ADVANCE OPERATIONS ON DATAFRAMES

ASSIGNMENTS

Q.1 Create a dataframe with following values

Indicator	Country	Year	Value
1	India	2005	6
2	India	2005	13
3	India	2005	10
4	India	2005	11
5	India	2005	5
1	India	2006	3
2	India	2006	2
3	India	2006	7
4	India	2006	3
5	India	2006	6

Now display the data in following manner

Country	Year	1	2	3	4	5
India	2005	6	13	10	11	5
India	2006	3	2	7	3	6

Q.2 Create a dataframe with following values

	Name	Position	City	Age	Sex
0	Maya	Manager	Mumbai	35	Female
1	Joy	Manager	Kolkata	37	Male
2	Lata	Manager	Lucknow	40	Female
3	Vibha	Programmer	Kolkata	29	Female
4	Sarita	Programmer	Mumbai	31	Female
5	Maya	Manager	Mumbai	26	Female
6	Lata	Manager	Kolkata	28	Female

a)Now display data in following pattern(Average of age is displayed)

City	Kolkata	Lucknow	Mumbai
Position			
Manager	32.5	40.0	30.5
Programmer	29.0	NaN	31.0

b)Now display data in following pattern(sum of age is displayed)

City	Kolkata	Lucknow	Mumbai
Position			
Manager	65.0	40.0	61.0
Programmer	29.0	NaN	31.0

c)Now display data in following pattern(First occurrence name is displayed)

City	Kolkata	Lucknow	Mumbai
Position			
Manager	Joy	Lata	Maya
Programmer	Vibha	NaN	Sarita

d)Now display data in following pattern

	Position	Kolkata	Lucknow	Mumbai
0	Manager	Joy, Lata	Lata	Maya, Maya
1	Programmer	Vibha	-	Sarita

Ans.

```
print (df.pivot_table(index='Position', columns='City', values='Name', aggfunc=',
'.join, fill_value='-')
      .reset_index()
      .rename_axis(None, axis=1))
```

Q.3 Create a dataframe with following values

	Brand	Price	Year
0	Samsung J7	22000	2015
1	Vivo V11	25000	2013
2	Honor play	27000	2018
3	Xiomi mi8	35000	2018

a) Sort the data on Brand name

b) Sort the data on Brand name in descending order

c) Sort the data on first year basis then price in ascending order

Q.4 What will be output after following program execution?

```
import pandas as pd
table = {
   "Name": ["anil", "vishal","manish","mohak"],
   "Age": [12,34,22,14],
}
df = pd.DataFrame(table)
print(df)
print (df.pivot_table(index="Name",columns="Name",values="Age"))
```

Q. 5 Create the following dataframe

	Name	Exam	Subject	Score
0	Abhay	Semester 1	Mathematics	62
1	Bhargav	Semester 1	Mathematics	47
2	Chitresh	Semester 1	Mathematics	55
3	Abhay	Semester 1	Science	74
4	Bhargav	Semester 1	Science	31
5	Chitresh	Semester 1	Science	77
6	Abhay	Semester 2	Mathematics	85
7	Bhargav	Semester 2	Mathematics	63
8	Chitresh	Semester 2	Mathematics	42
9	Abhay	Semester 2	Science	67
10	Bhargav	Semester 2	Science	89
11	Chitresh	Semester 2	Science	81

Now display data in following manner using pivot_table() function

		Score
Exam	Subject	
Semester 1	Mathematics	164
	Science	182
Semester 2	Mathematics	190
	Science	237

Score is result of sum of scores of that category(like semester 1 Mathematics total score)

Q.6 Based on the following data,write python program to find mean,median of grade and descriptive statistics of age and grade.

Name	age	favorite_color	grade
Viraj	20	blue	88
Mahak	19	blue	92
Amar	22	yellow	95
Sunit	21	green	70

Q. 7 Find the median, lower quartile, upper quartile of the following numbers. 12, 5, 22, 30, 7, 36, 14, 42, 15, 53, 25, 65 using python program.

Q,8 Find the first quartile for the following data set:
7, 9, 13, 4, 18, 3, 9, 10, 15, 8, 2, 6, 9

Q.9 Find the second quartile, or the median, for the following data set:
7, 9, 13, 4, 18, 3, 9, 10, 15, 8, 2, 6, 9

Q.10 Given the following distribution of returns, determine the lower quartile:
{10% 23% 12% 21% 14% 17% 16% 11% 15% 19%}

Q.11 Find the first and third quartiles of the set {3, 7, 8, 5, 12, 14, 21, 15, 18, 14}.

Q.12 Divide the following data set into quartiles: 2, 5, 6, 7, 10, 22, 13, 14, 16, 65, 45, 12.