Supermarket Sales Analysis: Exploratory Data Analysis and Data Visualization

Abstract:

In this project, we conduct exploratory data analysis (EDA) and data visualization on a dataset containing supermarket sales data. The aim is to extract insights and patterns from the data using various statistical techniques and visualizations. We employ Python as the primary tool for data analysis, utilizing libraries such as Pandas, Matplotlib, Seaborn, and Plotly for data manipulation, visualization, and interpretation.

The project follows a structured design flow, beginning with data preprocessing to clean and prepare the dataset for analysis. This includes handling missing values, converting data types, and ensuring data consistency. Following preprocessing, we conduct EDA and employ techniques like descriptive statistics, correlation analysis, and data summarization techniques to uncover patterns and trends in the data. We create plots such as histograms, box plots, scatter plots, and heat maps to showcase distributions, trends, correlations, and outliers within the data.

Through our analysis and visualization efforts, we aim to answer pertinent business questions, such as understanding sales trends over time, identifying bestselling products, analyzing customer purchasing behavior, and uncovering relationships between different variables. Insights derived from this analysis can inform strategic decision-making, marketing campaigns, and inventory management strategies for the supermarket.

In conclusion, this project demonstrates the power of Python for conducting exploratory data analysis and data visualization on real-world datasets. By leveraging Python libraries and following a structured analysis workflow, we uncover valuable insights from the data, enabling stakeholders to make informed decisions and drive business growth.

Keywords: Python, Exploratory Data Analysis, Data Visualization, Pandas, Matplotlib, Seaborn, Plotly, Supermarket Sales.

SUBMITTED BY:

V. SRI PADMA VILASINI	202U1A3359
M. SRINIJA	212U5A3303
V. SOWMYA LAKSHMI	202U1A3358
K. RANJANI	202U1A3322