

## Solutions for the Assignment given on 17-03-2023

**Name:** Dr. Nalla Nandakishore

**Batch:** DS2302

**Institute:** DataTrained

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

- a)  $\beta_0, \beta_1, \dots, \beta_r$  are the **regression coefficients**.
- b) Linear regression is about determining the **best predicted weights** by using the **method of ordinary least squares**.
- c)  $E$  is the random interval
- d) Both a and b

**Ans: Option D: Both a and b**

22) What indicates that you have a **perfect fit** in linear regression?

- a) The value  $R^2 < 1$ , which corresponds to  $SSR = 0$
- b) The value  $R^2 = 0$ , which corresponds to  $SSR = 1$
- c) The value  $R^2 > 0$ , which corresponds to  $SSR = 1$
- d) The value  $R^2 = 1$ , which corresponds to  $SSR = 0$

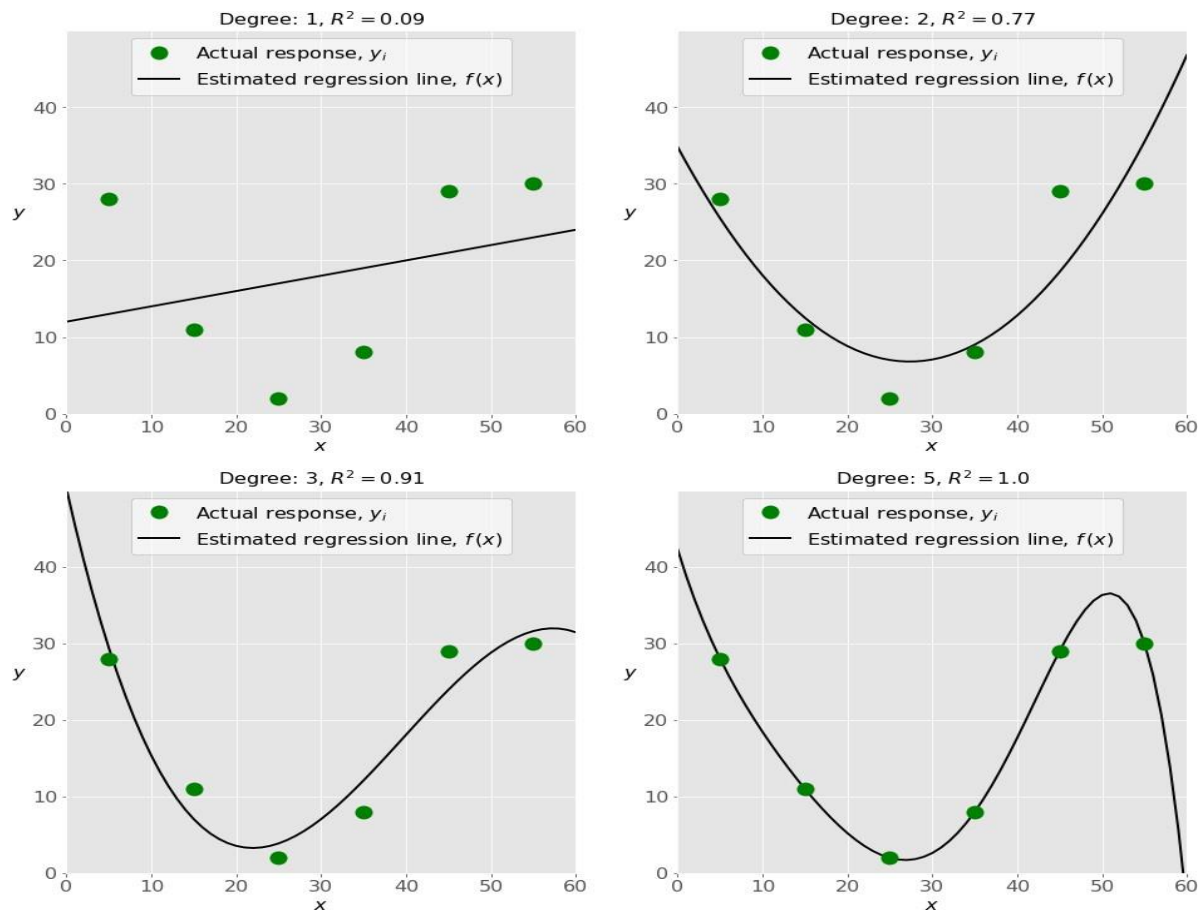
**Ans: Option D: The value  $R^2 = 1$ , which corresponds to  $SSR = 0$**

23) In simple linear regression, the value of **what** shows the point where the estimated regression line crosses the  $y$  axis?

- a)  $Y$
- b)  $B_0$
- c)  $B_1$
- d)  $F$

**Ans: Option B:  $B_0$**

24) Check out these four linear regression plots:



Which one represents an **underfitted** model?

- The bottom-left plot
- The top-right plot
- The bottom-right plot
- The top-left plot

Ans: Option D: The top-left plot

25) 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?

- e, c, a, b, d
- e, d, b, a, c
- d, e, c, b, a
- d, b, e, a, c

Ans: Option D: d, b, e, a, c

26 ) Which of the following are optional parameters to LinearRegression in scikit-learn?

- a) Fit
- b) fit\_intercept
- c) normalize
- d) copy\_X
- e) n\_jobs
- f) reshape

**Ans: Options b,c,d and e : fit\_intercept, normalize, copy\_X and n\_jobs**

27) 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  $x^2$ ?

- a) Multiple linear regression
- b) Simple linear regression
- c) Polynomial regression

**Ans: Option C : Polynomial Regression**

28) You should choose statsmodels over scikit-learn when:

- a) You want graphical representations of your data.
- b) You're working with nonlinear terms.
- c) You need more detailed results.
- d) You need to include optional parameters.

**Ans: Option a: You want graphical representation of your data**

29) \_\_\_\_\_ 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.

- a) Pandas
- b) Numpy
- c) Statsmodel
- d) scipy

**Ans: Option B: Numpy**

30 ) \_\_\_\_\_ 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 structures.

- a) Bokeh
- b) Seaborn
- c) Matplotlib
- d) Dash

**Ans: Option B: Seaborn**