**Google Data Analytics Certificate**

**Couse 8: Capstone**

**Case Study: Bike Purchase Insights**

Kumail Alhuwayji

May 20, 2024

Table of Contents

[Scenario 3](#_Toc167207169)

[Dataset Overview 3](#_Toc167207170)

[Objective 3](#_Toc167207171)

[Task 3](#_Toc167207172)

[Summary of the Business Task 4](#_Toc167207173)

[Description of All Data Sources Used 4](#_Toc167207174)

[List of Any Cleaning or Manipulation of Data 4](#_Toc167207175)

[Summary of Analysis 5](#_Toc167207176)

[Supporting Visualizations 5](#_Toc167207177)

[Recommendations Based on Analysis 8](#_Toc167207178)

Table of Charts

[Chart 1 Avg Income Per Purchased Bike 6](#_Toc167207265)

[Chart 2 Purchased Bike Per Age Brackets 6](#_Toc167207266)

[Chart 3 Purchased Bike by Gender 7](#_Toc167207267)

[Chart 4 Purchased Bike by Region 7](#_Toc167207268)

[Chart 5 Purchased Bike Per Commute Distance 8](#_Toc167207269)

# Scenario

Stakeholder: Jane Smith, Marketing Director at XYZ Retail Company.

# Dataset Overview

You have been provided with a dataset named "bike buyers" that contains information about customers and their bike purchasing behavior. The columns in the dataset are as follows:

* ID: Unique identifier for each customer.
* Marital Status: Customer's marital status (M, S).
* Income in $: Customer's annual income in dollars.
* Gender: Customer's gender (M, F).
* Number of Children: Number of children the customer has.
* Education: Customer's education level (Bachelor, High School, Partial College, Graduate Degree).
* Occupation: Customer's occupation (Skilled Manual, Clerical, Professional, Others).
* Home Owner: Whether the customer is a homeowner (Yes, No).
* Number of Cars Owned: Number of cars owned by the customer.
* Commute Distance: Distance the customer commutes to work (0-1 Miles, 2-5 Miles, 5-10 Miles, More than 10 Miles).
* Region: Customer's region (Pacific, Europe, Others).
* Age: Customer's age.
* Purchased Bike: Whether the customer purchased a bike (Yes, No).

# Objective

As the Marketing Director, Jane Smith wants to understand the key factors that influence bike purchasing decisions among customers.

# Task

1. A clear summary of the business task

2. A description of all data sources used

3. Documentation of any cleaning or manipulation of data

4. A summary of analysis

5. Supporting visualizations and key findings

6. Recommendations based on analysis

# Summary of the Business Task

The business task of this analysis is to identify key factors influencing bike purchasing decisions among customers. By understanding these factors, XYZ Retail Company can tailor its marketing strategies to effectively target potential bike buyers and increase sales.

# Description of All Data Sources Used

The primary dataset used in this analysis is the "bike buyers" dataset, which contains information on customer demographics, lifestyle factors, and bike purchasing behavior. The dataset includes the following columns: ID, Marital Status, Income, Gender, Number of Children, Education, Occupation, Home Owner, Number of Cars Owned, Commute Distance, Region, Age, and Purchased Bike. No additional data sources were used.

# List of Any Cleaning or Manipulation of Data

* Removed 26 duplicated rows to ensure the dataset's integrity and accuracy.
* Changed values from "S" and "M" to "Single" and "Married" in the 'Marital Status' column to enhance readability and clarity during visualization and analysis.
* Changed values from "M" and "F" to "Male" and "Female" in the 'Gender' column to make the data more understandable during visualization and analysis.
* Converted the values in the 'Income' column from plain numbers to currency format, improving the presentation and comprehension of income data.
* Created a new column called 'Age Brackets' to categorize ages into three groups: Adolescent, Middle Age, and Old. This categorization simplifies the analysis of age-related trends and insights.

# Summary of Analysis

* Middle-aged individuals are more likely to purchase bikes compared to adolescent and older adults.
* Bike purchase rates increase with higher income.
* Customers with a commute distance of 0-1 miles are the most likely to purchase bikes.
* Approximately 50% of customers in the Europe region have purchased bikes.
* Those with a commute distance of more than 10 miles have the lowest bike purchase rate.
* Homeowners with a short commute distance (0-5 miles) have the highest bike purchase rate.
* Non-homeowners with long commutes (more than 10 miles) are the least likely to buy bikes.

# Supporting Visualizations

Chart 1 Avg Income Per Purchased Bike

Chart 2 Purchased Bike Per Age Brackets

Chart 3 Purchased Bike by Gender

Chart 4 Purchased Bike by Region

Chart 5 Purchased Bike Per Commute Distance

Additionally, there is a simple dashboard included in the Purchased\_Bike\_Analysis.xlsx file on a separate sheet.

# Recommendations Based on Analysis

* Create marketing campaigns aimed at middle age with higher incomes, highlighting premium features and benefits.
* Promote the advantages of biking for short commutes (0-1 miles) to attract customers who live close to their workplaces.
* Segment the marketing strategy by income levels. Offer premium models and financing options to high-income customers, and more affordable models to lower-income customers.