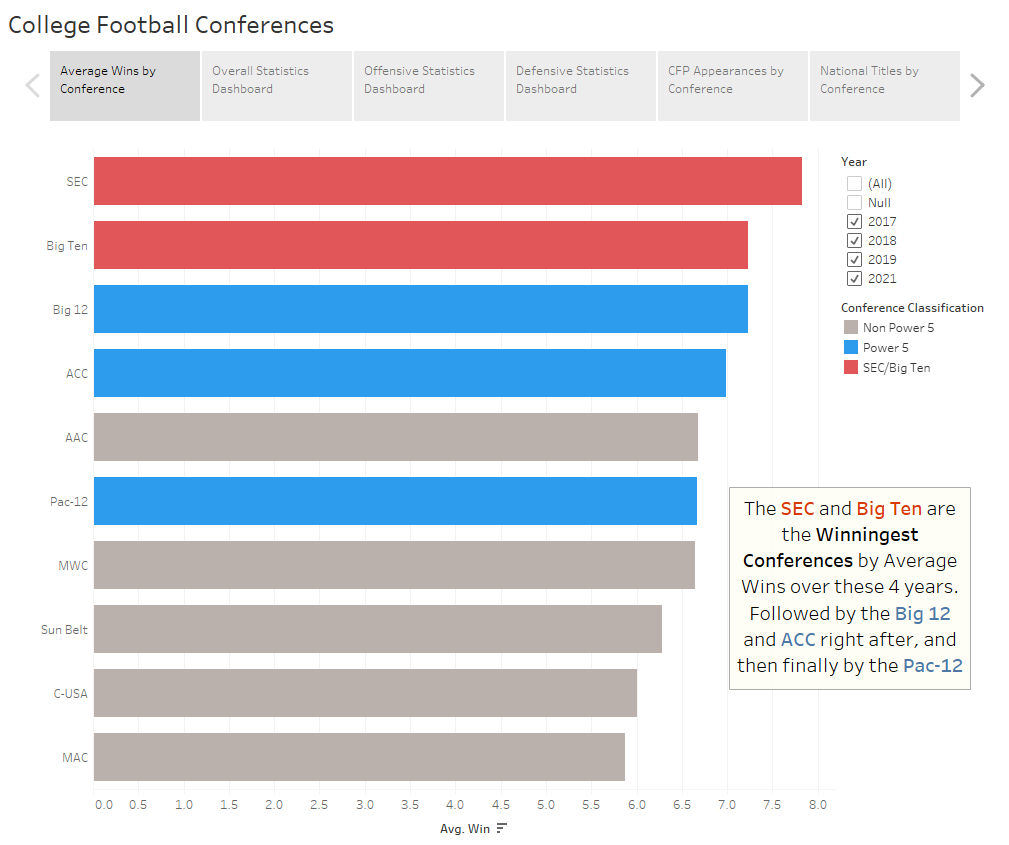
Project Final Report - Max Stevens & Zack Lasek

Implementation:

[**The Final Story and Dashboard**](https://public.tableau.com/app/profile/max.stevens2588/viz/ProjectGraphs_Final/FinalStory)

Write-Up:

**Executive Summary**



Our project deals with the success of college football conferences over the span of 4 seasons of data (2017, 2018, 2019, and 2021) and what goes into or leads to which conferences being the most successful. The goal was to create a project that can be used by the public, but more specifically college football fans or amateur analysts who want to see what conferences win the most, the different statistics that go into that success the most, and how that all has translated to postseason success.

* The first sheet, which can be seen above, shows which conferences had the most wins on average over the 4 years. We highlighted (in Blue) the Power 5 Conferences, on this graph/sheet and all others, because those 5 are the ones we wanted to single out the non-Power 5 Conferences. Additionally, we highlighted the SEC and Big Ten even more (in Red) because they were the top 2 performing conferences and we wanted to focus on them even more throughout the dashboard and the rest of the graphs made.
* The next 3 sheets are for the interactive dashboard. There is a main/miscellaneous one that shows stats like turnovers and penalties. Additionally, the next 2 sheets are for offensive statistics and defensive statistics, like rushing and passing offense and defense. The dashboards have navigation buttons that can be used easily (using ALT + click) to navigate between the 3 dashboards for interactivity. Finally, they all have a year filter for an extra layer of interactive elements and for more analysis tools. These 3 sheets serve as the main tool to show why conferences are having success and possibly outliers or statistics that are more or less important than others based on that success.
* The second to last sheet shows the conferences with the most CFP appearances over the 4 years. We wanted to have final slides that show how wins in the regular season and statistics correlate with the final goal of making the CFP and winning the National Championship, if at all.
* The final sheet shows the total National Championships won over the 4 years. This is similar to the CFP appearances sheet, but we just wanted to take the one a step further and show what conferences won the titles, since that is the end goal in college football.

**Basic Info**

The Winningest Conferences in College Football and the Factors Leading to It

Max Stevens and Zack Lasek

**Data**

The data that we used came from Kaggle and it deals with [College Football Data](https://www.kaggle.com/datasets/jeffgallini/college-football-team-stats-2019). We only used the data from the years of 2017-2021, and we excluded 2020 because it was the COVID season, and a lot of the statistics and games played are way different compared to the other seasons. We cleaned the data by removing columns that we thought were unnecessary or we had no plans of using at all, and we did this because the original data had so many columns. The final data had 40 columns and each sheet had 128 rows of data, so there are 512 final rows when combined in Tableau. The sheets were joined in Tableau using a Union, so that is why the data for different years are on different sheets. We added a Conference column because it did not have one at first for all the years, and we wanted conferences to be the main focus of our story, so we had to add one for all years. The other data cleaning steps that we had to complete was only making sure all column names, conference spellings, team spellings, and year columns were correct and consistent across all four years to avoid joining errors and any other errors once in Tableau. The final addition that was needed for the data was to add columns that showed College Football Playoff (CFP) Appearances and National Titles won during the span of the data, so we found that data online and added it to the data. We wanted to use this data as the final conclusion to the story, so it was necessary to add it towards the end of the project. We added it simply by adding the totals for all teams across all 4 years to the end of the 2021 season sheet of data on the excel sheet. It was easier to add it this way than doing the numbers for every year individually on each sheet, and since it didn’t mess the data up when in Tableau we went with this method in the end. After all those changes, the final data dictionary can be seen below and the data file itself will be submitted alongside this report.

Data Dictionary

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Type** | **Description** |
| TeamID | Numeric | Unique ID for each team |
| Team | Text | College/Team name |
| Conference | Text | Conference name |
| Games | Numeric | Games played |
| Win | Numeric | Games won |
| Loss | Numeric | Games lost |
| Off Rank | Numeric | Offensive rank among teams |
| Off TDs | Numeric | Total offensive touchdowns scored |
| Off Yards per Game | Numeric | Total offensive yards gained per game |
| Def Rank | Numeric | Defensive rank among teams |
| Yards Allowed | Numeric | Total yards given up all season on defense |
| Yards per Game Allowed | Numeric | Average yards given up per game on defense |
| Passing Off Rank | Numeric | The rank of a team’s passing offense among all teams |
| Pass Yards Per Game | Numeric | Passing yards on offense per game |
| Pass Def Rank | Numeric | The rank of a team’s defense in terms of passing among all teams |
| Pass Yards Per Game Allowed | Numeric | Average yards given up on defense per game |
| Penalty Rank | Numeric | Rank among all teams in terms of amount of penalties and yards committed |
| Penalty Yards Per Game | Numeric | Penalty yards against a team per game |
| Redzone Off Rank | Numeric | Rank among all teams based on how good offense is inside the 20 yard line (Redzone) |
| Redzone Def Rank | Numeric | Rank among all teams based on how good defense is inside the 20 yard line (Redzone) |
| Redzone Points Allowed | Numeric | Total points given up in the Red Zone (20 yards to endzone) by a defense |
| Rushing Off Rank | Numeric | Rank among all teams in terms of their rushing offense |
| Yards/Rush | Numeric | Number of rushing yards gained per every rushing attempt |
| Rushing Yards per Game | Numeric | Average rushing yards gained on offense per game |
| Rushing Def Rank | Numeric | Rank among all teams based on team’s defense against the run |
| Yds/Rush Allowed | Numeric | Number of yards given up by defense on rushing attempts |
| Rush Yards Per Game Allowed | Numeric | Average number of rushing yards given up by the defense per game |
| Scoring Def Rank | Numeric | Rank among all teams in terms of how good a team’s defense is at not allowing other team to score |
| Touchdowns Allowed | Numeric | Total number of touchdowns allowed by a team |
| Avg Points per Game Allowed | Numeric | Average amount of points given up by a team’s defense per game |
| Touchdowns | Numeric | Total touchdowns scored by a team (offense, defense, special teams) |
| Points Per Game | Numeric | Average points scored by a team in total per game |
| Time of Possession Rank | Numeric | Rank among all teams in terms of how much of the game a team has the ball for on offense |
| Average Time of Possession per Game | Time/Numeric | Average amount of time a team has the ball on offense per game |
| Turnover Rank | Numeric | Rank among all teams in terms of total offensive and defensive turnovers |
| Turnover Margin | Numeric | Turnovers committed by the offense minus turnovers gained by the defense |
| Avg Turnover Margin per Game | Numeric | Average turnovers forced minus turnovers lost per game |
| Year | Numeric | Year of the season played |
| Punt Returns | Numeric | Number of times a team forced the other team to punt |
| Opponent Punt Returns | Numeric | Number of times a team was forced to punt |
| Scoring Off Rank | Numeric | Rank among all teams in regard to how well their offense scores in total |
| Total CFP Appearances (2021 sheet only) | Numeric | Total number of appearances made in College Football Playoff over all 4 years |
| Total Titles (2021 sheet only) | Numeric | Total number of National Championships won over all 4 years |

**Visualizations**

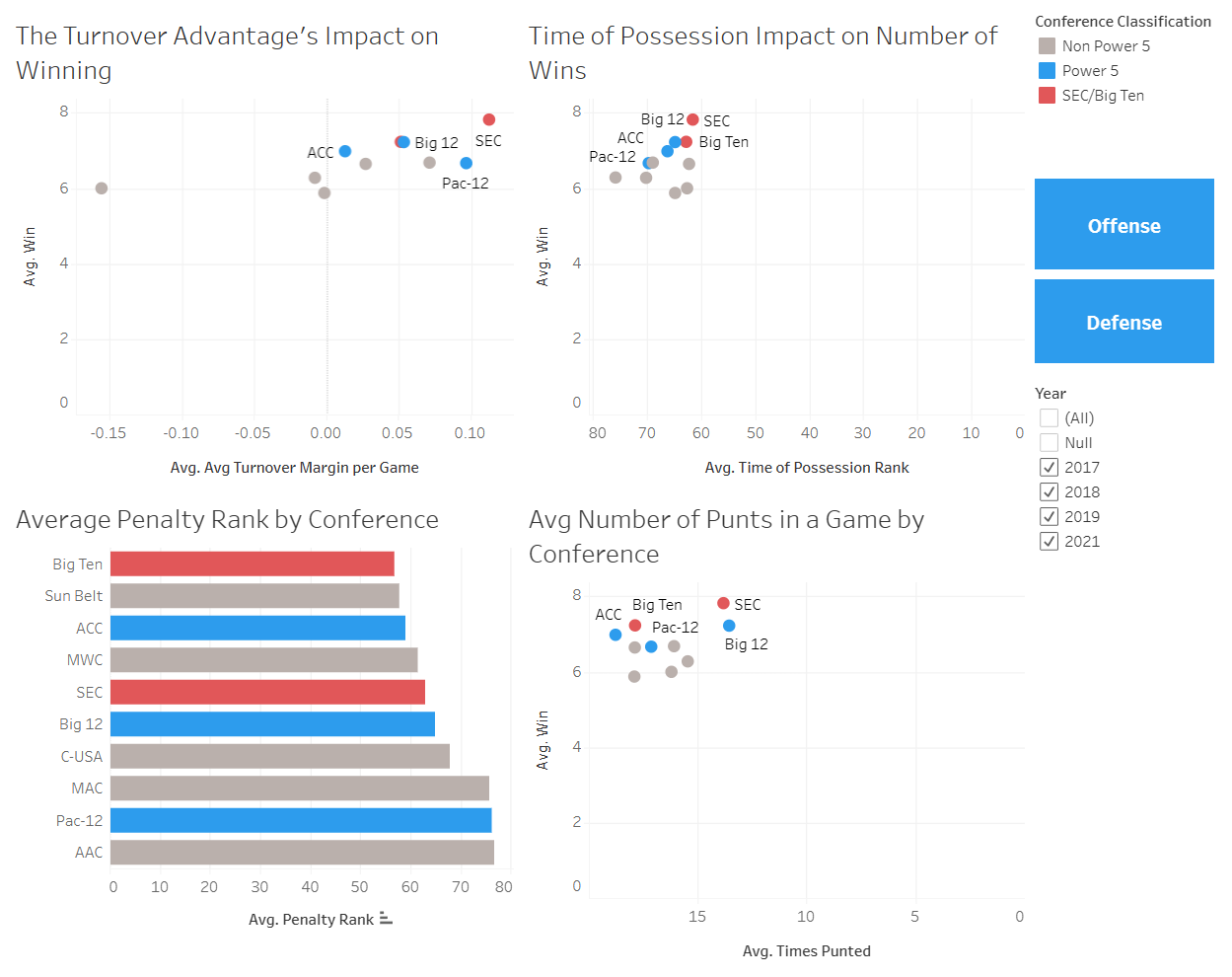
Interactive Dashboard(s)

For our interactive visualization we built a dashboard that explored overall team statistics for each conference that could then be switched to a defensive or offensive focus with the click of a button. The buttons on the right of each dashboard navigate to either the offensive or defensive dashboards, or back to the overall one. We felt having this interactivity with all visuals connected would be beneficial for our users, as it makes it very simple to view whichever kind of statistics you want to dive into. The second form of interactivity on our dashboard is the ability to switch the years each graph is looking at. With the year selection being available on the right side of each dashboard, users have the ability to look at a single season of interest or a range of which they choose. We felt this is important because coaches or fans who use our dashboard may have reasons for wanting to look at a specific season over others. The year filtering also allows for more detailed analysis, such as one singular year or a two-year span, or for more overall trend searching and analysis, such as three-year spans or all four years of the data. The year filter also is attached to and links all graphs on the dashboard, so that provides a way to interact with it and the views are linked.

All three of our dashboards are encoding the same way as well. We have set all marks for non-Power 5 conferences to gray, as our analysis is focused on Power 5. We decided to not exclude those marks however, so that if a coach or fan of a non-Power 5 school wanted to, they could use our dashboard to draw their own insights. Power 5 schools are encoded with a blue hue on all visualizations with our two conferences of focus, B1G and SEC, encoded with a red hue. We felt both of these choices were appropriate because they draw attention to these five school’s positions immediately and hopefully expedite how fast our users digest the information.

The choices of graph types came down to what we wanted to show. When two statistics were being compared, like wins and turnovers, a scatterplot was best and then adding conference in as a detail (color) was necessary. Additionally, bar charts were chosen when we wanted to compare all conferences in one stat, such as penalty rank or offensive/defensive rank.

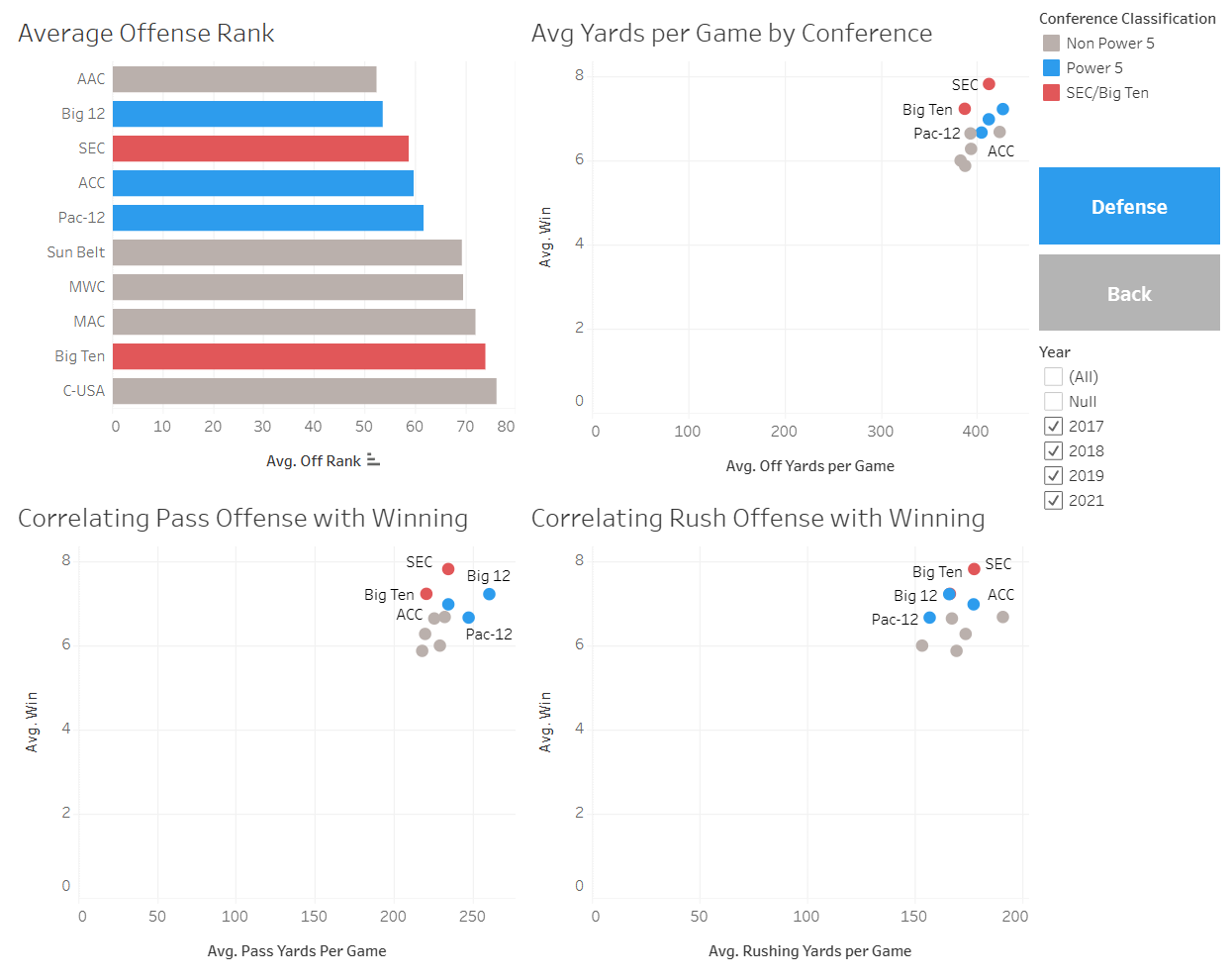
Overall Dashboard



**For each quadrant of the dashboard:**

|  |  |
| --- | --- |
| Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg turnover margin per game * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting B1G & SEC | Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg time of possession rank * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting B1G & SEC |
| Marks:   * Line: Represents a conference   Channels:   * Length: Encodes avg penalty rank * Hue: Encodes conference status as well as highlighting the B1G and SEC | Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg times punting per game * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting B1G & SEC |

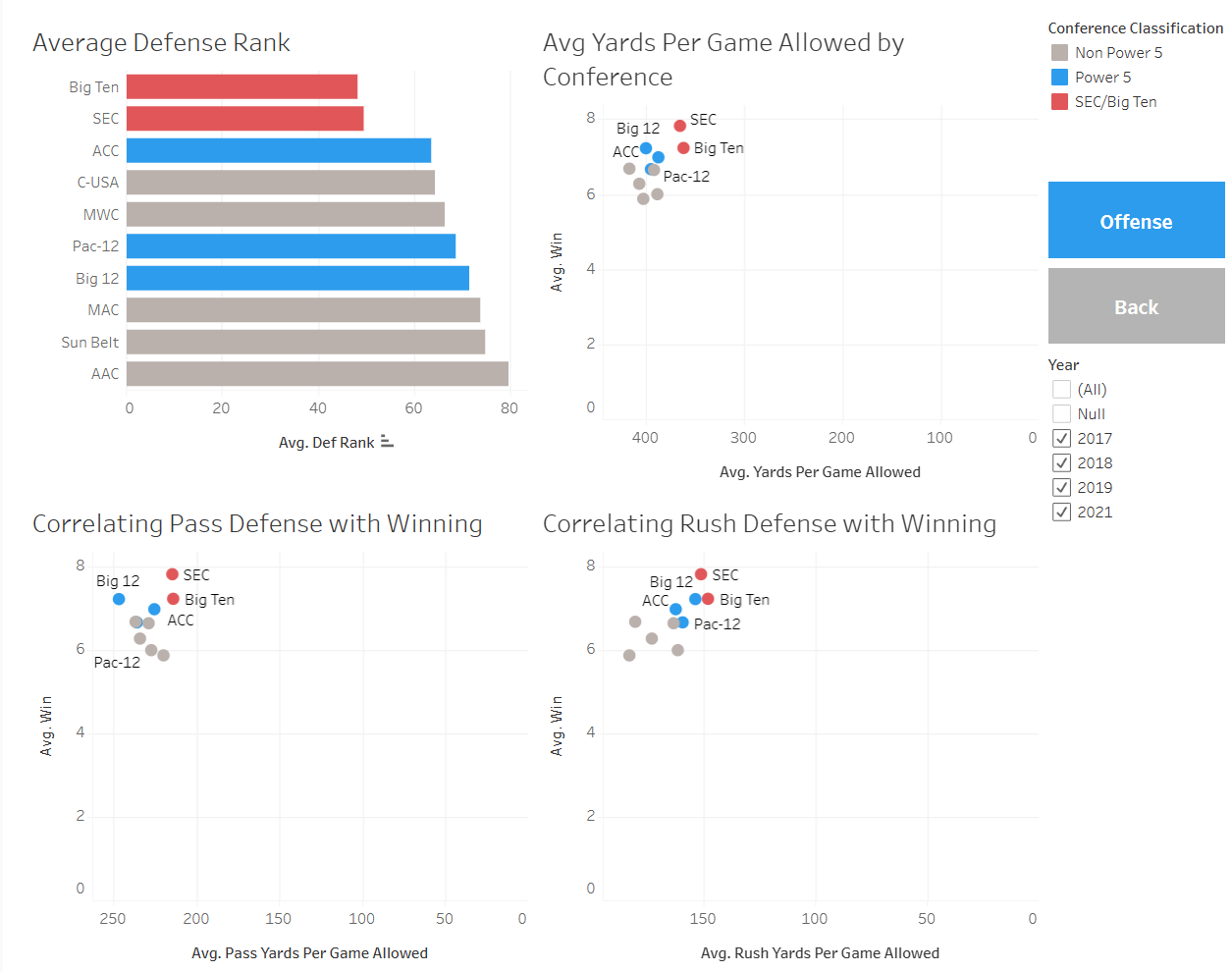
Offensive Focus Dashboard



**For each quadrant of the dashboard:**

|  |  |
| --- | --- |
| Marks:   * Line: Represent a conference   Channels:   * Length: Encodes avg penalty rank * Hue: Encodes conference status as well as highlighting the B1G and SEC | Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg offensive yds/game * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting B1G & SEC |
| Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg passing yds/game * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting B1G & SEC | Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg rushing yds/game * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting B1G & SEC |

Defensive Focus Dashboard



**For each quadrant of the dashboard:**

|  |  |
| --- | --- |
| Marks:   * Line: Represent a conference   Channels:   * Length: Encodes avg defense rank * Hue: Encodes conference status as well as highlighting the B1G and SEC | Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg yds/game allowed * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting the B1G and SEC |
| Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg pass yds/game allowed * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting the B1G and SEC | Marks:   * Points: Represent a conference   Channels:   * X-position: Encodes avg rush yds/game allowed * Y-position: Encodes avg wins per season * Hue: Encodes conference status as well as highlighting the B1G and SEC |

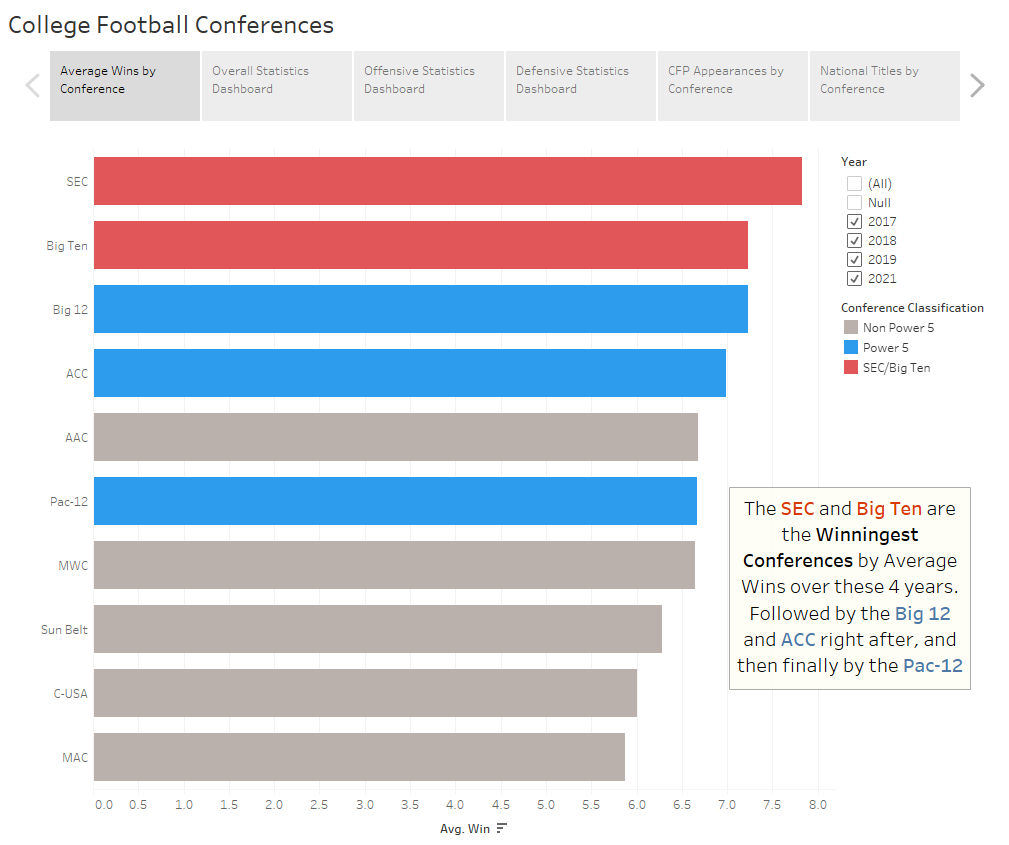
Static Visuals

Our three static visuals are encoding using the same rules as our three dashboards. The color-coding classification that has been used so far in the dashboard graphs is still being used here, so Power 5 conferences are Blue, and the SEC and Big Ten are singled out with Red.

The first static visual is the one shown in the executive summary, which is the conferences sorted by the highest wins per year over the four-year span of the data. It is the main reason and statistic behind our project and the goal it is serving, so that is why it was made and is the first and central visual of the story. The conferences that win the most are what we highlight throughout the dashboard and final visuals, so showing that right away and highlighting them on their own is vital to our story.

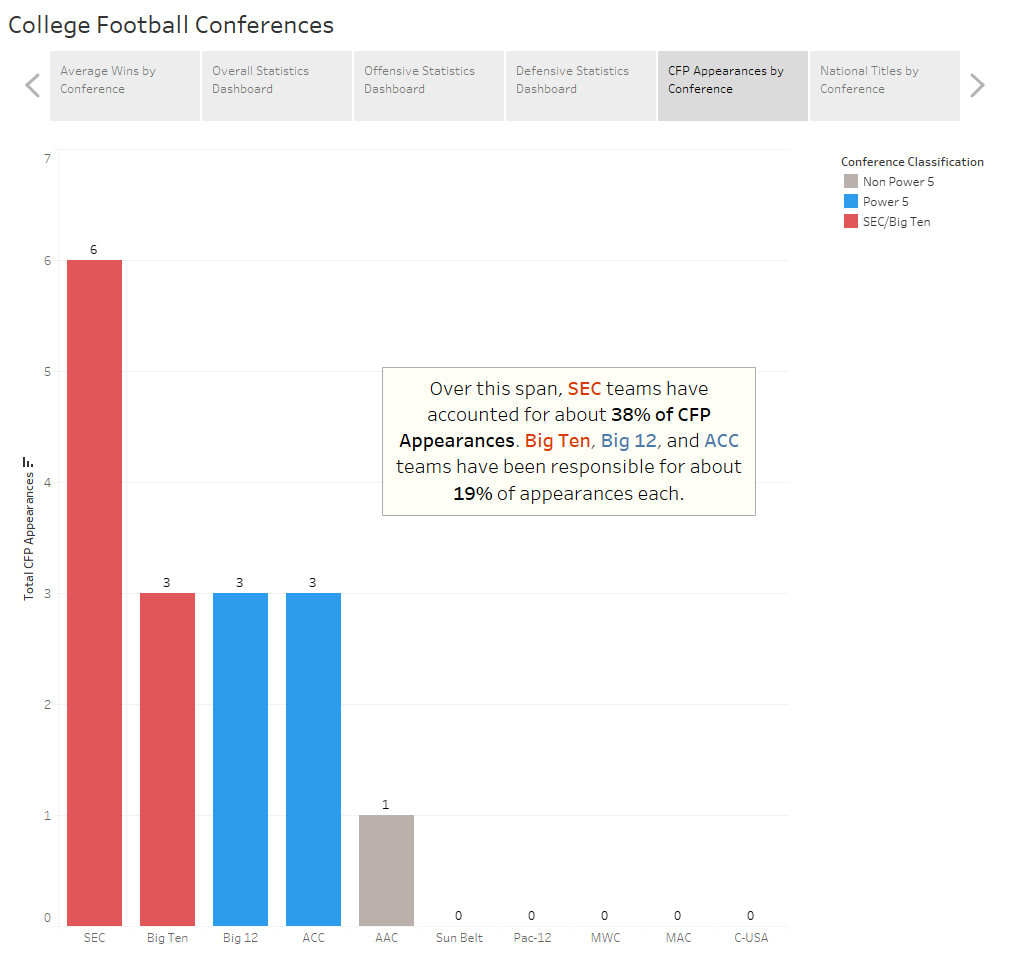
The second static visual shows the conferences with the most College Football Playoff (CFP) appearances over the span of the data. Making the CFP is a main goal of every Power 5 team who is competing, so showing how many appearances, out of 16 total, were made by what conferences seemed like a very important statistic and visual that could be used as a part of the conclusion to our overall story. It also is a way to show if the wins on average and team statistics are aligned with what conferences most often made the CFP.

The third and final static visual shows what conferences had the most National Championships won in the 4 years. It serves a very similar purpose as the previous static visual (CFP appearances) but just at a more detailed view, as in the only team, or conference, that actually won the playoff that year and not all the teams that made it. Even though there is not that much data being used (only the 4 years, so 4 champions), it still serves the purpose of showing the conference that won it all. Additionally, it turns out the conference with the most wins and that leads in many statistics, such as CFP appearances, also wins 3 out of 4 championships.



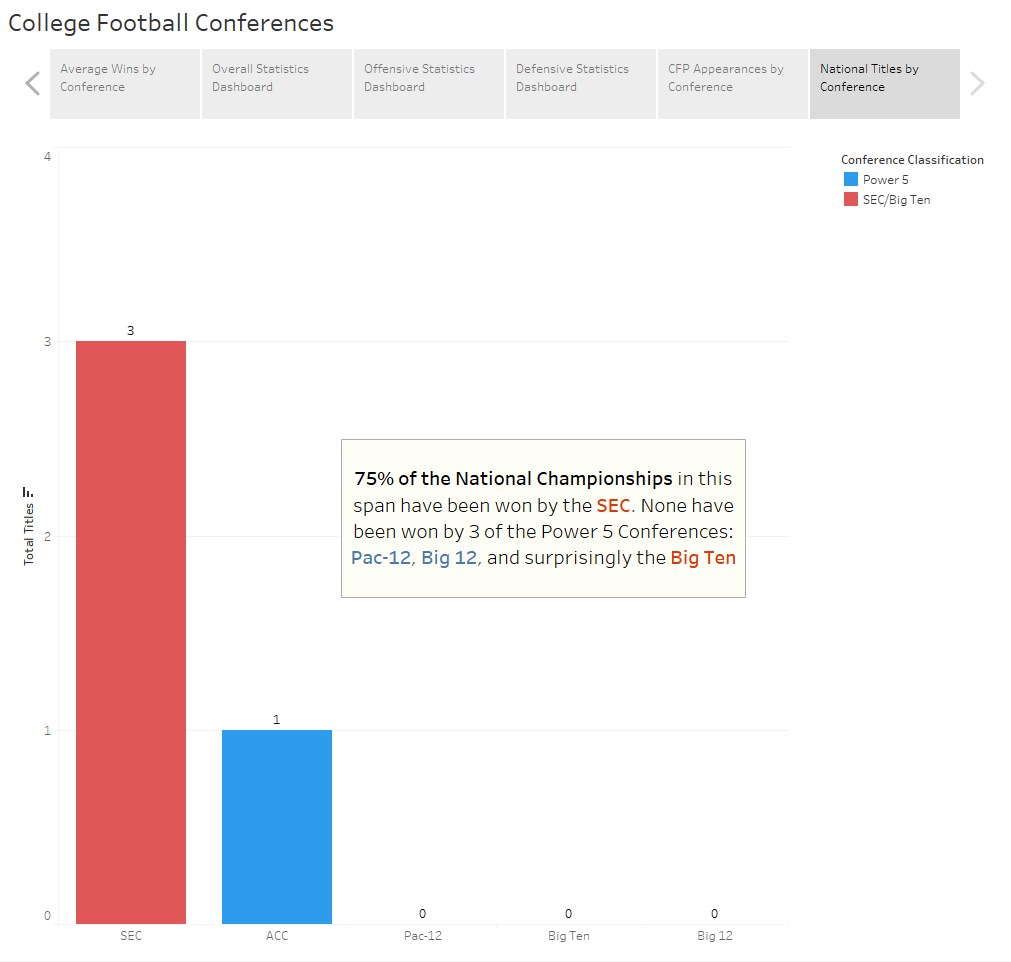
Marks: Line representing the conferences

Channels: Length-Encodes average wins and Hue-Encodes conference status (blue) as well as highlighting the B1G and SEC (red)



Marks: Line representing the conferences

Channels: Length-Encodes total CFP appearances and Hue-Encodes conference status (blue) as well as highlighting the B1G and SEC (red)



Marks: Line representing the conferences

Channels: Length-Encodes total National Championships won and Hue-Encodes conference status (blue) as well as highlighting the B1G and SEC (red)

**Usage Scenario**

Our project can be utilized as a strategic planning tool for college football teams and analysts. Teams can leverage the insights gleaned from the analysis to refine their gameplay, recruiting and training strategies. Understanding which specific statistics contribute to success enables teams to pinpoint areas for improvement and tailor their approaches to meet the unique challenges posed by their respective conferences. In essence, our project becomes a playbook for strategic decision-making, helping teams to optimize their performance on the field. Additionally, it can be used as a resource for fans and analysts who want to improve their knowledge on what leads to success on the field and use that in a personal or even professional way. The screenshots for all interactive elements can be seen in the Visualization section right above this one, since that section has the screenshots for all the graphs, sheets, and dashboards.

**Reflection**

The most enjoyable part of the project was being able to complete something that we are both proud of and took actual time and effort to complete. It is now a real resume builder. Also, working with data we enjoy and are familiar with was also a plus. The least enjoyable part was the roadblocks we faced at first with not having the right data to produce the full story that we wanted to tell. However, once we figured that out and added it to the project, it was a lot easier and made the end product even better and more rewarding.

This project successfully explored the winningest conferences in college football and the contributing factors. Leveraging Kaggle data from 2017 to 2021 (excluding 2020 due to COVID), we transformed and analyzed the dataset. The addition of a Conference column and key metrics like College Football Playoff Appearances expanded our arsenal of data. Our interactive dashboard was our focal point, offering a user-friendly exploration of team statistics by conference, with the flexibility to toggle between offensive and defensive focuses. The encoded visualizations, with Power 5 conferences in blue and B1G/SEC in red, enhanced clarity for our users. The utility of our project as a tool for college football teams and conferences was evident, providing actionable insights for gameplay, recruiting, and training strategies.

Project management was seamless, with each team member contributing to a similar extent. We believe we have delivered a project that combines data-driven analysis with accessible visualizations, showcasing our dedication and ability to translate complex data into practical insights.

The main changes that were made throughout the project lifespan were expanding the data to include more years, so there would be more data to use and further ways to explore and analyze the problem. We also added postseason success, which became a crucial part of the conclusion of the project. Our main idea, which was conference success, stayed the same throughout, but we just added new data or methods of analysis to further that initial goal and make it better.

Our initial proposal was very realistic and doable in Tableau because as we went along, we needed to further our goal/story even further from our initial one so we could utilize Tableau more and have an overall more insightful Tableau Story. We wanted to be able to implement recruiting rankings in some way but adding that data and trying to figure out what years to add it for or how to make it relevant and related to the rest of our graphs and story was difficult, so we decided to not pursue that further. Other than that, there are no other changes or things we would implement or do any different if we could start from scratch.

**Project Management & Team Assessment**

|  |  |  |  |
| --- | --- | --- | --- |
| **Task** | **Completed By** | **Date & Time Required** | **Notes** |
| Project Proposal | Zack & Max | 11/5 - 4 Hours | Completed the project proposal assignment in doc |
| Visual Exploration | Zack & Max | 11/8-12/6 - 8 Hours | Constantly building and editing our visuals to address questions and advance our story and improve it |
| Data Cleaning & Manipulation | Zack (first 2 sheets) & Max (second 2 sheets & union joining) | 11/8 - 1 Hour | Cleaned up multiple tables in excel & then joined them in Tableau using Union |
| Storyboarding | Zack & Max | 11/8-12/1 - 3 Hours | Constantly thinking about the story, how to tell it, and how to improve upon it throughout the entire project timeline |
| Project Update 1 | Zack & Max | 11/7-11/10 - 3 Hours | Completed the project update assignment in doc |
| Project Update 2 | Zack & Max | 11/25-11/28 - 3 Hours | Completed the second update for the project |
| Creating the Dashboards | Zack | 11/20-12/1 - 3 Hours | Created the dashboard that we have now and updating it along the way to be finalized |
| Creating final 3 Sheets for Story and Finalizing the Story | Max | 12/1-12/7 - 3 Hours | Created the visuals and sheets for the story, excluding the dashboard. Added everything to the story and made it look cohesive and polished. |
| Final Project | Zack & Max | 12-7-12/9 - 3 Hours | Created this Final Report |
| **-** | **-** | **-** | **-** |
| **Totals** | **-** | **31 Hours** | **-** |

**Credits**

No credits were used or needed.