Superstore Sales Analysis

Data Science Internship- Coding Samurai

by **Ridhwan S**

Executive Summary

This project aimed to analyze sales performance, profit trends, and discount strategies using real-world retail data from a US-based superstore. Python and its visualization libraries were used to uncover business insights and recommendations for improved decision-making.

Dataset Overview

- 9,994 rows × 21 columns
- Key fields: Order Date, Sales, Profit, Category, Region, Discount
- Source: Kaggle (Superstore Dataset)

Methodology

- Tools: Python, Pandas, Matplotlib, Seaborn
- Techniques: Data Cleaning, EDA, Visualizations
- Timeline: 1 week

Key Findings

- Technology is the most profitable product category
- New York, LA, and Seattle are top-performing cities
- December sees sales spikes — but not always in profit
- Discounts over 30% heavily reduce profitability

Recommendations

- Optimize discount strategy: avoid >30% discounts
- Promote Tech more aggressively
- Prioritize marketing in East and West regions

Appendix

- Link to GitHub repo
- Link to presentation PDF
- Visuals and charts from analysis