

Verslag datavisualisatie

Keuze dataset

Wij hebben een dataset (<https://www.kaggle.com/datasets/arevel/chess-games/data>) met ~2.5M schaakmatches gespeeld op online schaakplatform Lichess. Naast de gebruikelijke velden als movements en resultaat van games bevat deze dataset ook de ELO-ratings van de spelers en de openings. Zo kunnen wij kijken naar de populariteit van bepaalde openings binnen bepaalde ELO-gebieden en hun slaagkans, alsook invloed van de opening op de ELO-rating van de spelers zelf.

Brainstorm mogelijke visualisaties

- **Openings and Strategies**

- **Heatmap:** Distribution and winning chances of opening moves by popularity. In the shape of a chess board
- **Heatmap:** Winning chances based on different openings.
- **Interactive Feature:** Users can input an opening move and see its win/loss rate.

- **The legends of chess**

- **Table & Bar Chart:** Most successful players and their win rates.
- **Scatter Plot:** Relationship between Elo rating and win percentage.
- **Feature:** Click on a player to see a brief bio and their best games.

- **White vs black**

- **Win Percentage Chart:** Performance of White vs. Black distributed over time.
- **Stacked Bar Chart:** Common openings for White and Black and their outcomes.
- **Tooltip Feature:** Hovering over bars shows historical match examples.

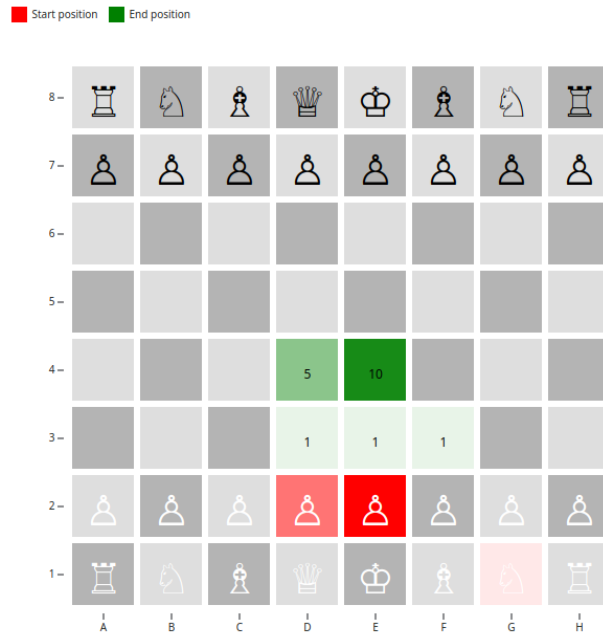
- **Flow of a chess game**

- **Histogram or Box plot:** Game length distribution of all the games
- **Line Graph or Heatmap:** Move Complexity over time during the game
- **Scatter Plot or Timeline Visualization:** Checkmate & draw moments

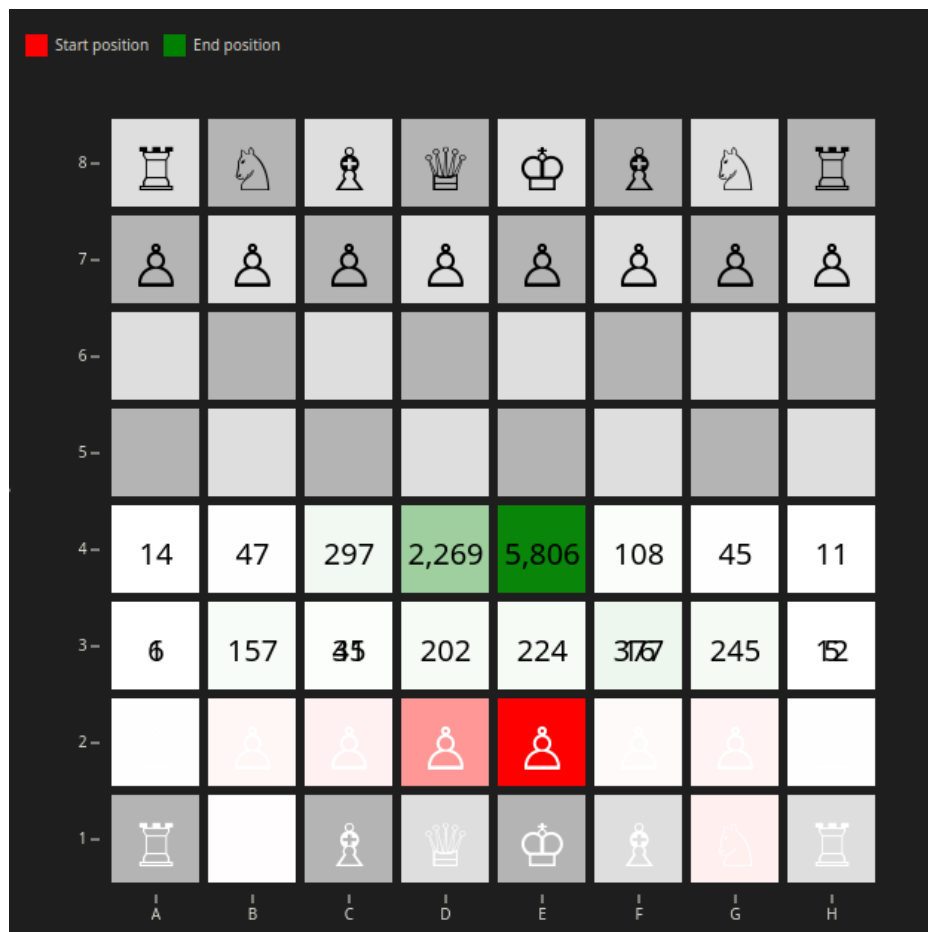
Schaakbord visualisatie

Uiteindelijk hebben wij gekozen voor kwaliteit in plaats van kwantiteit van de visualisaties. Wij hebben de focus gelegd op visualiseren van de meest populaire zetten.

We zijn begonnen met het maken van de heatmap voor de meest populaire openingszetten. Hiervoor hadden we het idee om dit in de vorm van een schaakbord te doen. Hiervoor hadden we een statisch schaakbord gemaakt.



In de figuur hierboven ziet men beperkte data omdat wij met beperkte data werkten tijdens ontwikkeling. Er zijn uiteraard meer dan 18 matches gespeeld.

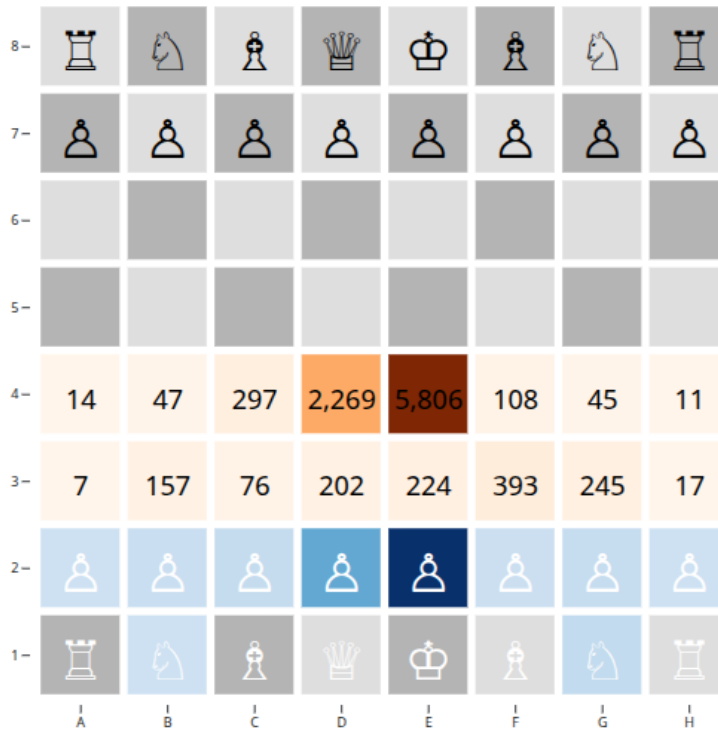


Een volgend idee dat we hadden was om dan een zet te kunnen doen en alle opeenvolgende tegenzetten te kunnen doen en dit voor de eerste vijf zetten voor elke speler.

Start position move count



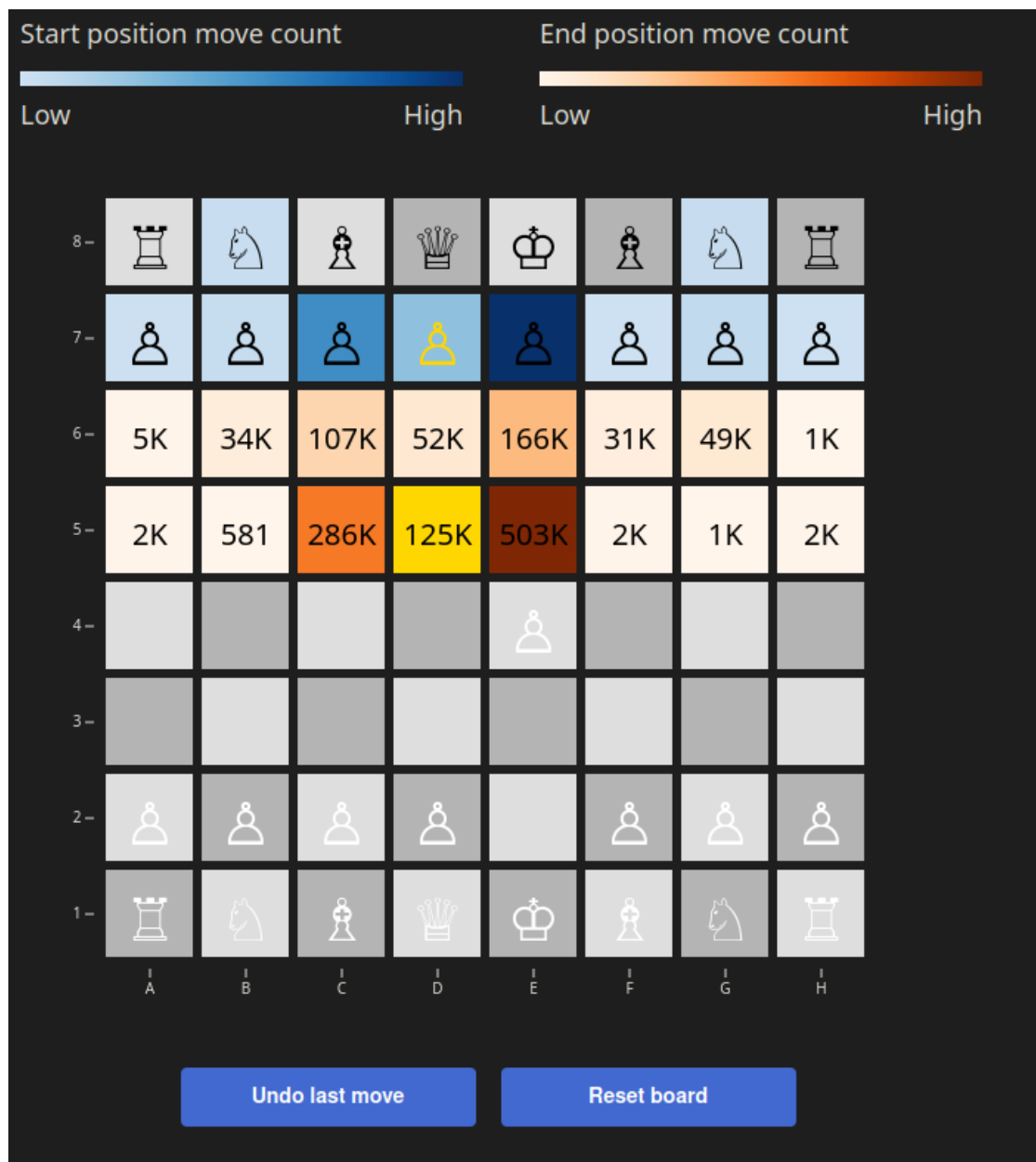
End position move count



Undo last move

Reset board

Na de feedback hebben we de kleuren aangepast van de visualisatie en twee knoppen toegevoegd zodat het bord gereset kan worden of een stap kan terug gezet worden. Ook is de legende verbeterd zodat deze duidelijk zichtbaar is.



Om deze visualisatie naar de volgende niveau te brengen hebben wij Stockfish geïntegreerd in onze visualisatie om de beste zet in de positie te tonen. De visualisatie toont nog steeds de meest populaire zetten, met de beste zet volgens Stockfish 17 getoond in goud. Wij zien hier ook direct dat de meestgekozen zet niet altijd het beste is.

Taakverdeling

We hebben aan alles redelijk door elkaar zitten werken en elkaar geholpen indien we vast zaten, maar de verdeling is ongeveer:

Jozef:

- Data (pre-)processing
- Tekst rond de visualisaties geschreven
- Integratie van Stockfish

Wouter:

- Schaakbord eerste lay-out
- Schaak logica en zetten
- Lay-out, kleuren en knoppen aanpassen