Internship No. : DSC2311

Question 1: When implementing linear regression of some dependent variable y on the set of independent variables $\mathbf{x} = (x_1, \ldots, x_r)$, where r is the number of predictors, which of the following statements will be true?

Ans. :- Both a and b

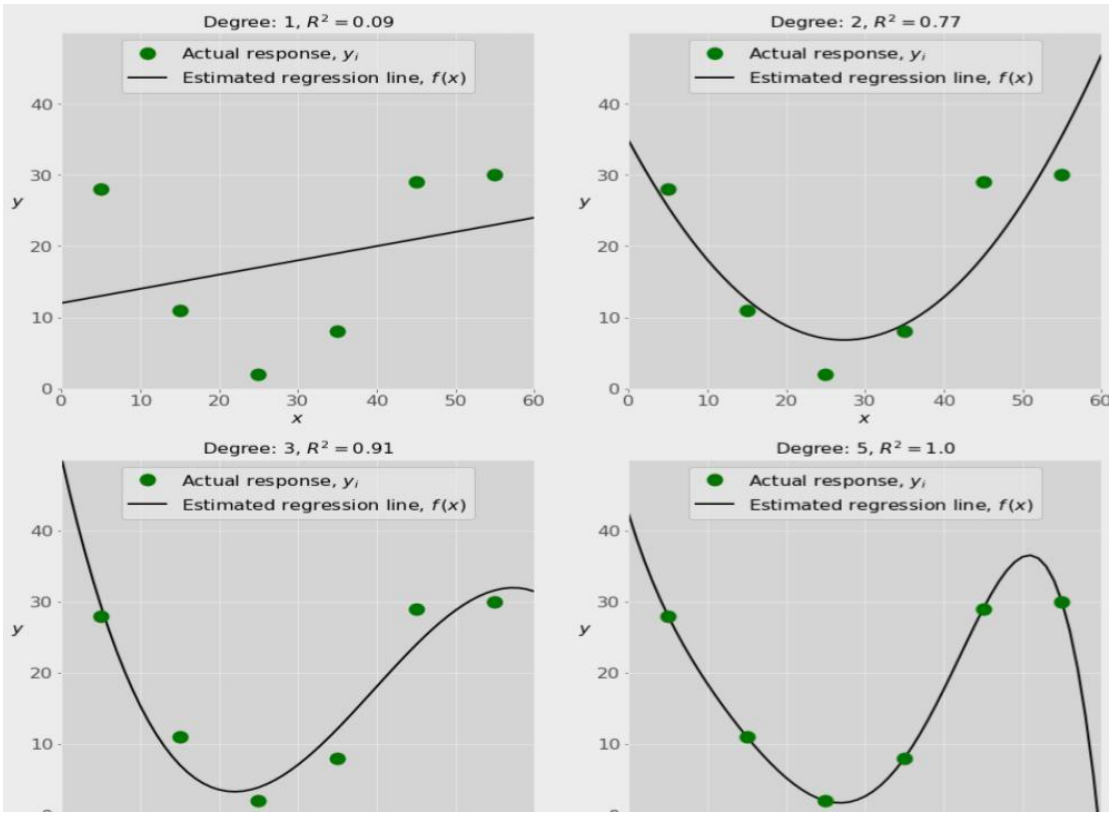
Question 2: What indicates that you have a perfect fit in linear regression?

Ans. :- **The value $R^2 = 1$, which corresponds to SSR = 0**

Question 3 :- In simple linear regression, the value of what shows the point where the estimated regression line crosses the y axis?

Ans. :- B0

Question 4: Check out these four linear regression plots:



Which one represents an underfitted model?

Ans. :- The top-left plot

Question 5 : There are five basic steps when you're implementing linear regression:

- a. Check the results of model fitting to know whether the model is satisfactory.
- b. Provide data to work with, and eventually do appropriate transformations.
- c. Apply the model for predictions.
- d. Import the packages and classes that you need.
- e. Create a regression model and fit it with existing data.

However, those steps are currently listed in the wrong order. What's the correct order?

Ans. :- d,b,e,a,c

Question 6 : Which of the following are optional parameters to LinearRegression in scikit-learn?

Ans. :- fit_intercept , normalize ,reshape ,copy_X , n_jobs, these are optional paramater to LinearRegression in scikit-learn.

Question 7 : While working with scikit-learn, in which type of regression do you need to transform the array of inputs to include nonlinear terms such as x^2?

Ans. :- Polynomial Regression

Question 8 : You should choose statsmodels over scikit-learn when:

Ans. :- You want graphical representation of your data

Question 9 : _____ is a fundamental package for scientific computing with Python. It offers comprehensive mathematical functions, random number generators, linear algebra routines, Fourier transforms, and more. It provides a high-level syntax that makes it accessible and productive.

Ans. :- Numpy

Question 10 : _____ is a Python data visualization library based on Matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics that allow you to explore and understand your data. It integrates closely with pandas data structure.

Ans. :- Seaborn