

NAME – ANKIT KUMAR
ROLL NUMBER – 414
DIVISION – D
BATCH – D1
PRN – 202201040161

EDS ASSINMENT 4

PANDAS OPERATIONS

```
import pandas as pd

df = pd.read_csv("dataset4.csv")
print("Original DataFrame:")
print(df)

#Filter employees from Pune
pune_employees = df[df['City'] == 'Pune']
print("Employees from Pune:")
print(pune_employees)

#Calculate the average salary
average_salary = df['Salary'].mean()
print("Average Salary:", average_salary)

#Count employees by gender
gender_counts = df['Gender'].value_counts()
print("Employee Gender Counts:")
print(gender_counts)

#Sort employees by salary in descending order
sorted_df = df.sort_values('Salary', ascending=False)
print("Employees Sorted by Salary:")
print(sorted_df)

#Group employees by designation and calculate the average salary for each group
designation_salary = df.groupby('Designation')['Salary'].mean()
print("Average Salary by Designation:")
print(designation_salary)

#Add a new column 'Age' with default value 0
df['Age'] = 0
print("DataFrame with 'Age' Column:")
print(df)
```

```
#Select only the 'Name' and 'Salary' columns
selected_columns = df[['Name', 'Salary']]
print("Selected Columns:")
print(selected_columns)

#Filter employees with a salary greater than 100,000
high_salary_employees = df[df['Salary'] > 100000]
print("Employees with High Salary:")
print(high_salary_employees)

# Calculate the total salary for each city
city_salary_total = df.groupby('City')['Salary'].sum()
print("Total Salary by City:")
print(city_salary_total)

# Calculate the age based on the current year and assuming all employees are born
in 1980
current_year = 2023
df['Age'] = current_year - 1980
print("DataFrame with Age Column:")
print(df)
```

Dataset:

ID	Name	City	Designation	Salary	Gender	Marital Status		
1	Sanvi	Pune	Manager	100000	female	single		
2	Mrunmaye	Pune	Sr. Manager	150000	male	married		
3	Jayesh	Nashik	Manager	90500	male	single		
4	Gouri	Nashik	Sr. Manager	100500	female	married		
5	Mahesh	Pune	Supervisor	85000	male	single		
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		
10	Radha	Pune	Supervisor	85000	female	divorced		

OUTPUT:

```
[Running] python -u "e:\_school\sem 2\EDS\assign_4.py"
```

Original DataFrame:

	ID	Name	City	Designation	Salary	Gender	Marital Status
0	1	Sanvi	Pune	Manager	100000	female	single
1	2	Mrunmayee	Pune	Sr. Manager	150000	male	married
2	3	Jayesh	Nashik	Manager	90500	male	single
3	4	Gouri	Nashik	Sr. Manager	100500	female	married
4	5	Mahesh	Pune	Supervisor	85000	male	single
5	6	Pranav	Pune	Manager	100000	male	divorced
6	7	Saksham	Pune	Sr. Manager	150000	male	single
7	8	Raja	Nashik	Manager	90500	male	married
8	9	Sunil	Nashik	Sr. Manager	100500	male	single
9	10	Radha	Pune	Supervisor	85000	female	divorced

Employees from Pune:

	ID	Name	City	Designation	Salary	Gender	Marital Status
0	1	Sanvi	Pune	Manager	100000	female	single
1	2	Mrunmayee	Pune	Sr. Manager	150000	male	married
4	5	Mahesh	Pune	Supervisor	85000	male	single
5	6	Pranav	Pune	Manager	100000	male	divorced
6	7	Saksham	Pune	Sr. Manager	150000	male	single
9	10	Radha	Pune	Supervisor	85000	female	divorced

Average Salary: 105200.0

Employee Gender Counts:

Gender

male 7

female 3

Name: count, dtype: int64

Employees Sorted by Salary:

	ID	Name	City	Designation	Salary	Gender	Marital Status
1	2	Mrunmayee	Pune	Sr. Manager	150000	male	married
6	7	Saksham	Pune	Sr. Manager	150000	male	single
3	4	Gouri	Nashik	Sr. Manager	100500	female	married
8	9	Sunil	Nashik	Sr. Manager	100500	male	single
0	1	Sanvi	Pune	Manager	100000	female	single
5	6	Pranav	Pune	Manager	100000	male	divorced
2	3	Jayesh	Nashik	Manager	90500	male	single
7	8	Raja	Nashik	Manager	90500	male	married
4	5	Mahesh	Pune	Supervisor	85000	male	single
9	10	Radha	Pune	Supervisor	85000	female	divorced

Average Salary by Designation:

Designation

Manager 95250.0

Sr. Manager 125250.0

Supervisor 85000.0

Name: Salary, dtype: float64

DataFrame with 'Age' Column:

	ID	Name	City	Designation	Salary	Gender	Marital Status	Age
0	1	Sanvi	Pune	Manager	100000	female	single	0
1	2	Mrunmayee	Pune	Sr. Manager	150000	male	married	0
2	3	Jayesh	Nashik	Manager	90500	male	single	0
3	4	Gouri	Nashik	Sr. Manager	100500	female	married	0
4	5	Mahesh	Pune	Supervisor	85000	male	single	0
5	6	Pranav	Pune	Manager	100000	male	divorced	0
6	7	Saksham	Pune	Sr. Manager	150000	male	single	0
7	8	Raja	Nashik	Manager	90500	male	married	0
8	9	Sunil	Nashik	Sr. Manager	100500	male	single	0
9	10	Radha	Pune	Supervisor	85000	female	divorced	0

Selected Columns:

	Name	Salary
0	Sanvi	100000
1	Mrunmayee	150000
2	Jayesh	90500
3	Gouri	100500
4	Mahesh	85000
5	Pranav	100000
6	Saksham	150000
7	Raja	90500
8	Sunil	100500
9	Radha	85000

Employees with High Salary:

	ID	Name	City	Designation	Salary	Gender	Marital Status	Age
1	2	Mrunmayee	Pune	Sr. Manager	150000	male	married	0
3	4	Gouri	Nashik	Sr. Manager	100500	female	married	0
6	7	Saksham	Pune	Sr. Manager	150000	male	single	0
8	9	Sunil	Nashik	Sr. Manager	100500	male	single	0

Total Salary by City:

City

Nashik 382000

Pune 670000

Name: Salary, dtype: int64

DataFrame with Age Column:

	ID	Name	City	Designation	Salary	Gender	Marital Status	Age
0	1	Sanvi	Pune	Manager	100000	female	single	43
1	2	Mrunmayee	Pune	Sr. Manager	150000	male	married	43
2	3	Jayesh	Nashik	Manager	90500	male	single	43
3	4	Gouri	Nashik	Sr. Manager	100500	female	married	43
4	5	Mahesh	Pune	Supervisor	85000	male	single	43
5	6	Pranav	Pune	Manager	100000	male	divorced	43

6	7	Saksham	Pune	Sr. Manager	150000	male	single	43
7	8	Raja	Nashik	Manager	90500	male	married	43
8	9	Sunil	Nashik	Sr. Manager	100500	male	single	43
9	10	Radha	Pune	Supervisor	85000	female	divorced	43

[Done] exited with code=0 in 0.613 seconds