1. 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?

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

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

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

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

   Ans: B0

4. Check out these four linear regression plots:

   Ans: The top-left plot

5. There are five basic steps when you're implementing linear regression:
   Ans: d, b, e, a, c

   Import the packages and classes that you need.
   Provide data to work with, and eventually do appropriate transformations.
   Create a regression model and fit it with existing data.
   Check the results of model fitting to know whether the model is satisfactory.
   Apply the model for predictions.

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

   Ans: fit_intercept, normalize, copy_X, n_jobs. (Optional Parameter)

7. 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?

   Ans: Polynomial regression

8. You should choose stats models over scikit-learn when:

   Ans: You need more detailed results.

9. _____ 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

10. _____ 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: Bokeh

11. Among the following identify the one in which dimensionality reduction reduces.

Ans: Collinearity

12. Which of the following machine learning algorithm is based upon the idea of bagging?

Ans: Random Forest

13. Choose a disadvantage of decision trees among the following.

Ans: Decision Tree are prone to overfit.

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

Ans: Training Data.

15. Which of the following machine learning techniques helps in detecting the outliers in data?

Ans: Anomaly Detection

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

Ans: Super Vector and Regression.

17. Analysis of ML algorithm needs

Ans: Both a and b [Statistical Learning Theory and Computational Learning Theory]

18. Identify the difficulties with the k-nearest neighbour algorithm.

Ans: Both a and b [Curse of dimensionality and Calculate the distance of test case for all training cases]

19. The total types of the layer in radial basis function neural networks is _____

    Ans: Three

20. Which of the following is not a supervised learning?

    Ans: PCA