

DSC630-T302: Brandon Mather - Week 1 assignment

```
In [1]: import pandas as pd  
import numpy as np  
import seaborn as sns
```

```
In [2]: stats = pd.read_csv('NFL_standings22.csv')
```

```
In [3]: stats
```

Out[3]:

		Tm	W	L	T	W-L%	PF	PA	PD	MoV	SoS	SRS	OSRS	DSRS
0		Buffalo Bills	13	3	0	0.813	455	286	169	10.6	0.4	10.9	7.1	3.8
1		Miami Dolphins	9	8	0	0.529	397	399	-2	-0.1	2.0	1.8	2.7	-0.9
2		New England Patriots	8	9	0	0.471	364	347	17	1.0	1.0	2.0	-0.3	2.3
3		New York Jets	7	10	0	0.412	296	316	-20	-1.2	2.0	0.8	-4.0	4.8
4		Cincinnati Bengals	12	4	0	0.750	418	322	96	6.0	0.9	6.9	5.5	1.4
5		Baltimore Ravens	10	7	0	0.588	350	315	35	2.1	1.1	3.1	-0.2	3.4
6		Pittsburgh Steelers	9	8	0	0.529	308	346	-38	-2.2	1.5	-0.8	-3.0	2.3
7		Cleveland Browns	7	10	0	0.412	361	381	-20	-1.2	1.1	-0.1	0.7	-0.9
8		Jacksonville Jaguars	9	8	0	0.529	404	350	54	3.2	-1.4	1.8	1.6	0.2
9		Tennessee Titans	7	10	0	0.412	298	359	-61	-3.6	0.0	-3.6	-4.7	1.1
10		Indianapolis Colts	4	12	1	0.265	289	427	-138	-8.1	-0.5	-8.6	-4.9	-3.7
11		Houston Texans	3	13	1	0.206	289	420	-131	-7.7	-0.8	-8.5	-5.4	-3.1
12		Kansas City Chiefs	14	3	0	0.824	496	369	127	7.5	-1.2	6.2	6.8	-0.6
13		Los Angeles Chargers	10	7	0	0.588	391	384	7	0.4	-1.3	-0.9	0.1	-1.0
14		Las Vegas Raiders	6	11	0	0.353	395	418	-23	-1.4	-1.1	-2.5	1.2	-3.7
15		Denver Broncos	5	12	0	0.294	287	359	-72	-4.2	-0.7	-5.0	-5.6	0.6
16		Philadelphia Eagles	14	3	0	0.824	477	344	133	7.8	-1.3	6.5	5.8	0.7
17		Dallas Cowboys	12	5	0	0.706	467	342	125	7.4	-0.8	6.5	5.3	1.3
18		New York Giants	9	7	1	0.559	365	371	-6	-0.4	0.0	-0.4	-0.8	0.4
19		Washington Commanders	8	8	1	0.500	321	343	-22	-1.3	0.4	-0.9	-3.5	2.6
20		Minnesota Vikings	13	4	0	0.765	424	427	-3	-0.2	0.1	-0.1	2.8	-2.9
21		Detroit Lions	9	8	0	0.529	453	427	26	1.5	0.7	2.2	4.7	-2.4
22		Green Bay Packers	8	9	0	0.471	370	371	-1	-0.1	0.4	0.3	-0.3	0.6
23		Chicago Bears	3	14	0	0.176	326	463	-137	-8.1	1.6	-6.4	-2.5	-4.0
24		Tampa Bay Buccaneers	8	9	0	0.471	313	358	-45	-2.6	0.4	-2.3	-3.3	1.1
25		Carolina Panthers	7	10	0	0.412	347	374	-27	-1.6	-0.6	-2.2	-1.3	-0.9
26		New Orleans Saints	7	10	0	0.412	330	345	-15	-0.9	-0.3	-1.2	-2.8	1.7
27		Atlanta Falcons	7	10	0	0.412	365	386	-21	-1.2	-0.9	-2.1	-0.1	-2.0
28		San Francisco 49ers	13	4	0	0.765	450	277	173	10.2	-2.3	7.9	3.3	4.6
29		Seattle Seahawks	9	8	0	0.529	407	401	6	0.4	-0.8	-0.5	1.9	-2.4
30		Los Angeles Rams	5	12	0	0.294	307	384	-77	-4.5	0.5	-4.0	-4.1	0.0
31		Arizona Cardinals	4	13	0	0.235	340	449	-109	-6.4	0.2	-6.2	-1.9	-4.3

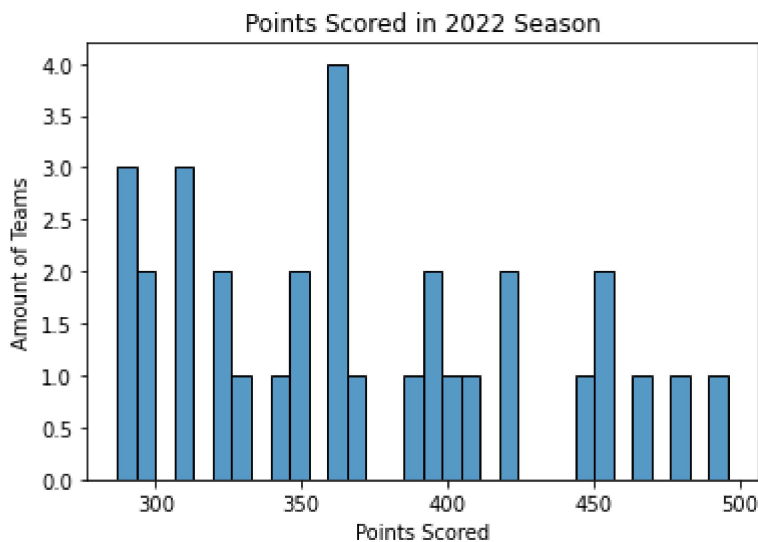
1) Write a summary of your data and identify at least two questions to explore visually with your data.

This data is from the NFL and was collected by Pro-Football-Reference. These are team stats that include the team, wins, losses, ties, win/loss percentage, points for, points against, point differential, margin of victory, strength of schedule, simple rating system, offensive simple rating system, and defensive simple rating system. The first question I want to answer is which has more affect on wins, points for or points against, to see which is more important, offense or defense. Second question I want to answer is to see what the relationship is between strength of schedule and winning/losing.

2) Create a histogram or bar graph from your data.

```
In [19]: histogram = sns.histplot(data = stats, x="PF", bins=32)
histogram.set(xlabel='Points Scored',
              ylabel='Amount of Teams',
              title='Points Scored in 2022 Season')
```

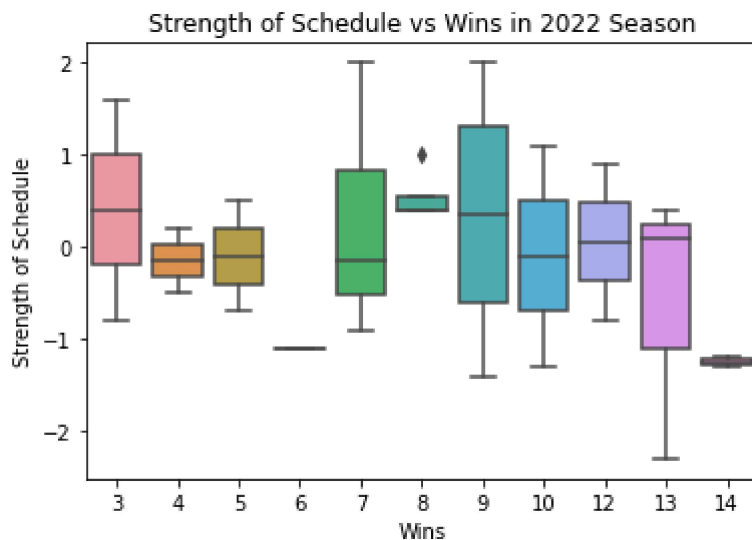
```
Out[19]: [Text(0.5, 0, 'Points Scored'),
Text(0, 0.5, 'Amount of Teams'),
Text(0.5, 1.0, 'Points Scored in 2022 Season')]
```



3) Create a boxplot from your data.

```
In [34]: plot = sns.boxplot(x='W', y='SoS', data=stats)
plot.set(xlabel='Wins',
        ylabel='Strength of Schedule',
        title='Strength of Schedule vs Wins in 2022 Season')
```

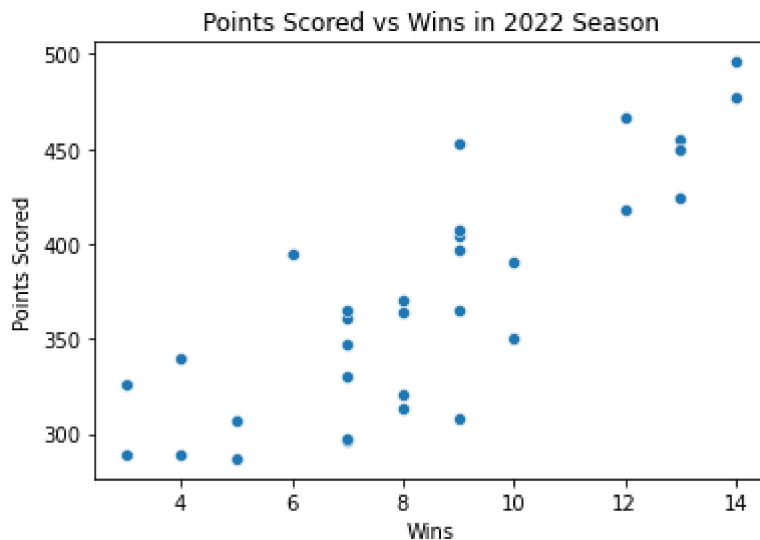
```
Out[34]: [Text(0.5, 0, 'Wins'),
Text(0, 0.5, 'Strength of Schedule'),
Text(0.5, 1.0, 'Strength of Schedule vs Wins in 2022 Season')]
```



4) Create a bivariate plot from your data.

```
In [38]: scatter = sns.scatterplot(data=stats, x="W", y="PF")
scatter.set(xlabel='Wins',
            ylabel='Points Scored',
            title='Points Scored vs Wins in 2022 Season')
```

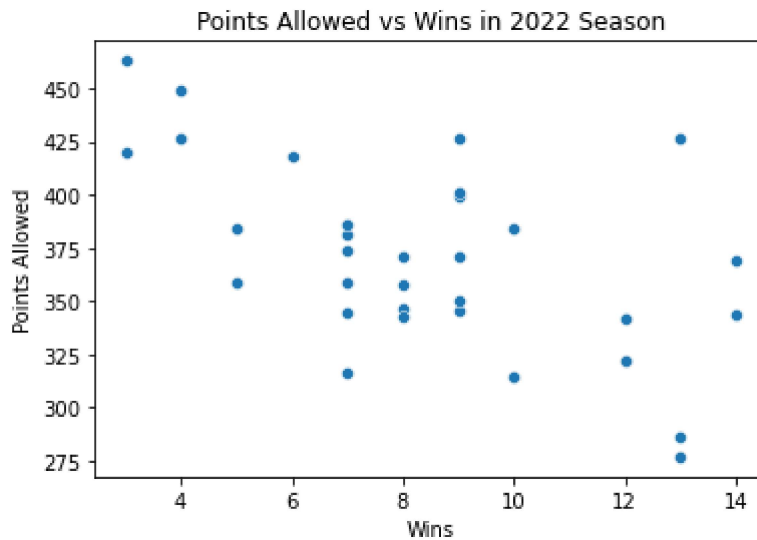
```
Out[38]: [Text(0.5, 0, 'Wins'),
Text(0, 0.5, 'Points Scored'),
Text(0.5, 1.0, 'Points Scored vs Wins in 2022 Season')]
```



5) Create any additional visualizations that will help to answer the question(s) you want to answer.

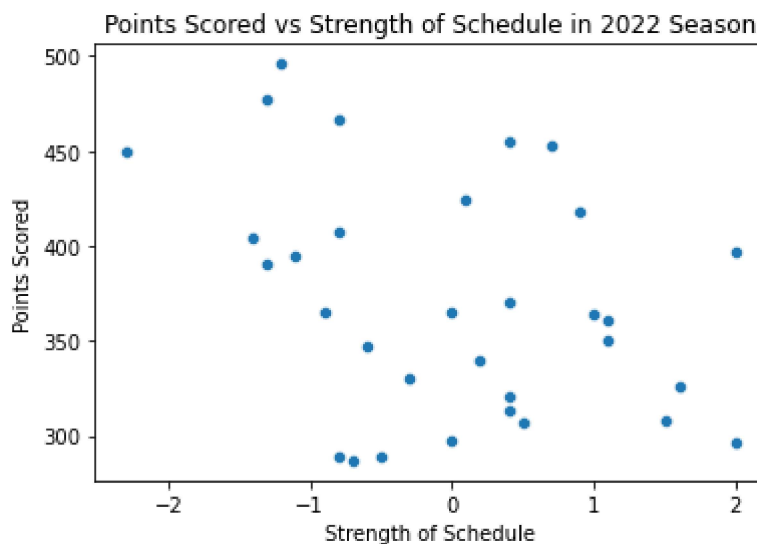
```
In [40]: scatter2 = sns.scatterplot(data=stats, x="W", y="PA")
scatter2.set(xlabel='Wins',
            ylabel='Points Allowed',
            title='Points Allowed vs Wins in 2022 Season')
```

```
Out[40]: [Text(0.5, 0, 'Wins'),
Text(0, 0.5, 'Points Allowed'),
Text(0.5, 1.0, 'Points Allowed vs Wins in 2022 Season')]
```



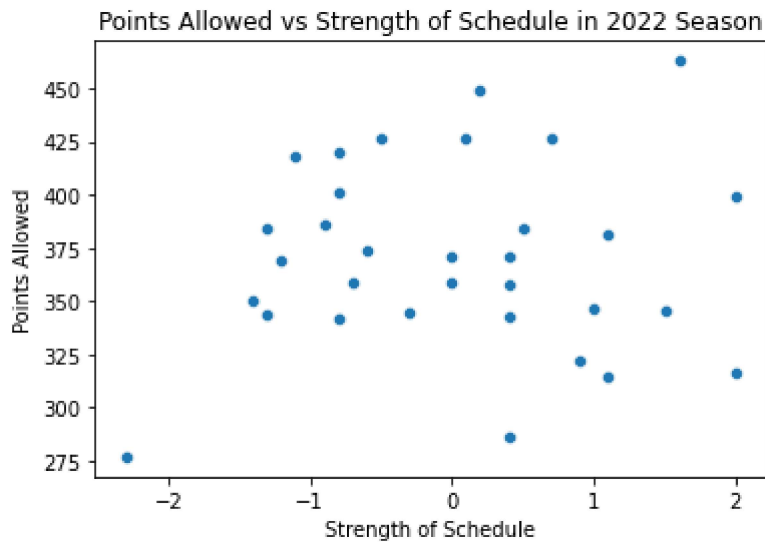
```
In [43]: scatter3 = sns.scatterplot(data=stats, x="SoS", y="PF")
scatter3.set(xlabel='Strength of Schedule',
            ylabel='Points Scored',
            title='Points Scored vs Strength of Schedule in 2022 Season')
```

```
Out[43]: [Text(0.5, 0, 'Strength of Schedule'),
Text(0, 0.5, 'Points Scored'),
Text(0.5, 1.0, 'Points Scored vs Strength of Schedule in 2022 Season')]
```



```
In [42]: scatter4 = sns.scatterplot(data=stats, x="SoS", y="PA")
scatter4.set(xlabel='Strength of Schedule',
            ylabel='Points Allowed',
            title='Points Allowed vs Strength of Schedule in 2022 Season')
```

```
Out[42]: [Text(0.5, 0, 'Strength of Schedule'),
Text(0, 0.5, 'Points Allowed'),
Text(0.5, 1.0, 'Points Allowed vs Strength of Schedule in 2022 Season')]
```



6) Summarize your results and make a conclusion. Explain how you arrived at this conclusion and how your visualizations support your conclusion.

My two questions were: 1) Which effects wins more, offensive scoring or defensive scoring. 2) Does strength of schedule play a role in wins or losses.

To answer the first question I used the histogram and scatter plots to find which effects wins more. I would say based off these visualizations it looks like Offensive Scoring has a bigger effect on winning. The plots slowly follow upwards as wins go up and the scatter plot with Points allowed is a little more random.

To answer the second question it doesn't look like there is a relationship with strength of schedule and winning. I decided to dig in a little deeper to see if Points Scored and Points Allowed were effected by strength of schedule. I don't believe there is anything concrete but we see a slight effect on points scored being a little bit lower with a higher SoS.

References: 1) 2022 NFL Standings & Team stats. Pro Football Reference. (n.d.). <https://www.pro-football-reference.com/years/2022/>