The title of the project is “Fan Impact on NBA Games”. This is somewhat consistent with the topic of the presentation as it does go over what impact the fans could have on NBA home winning percentages. I do think a more apt title be What Impact do Fans have on Home Teams Winning, or something to that effect as that is the real analysis that occurs in the rest of the presentation. This of course still covers fan impact but is more specific in what the impact being analyzed in the research question when compared to the original title.

The research question asks if fan attendance effects the results of NBA games. The website being scraped is Baskteball-Reference. The motivation for this project is to find out how important fan attendance is for determining the win rate of home teams so that those involved in watching and analyzing basketball can see how much of an impact fans actually have on games. The presenter wants to be able to better isolate the impact of fan attendance on games as this season has given us the best ability to analyze this due to most of the games taking place without any fans present. Some games have limited fans as well, so that can also be helpful for determining fan impact. The presenter is comparing this data from the current season to the last 5 normal seasons pre-covid since they had no fan limitations and can be used as a baseline to track what the normal home winning percentage is. The only outside information brought in was to supplement the presentation itself. This extra information did not really provide data to supplement the research question. This still added to the project however as it added flavor to the presentation and emphasized the purpose of the research question.

I think the presentation is fine overall. One thing that I think would improve the presentation is to do multiple takes of it. There are times where the presenter had some trouble with accidently advancing slides so a few extra takes might have allowed for a better flowing presentation. Another thing I would have liked to see more of is R-code being shown or some part of the RStudio process being mentioned in the presentation itself as that would better allow us to know how you found the data. The graphs in the presentation are also a bit confusing, as I know that the 2015-2019 data is supposed to be together on the first bar, but I am not sure what the bar before the last one is supposed to represent. If it just represented the previous season (2020) would it not be better to just remove it entirely as the line clogs up the data table. I do like the conclusion as it puts the information more into context, even if the conclusion is not decisive as it is in this case (not that it can be on something this difficult to test).