

Home Field Advantage in the NFL

Jason Schwartz, Praneeth Puli, Xuankai Zhu, Christine Lekishon

Introduction

If you ask any NFL fan whose fans are the loudest, they'll probably tell you its their team's fan base, and how passionate they are. They'll tell you all about how loud they were screaming when they were at the game, and how they felt like they were a part of the team. Part of what makes the NFL so popular is how involved the fans feel in the game, and how their cheering is helping their players win.

An example of this involvement is the Seattle Seahawks fan base, who refer to themselves as the "12th Man," and that they feel like they are so impactful towards the outcome of the game that it's almost like they're right there on the field. The fans create a very hostile environment for the away team, making it very difficult to communicate with their teammates and coaches.

From the factors we looked at, home field advantage in the NFL does have an effect on the outcome of the game, and the presence of fans at the game plays a significant role to the performance of the home team. However, factors such as the playing surface each team uses at their home stadium, and the time of the game (prime time or regular time) do not play a significant role when looking at home field advantage.

The Data We Are Examining

The data set we are using includes all NFL games from 2011 - 2020. The authors on github are: mrcaseb, tanho63, guga31bb, john-b-edwards. The data were collected through scraping the data off of the NFL's game database that has a history of every game played. The authors collected a wide variety of different NFL statistics, including player stats, games played (the data set we're using), play-by-play summaries and outcomes of each play for specific games, and much more. They collected all of this information to build models on trends and consistencies in the NFL. They provided each season as its own data set online, so we downloaded the seasons we wanted, and used code to combine each season into a giant database of games.

Here is a link to the data online: <https://github.com/nflverse/nflverse-pbp/blob/master/schedules>

Some of the key variables in this data set that we will be using in our inferences and graphs are game time (prime time, regular time), home team and away team, result of game, and surface of playing field. These variables all relate to how much of an advantage the home team supposedly gets against the visiting team. Each row of data in the data set represents a game that was played.

Our Models

The graphs we will be looking at relate to the overall home team's performance throughout each season we have downloaded, and any trends over each season. The inferences we will be performing include the Pre-Covid seasons vs Covid season, the away team playing on the same surface they usually play on vs the away team playing on a different type of surface from what they're used to, and lastly an inference looking at prime time games vs regular games.

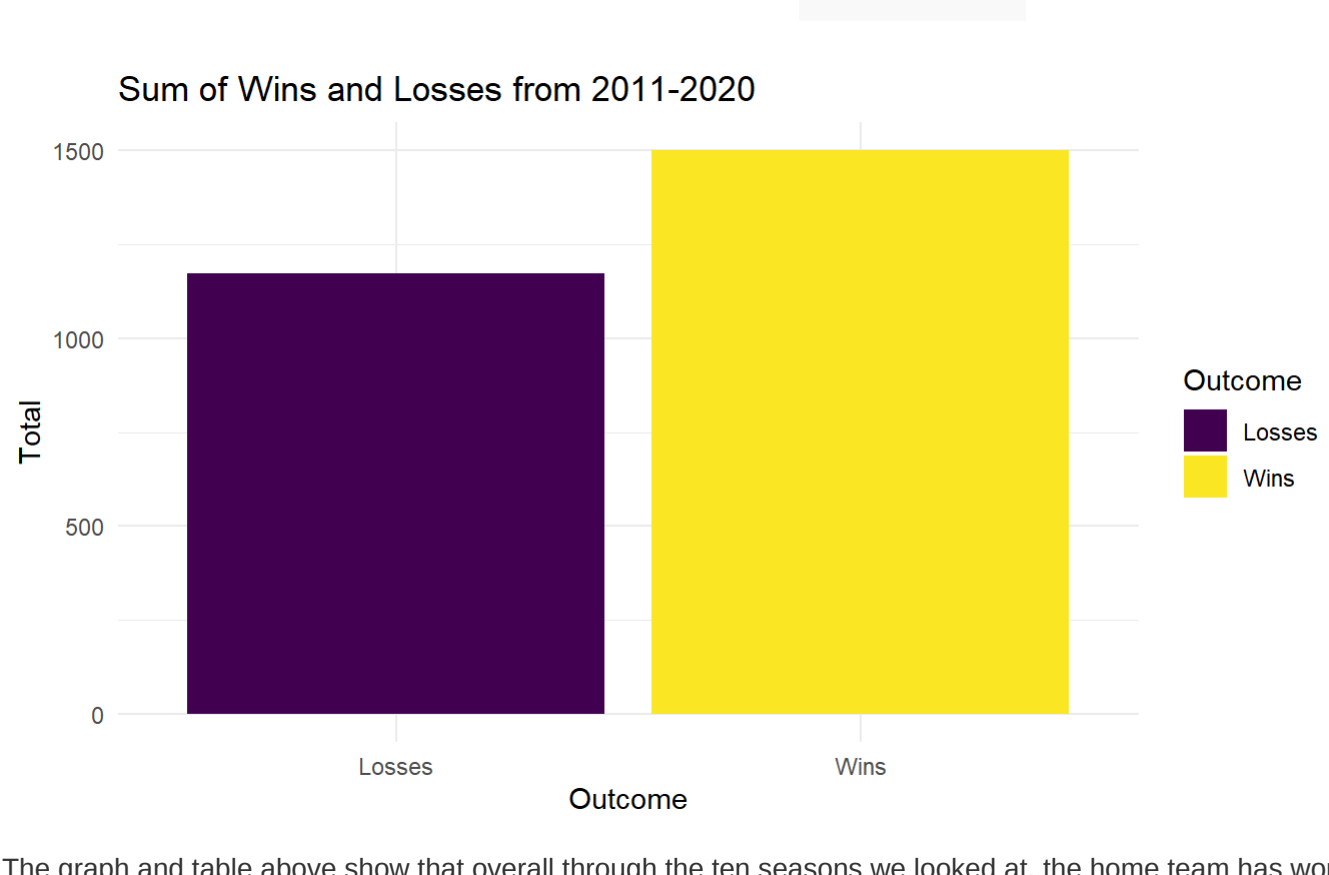
For each inference, we decided to use a one-sided z-test for difference in probabilities in order to compare the two probabilities each inference is questioning, and to try and find significant evidence that there is a difference between the two probabilities each inference is comparing.

We believe that the cumulative evidence (or insignificant evidence) of each graph and test will help us draw a conclusion and insightful analysis of the impact of home field advantage in the NFL.

The Following Models do not take into Account the Following Conditions:

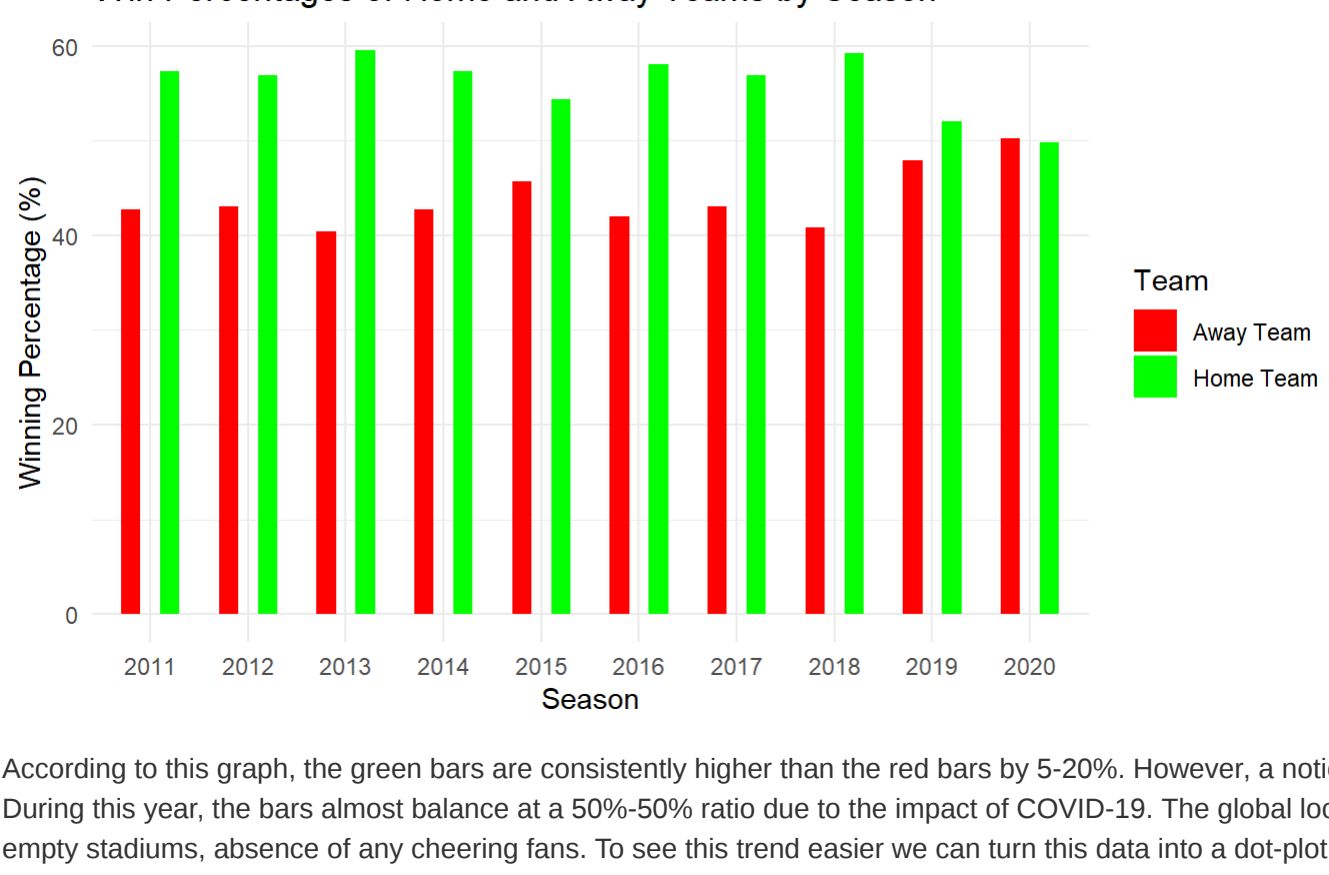
- A measurement of how "good" the home team is, i.e. their record could be undefeated, or maybe they've lost every game so far that season. We considered the probability of the home team winning the same for every game played in the data set to fit a binomial model (A true NFL fan would say this is actually a plausible assumption because all NFL players are world-class athletes, and you can never rule a team out, no matter how bad their record is, and that you can't 100% predict a team will win, no matter how good their record is. You might hear this being referenced as the phrase, "Any given Sunday!")
- We consider each game played, or trial, in the data set to be independent of each other
- If there is a noticeable impact on how playing on certain surfaces varied through different seasons
- Each individual teams home record (our model pools them all together)
- How each team might get better or worse throughout each season
- Whether or not the game is a regular season game or playoff game
- Almost every game ends in either a win or a loss, but in some rare cases, a game can end in a tie. Since the amount of ties is so few in our data (there were 9 games out of the 2672 that were played in the 2011-2020 NFL seasons that resulted in a tie), we considered them as "non-wins" to fit our binomial model of two possible outcomes of each game, and added the ties into the losses category for the home team
- We make all of these assumptions in order to focus our tests on the variables of question

Graphs

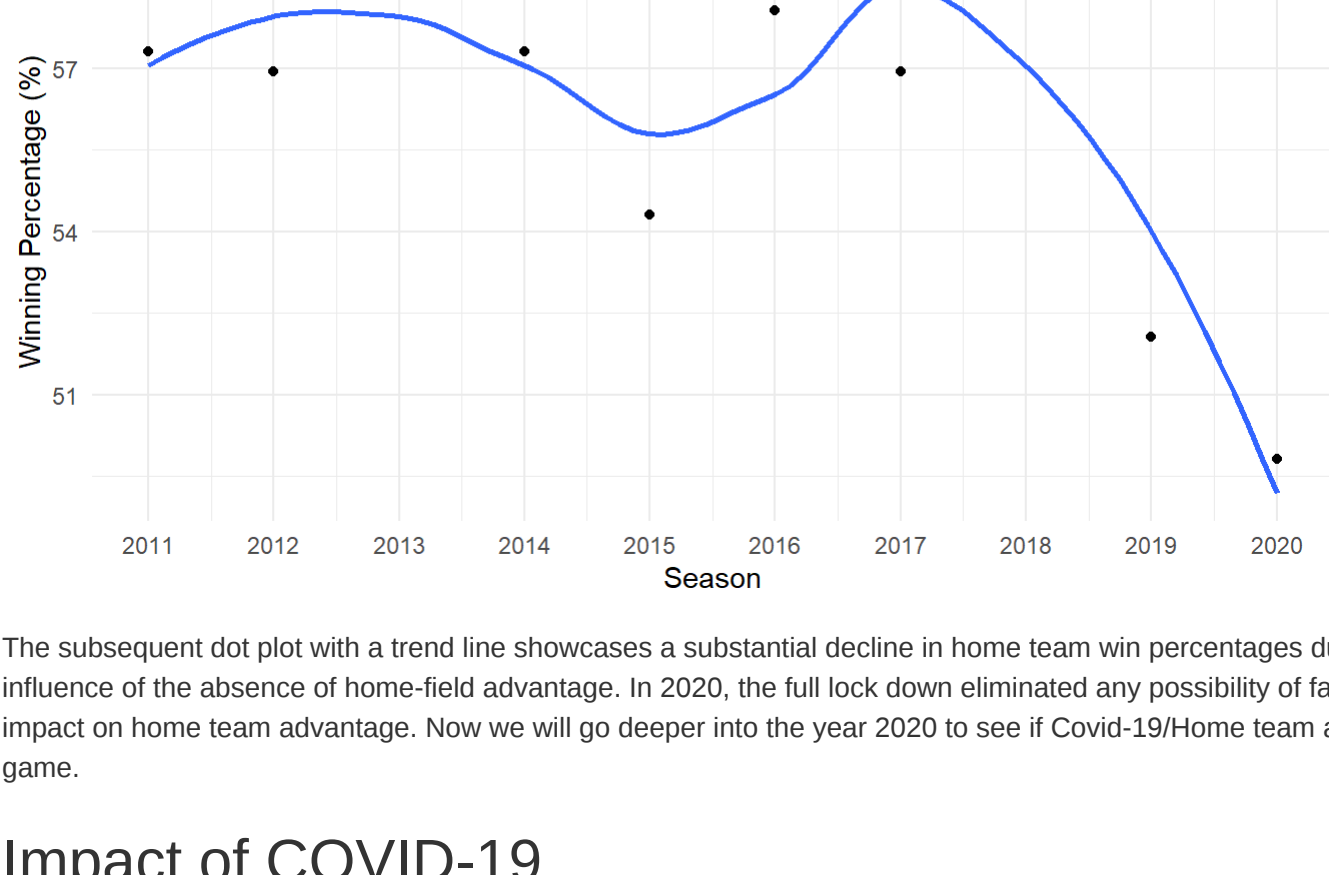


The graph and table above show that overall through the ten seasons we looked at, the home team has won many more games than the away team.

This next graph illustrates the wins and losses of home teams over a decade, from 2011 to 2020. The red bars denote home team losses, while the green bars represent their wins. X-axis has the years 2011-2020 and y-axis has the winning percentage. Clearly, the dominance of green over red shows that home teams secure more wins than losses, signaling the existence of a home-field advantage. This data spans years, so it encompasses the various team changes. To see this trend at a deeper level we can plot the win percentage of home and away teams for each year.



According to this graph, the green bars are consistently higher than the red bars by 5-20%. However, a noticeable shift emerges in the year 2020. During this year, the bars almost balance at a 50%-50% ratio due to the impact of COVID-19. The global lock down prompted teams to play in empty stadiums, absence of any cheering fans. To see this trend easier we can turn this data into a dot-plot with a trend line.



The subsequent dot plot with a trend line showcases a substantial decline in home team win percentages during this year, emphasizing the influence of the absence of home-field advantage. In 2020, the full lock down eliminated any possibility of fan presence, further intensifying the impact on home team advantage. Now we will go deeper into the year 2020 to see if Covid-19/Home team advantage truly has an affect on the game.

Impact of COVID-19

COVID-19 led to prolonged and unprecedented disruptions across various industries, and the NFL was no exception. As a global pandemic, it uniformly affected teams across the US, potentially diminishing the traditional home field advantage in sports. Typically, this advantage is attributed to factors such as familiar playing conditions, supportive home crowds, and the absence of travel fatigue.

However, the 2020 season's unique circumstances, primarily driven by the COVID-19 pandemic, potentially altered these dynamics. For instance, with stadiums either empty or at reduced capacity, the psychological boost of a home crowd, which often energizes the home team and can intimidate visiting teams, was largely missing. This change might have leveled the playing field between home and away teams, reducing the usual home advantage.

Furthermore, the pandemic-induced changes in daily routines, training schedules, and health protocols could have impacted players' performances and team dynamics. Players who tested positive or those exposed would need to isolate, hence missing future games. If a star player misses a game, then the team is likely to lose. The stress and uncertainty surrounding the pandemic, along with the strict health and safety protocols, might have affected players' mental and physical preparedness for games.

Inference

Is the winning percent of the home team more in the 2011-2019 seasons than the home team winning percent in the 2020 season?

Season	Games	Home Wins	Home Win Percent
2011-2019	2403	1366	0.5684561
2020	269	134	0.4981413

From our sample, we found that the home team wins 56.8% of the time in the 2011-2019 seasons, and the home team wins 49.8% of the time in the 2020 seasons. Is this difference in proportions significant?

For our model, we will use a binomial model, with p_H as the probability that the home team wins in the 2020 season, and the number of observations equal to 269 games. We will use p_A as the probability that the home team wins in the 2011-2019 seasons with the number of observations equal to 2403 games in this period.

Our null hypothesis for this inference is that the probability of the home team winning pre-Covid was the same as the probability of the home team winning during the Covid season, or that p_H equals p_A . Our alternative hypothesis for this inference is that the probability of the home team winning was greater pre-Covid than what it was during the Covid season, or that p_H is greater p_A .

Interpretation

Based on this sample, there is strong evidence to conclude that the home team in the NFL won more often in the 2011-2019 seasons than in the 2020 season ($p = 0.0138$, $z = 2.2039$, 1-sided z-test for difference in proportions).

Playing Surfaces: Does Grass and Turf Really Affect the Outcome of the Game?

Now, we are going to look at the significance of the different playing surfaces in the NFL. In the NFL, each team gets a choice of two options regarding how they want to build their field in their stadium: Grass, and Artificial Turf. As of today, the NFL is almost split in half with the number of teams that use grass and the number of teams that use turf. If you are unfamiliar with the difference between these two types of surfaces, here is a brief description:

Artificial turf is a type of synthetic grass, that is made to resemble natural grass. Along with the artificial grass, turf also has thousands of small, rubber pellets that squeeze in between each blade of turf. These pellets are designed to protect the players whenever they fall to the ground, and absorb some of the force as they hit the turf.

Grass fields are just as you would think: natural grass. Grass fields cost much more money and time to maintain than a turf field does, which is why many teams choose artificial grass instead.

Football teams and players across the NFL have complained recently about the diversity of playing surfaces, and have called for the NFL to mandate what type of surface each stadium has in order to have a uniform playing surface they can get accustomed to.

Many teams believe that the type of playing surface makes a huge deal about the outcome of the game. In this past Super Bowl, SB LVIII, many people believed the Philadelphia Eagles weren't as good as they normally were because of the artificial turf they were playing on, and that they repeatedly slipped while playing, which affected their performance and the outcome of the game. Another example of how serious each team takes the playing surface into account is the Wisconsin Badgers football team right here in Madison. At the start of each week, they find out what type of playing surface they will be playing on for their next upcoming game. If it's a grass field, then they will practice all week on the outdoor, grass practice field just north of Camp Randall, whereas if they are playing on turf, they will practice all week inside Camp Randall, which has a turf field, instead.

Inference

Does the home team win more often when the away team plays on a different type of surface at their respective home stadium?

Same Playing Surface	Games	Home Wins	Home Win Percent
FALSE	1387	791	0.5702956
TRUE	1285	709	0.5517510

From our sample, we found that the home team wins 57.03% of the time when the away team plays on a different type of playing surface at their home stadium, and the home team wins 55.18% of the time when the away team plays on the same type of playing surface at their home stadium. Is this difference in proportions significant?

To build our inference, we will be using a binomial model, with p_H as the probability that the home team wins when the visiting team plays on a different type of surface at their respective stadium, and the number of observations equal to 1387 games. We will use p_A as the probability that the home team wins when the visiting team plays on the same type of surface as the home team at their respective stadium, with the number of observations equal to 1285 games for this condition.

Our null hypothesis for this inference is that the probability of the home team winning when the away team plays on a different type of surface at their home stadium is the same as the probability of the home team winning when the away team plays on the same type of surface at their home stadium, or that p_H equals p_A . Our alternative hypothesis is that the probability of the home team winning when the away team plays on a different type of surface at their home stadium is greater than when the away team plays on the same style of surface, or that p_H is greater than p_A .

Interpretation

Based on this sample, there is not enough evidence to conclude that the home team in the NFL wins more often when the away team plays on a different type of playing surface at their respective stadium and when the away team plays on the same style of playing surface at their respective stadium ($p = 0.1672$, $z = 0.9652$, 1-sided z-test for difference in proportions).

Prime Time Games vs. Regular Time Games

Our last variable that we will be looking at is the impact of NFL games during a "prime time" slot. A prime time game in the NFL is broadcasted nationally, typically played at night under the lights, and is the only game that is being played in the NFL at that time. Every other game for that week is a "regular time" game, where each game is usually broadcasted only to the regions of the country where the home and away teams are based, played during the middle of the day, and multiple games are also being played at that same time.

Prime time games are usually "rowdier" and much louder than a regular time game, since the game always has a sold out crowd, at night, and has national attention from the media. This high energy crowd could potentially give the home team the motivational boost they need to win the game.

An example of a rowdier crowd at night is the "White Out" night game hosted by Penn State University each year, where over 100,000 people come to the game, and all dress in white to create a surreal atmosphere. An example from the NFL is Arrowhead Stadium, home of the Chiefs. The Chiefs have hosted the AFC Championship game, a prime time game, multiple times recently, and have some of the loudest fans, which many believe have helped them get to the Super Bowl twice in recent memory.

Inference

Does the home team win more often when the game is played during prime time compared to regular time?

Game Time	Games	Home Wins	Home Win Percent
Regular	2023	1126	0.5565991
Prime Time	649	374	0.5762712

This data frame shows that the home team wins 55.66% of the time when the game is played during regular time, and that the home team wins 57.63% of the time when the game is played during prime time. Is this difference in win percentages significant?

To conduct inference regarding the impact of prime time games, we will use a binomial model, and use p_P as the probability that the home team wins when the game is being played during a prime time slot, with the number of observations equal to 649 games. We will also set p_R as the probability that the home team wins when the game is being played during a regular time slot, with a total number of 2023 games.

Our null hypothesis is that these two proportions are equal to each other, or that p_P equals p_R . Our alternative hypothesis is that the probability of the home team winning during a prime time game is greater than the probability of the home team winning during a regular time game, or that p_P is greater than p_R .

Interpretation

Based on this sample, there is not significant evidence that the home teams wins more frequently during a prime time game then during a regular time game ($p = 0.1898$, $z = 0.8788$, 1-sided z-test for difference in proportions).

Non-Statistical Interpretations/Conclusions

Covid Season - 2011-2019 Seasons vs 2020 Season

We found that there was enough statistical evidence that the home team won more often during the 2011-2019 NFL seasons than during the 2020 NFL season.

This evidence can also be supported and argued with non-statistical calculations. When you're watching an NFL game, the "momentum" of the game matters towards the attitude and confidence of the home team. For example, scoring a touchdown or getting an interception already gets a team fired up and energized, swinging the "momentum" of the game in their favor. Adding in tens of thousands of fans to the equation, who are ecstatic at what the home team just did, increases the adrenaline, and even swagger of the home team astronomically. On top of the fans cheering for a positive play by the home team, negative plays by the away team can have a similar effect to the momentum of the game. Imagine how the away team's Quarterback feels after he just threw an interception, or fumbled the ball, and thousands of people are laughing and making fun of him. The frustration and lowered confidence can make all the difference in how the Quarterback will play the rest of the game. Now take all of that away, which happened in the 2020 NFL season, and the game feels much more emotionally balanced.

Playing Surfaces - Grass vs. Turf

Earlier in our report, we found that there was not enough evidence that the home team wins more often when the away team plays on a different style of surface at their home stadium than when the away team plays on the same style of surface as the home team.

This conclusion makes sense because when looking at the amount of preparation an NFL team goes through leading up to their game that week, each team can find out what types of employees and analysts finding small details about their opponent that will help them win. Anybody with access to a search engine can hunt out what type of playing surface (grass or turf) an NFL team's home stadium uses, and include adequate preparation leading up to the game to handle the conditions they will face. The Badgers, as mentioned earlier, utilize two separate practice fields that are different types of surfaces, which would cancel out any effect of playing on an unfamiliar type of surface.

Prime Time vs. Regular Time

As mentioned earlier in our report, we found that there is not enough evidence to conclude that the home team wins more often during prime time games than during regular time games.

A prime time game brings excitement to the atmosphere of the game - under the lights, national television, high stakes, and much more. We brought up during this inference how we believed this could lead to the fans getting more involved in the game, and the home players feeling more motivated while they're playing. With such an exciting environment, a similar claim could be said about the away team. With everyone watching you in the entire country, there's no room for error by either team, home or away. They know that their entire fan base is watching back home, and every fan that doesn't live in the area of the team that can't normally watch the game since on regular time, they're only broadcasted regionally, is watching them play now. For example, a Cincinnati Bengals fan living in Madison can't watch the Bengals play when they're on regular time, since their games aren't broadcasted to Madison (unless playing a team nearby, such as the Packers, Bears, or Vikings). When the Bengals are on prime time, no matter who they are playing or if they're the away team, their fans from everyone in the country can tune in and watch them play.

Short-Comings and Future Work

Our analysis looked at the NFL as a whole, and the overall view of the home team in the NFL. This could lead to a false assumption that all teams are equal in "skill" and as good as each other. We think that a way our tests could be improved could be by looking at a specific, single team in a specific season instead of looking at all teams throughout ten seasons (2011-2020) in order to eliminate the criticism of differing skill levels of the many NFL teams throughout the decade we examined.

From our three inferences conducted, only one of them provided significant evidence of a difference in proportion of the home team winning, that being the 2011-2019 seasons vs the 2020 season. We found evidence that in the 2020 NFL season, where all games had no fans present due to COVID, the home team won less often than in the 2011-2019 NFL seasons. Because of this finding, we believe that this could be used as a basis for future work revolving around the influence of spectators at NFL contests, and that maybe home field advantage has more to do with the atmosphere the home team's fan base creates for the home team.

A way to further improve the validity of this inference on the impact of fans at the games could be to collect the 2021 and 2022 NFL seasons (or wait a few more years to collect more data on NFL games), where fans returned to the stands and were once again present at the game - screaming, yelling, and doing anything they can to help their team win. Does the home team win percentage jump back up to what it was pre-Covid, or did the home team struggles continue even after fans were allowed back at the game?

Another question that could be studied using this data set would be looking at every teams record throughout the 2011-2020 NFL seasons, and seeing if each team improves when they finished the previous season with a negative record (when they have more losses than wins). The NFL awards draft picks at the end of the season based on how each team finished - i.e., the team with the worst record receives the first pick in each round of the draft, and the team that wins the Super Bowl gets the last pick in every round. A possible inference could be if this draft positioning truly helps a team that didn't do so well that year. New data that could further investigate this question could include the NFL Draft selections that correspond to each season in the current data set, and seeing if certain draft selections have a bigger impact in the change in record the following season. For example, the worst team each year that also gets the #1 overall pick typically selects a quarterback; does selecting a quarterback in the first round lead to a positive record?

Conclusion

Based on our analysis, there is not enough evidence to conclude that factors like playing surfaces and prime-time slots significantly influence the chances of a home team winning in the NFL, with our calculated p-values mentioned earlier as evidence. Our graphs indicated that home teams generally had a higher overall winning percentage compared to losing percentage. Additionally, the winning percentage for home teams consistently exceeded that of away teams for each year from 2011 to 2020. With this pattern we thought there is a possible advantage for home teams related to playing surface and prime-time slots.

The strong evidence from the z-test, indicating that home teams in the NFL won more frequently during the 2011-2019 seasons compared to the 2020 season, further reinforced our intuition, particularly because of the reduced fan attendance in 2020 due to Covid-19. However, a deeper investigation into the specific factors of playing surfaces and prime-time game slots revealed that there is not enough evidence that these elements significantly affect the home team's likelihood of winning. This outcome suggests that while the home field advantage could be a real phenomenon, as evidenced by the overall higher win rates for home teams, it is not substantially influenced by the type of playing surface or the game type (prime time vs. regular time).