

Advanced data visualization

Experiment-1

| | |
|------------|--------------|
| Name | Parth Gandhi |
| UID | 2021300033 |
| Batch | Batch H |
| Department | COMSP A |

Aim: Analyze e-commerce dataset using various charts like bar graph, pie chart, time series charts and observe product wise and region wise sales.

Description of dataset

The dataset used is an ecommerce dataset. The dataset has 8 attributes and 541,909 records. The detailed attributes are:

1. **InvoiceNo:** Its type is categorical and it is used as a unique identifier for each transaction (or invoice).
2. **StockCode:** Its type is categorical and it is used as a unique identifier for each distinct product.
3. **Description:** Its type is text and it describes the name or description of the product.
4. **Quantity:** Its type is numerical and it describes the number of units of the product. Negative values indicate product returns.
5. **InvoiceDate:** Its type is text but represents DateTime. It describes date and time when the transaction occurred.
6. **UnitPrice:** Its type is float (Numerical). It delineates the price per unit of the product in GBP (British Pound).
7. **CustomerID:** Its type is numerical, but represents ID. It is used as a unique identifier for each customer.
8. **Country:** Its type is categorical. It describes the country where the customer resides.

Product-wise Sales Analysis:

- Using StockCode, Description, Quantity, and UnitPrice, you can calculate total sales for each product.
- You can use bar charts and pie charts to visualize the sales distribution across different products.

Region-wise Sales Analysis:

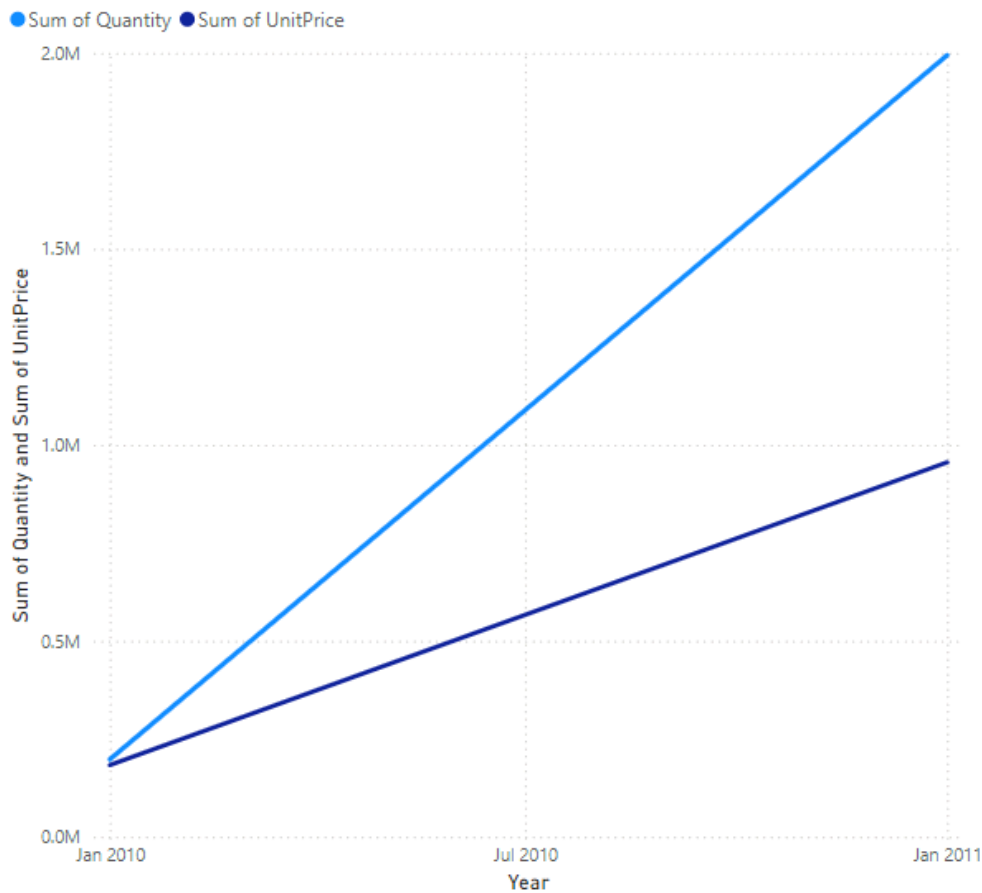
- The Country column will help you analyze sales distribution across different regions.
- Bar charts and pie charts can illustrate which regions contribute the most to sales.

Time Series Analysis:

- InvoiceDate allows you to perform time series analysis to observe sales trends over time.
- You can use time series charts to track how sales fluctuate daily, monthly, or seasonally.

Analysis of charts

Sum of Quantity and Sum of UnitPrice by Year

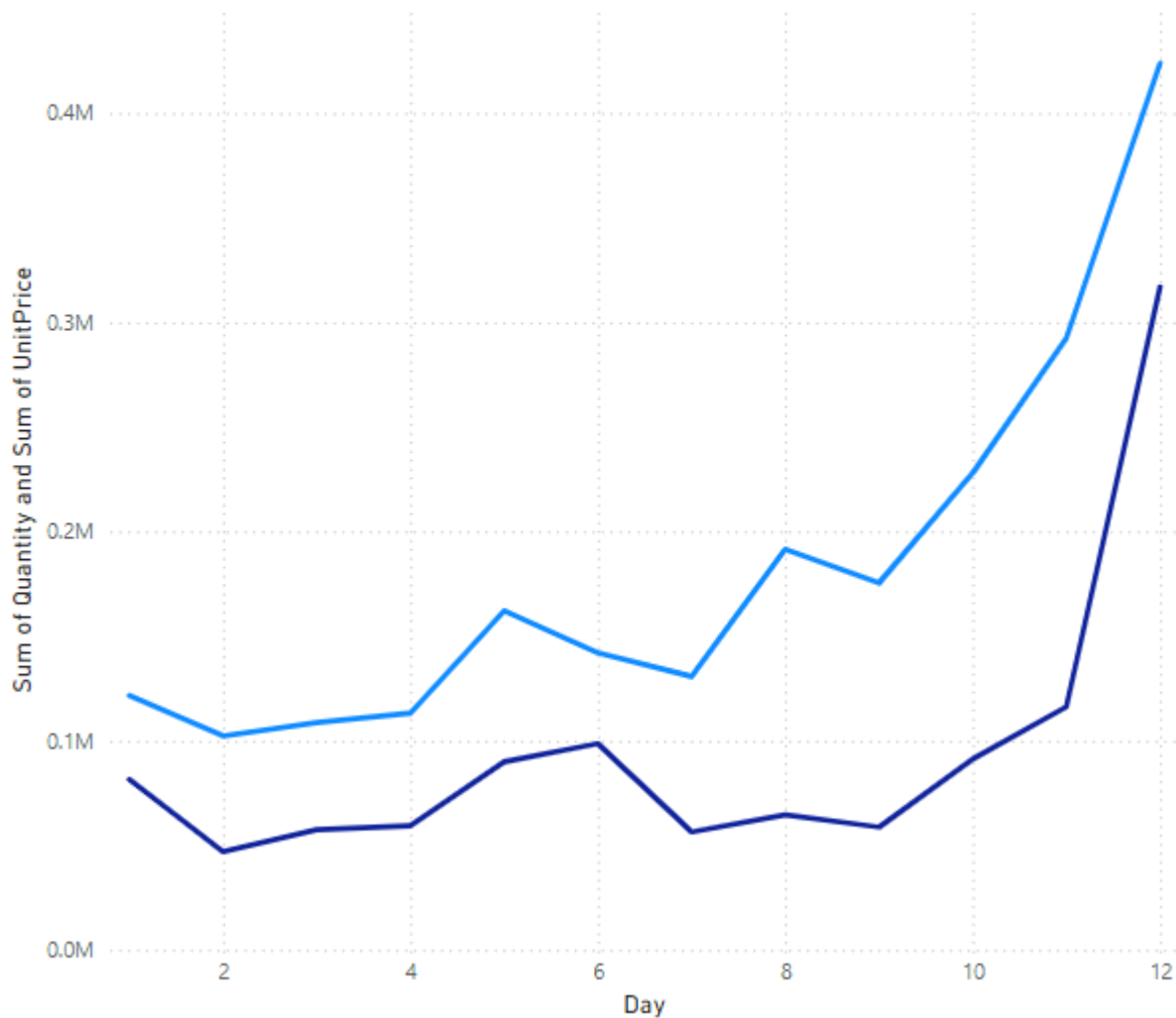


Line Chart (Sum of Quantity and Sum of UnitPrice by Year): This chart illustrates the trend of the total quantity sold and the total unit price over a given period, presumably from January 2010 to January 2011.

- The blue line represents the **Sum of Quantity** sold, while the darker blue line represents the **Sum of UnitPrice**.
- Both metrics show an increasing trend over time, with the Sum of Quantity increasing more rapidly than the Sum of UnitPrice.

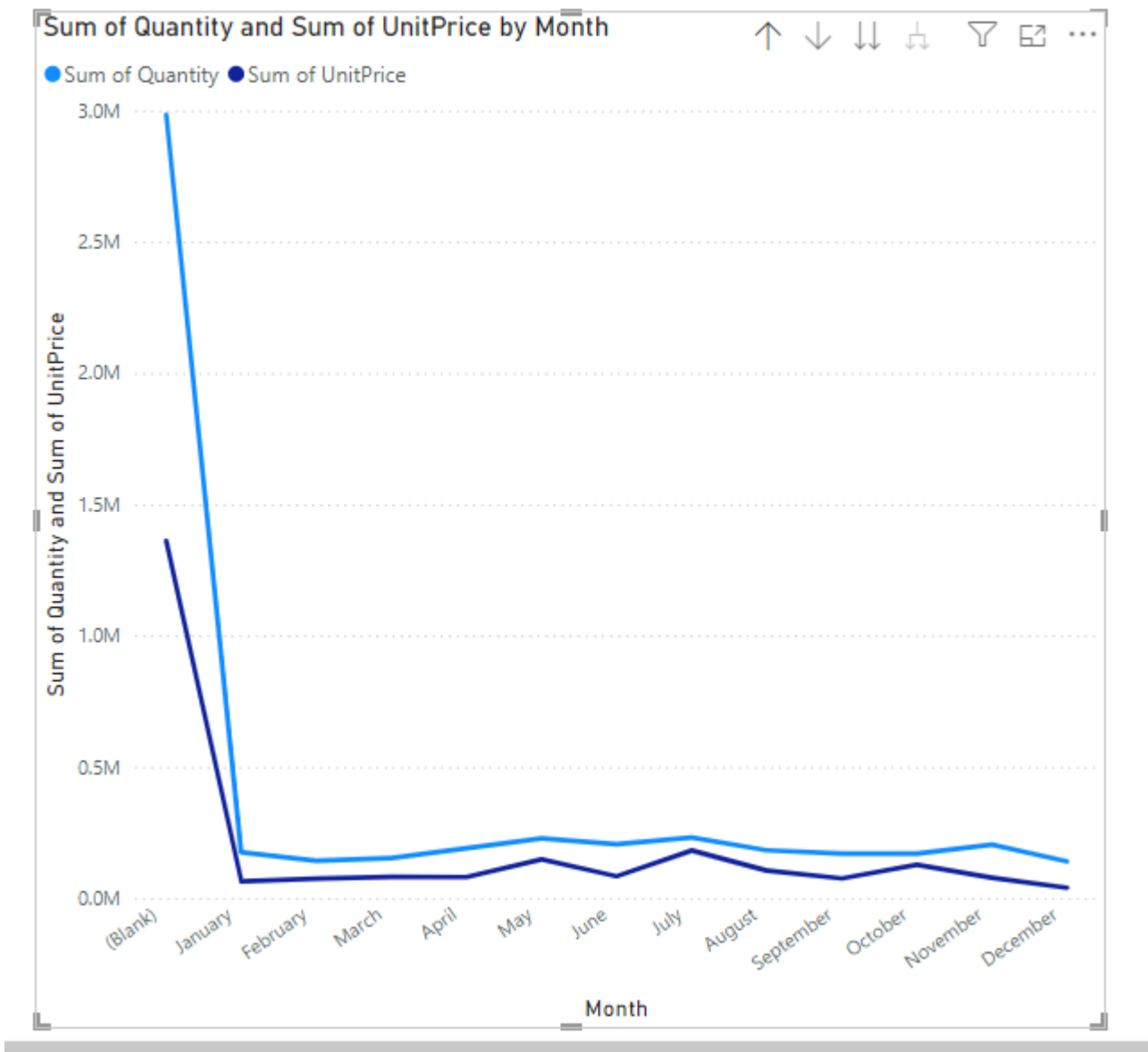
Sum of Quantity and Sum of UnitPrice by Day

● Sum of Quantity ● Sum of UnitPrice



1. By Day

- This line chart shows the daily trends for the **Sum of Quantity** and **Sum of UnitPrice**.
- The blue line representing the Sum of Quantity generally trends upwards, indicating a daily increase in the quantity of products sold.
- The darker blue line representing the Sum of UnitPrice follows a similar trend, with noticeable increases towards the end of the observed period.
- This could suggest a buildup in sales activity towards the end of the month, possibly due to promotions or consumer behavior patterns.



2. By Month

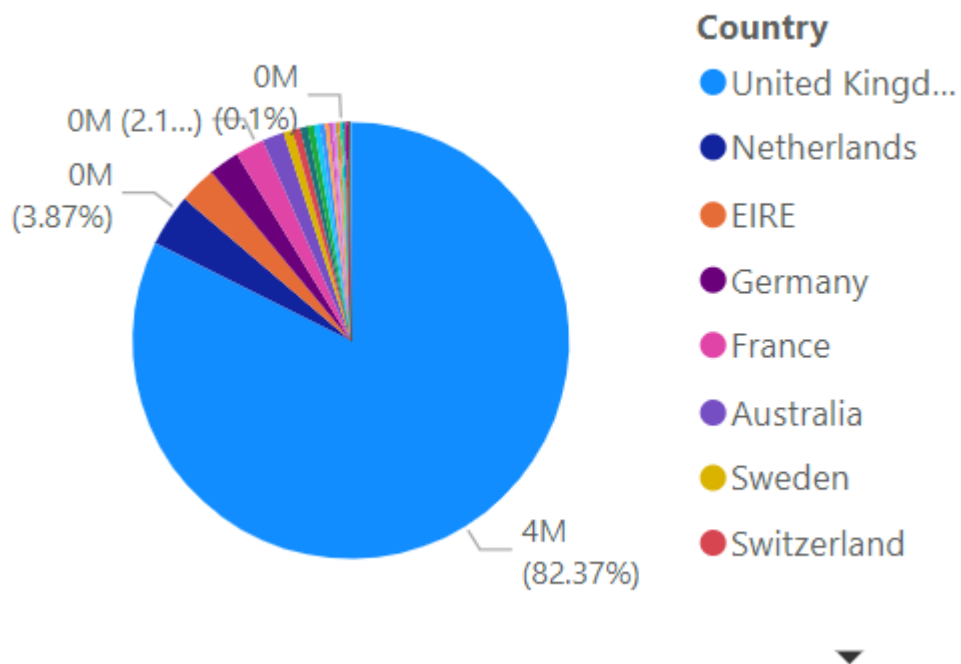
- This chart presents the monthly aggregated Sum of Quantity and Sum of UnitPrice.
- There is a significant spike in sales in January, followed by a sharp decline and stabilization in the following months.
- The large spike in January might indicate a major sales event or seasonal demand, such as post-holiday shopping or New Year's promotions.
- The rest of the months show much lower and relatively consistent sales, indicating typical monthly sales patterns without major fluctuations.



3. By Quarter

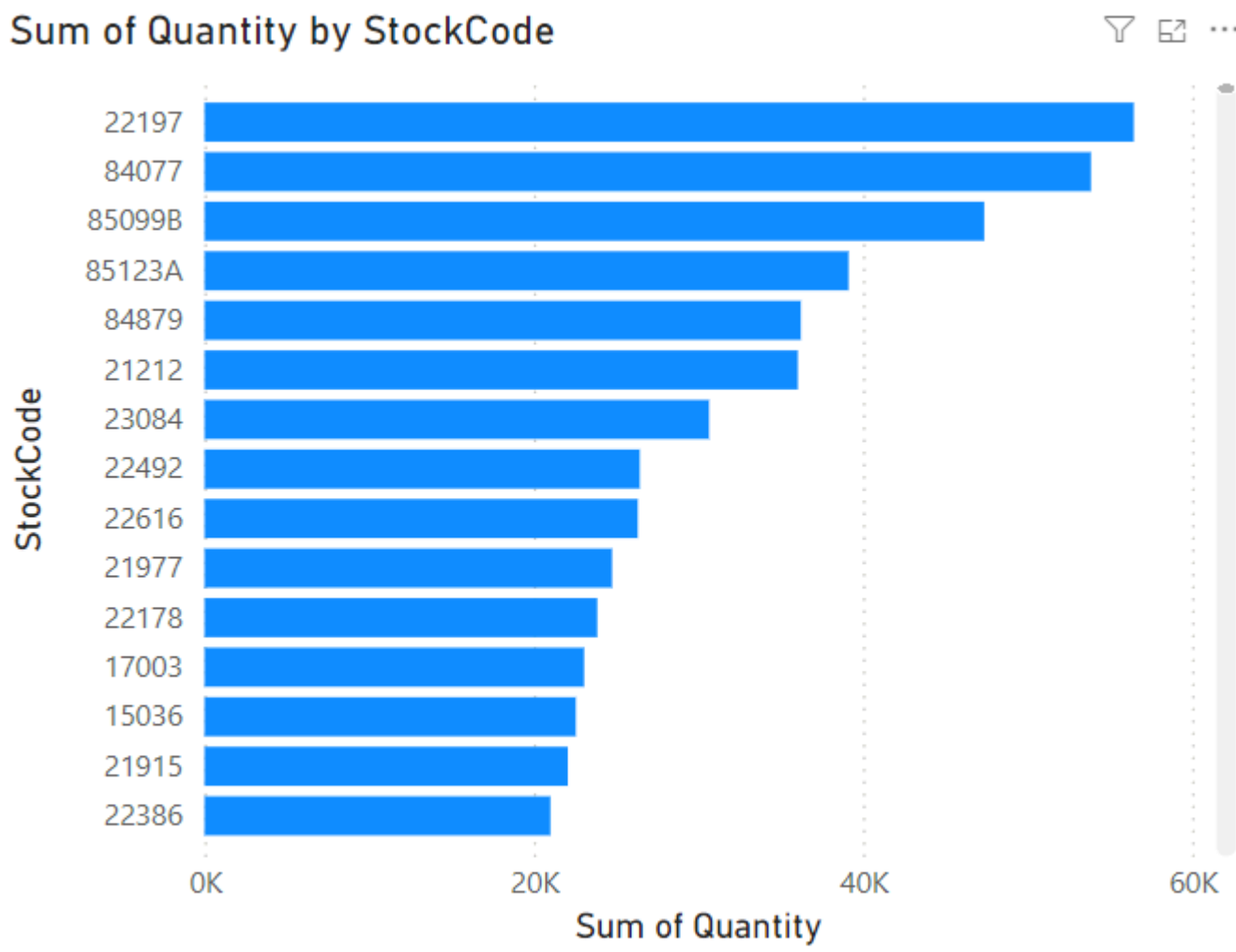
- The quarterly analysis also shows a significant peak in the first quarter (Qtr 1), followed by a notable drop in Qtr 2 and gradual decline through Qtr 4.
- This pattern reinforces the observation from the monthly chart, with strong sales activity at the beginning of the year that tapers off as the year progresses.
- The sharp decrease after Q1 could be due to the absence of major sales events or seasonal factors that drive consumer spending early in the year.

Sum of Quantity by Country



Pie Chart (Sum of Quantity by Country): This chart provides a visual breakdown of the total quantity of products sold by country.

- The United Kingdom dominates the sales, accounting for 82.37% of the total quantity sold, as indicated by the large blue segment.
- Other countries like the Netherlands, EIRE, Germany, France, Australia, Sweden, and Switzerland contribute much smaller portions, each represented by thinner slices of the pie.



Bar Chart (Sum of Quantity by StockCode): This chart highlights the sum of the quantity of products sold, segmented by different stockcodes.

- The stock code 22197 has the highest quantity sold, with just under 60,000 units. This suggests that the product associated with this stock code is highly popular or frequently purchased
- Stock codes 84077 and 85099B also have significantly high sales, with quantities approaching or slightly above 40,000 units. These are likely to be other top-performing products..
- The difference in quantity between each subsequent stock code decreases gradually, indicating that the top products are well-distributed in terms of sales, though with notable leaders.

Conclusion: In conclusion, the analysis of the e-commerce dataset has provided valuable insights into product and region-wise sales performance. The use of various visualizations, including bar charts, pie charts, and time series charts, has revealed key patterns in consumer behavior. Specifically, certain products consistently outperform others, as seen in the high quantities sold for top stock codes like `22197`. Region-wise analysis further highlights the concentration of sales in specific markets, with the UK dominating the ecommerce market and potential opportunities for growth in underperforming areas.

Overall, this experiment underscores the importance of data-driven decision-making in e-commerce. By identifying top-selling products and understanding regional sales trends, businesses can optimize their inventory management, tailor marketing strategies, and ultimately enhance customer satisfaction and profitability.