Project Overview and Summary

This project focuses on **analysing customer data** from a banking dataset to uncover insights into customer demographics, behaviour, and factors influencing retention. The analysis employs **Exploratory Data Analysis (EDA)** techniques, including **univariate, bivariate, and multivariate analysis**, to visualize trends and patterns. Additionally, data preprocessing techniques such as **handling imbalances** and **feature scaling** were applied to ensure accurate modelling.

A **Decision Tree Model** was built to predict customer behaviours, such as the likelihood of customer churn. The study further explores **feature importance analysis** to determine key variables influencing customer decisions. The insights derived from this project can help the bank make **data-driven decisions** to improve customer engagement and retention strategies.

Key Data Insights & Visualizations:

1. **Customer Demographics:**
    - The majority of customers (**50.14%**) are from **France** (**5,014 out of 10,000 customers**), followed by other regions.
    - **Visualization:** A **bar chart** showcases the distribution of customers across different geographies.
2. **Gender Distribution:**
    - **Male customers are more prevalent than female customers.**
    - **Visualization:** A **pie chart** illustrates the gender ratio, confirming the male-dominant customer base.
3. **Card Type Analysis:**
    - All four card types have a **nearly equal number of customers**, indicating no significant preference for any particular type.
    - **Visualization:** A **histogram or grouped bar chart** demonstrates the balanced card distribution.
4. **Feature Scaling & Imbalance Handling:**
    - Standardization techniques were applied to numerical features to ensure model consistency.
    - Imbalanced data (if present) was addressed to improve the predictive model's accuracy.
5. **Predictive Modelling:**
    - A **Decision Tree Model** was trained to analyse customer behaviour and predict churn likelihood.
    - **Model Evaluation:** Performance metrics were assessed to ensure reliability.
    - **Feature Importance Analysis:** The most influential features in determining customer behaviour were identified.

Skills Required for This Project:

To build this banking customer analysis project, you need a mix of **data science, machine learning, and analytical skills**:

1. **Data Analysis & Visualization** – Use **Python (pandas, NumPy, Matplotlib, Seaborn)** to analyse and visualize customer trends.
2. **Data Preprocessing** – Handle **missing values, feature scaling, and imbalanced data** to improve model accuracy.
3. **Machine Learning & Model Building** – Implement a **Decision Tree model** using **scikit-learn**, and evaluate it with metrics like **accuracy, precision, and recall**.
4. **Statistical Knowledge** – Apply **descriptive statistics, probability, and hypothesis testing** for better insights.
5. **Business & Banking Understanding** – Knowledge of **customer segmentation, churn analysis, and retention strategies** to make data-driven decisions.
6. **Programming & Tools** – Work with **Jupyter Notebook, Python libraries, and Git/GitHub** for version control.
7. **Problem-Solving & Critical Thinking** – Interpret trends, debug issues, and optimize models for better predictions.

These skills help in analysing customer behaviour, predicting churn, and recommending strategies for improved customer retention.

Recommendations:

- **Customer Retention Strategies:** Since France has the highest number of customers, targeted retention campaigns can focus on this region.
- **Personalized Offerings:** Male customers dominate the dataset, so gender-specific marketing strategies might be beneficial.
- **Data-Driven Decision Making:** Analysing customer preferences based on the most influential features can help in better customer segmentation and engagement.

Conclusion:

This analysis provides a **data-driven approach** to understanding customer demographics and behaviours. The **visualizations** effectively highlight key trends, and the **predictive model** helps in making strategic decisions. By leveraging these insights, businesses can improve customer retention and optimize financial offerings.