

# Part A

## Importing the dataset

```
import pandas as pd

#Import the file from github using the link given
url =
"https://vincentarelbundock.github.io/Rdatasets/csv/datasets/mtcars.csv"
mtcars = pd.read_csv(url)
```

## Displaying the head of the top 5 rows

```
mtcars.head()
```

		rownames	mpg	cyl	disp	hp	drat	wt	qsec	vs
am	gear	\								
0		Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0
1	4									
1		Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0
1	4									
2		Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1
1	4									
3		Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1
0	3									
4		Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0
0	3									

```
carb
```

0	4
1	4
2	1
3	1
4	2

## Calculate the Average MPG for Each cyl Value:

```
avgMpgByCyl = mtcars.groupby('cyl')['mpg'].mean()

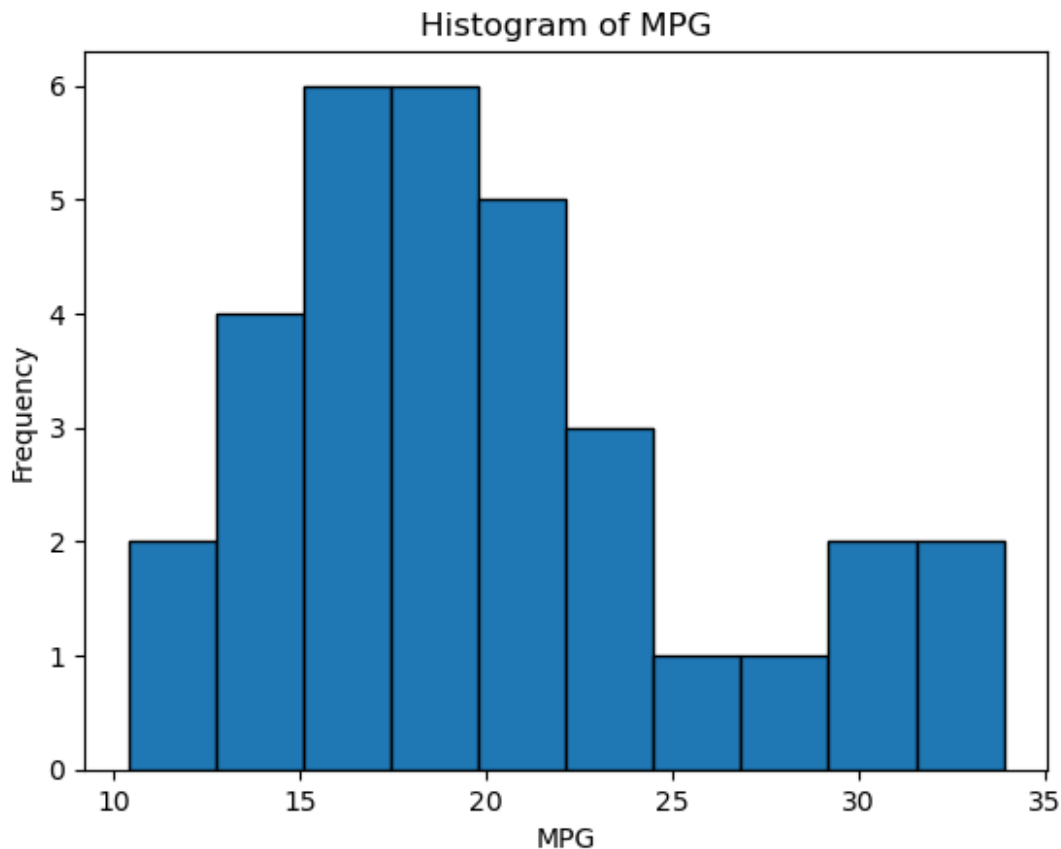
avgMpgByCyl

cyl
4    26.663636
6    19.742857
8    15.100000
Name: mpg, dtype: float64
```

## Create a Histogram

```
import matplotlib.pyplot as plt

plt.hist(mtcars['mpg'], bins=10, edgecolor='black')
plt.title('Histogram of MPG')
plt.xlabel('MPG')
plt.ylabel('Frequency')
plt.show()
```



## Part B

The chosen dataset is 'Road Casualties in Great Britain 1969–84' which contains the time series giving the monthly totals of car drivers in Great Britain killed or seriously injured Jan 1969 to Dec 1984.

Contains the following variables: DriversKilled: car drivers killed, drivers: same as UKDriverDeaths, front: front-seat passengers killed or seriously injured, rear: rear-seat passengers killed or seriously injured, kms: distance driven, PetrolPrice: petrol price, VanKilled: number of van ('light goods vehicle') drivers, law: 0/1: was the law in effect that month?

## Citation

Harvey, A. C. and Durbin, J. (1986). The effects of seat belt legislation on British road casualties: A case study in structural time series modelling. Journal of the Royal Statistical Society series A, 149, 187–227. [doi:10.2307/2981553](https://doi.org/10.2307/2981553).

```
seatbelt =  
"https://vincentarelbundock.github.io/Rdatasets/csv/datasets/Seatbelts  
.csv"  
ukDrivers = pd.read_csv(seatbelt)
```

## Displaying the head of the dataset

```
ukDrivers.head()
```

	rownames	DriversKilled	drivers	front	rear	kms	
	PetrolPrice	\					
0	1	107	1687	867	269	9059	0.102972
1	2	97	1508	825	265	7685	0.102363
2	3	102	1507	806	319	9963	0.102062
3	4	87	1385	814	407	10955	0.100873
4	5	119	1632	991	454	11823	0.101020

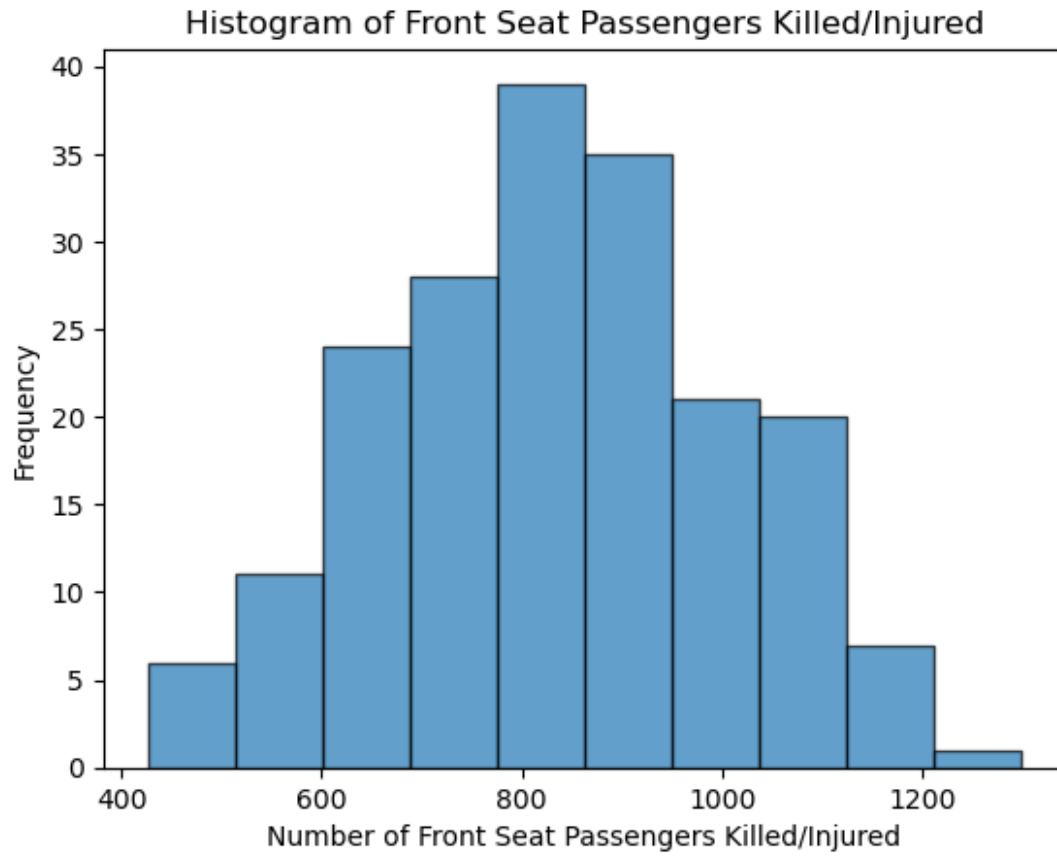
  

	VanKilled	law
0	12	0
1	6	0
2	12	0
3	8	0
4	10	0

## Visualizing the dataset

### Scatter Plot of petrol price vs Drivers killed

```
plt.hist(ukDrivers['front'], bins=10, edgecolor='black', alpha=0.7)  
plt.title('Histogram of Front Seat Passengers Killed/Injured')  
plt.xlabel('Number of Front Seat Passengers Killed/Injured')  
plt.ylabel('Frequency')  
plt.show()
```



## Rear Seat Passengers Killed/Injured

```
plt.hist(ukDrivers['rear'], bins=10, edgecolor='black', alpha=0.7, color='orange')
plt.title('Histogram of Rear Seat Passengers Killed/Injured')
plt.xlabel('Number of Rear Seat Passengers Killed/Injured')
plt.ylabel('Frequency')
plt.show()
```

