

Final Project

Data Visualization for Business Intelligence

<Superstore sales information >

<Specific data source>: <https://www.kaggle.com/datasets/vivek468/superstore-dataset-final>

Xinxue, Guo
xg2407
xg2407@nyu.edu

Mason, Liu
el4165
el4165@nyu.edu

Xinjing, Guo
xg2404
xg2404@nyu.edu

Department of Technology Management and Innovation
New York University, Tandon
Summer 2024

Overview

- What is the project about (in ~25 words)?
 - The project analyzes sales, profit, and customer data from a retail store to uncover business insights, trends, and patterns to support data-driven decision-making.
- Who is the audience and/or users?
 - The audience includes retail business managers, analysts, and decision-makers seeking to optimize sales strategies, improve customer relationships, and enhance profitability.
- How can they use the info/insights provided? (data exploration, business insights, decision support, fun,...)
 - They can use the insights for data exploration, business insights, and decision support, enabling better inventory management, targeted marketing, and improved overall business performance.

Group Summary

Group Member Name	Screen Slide #	Slide Titles	Idiom (Chart type)	# of Attributes (must be >2)	Dashboard? (Y/N) (2 or more relevant charts in single screen)
<i>Guo, Xinxue</i>	<slide 4,5>	Yearly Sales and Profit Trends by Product Category	Combination Chart (Line for Sales, Bar for Profit)	4 (Year, Product Category, Sales, Profit)	Yes
<i>Guo,Xinxue</i>	<slide 6,7>	Geographical Distribution of Profits Across the United States in 2017	Symbol Map with circle sizes representing profit	3 (Region, Profit, Year)	No
<i>Guo, Xinjing</i>	<slide 9,10>	Quarterly Sales Trend Over Time	Line Chart	3 (Year-Quarter, Sales, SUM (quarter order data))	No
<i>Guo, Xinjing</i>	<slide 11,12>	Top 10 Customers by Sales	Vertical Bar Chart	3 (Customer Name, Sales, Ranking)	No
<i>Mason Liu</i>	<slide 14,15>	Sales by Ship Mode	Scatter plot	3 (Profit, Discount, Category)	Yes
<i>Mason Liu</i>	<slide 16,17>	Sales by Ship Mode	Pie Chart	3 (Sales, Ship Mode,City)	No

Chart 1 Analysis - Yearly Sales and Profit Trends by Product Category

Role of User				Question Addressed	
Retail Business Manager				How have yearly sales and profits changed across different product categories?	
User Task <action, target>		Idiom Used		Reasoning for selection	
Compare yearly sales and profits by product category		Combination Chart (Line for Sales, Bar for Profit)		Combination charts effectively show trends and comparisons between different metrics over time.	
Datatype		Encoded item		Reasoning for selection	
Sales and profit data		Product categories		These metrics are essential for evaluating business performance across different product categories.	
Mark	Encoded item		Number of Items	Reasoning for selection	
Line for sales Bar for Profit	Yearly sales and profits		3 categories, 4 years	Combination charts clearly differentiate and compare the trends of sales and profit over time.	
Channel	Channel Type	Encoded attribute	Number of Values	Reasoning for selection	
Color	Categorical	Product categories	3	Colors help distinguish between different categories.	
Position	Quantitative	Sales, Profit	Continuous values	Positions of lines and bars represent the magnitude of sales and profits.	

Chart 1 - Yearly Sales and Profit Trends by Product Category

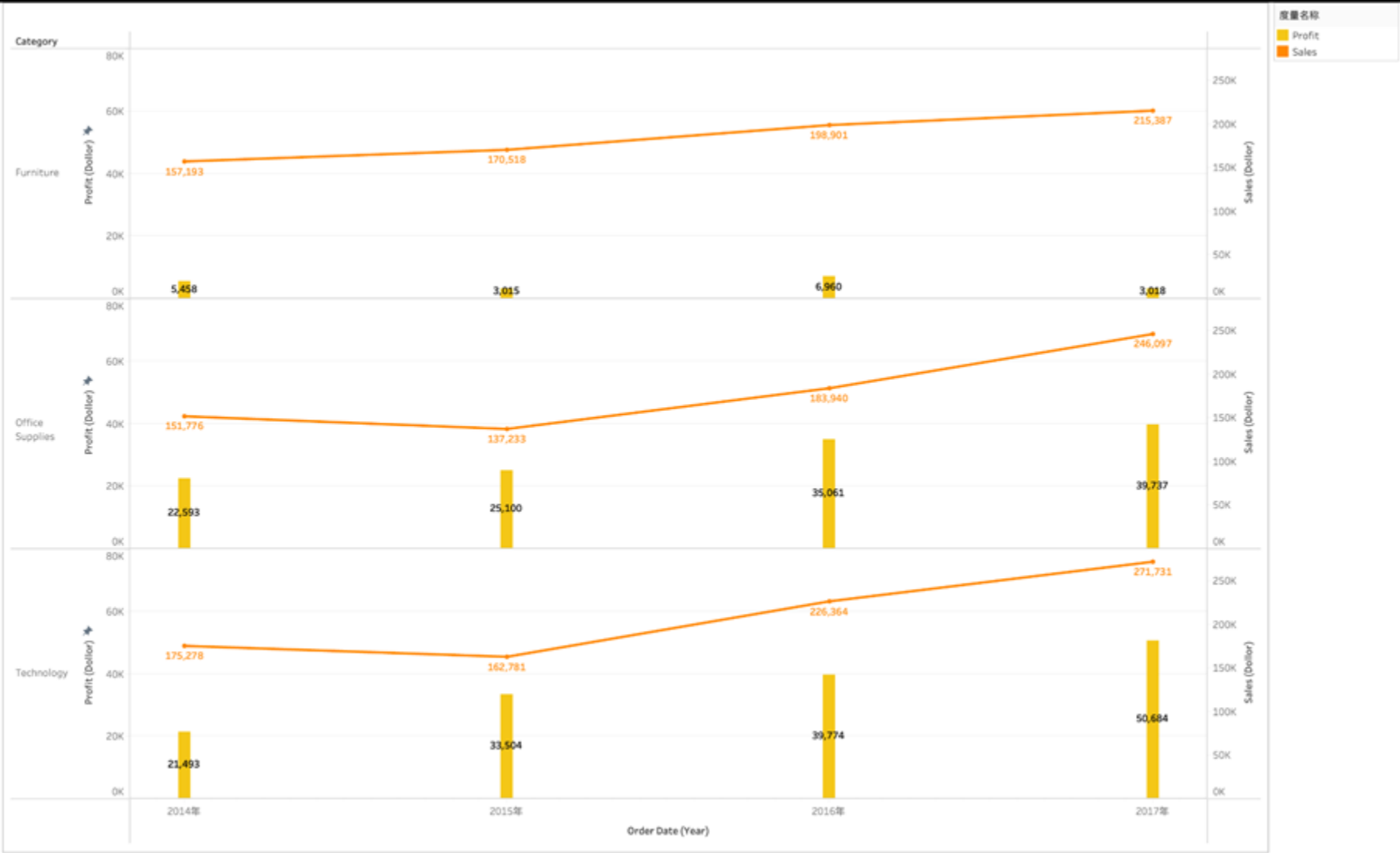
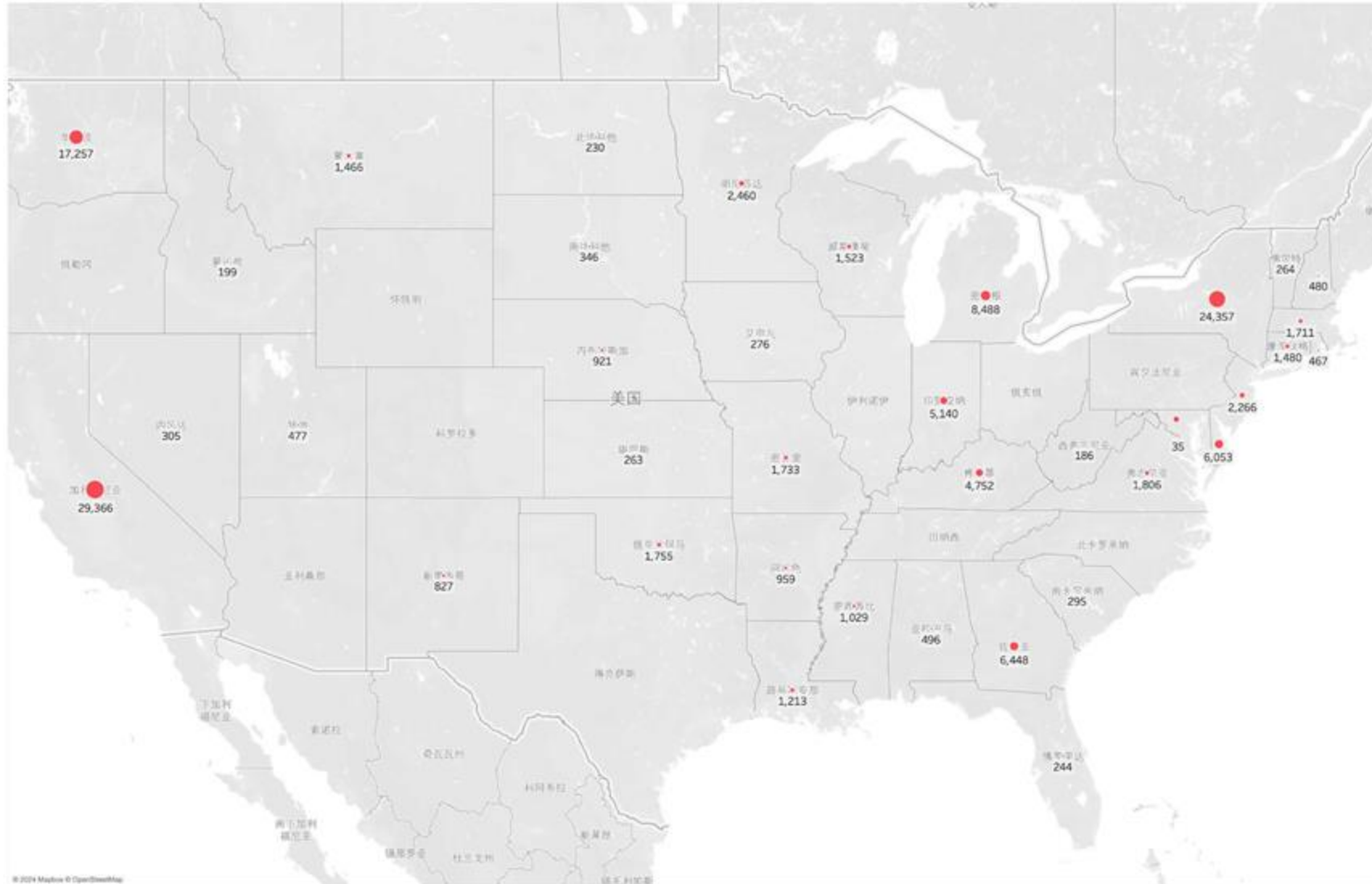


Chart 2 Analysis - Geographical Distribution of Profits Across the United States in 2017

Role of User		Question Addressed		
Retail Business Manager		How is the profit distributed geographically across the United States in 2017?		
User Task <action, target>		Idiom Used	Reasoning for selection	
Analyze geographical profit distribution in 2017		Symbol Map with circle sizes representing profit	Symbol maps are effective for visualizing geographical data, and circle sizes provide a clear indication of profit magnitude	
Datatype		Encoded item	Reasoning for selection	
Geographical data and profit data		Region and Profit	These metrics are essential for evaluating geographical performance and identifying high-profit areas.	
Mark		Encoded item	Number of Items	Reasoning for selection
Circle (Size of circle represents profit)		Regions, Profits	Multiple regions	Circle size effectively represents the magnitude of profit, making it easy to compare different
Channel	Channel Type	Encoded attribute	Number of Values	Reasoning for selection
Size	Quantitative	Profit	Continuous	Size variations clearly indicate profit differences across regions.
Position	Spatial	Region	Multiple regions	Positioning on the map shows the geographical location of each region.

Chart 2 - Geographical Distribution of Profits Across the United States in 2017



□ Chart 3 Analysis - Quarterly Sales Trend Over Time

Role of User		Question Addressed		
Sales Manager		What are the sales trends over time?		
User Task <action, target>		Idiom Used	Reasoning for selection	
Analyze sales trends over time		Line Chart	Line charts are effective for showing trends over continuous periods.	
Datatype		Encoded item	Reasoning for selection	
Time series data		Sales	Time series data needs a format that shows change over time.	
Mark	Encoded item	Number of Items	Reasoning for selection	
Line	Sales	Quarterly	Lines are suitable for representing time trends.	
Channel	Channel Type	Encoded attribute	Number of Values	Reasoning for selection
horizontal Position (X-axis)	Quantitative	Time (Quarter)	16 (4 quarters x 4 years)	Shows the progression of time clearly
vertical Position (Y-axis)	Quantitative	Sales	Continuous	Clearly shows the fluctuation of sales

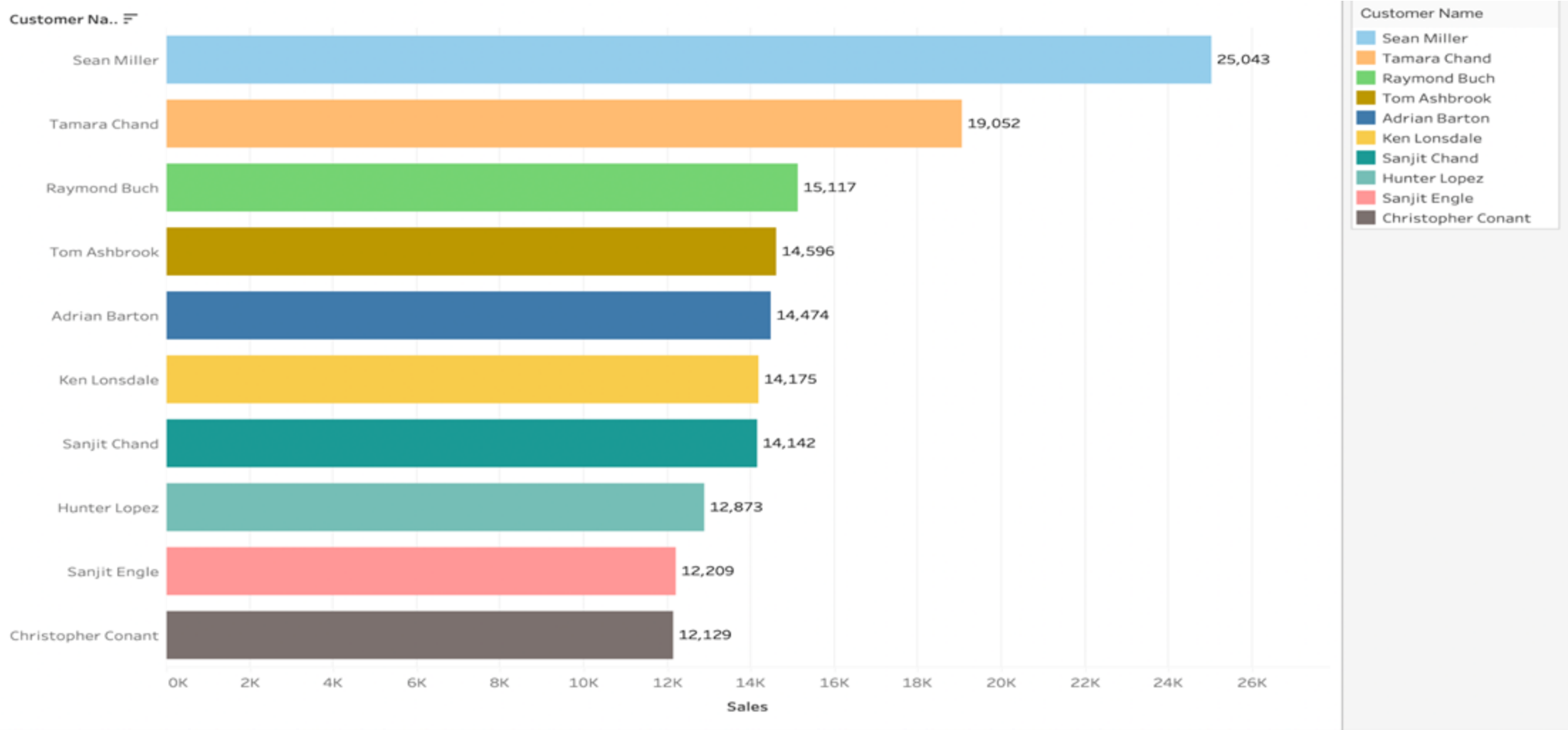
Chart 3 - Quarterly Sales Trend Over Time



Chart 4 Analysis - Top 10 Customers by Sales

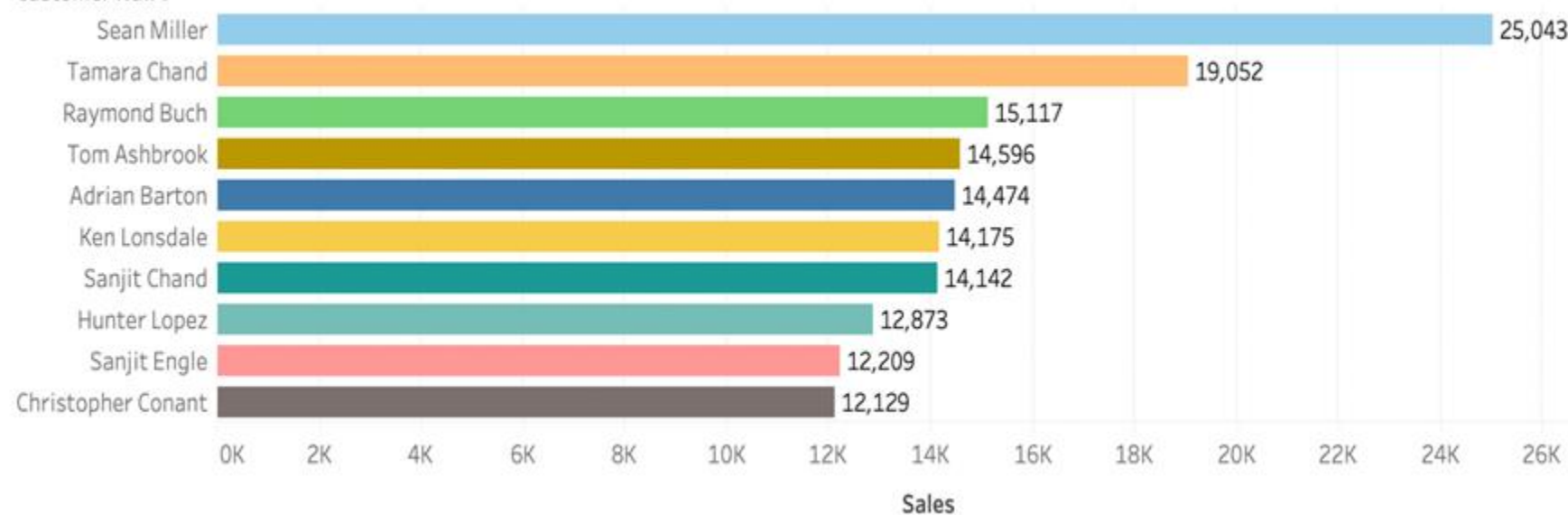
Role of User			Question Addressed	
sales manager			Who are the top 10 customers by total sales?	
User Task <action, target>		Idiom Used	Reasoning for selection	
Identify top customers by sales		Vertical Bar Chart	Bar charts effectively show comparative data among categories.	
Datatype		Encoded item	Reasoning for selection	
Categorical Data		Customer Name	Identifying specific customers by sales amount.	
Mark	Encoded item	Number of Items	Reasoning for selection	
Bar	Sales	10	Bars are suitable for comparing discrete categories.	
Channel	Channel Type	Encoded attribute	Number of Values	Reasoning for selection
Color hue	Categorical	Customer Name	10	Different colors distinguish each customer.
horizontal Position (X-axis)	Quantitative	Sales	Continuous	Clearly shows the amount of sales for each customer.
vertical Position (Y-axis)	Categorical	Customer Name	10	Identifies top customers by name.

Chart 4 - Top 10 Customers by Sales





Customer Name



Customer Name

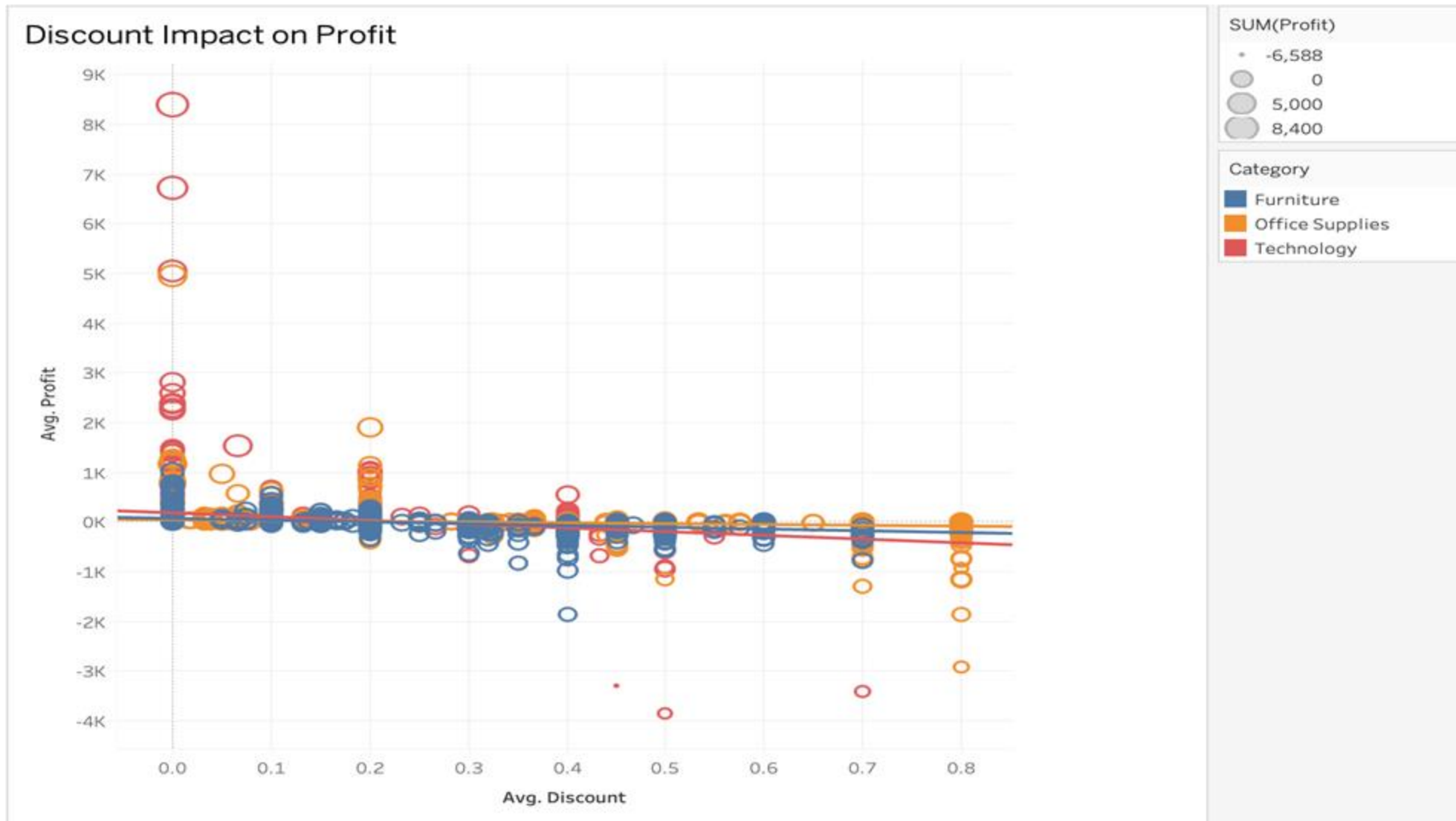


Chart 5 Analysis -Discount Impact on Profit

Student Name
Mason Liu

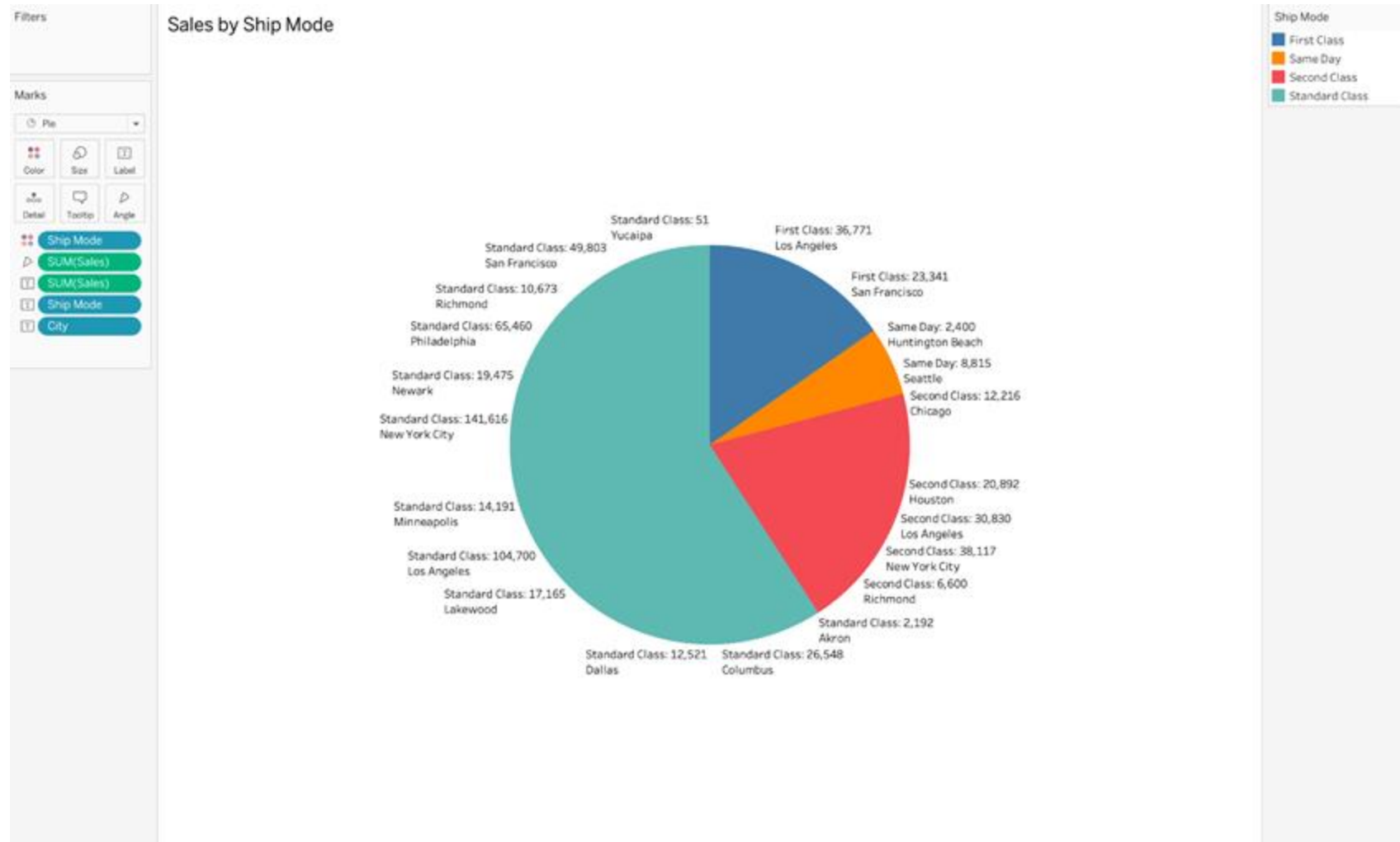
Role of User			Question Addressed	
Retail business managers and analysts			How does the discount level affect the profit margins on sales?	
User Task <action, target>		Idiom Used	Reasoning for selection	
Analyze the relationship between discounts given and the profit earned from sales.		Scatter plot	A scatter plot effectively illustrates the relationship between two quantitative variables, making it ideal for observing trends or patterns in discounts versus profits.	
Datatype	Encoded item		Reasoning for selection	
Sales data	Discount and profit values		The discount rate directly impacts profitability, and plotting these values will reveal whether higher discounts correlate with lower profits.	
Mark	Encoded item		Number of Items	Reasoning for selection
Point	Each point represents one order		Each point on the scatter plot	Good for represent individual data entries in a scatter plot, allowing each order's discount and profit to be visualized distinctly.
Channel	Channel Type	Encoded attribute	Number of Values	Reasoning for selection
Position	Quantitative	Discount rate on x-axis and profit on the y-axis	Continuous	The position channel effectively encodes quantitative changes, ideal for showing how changes in the discount might affect profit levels
Color hue	Categorical	Product Category	3 (Furniture, Office Supplies, Technology)	Coloring points by product category can identify if certain categories are more sensitive to discount changes in terms of profitability.

Chart 5 - Discount Impact on Profit

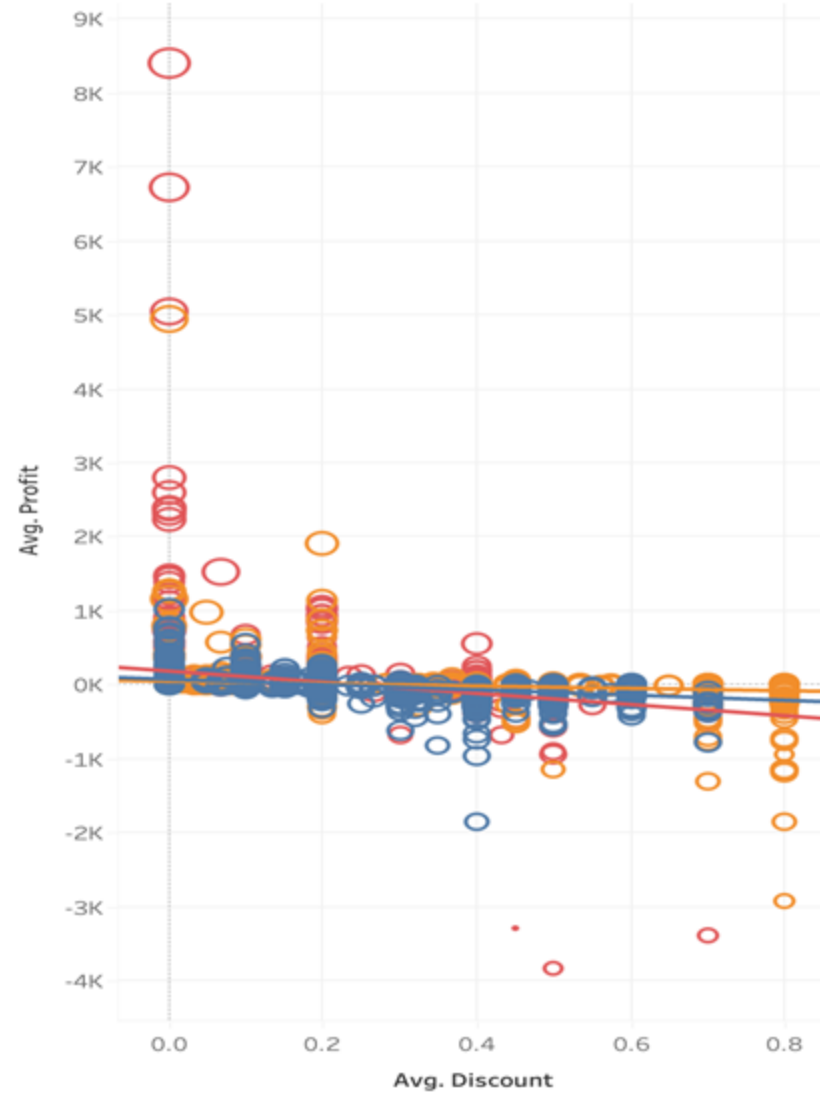


Role of User		Question Addressed		
Retail business managers and decision-makers		What is the proportion of sales volume contributed by each shipping mode?		
User Task <action, target>		Idiom Used	Reasoning for selection	
Compare sales distribution across different shipping modes		Pie Chart	It display the proportionate contributions of categories within a whole, making it easy to see which shipping modes dominate sales.	
Datatype	Encoded item		Reasoning for selection	
Sales data	Sales amount by shipping mode		Showing sales by shipping mode in a pie chart lets users quickly grasp the relative significance of each mode to overall sales	
Mark	Encoded item		Number of Items	Reasoning for selection
Sector (slice of the pie)	represents the total sales for a particular shipping mode		One slice per shipping mode	Slices in a pie chart directly represent parts of a whole, suitable for comparison of categories like shipping modes
Channel	Channel Type	Encoded attribute	Number of Values	Reasoning for selection
Angle and Area	Quantitative	Sales amount	Proportional to the sales volume	The size of each slice visually encodes the proportion of total sales, offering an intuitive grasp of each category's contribution
Color hue	Categorical	Ship Mode	4 (Standard Class, Second Class, First Class, Same Day)	Using different colors for each ship mode clearly differentiates the slices, allowing for immediate visual segmentation of the data.

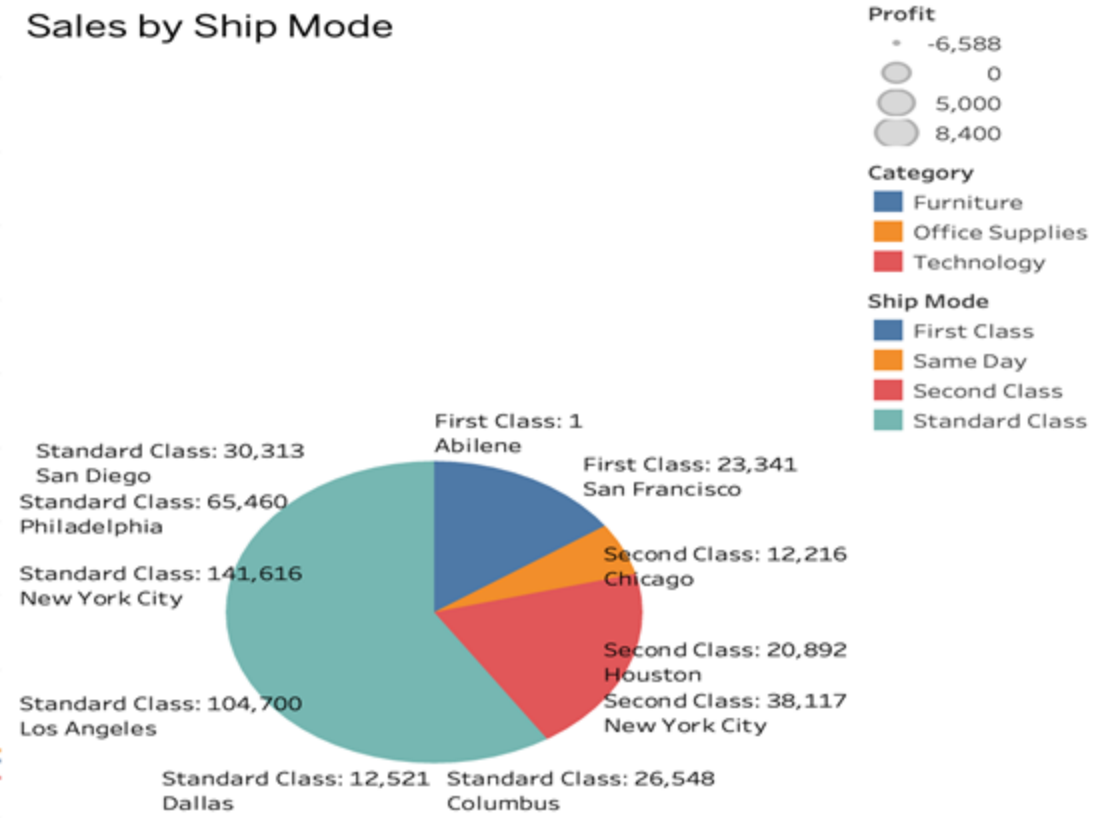
Chart 6 - Sales by Ship Mode



Discount Impact on Profit



Sales by Ship Mode



Summary/Conclusions

- Slide 5: The demand for technology, office supplies and furniture is increasing, but only the profits of office supplies and furniture are growing.
- Slide 7: The most profitable regions are concentrated in the western and eastern United States, especially New York and California.
- Slide 10: Major peaks occur in 2014 Q4, 2015 Q3, 2016 Q4, and the highest in 2017 Q4. Despite the variability, there is an overall upward trend in sales over the period.
- Slide 12: Among top 10 customers based on their total sales, Sean Miller leads with 25,043 in sales, followed by Tamara Chand with 19,052. The other top customers range from 15,117 to 12,129 in sales.
- Slide 15: Overall discounts have a minimal impact on profit, but high discounts in specific transactions can lead to significant profit changes.
- Slide 17: Standard shipping has the largest share, followed by second class, while first class and same-day shipping have smaller proportions, indicating customers' preference for more economical shipping option.