

Math 32 Course Project Proposal

by Armaan Kapoor

October 20, 2019

1. My data set that I will be using is a data frame of 53940 rows and 10 columns of diamonds.
2. I chose this data set because I find this to be interesting. I thought that diamonds had this standard equation that they would use to determine its worth. But this data set tells me that diamonds not only come in so many different shapes, sizes, and other qualities that it is difficult to determine the actual worth of a diamond. Especially with fluctuating economies across countries and the hemispheres, diamonds can be worth very different even though they have very similar characteristics. I hope this data set will be used to come to a conclusion of this problem.
3. I will use this data set to ultimately formulate an equation or some type of graph that can be used to predict the price of a diamond. I will use the price category in all of my analysis (Y-axis). I will test different categories individually as well as together (evenly split weight?) and see what kind of distribution do they hold against the price. Finally, I will determine the weight of each category to develop the graph of predicting the price of the diamonds.