Pandas Question:

1. List the first top 10 data in dataframe

    1.1 and Store them in a separate csv file.

    1.2 Show the directory also

    1.3 Read the stored csv

    1.4 data are separated by hash

    1.5 Using pd.read_csv work on features like "header",'names"

    1.6 Know about na_value feature in pd.read_csv and fill some na_values

2. Describe function in pandas

    2.1 . Get the description of the dataset

    2.2 Print all the columns / features of dataset using pandas

    2.3 Select a particular feature and display its values

    2.4 Print the type of particular feature

3. What are the types of datatypes in pandas ?
4. What is series and dataframes?
5. Select any feature of dataset and convert that series data in dataframe format

    5.1 Check the datatype of above conversion

6. fetch two or more data columns or records
7. Read some data from excel sheet

    7.1 Read data from a particular sheet in excel

8. Read data from HTML file , display that and analyze what is happening

Source:
https://www.basketball-reference.com/players/i/ibakase01.html?utm_source=direct&utm_medium=Share&utm_campaign=ShareTool

- calculate how many tables in this html page
- fetch the first table from the html page

9. Read data from github repository
10. Find the datatypes of all the features in the dataframe
11. Filter only object data types
12. Give me the column name of data whose data type is object
13. Get the complete data with values whose data type is object
14. Find the statistical info of NON-NUMERICAL DATA TYPES
15. From original dataset ,filter only the data whose datatype is float
16. Find the datatype of features which are in the dataset individually—-> type(df[feature[])
17. Get the list of all column names in a dataset and also use for loop to iterate through the column names
18. How to add a new column into a dataset
19. How to assign values of one column to another column
20. How to find the distinct values of a variable

Titanic dataset—->
https://raw.githubusercontent.com/Nitish14-MC/Python_notes/main/titanic_dataset.csv

21. Check for null values in "cabin" column
22. Get the index positions of null values in "Cabin" column
23. In titanic dataset, find what is the maximum value in "Fare" column
24. Get the data of whose fare is maximum
25. Get the Name of passenger whose fare is maximum
26. How to add a new column by adding the values of two columns
27. Series is somewhat similar to lists...We can perform slicing and most of the operations
28. In series we cannot access values using NEGATIVE INDEXING
29. In TITANIC DATASET, find out How many male passengers survived
30. How to generate a random numbers using numpy
31. Perform some operations on loc and iloc
32. Difference between loc and iloc
33. How to make a dataset using random numbers
34. Create a random dataset and convert it into the dataframe
35. After the above , give index and column values to the dataframe
36. Seed the random numbers
37. How to get some negative values in the dataset which is randomly generated
38. How to create a dataframe using DICTIONARY
39. How to create a dataframe using DICTIONARY with null values
40. Using dropna function remove the rows that contains NULL values

41. Using dropna function remove the columns that contains NULL values
42. Know about thresh in dropna function