GAME INSIGHT: ANALYSING TOURNAMENTS (CRICKET) WITH POWER BI

A Project Report

Submitted in partial fulfilment of the requirements for the

Award of the degree of

MASTER OF COMPUTER APPLICATIONS

By

Ashish Sharma

(322100287)

LPUOnline

Same Degree, Now Online.

Entitled by UGC

Centre for Distance and Online Education

LOVELY PROFESSIONAL UNIVERSITY

PHAGWARA, PUNJAB

2022-24

Table of Contents

Declaration by the Student

To whom-so-ever it may concern

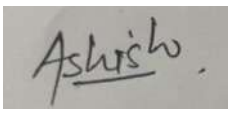
I, **Ashish Sharma, 322100287,** hereby declare that the work done by me on **Game Insight: Analysing Tournaments (Cricket) with Power BI,** is a record of original work for the partial fulfilmentof the requirements for the award of the degree, **MCA.**

Name of the Student (Registration Number)

Ashish Sharma (322100287)

Signature of the student

Ashish Sharma



Dated: 30 July 2024

Acknowledgement

I am grateful for the opportunity to present "Game Insight: Analysing Tournaments (Cricket) with Power BI," completed during my last semester. This project has greatly enhanced my skills in data analysis and visualization for cricket tournaments.

I want to thank Lovely Professional University for providing the resources and environment that made this project possible. I also appreciate the support and collaboration from my peers, whose insights have been invaluable.

A special thanks to my family and friends for their constant encouragement and belief in my abilities. Their support has been a source of motivation throughout this journey.

This project has been a significant milestone in my academic career, and I am excited to apply the knowledge and skills gained to future endeavours.

Abstract

This project, "Game Insight: Analyzing Tournaments (Cricket) with Power BI," evaluates IPL seasons from 2021 to 2023 using Power BI for comprehensive performance analysis. It highlights key insights into team dynamics and player performances, showcasing the Chennai Super Kings' and Gujarat Titans' achievements. The project provides detailed dashboards and visualizations on batting and bowling statistics, team performance, and tournament records. Metrics include top batsmen by runs, strike rates, and sixes, as well as leading bowlers by wickets, economy rates, and maiden overs. Additional analyses cover team win percentages, boundary statistics, and best all-rounders and playing XI. The use of Power BI facilitates effective data visualization and insightful cricket analytics.

List of Tables

1. dim_match_summary

Columns: team1, team2, winner, margin, matchDate, match_id, Year, Wickets.2, Chase win

Description: Contains summary information of each match, including teams, winner, match date, and relevant statistics.

team1	team2	winner	margin	matchDate	match_id	Year	Wickets.2	Chase WIN
Capitals	KKR	KKR	3 wickets	13 October 2021	T206442	2021	wickets	1
RCB	KKR	KKR	4 wickets	11 October 2021	T208597	2021	wickets	1
Capitals	Super Kings	Super Kings	4 wickets	10 October 2021	T201819	2021	wickets	1
Capitals	RCB	RCB	7 wickets	08 October 2021	T205942	2021	wickets	1
Super Kings	Punjab Kings	Punjab Kings	6 wickets	07 October 2021	T201866	2021	wickets	1
Royals	Mumbai	Mumbai	8 wickets	05 October 2021	T205599	2021	wickets	1
Super Kings	Capitals	Capitals	3 wickets	04 October 2021	T207682	2021	wickets	1
Sunrisers	KKR	KKR	6 wickets	03 October 2021	T207430	2021	wickets	1
Super Kings	Royals	Royals	7 wickets	02 October 2021	T205672	2021	wickets	1
Mumbai	Capitals	Capitals	4 wickets	02 October 2021	T209349	2021	wickets	1
KKR	Punjab Kings	Punjab Kings	5 wickets	01 October 2021	T201051	2021	wickets	1
Sunrisers	Super Kings	Super Kings	6 wickets	30 September 2021	T203061	2021	wickets	1
Royals	RCB	RCB	7 wickets	29 September 2021	T204742	2021	wickets	1
Punjab Kings	Mumbai	Mumbai	6 wickets	28 September 2021	T208262	2021	wickets	1
Capitals	KKR	KKR	3 wickets	28 September 2021	T206700	2021	wickets	1
Royals	Sunrisers	Sunrisers	7 wickets	27 September 2021	T206862	2021	wickets	1
KKR	Super Kings	Super Kings	2 wickets	26 September 2021	T202406	2021	wickets	1
RCB	Super Kings	Super Kings	6 wickets	24 September 2021	T207776	2021	wickets	1
Mumbai	KKR	KKR	7 wickets	23 September 2021	T202680	2021	wickets	1
Sunrisers	Capitals	Capitals	8 wickets	22 September 2021	T204364	2021	wickets	1
RCB	KKR	KKR	9 wickets	20 September 2021	T204307	2021	wickets	1
Punjab Kings	Capitals	Capitals	7 wickets	02 May 2021	T208407	2021	wickets	1
Super Kings	Mumbai	Mumbai	4 wickets	01 May 2021	T203694	2021	wickets	1
KKR	Capitals	Capitals	7 wickets	29 April 2021	T201595	2021	wickets	1
Royals	Mumbai	Mumbai	7 wickets	29 April 2021	T209461	2021	wickets	1
Sunrisers	Super Kings	Super Kings	7 wickets	28 April 2021	T205094	2021	wickets	1
Punjab Kings	KKR	KKR	5 wickets	26 April 2021	T209997	2021	wickets	1
KKR	Royals	Royals	6 wickets	24 April 2021	T205585	2021	wickets	1
Mumbai	Punjab Kings	Punjab Kings	9 wickets	23 April 2021	T204808	2021	wickets	1
Royals	RCB	RCB	10 wickets	22 April 2021	T201545	2021	wickets	1

Table: dim_match_summary (206 rows)

2. dim_players

Columns: team, batting style, bowling style, playingrole, custom batting order, name

Description: Contains player details including their team, playing roles, and styles.

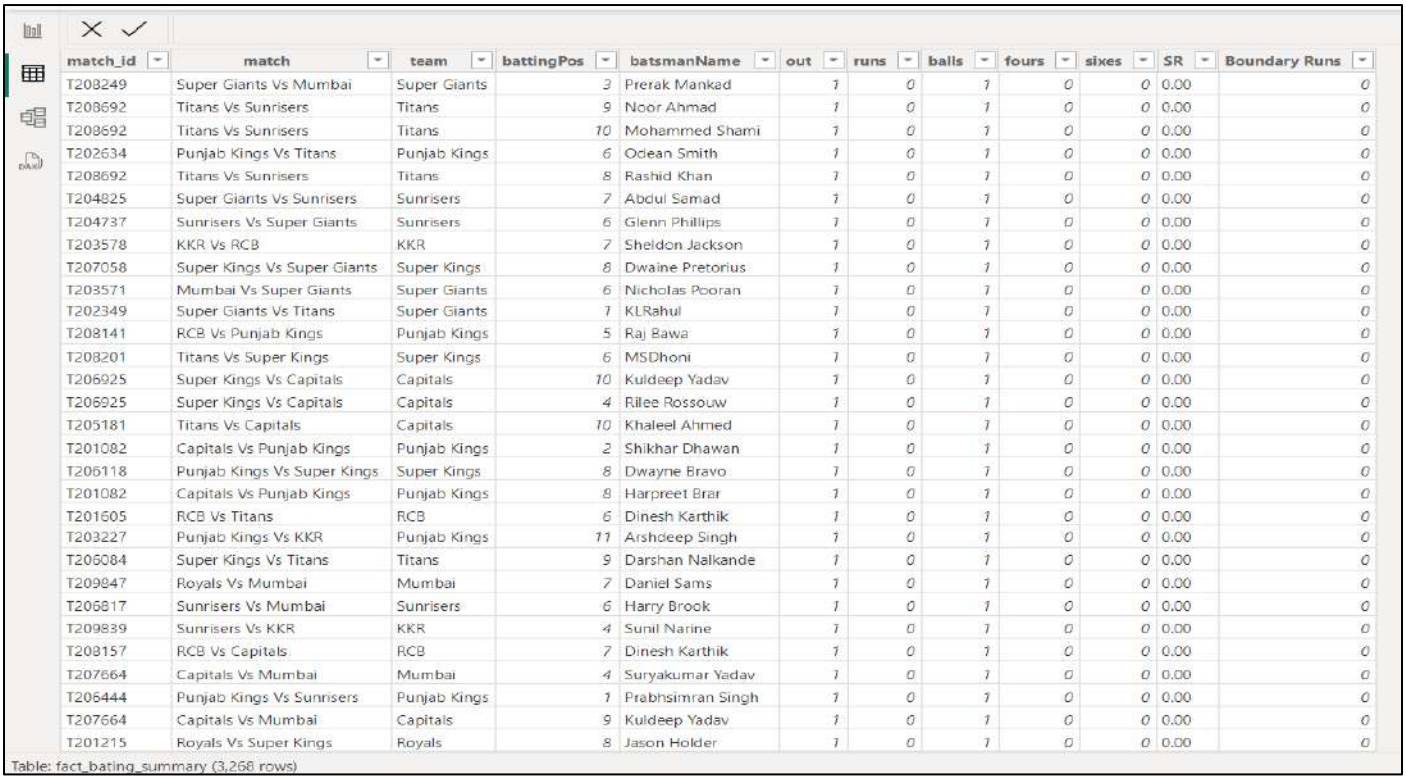
team	battingStyle	bowlingStyle	playingRole	Custom Batting Order	name
KKR	Right hand Bat	Right arm Fast	Bowler		Lockie Ferguson
KKR	Right hand Bat	Right arm Fast medium	Bowler		Shivam Mavi
KKR	Right hand Bat	Legbreak Googly	Bowler		Varun Chakravarthy
Super Kings	Right hand Bat	Right arm Medium	Bowler		Deepak Chahar
Super Kings	Left hand Bat	Right arm Fast medium	Bowler		Josh Hazlewood
Super Kings	Right hand Bat	Right arm Medium	Bowler		Shardul Thakur
Capitals	Right hand Bat	Right arm Fast	Bowler		Anrich Nortje
Capitals	Right hand Bat	Right arm Fast medium	Bowler		Avesh Khan
Capitals	Left hand Bat	Right arm Fast	Bowler		Kagiso Rabada
RCB	Right hand Bat	Right arm Medium	Bowler		Harshal Patel
RCB	Right hand Bat	Right arm Fast	Bowler		Mohammed Siraj
RCB	Right hand Bat	Legbreak Googly	Bowler	10	Yuzvendra Chahal
Mumbai	Right hand Bat	Right arm Fast	Bowler		Jasprit Bumrah
Mumbai	Right hand Bat	Left arm Fast medium	Bowler	8	Trent Boult
Sunrisers	Right hand Bat	Right arm Medium	Bowler		Siddarth Kaul
Sunrisers	Right hand Bat	Right arm Fast	Bowler		Umran Malik
Royals	Right hand Bat	Left arm Medium	Bowler		Jaydev Unadkat
Royals	Left hand Bat	Left arm Medium fast	Bowler		Chetan Sakariya
Royals	Left hand Bat	Left arm Fast medium	Bowler		Mustafizur Rahman
Punjab Kings	Right hand Bat	Right arm Fast	Bowler	11	Mohammed Shami
Punjab Kings	Left hand Bat	Slow Left arm Orthodox	Bowler		Harpreet Brar
Punjab Kings	Left hand Bat	Left arm Medium fast	Bowler		Arshdeep Singh
Punjab Kings	Right hand Bat	Right arm Fast medium	Bowler		Chris Jordan
Punjab Kings	Right hand Bat	Legbreak Googly	Bowler		Ravi Bishnoi
Sunrisers	Right hand Bat	Right arm Medium	Bowler	9	Bhuvneshwar Kumar
Royals	Left hand Bat	Left arm Fast medium	Bowler		Kuldip Yadav
Mumbai	Right hand Bat	Right arm Offbreak	Bowler		Jayant Yadav
KKR	Right hand Bat	Right arm Medium fast	Bowler		Tim Southee
Royals	Right hand Bat	Left arm Medium fast	Bowler		Akash Singh
Royals	Right hand Bat	Legbreak	Bowler		Mayank Markande

Table: dim_players (292 rows)

3. fact_batting_summary

Columns: match_id, match, team, battingpos, batsmanName, out, runs, balls, fours, sixes, SR, Boundary Runs

Description: Details individual batting performances for each match, including runs, balls faced, and boundary statistics.



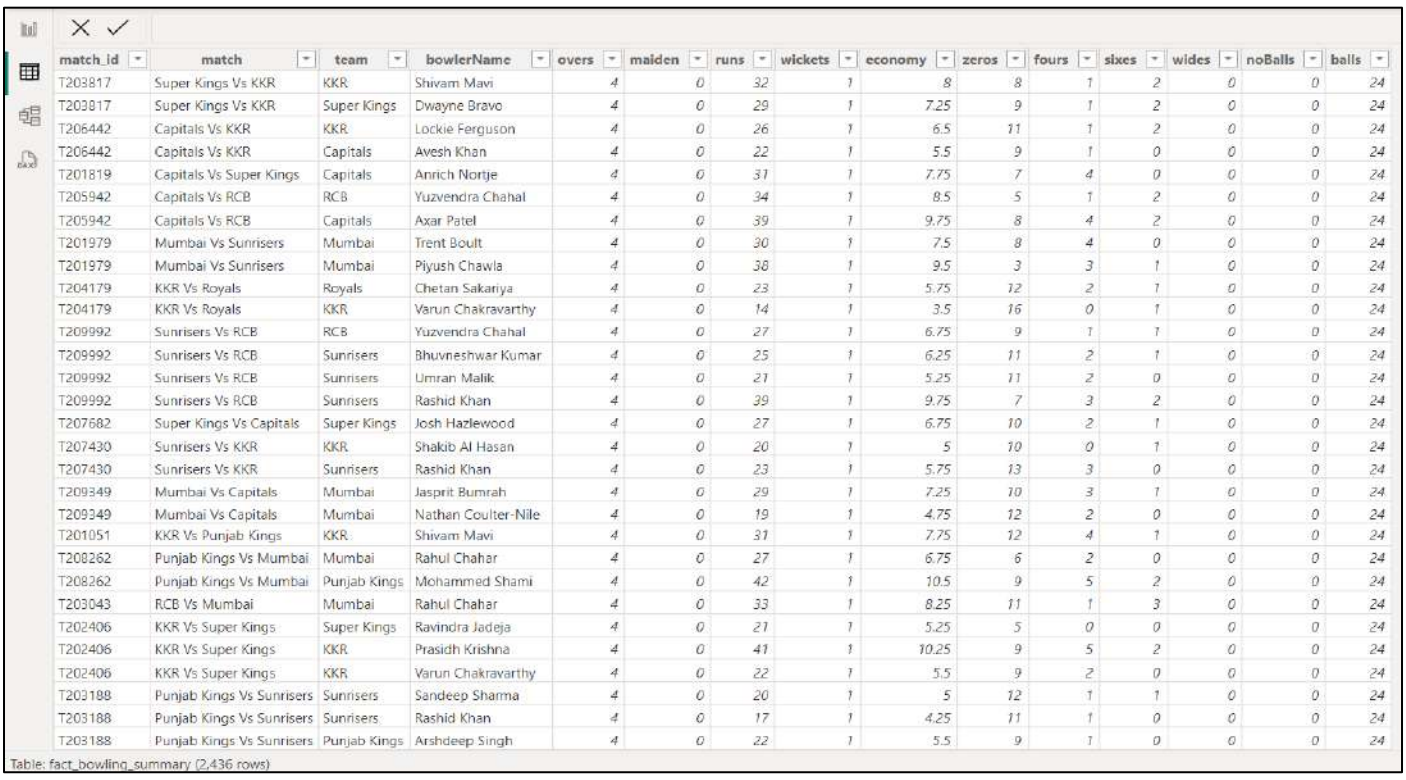
match_id	match	team	battingPos	batsmanName	out	runs	balls	fours	sixes	SR	Boundary Runs
T208249	Super Giants Vs Mumbai	Super Giants	3	Prerak Mankad	1	0	1	0	0	0.00	0
T208692	Titans Vs Sunrisers	Titans	9	Noor Ahmad	1	0	1	0	0	0.00	0
T208692	Titans Vs Sunrisers	Titans	10	Mohammed Shami	1	0	1	0	0	0.00	0
T202634	Punjab Kings Vs Titans	Punjab Kings	6	Odean Smith	1	0	1	0	0	0.00	0
T208692	Titans Vs Sunrisers	Titans	8	Rashid Khan	1	0	1	0	0	0.00	0
T204825	Super Giants Vs Sunrisers	Sunrisers	7	Abdul Samad	1	0	1	0	0	0.00	0
T204737	Sunrisers Vs Super Giants	Sunrisers	6	Glenn Phillips	1	0	1	0	0	0.00	0
T203578	KKR Vs RCB	KKR	7	Sheldon Jackson	1	0	1	0	0	0.00	0
T207058	Super Kings Vs Super Giants	Super Kings	8	Dwaine Pretorius	1	0	1	0	0	0.00	0
T203571	Mumbai Vs Super Giants	Super Giants	6	Nicholas Pooran	1	0	1	0	0	0.00	0
T202349	Super Giants Vs Titans	Super Giants	1	KLRahul	1	0	1	0	0	0.00	0
T208141	RCB Vs Punjab Kings	Punjab Kings	5	Raj Bawa	1	0	1	0	0	0.00	0
T208201	Titans Vs Super Kings	Super Kings	6	MSDhoni	1	0	1	0	0	0.00	0
T206925	Super Kings Vs Capitals	Capitals	10	Kuldeep Yadav	1	0	1	0	0	0.00	0
T206925	Super Kings Vs Capitals	Capitals	4	Rilee Rossouw	1	0	1	0	0	0.00	0
T205181	Titans Vs Capitals	Capitals	10	Khaleel Ahmed	1	0	1	0	0	0.00	0
T201082	Capitals Vs Punjab Kings	Punjab Kings	2	Shikhar Dhawan	1	0	1	0	0	0.00	0
T206118	Punjab Kings Vs Super Kings	Super Kings	8	Dwayne Bravo	1	0	1	0	0	0.00	0
T201082	Capitals Vs Punjab Kings	Punjab Kings	8	Harpreet Brar	1	0	1	0	0	0.00	0
T201605	RCB Vs Titans	RCB	6	Dinesh Karthik	1	0	1	0	0	0.00	0
T203227	Punjab Kings Vs KKR	Punjab Kings	11	Arshdeep Singh	1	0	1	0	0	0.00	0
T206084	Super Kings Vs Titans	Titans	9	Darshan Nalkande	1	0	1	0	0	0.00	0
T209847	Royals Vs Mumbai	Mumbai	7	Daniel Sams	1	0	1	0	0	0.00	0
T206817	Sunrisers Vs Mumbai	Sunrisers	6	Harry Brook	1	0	1	0	0	0.00	0
T209839	Sunrisers Vs KKR	KKR	4	Sunil Narine	1	0	1	0	0	0.00	0
T208157	RCB Vs Capitals	RCB	7	Dinesh Karthik	1	0	1	0	0	0.00	0
T207664	Capitals Vs Mumbai	Mumbai	4	Suryakumar Yadav	1	0	1	0	0	0.00	0
T206444	Punjab Kings Vs Sunrisers	Punjab Kings	1	Prabhsimran Singh	1	0	1	0	0	0.00	0
T207664	Capitals Vs Mumbai	Capitals	9	Kuldeep Yadav	1	0	1	0	0	0.00	0
T201215	Royals Vs Super Kings	Royals	8	Jason Holder	1	0	1	0	0	0.00	0

Table: fact_bating_summary (3,268 rows)

4. fact_bowling_summary

Columns: match_id, match, team, bowlerName, overs, maiden, runs, wickets, economy, zeroes, fours, sixes, wides, noBalls, balls

Description: Details individual bowling performances for each match, including overs bowled, wickets taken, and other bowling statistics.



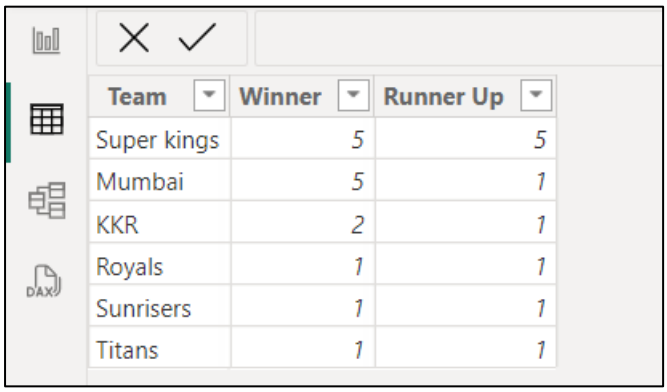
match_id	match	team	bowlerName	overs	maiden	runs	wickets	economy	zeros	fours	sixes	wides	noBalls	balls
T203817	Super Kings Vs KKR	KKR	Shivam Mavi	4	0	32	1	8	8	1	2	0	0	24
T203817	Super Kings Vs KKR	Super Kings	Dwayne Bravo	4	0	29	1	7.25	9	1	2	0	0	24
T206442	Capitals Vs KKR	KKR	Lockie Ferguson	4	0	26	1	6.5	11	1	2	0	0	24
T206442	Capitals Vs KKR	Capitals	Avesh Khan	4	0	22	1	5.5	9	1	0	0	0	24
T201819	Capitals Vs Super Kings	Capitals	Anrich Nortje	4	0	31	1	7.75	7	4	0	0	0	24
T205942	Capitals Vs RCB	RCB	Yuzvendra Chahal	4	0	34	1	8.5	5	1	2	0	0	24
T205942	Capitals Vs RCB	Capitals	Axar Patel	4	0	39	1	9.75	8	4	2	0	0	24
T201979	Mumbai Vs Sunrisers	Mumbai	Trent Boult	4	0	30	1	7.5	8	4	0	0	0	24
T201979	Mumbai Vs Sunrisers	Mumbai	Piyush Chawla	4	0	38	1	9.5	3	3	1	0	0	24
T204179	KKR Vs Royals	Royals	Chetan Sakariya	4	0	23	1	5.75	12	2	1	0	0	24
T204179	KKR Vs Royals	KKR	Varun Chakravarthy	4	0	14	1	3.5	16	0	1	0	0	24
T209992	Sunrisers Vs RCB	RCB	Yuzvendra Chahal	4	0	27	1	6.75	9	1	1	0	0	24
T209992	Sunrisers Vs RCB	Sunrisers	Bhuvneshwar Kumar	4	0	25	1	6.25	11	2	1	0	0	24
T209992	Sunrisers Vs RCB	Sunrisers	Umran Malik	4	0	21	1	5.25	11	2	0	0	0	24
T209992	Sunrisers Vs RCB	Sunrisers	Rashid Khan	4	0	39	1	9.75	7	3	2	0	0	24
T207682	Super Kings Vs Capitals	Super Kings	Josh Hazlewood	4	0	27	1	6.75	10	2	1	0	0	24
T207430	Sunrisers Vs KKR	KKR	Shakib Al Hasan	4	0	20	1	5	10	0	1	0	0	24
T207430	Sunrisers Vs KKR	Sunrisers	Rashid Khan	4	0	23	1	5.75	13	3	0	0	0	24
T209349	Mumbai Vs Capitals	Mumbai	Jasprit Bumrah	4	0	29	1	7.25	10	3	1	0	0	24
T209349	Mumbai Vs Capitals	Mumbai	Nathan Coulter-Nile	4	0	19	1	4.75	12	2	0	0	0	24
T201051	KKR Vs Punjab Kings	KKR	Shivam Mavi	4	0	31	1	7.75	12	4	1	0	0	24
T208262	Punjab Kings Vs Mumbai	Mumbai	Rahul Chahar	4	0	27	1	6.75	6	2	0	0	0	24
T208262	Punjab Kings Vs Mumbai	Punjab Kings	Mohammed Shami	4	0	42	1	10.5	9	5	2	0	0	24
T203043	RCB Vs Mumbai	Mumbai	Rahul Chahar	4	0	33	1	8.25	11	1	3	0	0	24
T202406	KKR Vs Super Kings	Super Kings	Ravindra Jadeja	4	0	21	1	5.25	5	0	0	0	0	24
T202406	KKR Vs Super Kings	KKR	Prasidh Krishna	4	0	41	1	10.25	9	5	2	0	0	24
T202406	KKR Vs Super Kings	KKR	Varun Chakravarthy	4	0	22	1	5.5	9	2	0	0	0	24
T203188	Punjab Kings Vs Sunrisers	Sunrisers	Sandeep Sharma	4	0	20	1	5	12	1	1	0	0	24
T203188	Punjab Kings Vs Sunrisers	Sunrisers	Rashid Khan	4	0	17	1	4.25	11	1	0	0	0	24
T203188	Punjab Kings Vs Sunrisers	Punjab Kings	Arshdeep Singh	4	0	22	1	5.5	9	1	0	0	0	24

Table: fact_bowling_summary (2,436 rows)

5. Top teams

Columns: Team, Winner, RunnerUp

Description: Provides a summary of the top teams from each year, including the winner and runner-up.



Team	Winner	Runner Up
Super kings	5	5
Mumbai	5	1
KKR	2	1
Royals	1	1
Sunrisers	1	1
Titans	1	1

List of Figures/Charts

List of Schemas/Algorithms

- Data Cleaning and Preprocessing - Prepare data by handling missing values, duplicates, and formatting issues.
- Calculating Batting Metrics - Compute batting average, strike rate, and boundary percentage.
- Calculating Bowling Metrics - Compute bowling average, economy rate, and dot ball percentage.
- Orange and Purple Cap Players - Use of filtering based on most runs scored and most of wickets taken by an individual player in each year.
- Selecting Best XI and Top 3 All-rounders - Optimize team selection and identify top all-rounders based on past performance.

Dax queries used for creating useful measures-

1-Batting Analysis

i)Batting Average min 60 balls =

```
VAR TotalRuns = [Total Runs]
VAR TotalOuts = [Total Inning Dismissed]
VAR SeasonsOver60Balls =
    CALCULATE(
        COUNTROWS(
            FILTER(
                VALUES(dim_match_summary[Year]),
                [Total Balls Faced] >= 60
            )
        ),
        dim_match_summary[Year] >= 2021 && dim_match_summary[Year] <= 2023
    )
RETURN
    IF(
```

```
    NOT ISFILTERED(dim_match_summary[Year]),
    IF(
        SeasonsOver60Balls = 3,
        DIVIDE(TotalRuns, TotalOuts, 0),
        BLANK()
    ),
    IF(
        [Total Balls Faced] >= 60,
        DIVIDE(TotalRuns, TotalOuts, 0),
        BLANK()
    )
)
```

ii) Batting Strike min 60 balls =

```
VAR TotalRuns = [Total Runs]
VAR TotalBalls = [Total Balls Faced]
VAR SeasonsOver60Balls =
    CALCULATE(
        COUNTROWS(
            FILTER(
                VALUES(dim_match_summary[Year]),
                dim_match_summary[Year] >= 2021 &&
                dim_match_summary[Year] <= 2023 &&
                [Total Balls Faced] >= 60
```

```
            )
        )
    )
RETURN
    IF(
        NOT ISFILTERED(dim_match_summary[Year]),
        IF(
            SeasonsOver60Balls = 3,
            DIVIDE(TotalRuns, TotalBalls, 0) * 100,
            BLANK()
        ),
        IF(
            SeasonsOver60Balls > 0,
            DIVIDE(TotalRuns, TotalBalls, 0) * 100,
            BLANK()
        )
    )
```

iii) Boundary Runs = SUM(fact_bating_summary[Boundary Runs])

iv) Boundary% =

```
VAR TotalRuns = [Total Runs]
VAR BoundaryRuns = [Boundary Runs]
VAR SeasonsOver60Balls =
    CALCULATE(
```

```
        COUNTROWS(
            FILTER(
                VALUES(dim_match_summary[Year]),
                [Total Balls Faced] >= 60
            )
        ),
        dim_match_summary[Year] >= 2021 && dim_match_summary[Year] <= 2023
    )
RETURN
    IF(
        NOT ISFILTERED(dim_match_summary[Year]),
        IF(
            SeasonsOver60Balls = 3,
            DIVIDE(BoundaryRuns,TotalRuns,0),
            BLANK()
        ),
        IF(
            [Total Balls Faced] >= 60,
            DIVIDE(BoundaryRuns,TotalRuns,0),
            BLANK()
        )
    )
```

iv) runs_new = CONCATENATE(SUM(fact_bating_summary[runs])," Runs")

v) Total Balls Faced = SUM(fact_bating_summary[balls])

vi) Total Inning Batted = COunt(fact_bating_summary[match_id])

vii) Total Inning Dismissed =

```
CALCULATE(
    COUNTROWS(fact_bating_summary),
    fact_bating_summary[out] = 1)
```

viii) Total Runs = SUM(fact_bating_summary[runs])

2- Bowling Analysis

i) Bowling Average min 60 balls each season =

```
VAR TotalRunsConceded = [Total runs conceded]
VAR TotalWickets = [Total Wickets]
VAR SeasonsOver60Balls =
    CALCULATE(
        COUNTROWS(
            FILTER(
                VALUES(dim_match_summary[Year]),
                [Total balls bowled] >= 60
            )
        ),

        dim_match_summary[Year] >= 2021 &&
```

```
        dim_match_summary[Year] <= 2023
    )
RETURN
    IF(
        NOT ISFILTERED(dim_match_summary[Year]),
        IF(
            SeasonsOver60Balls = 3,
            DIVIDE(TotalRunsConceded, TotalWickets,0),
        BLANK()
    ),
    IF(
        [Total balls bowled] >=60,
        DIVIDE(TotalRunsConceded, TotalWickets,0),
        BLANK()
    )
    )
```

ii) Dot Balls = SUM(fact_bowling_summary[zeros])

iii) Dot Balls % Min 60 Balls =

```
VAR TotalDotBalls = [Dot Balls]
VAR TotalBalls = [Total Balls Bowled]
VAR SeasonsOver60Balls =
    CALCULATE(
        COUNTROWS(
```

```
        FILTER(
            VALUES(dim_match_summary[Year]),
            [Total balls bowled] >= 60
        )
    ),

    dim_match_summary[Year] >= 2021 &&
    dim_match_summary[Year] <= 2023
)
RETURN
  IF(
    NOT ISFILTERED(dim_match_summary[Year]),
    IF(
      SeasonsOver60Balls = 3,
      Divide([Dot Balls], [Total Balls Bowled],0),
    BLANK()
  ),
  IF(
    [Total balls bowled] >=60,
    Divide([Dot Balls], [Total Balls Bowled],0),
    BLANK()
  )
  )
```

iv) Economy Rate min 60 Balls =

```
VAR TotalRunsC = [Total Runs Conceded]
VAR TotalBallsB = [Total Balls Bowled]
VAR SeasonsOver60Balls =
    CALCULATE(
        COUNTROWS(
            FILTER(
                VALUES(dim_match_summary[Year]),
                [Total Balls Bowled] >= 60
            )
        ),
        dim_match_summary[Year] >= 2021 &&
        dim_match_summary[Year] <= 2023
    )
RETURN
    IF(
        NOT ISFILTERED(dim_match_summary[Year]),
        IF(
            SeasonsOver60Balls = 3,
            Divide(TotalRunsC, TotalBallsB,0) * 6,
            BLANK()
        ),
        IF(
```

```
    [Total Balls Bowled] >= 60,
    DIVIDE(TotalRunsC, TotalBallsB,0) *6,
    BLANK()
  ))
```

v) Total Balls Bowled = Sum(fact_bowling_summary[balls])

vi) Total Runs Conceded = sUM(fact_bowling_summary[runs])

vii) Total Wickets = Sum(fact_bowling_summary[wickets])

viii) wickets_new = CONCATENATE(SUM(fact_bowling_summary[wickets])," Wickets")

List of Symbols

- RRR: Total Runs Scored by a Batsman.
- BBB: Total Balls Faced by a Batsman.
- WWW: Total Wickets Taken by a Bowler.
- OOO: Total Overs Bowled by a Bowler.
- 4s: Number of Fours Hit.
- 6s: Number of Sixes Hit.
- SR: Strike Rate - (Total Runs / Total Balls Faced) × 100.
- Econ: Economy Rate - Total Runs Conceded / Total Overs Bowled.
- Avg: Average - Batting Average = Total Runs / Total Times Out; Bowling Average = Total Runs Conceded / Total Wickets Taken.
- %: Percentage used in metrics like boundary and dot ball percentages.
- 4+6: Total Boundaries (Fours + Sixes).
- Dot Ball: Ball with no runs scored.
- Dot Ball Percentage: (Total Dot Balls / Total Balls Bowled) × 100.
- Boundary Percentage: (Total Boundaries / Total Balls Faced) × 100.

List of Abbreviations

- IPL: Indian Premier League - A major T20 cricket league in India.
- CSV: Comma-Separated Values - A format for storing tabular data.
- dim: Dimension - Descriptive tables in a data warehouse.
- fact: Fact - Quantitative tables in a data warehouse.
- SR: Strike Rate - Rate of runs scored by a batsman.
- Econ: Economy Rate - Rate of runs conceded by a bowler.
- Avg: Average - Batting or bowling average.
- T: Match Identifier Prefix - Unique identifier for matches.
- 4s: Fours - Boundaries worth 4 runs hit by a batsman.
- 6s: Sixes - Boundaries worth 6 runs hit by a batsman.
- O: Overs - Total overs bowled by a bowler.
- W: Wickets - Total wickets taken by a bowler.
- Dot Ball: Ball with no runs scored.
- Boundary %: Boundary Percentage - Percentage of runs scored from boundaries.
- Dot Ball %: Dot Ball Percentage - Percentage of dot balls bowled.
- Orange Cap: Highest Run Scorer - Awarded to the player with most runs.
- Purple Cap: Highest Wicket Taker - Awarded to the player with most wickets.
- Winner: Winning Team - The team that wins the IPL season.
- Runner-up: Runner-up Team - The team that finishes second.

Chapter 1: Introduction

The Indian Premier League (IPL) is a premier T20 cricket tournament that has captured the imagination of cricket fans worldwide. With its exciting matches, star-studded line-ups, and high-stakes games, the IPL is a significant event in the cricket calendar. This project, "Game Insight: Analysing Tournaments (Cricket) with Power BI," aims to provide an in-depth analysis of the IPL seasons from 2021 to 2023 using Power BI, a leading data visualization tool.

The objective of this project is to offer a comprehensive view of player and team performances across these seasons. By analysing data related to batting, bowling, and team statistics, this project seeks to uncover trends, highlight key performers, and evaluate overall team success. Key areas of focus include identifying top run-scorers, best bowlers, and analysing team performances in terms of wins, boundaries, and chasing targets.

Power BI is utilized to create detailed dashboards that visualize data effectively, allowing for easy interpretation and insight generation. The analysis encompasses various aspects, such as top batting and bowling performances, team dynamics, and tournament achievements. By presenting this data through engaging visualizations, the project aims to enhance the understanding of IPL trends and performance metrics, demonstrating the impact of data analytics in evaluating and appreciating cricketing excellence.

This project also delves into how teams strategize and perform under different conditions, such as batting first versus chasing targets. By analysing these scenarios, the project provides insights into which teams excel in specific situations and how individual performances contribute to overall outcomes. The use of Power BI enables the creation of interactive visuals that make it easy to explore these patterns and understand the strategies behind winning matches.

In addition, the project highlights standout players who have consistently delivered strong performances across the 2021 to 2023 IPL seasons. By focusing on metrics like total runs scored, wickets taken, and match-winning contributions, the analysis showcases the players who have had the most significant impact on their teams' success. This approach not only celebrates cricketing talent but also demonstrates how data-driven insights can deepen our appreciation of the game.

Chapter 2: Review of Literature

The review of literature explores various aspects of cricket analytics and data visualization, focusing on studies and methodologies relevant to analysing sports data, particularly in the context of the Indian Premier League (IPL).

1. Sports Analytics and Data Visualization

Sports analytics has gained prominence as teams and analysts increasingly rely on data-driven insights to enhance performance and strategy. It involves collecting and analysing performance data to make informed decisions and improve outcomes. Visualization tools, such as Power BI, play a crucial role in presenting this data in a comprehensible format. Effective visualization helps in understanding complex data sets and deriving actionable insights, making it an essential component of modern sports analytics.

2. Cricket Performance Analysis

Cricket performance analysis has evolved significantly, with numerous studies focusing on various performance metrics. For example, batting and bowling statistics can be analysed to evaluate player performance. Metrics such as strike rate, economy rate, and boundary percentage are commonly used to assess players' effectiveness. This approach aligns with the methods used in this project, where key performance indicators for batting and bowling are analysed to identify top performers.

3. IPL-Specific Studies

The IPL, as a major T20 tournament, has been the subject of various studies analysing team performance, focusing on win-loss ratios and match outcomes. Understanding team dynamics and individual performances is crucial in determining success. This project builds on such analyses by providing a detailed examination of team performances, including match results, boundary statistics, and chasing performance.

4. Data Analytics in Sports

The application of data analytics in sports is well-documented, with research emphasizing its impact on performance assessment and strategy development. Advanced analytics and visualization techniques can provide deeper insights into player and team performances. The use of tools like Power BI for creating interactive dashboards and reports is a key aspect of this approach, supporting more informed decision-making in sports analysis.

5. Visualization Techniques

Visualization techniques are essential for effective data presentation. Interactive dashboards and charts facilitate better understanding and interpretation of data. This project employs Power BI to create various visualizations, including batting and bowling analyses, team performance dashboards, and tournament achievements, ensuring that insights are presented clearly and effectively.

6. Player Performance Metrics

Player performance metrics are essential for understanding how individual players contribute to their teams. Metrics like batting average, strike rate, and economy rate help evaluate a player's effectiveness in different situations. This project uses these metrics to analyse how players perform and how their contributions impact their teams' success in the IPL.

7. Season Comparison

Comparing different IPL seasons can reveal important trends and changes. By looking at data from the 2021-2023 seasons, this project highlights shifts in team strategies, player performances, and overall tournament dynamics. This approach helps us understand how teams and players evolve over time and adapt to new challenges.

This review underscores the importance of data analytics and visualization in sports, particularly in cricket. By building on existing research and methodologies, this project aims to provide a comprehensive analysis of IPL data, offering valuable insights into player and team performances over the 2021-2023 seasons.

Chapter 3: Implementation of Project

The implementation of the project "**Game Insight: Analysing Tournaments (Cricket) with Power BI**" involves several stages, from data collection and preparation to the creation of insightful dashboards using Power BI.

3.1 System Requirements

To successfully implement this project, the following system requirements are essential:

Hardware:

- **Processor:** Intel Core i3 or higher
- **RAM:** 8 GB or higher
- **Storage:** 256 GB SSD or higher
- **Graphics Card:** Dedicated GPU (optional but recommended for better performance)

Software:

- **Operating System:** Windows 10 or later, macOS Mojave or later
- **Power BI Desktop:** Latest version installed
- **Microsoft Excel:** For data preprocessing
- **Database Software:** SQL Server or MySQL (optional)
- **Programming Language:** Python (Pandas and NumPy for advanced data handling)
- **Internet Connection:** Required for downloading data and updates

3.2 Data Collection

The first step in the implementation was to gather relevant data from reliable sources. The primary data sources included the official IPL website, ESPN Cricinfo, Kaggle and other cricket statistics platforms. The data collected covered match summaries, player details, batting and bowling performances, and team statistics for the IPL seasons from 2021 to 2023.

3.3 Data Preparation

Data preparation involved cleaning, transforming, and integrating the collected data to ensure it was suitable for analysis. This stage included the following steps:

- **Data Cleaning:** Removing duplicates, correcting errors, and handling missing values to ensure data quality and consistency.

- **Data Transformation**: Converting data into a suitable format for analysis. This included standardizing date formats, normalizing player names, and categorizing data into relevant fields.

- **Data Integration:** Merging data from various sources into a unified dataset. This involved creating relational tables and linking them through common identifiers such as match IDs and player names.

3.4 Database Schema Design

A relational database schema was designed to organize the data effectively. The key tables created were:

- **dim_match_summary:** Contains match-level data, including team names, winner, margin, match date, match ID, year, wickets, and whether the match was won by chasing.

- **dim_players:** Stores player information such as team, batting style, bowling style, playing role, custom batting order, and name.

- **fact_batting_summary**: Details individual batting performances, including match ID, match, team, batting position, batsman name, runs, balls faced, fours, sixes, strike rate, and boundary runs.

- **fact_bowling_summary**: Details individual bowling performances, including match ID, match, team, bowler name, overs bowled, maiden overs, runs conceded, wickets taken, economy rate, dot balls, fours, sixes, wides, no-balls, and balls bowled.

- Top 3 teams: Summarizes the top three teams each year, listing the team, winner, and runner-up.

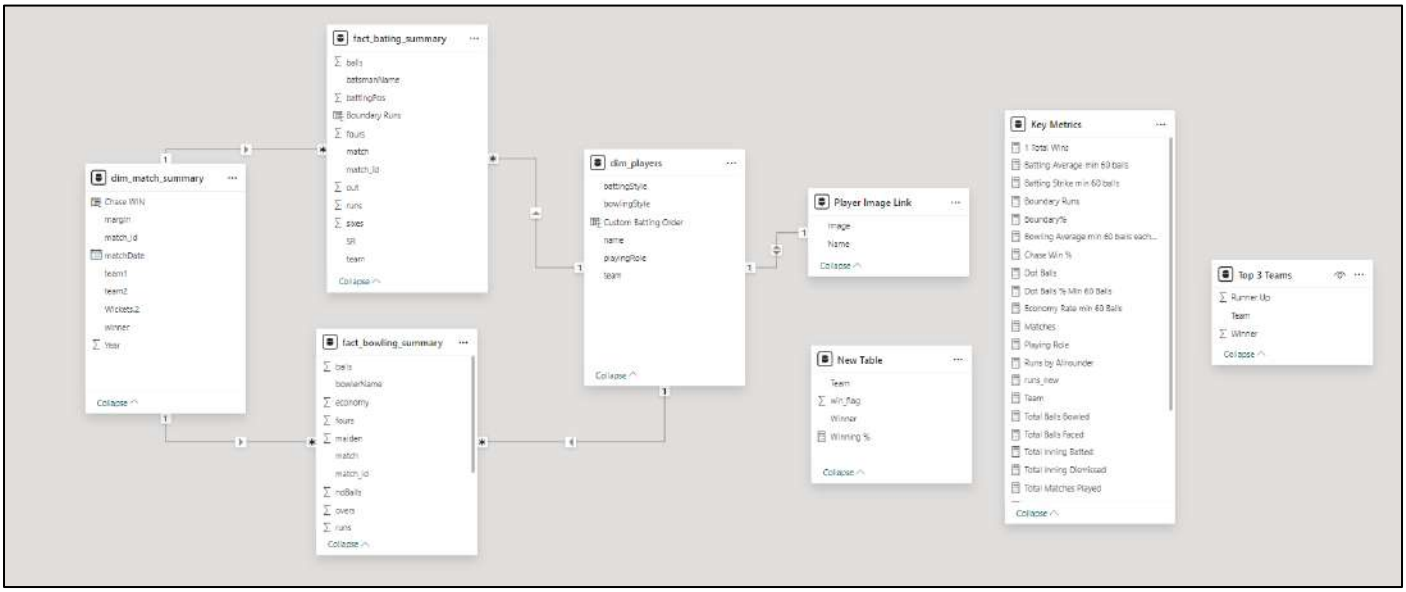


Figure 3.1

3.5 Creating Dashboards in Power BI

Power BI was used to create interactive and visually appealing dashboards to present the data insights. The following dashboards were developed:

1. IPL Game Insights 2021-23: Provides an overview of the key highlights and statistics from the IPL seasons.

2. Batting Analysis (2021, 2022, 2023):

- Top 10 Batsmen with Highest Runs: Displays the top run-scorers each season.

- Top 10 Batsmen with Highest Strike Rate: Highlights batsmen with the best strike rates.

- Top 10 Batsmen with Maximum Sixes: Lists the batsmen who hit the most sixes.

- Top 10 Batsmen with Highest Boundary Percentage: Shows batsmen with the highest percentage of boundary runs.

Criteria: Batsmen must have played a minimum of 60 balls each season.

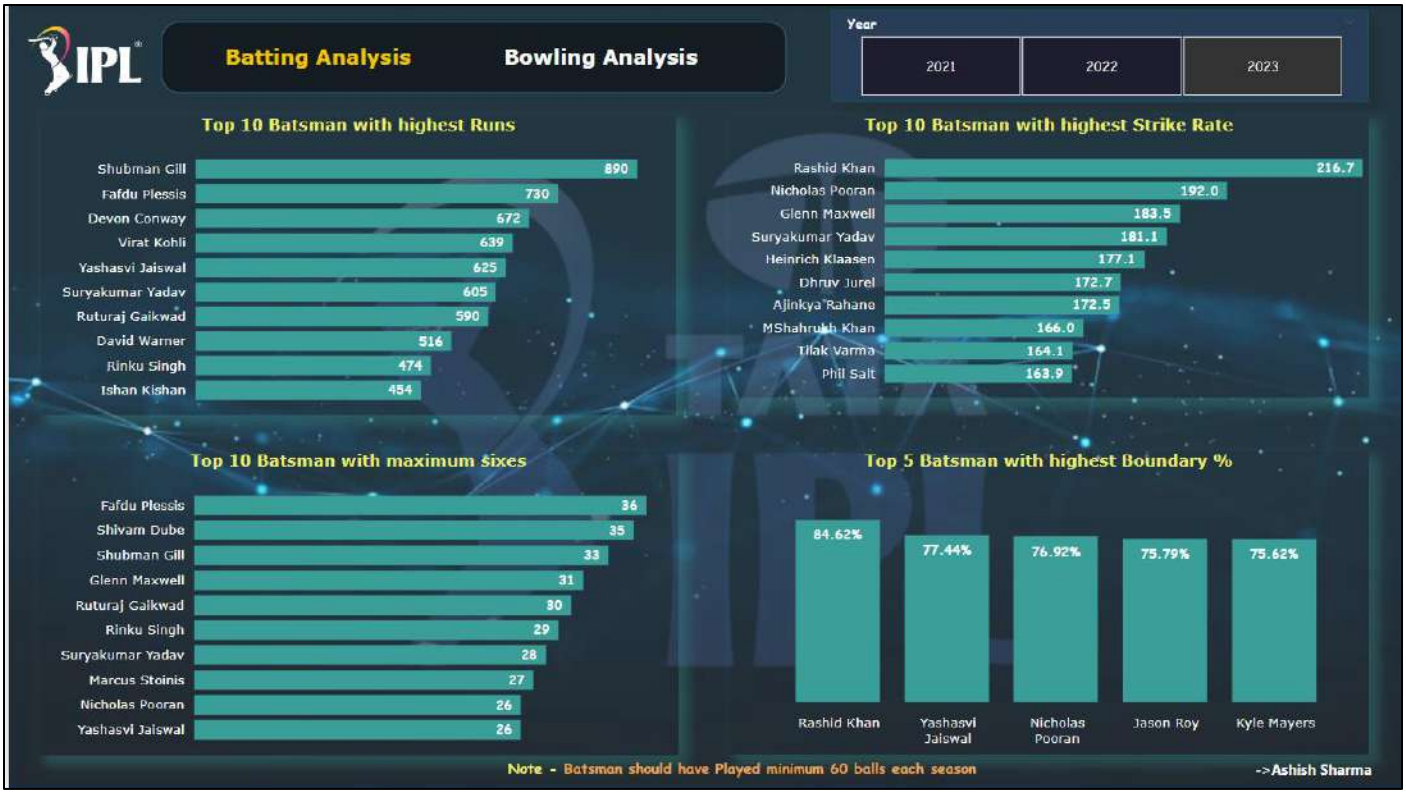


Figure 3.2

3. Bowling Analysis (2021, 2022, 2023):

- **Top 10 Bowlers with Highest Wickets**: Displays the top wicket-takers each season.

- **Top 10 Bowlers with Best Average**: Highlights bowlers with the best bowling averages.

- **Top 10 Bowlers with Best Economy Rate**: Shows bowlers with the best economy rates.

- **Highest Maiden Overs**: Lists bowlers who bowled the most maiden overs.

- **Highest Dot Ball Percentage**: Highlights bowlers with the highest percentage of dot balls.

Criteria: Bowlers must have bowled a minimum of 60 balls each season.

Figure 3.3

4. Team Analysis:

- Total Matches, Wins, and Winning Percentage: Analyzes team performance based on total matches played, wins, and win percentage.

- Total Sixes and Fours by Teams: Summarizes the total boundaries hit by each team.

- Matches, Wins, and Win Percentage While Chasing: Examines team performance while chasing targets.

- Total Season Wins by Teams: Lists the total number of seasons wins by each team.

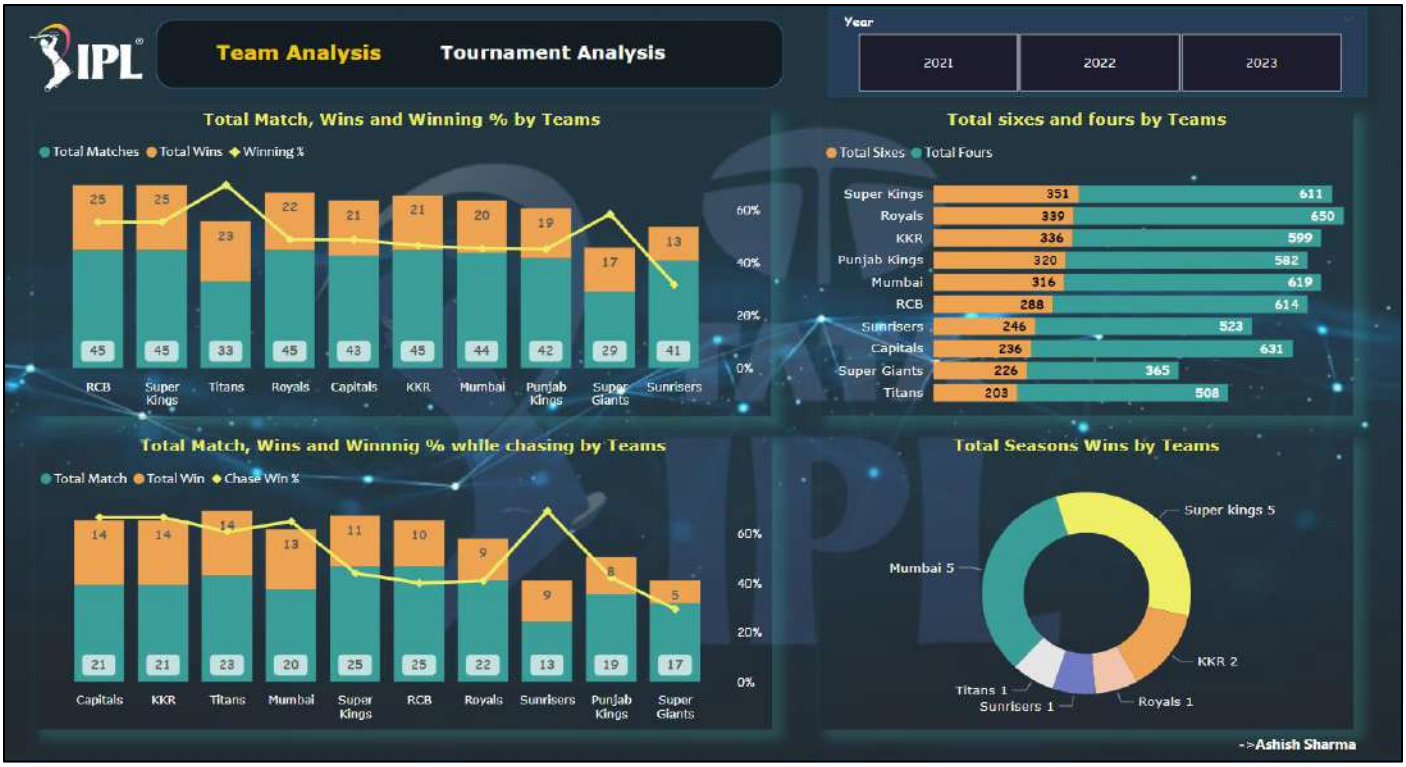


Figure 3.4

5. Tournament Analysis:

- **Title Winners**: Lists the title-winning teams for each year.

- **Orange Cap Winners**: Highlights the top run-scorers of each season.

- **Purple Cap Winners**: Shows the top wicket-takers of each season.

- **Tournament Sixes and Fours**: Summarizes the total number of sixes and fours hit each season.

Figure 3.5

6. Best 4 All-Rounders (2021, 2022, 2023): Identifies the best all-rounders based on performance in both batting and bowling.

Criteria: All-rounders must have played a minimum of 60 balls and bowled a minimum of 60 balls each season.

Figure 3.6

7. Best Playing XI (2021, 2022, 2023): Selects the best playing XI based on overall performance.

Criteria: Batsmen must have played a minimum of 60 balls, and bowlers must have bowled a minimum of 60 balls each season.

Figure 3.7

3.6 Visualization Techniques

Power BI's visualization capabilities were leveraged to create interactive and dynamic reports. Various charts, graphs, and tables were used to represent data insights effectively. These visualizations included bar charts for top performers, pie charts for team performance distribution, and line graphs for trend analysis over the seasons.

3.7 Insights and Interpretation

The final step involved interpreting the visualized data to derive meaningful insights. Key findings included identifying consistent top performers like Faf du Plessis and Virat Kohli, highlighting emerging stars like Shubman Gill, and understanding team dynamics and strategies that led to victories.

Figure 3.8

3.8 Conclusion

The implementation of this project demonstrated the power of data analytics and visualization in sports. By systematically collecting, preparing, and analysing IPL data, valuable insights were generated that enhance the understanding of player and team performances. Power BI proved to be an effective tool for presenting complex data in an accessible and engaging format, making it easier to appreciate the nuances of cricketing excellence in the IPL.

Chapter 4: Results and Discussions

4.1 Batting Analysis

The batting analysis focuses on identifying top performers based on various metrics such as runs scored, strike rate, sixes hit, and boundary percentage.

4.1.1 Top 10 Batsmen with Highest Runs

The analysis of the top run-scorers revealed consistent performance from key players:

- 2021: Ruturaj Gaikwad emerged as the leading run-scorer, followed by Faf du Plessis and KL Rahul.

- 2022: Jos Buttler topped the charts, with KL Rahul and Quinton de Kock continuing their strong performances.

- 2023: Shubman Gill made a significant impact, leading the run-scoring charts, with Faf du Plessis and Devon Conway following closely.

4.1.2 Top 10 Batsmen with Highest Strike Rate

The strike rate analysis highlighted aggressive batting performances:

- 2021: Shimron Hetmyer, Prithvi Shaw, and Evin Lewis were the top performers.

- 2022: Tim David, Dinesh Kartik, and Liam Livingstone showcased explosive batting.

- 2023: Rashid Khan, Nicholas Pooran, and Glenn Maxwell dominated with high strike rates.

4.1.3 Top 10 Batsmen with Maximum Sixes

The six-hitting ability was a key metric for evaluating power hitters:

- 2021: Faf du Plessis, Ruturaj Gaikwad, and KL Rahul led the charts.

- 2022: Jos Buttler, Liam Livingstone, and Andre Russel continued their dominance.

- 2023: Faf du Plessis, Shivam Dube, and Shubhman Gill were the top six-hitters.

4.1.4 Top 10 Batsmen with Highest Boundary Percentage

Boundary percentage indicated the effectiveness of batsmen in scoring through boundaries:

- 2021: Andre Russel, Yashasvi Jaiswal, and Evin Lewis excelled in this metric.

- 2022: Andre Russel, Tim David, and Dewald Brevis showcased their boundary-hitting prowess.

- 2023: Rashid Khan, Yashasvi Jaiswal, and Nicholas Pooran led in boundary percentage.

4.2 Bowling Analysis

The bowling analysis focused on identifying top performers based on wickets taken, average, economy rate, maiden overs, and dot ball percentage.

4.2.1 Top 10 Bowlers with Highest Wickets

The wicket-taking ability was crucial in evaluating bowler performance:

- 2021: Harshal Patel, Avesh Khan, and Jasprit Bumrah were the top wicket-takers.

- 2022: Yuzvendra Chahal, Wanindu Hasaranga and Kagiso Rabada led the charts.

- 2023: Mohammed Shami, Mohit Sharma, and Rashid Khan dominated with the highest wickets.

4.2.2 Top 10 Bowlers with Best Average

The average was a key metric for evaluating bowler efficiency:

- 2021: Moises henriques, Shahbaz Ahmed, and Harshal Patel excelled.

- 2022: Mohsin Khan, Andre Russel, and Wanindu Hasaranga showcased their efficiency.

- 2023: Mark Wood, Mohit Sharma, and Mitchell Marsh maintained top averages.

4.2.3 Top 10 Bowlers with Best Economy Rate

Economy rate highlighted bowlers' ability to control runs:

- 2021: Moises Henriques, Harpreet Brar and Anrich Nortje had the best economy rates.

- 2022: Sunil Narine, Mohsin Khan and David Willey excelled in controlling runs.

- 2023: Mitchell Santner, David Willey, and Axar Patel continued to dominate economy rates.

4.2.4 Highest Maiden Overs

Maiden overs indicated bowlers' ability to apply pressure:

- 2021: Ishant Sharma, Anrich Nortje, and Bhuvaneshwar Kumar led in maiden overs.

- 2022: Prasidh Krishna, Trent Boult and Bhuvaneshwar Kumar maintained their performance.

- 2023: Trent Boult, Khaleel Ahmed, and Mohammed Shami continued to excel.

4.2.5 Highest Dot Ball Percentage

Dot ball percentage highlighted bowlers' ability to restrict scoring:

- 2021: Moises Henriques, Chris Woakes, and Shivam Mavi led in dot ball percentage.

- 2022: Mohsin Khan, Simarjeet Singh, and Prasidh Krishna excelled.

- 2023: Mohammed Siraj, Mark Wood, and Mohammed Shami maintained top dot ball percentages.

4.3 Team Analysis

The team analysis provided insights into overall team performance, including matches played, wins, winning percentage, and boundary statistics.

4.3.1 Total Matches, Wins, and Winning Percentage

- 2021: Chennai Super Kings (CSK) dominated with the highest win percentage, followed by Delhi Capitals (DC) and Royal Challengers Bangalore (RCB).

- 2022: Gujarat Titans (GT) emerged as champions in their debut season, with high winning percentages, followed by CSK and DC.

- 2023: CSK regained the title, showcasing consistent performance, followed by DC and RCB.

4.3.2 Total Sixes and Fours by Teams

- 2021: CSK, KKR, and Punjab Kings led in hitting boundaries.

- 2022: Royals, Super Giants, and KKR continued the trend of boundary hitting.

- 2023: MI, CSK and KKR remained top in boundary statistics.

4.3.3 Matches, Wins, and Winning Percentage While Chasing

- 2021: CSK excelled in chasing, followed by DC and KKR.

- 2022: GT demonstrated strong chasing abilities, along with CSK and DC.

- 2023: CSK continued to lead in chasing performances, followed by Titans and Mumbai.

4.3.4 Total Season Wins by Teams

- 2021: CSK secured the highest number of wins.

- 2022: GT achieved the most wins in their debut season.

- 2023: CSK once again topped the list with the most wins.

4.4 Tournament Analysis

The tournament analysis highlighted key achievements, such as title winners, Orange Cap winners, Purple Cap winners, and boundary statistics.

4.4.1 Title Winners

- 2021: Chennai Super Kings clinched the title.

- 2022: Gujarat Titans emerged as champions.

- 2023: Chennai Super Kings triumphed again.

4.4.2 Orange Cap Winners

- 2021: Ruturaj Gaikwad emerged as the top run-scorer.

- 2022: Jos Buttler secured the Orange Cap.

- 2023: Shubman Gill claimed the Orange Cap.

4.4.3 Purple Cap Winners

- 2021: Harshal Patel led the wicket-taking charts.

- 2022: Yuzvendra Chahal secured the Purple Cap.

- 2023: Mohammed Shami took the top spot.

4.4.4 Tournament Sixes and Fours

- 2021: Faf du Plessis and KL Rahul led in hitting sixes and fours.

- 2022: Jos Buttler dominated boundary statistics.

- 2023: Faf du Plesis excelled again in boundary hitting.

Tournament sixes hit

- 2021: 679 Sixes.

- 2022: 1062 Sixes.

- 2023: 1120 Sixes.

Tournament Fours

- 2021: 1517 Fours.

- 2022: 2017 Fours.

- 2023: 2168 Fours.

4.5 Best All-Rounders

The analysis of all-rounders highlighted players who excelled in both batting and bowling:

- 2021: Ravindra Jadeja, Andre Russell, and Glenn Maxwell stood out.

- 2022: Andre Russell, Ravindra Jadeja, and Moeen Ali maintained strong performances.

- 2023: Andre Russell, Ravindra Jadeja, and Moeen Ali continued to excel.

4.6 Best Playing XI

The best playing XI was selected based on overall performance metrics:

- 2021: Key players included Faf du Plessis, Ruturaj Gaikwad, KL Rahul, Harshal Patel, and Jasprit Bumrah.

- 2022: Jos Buttler, KL Rahul, Faf du Plessis, Yuzvendra Chahal, and Kagiso Rabada were among the top selections.

- 2023: Shubman Gill, Faf du Plessis, David Warner, Yuzvendra Chahal, and Trent Boult were standout performers.

Final Chapter: Conclusion and Future Scope

5.1 Conclusion

The project "Game Insight: Analysing Tournaments (Cricket) with Power BI" successfully achieved its objective of providing an in-depth analysis of the IPL seasons from 2021 to 2023. Through meticulous data collection, preparation, and visualization, key insights were derived about player and team performances, offering a comprehensive understanding of the game.

Key Findings:

1. Consistent Performers: Players like Faf du Plessis, Virat Kohli, and Shubman Gill emerged as consistent top performers in batting, while Yuzvendra Chahal, Kagiso Rabada, and Jasprit Bumrah dominated in bowling.

2. Emerging Stars: Shubman Gill's rise as a significant player, especially in the 2023 season, highlights the emergence of new talent in the IPL.

3. Team Dynamics: Teams like Chennai Super Kings (CSK) and Gujarat Titans (GT) showcased strong performances, with CSK winning titles in 2021 and 2023 and GT emerging as champions in their debut season in 2022.

4. All-Rounder Impact: All-rounders like Andre Russell, Ravindra Jadeja, and Hardik Pandya played crucial roles, excelling in both batting and bowling departments.

5. Visual Insights: The use of Power BI enabled the creation of interactive dashboards that presented complex data in an easily interpretable format, facilitating better understanding and decision-making.

The project demonstrated the power of data analytics in sports, particularly in cricket, by providing valuable insights that can enhance team strategies, player development, and fan engagement. The visualizations created using Power BI effectively presented the data, making it accessible to a wider audience, including team analysts, coaches, and cricket enthusiasts.

5.2 Future Scope

While this project achieved its objectives, there are several areas for future work and enhancement that can further enrich the analysis and provide deeper insights into IPL and cricket analytics in general.

1. Expanded Data Sources:

- Incorporating Additional Seasons: Extending the analysis to include more IPL seasons will provide a broader perspective on performance trends and player development.

- Including External Factors: Analysing factors such as weather conditions, pitch reports, and player injuries can provide a more comprehensive understanding of match outcomes and player performances.

2. Advanced Analytics Techniques:

- Predictive Analytics: Implementing machine learning algorithms to predict player performances, match outcomes, and team strategies can offer actionable insights for teams and coaches.

- Sentiment Analysis: Analysing social media and fan sentiment can provide insights into public perception and engagement, aiding in marketing and fan interaction strategies.

3. Enhanced Visualization:

- Real-Time Analytics: Developing real-time dashboards that update with live match data can provide instant insights during matches, helping teams make strategic decisions on the fly.

4. Comparative Analysis:

- Cross-League Comparisons: Comparing IPL data with other T20 leagues such as the Big Bash League (BBL) and Caribbean Premier League (CPL) can offer insights into different playing styles and strategies.

- Historical Comparisons: Analysing historical cricket data alongside modern IPL data can highlight how the game has evolved over the years.

5. User Interaction:

- Interactive Reports: Developing more interactive and customizable reports where users can filter data based on their preferences (e.g., by player, team, season) can enhance user engagement.

- Mobile Integration: Creating mobile-friendly dashboards can increase accessibility, allowing users to access insights on-the-go.

Publication

Published IPL 2021-2023 analysis on the Power BI Server to securely share and collaborate on the insights with others while maintaining centralized control over the reports.

Project Link:

https://app.powerbi.com/view?r=eyJrIjoiMWEwYjNmNzUtODkxNS00Y2NlLTg4NzgtMzZhNTJiZWI4ODFkIiwidCI6ImM2ZTU0OWIzLTVmNDUtNDAzMi1hYWU5LWQ0MjQ0ZGM1YjJjNCJ9

References

Datasets:

https://www.kaggle.com/datasets/manasgarg/ipl ; Dated-05/07/2024

https://www.espncricinfo.com/series/ipl-2021-1249214 ; Dated-05/07/2024

https://www.espncricinfo.com/series/indian-premier-league-2022-1298423 ; Dated-05/07/2024

https://www.espncricinfo.com/series/indian-premier-league-2023-1345038 ; Dated-05/07/2024

For IPL stats and data validation

https://www.iplt20.com/stats/2021 ; Dated-15/07/2024

https://www.iplt20.com/stats/2022 ; Dated-15/07/2024

https://www.iplt20.com/stats/2023; Dated-15/07/2024

For Power BI and Dax

https://learn.microsoft.com/en-us/power-bi/ ; Dated-01/07/2024

https://learn.microsoft.com/en-us/dax/ ; Dated-08/07/2024