**SHIVAJI COLLEGE, UNIVERSITY OF DELHI**
**DEPARTMENT OF COMPUTER SCIENCE**
**INTERNAL TEST (Academic Year 2023-24)**

Name of the Course : B.Sc Physical Science with Computer Science DSE II
Semester: IV
Name of the Paper : Data Exploration and Visualization
Duration : Maximum Marks: 20 Date of Test : ARPRIL 25, 2024

**SET-A**

1. Consider the DataFrames First and Second given below : (5)

| One | Two |
|---|---|
| 0 | "A" |
| 2 | "B" |
| 5 | "D" |
| 6 | "C" |

first

| One | Two |
|---|---|
| 0 | "B" |
| 1 | "C" |
| 5 | "E" |
| 2 | "A" |

second

Consider the following py thon code segment :
right = pd.merge(first, second, how=right', on='One')
left = pd.merge(first, second, how-'inner, on=Two')
Show the contents of new dataframes right and left

2. Consider the series S1 and S2 given below what will be the output (6)

| | S1 | | S2 |
|---|---|---|---|
| A | 1 | A | 5 |
| B | 2 | B | 6 |
| C | 3 | D | 7 |
| D | 4 | E | 8 |

i. S1[: 3] * 10
ii. S1 + S2
iii. S2 [: : -1] * 5

3. Consider the following DataFrame df. (6)

| Items | SugarType | Price |
|---|---|---|
| Yogurt | Low Fat | 45 |
| Chips | Regular | 30 |
| Soda | Low Fat | 50 |
| Yogurt | High Fat | 70 |
| Cake | Regular | 140 |
| Chips | NA | 40 |
| Yogurt | NA | 50 |

Give commands to perform the following operations:
i) List the name of unique items sold.
ii) Save the dataframe as a csv with separator as ;
iii) Delete the rows which have duplicate values of Items.
iv) Give the average price of all Low Fat items.
v) Fill all the missing values in the dataframe with the value 'Regular'
vi) Create a hierarchical index on Items

4. Consider the series a given below and give the output (3)
a = pd.Series([4, 1, 7, 1, 8, 9, 0, 8, 2, 3, 9])
(i) a.rank() (ii) a.rank(method='first')

Abha Vasal

**(ABHA VASAL)**

**SHIVAJI COLLEGE, UNIVERSITY OF DELHI**
**DEPARTMENT OF COMPUTER SCIENCE**
**INTERNAL TEST (Academic Year 2023-24)**

Name of the Course : B.Sc Physical Science with Computer Science DSE II
Semester: IV
Name of the Paper : Data Exploration and Visualization
Duration : Maximum Marks: 20 Date of Test : ARPRIL 25, 2024

(iii) a.rank(ascending =False)

**SET-B**

a. Consider the DataFrame df given below (5)

| EmployeeID | Department | Salary | Age |
|---|---|---|---|
| 1001 | English | 1000 | 23 |
| 1002 | English | 1002 | 34 |
| 1003 | English | 1004 | 39 |
| 1004 | English | 1000 | 43 |
| 1003 | Maths | 1004 | 34 |
| 1004 | Maths | 1005 | 43 |
| 1001 | Maths | 1006 | 53 |
| 1002 | Maths | 1002 | 43 |

1. Create a hierarchical index on Department and Employee ID.
2. df.stack)
3. df.unstack()
4. Give the summary level statistics for each column.
5. write command to Find the average age column

b. Consider the DataFrame Frame. Write python commands to perform (7)

| Name | age | weight | height |
|---|---|---|---|
| Ram | 15 | 45.6 | 140 |
| Ravi | 23 | 34.9 | 160 |
| Reena | 32 | 45.6 | 145 |
| Rita | 20 | 60.7 | 155 |
| Rishi | 33 | 54.7 | 170 |
| Romi | 21 | 34.6 | 144 |

a) Compute correlation of age with both weight and height
b) Sort frame in descending order of age
c) Find the index for the row with minimum age
d) Calculate cumulative sum for weight for all students
e) Set the height of Rita and Romi to NA
f) Define map function to convert values of Name column to upper case
g) Replace the value 32 with18 and 33 with 19 in age column

c. Give the python commands to create a dictionary with 5 keys ‘A’,’B’, ‘C’, ‘D’, ‘E’ and values as follows (5)

**(ABHA VASAL)**

**SHIVAJI COLLEGE, UNIVERSITY OF DELHI**
**DEPARTMENT OF COMPUTER SCIENCE**
**INTERNAL TEST (Academic Year 2023-24)**

Name of the Course : B.Sc Physical Science with Computer Science DSE II
Semester: IV
Name of the Paper : Data Exploration and Visualization
Duration : Maximum Marks: 20 Date of Test : ARPRIL 25, 2024

| Key | Value |
|---|---|
| A | List of numbers from 1 to 10 skipping 2 at a time. |
| B | List of Strings from A to E. |
| C | List of 5 numbers obtained using random normal distribution function. |
| D | List of 5 random integers from 20 to 30. |
| E | Square root of 5 random numbers from 50 to 70. |

Give python commands to perform the following
a) Create dataframe data using the above dictionary
b) Convert column A to index
c) Rename rest of the columns as Area, Temperature, Longitude, latitude
d) Delete the column longitude from the data
e) Save data as csv file with separator as ;

d. Differentiate between cut and qcut giving examples (3)

Abha Vasal

**(ABHA VASAL)**