

```
import pandas as pd
import matplotlib.pyplot as plt
```

```
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
```

```
df = pd.read_csv('/content/HRPDataset (1).csv')
```

```
df.head()
```

	Area	Price	
0	400	2700000	
1	425	2837500	
2	450	2975000	
3	475	3112500	
4	500	3250000	

```
df.shape
```

```
(50, 2)
```

```
df.info
```

```
<bound method DataFrame.info of
0    400    2700000
1    425    2837500
2    450    2975000
3    475    3112500
4    500    3250000
5    525    3387500
6    550    3525000
7    575    3662500
8    600    3800000
9    625    3937500
10   650    4075000
11   675    4212500
12   700    4350000
13   725    4487500
14   750    4625000
15   775    4762500
16   800    4900000
17   825    5037500
18   850    5175000
19   875    5312500
20   900    5450000
21   925    5587500
22   950    5725000
23   975    5862500
```

24	1000	6000000
25	1025	6137500
26	1050	6275000
27	1075	6412500
28	1100	6550000
29	1125	6687500
30	1150	6825000
31	1175	6962500
32	1200	7100000
33	1225	7237500
34	1250	7375000
35	1275	7512500
36	1300	7650000
37	1325	7787500
38	1350	7925000
39	1375	8062500
40	1400	8200000
41	1425	8337500
42	1450	8475000
43	1475	8612500
44	1500	8750000
45	1525	8887500
46	1550	9025000
47	1575	9162500
48	1600	9300000
49	1625	9437500>

```
print('Defining x and y variable...')  
x= df['Area']  
y= df['Price']
```

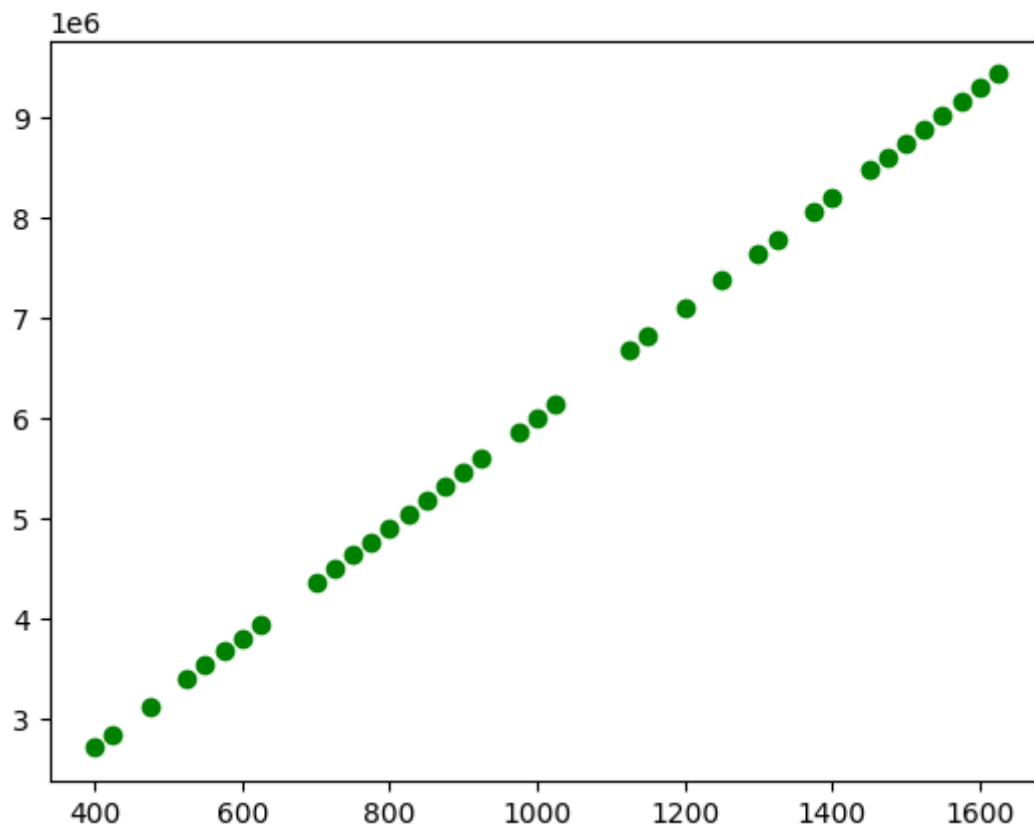
Defining x and y variable...

```
plt.scatter(x,y,color='blue')
```

```
<matplotlib.collections.PathCollection at 0x7c69266f9fc0>
```

```
1e6
x_train,x_test,y_train,y_test = train_test_split(x,y,random_state=0)
plt.scatter(x_train,y_train,color='green')
```

```
<matplotlib.collections.PathCollection at 0x7c6926551cf0>
```



```
plt.scatter(x_test,y_test,color='red')
```

<matplotlib.collections.PathCollection at 0x7c6923f23850>



Double-click (or enter) to edit

```
lr = LinearRegression()
lr.fit(x_train.values.reshape(-1,1),y_train)
```

▼ LinearRegression
LinearRegression()

```
print('Predicting using the trained model for the values of x_test')
y_pred=lr.predict(x_test.values.reshape(-1,1))
```

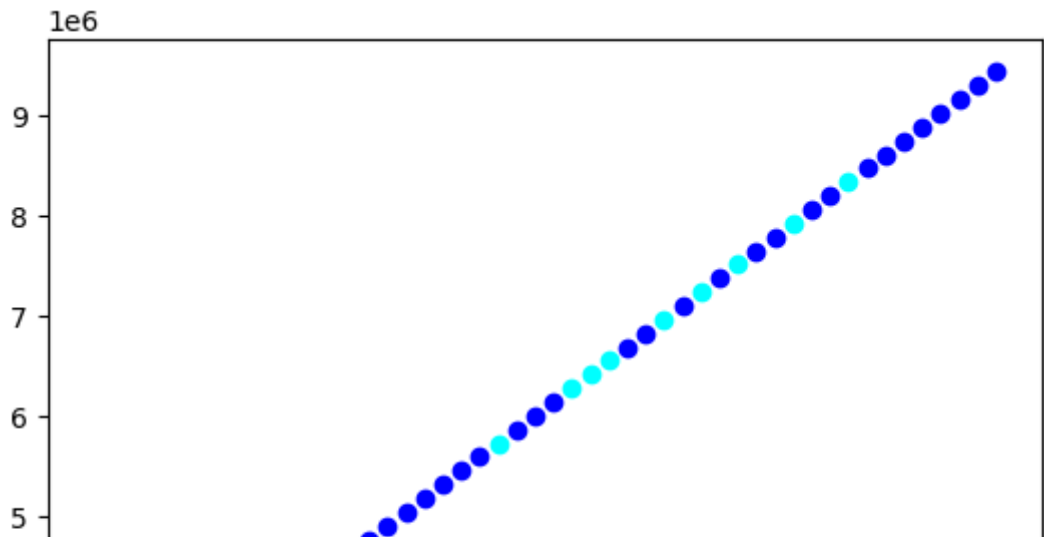
Predicting using the trained model for the values of x_test

```
print(y_test)
print(y_pred)
```

```
28    6550000
11    4212500
10    4075000
41    8337500
2     2975000
27    6412500
38    7925000
31    6962500
22    5725000
4     3250000
33    7237500
35    7512500
26    6275000
Name: Price, dtype: int64
[6550000. 4212500. 4075000. 8337500. 2975000. 6412500. 7925000. 6962500.
 5725000. 3250000. 7237500. 7512500. 6275000.]
```

```
plt.scatter(x_train,y_train, color='blue')
plt.scatter(x_test, y_pred, color='cyan')
```



<matplotlib.collections.PathCollection at 0x7c69263e6110>



```
df = pd.read_csv('/content/Placement (1).csv')
```

```
df
```

```
df.head()
```

	cgpa	package	
0	6.89	3.26	
1	5.12	1.98	
2	7.82	3.25	
3	7.42	3.67	
4	6.94	3.57	

```
plt.scatter(df['cgpa'],df['package'])
plt.xlabel('CGPA')
plt.ylabel('Package(in lpa)')
```

```
Text(0, 0.5, 'Package(in lpa)')
```



```
df.shape
```

```
(200, 2)
```

```
df
```

```
X = df.iloc[:,0:1]
```

```
Y = df.iloc[:,1]
```

```
X
```

```
Y
```

```
0      3.26
```

```
1      1.98
```

```
2      3.25
```

```
3      3.67
```

```
4      3.57
```

```
...
```

```
195    2.46
```

```
196    2.57
```

```
197    3.24
```

```
198    3.96
```

```
199    2.33
```

```
Name: package, Length: 200, dtype: float64
```

```
from sklearn.model_selection import train_test_split
```



```
X_train,X_test,Y_train,Y_test = train_test_split(X,Y,test_size=0.2,random_state=2)
```

```
lr.fit(X_train,Y_train)
```

```
LinearRegression
```

```
LinearRegression()
```

```
X_test
```

	cgpa	
112	8.58	
29	7.15	
182	5.88	
199	6.22	
193	4.57	
85	4.79	
10	5.32	
54	6.86	
115	8.35	
35	6.87	
12	8.94	
92	7.90	
13	6.93	
126	5.91	
174	7.32	
2	7.82	
44	5.09	
3	7.42	
113	6.94	
14	7.73	
23	6.19	
25	7.28	
6	6.73	
134	7.20	
165	8.21	
173	6.75	
45	7.87	
65	7.60	
48	8.63	
122	5.12	
178	8.15	
64	7.36	

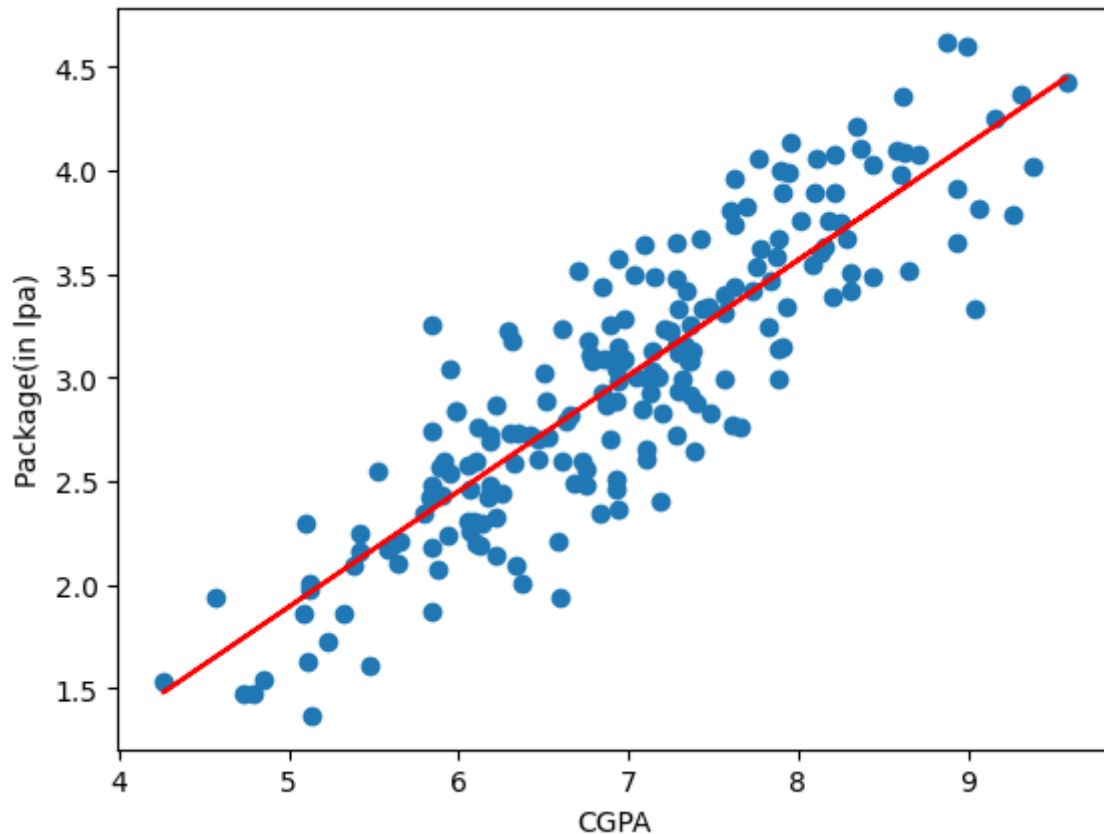
```
lr.predict(X_test.iloc[0].values.reshape(1,1))
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X c
warnings.warn(
array([3.89111601])
```

```
176 6.29
```

```
plt.scatter(df['cgpa'],df['package'])
plt.plot(X_train,lr.predict(X_train),color='red')
plt.xlabel('CGPA')
plt.ylabel('Package(in lpa)')
```

```
Text(0, 0.5, 'Package(in lpa)')
```



```
m = lr.coef_
m
```

```
array([0.55795197])
```

Double-click (or enter) to edit

```
b = lr.intercept_
b
```

```
-0.8961119222429144
```

```
m * 8.58 + b
```



```
array([3.89111601])
```

```
m * 6.87 + b
```

```
array([2.93701814])
```

```
m * 9.5 + b
```

```
array([4.40443183])
```

```
m * 100 + b
```

```
array([54.89908542])
```

[Colab paid products](#) - [Cancel contracts here](#)

✓ 0s completed at 17:18

