# Tennis\_Neo4J

BY ARUN KUMAR 17PD05 and AJAY DAMODAR 17PD03

TENNIS TOURNAMENT ANALYSIS USING NEO4J

Data : https://github.com/JeffSackmann/tennis\_wta/ https://github.com/JeffSackmann/tennis\_atp/

Exploring data:

row

{

"l\_1stWon": null,

"loser\_ioc": "USA",

"w\_ace": null,

"loser\_age": "32.2737850787",

"l\_df": null,

"winner\_id": "211901",

"winner\_hand": "R",

"score": "6-3 6-3",

"l\_bpFaced": null,

"winner\_rank\_points": "350",

"match\_num": "001",

"winner\_seed": null,

"best\_of": "3",

"loser\_name": "Jennifer Elie",

"w\_bpFaced": null,

"winner\_ht": null,

"l\_1stIn": null,

"tourney\_level": "S",

"draw\_size": "32",

"surface": "Hard",

"l\_svpt": null,

"minutes": null,

"tourney\_date": "20181231",

"l\_ace": null,

"loser\_rank": "450",

"l\_SvGms": null,

"l\_2ndWon": null,

"w\_1stIn": null,

"loser\_id": "202495",

"winner\_rank": "168",

"winner\_ioc": "MDA",

"winner\_age": "21.1498973306",

"tourney\_id": "2019-W-ITF-AUS-01A-2019",

"tourney\_name": "W25 Playford",

"loser\_rank\_points": "59",

"loser\_ht": null,

"winner\_entry": null,

"winner\_name": "Aliona Bolsova",

"w\_SvGms": null,

"loser\_hand": "R",

"l\_bpSaved": null,

"w\_svpt": null,

"loser\_seed": null,

"w\_1stWon": null,

"w\_bpSaved": null,

"round": "R32",

"w\_df": null,

"w\_2ndWon": null,

"loser\_entry": "Q"

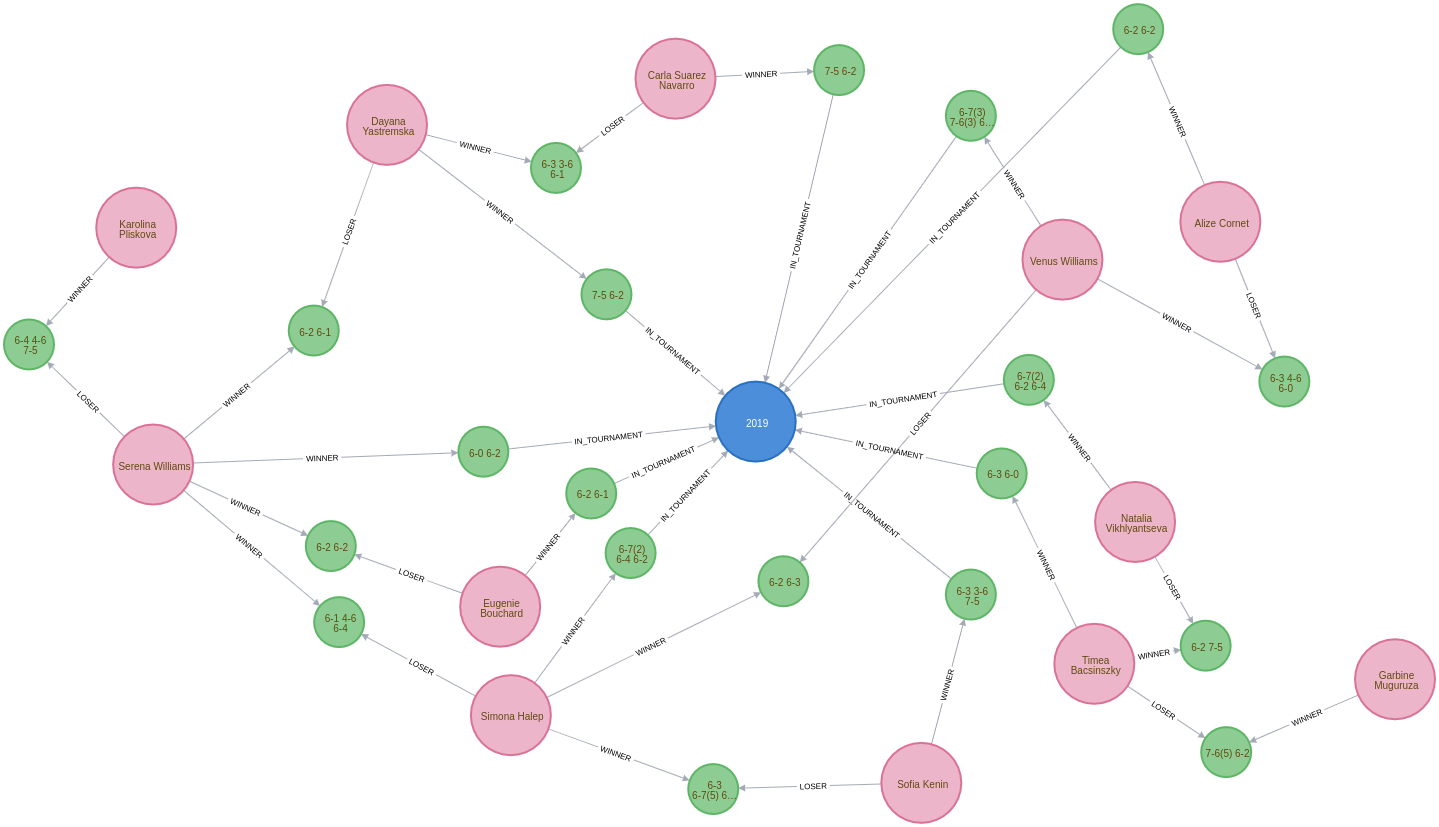
}

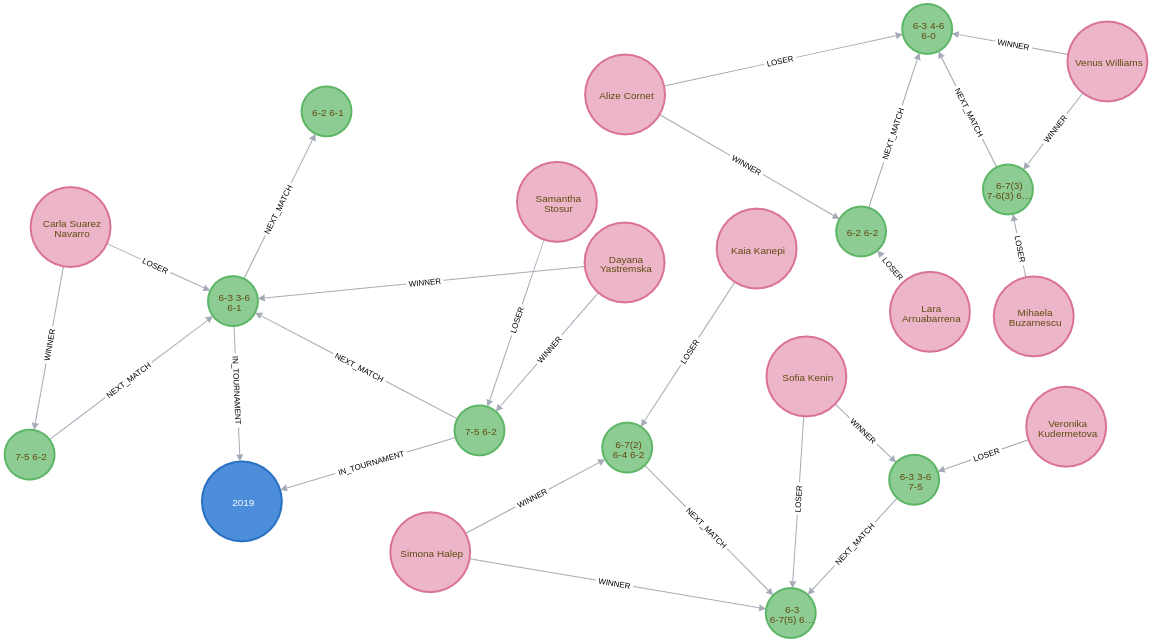
filter the data using the tourney\_name so that we only have matches from the Australian Open. winner\_id and loser\_id will act as the primary keys for our players and we can combine match\_num and tourney\_date as the primary key for matches. winner\_name and loser\_name give us the human readable version of the players and the score property tells us the result of the match.

Constraints :

a unique node property constraint on the Player label, id property and Match label, id property.

a node key constraint on the Tournament label name and year properties





QUERY OUTPUTS:

|  |  |  |  |
| --- | --- | --- | --- |
| **year** | **winner** | **loser** | **score** |
| 2000 | "Lindsay Davenport" | "Martina Hingis" | "6-1 7-5" |
| 2001 | "Jennifer Capriati" | "Martina Hingis" | "6-4 6-3" |
| 2002 | "Jennifer Capriati" | "Martina Hingis" | "4-6 7-6(7) 6-2" |
| 2003 | "Serena Williams" | "Venus Williams" | "7-6(4) 3-6 6-4" |
| 2004 | "Justine Henin" | "Kim Clijsters" | "6-3 4-6 6-3" |
| 2005 | "Serena Williams" | "Lindsay Davenport" | "2-6 6-3 6-0" |
| 2006 | "Amelie Mauresmo" | "Justine Henin" | "6-1 2-0 RET" |
| 2007 | "Serena Williams" | "Maria Sharapova" | "6-1 6-2" |
| 2008 | "Maria Sharapova" | "Ana Ivanovic" | "7-5 6-3" |
| 2009 | "Serena Williams" | "Dinara Safina" | "6-0 6-3" |
| 2010 | "Serena Williams" | "Justine Henin" | "6-4 3-6 6-2" |
| 2011 | "Kim Clijsters" | "Na Li" | "3-6 6-3 6-3" |
| 2012 | "Victoria Azarenka" | "Maria Sharapova" | "6-3 6-0" |
| 2013 | "Victoria Azarenka" | "Na Li" | "4-6 6-4 6-3" |
| 2014 | "Na Li" | "Dominika Cibulkova" | "7-6(3) 6-0" |
| 2015 | "Serena Williams" | "Maria Sharapova" | "6-3 7-6(5)" |
| 2016 | "Angelique Kerber" | "Serena Williams" | "6-4 3-6 6-4" |
| 2017 | "Serena Williams" | "Venus Williams" | "6-4 6-4" |
| 2018 | "Caroline Wozniacki" | "Simona Halep" | "7-6(2) 3-6 6-4" |
| 2019 | "Naomi Osaka" | "Petra Kvitova" | "7-6(2) 5-7 6-4" |

|  |  |  |  |
| --- | --- | --- | --- |
| **player.name** | **firstLoss.year** | **firstWin.year** | **theWait** |
| "Marat Safin" | 2002 | 2005 | 3 |

|  |  |  |  |
| --- | --- | --- | --- |
| **player.name** | **firstLoss.year** | **firstWin.year** | **theWait** |
| "Kim Clijsters" | 2004 | 2011 | 7 |
| "Na Li" | 2011 | 2014 | 3 |
| "Maria Sharapova" | 2007 | 2008 | 1 |

