### **Data Loading and Initial Exploration**

* Purpose: Load the dataset from a CSV file and display the first few rows and summary statistics to understand the data structure and initial quality.
* Actions:
  + Load data using pandas.
  + Display first few rows to get a preliminary view of the columns and values.
  + Generate descriptive statistics to understand the distribution of numerical features.
  + Provide an information summary including data types and the presence of null values.

### **Data Cleaning**

* Purpose: Prepare the dataset for analysis by removing unnecessary columns and handling missing values.
* Actions:
  + Drop columns that are not relevant to the prediction model, such as 'Retention' and 'CLV'.
  + Remove any rows that contain missing values to ensure model accuracy.
  + Confirm that all missing values have been addressed by checking if there are still any nulls left.

### **Data Visualization**

* Purpose: Visualize different aspects of the dataset to uncover patterns, detect outliers, and understand variable distributions.
* Visualizations Include:
  + Histograms to see the distribution of credit scores and check for skewness or bi-modality.
  + Box plots to visually assess the spread, central tendency, and outliers within the credit scores.
  + Correlation heatmaps to evaluate how variables interrelate, particularly how they might influence the credit score.

### **Feature Engineering and Preprocessing**

* Purpose: Transform raw data into a format suitable for modeling, enhancing the model's ability to learn effectively.
* Actions:
  + Identify and separate target variable and features.
  + Apply one-hot encoding to categorical variables to make them machine-readable.
  + Scale numeric features to normalize their range and variance.
  + Impute missing values in a manner appropriate to their distribution and importance.

### **Model Building and Compilation**

* Purpose: Construct a neural network to predict credit scores based on the processed features.
* Actions:
  + Define a sequential model with dense layers that learn to map features to the target.
  + Include normalization and dropout for better generalization.
  + Compile the model with an optimizer and loss function suited to regression tasks.

### **Model Training and Evaluation**

* Purpose: Fit the model to the training data to learn the relationships between features and the target, and evaluate its performance on unseen data.
* Actions:
  + Train the model using batches of data, validating its performance on a holdout set.
  + Assess the model's prediction error and accuracy using metrics such as mean absolute error.

### **Prediction and Categorization**

* Purpose: Use the trained model to make predictions on new data and categorize the results into meaningful groups.
* Actions:
  + Predict credit scores using the model and preprocessing steps on new data samples.
  + Classify predictions into categories like 'Poor', 'Average', or 'Wealthy' based on score thresholds.
  + Calculate the distribution of these categories across a dataset to understand demographic financial health.