Travel in Professional Women's Tennis
(Review and Analysis of the 2018 season)

Kiril Traykov
PGDV 5210: Major Studio 2
Professor Christian Swinehart
May 12, 2019

TABLE OF CONTENTS

Background and Data Analysis Chapters

I.   Introduction:

Tennis is unique from other sports with the fact that there are no "home games" for the players and all tournaments take place in different cities throughout the world. Players typically spend 10+ months of the year on the road and can easily accumulate fatigue that could negatively influence their performance. Medical professionals state that jet lag, sleep interruptions, and altitude changes, among other consequences, can resulting in miscoordinated performance if there is no adequate time for adjustment. The magnitude of this effect is unknown, but the degree of travel is addressed in this research through data analysis and visualization.

The author examines the 2018 full tennis season by the Women's Tennis Association and focuses on the full schedule of the year-end top 30 players. The data and the findings are presented from two angles: (1) unfiltered, which is inclusive of every tournament, and (2) filtered, which examines only the cases when players competed in back-to-back tournaments in consecutive weeks. The reason for this dual inspection is because travel from one tournament to another in successive weeks is the result of a different planning and decision-making process, and there could be a sub-trend in the data.

Based on the data, this research outlines the "travel efficiency" of top players (necessary kilometers to earn 1 ranking point) and finds that there is a very strong efficiency at the top. In order to be in the top 10, a player needs to 1) play at least 15 tournaments a year and 2) achieve an efficiency rate of 1 point for every 20km traveled in her best 17 tournaments used for ranking. It isn't surprising to find the best performers at the top, but this research confirms that the best players maintain a very focused schedule in which they play only the biggest tournaments to gain the maximum amount of points for the least amount of physical expenditure, which includes travel. Top players rarely deviate from the "top tournament path", which highlights that traveling for more tournaments than necessary isn't part of how the tour works at present. This research also finds indirect evidence that consecutive-week travel leads to diminishing returns.

II.   The WTA Tour:

The Women Tennis Association (WTA) tour comprises of 58 tournaments[1] played around the world over 46 weeks during the calendar year[2]. All 58 tournaments award prize money and ranking points to the competitors under a positive linear relationship: the higher the prize money, the higher the points for progressing deeper into the tournament[3]. The rationale for that is that a higher prize pool will attract more players of higher caliber and the competition will be more intense with the tournament consisting of more rounds of direct elimination. This infers that the prestige of a given tournament is correlated with the prize pool, although other factors such as tournament history, years of existence, and court facilities also have an important role in the classification of tournaments from the players' perspective.

The overall prize money that the 58 tournaments collectively awarded in 2018 was $165 million[4]. These 58 tournaments are divided into five categories with a sixth one for the Year-End Finals where only the top 20 from the year can compete divided in two tournaments. The exact categories are:

Category	No.of tournaments	Total Prize Pool for Each
Grand Slam	4	$18.8M - $25.3M
Premier Mandatory	4	$6.7M - $8.7M
Premier 5	5	$2.7M - $3.4M
Premier	12	$799K - $1M with Dubai as an exception offering $2.6M
International	30	$250K
Year-End Finals	2	$7.0M for players ranked No.1-8 and $2.4M for players ranked No.9-20
Total:	58	

Grand Slams: These are the 4 most prestigious tennis tournaments that have legendary status: the Australian Open, the French Open (Roland Garros), Wimbledon, and the US Open. They have 128 participants in direct elimination, which requires a player to win 7 matches in order to win the tournament. The Grand Slams are played over 2 week periods.

Premier Mandatory: These are the next 4 most prestigious tournaments, which the WTA has made mandatory for the players in order to ensure the high-profile status of the tournaments. They are Indian Wells, Miami Open, Madrid Open, and China Open. Indian Wells and Miami feature 96 players and are both played over 10 days' period. The top 32 players receive a "bye" in the first round and join the competition from the second round. This means that a top 32 player would need to win 6 matches in order to win the tournament, but a player outside the top 32 would need 7 victories. The Madrid Open and China Open have 64 players at the start and the winner needs to win 6 matches, regardless of rank. Both tournaments are played over a single week.

Premier 5: This is a group of 5 prestigious tournaments that offer half of the prize money of the Premier Mandatory ones. They are in Doha, Rome, Montreal, Cincinnati, and Wuhan. They are not mandatory and a player can decide to compete if she likes the facilities, the surface, or the location. This group of tournaments features 56 players and the top 8 are given a "bye" in the first round. This means that a top 8 player will need to win 5 matches in order to win the tournament, while a player outside of the top 8 needs to win 6. These tournaments are played during the course of a single week.

Premier: This is a group of smaller tournaments, which take place in Brisbane, Sydney, St. Petersburg (Russia), Dubai, Charleston, Stuttgart, Birmingham, Eastbourne, San Jose, New Haven, Tokyo, and Moscow. They can have different draw sizes ranging between 28, 32, or 56 players so this group has variations as to how many matches a player needs to win to become the champion. The minimum is 4, the maximum is 6. Yet, they award the same number of points. The differences in draw size are mostly due to the number of available tennis courts in the tournament's complex and the requirement to finish play within a single week.

International: These are tournaments that happen around the world at places, which are trying to popularize tennis as a sport. They often change year after year due to sponsorship withdrawals, lack of leadership, or waning interest. They are organized at tennis club or university facilities and do not involve significant infrastructure investments. There are many instances when a new tournament will apply for a license and get approval after the country has seen a few successful star-like players in the recent 5 years (e.g. the Bucharest Open in Romania, starting in 2014, when Simona Halep finished as No.2 having ascended in the rankings in the years prior to that). However, there are also a few

tournaments in this category, which have existed for decades (e.g. Linz, Luxembourg), but keep its small format due to limitations (e.g. indoor courts). All these tournaments feature either 28 or 32 players and wining them requires 4 or 5 wins depending on player's entry rank. They award the same amount of points and are played during a single week.

Year-End Finals: These are two tournaments reserved for players who finish in the top 20. The first one is only for the players ranked No.1-8 and takes place in Singapore, while the second one is for players ranked No.9-20 and takes place in Zhuhai, China. These tournaments have a round-robin format before the semifinals and the final and have a generous prize pool (per number of players participating) as an incentive to qualify for them.

To understand the players' motivations in constructing their schedule, it is interesting to look at the overall prize money (of $165M) split by category:

Figure 1: Tournament category share from the overall prize money in the WTA tour:
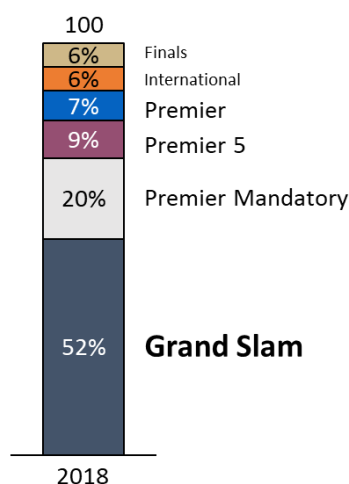
This graph clearly shows that Grand Slams are the bread and butter for earning a living as a professional tennis player. Therefore, most players try to pace and organize their schedule in a way that leads them to peak exactly during those Grand Slams, which are the most coveted trophies.

III.   The WTA Calendar:

In making its calendar, the WTA needs to **reconcile two dimensions: court surface and tournament location.**

Court Surface: There are 4 court surfaces - outdoor hard courts, clay courts, grass courts, and indoor hard courts[5]. The indoor hard courts are typically a carpet that is laid over arenas that host other sports (e.g. at Madison Square Garden in New York City). If a player needs to switch between court surfaces every week, she will lose the ability to play effectively on any of them as the body learns to adapt to a single surface over time. For example, sliding on clay courts is a skillful ability to retrieve balls and move around the court, while that is nearly impossible (and impractical) to replicate on the other court types. Attempting to do the same on grass may result in a fall and injury as the surface can be very slippery, while hard courts require plain running. Thus, for a player to remain healthy and competitive, it is necessary to play on a single-type court for at least several weeks at a time.

Tournament Location: Transitioning players gradually from one geography to another is another key element for keeping them focused and committed to the tour[6]. If a players needs to travel to different world regions every week, she will quickly build fatigue and exhaustion, which will lead to difficulties in realizing one's potential and playing with a fresh mind[7]. In particular, the Hospital for Specialty Surgery (HSS)[7] points that severe jet lag can affect a player's performance quite negatively. Travel fatigue can disrupt a player's normal routine, including sleeping and eating patterns and internal biorhythm. This can create a range of adverse effects ranging from decreased concentration to drowsiness to delayed reaction time to delayed cognition. The disruption to the player's bodies becomes more drastic with the number of time zones crossed - players can be slower at the start of the match and take longer to establish their rhythm. This can lead to more errors on the court and predispose the player to possible injury.

The first-hand experience of players also confirms that. CNN has reported how cross-country flights and travel delays affect players[8]. It is noteworthy to point that travel is seen as an adversary not only when it is international, but domestic as well. In the month of March, the players need to travel from Indian Wells, CA, to Miami, FL (two of the four Premier Mandatory tournaments), which are scheduled back-to-back. This coast-to-coast trip is described by the 2011 U.S. Open winner Samantha Stosur as "one of the worst ever"[8]. Sports journalist Nick McCarvel covering the WTA Tour further elaborates (*see full interview in Appendix*) that this particular leg of the tour is a change of climate conditions: from one elevation to another and from dry desert air to humidity.

Then, just 9 days after the Miami Open, there is an International-category tournament in Bogota, Colombia, for which many players have extensively remarked that it is a taxing change in altitude (at 8,860 feet), which leads them to experience dizziness on court, while running and breathing. The 2018 champion in that tournament, Anna Karolina Schmiedlova, offered the following comment in her post-match interview: "I played in Bogota one time before, I lost in first round […] This time when I came here I prepared longer for the altitude […] "[9]. Facing such conditions, it comes to no surprise that none of the top 30 year-end players chose to travel to compete in Bogota in 2018. This is one of the advantages of top players: they can cherry-pick their battlegrounds as they have proven themselves in earlier tournaments.

Despite the curveballs of travel, Nick McCarvel reveals from his experience interviewing players that the travel component is rarely talked about. He believes that the tennis players have become consummate professionals. They accept it as part of their job and do their best to be professional about it - staying hydrated, sleeping the hours they should in the time zone they have arrived in, and making other necessary adjustments. Nick shares that he has heard players trying to exercise in some way -- having a hit, going for a run or to the gym - when they arrive in a new place. The basic strategy is to "power through" the new time zone they have arrived in and stay up till bedtime in order to adjust more quickly.

When one hears them discuss travel, it is because of outstanding circumstances (as the CNN article8): delayed flights, missed connections, lost baggage or falling ill while in transit. Nick points that they are always very attentive to their schedules, their bodies and how they can best perform on the court itself. Most top players would arrive 4, 5, or 6 days before a Grand Slam's first day's play, while at a Premier Mandatory event, it would be usually 2, 3 or 4 days. Nick reports that it is a rare sight for a player to arrive only one day before they are set to play.

Taking such precautions into consideration, the WTA has developed a calendar, which attempts to be less disruptive, reduce stress (the "wear and tear"), and provide more structure for the players. In fact, it could be argued that the WTA has molded a few mini-tours moving from one geography to another, played on one type of court surface at a time.

The approximate sequence is:

Exhibit 1: The WTA calendar by court surface and world region:

	January	February	March	April	May	June	July	August	September	October	November	December
Court Type	Outdoor hard	Indoor hard	Outdoor hard	Clay	Clay	Grass	Grass	Outdoor hard	Outdoor hard	Outdoor hard + Indoor hard	**Off-season**	**Off-season**
World region	Australia	Middle East & Russia	USA	EU	EU	The U.K and EU	The U.K.	USA & Canada	USA → Asia	Asia & EU & Russia	**Off-season**	**Off-season**

➤ January: The season starts in Australia & New Zealand with warm-up tournaments on outdoor hard courts for the Australian Open

➤ February: After the Australian Open, the tour moves to Doha & Dubai in the Middle East and St. Petersburg in Russia, which are played on indoor courts. However, there are also a few International category tournaments dispersed around the world during this month in Taipei, Budapest, and Acapulco. Thus, the month of February can differ a lot in terms of travel schedule for the players.

➤ March: The two Premier Mandatory tournaments of Indian Wells and Miami Open are played on outdoor hard courts in the USA.

➤ April: The tour moves to Europe in preparation for the French Open (Roland Garros) where all tournaments are on clay courts (Lugano, Stuttgart, Madrid, Rome, Prague, Strasbourg). However, there are a few International category tournaments outside of Europe in Monterey, Bogota, and Rabat which are also on clay courts, but substantially far away from the main tour.

➤ May: The French Open (Roland Garros) takes place on clay courts after which the tour immediately moves to grass.

➤ June: There are warm-up grass court tournaments as preparation for Wimbledon that take place in the U.K. (Nottingham, Birmingham, Eastbourne), the Netherlands, and Mallorca.

➤ July: Wimbledon is played on grass in the U.K.

➤ August: The tour moves to outdoor hard courts in the USA & Canada (San Jose, Washington, D.C., Cincinnati, Montreal/Toronto, New Haven) as preparation for the US Open.

➤ September: The U.S. Open is played on outdoor hard in New York.
➤ October: The tour moves to Asia where all tournaments are on outdoor hard courts (Guangzhou, Seoul, Tokyo, Wuhan, Beijing, Hong Kong, Tianjin). However, these are followed by 3 tournaments in Europe (Luxembourg, Linz, and Moscow). The Year-end finals are played in Singapore and Zhuhai.

Additionally, the WTA is careful not to schedule any other tournaments over the weeks when Grand Slams and Premier Mandatory events are played (8 such tournaments over 14 weeks). In all other weeks, the WTA also spreads the Premier 5 and Premier events as to avoid any overlaps between them. Typically, a Premier event of any type will overlap only with an International event.

As could be deduced from the options in this calendar, there are many different travel itineraries that a player can complete in a given year, which could have an impact on her performance if her chosen tournaments are away from each other and played over successive weeks. To summarize the number of options available, the following histogram bar chart shows the number of tournaments on the schedule per week:

Figure 2: Number of tournaments available per week and exact destinations
Combined weeks 11+12, 13+14, etc show tournaments that are happening over 2-week periods


This chart shows WTA's attempt to "contain" the tournaments and create an equitable for the players and the tournaments schedule in which no week has more than 3 tournaments. Yet, this represents a vast sea of travel routes. In the theoretical case that a player goes to play every week in 1 tournament, there are exactly 1,769,472 combinations possible, which is quite staggering. Of course, a player will not be physically able to compete in all weeks of the year, so she needs to be smart about her schedule and weigh the tournaments according to her goals and abilities.

WTA's calendar also distributes the prize money of the tour across the regions, requiring players to travel overseas to earn their living. Theoretically, if a player confines herself to a single region (e.g. North America), she has gateway to at most only a third of the total prize money. Although approximately 65% of the prize money is still collected in Europe, the U.K. and North America, the past decade has seen the rise of Asia (primarily China) as a destination for WTA tour events. Asia now

accounts for 15% of the total prize pool, despite not having a Grand Slam (the biggest tournament in 2018 the China Open in Beijing, a Premier Mandatory event).

Figure 3: WTA's prize money by region
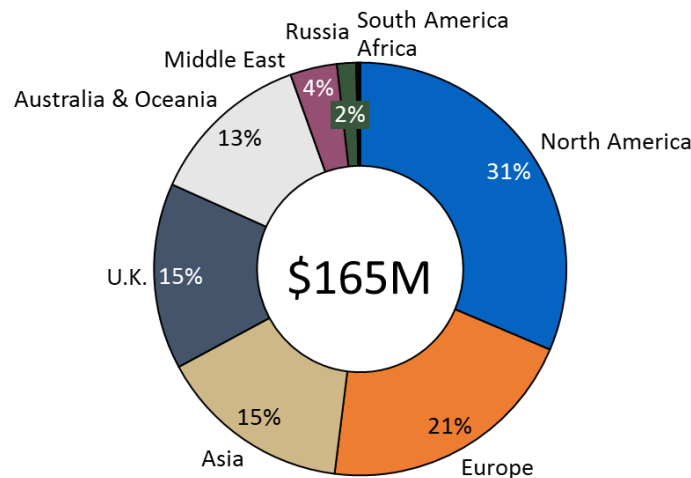
IV. Players' Schedule:

The game of tennis has become more physical and draining[10]. The depth of the women's field has improved and top players face serious competition even in the early rounds at every tournament[11]. In 2018, the year of research data for this paper, only 3 players from the top 10 reached the quarterfinals of French Open (Roland Garros) and even more strikingly, only 1 player from the top 10 reached the quarterfinals of Wimbledon. As a result, the points difference among the top 10 has shrunk and almost at every (Premier 5 or above) tournament, there could be a new world No.1, keeping the suspense high. After the end of the 2018 season, on the eve of the Australian Open in January 2019, there were 11 different women who could become No.1, which was unprecedented[12].

This increased competition and sporting drama is excellent for the fans[13], but has a physically exhausting effect on the players. It has become more imperative that a player makes a smart schedule and avoids excessive play. From a player's perspective, her schedule needs to be the following: 4 Grand Slams, 4 Premier Mandatory Tournaments, and 9 other tournaments in order to bring her total number of tournaments played to 17. This is because the WTA uses the best 17 tournaments to calculate a player's ranking[4]. If a player competes in less than 17 tournaments, there is a lost opportunity, while if she competes in more than 17, the WTA will pick the most advantageous ones to count for her ranking. Thus, an important conclusion is that every time a player loses in the first round of a tournament, it is only to her advantage to travel to another one and try to win points there. This replacement opportunity pushes some players to even compete at 25 tournaments per year, as will be shown below.

The choice how to fill the 9 other tournaments is left to the player. Typically, her choice will be dictated by any or a combination of the following factors:

- The tournament is in the player's home country or town
- The tournament is offering her appearance money (which is legal under the WTA rules [4])
- The player is coming back from injury and needs match play to get back in form

- The player suffered an unexpected early round loss and needs more match play, especially before a Grand Slam
- The player is chasing the Year-End WTA Finals (a phenomenon that only occurs in October when several players may be in contention for the coveted spots)
- The player is just curious about a place

As the season progresses, players often adjust their schedules to take into account unique developments described above - early round losses, injuries, or the need for more match play before Grand Slams. To ascertain these behavioral claims, we can examine the data for the year-end top 20 for 2018. We focus on the top 20 because they are in the best position to escape the travel fatigue due to three main factors:
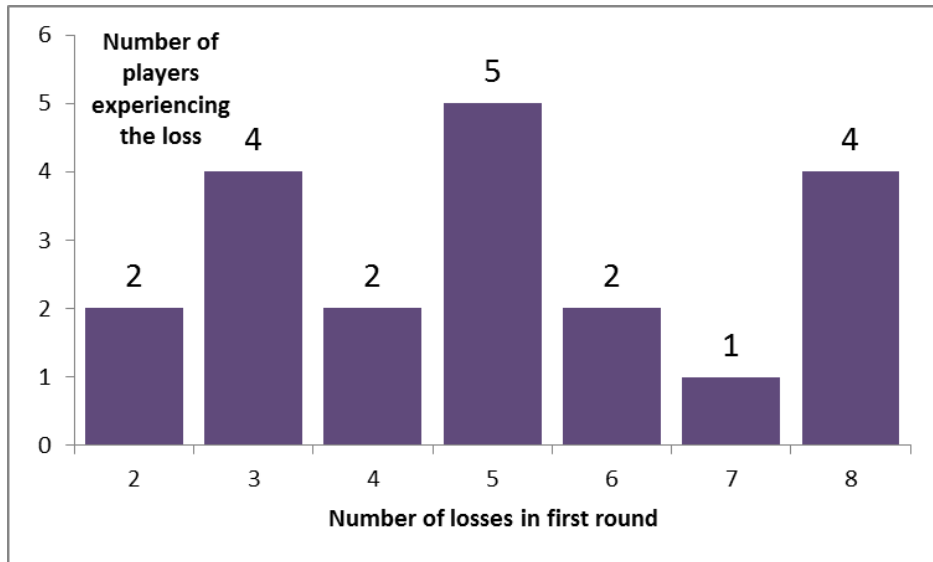
- They are the highest earners on the tour and can afford first-class travel, physiologists/fitness instructors, and nutrition counselors
- They typically go deep into tournaments and are therefore less likely to sign-up for tournaments in consecutive weeks
- They can cherry-pick among tournaments on a grand level and construct a schedule, which they believe would produce the best results (suggesting that "smart travel" is already pre-built in their schedules)

We will examine 6 claims:

Claim No.1	*All players,* whose ranking allows that, compete at the 4 Grand Slams and the 4 Premier Mandatory tournaments for a total of 8 required tournaments per year
Evidence	➢ 15 of the top 20 players in 2018 competed at all 8 events ➢ 4 players competed at 7 events (all withdrew from the 8th because of injuries) ➢ 1 player (Serena Williams) competed at just 5 (Serena Williams' schedule is an exception as she started the year only in March with no rank due to return from maternity leave)
Conclusion	**Correct (factoring injuries and discounting the one outlier due to special circumstances)**

Claim No.2	*Most players* supplement that with at least 9 additional tournaments because the WTA takes the best 17 tournaments to compute a players' ranking.
Evidence	➢ The top 20 played an average of 20.4 tournaments in 2018, with a maximum of 25. Removing Serena Williams, who played only 7, the average becomes 21.1. That is 13 in excess to the required 8. ➢ To check for distribution, 17 of the top 20 played at least 18 tournaments in the year, which meets the claim's threshold of 9 (over the required 8)

Conclusion	**Correct**

Claim No.3	*Most players* would travel to another tournament to try and gain points there, if they lost in the first round.
Evidence	➢ The average number of first-round losses for top 20 players is 5, distributed as follows:

  o 2 players lost 2 times in their first round
  o 4 players lost 3 times in their first round
  o 2 players lost 4 times in their first round
  o 5 players lost 5 times in their first round
  o 2 players lost 6 times in their first round
  o 1 player lost 7 times in her first round
  o 4 players lost 8 times in their first round

➢ If the top 20 players lose 5 times in their opening match, they should then play 17+5 = 22 tournaments. We saw from claim No.2 that the average without Serena Williams is 21. Collectively, they fall 1 tournament short.

➢ At individual level, 10 of the top 20 have at least one tournament that counts in their best 17 with a score of 0 [14] (a wasted slot), even though they still play additional tournaments to try and make up for that.

Conclusion	**Partially correct – players seek to fill all 17 slots with points, but 50% still**

	never reach that (too many first-round losses)

Claim No.4	*Most players* do not play in the week prior to a Grand Slam to avoid burnout and stay fresh and acclimate to the Grand Slam's location
Evidence	➢ 6 of the top 20 players played the week before the Australian Open ➢ 4 of the top 20 players played the week before the French Open ➢ 7 of the top 20 players played the week before Wimbledon ➢ 4 of the top 20 players played the week before the U.S. Open ➢ **Collectively**, only 11 of the top 20 ever played in the week before a Slam
Conclusion	**Correct**

Claim No.5	If No.4 if correct, those who play only play 1-2 matches for practice and do not put their best effort in tournaments in the week before a Grand Slam
Evidence	➢ The top 20 players lost in the first round 8 times from 21 collective attempts (38%) in tournaments preceding the Grand Slams ➢ The top 20 players collectively earned 123 points on average in the tournament before a Grand Slam, which translates to a finish between a quarter-final and a semi-final. Most would be expected to reach that stage at any regular tournament regardless of where it falls in the calendar. ➢ 3 of the top 20 players won those Slam "warm-up" tournaments (Angelique Kerber in Sydney, Caroline Wozniacki in Eastbourne, and Aryna Sabalenka in New Haven).
Conclusion	**Incorrect – players reach later rounds and also win those tournaments**

Claim No.6	*Most players* avoid playing consecutive weeks
Evidence	

➤ On average, 46% of the tournaments that the top 20 competed at were consecutive tournaments requiring the players to directly travel from one location to another

➤ However, the chart presents evidence to postulate that the top 5 players avoid playing consecutive tournaments (as well as Serena Williams at No.16, and Madison Keys at No.17)

Conclusion	**Incorrect – playing in consecutive weeks is more common than thought, but probably scheduled "smartly"**

Based on the last finding, we can conclude that travel is a factor in as many as half of the tournaments played by the top 20. As such, it would be an important metric to examine more closely.

V.   Data Overview:

Playing mostly away from home is a fact of life for women's tennis players. One study[15] tried to isolate "home play" in professional tennis and found that on the men's side, "home play" carries significant advantage (the best example is Roger Federer winning 8 times in Basel, his home town), but on the women's side, it has no effect on results. This is an interesting finding that could perhaps be explained with the fact that women start their professional career earlier than men and the feeling of home is not as developed.

Thus, we could proceed our review with the basic assumption that no woman has any advantage anywhere. Data for the 2018 season was collected directly from the WTA Tour web-site and the scope includes all tournaments in which the year-end top 30 competed. The author drew a cut-off at 30 for two primary reasons:

- This number is sufficient to create an "overall picture" given that this sample size is what is generally recommended in statistics

- Nobody, except one single player (year-end rank No.24), played in any tournaments from the lower-level tour of the International Tennis Federation (ITF). This cut-off at 30 provides such differentiation because the incentives for top-ranked players to play lower-echelon tournaments are very tenuous and the visualized data will be free of mixtures and will truly represent the WTA tour, which is the central aim of this project.

Data will be examined for two cases: (1) unfiltered, inclusive of every tournament, and (2) filtered for consecutive tournaments only. The reason for this dual-inspection is because **travel from one tournament to another in successive weeks predisposes a different decision-making process, and there could be a sub-trend in the data.**

Unfiltered Data (all tournaments):

We have a total of 613 observations across 30 players. 21 players (70%) have competed in at least 20 tournaments.

Figure 4: Number of tournaments and first-round losses:
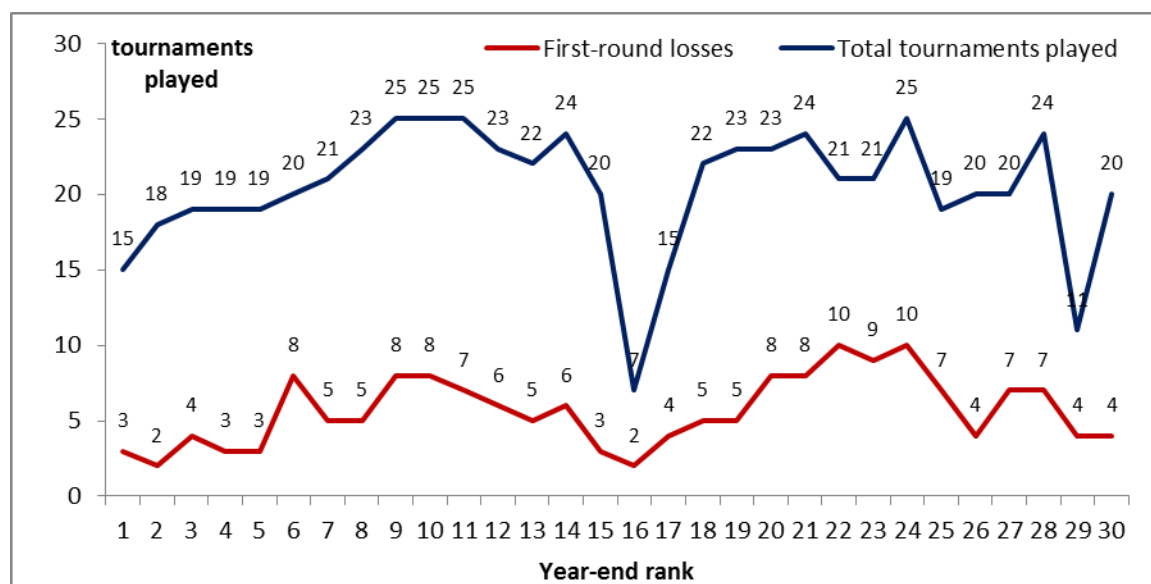
We see from this graph that many players suffer excessive first-round losses and do not reach 17 tournaments in which they have won their first match (which will positively contribute to their ranking). Three players from the top 10 have suffered 8 first-round losses, which under all circumstances could be described as unexpectedly high. This phenomenon testifies to the increased competition in the

game and the fact that players could be driven by different levels of motivation, given the unbalanced emphasis on a select few set of tournaments.

There are two outliers in this unfiltered case: the players finishing the year at No.16 and No.29, Serena Williams and Maria Sharapova, respectively. Both had different paths in 2018 with Serena Williams returning from pregnancy and playing only 7 tournaments, while Maria Sharapova battled injuries and had to cut her schedule short.

Filtered Data (consecutive tournaments only):

We have a total of 287 observations across 30 players. 21 players (70%) have competed in at least 10 consecutive tournaments. When evaluating the numbers, it should be pointed that because of the calendar construction, there are 4 pairs of consecutive tournaments in which almost every player participates: the Indian Wells-Miami, Madrid-Rome, Canadian Open-Cincinnati, and Wuhan-Beijing ones, which are either "Premier Mandatory" or "Premier 5" category tournaments, awarding big points for the players. So the base in this category should be considered to be 4. This is exactly how many the No.1 player of the season, Simona Halep, played.

Graphing the data, it is telling to observe how low the distribution of consecutive tournaments among the top 5 is: those players **have avoided competing at successive tournaments**. One explanation could be that they had a deep run into their current tournament and cancelled their participation in next week's one based on the good results they posted. Another could be that they just don't play week-in and week-out by original intent. However, both explanations lead to the hypothesis that lesser number of consecutive tournaments may be one of the factors for their success and the reason why they finished the year in the top 5. This indirectly supports the claim that lesser consecutive travel leads to greater overall success.

Figure 5: Number of consecutive tournaments and first-round losses:
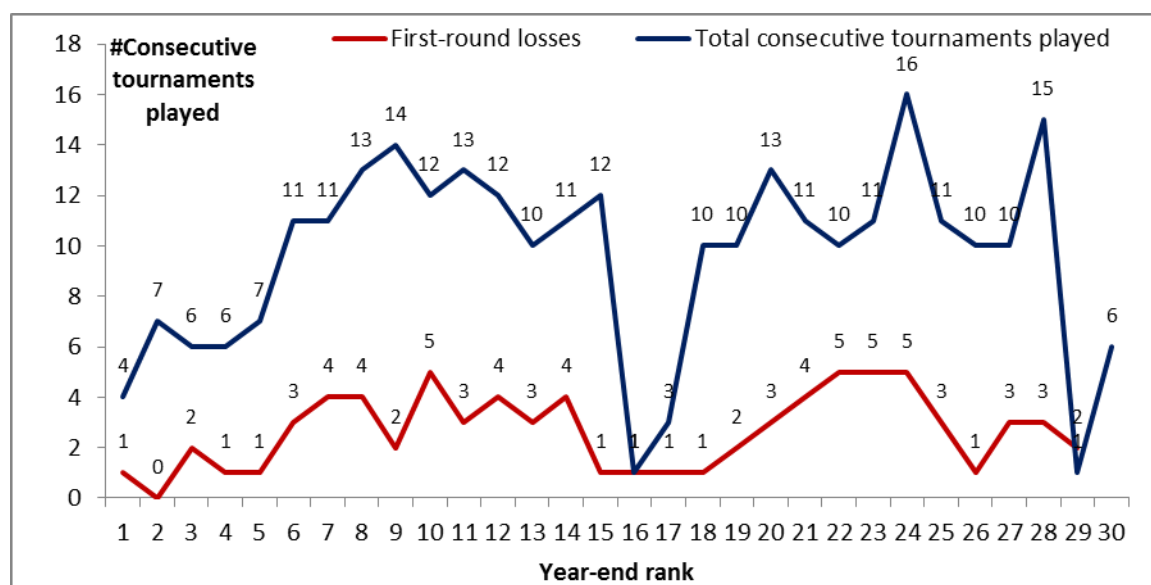
The filtered data set points to 3 outliers: players who participated in less than 4 consecutive tournaments (the base). Two of them the same as in the unfiltered case: Serena Williams and Maria

Sharapova, and the third one is American player Madison Keys. This is a normal working number of outliers for the data set, which would not distort our conclusions.

Of special scrutiny are the outliers on the opposite end, the players who competed more than everybody else in consecutive tournaments:

- First, at year-end rank No.24, we find Mihaela Buzarnescu, who competed in 16 consecutive tournaments. The return, as we see in Figure 6, has not been as good because she only earned 74 points per consecutive tournament, while the average for the top 30 stood at 158. Furthermore, just before the U.S. Open, Buzarnescu suffered an injury, which caused her to miss the next 2 months on the tour[17]. This could have come directly from physical exhaustion from competing in so many consecutive tournaments.

- Second, at year-end rank No.28, we find Su-Wei Hsieh, who competed in 15 consecutive tournaments. She also did not have a great return for all that travel because she only earned 69 points per consecutive tournament. Unlike Buzarnescu, she did not get injured, but she started the year at rank No.103 and managed to finish at No.28, which suggests that she originally had planned to participate in more tournaments in order to improve her ranking

If we examine the quality of performance in consecutive tournaments, we can graph the average points earned per consecutive tournaments played:

Figure 6: Average number of points earned when playing consecutive tournaments (with linear trendline):
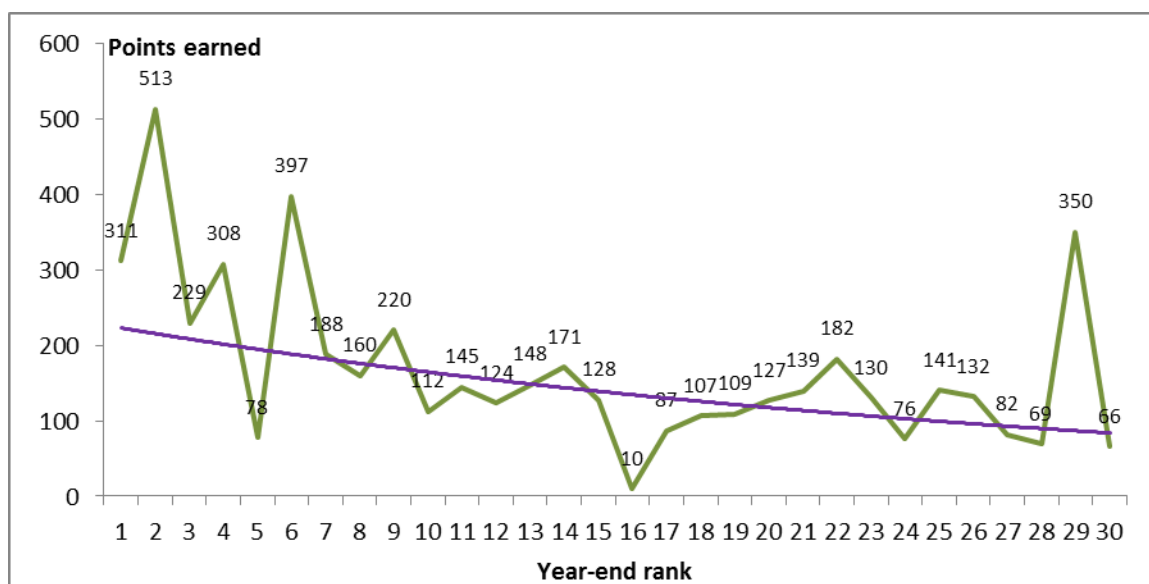

From this graph, we see that the majority of top 5 (except player ranked No.5) achieved better results *per consecutive tournament played* than their lower-ranked peers (one exception at rank No.29, Maria Sharapova, who played only 1 consecutive tournament). However, the number of consecutive

tournaments that the top 5 contested is lower, so their performance may be rooted in the fact that they are less exhausted from more back-to-back traveling than their lower-ranked peers.

Overall, based on these observations, it is interesting to combine Figures 5 and 6 and construct an "efficiency scatter plot" to examine the number of consecutive tournaments played (from Figure 5) to the average points earned (from Figure 6) for the entire top 30:

Figure 7: Average points earned by number of consecutive tournaments played (with linear trendline). The red circles indicate the outliers discussed above (Maria Sharapova, Serena Williams, and Madison Keys at the left end with 1, 2, and 3 consecutive tournaments, respectively, and Su-Wei Hsieh and Mihaela Buzarnescu on the right, with 15 and 16 consecutive tournaments).
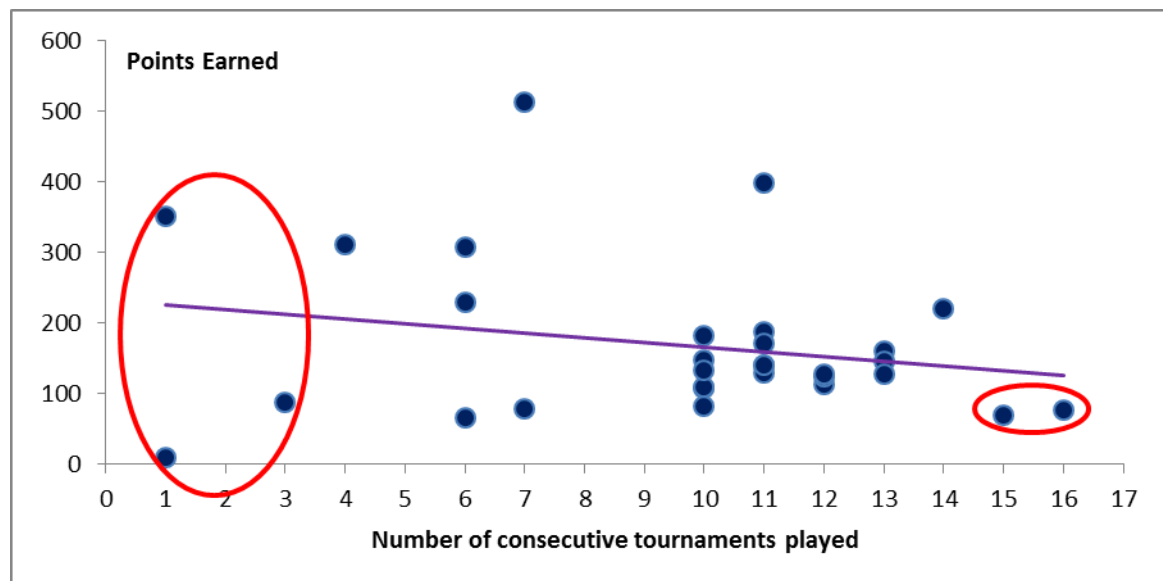
It is very interesting to observe how many players earned similar points at 10, 11, 12, and 13 consecutive tournaments played. The trendline, however, suggests that there are diminishing returns to playing more consecutive tournaments.

From this graph, it is worthwhile to look into 2 more outliers: the players who had the highest average return from playing consecutive tournaments from the top 30:

- Caroline Wozniacki (year-end rank No.3) who earned 513 points on average playing 7 consecutive tournaments: Caroline won Beijing (one of the four Premier Mandatory tournaments) after playing Wuhan.

- Sloane Stephens (year-end rank No.6) who earned 397 points on average playing 11 consecutive tournaments: Sloane won the Miami Open (one of the four Premier Mandatory tournaments) after playing Indian Wells

In both of these cases, Caroline Wozniacki and Sloane Stephens lost in their second match the week prior and then played the following week in the same country (China and USA), winning the tournament. For Sloane Stephens, this was also her home country, although we cannot attribute any

advantage to that. However, this could be food for thought that when travel is isolated to consecutive travel within the same country, results could be solid. This success could also stem from higher motivation to perform better after suffering an unexpected early loss.

VI.   Travel Data Analysis:

Again, we review the data from an unfiltered and filtered perspective.

Unfiltered Data (all tournaments):

We begin our review of travel insights with a holistic analysis of the unfiltered data (all tournaments). The total number of distinct routes "Tournament *A* – Tournament *B*" that the year-end top 30 players did in 2018 was 137. This number may sound surprisingly low given the vast pool of total possible travel combinations (as discussed in section II), but it points to that fact that the elite players do not deviate too much from a core set of tournaments (the Gland Slams/Premier Mandatory/Premier 5), which is exactly the goal of the WTA – to consistently bring the highest competition to its Premier events by incentivizing the players with points and prize money. Of these 137 routes, 55 (40%) were done in consecutive weeks, which will be of focus as these are the cases when players cannot take a break. The frequency spread of all 137 routes is as follows:

Figure 8: Frequency plot of routes by distance (in km) and number of players traveling
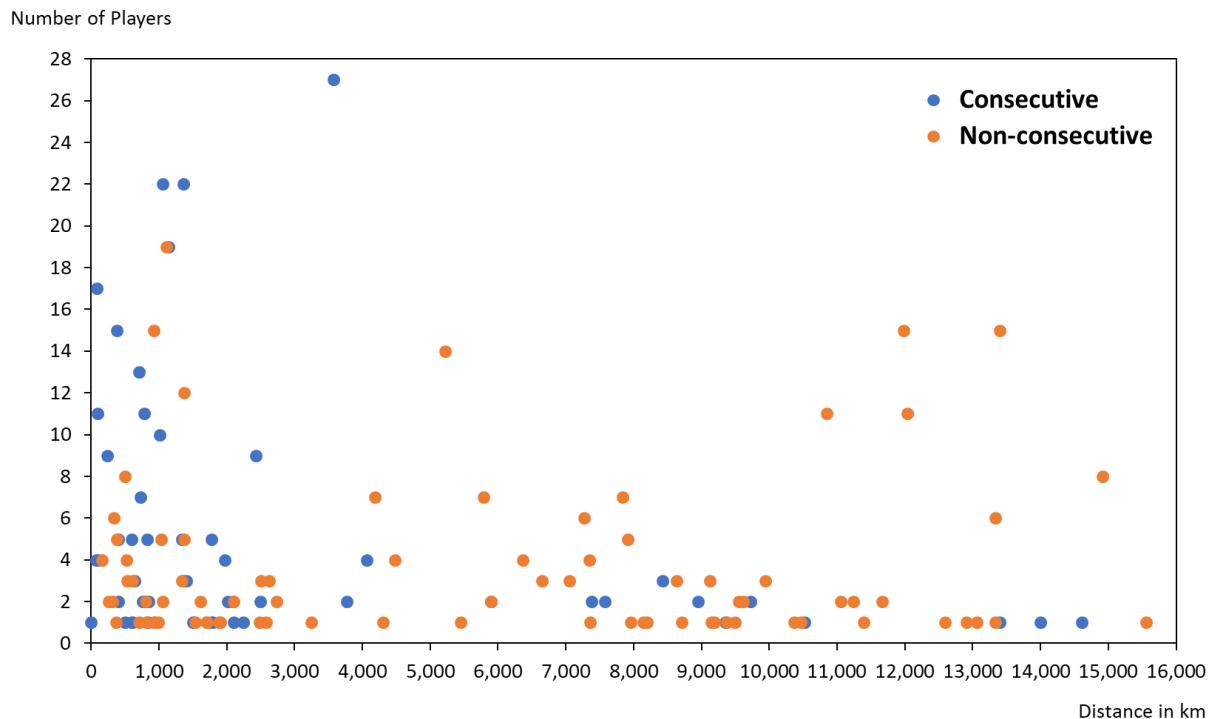
From this graph we clearly see that when players travel long distances, it is typically for non-consecutive tournaments.

We can also make an illustrative chart to show how much each player traveled throughout the year. This will be a theoretical construct to compare the total travel utilization in the top 30. This is a theoretical chart because in non-consecutive tournaments, players may choose to travel anywhere for a variety of reasons - for training, a doctor's appointment, an event, or to see family.

Figure 9: Total Travel Utilization by Player with average for the top 30 in red
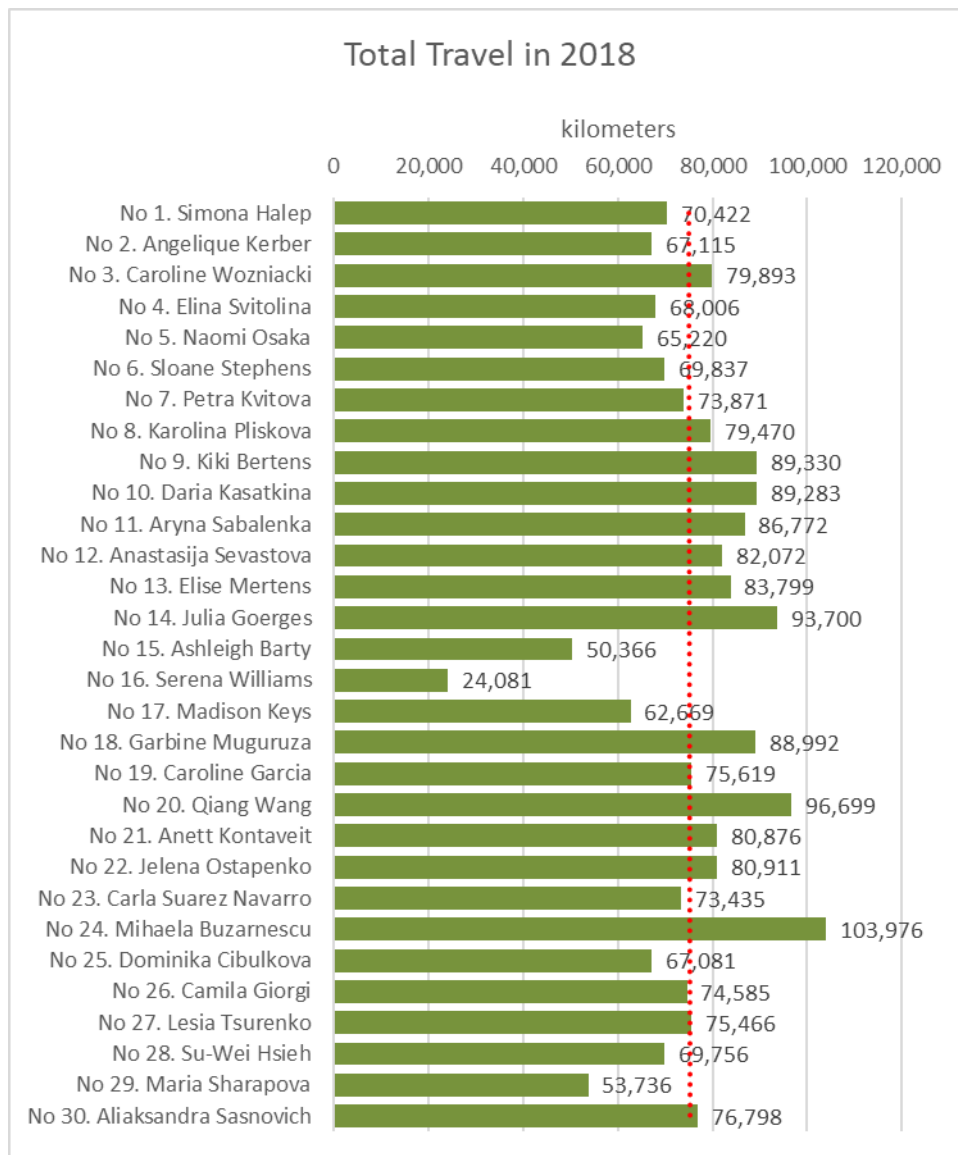
Intuitively, this total travel utilization will be correlated with the number of tournaments played. The magnitude of this correlation is very strong at 0.84. This means that for players who compete in relatively the same number of tournaments, the travel variance is not too big – which comes, again, as an extension of the fact that the elite players focus on a select group of tournaments driven by the WTA calendar construct. This is why we see that the top 7 players fall below the average number of kilometers traveled – they also play lesser number of tournaments (but go deeper in them).

Earlier, we discussed the cases of Serena Williams, Madison Keys, and Maria Sharapova, who, because of lower number of tournaments also have lower number of kilometers. A new outlier that emerges, however, is pla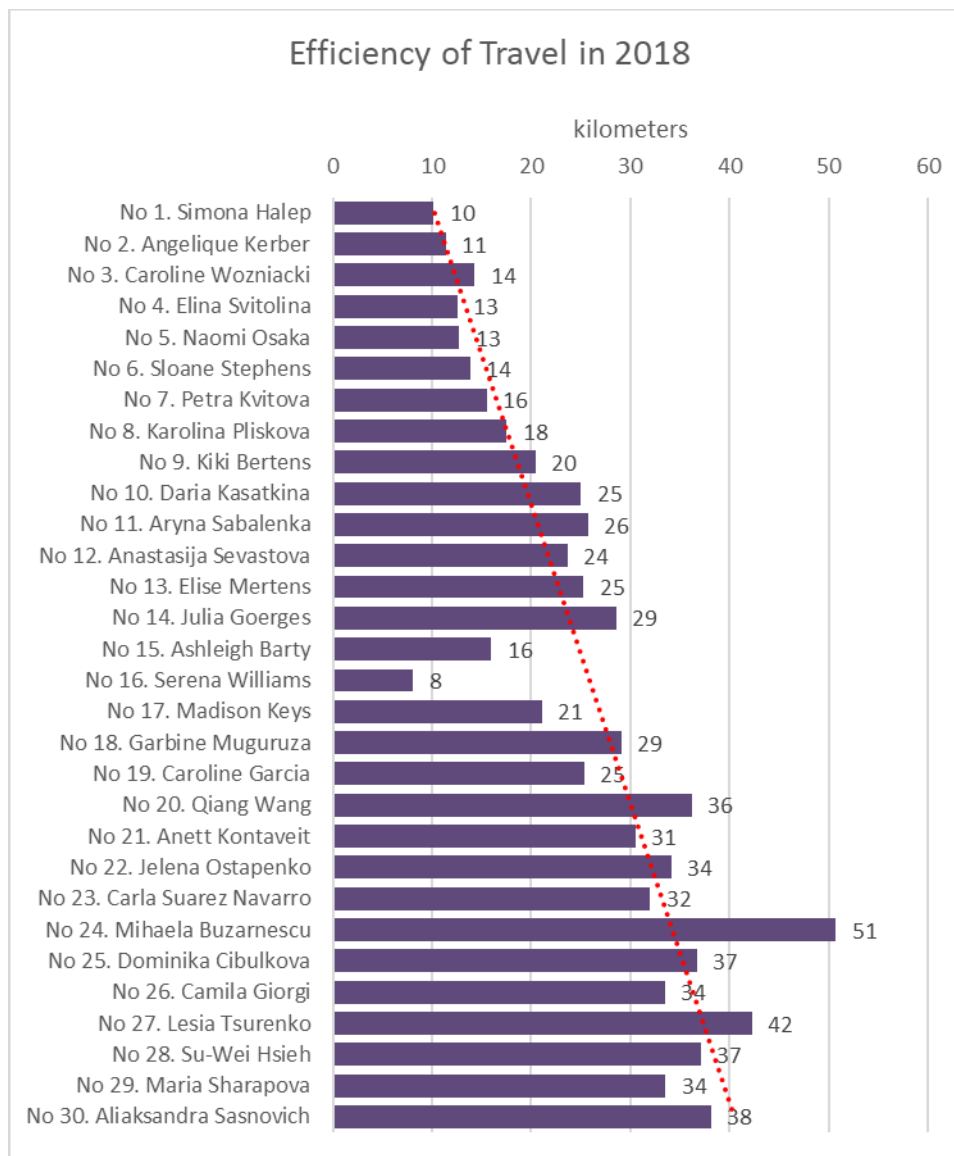yer No.15 Ashleigh Barty from Australia. She played in 20 tournaments in 2018, yet traveled 33% less than the average for all year-end 30 players. Examining her records, we see that she skipped the Middle East portion of the season (Doha and Dubai, which are played in February) and moved from Australia in January to the USA in March. She also played chose her International category tournaments to be all in the U.K. before Wimbledon: remarkably, Ashleigh played 4 consecutive grass tournaments in the U.K. in June, which is rare, but something she could have strategically chosen to do (to increase her experience on grass).

Another interesting observation is to compare the pair No.14 Julia Georges and No.28 Su-Wei Hsieh. They both played 24 tournaments, a high total utilization, but Su-Wei traveled 26% less. Julia still finished the year ranked higher but did not spare any travel sacrifices, going to the Middle East and twice to Moscow along with all other Premier category and above tournaments. Su-Wei, on the other hand, finished her year in Asia and did not travel back to Europe to play Luxembourg/Linz, which is something that Julia did.

The player who traveled the most, No.24 Mihaela Buzarnescu, did not go without an injury. As discussed above, she sustained an ankle injury before the U.S. Open, which prevented her from playing for 2 months.

Overall, we can construct an ***efficiency scale*** to measure who earned the most points for the least amount of travel.

Figure 10: Efficiency Scale for the year-end top 30 (with trend-line in red)

Efficiency of Travel in 2018

kilometers

Player	km
No 1. Simona Halep	10
No 2. Angelique Kerber	11
No 3. Caroline Wozniacki	14
No 4. Elina Svitolina	13
No 5. Naomi Osaka	13
No 6. Sloane Stephens	14
No 7. Petra Kvitova	16
No 8. Karolina Pliskova	18
No 9. Kiki Bertens	20
No 10. Daria Kasatkina	25
No 11. Aryna Sabalenka	26
No 12. Anastasija Sevastova	24
No 13. Elise Mertens	25
No 14. Julia Goerges	29
No 15. Ashleigh Barty	16
No 16. Serena Williams	8
No 17. Madison Keys	21
No 18. Garbine Muguruza	29
No 19. Caroline Garcia	25
No 20. Qiang Wang	36
No 21. Anett Kontaveit	31
No 22. Jelena Ostapenko	34
No 23. Carla Suarez Navarro	32
No 24. Mihaela Buzarnescu	51
No 25. Dominika Cibulkova	37
No 26. Camila Giorgi	34
No 27. Lesia Tsurenko	42
No 28. Su-Wei Hsieh	37
No 29. Maria Sharapova	34
No 30. Aliaksandra Sasnovich	38

This measure indicates that No.1 Simona Halep traveled 10 kilometers to earn 1 ranking point. We see that as a general rule, a player needs to achieve an efficiency of 20km for 1 ranking point in order to be in the top 10. Figure 10, however, takes into account all tournaments played by a given player, not just the best 17. This is why we observe players who still have very good efficiency, but finished lower than No.10 Daria Kasatkina. Based on this, we can formulate the following 2 directives about our travel efficiency measure:

- Travel efficiency can be compared only when players participate above a certain threshold number of tournaments. Looking at this chart, one may conclude that Serena Williams is the most efficient player, but such a conclusion would miss the fact that she only competed in 7 tournaments, which she cherry-picked. Had she competed in more, she could have accumulated more tiredness and her efficiency (and ranking) could be different. The author proposes this threshold to be 15, which is what 90% of the players in the top 30 cleared and which indicates full commitment to the tour.

- Travel efficiency does not mimic the rankings one-to-one because of the capped approach to the ranking calculation. Since the rankings take the 17 best tournaments, the ranking system would reward a few high peaks disproportionately high and could leave more extensive travel to smaller tournaments unaccounted for. In that sense, losses in smaller, far-off tournaments can have little effect in the rankings if a player has already won big at another stage, but could change the travel efficiency calculation. This is why we see No.10 Daria Kasatkina and No.19 Caroline Garcia end up with the same travel efficiency, but with different ranking: No.10 and No.19. No.10 Daria Kasatkina achieved a few higher peaks than No.19 Caroline Garcia and was able to cumulatively earn more points within the 17 capped slots that she had.

  This is precisely why we also see that No.15 Ashleigh Barty is as efficient as No.7 Petra Kvitova, but finished 8 places below: throughout the year Petra suffered 5 losses in the opening rounds vs 3 for Ashleigh, but at the same time achieved higher peaks in a number of important tournaments. Ashleigh's spread-out performance was travel efficient, but without big wins, she could not enter top 10. However, we can still describe Ashleigh's schedule as "smart" because she was able to produce a ranking this high with much lesser travel (as previously discussed).

Filtered Data (consecutive tournaments):

Switching to the filtered data for consecutive tournaments only, we construct the frequency spread for the 55 consecutive routes by distance and number of players who did them.

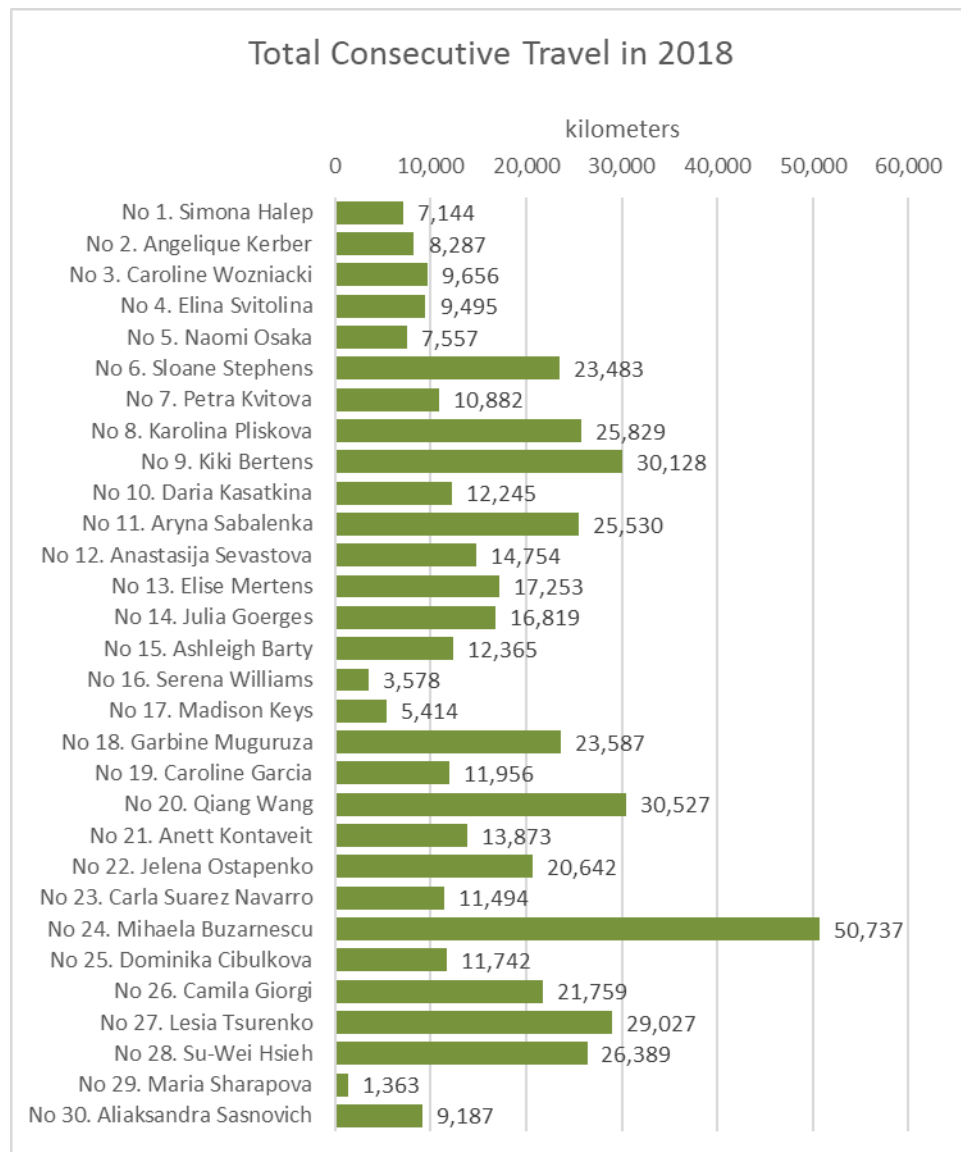*Figure 11: Frequency plot of **consecutive** routes by distance (in km) and number of players traveling*
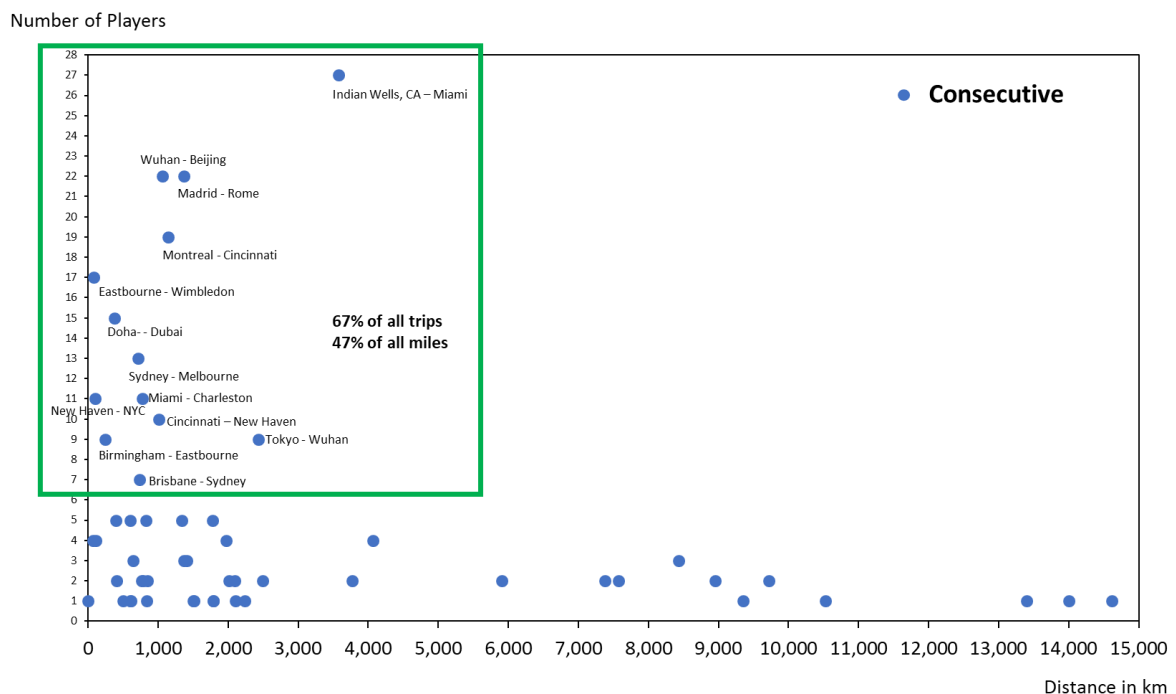
We observe that the consecutive routes are overall much shorter. The routes in the green rectangle represent the 13 most frequent consecutive routes. These 13 routes account for 67% of the

consecutive trips that the year-end top 30 made and represent 47% of the total kilometers traveled. This points to the fact that at the end of the tail, there are some long-distance consecutive trips that occurred. In almost all cases, these were undertaken by just 1 player.

In terms of consecutive travel, we have the following total kilometrage by player:

Figure 12: Total Consecutive Travel Utilization by Player

Total Consecutive Travel in 2018

kilometers

Player	Kilometers
No 1. Simona Halep	7,144
No 2. Angelique Kerber	8,287
No 3. Caroline Wozniacki	9,656
No 4. Elina Svitolina	9,495
No 5. Naomi Osaka	7,557
No 6. Sloane Stephens	23,483
No 7. Petra Kvitova	10,882
No 8. Karolina Pliskova	25,829
No 9. Kiki Bertens	30,128
No 10. Daria Kasatkina	12,245
No 11. Aryna Sabalenka	25,530
No 12. Anastasija Sevastova	14,754
No 13. Elise Mertens	17,253
No 14. Julia Goerges	16,819
No 15. Ashleigh Barty	12,365
No 16. Serena Williams	3,578
No 17. Madison Keys	5,414
No 18. Garbine Muguruza	23,587
No 19. Caroline Garcia	11,956
No 20. Qiang Wang	30,527
No 21. Anett Kontaveit	13,873
No 22. Jelena Ostapenko	20,642
No 23. Carla Suarez Navarro	11,494
No 24. Mihaela Buzarnescu	50,737
No 25. Dominika Cibulkova	11,742
No 26. Camila Giorgi	21,759
No 27. Lesia Tsurenko	29,027
No 28. Su-Wei Hsieh	26,389
No 29. Maria Sharapova	1,363
No 30. Aliaksandra Sasnovich	9,187

We observe that when it comes to consecutive travel, the picture is more fragmented within the ranks with some players choosing to travel a lot more frequently than others. A few notable cases:

- No.6 Sloane Stephens: Sloane enters consecutive tournaments after she loses in her opening matches (unexpectedly too early). In the Asian swing of the tour in

September/October, she lost her opening matches in Tokyo and Wuhan and requested a place in the tournament in Moscow in the week before the WTA Year-End Final Tournament in which she also lost in the first round. Thus, Sloane is one of the players who seeks to get more match practice after a loss, although she is not always able to achieve that goal.

- No.8 Karolina Pliskova: Karolina was in a three-way contention with No.6 Sloane Stephens and No.9 Kiki Bertens for the two final spots at the WTA Year-End Final Tournament. To ensure that she gets there, Karolina played 4 consecutive tournaments right before the WTA Year-End Final Tournament, also adding kilometers to her record. In the end, No.1 Simona Halep withdrew from the WTA Year-End Final Tournament due to injury, which placed all three women in the WTA Year-End Final Tournament tournament. In that sense, all the travel turned out to be unnecessary.

- No.9 Kiki Bertens: Same as No.8 Karolina Pliskova above.

Similar to the unfiltered data, we check the correlation of the filtered data to the number of consecutive tournaments played and although it also turns out very strong at 0.77, it sits below that of the filtered data. This means that for players who compete at the same number of consecutive tournaments, the travel is more unequal than in the unfiltered data set. This could lead to some differences in performance as travel becomes longer for some and shorter for others.

We also recreate the **efficiency scale** presented in the unfiltered analysis to measure who earned the most points with less travel for consecutive tournaments, but with the note that there is much less choice to construct a "smart schedule" when a player decides to add a consecutive tournament to her schedule – there are usually either 1 or 2 options in any given week, or 3 in just select 4 weeks (as presented in Section III).

Figure 13: Efficiency Scale between consecutive travel and points earned for the year-end top 30 (with trend-line in red)

Efficiency of Consecutive Travel in 2018
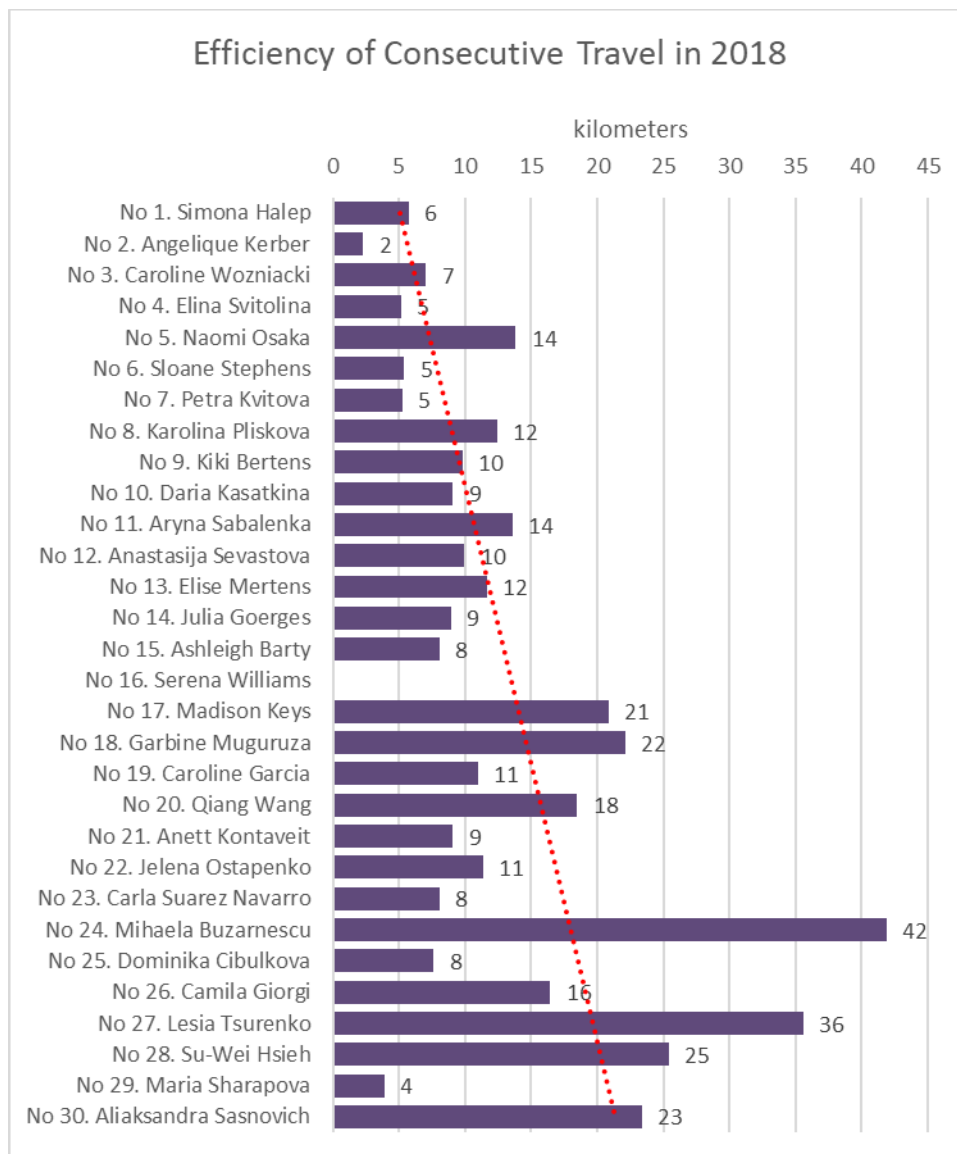
Note: No.16 Serena Williams was removed because she only competed in 1 consecutive tournament (Miami Open) and lost in her opening round.

We observe two important things:

1) Players travel less to earn their points than in the unfiltered case: No.1 Simona Halep, for example, traveled 6 kilometers to earn 1 point vs 10 kilometers for all her tournaments. This comes as a result of the fact that because consecutive travel occurs in closer geographies due to the WTA calendar construction.

2) There is a lot of variability in performance among the top 30. On a high level, Figure 13 shows who plays herself into form by participating in back-to-back tournaments and who suffers from the travel more than the rest. To illustrate that, it is noteworthy to observe the pair No.2 Angelique Kerber and No.5 Naomi Osaka. They both played 7 consecutive tournaments in 2018, but Angelique earned a lot more points from her consecutive tournaments and has an enviable efficiency of 2 kilometers for 1 point, the greatest return in the top 30. Although this is mostly

due to her Wimbledon win preceded by a week of play at Eastbourne, U.K., a closer examination of Angelique's record shows that she earned the same or higher number of points in the second tournament of all 7 instances when she played in two consecutive weeks. So on a grand level, consecutive travel benefits her as she plays herself into form. The opposite can be said about No.5 Naomi Osaka – she earned a lot less points in her consecutive tournaments, so travel could be more disruptive to her than others. In 2018, Naomi notably won the U.S. Open but she did not play in the 2 weeks preceding the tournament, which is contrary to what the majority of top 30 would do – play at least once in the 2-week period before the U.S. Open (26 out of the top 30 did that in 2018).

VII.   Conclusion

This research examined a variable – travel - that had not been considered in any previous study on tennis before. It sought to give a first read on the annual kilometrage accumulated by the top 30 players in 2018 and analyze the impact from travel by examining the data from different angles for consecutive (filtered) and combined consecutive and non-consecutive (unfiltered) travel. As a result of that, the research formulates a new "rule of thumb" that in order to be in the top 10, a player needs to 1) commit to at least 15 tournaments a year and 2) achieve travel efficiency rate of 1 point for every 20km traveled.

The conditions placed for the correct usage of this efficiency measure originate from the data:

- If a player competes in less than 15 tournaments, she will not accumulate enough points even if her travel efficiency is very high (as is the case of Serena Williams, who had the best travel efficiency of 8 kilometers for 1 point, but finished the 2018 season at No. 16 because she only played in 7 tournaments)

- If a player has already won a big title, a loss in a smaller tournament at a far-off location will have little impact on the rankings (especially if it's beyond the best 17), but could change the travel efficiency result. This is why we see players with good travel efficiency still not make it into the top 10: they show consistency, but did not achieve the same peaks in the top-rated tournaments that award disproportionate number of points as their rivals who end up in the top 10.

Other significant findings from this research that are related to consecutive travel are that:

- The top 5 players avoided playing consecutive tournaments. This could be a result from intentional or unintentional choice. The unintentional choice comes the fact that the top 5 players go deeper in the tournaments they play, on average, and possibly need a break after a week of intensive play (they are also older than other top 5 in the past). The intentional choice, however, is when they "cherry-pick" their tournaments with the explicit purpose to avoid consecutive travel and come more well-rested for each new tournament. This could also be the cause of why they finish in the top 5, having avoided too much complex travel.

- The data indirectly shows that there are diminishing returns to playing in more consecutive tournaments, which doesn't come as a surprise, but is an indirect confirmation of an intuitive hypothesis.

- Different players react differently to consecutive weeks' play – for some, it is a way to play themselves into form and post better results in the second week of play than the first one (such as No. 2 Angelique Kerber, as discussed above), but for others, it appears to be less successful (such as No. 5 Naomi Osaka, as discussed above). This shows that each player has a unique body system and results can really vary even at the top, depending on the schedule.

Overall, this research confirms that the top players maintain very high travel efficiency - they collect more points on average for their travels and are no "journeywomen", but focused professionals playing on a "smart schedule" in order to achieve the maximum returns at the least physical expenditure.

Because tennis is played on a single tour internationally and not in a conference/regional format, the players have developed very strong habits and defense mechanisms to offset travel-related predicaments from changing time zones, climates, and altitudes. This is why they also rarely bring issues of travel in their interviews, as reported by sports journalist Nick McCarvel (read full interview in Appendix).

---

[1] The $125K series are excluded, a category that is between ITF-level events (lower tour) and WTA Tour events. The Fed Cup is also excluded, as it is a special-format tournament, which may only apply to certain players if their country is playing and winning.

[2] WTA Calendar for 2018 at https://www.wtatennis.com/calendar/year/2018

[3] The WTA points system for 2018 at https://wta-playerzone.com/common/pdf/Rankings.pdf

[4] The WTA Rule Book for 2018 at http://wtafiles.wtatennis.com/pdf/publications/2018WTARulebook.pdf

[5] Faculty interview, Department of Physics at the University of Illinois at Urbana Champaign, October 2007, available at https://van.physics.illinois.edu/qa/listing.php?id=948&t=tennis-courts

[6] Editorial, "The Quick and the Dead – how to adapt to different surfaces and conditions", *The Guardian*, June 2009, available at https://www.theguardian.com/lifeandstyle/2009/jun/28/tennis-adapt-surfaces-conditions

[7] Felix, Ioonna, "How Playing Tennis in Different Regions Impact Joints", *official website for the Hospital for Specialty Surgery (HSS)*, January 2016, available at https://www.hss.edu/playbook/how-playing-tennis-in-different-regions-impact-joints/

[8] Rossingh, Danielle, "Travel Like a Tennis Pro: Jet Lag, Delays, and private Jets (If You Are Lucky)", *CNN online,* November 2016, available at https://www.cnn.com/2016/11/22/tennis/tennis-travel-story/index.html

[9] Livaudais, Stephanie, "Schmiedlova overcomes 'big pressure' to reach Bogota final", *WTA Online*, April 2018, aavailable at https://www.wtatennis.com/news/schmiedlova-overcomes-big-pressure-reach-bogota-final

[10] Kimmelman, Dennis, "How Power Has Transformed Women's Tennis", *The New York Times,* August 2010, available at https://www.nytimes.com/2010/08/29/magazine/29Tennis-t.html?mtrref=www.google.com&gwh=8618CDFB4D2D8DB9B895A133287AA6BE&gwt=pay

[11] Perrotta, Tom, "There Are No Easy Matches in Women's Tennis Any More", *FiveThirtyEight*, July 2017, available at https://fivethirtyeight.com/features/there-are-no-easy-matches-in-womens-tennis-anymore/

[12] Nguyen, Courtney, "Australian Open 2019: Eleven Players Vying for No.1 Spot in Melbourne", *WTA online,* January 2019, available at https://www.wtatennis.com/news/australian-open-2019-eleven-players-vying-no1-spot-melbourne

[13] Tignor, Steve, "The WTA in 2019: Are Fans Ready to Embrace Depth Rather Than Dominance", *Tennis.com¸*December 2018, available at http://www.tennis.com/pro-game/2018/12/depth-perception-are-fans-ready-embrace-variety-womens-game/78462/

[14] "0" is used here for stylistic purposes, in reality points are still given for first-round losses, but they are either 1 or 10 depending on the tournament category, which is negligibly small.

[15] Koning, Ruud, Home advantage in professional tennis, *Journal of Sports Sciences*, 29:1, pp. 19-27, 2011

[16] ITF schedules and tour information at https://www.itftennis.com/procircuit/tournaments/tour-info.aspx

[17] Kane, David. "Buzarnescu out of US Open Series after Montreal ankle injury", *WTA Tour online,* August 2018, available at https://www.wtatennis.com/news/buzarnescu-out-us-open-series-after-montreal-ankle-injury

Design Chapters:

VIII. Visualization Precedents

Comprehensive visualizations of women's tennis do not appear too frequently in the media, but there are a few compelling examples that draw attention and are excellent precedents to understand the research questions that others have tackled, possibly as a response to public interest:

1) A visual history of women's tennis, found [here][18]
   - This visualization is a hybrid blend between a bar-chart and a heatmap, displayed on a timeline. It shows all players who had been ranked as No.1 in history (up to 2016) and how their rank had progressed.

2) The increasing average age among women's top players, found [here][19]
   - This visualization is in the form of several line graphs that show how the median age of a top-50 women's player has jumped from 23 in 2008 to 27 in 2018, testifying to the change in fitness and physical recovery routines (described above in the paper). It also compares the sport to the median ages of female competitors in basketball, volleyball, swimming, and gymnastics, and finds that tennis has seen the most dramatic change.

3) Women's equality in prize money compared to men's, found [here][20]
   - This visualization is in the form of bar graphs that show a timeline by year how women's prize money caught up to men's in each of the four Grand Slams. We see that the U.S. Open has been the most equitable Grand Slam offering equal prize money to men and women since 1973, while it took more than three additional decades for The French Open (Roland Garros) and Wimbledon to do the same and catch up women to men's pay.

4) Women's distribution of prize money across the top 200 (and in comparison to golf), found [here][21]
   - This visualization is in the form of a line graph that shows the decreasing rate of prize money earnings by rank. We see that only 17 women earned more than $2M in one year (2017) and 37 earned more than $1M. The player ranked No. 100 earned $370K (before tax and all other travel-related expenses). This isn't as much as the men earn (since most tournaments outside of the Grand Slams still offer lower prize money for women's players than men's), but it is more than what women earn in golf.

In addition to the four precedents above, narrowing the scope to a single tournament and its statistics, one visualization professional from the U.K. has presented an interesting set of 10 visualizations to portray the myriad and creative ways there are to visualize a single tournament. His work focuses on Wimbledon 2013 and can be found [here][22].

IX. The Making of the Visualization

The visualization that the author created to support this research (viewable at https://kiriltraykov.github.io/thesis/ with all code available at https://github.com/kiriltraykov/kiriltraykov.github.io/tree/master/thesis) aims to equip the user with both macro and micro knowledge of the world of women's tennis and lead to a path for data exploration and critical evaluation. The visualization offers the user an opportunity to find stories in the data and make his or her own connections and conclusions – in that way the visualization can be thought of as a "tool" for studying this topic, not as a mere regurgitation of facts.

The data comes from the web-page chronicle of players' activity from the WTA web-site (see Appendix for the source web addresses). After scraping the web-site, the author did his own standardization and categorization by calendar weeks to differentiate between consecutive and non-consecutive travel. The author then gathered additional data on the distances between the tournament pairs and matched it back to the players' activity to create a spreadsheet by which to conduct all data analysis and charting presented above.

For the interactive web visualization, the author created a main page that has a title, a subtitle (short description), and a menu in the form of a tennis court (a grid configuration in CSS):

The font for the title and the subtitle was chosen to be narrow and slanted to convey a sporty feeling of speed. The exact chosen font-family was "Lato", sourced from Google Fonts. For all other text elements in the body of the visualization this font was concluded to be difficult to read (especially when a list of 30 foreign names are displayed), so the author employed the serif font-family "Sanchez", also sourced from Google Fonts.

The tennis court is green in color as a tribute to the origin of the sport – until 2005 the U.S. Open was played on green courts before it was determined that the blue color has better visibility on TV and provides a better contrast to the "optic yellow" color of the ball[23]. However, green courts are still in use at Wimbledon, as the natural grass on which the tournament is played is still green. For the background color, the author chose a paler pink (HEX #ff8b94) to provide lower intensity saturation on the page and make it more consumable, even when there is a lot of text.

The court serves as a portal towards the 4 chapters of the visualization: I. The Big Picture, II. The Weekly Picture, III. Annual Path and Efficiency, and IV. The New Rule of Thumb. The font-size of this text was chosen to be rather large (2.4 times the size of a regular <p> element) so that all 4 titles could fit in 2 lines for symmetricity. When the user hovers over these court sections, the color changes to the complementary one to signal that an event will happen. The complementary color in this case is a rich purple (HEX #8E388E).

"I. The Big Picture": This chapter serves as an introduction to the topic and provides a macro-level overview of the WTA tour to ground everyone in the basics. It is composed of static posters that serve as "information cards" for the user. They were imported as PNG images that were originally done in PowerPoint. Their content is fully developed in writing in sections II and III of this research paper.


Once the user goes into this chapter, he or she sees a "Return to Court" button, which has been styled in the same green color as the main-page tennis court and carries a white thick border just like the tennis court. This button appears in all other chapters and is a big and visible navigational element.

"II. The Weekly Picture": Following the macro-level calendar introduction, this chapter takes the user down at "ground zero" to show week-by-week the choices that players have. It quickly immerses the user into the world of the tennis player, prompting the user to think about travel, choices, and sequence. When a user goes into this chapter, he or she sees a weekly timeline with all 58 tournaments displayed. The structure for this was laid as another grid system in CSS and styled appropriately. The border-color chosen was the "optical yellow", which is the color of the professional tennis ball. When there are more than one tournament in a given week, the order of display is alphabetical.

Initially, there are only 4 tournaments that are highlighted (in the same green color as the main-page tennis court) and those are the four Grand Slams, the most pivotal tournaments in the year. There is a legend on the page to explain that. The intention is to quickly orient the user where in the year those fall and provide initial food for thought how a player might structure her calendar.

When each tournament is clicked, its color changes to the complementary purple, so that the user remembers which one he or she is currently viewing. Along with that, some of the next tournaments change color as well – they are the next destinations for all players who played the purple one.
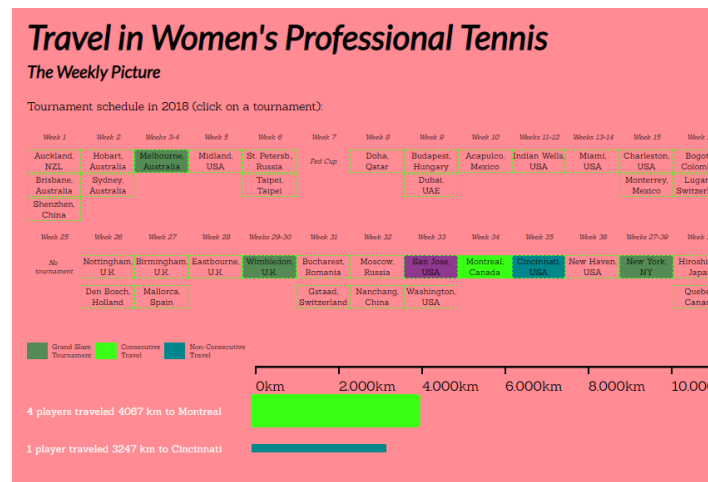
The tournaments that are consecutive to the purple one (e.g. "Montreal, Canada", in the example above) change their color to the "optical yellow" (and merge with the border of the box). This saturated color signals that this travel is different: it is consecutive, and the player needs to roll (just like the ball) to the next tournament in succession. All other next destinations that are not consecutive are highlighted in a shade of blue. There is a legend that appears to explain these color codings.

Once a user clicks on a tournament, there is an interactive bar graph that appears directly below to give information about the distance to the next tournament. The height of the bar is proportionate to the number of players traveling to that destination (1 player = 10px height, 4 players = 40px height, etc.) and there is written text on the side to spell out what the user is seeing. The order of display is: 1) consecutive tournaments (in "optical yellow") come before non-consecutive ones (in blue); 2) within each category, routes that are chosen by a higher number of players come first;

This bar chart uses all available dimensions: height to show distance traveled, width to show number of players who undertook the trip, and color to signal its category (consecutive or non-consecutive). From there, a user can gather a lot of visual clues about the popularity of some routes versus others. The interactivity of this bar chart is accomplished through CSS and JQuery – the code in CSS animates the bars, while the code in JQuery switches the tournaments upon a click event.

"III. Annual Path and Efficiency": this chapter presents the opportunity to explore each player's path in detail and compare travel between players. There are 3 sections:
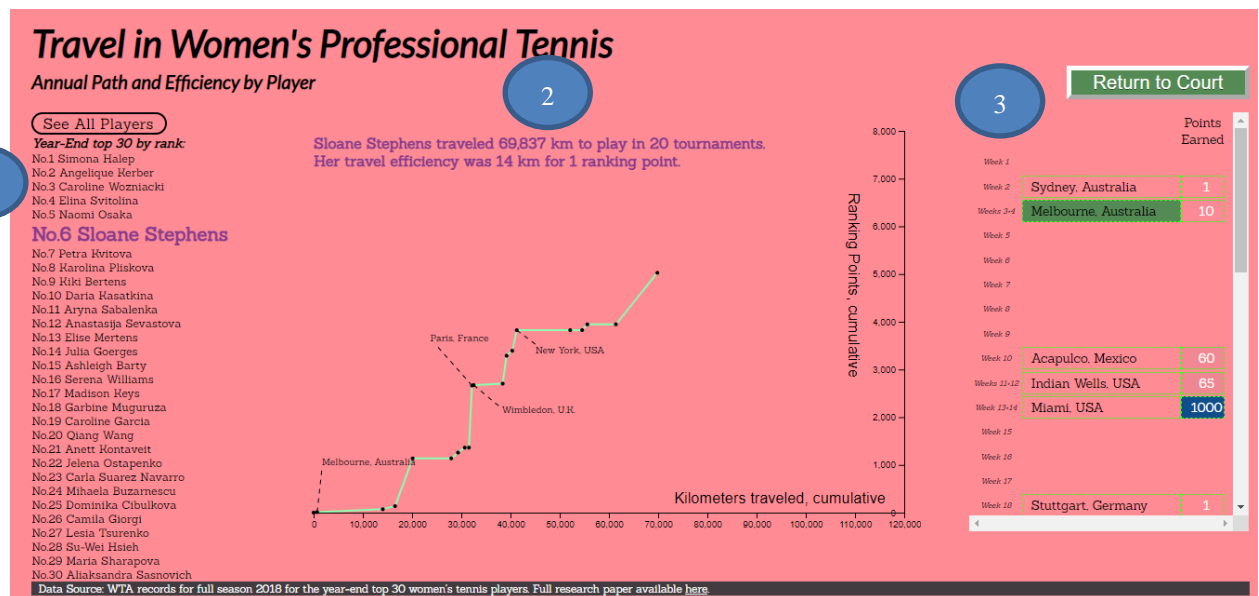
In **Section 1**, a user sees a list of all 30 year-end players for 2018. Through the use of CSS and JQuery, the text is animated (enlarged and converted to complementary purple: e.g. "No.6 Sloane Stephens" above) when the user hovers over. This formatting also holds from the time when a player's name get clicked until any other click event on another player's name. This ensures that the user remembers which player he or she is viewing.

Once a particular player is clicked, **Section 2** populates with dots: these are all tournaments that she had competed in. The dots are then quickly connected with a line (in the "optical yellow" color, keeping the dots in the foreground) to illustrate the progression of the player throughout the year. The x-axis measures cumulative kilometers traveled with each successive tournament, while the y-axis measures the cumulative points earned. This allows the user to see when a player traveled a lot but did not gain points and vice-versa. **The rule is that the steeper the path is, the more efficient the travel was** because the travel materialized in points earned for the rankings (i.e. victories). To provide reference points for the path, the 4 Grand Slams are shown via a dashed line legend (customized by each Grand Slam because of design fitting challenges). These labels orient the user how the players performed in each Grand Slam and what happened in the tournaments leading and following that. Unfortunately, these labels could not be applied to all tournaments because the graph becomes almost inelible and the overall design aggravates from density of text.

To tell the player's story to the user, there is interactive text that appears at the top of **Section 2**. The interactive text informs the user how many kilometers the player traveled and how many tournaments she played in. Next, it informs the user of her travel efficiency, the metric of special consideration in this research. The text appears in complementary purple to tie back to the formatting of the player's name in **Section 1** and be in unison.

To answer all remaining questions for the users about what happened in non-Grand Slam tournaments, the full player's schedule is introduced in **Section 3**. The box is aligned to the graph in **Section 2** and to fit, it provides a scroll-down option. In this section, a user can see the full details about where a player traveled to by week and how many points she earned. The formatting follows the same stylistics as those in "II. The Weekly Picture" and the user can flip back to "II. The Weekly Picture" to

even explore what other options the player had in certain weeks. In this **Section 3**, a user can see the points that a player collected in each tournament. To ease the visual inspection, the points are displayed with a heatmap-style background: the darker the color, the more points were gleaned. The chosen color is dark enough to contrast to the pale pink background. The four Grand Slam tournaments are, for consistency, once again colored in the tennis-court green color in order to be found quickly.

The chart in **Section 2** is driven by a d3.js script that plots the travel paths with a transition. The dots that represent the tournaments are also powered by d3.js and the interesting coding decision in this case is that the dots and the paths are both appended onto the same "svg" fitted to the same axis using the same data – i.e., the dots are not appended to the path or the path is not appended to the dots – they are each a class that is appended separately but appear as one.

Returning to **Section 1"**, there is a button for "See All Players". Upon clicking, the user is able to see all 30 paths in the data set visualized and can start comparing between players. (Note: there are no dots to represent each tournament because the visualization will suffer from the inclusion of 30x21 = 603 dots. Also, note that **Section 3** remains blank because it is no longer about an individual player). **This functionality is meant to visually answer the question of "Who traveled the most?"**
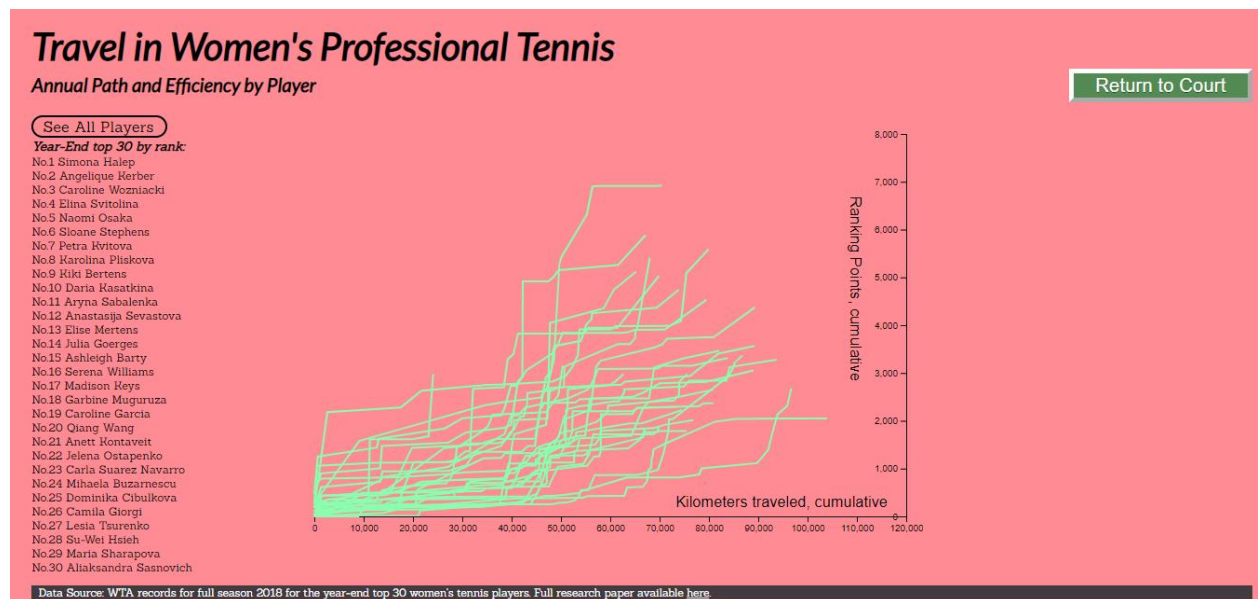

Looking by the x-axis, a user can quickly determine that the player who traveled the most is No. 24 Mihaela Buzarnescu:
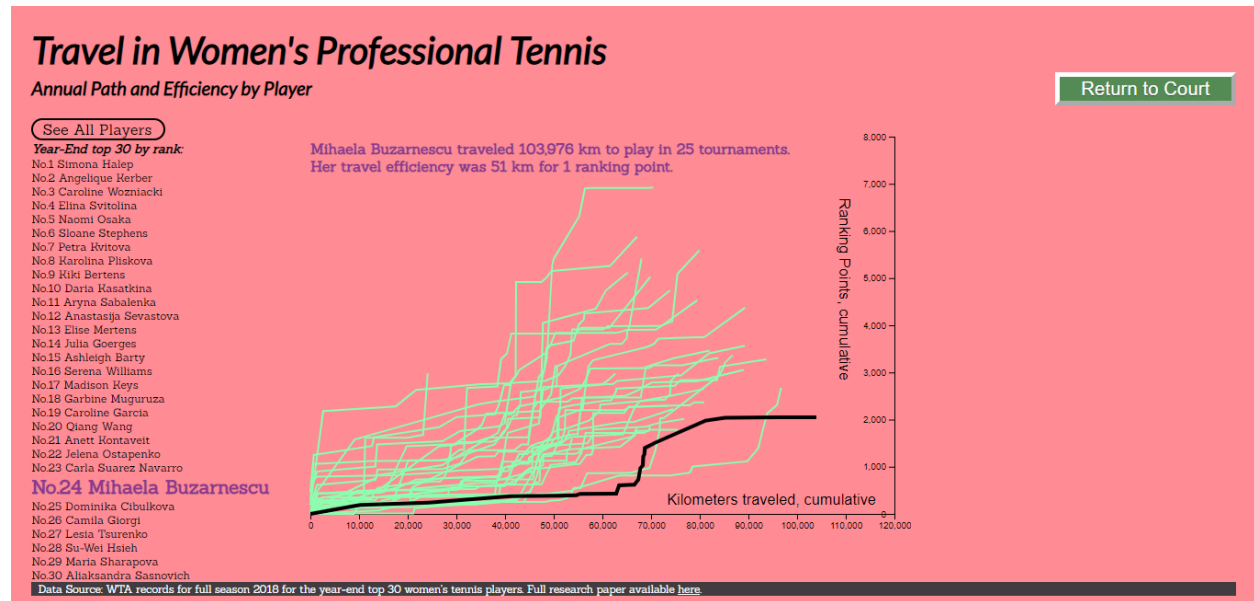
When a user hovers over a player's path, that path changes its color to black and both the player's name in **Section 1** is formatted (to stand out), as well as her story appears on top of **Section 2.** This allows the user to scroll over any player and know who it is and read the player's short story for 2018. If the user wants to see more details about that particular player, he or she can simply click back on the name in **Section 1.**

In that way, all three different sections in this chapter are connected and offer an ecosystem. A user can go back and forth between "See All Players" and individual selections and learn more by observing the data. This whole chapter uses d3.js, JQuery, and CSS to achieve everything.

"IV. The New Rule of Thumb": This chapter summarizes the takeaway from the research focusing on the travel efficiency metric. The text adds a few callouts for further consideration and it references the summary bar chart displayed to the side (with main color the complementary purple).
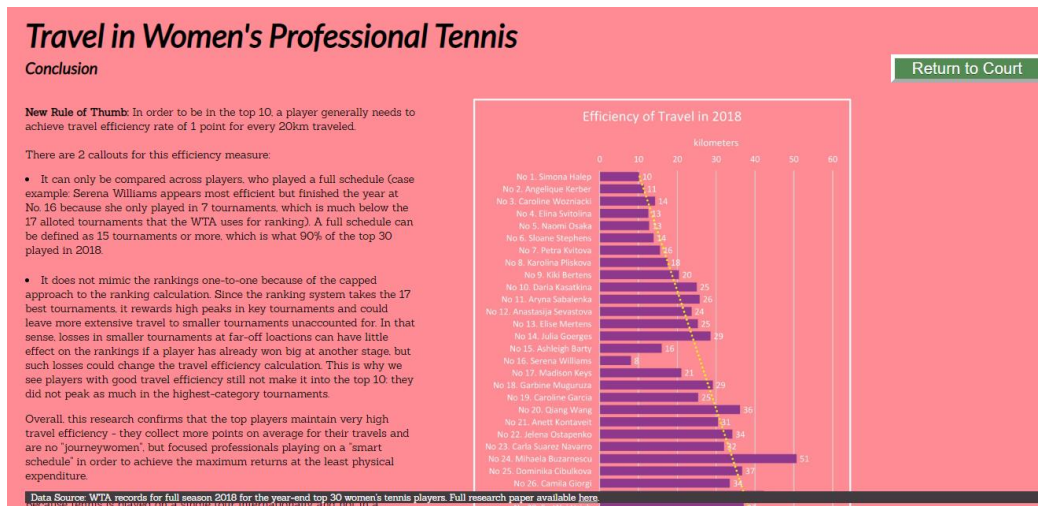
This chapter is identical to section "IX. Conclusion" in this paper. In the web visualization, there is also a link at the end of the conclusion to a PDF file with the interview conducted with an expert to support the research.

Overall, the web visualization was created with the intent to offer flexibility, intuitive design, and rich data platform for the users to engage in.

_____

18 Burn-Murdoch, John, "A visual history of women's tennis", *The Financial Times,* September 2016, available at https://ig.ft.com/sites/visual-history-of-womens-tennis/

19 Foley, Katherine Ellen, Dan Kopf September 2018, "Tennis has evolved and the 30s are the new 20s", *Quartz*, available at https://qz.com/1379932/the-2018-us-open-makes-it-clear-tennis-has-evolved-and-the-30s-are-the-new-20s/

20 Popovic, Nadja, "Battle of the sexes: charting how women in tennis achieved equal pay", *The Guradian,* September 2015, available at https://www.theguardian.com/sport/2015/sep/11/how-women-in-tennis-achieved-equal-pay-us-open

21 Jericho, Greg, "Tennis Players Want More Money? It's not as absurd as it sounds", *The Guardian*, January 2018, available at https://www.theguardian.com/business/grogonomics/2018/jan/18/tennis-players-want-more-money-its-not-as-absurd-as-it-sounds

22 Cook, Peter, "10 visualizations of Wimbledon 2013", available at https://charts.animateddata.co.uk/tennis/

23 Flink, Steve, "50 Moments That Mattered: US Open blue courts make their debut",
US Open web-site, March 2018, available at
https://www.usopen.org/en_US/news/articles/2018-03-30/2018-03-30_50_moments_that_mattered_us_open_blue_courts_make_debut.html

X.   APPENDIX:

Full Interview with Nick McCarvel, a sports journalist, live commentator, and host, with extensive experience traveling and reporting on the WTA Tour:

(1) How do the WTA players talk about travel in general? Do they like it? Do they hate it? Do they bring it up often in interviews?

    I think tennis professionals are pros - too - at travel. They have to be: The tour moves from Australia to the Middle East to the US to Europe, and that's just in the first few months of the season. WTA players talk about travel just the way the men do: They accept it as part of their jobs and do their best to be professional about it, staying hydrated, sleeping the hours they should in whatever time zone they have arrived in and the like. They don't bring it up in interviews, no, unless asked about it in particular. In small talk around interviews at the beginning or end of an event you might discuss travel, however.

(2) How do they deal with any travel-related fatigue?

    The basic understanding I know on tour is to "power through" whatever time zone you have arrived in. If you've landed in the morning in Australia from the U.S., you are best to stay up until "bedtime" in Australia to try and adjust as quickly as possible. I have heard players trying to exercise some way -- having a hit, going for a run or to the gym - when they arrive in a new place.

(3) Do they correlate their performance (e.g. losses) to long and exhausting travel (when asked to explain why they have lost)?

    I don't have a specific example, but the only time this would ever come up for a player is if they have gone long, long distances in the previous few days (ie more than their opponents). After a Fed or Davis Cup week (say, currently, some of the players went from Australia to Stuttgart this week) this might be pertinent, or when the tour is transitioning from one part of the world to the next. Only then would you hear a player bring it up... and/or if they had had travel issues while doing so.

(4) How do they feel about playing in consecutive tournaments which require travel from one place to another without a break?

    It can be hard for the players, but mostly I think the tours have tried to build in safeguards so that it doesn't impact them. I think when this does come into play it's about conditions: The time change might not be massive from LA to Miami (three hours), but add that to the elevation difference and dry air vs. humidity and it's quite the switch when they come from Indian Wells to play Miami. The same could be said about Rome, when they have just come from the high elevation in Madrid, though the travel - time-zone-wise - isn't very taxing.

(5) How early do they arrive for a given tournament?

    That really depends. But most of the top players will arrive 4, 5, 6 days before a Grand Slam's first day's play. At a Masters 1000/Mandatory event, it's usually 2, 3 or 4 days. Rarely do you see a player arrive only one day before they are set to play. I would imagine they would tell you they like to be on the ground at least 48 hours before they play.

(6) How do you think travel affects them overall?

You know that's a good question. I'm always impressed by the athletes: They rarely talk about travel or fatigue or jet lag. They are consummate professionals. When you do hear about it, it's because of outstanding circumstances: Delayed flights, missed connections, lost baggage or falling ill due to... well, being human. But they are always very attentive to their schedules, their bodies and how they can best perform on the court itself.

Data Sources:

No. 1 Simona Halep: https://www.wtatennis.com/players/player/314320/title/Simona-halep#matches
No. 2 Angelique Kerber: https://www.wtatennis.com/players/player/311470/title/Angelique-KERBER#matches
No. 3 Caroline Wozniacki: https://www.wtatennis.com/players/player/313402/title/Caroline-wozNIACKI#matches
No. 4 Elina Svitolina: https://www.wtatennis.com/players/player/316738/title/Elina-svitolina#matches
No. 5 Naomi Osaka: https://www.wtatennis.com/players/player/319998/title/naomi-osaka-0#matches
No. 6 Sloane Stephens: https://www.wtatennis.com/players/player/315683/title/sloane-stephens%20#matches
No. 7 Petra Kvitova: https://www.wtatennis.com/players/player/314206/title/Petra-Kvitova#matches
No. 8 Karolina Pliskova: https://www.wtatennis.com/players/player/313974/title/Karolina-Pliskova-0#matches
No. 9 Kiki Bertens: https://www.wtatennis.com/players/player/314584/title/Kiki-Bertens#matches
No. 10 Daria Kasatkina: https://www.wtatennis.com/players/player/322082/title/Daria-Kasatkina#matches
No. 11 Aryna Sabalenka: https://www.wtatennis.com/players/player/320760/title/Aryna-Sabalenka-0#matches
No. 12 Anastasija Sevastova: https://www.wtatennis.com/players/player/313987/title/Anastasija-Sevastova
No. 13 Elise Mertens: https://www.wtatennis.com/players/player/317964/title/elise-mertens-0#matches
No. 14 Julia Goerges: https://www.wtatennis.com/players/player/313381/title/julia-goerges-0#matches
No. 15 Ashleigh Barty: https://www.wtatennis.com/players/player/318033/title/Ashleigh-Barty-0#matches
No. 16 Serena Williams: https://www.wtatennis.com/players/player/230234/TITLE/SERENA-WILLIAMs#matches
No. 17 Madison Keys: https://www.wtatennis.com/players/player/316959/title/Madison-Keys#matches
No. 18 Garbine Muguruza: https://www.wtatennis.com/players/player/316239/title/Garbi%C3%B1e-Muguruza#matches
No. 19 Caroline Garcia: https://www.wtatennis.com/players/player/315391/title/CAROLINE-GARCIA#matches
No. 20 Qiang Wang: https://www.wtatennis.com/players/player/314429/title/Qiang-Wang#matches
No. 21 Anett Kontaveit: https://www.wtatennis.com/players/player/318425/title/anett-kontaveit%20#matches
No. 22 Jelena Ostapenko: https://www.wtatennis.com/players/player/319939/title/Jelena-Ostapenko-0#matches
No. 23 Carla Suarez Navarro: https://www.wtatennis.com/players/player/311338/title/Carla-Su%C3%A1rez-Navarro-0#matches
No. 24 Mihaela Buzarnescu: https://www.wtatennis.com/players/player/312268/title/Mihaela-BUZARNESCU#matches
No. 25 Dominika Cibulkova: https://www.wtatennis.com/players/player/312894/title/Dominika-Cibulkova-0#matches
No. 26 Camila Giorgi: https://www.wtatennis.com/players/player/314610/title/Camila-GIORGI#matches
No. 27 Lesia Tsurenko: https://www.wtatennis.com/players/player/315295/title/Lesia-Tsurenko-0#matches
No. 28 Su-Wei Hsieh: https://www.wtatennis.com/players/player/310053/title/Su-Wei-Hsieh#matches
No. 29 Maria Sharapova: https://www.wtatennis.com/players/player/310137/title/Maria-Sharapova#matches
No. 30 Aliaksandra Sasnovich: https://www.wtatennis.com/players/player/317790/title/Aliaksandra-SASNOVICH-0#matches