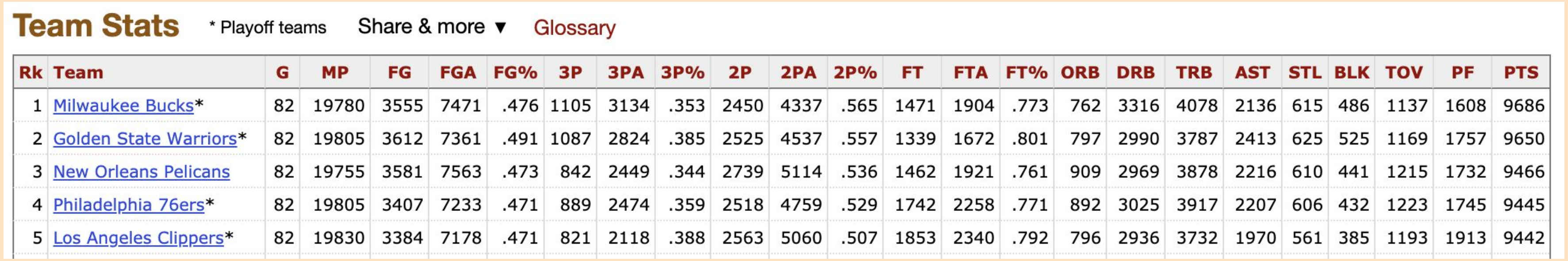
NBA in the Last 5 Years

Holden Bridge
Nikhil Daga
Tiger Hu

Our Data

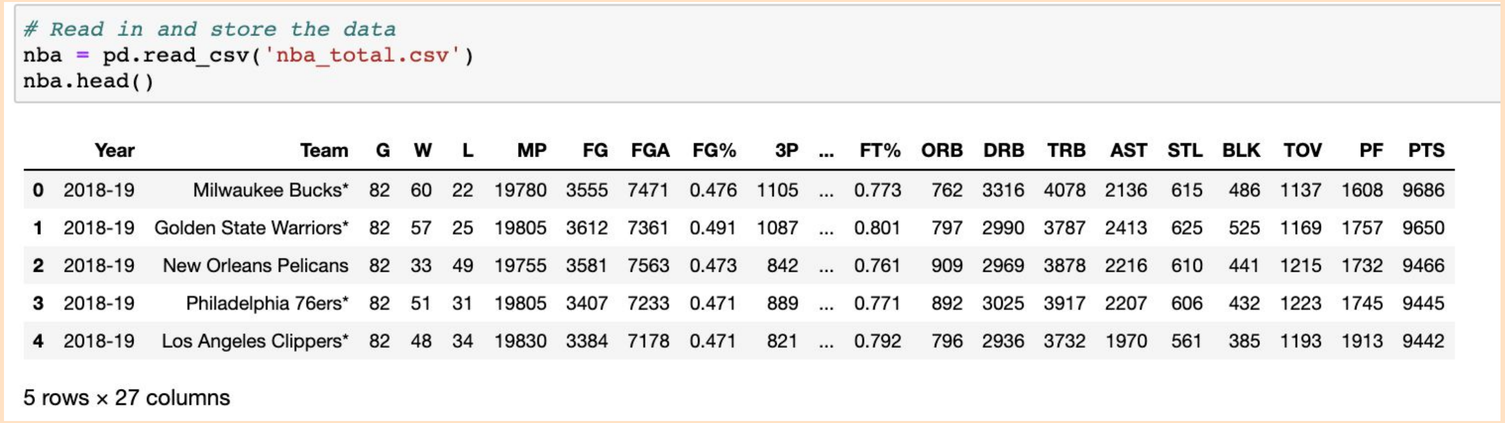


Team Stats * Playoff teams Share & more ▼ Glossary

Rk	Team	G	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS
1	Milwaukee Bucks*	82	19780	3555	7471	.476	1105	3134	.353	2450	4337	.565	1471	1904	.773	762	3316	4078	2136	615	486	1137	1608	9686
2	Golden State Warriors*	82	19805	3612	7361	.491	1087	2824	.385	2525	4537	.557	1339	1672	.801	797	2990	3787	2413	625	525	1169	1757	9650
3	New Orleans Pelicans	82	19755	3581	7563	.473	842	2449	.344	2739	5114	.536	1462	1921	.761	909	2969	3878	2216	610	441	1215	1732	9466
4	Philadelphia 76ers*	82	19805	3407	7233	.471	889	2474	.359	2518	4759	.529	1742	2258	.771	892	3025	3917	2207	606	432	1223	1745	9445
5	Los Angeles Clippers*	82	19830	3384	7178	.471	821	2118	.388	2563	5060	.507	1853	2340	.792	796	2936	3732	1970	561	385	1193	1913	9442

Pre-Processing

- Overall data was very clean
- Pulled multiple tables Basketball reference and merged them into one csv file



```
# Read in and store the data
nba = pd.read_csv('nba_total.csv')
nba.head()
```

	Year	Team	G	W	L	MP	FG	FGA	FG%	3P	...	FT%	ORB	DRB	TRB	AST	STL	BLK	TOV	PF	PTS
0	2018-19	Milwaukee Bucks*	82	60	22	19780	3555	7471	0.476	1105	...	0.773	762	3316	4078	2136	615	486	1137	1608	9686
1	2018-19	Golden State Warriors*	82	57	25	19805	3612	7361	0.491	1087	...	0.801	797	2990	3787	2413	625	525	1169	1757	9650
2	2018-19	New Orleans Pelicans	82	33	49	19755	3581	7563	0.473	842	...	0.761	909	2969	3878	2216	610	441	1215	1732	9466
3	2018-19	Philadelphia 76ers*	82	51	31	19805	3407	7233	0.471	889	...	0.771	892	3025	3917	2207	606	432	1223	1745	9445
4	2018-19	Los Angeles Clippers*	82	48	34	19830	3384	7178	0.471	821	...	0.792	796	2936	3732	1970	561	385	1193	1913	9442

5 rows × 27 columns

What do we want to discover?

- What teams have improved or worsened over the last 5 seasons in terms of wins

- What teams have had sustained success over the last 5 seasons in terms of wins

- Which teams have statistically led the league in recent years

- How has the league statistically trended in recent years

Functions

Create_df

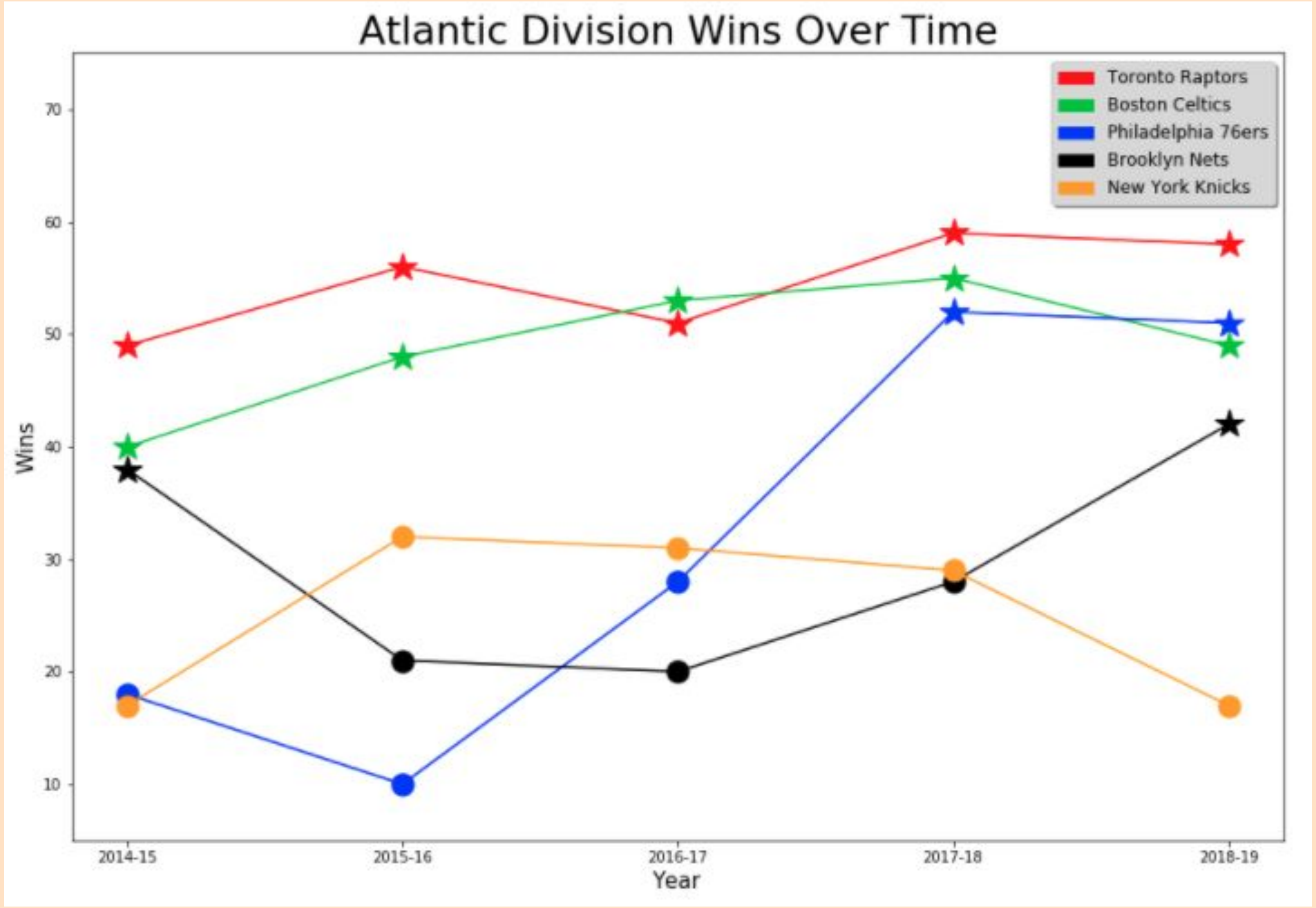
Input - team name

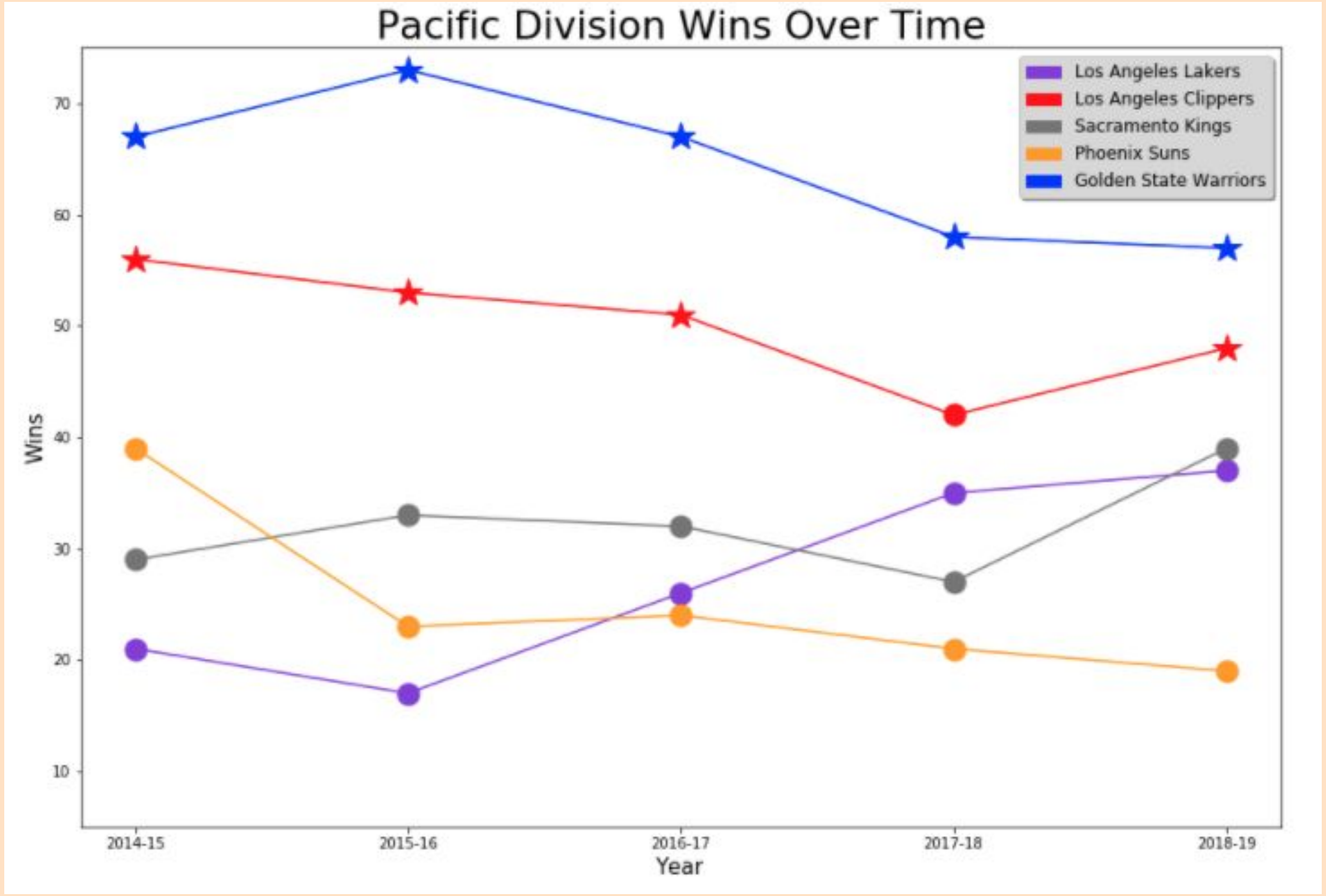
Output- new dataframe with all rows corresponding to a single team

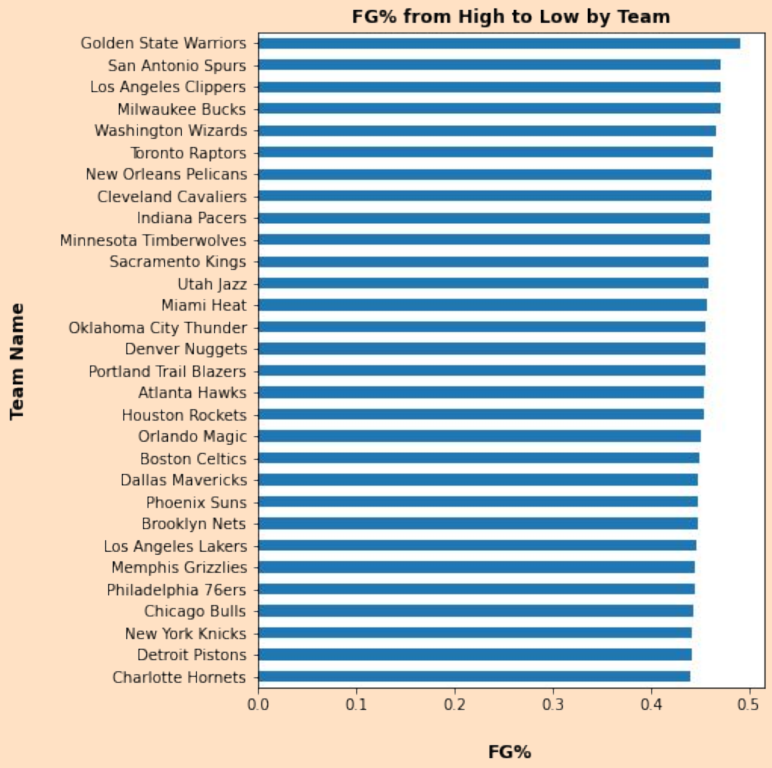
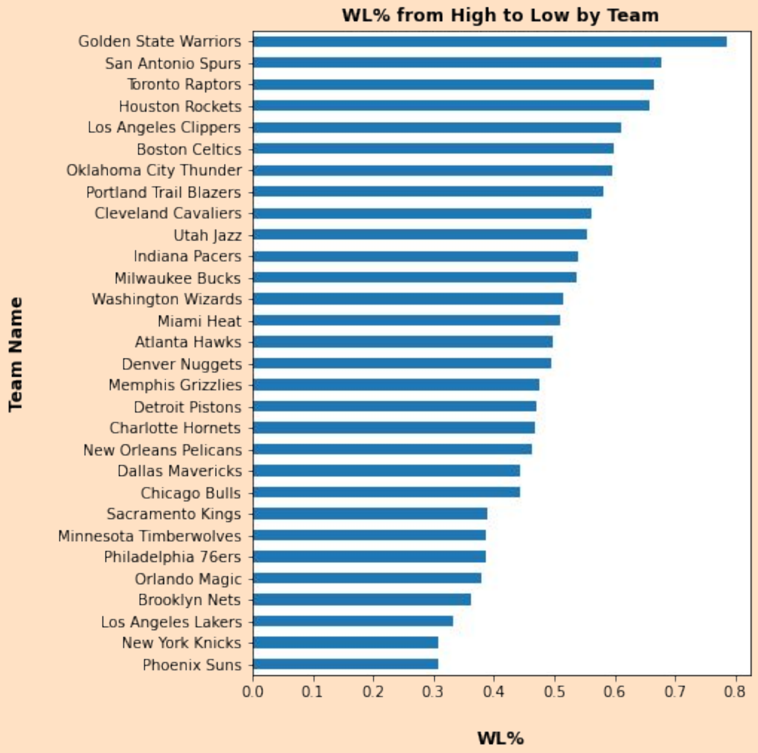
plot_division

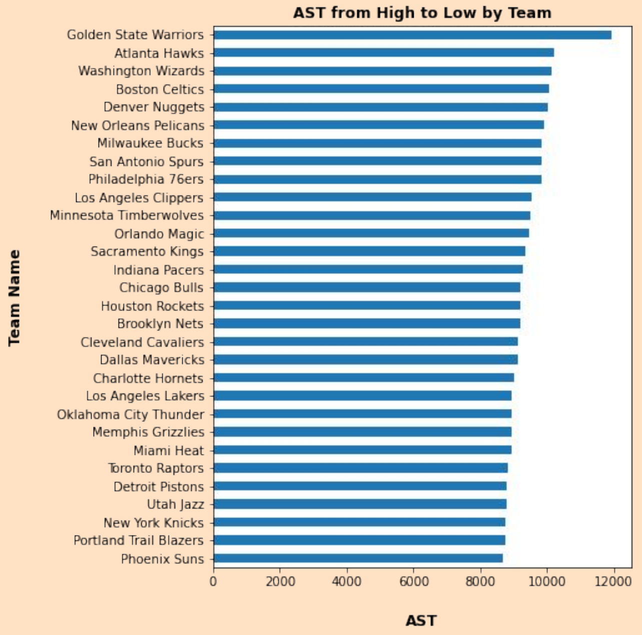
Input - 3 lists, team dataframes, team names, team colors

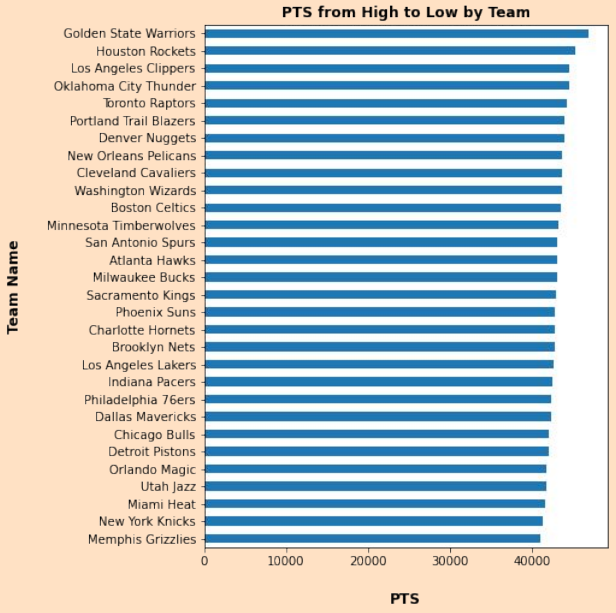
Output- plot of wins over time for all 5 teams

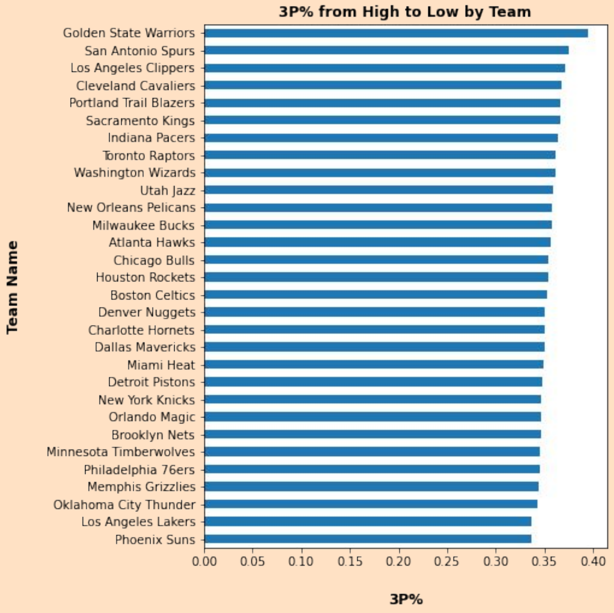

Atlantic Division Wins Over Time
Toronto Raptors
Boston Celtics
Philadelphia 76ers
Brooklyn Nets
New York Knicks
Wins
70
60
50
40
30
20
10
Year
2014-15
2015-16
2016-17
2017-18
2018-19

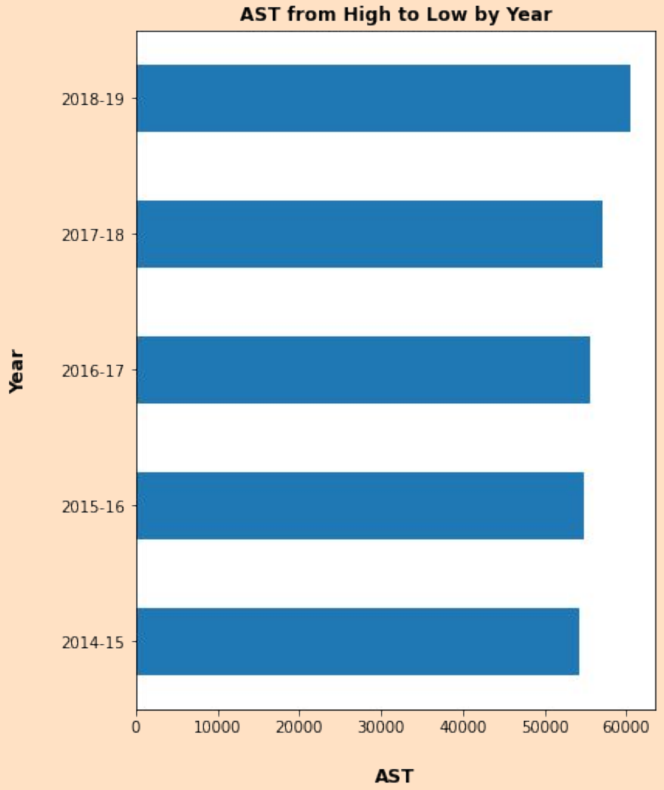

Pacific Division Wins Over Time
Los Angeles Lakers
Los Angeles Clippers
Sacramento Kings
Phoenix Suns
Golden State Warriors
Wins
Year
10
20
30
40
50
60
70
2014-15
2015-16
2016-17
2017-18
2018-19

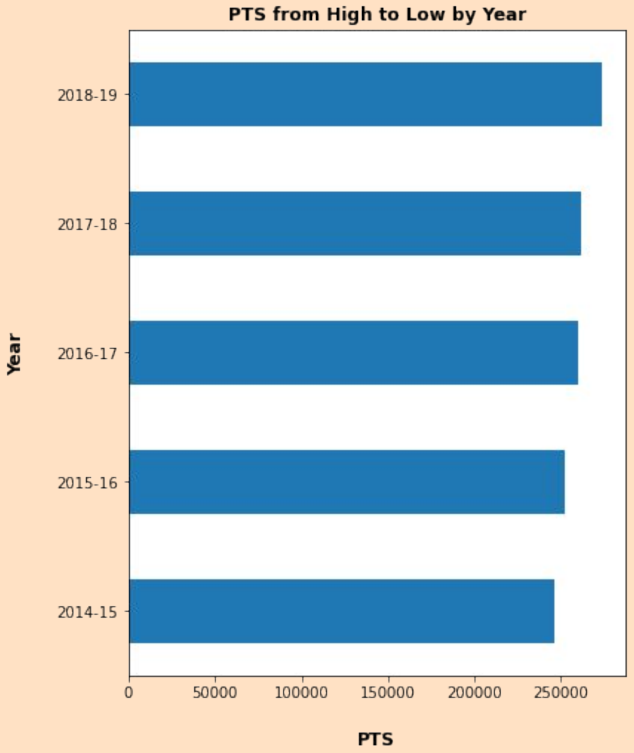

FG% from High to Low by Team
Team Name
FG%
Golden State Warriors
San Antonio Spurs
Los Angeles Clippers
Milwaukee Bucks
Washington Wizards
Toronto Raptors
New Orleans Pelicans
Cleveland Cavaliers
Indiana Pacers
Minnesota Timberwolves
Sacramento Kings
Utah Jazz
Miami Heat
Oklahoma City Thunder
Denver Nuggets
Portland Trail Blazers
Atlanta Hawks
Houston Rockets
Orlando Magic
Boston Celtics
Dallas Mavericks
Phoenix Suns
Brooklyn Nets
Los Angeles Lakers
Memphis Grizzlies
Philadelphia 76ers
Chicago Bulls
New York Knicks
Detroit Pistons
Charlotte Hornets
0.0
0.1
0.2
0.3
0.4
0.5

WL% from High to Low by Team
Team Name
WL%
Golden State Warriors
San Antonio Spurs
Toronto Raptors
Houston Rockets
Los Angeles Clippers
Boston Celtics
Oklahoma City Thunder
Portland Trail Blazers
Cleveland Cavaliers
Utah Jazz
Indiana Pacers
Milwaukee Bucks
Washington Wizards
Miami Heat
Atlanta Hawks
Denver Nuggets
Memphis Grizzlies
Detroit Pistons
Charlotte Hornets
New Orleans Pelicans
Dallas Mavericks
Chicago Bulls
Sacramento Kings
Minnesota Timberwolves
Philadelphia 76ers
Orlando Magic
Brooklyn Nets
Los Angeles Lakers
New York Knicks
Phoenix Suns
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

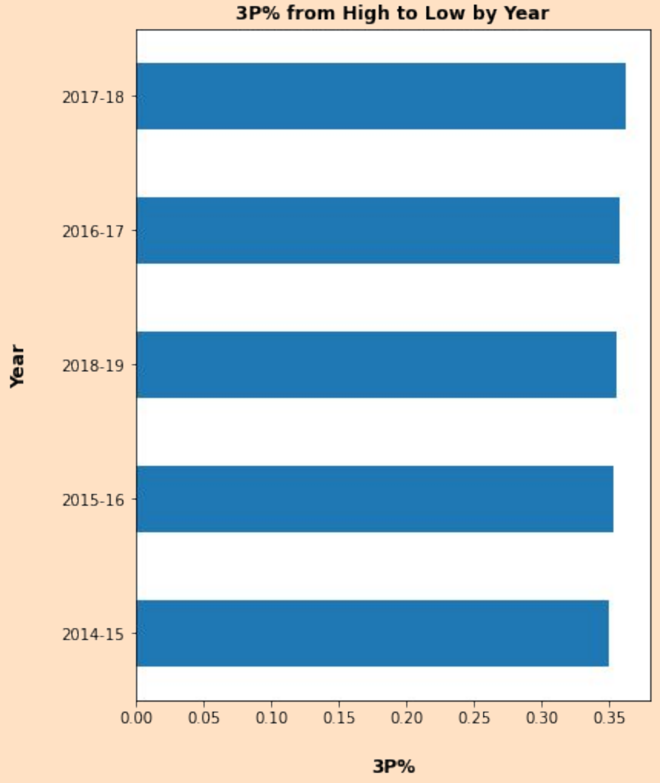

AST from High to Low by Team
Team Name
Golden State Warriors
Atlanta Hawks
Washington Wizards
Boston Celtics
Denver Nuggets
New Orleans Pelicans
Milwaukee Bucks
San Antonio Spurs
Philadelphia 76ers
Los Angeles Clippers
Minnesota Timberwolves
Orlando Magic
Sacramento Kings
Indiana Pacers
Chicago Bulls
Houston Rockets
Brooklyn Nets
Cleveland Cavaliers
Dallas Mavericks
Charlotte Hornets
Los Angeles Lakers
Oklahoma City Thunder
Memphis Grizzlies
Miami Heat
Toronto Raptors
Detroit Pistons
Utah Jazz
New York Knicks
Portland Trail Blazers
Phoenix Suns
0
2000
4000
6000
8000
10000
12000
AST

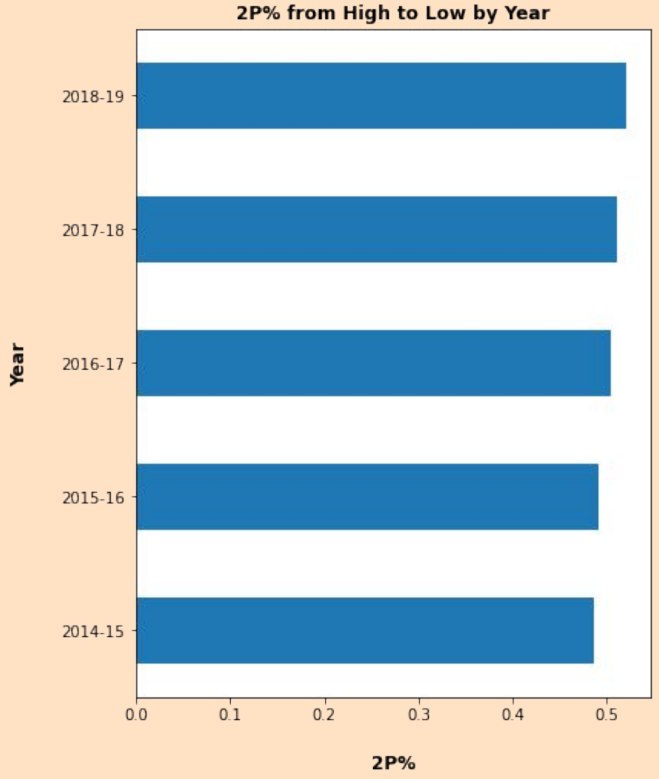

PTS from High to Low by Team
Team Name
Golden State Warriors
Houston Rockets
Los Angeles Clippers
Oklahoma City Thunder
Toronto Raptors
Portland Trail Blazers
Denver Nuggets
New Orleans Pelicans
Cleveland Cavaliers
Washington Wizards
Boston Celtics
Minnesota Timberwolves
San Antonio Spurs
Atlanta Hawks
Milwaukee Bucks
Sacramento Kings
Phoenix Suns
Charlotte Hornets
Brooklyn Nets
Los Angeles Lakers
Indiana Pacers
Philadelphia 76ers
Dallas Mavericks
Chicago Bulls
Detroit Pistons
Orlando Magic
Utah Jazz
Miami Heat
New York Knicks
Memphis Grizzlies
0
10000
20000
30000
40000
PTS


3P% from High to Low by Team
Team Name
Golden State Warriors
San Antonio Spurs
Los Angeles Clippers
Cleveland Cavaliers
Portland Trail Blazers
Sacramento Kings
Indiana Pacers
Toronto Raptors
Washington Wizards
Utah Jazz
New Orleans Pelicans
Milwaukee Bucks
Atlanta Hawks
Chicago Bulls
Houston Rockets
Boston Celtics
Denver Nuggets
Charlotte Hornets
Dallas Mavericks
Miami Heat
Detroit Pistons
New York Knicks
Orlando Magic
Brooklyn Nets
Minnesota Timberwolves
Philadelphia 76ers
Memphis Grizzlies
Oklahoma City Thunder
Los Angeles Lakers
Phoenix Suns
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
3P%


AST from High to Low by Year
Year
2018-19
2017-18
2016-17
2015-16
2014-15
0
10000
20000
30000
40000
50000
60000
AST


PTS from High to Low by Year
Year
2018-19
2017-18
2016-17
2015-16
2014-15
0
50000
100000
150000
200000
250000
PTS


3P% from High to Low by Year
Year
2017-18
2016-17
2018-19
2015-16
2014-15
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
3P%


2P% from High to Low by Year
Year
2018-19
2017-18
2016-17
2015-16
2014-15
0.0
0.1
0.2
0.3
0.4
0.5
2P%

Testing

1. **Unit testing main data frame and create_df**

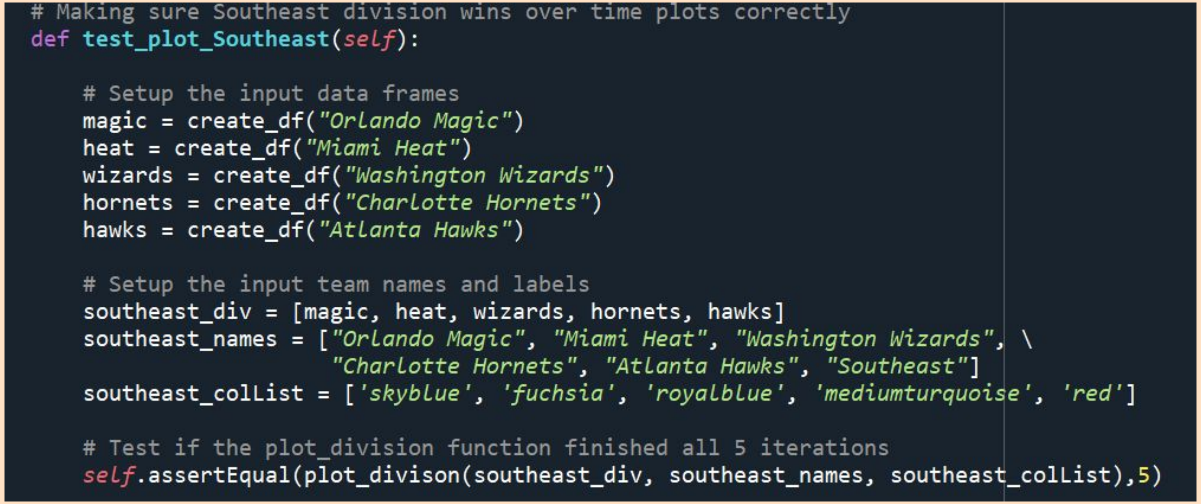
2. **Unit testing plot_division**

3. **Unit testing bestToWorst and bestToWorstYear**

4. **Visual testing all graphs**

The screenshot below shows a unit test of plot_division for Southeast division.

For this test, we first set up the necessary data frames for input through create_df, then the team names and color list. Then, we use assertEqual to test if the plot_division will return an integer of 5, which represents that the function finished all 5 iterations. The plot_division is written in a way that finishing all 5 iterations will ensure correct plot output.

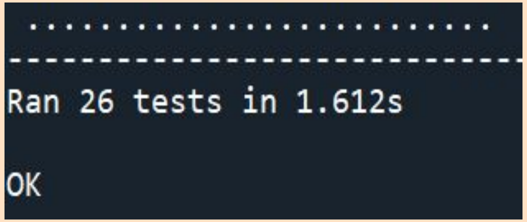


```
# Making sure Southeast division wins over time plots correctly
def test_plot_Southeast(self):

    # Setup the input data frames
    magic = create_df("Orlando Magic")
    heat = create_df("Miami Heat")
    wizards = create_df("Washington Wizards")
    hornets = create_df("Charlotte Hornets")
    hawks = create_df("Atlanta Hawks")

    # Setup the input team names and labels
    southeast_div = [magic, heat, wizards, hornets, hawks]
    southeast_names = ["Orlando Magic", "Miami Heat", "Washington Wizards", \
                       "Charlotte Hornets", "Atlanta Hawks", "Southeast"]
    southeast_colList = ['skyblue', 'fuchsia', 'royalblue', 'mediumturquoise', 'red']

    # Test if the plot_division function finished all 5 iterations
    self.assertEqual(plot_divison(southeast_div, southeast_names, southeast_colList),5)
```



```
 ..........................
---------------------------------
Ran 26 tests in 1.612s

OK
```

Thank you for joining us !