

Top Chess Players Interactive Visualization

Karthik Subramanian

Intro: Chess has recently undergone a surge in popularity. Many high-ranked players have garnered large followings and have inspired many to improve their chess skills and even pursue the game professionally. Many avid fans and chess enthusiasts want to track who the top players of the last decade has been and also analyze if certain factors generally lead to superior game strength. Questions that are asked to conduct this analysis include: “Which players have dominated the game over the years?”, “Which country outputs the most successful players?”, and “Are age and game strength correlated?”.

Data Domain: There is a dataset titled “Top 100 Chess Players Historical” (<https://www.kaggle.com/datasets/odartey/top-chess-players>) that includes many variables that can help answer these questions, including: name, rank, country, rating (the metric for a player’s skill level), ranking_date, and birth_year (from which age is derived). This dataset includes data from the years 2000 to 2017.

Storyboards:



Figure 1: Line Chart

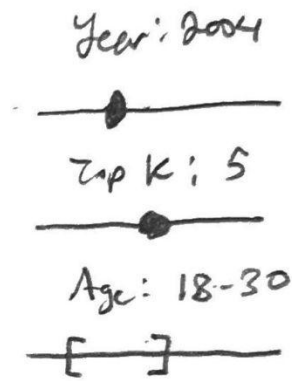


Figure 2: Filters

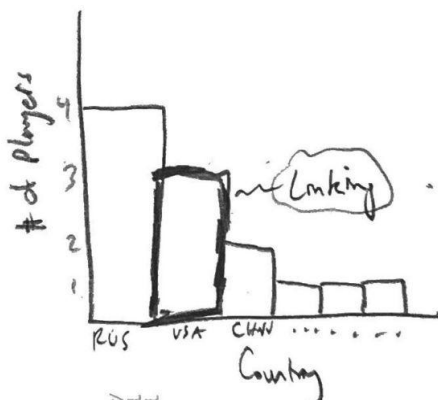


Figure 3: Countries Bar Chart

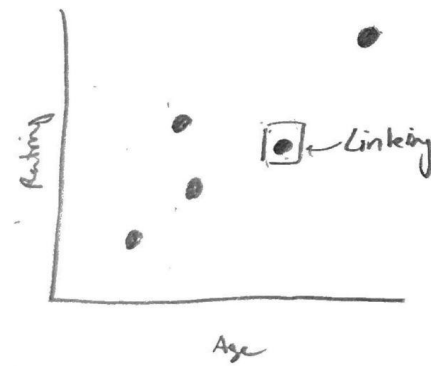


Figure 4: Age vs. Rating Scatter Plot

Final Interactive Visualization:

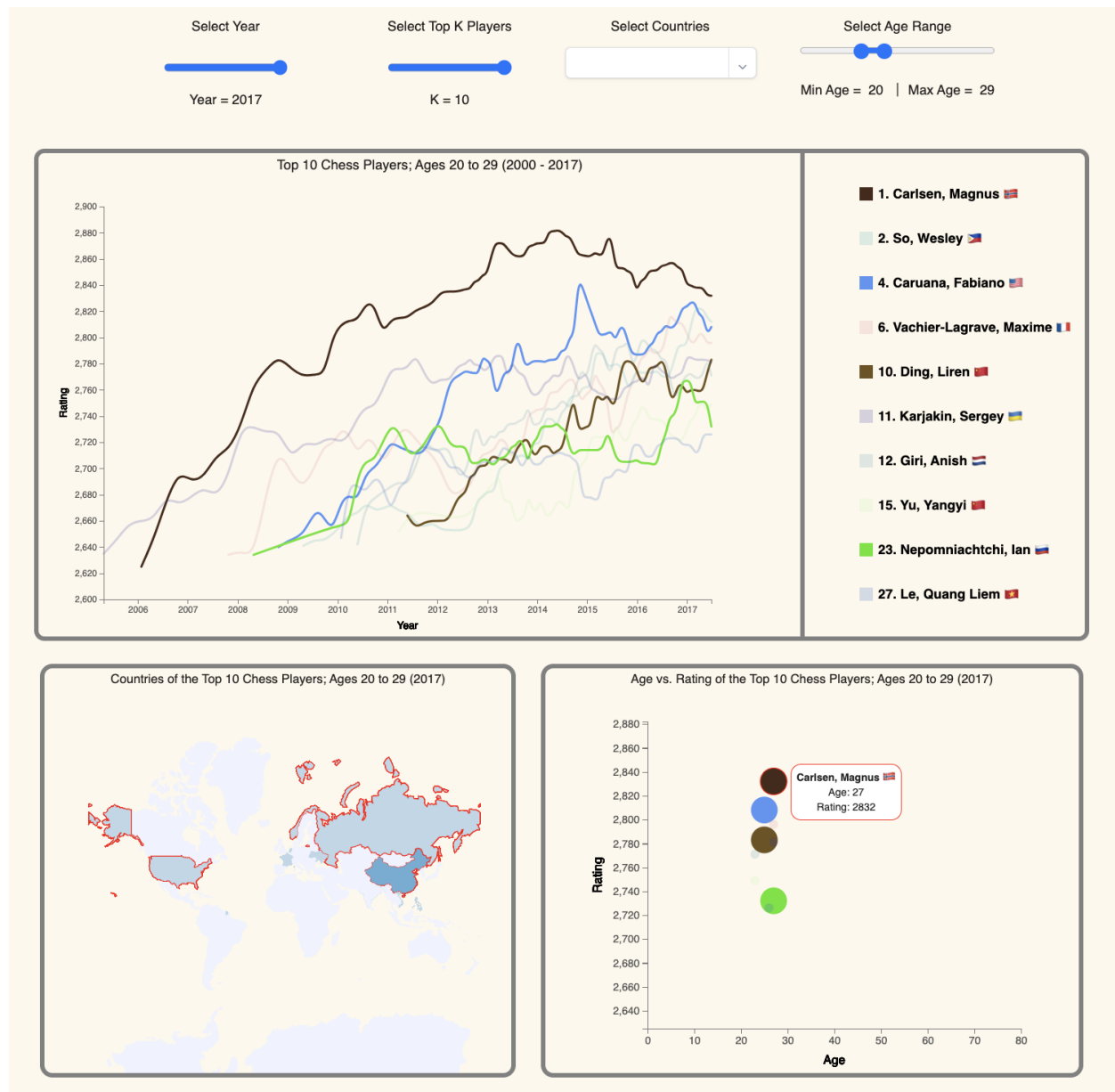


Figure 5: Final Interactive Visualization

Description: At the top of the visualization are filters that the user can toggle to modify the visualization to match the queries. These filters are to select the countries that are represented in the data, the age range of players represented, the year for which they want to see the rankings, and the number of players they want to be shown (up to 10).

The primary cell of the visualization is the time-series line chart that plots the top K players (I will be using “top K” for this report since K can be toggled) up until the current year. The x-axis represents the date and the y-axis represents the rating. To the right of the line chart is a legend

that specifies which player corresponds to which color. Each name has a unique color which is used in both the legend and the scatter plot. In addition, the order of the labels in the legend shows the relative order of ranking of the queried players for the chosen year. The absolute ranking among all players that year is written before each name too. Whenever the year slider changes, the order of the legend, as well as the lines on the chart change.

Underneath the line chart and legend are two different displays - one is a choropleth map and the other is a scatter plot. The choropleth map presents the number of distinct players a given country has produced at the given year. The darker the color (blue in this case), the more players that country had produced. The scatter plot maps age (at the current year) on the x-axis and rating on the y-axis, helping to answer the question of if age and game strength are correlated. Once again, the scatter plot is color coded to the corresponding player name.

Interactive Visualization Techniques: A user can modify the filters to transform the visualization, which utilizes the dynamic queries technique, since the data displayed changes based on filters. Then, users can select a player by clicking on the colored rectangle in the legend or a dot on the scatterplot. Using the technique of linking, when a player on the legend or dot on the scatterplot is selected three outcomes occur: 1) the line corresponding to this player on the line chart is highlighted while the rest of the lines drop in opacity, 2) the country to which this player belongs to is outlined in red, 3) the point on the scatter plot corresponding to this player is highlighted while the rest of the dots drop in opacity. A user can select multiple players on the legend and each of these three effects occur. A user can also select a country on the choropleth map and subsequently have all players from that country highlighted via linking as well. Additionally, tooltips are provided when the user hovers over both the map and the scatter plot cells. On the map, a user can hover over a particular country and read the name of the country and the number of distinct players from that country that made it to the top K players up until that year on the tooltip. On the scatterplot, a user can hover over each dot and read the player's name, age (at the current year) and their rating.

Changes from Storyboards: The original ideas for the line chart includes many features on the final version, but also a few features that did not make it. One such example is expanding the thickness of the line to indicate a player's experience/the number of times they appeared on the top K rankings (Figure 1). In reality, this visual encoding would be hard to decipher, already given the dense nature of the line chart, and was instead replaced by a number to the side of the player on the legend. The bar graph (Figure 3) was removed in favor of the choropleth map as this provided greater interactivity for the user and allowed them to investigate anomalies, geographic patterns, and search for their desired country. Finally, two last ideas removed were the ability to click on the line itself to trigger linking, and positioning the label next to the y-position of the corresponding line on the line chart (Figure 1). Both these features were coded, but in testing, I found the UX to be clunky and decided on the current format (only clicking on the legend rectangles and/or scatterplot dots for linking and having a fixed; spaced out legend).

Apart from that, many more features were added during development of the final interactive visualization, as the effect on user experience became clearer. On the line chart legend, the

absolute rankings are provided alongside the name of the player. That way, if the user queries by age or country, they are still aware of where the players land among all other players, while still being able to view the relative rankings between the players queried. Also, the choropleth map and all its interactive functions (mouseover, mouseout) were novel, as well as player highlighting from any of the cells. That is, users can now click on a scatter plot dot to highlight and link a player, or click on a country to highlight all players from that country. The ability to click on and highlight multiple countries/players on the legend was also demonstrated on the final version. Tooltips were another major feature that was not included on the storyboards but were implemented on both the choropleth map and scatter plot. Finally, filtering the data (dynamic query) by countries was also added. The scatter plot (Figure 4) remains largely the same, except for how the dots are highlighted, as they grow when selected in the final version.

Final Commentary: While learning and adapting to D3 was difficult, the hardest part of development was aggregating, manipulating, and attaching data to my svg's. Understanding how to utilize `enter()`, and `exit()` to properly alter the visualization, and even just tracking with components store to what data required a lot of care and debugging. In total, I spent over a week and a half learning D3, ideating, and developing, spending over 20 hours over that time span. Since I worked alone, I completed all the work myself.