

College Football Stats & Analytics

Harry Torres and Maurice Dure

Project 3



Motivation

Out of all the Conferences in Division I NCAA Football, who is the most dominant? What are the biggest statistical measures to find out who has the strongest and toughest conference?

Find robust dataset with plentiful data points.

Transform the dataset and load the data points into a visual that can be analyzed and interpreted.

Use predictive analysis to show what the data is telling us.



Data Cleanup / Exploration

| Year | Conference | Team | Games | Win | Loss | Off.Rank | Off.TDs | First.Down.Rank | Redzone.Off.Rank | Touchdowns.Allowed | Touchdowns | Total.Points | Points.Per.Game | Time.of.Possession.Rank | Win.Ratio | Scoring.Ra |
|------|---------------|-----------------------------------|-------|-----|------|----------|---------|-----------------|------------------|--------------------|------------|--------------|-----------------|-------------------------|-----------|------------|
| 2017 | Mountain West | Air Force (MWC) | 12 | 5 | 7 | 49 | 48 | 14 | 38 | 49 | 50 | 377 | 31.4 | 7 | 0.417 | 0.9 |
| 2017 | MAC | Akron (MAC) | 14 | 7 | 7 | 121 | 38 | 129 | 121 | 49 | 42 | 310 | 22.1 | 89 | 0.500 | 0.9 |
| 2017 | SEC | Alabama (SEC) | 14 | 13 | 1 | 29 | 64 | 17 | 57 | 18 | 66 | 519 | 37.1 | 70 | 0.929 | 0.8 |
| 2017 | Sun Belt | Appalachian St. (Sun Belt) | 13 | 9 | 4 | 31 | 55 | 20 | 67 | 30 | 55 | 435 | 33.5 | 32 | 0.692 | 0.8 |
| 2017 | Pac-12 | Arizona (Pac-12) | 13 | 7 | 6 | 12 | 67 | 128 | 34 | 59 | 71 | 537 | 41.3 | 92 | 0.538 | 0.9 |
| 2017 | Pac-12 | Arizona St. (Pac-12) | 13 | 7 | 6 | 37 | 48 | 110 | 5 | 54 | 51 | 414 | 31.8 | 29 | 0.538 | 0.8 |
| 2017 | SEC | Arkansas (SEC) | 12 | 4 | 8 | 94 | 42 | 67 | 38 | 56 | 46 | 345 | 28.8 | 28 | 0.333 | 0.9 |
| 2017 | Sun Belt | Arkansas St. (Sun Belt) | 12 | 7 | 5 | 10 | 55 | 32 | 122 | 40 | 59 | 454 | 37.8 | 87 | 0.583 | 0.9 |
| 2017 | FBS | Army West Point (FBS Independent) | 13 | 10 | 3 | 76 | 52 | 40 | 48 | 39 | 56 | 399 | 30.7 | 4 | 0.769 | 0.9 |
| 2017 | SEC | Auburn (SEC) | 14 | 10 | 4 | 26 | 55 | 32 | 36 | 31 | 58 | 474 | 33.9 | 31 | 0.714 | 0.8 |
| 2017 | MAC | Ball St. (MAC) | 12 | 2 | 10 | 110 | 26 | 46 | 105 | 65 | 27 | 215 | 17.9 | 10 | 0.167 | 0.8 |
| 2017 | Big 12 | Baylor (Big 12) | 12 | 1 | 11 | 59 | 33 | 74 | 101 | 52 | 33 | 292 | 24.3 | 68 | 0.083 | 0.7 |
| 2017 | Mountain West | Boise St. (MWC) | 14 | 11 | 3 | 57 | 51 | 58 | 88 | 42 | 57 | 455 | 32.5 | 35 | 0.786 | 0.8 |
| 2017 | ACC | Boston College (ACC) | 13 | 7 | 6 | 84 | 40 | 37 | 81 | 37 | 43 | 334 | 25.7 | 55 | 0.538 | 0.9 |
| 2017 | MAC | Bowling Green (MAC) | 12 | 2 | 10 | 85 | 34 | 121 | 17 | 61 | 36 | 304 | 25.3 | 117 | 0.167 | 0.8 |
| 2017 | MAC | Buffalo (MAC) | 12 | 6 | 6 | 36 | 42 | 68 | 106 | 39 | 43 | 342 | 28.5 | 69 | 0.500 | 0.8 |
| 2017 | FBS | BYU (FBS Independent) | 13 | 4 | 9 | 118 | 25 | 79 | 117 | 37 | 26 | 222 | 17.1 | 108 | 0.308 | 0.8 |

CSV of yearly data collected from Kaggle.

Data Cleanup / Exploration

```
df20 = grouped_20.sort_values(by=['Off.TDs', 'Touchdowns.Allowed'],
                               ascending=False)

plt.rcParams['figure.figsize'] = [20,8]

fig = plt.figure()
chart = df20.plot(kind="bar",y=["Off.TDs", "Touchdowns", "Touchdowns.Allowed"], color=['Blue','Green', 'Red'])
chart.set_xlabel('Conference')
chart.set_ylabel('Total Number of Touchdowns')
chart.set_title("Touchdown Metrics for 2020")
plt.xticks(rotation=45)
plt.show()
```

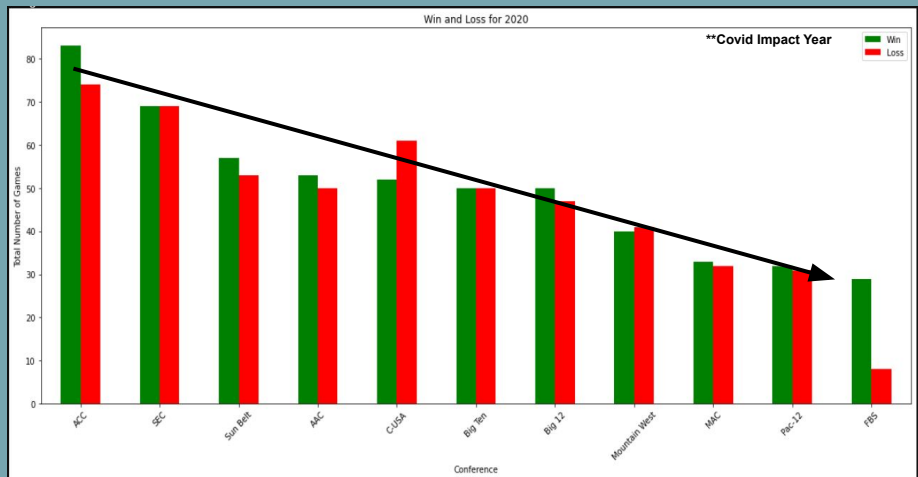
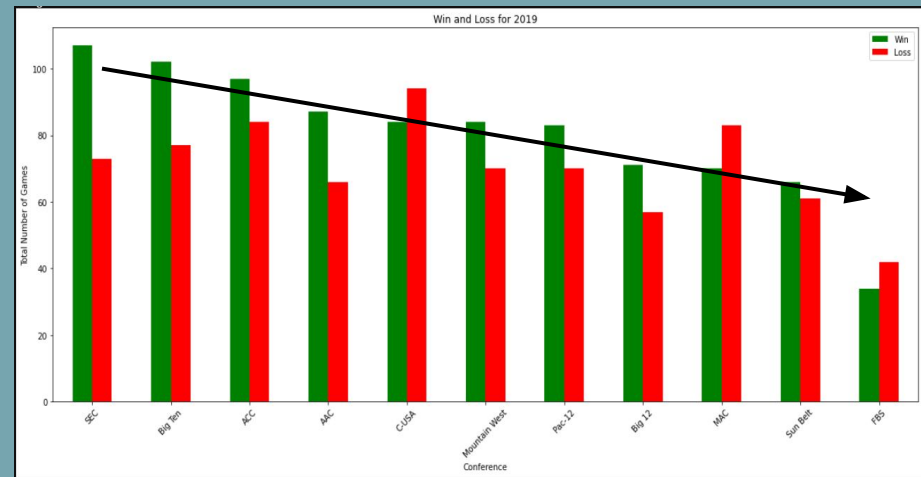
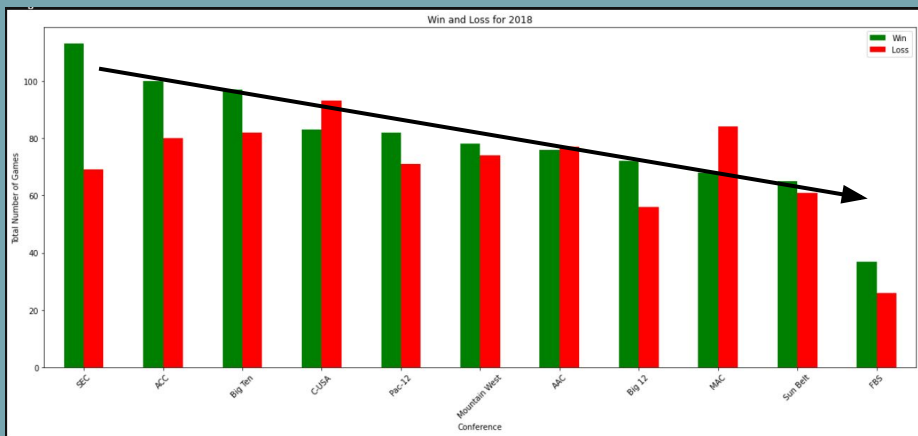
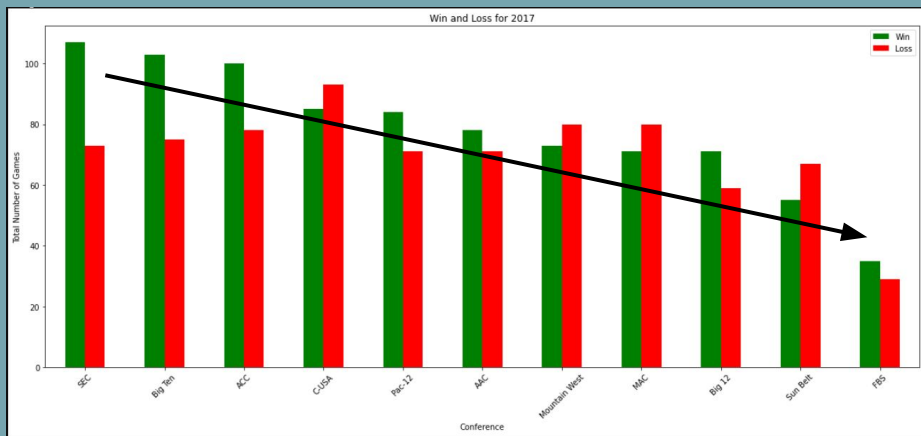
The code used for plotting the scoring and win/loss data.

Data Cleanup / Exploration

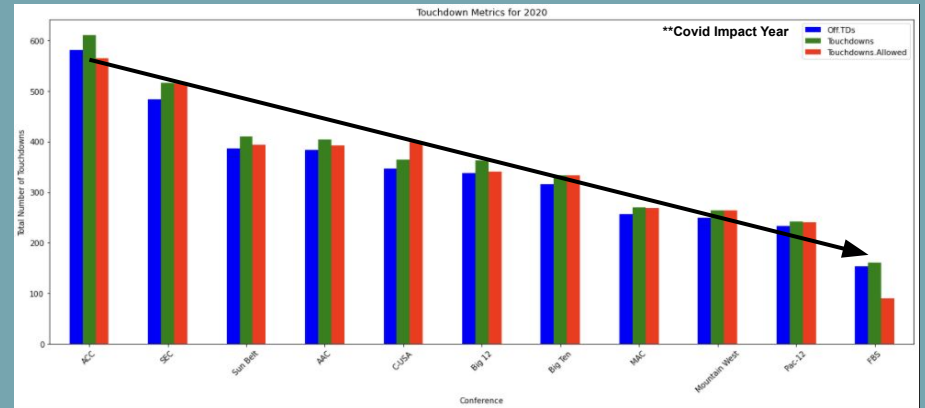
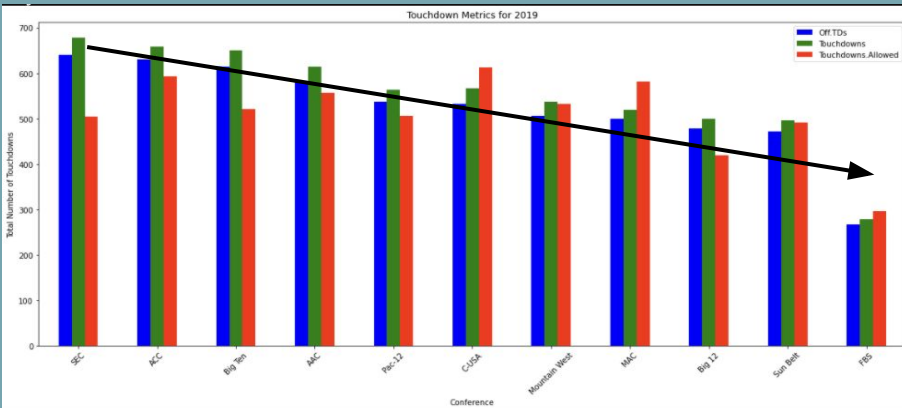
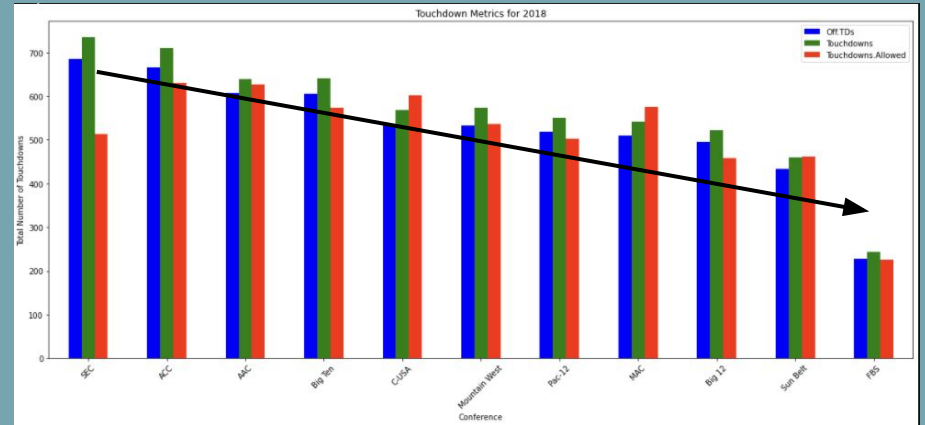
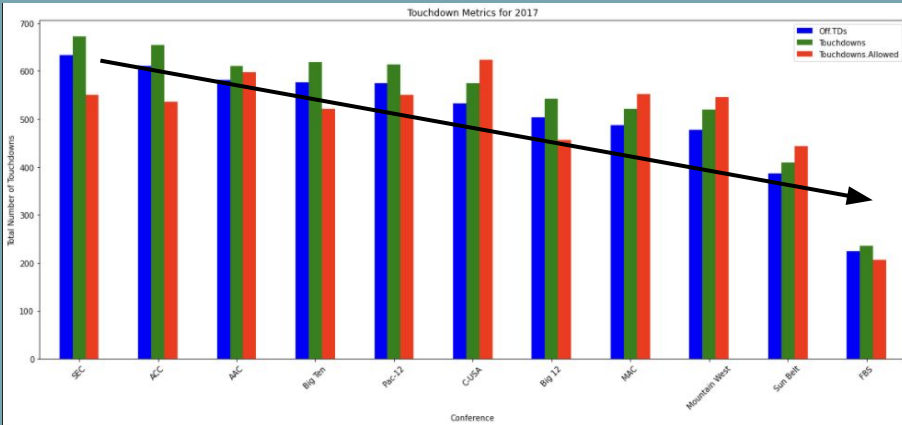


- Normalize yearly data
- Inserted certain metrics
- Summation and Averages
- Some metrics required the joining of multiple year's data

Data Analysis - Win and Loss

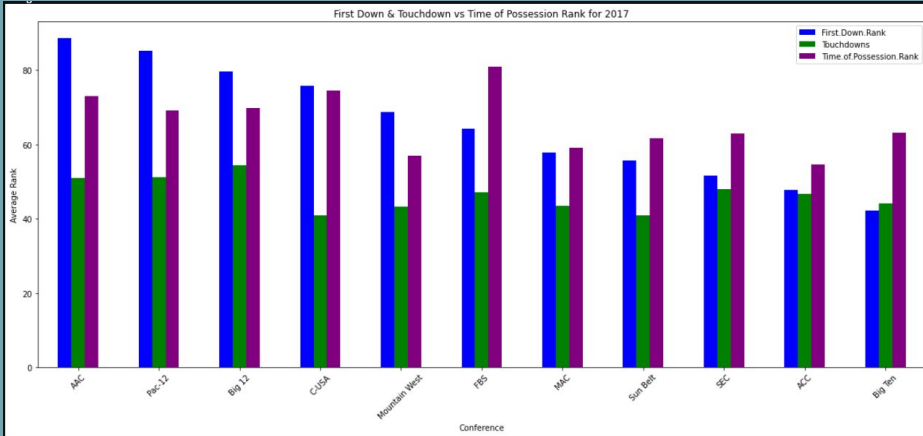


Data Analysis - Touchdown Metrics

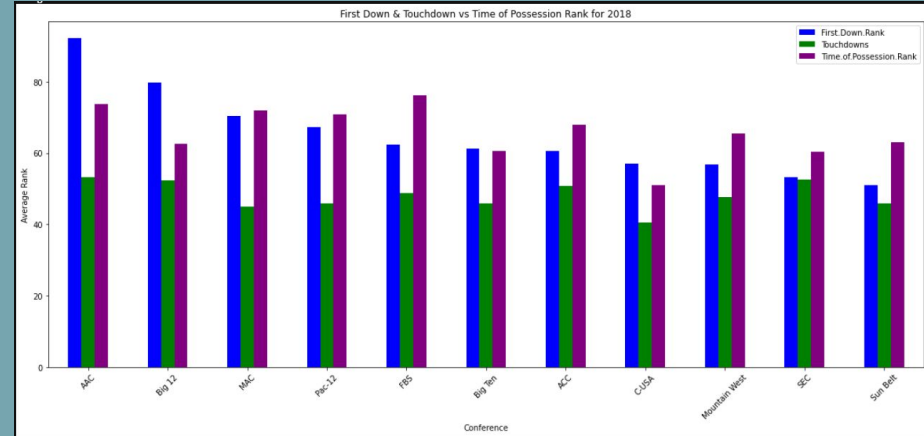


Data Analysis - Time of Possession vs First & Touchdown Ranks

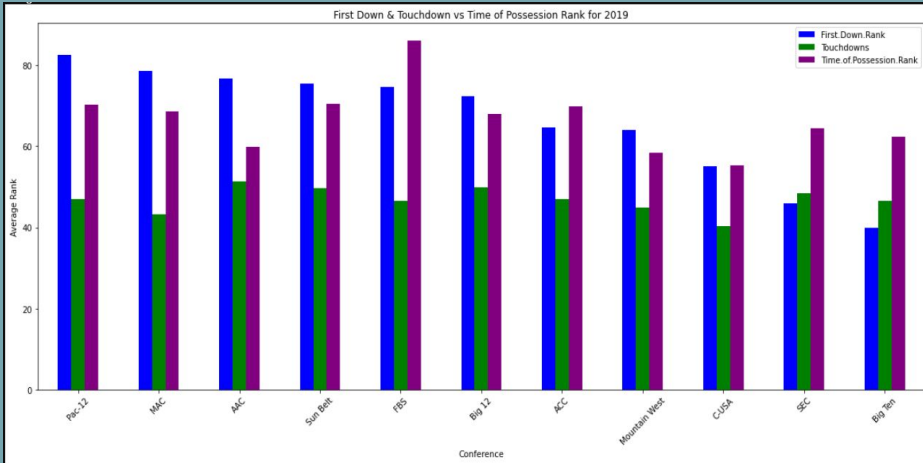
First Down & Touchdown vs Time of Possession Rank for 2017



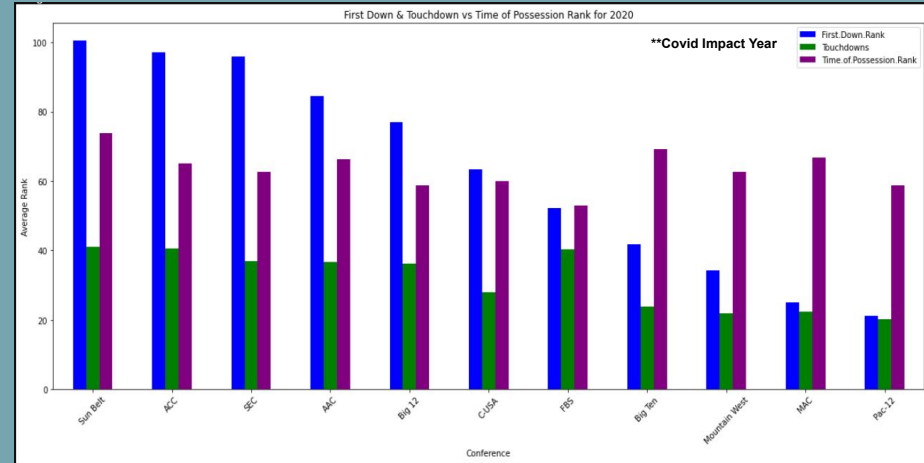
First Down & Touchdown vs Time of Possession Rank for 2018



First Down & Touchdown vs Time of Possession Rank for 2019



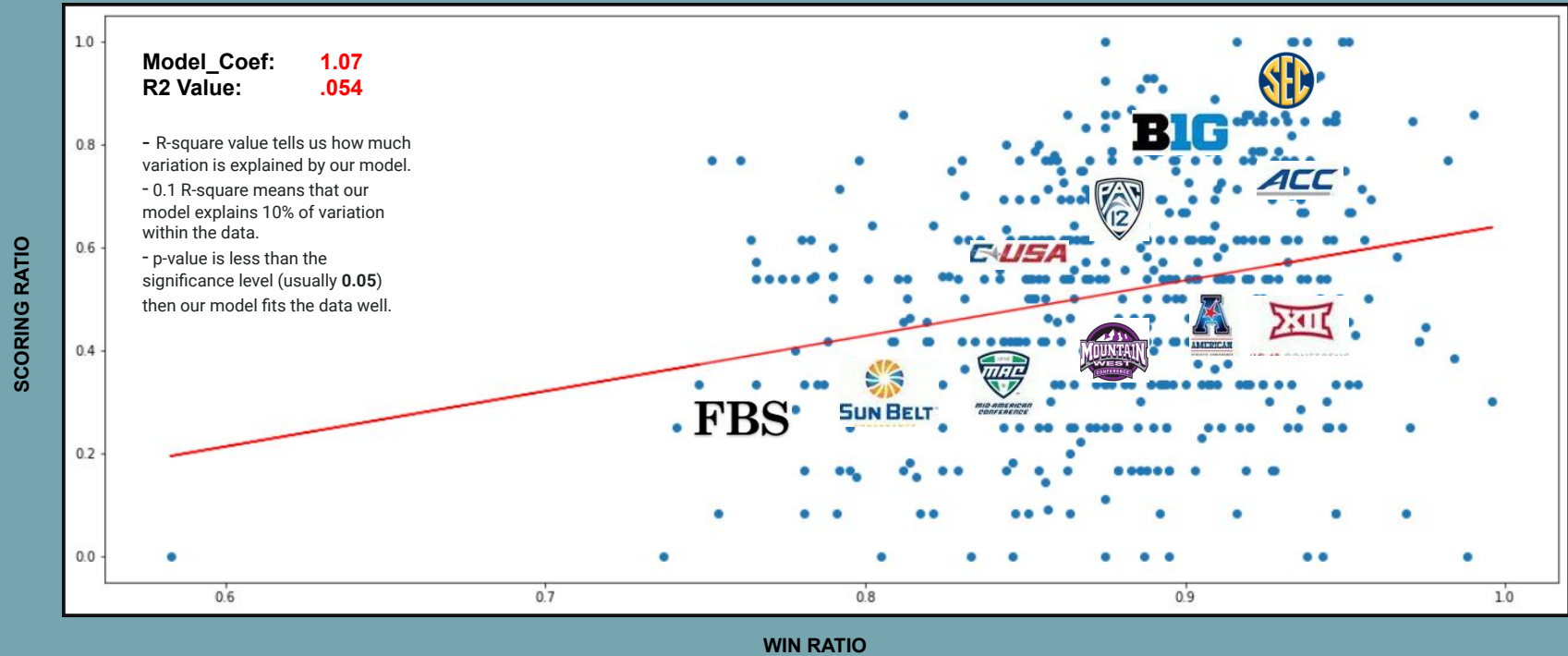
First Down & Touchdown vs Time of Possession Rank for 2020



AP POLL RANKINGS

| | 2017 | | 2018 | | 2019 | | 2020 | |
|-----------|-------------------|---------------|---------------|---------------|-------------|---------------|------------------|---------------|
| 1 | Alabama | SEC | Alabama | SEC | LSU | SEC | Alabama | SEC |
| 2 | Georgia | SEC | Clemson | ACC | Ohio State | Big Ten | Ohio State | Big Ten |
| 3 | Oklahoma | Big12 | Notre Dame | FBS | Clemson | ACC | Clemson | ACC |
| 4 | Clemson | ACC | Oklahoma | Big12 | Oklahoma | Big12 | TX A&M | Big12 |
| 5 | Ohio State | Big Ten | Georgia | SEC | Georgia | SEC | Notre Dame | FBS |
| 6 | UCF | AAC | Ohio State | Big Ten | Oregon | Pac12 | Oklahoma | Big12 |
| 7 | Wisconsin | Big Ten | Michigan | Big Ten | Baylor | Big12 | Georgia | SEC |
| 8 | Penn State | Big Ten | UCF | AAC | Wisconsin | Big Ten | Cincinnati | AAC |
| 9 | TCU | Big12 | Washington | Pac12 | Florida | SEC | Iowa State | Big12 |
| 10 | Auburn | SEC | Florida | SEC | Penn State | Big Ten | Northwestern | Big Ten |
| 11 | Notre Dame | FBS | LSU | SEC | Utah | Pac12 | BYU | FBS |
| 12 | USC | Pac12 | Penn State | Big Ten | Auburn | SEC | Indiana | Big Ten |
| 13 | Miami | ACC | Washington St | Pac12 | Alabama | SEC | Florida | SEC |
| 14 | Oklahoma State | Big12 | Kentucky | SEC | Michigan | Big Ten | Coastal Carolina | Sun Belt |
| 15 | Michigan State | Big Ten | Texas | Big12 | Notre Dame | FBS | Louisiana | Sun Belt |
| 16 | Washington | Pac12 | West Virginia | Big12 | Iowa | Big Ten | Iowa | Big Ten |
| 17 | Northwestern | Big Ten | Utah | Pac12 | Memphis | AAC | Liberty | FBS |
| 18 | LSU | SEC | MS State | SEC | Minnesota | Big Ten | North Carolina | ACC |
| 19 | Mississippi State | SEC | TX A&M | SEC | Boise State | Mountain West | Texas | Big12 |
| 20 | Stanford | Pac12 | Syracuse | ACC | App State | Sun Belt | Oklahoma St | Big12 |
| 21 | USF | AAC | Fresno State | Mountain West | Cincinnati | AAC | USC | Pac12 |
| 22 | Boise State | Mountain West | Northwestern | Big Ten | USC | Pac12 | Miami | ACC |
| 23 | NC State | ACC | Missouri | SEC | Navy | AAC | Ball State | MAC |
| 24 | Virginia Tech | ACC | Iowa State | Big12 | Virginia | ACC | San Jose State | Mountain West |
| 25 | Memphis | AAC | Boise State | Mountain West | Oklahoma St | Big12 | Buffalo | MAC |
| CFB CHAMP | CLEMSON | ACC | ALABAMA | SEC | CLEMSON | ACC | LSU | SEC |

Data Analysis - Linear Regression



Post Mortem

- Difficulties came when the data needed to be visualized, line plots would not work for telling the story of the data. The biggest issues were gaps in the data when it came to plotting.
- We tried different plotting libraries, and matplotlib was the only one that worked, with a bar chart.
- The SEC and ACC are the dominant conferences according to the plots and regression model.

QUESTIONS