PROJECT REPORT

T20 World Cup Data Analysis & Visualization

1. Introduction

Cricket is one of the most popular sports in the world, and the ICC T20 World Cup generates a massive amount of match and player data. Analyzing this data helps in understanding team performance, player consistency, match trends, and venue impact.

This project focuses on analyzing **T20 World Cup match data** and presenting meaningful insights through **interactive dashboards using Google Looker Studio**. The project covers batting, bowling, team performance, venue analysis, and player comparisons.

---

2. Objectives of the Project

The main objectives of this project are:

- To analyze team and player performance in the T20 World Cup
- To identify top-performing batsmen and bowlers
- To analyze match results across venues and countries
- To provide interactive visualizations for easy insights
- To enable player and team comparison using filters

---

3. Dataset Description

The project uses structured cricket datasets containing match-wise, batting, and bowling information.

3.1 Match Data

- `team1`, `team2`
- `winner`, `margin`
- `ground`, `matchDate`
- `match_id`

3.2 Batting Data

- `match`, `teamInnings`, `battingPos`
- `batsmanName`, `runs`, `balls`
- `4s`, `6s`, `strike rate`
- `match_id`

3.3 Bowling Data

- `match`, `bowlingTeam`, `bowlerName`
- `overs`, `runs`, `wickets`
- `wides`, `no balls`
- `match_id`

All datasets are joined using **match_id** and **player names** inside Looker Studio.

---

4. Tools & Technologies Used

- **Google Looker Studio** – Data visualization and dashboards
- **Google Colab / CSV** – Data Cleaning and preprocessing
- **Python** – Data cleaning and preprocessing
- **Excel** – Initial data formatting

---

5. Methodology

1. **Data Collection**
   - Data collected from cricket sources (ESPN Cricinfo / Wikipedia).
2. **Data Preprocessing**
   - Removed null values
   - Standardized team and player names
   - Verified match_id consistency
3. **Data Integration**
   - Joined Match, Batting, and Bowling datasets using `match_id`.
4. **Dashboard Creation**
   - Created calculated fields (strike rate, economy, totals)
   - Added filters for teams, players, venues
   - Designed multiple dashboard pages

---

6. Visualizations & Analysis

6.1 Team Performance

- Top 10 countries by number of wins
- Total runs scored by each team
- Match win distribution

6.2 Batting Analysis

- Top 10 batsmen by runs
- Strike rate vs runs comparison
- Balls faced vs total runs
- Player-wise performance filters

6.3 Bowling Analysis

- Top 10 bowlers by wickets
- Runs conceded vs wickets
- Overs bowled and extras analysis

6.4 Venue Analysis

- Best-performing teams by ground
- Matches won at specific venues

6.5 Player Comparison

- Side-by-side comparison of two batsmen
- Runs, strike rate, balls faced
- Match-wise performance trends

---

7. Dashboard Features

- Interactive filters (Team, Player, Match, Venue)
- Dynamic charts and tables
- Drill-down analysis
- Real-time insights without coding

---

8. Results & Insights

- Identified consistent top performers like **Virat Kohli** and **Suryakumar Yadav**
- Found key wicket-takers like **Arshdeep Singh** and **Ravichandran Ashwin**
- Highlighted venue-specific performance trends
- Enabled clear team-wise and player-wise comparison

---

9. Conclusion

This project successfully demonstrates how cricket data can be transformed into meaningful insights using data visualization tools. The interactive dashboards help analysts, fans, and selectors understand performance trends easily.

---

10. Future Enhancements

- Add predictive analysis (win probability)
- Include ball-by-ball data

- Integrate Power BI or Tableau
- Add real-time match updates

---

11. References

- ESPN Cricinfo
- ICC Official Website
- Google Looker Studio Documentation