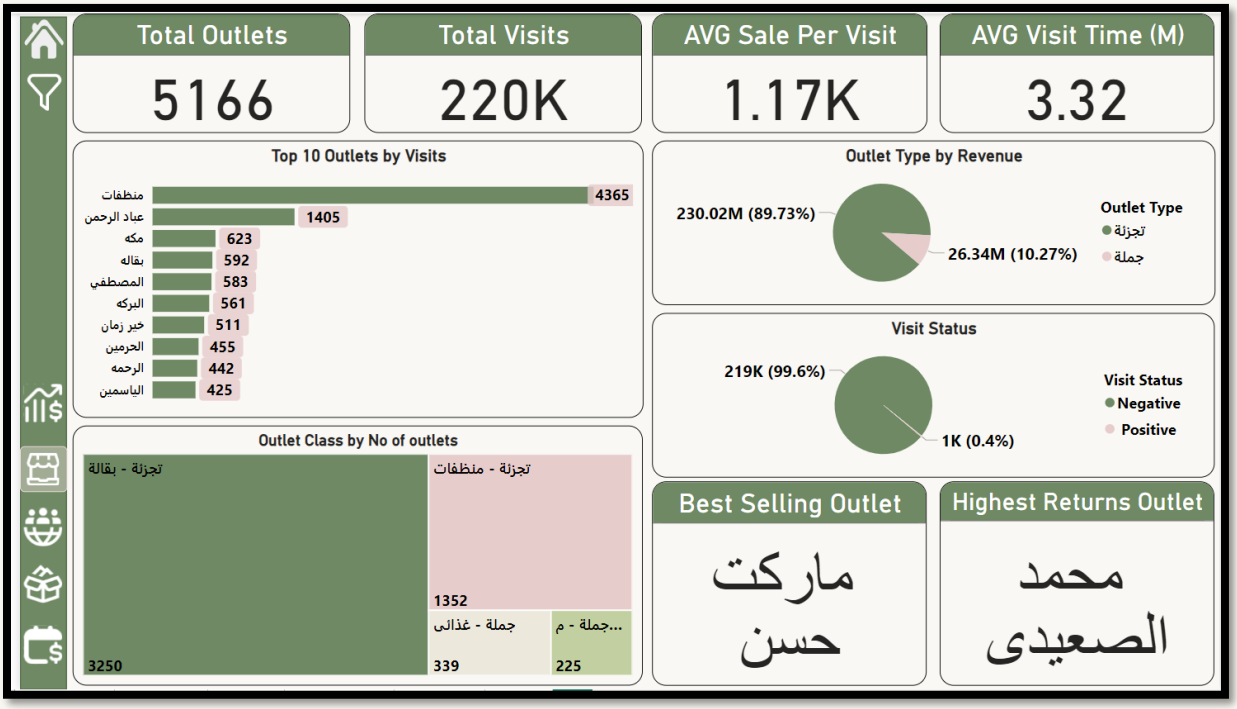


✓ Dashboard Implementation

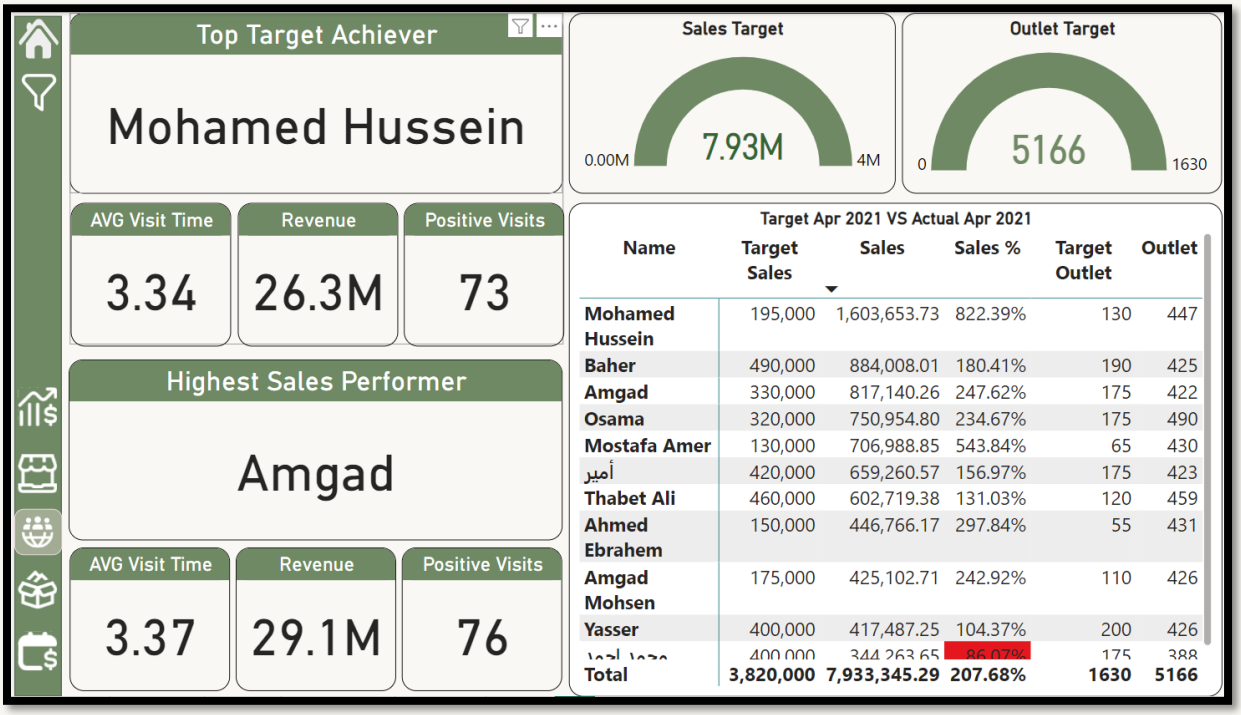
- Overview Page



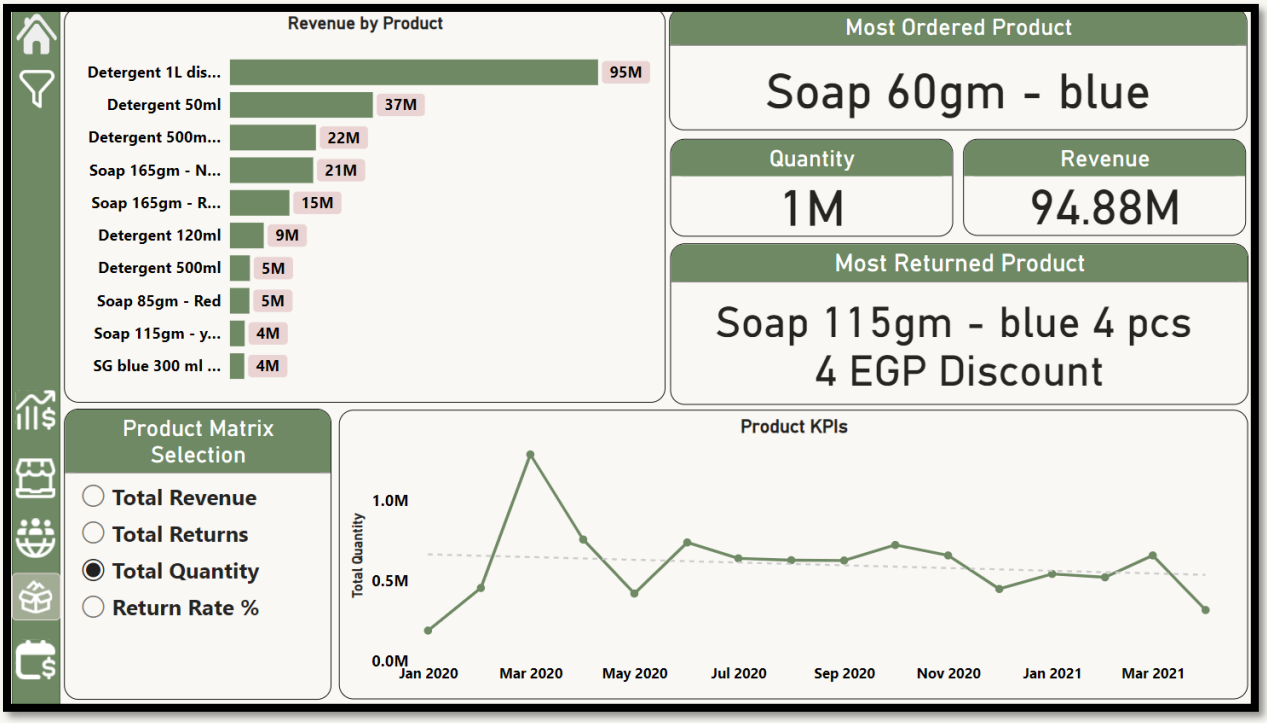
- Outlet Page



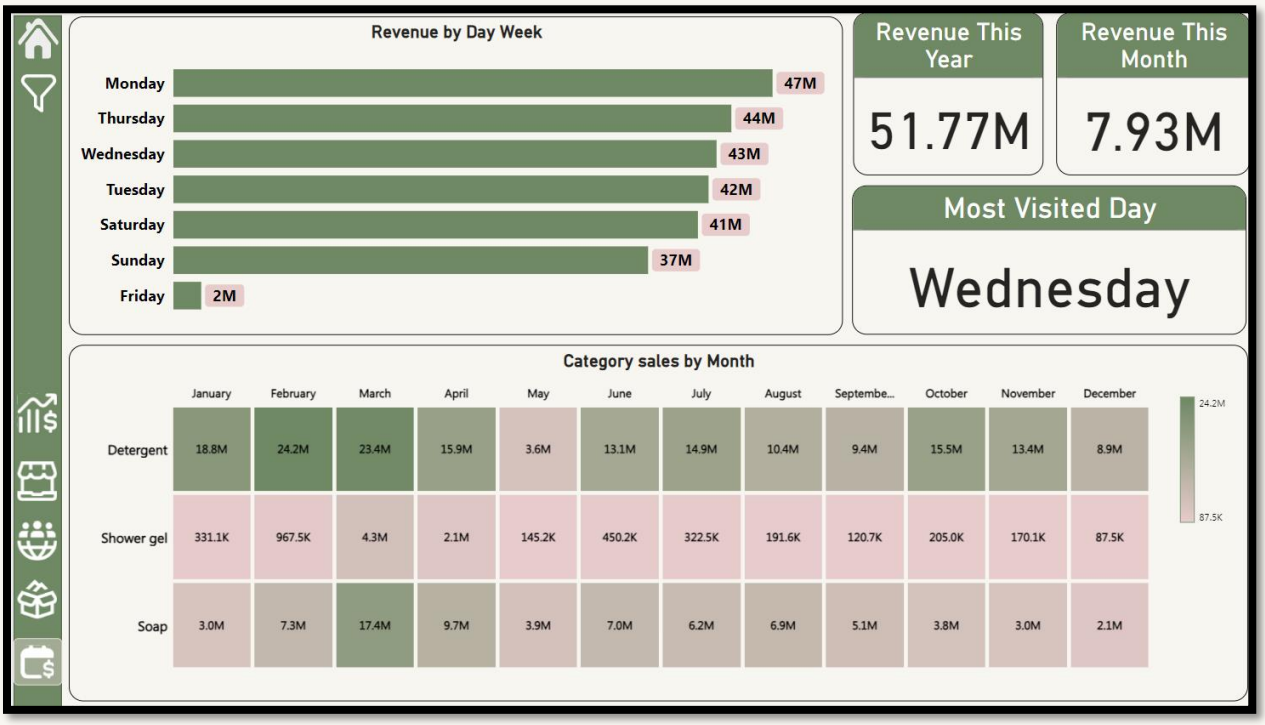
- Sales Person Page



- Product Page



- Date Page



4. Results and Key Findings

1. Sales Performance

- **March 2020** recorded the highest sales, generating 28 million.
- The **Delta region** emerged as the top performer, contributing 146 million to total sales.

2. Product Performance

- The **1-liter detergent** was the highest-selling product, contributing 66.9% of product sales.
- **Soap (115g)** had the highest return rate, prompting further investigation into product quality or customer preferences.

3. Outlet Visits

- **Retail outlets** accounted for 89% of total sales, with 3,250 out of 5,166 outlets being grocery retailers.
- The most visited outlet was **منظفات**, with 4,365 visits.
- **Mohamed Al-Saeedi** had the most product returns, warranting further training or review of sales strategies.

4. Employee Performance

- **Mohamed Hussein** was the top-performing employee, achieving 822% of his sales target.
- **Amgad** recorded the highest sales, totaling 29.1 million.

5. Conclusion

Our analysis provided crucial insights into the FMCG company's performance in the first quarter of 2021 and previous year. Key takeaways include:

- The **Delta region** and **March 2020** were standout performers.
- The **1-liter detergent** was the most successful product, while the **115g soap** had the highest return rate.
- **Retail outlets** were the dominant sales channel, particularly in the cleaning products category.
- Employee performance varied significantly, with some representatives exceeding targets while others struggled with returns.

These insights will help the company adjust its strategies for the second quarter, focusing on improving product quality, optimizing outlet support, and addressing employee performance gaps.

6. Recommendations

- Strengthen **quality control** for high-return products like soap (115g) to reduce returns.
- Focus on **outlet-specific support** for underperforming locations to boost sales.
- Continue incentivizing top-performing employees, while providing additional support and training for those facing challenges with returns.