# 03 Prove: kNN with More Interesting Data

1. Please provide the URL of your public GitHub repository.

<https://github.com/johnmwood/CS450>

1. What was the most difficult part of this assignment?

Figuring out the data so it worked with our existing classifier.

1. Briefly describe your process for handling non-numeric data.

We were able to handle the categorical data by filling them in with numeric values. We decided not to tackle the car names but we figured we could do it by manufacturer, however, that would fall under the assumption that the data was the same throughout the entire column.

1. Briefly describe your process for missing data.

We used df.replace() a lot to fill in column values by passing in a dictionary of the keys and values we need it.

1. Describe your results for the Car Evaluation dataset.

We got an 86% accuracy score on the cars dataset with a 10 fold CV.

1. Describe your results for the Pima Indian dataset.

We only got a 67% CV score from the diabetes dataset, which is lower than we thought we should get. We haven’t been able to explore specifically why it might be so low.

1. Describe your results for the Automobile MPG dataset.

For the mpg dataset, we only received a 54%, which was again very low. We don’t know why it was so low however, we still feel confident that the data was cleaned well. Possibly some of the categorical values could be improved.

1. Describe anything you did to go above and beyond the minimum standard requirements.

We felt proud of abstracting out our classes to reduce the amount of stress on the classifier and the shell. This helped keep the code clean and easy to use in the shell.

1. Please select the category you feel best describes your assignment:

Shows creativity and excels above and beyond requirements

1. Provide a brief justification (1-2 sentences) for selecting that category.

I feel like we really thought about our classes in a way that would make it easy for someone else to use our code. Although not perfect, I would say that we were definitely creative and went beyond what was expected.