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

In [11]:
```
df=pd.read_csv("Book1.csv", names=["name","income"], skiprows=1)
df
```

Out[11]:

| | name | income |
|---|---|---|
| 0 | Rob | 5000 |
| 1 | Rafiq | 6000 |
| 2 | Nina | 4000 |
| 3 | Sofia | 7500 |
| 4 | Mohan | 8000 |
| 5 | Tao | 7000 |
| 6 | Elon Musk | 10000000 |

In [12]:
```
df.income.quantile(0)
```

Out[12]:
```
4000.0
```

In [13]:
```
df.income.quantile(0.25,interpolation = "higher")
```

Out[13]:
```
6000
```

In [14]:
```
df.income.quantile(0.5,interpolation = "lower")
```

Out[14]:
```
7000
```

In [15]:
```
df.income.quantile(1)
```

Out[15]:
```
10000000.0
```

In [16]:
```
percentile_99 = df.income.quantile(0.99)
percentile_99
```

Out[16]:
```
9400479.999999994
```

In [17]:
```
df[df.income>percentile_99]
```

Out[17]:

| | name | income |
|---|---|---|
| 6 | Elon Musk | 10000000 |

In [18]:
```
df[df.income<=percentile_99]
```

Out[18]:

| | name | income |
|---|---|---|
| 0 | Rob | 5000 |
| 1 | Rafiq | 6000 |
| 2 | Nina | 4000 |
| 3 | Sofia | 7500 |
| 4 | Mohan | 8000 |
| 5 | Tao | 7000 |

In [20]:
```
df
```

Out[20]:

| | name | income |
|---|---|---|
| 0 | Rob | 5000 |
| 1 | Rafiq | 6000 |
| 2 | Nina | 4000 |
| 3 | Sofia | 7500 |
| 4 | Mohan | 8000 |
| 5 | Tao | 7000 |
| 6 | Elon Musk | 10000000 |

In [22]:
```
df1=df[df.income<=percentile_99]
df1
```

Out[22]:

| | name | income |
|---|---|---|
| 0 | Rob | 5000 |
| 1 | Rafiq | 6000 |
| 2 | Nina | 4000 |
| 3 | Sofia | 7500 |
| 4 | Mohan | 8000 |
| 5 | Tao | 7000 |

In [26]:
```
df['income'][3] = np.nan
print(df)
```
```
       name      income
0       Rob      5000.0
1     Rafiq      6000.0
2      Nina      4000.0
3     Sofia         NaN
4     Mohan      8000.0
5       Tao      7000.0
6  Elon Musk  10000000.0
```

In [27]:
```
df
```

Out[27]:

| | name | income |
|---|---|---|
| 0 | Rob | 5000.0 |
| 1 | Rafiq | 6000.0 |
| 2 | Nina | 4000.0 |
| 3 | Sofia | NaN |
| 4 | Mohan | 8000.0 |
| 5 | Tao | 7000.0 |
| 6 | Elon Musk | 10000000.0 |

In [28]:
```
df.income.mean()
```

Out[28]:
```
1671666.6666666667
```

In [30]:
```
df_new = df.fillna(df.income.mean())
df_new
```

Out[30]:

| | name | income |
|---|---|---|
| 0 | Rob | 5.000000e+03 |
| 1 | Rafiq | 6.000000e+03 |
| 2 | Nina | 4.000000e+03 |
| 3 | Sofia | 1.671667e+06 |
| 4 | Mohan | 8.000000e+03 |
| 5 | Tao | 7.000000e+03 |
| 6 | Elon Musk | 1.000000e+07 |

In [31]:
```
df_new = df.fillna(df.income.median())
df_new
```

Out[31]:

| | name | income |
|---|---|---|
| 0 | Rob | 5000.0 |
| 1 | Rafiq | 6000.0 |
| 2 | Nina | 4000.0 |
| 3 | Sofia | 6500.0 |
| 4 | Mohan | 8000.0 |
| 5 | Tao | 7000.0 |

6 Elon Musk 10000000.0

In [ ]: