In [1]:
```
import numpy as np
import pandas as pd
```

In [2]:
```
X = [i for i in range(1,8)]
Y =  [1.5,3.8,6.7,9.0,11.2,13.6,16.0]
```

In [3]:
```
dataset = pd.DataFrame(data={
    'X' : X,
    'Y' : Y
})
```

In [4]:
```
dataset
```

Out[4]:

	X	Y
0	1	1.5
1	2	3.8
2	3	6.7
3	4	9.0
4	5	11.2
5	6	13.6
6	7	16.0

In [5]:
```
dataset['XY'] = dataset['X'] * dataset['Y']
dataset['X2'] = dataset['X'] ** 2
```

In [6]:
```
dataset
```

Out[6]:

	X	Y	XY	X2
0	1	1.5	1.5	1
1	2	3.8	7.6	4
2	3	6.7	20.1	9
3	4	9.0	36.0	16
4	5	11.2	56.0	25
5	6	13.6	81.6	36
6	7	16.0	112.0	49

In [7]:
```
n = len(dataset)
sum_X = dataset['X'].sum()
sum_Y = dataset['Y'].sum()
sum_XY = dataset['XY'].sum()
sum_X2 = dataset['X2'].sum()
sum_X_h2 = sum_X**2
```

In [8]:
```
n,sum_X,sum_Y,sum_XY,sum_X2,sum_X_h2
```

Out[8]:
```
(7, 28, 61.800000000000004, 314.8, 140, 784)
```

In [9]:
```
numerator_m = (n*(sum_XY)) - (sum_X*sum_Y)
numerator_m
```

Out[9]:
```
473.1999999999998
```

In [10]:
```
denominator_m = ((n*sum_X2) - sum_X_h2)
denominator_m
```

Out[10]:
```
196
```

In [11]:
```
m = numerator_m / denominator_m
```

In [12]:
```
m
```

Out[12]:
```
2.4142857142857133
```

In [13]:
```
numerator_b = sum_Y - (m*sum_X)
denominator_b = n

b = numerator_b/denominator_b
b
```

Out[13]:
```
-0.8285714285714231
```

In [14]:
```
inputs = [i for i in range(1,8)]
outputs =[m * X+b for X in dataset['X']]

outputs
```

Out[14]:
```
[1.5857142857142903,
 4.0000000000000036,
 6.414285714285717,
 8.82857142857143,
 11.242857142857142,
 13.657142857142857,
 16.07142857142857]
```

In [15]:
```
dataset['outputs'] = outputs
```

In [16]:
```
dataset
```

Out[16]:

	X	Y	XY	X2	outputs
0	1	1.5	1.5	1	1.585714
1	2	3.8	7.6	4	4.000000
2	3	6.7	20.1	9	6.414286
3	4	9.0	36.0	16	8.828571
4	5	11.2	56.0	25	11.242857
5	6	13.6	81.6	36	13.657143
6	7	16.0	112.0	49	16.071429

In [17]:
```
dataset_metrics = dataset[['Y','outputs']]
```

In [18]:
```
!pip install scikit-learn
```

```
Requirement already satisfied: scikit-learn in c:\users\sankr\pictures\new fold
er\lib\site-packages (1.0.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in c:\users\sankr\pictures
\new folder\lib\site-packages (from scikit-learn) (2.2.0)
Requirement already satisfied: scipy>=1.1.0 in c:\users\sankr\pictures\new fold
er\lib\site-packages (from scikit-learn) (1.7.3)
Requirement already satisfied: numpy>=1.14.6 in c:\users\sankr\pictures\new fol
der\lib\site-packages (from scikit-learn) (1.21.5)
Requirement already satisfied: joblib>=0.11 in c:\users\sankr\pictures\new fold
er\lib\site-packages (from scikit-learn) (1.1.0)
```

In [19]:
```
from sklearn.metrics import r2_score,mean_absolute_error,mean_squared_error
```

In [20]:
```
r2_score(dataset_metrics['Y'],dataset_metrics['outputs'])
```

Out[20]:
```
0.99896818873402
```

In [21]:
```
mean_squared_error(dataset_metrics['Y'],dataset_metrics['outputs'])
```

Out[21]:
```
0.024081632653061246
```

In [ ]:
```

```