Visualizing Time Series Retail Sales Data

Sreelakshmi E U,

SREE SANKARA COLLEGE KALADY

MSc STATISTICS WITH DATA SCIENCE

Period of Internship: 25th August 2025 - 19th September 2025

Report submitted to: IDEAS – Institute of Data Engineering, Analytics and Science Foundation, ISI Kolkata

1. Abstract

The objective of this project is to perform an exploratory time series analysis and visualization of retail sales data to uncover patterns, trends, and seasonality in business performance. The dataset contains chronological sales records, which are analyzed to understand how sales vary over time, across months, and days of the week. Through comprehensive data preprocessing and visualization techniques in Python (using libraries such as pandas, matplotlib, seaborn, and statsmodels), key insights about sales behavior are derived. Line plots, moving averages, and seasonal decomposition are used to highlight underlying trends and periodic fluctuations. The study reveals seasonal peaks during specific months, consistent growth patterns over time, and weekday-level variations in customer purchasing behavior. These visual and statistical analyses provide valuable inputs for demand forecasting, inventory management, and strategic decision-making in retail operations.

2. Introduction

In today's competitive retail environment, businesses generate large volumes of transactional data every day. Analyzing this data effectively can help organizations understand customer behavior, monitor performance, and make data-driven decisions. Time series analysis plays a crucial role in this process by revealing how sales evolve over time and identifying patterns such as growth trends, seasonal peaks, and demand fluctuations.This project focuses on performing an exploratory data analysis (EDA) and visualization of retail sales data using Python. By examining sales trends across different periods— such as daily, monthly, and yearly intervals—the analysis aims to uncover meaningful insights into business performance. Techniques like moving averages and seasonal decomposition are applied to detect long-term trends and recurring seasonal effects.The study also explores relationships between key variables such as sales, profit, quantity, and discounts through visualizations like line charts, bar plots, and heatmaps. The insights gained from this analysis can assist retailers in

forecasting future sales, optimizing inventory, planning promotions, and improving strategic decision-making. Overall, this project demonstrates how time series visualization serves as a powerful tool for understanding and enhancing retail business operations.

Relevance

Analyzing time series retail sales data is highly relevant for understanding business performance and customer demand patterns. It helps retailers identify sales trends, seasonal variations, and high-demand periods, enabling better forecasting and inventory management. By visualizing sales behavior over time, businesses can make informed decisions on marketing strategies, pricing, and resource allocation, ultimately improving profitability and operational efficiency.

3. Project Objective

To analyze and visualize retail sales data over time and identify trends, seasonal patterns, and fluctuations in sales. Also explore relationships between key variables like sales, profit, and quantity and generate insights that support forecasting and business decision-making.

Technology Involved

- **Python** (for data analysis and visualization)
- **Pandas & NumPy** (for data manipulation)
- **Matplotlib, Seaborn, and Plotly** (for visualization)
- **Jupyter Notebook** (for implementation and documentation)

Background Material Survey

Before starting the project, a review of standard practices in data analysis and visualization was carried out. This included:

- Descriptive statistics to summarize datasets
- Handling missing values and duplicates

- Time-series trend analysis
- Category-wise grouping and visualization

Procedure Used

1. **Data Cleaning** – Removing duplicates, handling missing values, and formatting date fields.
2. **Exploratory Data Analysis** – Performing descriptive statistics and generating insights.
3. **Visualization** – Using line charts, bar charts, and categorical plots to show trends.
4. **Insights & Interpretation** – Identifying seasonal profit fluctuations and best-selling categories.
5. **Time Series Visualization Technique and Seasonal Decomposition   :**
   Line plots, bar charts, and moving average plots were created to observe trends, seasonality, and fluctuations in sales.
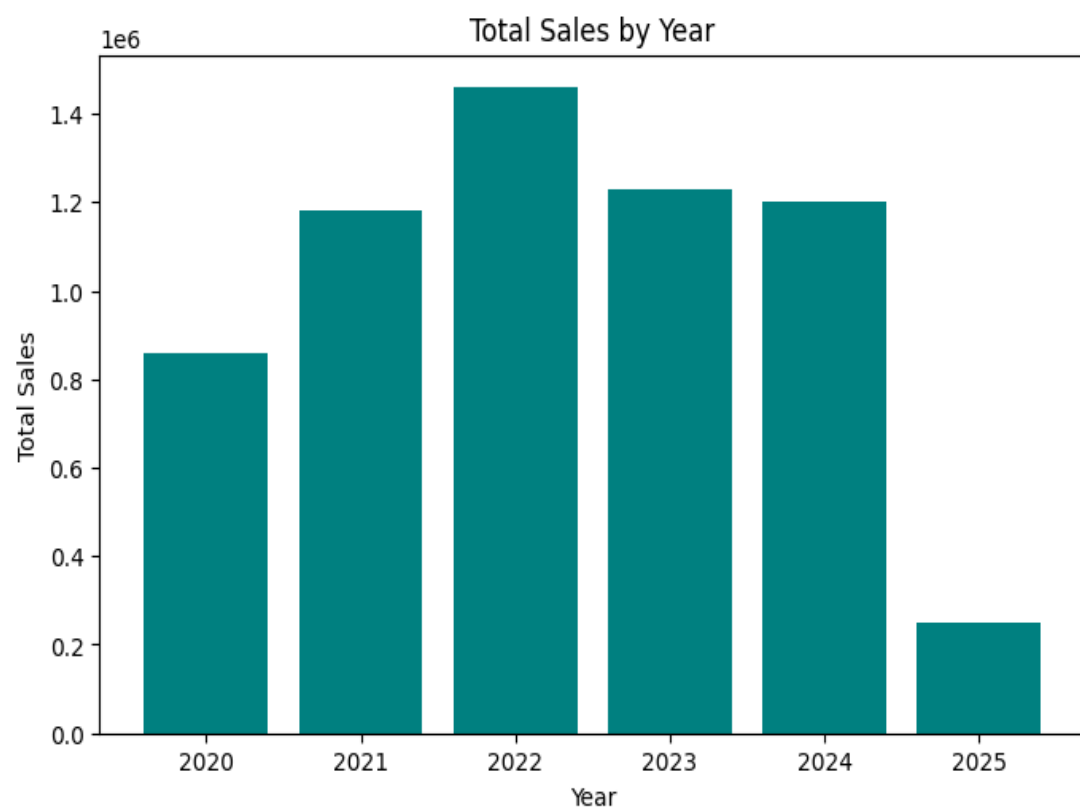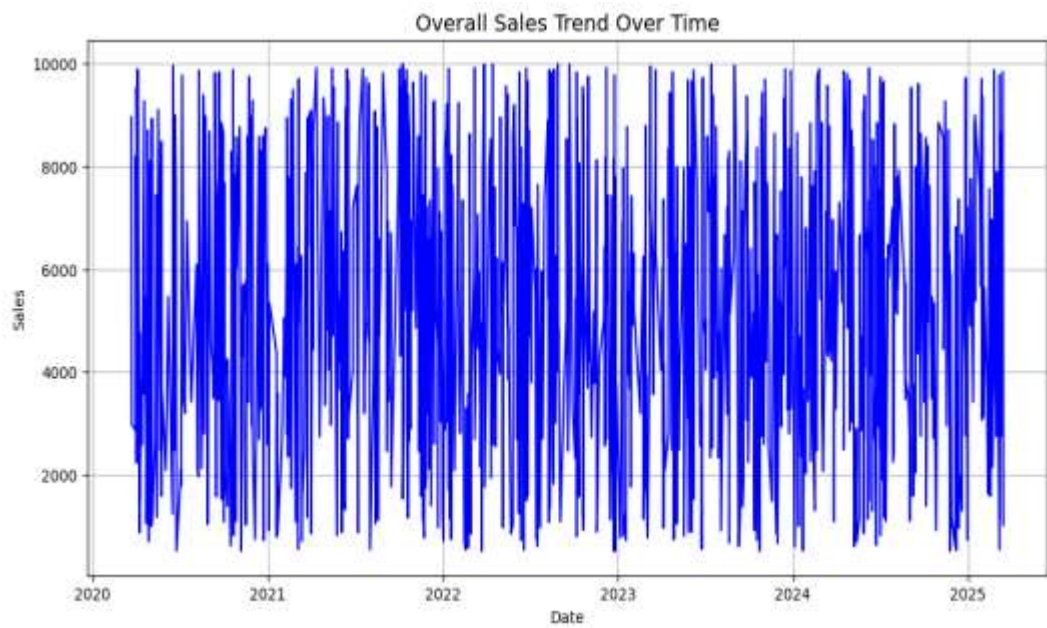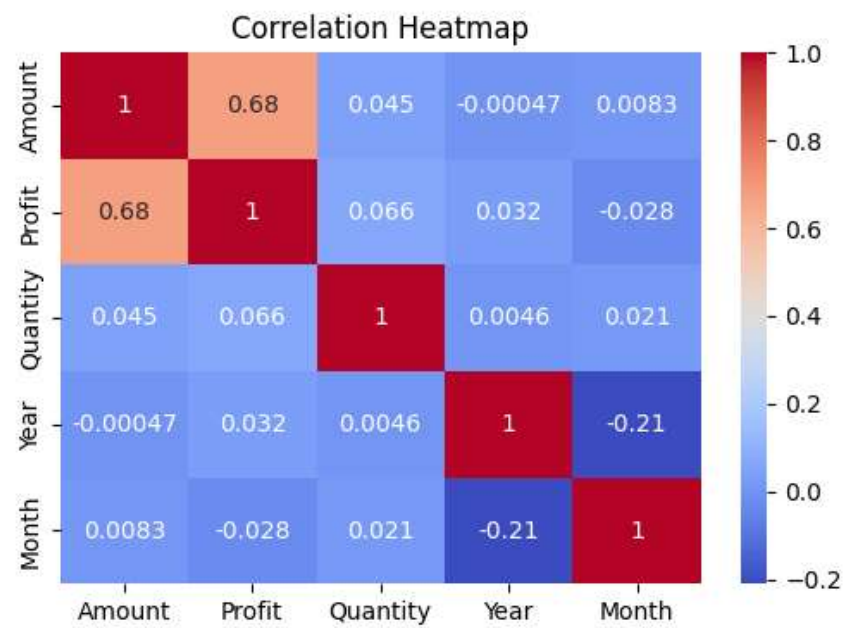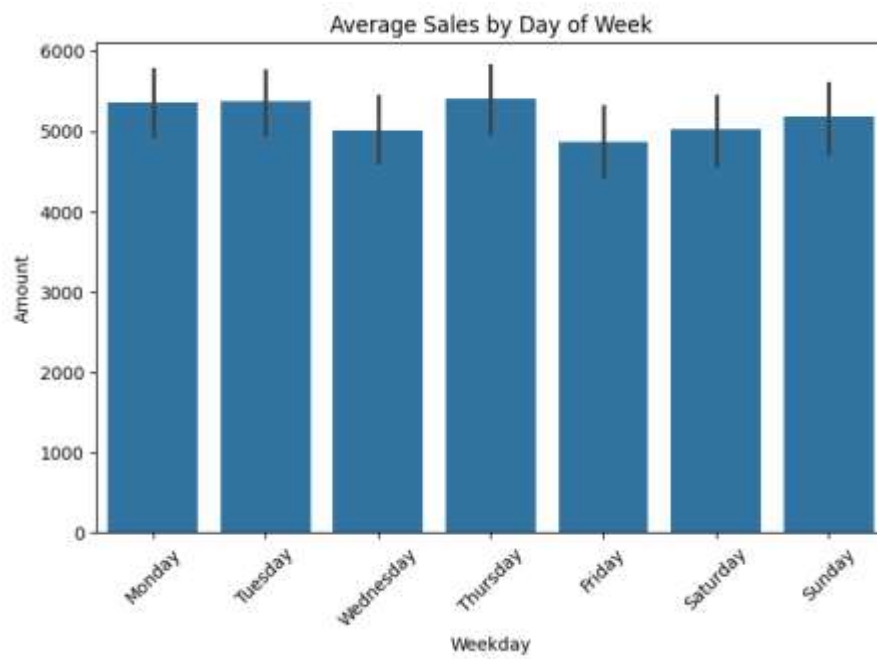
Purpose of the Project

The purpose of this project is to analyze and visualize retail sales data over time to identify trends, seasonal patterns, and key factors affecting sales performance, enabling data-driven business decisions and accurate forecasting.
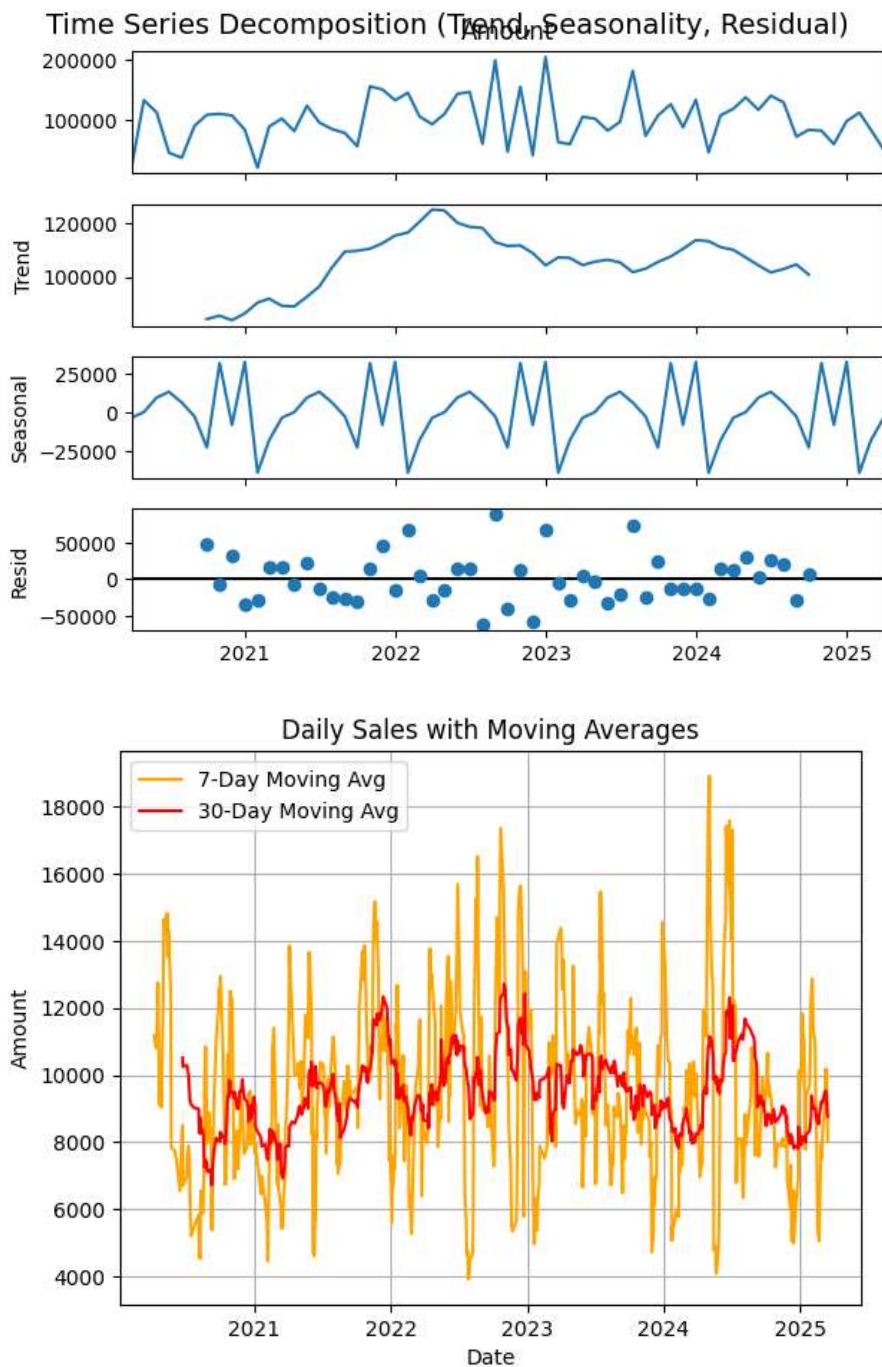
4. Methodology

The dataset was cleaned and preprocessed, followed by feature extraction from date variables. Exploratory Data Analysis (EDA) was performed using statistical measures and visualizations to identify trends, patterns, and seasonality. Time series decomposition and correlation analysis were conducted to derive meaningful insights for forecasting and business decisions.

5. Data Analysis and Results

Overall Sales Trend Over Time

Total Sales by Year

Average Sales by Day of Week

Correlation Heatmap

Time Series Decomposition (Trend, Seasonality, Residual)

Daily Sales with Moving Averages

6. Conclusion

The analysis of retail sales time series data revealed clear trends and seasonal patterns that influence sales performance. Visualizations helped identify peak sales periods and relationships between key

variables. These insights can support better forecasting, inventory planning, and strategic business decisions.

7. APPENDICES

References

1. Pratap Dangeti, *Statistics for Machine Learning*, Packt Publishing.
2. Tom M. Mitchell, *Machine Learning*, McGraw Hill.
3. Towards Data Science (https://towardsdatascience.com/)
4. Kaggle – Retail Sales Datasets (https://www.kaggle.com/)
5. Documentation of Python libraries: Pandas, NumPy, Matplotlib, Seaborn, Plotly.

Survey Questionnaire

1. How often do you purchase products from our store?
   - Weekly / Monthly / Occasionally
2. Which product categories do you prefer?
   - Electronics / Furniture / Clothing / Others
3. What is the usual budget range of your purchases?
   - < ₹1,000 / ₹1,000–₹5,000 / > ₹5,000
4. How satisfied are you with product quality?
   - Very Satisfied / Neutral / Dissatisfied
5. Would you recommend our store to others?
   - Yes / No

Document Links

Colab Notebook link :

https://colab.research.google.com/drive/1myk-dzczTFjYQQmPzjmuE3GDUDzO31ls?usp=sharing