In [1]:
```
import pandas as pd
```

In [2]:
```
import numpy as np
```

In [3]:
```
df = pd.read_csv("Salary_Data.csv")
```

In [5]:
```
df.head()
```

Out[5]:

| | Experience | Age | Income |
|---|---|---|---|
| **0** | 1.1 | 31-40 | 39343 |
| **1** | 1.3 | 21-30 | 46205 |
| **2** | 1.5 | 21-30 | 37731 |
| **3** | 2.0 | 31-40 | 43525 |
| **4** | 2.2 | 21-30 | 39891 |

In [6]:
```
df.isnull().sum()
```

Out[6]:
```
Experience    0
Age           0
Income        0
dtype: int64
```

In [19]:
```
mean = df['Income'].mean()
print(mean)
```

```
76003.0
```

In [12]:
```
median = df['Income'].median()
print(median)
```

```
65237.0
```

In [20]:
```
grouped_stats = df.groupby('Age')['Income']
```

In [21]:
```
grouped_stats.mean()
```

Out[21]:
```
Age
21-30    77229.133333
31-40    75602.500000
41-50    63218.000000
Name: Income, dtype: float64
```

In [22]:
```
grouped_stats.median()
```

Out[22]:
```
Age
21-30    67938.0
31-40    63570.0
41-50    63218.0
Name: Income, dtype: float64
```

In [23]:
```
grouped_stats.min()
```

Out[23]:
```
Age
21-30    37731
31-40    39343
41-50    63218
Name: Income, dtype: int64
```

In [24]:
```
grouped_stats.max()
```

Out[24]:
```
Age
21-30    121872
31-40    122391
41-50     63218
Name: Income, dtype: int64
```

In [25]:
```
grouped_stats.std()
```

Out[25]:
```
Age
21-30    28621.644721
31-40    27925.650364
41-50             NaN
Name: Income, dtype: float64
```

In [27]:
```
grouped_stats.agg(['mean'])
```

Out[27]:

| | mean |
|---|---|
| **Age** | |
| **21-30** | 77229.133333 |
| **31-40** | 75602.500000 |
| **41-50** | 63218.000000 |

In [28]:
```
grouped_stats.agg(['mean','median','min','max','std'])
```

Out[28]:

| | mean | median | min | max | std |
|---|---|---|---|---|---|
| **Age** | | | | | |
| **21-30** | 77229.133333 | 67938.0 | 37731 | 121872 | 28621.644721 |
| **31-40** | 75602.500000 | 63570.0 | 39343 | 122391 | 27925.650364 |
| **41-50** | 63218.000000 | 63218.0 | 63218 | 63218 | NaN |

In [29]:
```
df1 = pd.read_csv('iris.csv')
```

In [30]:
```
df1.isnull().sum()
```

Out[30]:
```
sepal.length    0
sepal.width     0
petal.length    0
petal.width     0
variety         0
dtype: int64
```

In [31]:
```
df1.head()
```

Out[31]:

| | sepal.length | sepal.width | petal.length | petal.width | variety |
|---|---|---|---|---|---|
| **0** | 5.1 | 3.5 | 1.4 | 0.2 | Setosa |
| **1** | 4.9 | 3.0 | 1.4 | 0.2 | Setosa |
| **2** | 4.7 | 3.2 | 1.3 | 0.2 | Setosa |
| **3** | 4.6 | 3.1 | 1.5 | 0.2 | Setosa |
| **4** | 5.0 | 3.6 | 1.4 | 0.2 | Setosa |

In [32]:
```
grouped_stats1 = df1.groupby('variety')['sepal.length']
```

In [37]:
```
grouped_stats1.agg(['mean','median','min','max','std'])
```

Out[37]:

| | mean | median | min | max | std |
|---|---|---|---|---|---|
| **variety** | | | | | |
| **Setosa** | 5.006 | 5.0 | 4.3 | 5.8 | 0.352490 |
| **Versicolor** | 5.936 | 5.9 | 4.9 | 7.0 | 0.516171 |
| **Virginica** | 6.588 | 6.5 | 4.9 | 7.9 | 0.635880 |

In [43]:
```
Setosa = df1[df1['variety'] == 'Iris-Setosa']
```

In [44]:
```
Versicolor = df1[df1['variety'] == 'Iris-Versicolor']
```

In [45]:
```
Virginica = df1[df1['variety'] == 'Iris-Virginica']
```

In [ ]:
```

```