Data Visualization Project – Milestone 2

Tennis Through Time: Visualizing Performance, Upsets, and Trends (2000–2024)
**Group:** SmashData

Project Goal

Our project aims to explore 24 years of professional tennis data through dynamic, game-inspired visualizations. Taking inspiration from fighting games like *Street Fighter*, we envision an interface where users can select two players as if they were combatants, instantly comparing their stats, ELO progression, and head-to-head records. This playful, competitive design makes data exploration feel intuitive and interactive.

To support this duel-based interface and provide a standardized way to compare players over time and across match contexts, we introduced **ELO ratings**. The ELO system offers a dynamic and interpretable measure of a player's relative strength, evolving with each match and reflecting the competitiveness of each encounter. It is particularly well-suited for visual storytelling and rivalry comparisons.

The main objective is to surface key patterns in performance, surfaces, and upsets—allowing users to dive into player trajectories and match histories without needing deep tennis expertise. The final experience will be delivered as an interactive website, combining Plotly for visualizations with HTML and JavaScript for seamless interaction and animation.

## 1. Player Performance & Head-to-Head Comparison (2000–2024)
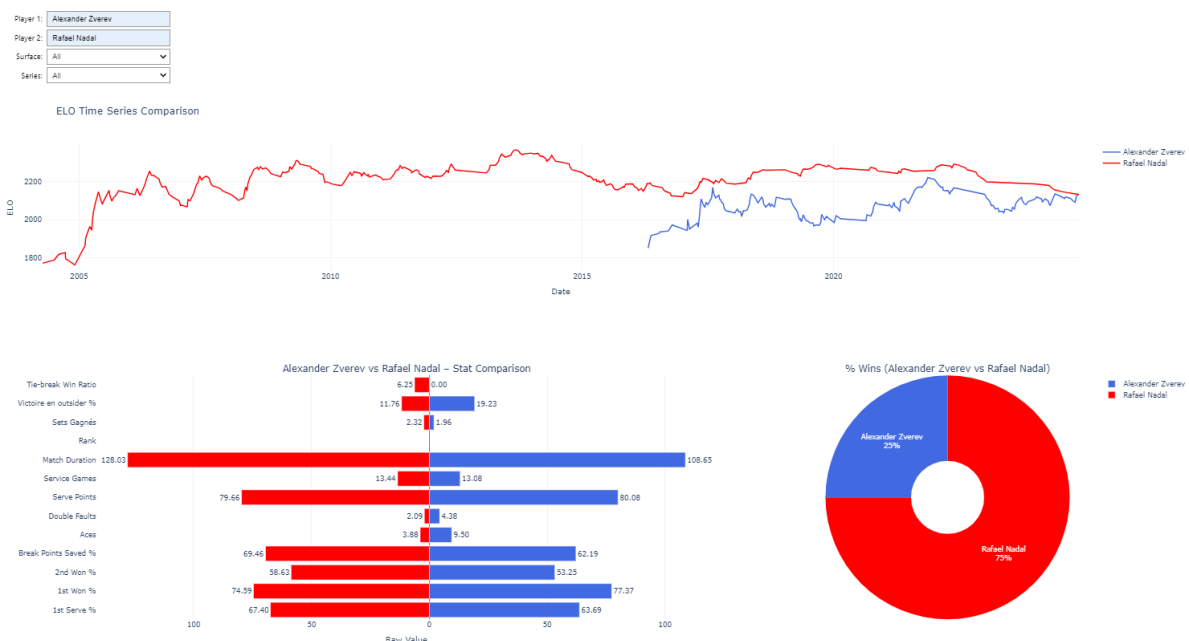
This visualization offers an interactive showdown between two of the top tennis players from the Open Era, modeled in the style of a fighting game character selection screen. On the top, the animated ELO timeline acts as a health bar, revealing each player's performance

trajectory year after year, with the option to filter by surface and tournament series for a more targeted rivalry.

Below, the duel continues through a split horizontal bar chart, comparing detailed match statistics such as serve efficiency, match duration, tie-break win rate, and more. Each stat becomes a power indicator, letting you assess player strengths and weaknesses at a glance—just like comparing fighters' attributes in a game.

To complete the face-off, a win percentage donut chart summarizes the direct encounters between the two selected players. Whether you're analyzing historical dominance or crafting your dream match-up, this feature-packed panel brings a gamified feel to tennis data exploration.

## 2. Evolution of Elo Ratings for Top 10 (2000–Present)
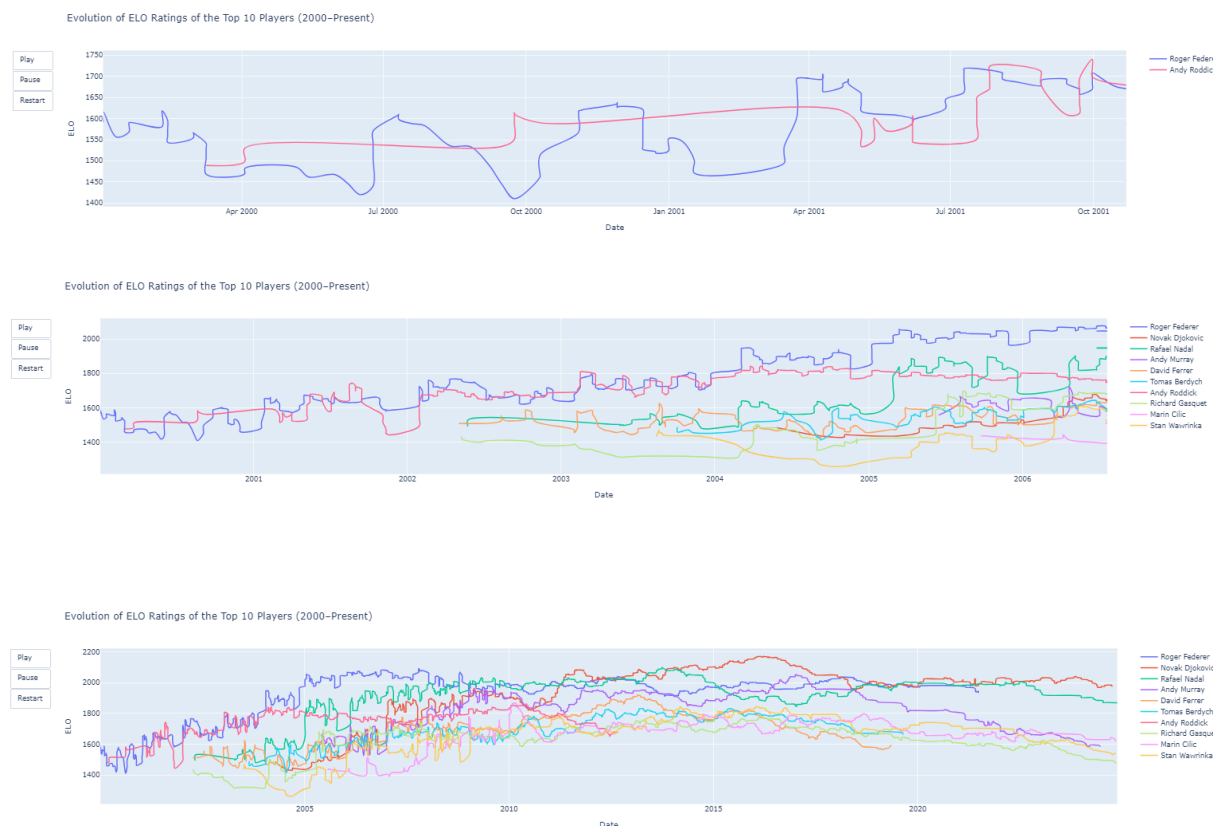
an animated timeline of player performance across two decades of professional tennis.

This dynamic visualization showcases the evolution of ELO ratings for the ten most dominant male tennis players from 2000 to the present. As the timeline progresses, each player's curve appears and grows organically, highlighting their rise, peak, and decline over time—mirroring the flow of their careers on tour.

The animated nature of the graph gives a strong videoludic aspect, reminiscent of in-game progress bars or ranked ladders in competitive games. Just like watching fighters level up, viewers can track momentum shifts, rivalries, and eras of dominance.

The legend updates in sync with the plot, reinforcing the visual story by gradually revealing player names alongside their color-coded trajectories. This visualization invites exploration of tennis history through a game-like lens, making data feel alive and engaging.

3. Animated Bar Chart – Top 10 Match Winners per Year
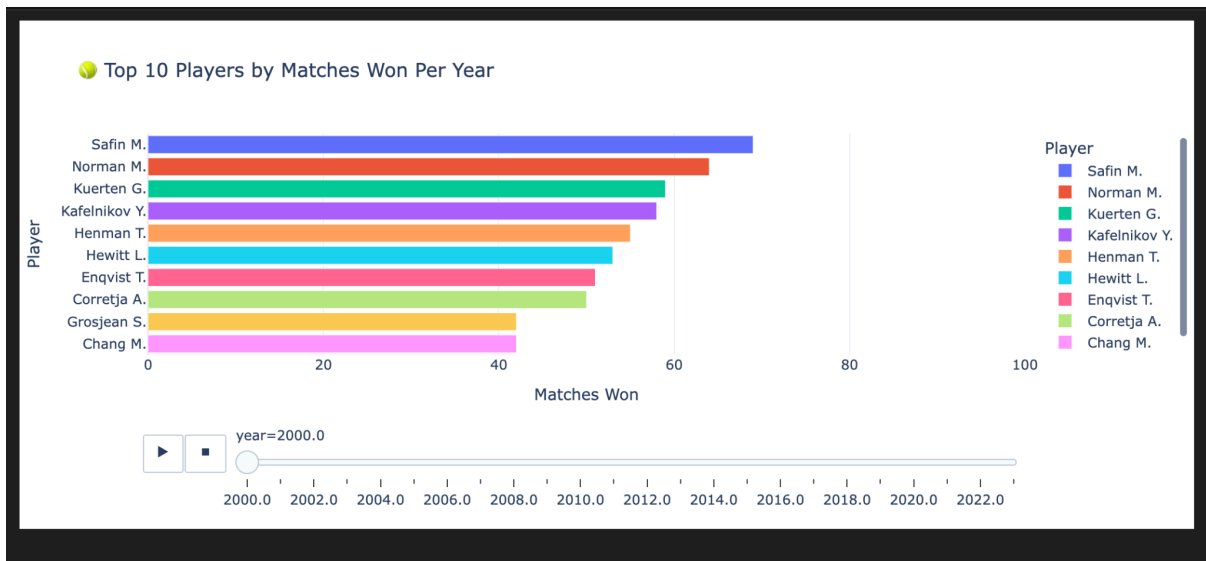
This animated bar chart race displays the top 10 players with the most matches won each year from 2000 to 2024. Users can interactively toggle between ATP and WTA tours. This plot helps visualize shifts in dominance over time and highlights consistency in player performance. It was implemented using Pandas for data aggregation and Plotly Express for the animation and interactive dropdown. Lectures referenced include *5.1 Interaction* for dropdown filtering, *6.1 Perception & Colors* for color consistency by player, and *7.1 Designing Visualizations* to ensure clarity and intuitive axis sorting. This component is fully functional and represents one of the project's minimal viable visualizations.

4. Underdog Upsets vs Top 10 Players

This interactive scatter plot captures every instance where an underdog (lower-ranked player) beat a top 10 opponent. The X-axis represents the underdog's world rank, while the

Y-axis shows the betting return on a $1 bet. The plot is color-coded by surface (clay, grass, hard) and allows filtering by the defeated top 10 player. Hover tooltips provide rich details including the names of both players, the surface, and the return. This visualization was built using Plotly Express and Pandas and was inspired by lecture *5.2 More Interaction* for dropdowns, *6.1 Perception & Colors* for the surface background scheme, and *7.2 Do's and Don'ts* to maintain a clean layout. It offers unique insight into which underdogs were the most profitable and which top players were most frequently upset.
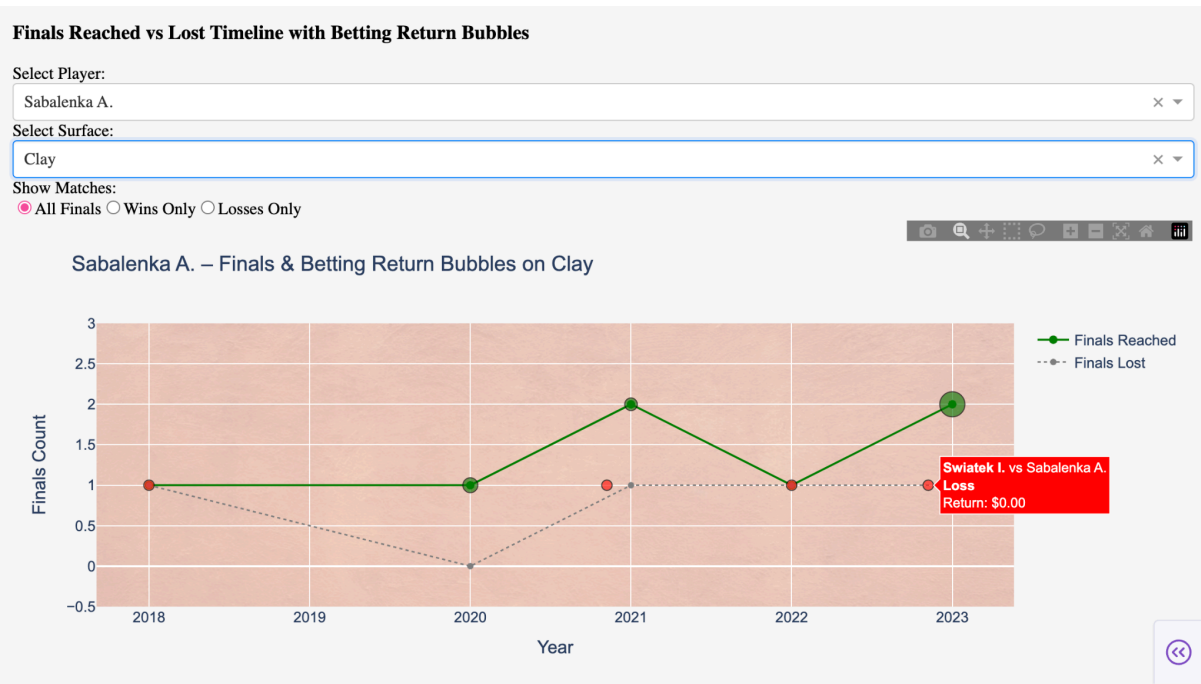
5. Finals Timeline + Betting Return Bubbles (e.g., Vekic D.)
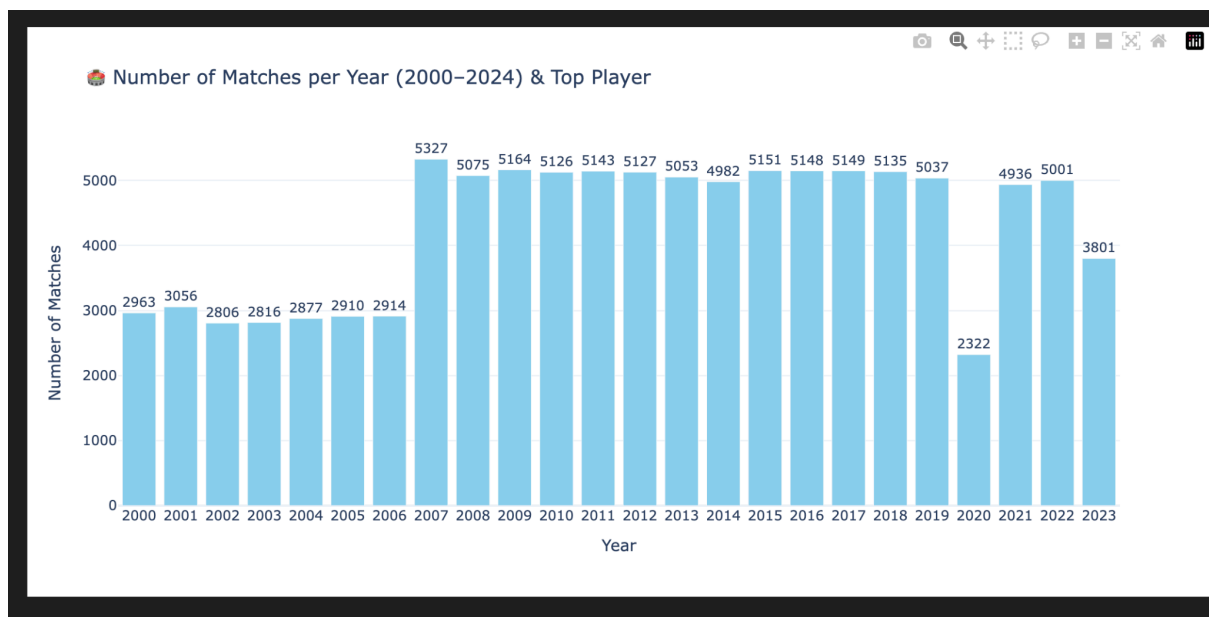

This plot shows a player's finals history over time, with green bubbles marking wins (scaled by betting return) and dashed gray markers for losses. The viewer can filter by player, surface, and outcome type (wins/losses/all). The core idea here is to merge timeline analysis with contextual betting return, helping users spot peak seasons or slumps. The background surface coloring adds subtle context, while interactivity lets users explore individual stories. This plot uses Plotly for custom bubble charts and hover tooltips, Dash for widgets, and was inspired by Lecture 5.2 (More Interactive D3) and 7.1 (Designing Visualizations), especially regarding minimal annotations and encoding values in visual forms like size and color.

6. Finals Timeline for Sabalenka (Highlighted Loss)

A specialized version of Plot 3, this focuses exclusively on Aryna Sabalenka's finals on clay courts. It emphasizes her rivalry with Swiatek by highlighting losses using large red labels with $0 return. The narrative quality of this view makes it ideal for storytelling and engaging a tennis-aware audience. It uses the same technologies as Plot 2 (Dash + Plotly), but the design focuses more on narrative clarity, in line with Lecture 7.2 (Do & Don't) and Lecture 6.1 (Perception), ensuring key moments (like finals lost to top opponents) stand out visually and emotionally.

Finals Reached vs Lost Timeline with Betting Return Bubbles

Select Player:
Sabalenka A.

Select Surface:
Clay

Show Matches:
◉ All Finals ○ Wins Only ○ Losses Only

Sabalenka A. – Finals & Betting Return Bubbles on Clay

Swiatek I. vs Sabalenka A.
Loss
Return: $0.00

7. Number of Matches per Year (2000–2024) & Top Player

This bar chart provides a high-level overview of the volume of matches played each year across the tour. It gives immediate visual feedback on anomalies, such as the drop in 2020 due to COVID-19. Subtle details like value labels on top and a neutral color scheme help focus the viewer's attention on the trend. This plot serves as a foundational context layer, anchoring the rest of the visualizations in tournament activity. Built with Plotly and pandas, it reflects best practices from Lecture 4.2 (D3 Basics) and Lecture 7.1 (Designing Visualizations) — especially the recommendation to reduce ink and guide attention toward the primary message.

8. Ranking Points Over Time – Djokovic N
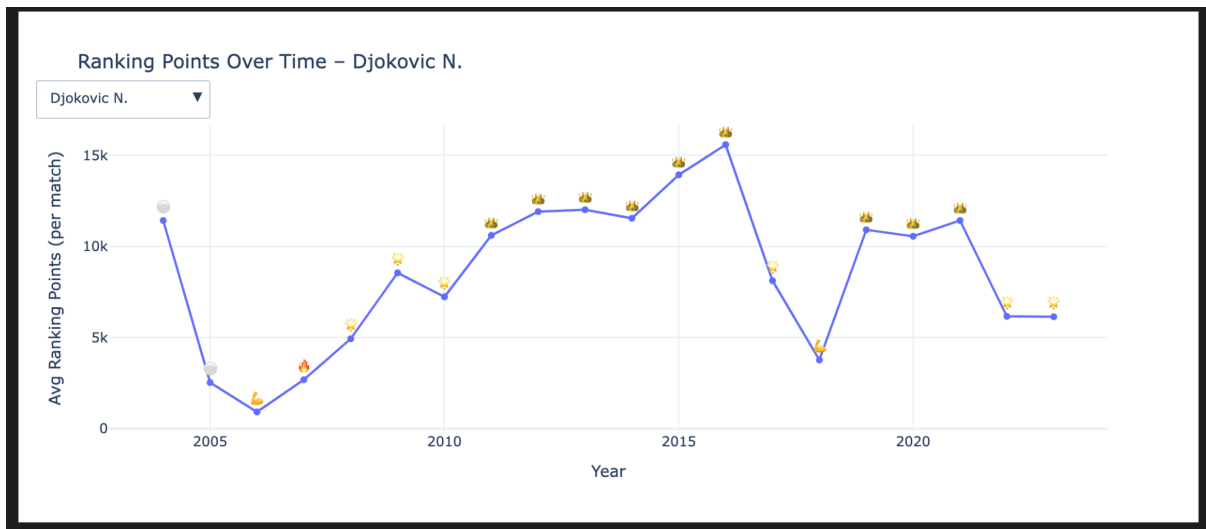
This line plot shows how a player's average ranking points per match evolve over the years. It includes emoji annotations to highlight dominant years (👑), comebacks (💪), or slumps (💤). The dropdown widget lets users switch between players instantly. This visualization is focused on *temporal performance analysis* and storytelling. It uses Plotly Express for smooth line rendering and ipwidgets for dropdown interaction, based on techniques from Lecture 5.1 (Interaction) and Lecture 6.1 (Color and Perception). Its simplicity and directness make it ideal as a complement to more complex plots.

Remarks

The first two plots will be kept in the final version of the project, as they provide a clear and interactive comparison between two players in terms of performance, statistics, and head-to-head results.

However, the inclusion of the following visualizations is still under discussion. We will evaluate their added value, clarity, and relevance to the main objectives of the project before deciding whether to include them.

Website : https://com-480-data-visualization.github.io/SmashData/