

**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?

**Answers:**

- 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**.

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

**Answer:**

- 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?

**Answer:**

- b)  $\beta_0$

**24.** Check out these four linear regression plots. Which one represents an **underfitted** model?

**Answer:**

- 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?

**Answer:**

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

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

**Answers:**

- b) `fit_intercept`
- d) `copy_X`
- e) `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  $x^2$ ?

**Answer:**

c) Polynomial regression

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

**Answer:**

c) You need more detailed results.

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.

**Answer:**

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.

**Answer:**

b) Seaborn