**Cricsheet Match Analysis**

# 1. Objective:

* Scrape cricket match data (Test, ODI, T20, IPL) from Cricsheet using Selenium.
* Process and structure the data using Python and Pandas.
* Store it in SQL databases.
* Analyze the data using SQL queries and Python EDA.
* Build a Power BI dashboard to visualize player and team performance trends.

# 2. Skills Applied:

* Data Cleaning & Processing: Python (pandas, JSON handling).
* SQL: SQLite for managing structured cricket data.
* EDA & Visualization: Python (matplotlib, seaborn, plotly).
* BI Tools: Power BI dashboard creation.

# 3. Dataset Overview:

* Source: Cricsheet.org.
* Format: Nested JSON files per match.
* Match Types: Test, ODI, T20, IPL.
* Match metadata (date, venue, teams, toss).
* Innings data (overs, runs, wickets).
* Player performance (batter, bowler stats).

# 4. Technical Stack:

* Processing: Python, Pandas.
* Storage: SQLite.
* Analysis: SQL queries, Python (EDA).
* Visualization: Power BI, Matplotlib, Seaborn, Plotly.

# 5. Methodology:

1. Data Transformation:

* Parsed JSONs and flattened nested structures into pandas Data Frames.
* Created separate Data Frames for Test, ODI, T20, IPL matches.

2. Database Management (SQL):

* Created SQLite tables for each Matches.
* Inserted cleaned data using SQLAlchemy.

3. SQL-Based Data Analysis:

* Wrote 10+ optimized SQL queries (e.g., top players, teams, win margins).

4. Exploratory Data Analysis (EDA):

* Built 10+ plots using matplotlib, seaborn, and plotly.

5. Power BI Dashboard:

* Connected Power BI to SQLite.
* Created filters, trend visuals for player/team analysis.

# 6. Key Insights:

Player Performance:

* Players like Virat Kohli, Babar Azam, Joe Root stood out in respective formats.
* Top bowlers varied across formats showing format specialization.

Team Insights:

* India, Australia, and England topped win %.
* Afghanistan and WI performed well in T20s.

Match Outcomes:

* More narrow victories in T20s.
* IPL matches showed highest match variance.

# 7. Deliverables:

* Project\_2.ipynb: Python script for scraping & transformation.
* Cricket Analysis Dashboard.pbix: Power BI dashboard file.
* Cricsheet Match Analysis Report: Documentation.
* SQLite DB: SQL tables for all formats.
* EDA visualizations: Python plots.

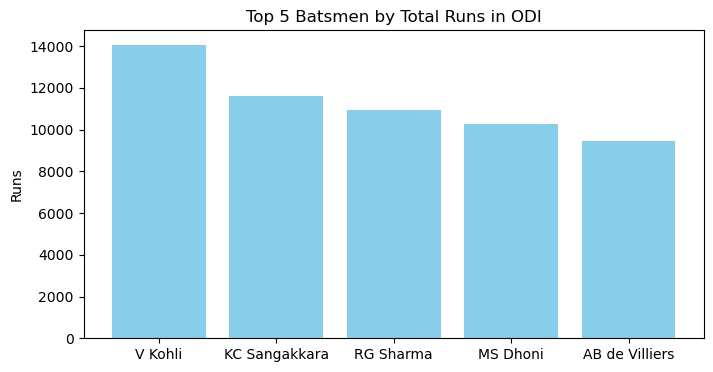
# 8. Conclusion:

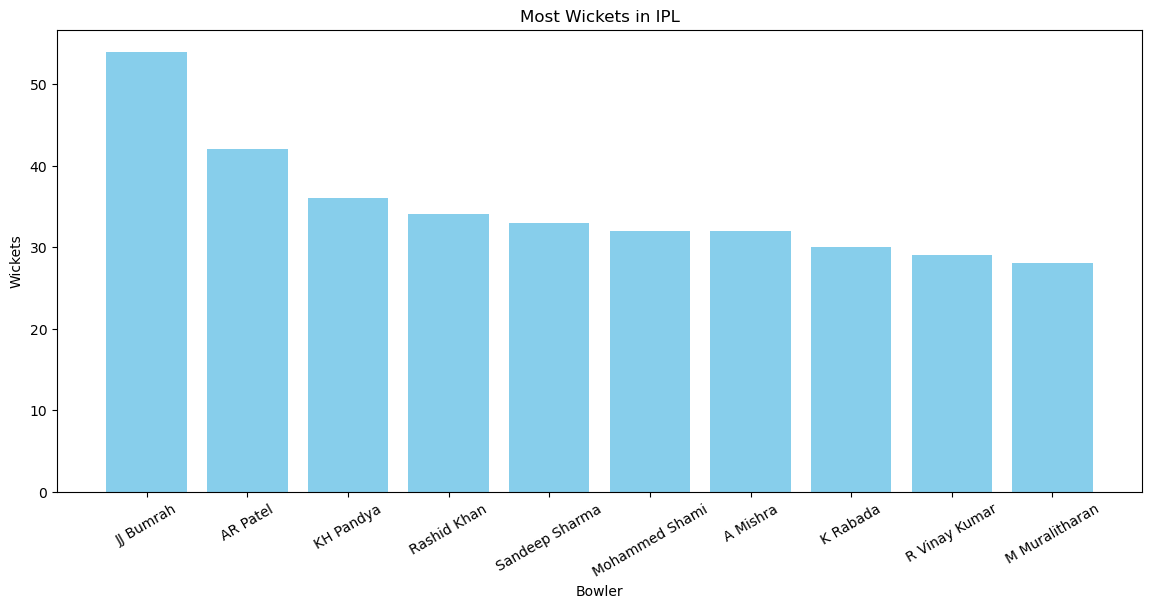
This project successfully demonstrates an end-to-end data analytics pipeline:

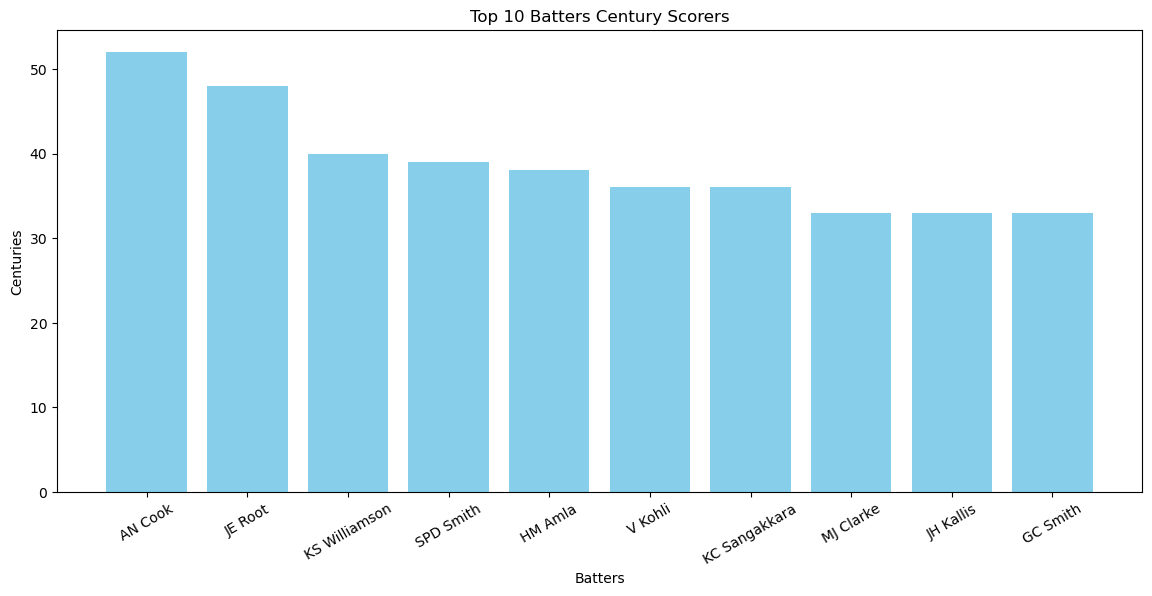
* Scraping real-world cricket data.
* Storing and analyzing using SQL.
* Visualizing insights via Power BI.
* The solution delivers actionable insights for analysts, teams, and fans.

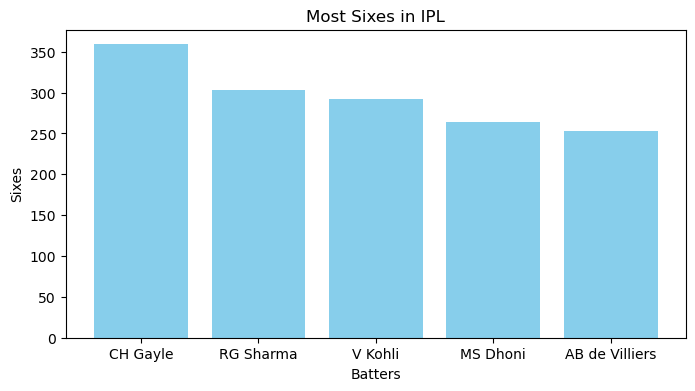
# 9. Visualizations from EDA:

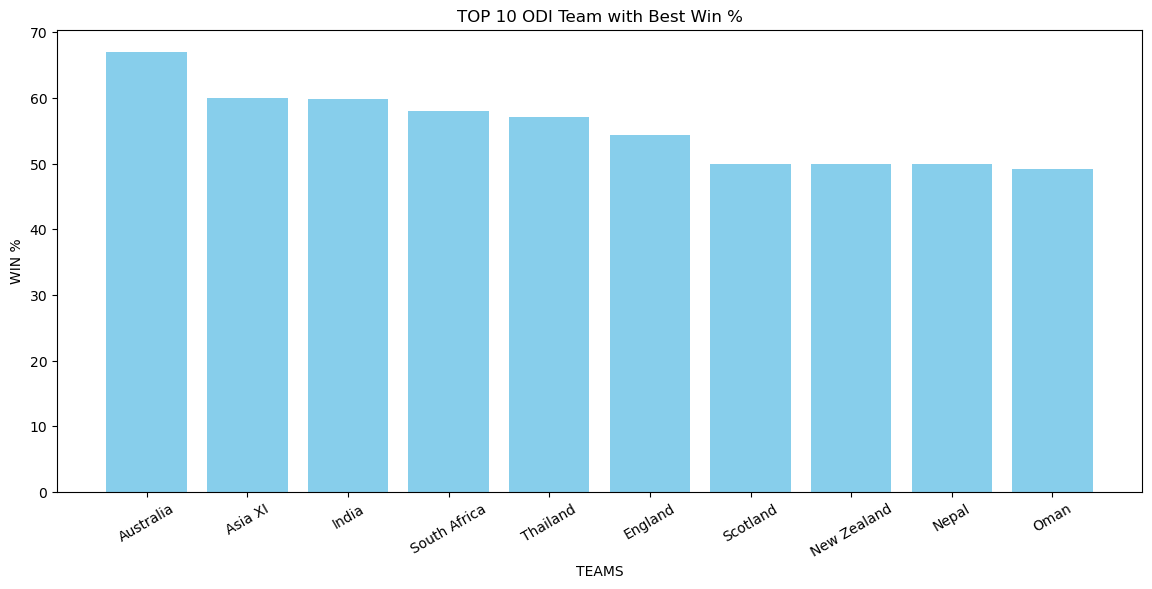
Below are sample visualizations generated using matplotlib, seaborn, and plotly from the EDA notebook.











# 10. Power BI Dashboard Screenshot:

