Project Report: Game Analytics – Unlocking Tennis Data

Prepared By: Team Analytics (Data Engineering & Analysis Group)

Date: January 12, 2026

Project Phase: Final Insights & Reporting

---

1. Executive Summary

This report consolidates the findings from our end-to-end data pipeline project using the SportRadar Tennis API. By ingesting data for **6,400+ competitions**, **3,800+ venues**, and **1,000+ top-ranked doubles players**, we have created a robust SQL database that provides a 360-degree view of the professional tennis landscape.

Our analysis reveals that while the **USA and Western Europe** dominate in terms of infrastructure (venues) and elite player output, the **ITF (International Tennis Federation)** circuits provide the critical volume of tournaments that sustain the global ecosystem.

---

2. Key Strategic Insights

Based on the analysis of 20 SQL queries across our dataset, we have identified five critical insights:

Insight 1: The "Engine Room" of Tennis is the ITF

- **Finding:** While ATP and WTA get the TV time, the **ITF Men's and Women's circuits** account for **65% of all competitions** in our database (4,230 out of ~6,400).

- **Implication:** For any predictive model or talent scouting system, data collection cannot be limited to top-tier events. The ITF level is where the vast majority of match data is generated.

Insight 2: Infrastructure Hegemony

- **Finding:** The **USA (600 venues)**, **Italy (279)**, and **France (268)** control nearly **30% of the world's recorded tennis infrastructure**.

- **Contrast:** Major tennis nations like **China (235)** and **Spain (209)** follow, but the gap between the USA and the rest of the world is significant (more than 2x the runner-up).

Insight 3: The "Great Britain" Doubles Anomaly

- **Finding:** In the current rankings snapshot, **3 of the Top 4 Doubles Players** are from Great Britain (Lloyd Glasspool, Julian Cash, Henry Patten).

- **Implication:** This suggests a highly specialized and successful doubles training program in the UK, potentially serving as a case study for other federations.

Insight 4: Efficiency vs. Volume

- **Finding: Taylor Townsend (USA)** is the most "efficient" player, averaging **606 points per competition**. Compare this to players who play 30+ tournaments but struggle to break the Top 50.

- **Takeaway:** A "Quality over Quantity" schedule strategy appears most effective for maintaining a Top 10 ranking.

Insight 5: Gender Parity in Volume

- **Finding:** The dataset shows a relatively balanced ecosystem with **3,677 Men's events** vs. **2,732 Women's events**. While men's events are more numerous (~57%), the women's circuit is robust compared to other global sports which often have wider gaps.

---

3. Detailed Data Breakdown (The "20 Queries" Summary)

A. Competition Analysis

1. **Top Category: ITF Men** (2,198 tournaments) is the largest single category.

2. **Gender Split: 57% Men / 43% Women**. Mixed doubles events are rare (only 13 recorded).

3. **Elite Level:** Grand Slam and ATP/WTA 1000 events make up less than **5%** of the total volume, highlighting their exclusivity.

B. Geographic & Venue Analysis

4. **Top City: London (49 venues)** is the global capital of tennis infrastructure in our dataset, followed by Buenos Aires and Rome.

5. **Average Complex Size:** The typical tennis complex has **6 to 7 documented courts**.

6. **Largest Complex:** We identified facilities with up to **30 distinct venues** (courts) listed, serving as major hubs for tournaments.

C. Player Performance (Doubles)

7. **#1 Ranked Player: Lloyd Glasspool (GBR)** with 8,610 points.

8. **Top Country by Points: USA** leads significantly with **84,950 total ranking points**, followed by Great Britain (61,651).

9. **Most Active:** Several lower-ranked players have competed in **35+ tournaments** in a 52-week period, indicating a "grind" to accumulate points.

10. **Neutral Athletes:** A significant portion of points (42,351) is held by athletes competing under a **"Neutral"** flag (primarily Russian/Belarusian athletes), effectively making them the 3rd strongest "nation" in tennis.

- Stremlit link:- https://tuqdmbggakfi23efmxweapp.streamlit.app/

Results for: 23. Competitor Search and Filtering

Share ☆ ✎ ⌂ ⋮

Competitor Search Filters

Search by Competitor Name (e.g., "Federer")

Filter by Rank Range

1   500

Filter by Country

Australia   ˅

Minimum Points Threshold

0   − +

	competitor_name	country	rank	points	movement
14	Hule, Petra	Australia	144	549	0
15	Birrell, Kimberly	Australia	151	527	0
16	Pearson, Kody	Australia	192	408	0
17	Walton, Adam	Australia	197	390	0
18	Charlton, Joshua	Australia	200	363	0
19	Schoolkate, Tristan	Australia	204	355	0
20	Micic, Elena	Australia	218	392	0
21	Osborne, Alexandra	Australia	232	350	2
22	Duckworth, James	Australia	239	290	0
23	Hulme, Matt	Australia	242	278	0

```
The above exception was the direct cause of t
Traceback (most recent call last):
  File "/home/adminuser/venv/lib/python3.13/s
    self._storage.get_file(absolute_path)
    ~~~~~~~~~~~~~~~~~~~~~~~^^^^^^^^^^^^^^^^
  File "/home/adminuser/venv/lib/python3.13/s
    raise MediaFileStorageError(
        f"Bad filename '{filename}'. (No medi
    ) from e
streamlit.runtime.media_file_storage.MediaFil
2026-01-21 14:31:48,656 - INFO - Attempting t
2026-01-21 14:31:48,658 - INFO - Database con
2026-01-21 14:31:48,658 - INFO - Loading data
SELECT
    c.competitor_name,
    c.country,
    r.rank,
    r.points,
    r.movement
FROM
    competitors c
JOIN
    competitor_rankings r ON c.competitor_id
WHERE
    (:name_filter IS NULL OR c.competitor_nam
    (:min_rank_filter IS NULL OR r.rank >= :m
    (:max_rank_filter IS NULL OR r.rank <= :m
    (:country_filter IS NULL OR c.country = :
    (:min_points_filter IS NULL OR r.points >
ORDER BY r.rank;
2026-01-21 14:31:48,661 - INFO - Data loaded
2026-01-21 14:31:48,661 - INFO - Database con

main   yash87015/sql/main/tennish_group_pr... ⋮ →
```

Share ☆ ✎ ⌂ ⋮

Competitor Search Filters

Search by Competitor Name (e.g., "Federer")

Filter by Rank Range

1   500

Filter by Country

All   ˅

Minimum Points Threshold

0   − +

	competitor_name	country	rank	points	movement
0	Glasspool, Lloyd	Great Britain	1	8610	0
1	Siniakova, Katerina	Czechia	1	8780	0
2	Cash, Julian	Great Britain	2	8520	0
3	Townsend, Taylor	USA	2	8485	0
4	Heliovaara, Harri	Finland	3	7980	0
5	Patten, Henry	Great Britain	3	7980	0
6	Errani, Sara	Italy	3	7195	0
7	Paolini, Jasmine	Italy	3	7195	0
8	Zeballos, Horacio	Argentina	5	7115	0
9	Mertens, Elise	Belgium	5	6730	0

```
    r.movement,
    r.competitions_played
FROM
    competitors c
JOIN
    competitor_rankings r ON c.competitor_id = r.competito
WHERE
    c.competitor_id = ?;
2026-01-19 10:07:58,056 - INFO - Data loaded successfully.
2026-01-19 10:07:58,057 - INFO - Database connection closed
2026-01-19 10:08:10,319 - INFO - Attempting to get a databa
2026-01-19 10:08:10,320 - INFO - Database connection obtain
2026-01-19 10:08:10,320 - INFO - Loading data with query:
SELECT
    c.competitor_name,
    c.country,
    r.rank,
    r.points,
    r.movement
FROM
    competitors c
JOIN
    competitor_rankings r ON c.competitor_id = r.competito
WHERE
    (:name_filter IS NULL OR c.competitor_name LIKE :name_
    (:min_rank_filter IS NULL OR r.rank >= :min_rank_filte
    (:max_rank_filter IS NULL OR r.rank <= :max_rank_filte
    (:country_filter IS NULL OR c.country = :country_filte
    (:min_points_filter IS NULL OR r.points >= :min_points_
ORDER BY r.rank;
2026-01-19 10:08:10,325 - INFO - Data loaded successfully.
2026-01-19 10:08:10,325 - INFO - Database connection closed

main   yash87015/sql/main/tennish_group_project/appnew3.py ⋮ →
```