# First Jupyter notebook

July 12, 2024

## 1 This is a heading

This is a text cell. Data scientists use text cells in a Jupyter notebook to write comments about a dataset or notes about findings. Double click this Markdown cell to change the text, and click Run to format the text.

### 1.1 This is a smaller heading

Text cells can contain:

- Bulleted lists
- Like this

Or

1. Numbered lists
2. Like this

[1]:
```
# This is a code cell
# In Python, hashtags are used to write comments, or text
# It's good practice to use comments to write notes above a function or command
```

[2]:
```
# Import seaborn library
import seaborn as sns

# Load the miles per gallon dataset and assign the dataset to mpg
mpg = sns.load_dataset('mpg')

# View the dataset (by default, the first and last five rows are displayed)
mpg
```

[2]:

| | mpg | cylinders | displacement | horsepower | weight | acceleration | \ |
|---|---|---|---|---|---|---|---|
| 0 | 18.0 | 8 | 307.0 | 130.0 | 3504 | 12.0 | |
| 1 | 15.0 | 8 | 350.0 | 165.0 | 3693 | 11.5 | |
| 2 | 18.0 | 8 | 318.0 | 150.0 | 3436 | 11.0 | |
| 3 | 16.0 | 8 | 304.0 | 150.0 | 3433 | 12.0 | |
| 4 | 17.0 | 8 | 302.0 | 140.0 | 3449 | 10.5 | |
| .. | … | … | … | … | … | … | |
| 393 | 27.0 | 4 | 140.0 | 86.0 | 2790 | 15.6 | |
| 394 | 44.0 | 4 | 97.0 | 52.0 | 2130 | 24.6 | |

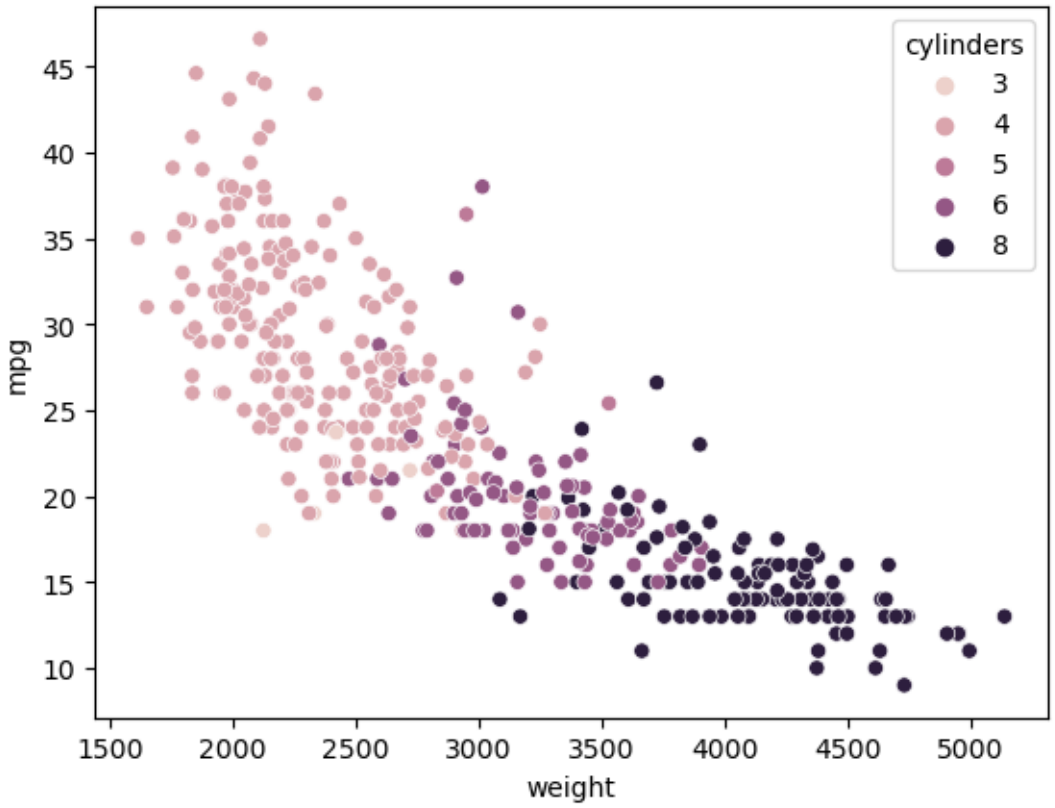
```
395  32.0          4         135.0        84.0    2295          11.6
396  28.0          4         120.0        79.0    2625          18.6
397  31.0          4         119.0        82.0    2720          19.4

     model_year  origin                       name
0            70     usa  chevrolet chevelle malibu
1            70     usa          buick skylark 320
2            70     usa         plymouth satellite
3            70     usa              amc rebel sst
4            70     usa                ford torino
..          …     …                          …
393          82     usa            ford mustang gl
394          82  europe                  vw pickup
395          82     usa              dodge rampage
396          82     usa                ford ranger
397          82     usa                 chevy s-10

[398 rows x 9 columns]
```

[3]:
```
# Make a scatter plot with weight on the x-axis and miles per gallon on the␣
  ↪y-axis
sns.scatterplot(data=mpg, x='weight', y='mpg', hue='cylinders')
```

[3]:
```
<Axes: xlabel='weight', ylabel='mpg'>
```




Based on the scatter plot, cars that are heavier (higher weight) have lower miles per gallon. Cars with larger engines, or more cylinders, also have lower miles per gallon.

[4]:
```
# Calculate summary statistics for the weight feature
mpg['weight'].describe()
```

[4]:
```
count     398.000000
mean     2970.424623
std       846.841774
min      1613.000000
25%      2223.750000
50%      2803.500000
75%      3608.000000
max      5140.000000
Name: weight, dtype: float64
```