1. **Which of the following is an invalid variable?**

   a) my_string_1   b) 1st_string   c) foo   d) _

2. **Which of these in not a core data type?**

   a) Lists   b) Dictionary   c) Tuples   d) Class

3. **What data type is the object below ?**

   L = [1, 23, „hello", 1].

   a) list   b) dictionary   c) array   d) tuple

4. **In order to store values in terms of key and value we use what core data type.**

   a) list   b) tuple   c) class   d) dictionary

5. **Which is the correct operator for power(x^y)?**

   a) X^y   b) X**y   c) X^^y   d) None of the mentioned

6. **Which one of these is floor division?**

   a) /   b) //   c) %   d) ^

7. **Mathematical operations can be performed on a string. State whether true or false.**

   a) True   b) False

8. **The expression int(x) implies that the variable x is converted to integer. State whether true or false.**

   a) True   b) False

9. **Input function in python**

   a) input()   b) read()   c) enter()   d) scanf()

10. **What arithmetic operators cannot be used with string?**

    a) +   b) *   c) –   d) a and b

11. **Suppose list1 is [2, 33, 222, 14, 25], What is list1[-1] ?**

   a) Error   b) None   c) 25   d) 2

12. **Suppose list1 is [1, 3, 2], What is list1 * 2 ?**

   a) [2, 6, 4]   b) [1, 3, 2, 1, 3]   c) [1, 3, 2, 1, 3, 2]   D) [1, 3, 2, 3, 2, 1]

13. **Tuple can be modified.**

      a) True   b) False

14. **List can not be modified.**

      a) True   b) False

15. **Which of the following is a Python tuple?**

   a) [1, 2, 3]   b) (1, 2, 3)   c) {1, 2, 3}   d) {}

16. **Suppose t = (1, 2, 4, 3), which of the following is incorrect?**

      a) print(t[3]) b) print(len(t))   c) print(max(t))   d) t[3] = 45

17. **Which of the following statements create a dictionary?**

   a) d = {}   b) d = {"john":40, "peter":45}

   c) d = {40:"john", 45:"peter"}   d) All of the mentioned

18. **Read the code shown below carefully and pick out the keys?**

         d = {"john":40, "peter":45}

      a) "john", 40, 45, and "peter"   b) "john" and "peter"

   c) 40 and 45   d) d = (40:"john", 45:"peter")

19. **Identity operator in python is**

      a) in   b) is   c) is not   d) both b & c

20. Membership operator is:

    a) in   b) is   c) not in   d) a and c

21. Proper indentation is required in python.

    a) True   b) False

22. Functions are declared using the following keyword:

    a) def   b) define   c) declare   d) start

23. Which statement brings the program control to the beginning of the loop.

    a) main   b) continue   c) begin   d) first

24. A function cannot return a value.

    a) True   b) False

25. Keyword argument can be used within a function calling

    a) True   b) False

26. Anonymous function objects in Python:

    a) lambda   b) try   c) loop   d) inline

27. Applies the given function to each item of a given iterable (list, tuple etc.)

    a) lambda   b) reduce   c) map   d) inline

28. A class constructor

    a) Same as class name   b)＿   c) different name   d) All
        init＿

29. The first parameter of any function in the class is:

    a) class reference   b) function name   c) datatype   d) None

30. **A special type of method (function) which is used to initialize the members of the class.**

a)   function   b) constructor   c) map   d) member

31. **Python supports method overloading**

a) True   b) False

32. **An abnormal condition in a program resulting in the disruption in the flow of the program.**

a) error   b) exception   c) condition   d) logical error

33. **The try block must be followed by the except statement.**

a) True   b) False

34. **The try block may have more than one except statement.**

a) True   b) False

35. **An exception can be raised by using**

a) throw   b) generate   c) create   d) raise

36. **A python module is**

a) folder   b) python program file   c) python function   d) None

37. **The extension of python program file**

a) .txt   b) .py   c) .jpynb   d) .python

38. **A one-dimensional labeled array capable of holding data of any type**

a) Series   b) Data frame   c) array   d) element

39. **A Data frame is a two-dimensional data structure**

a) True   b) False

40.A data frame is created using dictionary, list, or any other data structure.

a) True   b) False

41. Function used to add a new row at the end of the data frame

a)   append()   b) insert()   c) new()   d) create()

42.To access a group of rows and columns by label(s).

a) view()   b) search()   c) loc()   d) location()

43.Function for arithmetic median of values

a) middle()   b) mead()   c) median()   d) avg()

44.To access group of rows and columns by integer locations

a) view()   b) iloc()   c) loc()   d) location()

45.Which function computes a summary of statistics pertaining to the DataFrame column?

a) info()   b) stat()   c) describe()   d) summery()

46.Which machine learning algorithm solves both classification and regression problem?

a) Decision Tree   b) Naïve Bayes   c) K-means   d) All

47.Regression algorithms fall under the

a) Supervised M/L   b) Unsupervised M/L   c) Semi-supervised M/L d) None

48.Algorithm used to estimate real values based on continuous variable

a) median   b) mode   c) kNN   d) Linear Regression

49. **Machine learning library for the Python programming language**

    a) pandas   b) numpy   c) sklearn   d) matplotlib

50. **Algorithm which solves the clustering problem**

    a) Naïve Bayes   b) Linear Regression   c) K-Means   d) Decision Tree

51. **Which of the following object you get after reading CSV file?**

    a) DataFrame   b) Character Vector   c) Panel   d) All

52. **Which of the following is used for machine learning in python?**

    a) scikit-learn   b) seaborn-learn   c) stats-learn   d) none

53. **The _____ project builds on top of pandas and matplotlib to provide easy plotting of data.**

    a) yhat   b) Seaborn   c) Vincent   d) None

54. **Data that summarize all observations in a category are called _____ data.**

    a) frequency   b) summarized   c) raw   d) All

55. **Which of the following is characteristic of Raw Data?**

    a) Data is ready for analysis   b) Original version of data
    c) Easy to use for data analysis   d) All

56. **Which of the following thing can be data in Pandas?**

    a) A python dict   b) an ndarray   c) a scalar value   d) all

57. **Series is a one-dimensional labeled array capable of holding any data type.**

    a) True   b) False

58. **All pandas data structures are ___ mutable but not always _____ mutable.**

    a) size, value   b) semantic, size   c) value, size   d) none

59. What is called when a function is defined inside a class?

a) Module   b) Class   c) Another function   d) Method

60. Size of the pandas dataframe is mutable.

a) True   b) False

61. A pandas DataFrame can be created

a) pandas.DataFrame   b) pandas.dataFrame   c) pandas.dataframe   d) dataframe

62. To change the name of the column in a dataframe, use:

a) change()   b) ren()   c) rename()   d) None

63. To insert column into DataFrame at specified location, use:

a) enter()   b) insert()   c) update()   d) None

64. Rows can be selected by passing integer location to loc() function

a) True   b) False

65. An open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures

a) Numpy   b)Pandas   c) Matlab   d) Dataframe

66. Pandas is well suited for:

a) Tabular data with heterogeneously-typed columns, as in an SQL table or Excel spreadsheet

b) Ordered and unordered time series data

c) Arbitrary matrix data (homogeneously typed or heterogeneous) with row and column labels

d) All

67. **The most widely used metrics and tools to assess a classification model is:**

   a) Confusion matrix   b) Error   c) Mean Squared Error   d) None

68. Which of the following is a disadvantage of decision trees?

   a) Factor analysis   b) Decision trees are robust to outliers

   c) Decision trees are prone to be overfit   **d)** None of the above

69. How do you handle missing or corrupted data in a dataset?

a) Drop missing rows or columns   b)Replace missing values with mean/median/mode

b) Assign a unique category to missing values   d) All of the above

70. What is the purpose of performing cross-validation?

   a. To assess the predictive performance of the models

   b. To judge how the trained model performs outside the sample on test data

   c. Both A and B

   d. None

71. Which method used to read comma separated value file in pandas?

   a) read_excel()   b) read_csv()   c) csv_read()   d) none

72. The most widely used metrics and tools to assess a regression model is:

   a) Confusion matrix   b) Mean Squared Error   c) error   d) all

73. Handles non NaN values in a DataFrame:

   a) values   b) na_values   c) values_na   d) na

74. To check missing values in a data frame:

   a) null()   b) isnull()   c) missing()   d) values()

75. To drop Rows/Columns with Null values:

    a) dropna()   b) remove()   c) drop()   d) fill_na()

76. To fill missing values with specific criterion, use

    a) fillna()   b) fill()   c) fill_na   d)none

77. To count frequency of each category

    a) value_counts()   b) value_count()   c) value()   d) counts()

78. To get a concise summary of the dataframe, use

    a) info()   b) detail()   c) information()   d) none

79. To print the data types of every columns, use

    a) type   b) dtypes   c) datatypes   d) all

80. To change the data types of a dataframe, use

    a) atype   b) astype   c) sametype   d) type

81. **A function in Sklearn for splitting data arrays into two subsets: for training data and for testing data**

    a) split()   b) train_test_split()   c) train_split()   d) test_split()

82. ……… means to do some calculation and then do transformation.

    a) fit_transform   b) fit   c) transform   d) trans_fit

83. The ……… function is used to convert categorical variable into dummy/indicator variables.

    a) get_dummies()   b) dummy()   c) get_dummy()   d) none

84. Which method predict the label of a new set of data

    a) predict()   b) output()   c) fit()   d) score()

85. For supervised learning applications, this accepts two arguments: the data X and the labels y

a) model. fit()   b) model.predict()   c) model.learn()   d) medel.score()

86. ………… calculates the score of classification or regression problem.

a) model.score()   b) model.fit()   c) model.learn()   d) none

87. Decision tree splits into 2 parts at each node

a) True   b) False

88. Class imbalance can be measured by computing Gini.

a) True   b) False

89. A confusion matrix is a summary of prediction results on a classification problem.

a) True   b) False

90. Library for data visualization

a) plot   b) matplotlib   c) matplot   d)none