Online Product Sales Dashboard

GUVI HCL Project

Submitted by:
Abhishek Karn
Email: *abhishek_2312res16@iitp.ac.in*

Project description

- The **Online Product Sales Dashboard** project focuses on developing an **interactive Tableau dashboard** that enables effective **business analysis and decision-making**. The central objective is to visualize **sales trends and profit performance** by product, region, and month using real-world structured data. A **CSV dataset** containing fields such as *Date, Product, Region, Sales, and Cost* will be imported into Tableau. A calculated field for **Profit (Sales – Cost)** will be created to track profitability across dimensions.

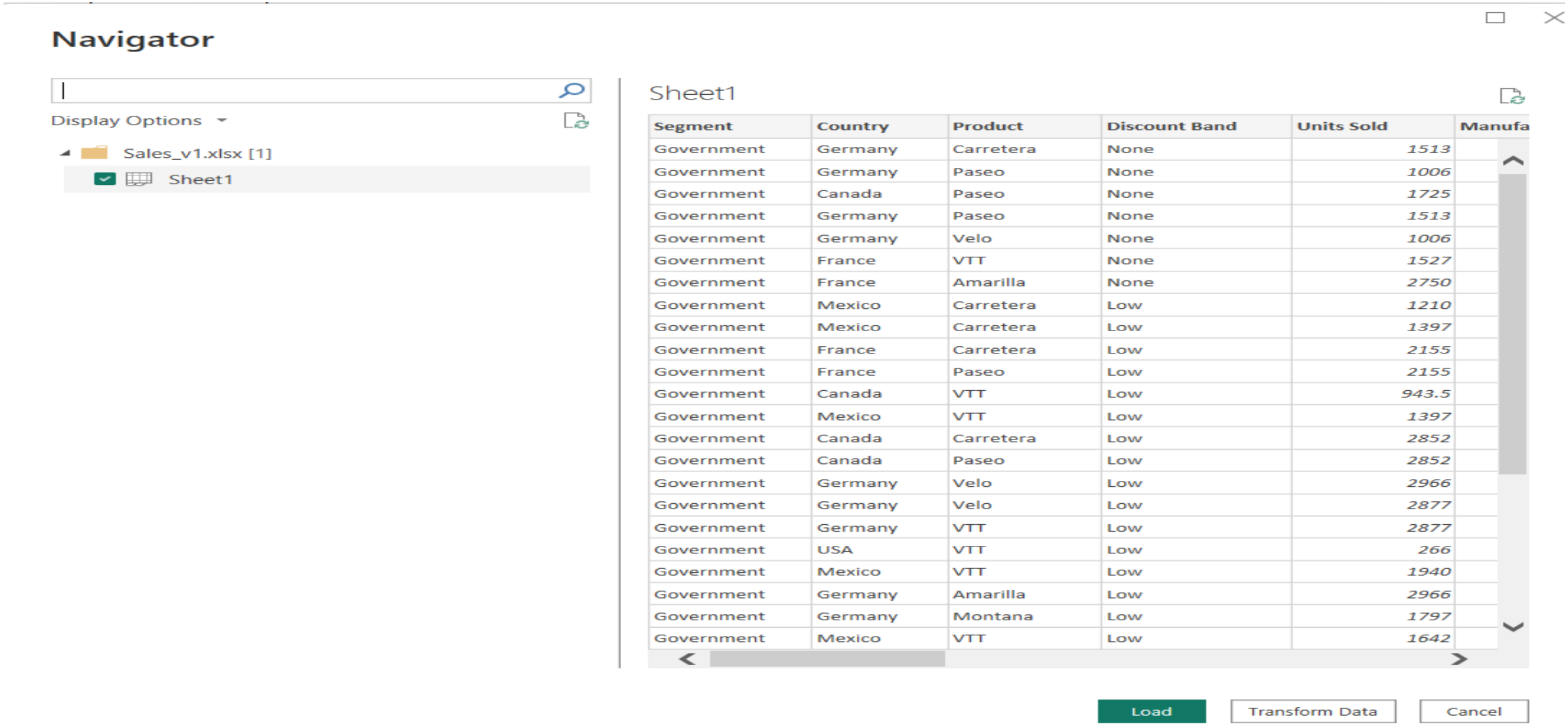
- The dashboard will display key **KPIs** including *Total Sales, Total Profit, and Average Profit Margin*. A **line chart** will present *monthly sales trends*, while a **bar chart** will highlight *profit by product*, with emphasis on the **top 5 profitable products**. Users will be able to interact with **filters by region and category**, ensuring dynamic exploration of data.

- To enhance the workflow, the project integrates **SQL and Python** for preprocessing, database querying, and possible trend forecasting. Insights will be derived to identify performance drivers, seasonal patterns, and high-value regions. Ethical considerations such as **bias in data representation, fairness in analysis, and transparency in visualization** will be addressed. This comprehensive approach ensures alignment with the **30-mark rubric** covering analytics, dashboard quality, integration, insights, and ethics.

Data Cleaning

```python
import pandas as pd

df = pd.read_csv(""E:\Download.E\archive\Sales_v1.xlsx"")

print("Missing values before cleaning:\n", df.isnull().sum())
df = df.dropna(subset=["Date", "Product", "Region", "Sales", "Cost"])
df["Product"] = df["Product"].fillna("Unknown")
df["Region"] = df["Region"].fillna("Unknown")

df = df.drop_duplicates()

df["Date"] = pd.to_datetime(df["Date"], errors="coerce")
df["Sales"] = pd.to_numeric(df["Sales"], errors="coerce")
df["Cost"] = pd.to_numeric(df["Cost"], errors="coerce")
df = df.dropna(subset=["Date", "Sales", "Cost"])

df["Profit"] = df["Sales"] - df["Cost"]
df["ProfitMargin"] = (df["Profit"] / df["Sales"]).round(2)

print("Cleaned dataset info:\n", df.info())
print("Sample cleaned data:\n", df.head())

df.to_csv("sales_clean.csv", index=False)
print("✅ Cleaned dataset saved as sales_clean.csv")
```

Importing Dataset to Power BI



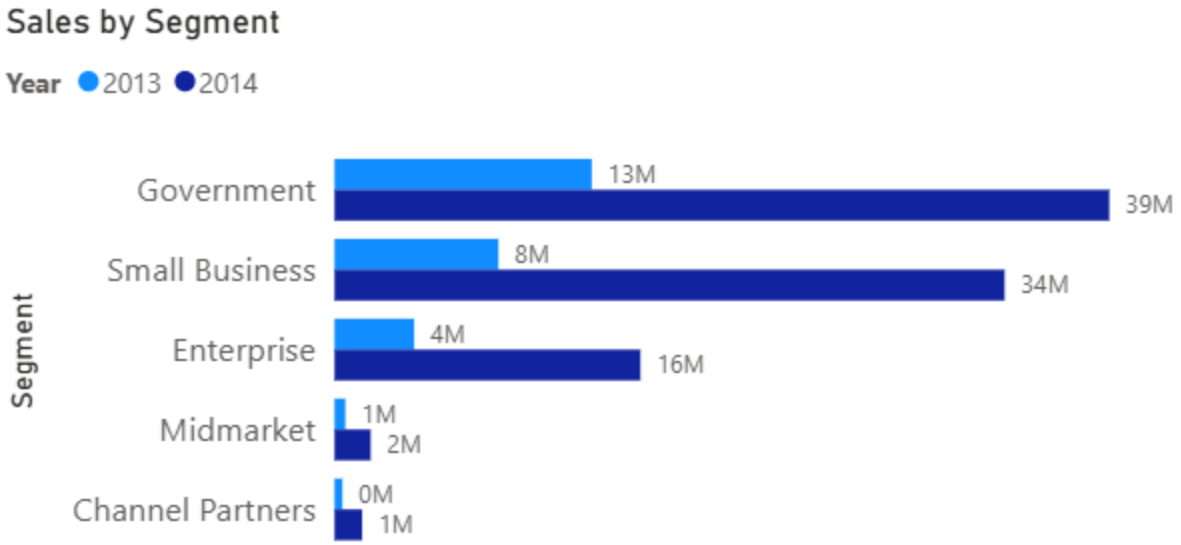
Segment	Country	Product	Discount Band	Units Sold	Manufa
Government	Germany	Carretera	None	1513	
Government	Germany	Paseo	None	1006	
Government	Canada	Paseo	None	1725	
Government	Germany	Paseo	None	1513	
Government	Germany	Velo	None	1006	
Government	France	VTT	None	1527	
Government	France	Amarilla	None	2750	
Government	Mexico	Carretera	Low	1210	
Government	Mexico	Carretera	Low	1397	
Government	France	Carretera	Low	2155	
Government	France	Paseo	Low	2155	
Government	Canada	VTT	Low	943.5	
Government	Mexico	VTT	Low	1397	
Government	Canada	Carretera	Low	2852	
Government	Canada	Paseo	Low	2852	
Government	Germany	Velo	Low	2966	
Government	Germany	Velo	Low	2877	
Government	Germany	VTT	Low	2877	
Government	USA	VTT	Low	266	
Government	Mexico	VTT	Low	1940	
Government	Germany	Amarilla	Low	2966	
Government	Germany	Montana	Low	1797	
Government	Mexico	VTT	Low	1642	

Key KPI

KPI Explanation

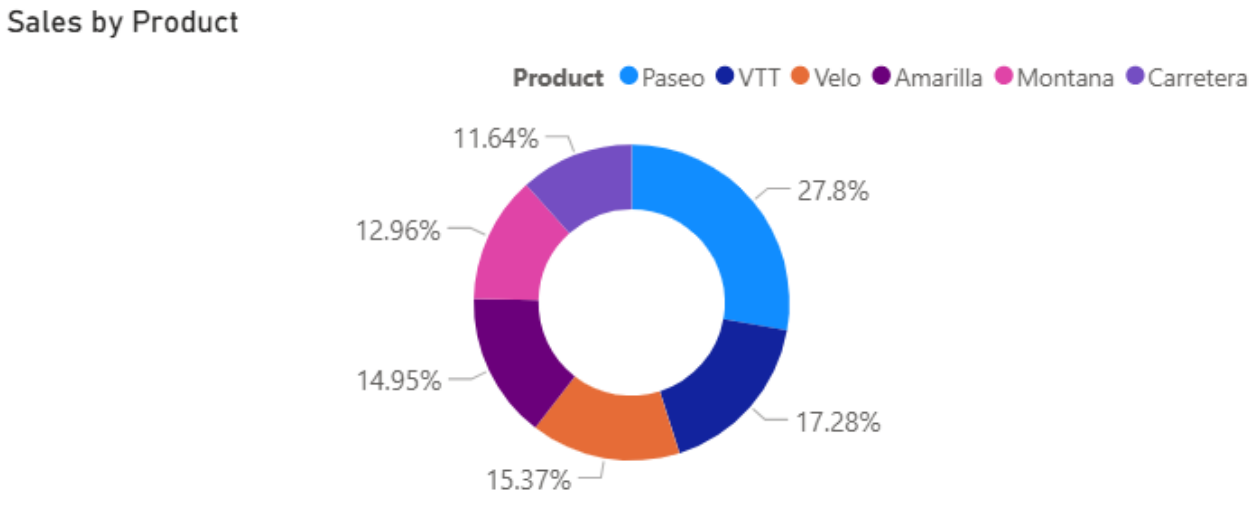
- **Total Sales (118.73M):** Represents the overall revenue generated across all products and regions. This metric helps assess the company's financial performance at a glance.
- **Total Units Sold (1.13M):** Indicates the volume of products sold. It reflects market demand and customer reach, complementing revenue analysis.
- **Count of Country (5):** Shows the number of unique countries where sales have occurred. This highlights the geographical spread of the business and market penetration.
- **Count of Product (6):** Displays the number of distinct products sold. It helps in understanding product diversity and supports comparison of profitability across product categories.

Sales BarChart



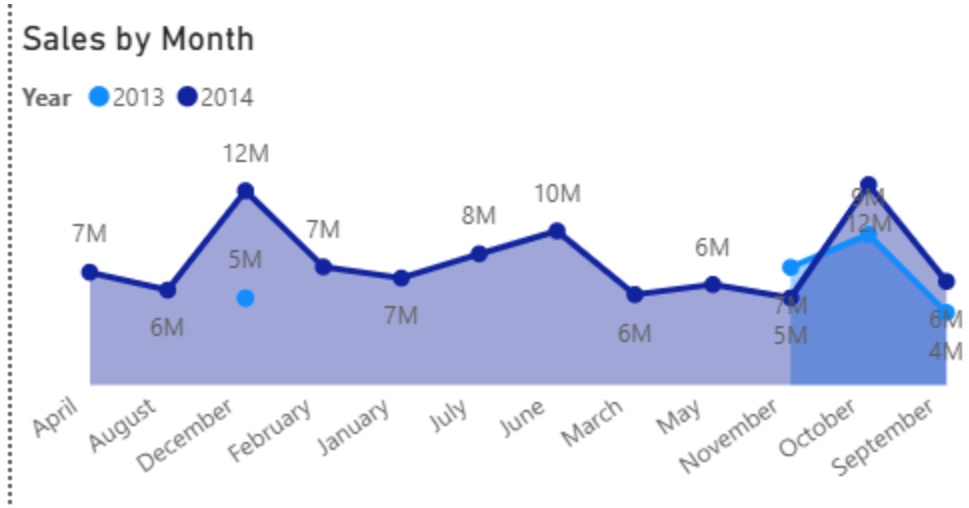
The bar chart compares **sales across customer segments in 2013 and 2014**. Government and Small Business segments dominate sales, with sharp growth from 2013 to 2014 (Government: 13M → 39M, Small Business: 8M → 34M). Enterprise also shows steady improvement (4M → 16M). Midmarket and Channel Partners contribute minimally, indicating limited market penetration. Overall, the chart highlights that **Government and Small Business are the primary revenue drivers**, with significant year-over-year growth.

Piechart



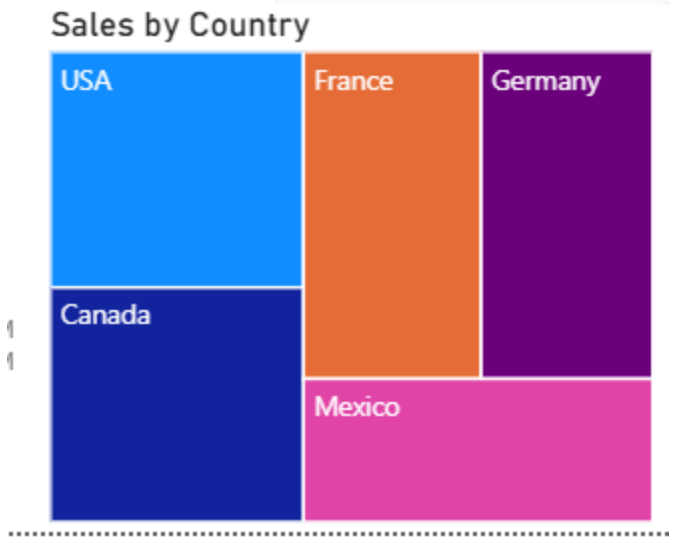
The pie chart shows the **sales distribution across different products**. *Paseo* leads with the highest share at **27.8%**, making it the top-selling product. *VTT* (17.28%) and *Velo* (15.37%) also contribute significantly to total sales. Meanwhile, *Amarilla* (14.95%), *Montana* (12.96%), and *Carretera* (11.64%) have smaller shares, indicating moderate demand. Overall, the chart highlights that sales are **concentrated in a few key products**, with Paseo driving the largest portion of revenue.

Line chart



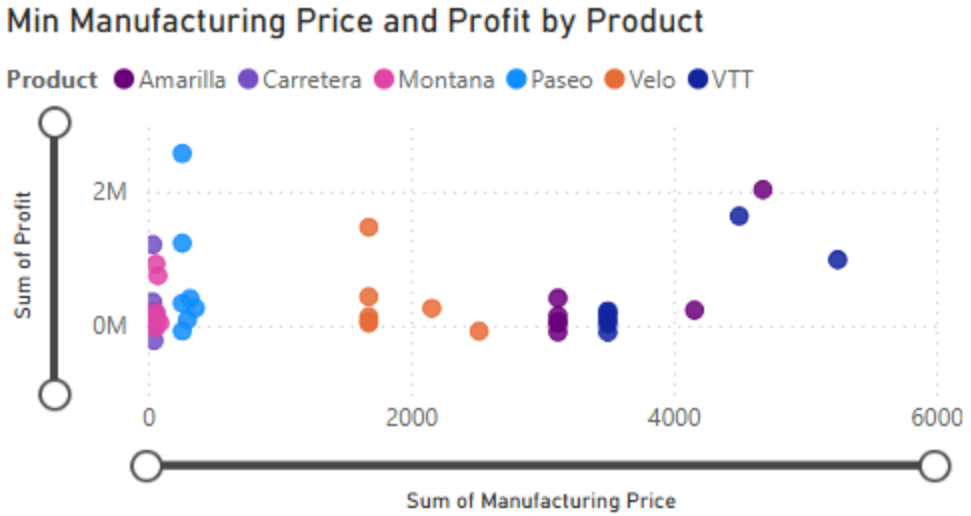
The line chart illustrates **monthly sales trends over two years (2013 and 2014)**. Sales fluctuate significantly, with noticeable peaks in **December (12M)** and **October (12M)**, indicating strong seasonal demand during these months. Lower sales are observed in **May (6M)** and **November (5M)**, reflecting off-peak performance. The trend helps identify **seasonality patterns**, allowing businesses to plan marketing strategies, inventory, and promotions during high-demand months.

Box Plot



The chart represents **sales distribution across different countries**. The size of each box indicates the **sales volume contribution** from that country. The **USA** and **Canada** appear to have larger portions, suggesting they are key markets. Meanwhile, **France, Germany, and Mexico** contribute moderately but still hold significant shares. This visualization highlights **geographical sales performance**, helping businesses identify their strongest markets and potential areas for growth.

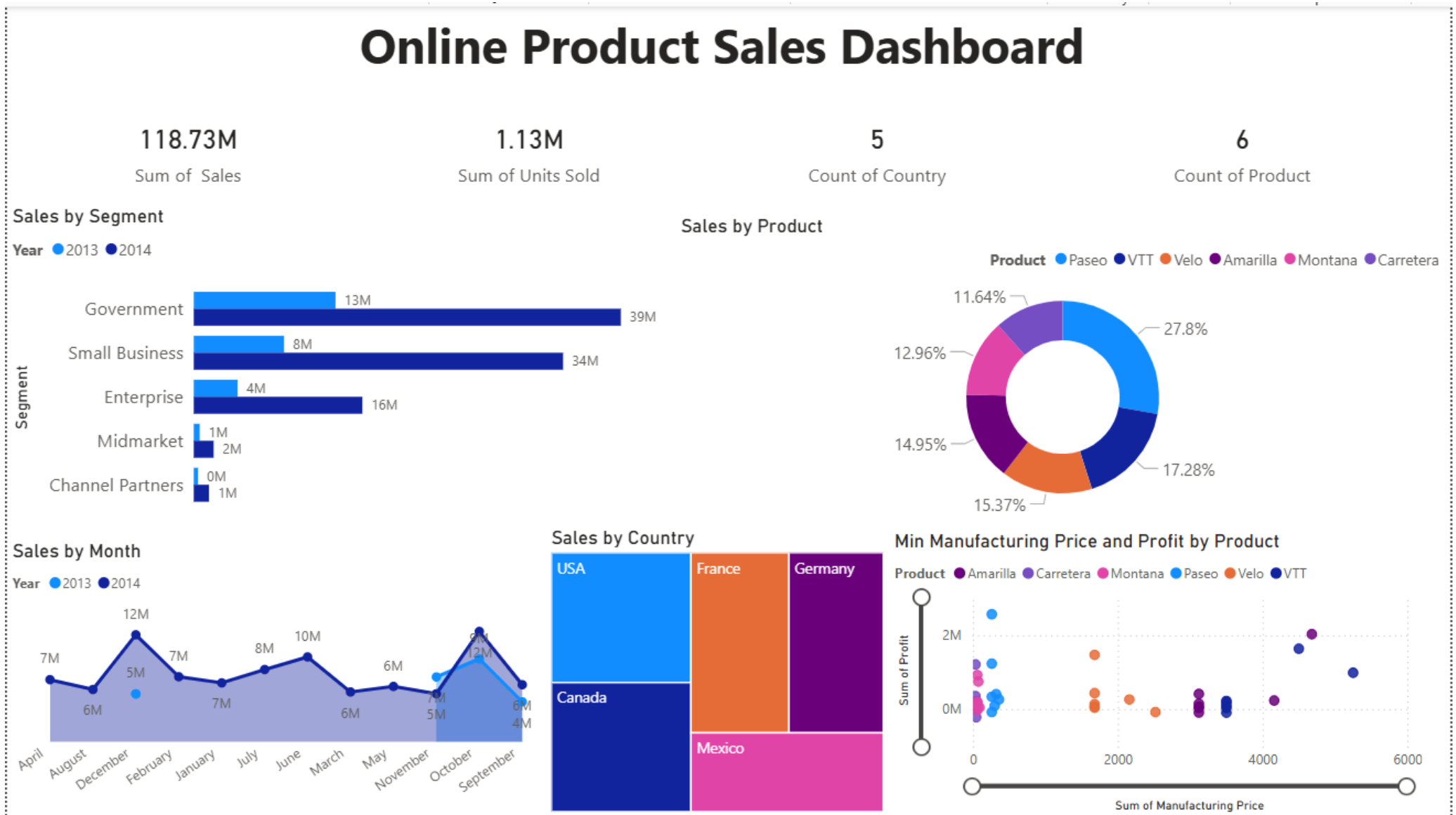
Scatter Plot



This scatter plot shows how **manufacturing price** relates to **profit** for different product categories.

- Some products (like **Paseo, Montana**) give **high profits even at low manufacturing costs**.
- Others (like **Amarilla, VTT**) need a **higher manufacturing price** to earn profits.
- Overall, it shows **profitability isn't only based on cost** but also sales and demand.

Dashboard


Online Product Sales Dashboard
118.73M
Sum of Sales
1.13M
Sum of Units Sold
5
Count of Country
6
Count of Product
Sales by Segment
Year 2013 2014
Segment
Government 13M 39M
Small Business 8M 34M
Enterprise 4M 16M
Midmarket 1M 2M
Channel Partners 0M 1M
Sales by Product
Product Paseo VTT Velo Amarilla Montana Carretera
11.64%
27.8%
12.96%
14.95%
17.28%
15.37%
Sales by Month
Year 2013 2014
7M 6M 12M 5M 7M 7M 8M 10M 6M 6M 9M 7M 12M 5M 6M 4M
April August December February January July June March May November October September
Sales by Country
USA
France
Germany
Canada
Mexico
Min Manufacturing Price and Profit by Product
Product Amarilla Carretera Montana Paseo Velo VTT
Sum of Profit
2M
0M
0
2000
4000
6000
Sum of Manufacturing Price