Business Data Analytics

Office Store Case Study

By

Sreenija Dharma

Bal Thirupathi Guddati

Sandeep Chanda

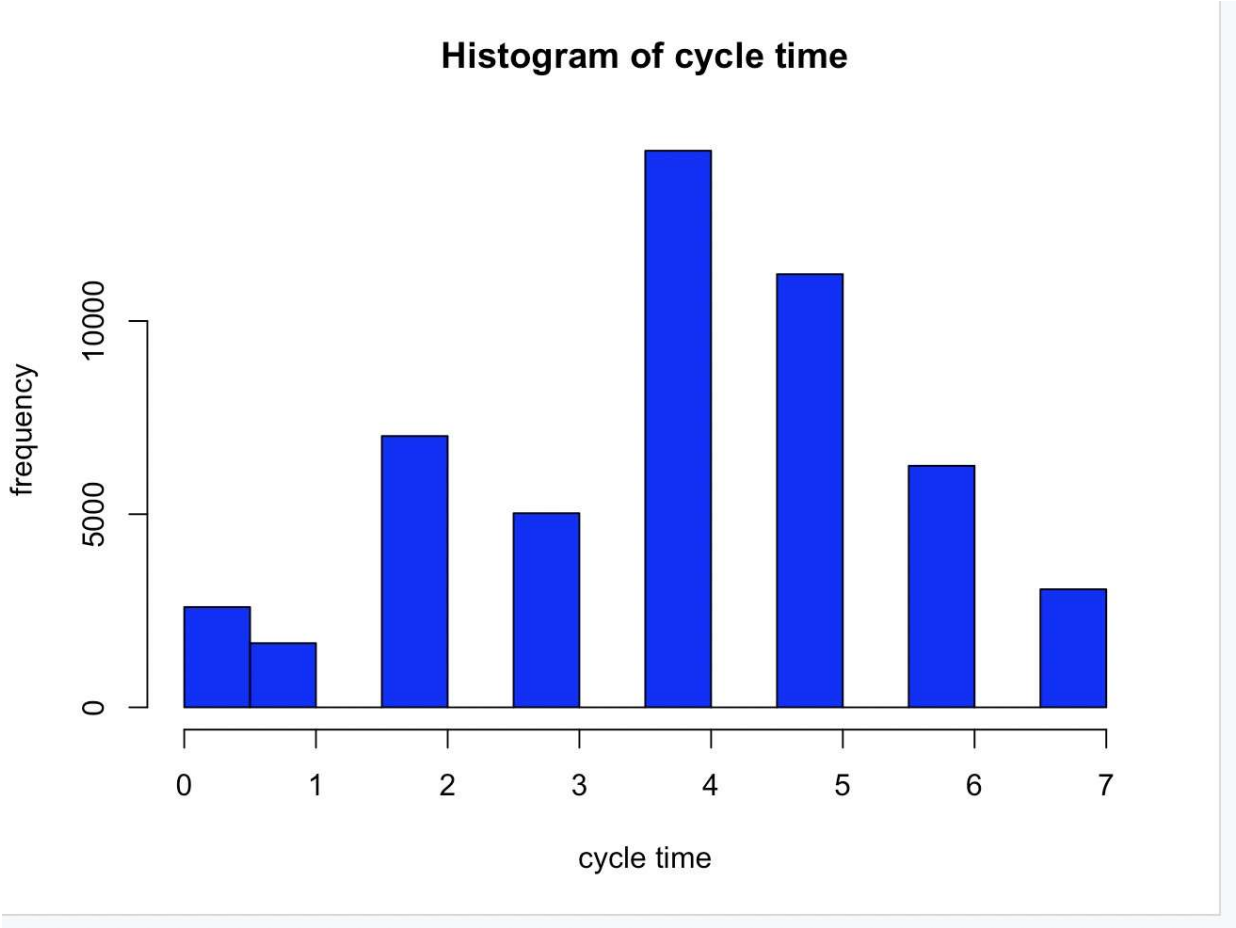
Professor: Santiago Aguirre

BIS581

Class:22450889

Table of contents

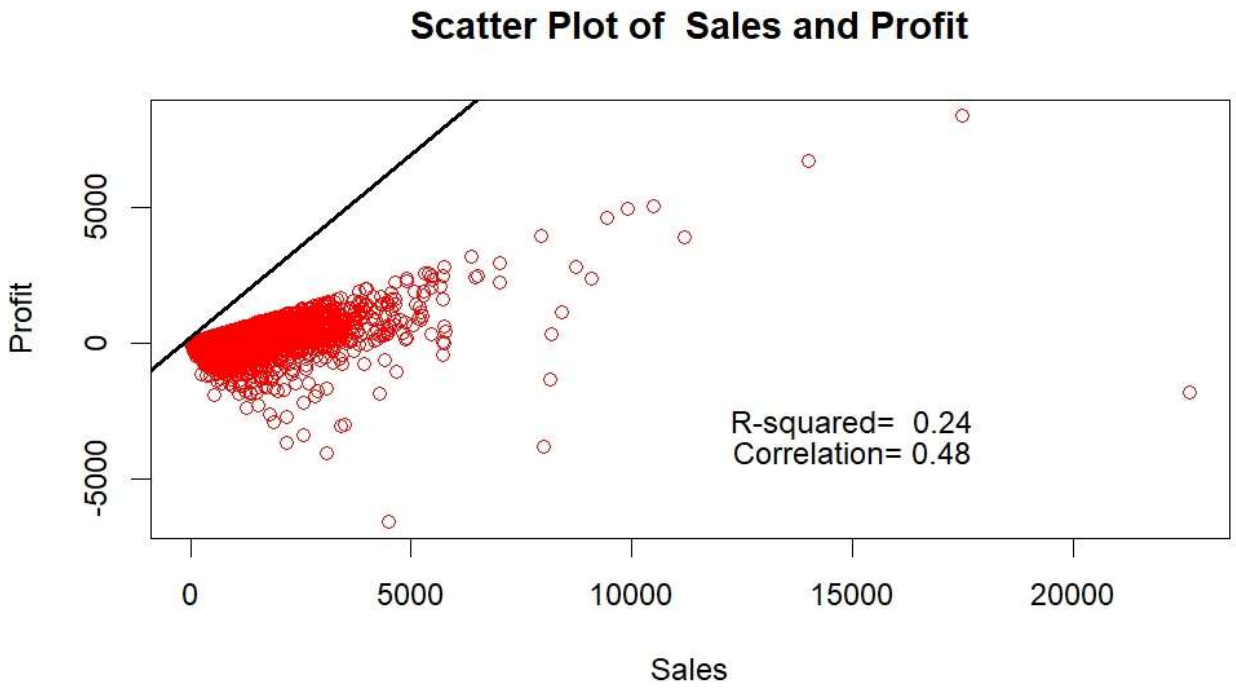
1) The Histogram of Cycle Time:



Analysis:

The histogram represents the distribution of cycle time from the Office Store dataset, with cycle time displayed on the x-axis and frequency on the y-axis. With a mean of 3.969651 and a median of 4, the graph exhibits a leftward skewness, indicating a tendency for the data to be concentrated on the higher values with a tail extending towards the lower values. The calculated skewness value of -0.431 further confirms this negative skewness. Despite showing one prominent peak, the graph also displays multiple curvier points, suggesting multidimensional properties within the dataset.

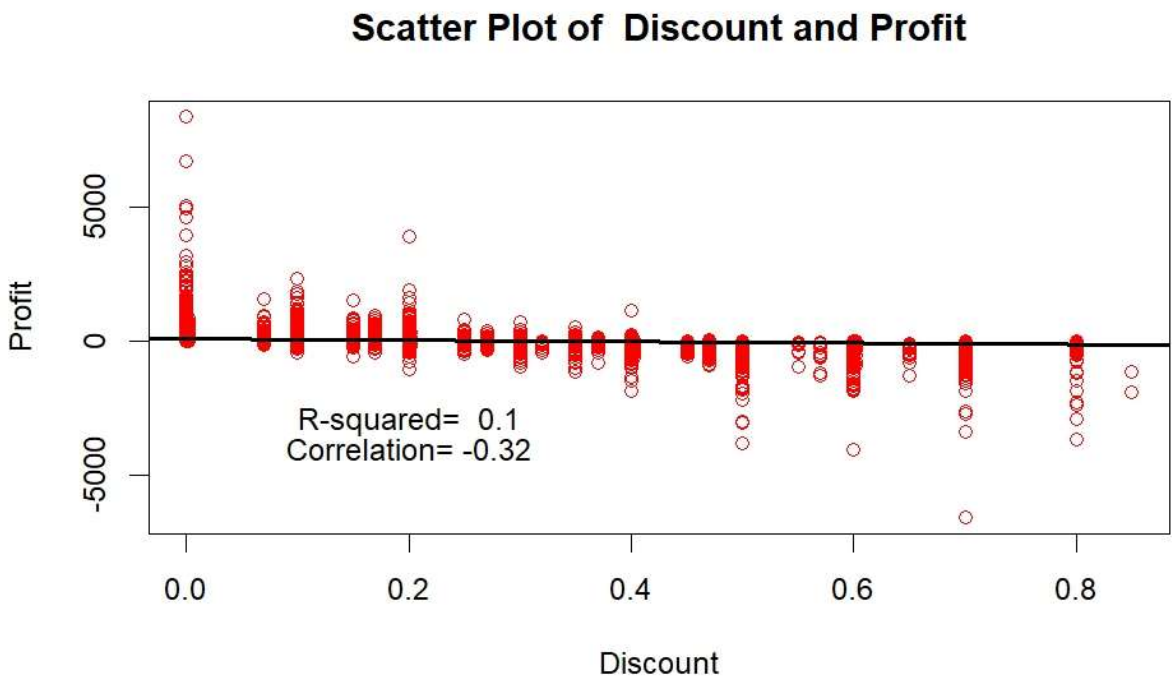
2)Scatter plot of Sales and profit:



Analysis:

In the above scatter plot, the x axis represents the sales, and the y axis represents the profit. The correlation between sales and profit is 0.48, which indicates that it has a low positive correlation. Also, in the graph, there is a correlation with the line of best fit. Here, profit is the dependent variable. The intercept that is b in y=mx+b is -14.1363. The coefficient of determination of 0.235 indicates that the dependency of profit on sales is less.

3) The Relationship between Discounts and Profit

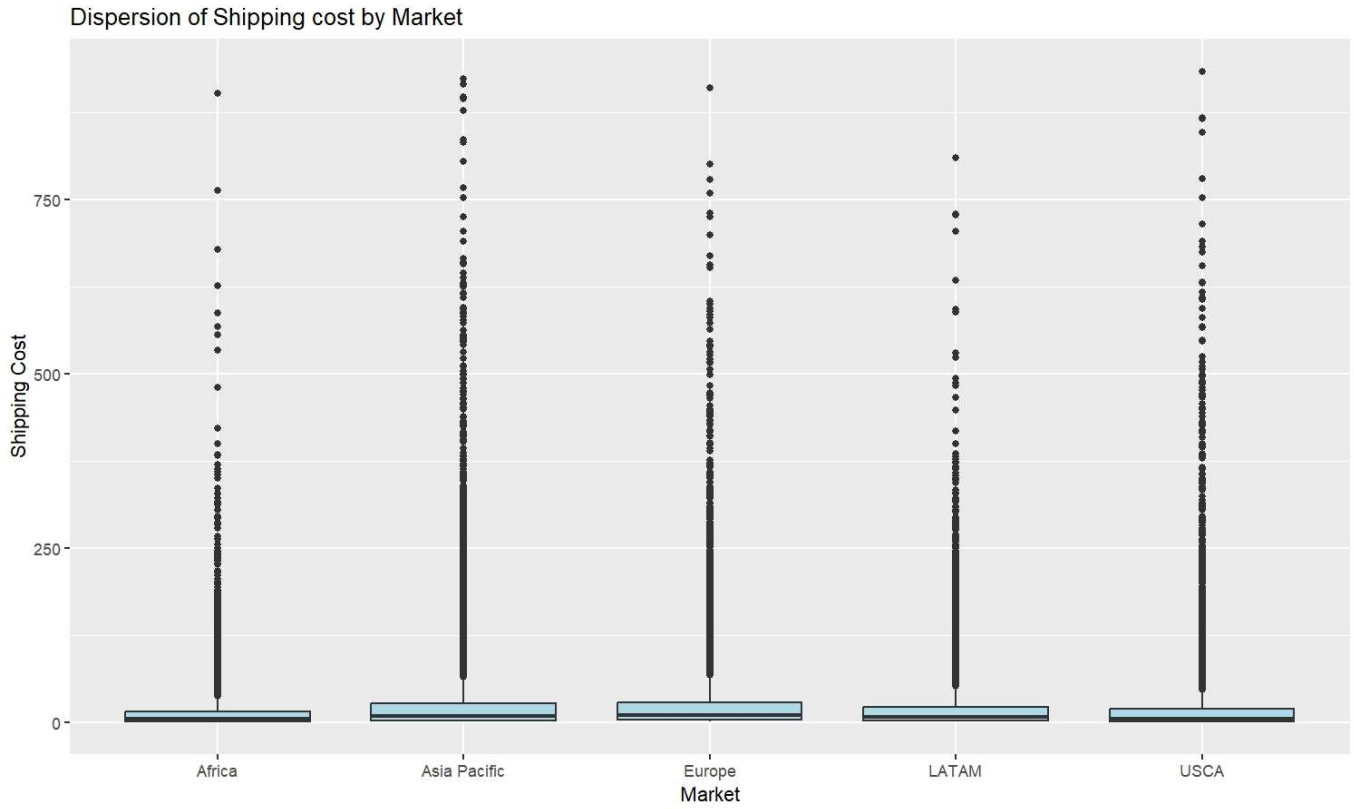


Analysis:

In the above scatter plot x axis represents the discount and y axis represents the profit. From the scatter plot we can notice that the correlation between the discount and profit is -0.32 which indicates the low negative correlation exists between discount and profit indicates that there is a slight tendency for profits to decrease as discounts increase. There are many data points that fall above the trend line, indicating that some discounts are associated with increased profits and coefficient of determination (r squared) is 0.1.

A small discount may not have a significant impact on profit, while a large discount could significantly reduce profit. Discounts on high-priced products will have a greater impact on profit than discounts on low-priced products.

4) Dispersion of Shipping Cost by Market:

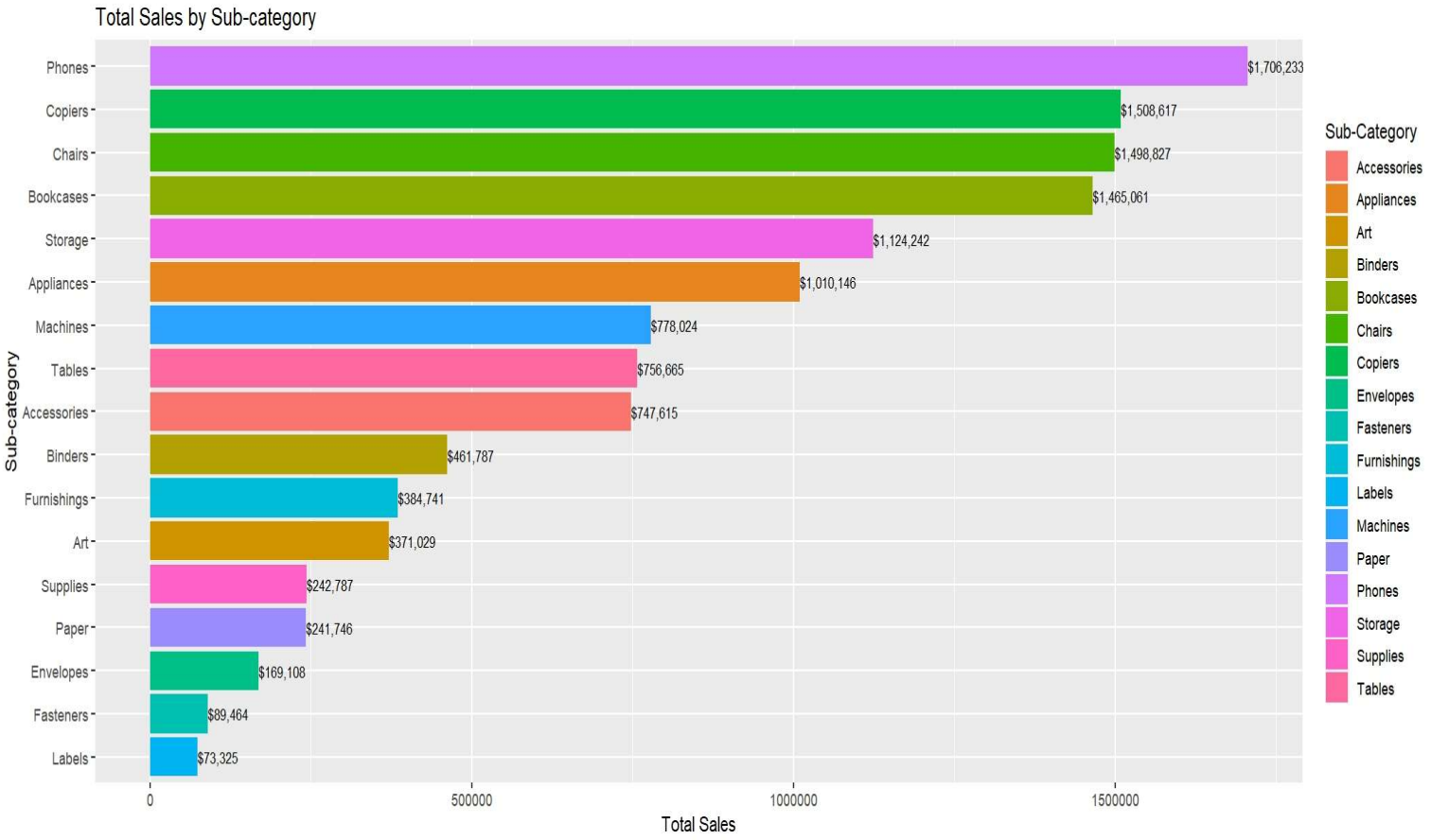


S.No	Market	Total Shipping Cost	Average Shipping Cost (Mean)	Median of Shipping Cost	Standard Deviation of Shipping Cost
1	Africa	88837.68	19.37	5.15	47.69
2	Asia Pacific	437647.99	30.62	9.14	64.77
3	Europe	348166.74	29.75	10.55	58.32
4	LATAM	235011.35	22.83	7.46	45.86
5	USCA	246991.88	23.83	5.13	58.43

Analysis:

The above box plot represents the shipping costs across different markets, it is evident that Asia Pacific stands out with the highest average shipping cost at $30.62, indicating it as the most expensive market, while Africa records the lowest average shipping cost of $19.37. Notably, all markets exhibit high standard deviations, indicating substantial variability in shipping costs within each region. The medians consistently fall below the means, highlighting the influence of high-cost outliers in the data.

5) Total Sales by Sub-Category:

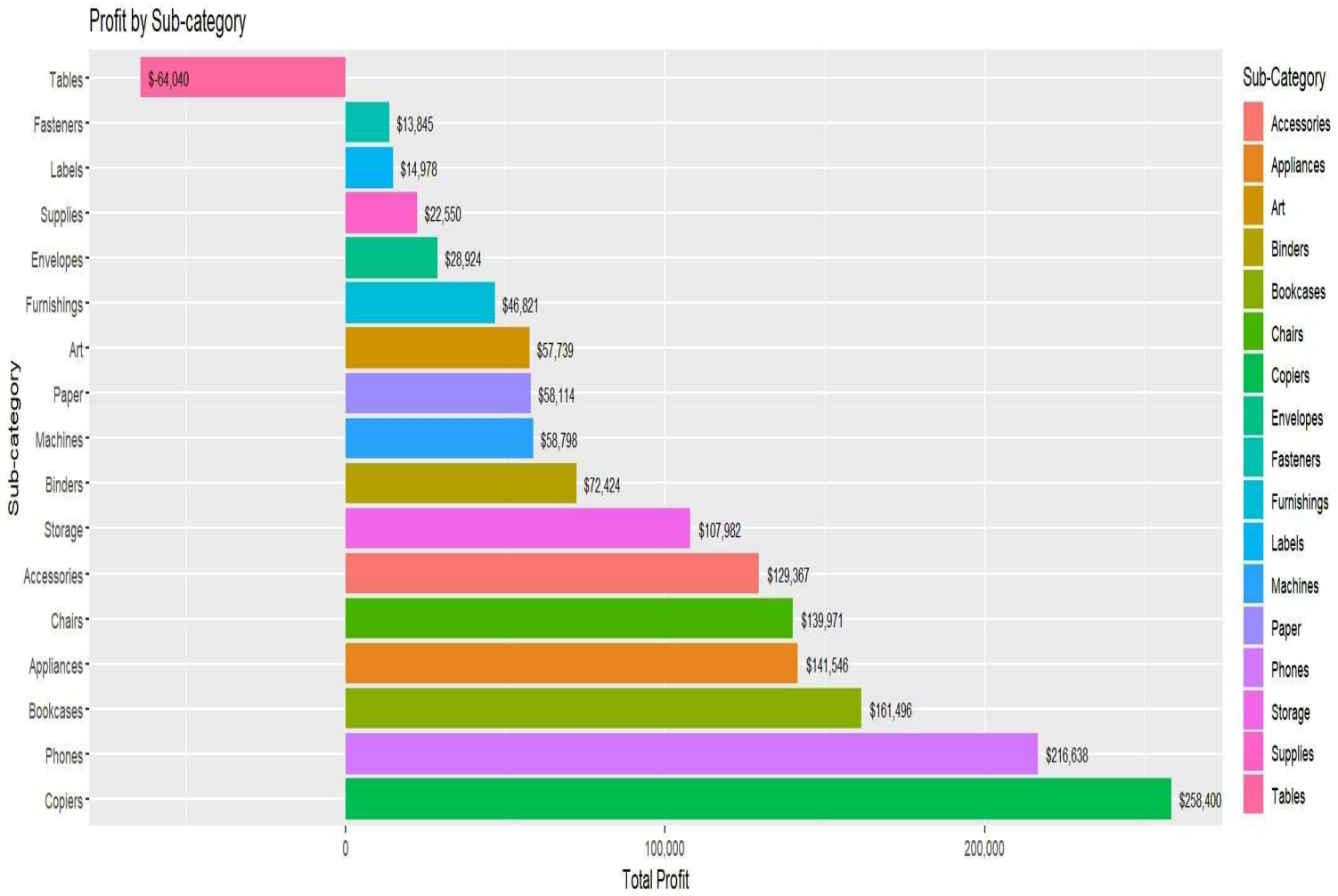


Analysis:

The above bar graph represents the total sales by sub-category x axis represents the total sales and y axis represents the sub- category. Phones are the top-selling subcategory, generating $1,706,233 in sales. while labels are the lowest-selling subcategory, generating $73325. The top 5 subcategories, including phones, copiers, chairs, bookcases, storage, contribute over $7 million in sales, constituting

approximately 57.8% of the total sales. The bottom 5 subcategories labels, fasteners, envelopes, paper and supplies collectively generate less than $816430.31 in sales.

Total Profit by Sub-Category:

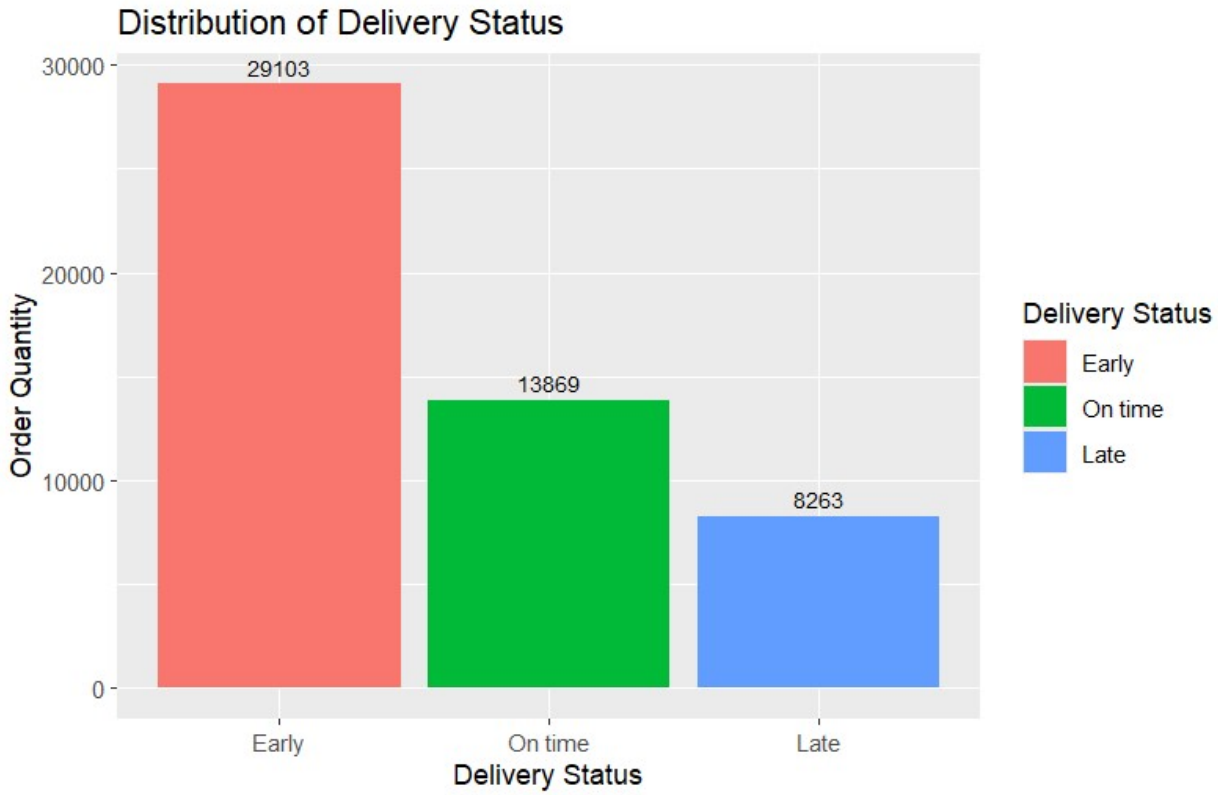


Analysis:

The above bar graph represents the total profit by sub-category. The x axis represents the total profit, and the y axis represents the sub-category. Copiers are the highest profit-gaining subcategory, generating $258400. while fasteners are the lowest-profit subcategory at $13845

The company should focus on selling more copiers because they make the most profit. They also need to figure out why tables are causing them to lose money, with a negative profit of -$64,040(loss). Finding out why this is happening will help them fix the problem, like changing prices or reducing costs.

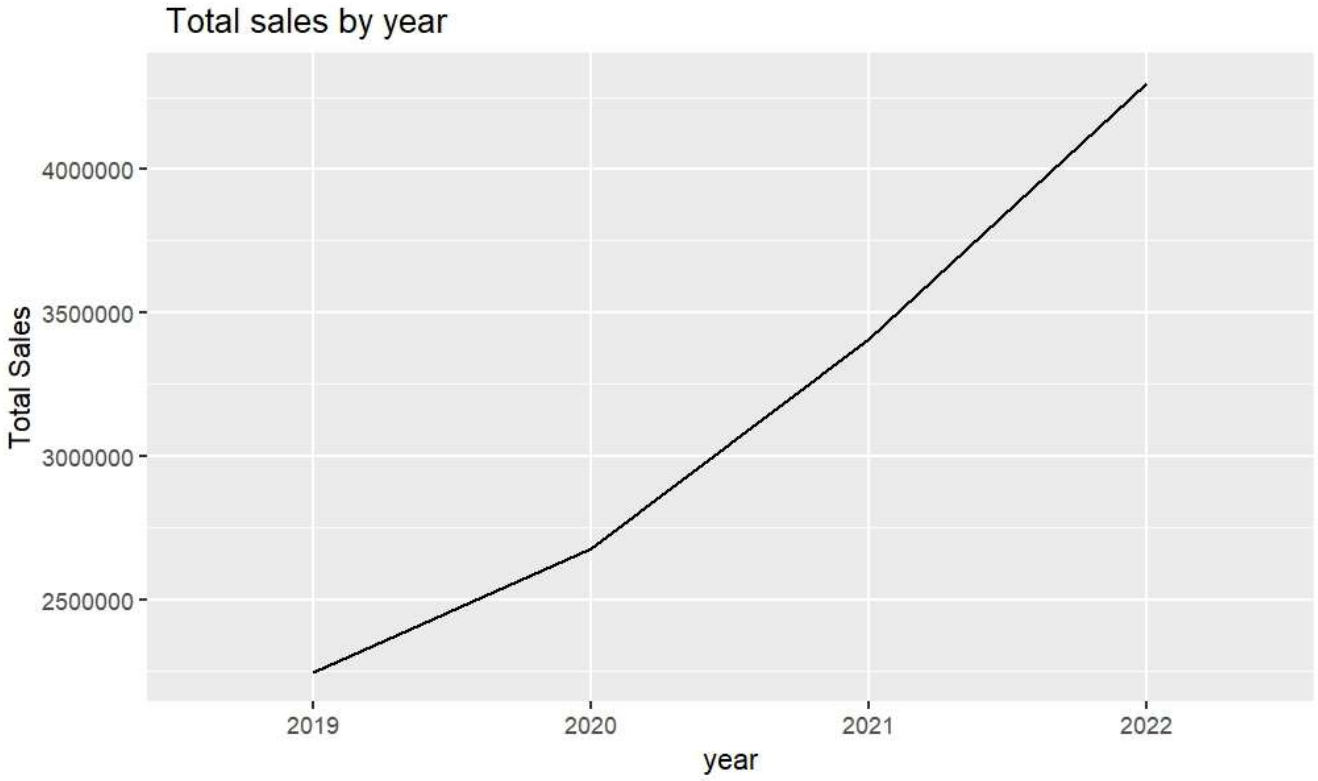
The Relationship Between Order Quantity and Late Deliveries:



Analysis:

The bar chart represents the relationship between delivery status and order Quantity. The X-axis represents the delivery status, and the Y-axis represents the order quantity. In the graphic, we can see that there are 8263 late deliveries, which comprise 16.1% of total deliveries. There are 13869 on-time deliveries, which contribute 27% of total deliveries. And finally, there are 29103 early deliveries, which contribute 56.8% of total deliveries. From the above data, it is very clear that the company needs to focus on late deliveries rather than early deliveries to improve sales and performance. There are 56% early deliveries, so there will not be any impact on early deliveries if the company keeps more concentration on late deliveries.

Sales Trend Using a Line Chart:



Analysis:

The line chart visualizes the total sales generated for each year. The x-axis indicates the years, and the y-axis indicates the total sales in dollars. The sales generated in 2019 were 17% of the total sales. There is an increase in sales of 4% in 2020, which is 21%. In 2021, the sales further increased to 25%. Notably, there was an 8% rise in sales in 2022 compared to the previous year. The line chart clearly indicates that there is a gradual increase in sales from 2019 to 2022.

Insights:

- Medians consistently fall below means, suggesting the influence of high-cost outliers in each region in box plot of Shipping Cost Dispersion by Market
- The company should focus on optimizing sales of high-profit subcategories, like copiers, and investigate the factors causing losses in subcategories like tables.
- Concentration on reducing late deliveries, which affect 16.1% of orders, could impact sales and overall performance.

Conclusion:

- Attention to late deliveries is crucial for enhancing customer satisfaction and overall performance.
- Optimization of discount strategies is needed, considering the low negative correlation between discounts and profit.
- Emphasis on high-profit subcategories, such as copiers, can contribute to overall profitability. Further investigation into the reasons behind losses in specific subcategories, like tables, is essential for corrective actions.
- The company should Investigate and identify the specific factors contributing to the high-cost outliers in each market. Addressing these outliers can bring down the overall average shipping cost across all the markets.