K-Means Hands-On

In this Hands-On exercise, you will create a project that will solidify your understanding of *k*-means and *k*-nearest neighbors. This Hands-On will be completed in Python, using your text editor or IDE of choice (e.g. VSCode, Jupyter Notebooks, Spyder, etc.).

Determine how cars are grouped together by using the `mpg` dataset built into Seaborn. Import it using the following code:

```
Mpg = sns.load_dataset('mpg')
```

If seaborn isn't working for you, **click here** to download the data.

Remember that you need continuous variables for these analyses, so you'll want to pinpoint columns such as `mpg`, `cylinders`, `displacement`, `horsepower`, `weight`, `acceleration` or `model_year` as variables. You'll also need to have those continuous variables as integers...so...hint, hint...there may be a little data wrangling involved.

Then use first the *k*-means machine learning algorithm and then the *k*-nearest neighbors algorithm to find the most appropriate *k* to examine, and provide the graph as well as add the cluster labels back into your dataframe. How are these groups being divided? What conclusions can you draw about the data?