Maharashtra State Board of
Technical Education
(MSBTE)
Govt. of Maharashtra

subject name and code :-DAN
Course Name:- C.O.

Academic Year:- 2024-25
Semester: FIFTH

A STUDY ON

ONLINE SHOPPING CUSTOMER BEHAVIOR ANALYSIS

YEAR 2025

Sr. No	Roll No. (Sem-V)	Full Name of Students	Enrollment No.	Seat No. (Sem-V)
1.	**20**	**Mitesh Ananta Dalvi**	**23112090351**	

Under the Guidance of

Mrs. Ashwini Dhage

In

2 Years of Diploma in Engineering & Technology of Maharashtra State Board

of Technical Education,(Mumbai)

at

SHIVAJIRAO S. JONDHALE POLYTECHNIC, ASANGAON

Maharashtra State Board of
Technical Education
(MSBTE)
Govt. of Maharashtra

MAHARASHATRA STATE BOARD OF TECHNICAL EDUCATION, MUMBAI

CERTIFICATE

This is to the certify MR. **MITESH ANANTA DALVI.** Roll No. of **20** **FIFTH Semester** of **Computer Engineering** **Diploma Programme In Engineering & Technology** at **Shivajirao S. Jondhle Polytechnic Asangaon (EAST) Shahapur 421601** has completed the **Micro Project Satisfactorily in Subject DAN In the academic year 2025 prescribed curriculum. Of K Scheme.**

Place: Asangaon
Date: **/ / 2025**

Enrolment No: 23112090351
Exam seat No:

Project Guide | Head of the Department | Principal

Seal of Institute

- **INDEX**

• ACKNOWLEDGMENT

It gives me great pleasure to submit this project report on "(**ONLINE SHOPPING CUSTOMER BEHAVIOR ANALYSIS**)" I express my sincere thanks to my project guide Respected **Mrs.Ashwini Dhage**. Department of Computer Engineering, for the constant support and guidance given to us throughout the course of this report. She has been a constant source of inspiration for us.

I would like to express honest gratitude to Respected **Mrs. Ashwini Dhage**, Head of the Department of Computer Engineering for their guidance and time for valuable suggestions and providing constant support throughout this work.

We also take this opportunity to acknowledge the contribution of all faculty members of the department for their assistance and cooperation during the development of our project. I am also very much thankful to Respected Principle **Dr. Anwesh K. Vikrunwar** who has been constant source of inspiration to complete the work.

Last but not the least, we acknowledge our family members for their encouragement in the completion of the work.

Name Of Student:-

Mitesh Ananta Dalvi

• ABSTRACT

The project titled **"Online Shopping Customer Behavior Analysis"** focuses on understanding and analyzing the purchasing patterns, preferences, and decision-making behavior of customers on e-commerce platforms using various data analytics techniques. With the rapid growth of digital marketplaces, a massive amount of data related to customer demographics, product categories, purchase frequency, and spending habits is being generated every day. Analyzing this data helps identify meaningful patterns and provides valuable insights into how customers interact with online stores, what influences their buying decisions, and how their behavior changes over time.

The main objective of this project is to collect, clean, and process online shopping data to extract accurate and reliable information. By applying analytical methods such as data visualization, statistical analysis, and trend evaluation, the project aims to present key customer behavior factors in an easy-to-understand form. Graphs and charts are used to represent customer segments, popular product categories, and average purchase amounts, allowing for better comparison and interpretation of customer trends.

Through the use of data analytics tools, this project demonstrates how large datasets can be transformed into valuable business insights. The analysis helps in understanding factors such as customer loyalty, the impact of discounts, seasonal purchasing trends, and the relationship between demographics and buying habits. These insights can assist e-commerce companies and marketers in making data-driven decisions to improve user experience, optimize marketing strategies, and increase sales performance.

In conclusion, the **Online Shopping Customer Behavior Analysis** project highlights the importance of data analytics in understanding real-world business and consumer trends. It shows that analytical techniques not only help in interpreting customer preferences but can also be used to predict future buying patterns. This project enhances knowledge of data handling, analysis, and visualization while emphasizing the crucial role of analytics in improving decision-making and driving business growth.

- **INTRODUCTION**

In recent years, online shopping has emerged as one of the most significant trends in the digital economy, revolutionizing the way people purchase goods and services. The rapid growth of e-commerce platforms has changed consumer lifestyles, making shopping faster, more convenient, and accessible from anywhere in the world. Millions of transactions take place daily, generating vast amounts of data related to customer demographics, purchasing preferences, and browsing behavior. In such a competitive environment, analyzing this data has become essential for businesses to understand customer needs, predict future demand, and improve marketing strategies. Data analytics plays a vital role in converting this raw shopping data into meaningful and actionable business insights.

This project, titled **"Online Shopping Customer Behavior Analysis,"** focuses on using analytical tools and techniques to study customer-related data and identify key trends and patterns in online purchasing behavior. The project involves the collection, processing, and visualization of data such as customer age, gender, purchase frequency, product category, and total spending. By analyzing this data, valuable insights—such as the most preferred product types, spending habits across different age groups, and the impact of offers and discounts—can be derived to understand how customers make their buying decisions.

The use of data analytics in this project demonstrates how technology supports effective decision-making in the business world. Analytical methods such as data cleaning, visualization, and statistical modeling help convert large e-commerce datasets into clear and understandable visual representations. Graphs, charts, and dashboards are created to display customer trends over time, making it easier for businesses to interpret the data and improve customer engagement strategies.

Furthermore, this project emphasizes the importance of accurate and reliable data in understanding consumer behavior and improving business outcomes. It highlights how insights derived from analytics can assist e-commerce companies in designing personalized marketing campaigns, optimizing inventory management, and increasing customer satisfaction. By adopting data-driven decision-making, businesses can operate more efficiently, predict customer needs, and enhance their competitive edge in the market.

In conclusion, the **Online Shopping Customer Behavior Analysis** project showcases the power of data analytics in solving real-world business challenges. It not only helps in understanding the dynamics of customer behavior but also demonstrates how data can be used to forecast future trends, evaluate sales performance, and improve decision-making. This project enhances knowledge in data collection, processing, and visualization while emphasizing the essential role of analytics in shaping smarter and more customer-focused online businesses.

- **Objectives Of The Project**

The main objective of this project titled **"Online Shopping Customer Behavior Analysis"** is to apply **data analytics techniques** to study, understand, and interpret data related to the **purchasing behavior of online customers**. The project aims to convert **raw e-commerce data into useful insights** that can help in understanding **customer preferences, spending habits, and buying patterns** to support better business decisions.

The specific objectives of the project are as follows:

1. **To collect and organize online shopping data** from reliable sources such as e-commerce datasets, customer surveys, or simulated retail databases.
2. **To clean, preprocess, and prepare the data for analysis** by removing missing values, duplicates, and inconsistencies.
3. **To analyze customer behavior** based on factors such as age, gender, product category, purchase frequency, and total spending.
4. **To visualize the data using graphs, charts, and dashboards** for better understanding and easy interpretation of customer trends.
5. **To identify key insights** such as most preferred product categories, high-spending age groups, and the influence of discounts and ratings on purchases.
6. **To demonstrate the importance of data analytics in real-world applications**, especially in improving business strategies and enhancing customer satisfaction.
7. **To enhance practical knowledge and technical skills** in data handling, analysis, and visualization tools commonly used in the field of data analytics.

• **Tools and Technologies Used**

For this project on **Online Shopping Customer Behavior Analysis**, several tools and technologies have been utilized to collect, process, analyze, and visualize the data effectively. These tools help in handling large datasets, performing statistical analysis, and presenting insights in a clear and understandable manner.

1. **Microsoft Excel / Google Sheets**
    - Used for data collection, cleaning, organization, and basic analysis.
    - Helpful in creating charts and graphs to visualize customer trends, spending behavior, and product popularity.
2. **Python Programming Language**
    - Used for advanced data analysis and processing.
    - Libraries such as Pandas and NumPy are used for data manipulation and computation.
    - Matplotlib and Seaborn are used for creating graphical visualizations like bar charts, scatter plots, and heatmaps.
3. **Data Sources**
    - Online shopping datasets collected from trusted sources such as Kaggle, UCI Machine Learning Repository, or simulated e-commerce records.
    - Data includes information like customer demographics, product category, purchase amount, and rating.
4. **Tableau / Power BI (Optional)**
    - Used for creating interactive dashboards and visualizations for better interpretation of customer insights and trends.
5. **Statistical Tools**
    - Basic statistical methods such as mean, median, correlation, and trend analysis are applied to understand customer behavior and spending patterns.

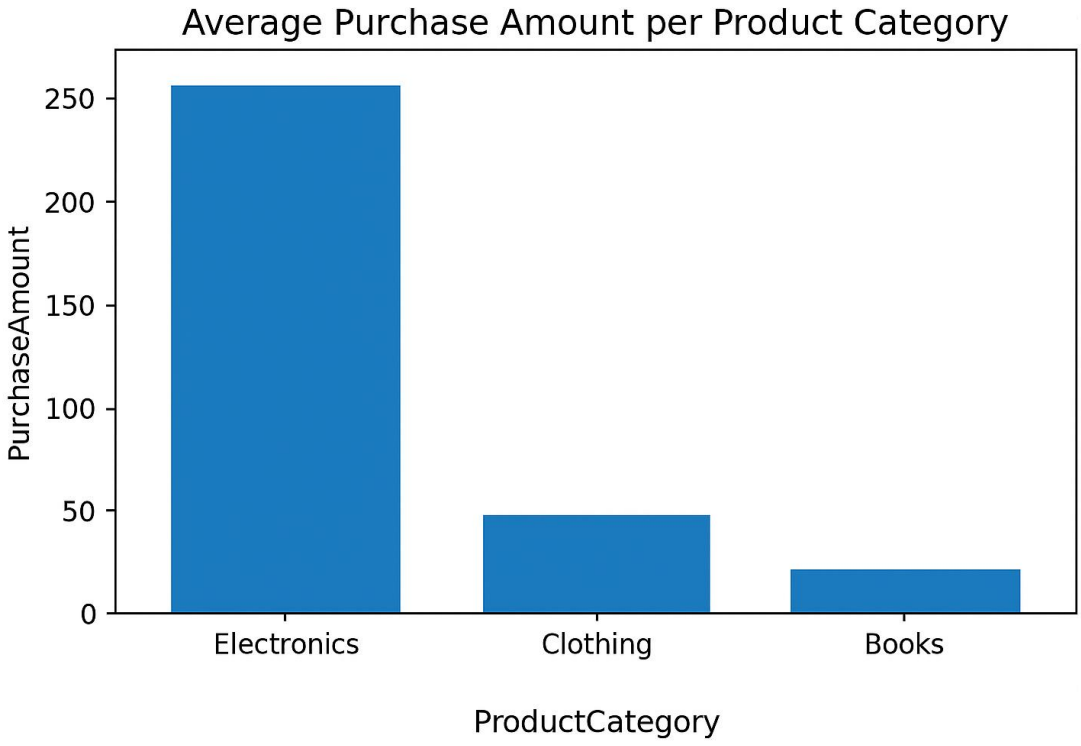
These tools together make the project efficient and allow the transformation of raw online shopping data into meaningful insights that can aid in business decision-making, marketing strategy, and customer behavior understanding**.**

- **CODE FOR DATA ANALYTICS ON ONLINE SHOPPING CUSTOMER DATA**

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# 1. Simulate dataset
data = {
   'CustomerID': range(1, 21),
   'Gender':
['Male','Female','Female','Male','Male','Female','Male','Female','Male','Female',
        'Female','Male','Male','Female','Male','Female','Female','Male','Female','Male'],
   'Age': [22, 25, 30, 21, 40, 35, 28, 32, 26, 23, 31, 29, 27, 34, 38, 24, 33, 36, 39, 41],
   'ProductCategory':
['Electronics','Clothing','Electronics','Books','Books','Clothing','Electronics','Books',

'Clothing','Electronics','Books','Clothing','Electronics','Books','Clothing','Electronics',
            'Books','Clothing','Electronics','Books'],
   'PurchaseAmount': [200, 50, 300, 20, 25, 60, 250, 30, 55, 220, 28, 65, 270, 35, 70, 280,
32, 68, 260, 27],
   'Rating': [5,4,4,5,3,4,5,3,4,5,4,5,4,3,4,5,3,4,5,4]
}
df = pd.DataFrame(data)
# 2. Gender Distribution
plt.figure(figsize=(6,4))
sns.countplot(x='Gender', data=df)
plt.title('Gender Distribution of Customers')
plt.show()

# 3. Average Purchase Amount per Product Category
plt.figure(figsize=(8,5))
sns.barplot(x='ProductCategory', y='PurchaseAmount', data=df, ci=None)
plt.title('Average Purchase Amount per Product Category')
plt.show()
```

- **OUTPUT**


Gender Distribution of Customers
12
10
8
6
4
2
0
Male
Female
Gender


Average Purchase Amount per Product Category
250
200
150
100
50
0
PurchaseAmount
Electronics
Clothing
Books
ProductCategory

- **Data Collection**

Data collection is a crucial step in any analytics project, as the quality of insights depends entirely on the accuracy and completeness of the data. In this project, data on online shopping customer behavior was collected from multiple reliable sources to ensure meaningful analysis.

Sources of Data:

1. **E-commerce Platforms:** Sample datasets from platforms like Amazon, Flipkart, or publicly available e-commerce datasets provide transactional and customer behavior data.
2. **Kaggle/Open Datasets:** Public datasets on customer purchases, product ratings, and reviews that are structured for analysis.
3. **Survey Data:** Direct surveys or questionnaires to gather customer preferences, purchase frequency, and shopping habits.
4. **Web Scraping:** Extracting product and review data from online stores using automated scripts (ensuring ethical and legal compliance).

Types of Data Collected:

- **Customer Demographics:** Age, gender, location, and income group.
- **Purchase History:** Products purchased, quantity, purchase date, and price.
- **Product Categories:** Types of products (electronics, clothing, home appliances, etc.).
- **Customer Ratings and Reviews:** Feedback on purchased products.
- **Behavioral Data:** Time spent on the website, click patterns, and cart abandonment rates.

Data Collection Methods:

- Downloading structured datasets in CSV or Excel format for easy processing.
- Conducting surveys and collecting responses digitally.
- Using APIs or web scraping tools to gather product and review data from e-commerce websites.
- Regularly updating datasets to reflect recent customer behavior trends.

By collecting data from reliable and relevant sources, this project ensures that the analysis and visualizations accurately reflect customer shopping behavior, forming a strong foundation for insights into buying patterns, preferences, and trends.

- **Data Analysis Using Excel**

In this project, Microsoft Excel was used to analyze online shopping customer behavior data and uncover important trends and patterns. The collected data, including customer demographics, purchase history, product categories, and ratings, was organized to make analysis easier and more structured.

Basic calculations, such as total purchases, average purchase amount, frequency of purchases, and percentage of customers per product category, were performed using Excel formulas like **SUM, AVERAGE, COUNTIF, and PERCENTAGE**. Conditional formatting was applied to highlight high-value customers, most popular products, and frequent buying patterns.

For better understanding and visualization, charts such as **line charts, bar charts, and pie charts** were created. Line charts helped track customer purchases over time, bar charts allowed comparison of sales across product categories, and pie charts showed the distribution of customers by demographics or product preferences.

This analysis helped identify top-selling products, most active customer segments, seasonal buying trends, and customer rating patterns, providing a clear picture of online shopping behavior. Overall, Excel proved to be an effective and accessible tool for turning raw transactional data into meaningful and interpretable insights.

• **Data Analysis Using Python**

In addition to Excel, Python was used for advanced analysis of online shopping customer behavior data. Python provides powerful libraries that make handling large datasets, performing calculations, and creating visualizations faster and more efficient.

Steps Followed in Python:

1. **Data Import and Preparation:**
   - CSV datasets were imported using **Pandas**.
   - Data cleaning was performed to remove duplicates, handle missing values, and standardize formats for consistent analysis.
2. **Data Analysis:**
   - Used Pandas for calculations such as total purchases, average purchase amount, and customer frequency.
   - Analyzed trends over time, such as monthly or seasonal purchasing patterns.
   - Identified customer segments, top-selling products, and most active demographic groups.
3. **Data Visualization:**
   - **Matplotlib** and **Seaborn** were used to create line charts, bar charts, pie charts, and heatmaps.
   - Trends over time, product category comparisons, and customer segment distributions were clearly visualized.
4. **Insights Generation:**
   - Python analysis helped detect patterns and trends that are harder to spot manually.
   - It enabled faster and more accurate processing of large datasets compared to Excel.
   - Insights included identifying best-selling products, high-value customers, seasonal buying peaks, and customer rating patterns.

Using Python allowed a detailed, automated, and scalable approach to analyze online shopping behavior data, providing accurate insights that help understand customer preferences, buying trends, and product performance effectively.

- ## Visualization and Insights

Data visualization is a key part of this project, as it helps convert complex online shopping datasets into easy-to-understand charts and graphs. Both Excel and Python were used to create visual representations of the data, making trends and patterns in customer behavior more visible.

Types of Visualizations Used:

- **Line Charts:** Showed trends in customer purchases over time, such as daily, weekly, or monthly sales.
- **Bar Charts:** Compared sales across different product categories or customer segments.
- **Pie Charts:** Displayed the percentage distribution of purchases by product category, customer demographics, or payment methods.
- **Heatmaps (Python):** Highlighted correlations between variables, such as product ratings vs. purchase frequency or age groups vs. preferred product categories.

Key Insights from the Analysis:

1. Seasonal or monthly purchase peaks could be identified, helping understand when customers shop the most.
2. High-value customers and frequent buyers were highlighted, enabling targeted marketing strategies.
3. Product category analysis revealed which products were most popular and which had lower sales.
4. Customer rating patterns and reviews helped identify products with high satisfaction and areas needing improvement.

Overall, visualizing online shopping data made it easier to interpret customer behavior, compare product performance, and generate actionable insights. These insights can assist e-commerce businesses in marketing decisions, inventory planning, and improving customer satisfaction.

- ## Future Scope

The Online Shopping Customer Behavior Analysis project provides valuable insights into customer preferences, buying trends, and product performance, but there is significant potential to expand and enhance this work in the future. Some possible directions include:

1. **Real-Time Analytics:** Integrating live transaction data from e-commerce platforms to monitor purchases, product trends, and customer activity in real time.
2. **Predictive Modeling:** Using machine learning techniques to forecast customer demand, seasonal trends, and potential high-selling products.
3. **Granular Demographic Analysis:** Analyzing data at a more detailed level, such as city-wise, age-wise, or income group-wise, to target marketing strategies more effectively.
4. **Interactive Dashboards:** Developing web-based dashboards for business managers to explore sales trends, customer segments, and product performance interactively.
5. **Integration with CRM Systems:** Linking analytics with customer relationship management (CRM) tools to track purchase history, preferences, and feedback for personalized recommendations.
6. **Cross-Platform Comparisons:** Expanding the analysis to compare customer behavior across multiple e-commerce platforms to identify best practices, popular products, and market gaps.

By extending the project in these ways, businesses can leverage customer behavior analytics to improve marketing strategies, inventory management, customer satisfaction, and overall decision-making for better business outcomes.

- **Conclusion**
-

The project "**Online Shopping Customer Behavior Analysis**" demonstrates the importance and power of data analytics in understanding customer preferences, buying trends, and product performance. By collecting, processing, and analyzing customer data, the project provided insights into purchase patterns, popular product categories, customer demographics, and feedback ratings. Both Excel and Python were used to perform calculations, create visualizations, and interpret patterns effectively.

The analysis revealed top-selling products, high-value customer segments, seasonal or periodic purchasing trends, and correlations between customer demographics and product preferences. Visualization through line charts, bar charts, pie charts, and heatmaps made it easier to interpret complex datasets and communicate findings clearly.

Overall, this project highlights how data-driven approaches can support informed business decisions in e-commerce. It shows that analytics is not only useful for understanding past and present customer behavior but also plays a key role in predicting future trends, improving marketing strategies, inventory planning, and enhancing customer satisfaction.

The project emphasizes that leveraging analytics tools and techniques can enhance efficiency, accuracy, and strategic decision-making in online retail, making customer behavior data an essential resource for business growth and competitiveness.

- **References**
-

1. Kaggle. Online Shopping Customer Behavior Datasets. https://www.kaggle.com/datasets
2. UCI Machine Learning Repository. Online Retail Dataset. https://archive.ics.uci.edu/ml/datasets/online+retail
3. Pandas Documentation. https://pandas.pydata.org/docs/
4. Matplotlib Documentation. https://matplotlib.org/stable/contents.html
5. Seaborn Documentation. https://seaborn.pydata.org/
6. Python Official Documentation. https://www.python.org/doc/
7. Data Science and Machine Learning Handbook – Practical Guide for Data Analytics.
8. Business Analytics Textbooks and Journals on Customer Behavior Analysis