**1. Project title.**

**Bike Purchase Analysis: The Influence of Factors such as Income, Gender, and Marital Status.**

**2. Introduction.**

* **Objective**: The purpose of this analysis is to explore how various demographic factors such as income, gender, region and marital status influence the decision to purchase a bike. The goal is to identify patterns in the data that could inform marketing and business strategies for bike sales.
* **Context**: In the context of this analysis, we are investigating a data set that includes information on individuals’ demographic factors.

**Data Sources**: The data set used for this analysis was obtained from an Excel file which contains several demographic columns such as Marital Status, Gender, and Income, among others.

**3. Data Understanding.**

* **Data Description**: The data set contains columns like:
  + **Marital Status**: Both married and single customers are well-represented.
  + **Gender**: The data set includes a mix of male and female customers.
  + **Income**: Income levels vary across customers.
  + **Age Bracket**: The customers are grouped into Youth, Middle Age, and Old brackets, with a higher concentration in the Middle Age group.
  + **Purchased Bike**: Shows whether the person has purchased a bike or not using boolean (Yes/No) etc.

The data set was examined for duplicate values (26), which were found and removed. Several columns, such as marital status and gender, were cleaned to replace abbreviations with full text (e.g., "M" with "Married" and "S" with "Single"). The income column was also rounded to 0 decimal places for consistency.

* **Exploratory Data Analysis (EDA)**:  
  Initial data exploration involved:
  + Checking for duplicates and removing them.
  + Cleaning the columns for better readability.

Visualizing the distribution of key factors like gender, marital status, and income levels.

**4. Methodology**

* **Data Processing**:
  + **Handling Missing Values**: Missing or incomplete data points were checked, though no specific mention of missing data handling was observed in the notebook.
  + **Data Transformation**: The gender and marital status columns were transformed from coded values ("M", "F",, “M”, "S") to more descriptive text labels ("Male", "Female", "Married", "Single"). The income data was also rounded for simplified analysis.
  + **Feature Engineering**: No additional features were created, but transformations were made to ensure proper data handling and interpretation.
* **Analytical Approach**:  
  The analysis focuses on basic exploratory statistics and visualizations. No advanced machine learning models or complex statistical tests were used. The analysis relies on descriptive statistics and visual examination to draw insights.
* **Assumptions**:
  + It is assumed that the data set is a representative sample of the population of interest.
  + It is assumed that the gender, marital status, and income factors are key drivers of bike purchase decisions, though other factors could also play a role.

1. **Analysis**

* **Results**:  
  After cleaning the data, the following observations were made:
  + **Income Distribution**: There is variability in income levels across individuals, most people who earn within the average of the demography purchase bike.
  + **Marital Status and Gender Influence**: From the cleaned data, patterns emerged showing potential differences in bike purchase behavior based on marital status and gender.
* **Interpretation**:  
  The initial EDA suggests that:
  + Married individuals might show different purchase behaviors compared to singles, potentially due to differences in financial situations.
  + Gender may also play a role in influencing bike purchases.
* **Validation**:  
  No specific validation techniques were used in this analysis.

**6. Conclusion**

**Key Insights:**

Target Demographic: Middle-aged individuals with moderate to high income seem like potential target customers for bike marketing.

Regional Focus: Regions with higher purchase rates should receive targeted marketing, while regions with lower rates might benefit from awareness campaigns.

Lifestyle Considerations: Short commute distances and a lack of car ownership are lifestyle factors that could be highlighted in marketing campaigns to attract more bike buyers.

* **Recommendations**:

The insights on demographic patterns can be used to inform targeted marketing strategies, such as promoting bikes to specific income group’s e.g the middle age bracket or tailoring messaging for married vs. single customers.

* **Limitations**:

Further research could include additional factors, such as geographic location or education level, to provide a more comprehensive analysis.