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.

First, I might generate a column that includes the total bill to see if that predicts whether people will pay on time. Larger bills might be paid later. KNN would then be used to fit continuous purchase features with late and on-time payer categories. The k variable could be tested at multiple sizes to see where the time to complete the program and the accuracy are both optimal. The distances between a test and train data with known features and classification as either late or on-time payer could be used to test the accuracy by dividing total – failed predictions by the total predictions.