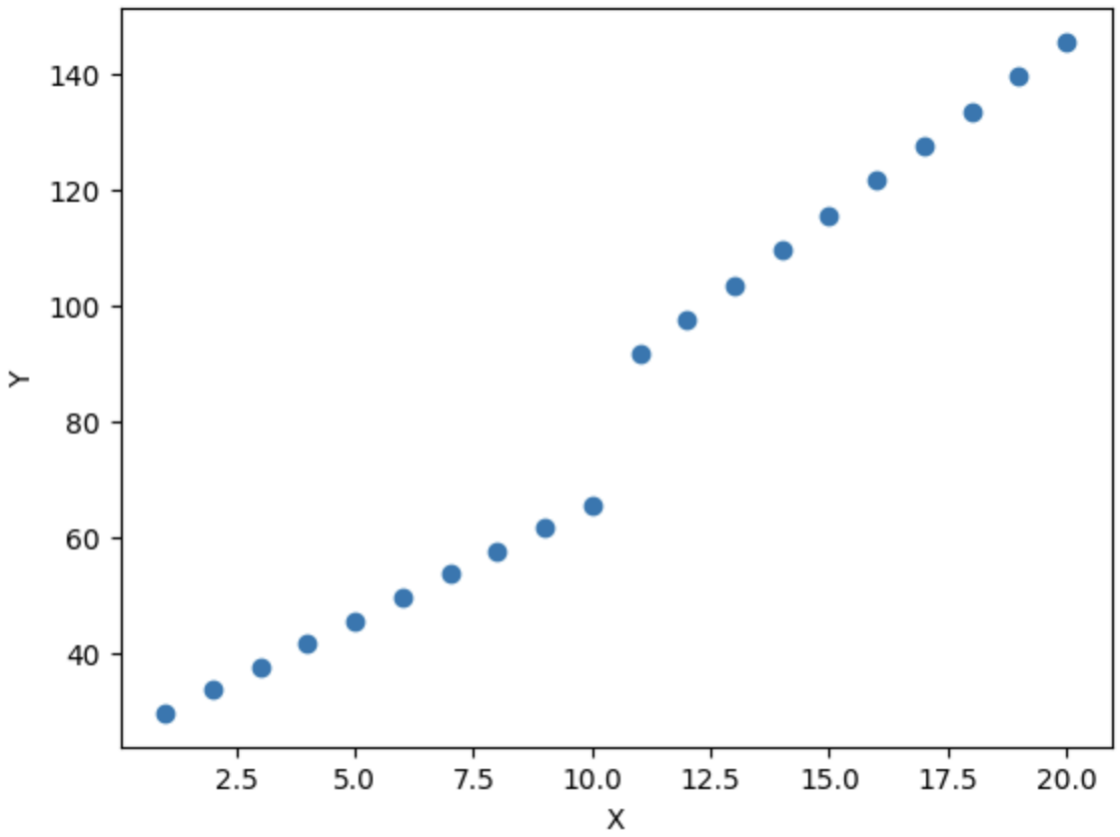
Swap Regression

Viral Chitlangia

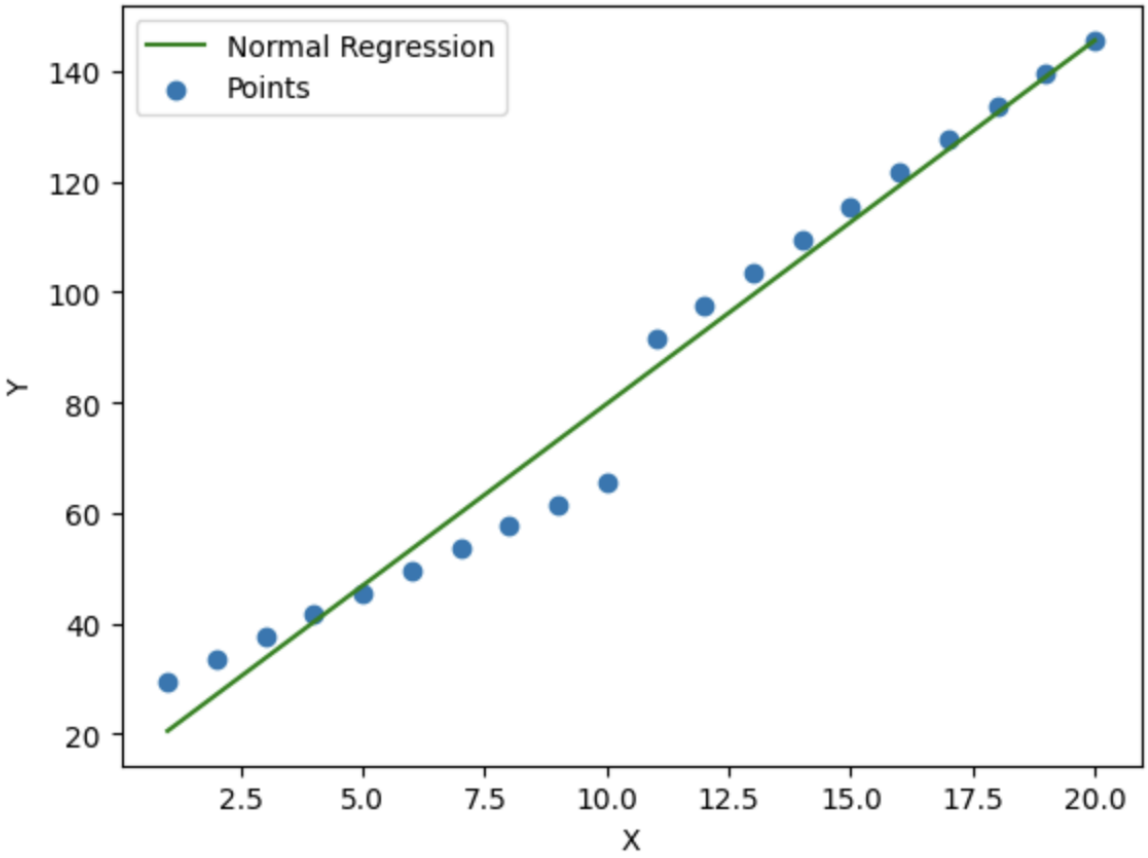
1 Data Used

I used the Random library in Python to generate a function with X going from 1 to 20, X1 from 1 to 10, and X2 from 11 to 20 and y changing as 4x + 6 for X1, and as 6x + 6 for X2. Y1 is to be regressed on X1, and X2 on Y2.



2 Normal Regression

Normal regression gives us the line y = 6.57894737x + 13.99677366



3 SWAP Regression

For SWAPping, I used Gradient Descent with Learning Rate 0.001, with Loss Function as -

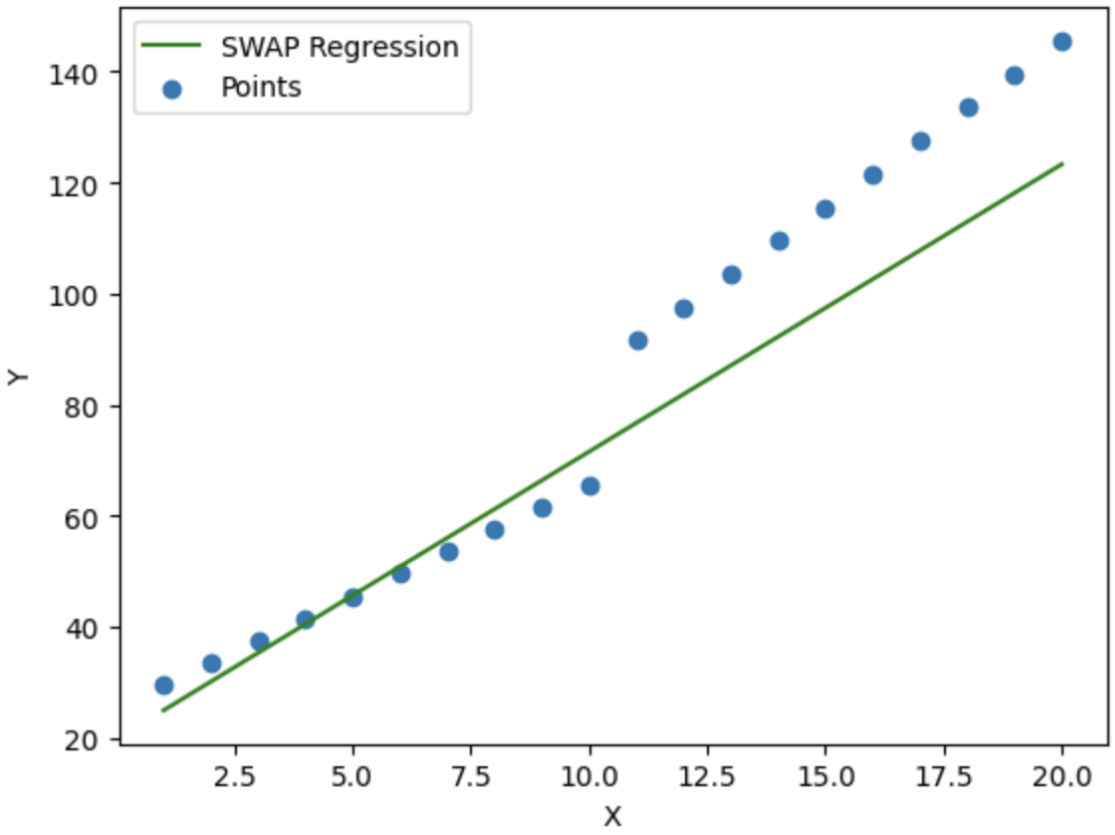
$L(m,c) = \Sigma_i(Y_{1i} - mX_{1i} - c)^2 + \Sigma_i(X_{2i} - \frac{(Y_{2i}-c)}{m})^2$

$\frac{\partial L}{\partial c} = \Sigma_i - 2(Y_{1i} - mX_{1i} - c) + \Sigma_i \frac{2}{m}(X_{2i} - \frac{(Y_{2i}-c)}{m})$

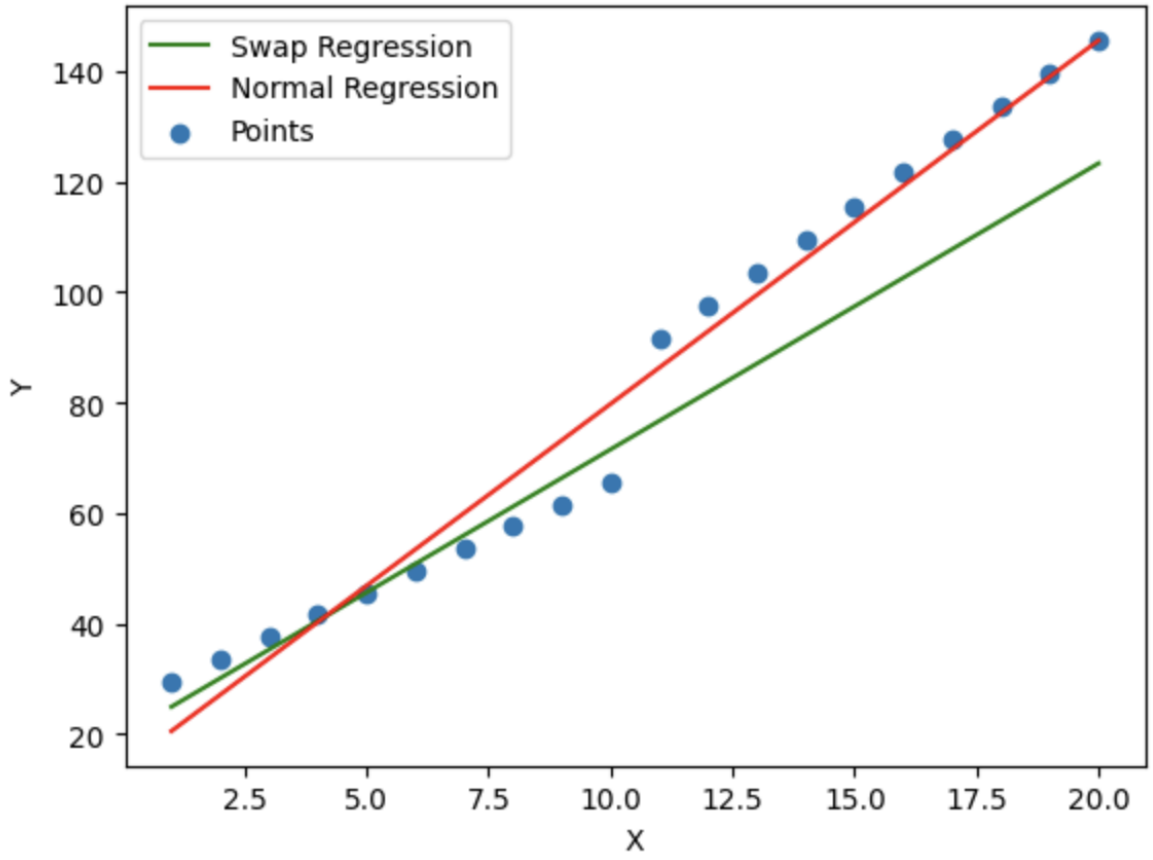
$\frac{\partial L}{\partial m} = \Sigma_i - 2X_{1i}(Y_{1i} - mX_{1i} - c) + \Sigma_i 2(\frac{Y_{2i}-c}{m^2})(X_{2i} - \frac{(Y_{2i}-c)}{m})$

Minimising the loss gives us the line - 5.175808605072054x + 19.80115568591847

$\sqrt{(\frac{\partial L}{\partial m})^2 + (\frac{\partial L}{\partial c})^2} = 0.000994026021313109$


SWAP Regression
Points
140
120
100
80
60
40
20
Y
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
X

4 Comparison and Analysis



	Normal Regression	SWAP Regression
Mean Squared Error on X1 and Y1	7.852254118203402	3.4474963806379595
Mean Squared Error on X2 and Y2	3.0907331782195526	18.699981027583256
Mean Squared Error on X and Y	5.967014593403866	13.445715334754146

5 Final

As of now, in this model, Normal Regression is giving better result than SWAP Regression, except when Y is regressed on X.