Final Writeup

Aiden Morrison

SECTION J

1. Description of Data Set:
   a. I am using the Car Information dataset. It is a basic dataset including name, miles per gallon, number of cylinders, displacement, horsepower, weight, acceleration, model year, and country of origin for a variety of vehicles.

2. Question Answered:
   a. I want the data to connect the correlation of number of cylinders to the fuel economy of a vehicle. The data can do this by performing a correlation analysis to determine this relationship. I can then create a few different visualizations to pain a clear picture of what this correlation is and communicate it in an effective manner. I am also doing a similar correlation to horsepower against acceleration.

3. Instructions to Run Your Project:
   a. Run command 'python submission.py'

4. Description of Required Input Format:
   a. A csv file named 'Car_Information.csv'. 9 different columns of data from 'name', 'mpg', 'cylinders', 'displacement', 'horsepower', 'weight', 'acceleration', model_year', and 'origin'.

5. List any extra modules imported by your program (if relevant):
   a. Pandas

6. A snippet of important code from your program and an explanation of what it does:
   a. The two functions 'plot_cylinders_vs_mpg' and 'plot_horsepower_vs_acceleration' create the two scatter plots. The first two lines import two libraries that make the code more efficient.

```python
1   import pandas as pd
2   import matplotlib.pyplot as plt
3
4   # read the csv file
5   data = pd.read_csv('Car_Information.csv')
6
7   def plot_cylinders_vs_mpg():
8       plt.figure(figsize=(8, 6))
9       plt.scatter(data['cylinders'], data['mpg'])
10      plt.title('Number of cylinders vs MPG')
11      plt.xlabel('Number of cylinders (qty)')
12      plt.ylabel('MPG (miles per gallon)')
13      plt.grid(True)
14      plt.savefig('cylinders_vs_mpg.png')  # save the plot as a png
15
16  def plot_horsepower_vs_acceleration():
17      plt.figure(figsize=(8, 6))
18      plt.scatter(data['horsepower'], data['acceleration'])
19      plt.title('Horsepower vs acceleration')
20      plt.xlabel('Horsepower (hp)')
21      plt.ylabel('Acceleration (m/s^2)')
22      plt.grid(True)
23      plt.savefig('horsepower_vs_acceleration.png')  # save the plot as a png
```

7. Description of a User-Defined Function:
   a. The function 'plot_cylinders_vs_mpg' is a good example. While it doesn't explicitly take input parameters, it works with data based on the imported csv file. The function reads the csv file, uses the 'plt.scatter' function to generate a scatter plot, labels the x and y axis with the corresponding values, establishes a title for the plot, uses 'plt.grid(true) to display gridlines, and then saves the generated plot as a png titled 'cylinders_vs_mpg.png'. Ultimately, the output of this function is the png showing the plot.

8. Description of your data visualization:
   a. Two scatterplots. The first plot shows Number of Cylinders vs Miles per Gallon, the second one shows Horsepower vs Acceleration.

9. A copy of the actual visualization generated by your program, saved and embedded in this report.

Number of cylinders vs MPG

Horsepower vs acceleration