POST DATA ANALYSIS PRESENTATION CONTENT/ FRAMEWORK (in progress)

1. **Data Collection**

   Personal financial data: income, expenses, credit history, existing debts.
   Behavioral data: payment habits, account activity.
   Alternative data (is it possible to get alternative data): utility payments, mobile transactions.

2. **Data Cleaning and Preprocessing**

   Handle missing or inaccurate data.
   Standardize variables (e.g., normalize income levels, create categorical variables).

3. **Feature/ Attribute Engineering (from my Python programming course) (could these be research questions/answers)**

   Create derived variables: debt-to-income ratio, credit utilization, and loan-to-value ratio.
   Incorporate historical trends: late payment frequency, credit growth.
   Credit Approvals
   Credit App Rejections

4. **Model Development**

   Logistic Regression for simplicity.
   Decision Trees/Random Forests for interpretability.

5. **Model Evaluation**

   Metrics: Precision
   Validation: Cross-validation, backtesting on historical data.

6. **Deployment?**

   Integrate with credit decision systems?
   Monitor for model drift or changes in borrower behavior?

7. **Post-Deployment Monitoring?**

   Periodically assess model performance?
   Regularly retrain using updated data?