## 

## 

# Data Visualization on IPL Dataset

Kamal Panchal

Department of Applied Data Science, San Jose State University

DATA 230: Data Visualization

Prof. Andrew H. Bond

December 11, 2022

## 

## 

## 

## 

**Abstract**

The Indian Premier League, or IPL, is cricket's most lucrative format because it allows young, gifted players to demonstrate their skills on a variety of fields. The primary users of all sports analytics framework tenets are decision-makers. Numerous athletes and teams across many sports have found success thanks to sports analytics. The top players for a squad may be chosen using sports analytics and data visualization. Indian Premier League (IPL), Indian professional Twenty20 (T20) cricket league established in 2008. The league, which is based on a Round- robin group and knockout format, has teams in major Indian cities. The IPL is the most-attended cricket league in the world and in 2014 was ranked sixth by average attendance among all sports leagues. Here some visualizations on the IPL Dataset have been done which was taken from Kaggle. This report discusses how data visualization and analysis related to toss may help decision-makers find natural players for their teams. Graphs for Highest Run Scorer, Most Wicket Taker, the outcome of the toss, biggest win by runs, wins by the most number of wickets, Highest total, etc are some of the other visualizations. Also some data visualization on different statistics which can be beneficial for the supporters and fans who want to analyze the IPL and make predictions using the statistics. It will make it a lot easier to understand the data if it is in the form of visualization.

Table of Contents

[**Data Visualization on IPL Dataset**](#_heading=h.1v6155pxhu4h) **1**

[1. INTRODUCTION](#_heading=h.gjdgxs) **4**

[i. Project Summary:](#_heading=h.30j0zll) 5

[ii. Purpose:](#_heading=h.1fob9te) 5

[iii. Intended Audience](#_heading=h.2et92p0) 5

[2. Project Background](#_heading=h.2s8eyo1) **5**

[i. Project Planning and Scheduling:](#_heading=h.3rdcrjn) 6

[ii. Project Development Approach](#_heading=h.35nkun2) 6

[a) Tableau](#_heading=h.1ksv4uv) 6

[b) Tableau Features](#_heading=h.fqzzkudr8873) 6

[**3. DATASET**](#_heading=h.4ho28s6yl65) **7**

[i. IPL dataset](#_heading=h.4i7ojhp) 8

[**IPL Complete Dataset (2008-2020)**](#_heading=h.z9150kb4n6gp) **8**

[ii. Preprocessing](#_heading=h.44my32gokx62) 8

[iii. Data Fields](#_heading=h.xz0arwsn6n45) 9

[4. DATA VISUALIZATION](#_heading=h.3as4poj) **10**

[i. Process](#_heading=h.1pxezwc) 10

[a) Importing Dataset to Tableau :](#_heading=h.49x2ik5) 10

[b) Creating different Worksheets:](#_heading=h.43ogz1w0rhmx) 10

[c) Worksheet 1:](#_heading=h.4gg8s9l7djcv) 11

[d) Worksheet 2:](#_heading=h.1tzbd2kbsc7i) 12

[e) Worksheet 3:](#_heading=h.vf75cobuhtqo) 13

[f) Worksheet 4:](#_heading=h.et6ig3hjnyld) 14

[g) Worksheet 5:](#_heading=h.sh0rnuayuet7) 15

[h) Worksheet 6:](#_heading=h.9ujiij953j2n) 16

[i) Worksheet 7:](#_heading=h.2vwkin83uzpv) 17

[j) Worksheet 8:](#_heading=h.ku6oow8hoj8v) 18

[k) Worksheet 9:](#_heading=h.4x9qm8i4z107) 19

[l) Worksheet 10:](#_heading=h.7fw93f2iqiss) 20

[m) Worksheet 11:](#_heading=h.lzg4lof66zzw) 21

[ii. Dashboards](#_heading=h.1hmsyys) 22

[a. Dashboard 1:](#_heading=h.bb3ifybp6ggk) 22

[b. Dashboard 2:](#_heading=h.8xbanzpdnzgf) 23

[c. Dashboard 3:](#_heading=h.4yhbxquy9un) 24

[5. Conclusion](#_heading=h.43ky6rz) **25**

[REFERENCE](#_heading=h.2iq8gzs) 25

# INTRODUCTION

## Project Summary:

* In this project, data visualizations regarding data of different IPL seasons has been done using Tableau.
* This project focuses on different visualizations like whether toss outcome matters the win/loss probability of the team, which ground was more effective for the batting side or the bowling side, which team won the most number of leagues also which team had the most win in a single season and so on.

## Purpose:

* IPL being one of the most competitive cricket leagues all over the world, the cricket pundits, coaches and the fans love to take notes and analysis of their favorite teams .
* To have their players perform better in each match, all the playing teams have been taking an analysis on different players' performance, the pitch conditions as well as pros and cons of different teams.
* The main purpose of this project is to analyze this through different forms of visualizations so that it can be beneficial to the teams and players as well as the coaches to emphasize this visualizations to get the best out of their teams

## Intended Audience

* Project Supervisor
* Class Students
* Die-Hard Cricket Fans .
* Coaches & Players.

# Project Background

## Project Planning and Scheduling:

* Tableau has been used in this project to visualize data related to data from several IPL seasons. This project focuses on various visualizations, such as whether the outcome of the coin toss affects the team's win/loss probability, which field was better for the batting side or the bowling side, which team won the most leagues as well as which team had the most victories in a single season, and so on.This project's major goal is to examine this through various visualizations so that teams, players, and coaches may benefit from emphasizing these visuals to get the most out of their teams.

## Project Development Approach

### Tableau

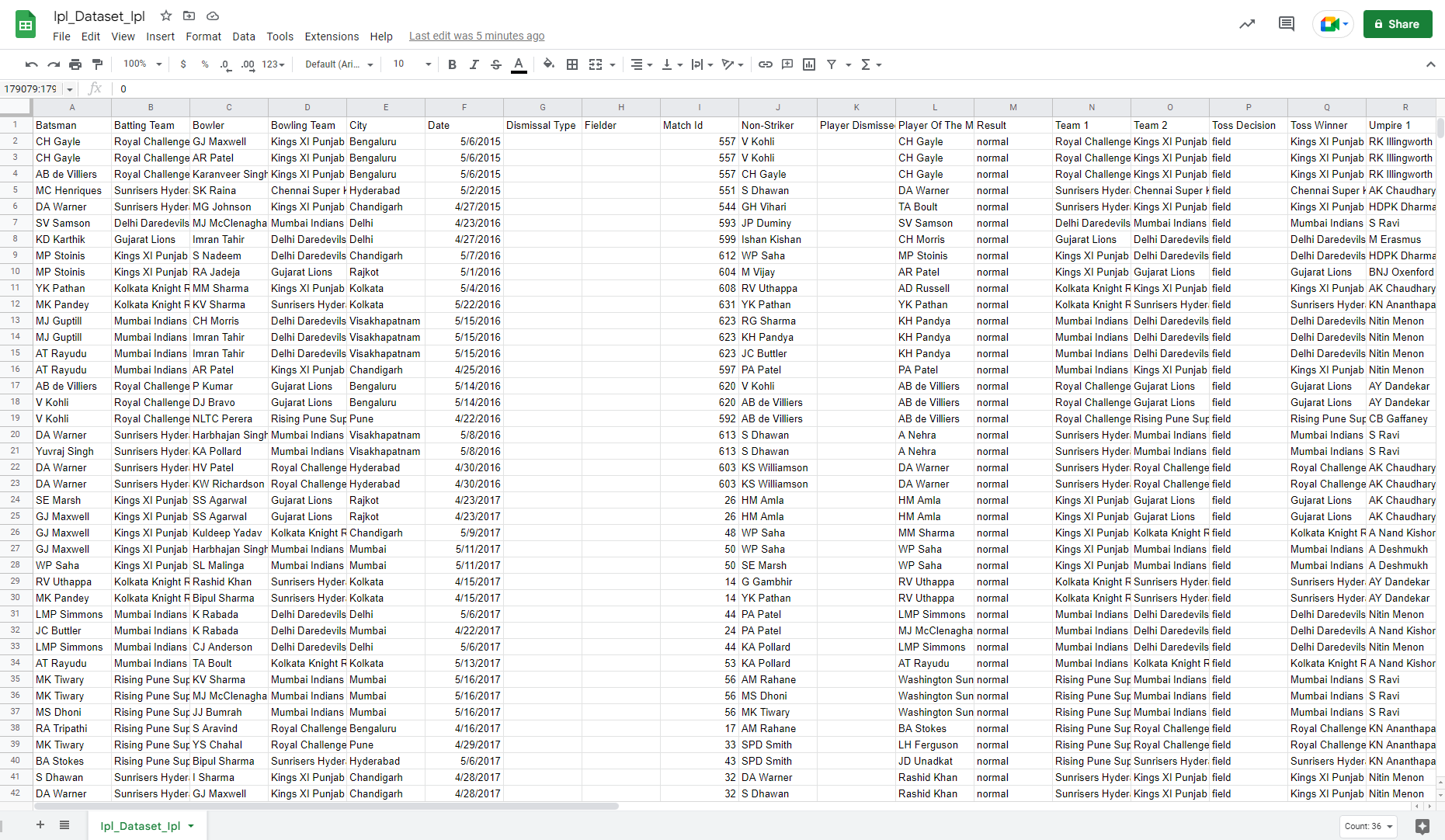
* For data analysis and business intelligence, Tableau is a top data visualization tool. Tableau is a fantastic business intelligence and data visualization application for reporting and analyzing huge amounts of data. It was founded in America in 2003, and in June 2019, Salesforce bought Tableau. It assists users in producing a variety of graphs, maps, dashboards, and stories for the purpose of visualizing and analyzing data to aid in corporate decision-making. Tableau is one of the most well-liked business intelligence tools since it offers so many interesting, distinctive features (BI).

### Tableau Features

* Users of Tableau can quickly find the answers to crucial questions because to the platform's robust data search and exploration capabilities.
* You don't need to know any programming beforehand.
* Users with no prior expertise in producing visualizations may get started right away using Tableau.
* It can link to a number of data sources that are not supported by other BI products. Users of Tableau may combine and blend various datasets to produce reports.
* A single place to manage all published data sources inside an organization is supported by Tableau Server.

# DATASET

## IPL dataset

* Indian Premier League (IPL)(2008-2020) dataset
* Source : Kaggle
* Dataset size : 49 MB
* Dataset Columns : 46
* Dataset Rows : 179079

## Preprocessing

* Total 4 Initial Datasets combined to 1 dataset
* Dropped Multiple column which were not required
* Missing values removed
* Not much data cleaning required as most of the data was cleaned
* Columns after preprocessing : 39

## Data Fields



# DATA VISUALIZATION

## Process

### Importing Dataset to Tableau :

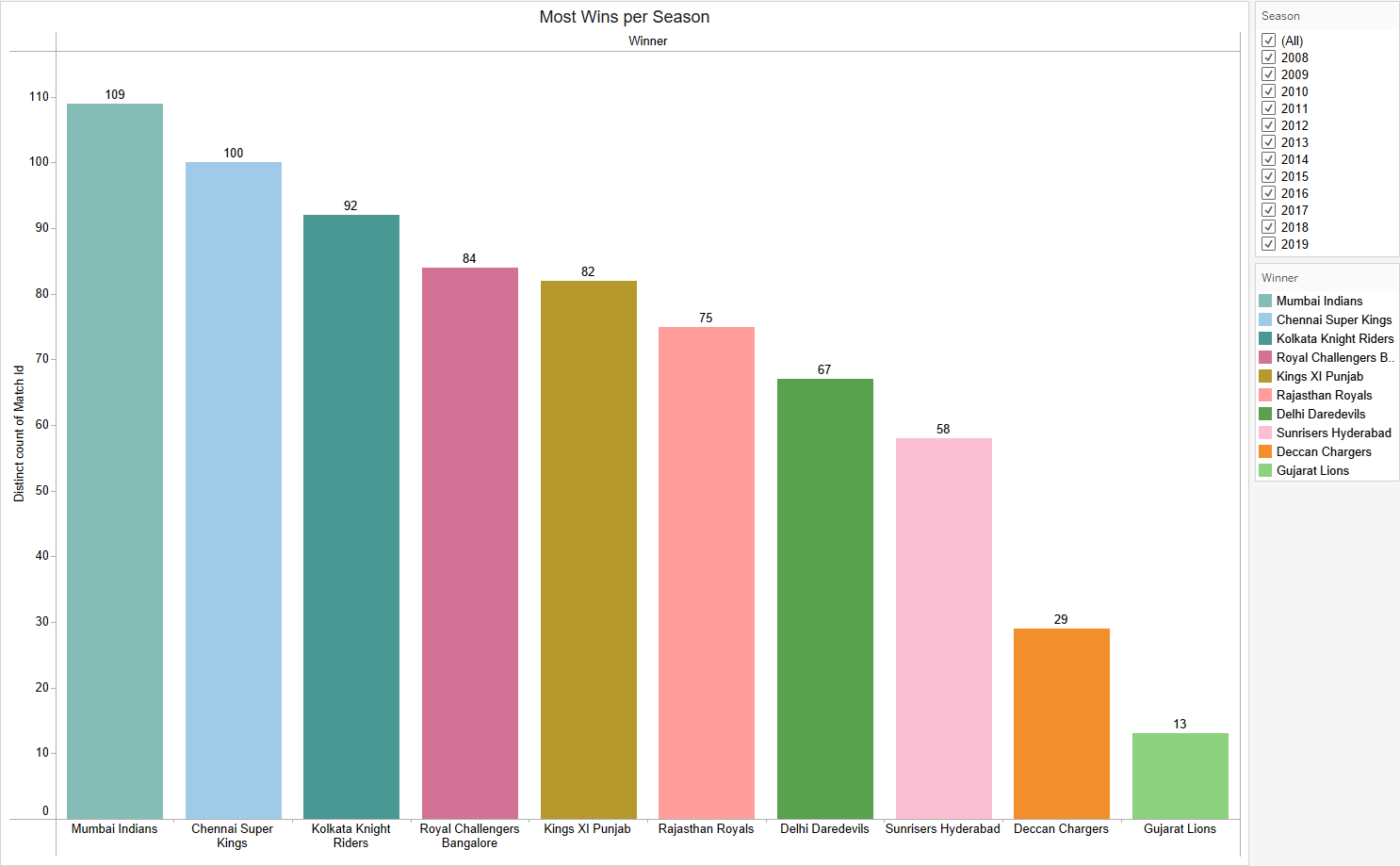
* Step 1: Locate the Tableau Text File Connector.
* Step 2: Choose the Tableau CSV Connector.
* Step 3: Import the Desired CSV File from your System.
* Step 4: Configure the Tableau CSV Text File Properties.

### Creating different Worksheets:

* Total 9 worksheets were visualized regarding different statistics of the teams playing in IPL throughout different seasons.
* Detailed explanation regarding each sheet has been given.
* Some major statistics visualized were most wins by any team, average win ratio of each team, team with most coin tosses won, how toss impacts the win rate on different occasions and many more.

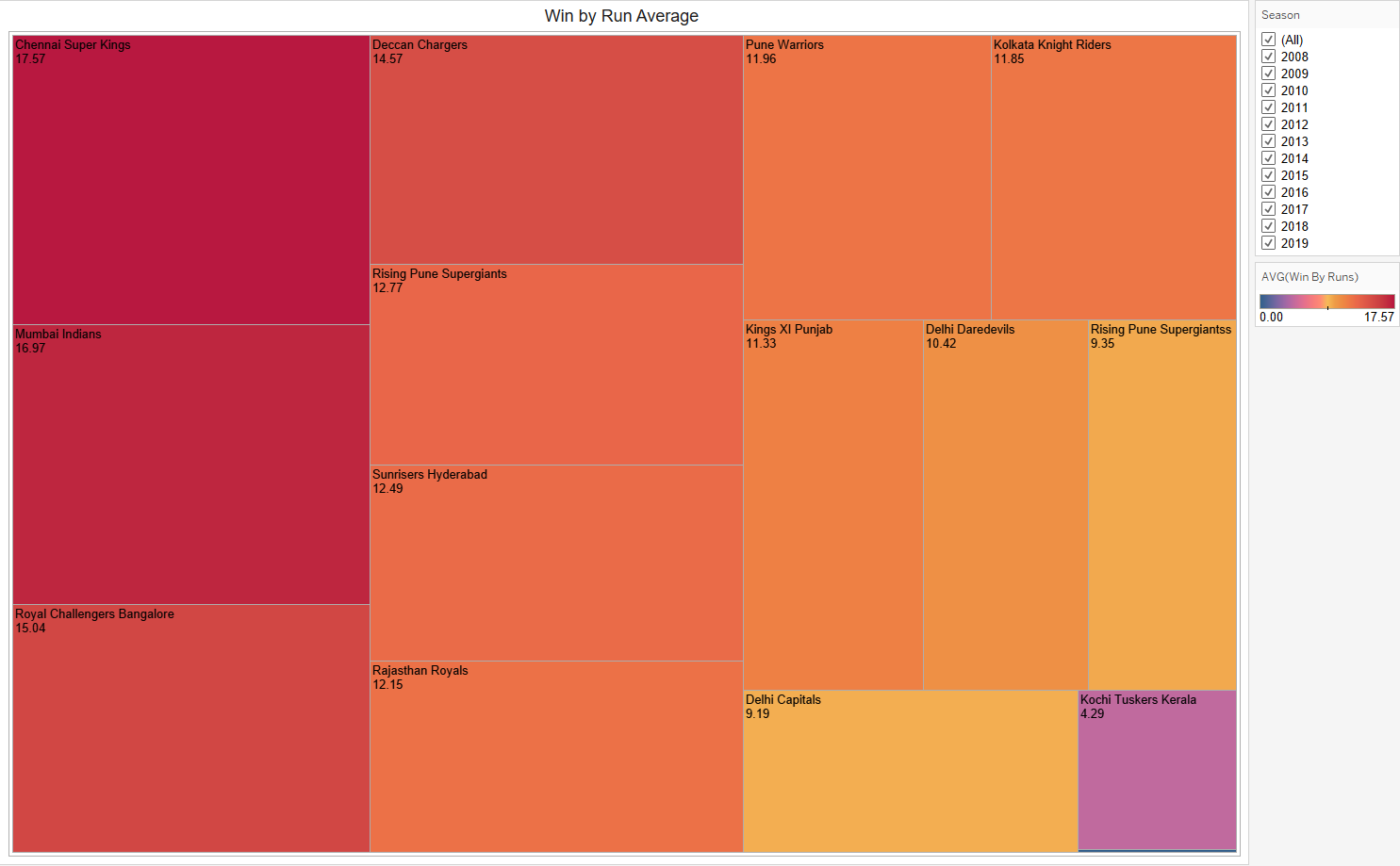
### Worksheet 1:

### The total number of wins by each team in all seasons out of which Team Mumbai Indians had had the highest wins with 109 wins in total.



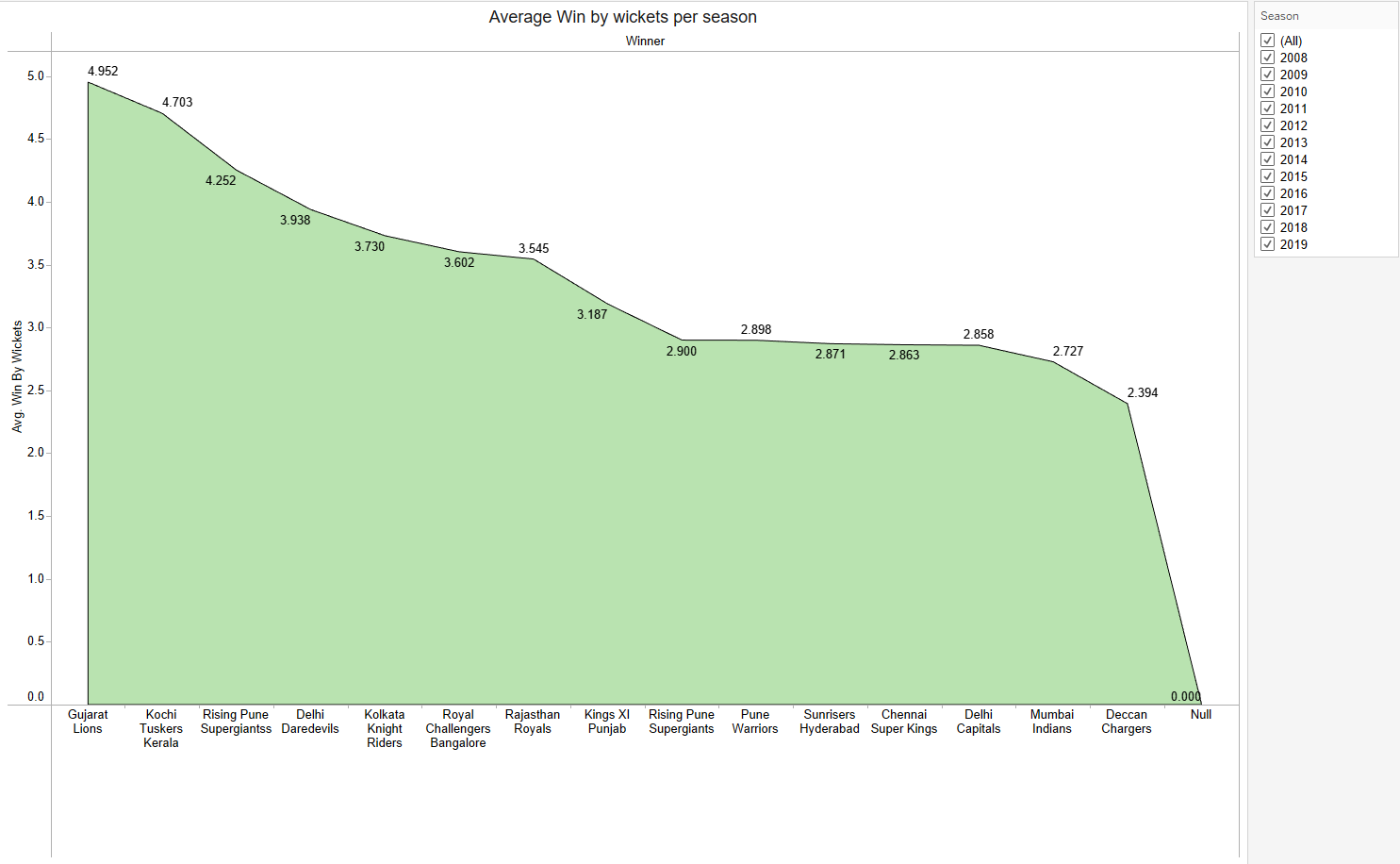
### Worksheet 2:

* This sheet shows the average by which each team won-by-runs in each game of different seasons.



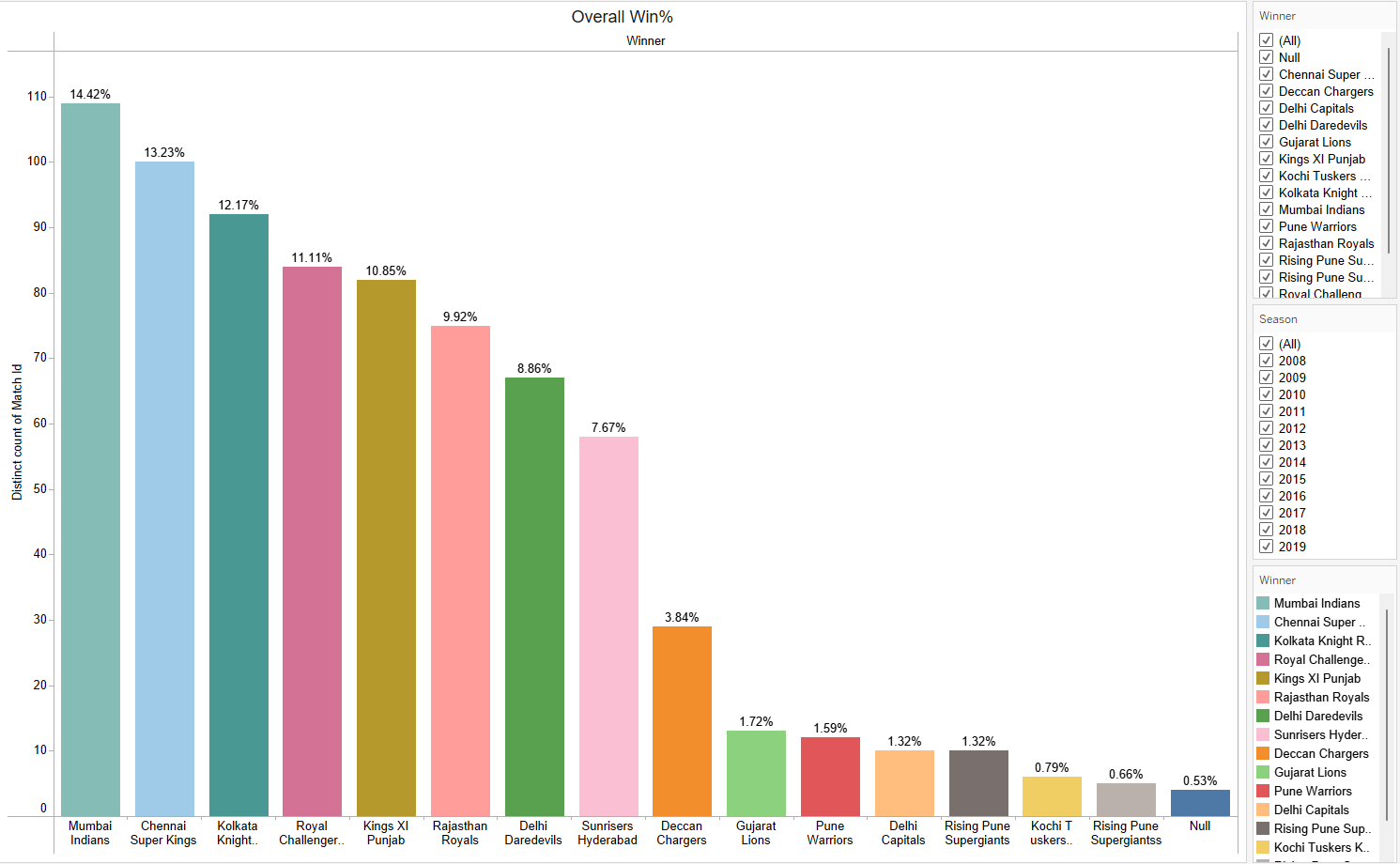
### Worksheet 3:

* This sheet shows the average by which each team won-by-wickets in each game of different seasons.



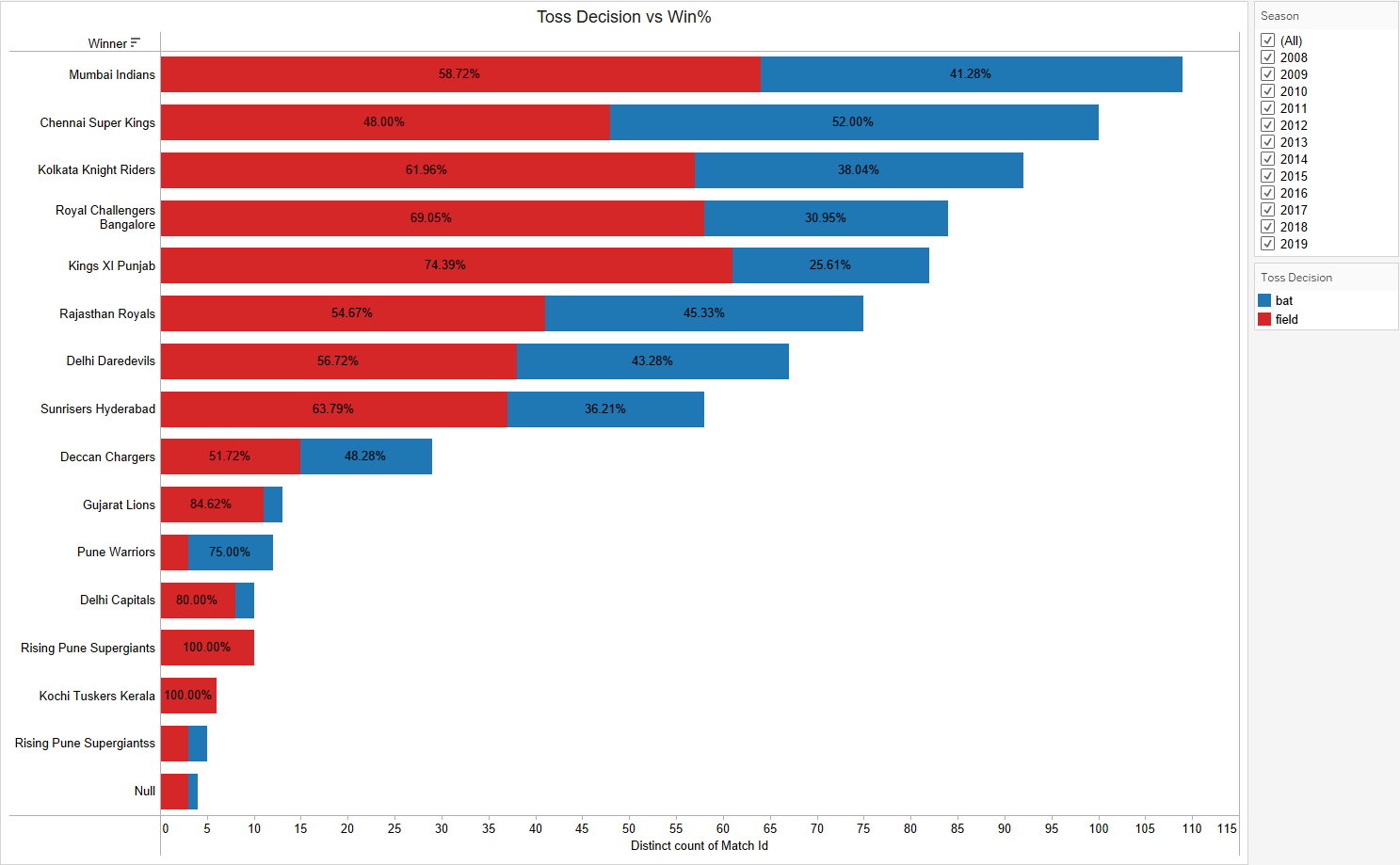
### Worksheet 4:

* This sheet shows the overall percentage of total wins by each team in a single IPL season. Here in this worksheet we can see that Team Mumbai Indians had won on an average of 14.42% matches out of all the seasons.



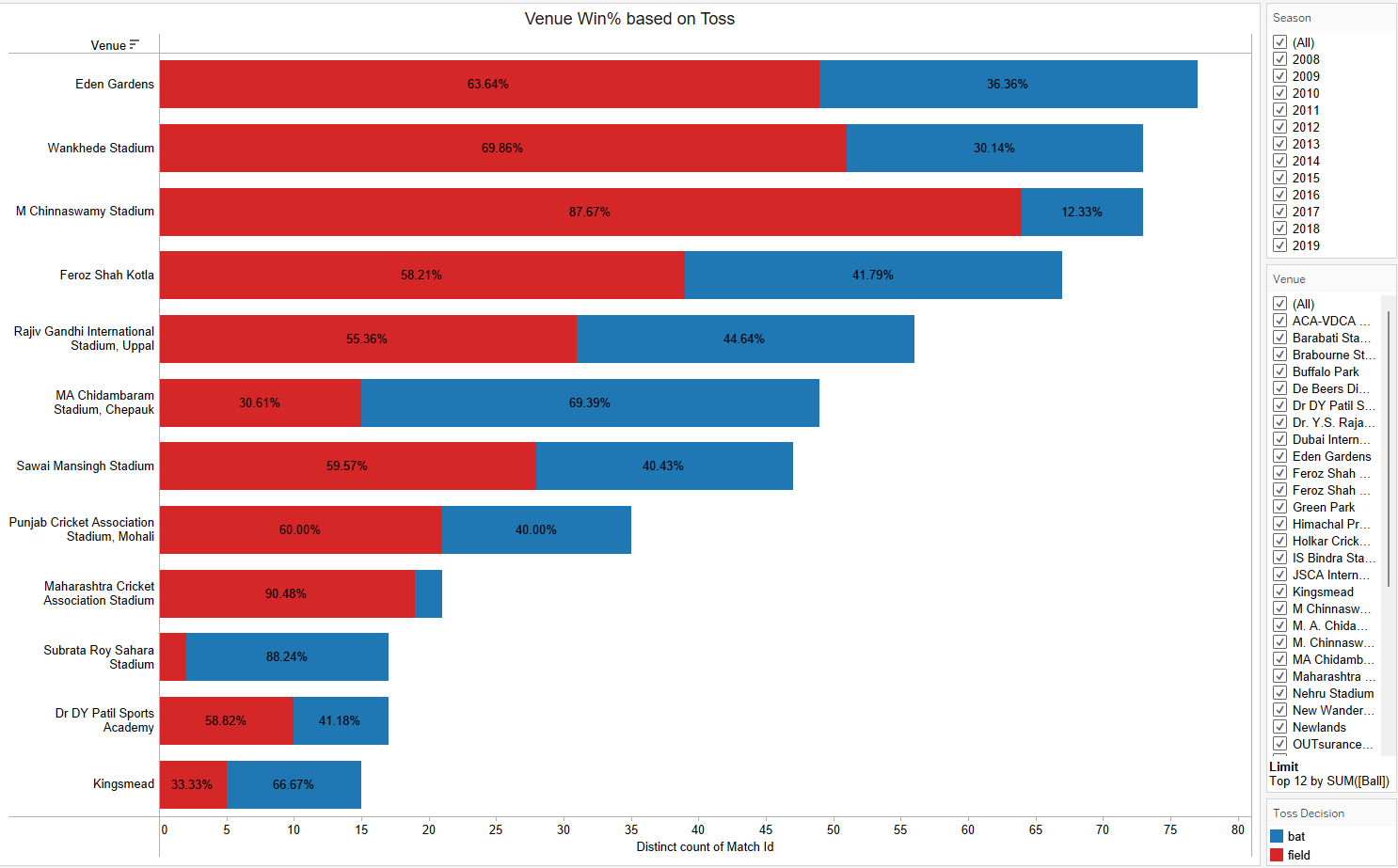
### Worksheet 5:

* This sheet shows the win percentage based on the toss decision taken by the toss winning teams, where the red indicates what percentage the team won the match bowling first whereas blue indicates what percentage the team won the match batting first .



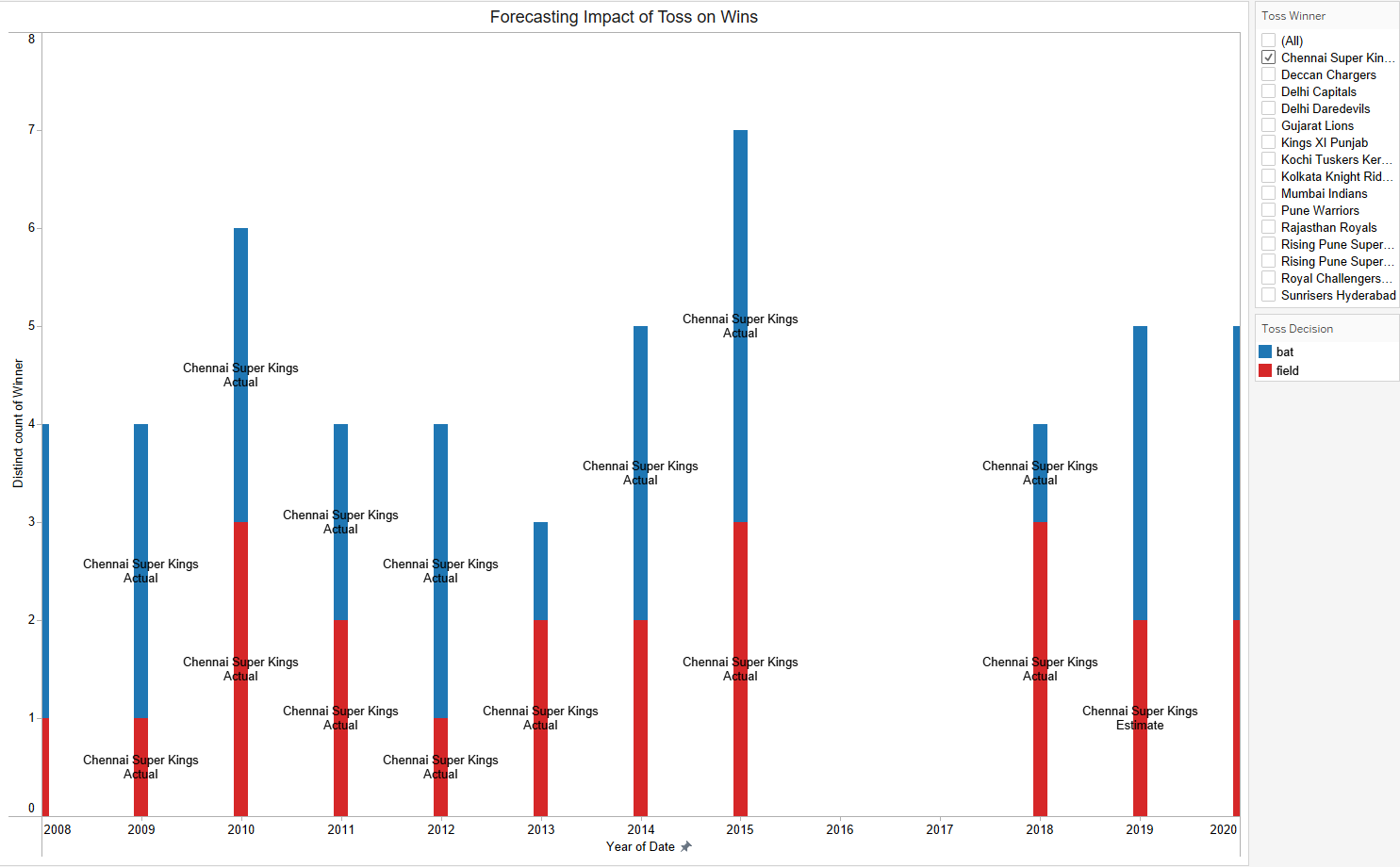
### Worksheet 6:

* This sheet shows the win percentage based on the toss decision taken by the toss winning teams on different venues where we can see that the team selecting to field first had a greater chance of winning the match at most of the venues with respect to the teams selecting to bat first.



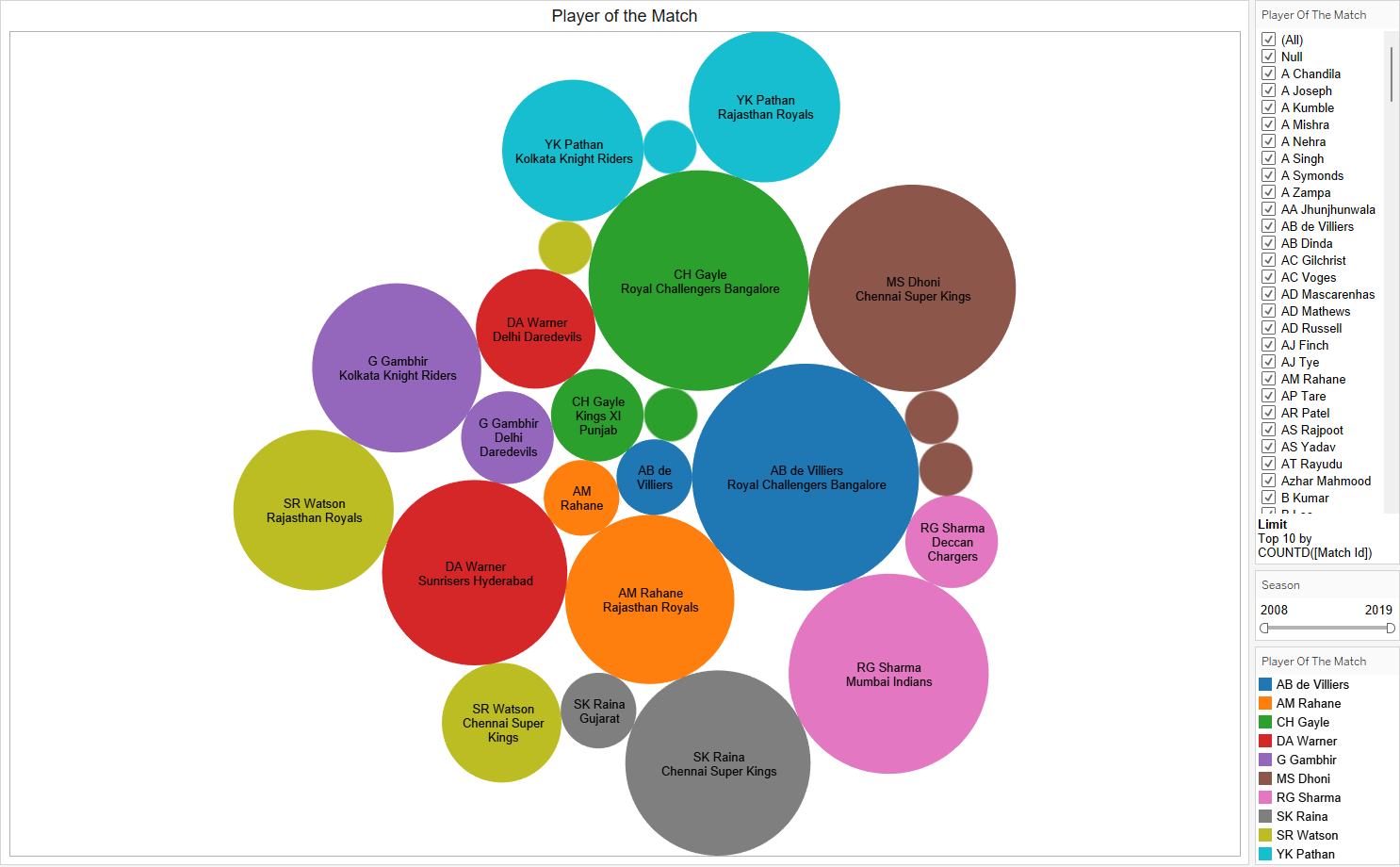
### Worksheet 7:

* This sheet shows the number of wins based on the toss decision taken by the toss winning teams in different IPL seasons.



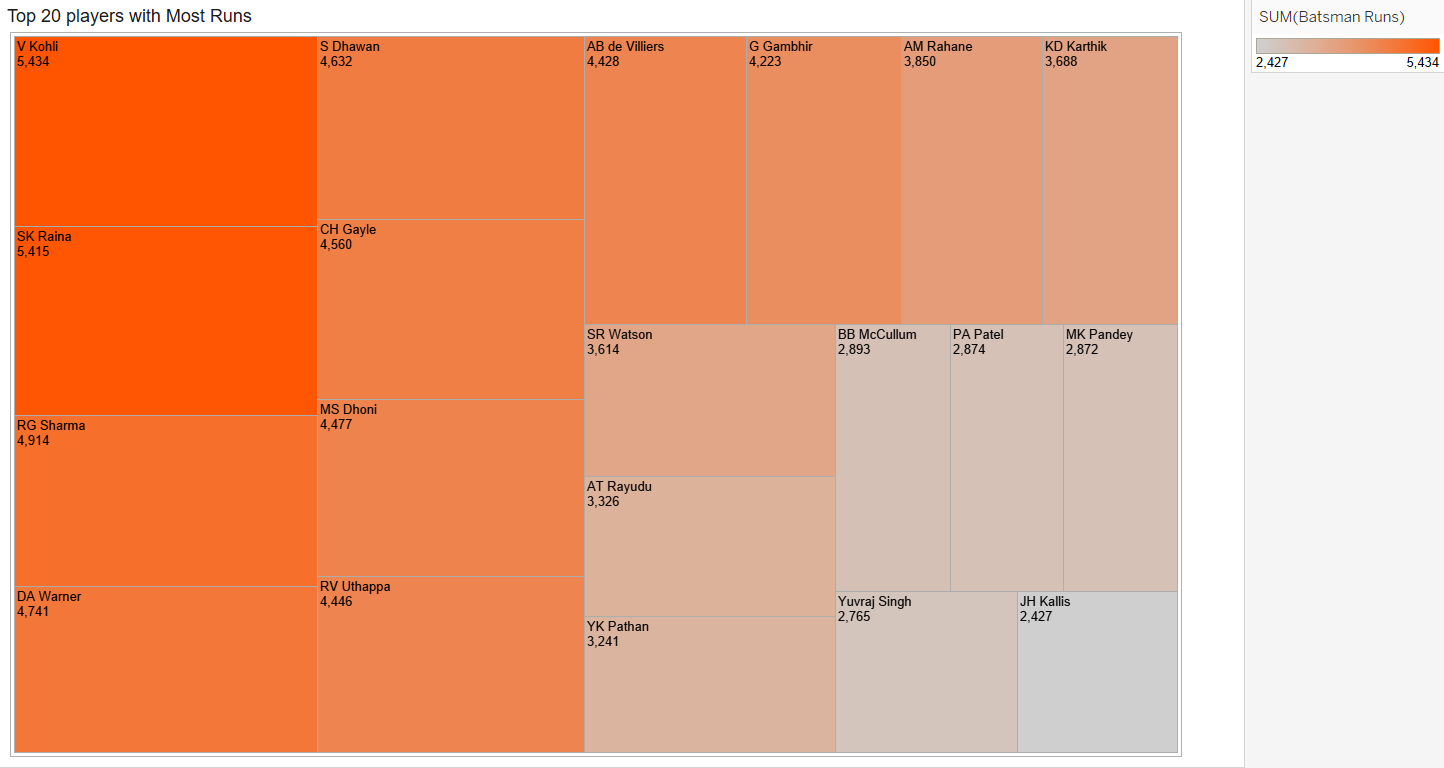
### Worksheet 8:

* This sheet shows the top players winning the most Man of the Match (MOTM) awards for their teams.



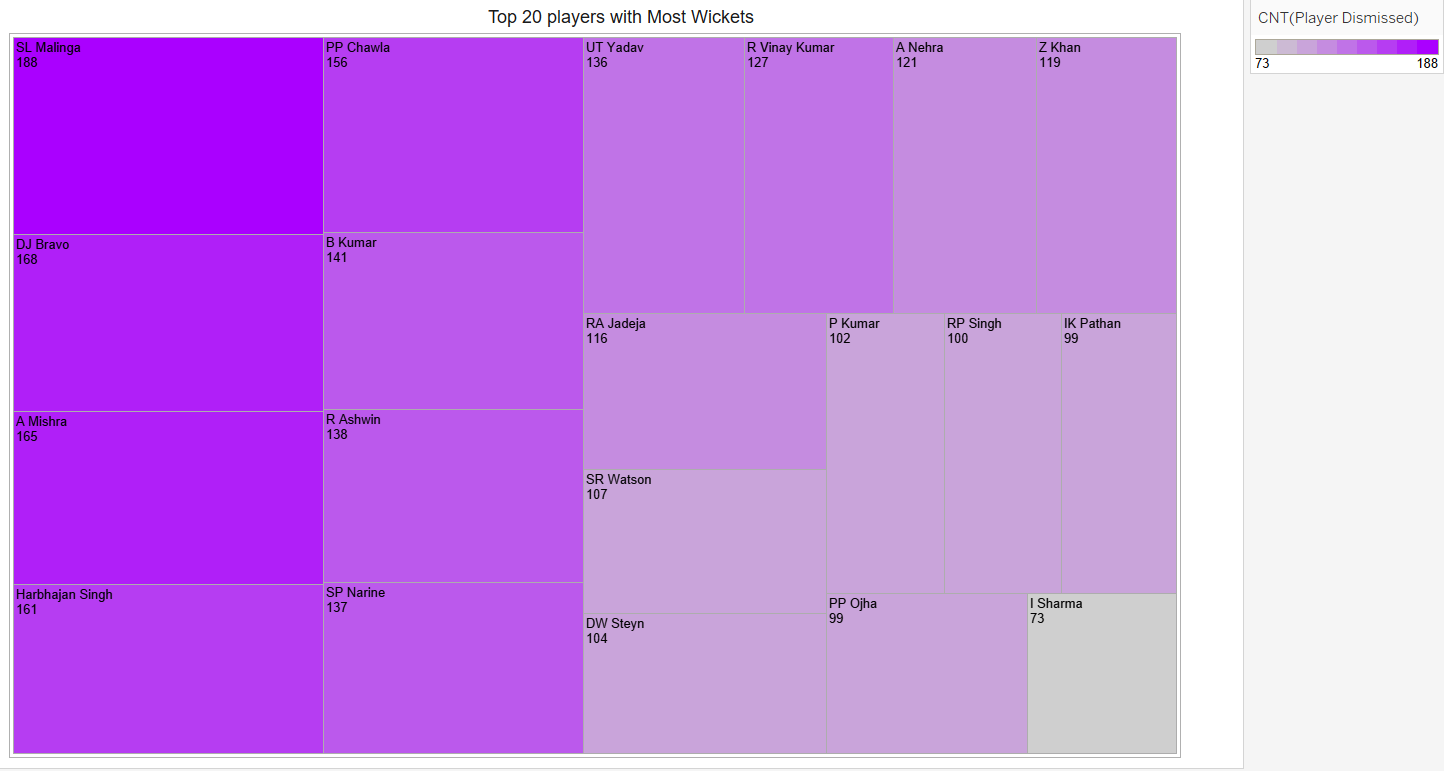
### Worksheet 9:

* This sheet shows the top 20 players with the most number of runs throughout all the IPL seasons where V Kohli has the most number of runs with 5434 runs in total.



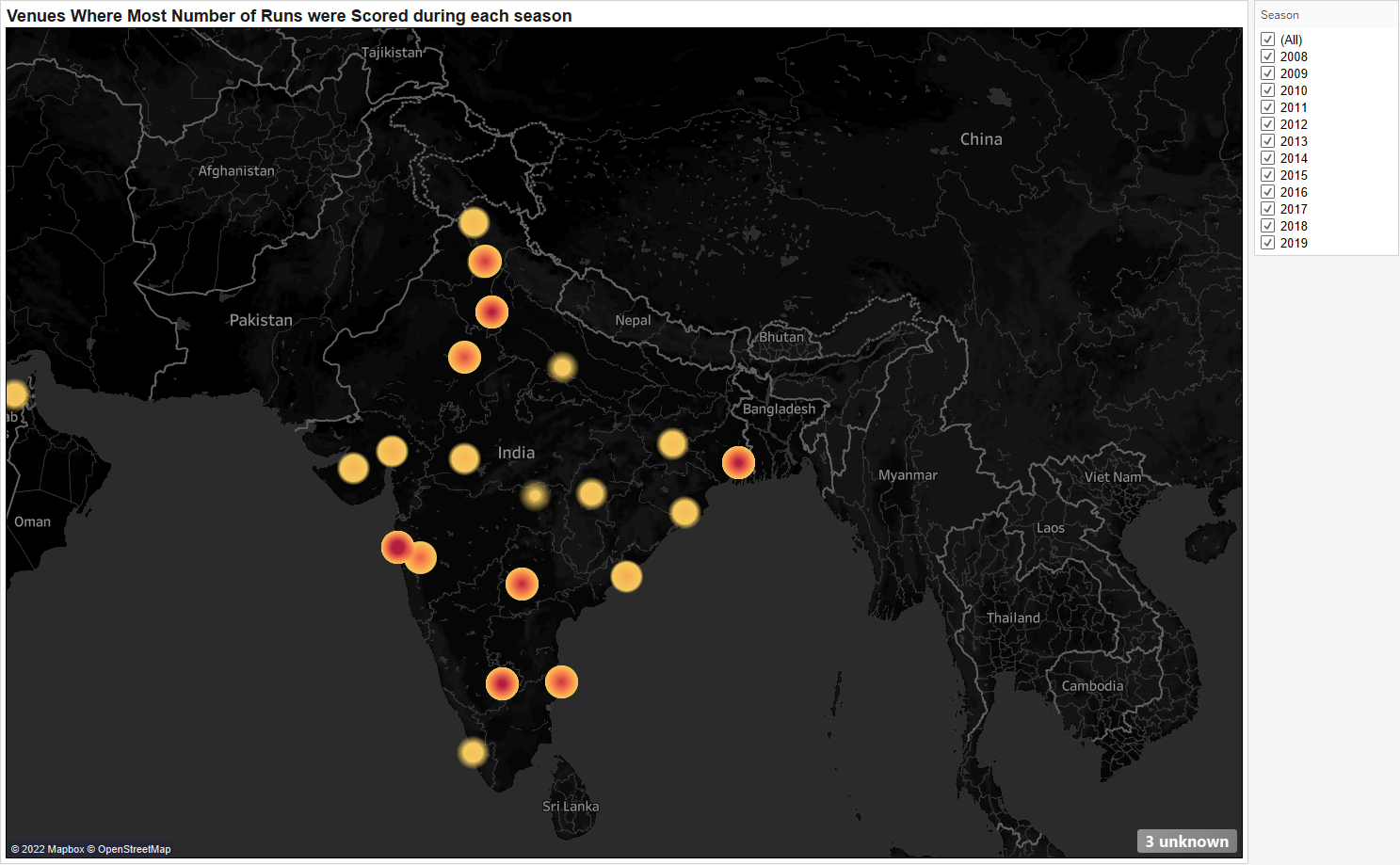
### Worksheet 10:

* This sheet shows the top 20 players with the most number of wickets throughout all the IPL seasons where SL Malinga has the most number of wickets with 188 wickets in total.



### Worksheet 11:

* This sheet shows the venues where the most number of runs were scored with respect to matches played on each venue.



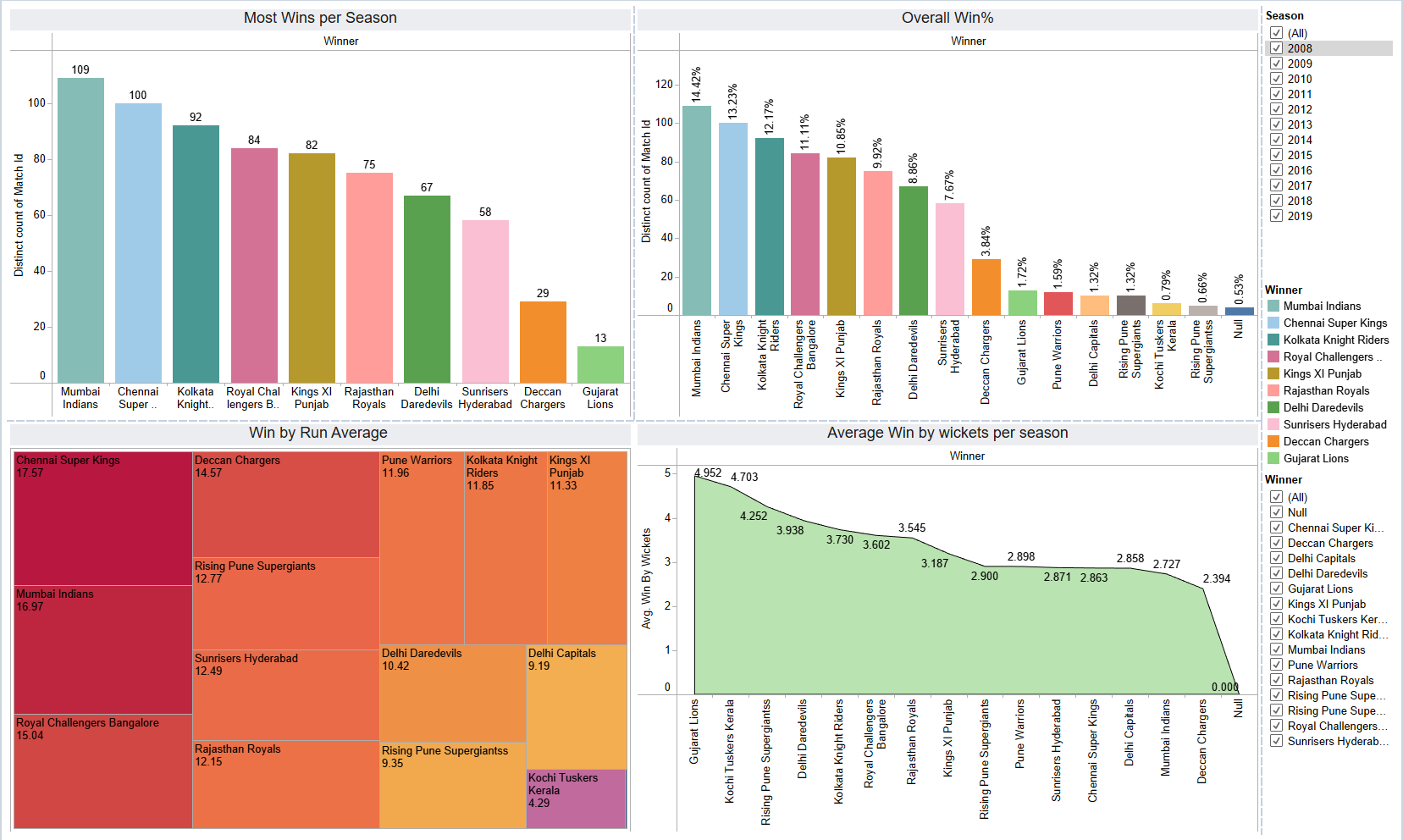
### 

### 

## Dashboards

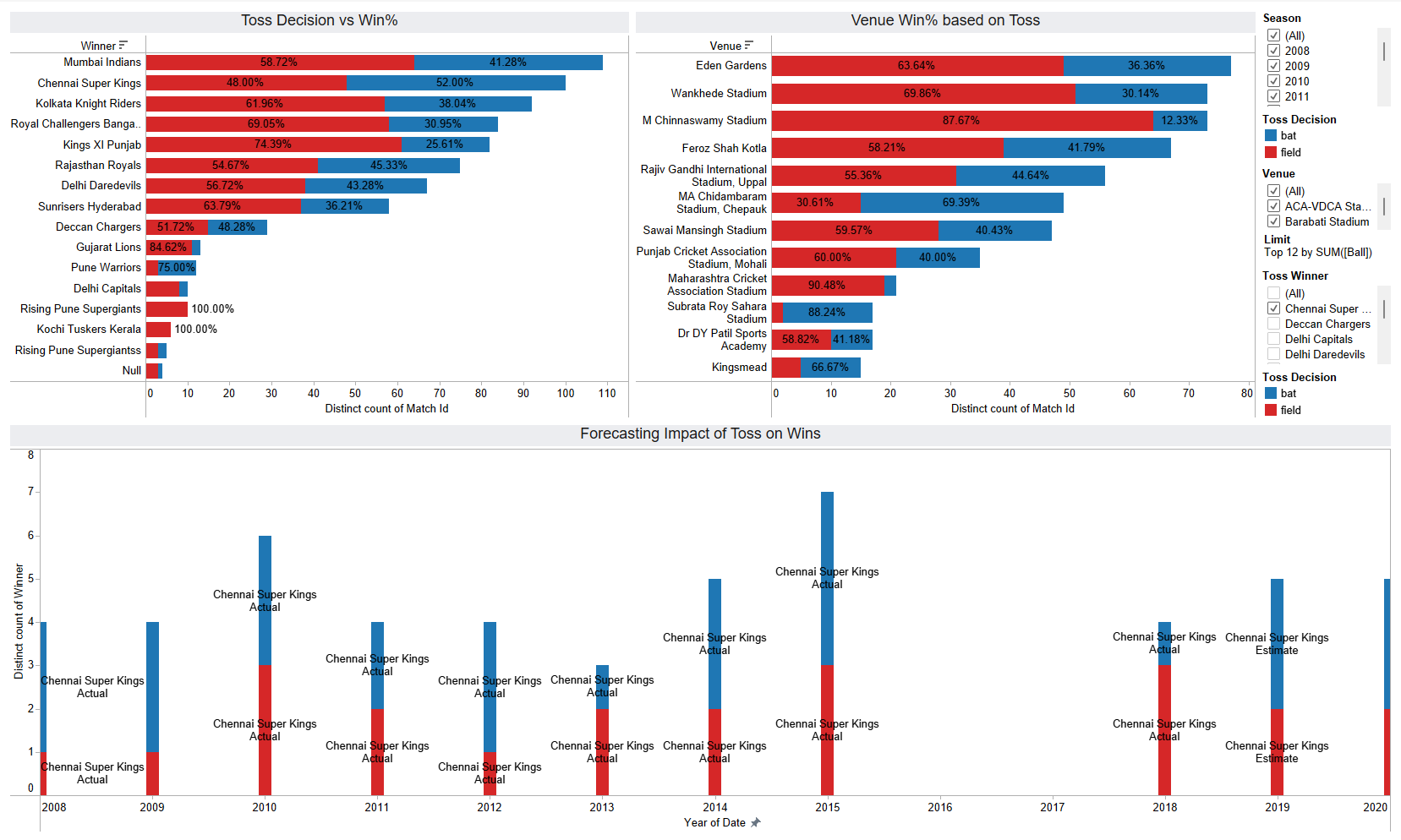
### Dashboard 1:

* This Dashboard shows the different match statistics of the teams playing in the IPL.



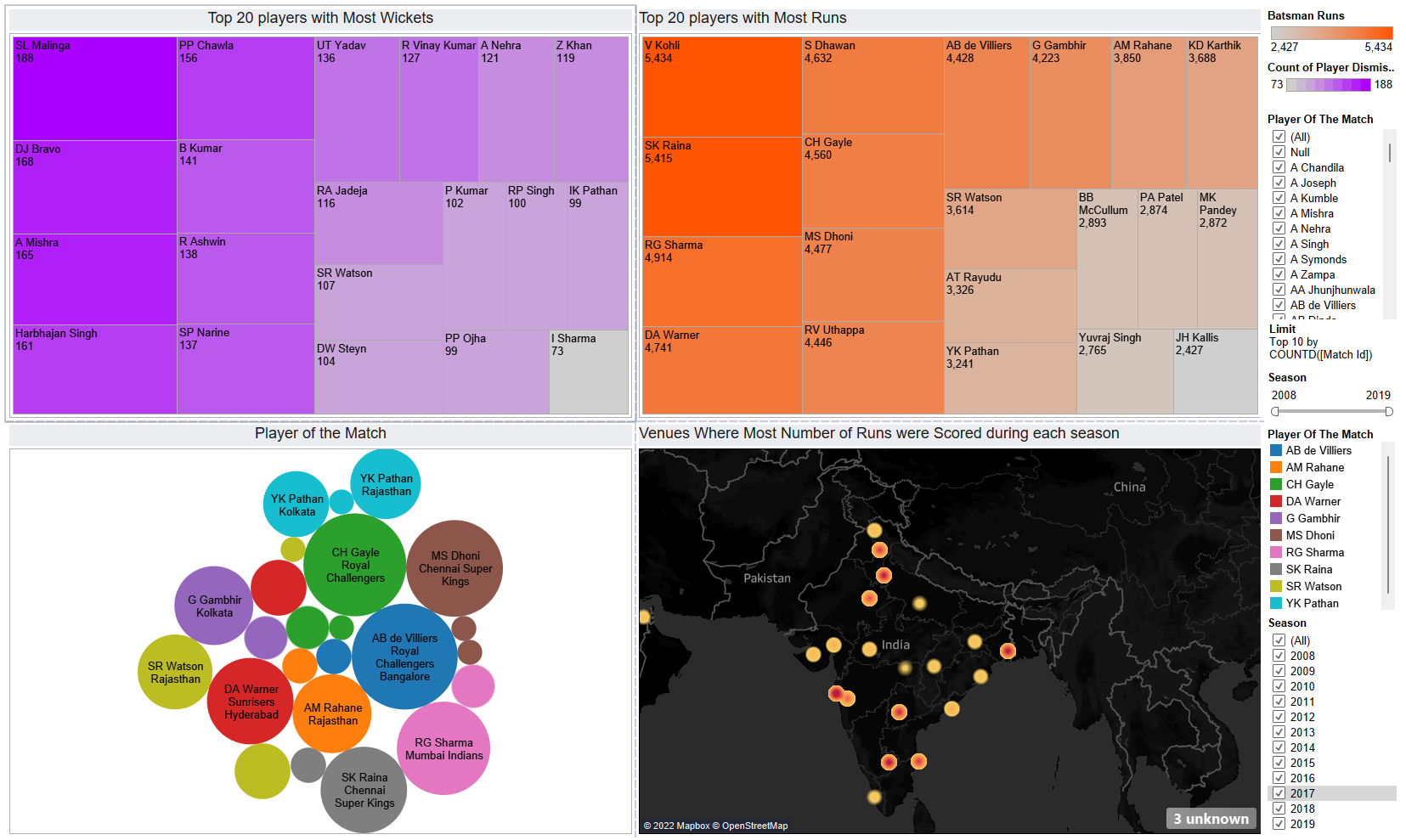
### Dashboard 2:

* This Dashboard shows the different match statistics of the teams playing in the IPL based on the toss decisions.



### Dashboard 3:

* This Dashboard shows the different match statistics of the different players playing in the IPL.



### 

## 

# Use Cases

* IPL Auction : Analysis of different player performance
* IPL Team Selection
* IPL performance analysis
* Analysis for Die-Hard Cricket Fans

# Conclusion

From the 2008–2019 IPL season, the performance of cricket players (batsmen) and toss-related information has been displayed. The team owners and selectors can identify stronger players by learning the secret variables, trends, and traits that influence a cricket match's result. This report examines player performance, particularly that of the batters, and discusses the analysis completed for the Top 20 Bowlers with the Most Wickets, IPL's top Players with the Most Runs, and Highest Man of the Matches.The experiment examined data from roughly 600 games, including toss-related analyses such the number of toss victories, the decisions made by each team after winning the toss, the toss decision season-by-season, and the toss decision team-by-team. The Indian batters are exceptionally skilled and the selectors' first pick, according to the aforementioned study. SK Raina is regarded as one of the best batters and is second on the list of batsmen with the most runs, with V Kohli taking the top spot with the most runs while on the bowling side SL Malinga has been named as the most effective wicket taker and one of the finest bowlers.

# Future Goals

* Data Visualization on IPL Auction.
* Visualization of player performance on the basis of their physicality.
* Ground analysis and pitch condition analysis

## REFERENCE

1. https://www.kaggle.com/datasets
2. https://public.tableau.com/app/discover