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

Avimta Ajit Shahane

shahane.avimta@gmail.com

Under the Guidance of

Jay Rathod

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deep sense of gratitude to all individuals who helped me directly or indirectly during this thesis work.

Firstly, I would like to thank my supervisor Mr. Jay Rathod, for being a great mentor and the best adviser I could ever have. His advice, encouragement and the critics are a source of innovative ideas, inspiration and causes behind the successful completion of this project. The confidence shown in me by him was the biggest source of inspiration for me. He always helped me during my project and many other aspects related to the program. His talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional.

ABSTRACT

This project focuses on utilizing data analysis techniques to uncover and analyze significant shopping trends within the contemporary retail landscape. By examining extensive datasets encompassing consumer purchase history, browsing behavior, social media engagement, and market demographics, we will gain valuable insights into the evolving preferences and purchasing patterns of consumers. Key objectives include identifying emerging trends such as preferred product categories, seasonal buying patterns, and the influence of external factors like economic conditions and social events on consumer behavior. Furthermore, we will segment customers based on their shopping habits, demographics, and preferences to tailor marketing strategies and deliver personalized experiences. The project will also involve developing predictive models to forecast future demand, anticipate inventory requirements, and optimize pricing strategies. Through competitive analysis, we aim to identify market gaps and formulate recommendations for businesses to gain a competitive edge. The project will employ a combination of data mining techniques, including clustering, classification, regression analysis, and time series analysis. The findings will be presented through visualizations and reports, providing actionable insights to empower businesses to enhance their operations, improve customer satisfaction, and drive revenue growth in the dynamic and competitive retail market.

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Problem Statement:

In today's dynamic retail landscape, understanding and anticipating consumer behavior is crucial for businesses to thrive. However, raw data alone provides limited insights. This project aims to leverage data analysis techniques to identify and analyze key shopping trends. We want to understand how people shop. By analyzing data from online stores and physical shops, we can uncover hidden patterns in what people buy and when. This will help businesses predict future trends, target the right customers, and stay ahead of the competition.

1.2 Motivation:

The motivation behind choosing this topic is to identify and analyze key shopping trends of consumers. And also to improve data analysis skills by solving challenging problems on working with real world project.

1.3 Objectives:

- To understand customer preferences and purchase habits
- To forecast demand for specific products and categories
- To present findings through clear and insightful visualizations
- To understand the factors influencing customer purchasing decisions

1.4 Scope of the Project:

1.4.1 Scope:

- Gathering relevant data from various sources.
- Summary statistics, data visualization, and initial trend identification (EDA).
- Presenting actionable insights and recommendations.

CHAPTER 2

Literature Survey

2.1Literature Review:

The paper "Study of Customer Shopping Trends Analysis" delves into the evolving landscape of consumer behavior, particularly in the clothing industry. The literature review synthesizes key themes and findings from the research, drawing from various contexts provided.

- **Evolution of Consumer Behavior:** The introduction highlights the significant transformation in consumer shopping behaviors due to digital technologies and socio-economic changes. Businesses are increasingly leveraging data-driven insights to understand these patterns, which is essential for maintaining competitiveness in the market.
- **Data-Driven Insights:** The research emphasizes the challenge of converting vast amounts of consumer data into meaningful insights. It discusses the application of statistical analysis, machine learning algorithms, and data mining techniques to uncover complex relationships that influence customer shopping trends.
- **Demographic Factors:** A critical aspect of the study is the consideration of demographic variables such as age, gender, and location. These factors play a significant role in shaping consumer preferences and purchasing decisions. The research underscores the importance of understanding these demographics to tailor marketing strategies effectively.
- **Cultural and Economic Influences:** The literature review also addresses the impact of cultural influences and economic conditions on shopping behaviors. These elements are crucial for businesses aiming to align their offerings with consumer expectations and demands.
- **Integration of E-commerce and Traditional Retail**: The rise of e-commerce and mobile shopping applications has blurred the lines between online and offline retail experiences. The paper advocates for a holistic approach to analyzing customer shopping trends, recognizing that consumers engage with products across various platforms.
- **Statistical and Machine Learning Techniques**: The research employs advanced statistical methods and machine learning algorithms to analyze a large dataset sourced from retail transactions. This approach aids in identifying patterns, correlations, and forecasting future shopping trends, providing a robust framework for understanding consumer behavior.

- **Practical Implications for Retailers:** The findings of this research offer valuable insights for retailers, enabling them to enhance their marketing strategies, optimize inventory management, and provide personalized customer experiences. By understanding shopping behaviors, businesses can better align their products and services with the evolving demands of consumers.

2.2Existing Methodology:

The methodology section outlines the approach taken to conduct the research, including data collection, analysis, and tools used. The methodologies provide a framework to continuously track and model customer shopping patterns for future purchase behavior. This can assist clothing retailers to optimize customer value.

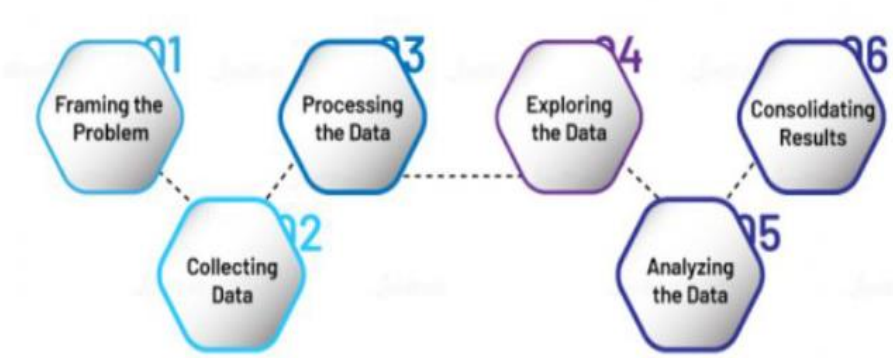

Fig.1. Flow diagram

To study customer shopping trends, the transaction data of 4,000 customers was obtained from the clothing retailer for their online and brick-and-mortar sales. The parameters included customer demographics like Age, Gender, Item Purchased, Category, Purchase Amount (USD), Location, Size, Color, Season, Review, Rating, Subscription Status, Payment Method, Shipping Type, Discount Applied, Promo Code Used, Previous Purchases, Preferred Payment Method, and Frequency of Purchases.

2.3Limitations:

- Inaccurate or incomplete data can lead to unreliable analysis and predictions.
- Changing consumer behavior due to external factors makes long-term trend prediction challenging.
- Collecting and analyzing customer data raises concerns about privacy and data security.
- Analyzing large and complex datasets requires significant computational power and resources, which may not be available to all businesses.
- The volume and velocity of data generated today can overwhelm traditional data analysis techniques.

CHAPTER 3

Proposed Methodology

3.1 System Design:

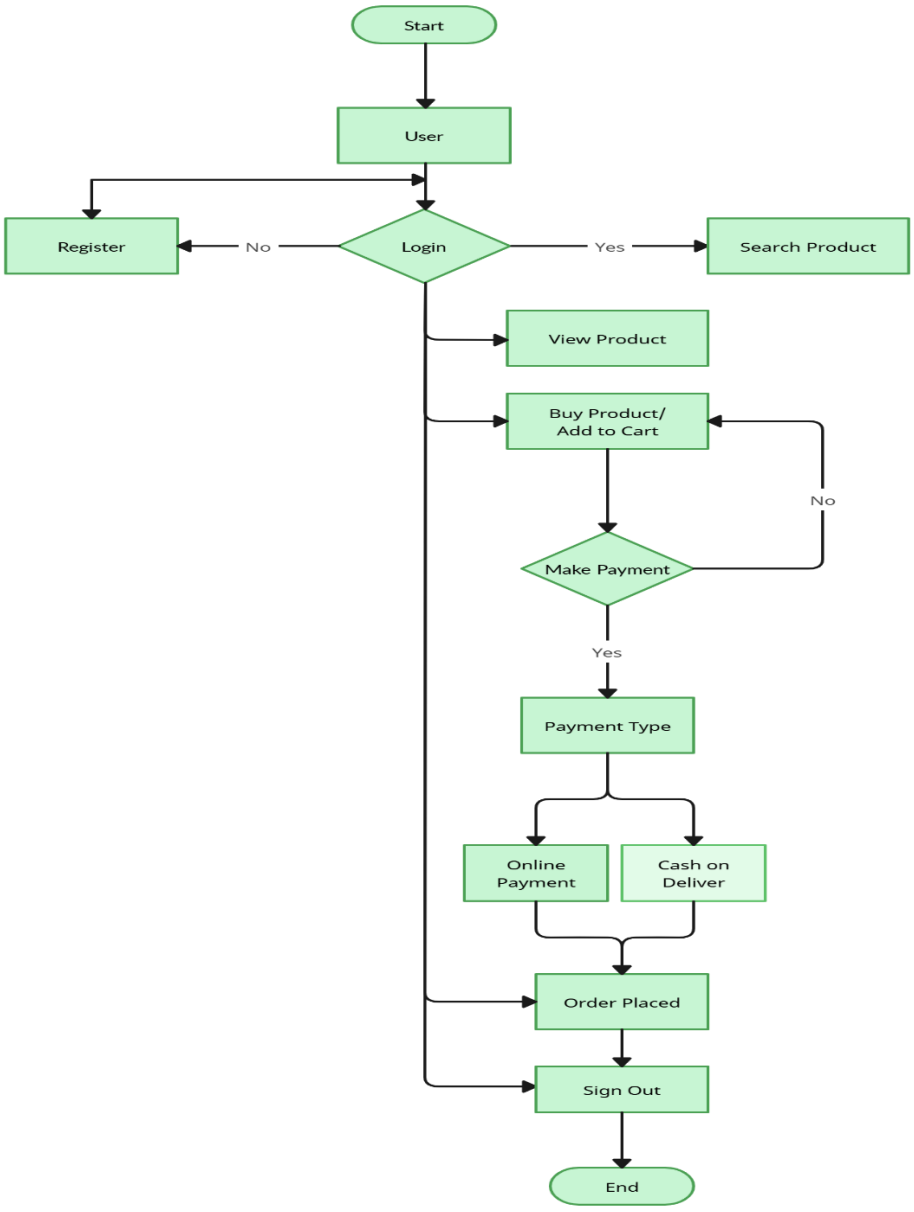

Fig.2. System design

3.2 Requirement Specification:

3.2.1 Hardware Requirements:

- **Microprocessor:** Intel Core i3
- **RAM:** 8 GB

3.2.2 Software Requirements:

- **Operating system:** Windows 64-bit operating system
- **Programming Languages:** Python, R and SQL
- **Data Analysis Libraries:** pandas, NumPy, Scikit-learn, Matplotlib/Seaborn

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:

```
[8]: data.to_excel('shopping_trends_updated.xlsx')
[9]: data.head()
```

[9]:

	Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Season	Review Rating	Subscription Status	Shipping Type	Discount Applied	Promo Code Used	Previous Purchases	Payn Met
0	1	55	Male	Blouse	Clothing	53	Kentucky	L	Gray	Winter	3.1	Yes	Express	Yes	Yes	14	Ve
1	2	19	Male	Sweater	Clothing	64	Maine	L	Maroon	Winter	3.1	Yes	Express	Yes	Yes	2	
2	3	50	Male	Jeans	Clothing	73	Massachusetts	S	Maroon	Spring	3.1	Yes	Free Shipping	Yes	Yes	23	
3	4	21	Male	Sandals	Footwear	90	Rhode Island	M	Maroon	Spring	3.5	Yes	Next Day Air	Yes	Yes	49	Pa
4	5	45	Male	Blouse	Clothing	49	Oregon	M	Turquoise	Spring	2.7	Yes	Free Shipping	Yes	Yes	31	Pa

Fig.3. head

[10] Collapse Output

[10]:

	Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Season	Review Rating	Subscription Status	Shipping Type	Discount Applied	Promo Code Used	Previous Purchases	Pa M
3895	3896	40	Female	Hoodie	Clothing	28	Virginia	L	Turquoise	Summer	4.2	No	2-Day Shipping	No	No	32	
3896	3897	52	Female	Backpack	Accessories	49	Iowa	L	White	Spring	4.5	No	Store Pickup	No	No	41	T
3897	3898	46	Female	Belt	Accessories	33	New Jersey	L	Green	Spring	2.9	No	Standard	No	No	24	
3898	3899	44	Female	Shoes	Footwear	77	Minnesota	S	Brown	Summer	3.8	No	Express	No	No	24	
3899	3900	52	Female	Handbag	Accessories	81	California	M	Beige	Spring	3.1	No	Store Pickup	No	No	33	

[ ]:

Fig.4. tail

```
[16]: data['Age_category'] = pd.cut(data['Age'], bins= [0,15, 18 , 30 , 50 , 70] , labels= ['child' , 'teen' , 'Young Adults' ,'Middle-Aged Adults'
                                                                                          , 'old'] )
[17]: fig = px.histogram(data, y = 'Age', x = 'Age_category')
      fig.show()
```

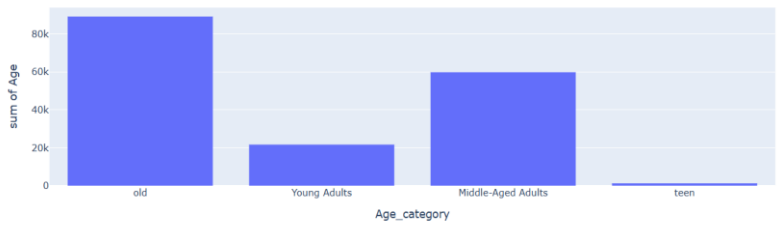

Fig.5. Age category

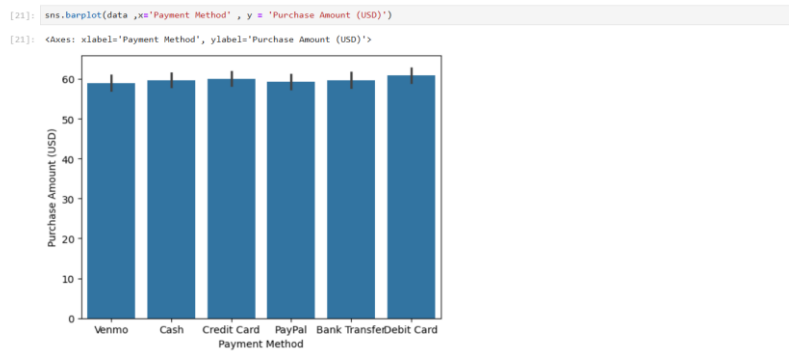

Fig.6. Payment method

Fig.7. Item Purchased

4.2 GitHub Link for Code:

https://github.com/AvimtaShahane/Shopping-Trends-Analysis

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

1. **Advanced AI/ML Techniques:**
   - Utilizing deep neural networks for more accurate predictions of individual customer behavior, including purchase likelihood, churn probability, and product preferences.
   - Leveraging generative models (like GANs or VAEs) to predict emerging trends and fashion styles, enabling proactive product development and marketing campaigns.
   - Creating hyper-personalized shopping experiences by generating unique product recommendations and customized marketing messages for each individual.

2. **Integration of Diverse Data Sources:**
   - Combining data from online, offline, and mobile channels to gain a holistic view of customer behavior across all touch points.
   - Analyzing social media data to understand consumer sentiment, identify emerging trends, and gain insights into customer preferences.
   - Integrating data from in-store sensors (like heat maps, shelf sensors) and wearable devices to track customer movement, preferences, and engagement in real-time.

3. **The Metaverse and Immersive Experiences:**
    - Analyze how users customize their avatars, which can reveal insights into their self-expression and social identity.
    - Track the purchase of virtual goods and in-game items, which can provide insights into consumer spending habits and preferences.
    - Analyze social interactions within the metaverse to understand how user communities form and influence purchasing decisions.

5.2 Conclusion:

In conclusion, by knowing the customer's shopping behaviors, businesses can significantly enhance their marketing strategies, also optimize inventory management, and offer personalized customer experiences. The findings of this research provide valuable implications for retailers, enabling them to stay ahead in the competitive market by aligning their products and services with the ever-changing demands and expectations of consumers.

REFERENCES

[1]. Gupta, Varun & Khanna, Vaibhav & Sahoo, Biswa. (2018). Analysis of shopping trends employing E-Commerce Applications: A Comparative Case Study. Procedia Computer Science. 132. 1728-1738. 10.1016/j.procs.2018.05.148.

[2]. Riwayat, Andik & Susilawati, Agnes & Naqiah, Zakiyyatun. (2024). Purchasing Patterns Analysis in E-commerce: A Big Data-driven Approach and Methodological. International Journal Software Engineering and Computer Science (IJSECS). 4. 148-164. 10.35870/ijsecs.v4i1.2384.

[3]. Business Research Insights (2023). Online Shopping Market Analysis: Trends, Growth, and Future Projections 2023 to 2028.

[4]. P. A. Todd, J. D. McKeen, and R. B. Gallupe, “The evolution of IS job skills: A content analysis of IS job advertisements from 1970 to 1990,” MIS Quarterly, vol. 19, no. 1, pp. 1–27, 1995.