

Assignment 3 - EDA

Nick Sun

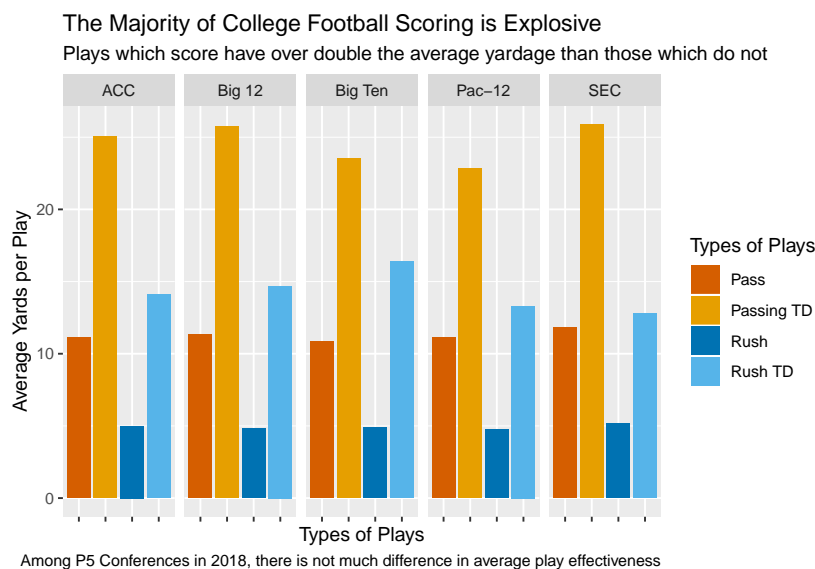
May 26, 2019

College Football Analysis - Introduction

The data I used is kept on a public API and contains everything from game scores, player stats, coaches salaries, talent rankings, to AP rankings (complete documentation can be found on <https://api.collegefootballdata.com/api/docs/>). I retrieved my data using the `jsonlite` R package and cleaned it using `dplyr`. I have also written several functions to hit certain API endpoints and have attached them as a .R file which I `source()` in this document. This API is maintained by Reddit user `/u/BlueSCar` who also serves as moderator for the `/r/CFBAnalysis` subreddit.

Play Call Effectiveness between Conferences

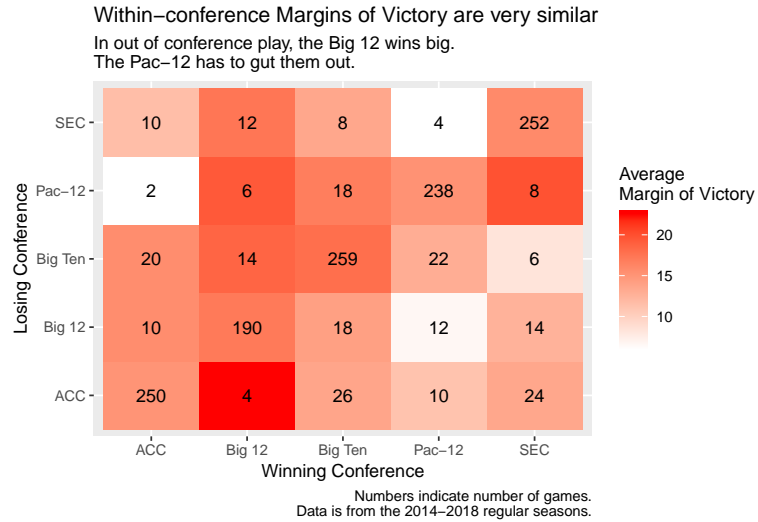
In my exploratory investigation, it quickly became apparent that **all** the conferences have roughly the same play-calling style - approximately 50% passing and 50% rushing. What was more interesting was the effectiveness of different types of plays among the different conferences.



Surprisingly, there is *little to no* practical difference between any of the conferences. Also in each conference *scoring plays have more than twice the average yardage*. I think this speaks to the immaturity of all college defenses to read offenses which results in most of the scoring in college football.

Conference competitiveness

Each of the Power 5 conferences claims to be balanced and competitive from top to bottom. I explored this using average margin of victory for matchups between and within the conferences.



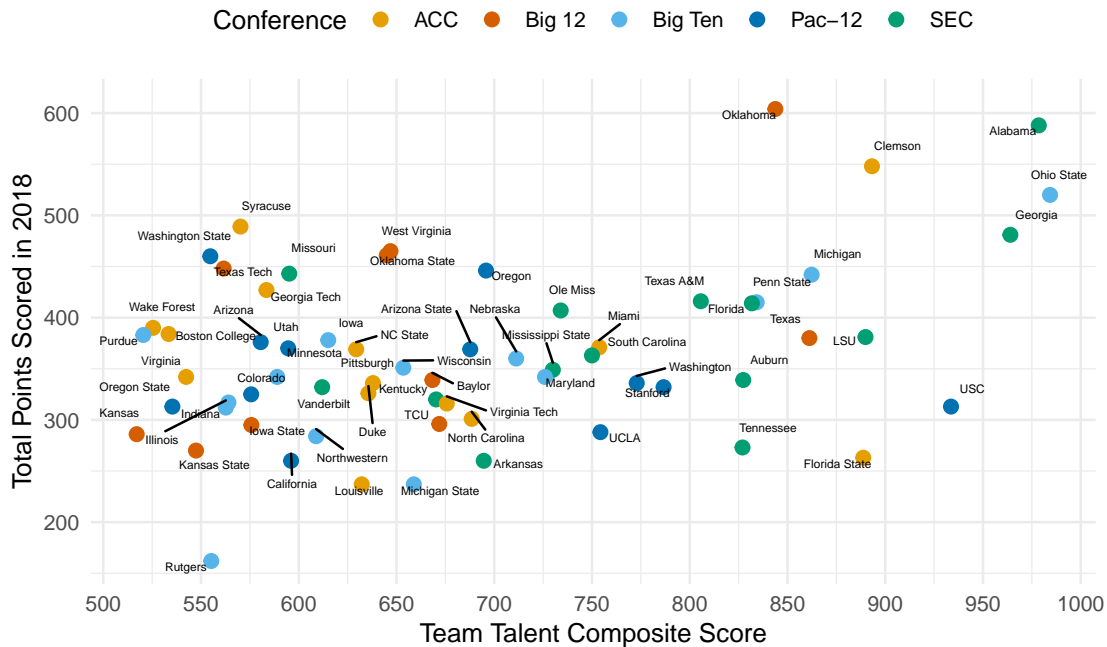
Intraconference margin of victory is pretty consistent between the conferences, so no conference stands out as being more unbalanced than the others. In inter-conference games, the Big 12 usually wins big while the Pac-12 usually only edge their opponents.

Does talent correspond to a better team?

A popular metric for recruiting is the *team talent composite* where a team with a high composite recruits the best players. I compared the talent composites of each P5 team with the points they scored in 2018.

For most schools, talent doesn't impact scoring

Only Oklahoma, Alabama, Clemson, and Ohio St are in a league apart



Talent alone doesn't build championships (as Alabama and company demonstrate, you also probably need money).