Walmart Customer Purchase Behaviours Analysis on Black Friday

About Walmart

Walmart is an American multinational retail corporation operating a chain of supercentres, discount departmental stores, and grocery stores globally. With over 100 million customers worldwide, Walmart seeks to understand purchase behaviour to optimize its offerings.

Business Problem

Walmart aims to analyse customer purchase behaviour, specifically the purchase amount, in relation to gender and various other factors. This analysis will help them understand spending habits and inform business decisions. Key questions include:

- Do female customers spend more than male customers on Black Friday?
- How do spending habits differ based on demographics and marital status?

Dataset

The dataset, `Walmart_data.csv`, contains transactional data of customers who purchased products at Walmart stores during Black Friday. It includes the following features:

- User_ID: Unique identifier for each user
- Product_ID: Unique identifier for each product
- Gender: Sex of the user (e.g., Male, Female)
- Age: Age group (e.g., 18-25, 26-35)
- Occupation: Masked occupation category
- City_Category: Category of the city (e.g., A, B, C)
- StayInCurrentCityYears: Number of years living in the current city
- Marital_Status: Marital status (e.g., Married, Single)
- ProductCategory: Masked product category
- Purchase: Purchase amount

Steps

1. **Data Import and Exploration:**
    - Import the `Walmart_data.csv` dataset.
    - Analyze data structure and characteristics using methods like `head()`, `info()`, and data visualization techniques.
2. **Data Cleaning:**

- Detect and handle missing values using techniques like imputation or deletion based on business context.
- Identify and address outliers using methods like boxplots or IQR analysis, considering their impact on the analysis.

3. **Data Analysis:**
    - **Gender-Based Spending:**
        - Calculate average purchase amount for male and female customers separately.
        - Compare average spending between genders and draw initial inferences.
        - Implement hypothesis testing (e.g., t-test) to determine statistically significant differences in spending habits between genders.
    - **Confidence Intervals:**
        - Construct confidence intervals (CIs) for the average purchase amount of both genders.
        - Experiment with different confidence levels (e.g., 90%, 95%, 99%) and observe the interval width.
        - Analyze overlapping or non-overlapping confidence intervals to understand if gender significantly impacts average spending.
    - **Marital Status and Age:**
        - Repeat steps similar to gender analysis for married vs. unmarried customers and different age groups.
        - Analyze purchase behaviour variations across these demographics.
4. **Recommendations and Action Items:**
    - Based on the findings, provide actionable insights and recommendations for Walmart, such as:
        - Targeted marketing campaigns based on gender, marital status, and age.
        - Product assortment optimization to cater to specific customer segments.

    - Personalized promotions and discounts based on purchase behaviour.
    - Store layout and product placement strategies based on customer demographics.

Expected Outcomes:

- Identify demographic groups with distinct spending habits.
- Understand if gender, marital status, and age significantly influence purchase behaviour.
- Provide data-driven recommendations to optimize marketing strategies, product offerings, and in-store experiences for improved customer engagement and sales.