Abstract

My team (IC22046) will be conducting our data analysis on data sets relating to the Payment Protection Program. The basic questions at hand were to identify why these loans were removed and how they relate to all the other existing loans.

The Payment Protection Program is a program governed by the United States government in supporting companies, mainly small businesses. One of the main goals was to keep employment running and stop the business from falling apart. Many loans were removed from the original data sets due to them being either withdrawn by the applicants, having problems with the application process or fraudulent data.

My team brainstormed ideas and methods on how we should conduct the project. I used the data given as well as outside sources. Data manipulation was key as we tried to make sense of why the loans were removed. Excel was our main program that we used to summarize the data and analyze them. We started by merging certain data sets. I also took certain values individually and compared them to see any specific characteristics. I also used visual graphs as support to answer the questions.

Through thorough data analysis we can figure out the main reasons for these loans to be removed and use our data analysis to predict characteristics of these loans.