TASK 3
Exploratory Data Analysis - Retail

By Ketki Kale

Intern in Data Science and Business Analytics at

The Sparks Foundation

June2021

Contents

- Objectives
- Business Problems
- Data Summary
- Insights and Demonstration
- Conclusions

Objectives

- Super Store Data Descriptive Analysis
- As a business manager, try to find out the weak areas where you can work to make more profit.
- What all business problems can be derived by exploring the data?

Business Problems

- What are the different profitable and loss making categories?
- Which region shows more sales, profit? Is profit same across all regions?
- Which are the top 5 profitable and non-profitable states?
- Is there any relationship between profit & discount offered across categories?

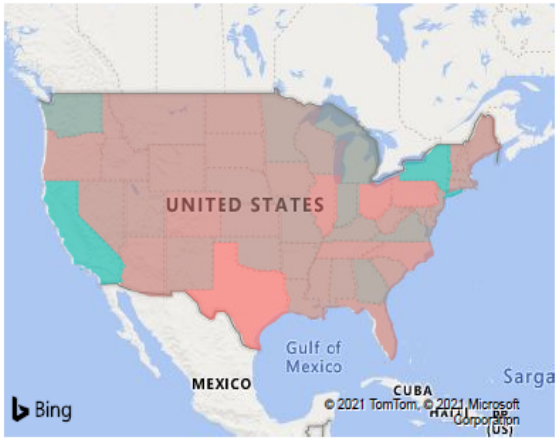
Data Summary

- **Is Data Clean? Yes :-data.isnull().sum()**
- Ship Mode 0
- Segment 0
- Country 0
- City 0
- State 0
- Postal Code 0
- Region 0
- Category 0
- Sub-Category 0
- Sales 0
- Quantity 0
- Discount 0
- Profit 0

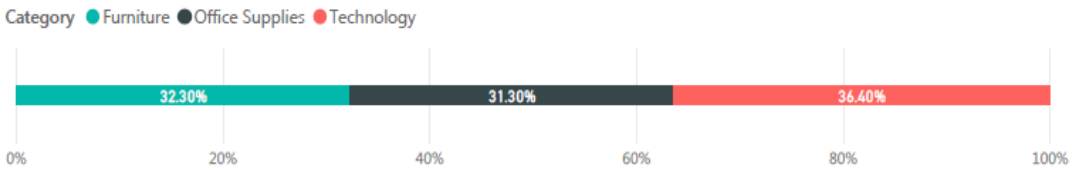
- **Data Information:- data.info()**
- RangeIndex: 9994 entries, 0 to 9993
- Data columns (total 13 columns):
- # Column Non-Null Count Dtype
- --- ------ -------------- -----
- 0 Ship Mode 9994 non-null object
- 1 Segment 9994 non-null object
- 2 Country 9994 non-null object
- 3 City 9994 non-null object
- 4 State 9994 non-null object
- 5 Postal Code 9994 non-null int64
- 6 Region 9994 non-null object
- 7 Category 9994 non-null object
- 8 Sub-Category 9994 non-null object
- 9 Sales 9994 non-null float64
- 10 Quantity 9994 non-null int64
- 11 Discount 9994 non-null float64
- 12 Profit 9994 non-null float64

Category Wise Analysis

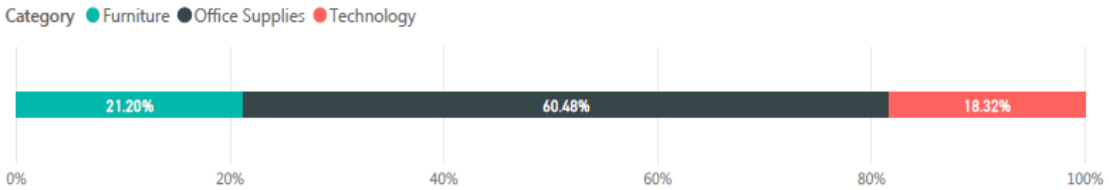
Profit by State



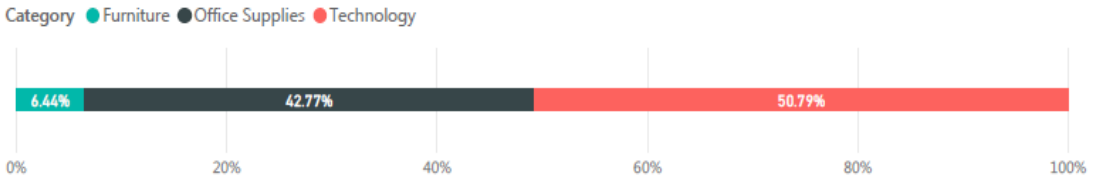
Sales by Category



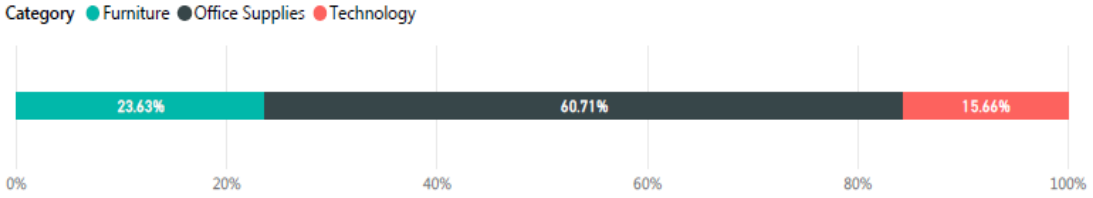
Quantity by Category



Profit by Category



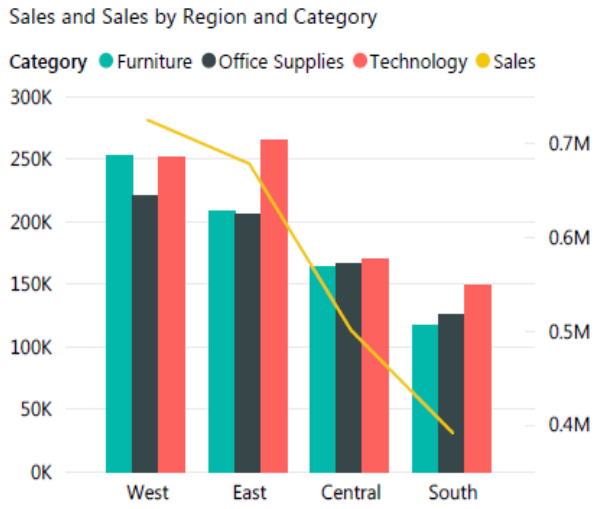
Discount by Category

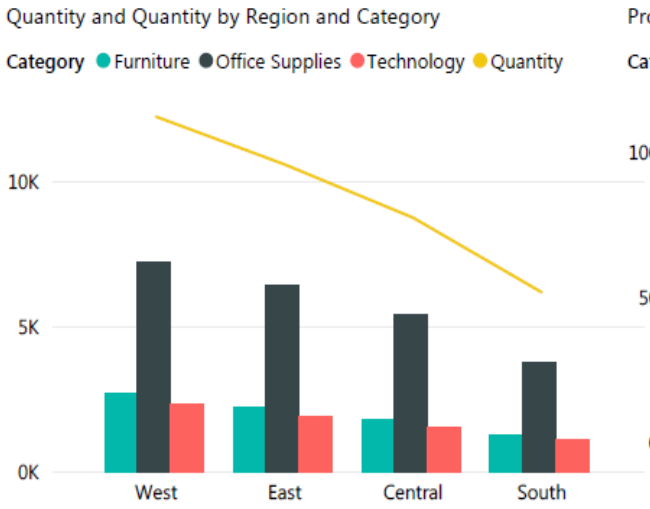


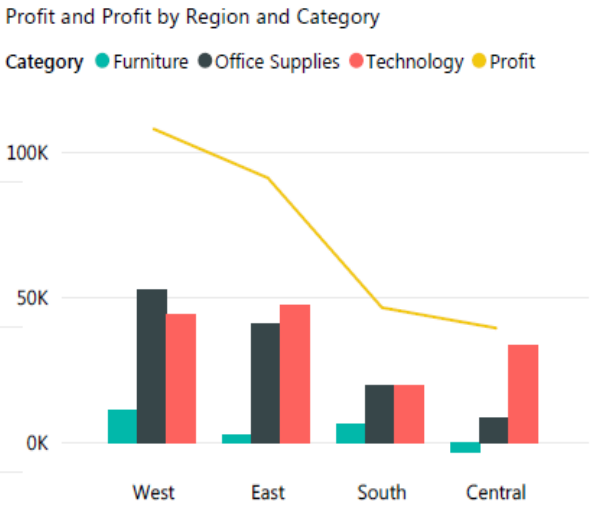
Insights:

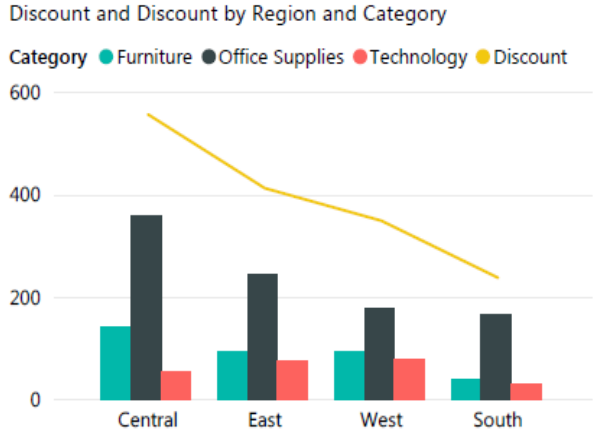
1. Statewise profit of US country is shown where green color-profit states and towards red-less profit states
2. There is slight difference in categorywise sales with highest in technology category.
3. Quantitiywise sales of category is highest in office supplies by 60 % but profit of technology is highest by 51% even though its quantities sold are less.
4. As Discount given is least on technology category, more profit is made in technology-51 %
5.As discount on office supplies is more,more quantites in that category is sold but due to discount, profit made is only 42% lesser than

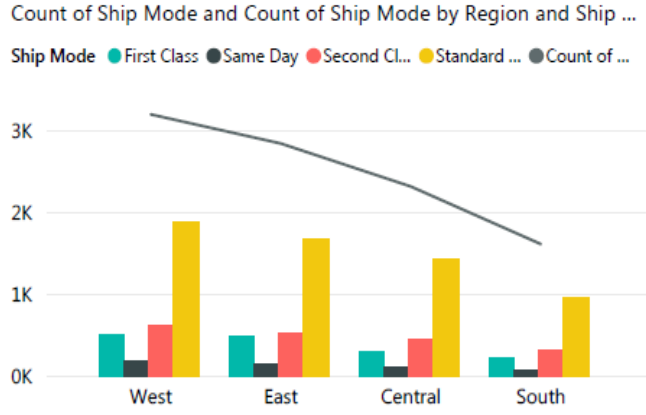
Region Wise Analysis









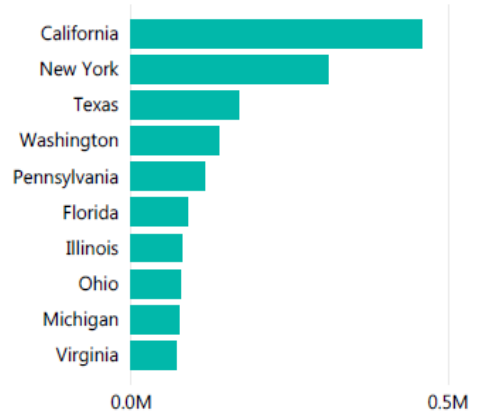


Insights:

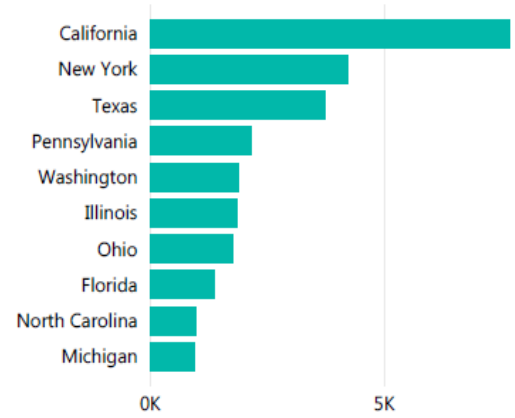
1. Top 3 graphs shows by all sales, quantity and profit wise ,west region is at top.
2. Though discounted items are more in central region, its quantity and sales wise count is lower as there may be less demand of those products
3. As sales and quantity are more in west region, count of ship mode is highest in west region.
4. Highest used ship mode is standard class in all regions.

States wise Analysis

Sales of Top 10 States



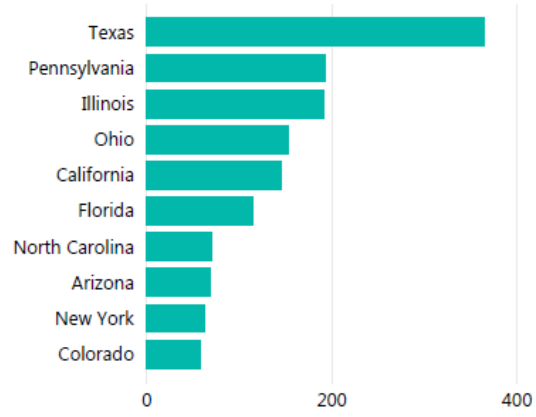
Quantity of Top 10 States



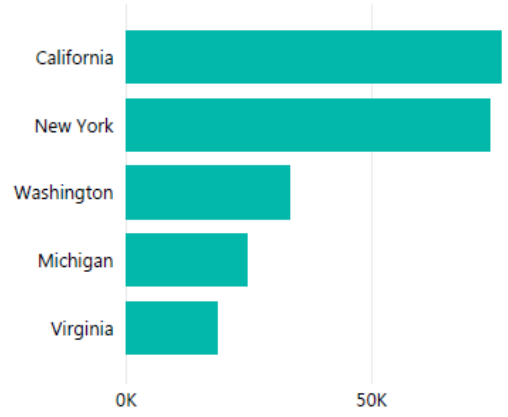
Insights:

1. Profit is highest in Califonia as well as it is also highest state quantity and sales wise which is in west region making highest profit.
2. Inspite of being 3rd in sales and quantity, Texas is top in loss making states .
3. As Texas is top in discount giving states, quantity and sales are more but is not profit making state.

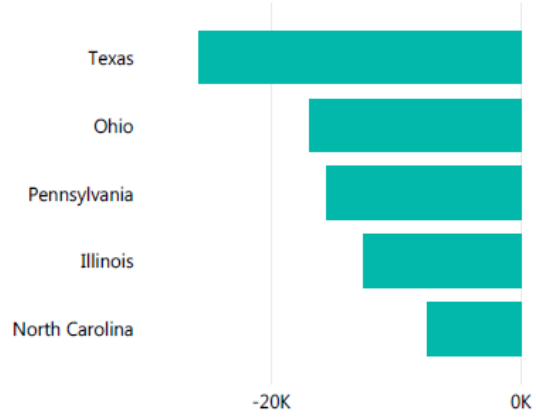
Top 10 Discount giving States



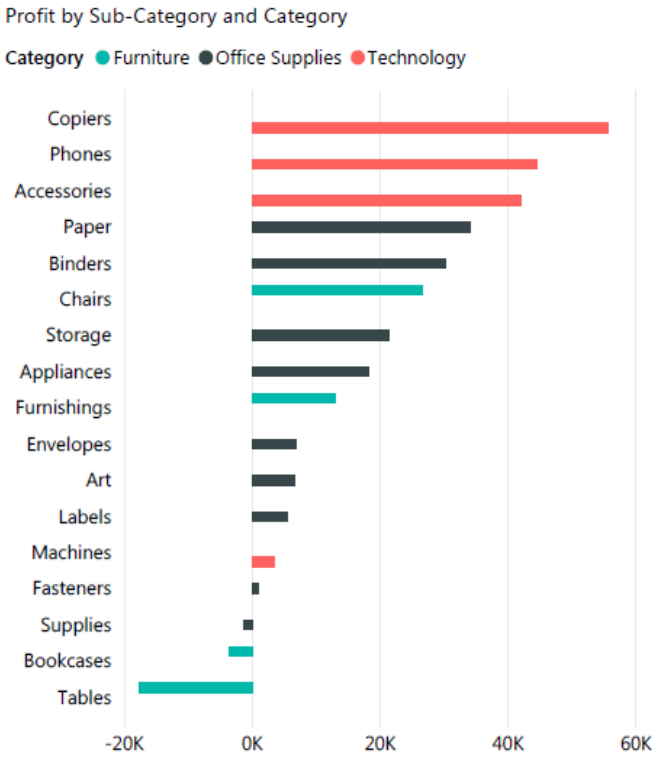
Top 5 Profit making States

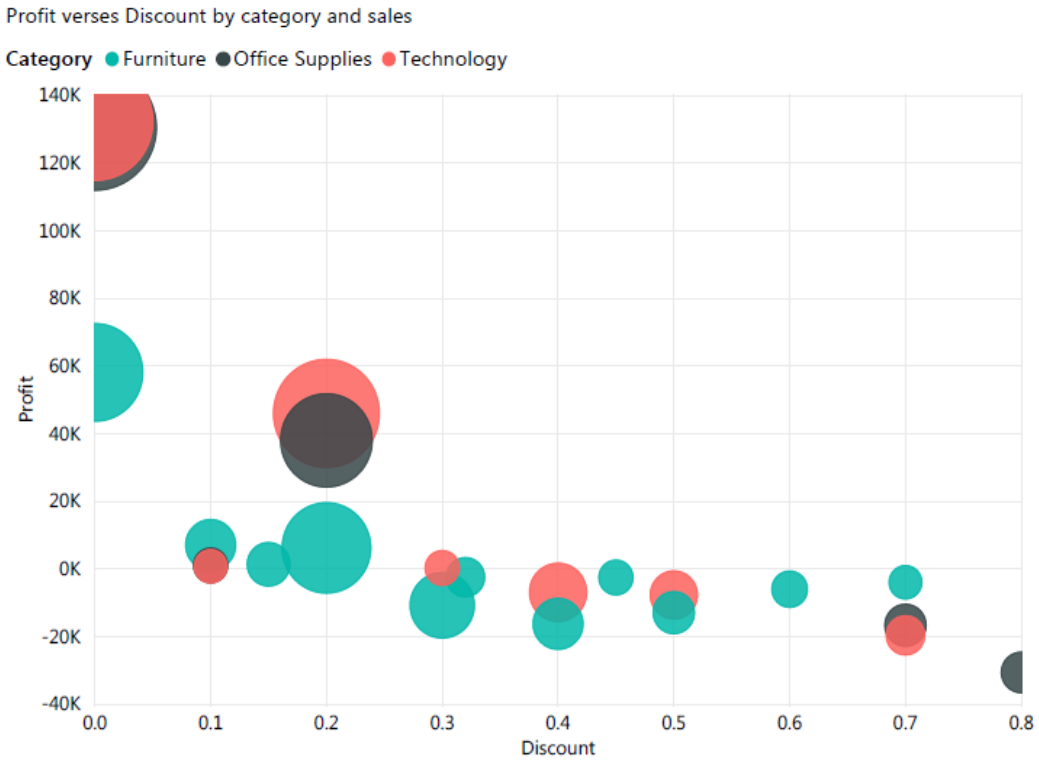


Top 5 Loss Making States



Profit Analysis





Insights:

1.Top 3 profit making subcategory are copiers, phones and accessories which are in technology category which is highest in profit making

2.Top 2 loss making subcategory are tables and bookcases which are in furniture category which is least making profit.

Conclusions

- Technology Category is the highest profit making category
- In US country, West region is the highest profit making region and California state has highest profit in West region.
- Standard class ship mode has highest sales, quantity and profit among all ship modes
- Several category products have less/no discounts. Still they are profit making due to good demand from customers
- Technology category has least discounted products, but it is still profitable

APPENDIX

Tools used

- Pycharm
- Power BI

Basic Profit Formulas

- Profit = Revenue – Costs
- Revenue = Quantity x Price
- Costs = Total Variable Costs + Total Fixed Costs
- Total Variable Costs = Quantity x Variable Costs

Profit = [(Price – Variable Costs) x Quantity] – Total Fixed Cost

Business problems derived using data

- Why there is less profit seen even though quantities and discounts are higher?
- How is the profit distributed in a particular product line, geography, or customer segment?
- Are we selling more low-prices products and fewer high-priced products?