



Project Title: Sales Data Visualization

- **Submitted By:** Ananya Sharma
- **Institution:** KIET Group of Institutions
 - **Course:** B. Tech CSE (AI)
 - **Date:** 03/11/2025

Introduction

Data visualization is an essential technique in the field of data science and business analytics. It helps in transforming raw data into graphical representations, making complex data easier to understand and analyze. In the context of sales analysis, visualization provides valuable insights into product performance, sales trends, and revenue distribution.

This project focuses on visualizing sales data to gain a deeper understanding of how different products perform over time. By employing effective data visualization techniques, the project aims to simplify data interpretation and support informed decision-making. With the use of bar charts and line plots, we aim to represent sales data clearly and concisely, allowing stakeholders to make data-driven decisions efficiently.

Problem Statement

The goal of this project is to analyze and visualize sales data to gain insights into product performance and revenue generation. Sales data from multiple products, including phones, laptops, monitors, and tablets, will be used to create visualizations that showcase revenue patterns and sales trends over time.

Methodology

1. **Data Collection:** Collected sales data from the provided CSV file.
2. **Data Preprocessing:** Cleaned and formatted the data for visualization.
3. **Data Visualization:** Used bar charts and line plots to visualize revenue and sales trends.
 - Bar Chart: Total revenue by product.
 - Line Plot: Units sold over time for each product.
4. **Analysis:** Analyzed the visualizations to identify patterns and trends.

Code

- ```
import pandas as pd
import matplotlib.pyplot as plt

Load the sales data from the CSV file
data = pd.read_csv('sales.csv')

Convert Date column to datetime format
data['Date'] = pd.to_datetime(data['Date'])

Total Revenue by Product (Bar Chart)
plt.figure(figsize=(8, 5))
product_revenue = data.groupby('Product')['Revenue'].sum()
product_revenue.plot(kind='bar', color='green')
plt.title('Total Revenue by Product')
```

```
plt.xlabel('Product')
plt.ylabel('Revenue')
plt.xticks(rotation=45)
plt.show()
```

- import pandas as pd  
import matplotlib.pyplot as plt

```
Load the sales data from the CSV file
```

```
data = pd.read_csv('sales.csv')
```

```
Convert Date column to datetime format
```

```
data['Date'] = pd.to_datetime(data['Date'])
```

```
Plotting Units Sold Over Time for Each Product
```

```
plt.figure(figsize=(10, 6))
```

```
Iterate over each unique product and plot separately
```

```
for product in data['Product'].unique():
```

```
 product_data = data[data['Product'] == product]
```

```
 plt.plot(product_data['Date'], product_data['UnitsSold'], marker='o',
 linestyle='-', label=product)
```

```
Adding plot details
```

```
plt.title('Units Sold Over Time by Product')
```

```
plt.xlabel('Date')
```

```
plt.ylabel('Units Sold')
```

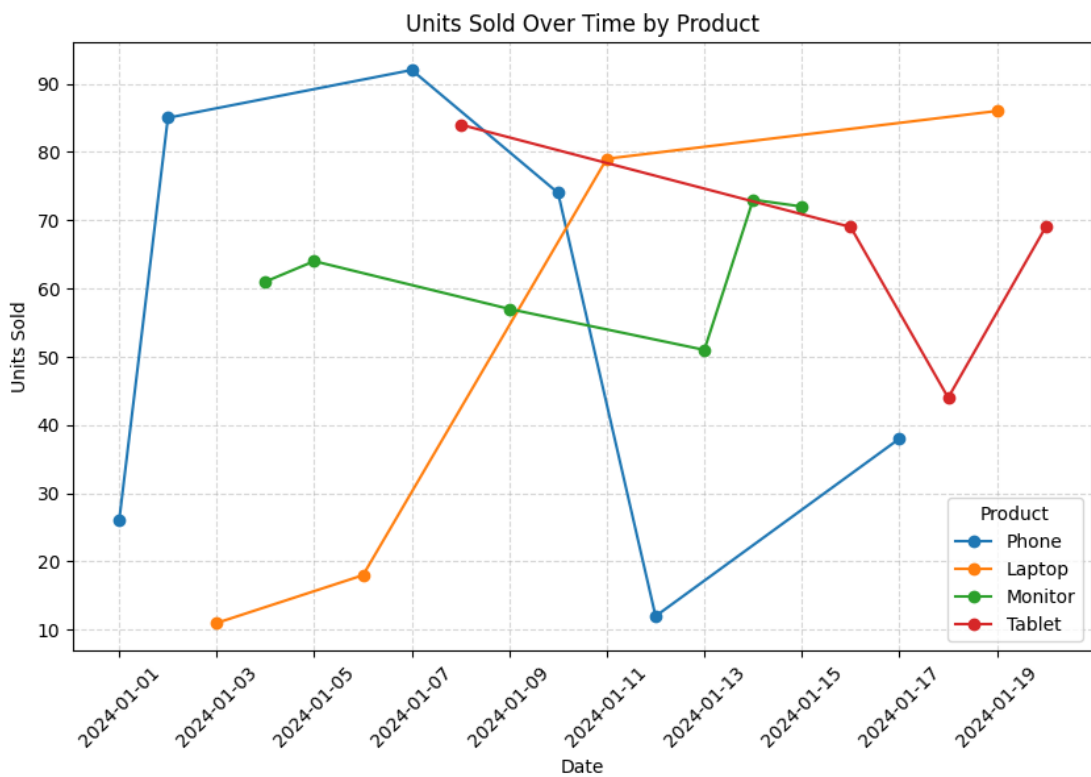
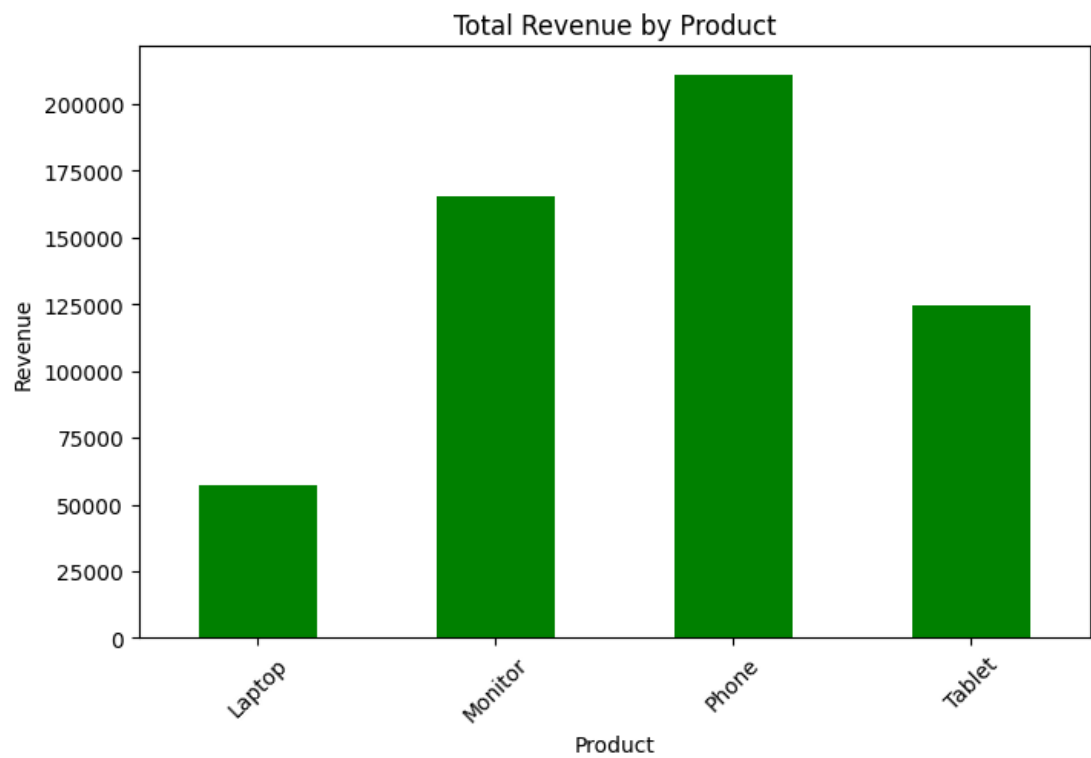
```
plt.legend(title='Product')
```

```
plt.grid(True, linestyle='--', alpha=0.5)
```

```
plt.xticks(rotation=45)
```

```
plt.show()
```

# Output/Results



## References

- Given dataset : “sales.csv” .
- Matplotlib Documentation: <https://matplotlib.org>
- Pandas Documentation: <https://pandas.pydata.org>