



# Analyzing Arsenal Data

CSE46801

Information Visualization

Assignment2 - 20171149 Beomsu Shin

demo-<https://bams11.github.io/IV/>

## Topic selection(Teaser)

I am a football fan and have been enjoying watching matches, including the 22/23 season. One team that stood out for me in the 22/23 season was Arsenal. They used to be a dominant team, even achieving an unbeaten season about 10 years ago, but they have struggled in the past decade. However, in this season, they showed great performances and took the lead in the standings. Although they finished as runners-up, it is still a remarkable achievement. I became curious about how Arsenal managed to achieve such results this season. Therefore, I thought it would be interesting to analyze their performance over the past 5 years. I collected data on their matches and player statistics to visualize and examine the trends. The visualization includes season-by-season match results, the overall win-loss record over the past 5 years, the key players for each season, and the details of each match.



## Data Used

### 1. English Premier League (EPL) Results-

(<https://www.kaggle.com/datasets/irkaal/english-premier-league-results?resource=download>)

The first dataset represents the overall match results of the English Premier League (EPL). It includes data such as the "Season," "Date of the Match," "Home Team," "Away Team," "Match Result," and "Shots and Goals." As I was interested in Arsenal's matches, I filtered the data to include only matches where Arsenal was either the home team or the away team. Additionally, I mapped whether Arsenal played as the home team and the result of the home team to preprocess the data and determine whether Arsenal won, drew, or lost each match. Using this preprocessed data, I created a dataset specifically for Arsenal's win-loss record in each season by iterating through the data in a loop. I also performed further preprocessing, separating the halftime results from the full-time results and visualizing the match details.

### 2. Arsenal EPL (2017/18 - 2022/23) -

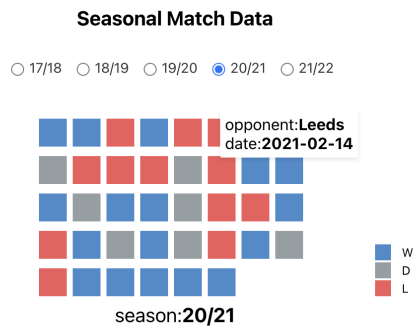
(<https://www.kaggle.com/datasets/rustemnagimov/arsenal-epl-dataset?select=README.md>)

The second dataset consists of individual player data for each match played by Arsenal. It includes information such as the "Match Date," "Starting Lineup Status," "Shots," "Passes," "Tackles," "Playing Time," and other relevant data regarding the performance of Arsenal players in each match. I aggregated the player data for each season by combining their match statistics and calculating the averages. Based on this, I identified the top-performing players in each season and visualized their match performances.

## Features & Design rationale

The first chart represents the match results for each season.

You can select the desired season using radio buttons. Each square in the chart is sorted by match date and represents the match result with

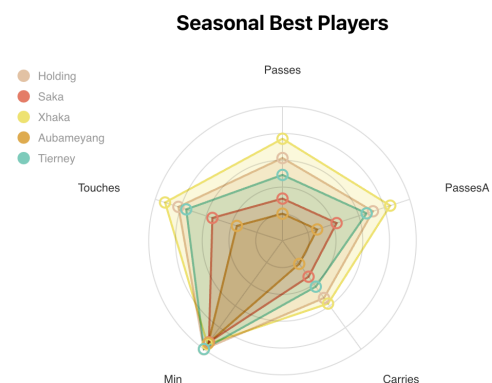


different colors indicating win, draw, or loss. When hovering over a square, a tooltip will appear displaying the match date and the opponent. By clicking on a square, it becomes selected, and it interacts with the "Selected Match Data" chart.

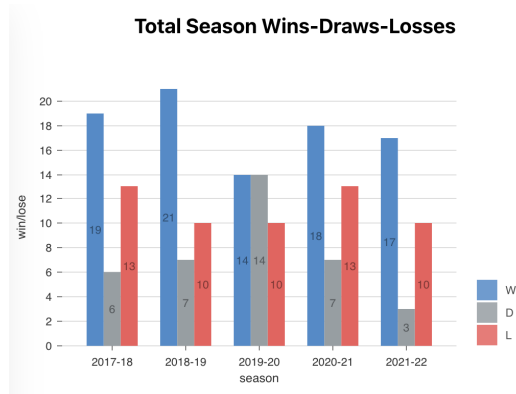
This design was chosen because the data includes dates and a single important attribute. Additional information is displayed on hover to provide more details. While a calendar layout could have been used, it would have resulted in unnecessary visual clutter since matches are not played every day. Therefore, this condensed format was selected.

The second chart represents the Best Players for each season. When a season is selected using the radio buttons above, it displays the Best Players for that specific season. When hovering over the chart, you can see specific numerical values, and you can compare different attributes and values for each key.

This design is suitable for displaying multiple players and their various attributes simultaneously. It allows for easy comparison of a single attribute across different players and facilitates a comprehensive comparison of the overall data.



The third chart represents the Win-Loss data over the course of 5 years. It

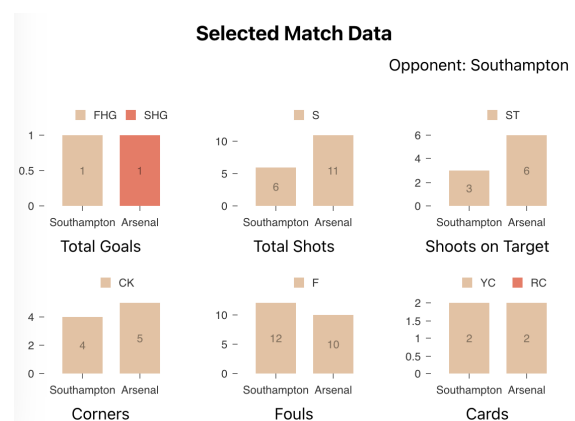


allows for easy and comparative visualization of wins, draws, and losses for each season. When hovering over the chart, tooltips provide additional information, and clicking on a specific season allows for selection. This chart also interacts with other charts in the visualization.

To provide a clear overview of the match results for each season, a Bar Chart was utilized. Specifically, a Grouped Bar Chart was employed to compare and showcase the wins, draws, and losses separately.

The fourth chart represents the detailed results of the selected match. When a match is selected from the first chart showcasing the season-wise match results, this chart provides a comprehensive view of the match details. It includes metrics such as total goals, total shots, shots on target, corner kicks, fouls committed, and disciplinary actions (warnings and dismissals).

To enable simultaneous comparison of multiple attributes between the two teams within a single match result, Small Multiples technique was employed. Four of the data attributes are visualized using a standard bar chart, while the remaining two are presented using a stacked bar chart. This design choice facilitates effective comparison and analysis of various metrics within a single match.



## Usage scenarios

1. Users can initially view the season-wise match details at a glance through the Total Season win, draw, and loss data.
2. They can then select a specific season from the Seasonal Match Data to delve into more detailed match results.
3. During the season selection process, users can also access information about the top players in each season.
4. When selecting individual matches within a season, users can examine the detailed data for each match.

## Observation

1. Users can easily understand how Arsenal's performance and player acquisitions have evolved over the past five years through this system.
2. Users can easily find out which players performed well in a specific season.
3. By comparing the detailed match data, users can gain insights into how the gameplay style has changed over time.

## Appendix

This assignment has been implemented using React, TypeScript, and the Nivo chart library.

You can follow these steps to run the assignment:

1. Clone the repository from <https://github.com/bams11/IV>.
2. Navigate to the "Arsenal" folder.
3. Run the command "npm install && yarn install" to install the necessary dependencies.
4. Run the command "npm start" to start the project.

Additionally, you can also view the demonstration using the link provided at the top.