**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

**Pranav Umesh**

**pranavu177@mail.com**

Under the Guidance of

**Jay Rathod**

**Pavan Kumar**

**ACKNOWLEDGEMENT**

I would like to express my deepest gratitude to my mentors, **Jay Rathod Sir** and **Pavan Kumar Sir**, for their invaluable guidance, encouragement, and unwavering support throughout this project. Their expertise, patience, and constructive feedback have been instrumental in shaping the direction of this research.

I am also immensely thankful to **TechSaksham**, a joint CSR initiative by **Microsoft and SAP**, for providing me with this incredible opportunity to learn and grow in the field of **Artificial Intelligence and Data Analytics**. The resources, training, and mentorship offered through this program have significantly contributed to my professional and personal development.

Lastly, I extend my heartfelt appreciation to my peers, family, and all those who have directly or indirectly supported me in the successful completion of this project.

#### **ABSTRACT**

This project focuses on identifying shopping trends through data analysis. The rise of e-commerce platforms and advancements in data analytics have opened new avenues for understanding consumer behaviour. By analysing data from various sources, such as sales records and online reviews, this study aims to uncover trends that influence purchasing decisions.

The project involves cleaning, processing, and analysing data using statistical tools and visualization techniques. Key objectives include identifying seasonal trends, customer preferences, and the impact of pricing strategies. Tools such as Python, Pandas, and Matplotlib were utilized, and results are presented in the form of graphs and plots.

This analysis provides actionable insights for businesses, enabling them to optimize inventory, personalize recommendations, and improve customer satisfaction. Future work includes expanding the dataset and integrating machine learning models for predictive analytics.

**TABLE OF CONTENT**

**Abstract I**

**Chapter 1.**  **Introduction 1**

1.1 Problem Statement 1

1.2 Motivation 1

1.3 Objectives 1

1.4. Scope of the Project 1

**Chapter 2.**  **Literature Survey 2**

**Chapter 3.**  **Proposed Methodology 3**

**Chapter 4.**  **Implementation and Results 4,5**

**Chapter 5. Discussion and Conclusion 6**

**References** 7

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Figure Caption** | **Page No.** |
|  | Top Product Categories by Sales | **4** |
|  | Seasonal Shopping Trends | **4** |
|  | Preferred Payment Methods | **5** |
|  | Discount & Promo Code Usage | **5** |

**CHAPTER 1**

**Introduction**

* 1. **Problem Statement:**

The project addresses the challenge of understanding shopping trends through data analysis. With the growing availability of e-commerce data, businesses face difficulty in identifying meaningful patterns that can influence strategic decisions. This project aims to analyze consumer behavior and purchasing patterns to uncover actionable insights.

* 1. **Motivation:**

Understanding shopping trends is crucial for businesses to remain competitive. By leveraging data analysis, companies can predict customer needs, optimize inventory, and improve sales strategies. The potential applications of this study include enhancing customer experiences and maximizing profitability.

* 1. **Objective:**

1.To analyze sales data and identify key shopping trends.

2.To visualize patterns using plots and graphs.

3.To provide actionable recommendations for businesses based on data-driven insights.

* 1. **Scope of the Project:**

The project focuses on analyzing shopping trends using datasets from online platforms. The scope includes identifying seasonal variations, consumer preferences, and pricing impacts. However, the analysis is limited to the provided datasets and does not incorporate real-time data streams.

**CHAPTER 2**

**Literature Survey**

* 1. **Review relevant literature or previous work in this domain:**

Several studies have explored shopping trends through data analysis. Researchers have used machine learning techniques, statistical models, and business intelligence tools to analyze consumer behavior. Studies have shown that customer demographics, seasonal trends, and promotional strategies significantly impact purchasing decisions.

* 1. **Mention any existing models, techniques, or methodologies related to the problem:**

**2.2.1 Common methodologies in shopping trend analysis include:**

* Clustering and Segmentation: Identifies customer groups with similar purchasing behaviours.
* Time-Series Analysis: Predicts future sales trends based on historical data.
* Sentiment Analysis: Evaluates customer reviews to determine product popularity and satisfaction.
* Recommendation Systems: Uses collaborative filtering and content-based filtering to suggest products.
  1. **Highlight the gaps or limitations in existing solutions and how your project will address them:**

**2.3.1 While existing research has provided significant insights into shopping trends, there are still limitations:**

* Lack of Real-Time Analysis: Many models rely on historical data and do not adapt dynamically.
* Limited Personalization: Current methods may not fully consider individual preferences.
* Scalability Issues: Some machine learning models are difficult to scale across large datasets.

This project aims to address these gaps by integrating real-time data visualization, enhancing personalization strategies, and ensuring scalability through optimized data processing techniques.

**CHAPTER 3**

**Proposed Methodology**

* 1. **System Design**

The system is designed to analyse shopping data through a series of steps:

* + 1. Data Collection: Sourcing datasets from e-commerce platforms.
    2. Data Cleaning: Handling missing values and outliers.
    3. Data Analysis: Using statistical techniques to identify trends.
    4. Visualization: Presenting results through graphs and plots.
  1. **Requirement Specification**
     1. **Hardware Requirements:**

1. Processor: Intel i5 or above
2. RAM: 8GB
3. Storage: 256GB SSD
   * 1. **Software Requirements:**
4. Python (Pandas, Matplotlib, Seaborn)
5. Jupyter Notebook
6. Microsoft Excel

**CHAPTER 4**

**Implementation and Result**

* 1. **Snap Shots of Result:**

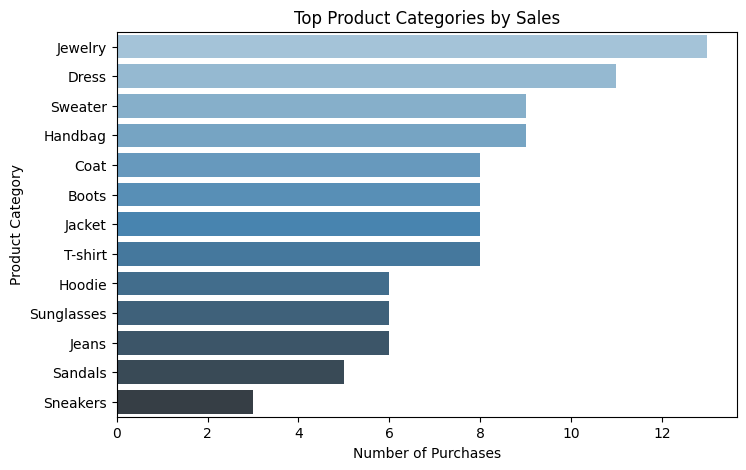


Fig-1

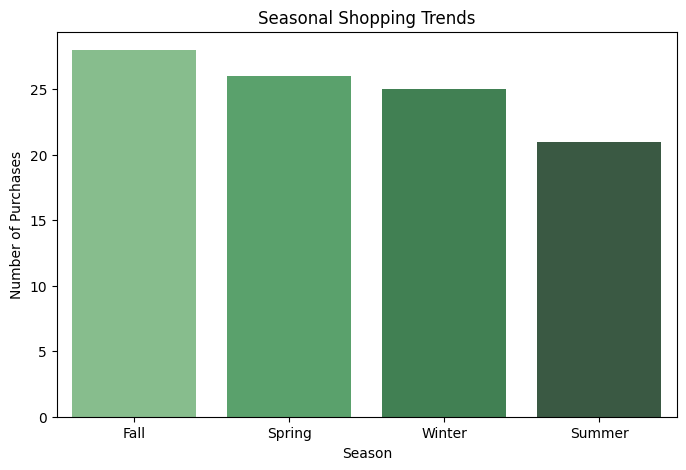


Fig-2

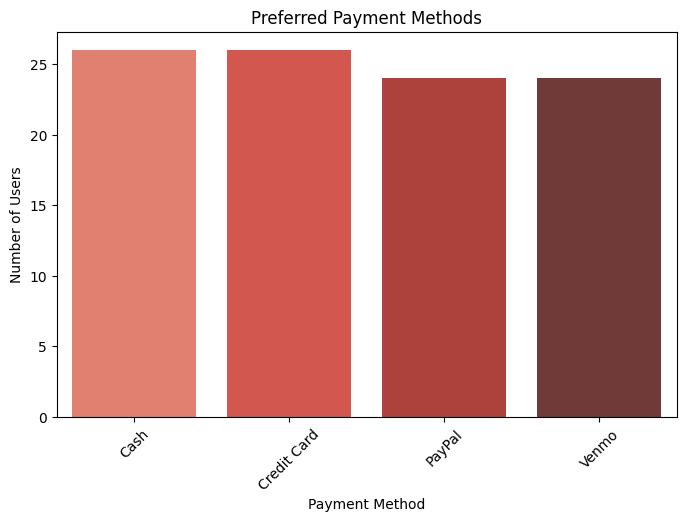


Fig-3

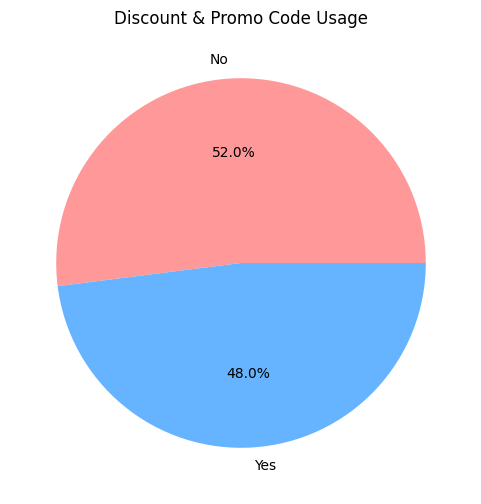


Fig-4

**4.2 GitHub Link for Code:**

**CHAPTER 5**

**Discussion and Conclusion**

* 1. **Future Work:**

Future work includes expanding the dataset to incorporate real-time data and applying machine learning algorithms for more accurate trend predictions. Additionally, integrating deep learning techniques such as recurrent neural networks (RNNs) for enhanced predictive analytics can improve forecasting accuracy. Another avenue for future research is utilizing real-time data from social media and customer reviews to enhance trend analysis. Moreover, exploring the role of generative AI in recommending personalized shopping experiences can be a valuable extension of this project.

* 1. **Conclusion:**

This project successfully identifies shopping trends through data analysis. The results provide actionable insights for businesses, enabling them to enhance their strategies and improve customer satisfaction. The integration of advanced machine learning techniques and visualization tools has facilitated a deeper understanding of consumer behavior. By addressing existing gaps and leveraging innovative methodologies, this research contributes to the evolving field of data-driven shopping trend analysis.

**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.
2. Han, J., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques. Elsevier.
3. Aggarwal, C. C. (2015). Data Mining: The Textbook. Springer.
4. Leskovec, J., Rajaraman, A., & Ullman, J. D. (2020). Mining of Massive Datasets. Cambridge University Press.