Exploratory Data Analysis Project

✅ **Problem and Detection:**
The volume of bicycle sales has dropped significantly on Amazon (main problem: drop in sales volume).

✅ **Objectives:**
Understand the reasons for the drop and identify insights to create a more efficient strategy to invest in sales campaigns for the audience with the greatest purchasing potential.

✅ **Action plan:**
Perform EDA to identify patterns, relationships, and anomalies in the data to define the target audience.

EDA development:

✅ **Data identification and cleaning:**
Replace: M: Male, F: Female, M: Married and S: Single.
Income: $

✅ **Understanding the characteristics of the data:**
Why is this drop occurring?
Is it momentary or is it a trend?
What are the factors that most impact the purchase decision?
What factors are influencing this drop. Are there outliers that may be significantly affecting the results?

✅ **Visualize the distribution:**
Age.
Number of cars.
Number of children.
Gender.
Distance between home and work.
Income.
Educational level.

✅ Identifying and cleaning data:

- Replace: M: Male, F: Female, M: Married and S: Single.
- Income: $

✅ Understanding the characteristics of the data:

- Why is this drop occurring?
- Is it momentary or is it a trend?
- What are the factors that most impact the purchasing decision?
- What factors are influencing this drop. Are there outliers that may be significantly affecting the results?

✅ Viewing the distribution:

- Age.
- Number of cars.
- Number of children.
- Gender.
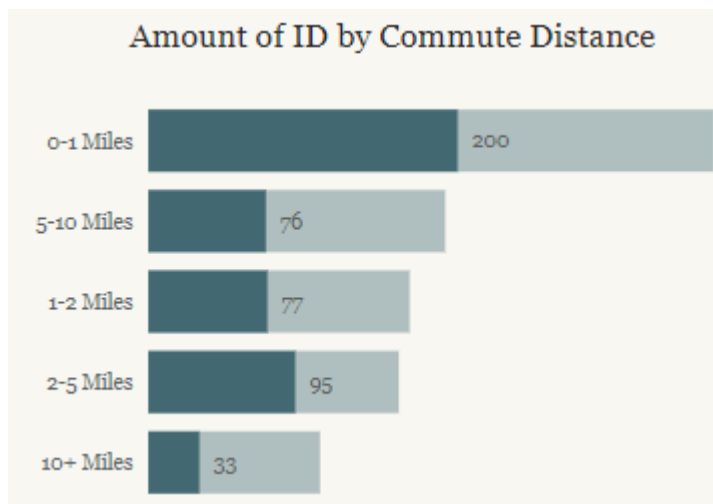- Distance between home and work.
- Income.
- Educational level.
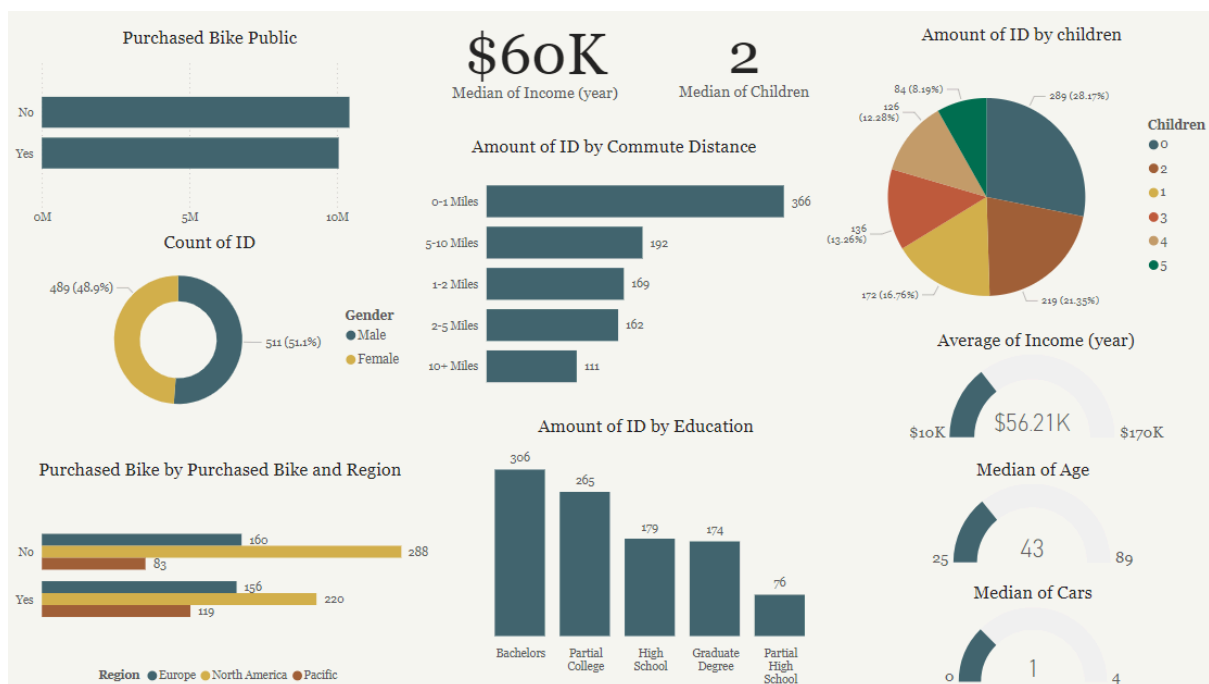
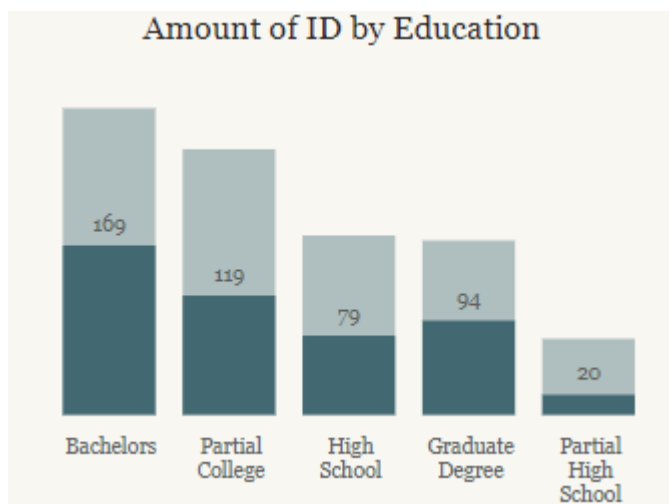## ✅ Define Benchmarks / Statistical Summary and Interpretation:

- Median of income: 60k
- Median of children: 2
- Median of age: 43 → 41 (buyer audience) and 44 (non-buyer audience).
- Median of cars: 1 → 1 (buyer audience) and 2 (non-buyer audience).

✅ **Analyze Purchasing Patterns / Trends Over Time:**

- Pattern identified 1: The audience that buys the most are people who work close to home '0-1 miles', '1-2 miles' and '2-5 miles'.

Amount of ID by Commute Distance

0-1 Miles	200
5-10 Miles	76
1-2 Miles	77
2-5 Miles	95
10+ Miles	33

- Identified pattern 2: People with Bachelor's and Partial College are the ones who buy the most.

Amount of ID by Education

| Bachelors	Partial College	High School	Graduate Degree	Partial High School
169	119	79	94	20

- Identified pattern 3: It is observed that those who have one child have greater purchasing potential than those who have more children, as observed in the graph.
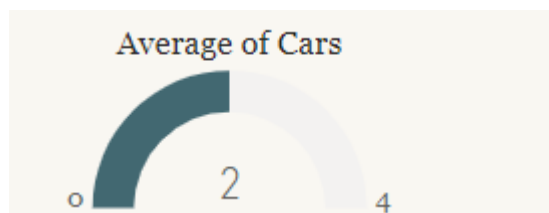
Amount of ID by children
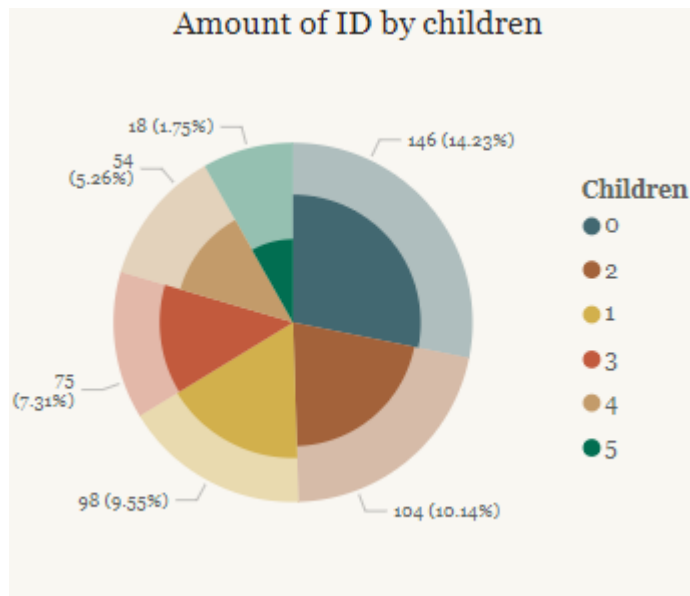
- Identified pattern 4: Analyzing the number of cars that people who buy bicycles have, on average, those who buy bicycles have only 1 car, and those who do not buy bicycles have two cars.

Bought:
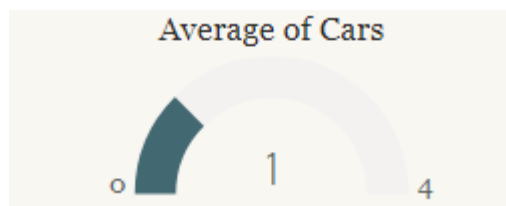
Did not buy:

✅Generate Hypotheses:

- Hypothesis 1: Those who have only one car are more likely to buy a bicycle.
- Hypothesis 2: Those who live 0-5 miles from work are more likely to buy a bicycle.
- Hypothesis 3: Those who have only 1 child are more likely to buy a bicycle.

✅ Examine Relationships (Identify Relationships Between Variables):

Since the graph shows the same dispersion between the purchased and non-purchased audiences, there is no relationship between the regions, this may just be a question of sample difference between the regions for research.
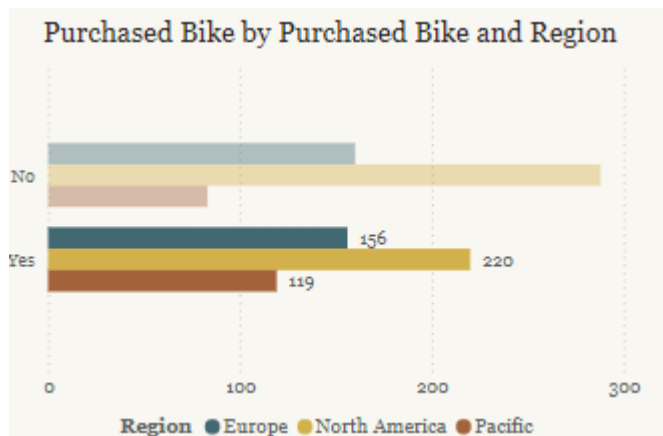

Purchased Bike by Purchased Bike and Region
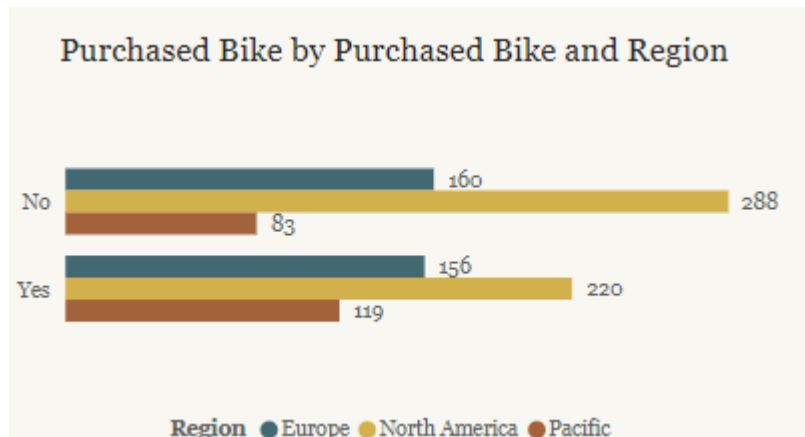
- There is no difference between genders for those with greater purchasing potential.
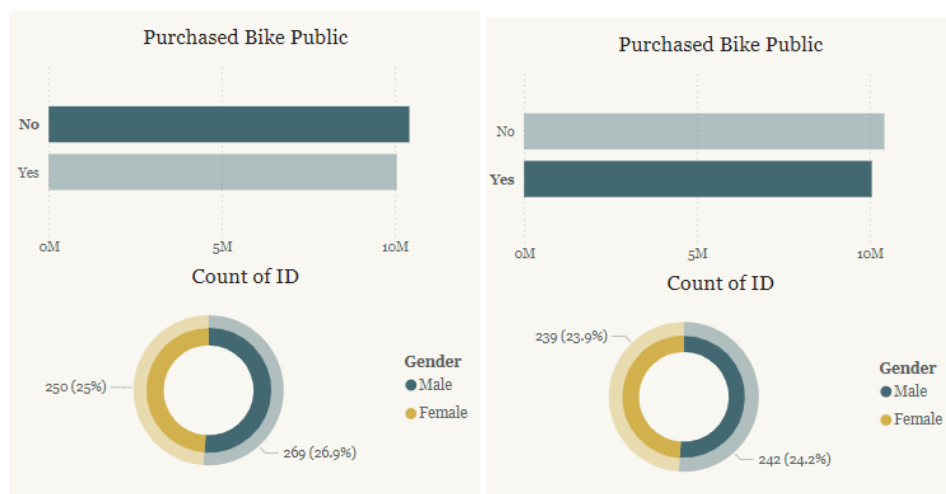
✅**Check Anomalies:**

Anomaly 1: Research was conducted with a larger number of people in Europe, which may change the results relatively due to cultural differences.

Purchased Bike by Purchased Bike and Region

Conclusão:

- **Why is this drop occurring?** Campaigns are not targeted to the right audience. It is important to consider the number of children, the distance between home and work, and the number of cars these people have.
- **Is this temporary or a trend?** It is a permanent trend, since people who have more children will have more cars, and in addition, those who live further from work do not use their bikes to go to work, for example (or if they live far from work, they probably do not have enough time to ride a bike).
- **What are the factors that most impact the purchasing decision?** Number of children, number of cars, and the distance between home and work.

- **Insight 1:** It is considered interesting to apply campaigns, for example, covering the health factor, so that people who live closer to work (distance between 0-5 miles) will be encouraged to go to work by bike.
- **Insight 2:** In addition, filtering campaigns for people with only one car is something worth exploring, so you can compare the price of buying a second car with just a simple decision to buy a bicycle.