The Untold Stories of IPL – Beyond the Numbers

Sai Nihal Konduti Yamini Srija Koulury Bhaavan Myneni

Abstract

This project offers an exploratory data analysis (EDA) of the Indian Premier League (IPL), aiming to uncover stories that go beyond wins and losses. It analyzes team dominance, player impact, underdog performances, playoff pressure, and fan-driven popularity trends. By combining match data with brand insights and interactive visualizations, the project reveals how leadership, consistency, rivalries, and game-day pressure shape the IPL's competitive and commercial landscape. Our goal is to turn complex statistics into compelling insights, making the IPL's narrative more engaging and data informed.

1. Introduction

The IPL is more than a cricket league—it's a battleground of dominance, drama, and un- der The Indian Premier League (IPL) is more than just a cricket league—it is a stage for dominance, drama, and underdog triumphs. This project seeks to uncover the deeper patterns that shape the IPL's evolving narrative, from high-pressure playoff performances to strategic overhauls and long-standing rivalries. Unlike conventional dashboards that stop at basic statistics like wins and titles, our analysis digs into granular match-level data to explore *why* teams succeed or falter. By examining trends in consistency, clutch performances, player retention, and even brand value, we aim to tell a richer, data-driven story of the IPL—one that reflects its complexity, unpredictability, and cultural impact across seasons.

2. Motivation

The Indian Premier League generates immense fan interest and media coverage, yet most analyses stop at aggregate statistics. We were motivated to understand what drives long-term success, the role of star players, and how pressure affects performance. By highlighting strategic insights and rivalries, this project makes IPL data more engaging and meaningful for analysts and fans alike.

3. Why this study ?

- To identify trends of dominance and underperformance across IPL teams and seasons.
- To analyze the impact of consistent performers and their influence on outcomes.
- To explore how playoff pressure affects team performance compared to league stages
- To uncover rivalry patterns and assess their influence on match outcomes.
- To link on-field performance with off-field popularity and brand growth.

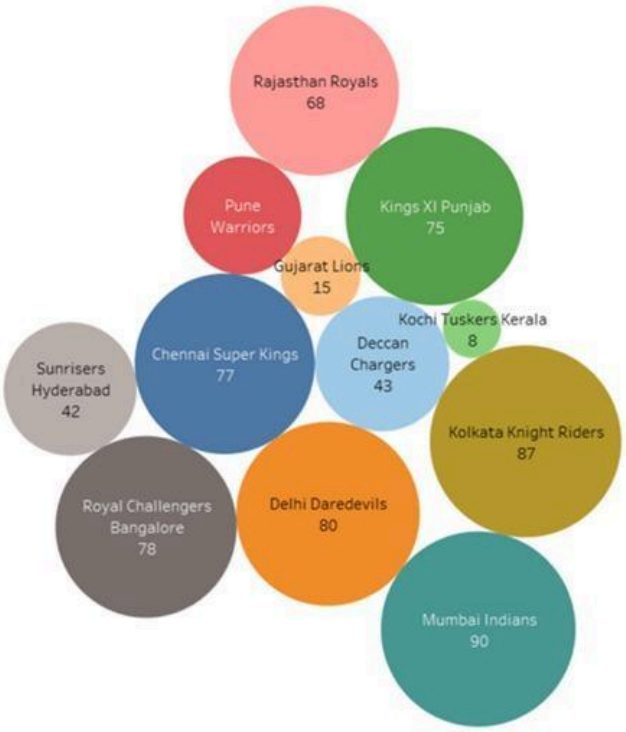
4. Background

Several public dashboards and analytical blogs have visualized IPL data using team-wise win counts, geolocation-based match maps, and title tallies. These visuals, while informative, often focus on isolated statistics without uncovering strategic narratives or contextual insights.

Team Wins Bubble Chart: Bubble size represents team win count, clearly highlighting teams like MI, CSK, and KKR.

Critique:

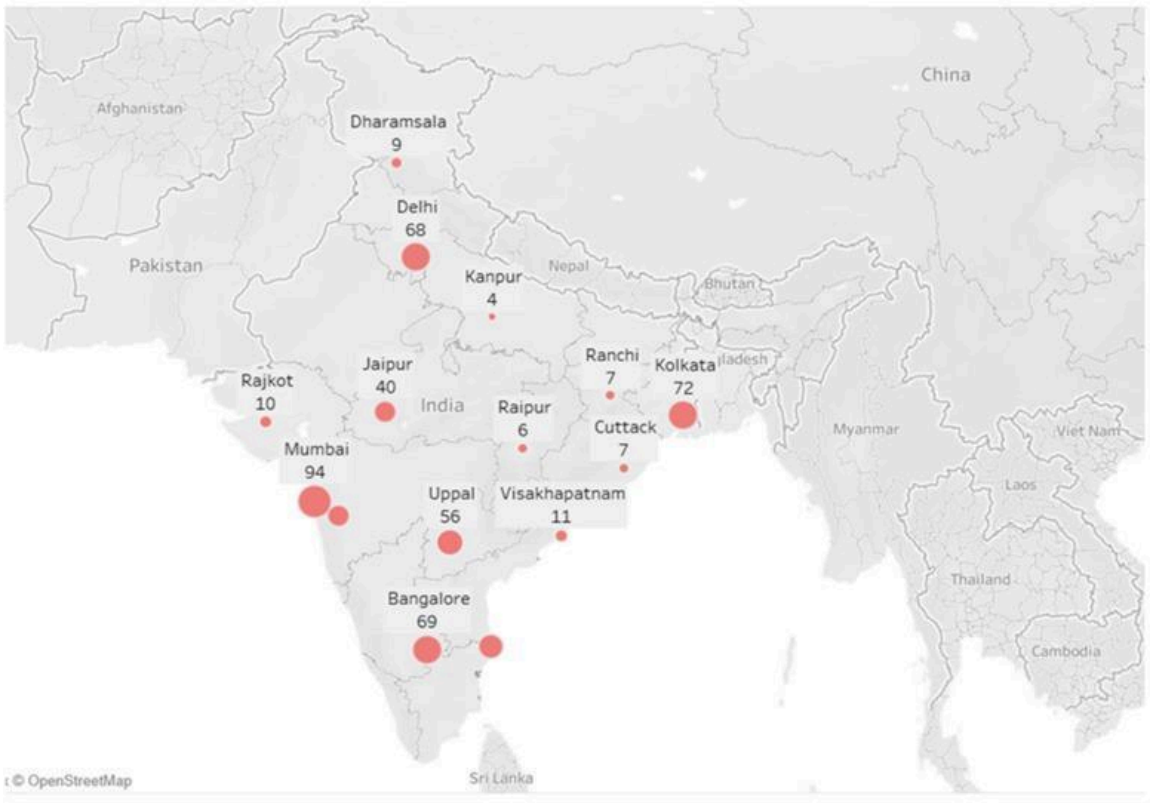
- Good for visual volume comparison.
- Misleading spatial layout—no meaningful proximity or ordering.
- Doesn't distinguish between active/defunct teams or season duration.



Stadium-wise Match Map: Displays city-wise IPL match counts using bubble overlays.

Critique:

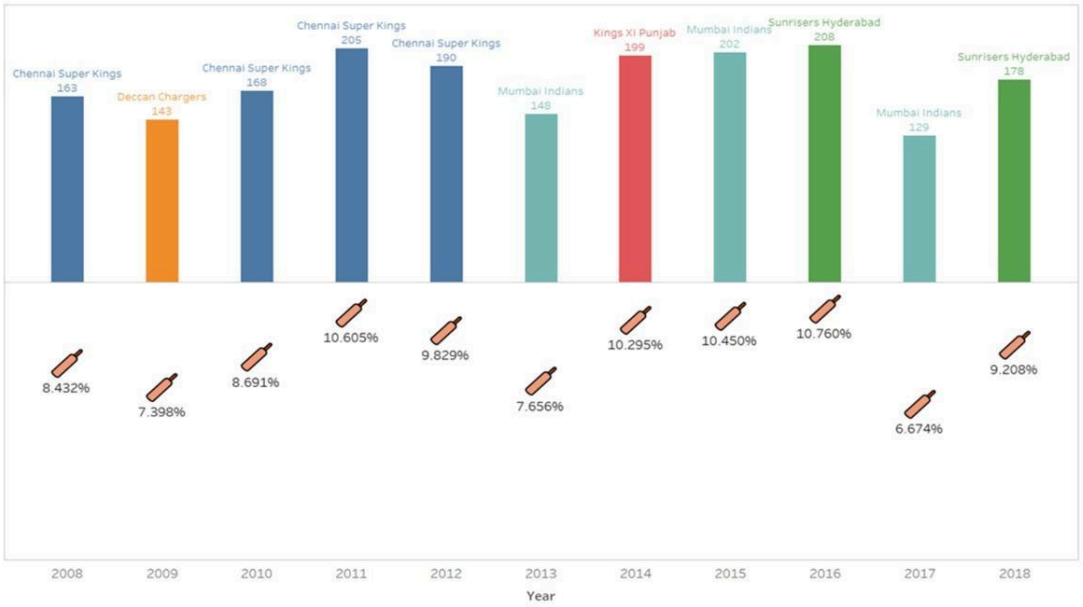
- Useful for regional spread analysis.
- Lacks win/loss context or home-ground advantage insights.
- No separation of regular season vs. playoffs.



Titles Won Horizontal Bar Chart: Uses bars and trophy icons to show title wins by team.

Critique:

- Visually intuitive and easy to grasp.
- Omits seasonal trends or gaps between titles.
- No normalization for seasons played per team.



Summary: Existing visualizations are polished but shallow in insight. Our project aims to reveal the "why" behind trends, not just the "what."

5. Objectives

This project aims to go beyond conventional statistics by uncovering deeper narratives within IPL history. While traditional dashboards focus on wins and titles, our analysis is driven by questions of consistency, pressure, strategy, and fan impact. We explore themes such as the long-term influence of top performers, strategic evolution, team resilience in playoffs, and the role of rivalries and matchups in shaping outcomes. Additionally, we investigate how performance correlates with popularity and identify teams that shine in crunch situations. These themes are explored through the following focused objectives:

- **The kings of runs** – identifying consistent top scorers over seasons.
- **The wicket-taking machines** – evaluating if bowlers have a bigger impact than batters.
- **Run rate evolution** – analyzing how batting strategies have changed over time.
- **The underdog story** – highlighting teams that exceeded expectations.
- **The playoff magic** – finding teams that thrive or crumble under pressure.
- **The cursed matchup** – uncovering teams that consistently struggle against specific opponents.
- **Last-over thrillers** – identifying teams that excel in crunch situations.
- **Popularity vs. performance** – investigating whether winning correlates with fan growth and brand value.

6. Dataset and Methods

For this project, we utilize the Kaggle IPL 2008–2024 Player History dataset, which offers rich, ball-by-ball data from over 15 seasons of the Indian Premier League. The dataset serves as the foundation for analyzing performance patterns, player consistency, team dominance, and brand value evolution. The primary file used for match-level data analysis is matches.csv.

In addition to gameplay data, we incorporate team performance metrics such as win-loss records, head-to-head statistics, and playoff histories to evaluate consistency and resilience. Furthermore, we supplement this data with IPL brand ranking information from Brandirectory, which helps us assess team popularity trends, brand value growth, and fan engagement over time.

The dataset contains a mix of numerical and categorical features, making it well-suited for tasks such as trend analysis, player comparisons, strategic evaluation, and performance forecasting.

Column Name	Description
match_id	Match identifier (linked to matches.csv)
inning	Inning number (1, 2, Super Over = 3/4)
batting_team	Team currently batting
bowling_team	Team currently bowling
over	Over number (1–20)
ball	Ball number in the over (1–6)
batter	Batter facing the delivery
bowler	Bowler delivering the ball
non_striker	Batter at the non-striker's end
batsman_runs	Runs scored off the bat
extra_runs	Runs from extras (e.g., wides, no-balls)
total_runs	Total runs for the delivery (bat + extras)
extras_type	Type of extra (if any)
is_wicket	1 if wicket fell, else 0
player_dismissed	Name of dismissed batter
dismissal_kind	Type of dismissal (e.g., bowled)
fielder	Fielder involved (if applicable)

Table 1: Ball-by-Ball Dataset: Column Descriptions

match_id	inning	batting_team	bowling_team	over	ball	batter	bowler	non_striker	batsman_runs	extra_runs	total_runs	extras_type	is_wicket	player_dismissed	dismissal_kind	fielder
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	1	SC Ganguly	P Kumar	BB McCullum	0	1	1	legbyes	0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	2	BB McCullum	P Kumar	SC Ganguly	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	3	BB McCullum	P Kumar	SC Ganguly	0	1	1	wides	0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	4	BB McCullum	P Kumar	SC Ganguly	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	5	BB McCullum	P Kumar	SC Ganguly	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	6	BB McCullum	P Kumar	SC Ganguly	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	7	BB McCullum	P Kumar	SC Ganguly	0	1	1	legbyes	0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	1	BB McCullum	Z Khan	SC Ganguly	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	2	BB McCullum	Z Khan	SC Ganguly	4	0	4		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	3	BB McCullum	Z Khan	SC Ganguly	4	0	4		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	4	BB McCullum	Z Khan	SC Ganguly	6	0	6		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	5	BB McCullum	Z Khan	SC Ganguly	4	0	4		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	6	BB McCullum	Z Khan	SC Ganguly	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	1	SC Ganguly	P Kumar	BB McCullum	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	2	SC Ganguly	P Kumar	BB McCullum	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	3	SC Ganguly	P Kumar	BB McCullum	0	1	1	legbyes	0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	4	BB McCullum	P Kumar	SC Ganguly	4	0	4		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	5	BB McCullum	P Kumar	SC Ganguly	1	0	1		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	6	SC Ganguly	P Kumar	BB McCullum	0	0	0		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	3	1	BB McCullum	AA Noffke	SC Ganguly	0	5	5	wides	0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	3	2	BB McCullum	AA Noffke	SC Ganguly	6	0	6		0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	3	3	BB McCullum	AA Noffke	SC Ganguly	0	1	1	legbyes	0	NA	NA	
335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	3	4	SC Ganguly	AA Noffke	BB McCullum	4	0	4		0	NA	NA	

Figure 1: Ball by Ball Data

result_margin	target_runs	target_overs	super_over	method	umpire1	umpire2
140.0	223.0	20.0	N		Asad Rauf	RE Koertzen
33.0	241.0	20.0	N		MR Benson	SL Shastri
9.0	130.0	20.0	N		Aleem Dar	GA Pratapkumar
5.0	166.0	20.0	N		SJ Davis	DJ Harper
5.0	111.0	20.0	N		BF Bowden	K Hariharan
6.0	167.0	20.0	N		Aleem Dar	RB Tiffin
9.0	143.0	20.0	N		IL Howell	AM Saheba
6.0	209.0	20.0	N		DJ Harper	GA Pratapkumar
3.0	215.0	20.0	N		Asad Rauf	MR Benson
66.0	183.0	20.0	N		Aleem Dar	AM Saheba
7.0	136.0	20.0	N		MR Benson	IL Howell
9.0	148.0	20.0	N		BF Bowden	AV Jayaprakash
10.0	155.0	20.0	N		Asad Rauf	SL Shastri
4.0	159.0	20.0	N		RE Koertzen	I Shivram
13.0	179.0	20.0	N		BR Doctrove	RB Tiffin
7.0	138.0	20.0	N		BF Bowden	AV Jayaprakash
10.0	192.0	20.0	N		Aleem Dar	I Shivram
7.0	165.0	20.0	N		BR Doctrove	RB Tiffin
45.0	197.0	20.0	N		RE Koertzen	GA Pratapkumar

Figure 3: Matches Dataset

id	season	city	date	match_type	player_of_match	venue	team1	team2	toss_winner	toss_decision	winner	result
335982	2007/08	Bangalore	2008-04-18	League	BB McCullum	M Chinnaswamy Stadium	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	field	Kolkata Knight Riders	runs
335983	2007/08	Chandigarh	2008-04-19	League	MEK Hussey	Punjab Cricket Association Stadium, Mohali	Punjab Kings	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings	runs
335984	2007/08	Delhi	2008-04-19	League	MF Maharoof	Feroz Shah Kotla	Delhi Capitals	Rajasthan Royals	Rajasthan Royals	bat	Delhi Capitals	wickets
335985	2007/08	Mumbai	2008-04-20	League	MV Boucher	Wankhede Stadium	Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	Royal Challengers Bangalore	wickets
335986	2007/08	Kolkata	2008-04-20	League	DJ Hussey	Eden Gardens	Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	bat	Kolkata Knight Riders	wickets
335987	2007/08	Jaipur	2008-04-21	League	SR Watson	Sawai Mansingh Stadium	Rajasthan Royals	Punjab Kings	Punjab Kings	bat	Rajasthan Royals	wickets
335988	2007/08	Hyderabad	2008-04-22	League	V Sehwag	Rajiv Gandhi International Stadium, Uppal	Deccan Chargers	Delhi Capitals	Deccan Chargers	bat	Delhi Capitals	wickets
335989	2007/08	Chennai	2008-04-23	League	ML Hayden	MA Chidambaram Stadium, Chepauk	Chennai Super Kings	Mumbai Indians	Mumbai Indians	field	Chennai Super Kings	runs
335990	2007/08	Hyderabad	2008-04-24	League	YK Pathan	Rajiv Gandhi International Stadium, Uppal	Deccan Chargers	Rajasthan Royals	Rajasthan Royals	field	Rajasthan Royals	wickets
335991	2007/08	Chandigarh	2008-04-25	League	KC Sangakkara	Punjab Cricket Association Stadium, Mohali	Punjab Kings	Mumbai Indians	Mumbai Indians	field	Punjab Kings	runs
335992	2007/08	Bangalore	2008-04-26	League	SR Watson	M Chinnaswamy Stadium	Royal Challengers Bangalore	Rajasthan Royals	Rajasthan Royals	field	Rajasthan Royals	wickets
335993	2007/08	Chennai	2008-04-26	League	JDP Oram	MA Chidambaram Stadium, Chepauk	Chennai Super Kings	Kolkata Knight Riders	Kolkata Knight Riders	bat	Chennai Super Kings	wickets
335994	2007/08	Mumbai	2008-04-27	League	AC Gilchrist	Dr DY Patil Sports Academy	Mumbai Indians	Deccan Chargers	Deccan Chargers	field	Deccan Chargers	wickets
335995	2007/08	Chandigarh	2008-04-27	League	SM Katich	Punjab Cricket Association Stadium, Mohali	Punjab Kings	Delhi Capitals	Delhi Capitals	bat	Punjab Kings	wickets
335996	2007/08	Bangalore	2008-04-28	League	MS Dhoni	M Chinnaswamy Stadium	Royal Challengers Bangalore	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings	runs
335997	2007/08	Kolkata	2008-04-29	League	ST Jayasuriya	Eden Gardens	Kolkata Knight Riders	Mumbai Indians	Kolkata Knight Riders	bat	Mumbai Indians	wickets
335998	2007/08	Delhi	2008-04-30	League	GD McGrath	Feroz Shah Kotla	Delhi Capitals	Royal Challengers Bangalore	Royal Challengers Bangalore	field	Delhi Capitals	runs

Figure 2: Matches Dataset

2020	2019	Name	Country	2020 - USD	2019 - USD	2020	2019
1	3	Mumbai Indians	India	69.56	65.65	AA+	AA
2	1	Chennai Super Kings	India	59.66	75.31	A+	AA
3	2	Kolkota Knight Riders	India	57.7	66.48	A+	AA-
4	4	Sunrisers Hyderabad	India	56.98	59.6	AA-	AA-
5	5	Delhi Capitals	India	51.58	54.21	AA	AA-
6	6	Royal Challengers Bangalore	India	49.3	50.76	AA-	AA-
7	8	Kings XI Punjab	India	37.51	41.7	A+	A+
8	7	Rajasthan Royals	India	35.58	43.41	A	A+

Figure 4: Brand Value Data

Team	BrandValue	Year
Kolkata Knight Riders	58.6	2017
Mumbai Indians	54.1	2017
Sunrisers Hyderabad	46.5	2017
Royal Challengers Bangalore	44.4	2017
Delhi Capitals	40.5	2017
Punjab Kings	36.2	2017
Chennai Super Kings	65.0	2018
Kolkata Knight Riders	62.0	2018
Sunrisers Hyderabad	54.0	2018
Mumbai Indians	53.0	2018
Royal Challengers Bangalore	49.0	2018
Delhi Capitals	44.0	2018
Rajasthan Royals	43.0	2018
Punjab Kings	40.0	2018
Chennai Super Kings	75.31	2019
Kolkata Knight Riders	66.48	2019
Mumbai Indians	65.65	2019

Figure 5: Brand Value Cleaned Data

Column Name	Description
id	Match ID
season	IPL season year
city	City where the match was played
date	Match date
match_type	Match type (League, Eliminator, etc.)
player_of_match	Awarded best player of the match
venue	Stadium name
team1, team2	Participating teams
toss_winner	Team that won the toss
toss_decision	Toss decision: Bat or Field
winner	Match winner
result	Match result type (normal, tie)
result_margin	Margin of win (runs or wickets)
target_runs	Target runs for the second innings
target_overs	Overs allocated for second innings
super_over	Indicates Super Over usage
method	Method used (DLS, etc.)
umpire1, umpire2	On-field umpires

Table 2: Match Metadata from matches.csv

Column Name	Description
team	IPL franchise name
year	IPL season year
brand value	Estimated team brand value
rank	Brand value rank among all franchises

Table 3: Brand Value Dataset Columns

7. Workflow

A. Visualization Techniques

To explore the deeper narratives of the IPL, this project leverages a diverse set of visualizations designed to highlight trends in performance, pressure, and popularity. Each technique is selected to answer specific research questions through intuitive, story-driven graphics:

- **Rose Chart (Polar Bar Chart) –** Visualize top run-scorers across seasons to analyze consistency.
- **Violin Plot** – Study the distribution of wickets across seasons to evaluate the impact of bowlers.
- **Box Plot** – Examine playoff vs. league-stage performance to identify pressure-handling ability.
- **Line Chart with Bubble Markers** – Shows if winning leads to fan base growth using brand value data.
- **Bar Charts** – Compare team performance and win tallies across seasons.
- **Chord Diagram** – Showcase IPL's fiercest rivalries and their evolution over time.
- **Heatmap** – Highlight one-sided matchups where teams consistently struggle against specific opponents.
- **Bubble Chart** – Explore the relationship between popularity and on-field success.

These visualizations were implemented using Pandas, NumPy, Matplotlib, Seaborn, and Plotly in a Jupyter Notebook environment.

Failed Experiments and Method Selection Rationale

In the early stages of our analysis, we explored the use of a scatter plot to investigate the relationship between team brand value and fan following metrics, such as the number of match wins. The objective was to determine whether improved on-field performance translates to greater off-field popularity.

However, this approach proved ineffective for the following reasons:

• Overlapping Data Points: Several teams shared similar brand values and win counts, which led to overlapping points and a cluttered visualization. This made it difficult to discern meaningful patterns or outliers.

• Temporal Disconnection: Scatterplots are not well-suited to representing changes over time. Since understanding how performance and popularity evolve across seasons was essential to our analysis, this method lacked the required temporal context.

To address these limitations, we transitioned to a line chart with bubble markers, which offered a clearer and more insightful representation:

• Temporal Trends: The line chart format allowed us to display the progression of brand value over multiple seasons, highlighting growth or decline.

• Performance Encoding: By using bubble size to represent the number of wins in each season, we were able to integrate both brand value and on-field success into a single, cohesive visualization.

Data Type Reasoning

- To ensure each visualization method aligned with the nature of the data, we carefully considered the data types involved:
- Brand value is a ratio-level variable—it has a true zero and equal intervals —making it well-suited for line charts that effectively illustrate trends over time.
- Win counts are also ratio-level, allowing for meaningful magnitude comparisons; thus, bar charts or bubble charts work well to showcase performance.
- Toss decisions and match types are nominal variables with no inherent order, so bar charts are appropriate for visualizing their frequency distributions.
- Player rankings (such as the top 10 batters) are ordinal, reflecting a meaningful order without equal intervals. These are best represented using polar or radial charts (e.g., rose diagrams) that highlight ranking consistency.
- Match outcomes (win/loss) between team pairs are nominal and directional, making a chord diagram ideal to depict the dominance relationships between teams.

B. Exploratory Data Analysis: Data Preprocessing

As a first step in analysis, categorical columns were examined to identify distinct values and remove inconsistencies:

matches_df['city'].unique() matches_df['venue'].unique()
matches_df['toss_decision'].unique()
matches_df['match_type'].unique()

Standardizations were applied to unify representations:

```
# Loading the dataset
import pandas as pd

matches_df = pd.read_csv('Updated_Matches.csv')

# Inspecting unique categories from categorical columns for preprocessing visualization
unique_cities = matches_df['city'].dropna().unique()
unique_venues = matches_df['venue'].dropna().unique()
unique_toss_decisions = matches_df['toss_decision'].dropna().unique()
unique_match_types = matches_df['match_type'].dropna().unique()

# Standardizing venue names to avoid duplication
matches_df['venue'] = matches_df['venue'].replace({
    'MA Chidambaram Stadium, Chepauk': 'MA Chidambaram Stadium',
    'Rajiv Gandhi International Stadium, Uppal': 'Rajiv Gandhi Stadium',
    'Punjab Cricket Association Stadium, Mohali': 'PCA Stadium, Mohali',
    'Feroz Shah Kotla Ground': 'Feroz Shah Kotla',
    'Sawai Mansingh Stadium, Jaipur': 'Sawai Mansingh Stadium'
})

# Standardizing toss decisions for consistent analysis
matches_df['toss_decision'] = matches_df['toss_decision'].replace({
    'field': 'Field',
    'bat': 'Bat'
})

# Clean city names
matches_df['city'] = matches_df['city'].str.strip()

# Standardizing match type naming
matches_df['match_type'] = matches_df['match_type'].replace({
    'Elimination Final': 'Eliminator',
    'Semi Final': 'Semifinal'
})

# Display cleaned sample categories
{
    "unique_cities": unique_cities[:5],
    "unique_toss_decisions": unique_toss_decisions,
    "unique_match_types": unique_match_types,
    "standardized_venues_sample": matches_df['venue'].unique()[:5]
}
```

```
{'unique_cities': array(['Bangalore', 'Chandigarh', 'Delhi', 'Mumbai', 'Kolkata'],
       dtype=object),
 'unique_toss_decisions': array(['field', 'bat'], dtype=object),
 'unique_match_types': array(['League', 'Semi Final', 'Final', '3rd Place Play-Off',
       'Qualifier 1', 'Elimination Final', 'Qualifier 2', 'Eliminator'],
       dtype=object),
 'standardized_venues_sample': array(['M Chinnaswamy Stadium', 'PCA Stadium, Mohali', 'Feroz Shah Kotla',
       'Wankhede Stadium', 'Eden Gardens'], dtype=object)}
```

This ensured consistent groupings in analyses related to city-level trends, match-type break- downs, and team behavior by toss decisions.

C. Feature Engineering

New features were introduced to enrich the analysis. Date columns were converted into temporal components:

- Extracted the season year from the match date.
- Merged season information with ball-by-ball records for player-level

analysis. Key features engineered:

- **Player Consistency** – Count of seasons a player was in the top 10 scorers.
- **All-Time Run Totals** – Aggregated scores per batter across all seasons.
- **Bowler Impact** – Number of wickets taken by bowlers and presence in title-winning seasons.
- **Season-Level Aggregates** – Team wins, toss outcomes, and playoff performances.

These attributes formed the foundation for advanced visual comparisons and statistical summaries.

```python
import pandas as pd
import numpy as np

# Load and clean the datasets
ipl_data = pd.read_csv("/Users/sainihalkonduti/Downloads/Ipl20.csv")
matches_df = pd.read_csv("/Users/sainihalkonduti/Downloads/Updated_Matches.csv")

ipl_data.columns = ipl_data.columns.str.strip()
matches_df.columns = matches_df.columns.str.strip()
matches_df['date'] = pd.to_datetime(matches_df['date'], errors='coerce')
matches_df['season_year'] = matches_df['date'].dt.year

# Merge match date info into ball-by-ball data
ipl_data = pd.merge(
    ipl_data,
    matches_df[['id', 'season_year', 'date']],
    left_on='match_id',
    right_on='id',
    how='left'
)

# Feature 1: Consistent Batters (top 10 in runs for multiple seasons)
runs_per_season = ipl_data.groupby(['season_year', 'batter'])['batsman_runs'].sum().reset_index()
top_scorers = runs_per_season.sort_values(['season_year', 'batsman_runs'], ascending=[True, False])
top_10_each_season = top_scorers.groupby('season_year').head(10)
batter_consistency = top_10_each_season['batter'].value_counts().reset_index()
batter_consistency.columns = ['batter', 'seasons_in_top10']

# Feature 2: All-Time Top Scorers
all_time_runs = ipl_data.groupby('batter')['batsman_runs'].sum().reset_index()
top_run_scorers = all_time_runs.sort_values(by='batsman_runs', ascending=False).head(10)

# Feature 3: Impact Bowlers in Title-Winning Teams
valid_wickets = ipl_data[
    (ipl_data['is_wicket'] == 1) &
    (~ipl_data['dismissal_kind'].str.lower().fillna('').str.contains('run out'))
]
bowler_team_season = valid_wickets.groupby(
    ['bowler', 'season_year', 'bowling_team']
)['is_wicket'].count().reset_index(name='wickets')

# Attach season-winning team info
season_winners = matches_df.sort_values('date').groupby('season_year').tail(1)[['season_year', 'winner']]
season_winners = season_winners.rename(columns={'winner': 'season_winner'})
bowler_team_season = bowler_team_season.merge(season_winners, on='season_year', how='left')

# Feature: Was this bowler part of the title-winning team?
bowler_team_season['won_title'] = bowler_team_season['bowling_team'] == bowler_team_season['season_winner']

# Get top 10 wicket-takers and analyze their impact
top_bowlers = bowler_team_season.groupby('bowler')['wickets'].sum().nlargest(10).index
top_bowler_data = bowler_team_season[bowler_team_season['bowler'].isin(top_bowlers)]
bowler_impact_summary = top_bowler_data.groupby('bowler').agg(
    seasons_played=('season_year', 'nunique'),
    title_winning_seasons=('won_title', 'sum'),
    total_wickets=('wickets', 'sum')
).reset_index()
```

D. Handling Null Values

Across all datasets—including Updated Matches.csv, Ipl20.csv, and brand value records—no significant null values were found. The data was well-structured and complete, with minimal missing entries in non-critical fields like city or player of match. These did not interfere with analysis, and therefore, no imputation or major cleaning was required.

8. Analysis

Total Wins by IPL Teams (2008–2024)

Description: This horizontal bar chart illustrates the cumulative number of matches won by each IPL franchise from 2008 to 2024. Teams are sorted by total wins, providing a historical comparison of on-field success.

Key Insights:

- **Mumbai Indians** lead the chart with the highest win count, followed closely by **Chennai Super Kings**, reflecting their dominance and consistency.
- **Kolkata Knight Riders** and **Royal Challengers Bangalore** follow in the next tier, maintaining strong performances across seasons.
- Defunct teams such as **Deccan Chargers** and **Kochi Tuskers Kerala** appear lower due to fewer seasons played.
- New franchises like the **Gujarat Titans** and **Lucknow SuperGiants** have accumulated a respectable number of wins in fewer seasons, indicating early success.

Interpretation: This chart offers a quick visual benchmark for franchise performance across IPL history. By combining historical longevity with performance data, it helps spotlight consistent winners, rising contenders, and short-lived teams. Highlighting the top team enhances the chart's storytelling value, making it effective for comparative analysis.

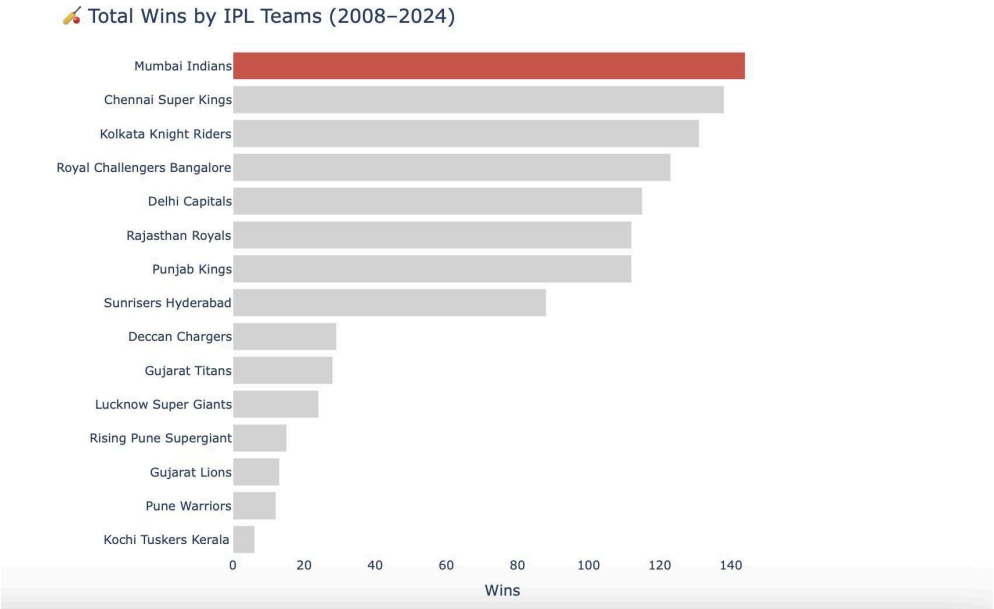


Figure 6: Bar Chart

Analysis: Player Impact on Team Success

Question Addressed: Do specific players have a significant impact on their teams' success? To evaluate the long-term impact of individual players, we used two key charts:

- **Chart 1: Consistency in Top 10 Run-Scorers Across Seasons** – This chart tracked how many times each player appeared in the seasonal top 10 run-getters from 2008 to 2024.
- **Chart 2: All-Time Total IPL Runs** – This chart visualized the aggregate career runs of top batters, providing a broader view of overall contribution.

Method: By comparing both visualizations, we identified players who not only delivered high aggregate runs but did so with remarkable season-to-season consistency. Players ap- pearing in both charts were flagged as having the most profound and sustained impact on the league.

Findings:

- **Virat Kohli, David Warner, Shikhar Dhawan, Rohit Sharma, AB de Villiers, Chris Gayle, and Suresh Raina** emerged as common names across both charts.
- These players consistently contributed large run totals, underlining their role as the backbone of their teams' batting strength.
- Their repeated presence among top performers suggests not only personal excellence but also a pivotal influence on match outcomes and team strategies.

Conclusion: Yes, specific players have shown sustained impact on their teams' success. Their ability to perform consistently across years cements their status as key influencers in IPL history.

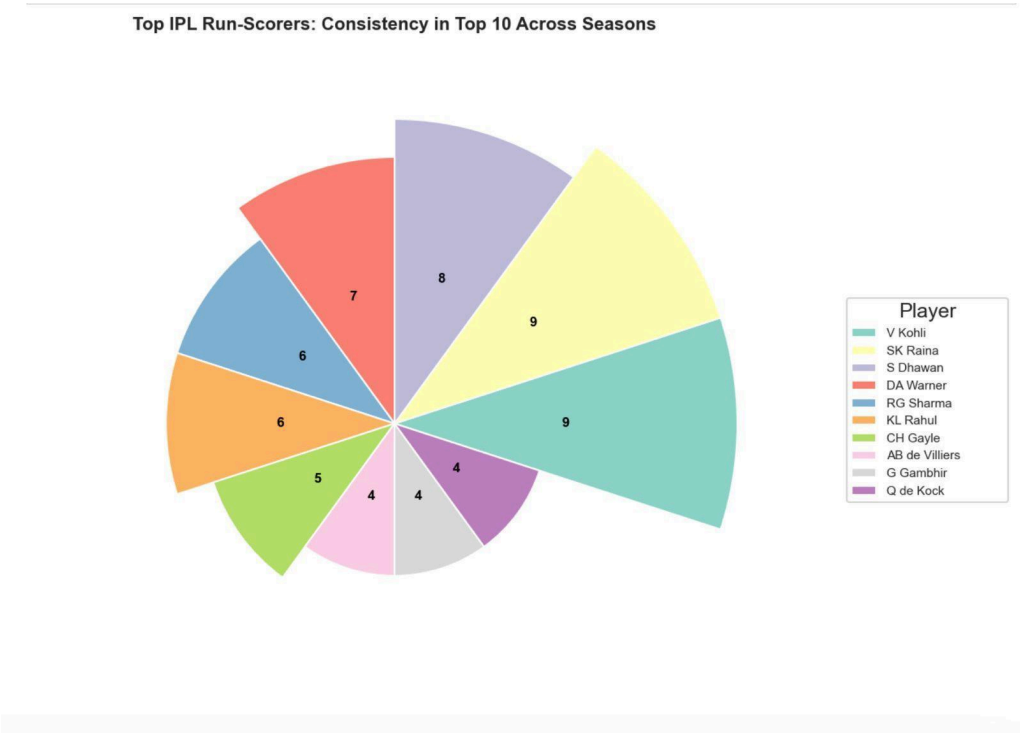


Figure 7: Rose Chart 1

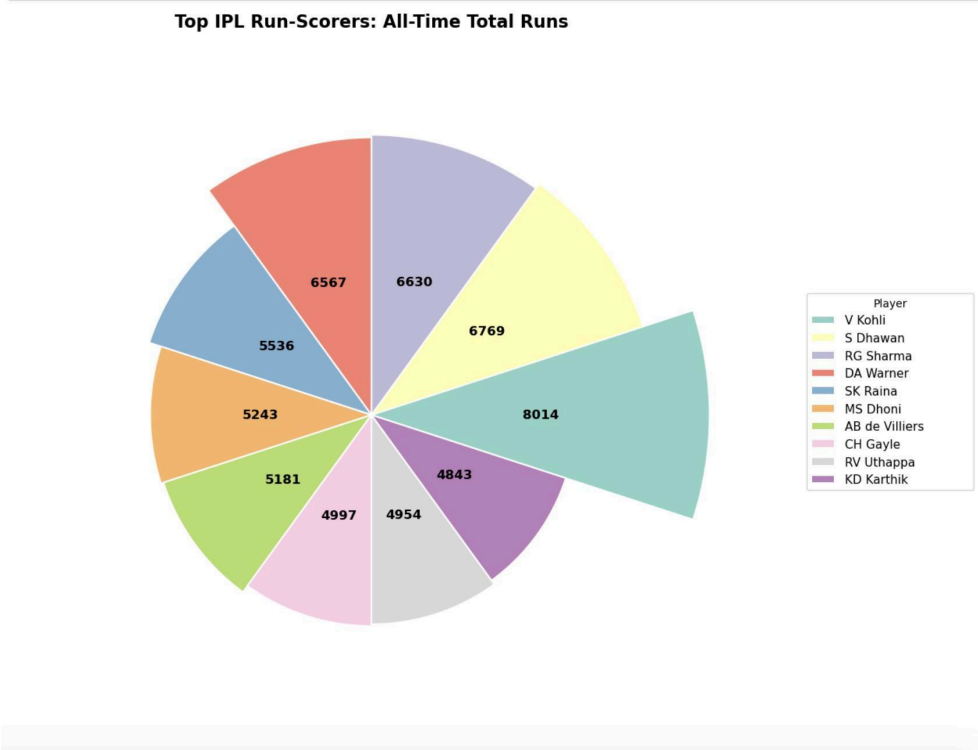


Figure 8: Rose Chart 2

Analysis: Player Impact on Team Success

Question Addressed:

Do specific types of players have a significant impact on their teams' success?

To evaluate the season-wise impact of individual performers, we used the following comparative analysis:

- **Chart 1: Wickets per Bowler (with Title Impact)** – This violin plot tracked season-wise wicket distributions for bowlers, highlighting those associated with title- winning teams. It helped isolate bowlers whose consistent performances contributed directly to team championships.
- **Chart 2: Runs per Batter (with Title Impact)** – A similar violin plot was used to visualize seasonal run totals of batters, with a focus on those in title-winning squads. Extreme run values were clipped at 700 to reduce skew from outliers.

Method:

By comparing the distribution of performance across seasons, we identified how bowling and batting contributions relate to title success. Players who consistently ranked high and appeared in winning teams were flagged as impactful. We also aggregated total wickets and winning seasons for top bowlers to support our findings.

Findings:

- **Jasprit Bumrah, DJ Bravo, SL Malinga, and SP Narine** emerged as standout bowlers, appearing repeatedly in title-winning seasons while maintaining high wicket counts.
- These bowlers were not only prolific but also consistent over time, suggesting they were pivotal to their teams' championship runs.
- Batters, while crucial, appeared more evenly spread across both winning and non-winning teams. Exceptional individual run tallies did not always translate into titles.

Conclusion:

Yes, specific players—especially bowlers—have had a significant impact on their teams' success. High-performing bowlers show a stronger correlation with title victories, making them central to IPL championship strategies.

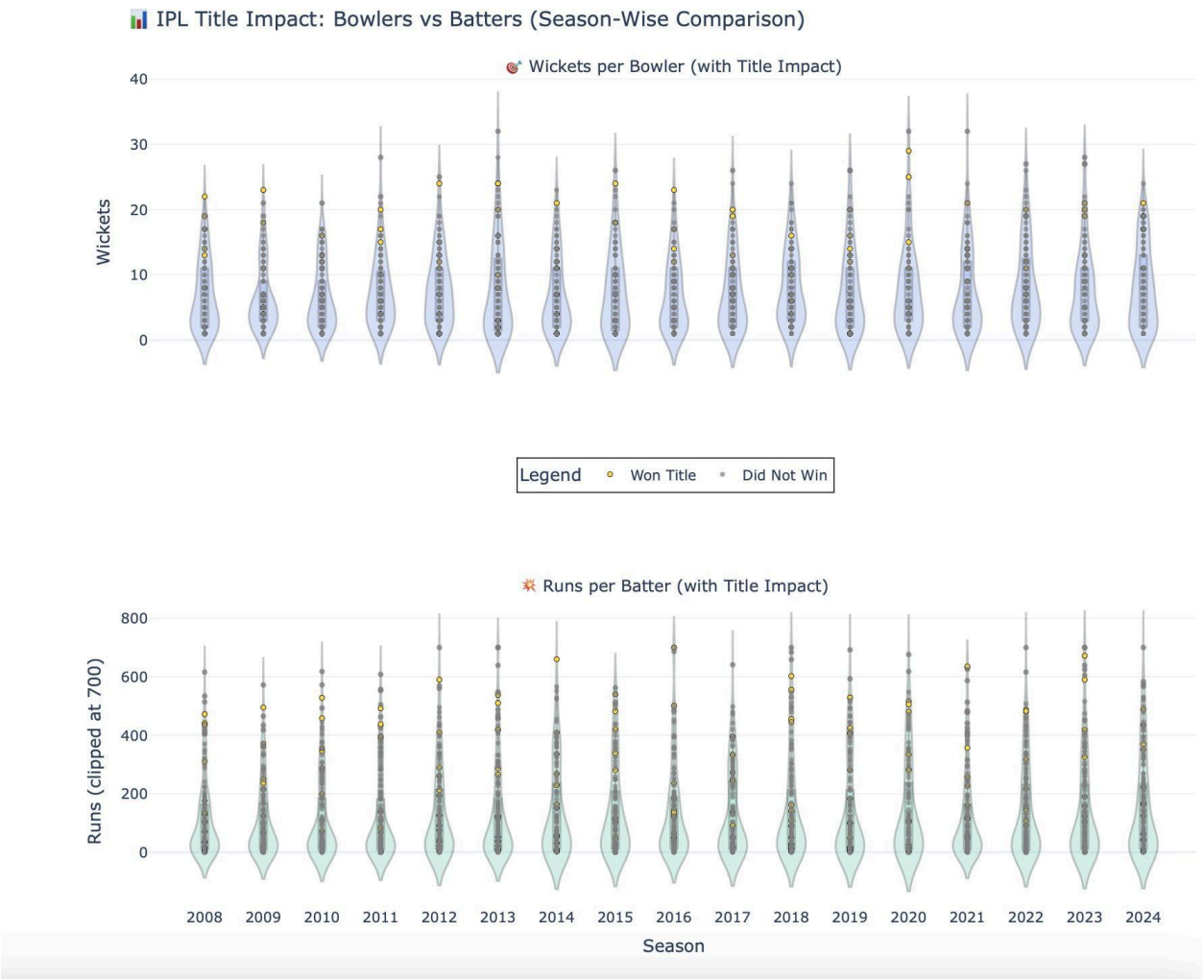


Figure 9: Violin Plots- Bowlers and Batters

Analysis: Playoff Performance Under Pressure

Question Addressed: How do teams perform in high-pressure playoff matches compared to league-stage games?

To evaluate teams' ability to handle pressure, we analyzed their scoring patterns in different match stages using a comparative box plot:

- **Chart: League vs. Playoff Run Distribution** – This box plot displays the total runs scored per match by each team during the league and playoff stages from IPL 2008 to 2024, highlighting shifts in performance under pressure.

Method: We used match-level scoring data and classified matches into 'League' and 'Playoff' stages. The box plot helped visualize the distribution, variability, and central tendencies of team performances in both contexts, allowing comparisons across franchises.

Findings:

- **Scoring tends to drop in the Playoffs** for teams like *Gujarat Titans*, *Sunrisers Hyderabad*, and *Royal Challengers Bangalore*, suggesting nerves or tougher competition in knockout games.
- *Punjab Kings* and *Rajasthan Royals* demonstrated **higher or consistent scoring** in playoffs despite irregular league-stage results—hinting at clutch potential.
- *Mumbai Indians* and *Chennai Super Kings* showed **tight, stable performance across both stages**, reflecting strategic depth and adaptability.
- **Newer franchises** like *Lucknow SuperGiants* showed higher variance due to smaller playoff sample sizes.
- **Overall playoff performance is less explosive**, reinforcing the narrative that tighter bowling and psychological pressure influence outcomes.

Conclusion: Playoff matches reshape team dynamics. While some teams maintain composure and deliver under pressure, others experience noticeable performance dips. Consistent playoff success often reflects deeper strategic competence, adaptability, and mental resilience.

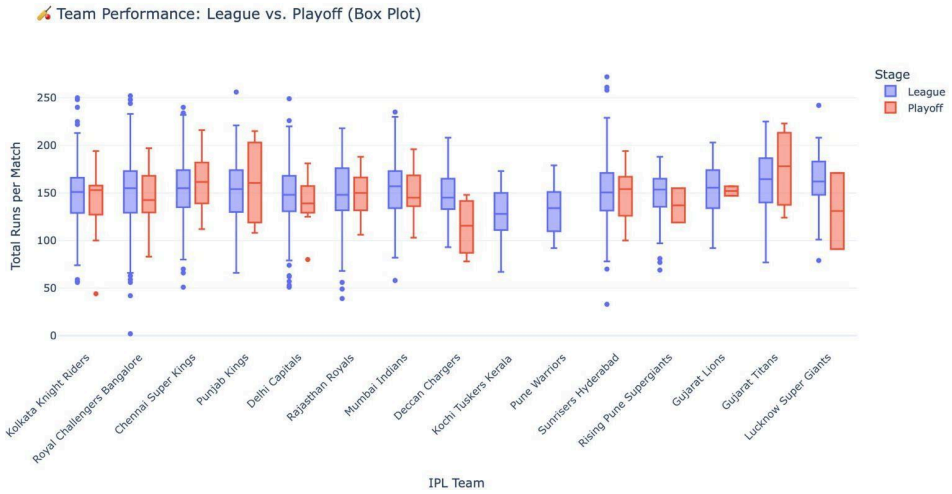


Figure 10: Box Plot for Playoffs and league stage

Analysis: Performance vs Popularity

Question Addressed: Which teams are gaining popularity, and is success the biggest factor driving fan following?

To investigate the relationship between on-field success and fan base growth, we analyzed two complementary charts:

- **Chart 1: Line Chart with Bubble Markers** – This chart plots each team's brand value over the years (2017–2024), with bubble sizes representing the number of wins in each season. It helps track how performance may drive brand value over time.
- **Chart 2: Bubble Chart – Wins vs Brand Value** – This bubble chart maps total wins (2008–2024) on the x-axis against latest brand value on the y-axis. Bubble size denotes brand growth percentage, revealing the overall popularity trajectory.

Method: We combined match results with annual brand valuation data to see if more wins translate to higher popularity. By observing chart overlaps and anomalies, we identified which teams succeeded both on and off the field, and which ones-built popularity despite inconsistent records.

Findings:

- **Mumbai Indians and Chennai Super Kings** consistently lead in both total wins and brand value, establishing their dominance in performance and commercial appeal.
- **Royal Challengers Bangalore**, despite fewer trophies, enjoy high brand value and fan loyalty, evident from their large bubble sizes.
- **Delhi Capitals and Kolkata Knight Riders** demonstrate a strong correlation between consistency and rising brand value.
- **Gujarat Titans and Lucknow SuperGiants** show disproportionately high brand growth despite limited match history, suggesting success in fan engagement and branding strategies.
- **Punjab Kings** struggle in both metrics, reflecting underachievement in converting legacy into influence.

Conclusion: While winning helps fuel popularity, it is not the sole driver. **Fan loyalty, visibility, and strategic marketing** significantly boost a team's public perception. Teams like RCB and GT show that even without consistent titles, effective branding and fan engagement can drive high brand value.

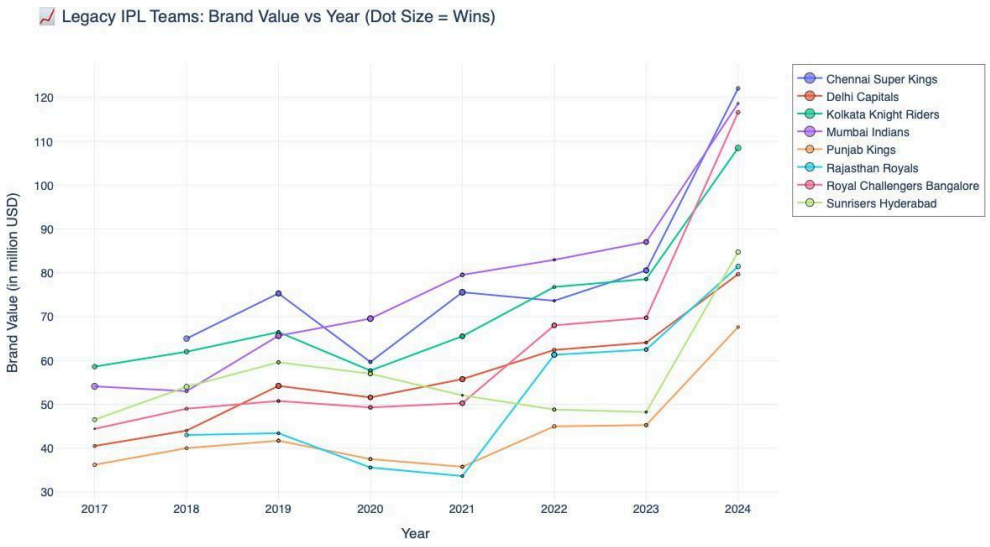


Figure 11: Line Chart with Bubble Markers

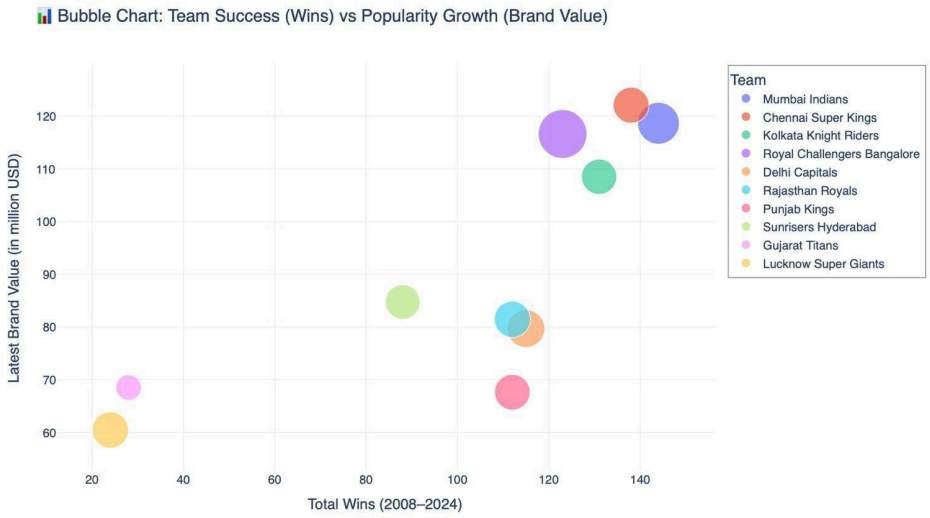


Figure 12: Bubble Chart

Analysis: One-Sided Rivalries – Heatmap

Question Addressed: Which matchups have been the most one-sided, with one team consistently dominating another?

This heatmap highlights one-sided matchups in IPL history, showing teams that consistently dominate or struggle against specific opponents. Darker cells indicate higher win margins or frequent losses, revealing intense rivalries and persistent underperformance patterns.

- **Chart: IPL Team vs Opponent Win Percentage Heatmap** – Visualizes head-to-head win percentages from IPL 2008 to 2024, with teams sorted and standardized.

Method: We calculated the win percentage of each team against every other opponent and plotted a matrix heatmap. We standardized team names and removed defunct teams to ensure cleaner analysis.

Findings:

- **One-sided rivalries:** *Sunrisers Hyderabad* dominated *Delhi Capitals* and *RCB* (100%), *Kolkata Knight Riders* beat *PBKS* (83%) and *DC* (77.8%), and *Rajasthan Royals* show strong records vs. both.
- **Balanced rivalries:** *MI vs CSK* (54.1%) and *RCB vs MI* (48.5%) reflect evenly matched competition.
- **Newer teams:** *GT* dominates early vs *RR* (80%) and *SRH* (75%), while *LSG* struggles vs *MI* and *RCB*.
- **Weak records:** *Punjab Kings* and *Sunrisers* show multiple deep-blue cells, indicating consistent underperformance.

Conclusion: The rivalry matrix captures long-standing dominance, evolving franchise dynamics, and emerging power shifts—critical for understanding IPL narratives beyond simple win counts.

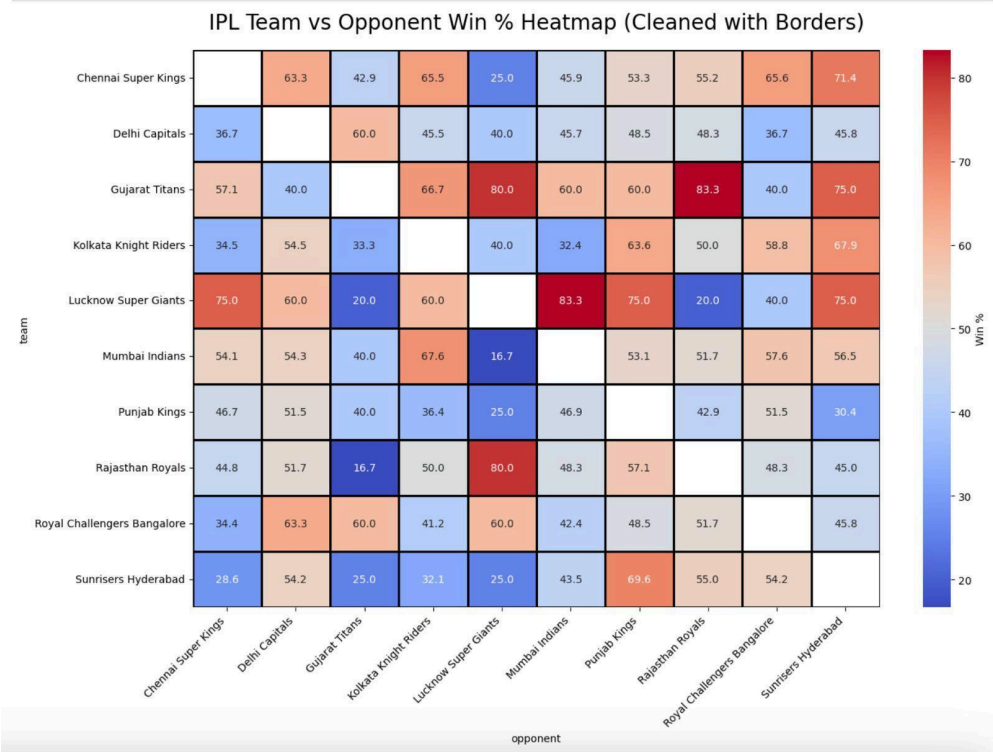


Figure 13: Heatmap for Legacy Teams

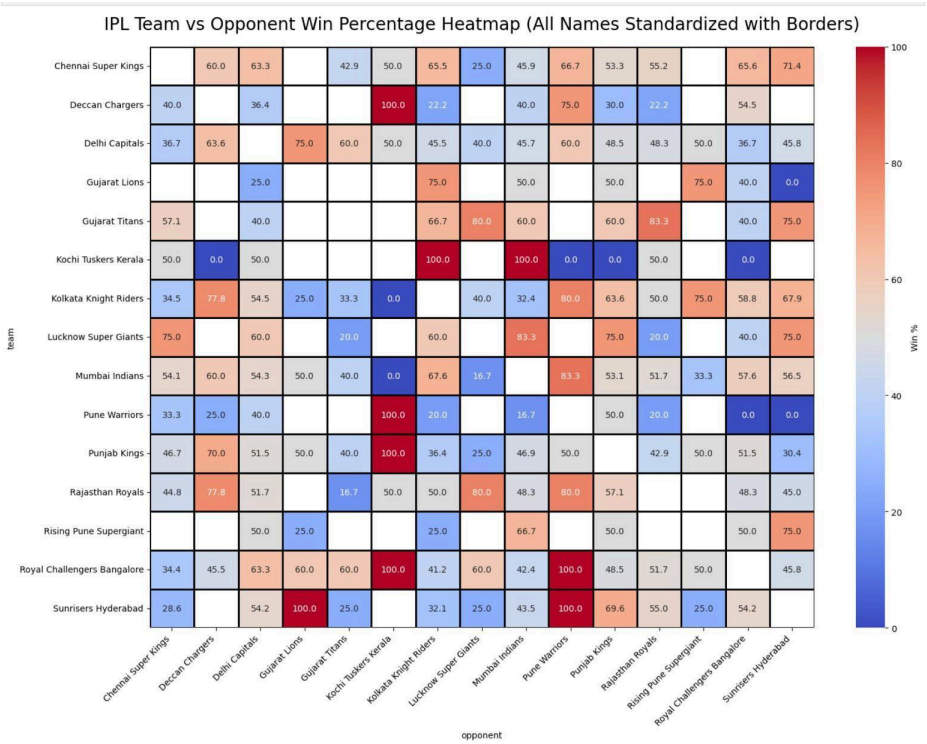


Figure 14: HeatMap for All teams

Analysis: Which Teams Have Dominated the IPL, and Why?

Question Addressed:

What makes a team dominant in the Indian Premier League? Is it about winning the most matches, sustaining high performance across seasons, retaining a strong core squad, or having standout individual performers?

To evaluate team dominance, we explored historical IPL data from 2008 to 2024 using various metrics and visual tools that measured both team and player-level impact.

Charts Used:

- **Bar Chart: Total Wins by IPL Teams** – Displays total match wins per franchise from 2008 to 2024, highlighting long-term success.
- **Facet Bar Charts: Win% by Season** – Shows seasonal win percentage trends for teams with over 10 seasons.
- **Treemap: Player Retention (5+ Seasons)** – Visualizes which teams retained players long-term, pointing to team stability.
- **Rose Chart: Run-Scoring Dominance** – Captures key batters' consistent impact over seasons using a radial layout.

Method:

To investigate which teams have dominated the IPL, we applied multi-faceted analytical techniques using match and player data:

- *Total Wins Aggregation:* Calculated total wins from 2008–2024, identifying teams like MI and CSK as top match-winners.
- *Win Percentage by Season:* Measured consistency across seasons for long-term franchises, using per-season win rates to reveal performance stability.
- *Player Retention Treemap:* Mapped how many players each team retained for 5+ seasons. Teams like CSK, KKR, and MI showed large, stable cores, supporting team chemistry and continuity.
- *Rose Chart – Individual Batting Dominance:* Captured which batters consistently led their teams in run-scoring. Players like Kohli, Raina, and Rohit emerged as core contributors to team success, reinforcing that individual brilliance often complements team dominance.

Data was processed and visualized using Python (pandas, seaborn, plotly), offering both clarity and depth.

Findings:

- Mumbai Indians and Chennai Super Kings clearly dominate in terms of **total wins**, with Mumbai leading the tally and CSK boasting consistent win percentages.
- **Stable team cores** are a recurring theme in dominant teams. Both CSK and MI retained key players for 5+ seasons (e.g., MS Dhoni – 15 years, Rohit Sharma – 14 years), creating strong internal dynamics.
- **Win consistency** across years differentiates elite teams from inconsistent ones. CSK and MI maintained high win rates in most seasons, while teams like Punjab Kings and Delhi Capitals fluctuated more.
- **Star players play a central role.** The rose chart confirms that high-performing individuals—like Kohli for RCB, Raina for CSK, and Rohit for MI—often contribute to a team's long-term competitive edge.
- **Underperforming franchises** such as PBKS and DC show a pattern of frequent player rotations and weaker season-by-season records, which may contribute to their lack of titles.

Conclusion:

Team dominance in the IPL is not solely about winning titles—it's a combination of **consistent performance**, **long-term player retention**, and **leadership from core stars**. Teams like Mumbai Indians and Chennai Super Kings demonstrate that *continuity breeds success*, while individual brilliance, strategic stability, and adaptable squad management help sustain performance across seasons. Meanwhile, teams lacking cohesion or cycling through frequent changes struggle to build lasting success.

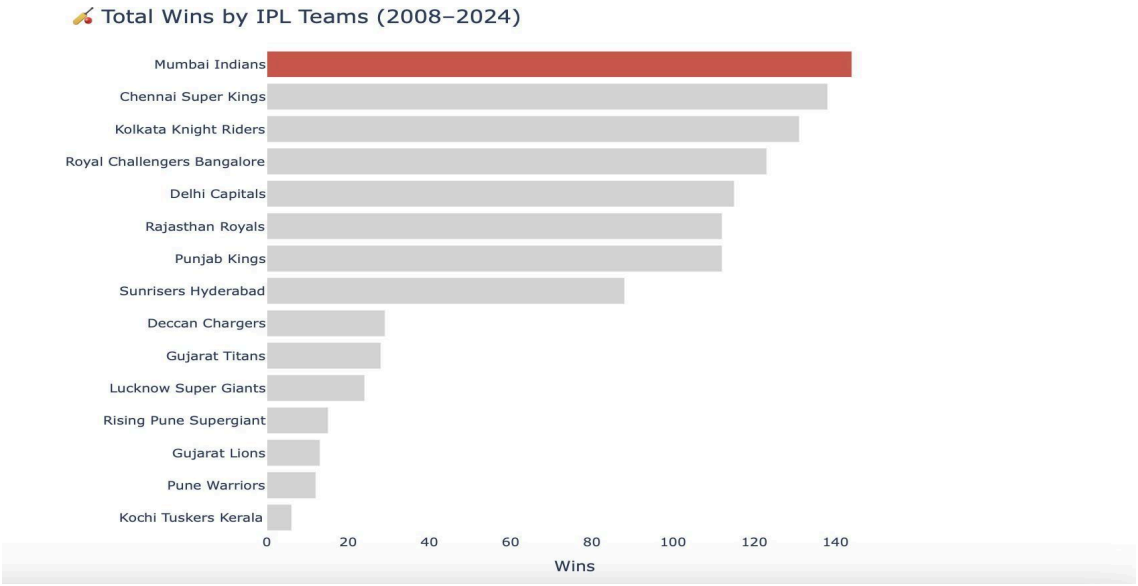


Figure 15: Bar Chart For IPL Wins

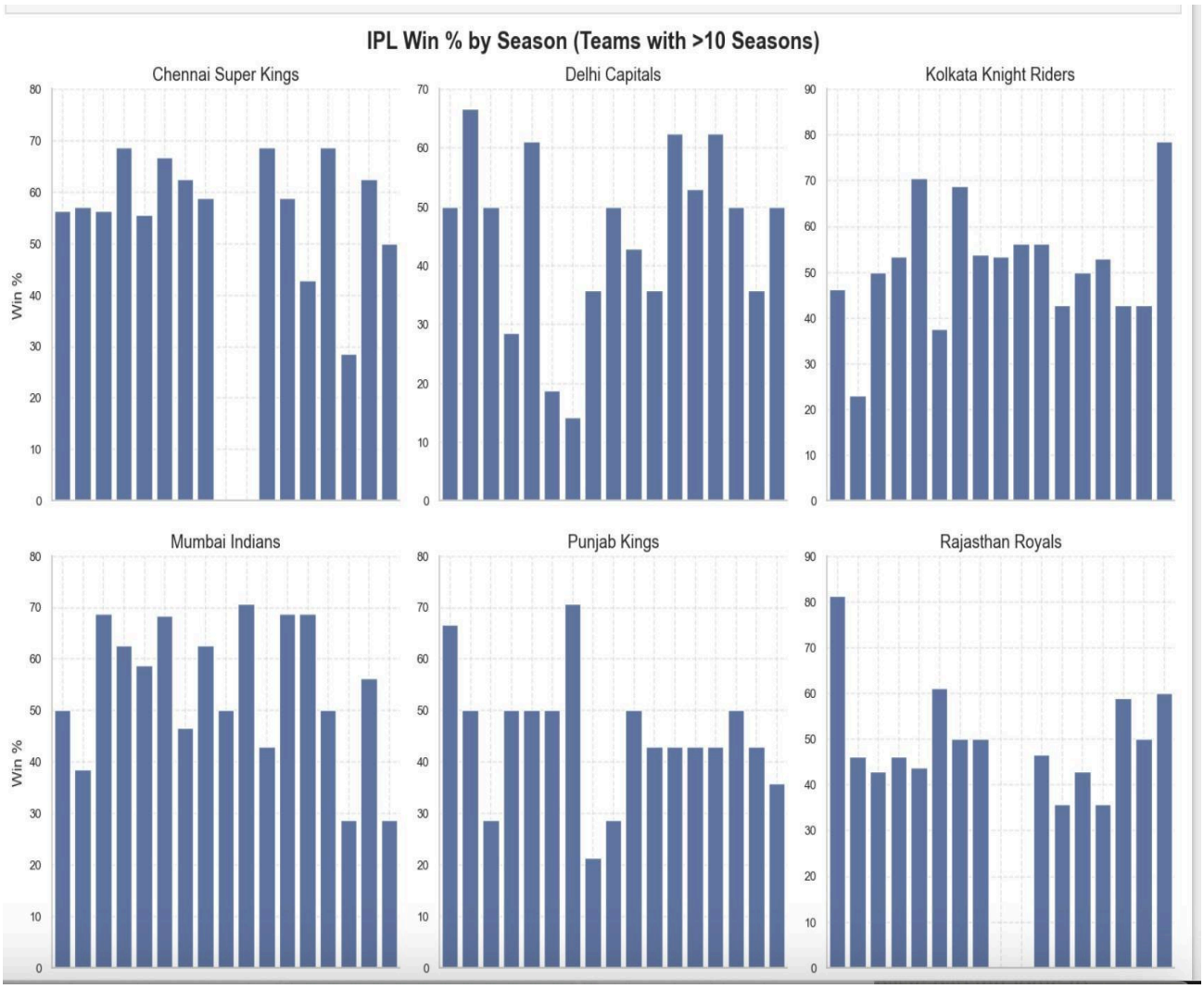


Figure 16: Facet Bar Chart 1

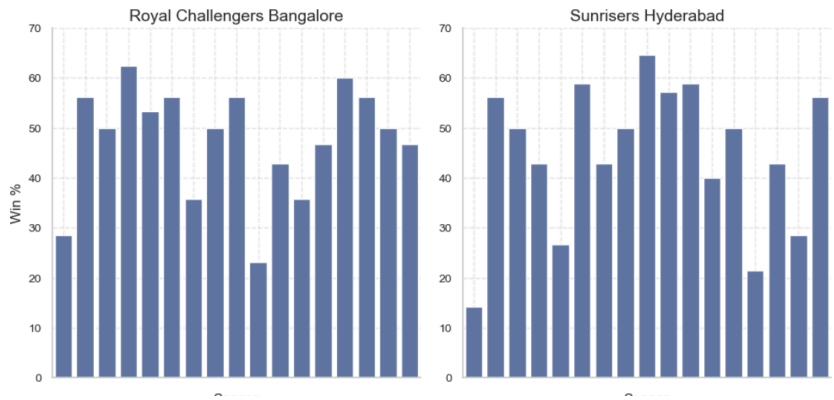


Figure 17: Facet Bar Chart 2

Figure 18: Retention TreeMap

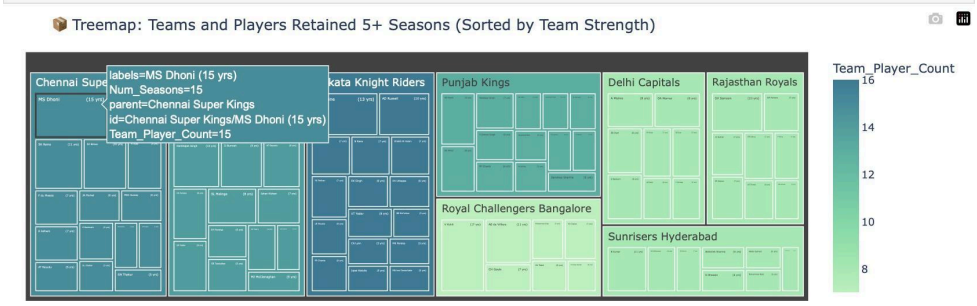


Figure 19: Retention TreeMap

Analysis: Venue Advantage and Match Conditions

Question Addressed: What impact do different match venues and conditions have on IPL team performances?

To investigate how geography and venue familiarity affect outcomes, we analyzed team wins by city and correlated them with home performance statistics of long-term franchises.

- **Chart 1: IPL Geography Map** – This interactive map visualizes team wins across Indian cities. Each circle's size represents the number of victories in that location, and color denotes team identity. Popups reveal per-team win distribution in each city, highlighting strongholds and neutral venues.
- **Chart 2: Home Win Percentage Table** – This data table ranks teams by their win percentage in home games, reflecting their effectiveness in leveraging local conditions over time.

Method: We filtered match data (2008–2024) to include only games with known cities. We created a team vs. city win map and computed each team's win percentage in matches held in their home cities. For multi-home teams like Punjab Kings and Rajasthan Royals, both cities were considered. Teams with fewer than 10 IPL seasons were excluded to ensure statistical stability.

Findings:

- **Kolkata Knight Riders** lead with a **71.2% home win rate**, showcasing strong utilization of Eden Gardens' conditions.
- **Sunrisers Hyderabad** (69.2%) and **Mumbai Indians** (67.4%) also exhibit high home success, supported by consistent pitch behavior and strong home crowds.
- **Chennai Super Kings**, often viewed as unbeatable at Chepauk, record a respectable **63.3% win rate**, though not the highest.
- **Punjab Kings** show a very low **18.8% home win rate**, indicating venue instability or ineffective use of home advantages.
- **Royal Challengers Bangalore** have a middling **55.8% win rate**, likely affected by their batting-friendly but unpredictable home pitch.

Conclusion: Match venues and local conditions significantly shape IPL outcomes. Teams that build strategies around home environments—considering pitch nature, boundary sizes, and crowd support—gain a measurable edge. Long-term success is often rooted in maximizing home advantage alongside tactical adaptability.

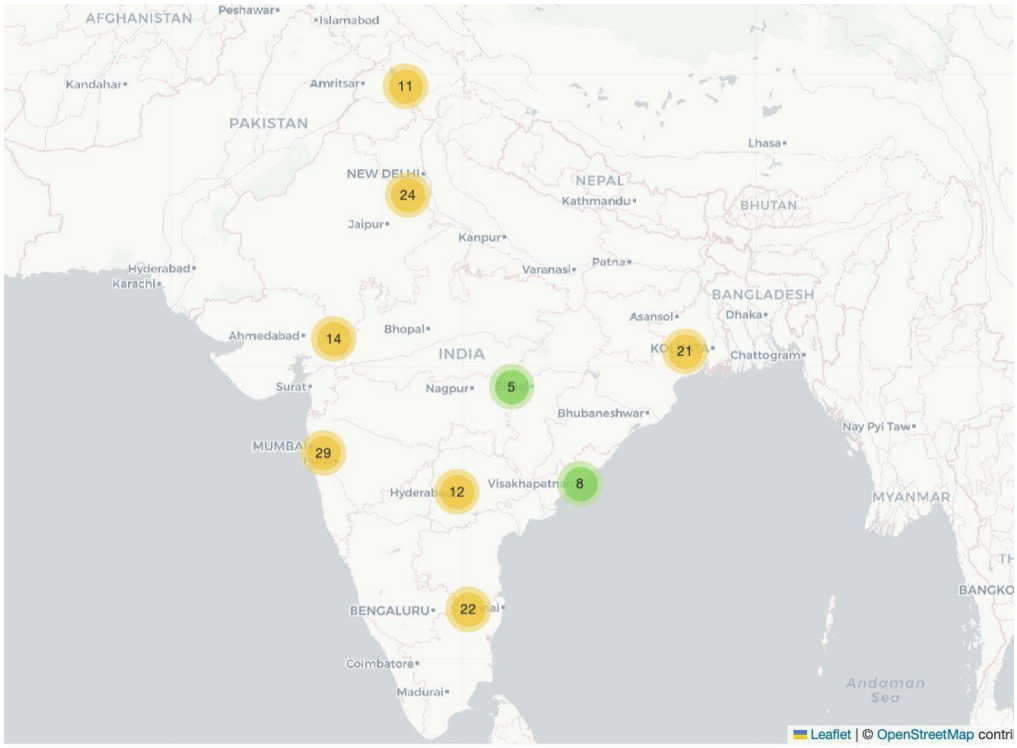


Figure 20: Geography Map -Venues Data Visualization

	winner	home_win_pct	home_wins	total_home_matches
2	Kolkata Knight Riders	71.232877	52	73
7	Sunrisers Hyderabad	69.230769	36	52
3	Mumbai Indians	67.368421	64	95
1	Delhi Capitals	67.272727	37	55
5	Rajasthan Royals	64.912281	37	57
0	Chennai Super Kings	63.291139	50	79
6	Royal Challengers Bangalore	55.769231	29	52
4	Punjab Kings	18.750000	6	32

Home conditions significantly influence IPL team success, as revealed by the home win statistics of long-term franchises:

Figure 21: Home vs Away win percent

Analysis: The Rivalry Chronicles – IPL's Fiercest Head-to-Head Battles

Question Addressed: Which head-to-head matchups define the IPL's most intense and historic rivalries?

Rivalries lie at the heart of the Indian Premier League's drama. They encapsulate not just competition, but emotion, loyalty, and legacy. To unearth the fiercest matchups, we analyzed win distributions between long-term IPL franchises and visualized them using a chord diagram.

- **Chart: Chord Diagram – IPL Matchups and Rivalries** — Each node in the diagram represents a long-standing IPL team. Curved, directional links between them

indicate the number of matches played and the number of wins secured by the dominant team. Rivalries are highlighted with custom colors and edge thickness based on magnitude and disparity.

Method: We filtered matches from 2008 to 2024 between teams with at least 10 seasons. For each team pair, we computed:

- Total matches played
- Wins by each team
- Win difference

Only rivalries meeting at least one of the following criteria were flagged as dominant:

- At least 15 wins by one team and a win difference of 5 or more
- Overall matchups count exceeding 35 games

Edges in the chord diagram were directed from the dominant team to the opponent, with color representing the dominant team and thickness indicating the win margin.

Findings:

- **Mumbai Indians vs Chennai Super Kings** emerges as the most played and balanced rivalry, though MI holds a slight edge.
- **RCB vs KKR** and **CSK vs RCB** stand out with clear dominance, as indicated by thick edges in favor of KKR and CSK, respectively.
- **Green edges** represent matchups without overwhelming dominance, typically involving teams with shorter histories or balanced win shares.
- Thick, unidirectional flows illustrate domination—e.g., MI and CSK consistently outperform PBKS and DC in head-to-head contests.
- Some rivalries (like **MI vs KKR**) remain highly competitive despite high frequency, suggesting historically close contests.

Conclusion: IPL rivalries are more than just statistical records—they reflect legacies, psychological edges, and fan-fueled intensity. Chord diagrams offer a powerful lens to visualize how certain teams have established dominance while others fight to balance historic records. These rivalries shape the narrative arcs of IPL seasons and are central to the league's enduring popularity.

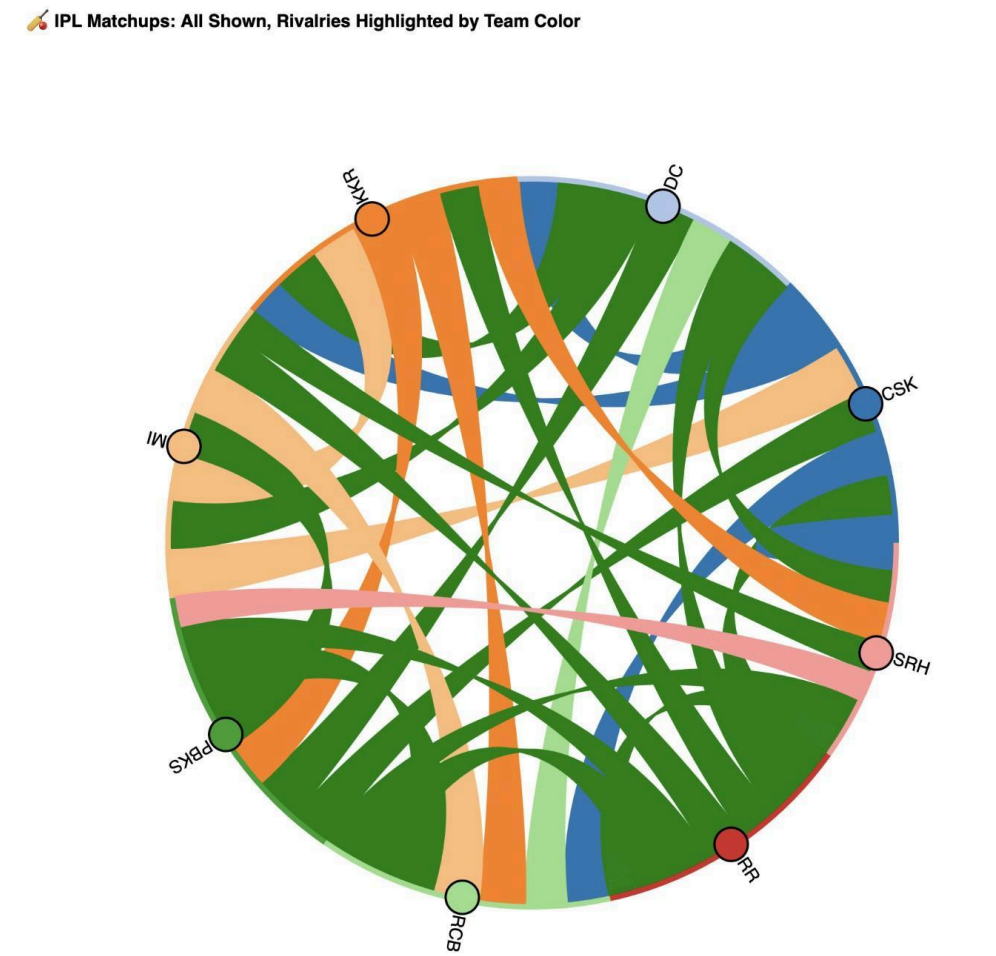


Figure 22: Chord Diagram Data Visualization

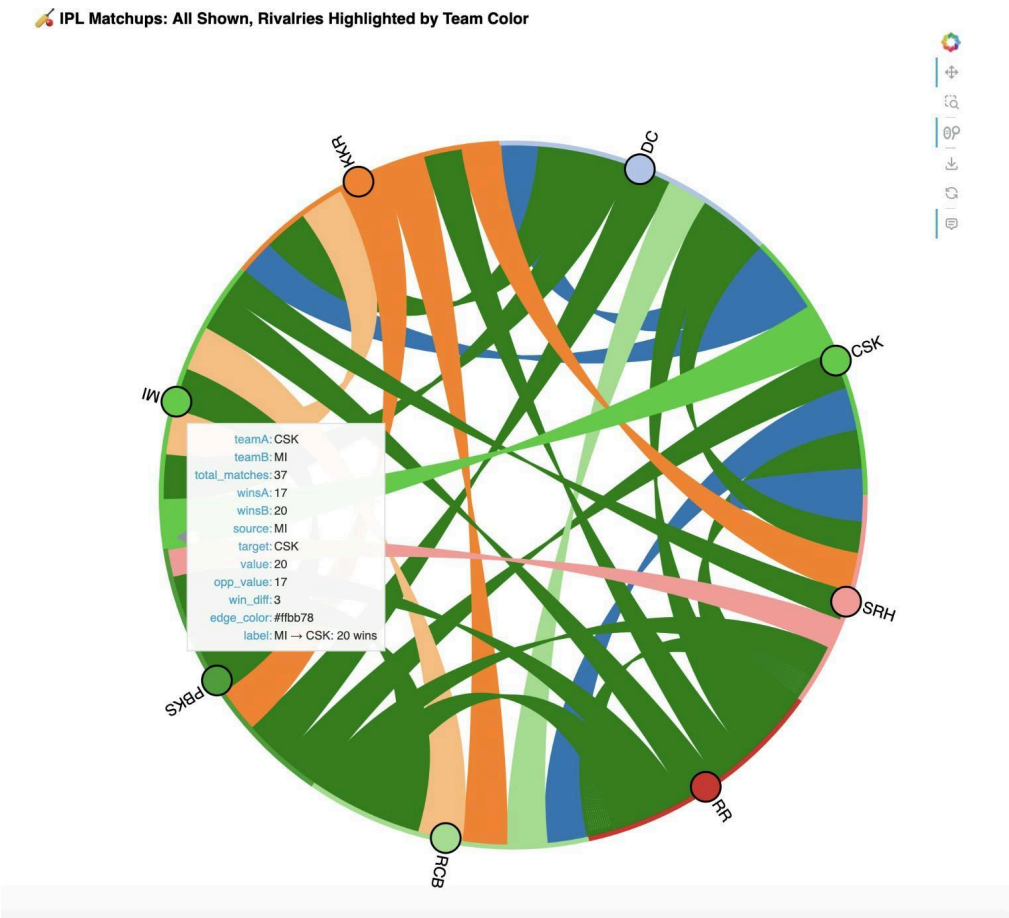


Figure 23: Chord Diagram Image 2

Analysis: The Underdog Story – Which Teams Have Overachieved or Underperformed?

Question Addressed

Which IPL teams have consistently exceeded or fallen short of performance expectations across seasons?

Chart

- **Performance Difference Bar Chart (All Teams)** – This stacked bar chart compares the actual win percentage to the expected (rolling 3-season average) for every IPL team from 2009 to 2024.
- **Filtered Chart for Long-Term Teams** – A second plot zooms in on franchises with over a decade of IPL participation, isolating consistent patterns in under- or overperformance.

Method

We calculated the expected win percentage using a 3-year moving average per team (excluding the current season), then subtracted it from the actual season-wise win rate. This created a **performance difference** metric:

$$\textit{Performance Difference} = \textit{Actual Win \%} - \textit{Expected Win \%}$$

Positive values signify overachievement, while negative values imply underperformance. Data was analyzed separately for all teams and long-standing franchises.

Findings

- **Frequent Overachievers:** Delhi Capitals and Punjab Kings had multiple breakout seasons (e.g., 2014, 2020) despite not being title favorites. These surges were often linked to revamped leadership or young rosters.
- **New Entrants Making a Mark:** Gujarat Titans had a stellar debut in 2022, significantly outperforming expectations. Lucknow Super Giants also exceeded the typical startup trajectory.
- **Consistent Performers:** Mumbai Indians and Chennai Super Kings generally showed stable alignment between expected and actual outcomes, reflecting strategic depth and leadership continuity.
- **Frequent Underperformers:** Royal Challengers Bangalore and Sunrisers Hyderabad appeared in the negative zone across multiple seasons, pointing to execution gaps despite talented squads. Rajasthan Royals also struggled post-2010.
- **Unpredictable Teams:** Kolkata Knight Riders and Sunrisers Hyderabad oscillated between outperforming and underperforming seasons, suggesting fluctuations driven by auction dynamics and internal change.

Conclusion

Success in the IPL is not solely determined by star power or legacy but often depends on internal strategy, talent cohesion, and long-term planning. Overachievement typically stems from smart recruitment and adaptability, while inconsistency, despite marquee names, highlights the complexity of succeeding in T20 formats.

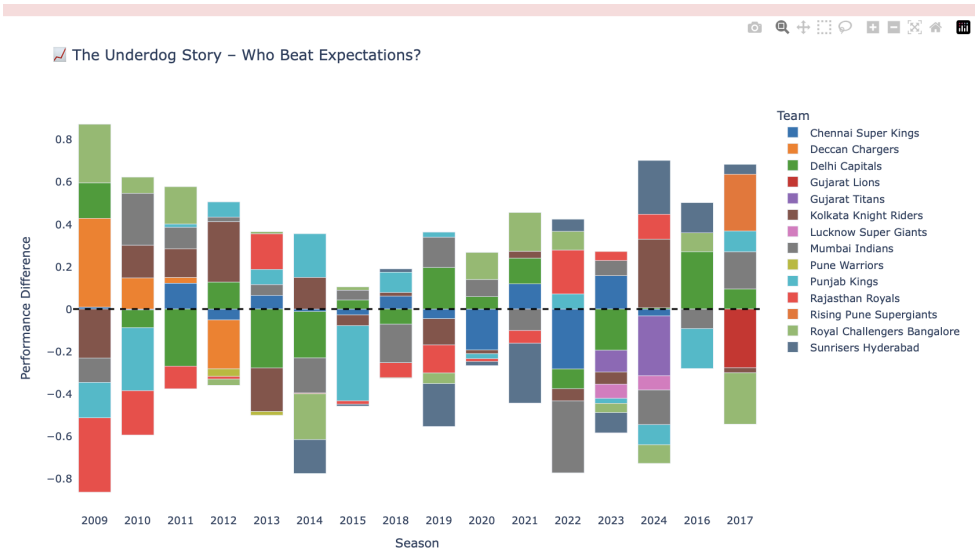


Figure 24: All teams (Overachieved or Underperformed)

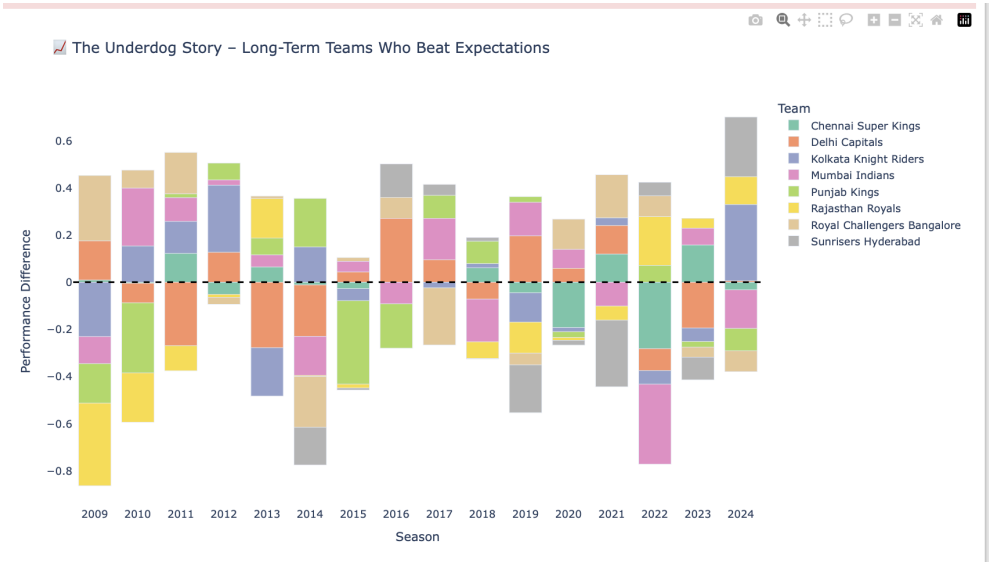


Figure 25: Underdog Story visualization

Analysis: Last-Over Thrillers – Identifying Teams That Excel in Crunch Situations

Question Addressed

Which IPL teams have demonstrated strong performance under pressure in final-over matches?

Chart

- **Horizontal Bar Chart** – Displays the win rate of each IPL team in close matches decided in the 20th over.
- Criteria: Only matches with a margin of ≤ 6 runs (defense) or ≤ 2 wickets (chase) are considered. Teams must have played a minimum of 3 such games to be included.

Method

We filtered matches that ended with:

- A win by *runs* with a margin ≤ 6
- A win by *wickets* with only 1 or 2 wickets remaining

We then calculated the win percentage for each team in these high-pressure scenarios, considering only those with at least 3 such encounters.

Findings

- **Top Performers:** Lucknow Super Giants lead the rankings despite being a new team, highlighting exceptional tactical execution and finishing ability.
- **Legacy Execution:** Chennai Super Kings and Sunrisers Hyderabad follow closely, likely benefiting from strong captaincy and experienced finishers.
- **Notable Performers:** RCB and Rajasthan Royals exhibit strong clutch-game performance, often pulling through tense finishes.
- **Surprising Midfield:** Mumbai Indians and Kolkata Knight Riders rank in the middle despite being traditionally strong, indicating occasional lapses under pressure.
- **Struggling Finishers:** Gujarat Titans and Rising Pune Supergiant occupy the bottom, potentially due to a lack of experience or weak death-over strategies.

Conclusion

Success in last-over thrillers is not just about talent—it hinges on calm decision-making, effective leadership, and end-game tactics. Some newer teams have quickly adapted and now outperform even the IPL's historic giants in crunch-time performance.

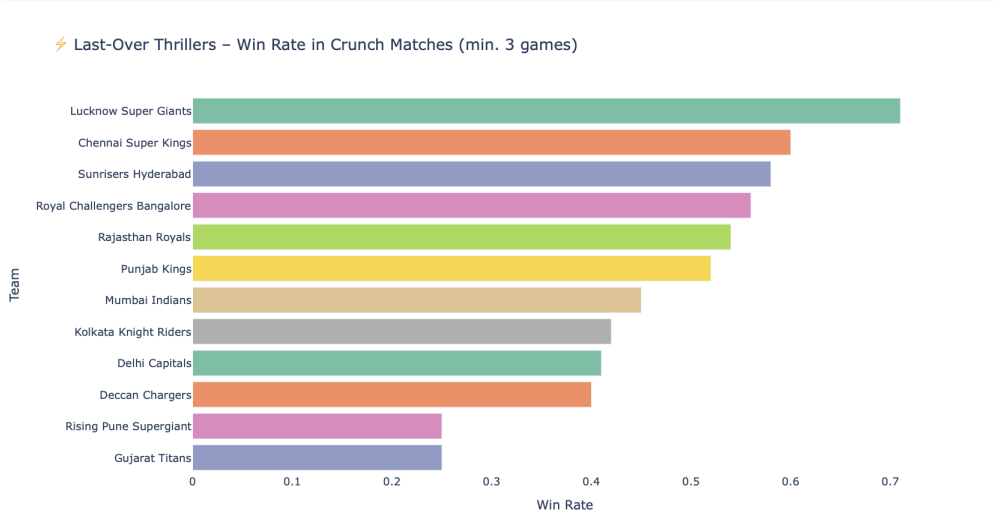


Figure 26: Win Rate Data Visualization

9. Discussion, Conclusion, and Future Work

Discussion

The Indian Premier League (IPL) is not just a compilation of match results, scores, and title counts—it is a multi-layered narrative filled with rivalries, comebacks, and defining moments. This project aimed to uncover the *untold stories* of the IPL that are often buried beneath the surface of raw numbers.

Rather than merely stating **what** happened, our approach focused on understanding **why** it happened. We explored high-pressure performances, consistency, team rivalries, and the evolving dynamics of fan support and brand value.

- Violin plots showed that title-winning teams tend to feature strong, high-performing bowling units, more so than batting stars.
- Box plots revealed which franchises thrive in high-stakes playoff matches and which struggle under pressure.
- Chord diagrams illustrated directional dominance in the league's most iconic rivalries.
- Heatmaps exposed one-sided head-to-head matchups, unveiling persistent underperformance patterns.
- Rose charts spotlight players who consistently deliver impact performances across seasons.
- Geographic maps displayed home dominance and neutral venue effects across India.
- Bubble charts linked team performance (wins) with popularity (brand value), raising questions about the role of loyalty versus results.

- Bar charts surfaced the overachievers—teams that exceeded expectations in specific seasons due to strong leadership, youth investment, or tactical cohesion.

Each chart told a sub-story, making it clear that the IPL is not one story, but many interwoven ones.

Conclusion

This project demonstrates that effective data storytelling can make IPL insights more engaging and human. By shifting the lens from static metrics to dynamic narratives, we uncover how teams handle pressure, how rivalries evolve, and how players step up in clutch moments.

Our overall approach was guided by the following principles:

- **Focus on why, not just what**
- **Highlight rivalries and big moments**
- **Explore pressure and trends**
- **Make IPL insights clear and engaging**

From strategic depth and leadership consistency (e.g., CSK and MI) to emerging narratives of underdogs and debutants (e.g., GT and LSG), the IPL is a living dataset of sport, emotion, and business.

Future Work

There is significant potential to extend this analysis:

- Develop ball-by-ball clustering models to isolate pressure phases such as death overs or Super Overs.
- Study player career arcs with auction history, role transitions, and retention impact.
- Perform sentiment analysis on social media to correlate fan reactions with match events and team form.
- Create real-time interactive dashboards that allow dynamic filtering and narrative tracking.
- Build predictive models incorporating venue, recent form, and matchups to simulate future outcomes.

  In summary, the IPL is not just a sports league—it is a vibrant ecosystem of strategy, psychology, legacy, and luck. By treating data as a medium for storytelling, we have only scratched the surface of what can be discovered.

References

[1] **S. Dhilip's IPL Data Analysis**
https://www.sdhilip.com/exploring-indian-premier-league-with-interactive-visualizations/

[2] **Official IPL Website – Match Statistics, Team Performances, and Historical Data**
https://www.iplt20.com/

[3] **Brand Directory IPL Report – Team Brand Values and Popularity Rankings**
https://brandirectory.com/reports/indian-premier-league/2020

[4] **Kaggle IPL 2008–2024 Player History Dataset – Detailed Ball-by-Ball Player Statistics**
https://www.kaggle.com/datasets/raj110/ipl-2008-2024-player-history

[5] **Data Visualization Catalogue – Guide for Selecting Appropriate Visualizations**
https://datavizcatalogue.com