Tennis Data Analytics Dashboard

Project Workflow

Introduction:

As part of my data engineering and analytics learning journey, I developed a Tennis Data Analytics Dashboard using the Sportradar Tennis API. My objective was to design a pipeline that ingests real-world sports data, stores it efficiently in a structured format, and presents it through an interactive and user-friendly dashboard.

This project involved three core components:

- **Python** for data ingestion and transformation
- **MySQL** for persistent storage
- **Streamlit** for interactive data visualization

Python – Data Collection, Cleaning & Processing:

API Integration

Using the key libraries: pandas, requests, I connected to the Sportradar Tennis API and retrieved structured JSON data from multiple endpoints:

- Categories and Competitions
- Complexes and Venues
- Doubles Competitor Rankings

Data Transformation

I applied the following steps using **pandas**:

- Flattened nested JSON structures
- Extracted relevant fields (e.g., category.id, competition.name, rank.points)
- Renamed columns for clarity and consistency
- Handled missing values (NaN, None)

Each cleaned dataset was then exported to **CSV** as an intermediary step before insertion into the database. This modular approach allowed for better debugging and reuse.

MySQL – Data Storage and Structuring:

To ensure persistence and scalability, I designed a relational schema in **MySQL**. I created three core tables:

- category_table
- complex_table
- competitor_table

I used **mysql.connector** in Python to automate data insertion from CSV into MySQL.

Key Features:

- Data type enforcement and cleaning before insertion
- Dynamic table and column referencing
- Graceful handling of special characters and nulls

Streamlit – Interactive Dashboard Development

I used **Streamlit** to build a fully interactive web-based dashboard that allows users to explore and filter tennis data with ease.

Dashboard Architecture

The dashboard is organized into three main sections:

- **Category View**: Explore competitions by gender, type, and region
- **Complex View**: View venue information by country and timezone
- **Competitor Rankings**: Analyze doubles competitors based on rank, country, and movement

Streamlit Features Implemented

- Sidebar filters with selectbox, multiselect, and slider
- Real-time table rendering with filtered data
- Download buttons for exporting filtered views
- Clean UI layout using st.tabs() and st.columns()

Tools Used:

- **Python** for data wrangling
- **MySQL** for backend storage
- **Streamlit** for frontend/dashboard
- **Pandas** for data manipulation
- **Sportradar Tennis API** for real-time data

Future Enhancements

- Add visual analytics (bar charts, pie charts by country)
- Integrate player performance history
- Add authentication layer to restrict access
- Enable live API refresh directly from Streamlit