In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn import linear_model
```

Creating a Data Frame

In [2]:
```
raw_data = pd.read_csv('raw_data.csv')
```

In [3]:
```
raw_data
```

Out[3]:

	area	bedrooms	age	price
0	2600	3.0	20	550000
1	3000	4.0	15	565000
2	3200	NaN	18	610000
3	3600	3.0	30	595000
4	4000	5.0	8	760000

In [4]:
```
raw_data.isnull()
```

Out[4]:

	area	bedrooms	age	price
0	False	False	False	False
1	False	False	False	False
2	False	True	False	False
3	False	False	False	False
4	False	False	False	False

In [5]:
```
raw_data.isnull().sum()
```

Out[5]:
```
area        0
bedrooms    1
age         0
price       0
dtype: int64
```

Managing Null Values (Data Preprocessing)

In [6]:
```
import math
```

In [7]:
```
median_bedrooms = math.floor(raw_data['bedrooms'].median())
```

In [8]:
```
median_bedrooms
```

Out[8]:
```
3
```

In [9]:
```
raw_data['bedrooms'].fillna(median_bedrooms)
```

Out[9]:
```
0    3.0
1    4.0
2    3.0
3    3.0
4    5.0
Name: bedrooms, dtype: float64
```

In [10]:
```
raw_data['bedrooms'] = raw_data['bedrooms'].fillna(median_bedrooms)
```

In [11]:
```
raw_data
```

Out[11]:

	area	bedrooms	age	price
0	2600	3.0	20	550000
1	3000	4.0	15	565000
2	3200	3.0	18	610000
3	3600	3.0	30	595000
4	4000	5.0	8	760000

Creating a linear regression model

In [12]:
```
lin_reg = linear_model.LinearRegression()                # Creating linear regre
ssion object
```

In [13]:
```
lin_reg.fit(raw_data[['area','bedrooms','age']], raw_data['price'])
```

Out[13]:
```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=Fals
e)
```

In [14]:
```
lin_reg.coef_
```

Out[14]:
```
array([   137.25, -26025.  ,  -6825.  ])
```

In [15]:
```
lin_reg.intercept_
```

Out[15]:
```
383724.9999999998
```

In [16]:
```
lin_reg.predict([[3000,3,40]])          # 3000 sq ft, 3 bedrooms, 40 year old h
ouse
```

Out[16]:
```
array([444400.])
```

In [17]:
```
lin_reg.predict([[2500,4,5]])          # 2500 sq ft, 4 bedrooms, 5 years old
```

Out[17]:
```
array([588625.])
```

In [ ]:
```

```