Referee Strictness (Placeholder)

Your Name(s)

December 8, 2025

1 Introduction

In the ecosystem of elite football, UEFA competitions represent the pinnacle of the sport. Referees selected for these tournaments, the Champions League, Europa League, and Conference League, are tasked with managing games involving diverse playing styles, high cultural stakes, and immense pressure. A single decision in these continental matchups can determine not just the outcome of a match, but the financial and historical trajectory of a club. However, despite the elite status of these officials, the discourse surrounding their performance often remains subjective. Fans and pundits frequently debate whether referees adjust their strictness based on the prestige of the tournament: are they more lenient in the "showcase" Champions League compared to the lower-tier Conference League?

The motivation for this project stems from the difficulty in validating these perceptions. While raw disciplinary statistics exist, they are rarely presented in a way that allows for a fair comparison across the three distinct tiers of UEFA competition. A simple count of yellow cards is insufficient to judge a referee's performance without context; one must account for the number of appearances and the specific tournament level to understand the referee's true "personality." Furthermore, varying nationalities of referees add another layer of complexity, do officials from certain regions enforce the rules more strictly when performing on the European stage?

To address this, our project develops an interactive visualization system that explores refereeing behaviors across the three major UEFA competitions spanning the 2021-2025 seasons. By synthesizing raw data (cards, penalties, appearances) into a derived "Strictness Score," we aim to quantify officiating patterns that are usually left to anecdote.

The goal of this visualization is to allow users to investigate the consistency of referees across the UEFA hierarchy. It enables the exploration of key questions: Do specific referees become "stricter" as they move down the tiers from the Champions League to the Conference League? Are there visible correlations between a referee's age or nationality and their likelihood to issue cards? By visualizing these metrics, we provide an objective, analytical lens on the men in the middle of Europe's biggest matches.

2 The Visualization Task

Goal: Define the data and the tasks formally (likely using Munzner's What-Why-How framework if it was covered in your lectures).

To design an effective visualization, we first analyzed the problem domain by abstracting the data into types and defining the abstract tasks the user needs to perform.

2.1 Data Abstraction (Lecture 5)

The underlying dataset is a tabular collection of disciplinary statistics for football referees officiating in UEFA competitions (Champions League, Europa League, and Conference League) across four seasons (2021–2025). The data consists of Items (individual referee records per season/league) and Attributes (characteristics of those records).

Categorical (Nominal)

- Name (Identifier).

- Nationality (e.g., "English", "German").

- League (The context of the Performance: UCL, UEL, UECL).

Quantitative (Ratio)

- Age (Sequential data allowing for correlation analysis).

- Appearances, Yellow Cards, Double Yellow Cards, Red Cards, Penalties (Counts start from zero).

Derived Attribute: The Strictness Score

Raw counts of cards are misleading without context (e.g., 5 red cards in 20 games is different from 5 red cards in 5 games). To solve this, we derived a quantitative attribute called Strictness Score. This provides a normalized single metric to rank officials.

The formula is defined as a weighted sum of disciplinary actions divided by the number of appearances:

$$\text{Strictness} = \frac{(W_y \times \text{Yellows}) + (W_{dy} \times \text{DoubleYellow}) + (W_r \times \text{Reds}) + (W_p \times \text{Penalties})}{\text{Appearances}}$$

(Note: In our implementation, we assigned default weights based on the severity of the punishment, e.g., Yellow=1, DoubleYellow=3, Red=5, Penalty=3. However they are user adjustable.)

2.2   Task Abstraction

The visualization is designed to support three primary high-level goals: Discover (finding unknown patterns), Derive (generating the strictness metric), and Compare (evaluating differences).

We broke these down into specific domain tasks:

1. **Summarize & Overview (Global Analysis):**

   - Abstract Task: Summarize the distribution of strictness across the entire dataset.
   - User Goal: The user needs to see the "All 3 Leagues" view to understand the general baseline of officiating in Europe before drilling down. This answers the question: "What is the average strictness of a European referee regardless of the tournament?"

2. **Compare Distributions:**

   - Abstract Task: Compare attributes between derived clusters.
   - User Goal: Compare the strictness levels between the Champions League and the Conference League. Are referees significantly harsher in lower-tier competitions?

3. **Identify Extremes (Outliers):**

   - Abstract Task: Identify items with extreme values.
   - User Goal: Spot the single "strictest" referee in the dataset or find a referee who gave an unusually high number of penalties in a single season.

4. **Correlate Attributes:**

   - Abstract Task: Lookup and Correlate quantitative values.
   - User Goal: Determine if Age correlates with Strictness. (e.g., "Do referees mellow out as they get older?").

By explicitly supporting the transition from a Summarized View (all leagues) to a Filtered View (specific league), the tool allows the user to verify if global trends hold true in local contexts

3   Related Work

Goal: show that we did our homework

- **Existing Solutions:** Briefly mention if there are other football stats dashboards (e.g., Transfermarkt, FBref).

- **The Gap:** Explain why those aren't enough. usually, they show raw tables (hard to read) or focus on players rather than referees. They might lack the cross-league comparison or the specific "Strictness" metric you calculated.

- **Academic/Visual Inspiration: Did you see a specific scatterplot or dashboard in a paper or the "Student Projects" gallery that inspired your layout? Cite it here.**

3.1   Design and Task References

- Reference articles and papers that support the visualization **task**.

- Reference **design references** or existing visualizations that influenced your approach.

- Mention any articles or world events that are highlighted by the data.

4   The Visualization Solution

Our solution is implemented as a single-page React application using D3.js for rendering complex vector graphics and Material UI for the control interface. The design follows Shneiderman's Information Seeking Mantra: "Overview first, zoom and filter, then details-on-demand."

4.1   Overview

Briefly describe the layout (e.g., "A dashboard consisting of a main scatterplot, a filtering sidebar, and a details-on-demand panel").

The dashboard is divided into three logical zones:

1. **Global Controls (Top Bar):** A persistent control panel allows users to filter the dataset by **Season** (2021-2025), **Nationality**, **Referee Name**, **Age Range**, and **Experience (Appearances)**. A global settings dialog allows users to adjust the weights of the "Strictness Formula" dynamically, empowering them to define what "strict" means to them (e.g., weighting red cards higher than penalties).

2. **Macroscopic Views (The Overview):** The upper section features two coordinated views:

   - **Geographic Strictness Map:** A choropleth map showing average strictness per country.
   - **Population Scatterplot:** A distribution plot placing every referee on an Age/Experience vs. Strictness plane.

3. **Microscopic/Comparative Views (Details-on-Demand):** The lower section adapts based on the user's selection:

   - No Selection: Shows general trends.
   - Multi-Selection: Activates the **League Comparison** Scatter (League X vs. League Y).
   - Single Selection: Activates the **Referee Passport** (Heatmap + Waffle Chart) for deep profiling.

4.2 Design Decisions & Justification

We grounded our design choices in visualization theory, specifically focusing on effective mark and channel usage.

A. **Strictness Encoding (The "Unifying Thread")** Since "Strictness" is a diverging quantitative attribute (ranging from "Lenient" to "Strict"), we consistently encoded it using a Diverging Color Scale (Blue → Yellow → Red).

- Justification: Blue represents leniency (cool/calm), while Red represents strictness (danger/cards). This semantic mapping reduces the cognitive load for football fans who already associate red with punishment.

B. **The Geographic View**

- **Encoding:** We used a choropleth map where countries are colored by the average strictness of their referees.
- **Interaction:** We implemented Linked Filtering. Clicking a country (e.g., Spain) acts as a global filter, instantly updating the scatterplots to show only Spanish referees.
- **Handling Sparsity:** To avoid misleading averages from countries with only 1 referee, we added a "Minimum Referees" slider and a "Region Mode" toggle (aggregating countries into regions like "Nordic" or "Balkans"). This addresses the Modifiable Areal Unit Problem (MAUP).

C. **The Correlation View**

- Marks & Channels: Each referee is a point.
  - **X-Position:** Mapped to Experience (Appearances) or Age (user-toggled). We chose position because it is the strongest channel for quantitative comparison.
  - **Y-Position:** Mapped to Strictness Index.
  - **Color:** Mapped to Strictness (redundant encoding to reinforce the pattern).
- **Occlusion Handling:** Non-selected referees are rendered with low opacity (0.3), while selected referees use high opacity (1.0) and a stroke border. This "Focus+Context" technique ensures the user can see their selection within the global distribution.

D. **The Cross-League Comparison**

- **Problem:** Users want to know if a referee changes their behavior between leagues (e.g., "Are they stricter in the Champions League than the Europa League?").
- **Solution:** A scatterplot where X = League A Strictness and Y = League B Strictness.
- **Reference Line:** We added a $y = x$ diagonal line. Points above the line indicate referees who are stricter in League B; points below are stricter in League A. This geometric aid allows for instant assessment of bias.

E. **Single Referee Profiling** When a single referee is selected, we provide a "Visual Identity Card":

(a) **The "Passport" (Heatmap):** A matrix view (Rows: Competitions, Columns: Seasons) colored by strictness. This reveals temporal trends (e.g., "They were lenient in 2021 but strict in 2023").

(b) **The "Card DNA" (Waffle Chart):** A 10 x 10 grid representing the percentage split of disciplinary actions (Yellows vs. Reds vs. Double Yellows).

- Justification: We chose a Waffle Chart over a Pie Chart because square grids are easier for the human eye to compare in terms of area and proportions.

5 Use Case

Goal: Walk the reader through a "story" to prove the tool works.

- Scenario: "Let's assume a user wants to find out if English referees are stricter when refereeing in Europe compared to the Premier League."

- Step 1: The user selects "Nationality: English" from the dropdown.

- Step 2: The user looks at the "League" distribution.

- Insight: "The visualization reveals that Referee X has a strictness score of 4.5 in the Europa League but only 2.0 in the Premier League."

- Conclusion: Include a screenshot of your visualization showing this specific state. This is crucial—show, don't just tell.

6 Limitations

Goal: Be honest about what is missing or could be better.

- **Data Limitations:** "Our data does not account for the severity of the foul, only the card given. A tactical yellow card looks the same as a violent yellow card in our data."

- **Visual Limitations:** "When all seasons are selected, the scatterplot suffers from occlusion (too many dots overlapping)."

- **Future Work:** "If we had more time, we would add a temporal view to show how a single referee's strictness changes over their career."

7 Division of Labor

Acknowledge all contributors and their specific roles.

Table 1: Project Contribution Matrix

Task	Contributor 1	Contributor 2	Contributor 3	Contributor 4
Visualization Design (drawings/discussion)	X		X	
Finding data	X	X		
Data cleaning and prep		X		
Coding the visualization	X	X	X	X
Fielding feedback		X	X	
Completing the report	X		X	

8 Bibliography

References