Pandas (Hands-on 1)

support@intellipaat.com

+91-7022374614

US: 1-800-216-8930 (Toll-Free)

Problem Statement:

Large datasets need to be structured and free of outliers in order to be used for further analysis and model building in cases of supervised or unsupervised learning. Python pandas is one such python package that deals with manipulation of data with various operations. Use the modules and processes available in the python pandas module to come up with useful insights about the data.

Tasks To be Performed:

1. Given a sample list, tuple and a set with elements as numbers from 1-10. Convert each of the data types into a pandas series. Create a separate list of index values for the series.
2. Create a python dictionary with three key value pairs. Keep the keys as the first names, and the respective last names as values in the dictionary. Use the dictionary to create a pandas series.
3. Create a pandas series, with the elements as numbers from 1-10. Use the index values starting from 0-10. For the above series:
    a. Convert it to a list.
    b. Convert it to a dictionary.
    c. Convert it to a tuple.
    d. Convert it to a set.
4. Create a sample pandas series with x integer elements. Perform the following operations on the same:
    a. Print the first 5 elements of the series
    b. Print the last 5 elements of the series
    c. Insert an integer value ‘50’ at the index 5.
    d. Update the last element in the series with the value 100.
    e. Delete the first element from the series.

    Note: The index values must not be affected by the operations and should be uniform for the series after the last operation.

5. For a given pandas series obtained in the 4th step. Write a sample code to find the element 100 in the series.
6. For the given series A and B, perform concatenation on the same.
    series_A = pd.Series([1,2,3,4,5])
    series_B = pd.Series([6,7,8,9,10])
7. Using loops in python, print all the elements within a pandas series.
    Sample = pd.Series([1,2,3,4,5,6,7,8,9,10])
8. For a given nested list and a sample dictionary, create two data frames using pandas.
    List = [['AA','BB','CC'], ['DD','EE','FF'], ['GG','HH','II']]
    Dictionary = { 'Kris': ['Jordan', 12,22],'Memphis': ['Depay', 23, 24]}

9. Create a pandas Dataframe with two columns = [‘Name’, ‘Age’]. Perform the following operations on the same.
    a. Print the first 3 rows of the Data Frame .
    b. Print the rows that have the age greater than 15.
    c. Insert 5 imaginary names and ages as values to the pandas Data Frame .
    d. Update the 5th row so that the name = john and age=56.
    e. Delete the first row from the Data Frame .
10. For the Data Frame obtained from step 9, convert the same to a list and a dictionary respectively.
11. Create two separate Data Frame s A and B, and perform the following operations on them.
    a. Concatenate the two Data Frame s to a new Data Frame C.
    b. Split the Data Frame C row wise.
    c. Split the Data Frame column wise based on a condition. For example - columns values greater than 10, column values less than 10, etc.
12. Create a Data Frame with columns Name and Age. Perform the following operations.
    a. Write a sample code to search for the Name/Names where Age is 20.
    b. Sort the Data Frame in ascending order of ages.
    c. Using Python loops, print the ages from the data frame and store it in a set.