**TARGET VS ACTUAL**

**SALES ANALYSIS REPORT**

Submitted By

*Penmetsa Navya*

[*navyapenmetsa94@gmail.com*](mailto:navyapenmetsa94@gmail.com)

**Table of Content**

1. Introduction
2. Objectives
3. Dataset Overview
   1. Actual Sales
   2. Target Sales
   3. Product Details
4. Data Pre-processing
5. Visualization
   1. Data Transformation
   2. Relationship Modelling
   3. DAX Calculation
   4. Visualizations Created
6. Outcomes and Recommendations
7. Conclusion
8. **Introduction**

Sales performance analysis is a critical task for optimizing revenue and profitability in any business. This project aims to analyse sales performance by comparing actual sales against targets, evaluating profitability, and identifying key trends across products, time periods, and geographies. The insights gained will support decision-making in optimizing sales strategies, inventory management, and pricing.

1. **Objectives**
   1. Compare actual sales against target sales to identify performance gaps and determine areas of improvement.
   2. Analyse sales trends over time to understand seasonality and growth patterns.
   3. Evaluate product-wise performance to identify high and low performers.
   4. Conduct profitability analysis to determine the most profitable products and regions.
   5. Assess the effectiveness of discount strategies on sales volume and profit margins.
   6. Analyse sales performance across different countries to identify the best-performing regions.
   7. Review product pricing strategies by comparing cost prices, sale prices, and profit margins.
2. **Dataset Overview**

The dataset consists of three sheets:

* 1. Actual Sales - Daily sales data with columns for
     1. *Date*: Daily sales dates.
     2. *Product Id*: Identifier for products.
     3. *Country*: The country where the sales occurred.
     4. *Product Category*: Category of the product (e.g., Standard, Elite).
     5. *Cost Price, Sale Price*: Unit cost and sale prices of products.
     6. *Quantity*: Number of units sold.
     7. *Cost Price Value, Sales Price Value*: Total cost and sales values.
     8. *Discount %*: Discount applied to the sales.
     9. *Final Sales Price*: Adjusted sales price after discounts.
     10. *Profit*: Profit earned on the sales.
  2. Target Sales - Provides sales targets for each product
     1. *Date*: Monthly dates for target sales.
     2. *Product Id*: Identifier for products.
     3. *Target Sale*: The target sales value for each product and month.
  3. Product Details - Includes product-specific data such as
     1. *Product Id*: Identifier for products.
     2. *Product Name*: Name of the product.
     3. *Product Category*: Category of the product.
     4. *Cost Price, Sale Price*: Unit cost and sale prices.

1. **Data Pre-processing**

Data pre-processing is a technique that is used to convert the raw data into a clean data set. The pre-processing pipeline was designed to clean and prepare the data for analysis and visualization. Python is used for pre-processing as it provides simple syntax, extensive libraries, compatibility, flexibility and integration and below are the steps taken

***Importing Libraries***

Essential libraries like Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Scipy and Pylab were used. These libraries provide robust tools for data cleaning, transformation, and visualization.

*Pandas*: Data manipulation and cleaning.

*NumPy*: Numeric computations.

*Scikit-learn*: Pre-processing (e.g., scaling, encoding).

*Matplotlib* and *Seaborn*: Visualizations for analysis.

*Scipy* and *Pylab*: Statistical analysis (e.g., Q-Q plots).

***Loading the Data***

All three sheets were loaded as separate DataFrames. Few methods like head, info, describe, shape were applied to gather information regarding each sheet.

The head() method returns a specified number of rows, string from the top. It returns the first 5 rows if a number is not specified. The info() method allows us to learn the shape of object types of our data. The describe() method gives us summary statistics for numerical columns in our DataFrame. The shape() method is used to fetch the dimensions of Pandas and NumPy type objects in python.

***Imputation***

Imputation is the process of replacing missing values with estimates. Missing data can distort analyses and models. Imputation ensures that the dataset is complete. Since neither of the datasets contained missing values no imputation was required.

***Removing Duplicates***

Duplicate records refer to repeated entries in the dataset. Duplicates can skew analysis and misrepresent the actual performance. Removed all identical rows to ensure accuracy and consistency.

***Handling Outliers***

Outliers are extreme values that lie outside the typical range of data. Outliers can mislead analyses, especially for metrics like averages or regression models. The Interquartile Range (IQR) method was used to identify and remove outliers in numeric columns such as quantity, cost price value, sales price value.

***Discretization***

Continuous data was converted into discrete bins or categories. This helps in creating more interpretable visualizations and analyses. Discount % were grouped into ranges like "Low," "Medium," and "High."

***Encoding***

Categorical columns were encoded into numeric formats. Machine learning models and statistical tools often require numeric inputs. Product Category column was encoding using Label Encoder. Label encoder encode target labels with value between 0 and n\_classes-1.

***Standardization***

Standardization scales the data to have a mean of 0 and a standard deviation of 1. It ensures all features contribute equally to analyses, especially in distance-based models. StandardScaler() function was used to scale Final Sales Price.

***Normalization***

Normalization scales the data to a range of [0, 1]. This is helpful for visualizations and models sensitive to input ranges. MinMaxScaler() function was used to normalize Discount % column.

***Q-Q Plots***

Quantile-Quantile (Q-Q) plots were used to check if numeric columns followed a normal distribution. Many statistical tests and models assume normality in the data. If deviations were found, transformations like logarithms and power were applied.

1. **Visualization**

Data visualization is the process of using visual elements to represent data, such as charts, graphs, maps, plots, infographics, and animations. It translates complex, high-volume, or numerical data into a visual representation that is easier to process.

Power BI provides dynamic and interactive visualizations, making it easier to identify patterns and trends. It also supports data transformation and relationship modelling, which is essential for this project.

* 1. Data Transformation
     1. Import Data: Loaded all three cleaned excel files into Power BI.
     2. Applied data types (e.g., Date, Decimal, Text)
     3. Removed similar columns (Cost Price, Sales Price)
  2. Relationship Modelling

Created relationships between tables:

Actual Sales[Product ID] → Product Details[Product ID]

Actual Sales[Date] → Target Sales[Date]

Target Sales[Product ID] → Product Details[Product ID]

* 1. DAX Calculations
     1. Measures Created:
        1. Total Actual Sales = SUM(ActualSales[SalesPrice])
        2. Total Target Sales = SUM(TargetSales[Target Sale])
        3. Total Profit = SUM(ActualSales[Profit])
        4. Average Discount = AVERAGE(ActualSales[Discount %])
        5. Sales vs Target (%) = DIVIDE([Total Actual Sales], [Total Target Sales], 0 )
     2. Calculated Columns:
        1. Profit Margin: (SalesPrice - CostPrice) / SalesPrice
  2. Visualizations Created
     1. Target vs Actual
        1. Chart Type: Clustered Column Chart
        2. X-axis: Product Categories
        3. Y-axis: Sales Amount (Actual and Target)
     2. Sales Target
        1. Chart Type: KPI
        2. Value: Total Actual Sales
        3. Trend Axis: Date
        4. Target: Total Target Sales
     3. Sales Trend
        1. Chart Type: Line Chart
        2. X-axis: Date
        3. Y-axis: Final Sales Price
     4. Cards
        1. Total Actual Sales
        2. Total Profit
     5. Sales by Category
        1. Chart Type: Area Chart
        2. X-axis: Date
        3. Y-axis: Final Sales Price
        4. Legend: Product Category
     6. Sales by Country
        1. Chart Type: Clustered Column Chart
        2. X-axis: Country
        3. Y-axis: Total Actual Sales
     7. Profit by Country
        1. Chart Type: Clustered Column Chart
        2. X-axis: Country
        3. Y-axis: Profit
     8. Sales & Profit by Discount
        1. Chart Type: Line & Stacked Column Chart
        2. X-axis: Discount Bin
        3. Y-axis: Final Sales Price
        4. Line Y-axis: Profit Margin
     9. Sales & Profit by Category
        1. Chart Type: Line & Clustered Column Chart
        2. X-axis: Product Category
        3. Y-axis: Final Sales Price
        4. Line Y-axis: Profit Margin
     10. Sales by Product
         1. Chart Type: Donut Chart
         2. Legend: Product Name
         3. Values: Final Sales Price
     11. Profit by Product
         1. Chart Type: Treemap
         2. Legend: Product Name
         3. Values: Profit
     12. Country & Category
         1. Chart Type: Decomposition Tree
         2. Analyze: Total Actual Sales
         3. Explain by: Country, Product Category
     13. Discount Impact
         1. Chart Type: Scatter Chart
         2. X-axis: Discount %
         3. Y-axis: Total Actual Sales
         4. Values: Country
         5. Size: Profit Margin
     14. Profit Trend
         1. Chart Type: Line Chart
         2. X-axis: Date
         3. Y-axis: Profit
     15. Sales Forecast
         1. Chart Type: Line Chart
         2. X-axis: Date
         3. Y-axis: Total Actual Sales
         4. Forecast: 4 Quarters
     16. Sales by Country
         1. Chart Type: Maps
         2. Location: Country
         3. Category: Product Category
         4. Bubble: Total Actual Sales
     17. Profit by Country
         1. Chart Type: Maps
         2. Location: Country
         3. Category: Product Category
         4. Bubble: Total Profit
     18. Slicers
         1. Date
         2. Discount Bin
         3. Product Category

1. **Outcomes and Recommendations**
   1. Key Insights
      1. Sales gaps were identified in specific regions and products.
      2. Discounts above 30% reduced profitability significantly.
      3. Certain regions consistently performed better, warranting further investment.
   2. Recommendations
      1. Focus on high-margin products and regions for promotions.
      2. Optimize discounts for volume without eroding profits.
      3. Reassess pricing strategies for low-performing products.
2. **Conclusion**

This project provided actionable insights into sales performance using a robust combination of Python pre-processing and Power BI visualization. The findings will help guide strategic decisions, improve operational efficiency, and enhance overall profitability.