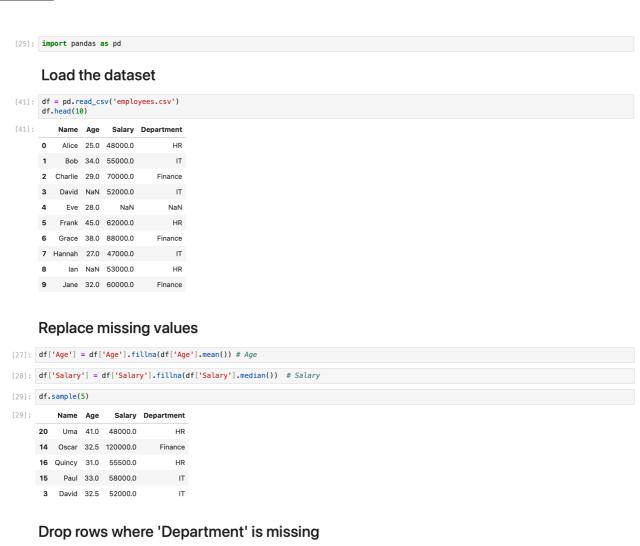
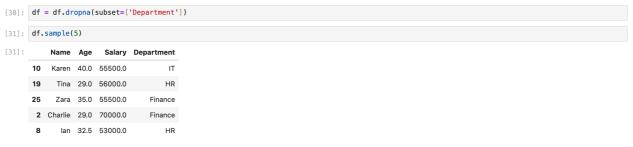
## **Assignment No. 11**

- B. Load the dataset Employees.csv into a Pandas DataFrame.
- i. Display the first 10 rows.
- ii. The dataset contains columns Age, Salary, and Department. Some rows have missing values. Write a script to: Replace missing values in Age with the mean age. Replace missing values in Salary with the median salary. Drop rows where Department is missing. iii. Create a new column in a DataFrame that categorizes employees based on their salary:
  - Low (Salary < 50,000)</li>
  - Medium (50,000 ≤ Salary < 100,000)
  - High (Salary > 100,000)
- iv. Group the DataFrame by the Department column and calculate the average salary for each department.

## Code:





## Categorize employees based on salary

```
[39]: def categorize_salary(salary):
    if salary < 50000:
        return 'Low'
    elif 50000 <= salary < 100000:
        return 'Medium'
    else:
        return 'High'

employees['Salary_Category'] = employees['Salary'].apply(categorize_salary)

print("\nDataset after handling missing values and categorization:")
print(employees)</pre>

Dataset after handling missing values and categorization:
```

```
Dataset after handling missing values and categorization:

Name Age Salary Department Salary_Category
Alice 25.0 48000.0 HR Low
55000.0 IT Medium
Charlie 29.0 70000.0 Finance Medium
         David 32.5
Frank 45.0
                                                  IT
HR
                                52000.0
                                                                            Medium
                                62000.0
                                                                           Medium
       Grace 38.0
Hannah 27.0
Ian 32.5
Jane 32.0
                                                Finance
IT
                                88000.0
                                                                           Medium
                                47000.0
                                                                               Low
                                                                           Medium
                                53000.0
                                                        HR
                                                 Finance
10
11
13
14
15
         Karen 40.0
                                55500.0
                                                       IT
                                                                           Medium
                                                 Finance
          Leo 32.5
Nina 30.0
                               78000.0
46000.0
                                                                           Medium
                                                                           Low
High
Medium
                                                        HR
                                                Finance
IT
         Oscar 32.5
Paul 33.0
                             120000.0
58000.0
       Quincy 31.0
Rachel 32.5
Steve 36.0
16
17
18
                                55500.0
                                                        HR
                                                                           Medium
                               72000.0
67000.0
                                                                            Medium
         Steve 36.0
Tina 29.0
Uma 41.0
                                                                           Medium
                                                        IT
19
20
                                56000.0
48000.0
                                                        HR
HR
                                                                           Medium
                                                                                Low
       Victor 39.0
Wendy 24.0
Yvonne 32.5
21
22
                                54000.0
                                                Finance
IT
                                                                           Medium
                                51000.0
                                                                           Medium
24
                                49000.0
                                                        HR
                                                                               Low
           Zara 35.0
                               55500.0
                                                                           Medium
```

## Average salary by department

```
[40]: average_salary_by_department = employees.groupby('Department')['Salary'].mean()
print("\nAverage salary by department:")
print(average_salary_by_department)
Average salary by department:
```

Department
Finance 74687.500000
HR 52187.500000
IT 55071.428571
Name: Salary, dtype: float64