

## IST 652 Scripting for Data Analysis



### The Extra Point

Prepared by:

Nabeel Khan  
Renjini Rajan  
Amanda Norwood

December 8<sup>th</sup> 2022

## SCOPE OF PROJECT

- The goal is to provide various NFL personnel and fans of the game with statistics and information that potentially impact game outcome and decision making in various areas, i.e.:
  - ❖ Stadium Type
  - ❖ Surface Type
  - ❖ Training Needs
- Specific questions were brought forth for potential hypothesis testing through the use of multiple data sources



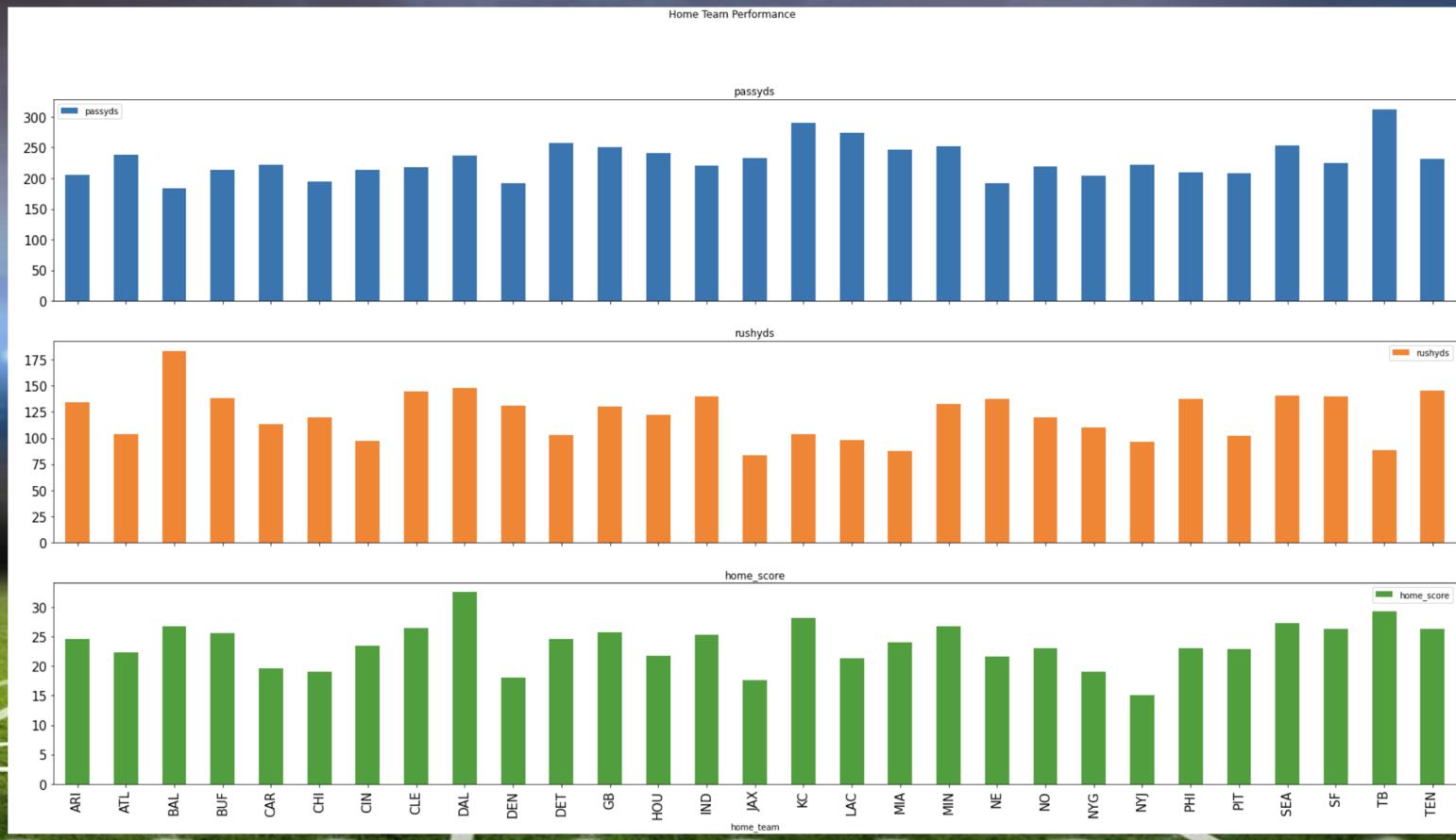
## Data Sources and Cleansing

- Data sources: Kaggle and Github
  - NFL Games (1999 to present) 6409 rows, 44 columns
  - NFL Statistics (2019 to present) 896 rows, 20 columns
  - NFL Teams
    - Team names, team abbreviation, division, conference
  - Special Teams (2000-2016) 568 rows and 26 columns
- Data Cleansing
  - Clean up null fields
  - Replace null fields with 0s
  - Merge NFL Games & NFL Statistics (on Teams)
  - Create boolean fields for certain statistics
  - Subset data to refine the scope of analysis (seasons included)



# Overall Home Team Performance

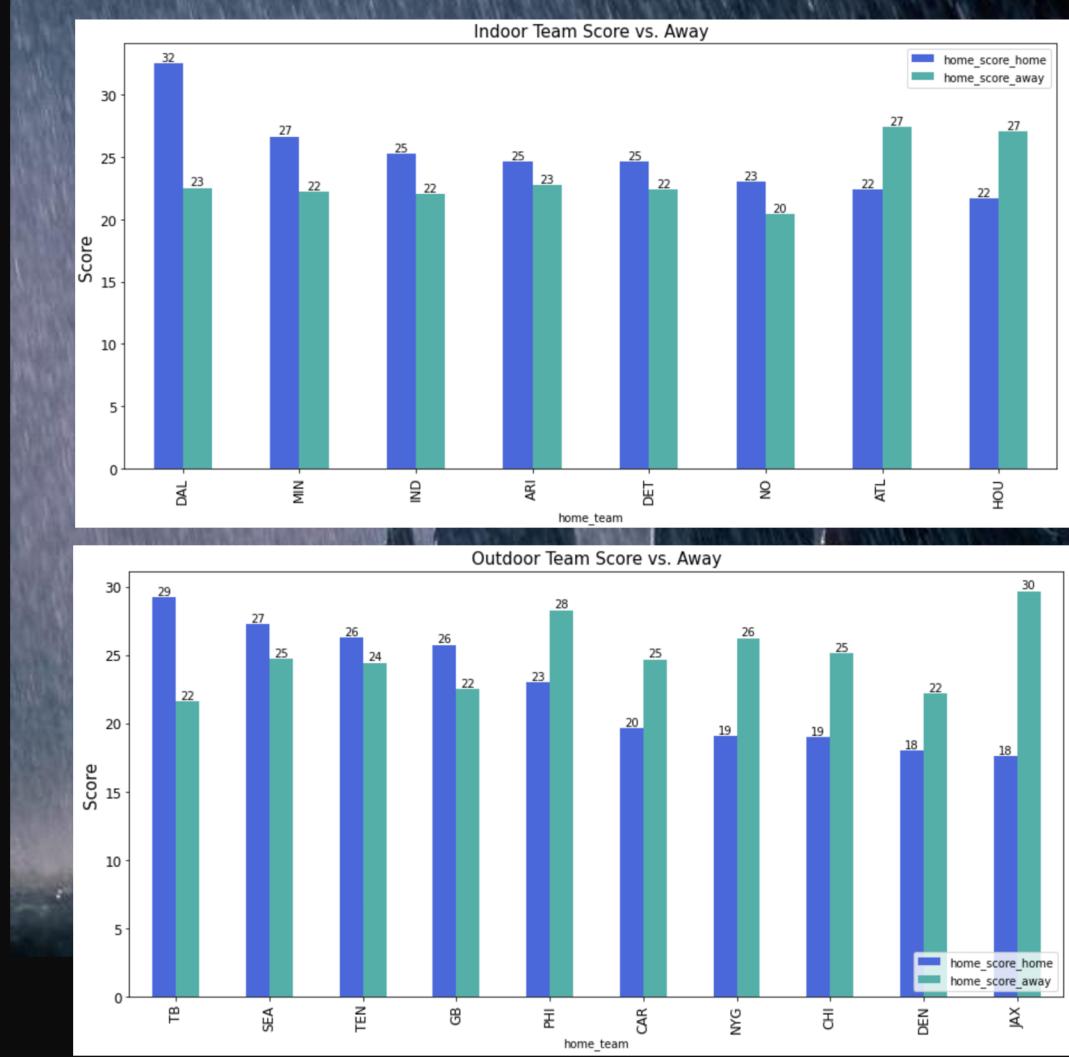
Passing yards, rushing yards and home score



## How do teams perform with Indoor/Outdoor environments?

Teams with an indoor stadium generally play better at home than away.

However, there is no data that shows that teams with an outdoor stadium perform better or worse when they are away.

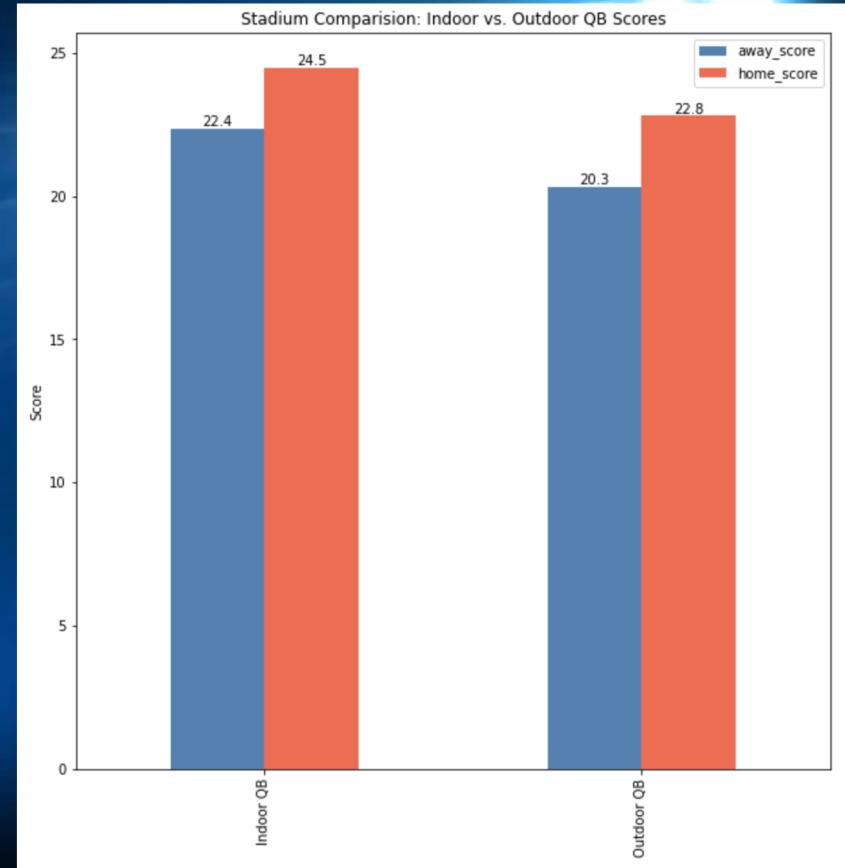
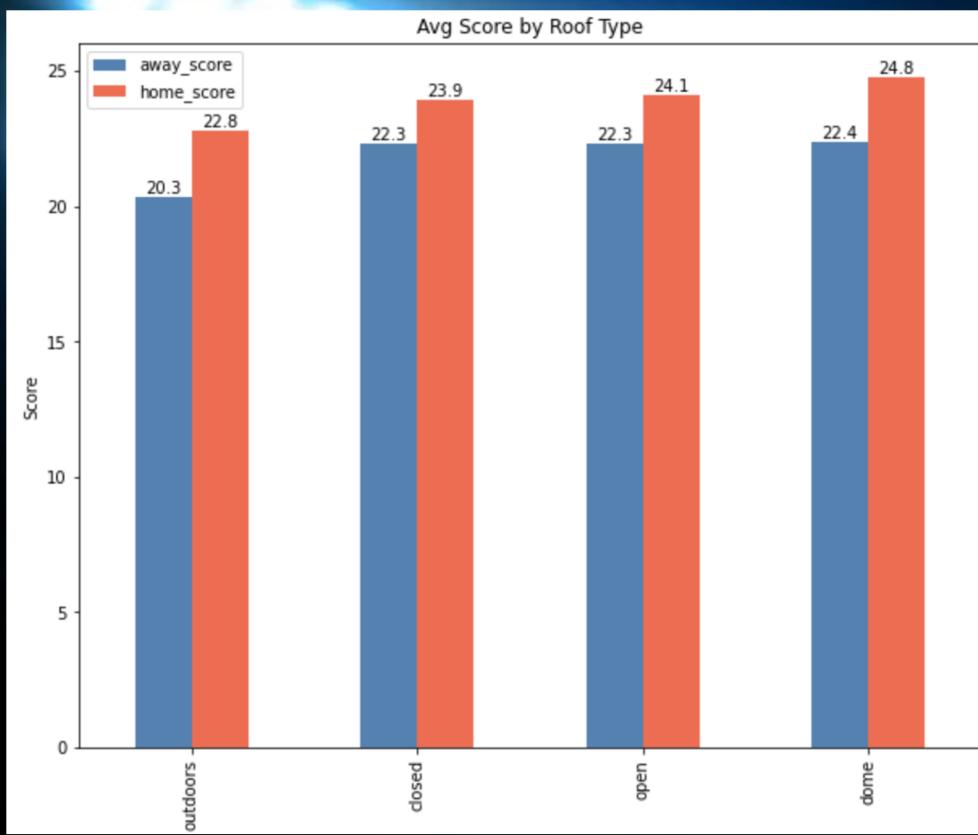


Data shows that outdoor stadiums yield less points overall than closed, open, or dome stadiums.

## Indoor vs. Outdoor Stadiums

Image: JPG file downloaded | Resolution: 5000x2750 px | www.psdgraphics.com

We can also see that when teams play outdoors, they tend to score less than playing indoors.

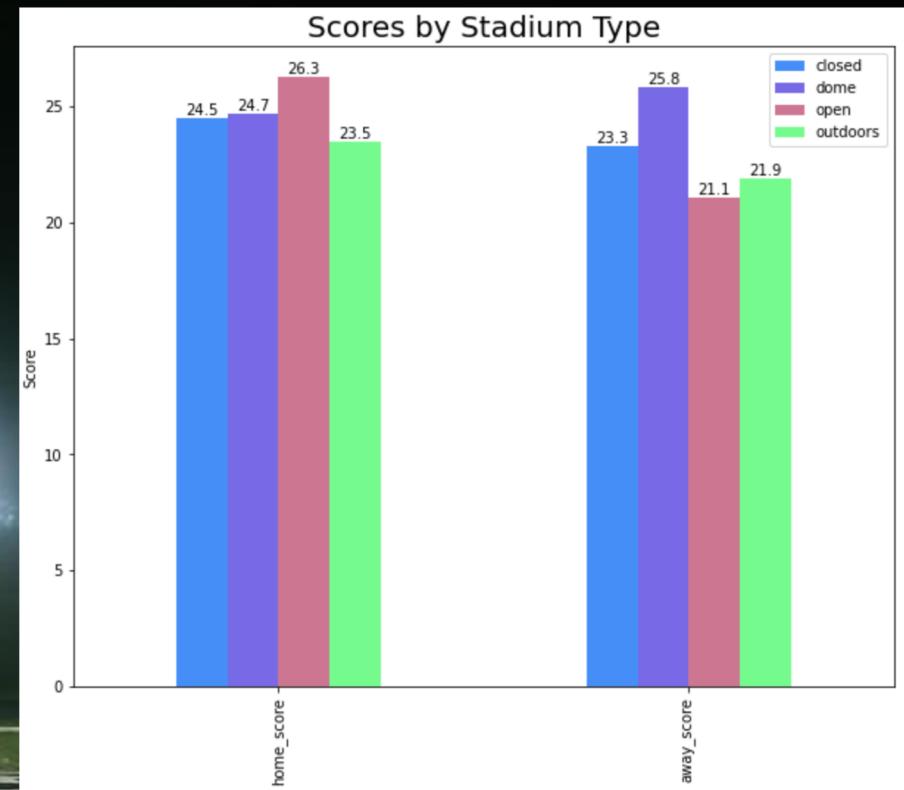
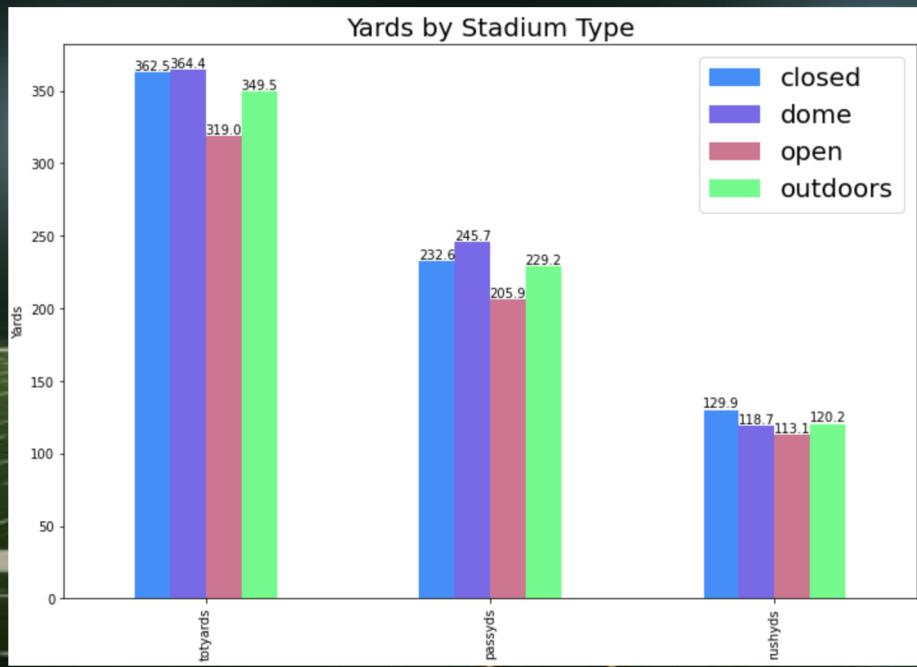


## Scores and Yards Comparison by Stadium Type

More yards does not equal more points.

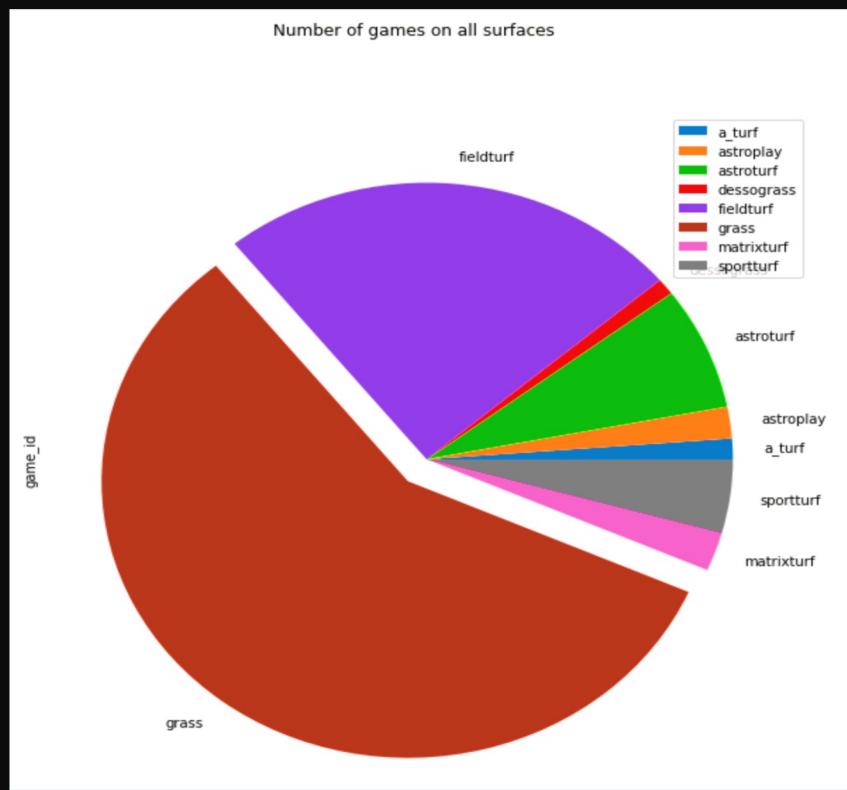
Total yards for games played in outdoor stadiums are ranked  $\frac{3}{4}$  but in terms of home score, they rank last.

As for open air stadiums, they have the fewest total yards but have the highest average home score.



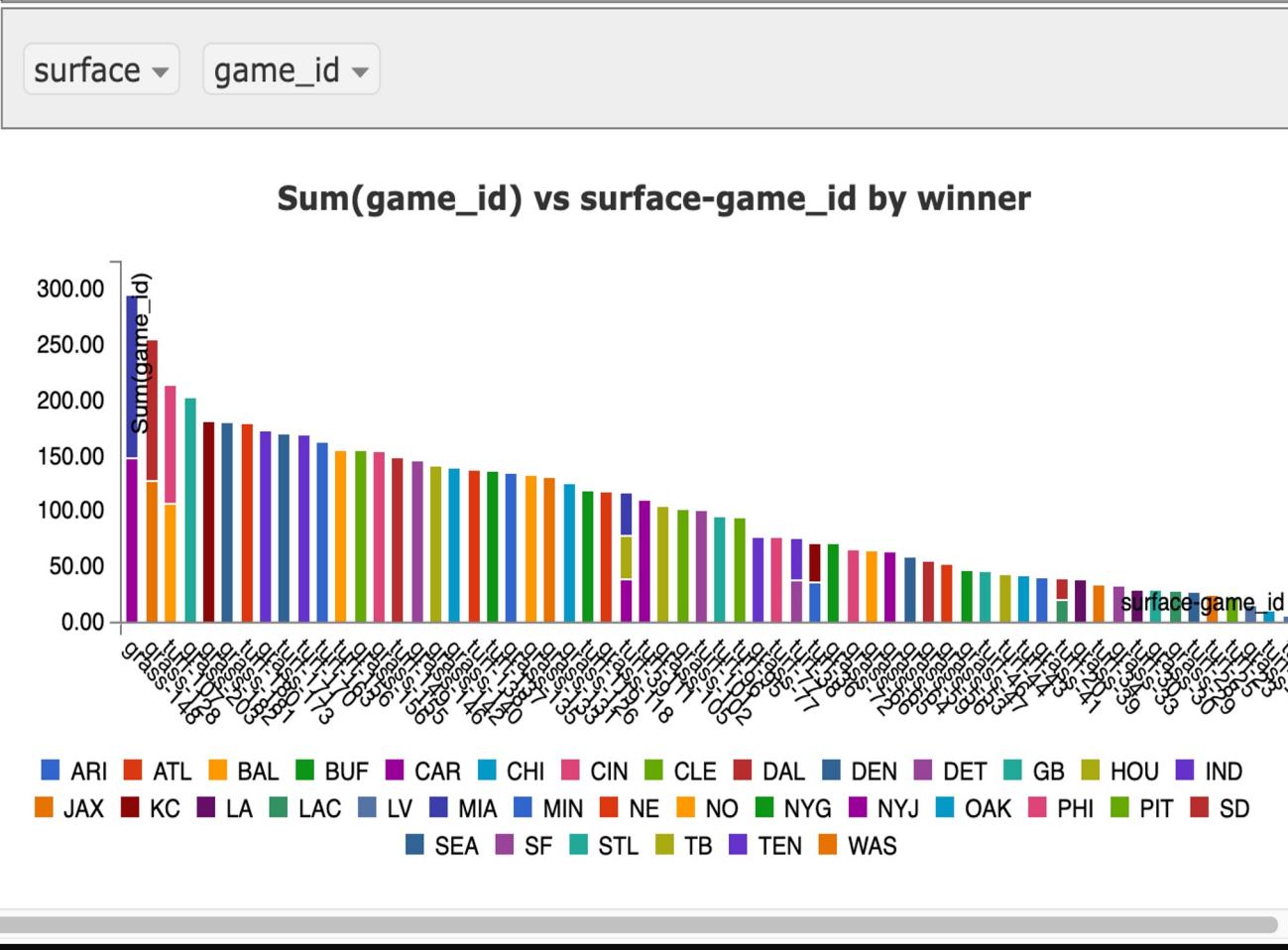
## EFFECTS OF FIELD SURFACE ON GAMES

Type of surface (grass vs turf) does not contribute significantly on the performance



# SOME SURFACE STATISTICS

Overall Team Stats



Turf Winners

Row Heatmap Sum ↑ → game\_id ▾

winner ▾ surface ▾

winner	surface	Totals
NE	turf	180.00
SEA	turf	171.00
IND	turf	170.00
MIN	turf	163.00
NO	turf	156.00
DAL	turf	149.00
ATL	turf	138.00
BUF	turf	137.00
NYG	turf	119.00
NYJ	turf	111.00
BAL	turf	107.00
CIN	turf	107.00
DET	turf	101.00
STL	turf	96.00
PIT	turf	95.00

Grass Winners

[pop\_out] Row Heatmap Sum ↑ → game\_id ▾

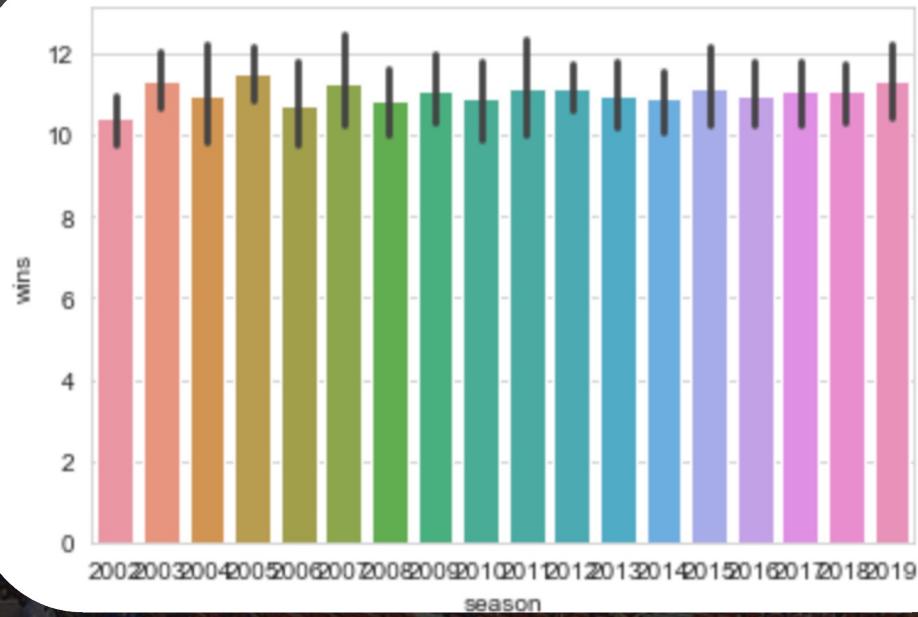
winner ▾ surface ▾

winner	surface	Totals
GB	grass	203.00
KC	grass	182.00
DEN	grass	181.00
TEN	grass	173.00
PIT	grass	156.00
PHI	grass	155.00
CAR	grass	148.00
MIA	grass	148.00
SF	grass	146.00
TB	grass	142.00
CHI	grass	140.00
ARI	grass	135.00
BAL	grass	133.00
WAS	grass	131.00
JAX	grass	128.00

## Playoff Statistics

Some Performance Stats on the teams who reached playoffs

To Assure 'Playoffs.' Approximate '12' wins during the Reg Seasons

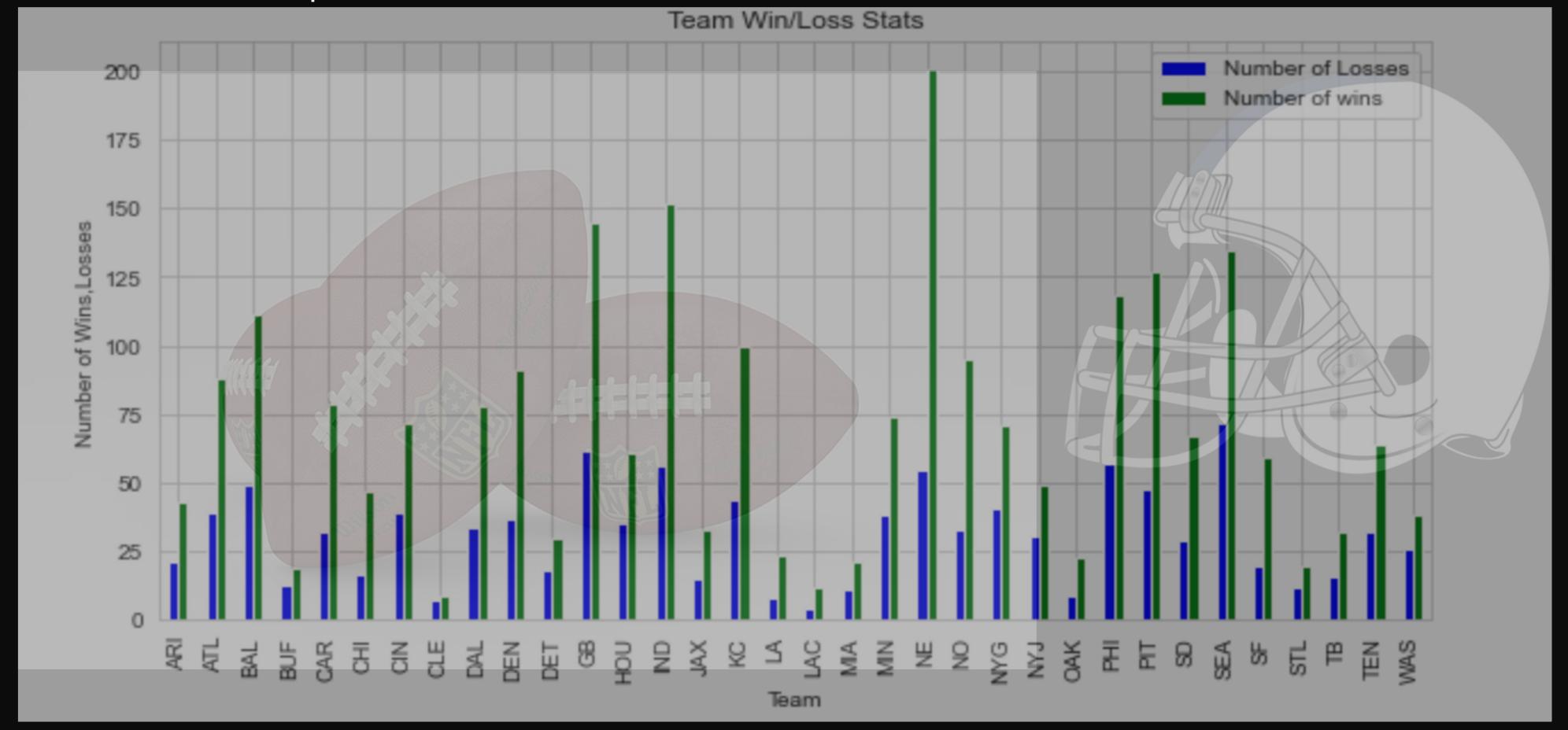


Correlation between rank and win rate(sos)

	div_rank	sos	seed
div_rank	1.000000	0.282809	0.803621
sos	0.282809	1.000000	0.291761
seed	0.803621	0.291761	1.000000

## Playoff Statistics

How did the teams performed over the Years??

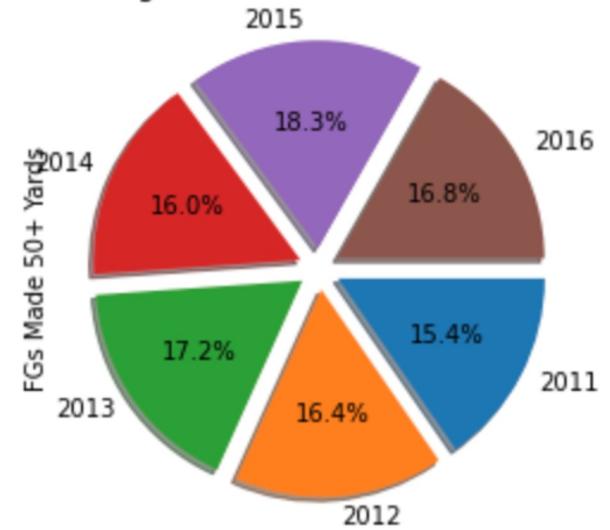


## Special Team Effects!! – Sometimes missing a Field Goal can cost you a Game!!

Who scored the longest FG?  
Per our analysis:  
Year: 2013  
Team: Denver  
Broncos  
Player: Matt Prater  
Distance: 64 yds



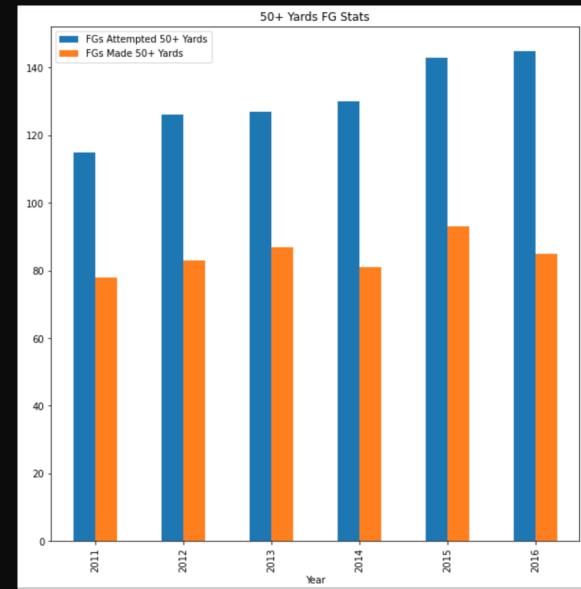
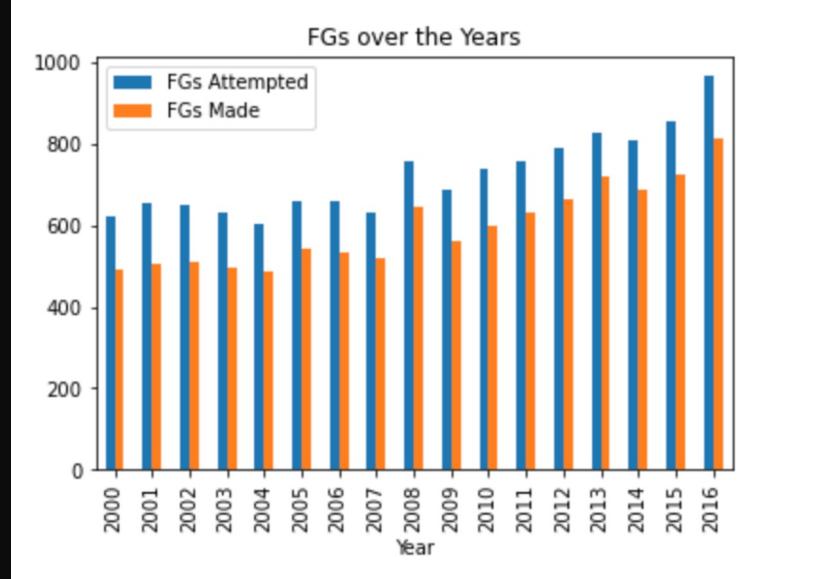
Percentage of 50+ Yard FGs over the Years



- Recent Field Goals Scoring Trends show that each team really want to try out for those extra points.

## Special Team Effects!! – Sometimes missing a Field Goal can cost you a Game!!

The trends are obviously favoring additional training and benefits that teams should invest into to sharpen the special skills



# HOW DOES WIND EFFECT THE HOME TEAM?

Wind does not seem to provide a strong correlation to scoring

Dataset normalization does not show any significant correlation

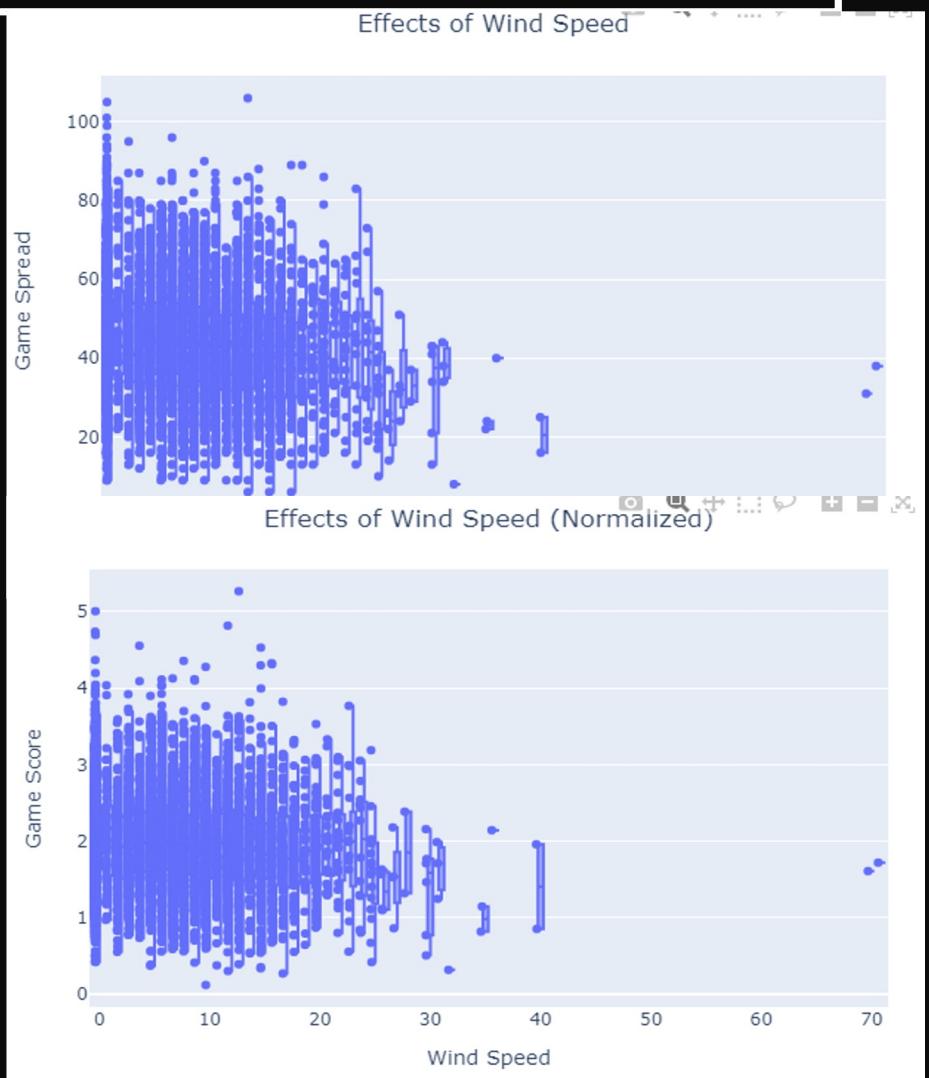
```
teams = games.home_team.unique()
asbs = pd.DataFrame()
for team in teams:
    tdf = games[(games.home_team == team) | (games.away_team == team)]
    tdf['team'] = team
    def get_score(ser):
        if ser.home_team == team:
            return ser.home_score
        else:
            return ser.away_score
    tdf['score'] = tdf.apply(get_score, axis=1) # axis=1 applies a function across each row
    tdf = tdf.groupby(['team', 'season']).score.mean().reset_index()
    asbs = asbs.append(tdf, ignore_index=True)
asbs

tgames = games.merge(asbs, left_on=['season', 'home_team'], right_on=['season', 'team'])
tgames = tgames.rename(columns={'score':'score_home_avg'})
tgames = tgames.merge(asbs, left_on=['season', 'away_team'], right_on=['season', 'team'])
tgames = tgames.rename(columns={'score':'score_away_avg'})

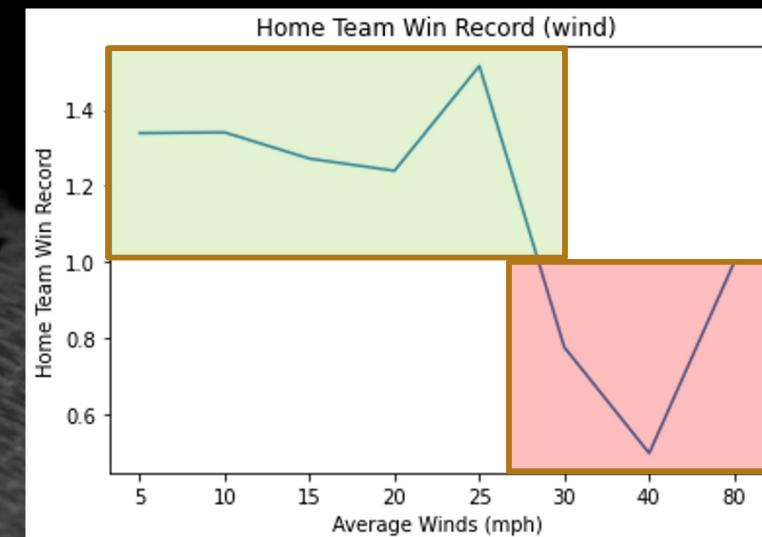
tgames['home_score_norm'] = tgames.home_score/tgames.score_home_avg
tgames['away_score_norm'] = tgames.away_score/tgames.score_away_avg

tgames['total_score_norm'] = tgames.home_score_norm + tgames.away_score_norm

tgames
```

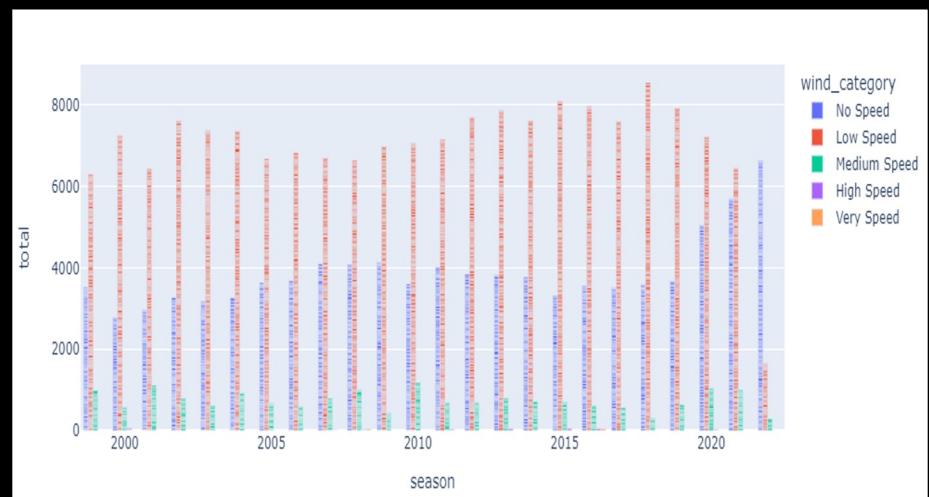


## HOW DOES WIND EFFECT THE HOME TEAM?



Running a correlation on scoring and winds, it reveals there is increased scoring in less windy conditions.

Additionally, the trend of low wind speeds to high scoring seems to be increasing in recent seasons.



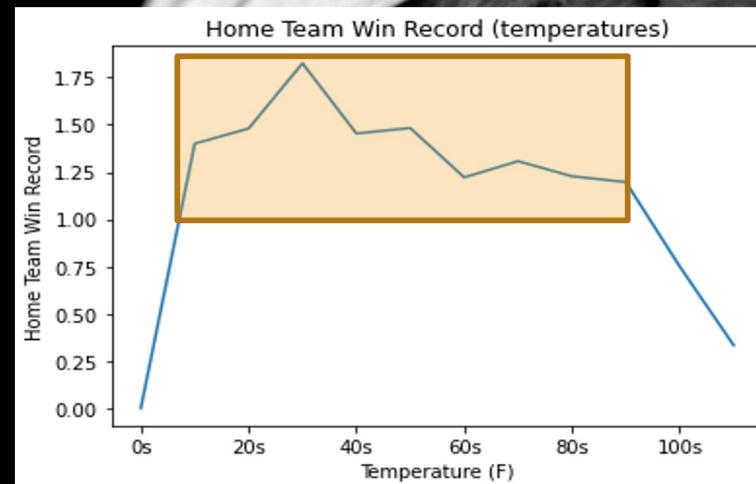
# HOW DOES TEMPERATURE EFFECT THE HOME TEAM?



Short Answer: Not often

Long Answer:

Environmental factors do not play a significant role in the spread of game points between teams. Temperature and field surface were found to only account for 2.5% of the variation in game point spread.



# PROJECTED WINS FOR CURRENT SEASON

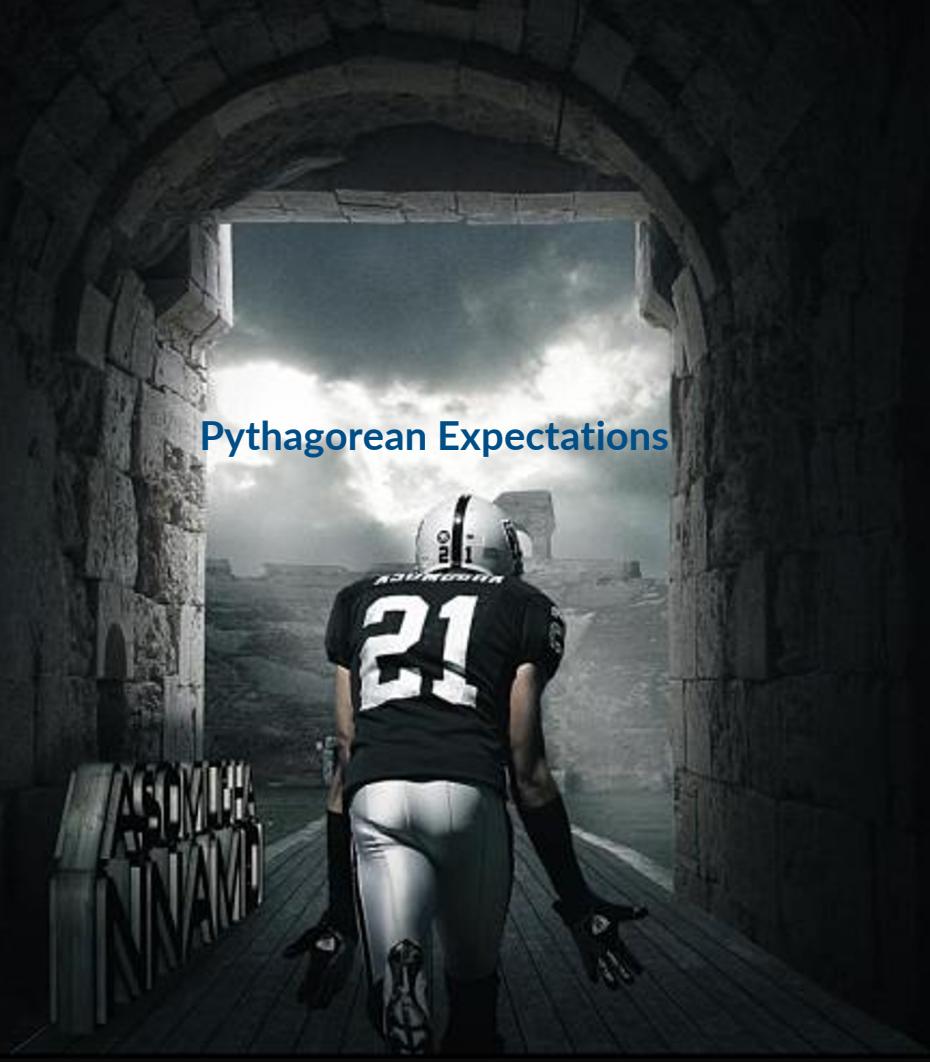
Projected Wins  
for AFC Teams

Team	Wins
BUF	12.04
KC	11.64
CIN	10.50
BAL	10.46
MIA	10.09
NYJ	9.29
TEN	8.78
NE	8.23
LAC	7.68
LV	6.97
CLE	6.91
PIT	6.39
JAX	5.87
IND	4.90
DEN	4.55
HOU	2.07

Projected Wins  
for NFC Teams

Team	Wins
PHI	13.93
MIN	12.09
DAL	12.07
SF	10.88
SEA	9.12
NYG	8.93
WAS	8.48
TB	7.98
DET	6.93
ATL	6.35
GB	6.25
CAR	5.65
ARI	5.51
NO	5.29
LA	4.24
CHI	4.12

Pythagorean Expectations



## HOW CAN THESE RESULTS BE USED?

- Indoor Stadiums tend to support teams for better wins while home and away
- Surface Type (grass vs. turf) do not yield any conclusive evidence on influencing performance
- Special Teams (Field Goals) tend to contribute more to overall scores/wins
- Regular Season Performances improves the Rankings which in turn adds to home ground advantages
- Temperature & Wind Speed do not yield significant benefits towards the performance

### **Future Enhancements to improve the Analysis Results:**

- Recent Studies conducted by NFL points out that Turf Surfaces are prone to less injuries.
  - Additional data on player injuries need to be added to enhance the analysis
- Special Team Statistics can be expanded to include other specialities (punts, kickers, fumbles etc.) to enhance the training needs
- Expand our datasets to including a modeling tool that depicts several factors to predict game outcome



# THANKS

Prepared by:

Nabeel Khan

Renjini Rajan

Amanda Norwood

## Credits

Images: <https://wallpapersafari.com/>

<https://www.nfl.com/>

END OF PRESENTATION