

Tennis Game Analytics

Dashboard Report

GitHub Link:

[Tennis Game Analytics Project](#)

Introduction:

The project "Game Analytics: Unlocking Tennis Data with SportRadar API" seeks to develop a complete solution to professional tennis data analytics by leveraging real-time data from the SportRadar API. The project seeks to help sports analysts, players, and organizations achieve greater insights into tournament dynamics, player performances, and event distributions. By taking leverage from structured data across categories, competitions, venues, competitors, and player rankings, the project creates a solid foundation for interactive analysis. The core technologies adopted are Python for data scraping, SQL for data management, and Streamlit for building an intuitive and interactive web application. The final outcome allows users to visualize data, filter based on certain parameters, and spot trends across world tennis events by turning raw sports data into operational insights.

Workflow:

The project was done using a structured multi-step process of data extraction, database creation and SQL querying using MySQL, and interactive dashboard creation using Streamlit with the help of SQLite. All the steps played an important role in converting raw tennis data into useful information.

1. Data Extraction and Normalization:

GitHub Link: [Tennis Game Analytics \(Python API Data Extraction\)](#)

The project started by retrieving tennis information from the SportRadar API through Python using Google Colab. Various endpoints were called to retrieve information on competitions, complexes, venues, players, and rankings in JSON format. The nested JSON results were parsed and converted to structured tabular data. The retrieved data was cleaned and normalized to fit relational database norms.

2. Exploratory Data Analysis (EDA) Using Python:

GitHub Link: [Tennis Game Analytics \(EDA\)](#)

Before loading the data into a database, exploratory data analysis in Python was conducted to discover the key trends and patterns. As per analysis, there were more numbers of male players under singles and doubles, with the USA leading the player count and venue contributions. Cities like London and countries like Japan, France, and Italy also showed good participation. The gender breakdown showed male players at 57.2%, female at 42.6%, and the mixed formats making up a mere 0.2%. These findings helped in shaping the analytical focus throughout the final dashboard.

3. Database Creation and Querying with MySQL:

GitHub Link: [Tennis Game Analytics \(MySQL Scripts\)](#)

After normalization, data were organized into relational tables in MySQL Workbench. Primary keys, foreign key constraints, and data types were used in creating tables like Categories, Competitions, Complexes, Venues, Competitors, and Competitor_Rankings. SQL queries were run in a sequence within MySQL Workbench to analyze key trends, e.g., competition types, ranking distributions, venue locations, and player performance measures.

4. Migration to SQLite for Streamlit Integration:

GitHub Link: [Tennis Game Analytics \(Exporting MySQL Database to SQLite\)](#)

In order to support lightweight and portable execution in the Streamlit environment, the MySQL database was exported to SQLite. This included rebuilding the schema and migrating all applicable data into an SQLite database file. The SQLite conversion provided for easy integration with the Streamlit application, which is designed to work most efficiently with file-based databases.

5. Streamlit Dashboard Development:

GitHub Link: [Tennis Game Analytics \(Streamlit Files\)](#)

An end-to-end The Streamlit dashboard was created from the SQLite database exported. The dashboard connects with SQLite in real time and runs SQL queries to present filtered information, summary statistics, and visualized insights. Leaderboards, competitor search and filtering, country-wise analysis, and interactive views categorized

by competitor name, rank, points, and country are a few of the important features. The application is convenient to use, giving stakeholders the means to explore and analyze tennis data with simplicity.

Challenges Faced:

1. API Data Extraction:

Getting data out of the SportRadar API involved dealing with authentication, rate limiting, and understanding complex endpoint structures, which were initial challenges.

2. Database Integration with Local Files:

Managing extracted data in CSV format and proper import into a local MySQL database included maintaining the format, schema, and primary-foreign key integrity.

3. Learning Streamlit from Scratch:

As Streamlit was a new technology for the team, some time was spent on learning more about it through online tutorials, documents, and websites.

4. Converting JSON to CSV:

Converting nested JSON data into neat, structured CSV files took custom parsing logic and validation to avoid losing or misrepresenting any data.

5. Team Coordination and Time Management:

Coordinating everyone's work, dividing tasks, and meeting deadlines together proved to be a challenge in a remote team setup.

Solutions to Overcome Challenges:

1. Improved Team Communication:

A dedicated WhatsApp group was created to facilitate better communication, to share information, and to coordinate tasks within the team efficiently.

2. Continuous Learning:

Team members actively upgraded themselves by referring to books, online tutorials, and YouTube tutorials to learn new tools and technologies.

3. Mentorship Support:

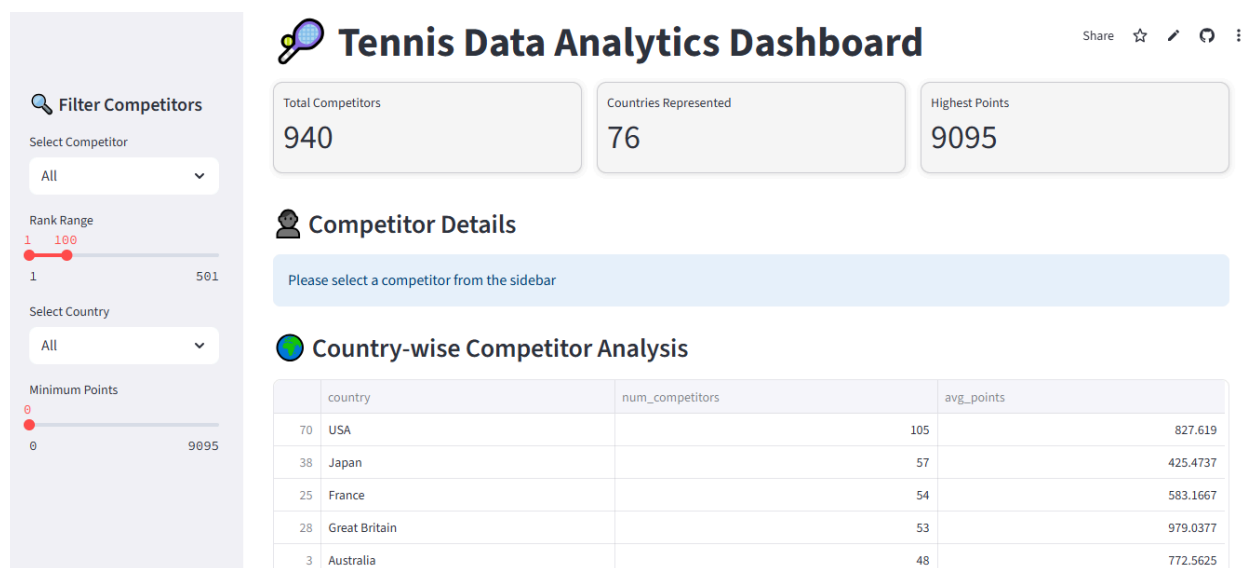
Regular attendance in doubt sessions led by mentors helped resolve technical issues, explain the strategy, and generate new ideas.

4. Leveraging AI Tools:

AI-driven tools were used for debugging, data validation, and logic code optimization, which is particularly useful towards the end for optimizing output and eliminating errors.

Streamlit Dashboard Snapshots:

Dashboard Link: [Tennis Game Analytics Dashboard \(Streamlit\)](#)



Filter Competitors

Select Competitor

All

Rank Range

1100

1501

Select Country

All

Minimum Points

09095

Filter Competitors

Select Competitor

All

Rank Range

1100

1501

Select Country

All

Minimum Points

09095

Filter Competitors

Select Competitor

All

Rank Range

1100

1501

Select Country

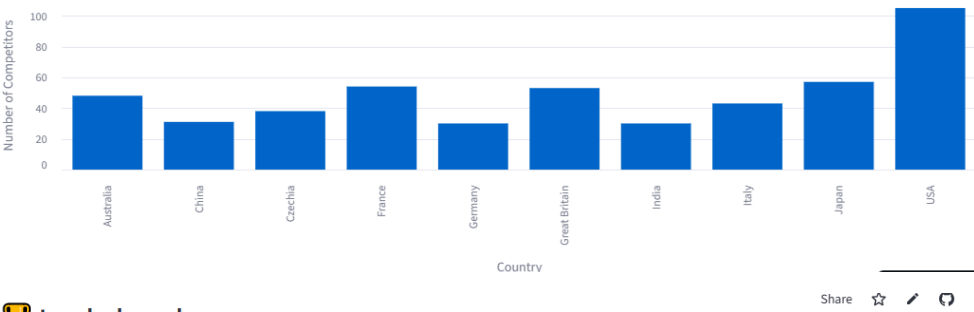
All

Minimum Points

09095

36	Italy	43	802.7209
18	Czechia	38	791.6316
14	China	31	759.2258
32	India	30	528.8
27	Germany	30	951.6333

Top 10 Countries by Number of Competitors



Leaderboards

Top Ranked Competitors

	name	rank	country	points
0	Pavic, Mate	1	Croatia	8280
500	Siniakova, Katerina	1	Czechia	9095
1	Arevalo-Gonzalez, Marcelo	1	El Salvador	8280
501	Townsend, Taylor	2	USA	8835
502	Routliffe, Erin	3	New Zealand	7610
2	Heliovaara, Harri	3	Finland	8060
3	Patten, Henry	3	Great Britain	8060
503	Ostapenko, Jelena	4	Latvia	6855
4	Krawietz, Kevin	5	Germany	6515
504	Errani, Sara	5	Italy	6480

Highest Point Scorers

	name	points	country
500	Siniakova, Katerina	9095	Czechia
501	Townsend, Taylor	8835	USA
0	Pavic, Mate	8280	Croatia
1	Arevalo-Gonzalez, Marcelo	8280	El Salvador
2	Heliovaara, Harri	8060	Finland
3	Patten, Henry	8060	Great Britain
502	Routliffe, Erin	7610	New Zealand
503	Ostapenko, Jelena	6855	Latvia
4	Krawietz, Kevin	6515	Germany
505	Paolini, Jasmine	6480	Italy

504	Errani, Sara	5	Italy	6480
-----	--------------	---	-------	------

505	Paolini, Jasmine	6480	Italy
-----	------------------	------	-------

Filtered Competitors

	name	rank	movement	points	country
0	Pavic, Mate	1	0	8280	Croatia
1	Arevalo-Gonzalez, Marcelo	1	0	8280	El Salvador
2	Heliovaara, Harri	3	0	8060	Finland
3	Patten, Henry	3	0	8060	Great Britain
4	Krawietz, Kevin	5	0	6515	Germany
5	Putz, Tim	6	0	6425	Germany
6	Granollers, Marcel	7	0	6135	Spain
7	Zeballos, Horacio	8	0	6135	Argentina
8	Thompson, Jordan	9	1	5440	Australia
9	Glasspool, Lloyd	10	1	5140	Great Britain

If we select competitor name from filter, we get the specific competitor details

Filter Competitors

Select Competitor

Aoyama, Shuko

Rank Range

1 100

1 501

Select Country

Japan

Minimum Points

0 9095

Total Competitors

1000

Countries Represented

77

Highest Points

9095

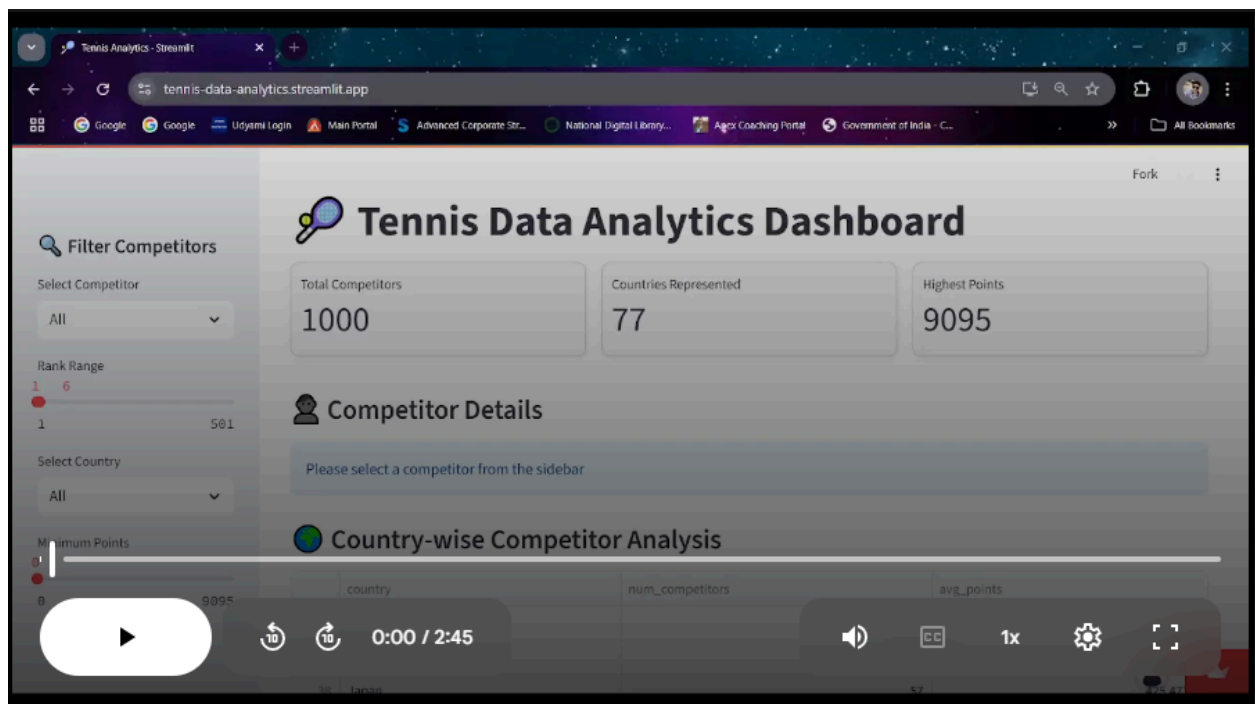
Competitor Details

	name	rank	movement	points	competitions_played	country
546	Aoyama, Shuko	47	0	1887	30	Japan

Country-wise Competitor Analysis

	country	num_competitors	avg_points
71	USA	105	827.619

Dashboard Demonstration Video:



Streamlit Dashboard Insights:

1. The total of 1000 players from 77 countries were being analyzed, representing broad global competition for professional tennis.

2. Countries with a majority of players are the USA, Great Britain, France, and Australia, which shows strong tennis infrastructure and player development programs in these nations.
3. Their highest score is 9095, which indicates significant power through the ranking system by top scorers.
4. The USA is leading with the most competitors and average points, which suggests both quantity and quality in terms of players. Other countries, such as Great Britain, Australia, and Italy, also show high average points, which belong to competitive players.
5. Leaderboard Top Picks:
 - Top-ranked Competitor: Pavic, Mate (Croatia) with 8280 points
 - Highest point Scorer: Siniakova, Katerina (Czechia) with 9095 pointsSeveral competitors share the top rank, reflecting tie conditions considered in ranking logic.
6. The dashboard can filter in real time by rank range, minimum points, country, and competitor name to help with defining particular player segments or regional trends.
7. The "Filtered Competitors" list contains most with zero rank movement, which shows consistent performance or level scores over the most recent period covered.
8. The visualizations (e.g., the top 10 countries by competitor count bar chart) show quick, intuitive insight into regional presence in global tennis.

Conclusion:

This project successfully uncovered the end-to-end process of transforming raw sports data into meaningful insights with the help of modern data tools and technologies. By way of extracting structured data from the SportRadar API, processing it through

MySQL and SQLite databases and presenting it in an interactive Streamlit dashboard, the project successfully met its target of developing interactive tennis data analysis. With ease, users can view player rankings, country-level performance, and event trends, and hence the platform is useful for analysts, enthusiasts, and sports business decision-makers. On the whole, the project developed technical and analytical skills while demonstrating the real-world application of data-driven storytelling in sports analytics.