In [113]:
```
#import library
import pandas as pd
```

In [114]:
```
#read trophies data
dt = pd.read_csv('trophies2.csv', header = [1])
dt_fix = dt
dt.head(11)
```

Out[114]:

	Football Team	2001-2019	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006	2006-2007	2007-2008	...	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016	2016-2017
0	FC Porto	75	48	NaN	52.0	56.0	NaN	59.0	60.0	61.0	...	NaN	68.0	71.0	73.0	NaN	NaN	NaN	NaN
1	S.L. Benfica	81	62	NaN	NaN	63.0	NaN	65.0	NaN	NaN	...	67.0	68.0	69.0	NaN	73.0	75.0	78.0	81.0
2	FC Bayern Munich	68	38	NaN	40.0	41.0	43.0	45.0	46.0	48.0	...	51.0	NaN	52.0	57.0	59.0	60.0	63.0	65.0
3	AFC Ajax	69	54	57.0	NaN	58.0	59.0	61.0	63.0	NaN	...	64.0	65.0	66.0	68.0	69.0	NaN	NaN	NaN
4	Manchester United F.C.	67	45	NaN	47.0	48.0	NaN	49.0	51.0	55.0	...	59.0	61.0	NaN	63.0	NaN	NaN	65.0	67.0
5	Arsenal F.C.	45	32	35.0	36.0	38.0	39.0	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	41.0	43.0	NaN	45.0
6	FC Barcelona	93	60	NaN	NaN	NaN	62.0	65.0	NaN	NaN	...	73.0	78.0	79.0	81.0	NaN	86.0	89.0	90.0
7	A.C. Milan	50	42	NaN	44.0	45.0	NaN	NaN	47.0	NaN	...	NaN	49.0	NaN	NaN	NaN	NaN	50.0	NaN
8	Real Madrid C.F.	90	65	68.0	70.0	NaN	NaN	NaN	71.0	73.0	...	NaN	74.0	76.0	NaN	80.0	NaN	83.0	88.0
9	Juventus F.C.	67	47	48.0	50.0	51.0	NaN	NaN	52.0	NaN	...	NaN	NaN	54.0	56.0	57.0	60.0	62.0	64.0
10	Liverpool F.C.	62	56	NaN	57.0	NaN	59.0	61.0	NaN	NaN	...	NaN	NaN	62.0	NaN	NaN	NaN	NaN	NaN

11 rows × 21 columns

In [115]:
```
#review data
print(dt.shape)
```

```
(11, 21)
```

In [116]:
```
print(dt.isnull().sum())
```

```
Football Team    0
2001-2019        0
2000-2001        0
2001-2002        7
2002-2003        3
2003-2004        3
2004-2005        6
2005-2006        4
2006-2007        4
2007-2008        7
2008-2009        8
2009-2010        6
2010-2011        4
2011-2012        3
2012-2013        5
2013-2014        5
2014-2015        6
2015-2016        4
2016-2017        4
2017-2018        6
2018-2019        5
dtype: int64
```

In [117]:
```
#fill null
dt_fix = dt_fix.fillna(0)
for row in range(11):
    for col in range(21):
        #print(dt_fix.iloc[row,col])
        if dt_fix.iloc[row,col] == 0:
            dt_fix.iloc[row,col] = dt_fix.iloc[row,col-1]
print(dt_fix.isnull().sum())
```

```
Football Team    0
2001-2019        0
2000-2001        0
2001-2002        0
2002-2003        0
2003-2004        0
2004-2005        0
2005-2006        0
2006-2007        0
2007-2008        0
2008-2009        0
2009-2010        0
2010-2011        0
2011-2012        0
2012-2013        0
2013-2014        0
2014-2015        0
2015-2016        0
2016-2017        0
2017-2018        0
2018-2019        0
dtype: int64
```

In [123]:
```
#review data type
print(dt.dtypes)
cols=[]
for col in dt_fix:
    if col != 'Football Team':
        cols.append(col)
dt_fix[cols] = dt_fix[cols].astype('int64')
season_name = cols
if '2001-2019' in season_name:
    season_name = season_name.remove('2001-2019')
team_name = dt_fix['Football Team'].tolist()
```

```
Football Team     object
2001-2019          int64
2000-2001          int64
2001-2002        float64
2002-2003        float64
2003-2004        float64
2004-2005        float64
2005-2006        float64
2006-2007        float64
2007-2008        float64
2008-2009        float64
2009-2010        float64
2010-2011        float64
2011-2012        float64
2012-2013        float64
2013-2014        float64
2014-2015        float64
2015-2016        float64
2016-2017        float64
2017-2018        float64
2018-2019        float64
dtype: object
```

In [119]:
```
#fix column-related issues
if '2001-2019' in dt_fix.columns:
    dt_fix = dt_fix.drop(columns = '2001-2019')

dt_fix.loc[1,'Football Team']  = 'Benfica'
dt_fix.loc[2,'Football Team']  = 'Bayern Munich'
dt_fix.loc[4,'Football Team']  = 'Man Utd'
dt_fix.loc[5,'Football Team']  = 'Arsenal'
dt_fix.loc[7,'Football Team']  = 'AC Milan'
dt_fix.loc[8,'Football Team']  = 'Real Madrid'
dt_fix.loc[9,'Football Team']  = 'Juventus'
dt_fix.loc[10,'Football Team'] = 'Liverpool'
```

In [120]:
```
dt_fix.head(11)
```

Out[120]:

	Football Team	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016	2016-2017	2017-2018
0	FC Porto	48	48	52	56	56	59	60	61	64	64	68	71	73	73	73	73	73	75
1	Benfica	62	62	62	63	63	65	65	65	65	67	68	69	69	73	75	78	81	81
2	Bayern Munich	38	38	40	41	43	45	46	48	48	51	51	52	57	59	60	63	65	68
3	AFC Ajax	54	57	57	58	59	61	63	63	63	64	65	66	68	69	69	69	69	69
4	Man Utd	45	45	47	48	48	49	51	55	57	59	61	61	63	63	63	65	67	67
5	Arsenal	32	35	36	38	39	39	39	39	39	39	39	39	39	41	43	43	45	45
6	FC Barcelona	60	60	60	60	62	65	65	65	71	73	78	79	81	81	86	89	90	93
7	AC Milan	42	42	44	45	45	45	47	47	47	47	49	49	49	49	49	50	50	50
8	Real Madrid	65	68	70	70	70	70	71	73	73	73	74	76	76	80	80	83	88	90
9	Juventus	47	48	50	51	51	51	52	52	52	52	52	54	56	57	60	62	64	67
10	Liverpool	56	56	57	57	59	61	61	61	61	61	61	62	62	62	62	62	62	62

In [124]:
```
dt_fix.groupby(['Football Team','2001-2002']).sum()
teams = []
seasons = []
Season = []
trophies = []
for team in team_name:
    for i in range(19):
        teams.append(team)
for i in range(11):
    seasons.append(season_name)
for i in seasons:
    for y in i:
        Season.append(y)
for row in range(11):
    for col in range(1,20):
        trophies.append(dt_fix.iloc[row,col])

Seasons = pd.DataFrame({'Seasons': Season})
Trophies = pd.DataFrame({'Trophies': trophies})
fb = pd.DataFrame({'Football Teams' : teams})
dt_fix2 = pd.concat([fb,Seasons,Trophies],axis = 1)
```

In [125]:
```
dt_fix2.describe()
```

Out[125]:

	Trophies
count	209.000000
mean	59.875598
std	12.901395
min	32.000000
25%	49.000000
50%	61.000000
75%	68.000000
max	94.000000

In [126]:
```
#check result
print(dt_fix2.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 209 entries, 0 to 208
Data columns (total 3 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Football Teams  209 non-null    object
 1   Seasons         209 non-null    object
 2   Trophies        209 non-null    int64 
dtypes: int64(1), object(2)
memory usage: 5.0+ KB
None
```

In [127]:
```
dt_fix2.to_csv('trophiesclean.csv',index = False)
```