**Business Case: Walmart - Confidence Interval and CLT Analysis**

**Company Overview:** Walmart Inc. is a leading American multinational retailer, operating an extensive network of supercenters, discount department stores, and grocery stores worldwide, serving over 100 million customers.

**Business Objective:** The management team at Walmart seeks to analyze customer purchase behavior, particularly the amount spent, based on gender and other relevant factors. The primary question is whether spending habits differ between male and female customers, specifically if women spend more than men during Black Friday. Insights from this analysis are intended to inform better business decision-making.

**Data Overview:** The dataset contains 550,068 records with 10 columns, of which 9 are categorical and 1 is numerical. There are no missing values; however, some outliers are present in the data. A summary of unique values for key attributes is as follows:

* **Product ID**: 3,631 unique values
* **Gender**: 2 unique values (Male and Female)
* **Occupation**: 21 unique values
* **City Category**: 3 unique values (A, B, C)
* **Stay Duration in Current City**: 5 unique values
* **Marital Status**: 2 unique values
* **Product Category**: 20 unique values

**Gender Distribution:**

* The dataset is divided into two gender groups: male and female.
* There are 414,259 male customers (75.3%) and 135,809 female customers (24.7%).

**Age Distribution:**

* 39.9% of customers are aged 26–35.
* 20.0% fall within the 36–45 range.
* 18.1% are aged 18–25.
* 8.3% are aged 46–50.
* 7.0% are aged 51–55.
* 3.9% are aged 55 and above.
* 2.7% are aged 0–17.

**City Category Distribution:**

* 42% of customers reside in city category B.
* 31% are from city category C.
* 26% are from city category A.

**Marital Status:**

1. 60% of customers are single.
2. 40% of customers are married.

**Product Category:**

1. A total of 20 product categories are available.
2. Product categories 5, 1, and 8 see the highest purchase volumes.

**User Demographics:**

1. Most users are male.
2. There are 20 unique occupations and product categories in the dataset.
3. The majority of users are from City Category B.
4. More users are single than married.
5. Product categories 1, 5, 8, and 11 have the highest purchase frequency.

**Average Spending:**

1. Average amount spent by each female customer: 712024.39
2. Average amount spent by each male customer: 925344.40
3. Male customers spend more on average than female customers.

**Confidence Intervals 📈**

**Gender-Wise:** Using the Central Limit Theorem:

1. The average amount spent by male customers is 925344.40.
2. The average amount spent by female customers is 712024.39

For a 99% Confidence Interval:

* We can estimate that, 99% of the time:
  + The average amount spent by male customers will be between 924631.85 to 925553.15
  + The average amount spent by female customers will be between 711532.08 to 712619.12

**Marital Status-Wise:** Using the Central Limit Theorem:

1. The average amount spent by married customers is 832644.12
2. The average amount spent by unmarried customers is 870448.39.

For a 99% Confidence Interval:

* While the confidence intervals for married and unmarried customers overlap, the overlap is noticeably reduced.
  + The average amount spent by married customers will lie between 784,352.68 and 880,373.15.
  + The average amount spent by unmarried customers will lie between 829,023.32 and 911,882.57.

**Age-Wise (99% Confidence Interval):**

1. Age 26-35: (973,422.17, 983,945.96)
2. Age 36-45: (865,476.76, 875,591.71)
3. Age 18-25: (839,048.97, 848,180.50)
4. Age 46-50: (774,367.04, 784,366.19)
5. Age 51-55: (743,802.44, 751,872.59)
6. Age 55+: (524,263.09, 530,526.91)
7. Age 0-17: (611,101.04, 618,088.49)

**Recommendations**

1. **Gender-Focused Strategy** 
   * **Observation:** Men tend to spend more than women.
   * **Recommendation:** Focus on retaining existing male customers and attracting new male customers through targeted marketing campaigns and offers.
2. **Product Category Optimization** 
   * **Observation:** Product categories 1, 5, 8, and 11 have the highest purchase frequency and are popular with customers.
   * **Recommendation:** Increase promotion and stock availability of these high-demand products. Additionally, explore ways to stimulate interest in less-purchased product categories through discounts or bundled offers.
3. **Marital Status Approach** 
   * **Observation:** Unmarried customers have higher average spending compared to married customers.
   * **Recommendation:** Design campaigns specifically targeting unmarried customers with tailored offers, loyalty programs, or special events to drive engagement and spending.
4. **Targeting Age Demographics** 
   * **Observation:** Customers aged 18-45 contribute a significant share of total spending.
   * **Recommendation:** Focus marketing and customer acquisition efforts on this age group, possibly with products or promotions that align with their interests and spending patterns.
5. **City Category Strategy** 
   * **Observation:** Male customers in City\_Category C demonstrate higher spending than those in City\_Categories B or A.
   * **Recommendation:** Increase the product offerings and advertising budget in City\_Category C to capture this higher-spending segment more effectively.
6. **Boosting Female Engagement** 
   * **Observation:** While male spending is higher, female customers still represent a substantial market share.
   * **Recommendation:** Develop personalized offers and campaigns that cater to female shopping preferences, especially around high-traffic shopping events.
7. **Enhance Digital Experience Across Key Age Groups** 
   * **Observation:** With significant spending coming from younger demographics, especially 18-45.
   * **Recommendation:** Ensure a seamless digital shopping experience and targeted online promotions, as younger customers may respond better to digital engagement.
8. **Marital Status-Based Promotions** 
   * **Observation:** Unmarried customers tend to spend more than married customers.
   * **Recommendation:** Create promotional offers that appeal to singles, such as social event partnerships or experience-based rewards, which could increase brand affinity and sales.
9. **Personalization for High-Purchase Categories** 
   * **Observation:** High-demand product categories (1, 5, 8, 11) are preferred.
   * **Recommendation:** Implement personalized recommendations for these categories on the website and app, encouraging repeat purchases with tailored suggestions based on previous buying behavior.