

```
In [2]: import numpy as np
import pandas as pd

# Data Visualisation
import matplotlib.pyplot as plt
```

```
In [3]: housing = pd.DataFrame(pd.read_csv("Housing.csv"))
housing.head()
```

```
Out[3]:
```

| | price | area | bedrooms | bathrooms | stories | mainroad | guestroom | basement | hotwaterheating |
|---|----------|------|----------|-----------|---------|----------|-----------|----------|-----------------|
| 0 | 13300000 | 7420 | 4 | 2 | 3 | yes | no | no | no |
| 1 | 12250000 | 8960 | 4 | 4 | 4 | yes | no | no | no |
| 2 | 12250000 | 9960 | 3 | 2 | 2 | yes | no | yes | no |
| 3 | 12215000 | 7500 | 4 | 2 | 2 | yes | no | yes | no |
| 4 | 11410000 | 7420 | 4 | 1 | 2 | yes | yes | yes | no |

```
In [4]: m = len(housing)
m
```

```
Out[4]: 545
```

```
In [5]: housing.shape
```

```
Out[5]: (545, 13)
```

```
In [8]: num_vars = ['area', 'bedrooms', 'bathrooms', 'stories', 'parking', 'price']
Newtrain = housing[num_vars]
Newtrain.head()
```

```
Out[8]:
```

| | area | bedrooms | bathrooms | stories | parking | price |
|---|------|----------|-----------|---------|---------|----------|
| 0 | 7420 | 4 | 2 | 3 | 2 | 13300000 |
| 1 | 8960 | 4 | 4 | 4 | 3 | 12250000 |
| 2 | 9960 | 3 | 2 | 2 | 2 | 12250000 |
| 3 | 7500 | 4 | 2 | 2 | 3 | 12215000 |
| 4 | 7420 | 4 | 1 | 2 | 2 | 11410000 |

```
In [ ]:
```