BATCH DS2306

FILE 2

Q1. 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? a) $\beta_0, \beta_1, \ldots, \beta_r$ 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 and b.

Ans- Both A and B

$\beta_0, \beta_1, \ldots, \beta_r$ are the regression coefficients.

Linear regression is about determining the best predicted weights by using the method of ordinary least squares.

Q2. What indicates that you have a perfect fit in linear regression?

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

Q3. In simple linear regression, the value of what shows the point where the estimated regression line crosses the y axis?

Ans- B0

Q4. Check out these four linear regression plots:

Ans- The top-left plot

Q5. 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

Q6. 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- B, C, D, E

Q7. 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- Polynomial regression

Q8. 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- You need more detailed results

Q9. _____ 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

Q10. _____ 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.

Ans- Seaborn