

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?

A21: d) Both a and b

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

A22: 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?

A23: b)  $B_0$

Q24: Which one represents an underfitted model?

A24: d) The top-left plot

Q25: There are five basic steps when you're implementing linear regression: However, those steps are currently listed in the wrong order. What's the correct order?

A25: d) d, b, e, a, c

Q26: Which of the following are optional parameters to LinearRegression in scikit-learn?

A26: 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  $x^2$ ?

A27: c) Polynomial regression

Q28: You should choose statsmodels over scikit-learn when:

A28: c) You need more detailed results.

Q29: \_\_\_\_\_ 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.

A29: 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.

A30: b) Seaborn