Question 1:

Match the following 1, 2 & 3 with A, B, C & D in the case of Linear Regression

Items

1) Increasing trend
2) Decreasing trend
3) Positive Intercept

A) Slope>0 or b1>0

B) Slope<0 or b1<0

C) Line meets the Y axis below origin

D) Line meets the Y axis above origin

Question 2:

Draw a scatter plot of the variables (Module_3_LR.csv) in the data frame with IV (X) on X-axis and the DV (Y) on the Y-Axis. How does the trend look like?

1) Decreasing
2) Increasing
3) Constant
4) neither increasing nor decreasing

Question 3:

Looking at the scatter plot, which of these can be a possible Linear Regression Line for the data in Module_3_LR.csv. (This question can be answered visually without using Python)

1) Y= 50+12.3 X
2) Y= -119 + 6X
3) Y=-50
4) Y= - 50 - 12.3 X
5) Y= -120+12.3X
6) Y= -120-12.3X

Question 4:

The equation of the LR line (Module_3_LR.csv) with IV as X and DV as Actual is Actual= -29.333+1.957X. For a value of X=0 what would you predict? _____ (Write the answer in 3

decimal places without rounding off)

Using this equation, find the predicted/fitted values for each point in the data-set. Store them in the data-frame as a new column Predicted. Do this using MS Excel. Python is not needed.
Now let us find the error the LR line is making for each point.

We can compute the error using df['error']= df['Actual']- df['Predicted'] , where error is the new column created to store the error, Predicted is the column created above.
What is the Total sum of errors of all Data points? _____ (round off to nearest integer without decimal)
What is the SSE for this data? _____ (round off to nearest Integer)

Note: Do this using MS Excel. Python is not needed.

#All codes are available at https://github.com/MyDataCafe/

#All Class Videos are at https://www.youtube.com/mydatacafe

#We are on Facebook https://www.facebook.com/mydatacafe/

#We are on LinkedIn https://www.linkedin.com/company/mydatacafe/