In [25]:
```python
#TASK1
import pandas as pd
df = pd.read_csv("data 1.csv")
df.head()
```

Out[25]:

	rank	discipline	phd	service	sex	salary
0	Prof	B	56	49	Male	186960
1	Prof	A	12	6	Male	93000
2	Prof	A	23	20	Male	110515
3	Prof	A	40	31	Male	131205
4	Prof	B	20	18	Male	104800

In [27]:
```python
import pandas as pd
df = pd.read_csv("cars.csv")
df.head(7)
```

Out[27]:

	HP	MPG	VOL	SP	WT
0	49	53.700681	89	104.185353	28.762059
1	55	50.013401	92	105.461264	30.466833
2	55	50.013401	92	105.461264	30.193597
3	70	45.696322	92	113.461264	30.632114
4	53	50.504232	92	104.461264	29.889149
5	70	45.696322	89	113.185353	29.591768
6	55	50.013401	92	105.461264	30.308480

In [26]:
```python
import pandas as pd
df = pd.read_csv("cars.csv")
df.tail(5)
```

Out[26]:

	HP	MPG	VOL	SP	WT
76	322	36.900000	50	169.598513	16.132947
77	238	19.197888	115	150.576579	37.923113
78	263	34.000000	50	151.598513	15.769625
79	295	19.833733	119	167.944460	39.423099
80	236	12.101263	107	139.840817	34.948615

In [17]:
```python
import pandas as pd
print("DataFrame dimensions:", df.shape)
```

```
DataFrame dimensions: (81, 5)
```

```python
In [18]: import pandas as pd
         print("Descriptive Statistics:\n")
         print(df.describe())
```

```
Descriptive Statistics:

               HP         MPG          VOL          SP          WT
count   81.000000   81.000000   81.000000   81.000000   81.000000
mean   117.469136   34.422076   98.765432  121.540272   32.412577
std     57.113502    9.131445   22.301497   14.181432    7.492813
min     49.000000   12.101263   50.000000   99.564907   15.712859
25%     84.000000   27.856252   89.000000  113.829145   29.591768
50%    100.000000   35.152727  101.000000  118.208698   32.734518
75%    140.000000   39.531633  113.000000  126.404312   37.392524
max    322.000000   53.700681  160.000000  169.598513   52.997752
```

```python
In [19]: #TASK2
         import pandas as pd

         data = {
             'Name': ['John', 'Jane', 'Babu', 'Peter', 'Leju'],
             'Age': [25, 30, 35, 40, 55],
             'City': ['New York', 'London', 'Paris', 'UK', 'Germany']
         }

         df = pd.DataFrame(data)
```

```python
In [20]: print("Original DataFrame:")
         print(df)
```

```
Original DataFrame:
    Name  Age       City
0   John   25   New York
1   Jane   30     London
2   Babu   35      Paris
3  Peter   40         UK
4   Leju   55    Germany
```

```python
In [21]: print("First 5 Rows of DataFrame:")
         print(df.head())
```

```
First 5 Rows of DataFrame:
    Name  Age       City
0   John   25   New York
1   Jane   30     London
2   Babu   35      Paris
3  Peter   40         UK
4   Leju   55    Germany
```

```python
In [22]: print("Descriptive Statistics:")
         print(df.describe())
```

```
Descriptive Statistics:
            Age
count   5.000000
mean   37.000000
std    11.510864
min    25.000000
25%    30.000000
50%    35.000000
75%    40.000000
max    55.000000
```

In [23]:
```python
print("DataFrame Shape (Rows, Columns):", df.shape)
```

DataFrame Shape (Rows, Columns): (5, 3)

In [24]:
```python
print("DataFrame Info:")
df.info()
```

```
DataFrame Info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5 entries, 0 to 4
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Name    5 non-null      object
 1   Age     5 non-null      int64
 2   City    5 non-null      object
dtypes: int64(1), object(2)
memory usage: 252.0+ bytes
```

In [4]:
```python
import pandas as pd

data = pd.DataFrame({
    'Department': ['Finance', 'Sales', 'Finance', 'IT', 'Finance', 'Sales', 'HR',
                   'HR', 'HR', 'Marketing', 'Operations', 'Finance', 'Sales', 'H
                   'HR', 'HR'],
    'EmployeeID': [1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008, 1009, 1010,
                   1011, 1012, 1013, 1014, 1015, 1016, 1017, 1018, 1019, 1020,
                   1021, 1022],
    'Salary': [48313, 60744, 75589, 70597, 74578, 51008, 78667, 52747, 65618, 54
               54853, 61529, 42592, 58034, 58035, 58036, 58037, 58038, 58039, 58
               58041, 58042],
    'YearsOfExperience': [6, 3, 4, 3, 7, 7, 5, 2, 4, 4,
                          8, 1, 1, 2, 3, 4, 5, 6, 7, 8,
                          9, 10],
    'Age': [42, 24, 34, 48, 29, 39, 30, 43, 26, 44,
            30, 34, 33, 34, 35, 36, 37, 38, 39, 40,
            41, 40]
})

data.head()
```

Out[4]:

	Department	EmployeeID	Salary	YearsOfExperience	Age
0	Finance	1001	48313	6	42
1	Sales	1002	60744	3	24
2	Finance	1003	75589	4	34
3	IT	1004	70597	3	48
4	Finance	1005	74578	7	29

In [5]:
```python
mean_salaries = data.groupby('Department')['Salary'].mean()
print("Average Salary by Department:\n")
print(mean_salaries)
```

```
Average Salary by Department:

Department
Finance       63362.000000
HR            61528.166667
IT            70597.000000
Marketing     42592.000000
Operations    58034.000000
Sales         56113.857143
Name: Salary, dtype: float64
```

In [6]:
```python
multi_group = data.groupby(['Department', 'Age']).agg({
    'Salary': ['mean', 'max', 'min'],
    'YearsOfExperience': ['mean', 'min']
})
print("\nGrouped by Department and Age with multiple aggregations:\n")
print(multi_group)
```

```
Grouped by Department and Age with multiple aggregations:

                    Salary                     YearsOfExperience
                    mean      max     min           mean    min
Department   Age
Finance      26     65618.0   65618   65618          4.0     4
             29     74578.0   74578   74578          7.0     7
             34     75589.0   75589   75589          4.0     4
             35     58035.0   58035   58035          3.0     3
             39     58039.0   58039   58039          7.0     7
             42     48313.0   48313   48313          6.0     6
HR           30     66760.0   78667   54853          6.5     5
             34     61529.0   61529   61529          1.0     1
             37     58037.0   58037   58037          5.0     5
             40     58042.0   58042   58042         10.0    10
             41     58041.0   58041   58041          9.0     9
IT           48     70597.0   70597   70597          3.0     3
Marketing    33     42592.0   42592   42592          1.0     1
Operations   34     58034.0   58034   58034          2.0     2
Sales        24     60744.0   60744   60744          3.0     3
             36     58036.0   58036   58036          4.0     4
             38     58038.0   58038   58038          6.0     6
             39     51008.0   51008   51008          7.0     7
             40     58040.0   58040   58040          8.0     8
             43     52747.0   52747   52747          2.0     2
             44     54184.0   54184   54184          4.0     4
```

In [ ]: