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

In [2]:
```
s2017_df = pd.read_csv('2017_season_data.csv')
```

In [3]:
```
players_df = pd.read_csv('player_data.csv')
```

In [4]:
```
s2017_df.head(1)
```

Out[4]:

	Year	Player	Pos	Age	Tm	G	GS	MP	PER	TS%	...	FT%	ORB	DRB	TRB
0	2017.0	Alex Abrines	SG	23.0	OKC	68.0	6.0	1055.0	10.1	0.56	...	0.898	18.0	68.0	86.0

1 rows × 52 columns

In [5]:
```
players_df.head(1)
```

Out[5]:

	name	year_start	year_end	position	height	weight	birth_date	college
0	Alaa Abdelnaby	1991	1995	F-C	6-10	240.0	June 24, 1968	Duke University

Data Wrangling Activities

1. Merge `s2017_df` and `players_df` with a left join

In [43]:
```
df=s2017_df.merge(players_df,how='left',left_on='Player',right_on='name')
```

In [44]:
```
df.head()
```

Out[44]:

	Year	Player	Pos	Age	Tm	G	GS	MP	PER	TS%	...	PF	PTS	name	y
0	2017.0	Alex Abrines	SG	23.0	OKC	68.0	6.0	1055.0	10.1	0.560	...	114.0	406.0	Alex Abrines	
1	2017.0	Quincy Acy	PF	26.0	TOT	38.0	1.0	558.0	11.8	0.565	...	67.0	222.0	Quincy Acy	
2	2017.0	Quincy Acy	PF	26.0	DAL	6.0	0.0	48.0	-1.4	0.355	...	9.0	13.0	Quincy Acy	
3	2017.0	Quincy Acy	PF	26.0	BRK	32.0	1.0	510.0	13.1	0.587	...	58.0	209.0	Quincy Acy	
4	2017.0	Steven Adams	C	23.0	OKC	80.0	80.0	2389.0	16.5	0.589	...	195.0	905.0	Steven Adams	

5 rows × 60 columns

In [45]:
```
df.head(1).T
```

Out[45]:

	0
Year	2017.0
Player	Alex Abrines
Pos	SG
Age	23.0
Tm	OKC
G	68.0
GS	6.0
MP	1055.0
PER	10.1
TS%	0.56
3PAr	0.724
FTr	0.144
ORB%	1.9
DRB%	7.1
TRB%	4.5
AST%	5.5
STL%	1.7
BLK%	0.6
TOV%	8.3
USG%	15.9
blanl	NaN
OWS	1.2
DWS	0.9
WS	2.1
WS/48	0.095
blank2	NaN
OBPM	-0.3
DBPM	-2.2
BPM	-2.5
VORP	-0.1
FG	134.0
FGA	341.0
FG%	0.393
3P	94.0
3PA	247.0

	0
3P%	0.381
2P	40.0
2PA	94.0
2P%	0.426
eFG%	0.531
FT	44.0
FTA	49.0
FT%	0.898
ORB	18.0
DRB	68.0
TRB	86.0
AST	40.0
STL	37.0
BLK	8.0
TOV	33.0
PF	114.0
PTS	406.0
name	Alex Abrines
year_start	2017
year_end	2018
position	G-F
height	6-6
weight	190.0
birth_date	August 1, 1993
college	NaN

In [ ]:
```
# Use it before modifying the `df` to have a copy
# just in case a modification doesn't go as expected
# df_copy = df.copy()
```

2. Are there misses (mismatches) in the resulting dataframe?

In [46]:
```
df['name'].isna().any()
```

Out[46]:
```
False
```

3. How many rows couldn't be matched?

In [47]:
```
df['name'].isna().sum()
```

Out[47]:
```
0
```

4. Extract the names of the players that couldn't be matched

In [48]:
```
df.loc[df['name'].isna()]
```

Out[48]:

Year	Player	Pos	Age	Tm	G	GS	MP	PER	TS%	...	PF	PTS	name	year_start	year_en

0 rows × 60 columns

In [49]:
```
player_misses = list(df.loc[df['name'].isna(),'Player'].values)
```

In [50]:
```
player_misses
```

Out[50]:
```
[]
```

5. Modify `players_df` with the correct names to re-try a successful merge

In [ ]:
```
# Use it before modifying the `df` to have a copy
# just in case a modification doesn't go as expected
# df_copy = df.copy()
```

In [51]:
```
players_df.head()
```

Out[51]:

	name	year_start	year_end	position	height	weight	birth_date	college
0	Alaa Abdelnaby	1991	1995	F-C	6-10	240.0	June 24, 1968	Duke University
1	Zaid Abdul-Aziz	1969	1978	C-F	6-9	235.0	April 7, 1946	Iowa State University
2	Kareem Abdul-Jabbar	1970	1989	C	7-2	225.0	April 16, 1947	University of California, Los Angeles
3	Mahmoud Abdul-Rauf	1991	2001	G	6-1	162.0	March 9, 1969	Louisiana State University
4	Tariq Abdul-Wahad	1998	2003	F	6-6	223.0	November 3, 1974	San Jose State University

In [52]:
```
names_mapping={
    'Luc Mbah a Moute':"Luc Mbah",
    'James Michael McAdoo':'James Michael',
    'Sheldon Mac':'Sheldon McClellan',
    'Metta World Peace':'Metta World',
}
```

In [53]:
```
players_df.loc[players_df['name'].str.contains('Mbah')]
```

Out[53]:

	name	year_start	year_end	position	height	weight	birth_date	college
2595	Luc Mbah	2009	2018	F	6-8	230.0	September 9, 1986	University of California, Los Angeles

In [54]:
```
for new_name,name_2017 in names_mapping.items():
    players_df.loc[players_df['name']==new_name,'name']=name_2017
```

In [55]:
```
player_misses = list(df.loc[df['name'].isna(),'Player'].values)
```

In [56]:
```
player_misses
```

Out[56]:
```
[]
```

7. Remove unnecessary columns

In [58]:
```
columns_to_drop = [
    "Year",
    "PER",
    "TS%",
    "3PAr",
    "FTr",
    "USG%",
    "blanl",
    "OWS",
    "DWS",
    "WS",
    "WS/48",
    "blank2",
    "OBPM",
    "DBPM",
    "BPM",
    "VORP",
    "FG%",
    "3P%",
    "eFG%",
    "FT%",
    "name",
]
```

In [59]:
```
df.drop(columns=columns_to_drop,inplace=True)
```

8. Rename teams to their full name

In [61]:
```
df.head()
```

Out[61]:

	Player	Pos	Age	Tm	G	GS	MP	ORB%	DRB%	TRB%	...	TOV	PF	PTS
0	Alex Abrines	SG	23.0	OKC	68.0	6.0	1055.0	1.9	7.1	4.5	...	33.0	114.0	406.0
1	Quincy Acy	PF	26.0	TOT	38.0	1.0	558.0	3.9	18.0	11.0	...	21.0	67.0	222.0
2	Quincy Acy	PF	26.0	DAL	6.0	0.0	48.0	4.6	15.2	9.7	...	2.0	9.0	13.0
3	Quincy Acy	PF	26.0	BRK	32.0	1.0	510.0	3.8	18.2	11.1	...	19.0	58.0	209.0
4	Steven Adams	C	23.0	OKC	80.0	80.0	2389.0	13.0	15.5	14.2	...	146.0	195.0	905.0

5 rows × 39 columns

In [62]:
```
team_mapping = {
    "OKC": "Oklahoma City Thunder",
    "DAL": "Dallas Mavericks",
    "BRK": "Brooklyn Nets",
    "SAC": "Sacramento Kings",
    "NOP": "New Orleans Pelicans",
    "MIN": "Minnesota Timberwolves",
    "SAS": "San Antonio Spurs",
    "IND": "Indiana Pacers",
    "MEM": "Memphis Grizzlies",
    "POR": "Portland Trail Blazers",
    "CLE": "Cleveland Cavaliers",
    "LAC": "Los Angeles Clippers",
    "PHI": "Philadelphia 76ers",
    "HOU": "Houston Rockets",
    "MIL": "Milwaukee Bucks",
    "NYK": "New York Knicks",
    "DEN": "Denver Nuggets",
    "ORL": "Orlando Magic",
    "MIA": "Miami Heat",
    "PHO": "Phoenix Suns",
    "GSW": "Golden State Warriors",
    "CHO": "Charlotte Hornets",
    "DET": "Detroit Pistons",
    "ATL": "Atlanta Hawks",
    "WAS": "Washington Wizards",
    "LAL": "Los Angeles Lakers",
    "UTA": "Utah Jazz",
    "BOS": "Boston Celtics",
    "CHI": "Chicago Bulls",
    "TOR": "Toronto Raptors"
}
```

In [63]:
```
df["Tm"].replace(team_mapping)
```

Out[63]:
```
0      Oklahoma City Thunder
1                        TOT
2           Dallas Mavericks
3              Brooklyn Nets
4      Oklahoma City Thunder
               ...
600         Charlotte Hornets
601            Boston Celtics
602             Orlando Magic
603             Chicago Bulls
604        Los Angeles Lakers
Name: Tm, Length: 605, dtype: object
```

In [65]:
```
df['Team']=df["Tm"].replace(team_mapping)
```

In [66]:
```
df[['Player','Tm','Team']].head()
```

Out[66]:

	Player	Tm	Team
0	Alex Abrines	OKC	Oklahoma City Thunder
1	Quincy Acy	TOT	TOT
2	Quincy Acy	DAL	Dallas Mavericks
3	Quincy Acy	BRK	Brooklyn Nets
4	Steven Adams	OKC	Oklahoma City Thunder

9. Convert birthday to a datetime object

In [68]:
```
df.head()
```

Out[68]:

	Player	Pos	Age	Tm	G	GS	MP	ORB%	DRB%	TRB%	...	PF	PTS	year_st
0	Alex Abrines	SG	23.0	OKC	68.0	6.0	1055.0	1.9	7.1	4.5	...	114.0	406.0	20
1	Quincy Acy	PF	26.0	TOT	38.0	1.0	558.0	3.9	18.0	11.0	...	67.0	222.0	20
2	Quincy Acy	PF	26.0	DAL	6.0	0.0	48.0	4.6	15.2	9.7	...	9.0	13.0	20
3	Quincy Acy	PF	26.0	BRK	32.0	1.0	510.0	3.8	18.2	11.1	...	58.0	209.0	20
4	Steven Adams	C	23.0	OKC	80.0	80.0	2389.0	13.0	15.5	14.2	...	195.0	905.0	20

5 rows × 40 columns

In [69]:
```
pd.to_datetime(df['birth_date'])
```

Out[69]:
```
0      1993-08-01
1      1990-10-06
2      1990-10-06
3      1990-10-06
4      1993-07-20
          ...
600    1992-10-05
601    1990-01-17
602    1996-09-09
603    1994-02-18
604    1997-03-18
Name: birth_date, Length: 605, dtype: datetime64[ns]
```

In [70]:
```
df['birth_date'] = pd.to_datetime(df['birth_date'])
```

10. Delete all players from the `TOT` team

In [72]:
```
df.loc[df['Player'].duplicated(keep=False),['Player','Tm','Team']].sort_values
```

Out[72]:

	Player	Tm	Team
58	Andrew Bogut	TOT	TOT
60	Andrew Bogut	CLE	Cleveland Cavaliers
59	Andrew Bogut	DAL	Dallas Mavericks
403	Andrew Nicholson	BRK	Brooklyn Nets
402	Andrew Nicholson	WAS	Washington Wizards
401	Andrew Nicholson	TOT	TOT
72	Anthony Brown	ORL	Orlando Magic
71	Anthony Brown	NOP	New Orleans Pelicans
70	Anthony Brown	TOT	TOT
384	Anthony Morrow	OKC	Oklahoma City Thunder

In [75]:
```
# if something goes wrong, just execute this line:
df_copy=df.copy()
```

In [76]:
```
df.loc[df['Tm']=='TOT']
```

Out[76]:

	Player	Pos	Age	Tm	G	GS	MP	ORB%	DRB%	TRB%	...	PF	PT
1	Quincy Acy	PF	26.0	TOT	38.0	1.0	558.0	3.9	18.0	11.0	...	67.0	222.
14	Justin Anderson	SF	23.0	TOT	75.0	10.0	1228.0	6.2	16.9	11.3	...	109.0	532.
32	Matt Barnes	SF	36.0	TOT	74.0	18.0	1777.0	4.6	19.9	12.4	...	185.0	527.
55	Bojan Bogdanovic	SF	27.0	TOT	81.0	54.0	2083.0	1.9	12.3	7.2	...	146.0	1113.
58	Andrew Bogut	C	32.0	TOT	27.0	21.0	583.0	10.6	33.9	21.7	...	86.0	79.
65	Corey Brewer	SF	30.0	TOT	82.0	11.0	1281.0	2.7	11.5	7.1	...	133.0	371.
	Anthony												

In [77]:
```
df=df.loc[df['Tm']!='TOT']
```

In [82]:
```
df.loc[df['Tm']=='TOT'].index
```

Out[82]:
```
Index([  1,  14,  32,  55,  58,  65,  70,  81,  90, 107, 110, 139, 140, 153,
       156, 166, 176, 188, 193, 200, 226, 236, 239, 259, 263, 278, 294, 306,
       314, 328, 355, 358, 361, 383, 401, 405, 411, 415, 435, 443, 446, 458,
       476, 492, 509, 517, 527, 535, 539, 563, 574, 577, 580, 586],
      dtype='int64')
```

In [83]:
```
df.drop(df.loc[df['Tm']=='TOT'].index,inplace=True)
```

Analysis

11. What's the team with the most players in the league?

In [88]:
```
df.groupby('Team')['Player'].size()
```

Out[88]:
```
Team
Atlanta Hawks            22
Boston Celtics           15
Brooklyn Nets            21
Charlotte Hornets        19
Chicago Bulls            18
Cleveland Cavaliers      22
Dallas Mavericks         24
Denver Nuggets           19
Detroit Pistons          15
Golden State Warriors    17
Houston Rockets          18
Indiana Pacers           17
Los Angeles Clippers     15
Los Angeles Lakers       19
Memphis Grizzlies        17
Miami Heat               15
Milwaukee Bucks          20
Minnesota Timberwolves   16
New Orleans Pelicans     27
New York Knicks          16
Oklahoma City Thunder    19
Orlando Magic            19
Philadelphia 76ers       22
Phoenix Suns             18
Portland Trail Blazers   15
Sacramento Kings         19
San Antonio Spurs        17
Toronto Raptors          17
Utah Jazz                15
Washington Wizards       18
Name: Player, dtype: int64
```

In [85]:
```
df['Team'].value_counts()
```

Out[85]:
```
Team
New Orleans Pelicans      27
Dallas Mavericks          24
Cleveland Cavaliers       22
Philadelphia 76ers        22
Atlanta Hawks             22
Brooklyn Nets             21
Milwaukee Bucks           20
Oklahoma City Thunder     19
Denver Nuggets            19
Charlotte Hornets         19
Los Angeles Lakers        19
Sacramento Kings          19
Orlando Magic             19
Phoenix Suns              18
Washington Wizards        18
Houston Rockets           18
Chicago Bulls             18
Golden State Warriors     17
Toronto Raptors           17
Memphis Grizzlies         17
Indiana Pacers            17
San Antonio Spurs         17
Minnesota Timberwolves    16
New York Knicks           16
Miami Heat                15
Los Angeles Clippers      15
Portland Trail Blazers    15
Detroit Pistons           15
Utah Jazz                 15
Boston Celtics            15
Name: count, dtype: int64
```

12. What's the team with the lowest FG ?

In [90]:
```
df.groupby('Team')['FG'].sum().sort_values()
```

Out[90]:
```
Team
Dallas Mavericks          2968.0
Memphis Grizzlies         2984.0
Utah Jazz                 3033.0
Charlotte Hornets         3093.0
Brooklyn Nets             3102.0
Sacramento Kings          3105.0
Orlando Magic             3139.0
Boston Celtics            3168.0
Chicago Bulls             3169.0
Milwaukee Bucks           3190.0
Miami Heat                3202.0
Toronto Raptors           3211.0
New Orleans Pelicans      3218.0
Minnesota Timberwolves    3235.0
Oklahoma City Thunder     3237.0
Los Angeles Clippers      3242.0
Portland Trail Blazers    3243.0
New York Knicks           3244.0
Detroit Pistons           3269.0
Phoenix Suns              3270.0
Houston Rockets           3305.0
Cleveland Cavaliers       3311.0
Philadelphia 76ers        3322.0
Denver Nuggets            3377.0
Indiana Pacers            3379.0
Washington Wizards        3388.0
Los Angeles Lakers        3414.0
San Antonio Spurs         3470.0
Golden State Warriors     3532.0
Atlanta Hawks             3595.0
Name: FG, dtype: float64
```

13. What's the team with the best FG% ?

In [93]:
```
fg_perc_per_team=df.groupby('Team')[['FG','FGA']].sum()
```

In [94]:
```
fg_perc_per_team['FG%']=fg_perc_per_team['FG']/fg_perc_per_team['FGA']
```

In [95]:
```
fg_perc_per_team.sort_values(by="FG%", ascending=False).head()
```

Out[95]:

Team	FG	FGA	FG%
Golden State Warriors	3532.0	7140.0	0.494678
San Antonio Spurs	3470.0	7284.0	0.476387
Los Angeles Clippers	3242.0	6819.0	0.475436
Washington Wizards	3388.0	7136.0	0.474776
Milwaukee Bucks	3190.0	6737.0	0.473505

In [96]:
```
fg_perc_per_team.sort_values(by="FG%").head()
```

Out[96]:

Team	FG	FGA	FG%
Memphis Grizzlies	2984.0	6854.0	0.435366
Dallas Mavericks	2968.0	6750.0	0.439704
Orlando Magic	3139.0	7133.0	0.440067
Philadelphia 76ers	3322.0	7545.0	0.440292
Charlotte Hornets	3093.0	7000.0	0.441857

14. What's the difference between the best and worst 3P shooters (by position)?

In [98]:
```
pos_3p_acc=df.groupby('Pos')[['3P','3PA']].sum()
```

In [99]:
```
pos_3p_acc['3P%']=pos_3p_acc['3P']/pos_3p_acc['3PA']
```

In [101]:
```
pos_3p_acc.sort_values(by='3P%',ascending=False)
```

Out[101]:

Pos	3P	3PA	3P%
SG	7776.0	21106.0	0.368426
PG	5662.0	15761.0	0.359241
C	1486.0	4210.0	0.352969
SF	5638.0	16043.0	0.351431
PF	3514.0	10210.0	0.344172

In [102]:
```
pos_3p_acc['3P%'].max()-pos_3p_acc['3P%'].min()
```

Out[102]:
```
0.024253659969040164
```

15. Find the best scorers in each team

In [103]:
```
df['Tm'].value_counts()
```

Out[103]:
```
Tm
NOP    27
DAL    24
CLE    22
PHI    22
ATL    22
BRK    21
MIL    20
OKC    19
DEN    19
CHO    19
LAL    19
SAC    19
ORL    19
PHO    18
WAS    18
HOU    18
CHI    18
GSW    17
TOR    17
MEM    17
IND    17
SAS    17
MIN    16
NYK    16
MIA    15
LAC    15
POR    15
DET    15
UTA    15
BOS    15
Name: count, dtype: int64
```

In [104]:
```
TEAM='BOS'
max_points_in_team=df.loc[df['Tm']==TEAM,'PTS'].max()
df.loc[(df['Tm']==TEAM)&(df['PTS']==max_points_in_team)]
```

Out[104]:

	Player	Pos	Age	Tm	G	GS	MP	ORB%	DRB%	TRB%	...	PF	PTS	yea
525	Isaiah Thomas	PG	27.0	BOS	76.0	76.0	2569.0	1.9	7.0	4.4	...	167.0	2199.0	

1 rows × 40 columns

In [105]:
```
for team in df['Tm'].unique():
    max_points_in_team=df.loc[df['Tm']==team,'PTS'].max()
    print(tuple(df.loc[(df['Tm']==team)&(df['PTS']==max_points_in_team),['Play
```

```
(array(['Russell Westbrook', 'Oklahoma City Thunder', 2558.0], dtype=objec
t),)
(array(['Harrison Barnes', 'Dallas Mavericks', 1518.0], dtype=object),)
(array(['Brook Lopez', 'Brooklyn Nets', 1539.0], dtype=object),)
(array(['DeMarcus Cousins', 'Sacramento Kings', 1528.0], dtype=object),)
(array(['Anthony Davis', 'New Orleans Pelicans', 2099.0], dtype=object),)
(array(['Karl-Anthony Towns', 'Minnesota Timberwolves', 2061.0],
      dtype=object),)
(array(['Kawhi Leonard', 'San Antonio Spurs', 1888.0], dtype=object),)
(array(['Paul George', 'Indiana Pacers', 1775.0], dtype=object),)
(array(['Marc Gasol', 'Memphis Grizzlies', 1446.0], dtype=object),)
(array(['Damian Lillard', 'Portland Trail Blazers', 2024.0], dtype=object),)
(array(['LeBron James', 'Cleveland Cavaliers', 1954.0], dtype=object),)
(array(['Blake Griffin', 'Los Angeles Clippers', 1316.0], dtype=object),)
(array(['Dario Saric', 'Philadelphia 76ers', 1040.0], dtype=object),)
(array(['James Harden', 'Houston Rockets', 2356.0], dtype=object),)
(array(['Giannis Antetokounmpo', 'Milwaukee Bucks', 1832.0], dtype=object),)
(array(['Carmelo Anthony', 'New York Knicks', 1659.0], dtype=object),)
(array(['Nikola Jokic', 'Denver Nuggets', 1221.0], dtype=object),)
(array(['Evan Fournier', 'Orlando Magic', 1167.0], dtype=object),)
(array(['Goran Dragic', 'Miami Heat', 1483.0], dtype=object),)
(array(['Devin Booker', 'Phoenix Suns', 1726.0], dtype=object),)
(array(['Stephen Curry', 'Golden State Warriors', 1999.0], dtype=object),)
(array(['Kemba Walker', 'Charlotte Hornets', 1830.0], dtype=object),)
(array(['Tobias Harris', 'Detroit Pistons', 1321.0], dtype=object),)
(array(['Dennis Schroder', 'Atlanta Hawks', 1414.0], dtype=object),)
(array(['John Wall', 'Washington Wizards', 1805.0], dtype=object),)
(array(['Jordan Clarkson', 'Los Angeles Lakers', 1205.0], dtype=object),)
(array(['Gordon Hayward', 'Utah Jazz', 1601.0], dtype=object),)
(array(['Isaiah Thomas', 'Boston Celtics', 2199.0], dtype=object),)
(array(['Jimmy Butler', 'Chicago Bulls', 1816.0], dtype=object),)
(array(['DeMar DeRozan', 'Toronto Raptors', 2020.0], dtype=object),)
```

In [107]:
```
df["Best Score Per Team"]=df.groupby('Team')['PTS'].transform('max')
```

In [109]:
```
df.loc[df['Tm']=='OKC',['Player','Tm','PTS','Best Score Per Team']].sort_value
```

Out[109]:

	Player	Tm	PTS	Best Score Per Team
567	Russell Westbrook	OKC	2558.0	2558.0
421	Victor Oladipo	OKC	1067.0	2558.0
301	Enes Kanter	OKC	1033.0	2558.0
4	Steven Adams	OKC	905.0	2558.0
468	Andre Roberson	OKC	522.0	2558.0
484	Domantas Sabonis	OKC	479.0	2558.0
202	Jerami Grant	OKC	421.0	2558.0
0	Alex Abrines	OKC	406.0	2558.0
315	Joffrey Lauvergne	OKC	286.0	2558.0
384	Anthony Morrow	OKC	230.0	2558.0

In [110]:
```
best_scorers_per_team =  df.loc[
    df['PTS'] == df["Best Score Per Team"],
    ['Player', 'Team', 'Pos', 'PTS']
].sort_values(by='PTS', ascending=False)
```

In [111]:
```
best_scorers_per_team
```

Out[111]:

	Player	Team	Pos	PTS
567	Russell Westbrook	Oklahoma City Thunder	PG	2558.0
214	James Harden	Houston Rockets	PG	2356.0
525	Isaiah Thomas	Boston Celtics	PG	2199.0
122	Anthony Davis	New Orleans Pelicans	C	2099.0
538	Karl-Anthony Towns	Minnesota Timberwolves	C	2061.0
331	Damian Lillard	Portland Trail Blazers	PG	2024.0
130	DeMar DeRozan	Toronto Raptors	SG	2020.0
120	Stephen Curry	Golden State Warriors	PG	1999.0
274	LeBron James	Cleveland Cavaliers	SF	1954.0
324	Kawhi Leonard	San Antonio Spurs	SF	1888.0
19	Giannis Antetokounmpo	Milwaukee Bucks	SF	1832.0
558	Kemba Walker	Charlotte Hornets	PG	1830.0
79	Jimmy Butler	Chicago Bulls	SF	1816.0
559	John Wall	Washington Wizards	PG	1805.0
185	Paul George	Indiana Pacers	SF	1775.0
62	Devin Booker	Phoenix Suns	SG	1726.0
20	Carmelo Anthony	New York Knicks	SF	1659.0
229	Gordon Hayward	Utah Jazz	SF	1601.0
336	Brook Lopez	Brooklyn Nets	C	1539.0
111	DeMarcus Cousins	Sacramento Kings	C	1528.0
31	Harrison Barnes	Dallas Mavericks	PF	1518.0
136	Goran Dragic	Miami Heat	PG	1483.0
180	Marc Gasol	Memphis Grizzlies	C	1446.0
488	Dennis Schroder	Atlanta Hawks	PG	1414.0
222	Tobias Harris	Detroit Pistons	PF	1321.0
209	Blake Griffin	Los Angeles Clippers	PF	1316.0
289	Nikola Jokic	Denver Nuggets	C	1221.0
101	Jordan Clarkson	Los Angeles Lakers	SG	1205.0
171	Evan Fournier	Orlando Magic	SG	1167.0
486	Dario Saric	Philadelphia 76ers	PF	1040.0

In [113]:
```
df.head()
```

Out[113]:

r	Pos	Age	Tm	G	GS	MP	ORB%	DRB%	TRB%	...	PTS	year_start	year_end	pos
‹ s	SG	23.0	OKC	68.0	6.0	1055.0	1.9	7.1	4.5	...	406.0	2017	2018	
/ /	PF	26.0	DAL	6.0	0.0	48.0	4.6	15.2	9.7	...	13.0	2013	2018	
/ /	PF	26.0	BRK	32.0	1.0	510.0	3.8	18.2	11.1	...	209.0	2013	2018	
ı s	C	23.0	OKC	80.0	80.0	2389.0	13.0	15.5	14.2	...	905.0	2014	2018	
ı)	SG	31.0	SAC	61.0	45.0	1580.0	0.7	8.4	4.6	...	515.0	2008	2018	

ł1 columns

In [115]:
```
df.groupby("Team")['birth_date'].mean().sort_values(ascending=False)
```

Out[115]:
```
Team
Portland Trail Blazers   1992-03-14 08:00:00.000000000
Toronto Raptors          1991-04-16 16:56:28.235294080
Boston Celtics           1991-04-04 19:12:00.000000000
Orlando Magic            1991-01-31 16:25:15.789473664
Denver Nuggets           1991-01-14 17:41:03.157894784
Detroit Pistons          1991-01-13 17:36:00.000000000
Phoenix Suns             1991-01-01 02:40:00.000000000
Washington Wizards       1990-12-08 18:40:00.000000000
Charlotte Hornets        1990-11-01 08:50:31.578947328
Brooklyn Nets            1990-09-23 17:08:34.285714304
Chicago Bulls            1990-09-05 00:00:00.000000000
Oklahoma City Thunder    1990-09-04 12:37:53.684210560
Houston Rockets          1990-08-14 06:40:00.000000000
Utah Jazz                1990-04-26 19:12:00.000000000
Philadelphia 76ers       1990-03-25 12:00:00.000000000
Miami Heat               1989-11-12 19:12:00.000000000
Sacramento Kings         1989-11-05 07:34:44.210526336
New York Knicks          1989-10-21 22:30:00.000000000
New Orleans Pelicans     1989-10-18 04:26:40.000000000
Dallas Mavericks         1989-10-02 13:00:00.000000000
Milwaukee Bucks          1989-07-26 15:36:00.000000000
Memphis Grizzlies        1989-06-18 18:21:10.588235264
Minnesota Timberwolves   1989-01-20 04:30:00.000000000
Indiana Pacers           1988-09-05 02:49:24.705882368
Golden State Warriors    1988-08-31 07:03:31.764705920
Los Angeles Lakers       1988-03-27 13:53:41.052631552
Los Angeles Clippers     1986-12-11 09:36:00.000000000
Atlanta Hawks            1985-09-16 08:43:38.181818176
San Antonio Spurs        1985-03-05 21:10:35.294117632
Cleveland Cavaliers      1985-01-02 09:49:05.454545472
Name: birth_date, dtype: datetime64[ns]
```

In [116]:
```
df['Age in Days'] = (pd.Timestamp.now() - df['birth_date']).dt.days
```

In [118]:
```
df.groupby("Team")['Age in Days'].mean().sort_values()
```

Out[118]:
```
Team
Portland Trail Blazers    11441.666667
Toronto Raptors           11774.294118
Boston Celtics            11786.200000
Orlando Magic             11849.315789
Denver Nuggets            11866.263158
Detroit Pistons           11867.266667
Phoenix Suns              11879.888889
Washington Wizards        11903.222222
Charlotte Hornets         11940.631579
Brooklyn Nets             11979.285714
Chicago Bulls             11998.000000
Oklahoma City Thunder     11998.473684
Houston Rockets           12019.722222
Utah Jazz                 12129.200000
Philadelphia 76ers        12161.500000
Miami Heat                12294.200000
Sacramento Kings          12301.684211
New York Knicks           12316.062500
New Orleans Pelicans      12319.814815
Dallas Mavericks          12335.458333
Milwaukee Bucks           12403.350000
Memphis Grizzlies         12441.235294
Minnesota Timberwolves    12590.812500
Indiana Pacers            12727.882353
Golden State Warriors     12732.705882
Los Angeles Lakers        12889.421053
Los Angeles Clippers      13361.600000
Atlanta Hawks             13812.636364
San Antonio Spurs         14007.117647
Cleveland Cavaliers       14069.590909
Name: Age in Days, dtype: float64
```