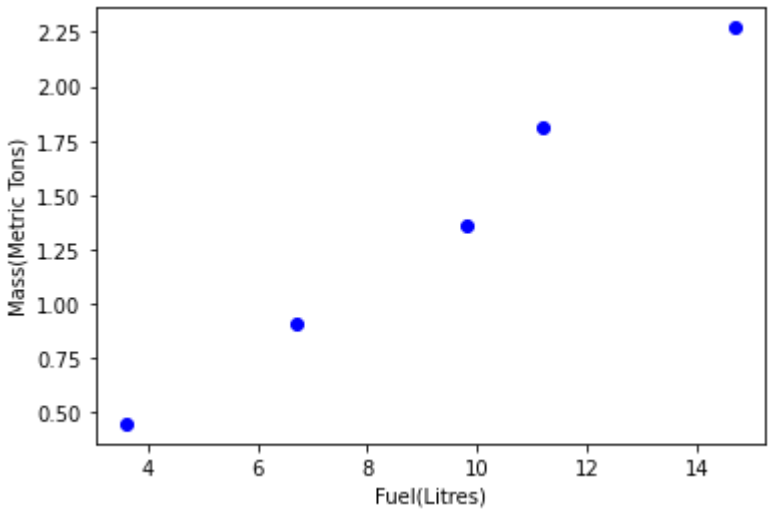
Question 4 - A car company wants to predict how much fuel different cars will use based on their masses. They took a sample of cars, drove each car 100km, and measured how much fuel was used in each case (in litres). Visualize the data using scatterplot and also find co-relation between the 2 variables (eg. Positive//Negative, Linear/ Non- linear co-relation) The data is summarized in the table below. (Use a reasonable scale on both axes and put the explanatory variable on the x-axis.) Fuel used (L) 3.6 6.7 9.8 11.2 14.7 Mass (metric tons) 0.45 0.91 1.36 1.81 2.27

In [ ]:
```
from matplotlib import pyplot as plt
import pandas as pd
import numpy as np
from scipy.stats import pearsonr
```

In [ ]:
```
fuel = [3.6, 6.7, 9.8, 11.2, 14.7]
mass = [0.45, 0.91, 1.36, 1.81, 2.27]
```

In [ ]:
```
plt.scatter(fuel, mass, c="blue")
plt.xlabel("Fuel(Litres)")
plt.ylabel("Mass(Metric Tons)")
plt.show()
```



In [ ]:
```
Correlation, _ = pearsonr(fuel, mass)
```

In [ ]:
```
Correlation
```

Out[ ]:
```
0.9938681082455859
```

The correlation between the 2 variables is positive. As the correlation coefficient is very close to 1, this suggests a highly linear relationship.