**Project Proposal:** **Consumer Electronics Sales Data Analysis and Prediction**

**Introduction:** The dataset consists of **9000** rows and **9** columns, providing comprehensive information. The columns in the dataset include:

ProductID, ProductPrice, CustomerAge, ProductCategory, ProductBrand, PriceRange, CustomerGender, PurchaseFrequency, CustomerSatisfaction, PurchaseIntent.

This project aims to analyze sales data from a consumer electronics dataset to identify key trends and predict future purchasing patterns.

**Methodology:**

1. **Data Collection:**
   * Load and inspect the dataset to understand its structure and contents.
   * Clean the data by handling missing values and removing irrelevant columns (e.g., ProductID).
2. **Data Preprocessing:**
   * Bin continuous variables (e.g., ProductPrice, CustomerAge) into categorical ranges to simplify analysis.
   * Encode categorical variables (e.g., ProductCategory, ProductBrand, PriceRange) using Label Encoding.
3. **Exploratory Data Analysis (EDA):**
   * Visualize the distribution of product prices, customer ages, and gender distribution to identify trends and patterns.
4. **Model Training:**
   * Select features (CustomerAge, PriceRange, CustomerGender) and target variables (ProductCategory, ProductBrand).
   * Split the data into training and testing sets for both prediction tasks.
   * Train Random Forest Classifiers to predict product category and brand.
5. **Prediction Function:**
   * Develop a function that takes customer age, customer gender and price range as inputs and returns the predicted product category and brand.

**Why This Project:**  I chose this project because it allows me to apply both Exploratory Data Analysis (EDA) and Machine Learning techniques. This will help me revise the previous module and practice new skills.

Muhammad Shaffan