VAR Optimizer

VAR Dashboard

The point of the dashboard is for a team to be able to estimate the chance of a VAR decision going in their favor. While this is intended to be used primarily in the event that VAR challenges are added to MLS, it could also be leveraged right now in the preparation for a match to determine what calls a referee is most likely to give in select circumstances.

- A decision "going in the team's favor" is defined as: a penalty being given, a goal being given, or a red card not being given
- Relevant Team: the team that will be awarded whatever the outcome of the decision is

The optimizer works by first searching through all VAR events from 2023, 2024, and 2025 in MLS that are cited on FotMob. It then selects each of the reviews that match the filters that are currently activated. It then calculates the proportion of those similar VAR reviews that went in the relevant team's favor. The optimizer is always from the relevant team's perspective. Again, this is the team that scored, won the penalty, or received the red card, for example.

The dashboard enables the user to filter on up to 9 different conditions to match their VAR review situation. While it's technically possible to select conditions for every filter, it's important to remember that the sample size gets smaller as the conditions become more specific, so there's a balance between matching the exact conditions, and having a broader sample space. The 9 filters are split into general and specific categories, with the general ones on the left and the specific ones on the right. The general ones each have a much larger sample size, and can more effectively be used in conjunction with each other. The specific ones provide much more information on the situation of the match, but each have less preceding examples to base analysis off of. The 4 general filters are:

- Review Type: in MLS, VAR reviews can be triggered for any of 4 match altering events: Goals, penalties, red cards, and mistaken identity. Since there have been no cases of mistaken identity in the last 3 seasons, it's not seen as an option in the filter. Review type is simply the event that triggered the review
- We are the Home Team: is the relevant team the home team? The filter is phrased like this since the optimizer is always from the relevant team's perspective. Again, this is the team that will be awarded the outcome of the review (goal/no goal, penalty/no penalty, red card/no red card)
- Gamestate: this is the state that the relevant team is in - winning, losing, or drawing - at the time of the review

- Off a Set Piece?: another true/false filter that asks if the event that triggered the review stemmed from a set piece or not (Ex. a headed goal off a corner that's being reviewed for a potential foul during the play)

The 5 specific filters are:

- Minute Bin: the game time line is split into seven 15 minute segments that you can select to match the time of the game that the review is happening in
- Referee: the head referee for the match
- Team For: the relevant team for the review
- Home Team/Away Team: the home and away teams in the match

Selecting from each of these filters will automatically update the percentage chance of getting the decision in the relevant team's favor. It displays Blank if there is a situation that's never been seen before in the data.

Referee Strictness: if a referee is selected, the dashboard will display their name and their "strictness" evaluation below the chance of the decision being overturned. The strictness is based on their tendency to award favorable decisions to the relevant team. Strictness is based on their overall percentage of favorable decisions. Adding more filters will narrow down to that ref's tendencies for those specific conditions, but the strictness is only calculated for their overall tendencies. The categories for strictness are:

- Strict: the ref has awarded 0-25% of VAR reviews in the relevant team's favor since 2023
- Average: the ref has awarded 26-50% of VAR reviews in the relevant team's favor since 2023
- Lenient: the ref has awarded 51-75% of VAR reviews in the relevant team's favor since 2023
- Generous: the ref has awarded 76-100% of VAR reviews in the relevant team's favor since 2023

Referee strictness offers an insight into the match referee's tendencies regardless of the conditions of the match/current situation.

VAR Dash Video 1

- This video displays an example case of a penalty review for the away team during a tied game in which the penalty was not won off a set piece. Notice how the chance of getting the decision updates with each condition that makes the situation more specific

<u>VAR Dash Video 2</u>

- This video displays the strictness ratings and overall chance of getting the decision from various different refs. The percentage displayed is the base percentage, as there are no other conditions selected. This is just to provide an example of the varying tendencies among different refs

<u>VAR Dash Video 3</u>

- This video simply shows a variety of different general and specific conditions being selected with each other. In this case, a referee remained selected the whole time to provide an example of what the different situations in matches under this ref look like

<u>FotMob Scraper</u>

Using selenium, the python script is able to scrape all the necessary VAR data from FotMob. The first thing the scraper does is gather all match urls for the given season. Since FotMob has a "Matches" tab for each season, it's easy to grab all the urls from this one place. By going to each page of the Matches tab, the scraper can gather the url for each game that season.

After the urls are compiled into a list, we can loop through each match url to gather the VAR data. The data is taken from the "Commentary" tab on each individual match page. After accessing the Commentary tab (which is just the match url with "tab=ticker" added on to the end), the loop uses the CSS selectors of specific elements on the page to know which items to pull. These items are: home team, away team, minute, var text, previous 1, 2, and 3, home score at time, away score at time, and referee. Each item is defined as follows:

- Home/Away team: the home and away team in that match. These are taken from the title of the webpage (which in this case is just the header at the top)
- Minute: the minute that the VAR review occurred (ended)
- Var Text: this is the description that FotMob provides next to each VAR event in the commentary. All these events are listed as "VAR Decision: …". The scraper looks through the commentary for the part that says Decision, and then scrapes all the text in that entry after it. This is what I use to figure out what type of review it is, who it's for, etc
- Previous 1, 2, 3: after the scraper finds a VAR event, it also scrapes the previous 3 entries in the commentary feed. This is used primarily to determine if the event that triggered the review was off a set piece or not

- Home/Away score at time: this is the score each team had before the decision of the VAR review. This is found by locating a VAR review and then going down the commentary feed until numbers appear. FotMob states the score in the commentary feed as "Home team X, Away team Y". This is the only time that any numeric values will appear in the feed, so we find the most recent mention of numbers and use that as the score. This is later used to derive gamestate for the teams
- Referee: the head referee for the match. The ref is displayed on FotMob for each match

After this, the scraping is done and some minor cleaning is needed to get the dataframe into its final, usable form. Details on these transformations are described in the python notebooks. The final data frame has columns:

- Home/Away Team: home and away team in the match
- Minute: minute the review ended
- Var_text: FotMob commentary's description of the VAR review
- Home/Away score at time: the score each team had when the review ended. If the decision was to award/take away a goal, the score at time reflects the score from before the goal that got reviewed. So if a goal was awarded after a VAR review to make it 2-1, the score at time for home/away is just 1-1
- Referee: the referee of the match
- Var_decision: the stripped version of var text that just contains the most relevant information (team and review type)
- Set Piece: true/false based on whether the event that triggered the review resulted from a set piece or not
- Winning: home/away/tied - describes which team is winning the match at the time of the review (or technically just before it). If the review is about a goal, this column displays the score from before that goal
- Team_For: the relevant team that the review is for
- For_Home: true/false if the relevant team is the home team
- Given: true/false if the penalty/goal/red card was awarded
- Good: true/false. Good is defined as getting awarded a goal, getting awarded a penalty, or not getting a red card
- GS_For: the game state of the relevant team
- Season: the season that the match took place in
- Minute_bin: the 15 minute interval that the end of the review fell into
- Review_type: goal/penalty/red card