REPORT

for

DV ESE: Dashboard Project

on

IPL DASHBOARD

SUBMITTED BY

Group 4

GROUP MEMBERS:

DEEKSHA MANDAL **(22070126029)**
DHWANI BHAVANKAR **(22070126034)**
DHRUVI RANWALA **(22070126035)**

SYMBIOSIS INSTITUTE OF TECHNOLOGY (SIT), PUNE 412115

A CONSTITUENT OF SYMBIOSIS INTERNATIONAL (DEEMED UNIVERSITY)

2023-24

TABLE OF CONTENTS

INTRODUCTION

The Indian Premier League (IPL), also known as the TATA IPL, is a men's Twenty20 cricket league held annually in India since 2007. Organized by the BCCI, the league features ten franchise teams and is the most popular cricket league globally, usually taking place from March to May. In 2014, it ranked sixth in average attendance among all sports leagues and was the first sporting event streamed live on YouTube in 2010. By 2022, its brand value was estimated at ₹90,038 crore (US$11 billion). In 2023, its media rights were sold for US$6.4 billion, valuing each match at $13.4 million.

The IPL is a massive business venture beyond cricket, with revenue generated through broadcast rights, sponsorships, franchise ownership, merchandise sales, ticket sales, and advertising. It boosts local economies and has a strong digital presence, making it a multi-billion-dollar industry. In 2022, the league's brand value was estimated at ₹90,038 crore (US$11 billion). According to the BCCI, the 2015 IPL season contributed ₹1,150 crore (US$140 million) to India's GDP.

In December 2022, the IPL achieved a valuation of US$10.9 billion, becoming a decacorn and registering a 75% growth in dollar terms since 2020, when it was valued at $6.2 billion, according to a report by the consulting firm D and P Advisory.

PROBLEM STATEMENT

How can historical data from the Indian Premier League (IPL) be leveraged to analyze and predict match outcomes based on key technical factors such as toss results, home ground advantage, pitch conditions, and team performance metrics? By identifying patterns and trends in this data, the goal is to enable companies and sponsors to make more informed and strategic investment decisions in IPL teams, optimizing their sponsorship deals, and maximizing returns on investment.

OBJECTIVES

1. **Analyze Historical Performance Trends:** Historical IPL match data analysis helps identify patterns in team and player performances, including win rates, player contributions, and the impact of match factors like toss results and venue conditions.

2. **Predict Match Outcomes:** Historical data can be utilized to predict match outcomes based on factors like team composition, performance, pitch conditions, and home ground advantage, aiding stakeholders in making data-driven predictions.

3. **Identify Key Investment Opportunities**: emphasize the importance of identifying key investment opportunities by analyzing trends to assist companies and sponsors in making strategic decisions in sponsorship and advertising.

4. **Maximize Return on Sponsorship Deals**: Analyze key metrics to optimize sponsorship strategies with IPL teams, enhancing brand visibility and maximizing ROI.

5. **Understand Fan Engagement Factors**: It aims to identify factors that boost fan engagement, such as high-scoring games and close finishes, to enhance marketing strategies and fan experience.

6. **Enhance Data-Driven Decision-Making for Teams**: Provide IPL teams useful information for tactical choices like player selection, preferred pitches, and in-game tactics to help them perform better as a team and engage fans.

DATASET DESCRIPTION

Data Source: The dataset was obtained from Kaggle, shared by the creator, Ashray Kothari. You can access the dataset at [IPL Dataset on Kaggle] (https://www.kaggle.com/datasets/ashraykothari/ipldataset).

The IPL is a professional Twenty20 cricket league in India contested during March or April and May of every year by eight teams representing eight different cities in India. The league was founded by the Board of Control for Cricket in India (BCCI) in 2008.

a. Use of the Dataset:

The dataset consists of two separate CSV files: matches and deliveries.
These files contain the information of each match summary and ball-by-ball details, respectively.
This dataset is primarily geared toward applications in sports analytics, specifically focused on cricket. It contains detailed statistics and match information from the Indian Premier League (IPL), which can help in analyzing player and team performances, predicting match outcomes, and even creating fantasy cricket teams. Beyond sports, the dataset can be valuable in fields like business (sponsorship analysis, fan engagement), media (content and viewership trends), and potentially healthcare (player fitness and injury patterns based on performance data).

b. Types of Tables or Columns:

The dataset consists of various data types, which include:

1. Numerical data: Columns like scores, wickets, overs, strike rates, etc.
2. Categorical data: player names, team names, match venues, and outcome types.
3. Textual data: descriptions of match events, player roles, and team strategies.
4. Time-related data: dates of matches, tournament seasons, and time intervals within games (e.g., over-by-over data).
5. Geographical data: information about match venues, often tied to cities or countries where IPL matches were held.

This combination of data types supports a broad range of analyses, from statistical assessments of player and team performance to time-series forecasting and regional analysis of venues.

c. File Type, Number of Samples:

1. File Type: This dataset is available as a CSV file.
2. Number of Samples: It consists of 21 columns and 756 samples. The number of rows (samples) in the dataset represents each match or player-specific event.

d. Columns in deliveries.csv:

1. match_id: A unique identifier for the specific cricket match.
2. inning: Refers to the current inning of the match (1 or 2).
3. batting_team: The name of the team currently batting.
4. bowling_team: The name of the team currently bowling.
5. over: The over number within the inning, representing a set of six deliveries.
6. ball: The ball number within the over (from 1 to 6).
7. batsman: The name of the batsman facing the current ball.
8. non_striker: The name of the batsman who is not on strike (standing at the other end).
9. bowler: The name of the bowler delivering the ball.
10. is_super_over: Indicates if the ball was part of a super over (1 for yes, 0 for no).
11. wide_runs: The number of runs conceded from a wide ball.
12. bye_runs: Runs scored when the ball passes the batsman without hitting the bat or body, and no dismissal occurs.
13. legbye_runs: Runs scored when the ball deflects off the batsman's body (not the bat).
14. noball_runs: Runs awarded due to an illegal delivery (no-ball).
15. penalty_runs: Extra runs awarded to the batting team due to penalties (e.g., slow over rate).
16. batsman_runs: The number of runs scored by the batsman from that particular ball.
17. extra_runs: Total number of extra runs awarded (sum of wide, bye, legbye, noball, and penalty runs).
18. total_runs: Total number of runs scored on that delivery, including batsman and extra runs.
19. player_dismissed: Name of the player dismissed on the current ball (if any).
20. dismissal_kind: Type of dismissal (e.g., bowled, caught, run out).
21. fielder: Name of the fielder involved in the dismissal (if applicable).

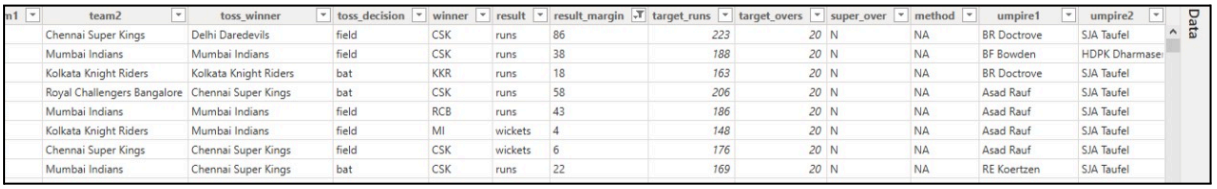
match_id	inning	batting_team	bowling_team	over	ball	batter	bowler	non_striker	batsman_runs	extra_runs	total_runs	extras_type	is_wicket
548312	1	Rajasthan Royals	Kolkata Knight Riders	11	1	BJ Hodge	SP Narine	AL Menaria	1	0	1		
548320	1	Rajasthan Royals	Kolkata Knight Riders	4	1	AM Rahane	SP Narine	R Dravid	1	0	1		
548320	1	Rajasthan Royals	Kolkata Knight Riders	12	1	SP Goswami	SP Narine	OA Shah	1	0	1		
548320	1	Rajasthan Royals	Kolkata Knight Riders	14	1	OA Shah	SP Narine	SP Goswami	1	0	1		
548323	1	Kings XI Punjab	Kolkata Knight Riders	17	1	PP Chawla	SP Narine	Bipul Sharma	1	0	1		
548323	1	Kings XI Punjab	Kolkata Knight Riders	19	1	PP Chawla	SP Narine	Harmeet Singh	1	0	1		

er	ball	batter	bowler	non_striker	batsman_runs	extra_runs	total_runs	extras_type	is_wicket	player_dismissed	dismissal_kind	fielder	Overs
11	1	BJ Hodge	SP Narine	AL Menaria	1	0	1		0	NA	NA	NA	0.166666666666667
4	1	AM Rahane	SP Narine	R Dravid	1	0	1		0	NA	NA	NA	0.166666666666667
12	1	SP Goswami	SP Narine	OA Shah	1	0	1		0	NA	NA	NA	0.166666666666667
14	1	OA Shah	SP Narine	SP Goswami	1	0	1		0	NA	NA	NA	0.166666666666667
17	1	PP Chawla	SP Narine	Bipul Sharma	1	0	1		0	NA	NA	NA	0.166666666666667

e. Columns in matches.csv:

1. id: A unique identifier for the specific match.
2. season: The year in which the match took place.
3. city: The city where the match was held.
4. date: The date on which the match occurred.
5. team1: The name of the first team participating in the match.
6. team2: The name of the second team participating in the match.
7. toss_winner: The team that won the toss before the match.
8. toss_decision: The decision made by the toss-winning team (bat or bowl first).
9. result: The outcome of the match (e.g., normal win, tie, no result).
10. dl_applied: Indicates if the Duckworth-Lewis (D/L) method was applied (1 for yes, 0 for no).
11. winner: The name of the team that won the match.
12. win_by_runs: If the match was won by runs, this column shows the margin (in runs).
13. win_by_wickets: If the match was won by wickets, this column shows the number of wickets remaining.
14. player_of_match: The name of the player awarded the "Player of the Match" for their outstanding performance.
15. venue: The name of the stadium or ground where the match was played.
16. umpire1: The name of the first on-field umpire.
17. umpire2: The name of the second on-field umpire.
18. umpire3: If present, the name of the third (TV) umpire, usually for review decisions.

id	season	city	date	match_type	player_of_match	venue	team1	team2	toss_winner	toss_
548380	2012	Chennai	Friday, May 25, 2012	Qualifier 2	M Vijay	MA Chidambaram Stadium, Chepauk	DD	Chennai Super Kings	Delhi Daredevils	fielc
548379	2012	Bangalore	Wednesday, May 23, 2012	Elimination Final	MS Dhoni	M Chinnaswamy Stadium	CSK	Mumbai Indians	Mumbai Indians	fielc
548378	2012	Pune	Tuesday, May 22, 2012	Qualifier 1	YK Pathan	Subrata Roy Sahara Stadium	DD	Kolkata Knight Riders	Kolkata Knight Riders	bat
501271	2011	Chennai	Saturday, May 28, 2011	Final	M Vijay	MA Chidambaram Stadium, Chepauk	CSK	Royal Challengers Bangalore	Chennai Super Kings	bat
501270	2011	Chennai	Friday, May 27, 2011	Qualifier 2	CH Gayle	MA Chidambaram Stadium, Chepauk	RCB	Mumbai Indians	Mumbai Indians	fielc
501269	2011	Mumbai	Wednesday, May 25, 2011	Elimination Final	MM Patel	Wankhede Stadium	MI	Kolkata Knight Riders	Mumbai Indians	fielc
501268	2011	Mumbai	Tuesday, May 24, 2011	Qualifier 1	SK Raina	Wankhede Stadium	RCB	Chennai Super Kings	Chennai Super Kings	fielc
419165	2010	Mumbai	Sunday, April 25, 2010	Final	SK Raina	Dr DY Patil Sports Academy	CSK	Mumbai Indians	Chennai Super Kings	bat



n1	team2	toss_winner	toss_decision	winner	result	result_margin	target_runs	target_overs	super_over	method	umpire1	umpire2
	Chennai Super Kings	Delhi Daredevils	field	CSK	runs	86	223	20	N	NA	BR Doctrove	SJA Taufel
	Mumbai Indians	Mumbai Indians	field	CSK	runs	38	188	20	N	NA	BF Bowden	HDPK Dharmase
	Kolkata Knight Riders	Kolkata Knight Riders	bat	KKR	runs	18	163	20	N	NA	BR Doctrove	SJA Taufel
	Royal Challengers Bangalore	Chennai Super Kings	bat	CSK	runs	58	206	20	N	NA	Asad Rauf	SJA Taufel
	Mumbai Indians	Mumbai Indians	field	RCB	runs	43	186	20	N	NA	Asad Rauf	SJA Taufel
	Kolkata Knight Riders	Mumbai Indians	field	MI	wickets	4	148	20	N	NA	Asad Rauf	SJA Taufel
	Chennai Super Kings	Chennai Super Kings	field	CSK	wickets	6	176	20	N	NA	Asad Rauf	SJA Taufel
	Mumbai Indians	Chennai Super Kings	bat	CSK	runs	22	169	20	N	NA	RE Koertzen	SJA Taufel

DATA PREPROCESSING

- As a result of the data being real-time, it is extremely clean, meaning that it does not contain any null values. (The columns, although they have NA, the player_dismissed, dismissal_kind, and fielder columns are indicative. Consequently, it remains unchanged.)
- The datasets contain a variety of features: bowler details, batsmen details, per over details, wicketkeeper, etc.
- The dataset, which spans the years 2008–2019, provides insight into the general trend and pattern of each team's play and victories throughout that time. As a result, there isn't much data preprocessing.
- However, there are a lot of data insights that must be comprehended from the substantial amount of data.
- To address the questions a sponsor might have before funding any team, we will use DAX queries.

DATA ANALYSIS

```
1 Total Matches Played =
2 CALCULATE(
3     COUNT(matches[id]),
4     ALLEXCEPT(matches, matches[team1], matches[team2])
5 )
```

This DAX Query is used to give us the total number of matches played. When a slicer is used, it gives us the total matches played by a particular team.

```
1 Distinct Cities = DISTINCTCOUNT(matches[city])
2
```

This Dax Query is used to find out the total number of cities wherein the matches are held.

```
1 Distinct Seasons = DISTINCTCOUNT(matches[season])
```

This DAX Query is used to calculate the total number of seasons conducted between 2008-2019.

```
1 Total Wickets = SUM(deliveries[is_wicket])
```

This is used to calculate the total number of wickets drawn.

```
Distinct Venue = DISTINCTCOUNT(matches[venue])
```

This DAX query gives us the number of distinct venues wherein the matches are held.

```
Total Runs = SUM(deliveries[batsman_runs])
```

This gives us the number of runs taken.

```
Total Overs = SUM(deliveries[Overs])
```

This gives us the count of the total number of overs.

```
Total Extras = SUM(deliveries[extra_runs])
```

This gives us the number of extras given in the matches.

```
Average Result Margin =
AVERAGEX(
    FILTER(
        matches,
        NOT(ISERROR(VALUE(matches[result_margin])))
    ),
    VALUE(matches[result_margin])
)
```

This gives us the average margin of winning the match.

```
Total Matches Played =
CALCULATE(
    COUNT(matches[id]),
    ALLEXCEPT(matches, matches[team1], matches[team2])
)
```

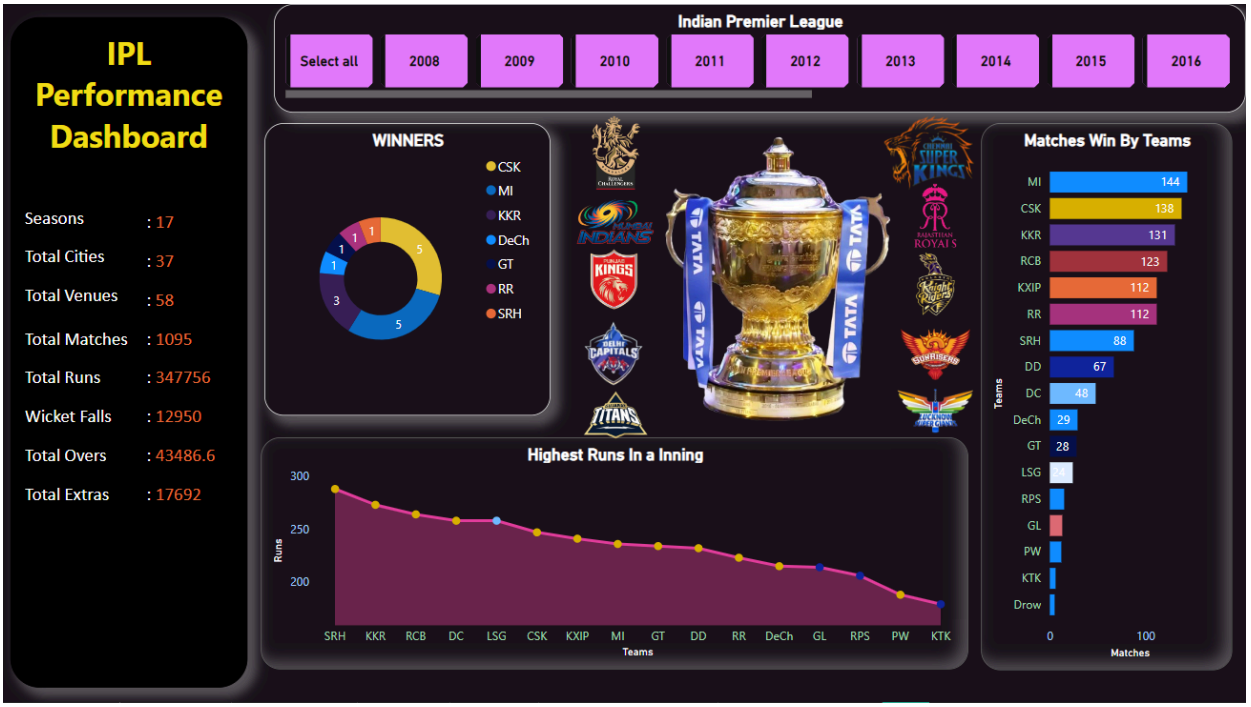
This gives us the count of total matches played (between 2008-2019).

```
Total Wins =
CALCULATE(
    COUNTROWS(matches),
    matches[winner] = SELECTEDVALUE(matches[team1]) || matches[winner] = SELECTEDVALUE(matches[team2]),
    NOT(ISBLANK(matches[winner]))
)
```

This gives us the count of the total number of matches won.

REPORT INSIGHTS

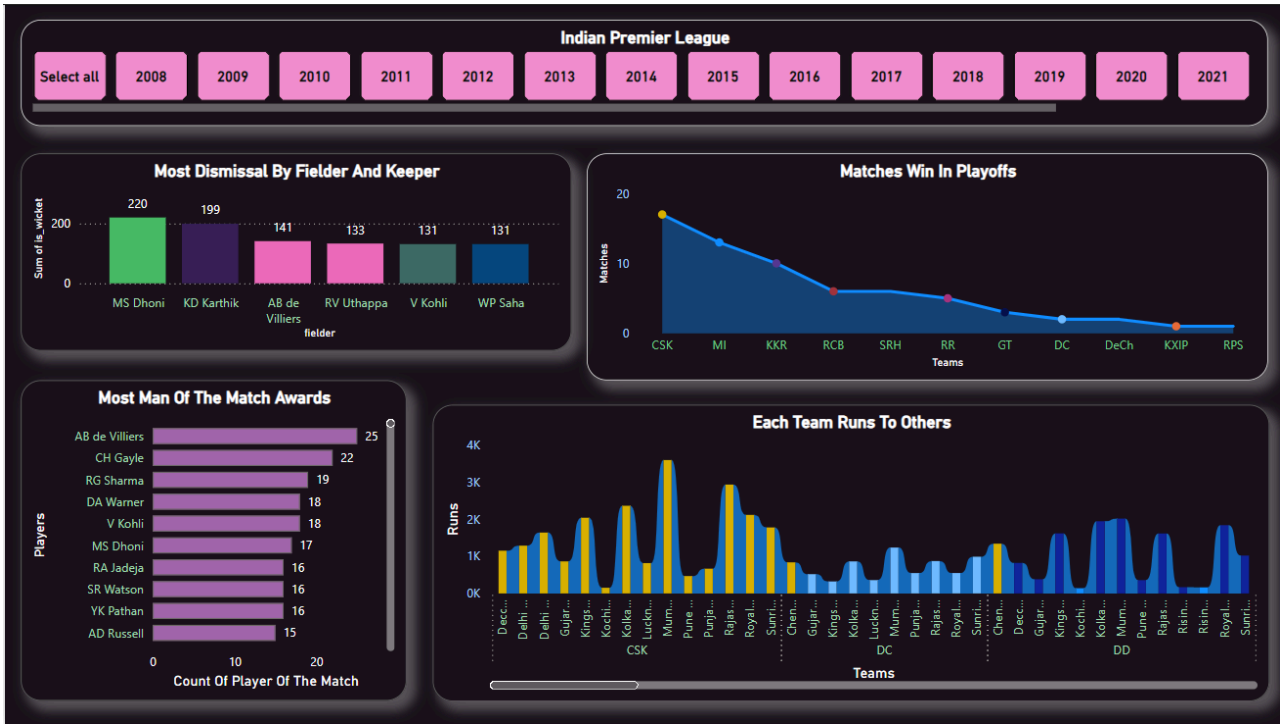
Page: IPL Home



This page is useful in giving us an idea about the overall IPL between 2008-2019. We can see the number of matches played, seasons, wickets, cities, venues, etc. about the IPL.
This page also tells us about the facts of overall IPL over the period of 2008-2019, wherein:

1. CSK and MI have won most of the matches.
2. The highest runs in the innings were made by SRH.

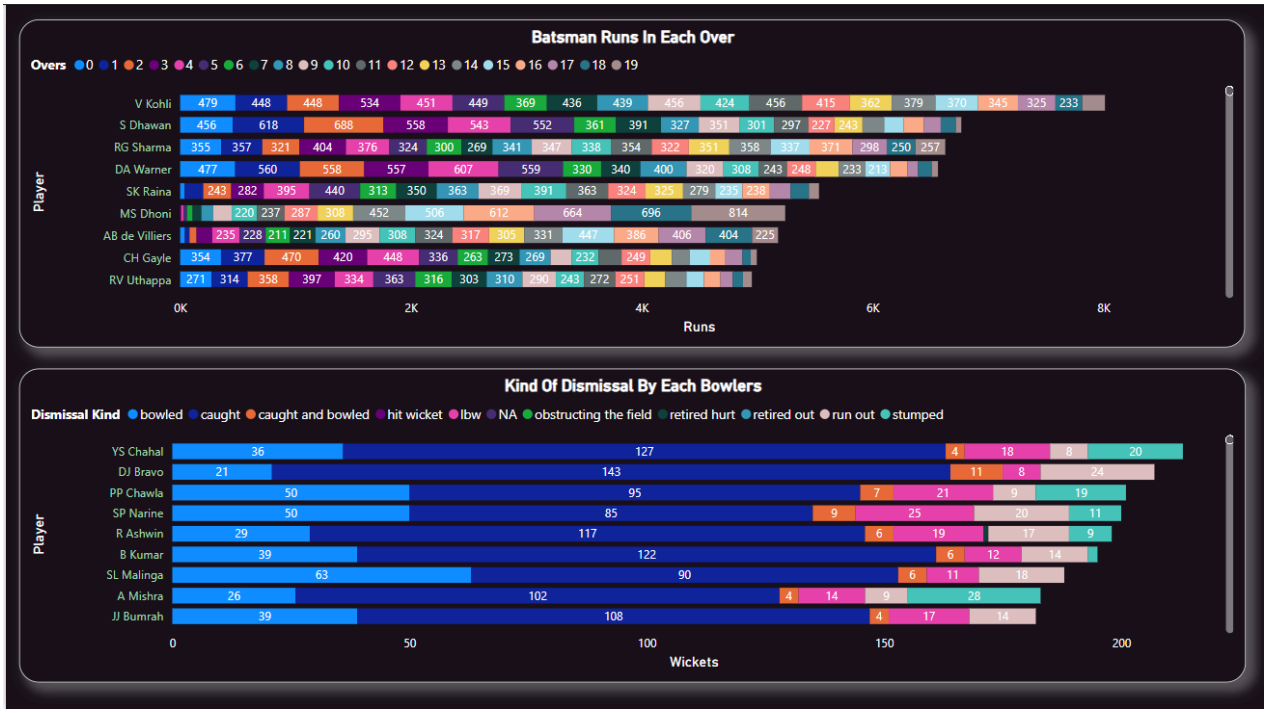
Page: Team-Player



This page gives us an idea about the overall team and players in the team. Some of the insights we can draw are:

1. Most dismissals by fielders and keepers are done by MS Dhoni, followed by KD Kartik.
2. CSK has won maximum matches in playoffs.
3. We can understand the runs made by each team versus the other team, giving us the idea about mindplay also.
4. Most of the Man of the Match awards were given to AB de Villers.

Page: Players



This page gives us an idea about the gameplay of different players. Some of the insights we can draw are:

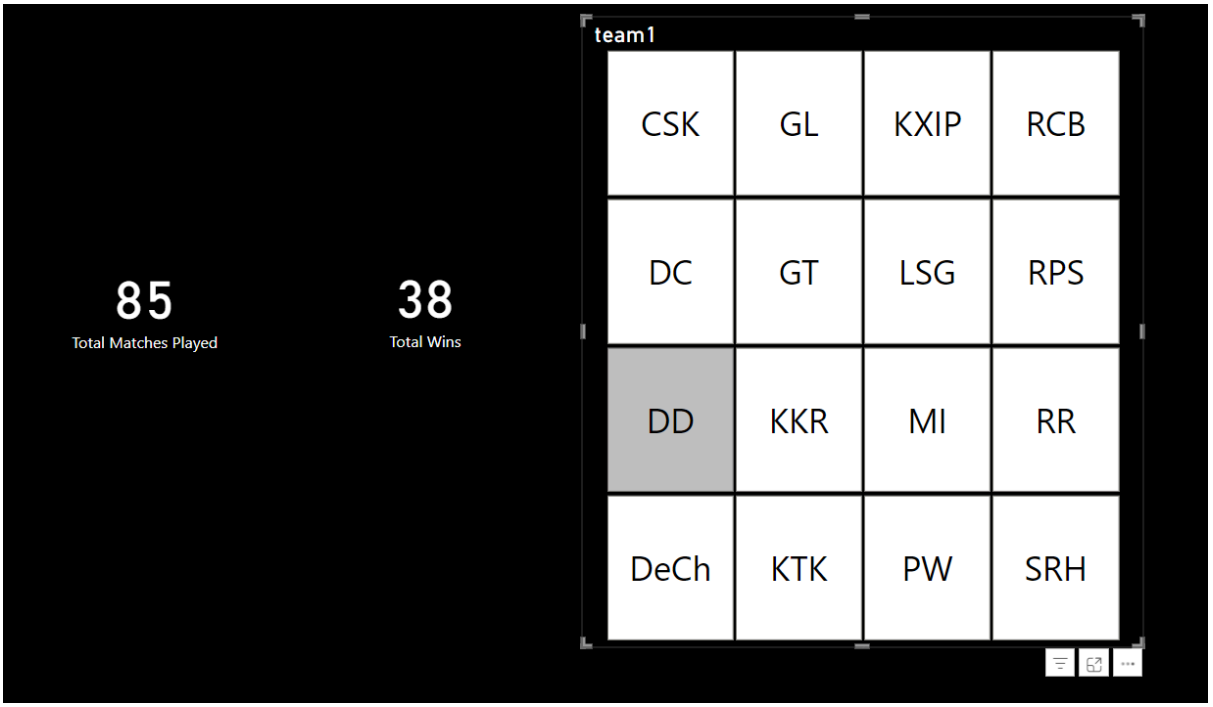
1. Most of the players are performing well while nearing the end of the game, while some players like Virat Kohli, S Dhawan are able to play well throughout the game.
2. Most of the dismissal is given via the catching of the ball.

Page: Match

This page is used to give us an idea about the overall match. Some of the insights that can be drawn from it are:

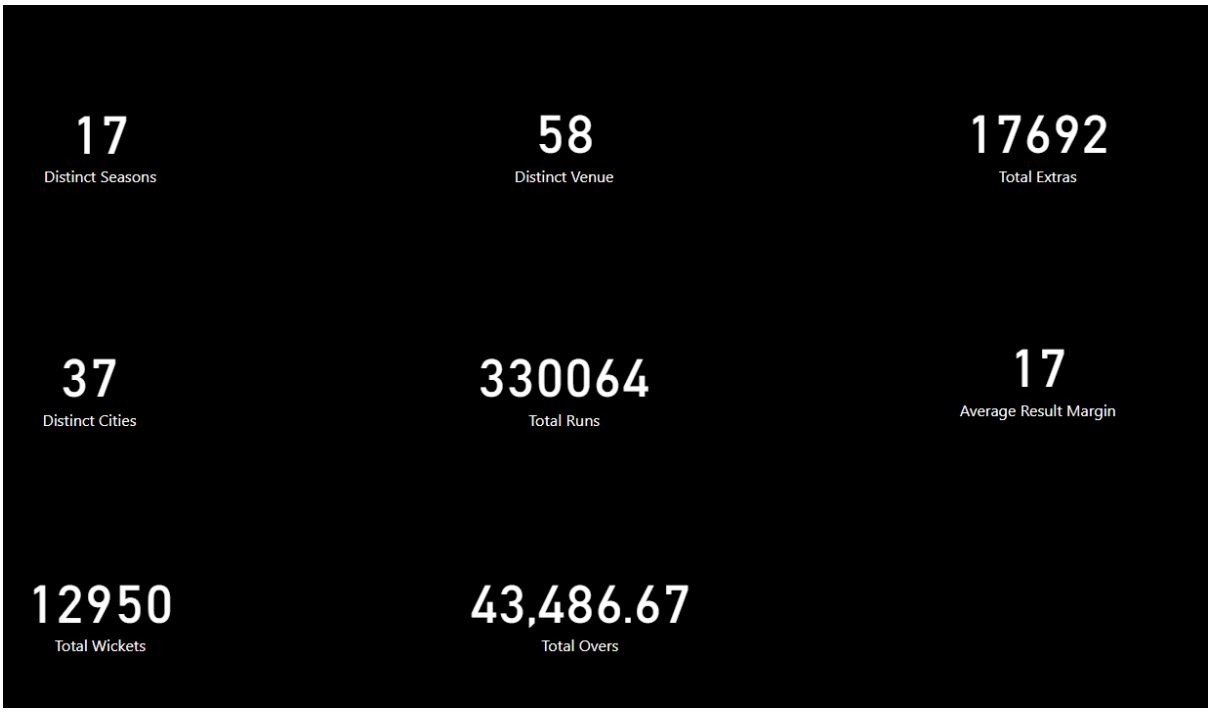
1. Most of the teams are performing almost equally in both innings.
2. MI, KKR, and CSK are outperforming the other teams.
3. In many matches, AK Chowdary has been the umpire.
4. A large number of matches are played in the Eden Garden and Wankhade Stadium.
5. In the entire IPL, there has been a trend of increasing runs till the 5th over, then a drop and then an increase again in the later overs.

Page: Team Understanding



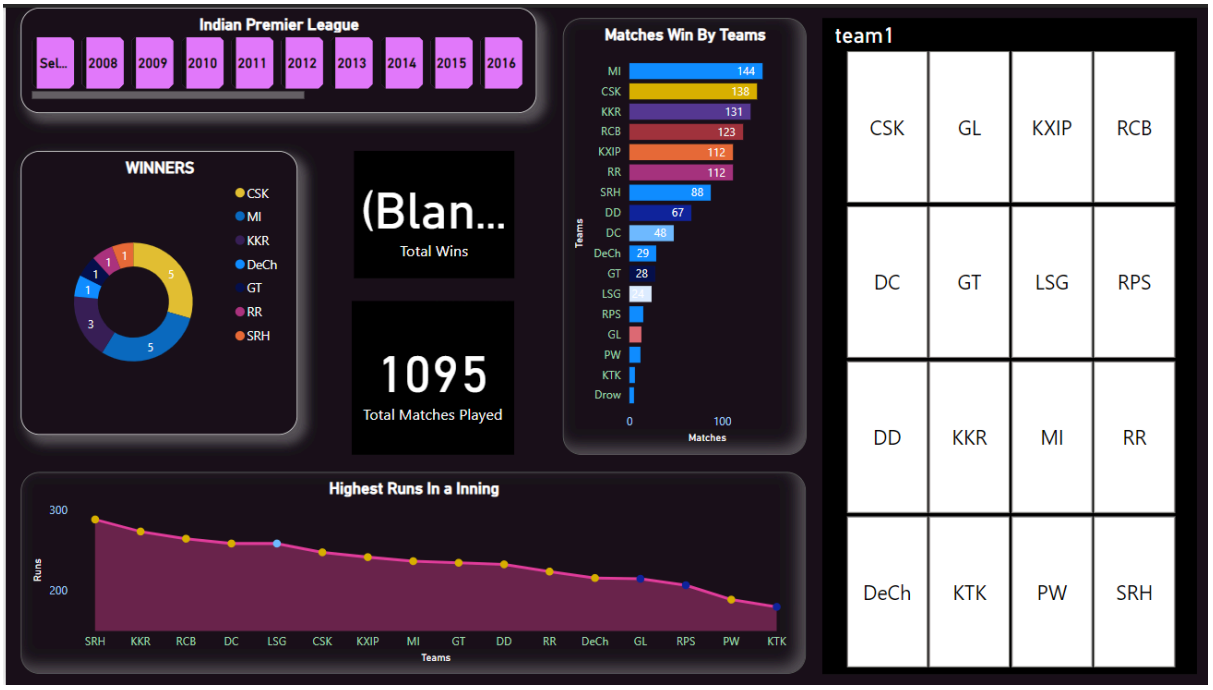
This page is dedicated to the sponsor understanding the team as in, if they are capable of winning, how many they have played, and how many they have won.

Page: IPL Understanding



This page is used for giving the base on the different statistical values that were used for the IPL Introduction in the Home page.

DashBoard



This image shows insights for a particular team

The main objective of our dashboard is to yield sponsors with a comprehensive overview, enabling them to quickly assess which teams are likely to perform well under various circumstances and yield profitable returns. This visually intuitive dashboard highlights key insights about each team, making it easy for sponsors to make informed decisions at a glance.

CONCLUSION

To sum up, this project provides thorough insights into IPL match dynamics, emphasizing aspects like toss results, pitch conditions, and team performance measures that have a big impact on results.

By leveraging historical data, the analysis ensures investors and sponsors make well-informed choices, improving returns on investment and sponsorship plans. The dashboard's predictive capabilities also support teams and analysts in strategic planning, ultimately contributing to the IPL's growth and fan engagement.

REFERENCES

1. https://www.kaggle.com/datasets/ramjidoolla/ipl-data-set
2. https://learn.microsoft.com/en-us/training/powerplatform/power-bi
3. https://www.youtube.com/watch?v=e6QD8lP-m6E
4. https://www.coursera.org/articles/what-is-power-bi
5. https://k21academy.com/microsoft-azure/data-analyst/data-analysis-in-power-bi/
6. https://learn.microsoft.com/en-us/power-bi/create-reports/sample-opportunity-analysis
7. https://learn.microsoft.com/en-us/power-bi/create-reports/sample-competitive-marketing-analysis
8. https://www.reddit.com/r/PowerBI/comments/1af5jd8/question_is_power_bi_for_data_visualisation_or/?rdt=33552
9. https://learn.microsoft.com/en-us/power-bi/consumer/end-user-analyze-visuals
10. https://www.tutorialspoint.com/power_bi/index.htm
11. https://www.geeksforgeeks.org/power-bi-tutorial/