

Exploratory Data Analysis on a Supermarket Sales Dataset

Problem Statement:

Supermarkets play a critical role in the retail sector, providing essential goods and services to customers. Effective decision-making based on data can help businesses optimize operations, understand customer behavior, and improve profitability. This project aims to analyze a dataset of supermarket sales, which contains records of sales transactions, customer demographics, and other key features.

The objective is to perform Exploratory Data Analysis (EDA) on the supermarket sales dataset to uncover actionable insights that can help the supermarket management make data-driven decisions. Insights could focus on identifying high-performing branches, understanding customer preferences, analyzing payment methods, and determining factors influencing sales and customer ratings.

Dataset:

<https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales>

Steps to Complete the Project:

1. Load and Understand the Data:

- Import the dataset using pandas.
- Inspect the data using `.head()`, `.info()`, `.describe()`, and `.shape()`.
- Check for null or missing values.

2. Data Cleaning:

- Handle missing values if any.

- Format columns (e.g., date format or categorical data).
- Remove duplicates if necessary.

3. Exploratory Data Analysis (EDA):

- Analyze sales by gender, product category, branch, and payment method.
- Study customer demographics (e.g., gender, age group) and their spending behavior.
- Examine the relationship between different features, such as:
 - Total Sales vs. Branches
 - Sales vs. Payment Method
 - Ratings vs. Sales

4. Data Visualization:

Use libraries like Matplotlib and Seaborn for visualization. Examples:

- Pie chart: Sales distribution by branch.
- Bar chart: Payment methods vs. total sales.
- Heatmap: Correlation between numerical features.
- Boxplot: Sales distribution by customer gender.

5. Summary and Insights:

- Write a summary of your findings.
- Highlight interesting patterns like:
 - Which branch generates the most revenue?
 - Which payment method is most popular?
 - Does gender influence the average sales amount?