

Harley Ryan

December 7, 2024

DS 4002

Predicting Nationally Televised College Football Games Based on Team Quality and Fanbase Size

In recent years, college football has become an increasingly profitable program for television networks. With the rise of streaming, live sports are one of the few areas where traditional television networks have an advantage, and the size and popularity of college football makes it a key target in the race for networks to snap up sports programming. The competitive nature of CFB programming makes it critical for television networks to carefully schedule their games to maximize viewership. The goal of this case study is for you to use a probit regression model to predict what types of games are most likely to be featured on national flagship networks (ESPN, ABC, CBS, NBC, and FOX), taking inputs both from the collegefootballdata.com API and from Reddit's college football subreddit. The collegefootballdata.com API will provide you with information about the quality of teams at the time of any given game, while the Reddit data will give you information about the size of the fan bases involved in each game. Through this model, you will be able to see the effects that fanbase size and team quality have on major networks deciding what games to show, and you will be able to predict how networks will act in the future. This work is very important to the television industry, as the results of this case study will help inform advertisers and networks how teams are selected for major games, and can provide insight alongside viewership statistics as to what networks might be able to tweak to maximize viewership.

<https://github.com/HarleyRyan3/DS4002-Case-Study>