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

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?

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?

b) B0

24) Which one represents an **underfitted** model?

a)The bottom-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?

d) d, b, e, a, c

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

c) normalize

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?

c) Polynomial regression

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

.

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.

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.

b) Seaborn

41) Among the following identify the one in which dimensionality reduction reduces.

d) Collinearity

42) Which of the following machine learning algorithm is based upon the idea of bagging?

b) Random Forest

43) Choose a disadvantage of decision trees among the following.

c) Decision Tree are prone to overfit

44)What is the term known as on which the machine learning algorithms build a model based on sample data?

c) Training data

45) Which of the following machine learning techniques helps in detecting the outliers in data?

c) Anomaly detection

46) Identify the incorrect numerical functions in the various function representation of machine learning.

c) Case based

47)Analysis of ML algorithm needs

d) Both a and b

48) Identify the difficulties with the k-nearest neighbor algorithm.

c) Both a and b

49) The total types of the layer in radial basis function neural networks is _____

b) 2

50)Which of the following is not a supervised learning

d) KMeans