SUMMER ANALYTICS 2024

## Week-1 Assignment

## Data Grand Prix!

**Welcome to your first assignment of Summer Analytics 2025! We hope you are excited to implement and test everything you have learnt up until now. The dataset which you'll use includes information about cars.**

**We've got an interesting set of questions for you to get a basic understanding of pandas and data visualization libraries. GOOD LUCK!**

***Let's get started with importing numpy, pandas, seaborn and matplotlib!***

Note - matplotlib should be imported with the command :

```
import matplotlib.pyplot as plt
```


**So lets get started!! Buckle up your belts for this exciting ride!!**

## 1) Start by importing all important libraries

For eg, "import numpy as np"

```
import matplotlib.pyplot as plt
import numpy as np
```

## 2) Read the csv file and assign it to a variable .

```
#your code here
```

## 3) Display shape of dataframe

Expected Output - (398, 9)

+ Code + Text

```
#your code here
```

## 4) Print all columns of dataframe

Return an array containing names of all the columns.

```
#your code here
```

### 6) Set the 'name' column as the index of dataframe

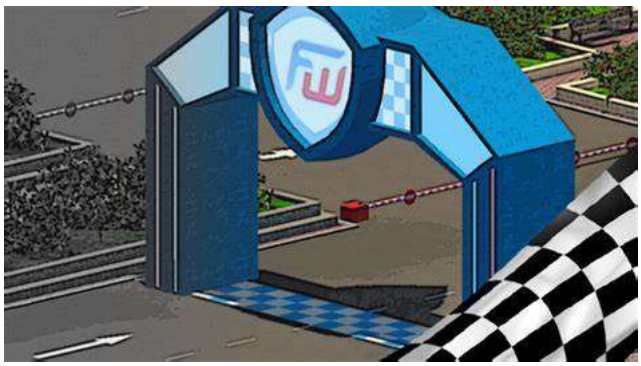
```
#your code here
```

### 7) Print a list of all the unique mpg values

```
#your code here
```

### 8) Create a column which contains the horsepower divided by weightas its metric and make this new column the index.

```
#your code here
```



**Checkpoint!! Congratulations on making it this far. You are really keeping up in Data Grand Prix. Now starts the real race i.e. graded questions of the quiz.**

## GRADED Questions (To be answered in the quiz)

Try to retrieve some information from the data and answer the questions below . BEST OF LUCK !!

### 1. What is name of car that has the highest horsepower?

```
#your code here
```

### 2. How many cars have mpg ≥ 35?

```
#your code here
```

### 3. What is the most common origin for cars with horsepower > 100 and weight < 3000?

```
#your code here
```

### 4. What is the mean acceleration of cars from Japan? (rounded to 2 decimals)

```
#your code here
```

5. Which year had the highest average mpg?

```
#your code here
```

*Congratulations on coming this far! Since we were having so much fun playing with this dataset, let's move towards finish line by attempting some Ungraded questions!*

**Note:** These questions are UNGRADED, and are given as an extra exercise.

Find the car (or cars) with the best ratio of horsepower to weight among all cars that also have above-median mpg.

```
#your code here
```

Design a multi-line plot using Matplotlib or Seaborn that shows the evolution of average mpg over the years, separately for each origin

```
#your code here
```

Create a Seaborn scatterplot (or PairGrid) where:

X = horsepower

Y = weight

Color by: origin

Size by: mpg

Hue order = ['japan', 'europe', 'usa']

Add meaningful plot titles and axis titles.

```
#your code here
```

We define a “consistent” car model as one that was produced over multiple years and had very low variation in mpg across those years (standard deviation < 1.0).

Tasks:

Identify car names that appear in more than one model_year.

For each such name, compute the standard deviation of mpg across years.

Return the car(s) with the lowest variation in mpg, among those with at least 2 appearances and std(mpg) < 1.0.

Report the model name(s), number of appearances, and the average mpg.

Bonus: Sort the result by number of appearances (descending), then mpg (descending).

```
#your code here
```