

Team 09 Progress Report

Project: Credit Card Default

Members: Jesus Gomez, Jammy Loeur

Abstract:

The term project, “Credit Card Default”, will strive to estimate the probability people will default on credit card payments. Our team will use various classification models to determine which attribute, or set of attributes, most likely contributes to a person defaulting on a payment.

About the Dataset

The Credit Card Default dataset contains 30,000 records of customer information for a Taiwanese bank. The customer data included in the dataset are education, gender, age, marriage status, and the payment actions for a six-month period.

Source: UCI Machine Learning Repository

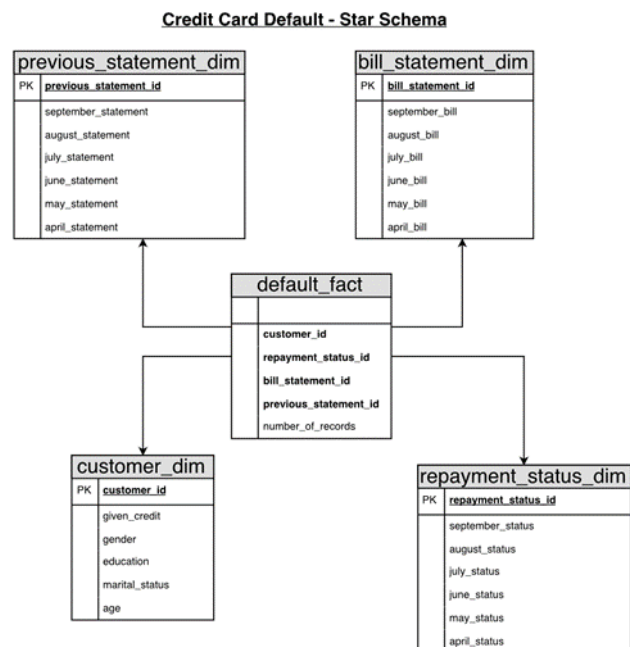
<https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients>

Changes from Project Proposal

Since our project proposal presentation, we have decided to focus on a singular dataset for both the data mart and data mining instead of the separate datasets we initially proposed. Also, we have some deviation for our initial starting date of the project due to class work and studies; but the date was set to allow room for such occurrences for our term project. Other than that, no other deviations occurred.

Current State of Project

Team 09 has been heavily engaged in learning the applications and tools needed for our DW. We have been utilizing the provided RStudio tutorials from SacCT as well as using online resources to aid in our development of our project. We are currently in preparing our dataset for use in a MySQL database for use in our data mart and expect to be finished by week 12.



Data Warehouse Questions to Answer:

- What effect does gender, age, or education have the likelihood that customers will default on their payments?
- What combination of customer characteristics will maximize the default chance? which minimize the default chance?
- Will my friend pay me back?

Schedule for Remaining Tasks

Week 11: Complete any last-minute preprocessing, work on creation of tables for DW

Week 12: Complete data mart, start analyzing the data for data mining project.

Week 13: Complete data mining project by end of week.

Week 14: Complete any last-minute stuff, create presentation slides

Presentation Day Preference: Week 15