Experiment Number: 6 & 7

Student Name: Garv Khurana **UID: 21BCS6615**

Branch: CSE-AIML **Section/Group: 21AML-9 (A)**

Semester: 3rd **Date of Performance: 12th October,2022**

Subject Code: 21CSH-238 **Subject Name: Python for Machine Learning**

---

1. **Aim/Overview of the practical:**

i). Create a dataframe that contains the student details (10 rows). It must contain 6 columns - UId, Name, Name_of_Program, Phone_No, City and E-mail_id. Create the values using Dictionary of series.

ii). Add a new Column 'marks' in the existing dataset and values in all the rows. Create 4 division A, B, C and D and show the distribution of marks for the same.

iii). Create a dataset of country details and its population (5 rows) and represent them in the form of a pie chart.

2. **Software Used:**
Google Colab and VS Code.

3.

Program: i

A. **Aim: Create a dataframe that contains the student details (10 rows). It must contain 6 columns - UId, Name, Name_of_Program, Phone_No, City and E-mail_id. Create the values using Dictionary of series.**

B. Program Code:

```
import csv import pandas as pd print("Creating dataframe using a list\n") data
= list(csv.reader(open('data.csv'))) df1 = pd.DataFrame(data[1:], columns=data[0]) print(df1)

print('-----------------------')

print("Creating dataframe using a dictionary\n") dic1
= {}

for count, key in enumerate(data[0]): dic1[key] =
   [value[count] for value in data[1:]]

df2 = pd.DataFrame(dic1) print(df2)
```

C. Output:

```
PS C:\VS Code\main.py> python -u "c:\VS Code\main.py\table.py"
Creating dataframe using a list

    uid     name  name of programn    phone no        city      email id marks
0  6659    Akrit              AIML  9805240621   California   a@gmail.com    98
1  6670  Sanchit               CSE      019722        Delhi   s@gmail.com    89
2  6570     Uday              AIML      543345       Jaipur   u@gmail.com    95
3  6626  Shubham       General CSE      788963   Chandigarh   d@gmail.com    99
4  6632    Anmol              AIML      267897     New York  as@gmail.com    92
5  6664   Rishav               MBA      438875        Delhi   r@gmail.com    98
6  6685  Millium  Hotel Management      569292  Los angeles   m@gmail.com    90
7  3791    Sahil                BD      560923   Chandigarh  sv@gmail.com    92
8  3850    Arpit                BD      287652        Delhi  sa@gmail.com    93
9  6394  Yashika              AIML      329843   Chandigarh   y@gmail.com    97
----------------------
Creating dataframe using a dictionary

    uid     name  name of programn    phone no        city      email id marks
0  6659    Akrit              AIML  9805240621   California   a@gmail.com    98
1  6670  Sanchit               CSE      019722        Delhi   s@gmail.com    89
2  6570     Uday              AIML      543345       Jaipur   u@gmail.com    95
3  6626  Shubham       General CSE      788963   Chandigarh   d@gmail.com    99
4  6632    Anmol              AIML      267897     New York  as@gmail.com    92
5  6664   Rishav               MBA      438875        Delhi   r@gmail.com    98
6  6685  Millium  Hotel Management      569292  Los angeles   m@gmail.com    90
7  3791    Sahil                BD      560923   Chandigarh  sv@gmail.com    92
8  3850    Arpit                BD      287652        Delhi  sa@gmail.com    93
9  6394  Yashika              AIML      329843   Chandigarh   y@gmail.com    97
PS C:\VS Code\main.py>
```

Program: ii

A. Aim: Add a new Column 'marks' in the existing dataset and values in all the rows.
Create 4 division A, B, C and D and show the distribution of marks for the same.

B. Program Code: import csv import pandas as pd
data = list(csv.reader(open('data.csv')))

dic1 = {}

```
for count, key in enumerate(data[0]): dic1[key] =
[(value[count]) for value in data[1:]] dic1['marks']
= list(map(int, dic1['marks']))

df = pd.DataFrame.from_dict(dic1, orient='index') print(df)

print('--------------------------')
above80 = []

print('Data for students with above marks 80') for
i in range(len(df.columns)):
    if int(df[i]['marks']) > 80:
        above80.append(df[i])

print(pd.concat(above80, axis=1))

print('--------------------------')
df = pd.DataFrame.from_dict(dic1)
a = df[(df['marks']) > 95]
print(f"Data for students with marks above 95:\n{a}") print('--------------------------')
b = df[(df['marks'] > 80) & (df['marks'] <= 95)]
print(f"Data for students with marks above 80 and below 95:\n{b}") print('---------
-----------------')
c = df[(df['marks'] > 60) & (df['marks'] <= 80)]
print(f"Data for students with marks  above 60 and below 80:\n{c}") print('---------
-----------------')
d = df[(df['marks'] > 40) & (df['marks'] <= 60)]
print(f"Data for students with marks  above 40 and below 60:\n{d}")
```

C. Output:

```
PS C:\VS Code\main.py> python -u "c:\VS Code\main.py\result_by_csv.py"
                        0           1           2  ...           7           8           9
uid                  6659        6670        6570  ...        3791        3850        6394
name                Akrit      Sanchit        Uday  ...       Sahil       Arpit      Yashika
name of programn     AIML         CSE        AIML  ...          BD          BD        AIML
phone no       9805240621      019722      543345  ...      560923      287652      329843
city           California       Delhi      Jaipur  ...   Chandigarh       Delhi   Chandigarh
email id      a@gmail.com  s@gmail.com  u@gmail.com  ...  sv@gmail.com  sa@gmail.com  y@gmail.com
marks                98          89          94  ...          78          59          97

[7 rows x 10 columns]
-------------------------
Data for students with above marks 80
                        0           1           2  ...           5                6           9
uid                  6659        6670        6570  ...        6664             6685        6394
name                Akrit      Sanchit        Uday  ...       Rishav          Millium      Yashika
name of programn     AIML         CSE        AIML  ...          MBA  Hotel Management        AIML
phone no       9805240621      019722      543345  ...      438875           569292      329843
city           California       Delhi      Jaipur  ...       Delhi       Los angeles   Chandigarh
email id      a@gmail.com  s@gmail.com  u@gmail.com  ...  r@gmail.com      m@gmail.com  y@gmail.com
marks                98          89          94  ...          98               90          97

[7 rows x 8 columns]
-------------------------
```

```
Data for students with marks above 95:
    uid     name name of programn    phone no        city     email id  marks
0  6659    Akrit            AIML  9805240621  California  a@gmail.com     98
3  6626  Shubham     General CSE      788963  Chandigarh  d@gmail.com     99
5  6664   Rishav             MBA      438875       Delhi  r@gmail.com     98
9  6394  Yashika            AIML      329843  Chandigarh  y@gmail.com     97
-------------------------
Data for students with marks above 80 and below 95:
    uid     name  name of programn phone no         city      email id  marks
1  6670  Sanchit               CSE   019722        Delhi   s@gmail.com     89
2  6570     Uday              AIML   543345       Jaipur   u@gmail.com     94
4  6632    Anmol              AIML   267897     New York  as@gmail.com     92
6  6685  Millium  Hotel Management   569292  Los angeles   m@gmail.com     90
-------------------------
Data for students with marks  above 60 and below 80:
    uid   name name of programn phone no        city      email id  marks
7  3791  Sahil              BD   560923  Chandigarh  sv@gmail.com     78
-------------------------
Data for students with marks  above 40 and below 60:
    uid   name name of programn phone no   city      email id  marks
8  3850  Arpit              BD   287652  Delhi  sa@gmail.com     59
PS C:\VS Code\main.py>
```

Program: iii

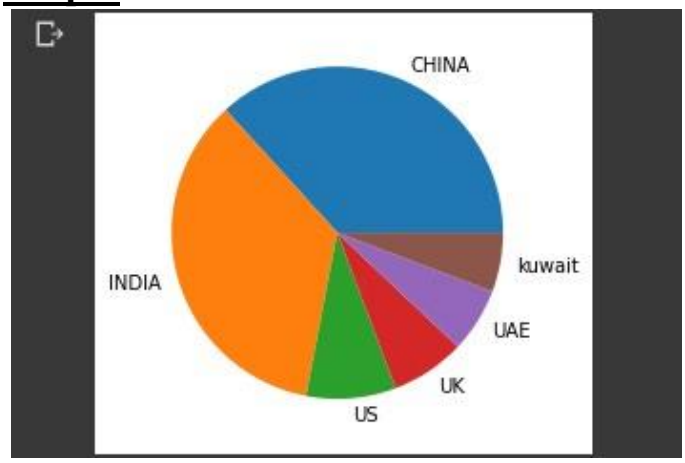
DEPARTMENT OF ACADEMIC AFFAIRS
Discover. Learn. Empower.
NAAC GRADE A+ ACCREDITED UNIVERSITY

A. <u>Aim</u>: Create a dataset of country details and its population (5 rows) and represent them in the form of a pie chart. B. <u>Program Code</u>:

```
import pandas as pd import
matplotlib.pyplot as plt import
numpy as np

country=['CHINA', 'INDIA', 'US', 'UK', 'UAE', 'kuwait'] data=[1397897728,
1339330514, 332475723, 275122131,238181034, 219463862]; plt.pie(data, labels
=country) plt.show()
```

C. <u>Output</u>:



<u>Learning outcomes (What I have learnt)</u>:

- **Knowledge of pandas library.**
- **Pandas Series.**
- **Pandas Dataframe.**

Evaluation Grid (To be created as per the SOP and Assessment guidelines by the faculty):

Sr. No.	Parameters	Marks Obtained	Maximum Marks
1.	Viva		10
2.	Performance		12
3.	Worksheet		8