DSC 530 Final Project

Matthew Litschewski

Bellevue University

Aug 6, 2020

The question that I decided to ask of my dataset was can we predictive a positive outcome given the information in the data set regardless of which campaign it was for. I looked manly at some of the demographics, balance and time that each respondent was engaged by bank personnel. It is not so much that I missed it but thought it might be irrelevant given that a high percentage of the campaigns used were 1-3 even though they had campaigns that numbered up to the 60s. It might have been better to look at each campaign separately and evaluate the efficacy of the campaign itself. It really isn’t clear if any other data would help with the analysis but then again some time you don’t know what you don’t know. The big assumption I made and already admitted to was not slicing out each campaign separately. I feel that maybe there would have been better correlations that way and not lose the strength of that correlation by diluting it with noise from the other campaigns. The big hurtle was computing PMF and figuring how to use Allen Downey’s code. There really isn’t a library I found that has a built in PMF function I found that surprising. The other thing that I am still struggling with is the null hypothesis. This something that will just take time and more practice with to have the concept really stick in my head. The other concept that tripped me up a bit was p-value and R^2 but I feel like I am gaining a deeper understanding the more I work with them.