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 tackle this task would be to use expenditure in each of these categories as a dimension. Default (non-payment of the bill) would then be the classifying criterion, e.g. has the borrower been late at least once in the last 12 months.

Another approach would be to use percentage of one’s total spending as the dimension, so that for a given person the breakdown could be: 20% groceries, 15% utilities, 15 % dining out, and 20 % entertainment (and the rest would be ‘other.’) There could be useful insights to be gained based on the proportion of one’s income that is spent on the respective categories. This could be even more useful if household income were added into the analysis, which is probably why credit card companies ask this question of their borrowers.