US Superstore

Plan

Data Cleaning

Exploratory Data Analysis and Pandas Profile Report

Dashboard

Presentation

Tools

Data Analysis: Google Colab (Python: pandas, matplotlib)
Visualization: Power BI

Dataset link: https://www.kaggle.com/datasets/juhi1994/superstore

Colab Link:
https://colab.research.google.com/drive/1teTWNCb1qhndquovjRc3JLiWLgl3ZqQO?usp=sharing

Dataset in Excel File:
🗴 US Superstore data (1).xlsm

Business questions

1. What is the average time between order date and shipping date done
2. How much profit is gained for the top 3 categories and top 5 products?
3. People from city/state shop the most?(max orders)/ **Sales by region(map)**
4. Which customer is regular/loyal over the years? (William Brown: WB-21850)
5. Which products are often bought together?
6. Top 10 loyal customers from each segment?(retained customers)
7. Which segment (consumer, corporate, home office) is generating most profit/sales separately?/ Sales by customer segment
8. Which category is often bought by each segment? Top 5 for ecg
9. Top 10 selling states
10. Popular products in each sub-category?
11. Least 5 profit products(How many distinct products)
12. Most popular month/time of the year for sales?done

    Most orders placed (Sep, Nov, Dec)(Q: order and ship date have the same trends. Can I just call it sales?)(hypothesis)

    Months with least orders placed(Jan Feb)

Insights

1.Bundle up mostly sold with least sold in a package

2. Find out most frequent buyers over the years(loyal customer) and ask them about improvements because they'll mostly be nearby residents and have better idea of what needs to be improved

3. I have 3 years of data and in all the years sales have increased. In the last 2 years sales increased by $100,000 so the stakeholders next target can be to target the increase of the same amount in the coming year. That can happen if I calculate average purchase value and give an estimate of how many customers are needed for coming years to generate target sales.

4. Technology category generates the most profit. That can be bundled up with most bought charging cables and batteries to increase revenue.

5. Sep Nov Dec are high selling months. Marketing team should actively promote 2 weeks prior to peak time and make sure the stock is available.

6. Tables and Bookcase sub categories have been generating loss. The store can consider dropping them from the catalog or reduce discounts on such items.

7. California and New York are among the top two cities in sales. Products should be marketed more here and a survey should be conducted for the cities with least sales. (Regions wise there are 4 regions east has highest sales and that can specifically be targeted for marketing as well)

8. Discounts more than 50% have a correlation with loss. The store can try not giving hefty discounts on such products.

Metrics for Dashboard:

On time shipping

Sales to date

Sales growth(rise of sales per year)

Regional Sales

Average Purchase Value

```
Total ($) Value of Orders / (#) of Orders over a Defined Period = ($) Average
Purchase Value
```