Sales Trends Analysis of a Superstore - Week 1

Project Overview

The project entails investigating into Superstore Sales data to determine business trends. Week 1 objective is to clean, explore, and visualize any dataset to get the useful information.

Objectives

Load and clean the data and preprocess it.

Learn how to understand the data structure, missing values and parse dates.

Look at the trend in sales and profit in the categories and geographies.

Visualize knowledge in bar charts, line graphs and heatmaps.

Make three to five main points found through visual evidence.

Steps Performed

1. Downloaded important libraries such as pandas, numpy, seaborn and matplotlib.
2. Imported the Superstore dataset with encoding.
3. Interpreted the dataset with .head(), .info() and .describe().
4. Verified the missing values and outliers.
5. Added training 5% random missing values in Sales and profit column.
6. Treated missing data through imputing with mean.
7. Ordered Date and Ship Date converted to datetime.
8. Group data 'Category' and 'Region' to use exploratory level of analysis.
9. Applied bar plots, line graphs and a heatmap to visualize data.

Important Business Insights

Best-Selling Category: Technology takes first place in the number of sales (Bar Chart).

Most Profitable Region: region East has the greatest profit (Bar Chart).

Sales Trend: November and December offer the highest times of sales (Line Graph).

Discount Impact: Discount and Profit have negative -0.22 correlation (Heatmap).

Conclusion

Week 1 - Sales Trends Analysis of a Superstore
The following analysis gives business intelligence to the decision-makers of the Superstore. The week 1 is an effective demonstration on preparing and investigating the real world data.