BUS 696

Final Project Outline

a) identify a dataset you will use

We will be using the Google Play Store Apps data found on Kaggle. The dataset can be found at <https://www.kaggle.com/lava18/google-play-store-apps>

b) the outcome you are trying to predict, and what variables you will use to predict it

Number of user downloads/installs will be our dependent variable and the variable we are trying to predict. Our initial independent variables include Category, Rating, Reviews, Size, Type (Free or Paid), Content Rating, and Genre. We will also transform the variables as we see fit as we proceed with the project.

c) motivation to your project -- as in the business or practical management use case of such a prediction

We want to discover the factors that make a popular app. We also want to discover if smaller sized apps are desired.

d) three methods you will use to analyze your question of interest

We will be using linear regression, elasticnet, and up-sampling when the data points are sparse.

e) the names of the students who will be part of your group

Ranjith, Theresa, Ken