A/B Testing on Retail Dataset

A. Dataset Information

Dataset Information

The "Online Retail Dataset" contains transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based online retail store. The company mainly sells unique all-occasion gifts. Many customers of the company are wholesalers.

Column Information

The dataset consists of the following columns:

1. **InvoiceNo**
   - **Description:** A unique identifier for each invoice (transaction).
   - **Type:** String (Alphanumeric).
   - **Example:** 536365, C536379 (Invoices starting with 'C' indicate a canceled transaction).
2. **StockCode**
   - **Description:** A unique identifier for each product.
   - **Type:** String.
   - **Example:** 85123A, 71053.
3. **Description**
   - **Description:** Description of the product.
   - **Type:** String.
   - **Example:** WHITE HANGING HEART T-LIGHT HOLDER, REGENCY CAKESTAND 3 TIER.
4. **Quantity**
   - **Description:** The number of units of the product purchased.
   - **Type:** Integer.
   - **Example:** 6, 32.
5. **InvoiceDate**
   - **Description:** The date and time when the transaction was generated.
   - **Type:** DateTime.
   - **Example:** 12/1/2010 8:26, 12/9/2011 14:33.
6. **UnitPrice**
   - **Description:** The price per unit of the product.
   - **Type:** Float.
   - **Example:** 2.55, 3.39.
7. **CustomerID**
   - **Description:** A unique identifier for each customer.

- ○ **Type:** Integer (often has missing values).
- ○ **Example:** 17850, 13047.
8. **Country**
   - ○ **Description:** The country where the customer resides.
   - ○ **Type:** String.
   - ○ **Example:** United Kingdom, France.

Summary

- **InvoiceNo:** Transaction ID.
- **StockCode:** Product ID.
- **Description:** Product description.
- **Quantity:** Number of units sold.
- **InvoiceDate:** Date and time of the transaction.
- **UnitPrice:** Price per unit.
- **CustomerID:** Customer ID.
- **Country:** Country of the customer.

Understand and preprocess the data:

- ○ Handle missing values in CustomerID.
- ○ Filter for relevant transactions (e.g., focus on UK customers).
- ○ Convert InvoiceDate to datetime format.
- ○ Add calculated columns (e.g., TotalAmount = Quantity * UnitPrice).

By using this dataset and following proper preprocessing steps, you can conduct meaningful analyses and A/B tests to derive valuable insights.

B. Methodology:

1. Import Libraries and Dataset

Objective: Import the necessary Python libraries and the dataset.

Explanation:

- **Libraries:** Python libraries such as pandas, numpy, matplotlib, seaborn, and scipy provide essential tools for data manipulation, visualization, and statistical analysis.
- **Dataset:** The dataset used here is the "Online Retail Dataset" from the UCI Machine Learning Repository, which contains transactional data from a UK-based online retail store. Loading the dataset into a pandas DataFrame allows for easy data manipulation and analysis.

2. Data Preprocessing

Objective: Clean and prepare the data for analysis.

Explanation:

- **Handling Missing Values:** Dropping rows with missing CustomerID ensures that each transaction is linked to a specific customer, which is crucial for customer-level analysis.
- **Filtering Transactions:** Focusing on transactions from UK customers ensures a more homogenous dataset, which can help in reducing variability due to different customer behaviors across countries.
- **Date Conversion:** Converting InvoiceDate to a datetime format allows for easier time-based analysis.
- **New Column for Total Amount:** Adding a TotalAmount column (quantity multiplied by unit price) helps in analyzing the monetary value of transactions.

3. Define the A/B Test

Objective: Set up the control and treatment groups for the A/B test.

Explanation:

- **Random Assignment:** Randomly assigning customers to either the control (Group A) or treatment (Group B) groups ensures that the groups are statistically similar, minimizing bias.
- **Customer-Level Aggregation:** Calculating the total amount spent by each customer allows us to compare overall spending between the two groups.
- **Separation of Groups:** Separating the control and treatment groups allows for straightforward comparison and statistical testing.

4. Analyze the Results

Objective: Use statistical tests to determine if there is a significant difference between the control and treatment groups.

Explanation:

- **Descriptive Statistics:** Calculating the mean and standard deviation of the total amount spent provides a summary of the central tendency and variability within each group.
- **T-Test:** The independent t-test compares the means of two independent groups (control and treatment) to see if there is a statistically significant difference. The t-statistic measures the difference relative to the variability in the data, and the p-value indicates the probability of observing the data assuming the null hypothesis (no difference) is true.

5. Draw Conclusions

Objective: Interpret the results of the statistical test to make data-driven decisions.

Explanation:

- **Significance Level (α):** The chosen significance level (typically 0.05) represents the threshold for determining statistical significance. If the p-value is less than α, we reject the null hypothesis, indicating a significant difference between the groups.
- **Conclusion:** Based on the p-value, we determine whether the observed difference in the total amount spent between the control and treatment groups is statistically significant. This helps decide if the change (e.g., a new website design) had a meaningful impact on customer spending.

Summary

- **Setup:** Configure the environment.
- **Import:** Load necessary libraries and the dataset.
- **Preprocess:** Clean and prepare the data.
- **Define Test:** Set up control and treatment groups.
- **Analyze:** Perform statistical analysis to compare groups.
- **Conclude:** Interpret the results and make decisions.

By following these steps, you conduct a comprehensive A/B testing project, which involves data preparation, experimental design, and statistical analysis to derive actionable insights.

c. **Results**

Let's break down and explain the results from your A/B test:

Descriptive Statistics

1. **Mean Total Amount (Control):** 901.44
   - This is the average amount spent by customers in the control group (Group A).
   - The mean value indicates that, on average, customers in the control group spent 901.44 units of currency.
2. **Std Total Amount (Control):** 4660.87
   - This is the standard deviation of the total amount spent by customers in the control group.
   - A high standard deviation indicates that there is a large variability in the spending behavior of customers in the control group.
3. **Mean Total Amount (Treatment):** 837.53
   - This is the average amount spent by customers in the treatment group (Group B).
   - The mean value indicates that, on average, customers in the treatment group spent 837.53 units of currency.
4. **Std Total Amount (Treatment):** 4254.27
   - This is the standard deviation of the total amount spent by customers in the treatment group.
   - Similar to the control group, the high standard deviation indicates large variability in spending behavior among customers in the treatment group.

Inferential Statistics

1. **T-statistic:** 0.6318
   - The t-statistic measures the difference between the two group means relative to the variability in the data.
   - A t-statistic close to zero suggests that the difference between the group means is small relative to the variability.
2. **P-value:** 0.5276
   - The p-value indicates the probability of observing the data assuming the null hypothesis is true (i.e., there is no difference between the control and treatment group means).
   - A p-value of 0.5276 is much higher than the typical significance level of 0.05.

Interpretation

Given the results:

- **Means Comparison:** The mean spending in the control group (901.44) is slightly higher than in the treatment group (837.53), but this difference is not substantial.

- **Variability:** Both groups exhibit high variability in spending (standard deviations of 4660.87 for control and 4254.27 for treatment), which suggests that individual spending amounts vary widely within each group.
- **Statistical Significance:** The t-statistic of 0.6318 and p-value of 0.5276 indicate that the difference in mean spending between the control and treatment groups is not statistically significant. The p-value (0.5276) is much greater than the significance level (α = 0.05), meaning we fail to reject the null hypothesis.

Conclusion

The analysis suggests that the change tested in the A/B experiment (presumably a new website design or some other modification) did not have a significant impact on customer spending. The observed difference in mean spending between the control and treatment groups is not statistically significant. Therefore, based on this data, it cannot be concluded that the treatment had any effect on the spending behavior of customers. Further experimentation or additional data might be required to draw more conclusive insights.