Perform Linear Regression by Hand

- Simple linear regression is a statistical method used to quantify the relationship between a predictor variable and a response variable.

Weight (lbs)	Height (inches)
140	60
155	62
159	67
179	70
192	71
200	72
212	75

Use the following steps to fit a linear regression model to this dataset, using weight as the predictor variable and height as the response variable.

Step 1: Calculate X*Y, X^2, and Y^2

Weight (lbs)	Height (inches)	X*Y	X^2	Y^2
140	60	8400	19600	3600
155	62	9610	24025	3844
159	67	10653	25281	4489
179	70	12530	32041	4900
192	71	13632	36864	5041
200	72	14400	40000	5184
212	75	15900	44944	5625

Step 2: Calculate ΣX, ΣY, $\Sigma X*Y$, ΣX^2, and ΣY^2

	Weight (lbs)	Height (inches)	X*Y	X^2	Y^2
	140	60	8400	19600	3600
	155	62	9610	24025	3844
	159	67	10653	25281	4489
	179	70	12530	32041	4900
	192	71	13632	36864	5041
	200	72	14400	40000	5184
	212	75	15900	44944	5625
Σ	**1237**	**477**	**85125**	**222755**	**32683**

Step 3: Calculate a

- The formula to calculate “a” is:

$$\mathbf{[n(\Sigma XY) - (\Sigma X)(\Sigma Y)] \ / \ [n(\Sigma X^2) - (\Sigma X)^2]}$$

- In this example,

$a = [7(85125) - (1237)(477)] \ / \ [7(222755) - (1237)^2]$

$a = \mathbf{0.2001}$

Step 4: Calculate b

- The formula to calculate b is:

$$\mathbf{([(\Sigma Y)(\Sigma X^2) - (\Sigma X)(\Sigma XY)] \ / \ [n(\Sigma X^2) - (\Sigma X)^2])}$$

- In this example

$b = [(477)(222755) - (1237)(85125)] \ / \ [7(222755) - (1237)^2]$

$b = \mathbf{32.783}$

Step 5: Place a and b in the estimated linear regression equation.

- The estimated linear regression equation is:

$$y = a*x + b$$

- In our example, it is **y = (0.2001)*x + 32.783**

```
import matplotlib.pyplot as plt
from scipy import stats

x = [140, 155, 159, 179, 192, 200, 212]
y = [60, 62, 67, 70, 71, 72, 75]
slope, intercept, r, p, std_err = stats.linregress(x, y)
def getY(x):
  return slope * x + intercept
line = list(map(getY, x))
print(r)
print(slope)
print(intercept)
plt.scatter(x, y)
plt.plot(x, line)
plt.xlim(0)
plt.ylim(0)
plt.show()
```

R=0.964944778696021
Slope=0.2000961670559142 4
Intercept=32.78300590740486