Code Alpha Internship Submit Report

𝗜𝗻𝘁𝗲𝗿𝗻 𝗗𝗲𝘁𝗮𝗶𝗹𝘀:  
- 𝗡𝗮𝗺𝗲: Mrudula Kamdi  
- 𝗦𝘁𝘂𝗱𝗲𝗻𝘁 𝗜𝗗: CA/AU1/4643  
- 𝗗𝗼𝗺𝗮𝗶𝗻: Data Analytics  
- 𝗗𝘂𝗿𝗮𝘁𝗶𝗼𝗻: 1st August 2025 to 30th August 2025

Two Task Completed:

1.Exploratory Data Analysis (EDA) of Bike Purchase Behaviour

2.Data Visualization Project Report

**1.Exploratory Data Analysis (EDA) of Bike Purchase Behaviour**

1. Executive Summary

This project focuses on performing an Exploratory Data Analysis (EDA) on a dataset to understand the factors that influence a customer's decision to purchase a bike. By using Python, specifically the Pandas, Matplotlib, and Seaborn libraries, I analyzed the data's structure, identified key demographic trends and patterns, and uncovered potential data issues. The analysis provides actionable insights into customer behaviour, which can be used to inform marketing strategies and business decisions.

2. Dataset Overview

* Source: Data1
* Description: The dataset contains 366 rows and 14 columns of customer information, including demographics (Gender, Age, Married Status), financial data (Income), lifestyle choices (Home Owner, Cars), and regional information (Region). The primary objective of the analysis was to understand the target variable, Purchased Bike.

3. Key Findings & Insights

* Customer Profile: The analysis revealed that the typical bike purchaser is most likely a middle-aged single female.
* Income & Purchase: While the average income of bike purchasers is slightly higher than non-purchasers, the median income is the same for both groups. This suggests that a higher income is not a strict requirement for a purchase, but bikes are more popular among high-income earners.
* Home & Car Ownership: A strong correlation was found between bike purchase and home and car ownership. A majority of bike purchasers are homeowners, and the largest group of purchasers owns zero cars, suggesting that bikes may serve as a primary mode of transport for this demographic.
* Regional Trends: The Europe region showed the highest number of bike purchasers, indicating a significant market opportunity in that area.

4. Technical Skills Demonstrated

* Data Loading & Inspection: Used Pandas to load the CSV file and inspect its structure, data types, and identify missing values.
* Data Cleaning: Confirmed the dataset was clean, with no duplicate rows or inconsistent data formatting.
* Exploratory Data Analysis (EDA): Formulated and answered meaningful business questions by calculating summary statistics and creating visualizations.
* Data Visualization: Utilized Matplotlib and Seaborn to create a variety of plots (bar charts, box plots, scatter plots) to identify trends, patterns, and anomalies.
* Hypothesis Testing: Framed initial findings as hypotheses and used analytical methods to either support or challenge them, demonstrating a solid understanding of data-driven conclusions.

5. Conclusion

This project successfully applied a comprehensive EDA framework to a real-world dataset. The insights gained provide a clear understanding of the customer demographics and key factors influencing bike purchases. This work highlights my ability to transform raw data into a coherent narrative that can be used to drive strategic business decisions.

**2.Data Visualization Project Report**

1. Project Objective

The main objective was to complete a data visualization task by:

* Transforming raw data into visual formats.
* Using Python libraries like pandas and matplotlib.
* Designing visuals that enhance understanding.
* Crafting a compelling data story.

2. Dataset Overview

The project utilized a CSV dataset containing financial records, including:

* Reporting Date: A time-series variable.
* Product Description: A categorical variable.
* Balance Outstanding: A numerical variable.
* Interest Rate: A numerical variable.
* Client Name and Client Code: Categorical variables.

3. Visualizations and Key Findings

Four distinct visualizations were created, each designed to answer a specific business question and reveal a different aspect of the data.

Line Chart: Total Balance Outstanding Over Time

* Purpose: To show trends and patterns of the total outstanding balance over time.
* Key Finding: The visualization revealed a consistent upward trend in the outstanding balance, suggesting a positive growth in the company's loan or product portfolio.

Bar Chart: Total Balance Outstanding by Product

* Purpose: To compare the total outstanding balance across different product categories.
* Key Finding: The bar chart effectively ranks each product by its contribution to the total balance, making it easy to identify top-performing products at a glance.

Scatter Plot: Balance Outstanding vs. Interest Rate

* Purpose: To explore the relationship or correlation between the outstanding balance and the assigned interest rate.
* Key Finding: This visualization can help identify any potential trends, such as whether higher outstanding balances are associated with higher or lower interest rates.

Pie Chart: Percentage of Total Balance by Product

* Purpose: To show the proportional contribution of each product to the total outstanding balance, giving a "part-to-whole" perspective.
* Key Finding: The pie chart provides a quick, clear summary of how the total balance is distributed among the different products, highlighting the most significant contributors in terms of percentage.

4. Conclusion

Through these visualizations, the raw financial data was transformed into actionable insights. Each chart tells a part of the story, from long-term growth trends to the specific performance of individual products. This comprehensive approach demonstrates proficiency in key data visualization skills, fulfilling all the project requirements.