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

In [2]:
```
data = pd.read_csv('nba.csv')
```

In [3]:
```
data.head()
```

Out[3]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
0	Avery Bradley	Boston Celtics	0.0	PG	25.0	6-2	180.0	Texas	7730337.0
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6-6	235.0	Marquette	6796117.0
2	John Holland	Boston Celtics	30.0	SG	27.0	6-5	205.0	Boston University	NaN
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6-5	185.0	Georgia State	1148640.0
4	Jonas Jerebko	Boston Celtics	8.0	PF	29.0	6-10	231.0	NaN	5000000.0

In [4]:
```
data['Salary'].isnull().sum()
```

Out[4]:
```
12
```

In [5]:
```
data['Age'].isnull().sum()
```

Out[5]:
```
1
```

In [6]:
```
data.dtypes
```

Out[6]:
```
Name          object
Team          object
Number       float64
Position      object
Age          float64
Height        object
Weight       float64
College       object
Salary       float64
dtype: object
```

In [ ]:
```
```

In [7]:
```
df = data['Salary'].fillna(data['Salary'].mean())
data['Salary'] = df;
df = data['Age'].fillna(data['Age'].median())
data['Age'] = df
```

In [8]:
```
data.dtypes
```

Out[8]:
```
Name         object
Team         object
Number      float64
Position     object
Age         float64
Height       object
Weight      float64
College      object
Salary      float64
dtype: object
```

In [9]:
```
data['Age'] = data['Age'].astype(int)
```

In [10]:
```
data.dtypes
```

Out[10]:
```
Name         object
Team         object
Number      float64
Position     object
Age           int64
Height       object
Weight      float64
College      object
Salary      float64
dtype: object
```

In [ ]:
```

```

In [17]:
```
data.groupby('Age')['Salary'].describe()
```

Out[17]:

Age	count	mean	std	min	25%	50%	75%
19	2.0	1.930440e+06	2.791658e+05	1.733040e+06	1.831740e+06	1930440.0	2.029140e+06
20	19.0	2.725791e+06	1.510913e+06	5.250930e+05	1.636920e+06	2481720.0	3.954720e+06
21	19.0	2.067380e+06	1.412350e+06	5.250930e+05	1.235480e+06	1584480.0	2.084940e+06
22	26.0	2.357963e+06	1.517378e+06	5.250930e+05	1.169190e+06	1793880.0	3.366733e+06
23	41.0	2.171719e+06	2.728808e+06	3.088800e+04	6.500000e+05	1201440.0	2.891760e+06
24	47.0	3.830295e+06	4.702753e+06	5.250930e+05	9.236380e+05	1535880.0	4.751583e+06
25	45.0	3.951130e+06	4.508414e+06	5.572200e+04	9.472760e+05	1358880.0	5.543725e+06
26	37.0	6.811867e+06	6.024341e+06	1.342150e+05	1.415520e+06	3750000.0	1.225000e+07
27	41.0	6.632008e+06	6.752122e+06	5.572200e+04	1.035000e+06	3425510.0	1.123596e+07
28	31.0	5.101559e+06	4.244345e+06	2.061920e+05	1.843421e+06	4389607.0	6.190017e+06
29	28.0	6.174838e+06	4.786782e+06	1.114440e+05	2.205000e+06	5271862.5	9.322612e+06
30	31.0	6.990272e+06	5.332677e+06	9.472760e+05	3.099470e+06	5675000.0	9.000000e+06
31	22.0	8.511397e+06	7.170163e+06	9.472760e+05	3.000000e+06	5439487.0	1.347500e+07
32	13.0	7.716958e+06	7.451336e+06	2.953270e+05	3.135000e+06	5200000.0	1.121739e+07
33	14.0	3.930739e+06	4.354293e+06	2.006000e+05	1.119326e+06	2300000.0	4.040084e+06
34	10.0	7.606030e+06	5.653035e+06	2.618940e+05	4.512750e+06	6280675.5	9.103916e+06
35	9.0	3.615178e+06	2.259354e+06	9.472760e+05	1.499187e+06	3750000.0	4.842684e+06
36	10.0	2.238120e+06	1.550061e+06	9.472760e+05	9.472760e+05	1809826.0	2.888735e+06
37	4.0	1.079400e+07	9.606489e+06	4.842684e+06	4.960671e+06	6666667.0	1.250000e+07
38	4.0	1.840041e+06	1.496661e+06	2.228880e+05	7.661790e+05	1880638.0	2.954500e+06
39	2.0	2.517872e+06	2.220522e+06	9.477260e+05	1.732799e+06	2517872.5	3.302946e+06
40	3.0	4.666917e+06	4.155421e+06	2.507500e+05	2.750375e+06	5250000.0	6.875000e+06

In [11]:
```
data.groupby('Age')['Salary'].size()
```

Out[11]:
```
Age
19     2
20    19
21    19
22    26
23    41
24    47
25    45
26    37
27    41
28    31
29    28
30    31
31    22
32    13
33    14
34    10
35     9
36    10
37     4
38     4
39     2
40     3
Name: Salary, dtype: int64
```

In [25]:
```
data.groupby('Age')['Salary'].agg(['mean','median','std','min','max'])
```

Out[25]:

Age	mean	median	std	min	max
19	1.930440e+06	1930440.0	2.791658e+05	1.733040e+06	2127840.0
20	2.725791e+06	2481720.0	1.510913e+06	5.250930e+05	5703600.0
21	2.067380e+06	1584480.0	1.412350e+06	5.250930e+05	5758680.0
22	2.357963e+06	1793880.0	1.517378e+06	5.250930e+05	6331404.0
23	2.171719e+06	1201440.0	2.728808e+06	3.088800e+04	16000000.0
24	3.830295e+06	1535880.0	4.702753e+06	5.250930e+05	16407501.0
25	3.951130e+06	1358880.0	4.508414e+06	5.572200e+04	15851950.0
26	6.811867e+06	3750000.0	6.024341e+06	1.342150e+05	17120106.0
27	6.632008e+06	3425510.0	6.752122e+06	5.572200e+04	20158622.0
28	5.101559e+06	4389607.0	4.244345e+06	2.061920e+05	19689000.0
29	6.174838e+06	5271862.5	4.786782e+06	1.114440e+05	16407500.0
30	6.990272e+06	5675000.0	5.332677e+06	9.472760e+05	22359364.0
31	8.511397e+06	5439487.0	7.170163e+06	9.472760e+05	22970500.0
32	7.716958e+06	5200000.0	7.451336e+06	2.953270e+05	22875000.0
33	3.930739e+06	2300000.0	4.354293e+06	2.006000e+05	13000000.0
34	7.606030e+06	6280675.5	5.653035e+06	2.618940e+05	20000000.0
35	3.615178e+06	3750000.0	2.259354e+06	9.472760e+05	7448760.0
36	2.238120e+06	1809826.0	1.550061e+06	9.472760e+05	5675000.0
37	1.079400e+07	6666667.0	9.606489e+06	4.842684e+06	25000000.0
38	1.840041e+06	1880638.0	1.496661e+06	2.228880e+05	3376000.0
39	2.517872e+06	2517872.5	2.220522e+06	9.477260e+05	4088019.0
40	4.666917e+06	5250000.0	4.155421e+06	2.507500e+05	8500000.0

In [26]:
```
data.groupby('Age')['Salary'].mean()
```

Out[26]:
```
Age
19    1.930440e+06
20    2.725791e+06
21    2.067380e+06
22    2.357963e+06
23    2.171719e+06
24    3.830295e+06
25    3.951130e+06
26    6.811867e+06
27    6.632008e+06
28    5.101559e+06
29    6.174838e+06
30    6.990272e+06
31    8.511397e+06
32    7.716958e+06
33    3.930739e+06
34    7.606030e+06
35    3.615178e+06
36    2.238120e+06
37    1.079400e+07
38    1.840041e+06
39    2.517872e+06
40    4.666917e+06
Name: Salary, dtype: float64
```

In [13]:
```
data.groupby('Age')['Salary'].median()
```

Out[13]:
```
Age
19    1930440.0
20    2481720.0
21    1584480.0
22    1793880.0
23    1201440.0
24    1535880.0
25    1358880.0
26    3750000.0
27    3425510.0
28    4389607.0
29    5271862.5
30    5675000.0
31    5439487.0
32    5200000.0
33    2300000.0
34    6280675.5
35    3750000.0
36    1809826.0
37    6666667.0
38    1880638.0
39    2517872.5
40    5250000.0
Name: Salary, dtype: float64
```

In [14]:
```
data.groupby('Age')['Salary'].std()
```

Out[14]:
```
Age
19    2.791658e+05
20    1.510913e+06
21    1.412350e+06
22    1.517378e+06
23    2.728808e+06
24    4.702753e+06
25    4.508414e+06
26    6.024341e+06
27    6.752122e+06
28    4.244345e+06
29    4.786782e+06
30    5.332677e+06
31    7.170163e+06
32    7.451336e+06
33    4.354293e+06
34    5.653035e+06
35    2.259354e+06
36    1.550061e+06
37    9.606489e+06
38    1.496661e+06
39    2.220522e+06
40    4.155421e+06
Name: Salary, dtype: float64
```

In [15]:
```
data.groupby('Age')['Salary'].min()
```

Out[15]:
```
Age
19    1.733040e+06
20    5.250930e+05
21    5.250930e+05
22    5.250930e+05
23    3.088800e+04
24    5.250930e+05
25    5.572200e+04
26    1.342150e+05
27    5.572200e+04
28    2.061920e+05
29    1.114440e+05
30    9.472760e+05
31    9.472760e+05
32    2.953270e+05
33    2.006000e+05
34    2.618940e+05
35    9.472760e+05
36    9.472760e+05
37    4.842684e+06
38    2.228880e+05
39    9.477260e+05
40    2.507500e+05
Name: Salary, dtype: float64
```

In [16]:

```
d=data.groupby(['Age','Salary']).size()
print(d)
```

```
Age  Salary
19   1733040.0    1
     2127840.0    1
20   525093.0     2
     1131960.0    1
     1282080.0    1
                 ..
39   947726.0     1
     4088019.0    1
40   250750.0     1
     5250000.0    1
     8500000.0    1
Length: 403, dtype: int64
```