College Football Analysis

Group 6: Maximus Sambucetti, Ricky Li, Hessa Alnuaimi

Introduction

We explored the latest US college football season to understand which team aspects help generate the most wins possible. The dataset used included many variables that allowed for a deeper assessment of each football team. Overall, our investigation seeks to address fundamental questions that dive into the sport's core. By unraveling the dynamics within each team, we aim to contribute valuable insights that can inform strategic decisions and enhance the overall competitiveness of college football.

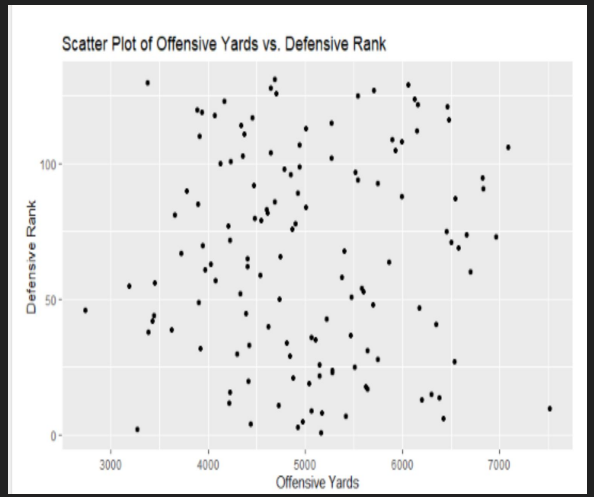
Research Aims

1. Evaluate the key factors influencing the success of teams in the 2022 college football season
2. Exploring team dynamics and strategic elements
3. Effect of special teams on season performance
4. Correlation between offensive and defense

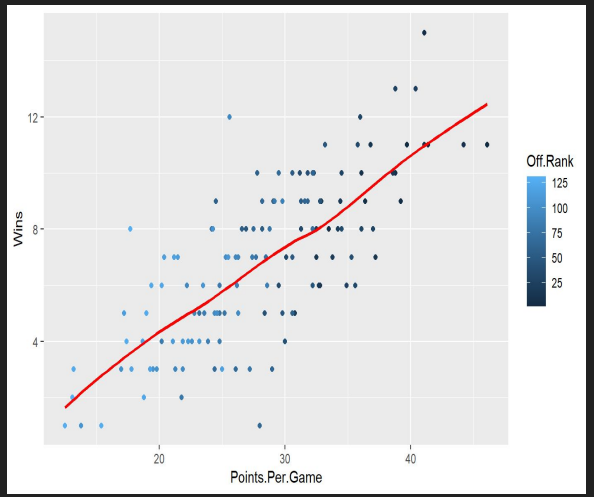
What's the correlation between offensive and defensive performance?

- Both offense and defense performances are crucial for gaining wins
- Below demonstrates that the more points a team gets per game the more wins
- It also shows that the less opponent yards allowed per play, the more wins a team gets

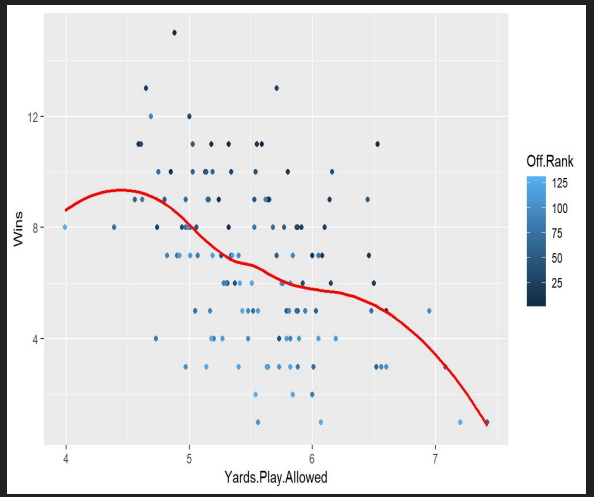
Both



Offense



Defense



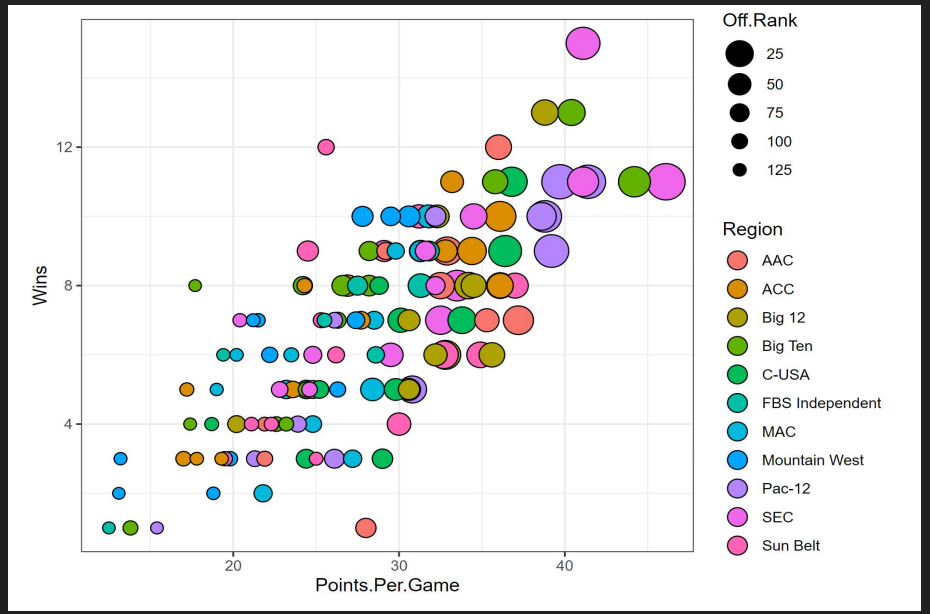
Which Collegiate Conference is the best?

- The average win to loss ratio for each Conference suggested that SEC was the best

- This is also seen within the plot of wins to points per game

- SEC had on average more points per game and more wins

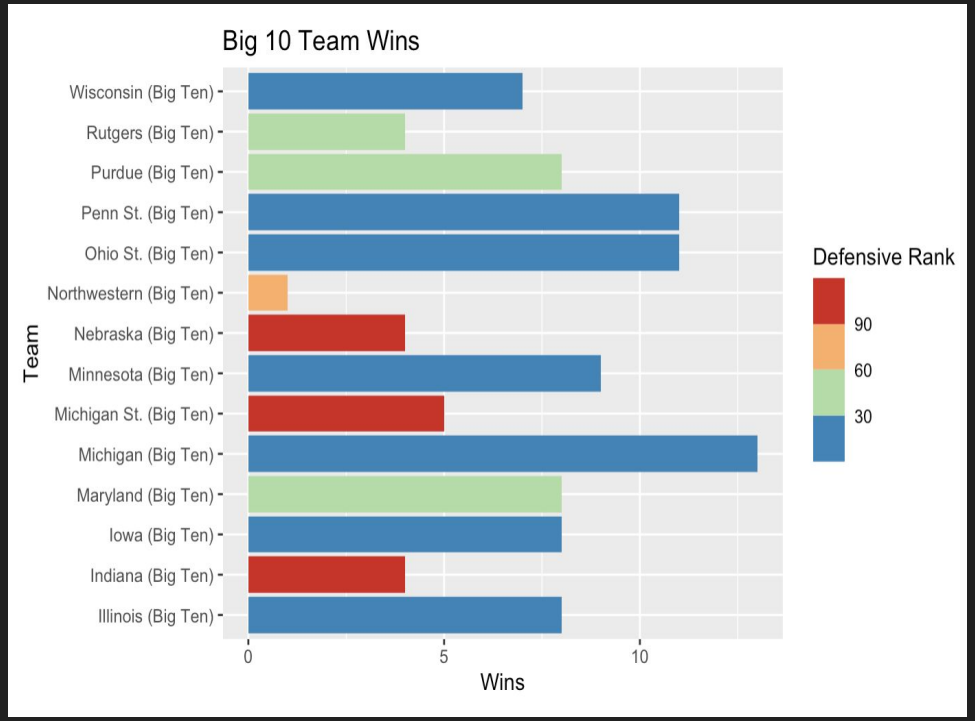
	Region	avg
	<chr>	<dbl>
1	SEC	3.14
2	Big Ten	1.64
3	Big 12	1.5
4	ACC	1.43
5	Pac-12	1.25
6	Sun Belt	0.5
7	AAC	0.455
8	FBS Independent	0.143
9	C-USA	-0.636
10	Mountain West	-0.75
11	MAC	-0.833

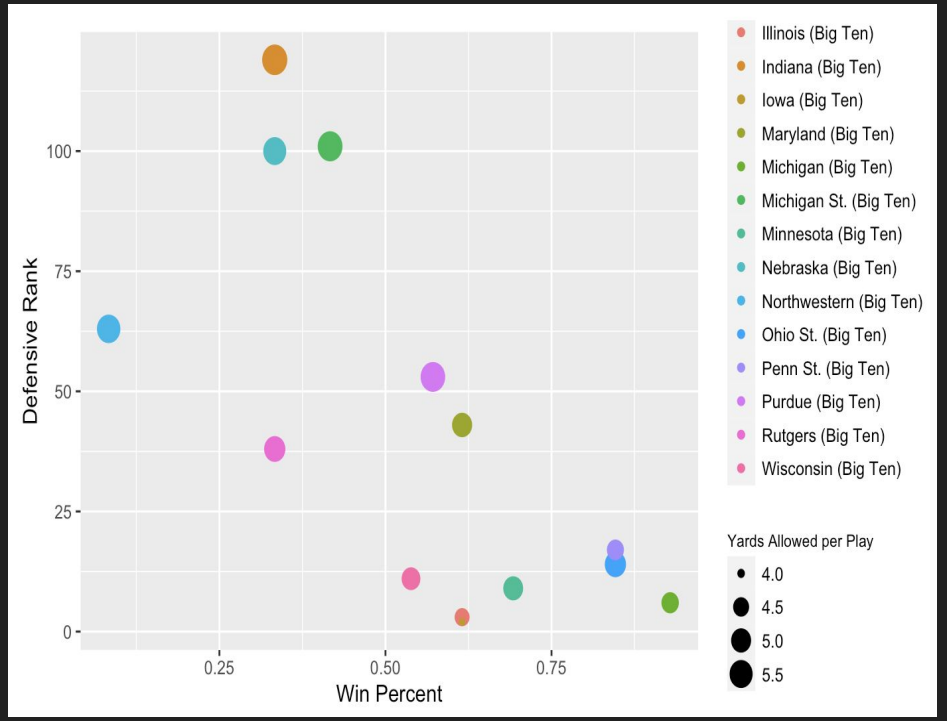


Do Better Defensive Teams Get More Wins?

- look only at teams in our conference (Big 10)

- higher defensive rank means more wins and higher win percentage





What Defensive Features Translate into Wins?

- used multivariable regression with lm()

- fit all defensive features to initial model

R^2 = 0.8076 adj R^2 = 0.788

```
Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                   12.4694664  3.8141527   3.269  0.00141
Def.Rank                      -0.0042148  0.0158200  -0.266  0.79038
Yards.Allowed                 -0.0002169  0.0048883  -0.044  0.96468
Yards.Play.Allowed            -0.9012509  0.5733958  -1.572  0.11868
Total.TDs.Allowed              0.0035156  0.0855946   0.041  0.96731
Yards.Per.Game.Allowed        -0.0005873  0.0631932  -0.009  0.99260
Opp.Completions.Allowed       -0.0050924  0.0084467  -0.603  0.54774
Opp.Pass.Yds.Allowed           0.0103251  0.0077026   1.340  0.18267
Pass.Yards.Per.Game.Allowed   -0.1019829  0.0949713  -1.074  0.28509
Avg.Points.per.Game.Allowed   -0.1531924  0.1667413  -0.919  0.36010
Sacks                          0.0156672  0.1723794   0.091  0.92774
Sack.Yards                     0.0017611  0.0047901   0.368  0.71379
Average.Sacks.per.Game        -0.7055138  2.2195489  -0.318  0.75115
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.329 on 118 degrees of freedom
Multiple R-squared:  0.8076,    Adjusted R-squared:  0.788
F-statistic: 41.28 on 12 and 118 DF,  p-value: < 2.2e-16
```

- backward selection to reduce features

R^2 = 0.8038 adj R^2 = 0.796

```
Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                   12.4257686  1.5016388   8.275 1.64e-13
Yards.Play.Allowed            -0.9216321  0.4259650  -2.164  0.03239
Opp.Pass.Yds.Allowed           0.0097181  0.0007242  13.419  < 2e-16
Pass.Yards.Per.Game.Allowed   -0.1024506  0.0110833  -9.244 8.02e-16
Avg.Points.per.Game.Allowed   -0.1764463  0.0432690  -4.078 8.03e-05
Average.Sacks.per.Game        -0.3667049  0.1372245  -2.672  0.00854
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.303 on 125 degrees of freedom
Multiple R-squared:  0.8038,    Adjusted R-squared:  0.796
F-statistic: 102.4 on 5 and 125 DF,  p-value: < 2.2e-16
```

Most Important Features:

Passing Yards

Sacks

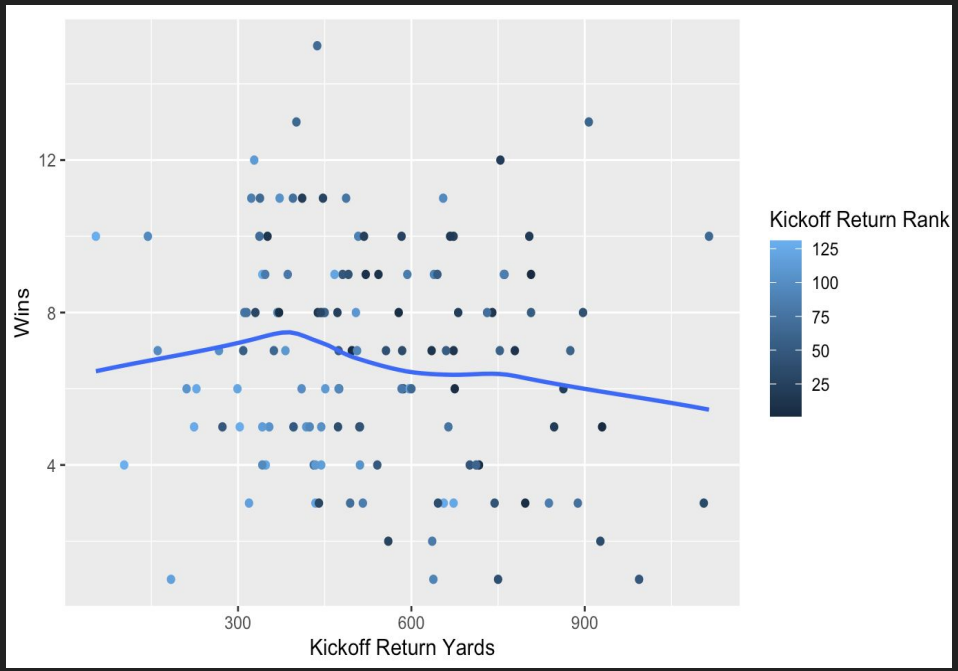
Average Points Allowed per Game

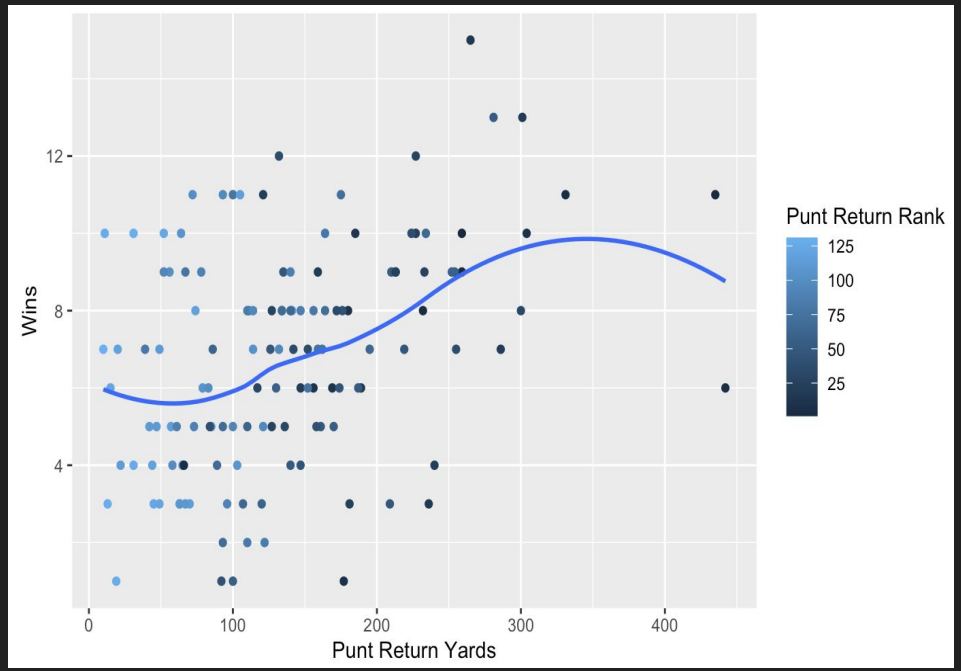
Do Special Teams Contribute to Wins?

- considered only kickoff return yards and punt return yards

- kickoff return yards do not seem to be correlated with wins but more yards yields higher kickoff return rank

- more punt return yards associated with more wins and higher punt return rank





Conclusion

- Although there was high evidence that defence does indeed increase win percentage, offence also played a significant role and increased the chances for having a higher win percentage
- After a deeper look into the defensive stats, it was concluded that the most important stats for a defensive team to win was the passing yards allowed per game, the amount of sacks, and the average points allowed per game
- It was also concluded that when it comes to special teams punt return yards helped teams gain more wins
- Based on all the data that was collected and the answers found, it appears that there are many different features of a college football team that impacts the performance they have for that season

Limitations

- Models cannot capture the unpredictability of injuries, team dynamics, fouls, ect.
- No individual player stats, players can play a massive role in how well the team performs for the season

References

1. College Football Team Stats Seasons 2013 to 2022 (kaggle.com)
2. College Football Exploratory Analysis | Kaggle