

Bike Purchase Data Analysis

Bike Purchase Data Analysis

About the dataset

This report summarizes data related to bike purchases, focusing on customer demographics and purchasing behavior. The primary dataset contains 1026 rows with 13 columns of raw data and 18 columns of cleaned data, including customer ID, marital status, gender, income, number of children, education level, occupation, home ownership, number of cars, commute distance, region, age, and whether they purchased a bike. The cleaned data introduces additional columns such as 'Income_0', 'Income Bracket', 'Age Group', and 'Purchase_Flag' which is a binary representation of 'Purchased Bike'.

Initial observations from the first five sample entries reveal a mix of married and single individuals, both genders, varying income levels, and different education and occupation backgrounds. The average income is approximately \$56,208.58, with a standard deviation of \$31,278.03, indicating a wide range of income levels among customers. The average number of children is 1.89. The average age is 44.14 years. The 'Commute Distance' column shows that most of the first five customers have a short commute (0-1 Miles).

The 'Bike_Data_Dashboard_Pivot Table_0.csv' dataset provides aggregated insights, showing counts of IDs and average purchase flags broken down by various categories such as occupation, region, and marital status. This table contains 22 rows and 10 columns. The 'Sales Analysis Dashboard' dataset appears to be a placeholder or template, containing only the title repeated across all columns and rows.

Relevant Inquiries

Q1.What is the proportion of bike purchases within each income bracket and age group?



Purchase Proportion Data

- Data Overview:** The proportion_data table provides the mean Purchase_Proportion for each combination of 'Income Bracket' and 'Age Group'. The purchase proportion ranges from **0.18 to 0.61**.
- Income Bracket Categories:** The income brackets include 'Low', 'Medium', and 'High'.
- Age Group Categories:** The age groups include 'Under 30', '30-45', '46-60', and 'Over 60'.

Detailed Purchase Proportions

- **High Income Bracket:**
 - **30-45 Age Group:** The purchase proportion is **0.55**.
 - **46-60 Age Group:** The purchase proportion is **0.42**.
 - **Over 60 Age Group:** The purchase proportion is **0.40**.

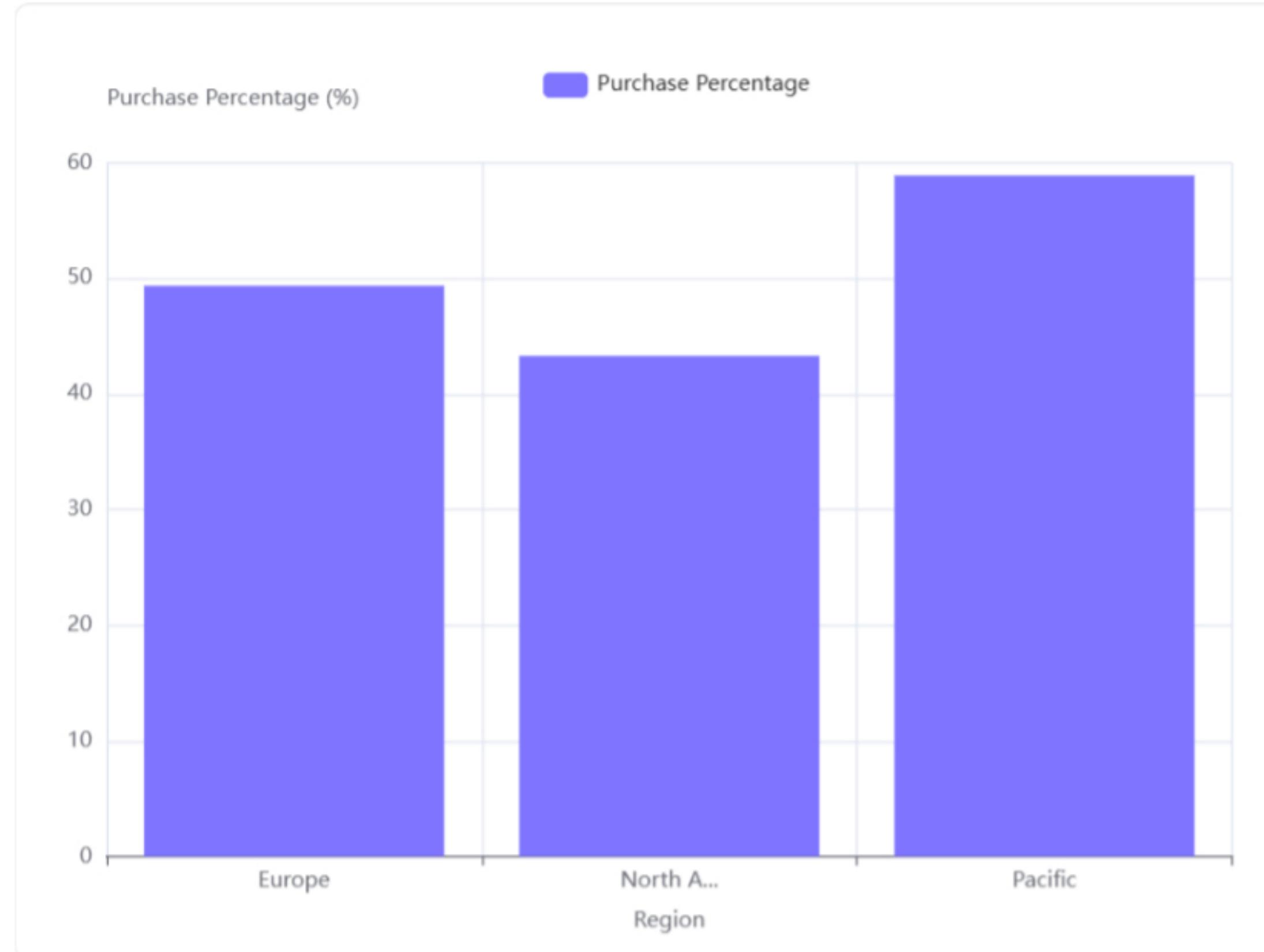
- **Low Income Bracket:**
 - **Under 30 Age Group:** The purchase proportion is **0.39**.
 - **30-45 Age Group:** The purchase proportion is **0.43**.
 - **46-60 Age Group:** The purchase proportion is **0.47**.
 - **Over 60 Age Group:** The purchase proportion is **0.18**.

- **Medium Income Bracket:**
 - **Under 30 Age Group:** The purchase proportion is **0.33**.
 - **30-45 Age Group:** The purchase proportion is **0.61**.
 - **46-60 Age Group:** The purchase proportion is **0.48**.
 - **Over 60 Age Group:** The purchase proportion is **0.33**.

Conclusion and Insights

- **Highest Purchase Proportion:** The highest proportion of bike purchases is observed in the **Medium Income Bracket for the 30-45 Age Group**, with a proportion of **0.61**. This suggests that individuals in this demographic are most likely to purchase bikes.
- **Lowest Purchase Proportion:** The lowest proportion of bike purchases is found in the **Low Income Bracket for the Over 60 Age Group**, with a proportion of **0.18**. This indicates a significantly lower likelihood of bike purchases within this segment.
- **Age Group Trends:**
 - For the **Low Income Bracket**, purchase proportion generally decreases with age, from 0.43 (30-45) to 0.18 (Over 60), with an exception for the 46-60 age group (0.47).
 - For the **Medium Income Bracket**, the 30-45 age group shows the highest proportion (0.61), which then decreases for older age groups (0.48 for 46-60 and 0.33 for Over 60).
 - For the **High Income Bracket**, the 30-45 age group has the highest proportion (0.55), followed by a slight decrease for older age groups (0.42 for 46-60 and 0.40 for Over 60).
- **Income Bracket Impact:**
 - The **Medium Income Bracket** shows the highest overall purchase proportion, particularly for the 30-45 age group.
 - The **Low Income Bracket** has the lowest purchase proportion for the Over 60 age group.
 - The **High Income Bracket** shows relatively consistent purchase proportions across age groups, with a slight decline for older demographics.

Q2.What is the percentage of bike purchases in each region?



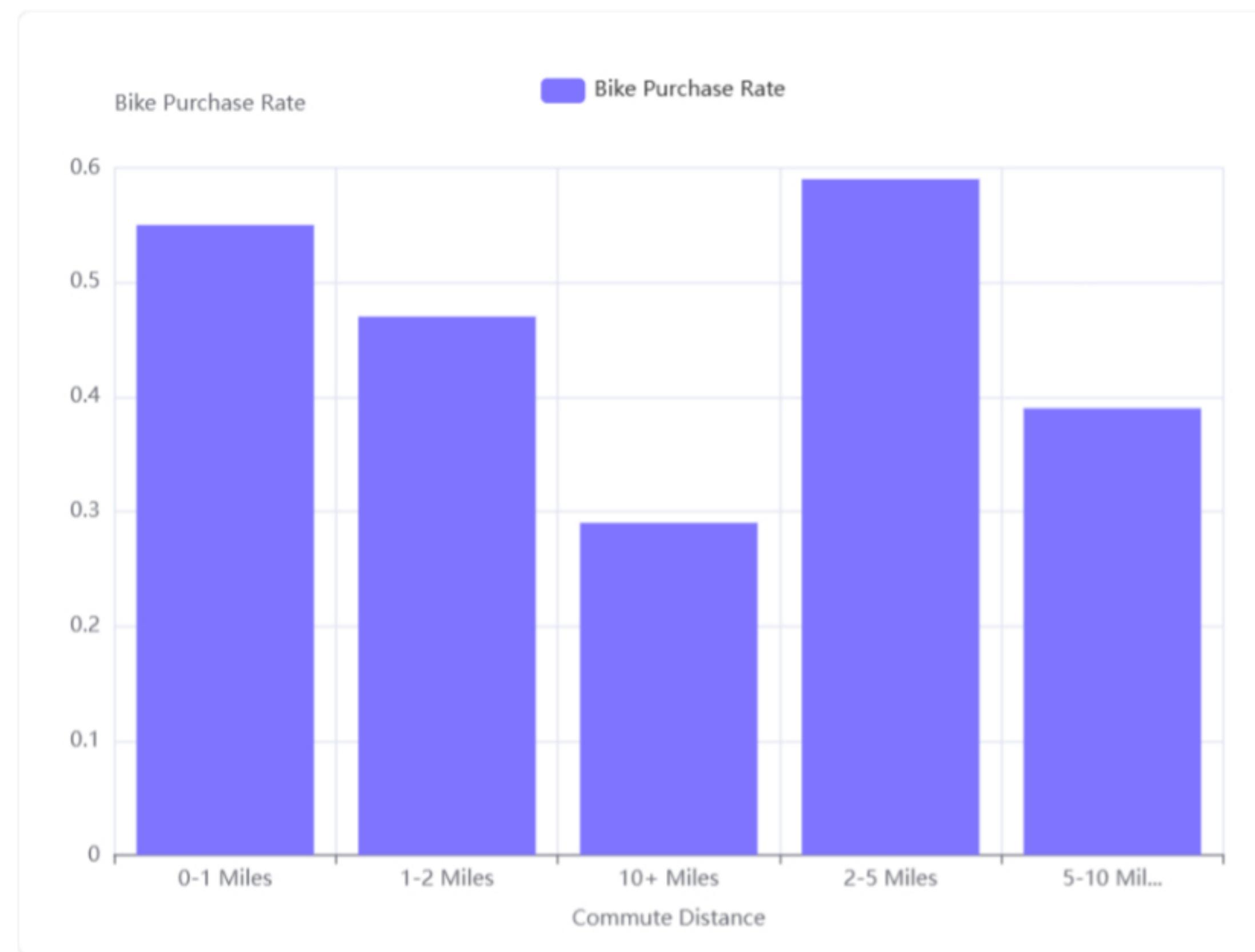
Regional Purchase Percentages

- **Europe:** The percentage of bike purchases in Europe is **49.37%**.
- **North America:** The percentage of bike purchases in North America is **43.31%**.
- **Pacific:** The percentage of bike purchases in the Pacific region is **58.91%**.

Conclusion and Insights

- **Regional Variation:** There is a notable variation in bike purchase percentages across different regions.
- **Highest Purchase Rate:** The **Pacific** region exhibits the highest bike purchase percentage at **58.91%**, indicating a strong market or higher propensity for bike purchases in this area.
- **Lowest Purchase Rate:** **North America** has the lowest bike purchase percentage at **43.31%**, suggesting potential areas for market growth or different consumer behaviors compared to other regions.
- **Moderate Purchase Rate:** **Europe** shows a moderate purchase percentage of **49.37%**, falling between North America and the Pacific region.

Q3. How does the purchase rate of bikes vary based on commute distance?



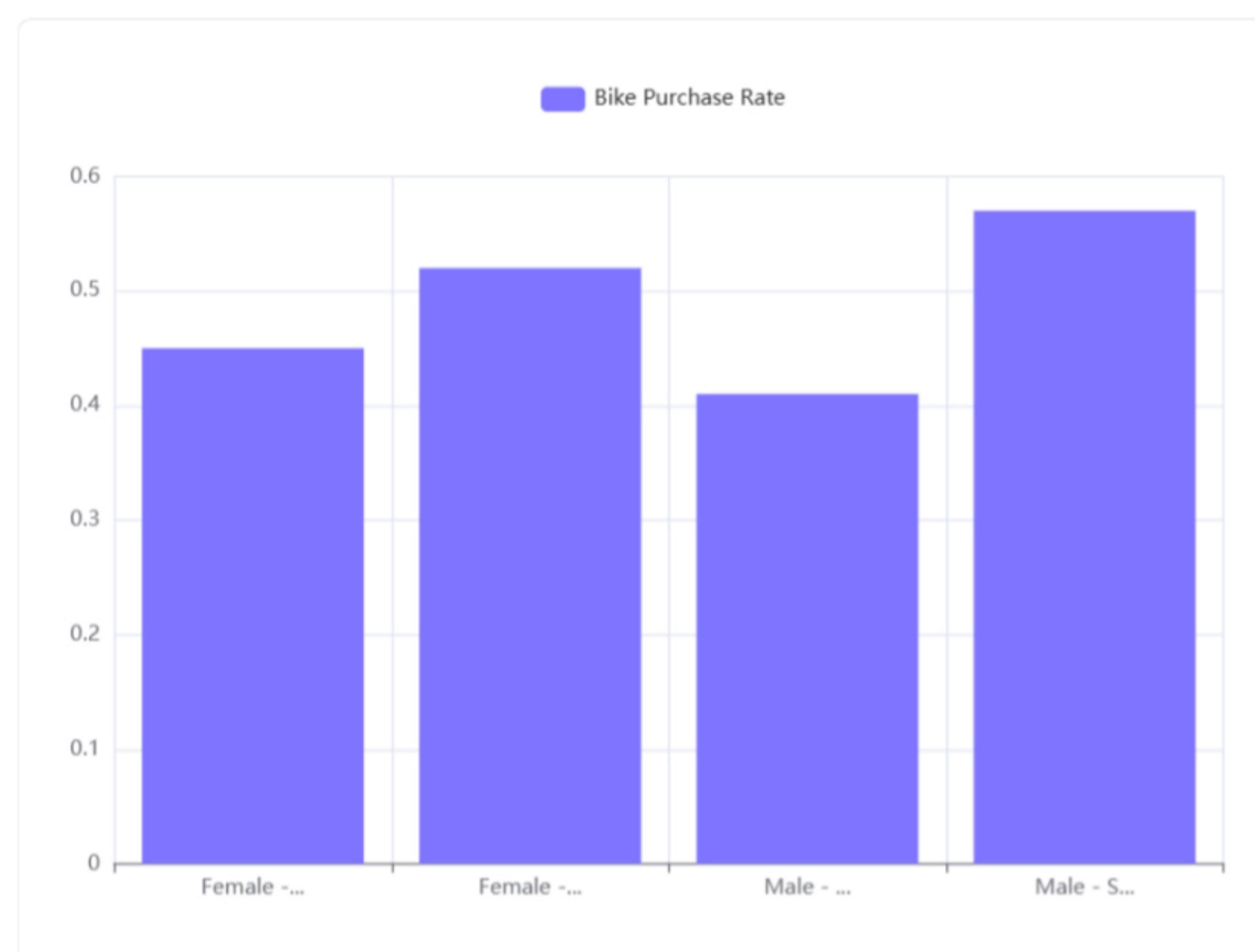
Purchase Rate by Commute Distance Category

- **0-1 Miles:** The bike purchase rate for individuals with a commute distance of 0-1 miles is approximately **0.55** (55%).
- **1-2 Miles:** The bike purchase rate for individuals with a commute distance of 1-2 miles is approximately **0.47** (47%).
- **10+ Miles:** The bike purchase rate for individuals with a commute distance of 10+ miles is the **lowest** at approximately **0.29** (29%).
- **2-5 Miles:** The bike purchase rate for individuals with a commute distance of 2-5 miles is the **highest** at approximately **0.59** (59%).
- **5-10 Miles:** The bike purchase rate for individuals with a commute distance of 5-10 miles is approximately **0.39** (39%).

Conclusion and Insights

- **Optimal Commute Distance for Purchase:** The **highest bike purchase rate** is observed among individuals with a commute distance of **2-5 miles** (approximately 59%), suggesting this range might be the most conducive for bike commuting.
- **Decreased Purchase for Longer Commutes:** There is a clear trend of **decreasing bike purchase rates as commute distance increases beyond 5 miles**. The lowest purchase rate is for those commuting **10+ miles** (approximately 29%), indicating that longer distances significantly reduce the likelihood of bike purchase.
- **Moderate Purchase for Very Short Commutes:** While not the highest, very short commutes (0-1 miles) still show a relatively high purchase rate (approximately 55%), indicating that bikes are popular for short distances as well.
- **Variability Across Distances:** The bike purchase rate varies significantly across different commute distance categories, ranging from a low of 29% for 10+ miles to a high of 59% for 2-5 miles.

Q4. What is the bike purchase rate for different combinations of gender and marital status?



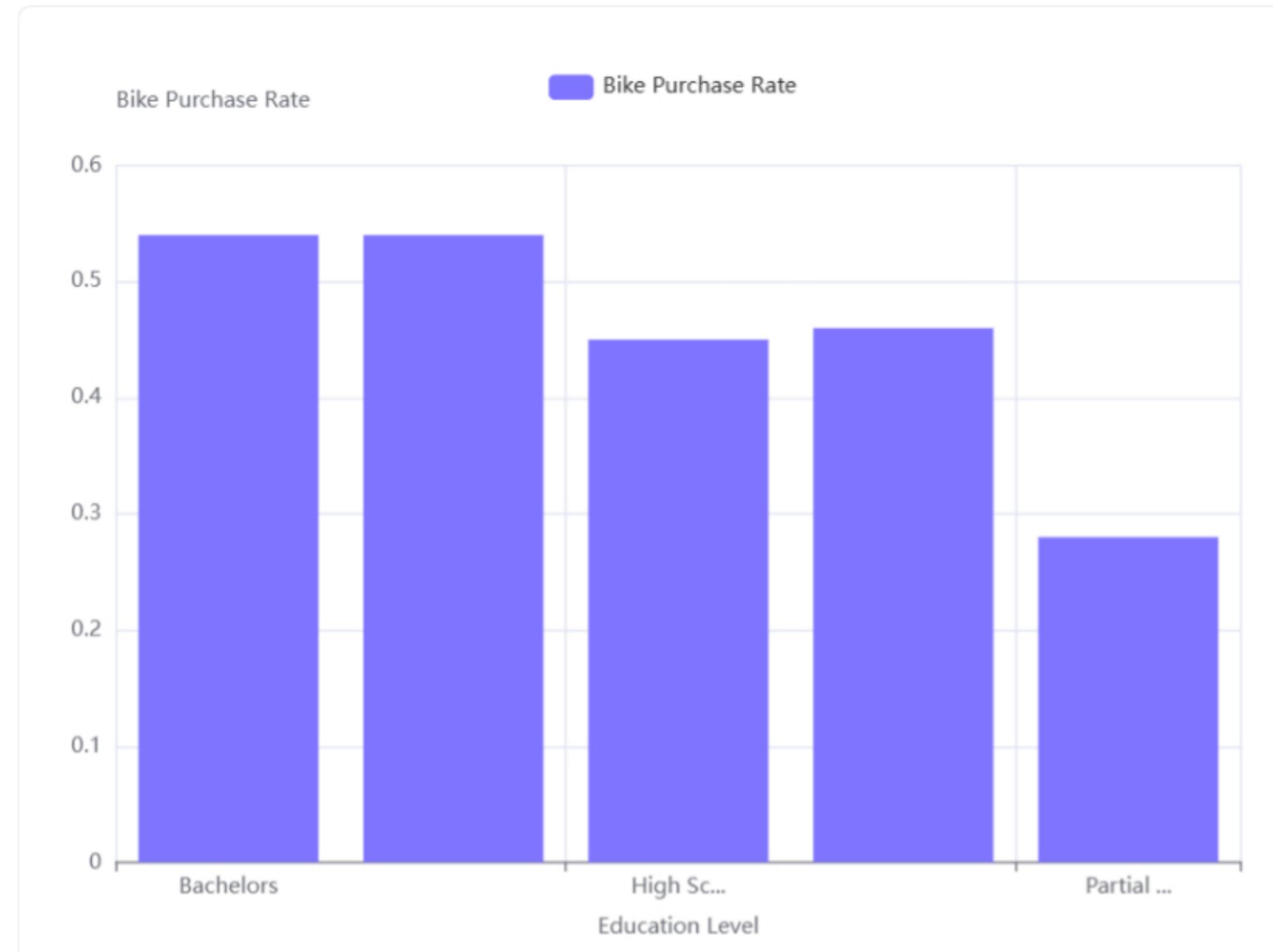
Bike Purchase Rates

- **Female - Married:** The bike purchase rate for **Female and Married** individuals is approximately **0.45** (or 44.9%).
- **Female - Single:** The bike purchase rate for **Female and Single** individuals is approximately **0.52** (or 52.0%).
- **Male - Married:** The bike purchase rate for **Male and Married** individuals is approximately **0.41** (or 41.4%).
- **Male - Single:** The bike purchase rate for **Male and Single** individuals is approximately **0.57** (or 57.0%).

Conclusion and Insights

- **Highest Purchase Rate:** **Single Males** exhibit the highest bike purchase rate at **57.0%**.
- **Lowest Purchase Rate:** **Married Males** show the lowest bike purchase rate at **41.4%**.
- **Impact of Marital Status:** For both genders, **single individuals have a higher bike purchase rate** compared to their married counterparts. Single females purchase bikes at 52.0% versus 44.9% for married females, and single males purchase at 57.0% versus 41.4% for married males.
- **Gender Differences:** Among single individuals, males have a higher purchase rate than females (57.0% vs. 52.0%). Among married individuals, females have a slightly higher purchase rate than males (44.9% vs. 41.4%).

Q5.What is the relationship between education level and bike purchase rate?



Detailed Purchase Rates

- **Bachelors:** The bike purchase rate for individuals with a Bachelor's degree is approximately **0.543**.
- **Graduate Degree:** The bike purchase rate for individuals with a Graduate Degree is approximately **0.543**.
- **High School:** The bike purchase rate for individuals with a High School education is approximately **0.446**.
- **Partial College:** The bike purchase rate for individuals with Partial College education is approximately **0.457**.
- **Partial High School:** The bike purchase rate for individuals with Partial High School education is approximately **0.282**.

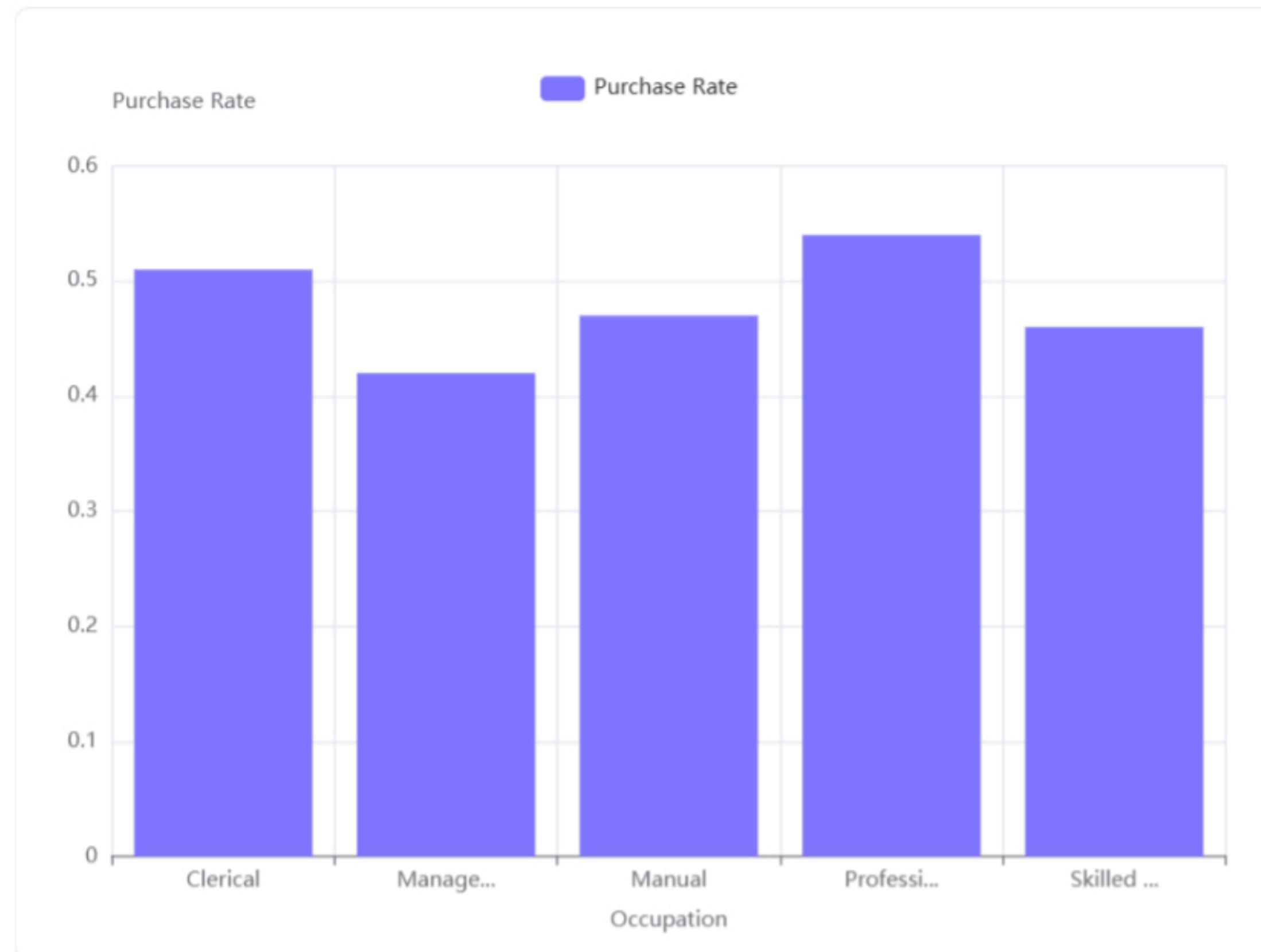
Visual Representation of Purchase Rates

- **Highest Purchase Rates:** Individuals with **Bachelors** and **Graduate Degree** education levels show the highest bike purchase rates, both at approximately **0.54**.
- **Mid-Range Purchase Rates:** Individuals with **High School** and **Partial College** education levels have mid-range purchase rates, at approximately **0.45** and **0.46** respectively.
- **Lowest Purchase Rate:** Individuals with **Partial High School** education have the lowest bike purchase rate, at approximately **0.28**.

Conclusion and Insights

- **Positive Correlation with Higher Education:** There is a clear positive correlation between higher levels of education and a higher bike purchase rate. Individuals with **Bachelors** and **Graduate Degrees** exhibit the highest propensity to purchase bikes.
- **Significant Drop for Lower Education Levels:** The bike purchase rate significantly decreases for individuals with lower education levels, with **Partial High School** showing the lowest rate, indicating a substantial difference compared to those with college or graduate degrees.
- **Tiered Purchase Behavior:** The data suggests a tiered purchasing behavior based on education: the highest rates are seen in advanced degrees, followed by high school/some college, and the lowest rates are observed in those with less than a high school education.

Q6.What is the bike purchase rate for each occupation category?



Purchase Rate by Occupation Category

- **Professional:** The highest bike purchase rate is observed among **Professional** individuals, with a rate of **0.54** (or approximately 53.57%).
- **Clerical:** **Clerical** occupations show a significant purchase rate of **0.51** (or approximately 50.80%).
- **Manual:** The bike purchase rate for **Manual** occupations is **0.47** (or approximately 46.83%).
- **Skilled Manual:** **Skilled Manual** occupations have a purchase rate of **0.46** (or approximately 45.56%).
- **Management:** The lowest bike purchase rate is found within **Management** occupations, at **0.42** (or approximately 41.95%).

Conclusion and Insights

- **Varying Purchase Rates:** Bike purchase rates vary across different occupation categories, ranging from a high of 54% for Professionals to a low of 42% for Management.
- **Targeted Marketing Opportunities:** Understanding these differences can help in tailoring marketing strategies. For instance, marketing efforts could be intensified towards Professional and Clerical groups, who show higher propensity to purchase bikes.
- **Areas for Improvement:** For occupations with lower purchase rates, such as Management and Skilled Manual, further analysis might be needed to understand the underlying reasons and develop specific campaigns to boost their purchase rates.

Q7. Analyze the likelihood of purchasing a bike for customers with 0 cars, 1 car, and 2 or more cars, to see if there's a specific threshold or pattern.



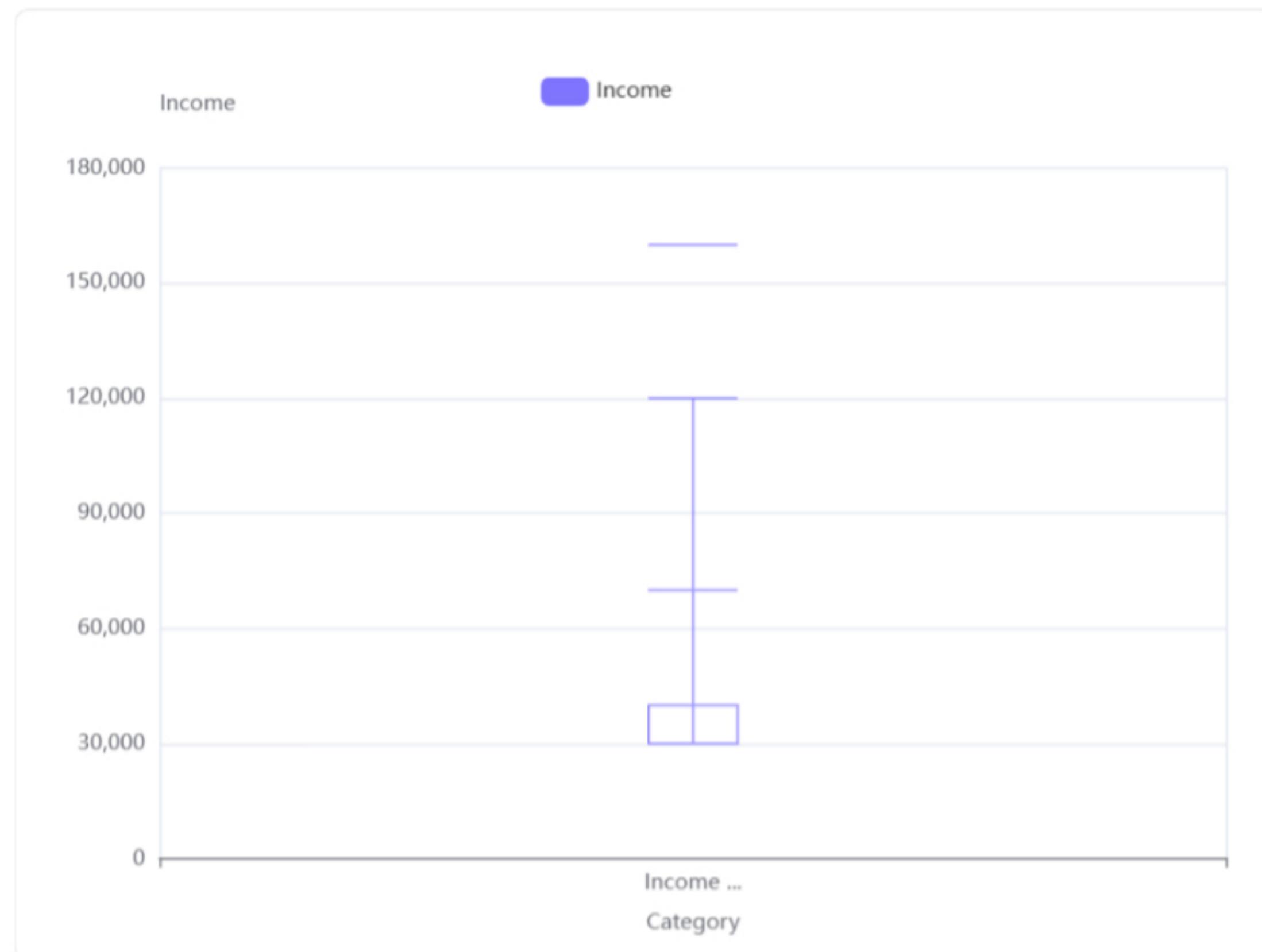
Average Bike Purchase Likelihood by Car Ownership Category

- **0 cars:** Customers owning **0 cars** have the **highest average bike purchase likelihood** at approximately **0.60** (or 60.24%).
- **1 car:** Customers owning **1 car** have a slightly lower average bike purchase likelihood than those with no cars, at approximately **0.57** (or 57.40%).
- **2 or more cars:** Customers owning **2 or more cars** have the **lowest average bike purchase likelihood** at approximately **0.37** (or 36.97%).

Conclusion and Insights

- **Inverse Relationship:** There is a clear **inverse relationship** between the number of cars owned and the likelihood of purchasing a bike. As car ownership increases, the average bike purchase likelihood decreases.
- **Significant Drop for Multiple Cars:** The most significant drop in bike purchase likelihood occurs when customers own **2 or more cars**, indicating a potential threshold where car ownership significantly reduces the need or desire for a bike.
- **Target Audience Identification:** Customers with **0 cars** represent the **most promising segment** for bike purchases, followed closely by those with **1 car**. This suggests that marketing efforts for bikes should primarily focus on individuals or households with fewer cars.

Q8. Are there any income outliers among customers who purchased bikes?



Outlier Identification Methodology

- **Outlier Definition:** Income outliers were identified for customers who purchased bikes based on the **Interquartile Range (IQR) method**. Incomes falling below $(Q1 - 1.5 * IQR)$ or above $(Q3 + 1.5 * IQR)$ were flagged as outliers.
- **Outlier Flagging:** A new column, '**Is_Outlier**', was added to the dataset, explicitly marking customers with outlier incomes as **True**.

Evidence of Outliers

- **Presence in Data:** The `bike_purchasers_income_with_outlier_flag` dataset contains records where the '**Is_Outlier**' column is marked as **True**, indicating that outliers were successfully identified.
- **Sample Outliers:** Examples from the sample data include customers with incomes of **\$160,000** and **\$120,000**, both of which were flagged as outliers.

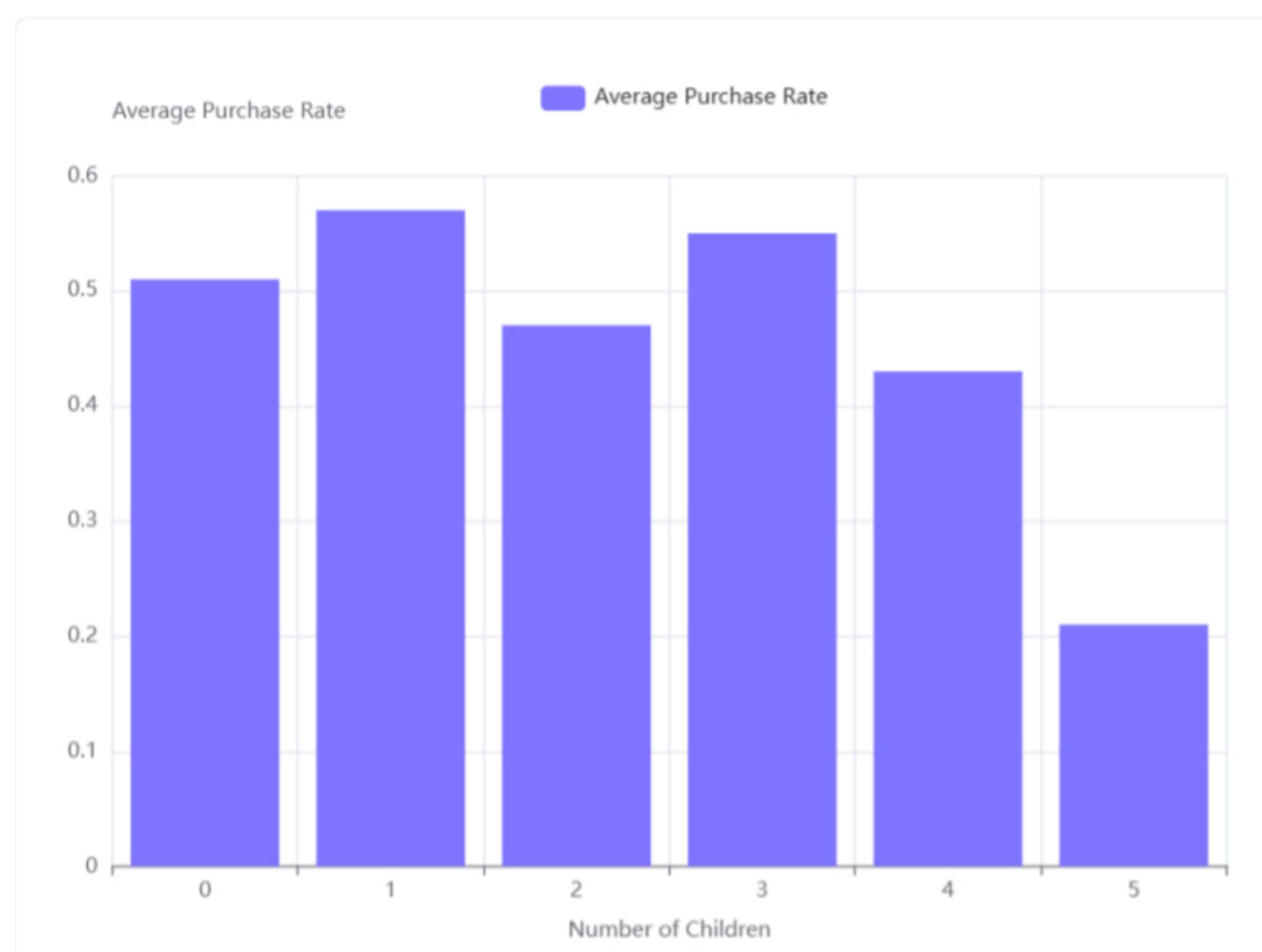
Visual Confirmation of Outliers

- **Box Plot Visualization:** The box plot titled "Income Distribution of Bike Purchasers" visually confirms the presence of income outliers. These outliers are represented by individual points or lines extending beyond the upper whisker of the box plot, specifically around the **\$160,000** income level.

Conclusion and Insights

- **Confirmed Outliers:** Yes, there are **income outliers** among customers who purchased bikes.
- **Methodological Basis:** The identification of these outliers was based on a standard statistical method using the **Interquartile Range (IQR)**, which defines outliers as values significantly deviating from the central income distribution.
- **Visual Validation:** The box plot visualization provides clear graphical evidence of these high-income outliers, showing them as distinct data points above the typical income range for bike purchasers.

Q9. How does the number of children a customer has affect their likelihood of purchasing a bike?



Detailed Purchase Rates

- **0 Children:** Customers with **no children** have an average bike purchase rate of approximately **50.5%**.
- **1 Child:** Customers with **one child** show the **highest** average bike purchase rate at approximately **57.0%**.
- **2 Children:** Customers with **two children** have an average bike purchase rate of approximately **47.5%**.
- **3 Children:** Customers with **three children** have a high average bike purchase rate of approximately **55.1%**.
- **4 Children:** Customers with **four children** have a lower average bike purchase rate of approximately **42.9%**.
- **5 Children:** Customers with **five children** have the **lowest** average bike purchase rate at approximately **21.4%**.

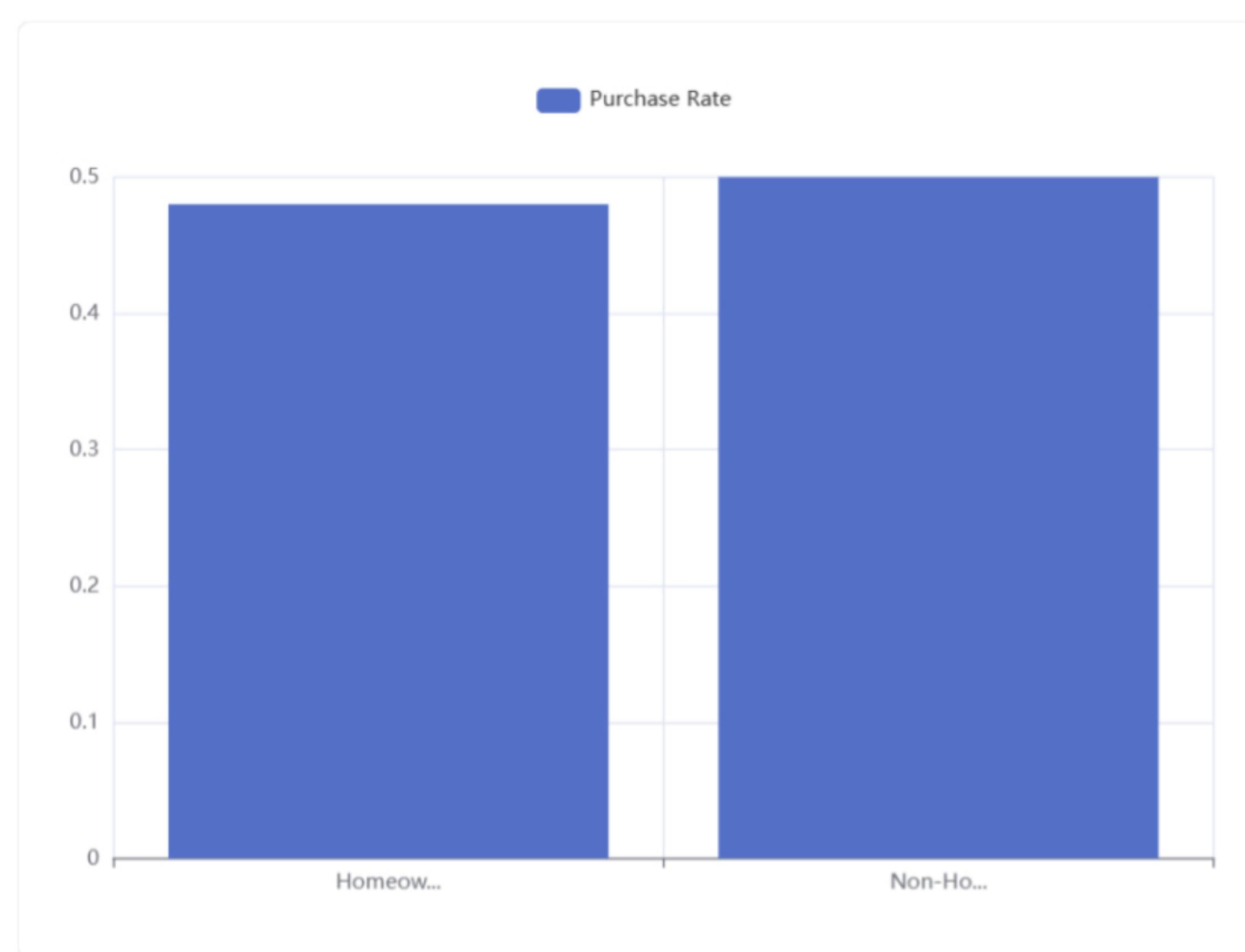
Visual Representation of Purchase Rates

- The bar chart visually confirms the average bike purchase rates for each number of children, aligning with the numerical data.
- The highest bars correspond to customers with **1 child (0.57)** and **3 children (0.55)**.
- The lowest bar corresponds to customers with **5 children (0.21)**.
- There is a noticeable **decline in purchase rate as the number of children increases beyond 3**.

Conclusion and Insights

- Optimal Purchase Likelihood:** Customers with **1 child** and **3 children** exhibit the **highest likelihood of purchasing a bike**, with average purchase rates of approximately 57% and 55% respectively. This suggests that having a small to moderate number of children might correlate with a higher propensity to purchase bikes.
- Decreasing Trend with More Children:** There is a clear **negative correlation** between the number of children and bike purchase likelihood for families with **four or more children**. The purchase rate significantly drops for customers with 4 children (42.9%) and is lowest for those with 5 children (21.4%). This could indicate that larger families might face increased financial constraints or have less disposable income for discretionary purchases like bikes, or perhaps have different lifestyle priorities.
- Moderate Likelihood for No Children:** Customers with **no children** show a moderate purchase rate (50.5%), which is lower than those with 1 or 3 children but higher than those with 4 or 5 children.
- Strategic Marketing Implications:** Marketing efforts could be most effectively targeted towards individuals or families with **1 to 3 children**, as they represent the segments with the highest bike purchase rates. Conversely, strategies for families with 4 or more children might need to address potential barriers such as cost or perceived utility.

Q10. What is the difference in bike purchase rates between homeowners and non-homeowners?



Purchase Rates

- Homeowner Purchase Rate:** The bike purchase rate for homeowners is approximately **0.48** (or 48%).
- Non-Homeowner Purchase Rate:** The bike purchase rate for non-homeowners is approximately **0.50** (or 50%).

Difference in Purchase Rates

- Calculated Difference:** The difference in bike purchase rates between homeowners and non-homeowners is **-0.0191** (approximately -0.02). This indicates that non-homeowners have a slightly higher purchase rate than homeowners.

Conclusion and Insights

- Slight Disparity:** There is a **minor difference** in bike purchase rates between homeowners and non-homeowners.
- Non-Homeowners Purchase More:** Non-homeowners exhibit a **slightly higher propensity to purchase bikes** compared to homeowners, with a purchase rate of 50% versus 48%.
- Marginal Impact:** The difference of approximately **2 percentage points** is relatively small, suggesting that homeownership status is not a major determinant of bike purchase behavior in this dataset.