

## ASSIGNMENT 2

**Q21.** 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?

**Answer: d) Both a) & b)**

**Q22.** What indicates that you have a perfect fit in linear regression?

**Answer: d) The value  $R^2 = 1$ , which corresponds to SSR = 0**

**Q23.** In simple linear regression, the value of what shows the point where the estimated regression line crosses the  $y$  axis?

**Answer: b) B0**

**Q24.** Check out these four linear regression plots: Which one represents an under fitted model?

**Answer: d) The top left plot**

**Q25.** 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?

**Answer: d) d, b, e, a, c**

**Q26.** Which of the following are optional parameters to Linear Regression in Scikit-learn?

**Answer: b) fit\_intercept**

**c) normalize**

**d) copy\_X**

**e) n\_jobs**

**Q27.** 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  $x^2$ ?

**Answer: c) Polynomial regression**

**Q28.** You should choose stats models over Scikit-learn when:

**Answer: c) You need more detailed results.**

**Q29.** The \_\_\_\_\_ 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.

**Answer: b) NumPy**

**Q30.** ) \_\_\_\_\_ 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.

**Answer: b) Seaborn**