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

T SAI GANESH, saiganeshtammannagari@gmail.com

Under the Guidance of

JayRathod and P Raja

ACKNOWLEDGEMENT

We would like to take this opportunity to express our gratitude to all individuals who contributed to this project. We extend our sincere thanks to our mentor, Jay Rathod and P Raja sir, for their guidance and support throughout this study. Their valuable insights and expertise played a crucial role in the successful completion of this project.

ABSTRACT

This project focuses on analyzing shopping trends using an extensive dataset. The study aims to understand customer behavior, purchasing patterns, and preferences by performing Exploratory Data Analysis (EDA). Various factors such as age distribution, purchase frequency, preferred payment methods, and seasonal trends were examined. The results provide valuable insights for businesses to enhance marketing strategies and improve customer experiences. The findings reveal significant trends, such as peak spending seasons and the impact of discounts on purchases, which can aid in strategic decision-making.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Problem Statement:

Understanding customer shopping behavior is crucial for businesses to optimize their marketing strategies and inventory management. This project aims to analyze shopping trends and uncover key insights from a given dataset.

1.2Motivation:

The rise of e-commerce and digital transactions has led to an abundance of consumer data. Analyzing this data provides businesses with actionable insights to improve sales, enhance customer satisfaction, and optimize operations.

1.3Objective:

To analyze customer demographics and purchasing behavior.

To identify the most purchased product categories.

To determine seasonal trends in shopping.

To evaluate the impact of discounts and promotional offers on sales.

To compare purchasing behavior based on different factors like age, gender, and payment methods.

1.4Scope of the Project:

This project involves the use of Python-based exploratory data analysis (EDA) techniques to extract insights from a shopping dataset. The study focuses on customer behavior analysis, trend detection, and predictive insights to aid decision-making in e-commerce and retail industries

CHAPTER 2

Literature Survey

Several studies have been conducted on consumer shopping behavior and purchasing trends. Researchers have utilized data-driven approaches to analyze seasonal trends, payment preferences, and the influence of discounts on shopping behavior. However, existing studies often lack real-time insights into how different customer demographics impact purchasing decisions. This project aims to fill that gap by conducting an in-depth analysis of customer purchase behavior.

CHAPTER 3

Proposed Methodology

3.1 System Design

The analysis involves data preprocessing, visualization, and trend identification using Python libraries such as Pandas, Matplotlib, and Seaborn.

3.2 Requirement Specification

Mention the tools and technologies required to implement the solution.

Hardware Requirements: Computer with at least 8GB RAM Processor: Intel i5 or higher processor

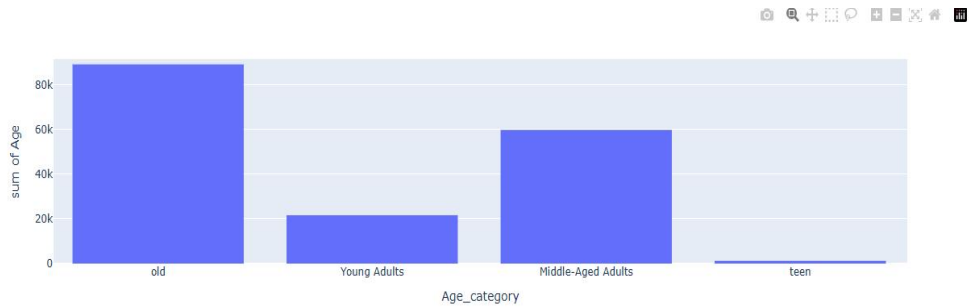
- **Software Requirements:** Jupyter Notebook

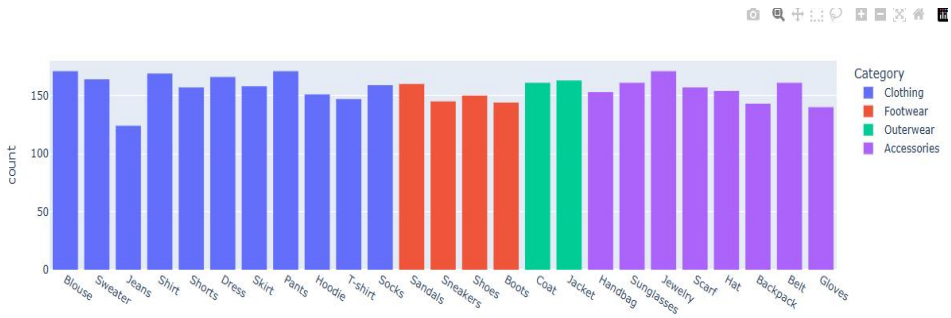
Python (Pandas, NumPy, Seaborn, Matplotlib, Scikit-learn) Excel or CSV data handling tools

CHAPTER 4

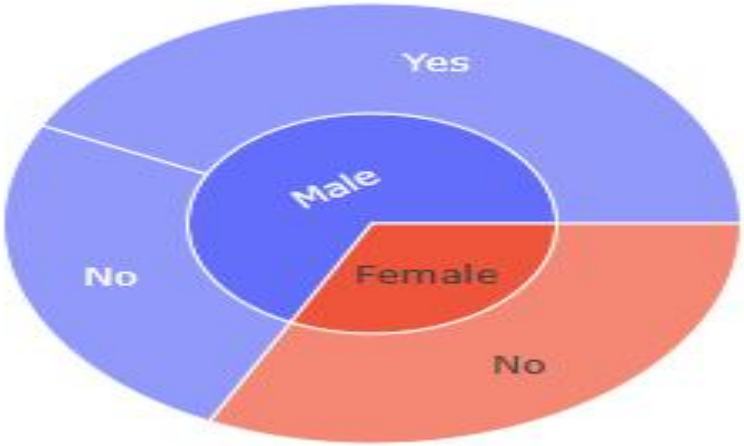
Implementation and Result

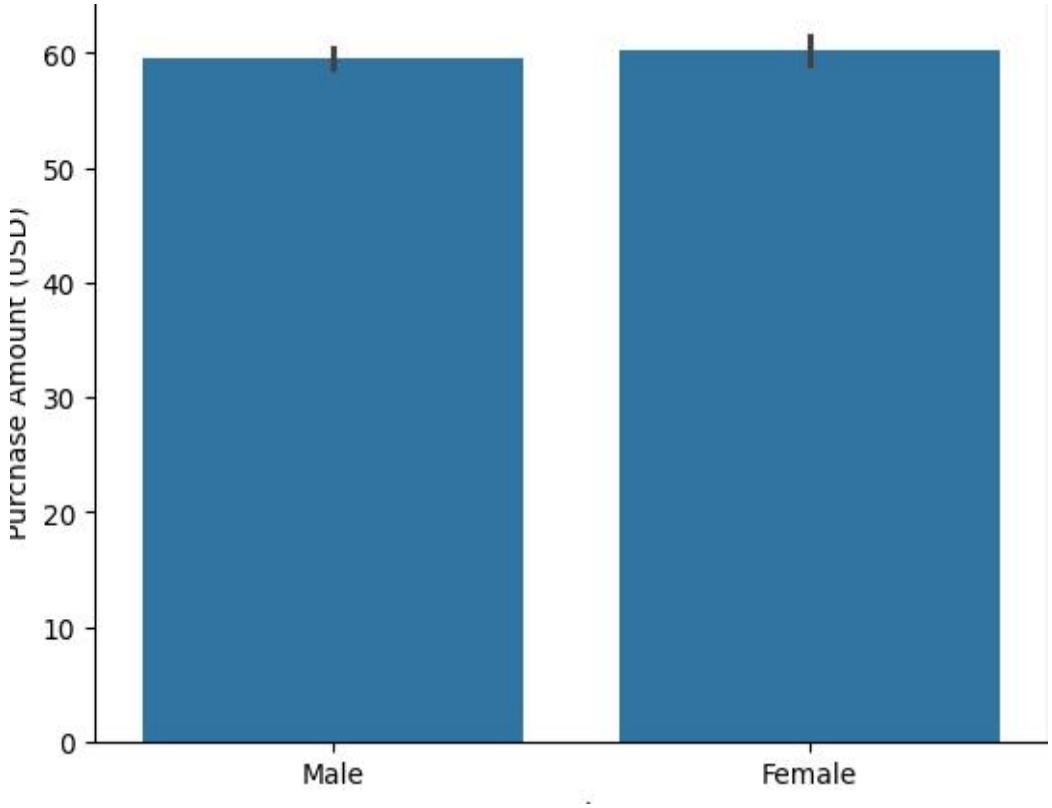
4.1 Snap Shots of Result:





4.2

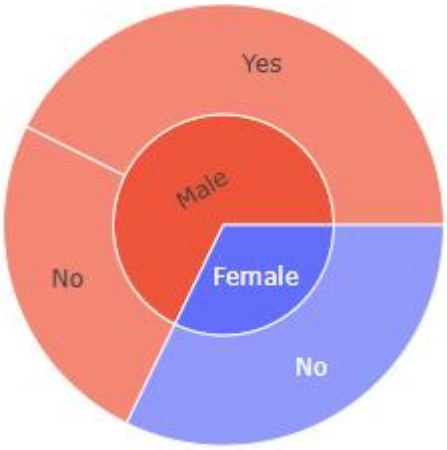



Purchase Amount (USD)
60
50
40
30
20
10
0
Male
Female

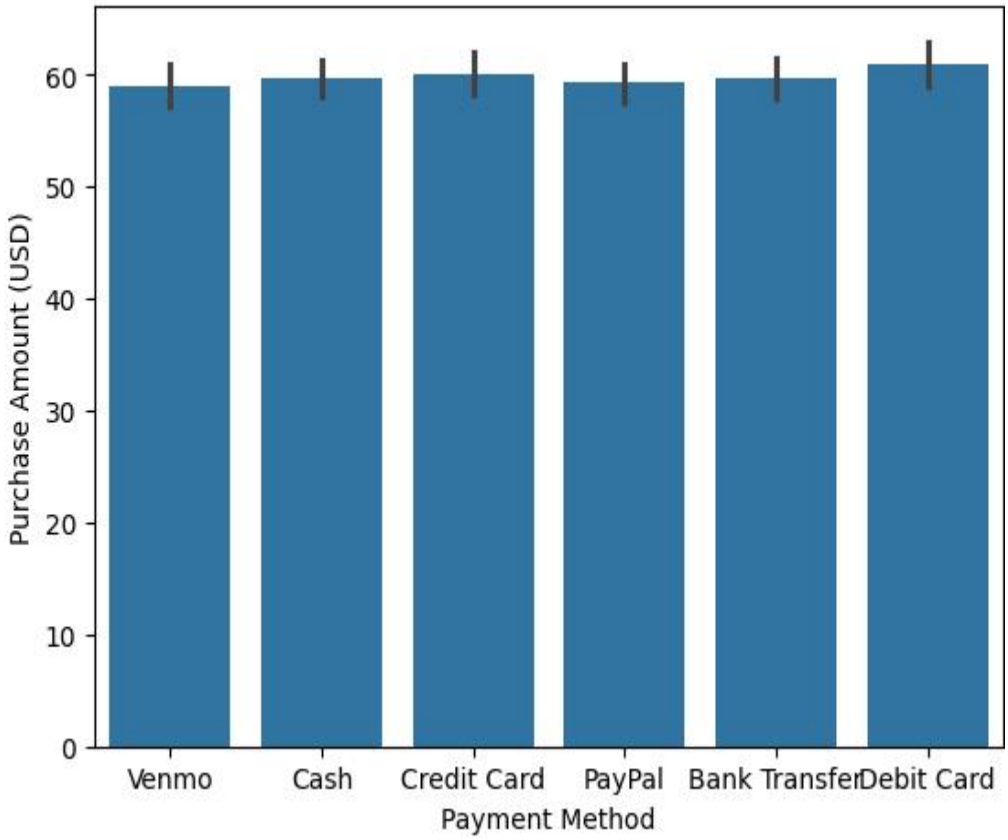
4.3

SunBurst:


old
old
old
old
old
old
Quarterly
Annually
Monthly
Weekly
Bi-Weekly
Fortnightly


Yes
No
No
Male
Female

4.4



4.5

4.6 GitHub Link for Code:

1. What is the overall distribution of customer ages in the dataset?

- The average customer age is **44 years**, with most customers aged between **31 and 57 years**. The youngest customer is **18**, and the oldest is **70**.

2. How does the average purchase amount vary across different product categories?

- **Accessories**: $59.84
- **Clothing**: $60.02
- **Footwear**: $60.25
- **Outerwear**: $57.17

3. Which gender has the highest number of purchases?

- **Male customers** have made more purchases (**2652** purchases) than female customers (**1248** purchases).

4. What are the most commonly purchased items in each category?

- **Accessories**: Jewelry
- **Clothing**: Blouse
- **Footwear**: Sandals
- **Outerwear**: Jacket

5. Are there any specific seasons or months where customer spending is significantly higher?

- The highest spending occurs in **Fall ($60,018)**, followed by **Spring ($58,679)** and **Winter ($58,607)**.

6. What is the average rating given by customers for each product category?

- **Accessories**: 3.77
- **Clothing**: 3.72
- **Footwear**: 3.79
- **Outerwear**: 3.75

7. Are there any notable differences in purchase behavior between subscribed and non-subscribed customers?

- **Non-subscribed customers** have a slightly higher average purchase amount ($59.87) compared to **subscribed customers** ($59.49).

8. Which payment method is the most popular among customers?

- The most used payment methods are:
    - **PayPal** (677 transactions)
    - **Credit Card** (671 transactions)
    - **Cash** (670 transactions)

9. Do customers who use promo codes tend to spend more than those who don't?

- **No**, customers who use promo codes spend **less** ($59.27) compared to those who don't ($60.13).

10. How does the frequency of purchases vary across different age groups?

- **18-24 years**: Most common frequency: *Every 3 months*
- **25-34 years**: Most common frequency: *Bi-weekly*
- **35-44 years**: Most common frequency: *Annually*
- **45-54 years**: Most common frequency: *Bi-weekly*
- **55-64 years**: Most common frequency: *Annually*
- **65+ years**: Most common frequency: *Quarterly*

11. Are there any correlations between the size of the product and the purchase amount?

- Customers purchasing **size S** spend the most ($61.04), while **size L** purchases have the lowest average ($58.56).

12. Which shipping type is preferred by customers for different product categories?

- **Accessories**: Store Pickup
- **Clothing**: Standard
- **Footwear**: Free Shipping
- **Outerwear**: Free Shipping

13. How does the presence of a discount affect the purchase decision of customers?

- Customers **without discounts** spend slightly **more** ($60.13) compared to those with discounts ($59.27).

14. Are there any specific colors that are more popular among customers?

- The top **5 most purchased colors** are:
    - Olive (177 purchases)
    - Yellow (174 purchases)
    - Silver (173 purchases)

- Teal (172 purchases)
- Green (169 purchases)

15. What is the average number of previous purchases made by customers?

- The average number of previous purchases is **25.35**.

16. How does the purchase amount differ based on the review ratings given by customers?

- Products with **higher ratings (4.9-5.0)** tend to have **higher** purchase amounts (~$63-$64).
- Products with **lower ratings (2.5-2.9)** have **lower** purchase amounts (~$56-$57).

17. Are there any noticeable differences in purchase behavior between different locations?

- The **top 5 locations** with the highest total spending are:
  - **Alabama**: $5,261
  - **Alaska**: $4,867
  - **Arizona**: $4,326
  - **Arkansas**: $4,828
  - **California**: $5,605

18. Is there a relationship between customer age and the category of products they purchase?

- Across **all age groups**, **Clothing** is the most purchased category.

19. How does the average purchase amount differ between male and female customers?

- **Females** spend slightly **more** per purchase ($60.24) compared to **males** ($59.53).

CHAPTER 5

Discussion and Conclusion

5.1 Future Work

- Implement predictive modeling to forecast customer preferences.
- Analyze more factors such as brand influence and seasonal promotions.
- Expand the dataset to include global trends.

5.2 Conclusion

The analysis provides valuable insights into shopping trends, helping businesses make data-driven decisions. Findings show variations in spending based on demographics, seasonality, and product categories. Businesses can use these insights to tailor marketing strategies and optimize customer engagement.

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.