Identifying Shopping Trends using Data Analysis

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning
with
TechSaksham – A joint CSR initiative of Microsoft & SAP

by

Arup Ranjan Dalai, arup240902@gmail.com

Under the Guidance of

Jay Rathod

ACKNOWLEDGEMENT

I am grateful to all who supported me throughout this project. Special thanks to my supervisor, Jay Rathod, for their guidance and encouragement, which were instrumental in completing this work successfully.

ABSTRACT

This project analyzes shopping trends using a dataset of 3,900 customer transactions. Key objectives include understanding demographic influences on purchases, identifying popular items, and evaluating seasonal and promotional impacts. Using Python for data analysis, findings highlight behavioral patterns, such as preferred payment methods and the role of subscription status. Insights provide actionable recommendations for targeted marketing and inventory management.

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1

Introduction

1.1Problem Statement:

Understanding customer behavior and shopping trends is crucial for businesses to optimize operations and enhance customer satisfaction. This project addresses the need to analyze patterns in a shopping dataset.

1.2Motivation:

The motivation for this project stems from the growing demand for data-driven decision-making in retail. Analyzing customer preferences enables businesses to deliver personalized experiences.

1.3Objective:

- Explore demographic and behavioral trends.
- Analyze the impact of discounts, seasons, and subscriptions.
- Identify opportunities for business optimization.

1.4Scope of the Project:

The analysis focuses on descriptive statistics and actionable insights. Limitations include potential biases in the dataset and lack of real-time updates.

CHAPTER 2

Literature Survey

2.1 Review relevant literature or previous work in this domain.

Data analytics has revolutionized the retail industry, enabling businesses to optimize operations and personalize customer experiences. Previous studies have focused on customer segmentation, demand forecasting, and trend analysis to enhance decision-making.

2.2 Mention any existing models, techniques, or methodologies related to the problem.

Common techniques include clustering for customer segmentation, regression models for sales prediction, and association rule mining for market basket analysis. Visualization tools like heatmaps and dashboards are often employed for trend insights.

2.3 Highlight the gaps or limitations in existing solutions and how your project will address them.

Existing solutions often overlook holistic trend analysis, focusing on individual aspects like sales or segmentation. This project bridges these gaps by combining demographic, seasonal, and behavioral trends, offering a comprehensive view to inform marketing and inventory strategies.

CHAPTER 3

Proposed Methodology

3.1 System Design

The proposed solution follows a structured workflow:

- Data Collection: The dataset is loaded and cleaned to ensure accuracy and consistency.
- Data Analysis: Key questions are addressed using statistical and visual analysis with Python libraries like Pandas, Matplotlib, and Seaborn.
- Insights Generation: Patterns and trends are identified to provide actionable insights for businesses.Requirement Specification.

System Design Diagram:-

[Data Source] ➔[Data Preprocessing] ➔[Analysis] ➔[Insights & Reporting]

Mention the tools and technologies required to implement the solution.

3.2.1 Hardware Requirements:-

Standard computing device with at least 4GB RAM.

- Intel Core i3 processor or higher.
- 10GB free disk space for data storage and processing.

3.1.1 Software Requirements:-

- Python 3.8 or higher.
- Jupyter Notebook for interactive coding.
- Libraries: Pandas, Matplotlib, Seaborn, NumPy.
- GitHub for version control and code sharing.

CHAPTER 4

Implementation and Result

Snap Shots of Result:-

Age Distribution: Average age is ~44 years, with customers ranging from 18 to 70 years.

Category Analysis: Footwear has the highest average purchase amount ($60.26).

Gender Trends: Males make 68% of purchases.

Seasonal Spending: Fall sees the highest spending.

Payment Methods: PayPal is the most popular.

Promotional Impact: Non-promo purchases yield slightly higher spending.

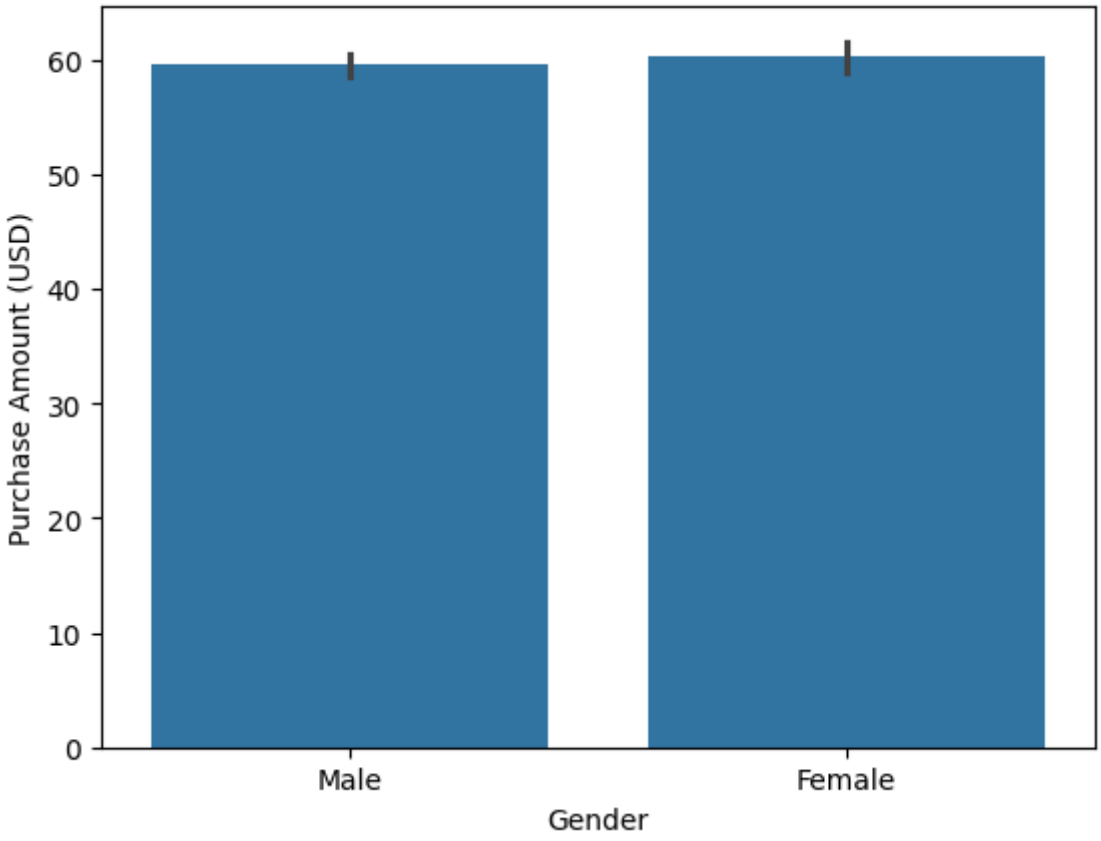

Figure 1: Gender-wise Distribution of Purchases.

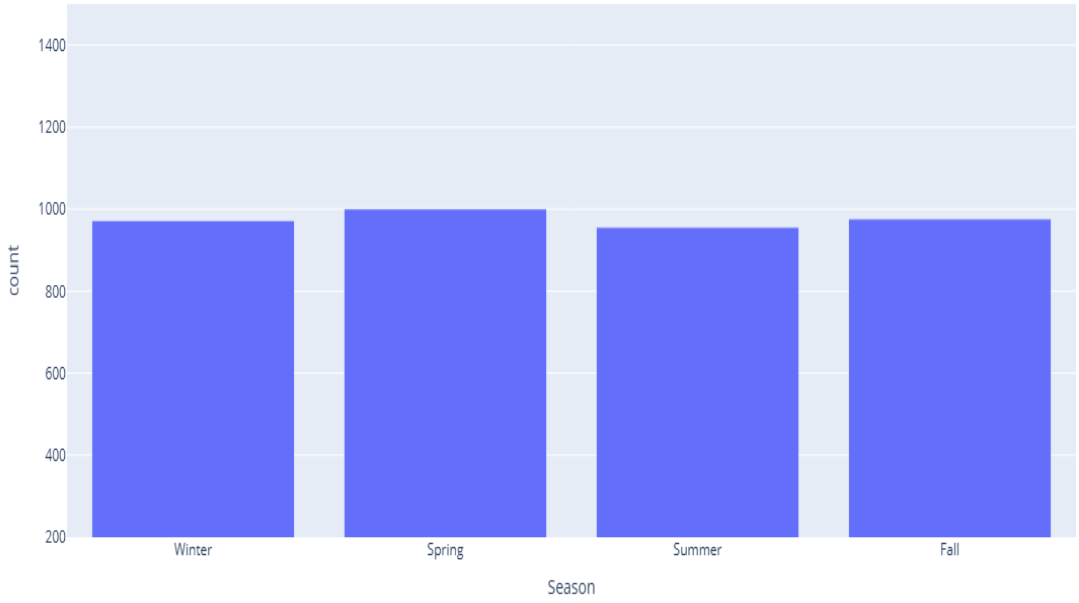

Figure 2: Distribution of Purchase Amounts by Season.

GitHub Link for Code:

https://github.com/ARUP-24/Identifying-Shopping-Trends-using-Data-Analysis/blob/main/shopping_trends_updated.csv

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

Incorporate machine learning models for predictive analysis and explore real-time trends.

5.2 Conclusion:

The project demonstrates how data analytics can uncover valuable insights into shopping trends. These insights can inform marketing strategies, inventory planning, and customer engagement initiatives.

REFERENCES

[1]. Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, “Detecting Faces in Images: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume. 24, No. 1, 2002.