Bike Buyers Data Analysis Report

Introduction

This project set out to explore customer demographics and lifestyle patterns to uncover what drives people to purchase bikes. Using Microsoft Excel, I carried out a full cycle of data cleaning, exploratory analysis, and dashboard creation. The goal was to turn raw data into clear insights that businesses can act on.

Dataset Overview

The dataset, known as the Bike Buyers dataset, contains information about nearly 1,000 customers. Each record captures details about who they are, how they live, and whether they purchased a bike.

Attributes:

- Gender and Marital Status
- Income and Age
- Education and Occupation
- Commute Distance and Region
- Home Ownership and Number of Cars
- Bike Purchase (Yes/No)

Tools Used

The entire analysis was conducted in Microsoft Excel:

- Data Cleaning functions
- Pivot Tables for quick summaries
- Charts for visualization
- An interactive dashboard with slicers

Data Cleaning

Before diving into analysis, the dataset was carefully cleaned to ensure accuracy:

- Duplicates removed using unique customer IDs.
- Missing values handled:
    - Categorical blanks (Gender, Marital Status, Home Owner) replaced with “Unknown”.
    - Numerical blanks (Income, Age, Cars, Children) filled using average values.
- Inconsistent labels (like “Y/N”) were converted to “Yes/No”, and text formatting issues were fixed.
- Feature engineering:
    - Age grouped into categories (Under 30, 30–40, 40–50, Above 50).
    - Income grouped into Low, Medium, High. These new features made the analysis more intuitive and easier to segment.

Exploratory Data Analysis (EDA)

Using Pivot Tables and Charts, several patterns emerged:

- Bike Purchase Distribution: Most customers did not buy a bike, suggesting purchases are selective.
- Income vs Purchase: Bike buyers tend to have higher incomes.
- Age vs Purchase: The 30–50 age group showed the highest purchase rates mostly working professionals with stable earnings.
- Commute Distance vs Purchase: Shorter commutes correlated strongly with bike ownership, maybe convenient as a driver.
- Region vs Purchase: Europe and Pacific regions stood out with higher adoption, hinting at cultural and infrastructure influences.

Dashboard Design

An interactive Excel dashboard was built to bring these insights to life.

- Charts: Distribution of purchases, Income vs Purchase, Age Group Analysis, Commute Distance, Region-wise breakdown.
- Filters (Slicers): Gender, Age Group, Region, Income Group.

This allowed users to slice and dice the data, exploring customer behavior from multiple angles.

Business Insights

From the analysis, several clear takeaways emerged:

- Higher-income customers are more likely to purchase bikes.
- Middle-aged customers (30–50) form the core buyer segment.
- Short commute distances significantly increase the likelihood of bike purchases.
- Regional differences matter: Europe and Pacific regions show stronger adoption.

Conclusion

This project demonstrates how Excel can be used for end-to-end data analysis from cleaning messy data to uncovering meaningful insights and presenting them in a dashboard. The findings provide a roadmap for businesses: target middle-aged, higher-income customers with short commutes, and strategies to regions with stronger cycling and presenting them in a dashboard.