

Exploring ATP Tournament Results

Lyle Paukner

Department of Statistics & Actuarial Science
University of Iowa

October 25, 2023



- The Association of Tennis Professionals is the governing body for men's professional tennis, and organizes circuits of tournaments at varying levels
 - ATP Tour
 - ATP Challenger Tour
 - ATP Champions Tour

- Consists of approx. 60 tournaments yearly

Level	# (2023)	Max Points
Grand Slam	4	2000
ATP Finals	1	1500
ATP Masters 1000	9	1000
ATP 500	13	500
ATP 250	39	250

The ATP Tour

- Tournaments are played on a variety of surfaces (hard court, grass, clay)
- Grand Slams and Masters tournaments are stand-alone events
- Usually multiple ATP 500 and 250 events happen simultaneously at various locations
- Ranking points are awarded for participation in the tournaments, with more points being awarded the further into a tournament a player gets.
 - Points earned count towards a player's ATP rating for 12 months
 - Maximum points coincide with the size of the draw at a tournament (and prestige and prize money).

Tournament Structure

- ATP tournaments have 32, 64, or 128-player draws, with the top quarter of players being awarded seeds
 - Unseeded players are randomly assigned to the bracket
- Grand Slams have full 128-player draws
- Masters events have either 64 or 128-player draws
 - The top 8 or top 32 seeds received a first-round bye, respectively (in 2023)
- All ATP 250 and most ATP 500 tournaments have 32-player draws
- Grand Slam events are played as best of 5 sets, while all other tournaments are best of 3

Tournament Seeding

- Seeding for tournaments is done primarily based on ATP ranking points
- Most tournaments use players' current ratings, though some look at ranking based on a set number of previous tournaments
 - Helps deal with players who have played far more/fewer tournaments in past 12 months (e.g., players returning from injury)
- Some tournaments include additional criteria in seeding
 - For example, Wimbledon adds extra weight to players performance in tournaments held on grass courts

Questions of Interest

- Do tournament results follow what we expect to see, especially at the smaller tournaments where participation from top players is lower?
- Is there any evidence of a middle-seed anomaly at the small tournaments?
 - If so, why?
 - Could some players take advantage of this by electing to go to one tournament over another?
- How does player strength by seed compare to other tournament types?

The Data

- To answer these questions, we need information on the tournaments (draw size, date), matches (players, seeds, round, ATP rankings) data and strength ratings for players (ELO, date)
- Fortunately, a few other tennis and data enthusiasts compiled much of this information:
 - Tournament and match data was retrieved from a GitHub repository: [Jeff Sackman - Tennis ATP](#)
 - ELO ratings were scraped from [ultimatetennisstatistics.com](#)

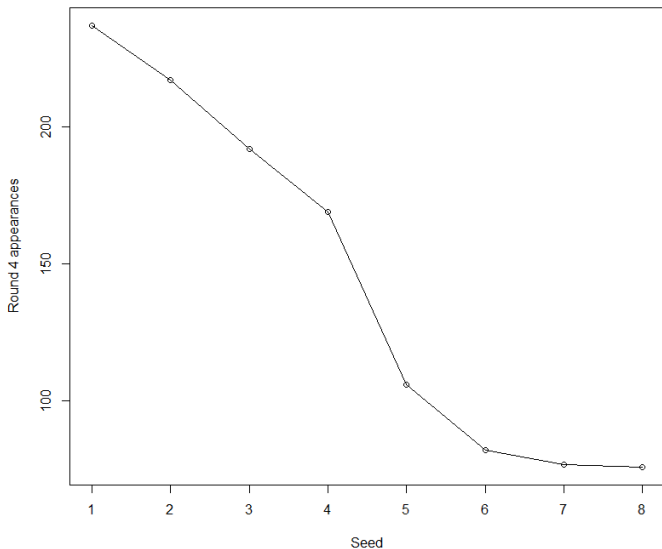
- Working with data from GitHub was straight-forward
 - Download match data from each year back to 2013, combine the files, and filter to desired columns
- Gathering the ELO ratings was more difficult
 - Use python to pass parameters to scrape the web page
 - Updated ELO ratings on the site are calculated/available on a weekly basis, so many dates needed to be iterated over
- Finally, merge the data into a single usable data frame
 - Requires matching on date and player names (which are not always identical in these two data sets)

- The end result is a data set containing details from 1376 tournaments and nearly 30,000 matches
- The data set used for analysis contains these columns:
 - Tournament ID
 - Tournament Name
 - Tournament Level
 - Draw Size
 - Date
 - Match Round
 - Player Name
 - Seed
 - ATP Ranking Points
 - ELO
 - W-L Outcome

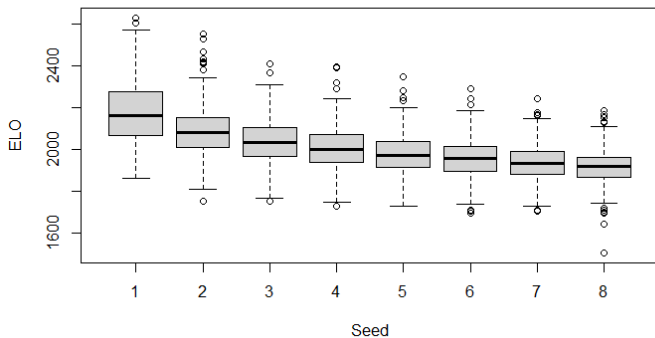
Results for 32-player draws

Seed	Round 3	Round 4
1	328	237
2	305	217
3	298	192
4	281	169
5	216	106
6	200	82
7	174	77
8	178	76
(None)	1972	820

Results for 32-player draws



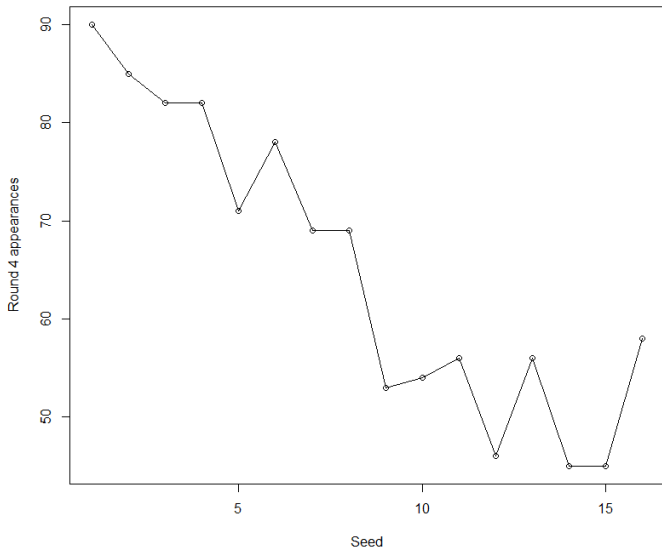
Results for 32-player draws



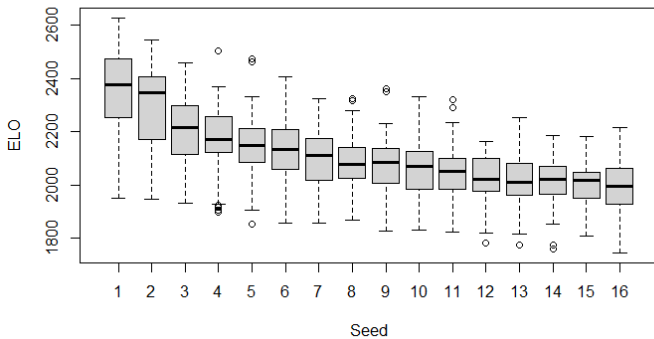
Results for 64-player draws

Seed	Round 3	Round 4
1	90	80
2	85	74
3	82	50
4	82	60
5	71	44
6	78	52
7	69	42
8	69	41
9	53	23
10	54	32
11	56	20
12	46	23
13	56	21
14	45	13
15	45	15
16	58	23
(None)	684	251

Results for 64-player draws



Results for 64-player draws



- Nothing seems out of the ordinary with the smallest ATP tournaments
- That is *not* the case with mid-size tournaments
 - Perhaps more older players with high ELO but low ATP ranking appear at these tournaments than the smaller ones
 - Does seeding at some mid-level tournaments differs from standard practice?
 - Does surface have an impact (e.g., is there a much greater proportion of grass/clay tournaments in this group)?

Future Things to Consider

- Include more data
 - Look at results pre-"Big Three"
- Explore outcomes in tournaments where seeding perfectly matches ELO
- Compare ATP results against WTA results
 - Anecdotally, WTA results have been more variable in recent years