Fanny Missillier - 346565 - fanny.missillier@epfl.ch
Rodrigue de Guerre - 327683 - rodrigue.deguerre@epfl.ch
Zacharie Bourlard - 329920 - zacharie.bourlard@epfl.ch
Team-GB

Data Visualisation (COM - 480)
Process Book

'The Shape of Strategy: A Data-Driven Tour into Chess Games'

Spring Semester 2025
Ecole Polytechnique Fédérale de Lausanne

I. Introduction

Chess is a game of strategy, patience, and precision, but it is also a game rich in data. With the rise of online platforms such as Lichess.com or Chess.com, millions of games are recorded every day, offering a unique opportunity to analyze chess not only as a game, but as a dynamic system of patterns, habits, and decisions.

This project, developed as part of **EPFL's Data Visualization course** (COM-480), aims to explore a dataset of 22,000 chess games played on Lichess. Our objective is to make this dataset accessible and insightful through interactive visualizations. Our main goal is to make our analyses appealing to both casual and experienced players.

The website we built provides a multi-faceted interface to navigate the chess world. It highlights the role of individual pieces on the board, trends in players' behavior and the main differences in strategies. Using tools such as D3.js and Node.js, we designed a responsive and educational platform that blends statistical analysis with intuitive visual storytelling.

Our work is guided by a central question:

How can we structure and visualize chess data to offer meaningful insights to players of all levels?

II. Methodology

Our project started with a dataset selection phase. Each team member proposed three potential datasets, and we collectively narrowed the list down to two: one based on global air traffic data, and another focused on chess games played on Lichess. We began exploring both in parallel, but we soon realized that the chess dataset had greater potential to support playful, engaging, and widely accessible visualizations. It also offered more room for creative storytelling, which aligned well with the goals of the Data Visualization course.

Once we had selected the chess dataset, we dedicated time to cleaning and formatting the data. This step was crucial, as several fields were not analysis-ready. One notable issue was the format of the start and end times of the games: they were stored as large numerical values in scientific notation, which represent Unix epoch timestamps in milliseconds. To use them effectively, we had to convert them into standard datetime objects. This transformation was done using Python, which allowed us to extract meaningful time features (such as hour of the day, duration in minutes, etc.).

After cleaning, we performed an exploratory analysis in a Jupyter notebook to identify promising directions for our visualizations. This helped us decide which questions we wanted to answer. We then translated our ideas into rough sketches and diagrams, many of which were refined and formalized during Milestone 3. These visual drafts served as a skeleton for the structure of the website and helped the team coordinate development efforts efficiently.

For the implementation phase, we collaborated through a shared GitHub repository. We did a simple backend in Nodejs + express to serve our files. We deployed the image in a docker container on a vps. Each visualization was integrated progressively and connected to the cleaned data, with careful attention paid to usability, responsiveness, and performance.

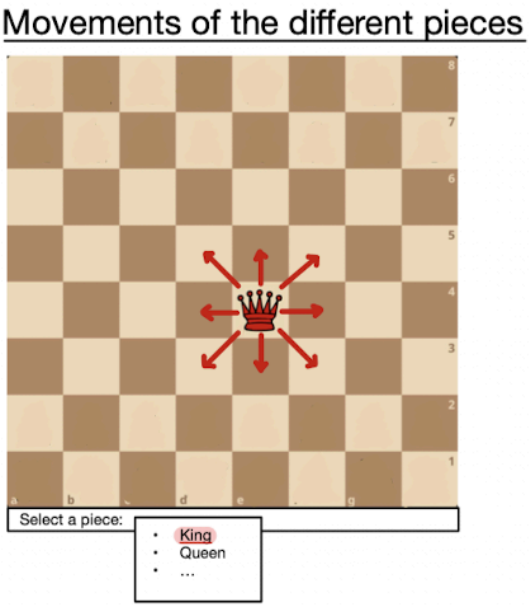
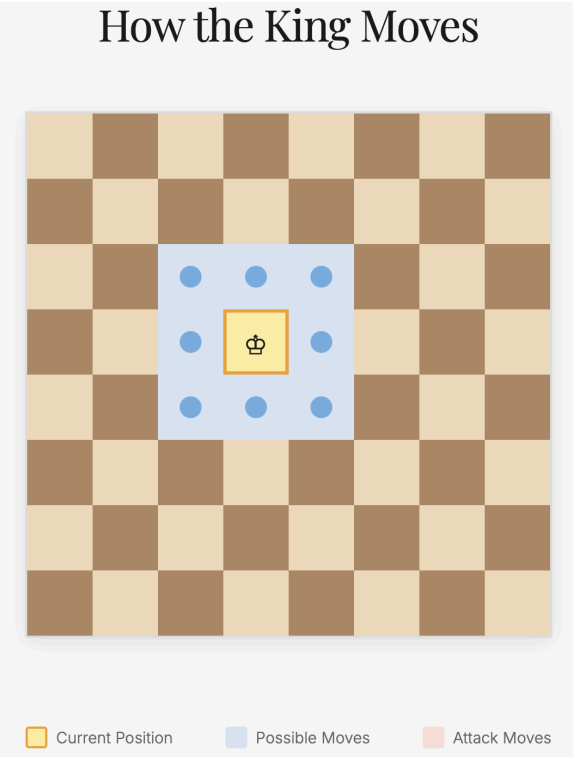
The website is available here: https://www.team-gb.org/

III. Design choices

To make our website both engaging and accessible to a wide audience, we carefully considered how to structure the user experience around progressive layers of complexity. Rather than overwhelming users with dense statistics or advanced chess theory from the outset, we designed the interface to guide them gradually from basic concepts to more in-depth analytical tools. This section presents the key design choices that shaped the three main pillars of our website:

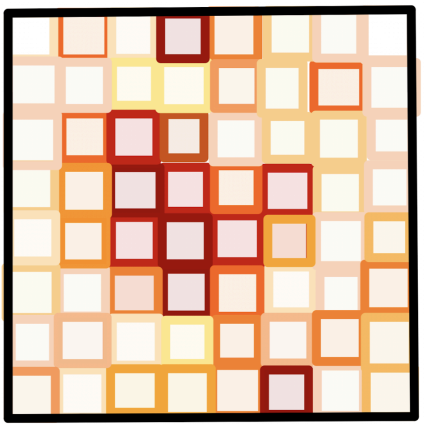
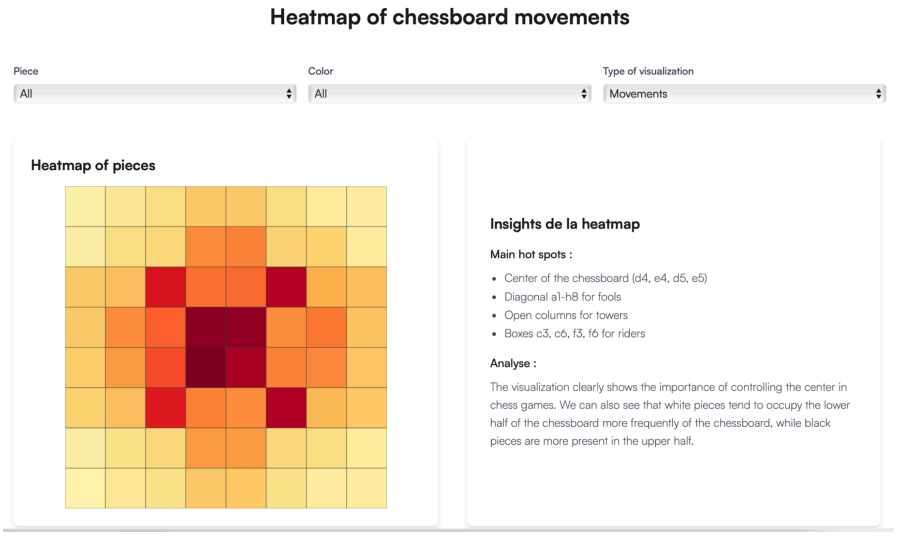
A. **The Pieces Section** – designed primarily for beginners, to explore how different pieces are used across thousands of games in a playful and visual way.
B. **The Game Analysis Section** – offering intermediate to advanced users tools to analyze common patterns, strategic moves, and opening dynamics across games.
C. **The Player Analysis Section** – allowing users to explore personalized statistics from our dataset by selecting individual players.

A. The Pieces Section

The idea	The outcome
	

We are very satisfied with the result of the Pieces section. Overall, we stayed true to the original spirit of what we wanted to develop. If you visit the website, you'll notice that we chose to let users navigate between the different pieces using arrows, as we found this solution more convenient and visually appealing than a dropdown menu. In addition, we added descriptive texts for each piece to make the section more informative and accessible.

B.The Game Analysis Section

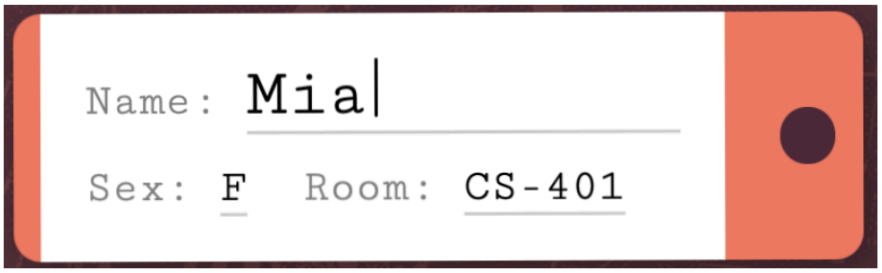
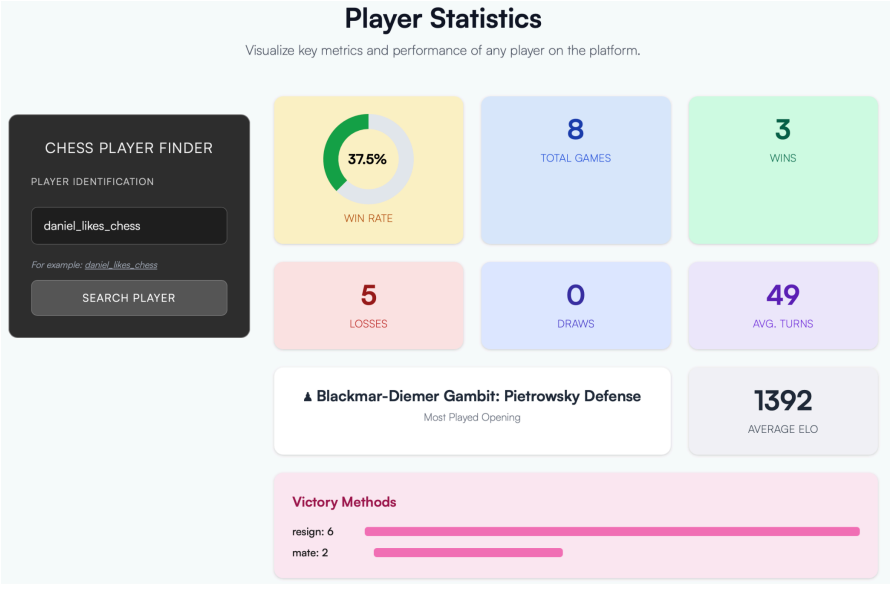
<table>
<tr>
<td>The inspiration</td>
<td>Heat map of movements on the chessboard
  </td>
</tr>
<tr>
<td>The outcome</td>
<td> 
</td>
</tr>
</table>

Regarding the Game Analysis section, we had several interesting sketches, and most of them are reflected in visual implementations on our website. However, we chose to highlight the

heatmap in our report, as we see it as one of our most successful outcomes, a clear example of how a simple idea can evolve and mature into a highly satisfying final result.

We originally set out to visualize the most frequent positions of certain pieces on the board. Building on that foundation, we decided to extend the concept by adding multiple dropdown menus, allowing users to refine the analysis even further. In the final version, users can select the piece type, its color, and the specific context in which it appears on a square — whether it's part of a **capture**, a **check**, or an **opening move**. This added depth makes the heatmap particularly valuable for more advanced players who wish to dive deeper into strategic patterns.

C. The Player Analysis Section

The inspiration	
The outcome	

Finally, one of the sections we cared most about was the Player Analysis. Our goal was to create something both visually appealing and user-friendly, bringing together all relevant

information in a compact and accessible format. The idea was to simplify the user experience by avoiding the need to search through multiple sections to find key insights about a player.

We initially faced some challenges in organizing the layout to keep everything clear and concise, but after some iteration, we managed to achieve a design that works well and aligns with our vision.

IV. Conclusion

A. Key takeaways

We hope this report has helped clarify our overall approach and shed light on how we arrived at the final version of our website. This project was not only an opportunity to strengthen our data analysis skills, but also, and perhaps more importantly, to reflect on how we want data to be presented, and for whom. Conducting high-quality analyses is essential, but they lose all their value if they are not effectively communicated to the audience. Learning how to present data clearly and meaningfully is a skill that will be immensely valuable to us in the future.

Looking back, there are several ideas we would have liked to explore, but either thought of too late or didn't have time to develop fully. For example, we wanted to go deeper into the analysis of openings, as they play a key role in shaping the game. We also had the idea to highlight some of the most remarkable players from our dataset, which could have made the site more engaging for chess fans.

The final result is quite minimalist, we would have liked to add more creativity and animation to make the design more dynamic and visually rich.

B. Peer assessment

Each member of our team made significant progress across multiple domains. During Milestones 1 and 2, we worked closely together to brainstorm ideas and explore directions that excited everyone. This collaborative phase was key to aligning our goals and identifying the narrative structure we wanted for the final site.

For Milestone 3, the workload was more distributed. Fanny contributed heavily to sketching out and refining the visual structure of the site to ensure we all had a clear roadmap. She also handled several parts of the site implementation, especially the Player Analysis section. Rodrigue focused on building the page structure and overall web architecture, securing and configuring the hosting server, and leading the analysis of global trends on the website. Zacharie took charge of the data preprocessing, contributed to the website development on the game overview part, and led the redaction of the written report.

We are all very satisfied with the final result. Our collaboration worked smoothly, and every contribution was essential to shaping a coherent, informative, and visually engaging exploration of chess through data.