Data Loading and Overview  
1. Read a CSV file into a Pandas DataFrame

➤ Load data.csv and display the first 5 rows.

import pandas as pd

dt = pd.read\_excel('ASS.xlsx')

print(dt.head(5))

2. Check the shape of a DataFrame

➤ How many rows and columns are present?

import pandas as pd

dt = pd.read\_excel('ASS.xlsx')

print(dt.shape)

3. Get summary statistics

➤ Use a method to get min, max, mean, etc., for numeric columns. Data Selection and Filtering

import pandas as pd

dt = pd.read\_excel('ASS.xlsx')

print(dt.describe())

4. Select a single column  
➤ Extract the "Age" column as a Series.

import pandas as pd

dt = pd.read\_excel('ASS.xlsx')

print(dt.[“age”])

5. Filter rows based on condition  
➤ Show rows where "Salary" > 50000.

Print(dt[“salary”]>50000)

6. Filter multiple conditions  
➤ Display rows where "Department" == 'HR' and "Age" > 30.

Hr = df[(df["Department"] == "HR") & (df["Age"] > 30)]

print(Hr)

Data Cleaning  
7. Check for missing values

➤ Find which columns have NaN values and how many.

missing\_values = df.isnull().sum()

print(missing\_values)

8. Replace missing values

➤ Fill NaN values in "Salary" with 0.

df["salary"].fillna(0, inplace=True)

9. Remove duplicate rows  
➤ Drop duplicates and reset the index.

da = dt.drop\_duplicates()

print(da)

Data Aggregation and Sorting  
10. Sort the DataFrame by a column

➤ Sort rows by "Age" in descending order.

sorted\_df = dt.sort\_values(by="age", ascending=True)

print(sorted\_df)

11. Group by and aggregate

➤ Group by "Department" and find the average "Salary".

avg\_salary\_by\_dept = dt.groupby("departemnt")["salary"].mean()

print(avg\_salary\_by\_dept)

12. Count unique values  
➤ How many unique departments are there in the "Department" column?

unique\_dep = dt["departemnt"].nunique()

print("Number of unique departments:", unique\_dep)