Project Proposal: MATH 531 Regression I
Due: Tuesday, Oct 23, 2018

Sara Bey, Wei Yang, Jayanth Sivakumar

October 22, 2018

Questions: What factors contribute most to our happiness? Is it money, sex, love, or work? Or, is it a combination of these factors? And, how do each of these factors, and combinations of them, contribute to our happiness? Beyond a certain point, say 100k, does any additional income generate more happiness? Or, is it possible that having too much money can negatively effect our rating of happiness?

Data Set: We will use the dataset happy, in the faraway package, to carry out our analysis. This data set is a collection of data from 39 students in the MBA class at University of Chicago. Data about the students' happiness, money, sexual activity, love and work are provided. Each student rates their happiness on a scale of 1 to 10. Money is measured by looking at family income of the student in thousands of dollars. Sex is measured by responding "1" for "satisfactory sexual activity" and "0" for "not" satisfactory sexual activity. Love is measured by responding "1" for "lonely", "2" for "secure relationships" or "3" for "deep feeling of belonging and caring". Work is measured on a 5 point scale with "1" representing "no job", "3" representing "OK job" and "5" representing "great job". Overall, there are 39 observations with 5 variables.

Hypothesis: We are happier when we have more money, satisfactory sexual activity, deeper feelings of belonging and caring and a great job. However, past a certain earning, money has a diminishing effect on that happiness.

Analysis: We will use nested F-testing. In doing so, we will determine which variable, sex, love, money or work, has the highest association with an individual student's rating of their happiness. We will then test to see whether there is some combination of 2 variables that increases this association. Last, we will determine if our happiness is given the highest rating when all variables are considered. This will determine which of the variables, or combinations, have the highest association with the students' rating of happiness. Then we will investigate whether earnings, past a certain point, have a negative impact on the rating of happiness. We hypothesize that there will be a high correlation between the variable sex and love, as well as between the variables money and work. This high correlation can be detected using the Variance Inflation Factor (VIF). If this is indeed the case, and the VIF values prove to be above the threshold of 4, we must proceed with caution. If the VIF values are above 10, we need to take action to remove the multicollinearity that exists between the variables. In doing so, we can get a more accurate model for predicting happiness based on the other variables of love, work, sex and money.

Rationale: There are many people that say "money can not buy you happiness". Although this may be true that money can not physically buy you happiness, we want to discover if there is evidence to suggest that having more money does in fact have an association with a higher rating of happiness. We wish to investigate this relationship between happiness and each of the other four variables, and combinations of them. We hypothesize that there is a strong correlation between money and work, as those who have a great job tend to be earning more, which increases the family income level. We hypothesize that past a certain earnings, the effect on the happiness rating due to money may not be as great. This could be that there is a certain level of income that someone is no longer worried about money, but more so about the other variables, such as close relationships, amongst other factors. We also believe there to be a strong correlation between love and sexual activity since someone in a sexually active relationship tends to feel deeper connections and have a stronger relationship.