

In [1]:

```
import pandas as pd
```

In [22]:

```
nba=pd.read_csv("nba.csv")
```

Shared methods and attributes b/n Series and Dataframes

In [21]:

```
nba.head(5)
```

Out[21]:

	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 [23]:

```
nba.tail()
```

Out[23]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
453	Shelvin Mack	Utah Jazz	8.0	PG	26.0	6-3	203.0	Butler	2433333.0
454	Raul Neto	Utah Jazz	25.0	PG	24.0	6-1	179.0	NaN	900000.0
455	Tibor Pleiss	Utah Jazz	21.0	C	26.0	7-3	256.0	NaN	2900000.0
456	Jeff Withey	Utah Jazz	24.0	C	26.0	7-0	231.0	Kansas	947276.0
457	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [24]:

```
nba.index
```

Out[24]:

```
RangeIndex(start=0, stop=458, step=1)
```

In [25]:

```
nba.values
```

Out[25]:

```
array([[ 'Avery Bradley', 'Boston Celtics', 0.0, ..., 180.0, 'Texas',  
        7730337.0],  
       [ 'Jae Crowder', 'Boston Celtics', 99.0, ..., 235.0, 'Marquette',  
        6796117.0],  
       [ 'John Holland', 'Boston Celtics', 30.0, ..., 205.0,  
        'Boston University', nan],  
       ...,  
       [ 'Tibor Pleiss', 'Utah Jazz', 21.0, ..., 256.0, nan, 2900000.0],  
       [ 'Jeff Withey', 'Utah Jazz', 24.0, ..., 231.0, 'Kansas', 947276.0],  
       [nan, nan, nan, ..., nan, nan, nan]], dtype=object)
```

In [26]:

```
nba.shape
```

Out[26]:

```
(458, 9)
```

In [27]:

```
nba.dtypes
```

Out[27]:

```
Name          object  
Team           object  
Number        float64  
Position       object  
Age           float64  
Height        object  
Weight        float64  
College        object  
Salary        float64  
dtype: object
```

In [28]:

```
nba.value_counts()
```

Out[28]:

Name	Team	Number	Position	Age	Height
Aaron Brooks	Chicago Bulls	0.0	PG	31.0	6-0
161.0 Oregon	2250000.0	1			
Mike Muscala	Atlanta Hawks	31.0	PF	24.0	6-11
240.0 Bucknell	947276.0	1			
Mike Dunleavy	Chicago Bulls	34.0	SG	35.0	6-9
230.0 Duke	4500000.0	1			
Mike Conley	Memphis Grizzlies	11.0	PG	28.0	6-1
175.0 Ohio State	9588426.0	1			
Michael Kidd-Gilchrist	Charlotte Hornets	14.0	SF	22.0	6-7
232.0 Kentucky	6331404.0	1			
..					
Hassan Whiteside	Miami Heat	21.0	C	26.0	7-0
265.0 Marshall	981348.0	1			
Harrison Barnes	Golden State Warriors	40.0	SF	24.0	6-8
225.0 North Carolina	3873398.0	1			
Greivis Vasquez	Milwaukee Bucks	21.0	PG	29.0	6-6
217.0 Maryland	6600000.0	1			
Greg Monroe	Milwaukee Bucks	15.0	C	26.0	6-11
265.0 Georgetown	16407500.0	1			
Zach Randolph	Memphis Grizzlies	50.0	PF	34.0	6-9
260.0 Michigan State	9638555.0	1			

Length: 364, dtype: int64

In [30]:

```
nba.count()
```

Out[30]:

```
Name      457
Team       457
Number     457
Position   457
Age        457
Height     457
Weight     457
College    373
Salary     446
dtype: int64
```

In [31]:

```
nba.dtypes.value_counts()
```

Out[31]:

```
object      5
float64     4
dtype: int64
```

In [32]:

```
nba.columns
```

Out[32]:

```
Index(['Name', 'Team', 'Number', 'Position', 'Age', 'Height', 'Weight',  
      'College', 'Salary'],  
      dtype='object')
```

In [33]:

```
nba.axes
```

Out[33]:

```
[RangeIndex(start=0, stop=458, step=1),  
 Index(['Name', 'Team', 'Number', 'Position', 'Age', 'Height', 'Weight',  
      'College', 'Salary'],  
      dtype='object')]
```

In [35]:

```
nba.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 458 entries, 0 to 457  
Data columns (total 9 columns):  
#   Column      Non-Null Count  Dtype  
---  -  
0   Name        457 non-null    object  
1   Team        457 non-null    object  
2   Number      457 non-null    float64  
3   Position    457 non-null    object  
4   Age         457 non-null    float64  
5   Height      457 non-null    object  
6   Weight      457 non-null    float64  
7   College     373 non-null    object  
8   Salary      446 non-null    float64  
dtypes: float64(4), object(5)  
memory usage: 32.3+ KB
```

In [37]:

```
nba.count()
```

Out[37]:

```
Name        457  
Team        457  
Number      457  
Position    457  
Age         457  
Height      457  
Weight      457  
College     373  
Salary      446  
dtype: int64
```

difference between shared methods

In [41]:

```
rev=pd.read_csv("revenue.csv",index_col="Date")  
rev.head(3)
```

Out[41]:

	New York	Los Angeles	Miami
Date			
1/1/16	985	122	499
1/2/16	738	788	534
1/3/16	14	20	933

In [42]:

```
rev.sum()
```

Out[42]:

```
New York      5475  
Los Angeles   5134  
Miami         5641  
dtype: int64
```

In [43]:

```
rev.sum(axis=0)
```

Out[43]:

```
New York      5475  
Los Angeles   5134  
Miami         5641  
dtype: int64
```

In [44]:

```
rev.sum(axis="index")
```

Out[44]:

```
New York      5475  
Los Angeles   5134  
Miami         5641  
dtype: int64
```

In [45]:

```
rev.sum(axis=1)
```

Out[45]:

Date
1/1/16 1606
1/2/16 2060
1/3/16 967
1/4/16 2519
1/5/16 438
1/6/16 1935
1/7/16 1234
1/8/16 2313
1/9/16 2623
1/10/16 555
dtype: int64

In [47]:

```
rev.sum(axis="columns")
```

Out[47]:

Date
1/1/16 1606
1/2/16 2060
1/3/16 967
1/4/16 2519
1/5/16 438
1/6/16 1935
1/7/16 1234
1/8/16 2313
1/9/16 2623
1/10/16 555
dtype: int64

Select one coulmm from DF

In [49]:

```
nba=pd.read_csv("nba.csv")  
nba.head(3)
```

Out[49]:

	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

In [53]:

```
nba.Name.head(5)
```

Out[53]:

```
0    Avery Bradley
1      Jae Crowder
2    John Holland
3     R.J. Hunter
4   Jonas Jerebko
Name: Name, dtype: object
```

In [55]:

```
nba["Name"].head()
```

Out[55]:

```
0    Avery Bradley
1      Jae Crowder
2    John Holland
3     R.J. Hunter
4   Jonas Jerebko
Name: Name, dtype: object
```

In [59]:

```
type(nba["Name"].head())
```

Out[59]:

```
pandas.core.series.Series
```

Select two or more coulm from DF

In [62]:

```
nba[["Team", "Number"]].head()
```

Out[62]:

	Team	Number
0	Boston Celtics	0.0
1	Boston Celtics	99.0
2	Boston Celtics	30.0
3	Boston Celtics	28.0
4	Boston Celtics	8.0

In [63]:

```
type(nba[["Team", "Number"]].head())
```

Out[63]:

pandas.core.frame.DataFrame

In [64]:

```
select=["Name", "Team", "Number"]
```

In [66]:

```
nba[select]
```

Out[66]:

	Name	Team	Number
0	Avery Bradley	Boston Celtics	0.0
1	Jae Crowder	Boston Celtics	99.0
2	John Holland	Boston Celtics	30.0
3	R.J. Hunter	Boston Celtics	28.0
4	Jonas Jerebko	Boston Celtics	8.0
...
453	Shelvin Mack	Utah Jazz	8.0
454	Raul Neto	Utah Jazz	25.0
455	Tibor Pleiss	Utah Jazz	21.0
456	Jeff Withey	Utah Jazz	24.0
457	NaN	NaN	NaN

458 rows × 3 columns

Add New column to DF

In [67]:

```
nba=pd.read_csv("nba.csv")
nba.head(3)
```

Out[67]:

	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

In [68]:

```
nba["Sport"]="Cricket"
```

In [70]:

```
nba.head()
```

Out[70]:

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

In [71]:

```
nba=pd.read_csv("nba.csv")
nba.head(3)
```

Out[71]:

	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

In [72]:

```
nba.insert(2,"Sport","Cricket")
```

In [73]:

```
nba.head()
```

Out[73]:

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

Broadcasting operations on DF

In [74]:

```
nba=pd.read_csv("nba.csv")
nba.head(3)
```

Out[74]:

	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

In [76]:

```
nba["Age"].add(5).head()
```

Out[76]:

```
0    30.0
1    30.0
2    32.0
3    27.0
4    34.0
Name: Age, dtype: float64
```

In [78]:

```
(nba["Age"]+5).head()
```

Out[78]:

```
0    30.0
1    30.0
2    32.0
3    27.0
4    34.0
Name: Age, dtype: float64
```

In [79]:

```
nba["Age"].sub(5).head()
```

Out[79]:

```
0    20.0
1    20.0
2    22.0
3    17.0
4    24.0
Name: Age, dtype: float64
```

In [82]:

```
nba["wgt in kgs"]=nba["Weight"]*0.453592
```

In [83]:

```
nba.head()
```

Out[83]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary	wgt in kgs
0	Avery Bradley	Boston Celtics	0.0	PG	25.0	6-2	180.0	Texas	7730337.0	81.646560
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6-6	235.0	Marquette	6796117.0	106.594120
2	John Holland	Boston Celtics	30.0	SG	27.0	6-5	205.0	Boston University	NaN	92.986360
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6-5	185.0	Georgia State	1148640.0	83.914520
4	Jonas Jerebko	Boston Celtics	8.0	PF	29.0	6-10	231.0	NaN	5000000.0	104.779752

In [85]:

```
nba["Salary"]/1000000
nba["sal in milns"]=nba["Salary"]/1000000
```

In [86]:

```
nba.head()
```

Out[86]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary	wgt in kgs
0	Avery Bradley	Boston Celtics	0.0	PG	25.0	6-2	180.0	Texas	7730337.0	81.646560
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6-6	235.0	Marquette	6796117.0	106.594120
2	John Holland	Boston Celtics	30.0	SG	27.0	6-5	205.0	Boston University	NaN	92.986360
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6-5	185.0	Georgia State	1148640.0	83.914520
4	Jonas Jerebko	Boston Celtics	8.0	PF	29.0	6-10	231.0	NaN	5000000.0	104.779752

A review on .value_counts() Mtd

In [91]:

```
nba=pd.read_csv("nba.csv")
nba.head(3)
```

Out[91]:

	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

In [90]:

```
nba["Team"].value_counts()
```

Out[90]:

New Orleans Pelicans	19
Memphis Grizzlies	18
New York Knicks	16
Milwaukee Bucks	16
Boston Celtics	15
Brooklyn Nets	15
Portland Trail Blazers	15
Oklahoma City Thunder	15
Denver Nuggets	15
Washington Wizards	15
Miami Heat	15
Charlotte Hornets	15
Atlanta Hawks	15
San Antonio Spurs	15
Houston Rockets	15
Dallas Mavericks	15
Indiana Pacers	15
Detroit Pistons	15
Cleveland Cavaliers	15
Chicago Bulls	15
Sacramento Kings	15
Phoenix Suns	15
Los Angeles Lakers	15
Los Angeles Clippers	15
Golden State Warriors	15
Toronto Raptors	15
Philadelphia 76ers	15
Utah Jazz	15
Orlando Magic	14
Minnesota Timberwolves	14

Name: Team, dtype: int64

In [95]:

```
nba["Position"].value_counts(ascending=True)
```

Out[95]:

C	78
SF	85
PG	92
PF	100
SG	102

Name: Position, dtype: int64

Drop rows with Null values

In [96]:

```
nba=pd.read_csv("nba.csv")
nba.head(3)
```

Out[96]:

	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

In [97]:

```
nba.tail(3)
```

Out[97]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
455	Tibor Pleiss	Utah Jazz	21.0	C	26.0	7-3	256.0	NaN	2900000.0
456	Jeff Withey	Utah Jazz	24.0	C	26.0	7-0	231.0	Kansas	947276.0
457	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [99]:

```
nba.dropna().head()
```

Out[99]:

	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 [100]:

```
nba.dropna().tail()
```

Out[100]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
449	Rodney Hood	Utah Jazz	5.0	SG	23.0	6-8	206.0	Duke	1348440.0
451	Chris Johnson	Utah Jazz	23.0	SF	26.0	6-6	206.0	Dayton	981348.0
452	Trey Lyles	Utah Jazz	41.0	PF	20.0	6-10	234.0	Kentucky	2239800.0
453	Shelvin Mack	Utah Jazz	8.0	PG	26.0	6-3	203.0	Butler	2433333.0
456	Jeff Withey	Utah Jazz	24.0	C	26.0	7-0	231.0	Kansas	947276.0

In [107]:

```
nba.dropna(how="all", inplace=True)#removes rows only all columns in a row are set to zero
```

In [110]:

```
nba.tail()
```

Out[110]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
452	Trey Lyles	Utah Jazz	41.0	PF	20.0	6-10	234.0	Kentucky	2239800.0
453	Shelvin Mack	Utah Jazz	8.0	PG	26.0	6-3	203.0	Butler	2433333.0
454	Raul Neto	Utah Jazz	25.0	PG	24.0	6-1	179.0	NaN	900000.0
455	Tibor Pleiss	Utah Jazz	21.0	C	26.0	7-3	256.0	NaN	2900000.0
456	Jeff Withey	Utah Jazz	24.0	C	26.0	7-0	231.0	Kansas	947276.0

In [112]:

```
nba.dropna(axis=1).head()#filter in columns
```

Out[112]:

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

In [113]:

```
nba.dropna(axis="columns").head()#filter in columns
```

Out[113]:

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

In [114]:

```
nba.head()
```

Out[114]:

	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 [117]:

```
nba.dropna(subset=["Salary", "College"]).head()#filter in only selected columns
```

Out[117]:

	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

Fill null values with predfeined value

In [118]:

```
nba=pd.read_csv("nba.csv")
```

In [120]:

```
nba.head()
```

Out[120]:

	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 [126]:

```
nba.fillna(value="new").head()
```

Out[126]:

	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	new
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	new	5000000.0

In [127]:

```
nba["Salary"].fillna(0,inplace=True)
```

In [130]:

```
nba.head()
```

Out[130]:

	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	0.0
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 [131]:

```
nba["College"].fillna(value="No College",inplace=True)
```

In [132]:

```
nba.head()
```

Out[132]:

	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	0.0
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	No College	5000000.0

The .astype() mtd

In [133]:

```
nba=pd.read_csv("nba.csv").dropna(how="all")
nba["Salary"].fillna(0,inplace=True)
nba["College"].fillna("No College",inplace=True)
```

In [135]:

```
nba.head()
```

Out[135]:

	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	0.0
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	No College	5000000.0

In [136]:

```
nba.dtypes
```

Out[136]:

```
Name      object
Team       object
Number     float64
Position   object
Age        float64
Height     object
Weight     float64
College    object
Salary     float64
dtype: object
```

In [137]:

```
nba.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 457 entries, 0 to 456
Data columns (total 9 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Name        457 non-null   object
1   Team        457 non-null   object
2   Number      457 non-null   float64
3   Position    457 non-null   object
4   Age         457 non-null   float64
5   Height      457 non-null   object
6   Weight      457 non-null   float64
7   College     457 non-null   object
8   Salary      457 non-null   float64
dtypes: float64(4), object(5)
memory usage: 35.7+ KB
```

In [138]:

```
nba["Salary"] = nba["Salary"].astype("int")
```

In [139]:

```
nba.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 457 entries, 0 to 456
Data columns (total 9 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Name        457 non-null   object
1   Team        457 non-null   object
2   Number      457 non-null   float64
3   Position    457 non-null   object
4   Age         457 non-null   float64
5   Height      457 non-null   object
6   Weight      457 non-null   float64
7   College     457 non-null   object
8   Salary      457 non-null   int32
dtypes: float64(3), int32(1), object(5)
memory usage: 33.9+ KB
```

In [142]:

```
nba["Number"] = nba["Number"].astype("int")
nba["Age"] = nba["Age"].astype("int")
```

In [143]:

```
nba.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 457 entries, 0 to 456
Data columns (total 9 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Name        457 non-null   object
1   Team        457 non-null   object
2   Number      457 non-null   int32
3   Position    457 non-null   object
4   Age         457 non-null   int32
5   Height      457 non-null   object
6   Weight      457 non-null   float64
7   College     457 non-null   object
8   Salary      457 non-null   int32
dtypes: float64(1), int32(3), object(5)
memory usage: 30.3+ KB
```

In [144]:

```
nba["Team"].nunique()
```

Out[144]:

30

In [150]:

```
nba["Position"].nunique()#To find unique values in an row
```

Out[150]:

30

In [147]:

```
nba["Team"] = nba["Team"].astype("category")
```

In [148]:

```
nba["Position"] = nba["Team"].astype("category")
```

In [149]:

```
nba.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 457 entries, 0 to 456
Data columns (total 9 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   Name        457 non-null   object
1   Team        457 non-null   category
2   Number      457 non-null   int32
3   Position    457 non-null   category
4   Age         457 non-null   int32
5   Height      457 non-null   object
6   Weight      457 non-null   float64
7   College     457 non-null   object
8   Salary      457 non-null   int32
dtypes: category(2), float64(1), int32(3), object(3)
memory usage: 26.7+ KB
```

Sorting using .sort_values() mtd

In [151]:

```
nba=pd.read_csv("nba.csv")
nba.head()
```

Out[151]:

	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 [154]:

```
nba.sort_values("Name").head()
```

Out[154]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
152	Aaron Brooks	Chicago Bulls	0.0	PG	31.0	6-0	161.0	Oregon	2250000.0
356	Aaron Gordon	Orlando Magic	0.0	PF	20.0	6-9	220.0	Arizona	4171680.0
328	Aaron Harrison	Charlotte Hornets	9.0	SG	21.0	6-6	210.0	Kentucky	525093.0
404	Adreian Payne	Minnesota Timberwolves	33.0	PF	25.0	6-10	237.0	Michigan State	1938840.0
312	Al Horford	Atlanta Hawks	15.0	C	30.0	6-10	245.0	Florida	12000000.0

In [155]:

```
nba.sort_values("Age",ascending=False).head()
```

Out[155]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
304	Andre Miller	San Antonio Spurs	24.0	PG	40.0	6-3	200.0	Utah	250750.0
400	Kevin Garnett	Minnesota Timberwolves	21.0	PF	40.0	6-11	240.0	NaN	8500000.0
298	Tim Duncan	San Antonio Spurs	21.0	C	40.0	6-11	250.0	Wake Forest	5250000.0
261	Vince Carter	Memphis Grizzlies	15.0	SG	39.0	6-6	220.0	North Carolina	4088019.0
102	Pablo Prigioni	Los Angeles Clippers	9.0	PG	39.0	6-3	185.0	NaN	947726.0

In [158]:

```
nba.sort_values("Salary",ascending=False).head()
```

Out[158]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
109	Kobe Bryant	Los Angeles Lakers	24.0	SF	37.0	6-6	212.0	NaN	25000000.0
169	LeBron James	Cleveland Cavaliers	23.0	SF	31.0	6-8	250.0	NaN	22970500.0
33	Carmelo Anthony	New York Knicks	7.0	SF	32.0	6-8	240.0	Syracuse	22875000.0
251	Dwight Howard	Houston Rockets	12.0	C	30.0	6-11	265.0	NaN	22359364.0
339	Chris Bosh	Miami Heat	1.0	PF	32.0	6-11	235.0	Georgia Tech	22192730.0

In [159]:

```
nba.sort_values("Salary",ascending=False,na_position="first").head()
```

Out[159]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
2	John Holland	Boston Celtics	30.0	SG	27.0	6-5	205.0	Boston University	NaN
46	Elton Brand	Philadelphia 76ers	42.0	PF	37.0	6-9	254.0	Duke	NaN
171	Dahntay Jones	Cleveland Cavaliers	30.0	SG	35.0	6-6	225.0	Duke	NaN
264	Jordan Farmar	Memphis Grizzlies	4.0	PG	29.0	6-2	180.0	UCLA	NaN
269	Ray McCallum	Memphis Grizzlies	5.0	PG	24.0	6-3	190.0	Detroit	NaN

In [161]:

```
nba.sort_values(["Team", "Name"]).head()
```

Out[161]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
312	Al Horford	Atlanta Hawks	15.0	C	30.0	6-10	245.0	Florida	12000000.0
318	Dennis Schroder	Atlanta Hawks	17.0	PG	22.0	6-1	172.0	NaN	1763400.0
323	Jeff Teague	Atlanta Hawks	0.0	PG	27.0	6-2	186.0	Wake Forest	8000000.0
309	Kent Bazemore	Atlanta Hawks	24.0	SF	26.0	6-5	201.0	Old Dominion	2000000.0
311	Kirk Hinrich	Atlanta Hawks	12.0	SG	35.0	6-4	190.0	Kansas	2854940.0

In [162]:

```
nba.sort_values(["Team", "Name"],ascending=False).head()#desc,desc
```

Out[162]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
379	Ramon Sessions	Washington Wizards	7.0	PG	30.0	6-3	190.0	Nevada	2170465.0
378	Otto Porter Jr.	Washington Wizards	22.0	SF	23.0	6-8	198.0	Georgetown	4662960.0
375	Nene Hilario	Washington Wizards	42.0	C	33.0	6-11	250.0	NaN	13000000.0
376	Markieff Morris	Washington Wizards	5.0	PF	26.0	6-10	245.0	Kansas	8000000.0
381	Marcus Thornton	Washington Wizards	15.0	SF	29.0	6-4	205.0	LSU	200600.0

In [163]:

```
nba.sort_values(["Team", "Name"], ascending=[True, False]).head()#asc, desc
```

Out[163]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
322	Walter Tavares	Atlanta Hawks	22.0	C	24.0	7-3	260.0	NaN	1000000.0
310	Tim Hardaway Jr.	Atlanta Hawks	10.0	SG	24.0	6-6	205.0	Michigan	1304520.0
321	Tiago Splitter	Atlanta Hawks	11.0	C	31.0	6-11	245.0	NaN	9756250.0
320	Thabo Sefolosha	Atlanta Hawks	25.0	SF	32.0	6-7	220.0	NaN	4000000.0
315	Paul Millsap	Atlanta Hawks	4.0	PF	31.0	6-8	246.0	Louisiana Tech	18671659.0

sort DF with .sort_index() Mtd

In [164]:

```
nba=pd.read_csv("nba.csv")
nba.head()
```

Out[164]:

	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 [166]:

```
nba.sort_values(["Number", "Salary", "Name"], inplace=True)
```

In [167]:

```
nba.head()
```

Out[167]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
291	Orlando Johnson	New Orleans Pelicans	0.0	SG	27.0	6-5	220.0	UC Santa Barbara	55722.0
248	Andrew Goudelock	Houston Rockets	0.0	PG	27.0	6-3	200.0	Charleston	200600.0
347	Josh Richardson	Miami Heat	0.0	SG	22.0	6-6	200.0	Tennessee	525093.0
266	JaMychal Green	Memphis Grizzlies	0.0	PF	25.0	6-9	227.0	Alabama	845059.0
47	Isaiah Canaan	Philadelphia 76ers	0.0	PG	25.0	6-0	201.0	Murray State	947276.0

In [168]:

```
nba.tail()
```

Out[168]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary
372	Drew Gooden	Washington Wizards	90.0	PF	34.0	6-10	250.0	Kansas	3300000.0
5	Amir Johnson	Boston Celtics	90.0	PF	29.0	6-9	240.0	NaN	12000000.0
68	Lucas Nogueira	Toronto Raptors	92.0	C	23.0	7-0	220.0	NaN	1842000.0
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6-6	235.0	Marquette	6796117.0
457	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [173]:

```
nba.sort_index(inplace=True)
nba.head()
```

Out[173]:

	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

Rank values with .rank(mtd)

In [185]:

```
nba=pd.read_csv("nba.csv").dropna(how="all")
nba["Salary"]=nba["Salary"].fillna(0).astype("int")
nba.head()
```

Out[185]:

	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
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6-6	235.0	Marquette	6796117
2	John Holland	Boston Celtics	30.0	SG	27.0	6-5	205.0	Boston University	0
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6-5	185.0	Georgia State	1148640
4	Jonas Jerebko	Boston Celtics	8.0	PF	29.0	6-10	231.0	NaN	5000000

In [187]:

```
nba["Salary Rank"]=nba["Salary"].rank(ascending=False).astype("int")
```

In [188]:

```
nba.head()
```

Out[188]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary	Salary Rank
0	Avery Bradley	Boston Celtics	0.0	PG	25.0	6-2	180.0	Texas	7730337	97
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6-6	235.0	Marquette	6796117	110
2	John Holland	Boston Celtics	30.0	SG	27.0	6-5	205.0	Boston University	0	452
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6-5	185.0	Georgia State	1148640	322
4	Jonas Jerebko	Boston Celtics	8.0	PF	29.0	6-10	231.0	NaN	5000000	147

In [190]:

```
nba.sort_values("Salary",ascending=False).head()
```

Out[190]:

	Name	Team	Number	Position	Age	Height	Weight	College	Salary	Salary Rank
109	Kobe Bryant	Los Angeles Lakers	24.0	SF	37.0	6-6	212.0	NaN	25000000	1
169	LeBron James	Cleveland Cavaliers	23.0	SF	31.0	6-8	250.0	NaN	22970500	2
33	Carmelo Anthony	New York Knicks	7.0	SF	32.0	6-8	240.0	Syracuse	22875000	3
251	Dwight Howard	Houston Rockets	12.0	C	30.0	6-11	265.0	NaN	22359364	4
339	Chris Bosh	Miami Heat	1.0	PF	32.0	6-11	235.0	Georgia Tech	22192730	5