#Name: Nikhil Irkar

#Roll No: 518

#PRN No:202201090076

## **#\*THFORY ASSIGNMENT DATASET4\***

- "4) Find the statistical Analysis on Employee Info (Refer Data Set 4)
- a. Find the total no of male and female employee
- b. Find the total no of single married and divorced employee
- c. Find the employee ID who is manager
- d. Find the all employee ID who is Supervisor
- e. Clean the dataset if record is empty with 0 values or delete incomplete data row
- f. Find the name of employee who is working as manager and from Pune
- g. Find the employee who's salary is greater than 1,00,000/-"

import pandas as pd import
numpy as np
df=pd.read\_csv("dataset4.cs
v")
print(df)

#Q1 Total no of male & female in dataset

female\_count=df[df['Gender']=='female'].value\_counts().s

um() print("Total no of female in company

are:",female\_count) male\_count=10-female\_count

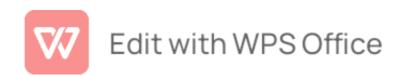
print("Total no male in company are: ",male\_count)

OUTPUT: Total no of female in company are: 3

Total no male in company are: 7

#Q2 Total no of single married and divorced employee single\_emp\_count=df[df['Status']=='single'].value\_counts().sum() married\_emp\_count=df[df['Status']=='married'].value\_counts().sum() divorced\_emp\_count=10-single\_emp\_count-married\_emp\_count

print("Total no of single, married and divorced employee are



```
:",single emp count,married emp count,divorced emp count)
OUTPUT: Total no of single, married and divorced employee are: 5 3 2
#Q3 employee ID who is manager
emp=df[df['Designation']=='Manager']['Emp id']
emp details=df[df['Designation']=='Manager']
print("The employee lds who is manager are:\n ",emp)
print("\n& their details is:\n",emp details)
OUTPUT:
The employee lds who is manager are:
2 3
5 6
7 8
Name: Emp_id, dtype: int64
& their details is:
  Emp id Employee name City Designation Salary Gender Status
          Sanvi Pune Manager 100000 female single
0
   1
2
          Jayesh Nashik Manager 90500 male single
          Pranav Pune Manager 100000 male divorced 7 8
                                                                        Raja Nashik
Manager 90500 male married
#Q4 All employee ID who is Supervisor
#1st way using boolean index
supervisor_ids = df[df['Designation'] == 'Supervisor']['Emp_id'] print("All the employee lds who is
supervisor are:
\n",supervisor ids.to string(index=False))
#2nd way using loc function
supervisor ids = df.loc[df['Designation'] == 'Supervisor', 'Emp id'] print("\n All the employee Ids who is
supervisor are:
\n",supervisor_ids.to_string(index=False))
#3rd way using query function
supervisor ids = df.query("Designation == 'Supervisor'")['Emp id'] print("\n All the employee Ids who is
supervisor are:
\n",supervisor ids.to string(index=False))
OUTPUT:
All the employee lds who is supervisor are:
5
10
All the employee lds who is supervisor are: 5
10
All the employee Ids who is supervisor are: 5
10
```



#Q5 Clean the dataset if record is empty with 0 values or delete incomplete data row

- # Replace empty records with 0 values df.fillna(0, inplace=True)
- # Delete incomplete data rows df.dropna(inplace=True)

print("Dataset after cleaning & deleting incompleting data row is:\n ",df) OUTPUT:

Dataset after cleaning & deleting incompleting data row is:

4 Gouri Nashik Sr. Manager 100500 female married 5 Mahesh Pune Supervisor 85000 male single 6 Pranav Pune Manager 100000 male divorced

6 Pranav Pune Manager 100000 male divorced 7 Saksham Pune Sr. Manager 150000 male single

8 Raja Nashik Manager 90500 male married
9 Sunil Nashik Sr. Manager 100500 male single 9 10 Radha Pune Supervisor 85000 female divorced

#Q6 the name of employee who is working as manager and from Pune emp\_name=df[(df['Designation']=='Manager') & (df['City']=='Pune')]['Employee\_name'] print("The empoyee who is workinng as manager & from pune are:\n",emp\_name.to\_string(header=False,index=False)) OUTPUT: the empoyee who is workinng as manager & from pune are:

Sanvi Pranav

#Q7 the employee who's salary is greater than 1,00,000/-

emp\_grt\_sal=df[df['Salary']>100000]['Employee\_name']

print("The employee who's salary is greater than 1,00,000 are:\n",emp\_grt\_sal)

OUTPUT:

The employee who's salary is greater than 1,00,000 are:

- 1 Mrunmayee
- 3 Gouri
- 6 Saksham
- 8 Sunil

Name: Employee\_name, dtype: object



