

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

```
In [2]: #reading data
data=pd.read_csv('headbrain.csv')
data.head()
```

Out[2]:

	Gender	Age Range	Head Size(cm^3)	Brain Weight(grams)
0	1	1	4512	1530
1	1	1	3738	1297
2	1	1	4261	1335
3	1	1	3777	1282
4	1	1	4177	1590

In [ ]:

headrain.csv data file\_

Show all

```
In [4]: #collecting of x and y
X=data['Head Size(cm^3)'].values
Y=data['Brain Weight(grams)'].values
```

```
In [5]: #mean of x and y
mean_x=np.mean(X)
mean_y=np.mean(Y)
print(mean_x)
print(mean_y)
```

```
3633.9915611814345
1282.873417721519
```

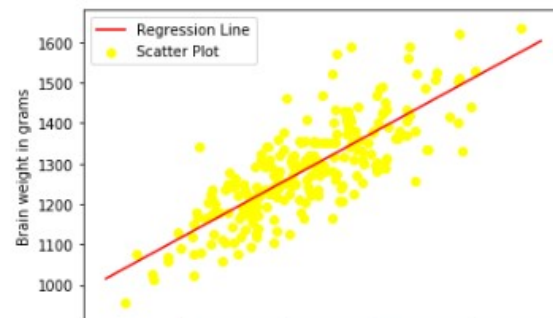
```
In [7]: #Total number of values
n=len(X)
n
```

```
Out[7]: 237
```

```
In [ ]:
```

```
In [8]: numer,denom=0,0
for i in range(n):
    numer+=(X[i]-mean_x)*(Y[i]-mean_y)
    denom+=(X[i]-mean_x)**2
b1=numer/denom
b0=mean_y-(b1*mean_x)
```

```
In [12]: max_x=np.max(X)+100
min_x=np.min(X)-100
#calculating line values of x and y
x=np.linspace(min_x,max_x,1000)
y=b0+b1*x
#plotting line
plt.plot(x,y,color='red',label='Regression Line')
#plotting scatter points
plt.scatter(X,Y,color='yellow',label='Scatter Plot')
plt.xlabel('Head size in cm3')
plt.ylabel('Brain weight in grams')
plt.legend()
plt.show()
```



Head size in cm3

```
In [14]: #calculating root mean square error
rmse=0
for i in range(n):
    y_pred=b0+b1*X[i]
    rmse+=(Y[i]-y_pred)**2
rmse=np.sqrt(rmse/n)
print("Root Mean Sqaure Error is: ",rmse)

Root Mean Sqaure Error is: 72.1206213783709
```

```
In [15]: #trying some values of x to get y
x=[10,20,30,40,50]
y=[]
for i in x:
    y.append(b0+b1*i)
print("X and Y Values are:")
for i in range(len(x)):
    print(x[i],y[i])
```

```
X and Y Values are:
10 328.2077144443362
20 330.84200783923023
30 333.4763012341242
40 336.11059462901824
50 338.7448880239122
```

In [ ]: