

Originally when I had submitted my project specification, I had no intention of performing regression analysis on the data set. I wanted to create a point system based on the data entries for each company that would sum over all reviews for a company. I wanted to simply find out which company was the best through this method. This had absolutely no optimization other than trying to find out which company had the best score. After talking with Prof. Ferris, I decided to change my project to fit more in line with what we had been asked to do. With Prof. Ferris' suggestion, I decided to move forward with a classification problem.

I decided to drop the point system and use the 'overall_rating' that was already provided inside the csv file since that should be a good indicator for how the passenger felt about the trip. I also decided that 'recommended' would also be an important indicator since that means they would recommend that flight to others.