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 solve this problem, we are trying to classify the spending behavior of the credit card companies customers.Spending behavior would be the ratio of the type of purchases overtime. Our outcome of interest is yes, they will pay on time, or not, they will not. From our dataset containing everyone's purchases, we can create a variable that takes a ratio of all of the purchases of the different classes. Next, we arrange an additional dataset of our outcome variable, whether customers of the credit card company have paid their bills on time that corresponds to their purchasing data.

KNN is useful because a single nearest data point does not determine the outcome, but rather multiple nearest data points, which is very important considering the variance among spending behavior and in the prevention of overfitting. From trial and error, or k tuning, we will find the proper k for the model.