

比賽之外，從另外一個角度觀賞網球

用數據看網球：世界男單Top 100選手分析



Why this matters?

世界男單網壇長期被**BIG 4**佔據前幾名接近20年，隨著他們退休以及新秀崛起，排名的版圖也有所變化。

不只是觀看世界級的網球比賽，這項入門不易的運動也是我的休閒活動之一，看著電視轉播的時候，不禁想要知道為什麼「怎麼又是Alcaraz拿到冠軍！」、「Sinner也太強了吧！」，除了想要讓自己變厲害，也好奇選手們之間的不同，所以嘗試使用網路上可以找到的公開數據來分析他們的勝利方程式。

What we wanted to achieve?

1. 建立一個自動化的選手數據資料庫
2. 透過關鍵數據指標，量化選手的比賽風格
3. 為選手進行科學化的分類，提供不一樣的洞見



ANDY MURRAY



ROGER FEDERER



NOVAK DJOKOVIC



RAFAEL NADAL



Approach

數據抓取

資料清理與整合

探索式資料分析(EDA)

象限分析

選手分群

使用 tennisabstract.com 當作資料來源

1. selenium模擬瀏覽器、模擬點擊 + BeautifulSoup
2. 將數據中帶有“%”的字串，轉為可以分析使用的float。

核心指標欄位：

Rk (Rank): 選手的排名

M W% (Match Win Percentage): 總比賽勝率

SPW-InP (Serve Points Won - In Play): 發球得分率 (不含 Aces 和 Double Faults)

Ace%: 發球直接得分率

RPW (Return Points Won): 接發球得分率

Brk% (Breaks Converted Percentage): 破發成功率

Gms (Games): 總局數

利用不同資料欄位之間的交互關係，判斷哪些是區分選手類型的重要條件，以及哪些欄位之間是相關性不夠強的。

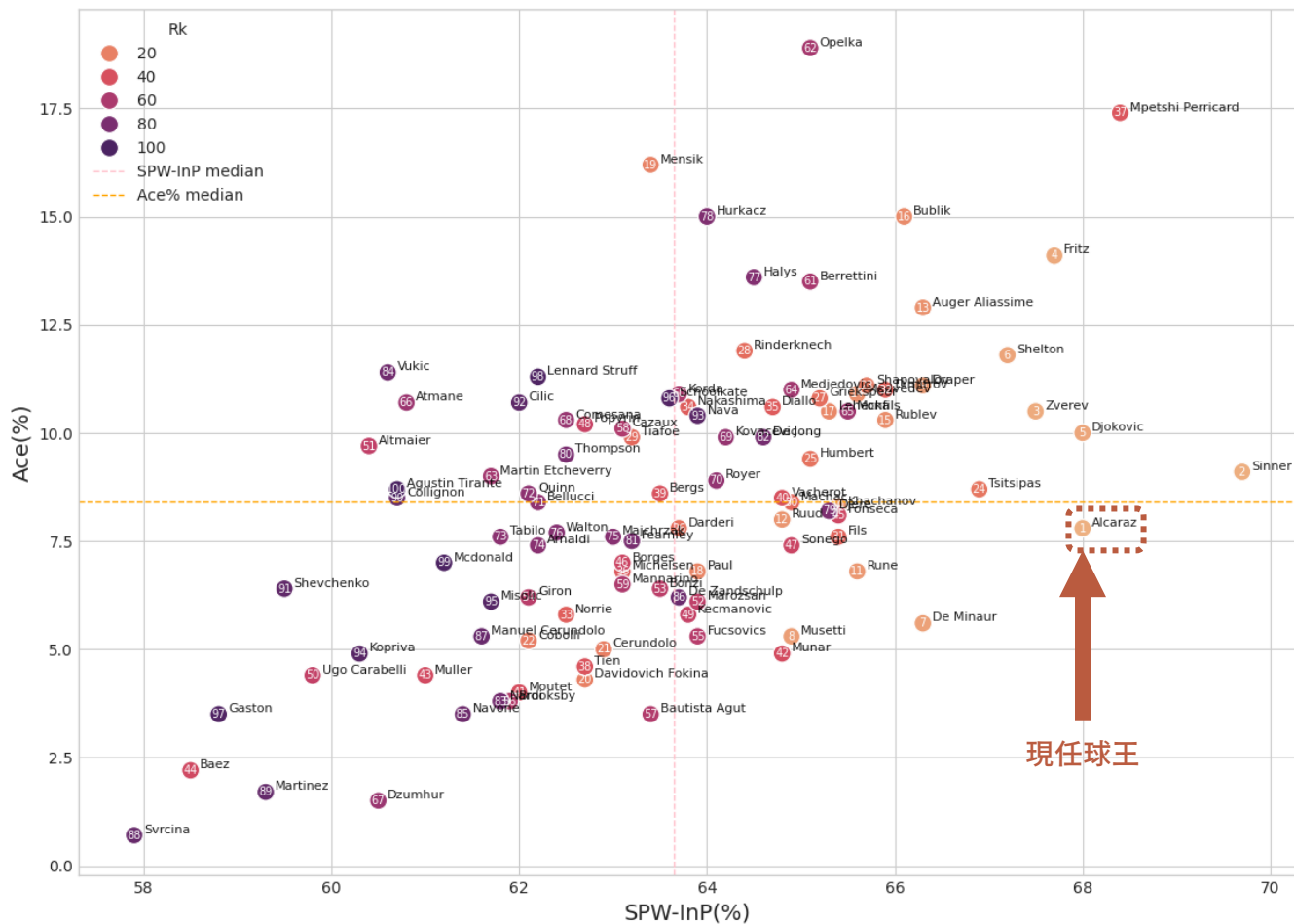
例如：SPW-InP與Ace%的相關性沒有想像中大。

最終選用SPW-InP以及RPW作為選手分群的基底，加上一些能夠協助判斷的欄位(例如經驗、得分效率相關的欄位)，作為分群的條件。

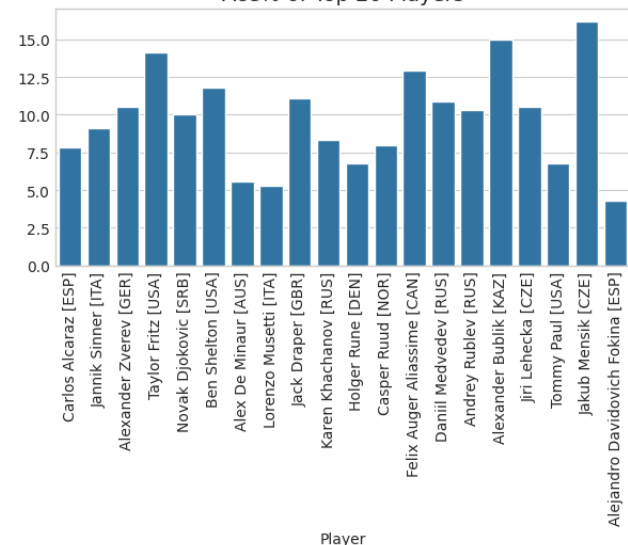


EDA for Serve

Server points won (in play) vs. Ace%



Ace% of Top 20 Players

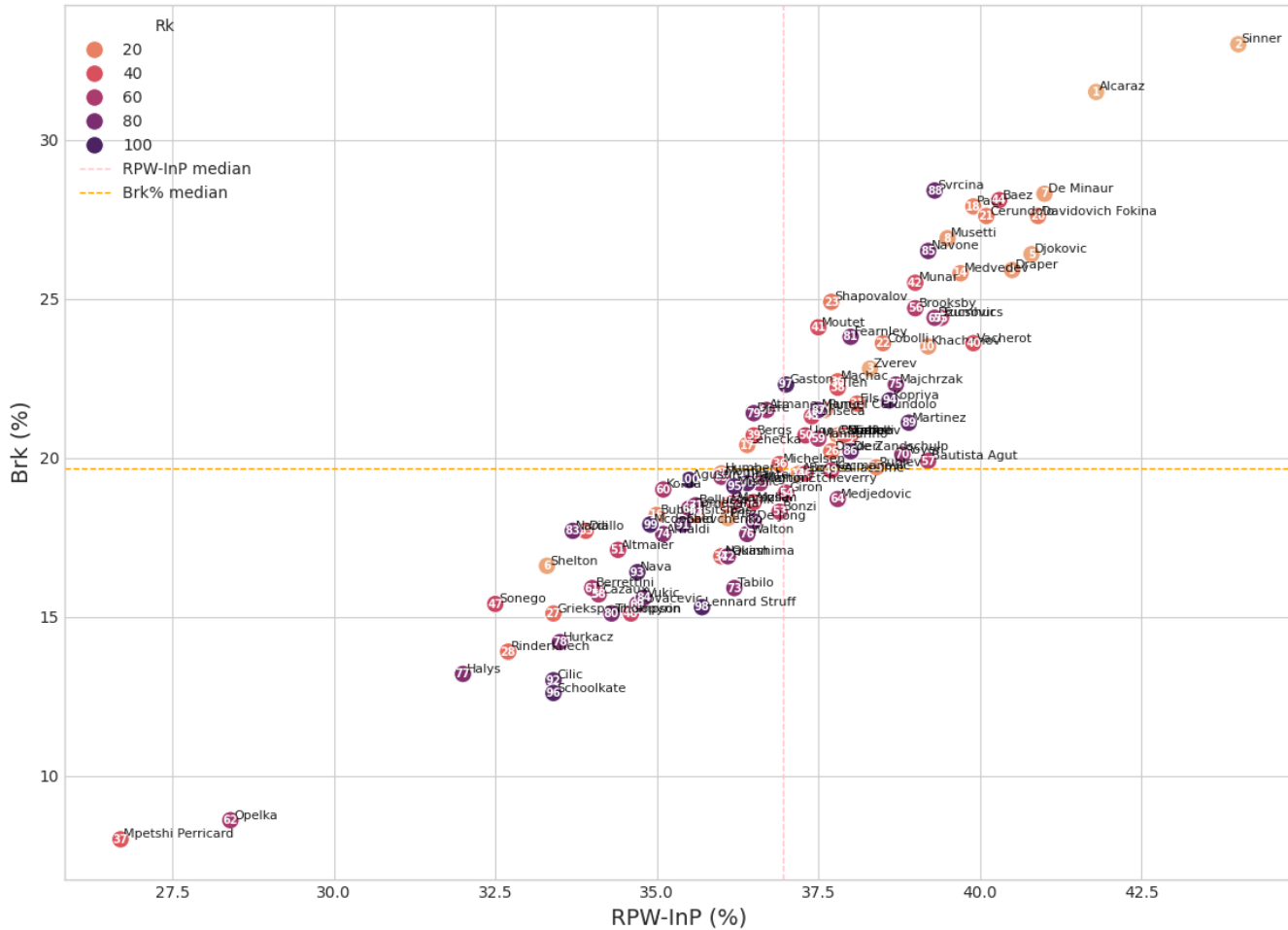


- **SPW vs. SPW-InP**：後者的計算扣除了發球直接得分與雙發失誤，更可以衡量發球發出後的回合能力，通常代表在發球局的進攻性。
- **Ace%**與排名前後沒有絕對的相關性，排名靠前面的選手(顏色偏橘色)未必有更強的發球直接得分能力。
- 擁有好的發球直接得分能力(Ace%)，也不代表在發球局的整體得分率就是高的。
- 網球是對於發球者有利的比賽(發球局更容易取得分數)。

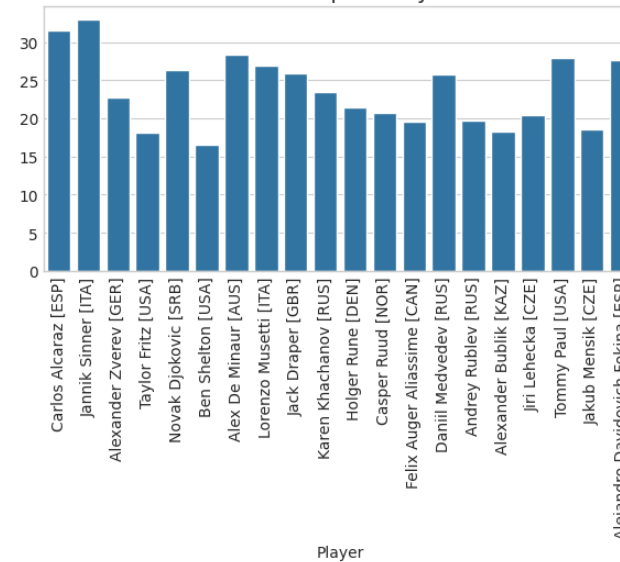


EDA for Return

Return points won (in play) vs. Break (%)



Brk% of Top 20 Players

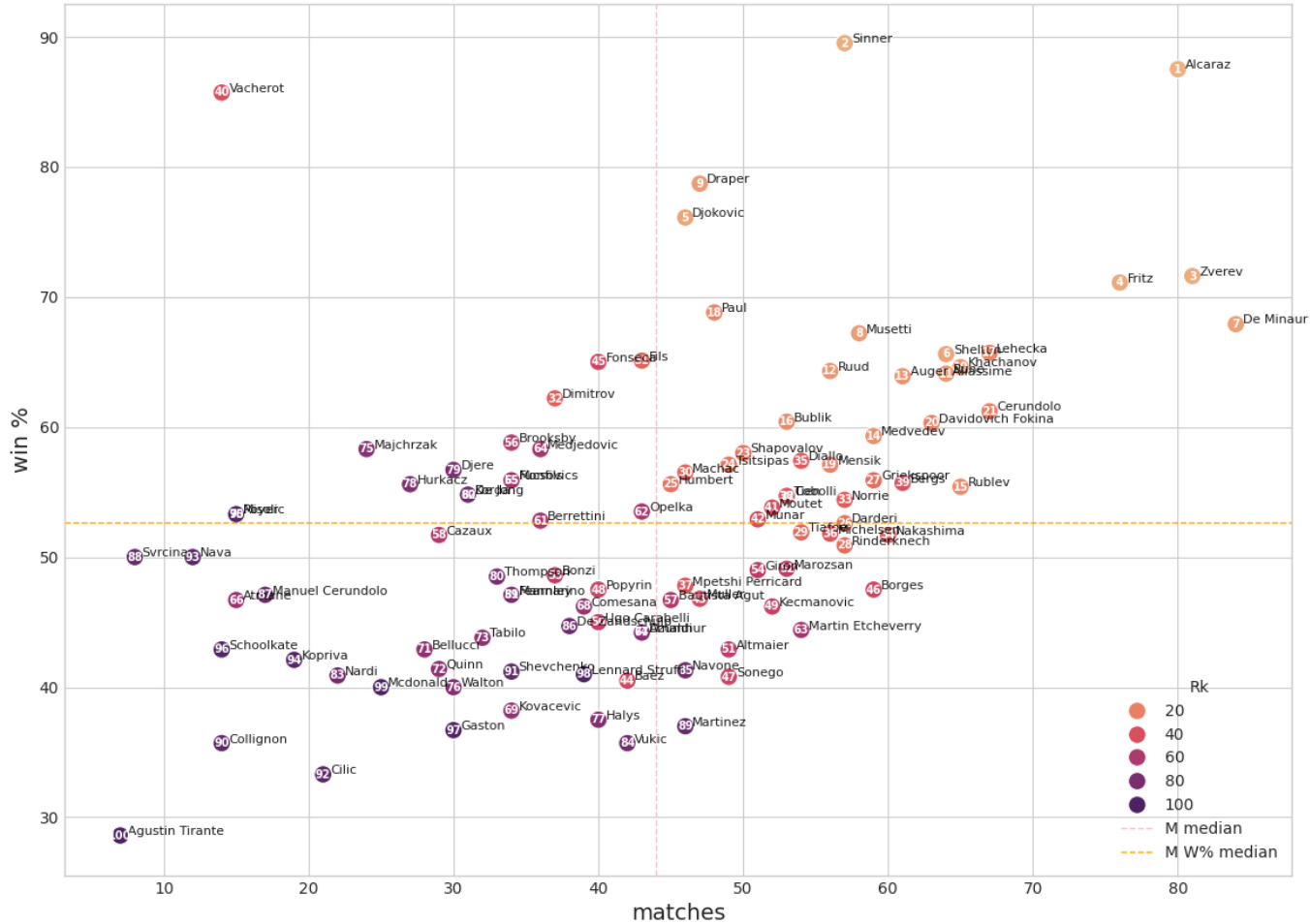


- **RPW vs. RPW-InP**：後者的計算扣除了對手發球直接得分與雙發失誤，更可以衡量接發球進入多拍回合後的底線能力。
- **破發率(Brk%)**與排名前後則有正相關性，趨勢上而言，排名靠前面的選手往往有更好的破發能力，但部分選手有不一樣的得分方式。
- 能夠在接發球局得到更多的分數(RPW-InP)當然就有更強的破發能力。



EDA for Matches

Playing more is the winner? the Number of matches vs. win %

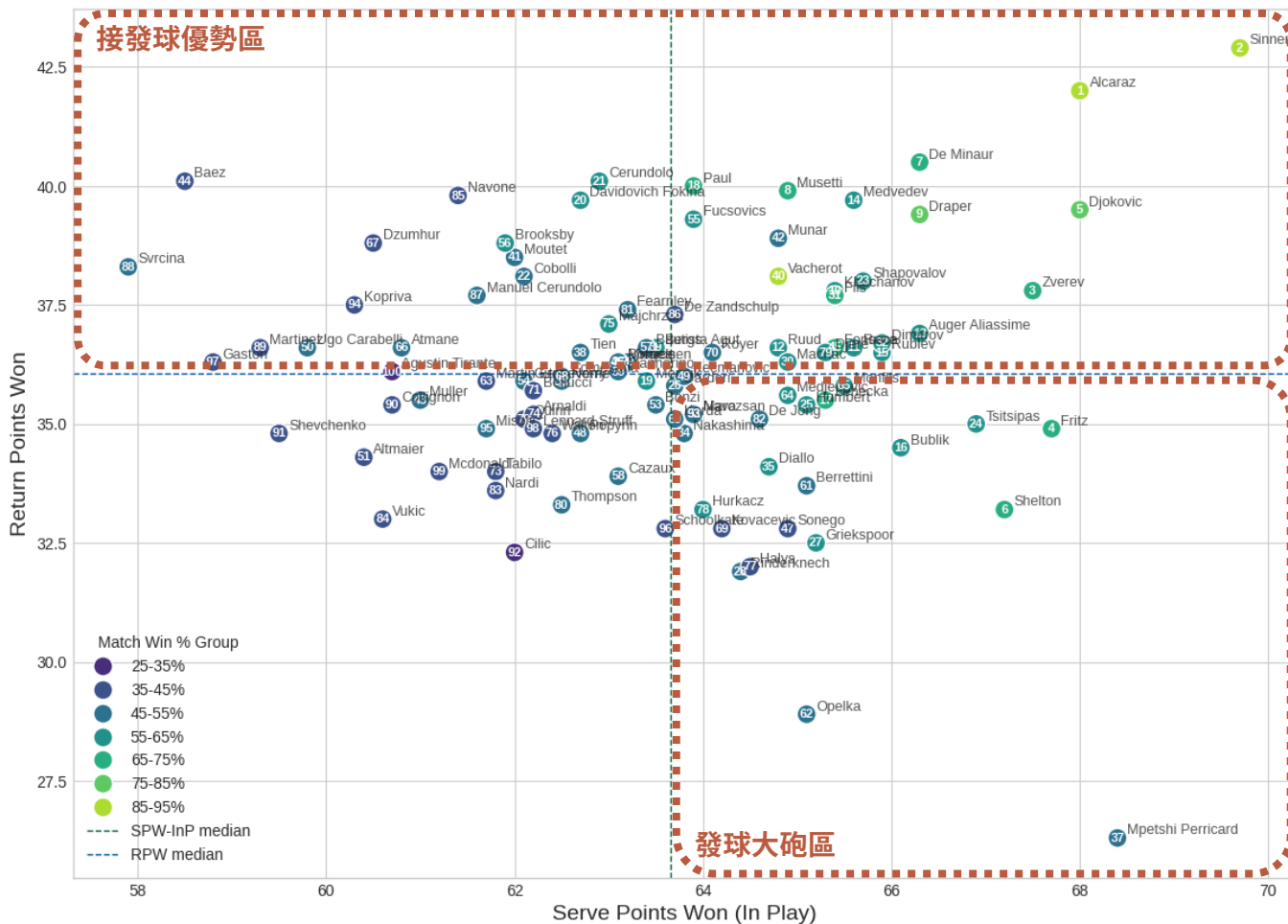


- 排名靠前 → 勝率越高 → 打越多比賽
- 只有極少數選手能以少場數、高勝率贏得大型比賽的排名分數
 - Rk40-Vacherot在2025上海大師賽從挑戰賽連勝八場，最終打贏冠軍戰

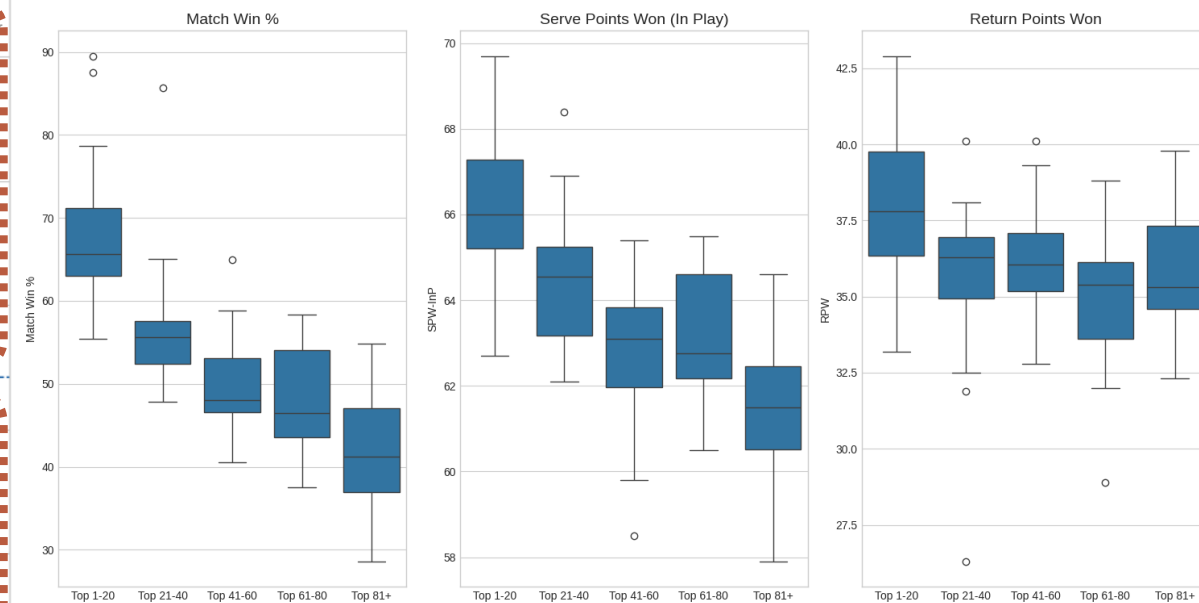


Serve vs. Return with Win Rate

Player Performance Analysis: Serve vs. Return with Win Rate



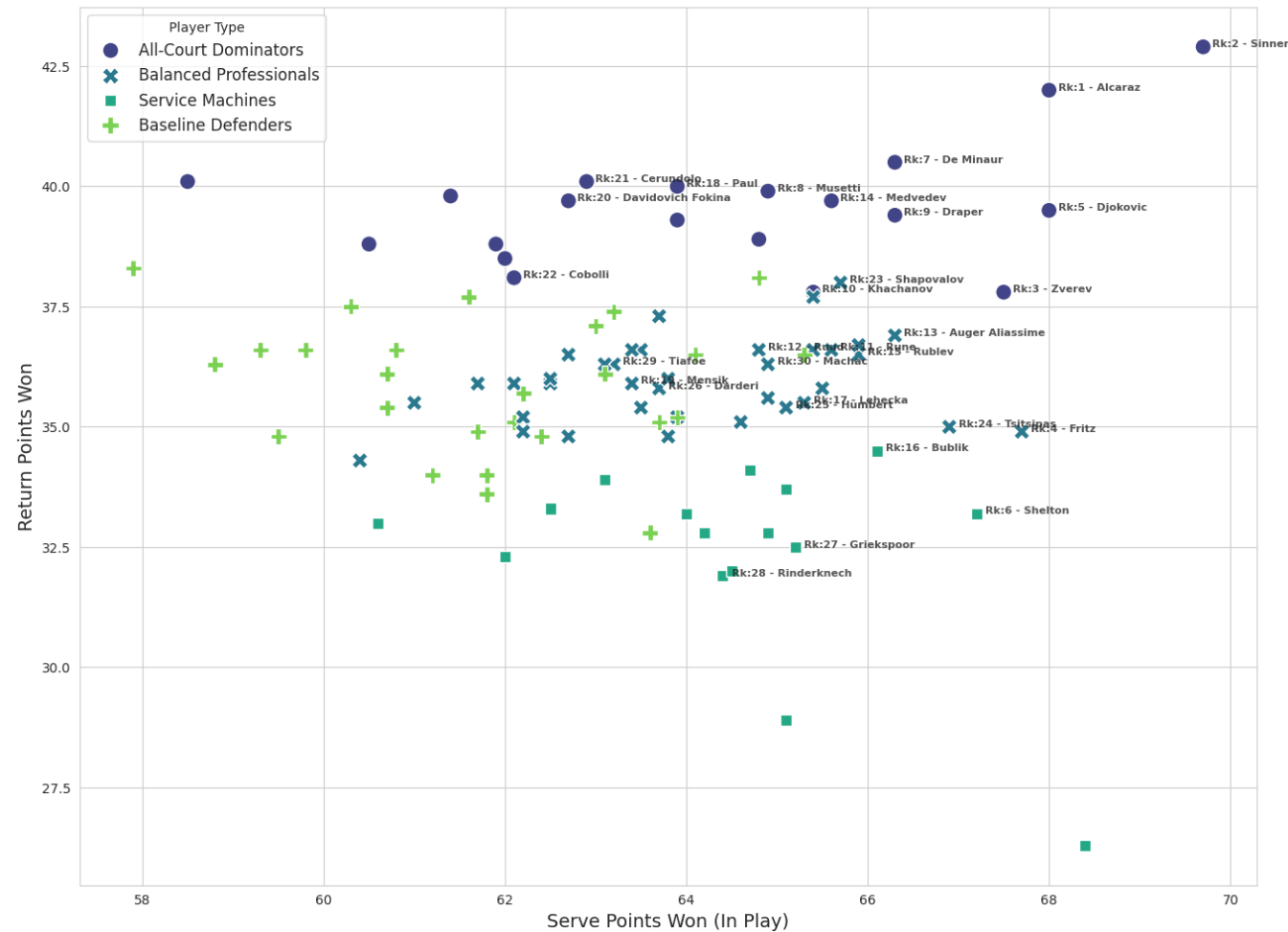
Performance Metrics by Player Rank Group



- SPW vs. SPW-InP、RPW vs. RPW-InP的選擇，結果相差不多，是根據對勝率(M W%)的相關係數挑選影響力大的。
- 頂尖選手的成功模式：攻防平衡或接發球突出
- 最頂尖的選手幾乎都擁有高SPW-InP、高RPW，但排名組別內仍是不同的能力分佈。



Player Classification by Ability Type



‘SPW-InP’, ‘Ace%’, ‘DF%’

發球相關 發球局得分率、Ace得分率、雙發失誤率

‘RPW’, ‘Brk%’, ‘vAce%’

接發球相關 接發球局分率、破發率、對手發球局Ace分率

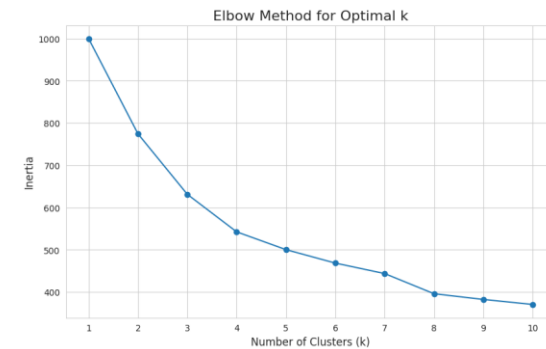
‘Gms’, ‘M’

經驗相關 局數、場數

‘Pts/RG’, ‘PtsW/RG’

效率相關 每場接發球局的得分數、每場接發球局的得分率

除了基本的SPW-InP、RPW欄位之外，再新增一些經驗相關、效率相關的欄位當作分群依據。最後使用K-means分群以及 $k=4$ 作為分群數量。



分爲四個選手類型：

- 全能型統治者 (All-Court Dominators)
- 均衡型中堅 (Balanced Professionals)
- 底線防守者 (Baseline Defenders)
- 發球機器 (Service Machines)

以六個向度來評估、視覺化選手能力值：

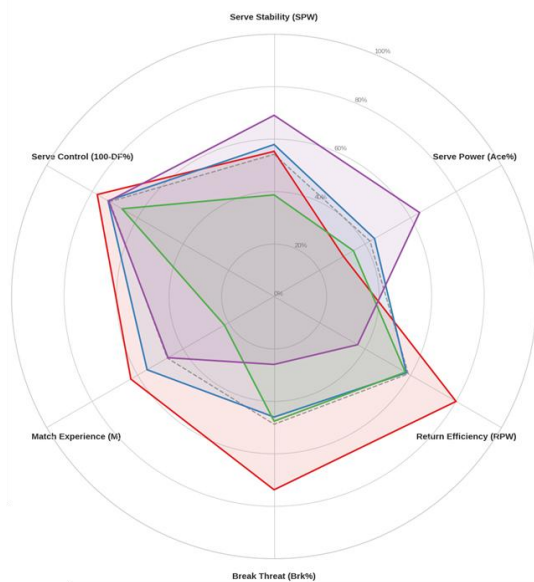
- 'Serve Stability (SPW)'
- 'Serve Power (Ace%)'
- 'Return Efficiency (RPW)'
- 'Break Threat (Brk%)'
- 'Match Experience (M)'
- 'Serve Control (100-DF%)'

Type: All-Court Dominators

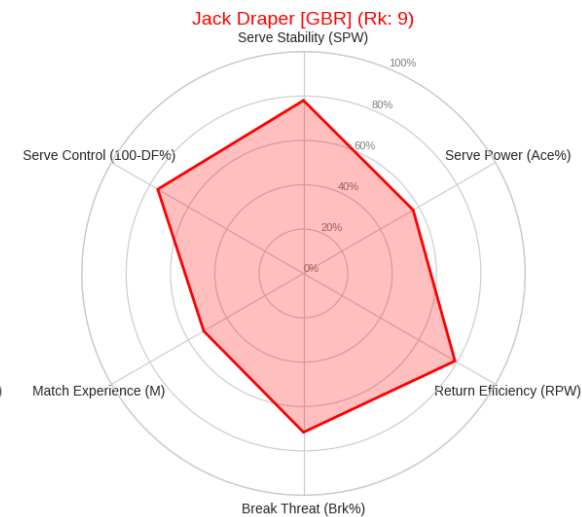
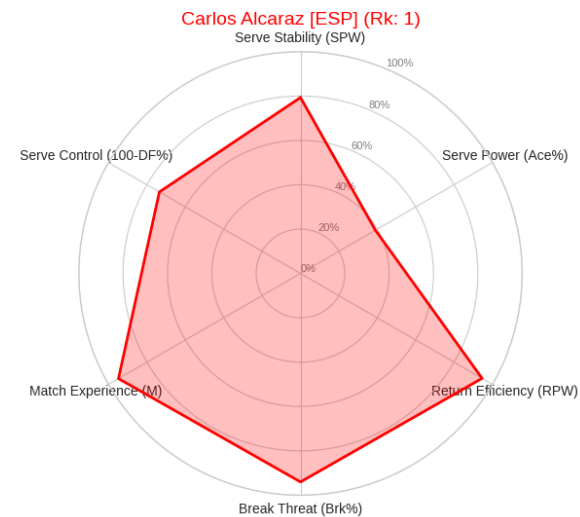
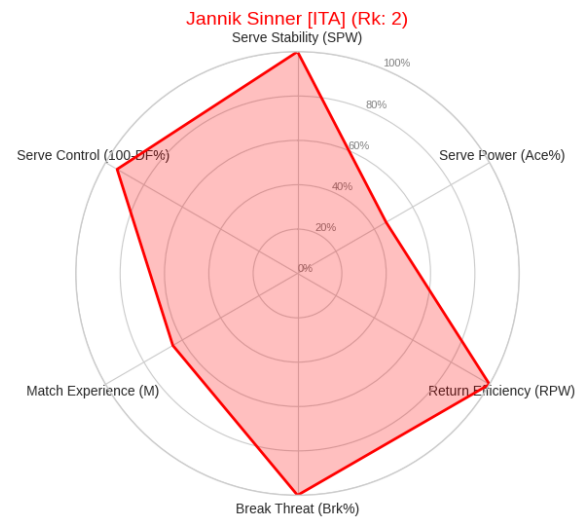
- 特徵洞見：攻防能力無懈可擊，在發球與接發球兩端均屬頂尖。
- 數據趨勢：輪廓面積最大，五項核心指標線條超過整體平均線。



Average Ability Profile of Player Types (with Overall Average Baseline)



--- Overall Average
 --- All-Court Dominators
 --- Balanced Professionals
 --- Baseline Defenders
 --- Service Machines

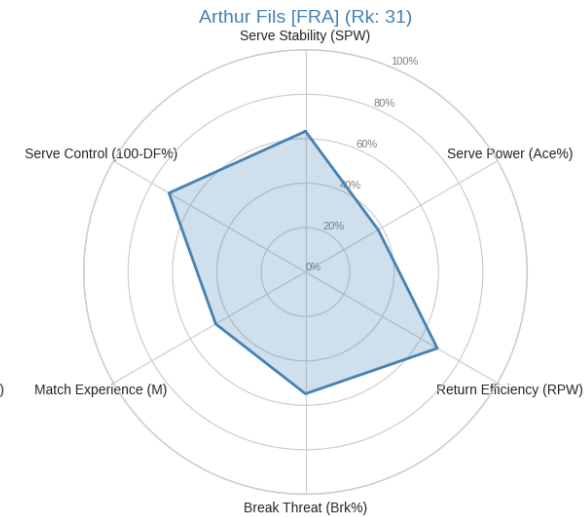
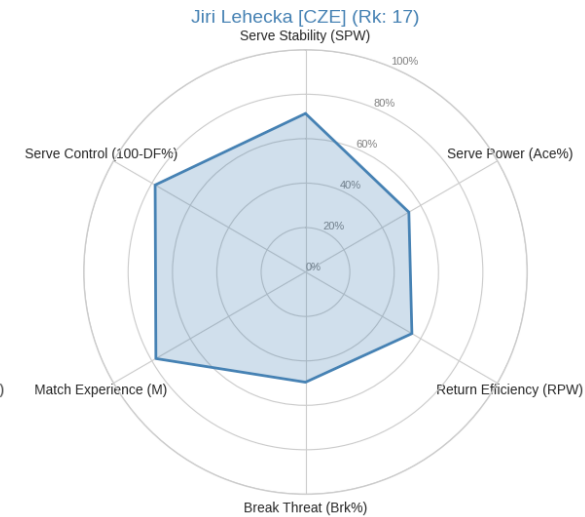
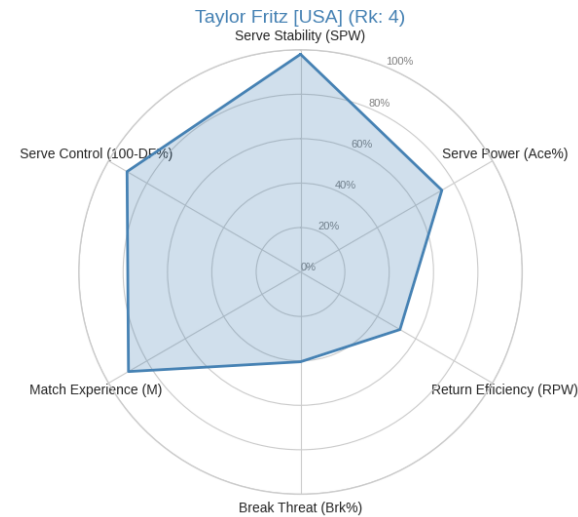
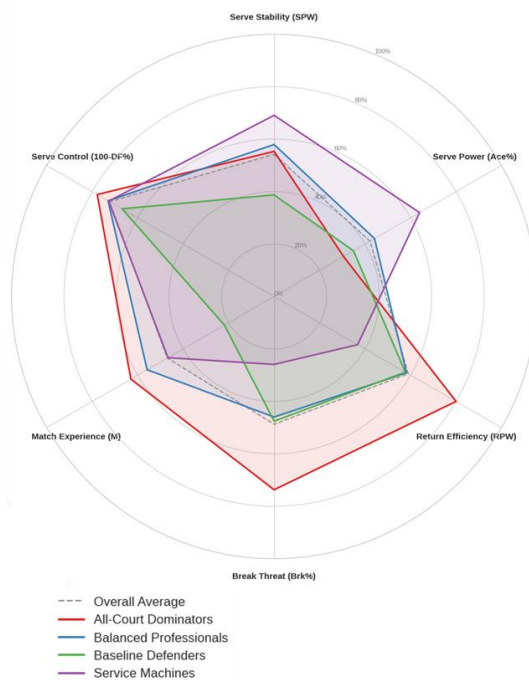


Type: Balanced Professionals

- 特徵洞見：各項指標表現中規中矩，攻防兩端均無明顯長處或短板。
- 數據趨勢：輪廓線條平穩且緊貼整體平均線，是統計學上最中庸的類型。



Average Ability Profile of Player Types (with Overall Average Baseline)

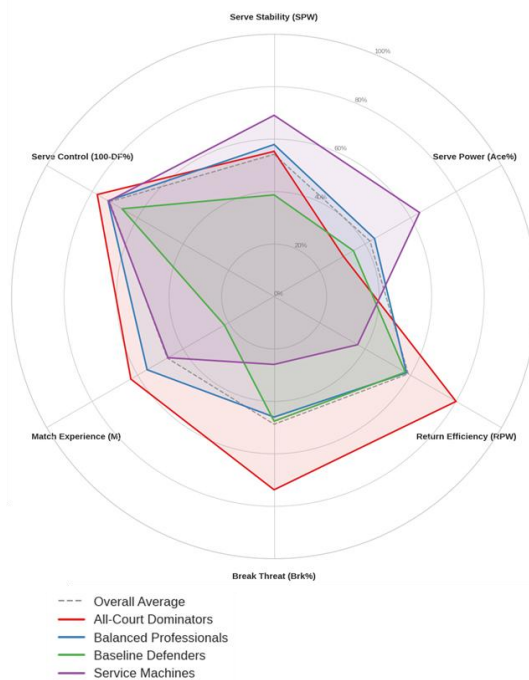


Type: Baseline Defenders

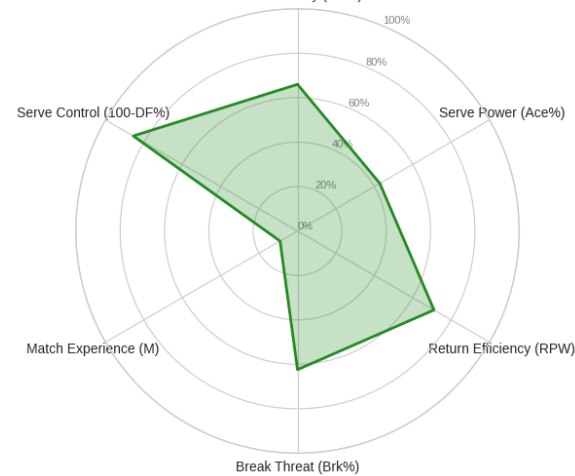
- 特徵洞見：發球力量偏弱，但擁有較好的接發效率與底線控制能力。
- 數據趨勢：輪廓在「接發球效率」和「破發威脅度」軸上較突出，發球力量軸偏低。



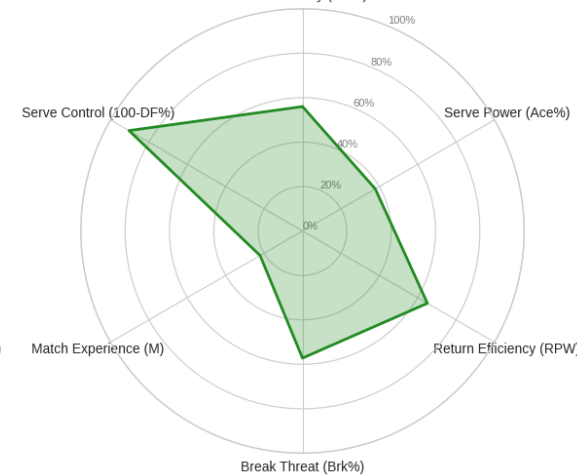
Average Ability Profile of Player Types (with Overall Average Baseline)



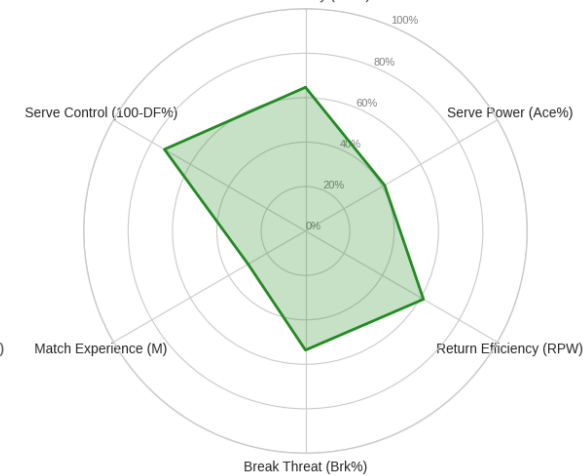
Valentin Vacherot [MON] (Rk: 40)
Serve Stability (SPW)



Kamil Majchrzak [POL] (Rk: 75)
Serve Stability (SPW)



Laslo Djere [SRB] (Rk: 79)
Serve Stability (SPW)

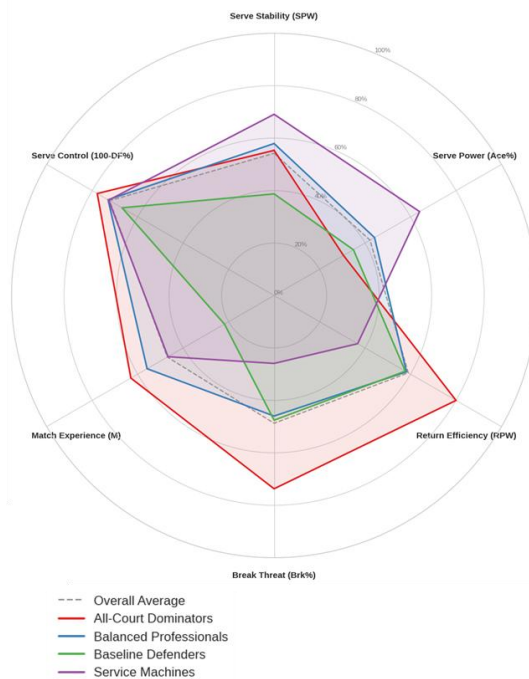


Type: Service Machines

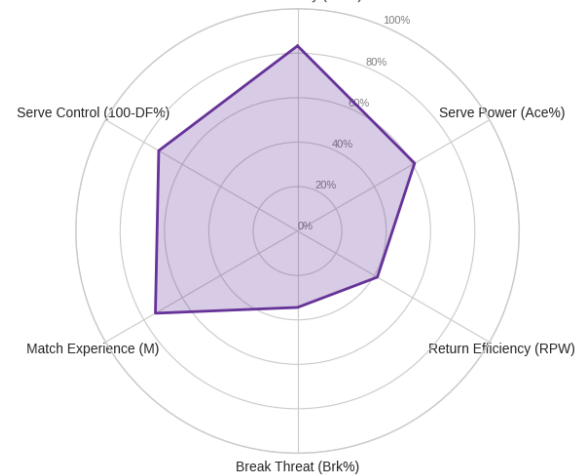
- 特徵洞見：極度依賴力量與速度，Ace 效率最高，但犧牲了接發能力。
- 數據趨勢：圖形在發球力量軸上最爲尖銳突出，接發球和破發軸向圓心凹陷。



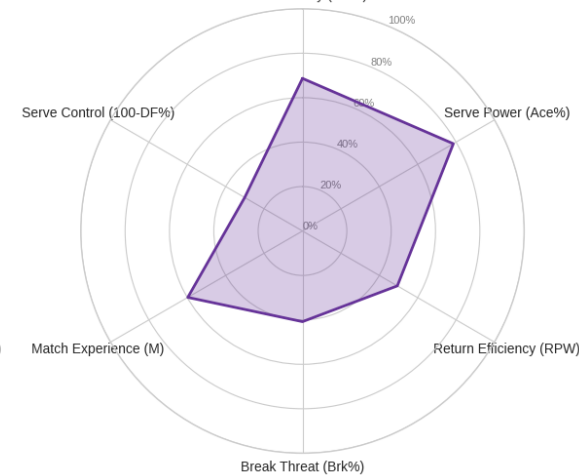
Average Ability Profile of Player Types (with Overall Average Baseline)



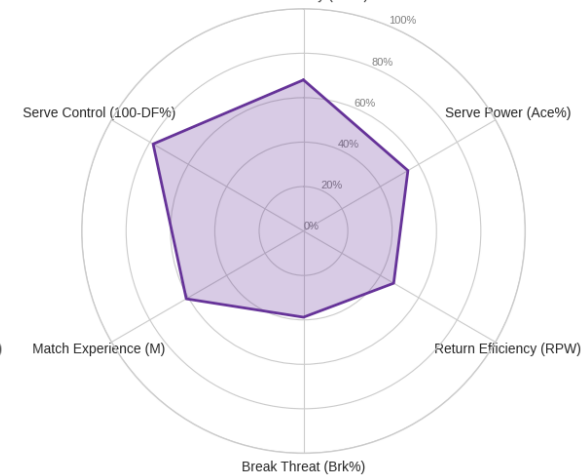
Ben Shelton [USA] (Rk: 6)
Serve Stability (SPW)



Alexander Bublik [KAZ] (Rk: 16)
Serve Stability (SPW)

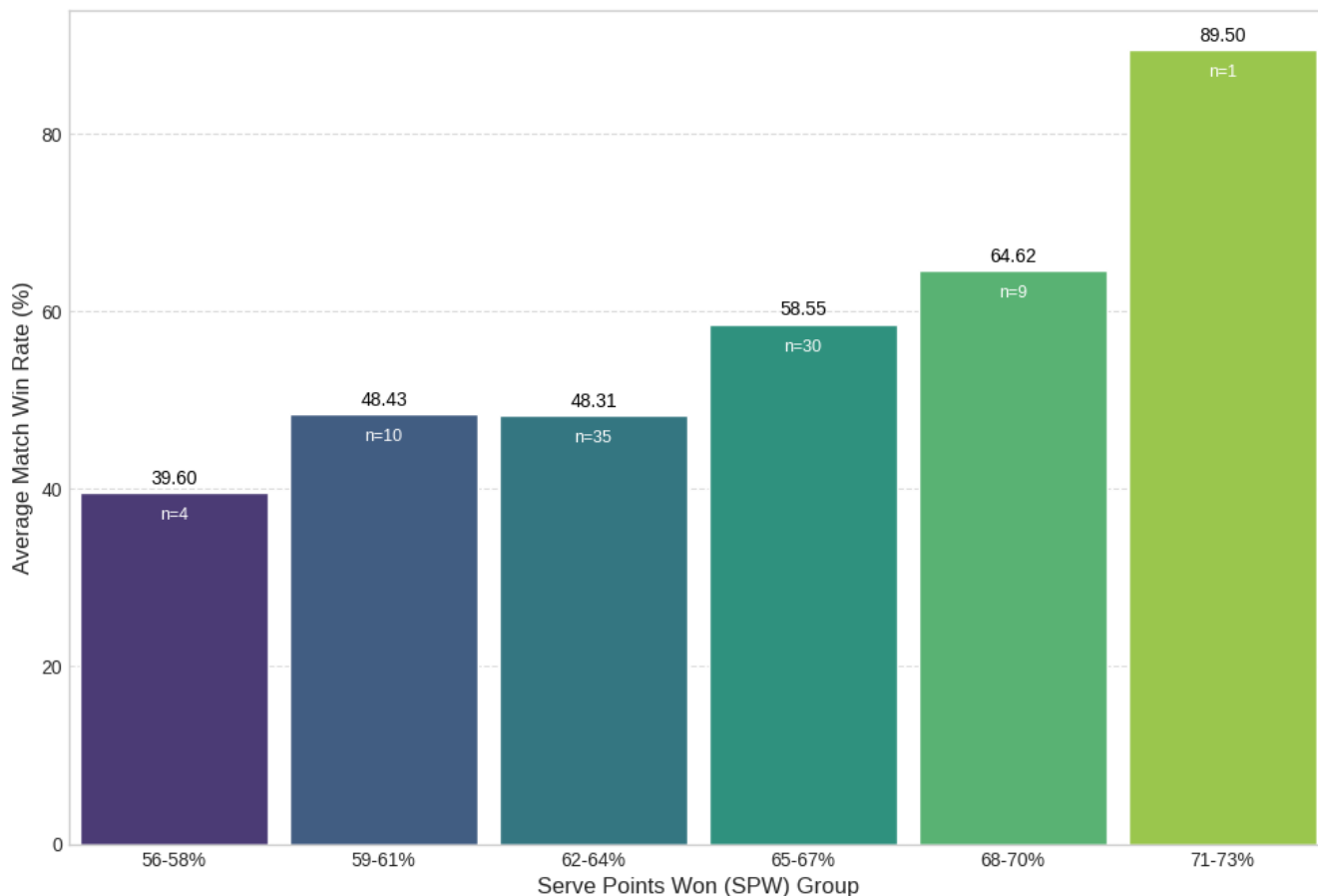


Gabriel Diallo [CAN] (Rk: 35)
Serve Stability (SPW)



Serve to Win: Is There a Tipping Point?

Average Match Win Rate by Serve Points Won (SPW) Group



想要贏球不需要當六邊形戰士！根據我們的分析結果，只要保持發球局、接發球局的穩定性就可以掌握制勝的關鍵了，**那是否存在一個魔法數字，只要突破了就可以大幅提升勝率呢？**

1. 穩步上升的趨勢：隨著 SPW 的提升，選手的平均勝率呈現穩定的上升趨勢。這證實了在發球局的得分能力越強，贏得比賽的可能性就越高的基本假設。
2. 重要的「臨界點」：在 65-67% 的區間，圖表出現顯著的跳躍。該區間的平均勝率達到 58.55%，比前一個區間增加了約 10%。這是一個關鍵洞見：**當在發球局的得分率超過 65% 時，選手的勝率會進入一個更高的層次。**
3. 頂尖選手的統治力：在 68-70% 和 71-73% 的區間，平均勝率持續攀升。特別值得注意的是，71-73% 的區間雖然只有一名選手（n=1），但他的平均勝率高達 89.50%。



Conclusion

1. 發球與接發球是定義現代網球選手風格的兩大基石。
2. 成功將選手劃分為四種具有顯著數據特徵的類型。
3. 「全能型統治者」選手在數據上展現了最全面的統治力。

Insights

- 對教練而言：
 - 可用於分析對手弱點，制定比賽策略。
- 對媒體/球迷而言：
 - 提供了一種新的、更客觀的視角來討論和評價選手。
- 對商業應用而言：
 - 可用於球員潛力評估、賽事轉播的看點包裝。

Future Work

1. **深化數據分析，讓分類更立體** 目前分析的是綜合表現，但下一步可以加入場地（紅土/草地）、對手強弱等數據。這能讓我們看出選手在不同壓力下的風格變化與適應性，讓分類更貼近真實戰況。
2. **打造互動工具，讓成果普及化** 將這份報告做成一個互動式的網站或儀表板。讓教練、球迷等使用者能自行查詢選手資料，觀看他們的能力雷達圖，並直觀地比較不同選手間的戰術優劣，把分析成果變成實用工具。
3. **導入AI預測，讓洞察產生價值** 建立一個AI模型，用我們分出的選手類型和各項數據去預測比賽勝負。這不僅能驗證我們分析的準確性，更能將數據洞察轉化為有價值的賽前情報，真正發揮數據的影響力。

