+ Code + Text

```
import pandas as pd
```

```
df=pd.read_csv("weight-height.csv")
```

```
df.head()
```

	Height	Weight
0	73.847017	241.893563
1	68.781904	162.310473
2	74.110105	212.740856
3	71.730978	220.042470
4	69.881796	206.349801

```
df=df.drop("Gender", axis=1)
```

```
from scipy import stats
```

```
import numpy as np
```

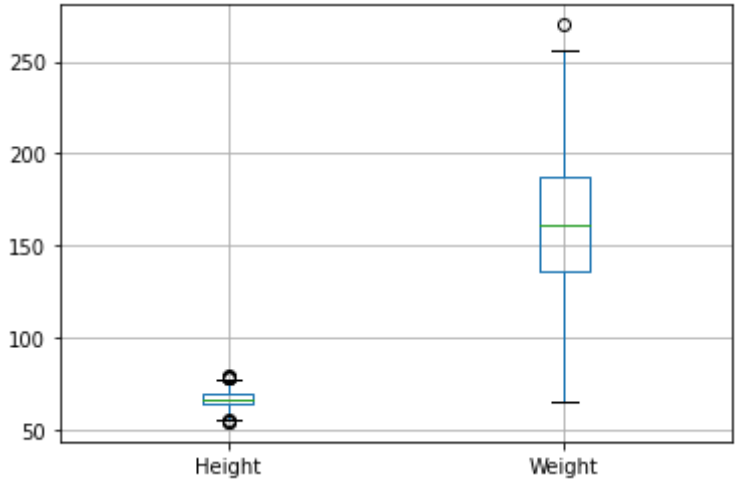
DATASET OUTLIERS

```
z=np.abs(stats.zscore(df))
```

```
df1=df.copy()
```

```
df.boxplot()
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fcb9d0e0550>
```



```
from scipy import stats
import numpy as np


df2=df.copy()


q1 = np.percentile(df, 25, interpolation = 'midpoint')
q2 = np.percentile(df, 50, interpolation = 'midpoint')
q3 = np.percentile(df, 75, interpolation = 'midpoint')

print('Q1 25 percentile of the given data is, ', q1)
print('Q1 50 percentile of the given data is, ', q2)
print('Q1 75 percentile of the given data is, ', q3)

IQR = q3 - q1
print('Interquartile range is', IQR)
```

```
     Q1 25 percentile of the given data is,  66.3178274068096
     Q1 50 percentile of the given data is,  77.72287280637059
     Q1 75 percentile of the given data is,  161.21292769948298
     Interquartile range is 94.89510029267338
```

Double-click (or enter) to edit

```
low_lim = q1 - 1.5 * IQR
up_lim = q3 + 1.5 * IQR
print('low_limit is', low_lim)
print('up_limit is', up_lim)
```

```
     low_limit is -76.02482303220047
     up_limit is 303.55557813849305
```

```
df2_new=df2[((df2>=q1-1.5*IQR)&(df2<=q3+1.5*IQR)).all(axis=1)]




df2_new.boxplot()
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fcb9aa2fd10>
```