```
import pandas as pd
```

```
data=pd.read_csv("Data.csv")
```

```
data
```

| | Country | Age | Salary | Purchased |
|---|---|---|---|---|
| 0 | India | 40.0 | 90000.0 | Yes |
| 1 | Georgia | 44.0 | 69490.0 | No |
| 2 | France | 44.0 | 72000.0 | No |
| 3 | Spain | 27.0 | 48000.0 | Yes |
| 4 | Germany | 30.0 | 54000.0 | No |
| 5 | Spain | 38.0 | 61000.0 | No |
| 6 | Germany | 40.0 | NaN | Yes |
| 7 | France | 35.0 | 58000.0 | Yes |
| 8 | Spain | NaN | 52000.0 | No |
| 9 | France | 48.0 | 79000.0 | Yes |
| 10 | Germany | 50.0 | 83000.0 | No |
| 11 | France | 37.0 | 67000.0 | Yes |

```
#slicing
data.head(3)
```

| | Country | Age | Salary | Purchased |
|---|---|---|---|---|
| 0 | India | 40.0 | 90000.0 | Yes |
| 1 | Georgia | 44.0 | 69490.0 | No |

```
data1=pd.read_csv("country.csv")
```

```
data1
```

| | Country | Surface area | Population | Population density | Sex ratio |
|---|---|---|---|---|---|
| 0 | India | 3287263 | 1339180 | 450.4 | 107.6 |
| 1 | Georgia | 69700 | 3912 | 56.3 | 91.4 |
| 2 | France | 551500 | 64980 | 118.7 | 96.7 |
| 3 | Spain | 505944 | 46354 | 92.9 | 96.2 |

```
pd.merge(data.head(4),data1,on="Country")
```

| | Country | Age | Salary | Purchased | Surface area | Population | Population density | Sex ratio |
|---|---|---|---|---|---|---|---|---|
| 0 | India | 40.0 | 90000.0 | Yes | 3287263 | 1339180 | 450.4 | 107.6 |
| 1 | Georgia | 44.0 | 69490.0 | No | 69700 | 3912 | 56.3 | 91.4 |
| 2 | France | 44.0 | 72000.0 | No | 551500 | 64980 | 118.7 | 96.7 |
| 3 | Spain | 27.0 | 48000.0 | Yes | 505944 | 46354 | 92.9 | 96.2 |

```
data2=pd.DataFrame({'Int_Rate':[2,1,2,3],'Ind_Gdp':[50,45,35,65],'year':[2016,2017,2018,2019]})
data3=pd.DataFrame({'Low_Tier':[50,45,35,65],'Unemployment':[1,3,5,7],'year1':[2006,2021,2013,2022]})
```

```
data2
```

| | Int_Rate | Ind_Gdp | year |
|---|---|---|---|
| **0** | 2 | 50 | 2016 |
| **1** | 1 | 45 | 2017 |
| **2** | 2 | 35 | 2018 |
| **3** | 3 | 65 | 2019 |

```
data3
```

| | Low_Tier | Unemployment | year1 |
|---|---|---|---|
| **0** | 50 | 1 | 2006 |
| **1** | 45 | 3 | 2021 |
| **2** | 35 | 5 | 2013 |
| **3** | 65 | 7 | 2022 |

```
data2.join(data3)
```

| | Int_Rate | Ind_Gdp | year | Low_Tier | Unemployment | year1 |
|---|---|---|---|---|---|---|
| **0** | 2 | 50 | 2016 | 50 | 1 | 2006 |
| **1** | 1 | 45 | 2017 | 45 | 3 | 2021 |
| **2** | 2 | 35 | 2018 | 35 | 5 | 2013 |
| **3** | 3 | 65 | 2019 | 65 | 7 | 2022 |

```
newdata=pd.concat([data,data1])
```

```
newdata
```

| | Country | Age | Salary | Purchased | Surface area | Population | Population density | Sex ratio |
|---|---|---|---|---|---|---|---|---|
| **0** | India | 40.0 | 90000.0 | Yes | NaN | NaN | NaN | NaN |
| **1** | Georgia | 44.0 | 69490.0 | No | NaN | NaN | NaN | NaN |
| **2** | France | 44.0 | 72000.0 | No | NaN | NaN | NaN | NaN |
| **3** | Spain | 27.0 | 48000.0 | Yes | NaN | NaN | NaN | NaN |
| **4** | Germany | 30.0 | 54000.0 | No | NaN | NaN | NaN | NaN |
| **5** | Spain | 38.0 | 61000.0 | No | NaN | NaN | NaN | NaN |
| **6** | Germany | 40.0 | NaN | Yes | NaN | NaN | NaN | NaN |
| **7** | France | 35.0 | 58000.0 | Yes | NaN | NaN | NaN | NaN |
| **8** | Spain | NaN | 52000.0 | No | NaN | NaN | NaN | NaN |
| **9** | France | 48.0 | 79000.0 | Yes | NaN | NaN | NaN | NaN |
| **10** | Germany | 50.0 | 83000.0 | No | NaN | NaN | NaN | NaN |
| **11** | France | 37.0 | 67000.0 | Yes | NaN | NaN | NaN | NaN |
| **0** | India | NaN | NaN | NaN | 3287263.0 | 1339180.0 | 450.4 | 107.6 |
| **1** | Georgia | NaN | NaN | NaN | 69700.0 | 3912.0 | 56.3 | 91.4 |
| **2** | France | NaN | NaN | NaN | 551500.0 | 64980.0 | 118.7 | 96.7 |
| **3** | Spain | NaN | NaN | NaN | 505944.0 | 46354.0 | 92.9 | 96.2 |

```
null_check = newdata.isnull()
```

```
null_check
```

| | Country | Age | Salary | Purchased | Surface area | Population | Population density | Sex ratio |
|---|---|---|---|---|---|---|---|---|
| **0** | False | False | False | False | True | True | True | True |
| **1** | False | False | False | False | True | True | True | True |
| **2** | False | False | False | False | True | True | True | True |
| **3** | False | False | False | False | True | True | True | True |
| **4** | False | False | False | False | True | True | True | True |
| **5** | False | False | False | False | True | True | True | True |
| **6** | False | False | True | False | True | True | True | True |
| **7** | False | False | False | False | True | True | True | True |
| **8** | False | True | False | False | True | True | True | True |
| **9** | False | False | False | False | True | True | True | True |
| **10** | False | False | False | False | True | True | True | True |
| **11** | False | False | False | False | True | True | True | True |
| **0** | False | True | True | True | False | False | False | False |
| **1** | False | True | True | True | False | False | False | False |

```
data4 = {
    'A':['A1', 'A2', 'A3', 'A4', 'A5'],
    'B':['B1', 'B2', 'B3', 'B4', 'B4'],
    'C':['C1', 'C2', 'C3', 'C3', 'C3'],
    'D':['D1', 'D2', 'D2', 'D2', 'D2'],
    'E':['E1', 'E1', 'E1', 'E1', 'E1'] }
df_data4 = pd.DataFrame(data4)


df_data4.B.unique()
```

```
array(['B1', 'B2', 'B3', 'B4'], dtype=object)
```

```
data.Country.unique()
```

```
array(['India', 'Georgia', 'France', 'Spain', 'Germany'], dtype=object)
```

```
new_data = data.rename(columns={'Country': 'Country_Names'}, index={'11': 'Last_Row'})
```

```
new_data
```

| | Country_Names | Age | Salary | Purchased |
|---|---|---|---|---|
| **0** | India | 40.0 | 90000.0 | Yes |
| **1** | Georgia | 44.0 | 69490.0 | No |
| **2** | France | 44.0 | 72000.0 | No |
| **3** | Spain | 27.0 | 48000.0 | Yes |
| **4** | Germany | 30.0 | 54000.0 | No |
| **5** | Spain | 38.0 | 61000.0 | No |
| **6** | Germany | 40.0 | NaN | Yes |
| **7** | France | 35.0 | 58000.0 | Yes |
| **8** | Spain | NaN | 52000.0 | No |
| **9** | France | 48.0 | 79000.0 | Yes |
| **10** | Germany | 50.0 | 83000.0 | No |
| **11** | France | 37.0 | 67000.0 | Yes |

```
brand_new_data = new_data.rename(index={11: 'Last_Row'})
```

```
brand_new_data
```

| | Country_Names | Age | Salary | Purchased |
|---|---|---|---|---|
| **0** | India | 40.0 | 90000.0 | Yes |
| **1** | Georgia | 44.0 | 69490.0 | No |
| **2** | France | 44.0 | 72000.0 | No |
| **3** | Spain | 27.0 | 48000.0 | Yes |
| **4** | Germany | 30.0 | 54000.0 | No |
| **5** | Spain | 38.0 | 61000.0 | No |
| **6** | Germany | 40.0 | NaN | Yes |
| **7** | France | 35.0 | 58000.0 | Yes |
| **8** | Spain | NaN | 52000.0 | No |
| **9** | France | 48.0 | 79000.0 | Yes |
| **10** | Germany | 50.0 | 83000.0 | No |
| **Last_Row** | France | 37.0 | 67000.0 | Yes |

```
data2.mean()
```

```
Int_Rate       2.00
Ind_Gdp       48.75
year        2017.50
dtype: float64
```

```
data3.mode()
```

| | Low_Tier | Unemployment | year1 |
|---|---|---|---|
| **0** | 35 | 1 | 2006 |

```
data6 = pd.DataFrame({"A":[12, 4, 5, 44, 1],
                      "B":[5, 2, 54, 3, 2],
                      "C":[20, 16, 7, 3, 8],
                      "D":[14, 3, 17, 2, 6]})
```

```
data6
```

| | A | B | C | D |
|---|---|---|---|---|
| **0** | 12 | 5 | 20 | 14 |
| **1** | 4 | 2 | 16 | 3 |
| **2** | 5 | 54 | 7 | 17 |
| **3** | 44 | 3 | 3 | 2 |
| **4** | 1 | 2 | 8 | 6 |

```
data6.median(axis = 0)
```

```
A    5.0
B    3.0
C    8.0
D    6.0
dtype: float64
```