**Problem Statement**

Walmart aims to improve its customer targeting and operational efficiency by understanding spending behaviour across different demographics. The objectives include:

1. Analysing patterns in customer spending across genders, marital statuses, and age groups.
2. Detecting outliers in spending data and handling them to ensure robust analysis.
3. Identifying demographic preferences for product categories to refine inventory and marketing strategies.
4. Evaluating the impact of gender, marital status, and age on overall spending and confidence intervals to guide strategic decision-making.

**Initial Analysis**

* **Data Types of Columns**:  
  The dataset contains 10 columns with the following data types:
  + User\_ID, Occupation, Marital\_Status, Product\_Category, and Purchase are numerical (int64).
  + Product\_ID, Gender, Age, City\_Category, and Stay\_In\_Current\_City\_Years are categorical (object).
* **Dataset Shape**:  
  The dataset consists of **550,068 rows** and **10 columns**.
* **Missing Values**:  
  No missing values were identified in any column, making the dataset complete and ready for further analysis.

**Detect Null values and outliers**

* **Outliers Detection:**
  + Outliers were observed in the Purchase column, as shown in the boxplot. These outliers indicate abnormally high or low spending values.
* **Outlier Handling:**
  + To manage outliers, data was clipped to the 5th and 95th percentiles using np.clip(). This approach ensures extreme values do not skew further analyses.

**Data Exploration**

* **Products Purchased by Age Groups:**

Different age groups exhibit varying preferences. Younger groups favor gadgets and electronics, while older groups purchase household and utility items.

* **Age, Marital Status, and Spending:**

Married individuals generally spend more, especially in the middle age groups (35-50). Younger and single individuals spend less on average.

* **Preferred Product Categories by Gender:**

Males prefer high-value items like electronics, while females lean towards household and personal care products.

**gender affect on the amount spent**

**Confidence Intervals for Genders:**

Â·         Males and females have overlapping confidence intervals for average spending, indicating no significant gender-based difference in spending habits.

Â·         Smaller sample sizes result in wider intervals, but larger samples stabilize the confidence intervals.

**Marital Status Effect on Spending**

Â·         Married individuals tend to spend more compared to unmarried individuals.

Â·         Confidence intervals become narrower with larger sample sizes, reducing variability in spending estimates.

**Age affect the amount spent**

Â·         Middle-aged individuals (35-50) exhibit higher spending compared to younger groups.

Â·         Confidence intervals across age groups overlap slightly, but spending preferences differ significantly.

**Reporting**

**a. Gender-Based Spending Confidence Intervals**

Â·         The confidence intervals for males and females overlap significantly, suggesting no substantial difference in spending between genders.

Â·         **Walmart's Actionable Insight**: Focus on gender-neutral promotions and campaigns, as spending behavior does not vary significantly by gender.

**b. Marital Status-Based Spending Confidence Intervals**

Â·         The confidence intervals for married and unmarried individuals show slight overlap, but married individuals tend to spend more.

Â·         **Walmart's Actionable Insight**: Introduce family-oriented promotions and discounts to attract married customers, who spend more on average.

**c. Age Group-Based Spending Confidence Intervals**

Â·         Confidence intervals for different age groups overlap slightly, but clear differences in average spending exist. Middle-aged customers spend the most.

Â·         **Walmart's Actionable Insight**: Design age-specific marketing campaigns, such as:

o    Promotions on electronics and gadgets for younger buyers.

o    Household and utility product deals for middle-aged and older customers.

**Recommendations**

1.      **Target Marketing**:

o    Focus on family-oriented promotions to attract married customers.

o    Design age-specific campaigns (e.g., electronics for younger buyers, household products for older customers).

2.      **Gender Neutrality**:

o    Since spending does not differ significantly by gender, maintain universal promotions rather than gender-specific campaigns.

3.      **Data-Driven Pricing**:

o    Adjust pricing strategies to account for high-spending outliers without alienating budget-conscious customers.

4.      **Category Insights**:

o    Promote high-demand products based on demographic preferences (e.g., gadgets for males, personal care for females).

5.      **Future Strategies**:

o    Regularly analyze data trends to refine inventory and marketing strategies.