

PROBLEM SET № 5

Grant Naberhaus, University of Oklahoma

02/20/2020

Web Scraping

The table I scrapped was from Wikipedia. The table lists all NBA teams that had a win percentage over 0.750 on the season, and the playoff outcomes for that season for the specified team. Although the data is not what I need for creating a model for current NBA games, it still provides insights about the NBA, especially the likelihood of a team's performance in the playoffs based on exceptional regular season performance.

NBA API

The API I used was ballr, which included imports of dplyr ($\geq 0.7.0$), janitor, lubridate, magrittr, rlang, rvest, and xml2. This API uses functions to pull data from basketball-reference.com for NBA games. The table I created provided the outcomes of all 82 regular season games for a specified team and season. Additionally, the days of rest the team had was included in this table. Analyzing how a team performs based on the days rest they received could offer insight into how that team will perform, especially on back to back games or three games within four nights. Additionally, I found the API nbastatsr, but I am still tampering with it to find data that would be relevant to what I want to do. Hopefully, between these two APIs, I should have plenty of NBA data to use.