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

In [2]:
```
df = pd.read_csv("nba.csv")
```

In [3]:
```
df.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]:
```
df.dropna(inplace = True)
```

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

Out[5]:

	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
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6-5	185.0	Georgia State	1148640.0
6	Jordan Mickey	Boston Celtics	55.0	PF	21.0	6-8	235.0	LSU	1170960.0
7	Kelly Olynyk	Boston Celtics	41.0	C	25.0	7-0	238.0	Gonzaga	2165160.0

In [6]:
```
df['Height'].value_counts()
```

Out[6]:
```
6-9     49
6-7     37
6-8     36
6-6     36
6-10    32
6-11    29
6-5     29
6-4     28
6-3     25
7-0     20
6-2     13
6-1     12
6-0     10
5-11     3
7-1      3
5-9      1
7-2      1
Name: Height, dtype: int64
```

In [7]:
```
heightgrp = df.groupby(df['Height'])
```

In [8]:
```
heightgrp.get_group('6-0')
```

Out[8]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
47	Isaiah Canaan	Philadelphia 76ers	0.0	PG	25.0	6-0	201.0	Murray State	947276.0
57	Ish Smith	Philadelphia 76ers	1.0	PG	27.0	6-0	175.0	Wake Forest	947276.0
67	Kyle Lowry	Toronto Raptors	7.0	PG	30.0	6-0	205.0	Villanova	12000000.0
100	Chris Paul	Los Angeles Clippers	3.0	PG	31.0	6-0	175.0	Wake Forest	21468695.0
142	Darren Collison	Sacramento Kings	7.0	PG	28.0	6-0	175.0	UCLA	5013559.0
152	Aaron Brooks	Chicago Bulls	0.0	PG	31.0	6-0	161.0	Oregon	2250000.0
228	J.J. Barea	Dallas Mavericks	5.0	PG	31.0	6-0	185.0	Northeastern	4290000.0
305	Patty Mills	San Antonio Spurs	8.0	PG	27.0	6-0	185.0	Saint Mary's	3578947.0
384	D.J. Augustin	Denver Nuggets	12.0	PG	28.0	6-0	183.0	Texas	3000000.0
394	Jameer Nelson	Denver Nuggets	1.0	PG	34.0	6-0	190.0	Saint Joseph's	4345000.0

In [9]:
```
bins = [19,25,31,36,40]
labels =['19-24','25-30','31-35','36-40']
df['AgeGroup'] = pd.cut(df['Age'],bins = bins,labels = labels,right = False)
```

In [10]:
```
df.head()
```

Out[10]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary	AgeGroup
0	Avery Bradley	Boston Celtics	0.0	PG	25.0	6-2	180.0	Texas	7730337.0	25-30
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6-6	235.0	Marquette	6796117.0	25-30
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6-5	185.0	Georgia State	1148640.0	19-24
6	Jordan Mickey	Boston Celtics	55.0	PF	21.0	6-8	235.0	LSU	1170960.0	19-24
7	Kelly Olynyk	Boston Celtics	41.0	C	25.0	7-0	238.0	Gonzaga	2165160.0	25-30

In [11]:
```
df['AgeGroup'].value_counts()
```

Out[11]:
```
25-30    172
19-24    128
31-35     48
36-40     14
Name: AgeGroup, dtype: int64
```

In [12]:
```
df.groupby('AgeGroup')['Salary'].max()
```

Out[12]:
```
AgeGroup
19-24    16407501.0
25-30    20158622.0
31-35    22875000.0
36-40     5675000.0
Name: Salary, dtype: float64
```

In [13]:
```
df.groupby('AgeGroup')['Salary'].min()
```

Out[13]:
```
AgeGroup
19-24     83397.0
25-30     55722.0
31-35    200600.0
36-40    222888.0
Name: Salary, dtype: float64
```

In [14]:
```
df.groupby('AgeGroup')['Salary'].count()
```

Out[14]:
```
AgeGroup
19-24    128
25-30    172
31-35     48
36-40     14
Name: Salary, dtype: int64
```

In [15]:
```
df.groupby('AgeGroup')['Salary'].std()
```

Out[15]:
```
AgeGroup
19-24    3.402048e+06
25-30    5.639105e+06
31-35    6.092403e+06
36-40    1.732754e+06
Name: Salary, dtype: float64
```

In [16]:
```
df.groupby('AgeGroup')['Salary'].mean()
```

Out[16]:
```
AgeGroup
19-24    2.829308e+06
25-30    5.749697e+06
31-35    6.084955e+06
36-40    2.365384e+06
Name: Salary, dtype: float64
```

In [17]:
```
df.groupby('AgeGroup')['Salary'].describe()
```

Out[17]:

	count	mean	std	min	25%	50%	75%	max
AgeGroup								
19-24	128.0	2.829308e+06	3.402048e+06	83397.0	947276.00	1615440.0	3116157.00	16407501.0
25-30	172.0	5.749697e+06	5.639105e+06	55722.0	1015421.00	3479421.5	8991573.75	20158622.0
31-35	48.0	6.084955e+06	6.092403e+06	200600.0	2096417.75	4171723.0	6646250.00	22875000.0
36-40	14.0	2.365384e+06	1.732754e+06	222888.0	947276.00	1809826.0	3500875.00	5675000.0

In [18]:
```
Salary_list = list(df.groupby('AgeGroup')['Salary'])
```

In [19]:
```
Salary_list
```

Out[19]:
```
[('19-24',
  3       1148640.0
  6       1170960.0
  8       1824360.0
  9       3431040.0
  10      2569260.0
            ...
  443     2658240.0
  444     9463484.0
  446    12000000.0
  449     1348440.0
  452     2239800.0
  Name: Salary, Length: 128, dtype: float64),
 ('25-30',
  0       7730337.0
  1       6796117.0
  7       2165160.0
  11      6912869.0
  12      3425510.0
            ...
  442     4775000.0
  448    15409570.0
  451      981348.0
  453     2433333.0
  456      947276.0
  Name: Salary, Length: 172, dtype: float64),
 ('31-35',
  19      6300000.0
  31      1635476.0
  33     22875000.0
  52      6500000.0
  78     13800000.0
  83     11710456.0
  100    21468695.0
  103     7085000.0
  106     3000000.0
  120     5219169.0
  131      947276.0
  133     5500000.0
  152     2250000.0
  154     4500000.0
  162    13400000.0
  167     8193029.0
  170      947276.0
  172      947276.0
  173     1276000.0
  180     2100000.0
  181     2500000.0
  195     3000000.0
  221      295327.0
  228     4290000.0
  230     3950313.0
  231     4053446.0
  232     2085671.0
  240      947276.0
  241     5378974.0
  258     5158539.0
  271     9638555.0
  303      200600.0
  308     1499187.0
  311     2854940.0
  313     1000000.0
  314     5746479.0
  315    18671659.0
  339    22192730.0
  340    10151612.0
  344      261894.0
  349    20000000.0
  367     5000000.0
  368     4000000.0
  372     3300000.0
  394     4345000.0
  413     3750000.0
  415     3135000.0
  434     5016000.0
  Name: Salary, dtype: float64),
 ('36-40',
  93      5675000.0
  101     3376000.0
  119      947276.0
  139     1449187.0
  183     2170465.0
  256      947276.0
  259     5000000.0
  260     3542500.0
  261     4088019.0
```

```
296     947276.0
343    2854940.0
392     947276.0
406     947276.0
420     222888.0
Name: Salary, dtype: float64)]
```

In [ ]: