ASSESSMENT 2 REPORT

A Visual Analysis of the Australian Open Tennis Championship

OCTOBER 1, 2025

COURSE: 32146 ADVANCED DATA VISUALISATION AND VISUAL ANALYTICS

Student ID: 26130563

Table of Contents

1. Introduction and Summary

The Australian Open stands as one of the most prestigious events in international tennis, with a storied history spanning 120 years. This vast timeline offers a rich dataset through which to explore the evolution of the sport and the defining characteristics of its greatest champions. The primary purpose of this report is to use advanced data visualization techniques to mine this dataset for meaningful patterns, translating raw numbers into an accessible and insightful narrative for a broad audience interested in tennis.

Through the use of advanced visualization techniques in Tableau, key historical trends, national dominance patterns, and the unique signatures of elite performance have been identified and explained. The analysis reveals Australia's overwhelming historical dominance, which has gradually given way to a more globalized field of champions. A focused investigation into players with five or more titles - including Novak Djokovic, Serena Williams, and Margaret Court - highlights that sustained excellence requires a combination of extraordinary skill, remarkable longevity, and often, a favorable competitive ecosystem. Techniques such as geographic maps, treemaps, parallel coordinates, and scatter plots were employed to deconstruct multi-dimensional match dynamics, and give a clear relationship between win rate and other metrics in securing victory. While Tableau assists in creating interactive and insightful visualizations, limitations in customization and data preparation were noted. Overall, this report not only highlights the evolution of a sporting event but also demonstrates how modern visual analytics can transform complex historical data into actionable insights for tennis enthusiasts and analysts alike.

2. Data Preparation and Setup

2.1 Data Source and Initial Overview

The analysis was conducted on a provided dataset detailing the finals of the Australian Open tennis championships from 1905 to 2025. The original dataset contained both categorical and numerical variables, including: Year, Gender, Champion Name, Champion Nationality, Champion Seed, Runner-up Name, Runner-up Nationality, Runner-up Seed, and Score. A significant feature was the pre-processing of the Score column, which had been parsed into separate columns for each set.

2.2 Data Collection

A crucial step in data preparation involved enriching the dataset with match duration. As specified in the task requirements, the Mins column was populated by researching historical match records using online resources, including structured queries to AI assistants like ChatGPT and DeepSeek, followed by manual verification where possible.

I proceeded by prompting to the DeepSeek AI assistant, with explicit instructions to source information only from authoritative and prestigious outlets. There are some data limitations encountered. For a substantial portion of early championship matches, particularly those from the pre-Open Era, duration metrics were not accessible, so I kept them as data gaps. Furthermore, for

the records that were available, the accuracy and consistency of timing methodologies across different historical periods remain unverifiable. Consequently, while the collected duration data provides valuable insights for trend analysis, it must be interpreted with appropriate caution regarding its completeness and reliability for the earliest decades of the tournament's history.

2.3 Data Transformation and Calculation

To facilitate the required analysis, the raw dataset underwent significant transformation through the creation of multiple calculated fields and data cleaning procedures within Tableau. The key operations are detailed below.

2.3.1 Calculations

- Sets Played: A formula was created to count the number of sets in each match by analyzing the structure of the Score field:

  ```
  LEN([Score]) - LEN(REPLACE(REPLACE([Score], ",", ""), ".", "")) + 1
  ```

  This works by counting the number of set delimiters (commas or periods) in the Score string and adding one.

- Games Won and Total Games Played:
  - Total Games Won:

    ```
    ZN([1St-Won]) + ZN([2Nd-Won]) + ZN([3Rd-Won]) + ZN([4Th-Won]) + ZN([5Th-Won])
    ```

  - Total Games Played (Total Loss was calculated the same way as Total Won):

    ```
    [Total_Won] + [Total_Loss]
    ```

- Match Win Rate: `[Total Games Won] / [Total Games Played]`
- First Set Win Rate: `ZN([1St-Won]) / (ZN([1St-Won]) + ZN([1St-Loss]))`
- Championship Count: A simple count of records where a player's name appeared as the Champion was used to determine their total number of wins. This field was the filter for isolating the "Top Players."

2.3.2 Normalizations

To be able to use Parallel Coordinates plot, which requires dimensions to be on a comparable scale, key numerical attributes were normalized to a 0-1 range. I applied normalization to the 1st-won through 5th-won and 1st-loss through 5th-loss fields using the standard min-max normalization formula for each axis, a sample is as follow:

```
([Nth-won] - MIN([Nth-won])) / (MAX([Nth-won]) - MIN([Nth-won]))
```

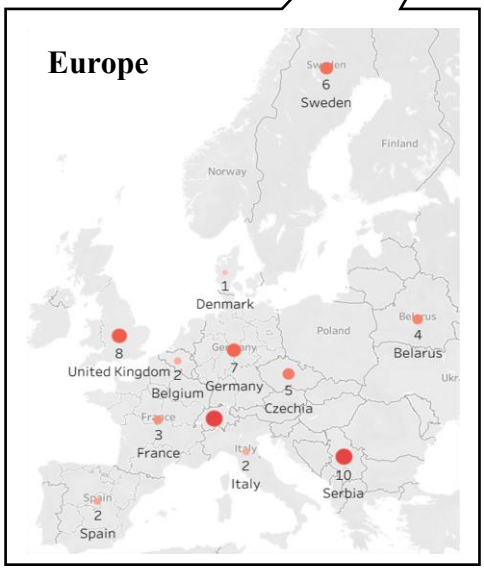
2.3.3 Data Cleaning

The Year field was converted from a string data type to an integer using `INT([Year])`, then divided into bins of 3 years to allow for proper chronological sorting and time-series analysis.

3. Visualization and Insights

3.1 Global Tennis Big Picture

3.1.1 World Tennis Power Map



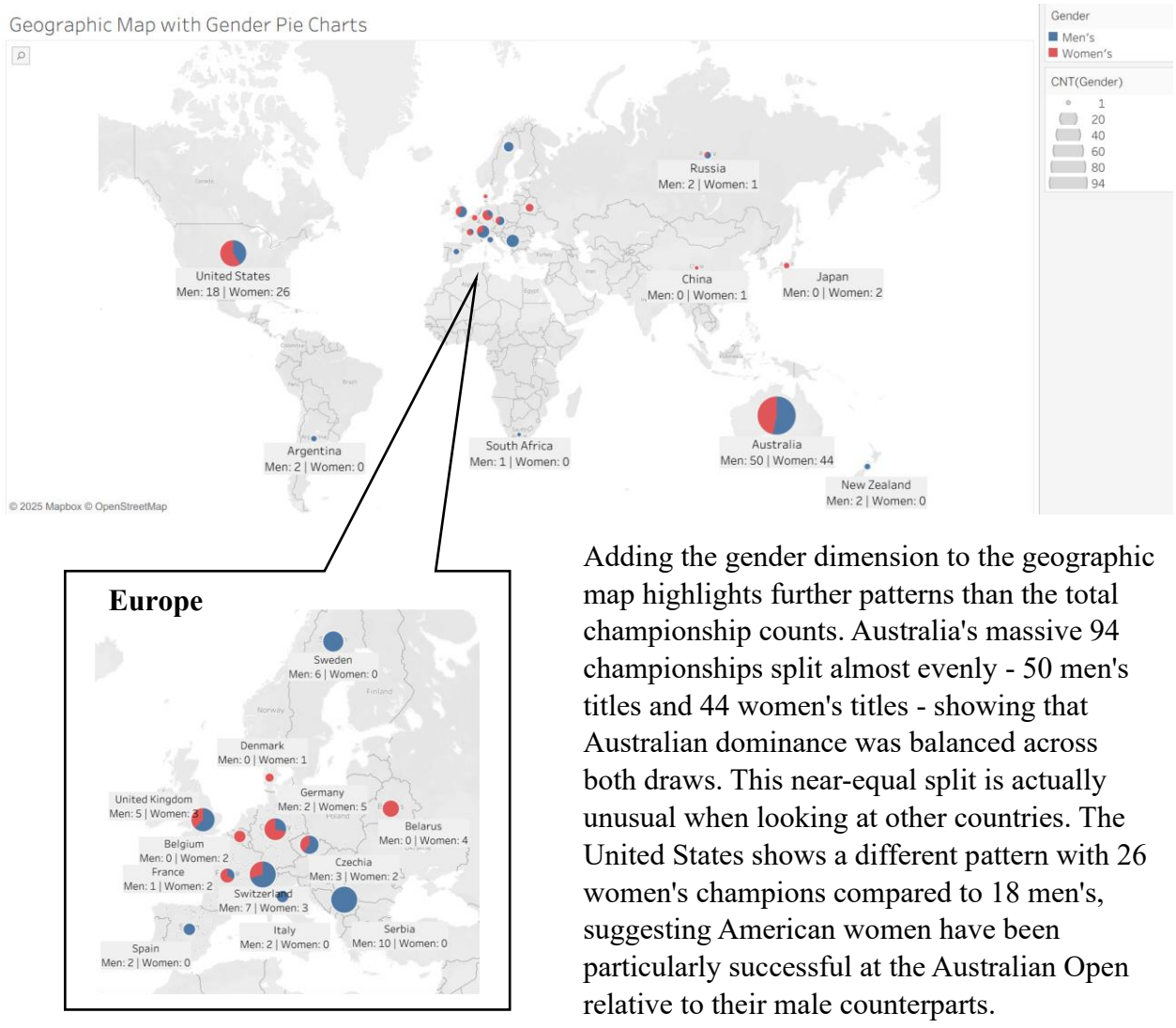
The geographic map immediately reveals where tennis power has been concentrated over the Australian Open's 120-year history. Australia dominates with 94 championships, shown by the massive circle on the map. This isn't just dominance, it's overwhelming home-court advantage that lasted for decades. The United States comes second with 44 championships, less than half of Australia's total but still significantly ahead of other nations. These two countries alone account for the majority of Australian Open titles throughout history.

Europe shows a scattered but strong presence across multiple nations. Germany has 7 championships, Serbia 10, and Belarus 10, while Sweden contributed 6 titles. The United Kingdom has 8, France 3, and Spain 2. What's interesting here is that European success is spread across many countries rather than concentrated in one or two nations. Each country seems to have produced a few exceptional players who broke through at different times, unlike Australia's sustained multi-generational dominance.

Russia and Japan each have modest showings with 3 and 2 championships respectively, while China, South Africa, Argentina, and New Zealand appear with 1-2 titles each. These smaller circles represent breakthrough moments with individual champions who put their countries on the Australian Open map.

The map immediately shows geographic patterns that numbers alone can't convey. It can be seen at a glance that tennis success has been a primarily Western phenomenon, with strong representation from Europe, North America, and Australia, while vast regions of Africa, South America, and Asia remain largely absent. The size difference between circles makes Australia's historical dominance impossible to miss, their home tournament truly was their domain for most of its history.

3.1.2 Gender Distribution of Champions by Nation



Adding the gender dimension to the geographic map highlights further patterns than the total championship counts. Australia's massive 94 championships split almost evenly - 50 men's titles and 44 women's titles - showing that Australian dominance was balanced across both draws. This near-equal split is actually unusual when looking at other countries. The United States shows a different pattern with 26 women's champions compared to 18 men's, suggesting American women have been particularly successful at the Australian Open relative to their male counterparts.

Europe presents the most varied gender patterns across the continent. Some countries are entirely single-gender success stories. Sweden produced 6 male champions but zero women's titles, dominated by players like Björn Borg and Stefan Edberg in their eras. Serbia shows the same

pattern with 10 men's titles and no women's champions, with Novak Djokovic's dominance. Spain (2 men, 0 women) and Italy (2 men, 0 women) follow this male-only pattern as well. On the flip side, Belarus achieved all 4 of its championships through women, while Belgium's 2 titles and Denmark's single championship also came entirely from women's tennis.

Some European nations show more balance. Switzerland has 7 men's titles and 3 women's, Germany split 2 men's and 5 women's (with the women actually outperforming the men), and Czechia divided 3 men's and 2 women's. The United Kingdom's 5 men's and 3 women's titles show reasonable representation on both sides. France achieved 1 men's title and 2 women's titles. This scattered European pattern suggests that championship success depended heavily on individual exceptional players rather than systematic national programs producing champions across both genders.

The Asia-Pacific region shows interesting contrasts. Japan's 2 championships both came from women's tennis, while China's single title was also a woman's achievement. This reflects the broader pattern in Asian tennis where women have broken through more successfully than men at the Grand Slam level. Russia's 2 men's and 1 women's title shows modest representation on both sides. Meanwhile, Argentina (2 men), South Africa (1 man), and New Zealand (2 men) produced only male champions, suggesting limited women's tennis infrastructure or opportunity in those regions during their competitive eras.

3.1.3 The Frequency of Australian Open Champions

This word cloud shifts focus from countries to individual legends, and the size differences tell the story of dominance. Margaret Court name appears largest in dark red, reflecting her record-breaking 11 Australian Open titles, though it's worth noting many came in the amateur era before 1968. Novak Djokovic's stands almost as large, his 10 titles making him the most successful man at this tournament. These two names tower over the cloud. The next tier of names includes Serena Williams with her 7 Australian Open, Roy Emerson appears prominently with his 6 titles from the 1960s, and Roger Federer's name stands large with his 6 championships.

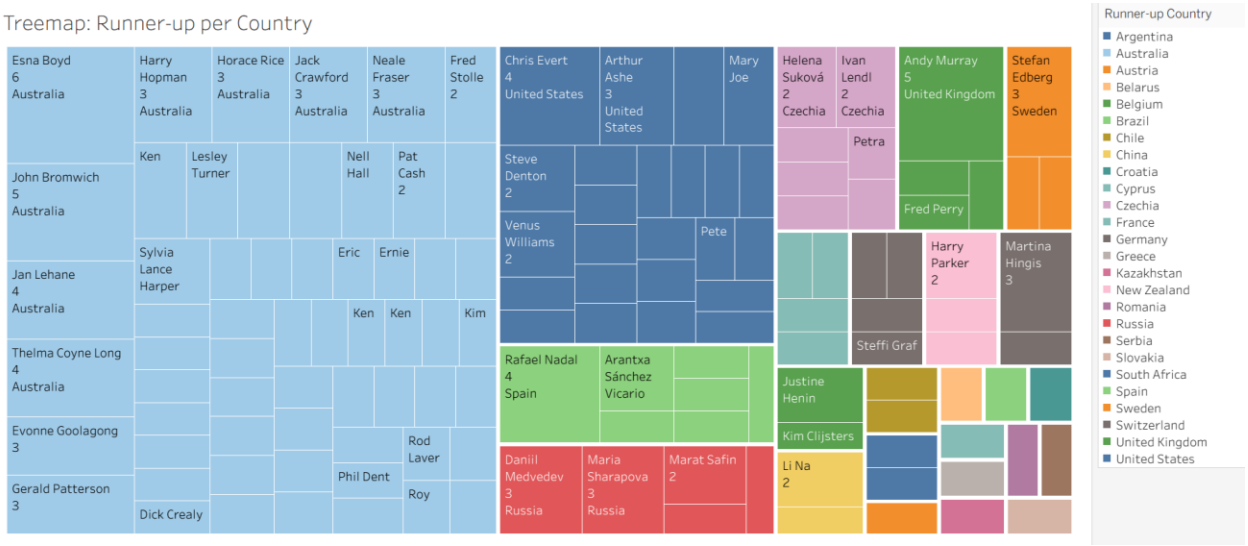
Mid-sized names reveal the next level of excellence. Steffi Graf, Monica Seles, Martina Navratilova, and Evonne Goolagong each won multiple titles and appear prominently. Andre Agassi, Mats Wilander, and Rod Laver also feature prominently, each with 3-4 titles. The mix of

nationalities like American, Australian, Czech, and Swedish names all appear at similar sizes, showing how the tournament's competitive balance shifted over time.

Smaller names scattered throughout represent single or double champions who still made their mark. In general, the cloud includes players from vastly different eras, from Jack Crawford and Adrian Quist in the 1930s to Daphne Akhurst in the 1920s and 1930s, all the way to modern champions like Aryna Sabalenka.

While the color gradient from light pink to dark red adds visual hierarchy to the word cloud, it has limitations as an analytical tool. It's hard to tell exact championship counts from text size alone, and the random positioning means no geographic or temporal patterns emerge. But as a complement for other visualization analysis, it works perfectly.

3.1.4 Most Frequent Runner-up Nations



This treemap shows the other side of tennis history - the players who reached finals but didn't win. Australia still dominates this chart with large blue rectangles, which is interesting when thinking about its overwhelming presence of champions. Players like Esna Boyd (6 times), John Bromwich (4 times), Jan Lehane (4 times), and Thelma Coyne Long (4 times) kept reaching finals but couldn't win the title. This suggests that during Australia's golden era, their players weren't just winning - they were filling up both sides of the finals draw.

The United States shows up strongly too, with Chris Evert appearing four times as a runner-up. This is surprising because Evert is considered one of the greatest players ever, yet she lost four Australian Open finals. Arthur Ashe, Venus Williams, and Steve Denton also add to the American runner-up count, which shows that even legendary players had their weak spots at certain tournaments.

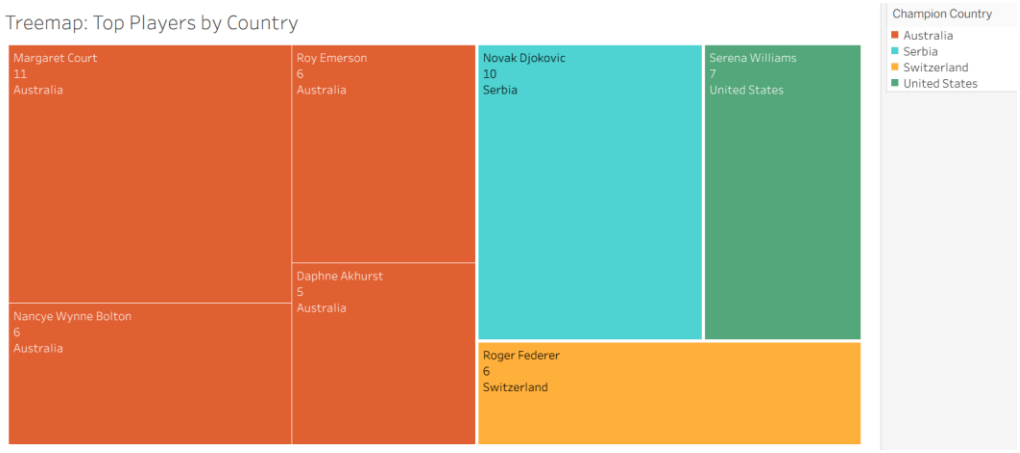
Russia stands out in red with three modern players who reached multiple finals: Daniil Medvedev (3 times), Maria Sharapova (3 times), and Marat Safin (2 times). This clustering in recent years shows that Russian players have been good enough to reach finals consistently but struggled to close them out. Spain's Rafael Nadal appears three times, Sweden's Stefan Edberg

three times, and Czech players like Helena Suková and Ivan Lendl also show up multiple times. These are all great players who simply couldn't crack through at this particular tournament.

The treemap's size-based layout makes it easy to see which countries had the most near-misses. Comparing this to the champions distribution, it could be seen that Australia appears huge in both charts, showing they had depth across many players. On ther other hand, countries like Switzerland barely shows up here despite having champions like Federer, which means their success came from just a few dominant players rather than a deep pool.

3.2 The Elite Champions Analysis

3.2.1 Elite Champions by Nationality



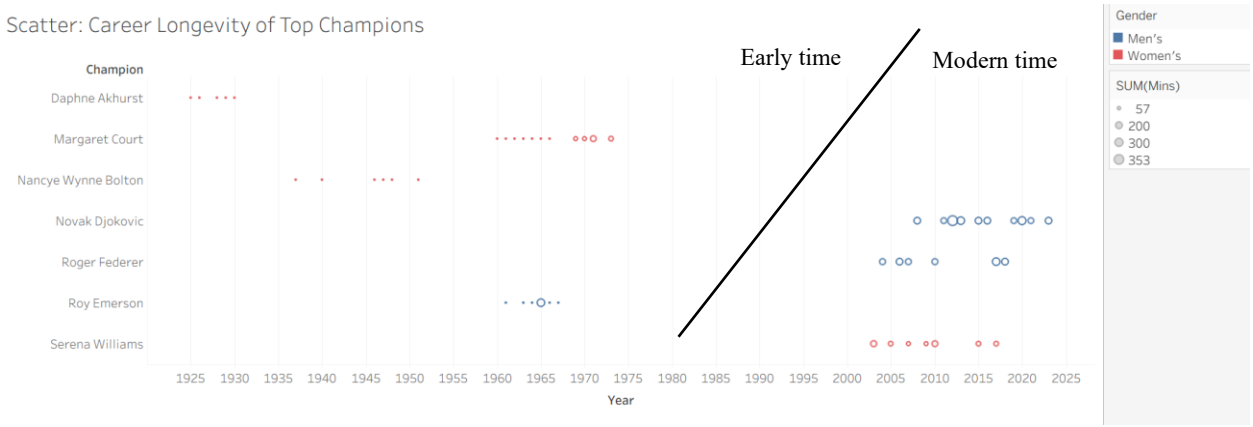
This treemap focuses exclusively on the most dominant players in Australian Open history - those who won 5 or more titles. It can be easily seen that only seven players achieved this elite threshold, and they come from just four countries. Australia again dominates the space with three massive orange rectangles representing Margaret Court (11 titles), Roy Emerson (6 titles), Nancye Wynne Bolton (6 titles) and Daphne Akhurst (5 titles). Although these Australian legends benefited from home-court advantage, their combined winnings reinforces Australia's golden era the same way as the geographic maps already suggested.

The modern era tells a different story. Novak Djokovic's turquoise rectangle representing his 10 titles stands nearly as large as Margaret Court's, but his achievements came against far deeper competition in the Open Era. Every one of his titles required beating the best players in the world who had access to professional coaching, year-round training, and modern sports science. Serena Williams' green rectangle shows 7 titles for the United States, making her the most successful woman at this tournament since the professional era began. Roger Federer's orange rectangle from Switzerland shows 6 titles, matching Roy Emerson's total but achieved in a completely different competitive landscape.

What stands out is how few countries have produced these ultra-dominant players. Only Australia, Serbia, Switzerland, and the United States appear on this treemap. Despite the geographic maps showing champions from 20+ countries, sustained dominance at the Australian Open has been rare and geographically concentrated. No European country besides Switzerland

and Serbia produced a player with 5+ titles. No Asian nation appears. No South American country makes the cut. This suggests that while breaking through for a single championship is achievable for many nations, building the sustained excellence needed for multiple titles requires resources, infrastructure, and often home-court advantage that only a handful of countries could provide.

3.2.2 Career Longevity of Top Champions



This scatter plot shows when these elite champions won their titles and how long their dominance lasted. The timeline spans exactly 100 years, from Daphne Akhurst's cluster of red dots in the mid-1920s to Novak Djokovic's blue dots stretching into 2025. Each dot represents a championship victory, and the horizontal spread shows career longevity - how many years separated a player's first and last title.

The early champions show tight clusters. Daphne Akhurst's five red dots bunch together between 1925 and 1930, representing a brief but intense period of dominance that ended with her early death at age 29. Margaret Court's red dots span from the late 1950s through the early 1970s, showing an impressive 13-year reign. Her championships came in two distinct phases with gaps between them, possibly reflecting breaks for childbirth or shifting focus between singles and other events. Nancye Wynne Bolton's red dots appear in the 1930s and 1940s, spanning roughly 15 years. Her pattern is interesting because it shows the longest career span among the historical champions, likely reflecting the disruption of World War II which limited international competition. Roy Emerson's blue dots cluster tightly in the early-to-mid 1960s, showing six titles compressed into just seven years (1961-1967).

The modern champions tell a completely different story about longevity. Roger Federer's blue dots stretch from 2004 to 2018, a remarkable 14-year span between first and last title. His dots show clusters in the mid-2000s and then a gap before returning to win again in 2017-2018, demonstrating the ability to remain competitive into his mid-30s. Serena Williams' red dots span an even more impressive range from 2003 to 2017, with championships scattered across 14 years and showing she could win in her early twenties and again in her mid-thirties after becoming a mother.

Novak Djokovic's dominance is unmistakable. His blue dots begin in 2008 and continue all the way to 2025, representing an astounding 17-year championship span. The density of his dots is significant with 10 titles spread across nearly two decades, particularly heavy clustering in the 2010s and early 2020s. Unlike Federer who had a long gap, Djokovic's championships show more consistent frequency throughout his career. This visualization makes it clear why many consider him the greatest Australian Open player ever.

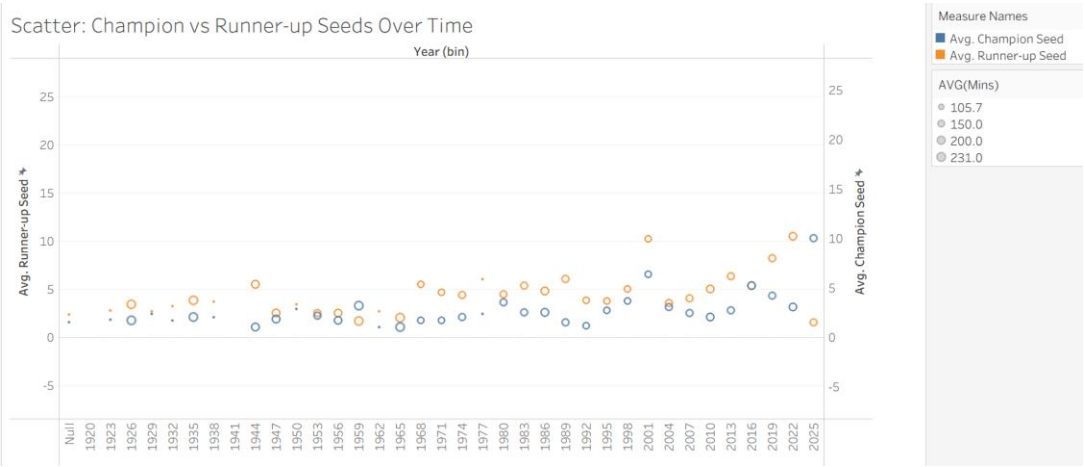
The dot sizes, which represent match duration, show some patterns worth noting. The larger circles appear predominantly in the modern era - Djokovic, Federer, and Serena Williams all have noticeably bigger dots, suggesting their championship matches lasted longer. The historical players like Daphurst, Court, and Emerson show smaller dots or missing data entirely, which makes sense given that match duration records from the 1920s-1960s are incomplete or weren't systematically tracked.

The gender split is visually clear with red (women's) and blue (men's) dots. What's striking is that among these elite champions, the women's dots (Court, Akhurst, Bolton, Williams) show patterns just as impressive as the men's in terms of career longevity and title frequency. Serena Williams' spread rivals or exceeds Federer's, while Margaret Court's span matches or exceeds many male champions of her era.

This scatter plot format works better than a simple table because it shows time visually, we can immediately see that tennis dominance has evolved: early champions had short and intense periods of winning, while modern champions like Djokovic sustain excellence across nearly two decades. The gaps between dots tell stories too: injuries, personal breaks, or simply getting beaten by better players that year.

3.3 Match Dynamic Analysis

3.3.1 Seed Dynamics between Champion vs Runner-up



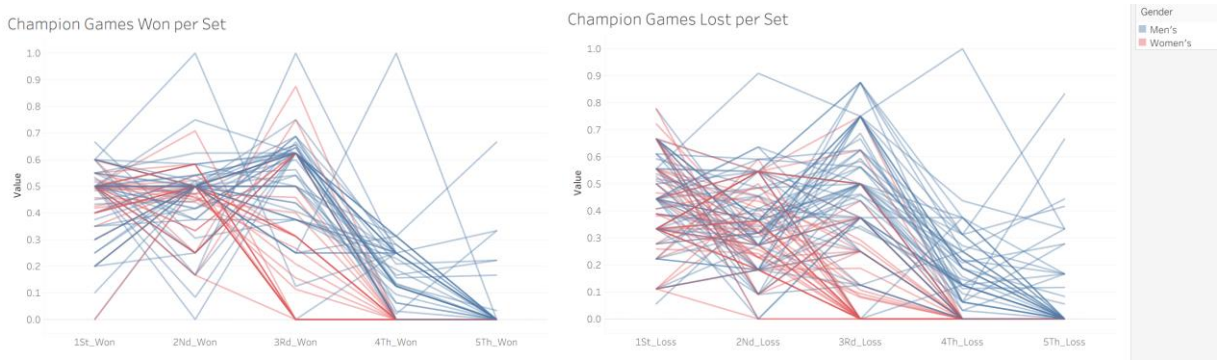
The scatter plot tracks champion and runner-up seedings over 120 years, with circle size showing average match duration. The data reveals how the Australian Open evolved from an unseeded amateur tournament to a structured and professional competition. The early years show no data points because seeding wasn't introduced until the Open Era began in 1968. Before that, players

simply entered and competed without any ranking-based structure, which explains why Margaret Court and Roy Emerson's dominance came in an era without the competitive organization we see today.

Once seeding data appears, the pattern is clear: champions typically come from lower seed numbers (better rankings) than their opponents. Most blue circles (champion seeds) cluster between 1-5, while orange circles (runner-up seeds) spread more widely from 1-10. This confirms our expectation that better-ranked players usually win. However, the spread of orange circles shows that upsets happen regularly. Runner-ups seeded 6th, 7th, or even 10th reaching the final demonstrates that the tournament structure allows for breakthrough performances, even if those underdogs rarely complete the fairy tale.

The size variation in the circles adds another layer to the story. Larger circles represent longer matches, and these appear scattered throughout the timeline without a clear pattern tied to seeding. Some finals between top seeds lasted over 200 minutes (shown by the larger circles), while others finished quickly. A final between the #1 and #2 seeds doesn't automatically mean a longer, more competitive match. Similarly, some matches where a lower-seeded player reached the final also produced extended battles, suggesting that upset runs can lead to either quick collapses or genuine competitive finals.

3.3.2 Anatomy of a Championship Match



The two parallel coordinates charts reveal the performance patterns of champions across the five sets of their matches, with one chart tracking games won per set and the other tracking games lost per set. Each flowing line represents an individual champion's journey through their match, with blue lines showing men's champions and red lines showing women's champions.

Looking at the games won chart on the left, most champions cluster between 0.5 and 0.7 in the first two sets, showing they typically establish control early. A few exceptional lines shoot up to 1.0 in the second or third set, while the fourth and fifth sets show dramatic variation with many lines plummeting toward zero. This pattern makes sense because not every match goes the full five sets - when a player wins in three or four sets, the remaining set columns naturally show zeros. On the other hand, the games lost chart on the right tells the complementary story. Champions naturally lose fewer games overall, but the variation in their lines reveals different playing styles and match narratives. Some lines stay consistently low around 0.2 to 0.4 across all sets, other lines show dramatic spikes reaching 0.8 or 0.9 in particular sets.

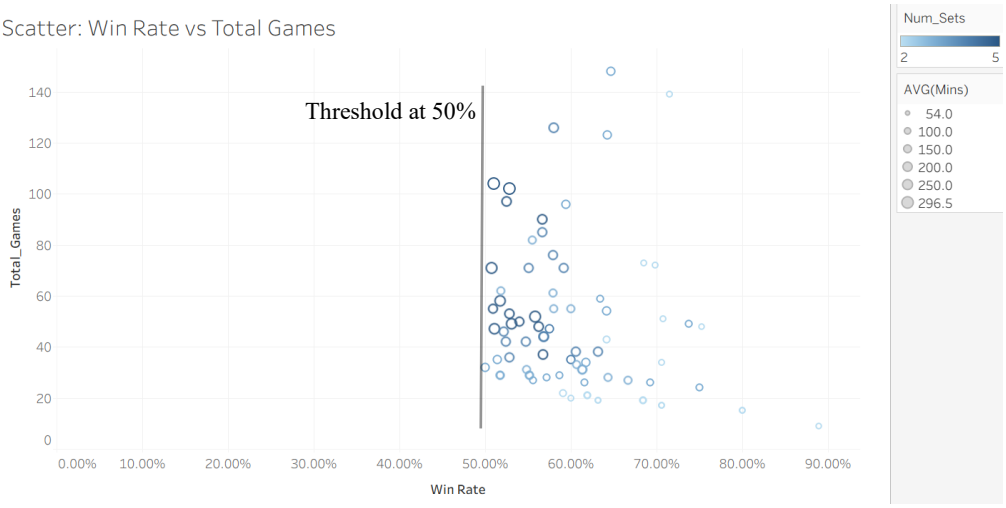
The gender comparison reveals interesting differences in match dynamics. Men's champions show tighter clustering in the early sets, with their blue lines forming a denser band between 0.5 and 0.6 in sets one through three. Women's champions display more spread across all sets, with their red lines showing greater variation. This doesn't indicate that one gender performs better than the other, but rather reflects different competitive dynamics in men's and women's tennis. Men's matches might feature more consistent serving dominance, while women's matches might involve more momentum swings and comeback attempts.

What makes parallel coordinates particularly effective for this analysis is its ability to show five dimensions of data simultaneously while preserving the connection between each dimension. Still, individual player tracking becomes difficult in areas where many lines overlap. The middle sections of both charts show such dense clustering that following a specific player's path through all five sets is nearly impossible without interactive highlighting.

3.4 Win Rate Analysis

Win rate serves as the fundamental measure of champion performance in tennis, capturing the percentage of games won throughout a match. However, win rate alone doesn't tell the complete story of how champions achieve victory. To understand the fuller picture, this section examines win rate in relation to three key dimensions: the total number of games played, match duration, and first set performance. These relationships may reveal different aspects of championship attainments.

3.4.1 Efficiency vs Endurance

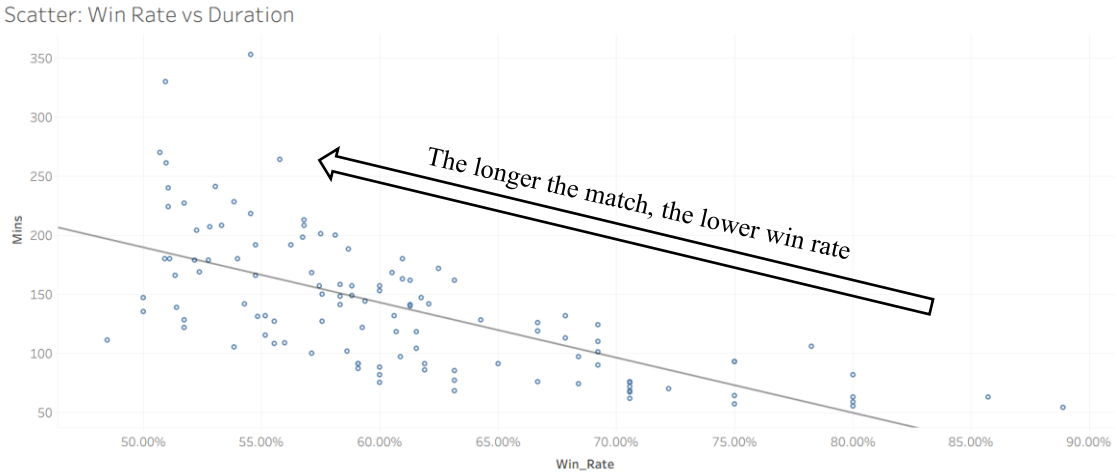


The first scatter plot positions win rate along the horizontal axis and total games played on the vertical axis, with circle size representing average match duration and color intensity indicating whether matches went to two or five sets. The data points cluster mainly between 50% and 70% win rate, which reflects the competitive nature of championship matches where both finalists have already proven themselves by reaching the final. No champion appears below 48% win rate, which makes sense because one cannot win a tennis match while losing more than half the games.

Moreover, the color encoding reveals more meaningful patterns. Darker blue circles representing five-set matches concentrate more heavily in the 50-60% win rate range and higher up on the game count axis. This aligns with our expectations: matches that go the full five sets are by definition closer contests, which pulls win rates toward 50%. These marathon matches also naturally accumulate more total games. Lighter colored circles representing shorter matches appear more frequently at higher win rates around 65-75%, showing that dominant champions often finish in three or four sets without needing to battle through a fifth set decider.

Some outliers deserve attention. One data point sits near 150 total games with a win rate around 65%, showing an extraordinarily long match where the champion still managed to win nearly two-thirds of games. On the opposite extreme, many small circles appear below 30 total games with win rates around 70%, representing quick straight-set victories where the champion dominated from start to finish. To conclude, there is no single path to championship victory for the tennis player.

3.4.2 The Efficiency Paradox



The second scatter plot flips the axes to show win rate on the horizontal axis and match duration in minutes on the vertical axis, with each circle representing an individual champion's performance. A gray trend line cuts diagonally downward across the chart, showing a clear negative correlation between these two variables. This relationship might seem counterintuitive at first but reveals a fundamental truth about tennis dominance.

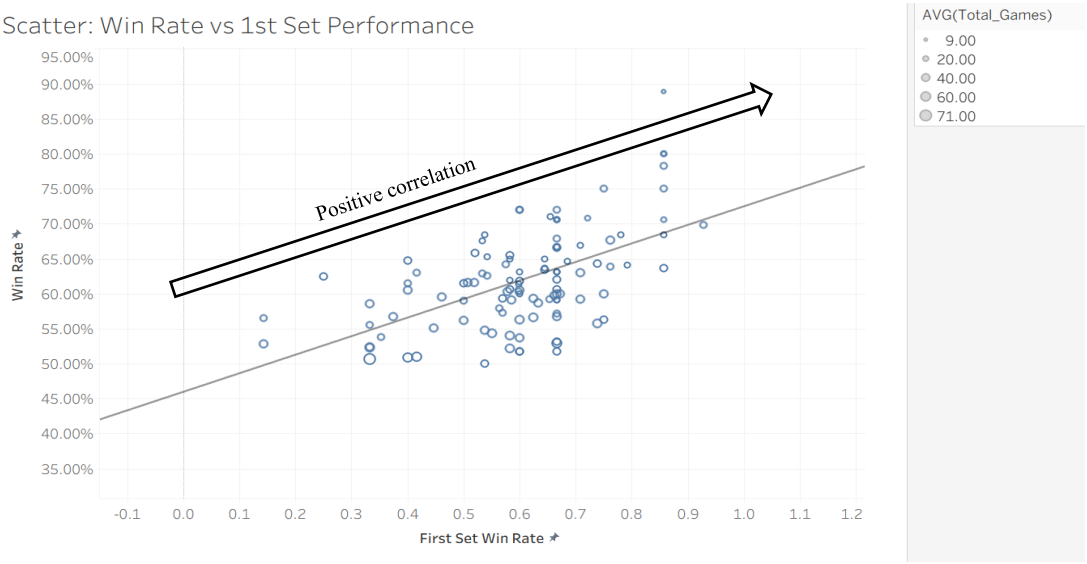
The trend line drops from approximately 200 minutes at 50% win rate down to around 50 minutes at 90% win rate. This downward slope demonstrates that champions who win a higher percentage of games finish their matches significantly faster. The logic becomes clear when considering tennis scoring - a player who wins 70% of games will likely win most sets 6-4 or 6-3, closing out the match in three or four sets. A player winning only 52% of games might squeak through sets 7-5 or win tiebreaks 7-6, requiring five full sets to accumulate enough games for victory. More games means more sets, and more sets means longer matches.

The data points scatter quite widely around this trend line, particularly in the 55-65% win rate range where match durations vary enormously from 80 minutes up to 270 minutes. This variation

reveals that factors beyond win rate heavily influence match length. Playing style, weather conditions, and the specific dynamics between two players all contribute to pace of play.

Several outliers punctuate the upper left portion of the chart. Two data points sit above 300 minutes at win rates around 50-53%, possibly the the legendary marathon matches that become part of tennis folklore. On the opposite extreme, several data points cluster below 70 minutes with win rates above 78%, these are the most dominant championship performances where the winner thoroughly outclassed their opponent. The negative correlation shown by the trend line reveals that efficiency in tennis, measured by win rate, can directly translates to time efficiency.

3.4.3 Impact of a Strong Start



The third scatter plot examines the relationship between first set win rate on the horizontal axis and overall match win rate on the vertical axis, with circle size representing the total number of games played. An upward-sloping trend line cuts across the chart giving a strong positive correlation that confirms what tennis strategists have long believed - the first set sets the tone for the entire match.

The trend line rises steadily from around 35% overall win rate at 30% first set win rate up to approximately 82% overall win rate at 90% first set win rate. This positive relationship demonstrates that champions who dominate the opening set tend to maintain that dominance throughout the match, and makes both statistical and psychological sense. Statistically, winning more games in the first set often means winning the set itself, giving the player a 1-0 lead in sets and requiring them to win only one or two more sets rather than three. Psychologically, winning the first set builds confidence and puts pressure on the opponent who now faces the prospect of fighting back from behind.

The data points cluster tightly around the trend line in the middle range between 55-70% for both metrics, showing that most champions follow the expected pattern. Several outliers appear well above the trend line in the 70-80% first set win rate range, representing players who converted strong first set performances into even better overall results. These champions likely won their

first sets convincingly, demoralized their opponents, and then cruised through the remaining sets with even more dominant play, pushing their overall win rates above 75%.

The circle sizes add another dimension to the analysis. Larger circles representing higher total game counts appear scattered throughout the chart but show slight concentration in the middle range around 60% for both metrics. This suggests that players who faced longer matches with more total games tended to have more balanced performance across sets rather than extreme dominance or struggle in the opening set. The largest circles appear in the 55-65% range, showing that marathon matches with 80-100+ total games typically feature relatively even first set performance that mirrors overall match performance. Conversely, several very small circles appear at the extremes - tiny dots at 85% first set win rate with 78% overall win rate, representing quick straight-set victories where the champion dominated from the opening game and never let up.

The strategic implications of this scatter plot are significant. The strong positive correlation suggests that championship matches are often decided by first set performance. Players who establish dominance early tend to carry that momentum through the entire match. What this visualization ultimately demonstrates is that while the cliché "it's not how you start, it's how you finish" might apply to many aspects of life, in championship tennis, how you start would strongly predicts how you'll finish.

4. Conclusion

This project has successfully used advanced data visualization to decode 120 years of Australian Open history, to transform a complex historical dataset into a clear narrative of evolution and performance. The application of specific visualization techniques was paramount to this success. The *Geographic Map* instantly communicated the shift from Australian hegemony to a global competitive landscape, while *Treemaps* effectively illustrated the concentration of success among nations and elite players. The *Parallel Coordinates Plot* provided a unique, multi-dimensional view of match dynamics, and *Scatter Charts* revealed critical relationships between variables like win rate, duration, and career longevity.

A deliberate storytelling and labeling approach guided the analysis, to ensure each visualization was not just a chart but a narrative device. Axes were carefully titled, colors were consistently encoded by gender, and annotations were used to direct the reader to the most significant insights, turning data points into compelling stories of athletic achievement.

The analysis yielded several definitive trends, with key breakthrough points in the comparison of men's and women's performance. While both genders have produced legends of equal dominance in terms of title counts, their competitive pathways have differed. The visualization of career longevity revealed that elite female champions like Serena Williams have sustained championship-level performance over spans rivaling or exceeding their male counterparts. Furthermore, the parallel coordinates plot suggested greater variance in set-by-set performance in women's matches, indicating potentially different tactical flows and comeback dynamics, while men's matches showed tighter clustering in early sets, pointing to more consistent serving dominance.

The comparison of top players' performance shows the multi-faceted nature of tennis greatness. It distinguishes between the raw volume of Margaret Court's 11 titles, the historic-era dominance of Roy Emerson, and the modern-era longevity of Novak Djokovic, whose 17-year championship span is unprecedented. Serena Williams' career shows a combination of high title count and dominance in the professional era. This comparison shows that there is no single blueprint for success; rather, elite performance is a combination of peak dominance and the ability to adapt and endure across different competitive eras.

Finally, this project underscored the distinct advantages of Tableau as a visualization platform. Its drag-and-drop interface enabled rapid prototyping and iterative design, allowing for the quick exploration of different visual narratives. Its powerful built-in functions for creating maps, treemaps, and calculated fields were indispensable. The ability to create interactive dashboards where filters can update multiple visualizations simultaneously provides a depth of engagement that static reports cannot match. In conclusion, by marrying these robust visual analytics techniques with a rich historical dataset, this report not only documents the past of the Australian Open but also provides a structured framework for understanding the very patterns of excellence in professional tennis.