

Business Question

- Problem: Increased default rates on loans from Credit One clients is having a negative impact on the company's reputation.
- We need a new method of determining credit worthiness.
- At minimum, we need a way to determine the probability of one of our client's customers defaulting on a loan.

Analysis Plan

- The goal of this analysis will be to determine if one factor or multiple factors in a customer's data profile is predictive of whether that customer is likely to default.
- We will evaluate parameters such as customer age, billing and payment history, education level, marital status, etc. as possible predictors.
- Machine Learning models will be created and evaluated for predictive accuracy.
- Project plan will be to report initial findings and provide insights to management as they become available.

Data Collection

- Data will be sourced from external customer database.
- Data Science team with leverage data cleansing techniques using Python to prepare the data for analysis and modeling.

Insights

- Initial data exploration will be to determine if there is any meaningful pattern that can be derived from a cursory review of the data.
- Further statistical analysis will be performed and any findings that are relevant to the business problem will be provided to management.
- Reliability of machine learning models will be evaluated for suitability to support our business case.

Recommendations

- Upon completion of the Data Science project, the team will provide a formal recommendation to management with the following:
 - Basis for recommendation if definitive results from machine learning models can be achieved
 - Plan for implementation of new business process to determine credit risk

Data Science Process

