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.

One way to use KNN here is to normalize the data. I would rescale the variables for groceries, dining out, utilities, and entertainment, so that they are between 0 and 1. Another way would be to calculate the z-score for the variables.

A useful aspect of KNN is the k-value. Since we have four categories, I would pick a lager k. Having a lager k-value means more data point are used in our prediction for on time payment. With a smaller k-value I would be afraid of overfitting.