**🏏 Cricsheet Match Data Analysis**

**Domain:** Sports Analytics / Data Analysis  
**Duration:** 1 Week  
**Tools & Technologies:** Selenium, Python, Pandas, SQL (MySQL/SQLite), Power BI, Matplotlib, Seaborn, Plotly

**🎯 Objective**

To automate the extraction, transformation, analysis, and visualization of cricket match data from Cricsheet. The project covers scraping JSON files, storing structured data in SQL databases, performing analytical queries, and building interactive dashboards.

📌 Problem Statement

Scrape JSON match data (Test, ODI, T20) from Cricsheet, store it in SQL tables, and analyze it using SQL and Python. Build a Power BI dashboard to visualize player and team performance metrics.

💼 Business Use Cases

* Player performance across formats
* Team comparisons and win/loss trends
* Match outcome analysis
* Strategic insights for coaches and analysts
* Fan engagement through interactive dashboards

🔍 Approach

1. Data Scraping

* Use Selenium to navigate and download JSON files from [Cricsheet](https://cricsheet.org/matches/)
* Target formats: Test, ODI, T20, IPL

2. Data Transformation

* Parse JSON files using Pandas
* Create separate DataFrames for each match type

3. Database Management

* Design SQL tables: test\_matches, odi\_matches, t20\_matches
* Insert cleaned data using SQLAlchemy or connectors

4. SQL Analysis

* Write 20 SQL queries to extract insights
* Ex: Top batsmen, leading wicket-takers, win percentages, century counts, close matches

5. EDA with Python

* Generate 10 visualizations using matplotlib, seaborn, plotly
* Present insights in a slide deck

6. Power BI Dashboard

* Connect to SQL database
* Visualize player trends, match outcomes, team comparisons

📊 Results

* Automated scraping and structured storage of match data
* Insightful SQL queries for performance analysis
* Python-based EDA and presentation
* Interactive Power BI dashboard

📦 Deliverables

* Python scripts for scraping and transformation
* SQL schema and query file
* EDA presentation
* Power BI dashboard (.pbix)
* Project documentation