# Customer Purchase Prediction - Project Summary

## Objective

The goal of this project was to predict whether a customer will make a purchase based on their demographics and behavior ( time on website, number of past purchases, loyalty status ). This helps businesses better target marketing and improve sales strategies.

## What I Did (Step-by-Step)

* Loaded and explored customer data  
  - Checked for missing values, data types, and class distribution.  
  - Dropped unnecessary columns like Gender and ProductCategory.
* Exploratory Data Analysis (EDA)  
  - Used histograms and boxplots to understand feature distributions and spot outliers.  
  - Found no missing data, slight class imbalance, and some outliers in income and purchases.
* Outlier Handling  
  - Used percentile capping (1st and 99th percentile) to reduce the influence of extreme values.
* Feature Scaling & Preprocessing  
  - Used StandardScaler to normalize numerical features for better model performance.
* Modeling (Classification)  
  - Trained three machine learning models:  
   • Logistic Regression (Baseline)  
   • Decision Tree Classifier  
   • Random Forest Classifier  
  - Evaluated performance using accuracy, precision, recall, F1-score, and confusion matrix.
* Model Evaluation  
  - Random Forest performed best with:  
   • Accuracy: 95.7%  
   • High precision and recall for identifying actual purchasers.
* Feature Importance  
  - Found that Time Spent on Website, Age, and Annual Income were top drivers of purchasing behavior.

## Business Takeaway

Time spent on the website is the strongest indicator of a potential purchase. Businesses can focus on increasing site engagement and targeting users who linger longer.

## Summary

This project showcases a complete machine learning pipeline from data exploration and preprocessing to model evaluation and feature importance visualization. It is intended as a portfolio project to demonstrate skills in classification modeling, EDA, and communicating insights using both code and visualizations.