NAME = Om Santosh Bhutkar.

Div = G

Batch = G-1

ROll No. = 707.

PRN No. :- 202201040111.

Subject :- EDS (Assignment No.4)

Double-click (or enter) to edit

```
import pandas as pd


df=pd.read_csv("/content/dataset4.csv")
print(df)
data = {
    'ID': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],
    'Name': ['Sanvi', 'Mrunmayee', 'Jayesh', 'Gouri', 'Mahesh', 'Pranav', 'Saksham', 'Raja', 'Sur
    'City': ['Pune', 'Pune', 'Nashik', 'Nashik', 'Pune', 'Pune', 'Pune', 'Nashik', 'Nashik', 'Pur
    'Position': ['Manager', 'Sr. Manager', 'Manager', 'Sr. Manager', 'Supervisor', 'Manager', 'Sr
    'Salary': [100000, 150000, 90500, 100500, 85000, 100000, 150000, 90500, 100500, 85000],
    'Gender': ['female', 'male', 'male', 'female', 'male', 'male', 'male', 'male', 'male', 'fema
    'Marital_Status': ['single', 'married', 'single', 'married', 'single', 'divorced', 'single',
}
df = pd.DataFrame(data)
```

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

```
# 1. How many records are there in the dataset?
num_records = len(df)
print("Number of records:", num_records)
```

```
Number of records: 10
```

```
# 2. What are the unique cities in the dataset?
unique_cities = df['City'].unique()
print("Unique cities:", unique_cities)
```

```
Unique cities: ['Pune' 'Nashik']
```

```
# 3. What is the average salary of all employees?
average_salary = df['Salary'].mean()
print("Average salary:", average_salary)
```

```
Average salary: 105200.0
```

```
# 3. What is the average salary of all employees?
average_salary = df['Salary'].mean()
print("Average salary:", average_salary)
```

```
Average salary: 105200.0
```

```
# 4. How many male employees are there in the dataset?
male_count = df[df['Gender'] == 'male'].shape[0]
print("Number of male employees:", male_count)
```

```
Number of male employees: 7
```

```
# 5. What is the highest salary in the dataset?
highest_salary = df['Salary'].max()
print("Highest salary:", highest_salary)
```

```
Highest salary: 150000
```

```
# 6. How many employees are single and earn more than 100,000?
high_earning_single_count = df[(df['Marital_Status'] == 'single') & (df['Salary'] > 100000)].shap
print("Number of single employees earning more than 100,000:", high_earning_single_count)
```

```
Number of single employees earning more than 100,000: 2
```

```
# 7. What is the average salary of female employees in Nashik?
average_salary_nashik_female = df[(df['City'] == 'Nashik') & (df['Gender'] == 'female')]['Salary
print("Average salary of female employees in Nashik:", average_salary_nashik_female)
```

```
Average salary of female employees in Nashik: 100500.0
```

```
# 8. How many divorced employees are there in Pune?
divorced_count_pune = df[(df['City'] == 'Pune') & (df['Marital_Status'] == 'divorced')].shape[0]
print("Number of divorced employees in Pune:", divorced_count_pune)
```

```
Number of divorced employees in Pune: 2
```

```
# 9. What is the total salary expense for the company?
total_salary_expense = df['Salary'].sum()
print("The total salary expense for the company is:", total_salary_expense)
```

```
The total salary expense for the company is: 1052000
```

```
# 10. What is the average salary of all employees?
average_salary = df['Salary'].mean()
print("The average salary of all employees is:",average_salary)
```

```
The average salary of all employees is: 105200.0
```