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.

To use this data most efficiently with KNN, I would first perform PCA and try and come out with 2 components to represent this data set. The predicted variable will be categorical, whether or not the credit card holder paid his/her bill. Assuming PCA, can give us two components that represent most of the variance, I would plot those two points in a scatter plot denoting which paid the bill, and which did not. KNN would now be useful to predict whether or not a credit card holder will pay his/her bill since the dimension of this data set is now manageable.