Let's say we work at a credit card company and we're trying to figure out if people are going to pay their bills on time. We have everyone's purchases, split into four main categories: groceries, dining out, utilities, and entertainment. What are some ways you might use KNN to create this model? What aspects of KNN would be useful? Write up your thoughts in submit a link below.

I would use all four variables to build the model because they are a unique variable, but if for example variables groceries and dining out are strongly correlated than I would drop one of the variables. Multicolinearity causes overfitting and it would be redundant to include a variable that does not add to the information. Since each variable has the same unite, I do not need to correct for it. There are only four variables, more chance of neighbors not being close together so it would be good idea to weighting the neighbors.

KNN is build to answer yes/no type of questions. Its simple to understand, use and implement and very effective. We don’t need training and testing data, all the data can be used for building the model.