**Table of Contents**

1. [INTRODUCTION 4](#_Toc33120505)
2. [BACKGROUND 5](#_Toc33120506)

[2.1 Basics of Cricket 5](#_Toc33120507)

[2.2 Introduction of IPL and Its Structure 6](#_Toc33120508)

[2.3 Need and Significance of Research 8](#_Toc33120509)

1. [THE SCOPE OF THE PROJECT 8](#_Toc33120510)

[3.1 Project objectives 8](#_Toc33120511)

[3.2 Project questions 9](#_Toc33120512)

[3.3 Project design 9](#_Toc33120513)

1. [IMPLEMENTATION AND METHODOLOGY FOR CATEGORIAL ANALYSIS 10](#_Toc33120514)

[4.1 Toss impact on match result: 10](#_Toc33120515)

[4.2 Finding Highest Run Scorer Batman for the team 11](#_Toc33120516)

[4.3 Selecting one of the reliable batting partner (who faced most balls) 12](#_Toc33120517)

[4.4 Selecting one of the most destructive batsman’s (More runs in less balls) 13](#_Toc33120518)

[4.5 Selecting a batsman with highest average runs (More runs in less balls) 14](#_Toc33120519)

[4.6 Selecting one of the most not out batsman to keep wickets for long 15](#_Toc33120520)

[4.7 Selecting one of the batsman based on his performance at specific ground 16](#_Toc33120521)

[4.8 Finding outa batsman who has hit more boundaries and sixes (optional) 18](#_Toc33120522)

[4.9 Select an individual batsman by making whole card of their performance 20](#_Toc33120523)

[4.10 Selecting the most wicket taking bowler 20](#_Toc33120524)

[4.11 Select the most economical bowler 21](#_Toc33120525)

[4.12 Finding out the best strike rate bowler (Less strike rate will be better) 23](#_Toc33120526)

[4.13 Now calculate average of bowler (Less average will be better) 23](#_Toc33120527)

[4.14 Let's pick the bowlers for different phases of match: 24](#_Toc33120528)

[4.15 Top Bowlers Based on the stadium 26](#_Toc33120529)

[4.16 Finding out bowlers with maximum number of dot balls 30](#_Toc33120530)

[4.17 Selecting the best wicket keeper out of available options 30](#_Toc33120531)

[4.18 Finding out the best fielders on the field 32](#_Toc33120532)

[4.19 Finding out the game changes (Man of the match) 33](#_Toc33120533)

[4.20 Umpire Analysis (important when you are thinking to take review) 33](#_Toc33120534)

1. [CHALLENGES FOR THE STUDY 35](#_Toc33120535)
2. [SUMMARY OF REQUIREMENT 36](#_Toc33120536)

[6.1 Prerequisite skills and knowledge 36](#_Toc33120537)

[6.2 General hardware requirements 36](#_Toc33120538)

[6.3 General software requirements 37](#_Toc33120539)

1. [CONCLUSION 37](#_Toc33120540)
2. [REFERENCES 38](#_Toc33120541)

**Index**

**Title**

**Acknowledgement**

**Declaration**

**Certificate**

**Contents**

Abstract

List of Table

List of figures

1. Introduction- Game Cricket, IPL, Aims and Method outline/statistics and team selection

1.1-Basics of Cricket

1.2-Introduction of IPL and It’s Structure

1.3- Need and Significance of Research

For later use:-

1.4-Involvement of different States/Cities in IPL

1.5-Involvement of Foreign and Indian players in IPL

1.6-How does IPL works financially.

1. Background-

2.1-Importance of Cricket Analysis

2.2-Sports Analytics

1. THE SCOPE OF THE PROJECT & Literature Review

The listed below objectives give an overview of the research:

-Project objectives

-project questions

-project design

1. Preliminary Data Analysis
2. Methodology -> master\_batool.pdf
3. Analysing & exploring Individual Category for team selection
4. Use of the Team selection analysis or research objectives
5. Challenges for team selection and league
6. Conclusion/Summary

Introductions

In this study, we used Standard Deviation to predict batsmen and bowlers’ performances in a match. We predict how many runs a batsman will score and how many wickets a bowler will take in the upcoming match. We targeted both the problems as classification problems where we classified runs and wickets into different ranges. We experimented with four supervised machine learning algorithms and compared their performance. The models generated by these algorithms can be used to predict the players’ performance in future matches.

Appendixes

Bibliography

References

**Inclusion of standard deviation**

**Content**

Acknowledgements-

I am fortunate to have had two exceptional people help me through this thesis that I need to thank. The first, my supervisor, Dr Vijay Babu Gupta with whom I’m very grateful to have been taught, guided and inspired by in the pursuit of knowledge. He has been supervising my work throughout the process and led me to the right direction with valuable and informative discussions.

This certainly would not have been possible without his assistance, and I could not express how thankful I am in this small space.

Lastly, I would like to thank my family for their encouragement and support.

**Cricket Analysis based on IPL Statistic.**

#### Chapter 1

# INTRODUCTION

In this thesis we are trying to select the winning team for winning an IPL Match .For winning any match ground specifications toss decision etc. also plays important role besides an individual player performance. The Performance of the players varies with the opposite team they play with and ground on which they play the match. Selection of the player is very important as whoever player is selected in the team of 11 players in the beginning of the match are fixed besides if there is any case of injury, but the substitutes have limited privileges, in order to predict any players performance, their past statistics and characteristics need to be analyse

Selecting the best team to win a match in IPL involves different things to be considered as 11 players who will part of the match will fall into different categories.

The categories are like bowling, bating, wicket keeping and fielding. Also Players’ performance varies with the team they play against and the ground on which they play the match. Selecting the best players for a particular match in any sport involves predicting the players’ performance.

We will also be considering other factors like ground conditions, toss decision and players performance on specific ground along player’s fitness level for individual matches.

Cricket players’ abilities and performance can be measured in terms of different stats. Batsmen’s statistics include batting average, batting strike rate, number of centuries etc. Whereas bowlers’ statistics are measured by bowling average, bowling strike rate, economy rate etc. Other characteristics of batsmen include, batting hand of the batsman, the position at which the batsman bats etc. and those of bowlers include, the type of bowler, bowling hand of the bowler etc.

Moreover, recent performances of the batsman/bowler, the performance of the batsman/bowler against a particular team and the performance of the batsman/bowler at a given venue are also taken into account for predicting his performance in the upcoming match. The team management, the coach and the captain utilize these facts and their own experience to select the team for a given match

Cricket is a game of numbers. The very core of the sport is entwined with numerical values that translate ultimately to a match result. These sport statistics are a natural by-product of competitive sport and have been around along as contested sport has existed. Currently sport reporters and commentators bombard observers with a vast array of numerical values designed to describe an individual's performance at a particular skill. These added extras contribute to the entertainment value of professional sport.

#### Chapter 2

# BACKGROUND

## 2.1 Basics of Cricket

The Game Cricket started from British Era in India and kept on involving different countries/teams across the world. It is started from limited days to limited overs match and now from 60 overs match to 50 and 20 overs matches. It is commercially played within the countries and across the countries and now as it is got a lot famous around it is getting more commercialized.

It is played by two teams with each side having eleven players. One team bats and the other team bowls (fields) at a time and one such session is called an innings. In the centre of the field, there is a 22-yard long pitch where most of the action takes place. Both ends of the pitch will have a wicket which has three wooden stumps and two cross pieces called the bails. Each team consists of batsmen, bowlers and a wicket-keeper.

All the players from the bowling team are on the field; one of them is behind the wickets, one of them bowls (throws the ball) from one end of the pitch and the other players are fielding, arranged in a particular fashion decided by the captain of the team. Two players from the batting team are on the field, alternating batting at a time. One of them bats from one end while the other one waits at the other end where the bowler is bowling from. The batsmen can be dismissed in many different ways with each ball bowled and this is called a wicket. As at a given time, there need to be exactly two batsmen on the field, the batting team has 10 wickets at the beginning of their innings. The batting team has to defend their wickets and score maximum runs possible and the bowling team has to get wickets as soon as possible and restrict the batting team from scoring runs. The team scoring the most runs wins at the end of the match.

Cricket is usually considered as representation and symbol of tradition with slow evolution, but it has adopted reasonable changes to survive and sustain its growth over last 50 years (Oliver 2018). It would not be wrong to say that cricket had to adopt reasonable amount of changes to survive and grow in this faster and more competitive world of sport.

## 2.2 Introduction of IPL and Its Structure

**IPL Overview**  
Indian Premier League (IPL) tournament is the league of the cricket of India. The Indian Premier League, established by the Board of Control for Cricket in India (BCCI) in 2007. The premier league is the shortest type of Cricket. A T20 match only has 20 over’s per inning. Indian and International Players take part in IPL and contributing there what is the world’s “richest cricket tournament”. IPL tournament is basically a cricket tournament and in this tournament all team plays exact two matches with all other teams one on the home town stadium and second on opponent home town. The Indian Premier League (IPL) is a professional league for Twenty20 cricket championship in India. It was initiated by the Board of Control for Cricket in India (BCCI), headquartered in Mumbai and is supervised by BCCI Vice President Rajeev Shukla, who serves as the league's Chairman and Commissioner.

**Tournament Format**

Currently, with eight teams, each team plays each other twice in a home-and-away round-robin tournament. At the conclusion of the league stage, the top four teams qualify for the finals. The top two teams play each other in the first preliminary round match, with the winner going straight to the final while the loser goes on to the second preliminary round. The third and fourth place teams play each other to decide who will take on the loser from the first preliminary round match. The winner of the second preliminary round match will move onto the final where the winner will be crowned the Indian Premier League champions.

**Local and Foreign Players**

The team can concise only 4 foreign nationals and 7 Indian players.

The players can be international player or local players from different countries who have shown their performance over a number occasions/tournaments and have been eye catchers for team selectors. They are selected through IPL bidding process with a cap amount and highest bid for individuals.The process of bidding is very simple that starts with a initial cap for player and increased with the bid.

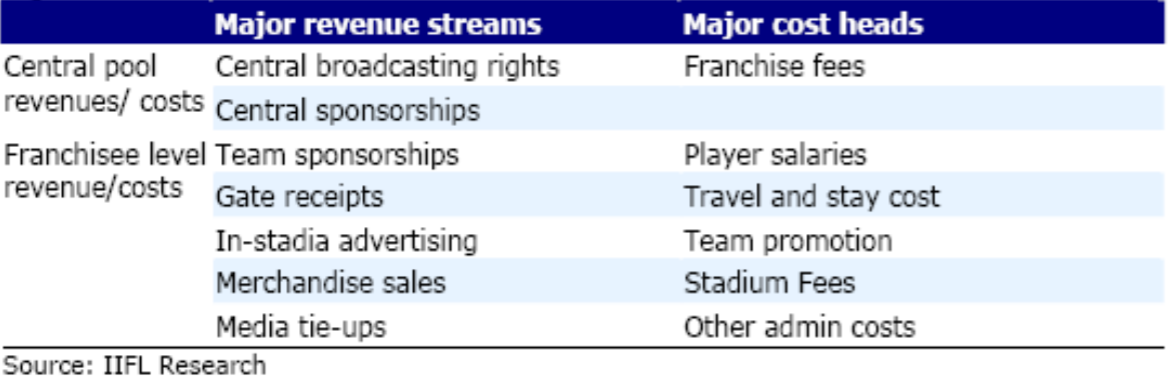
**Revenue Model**

IPL has two revenue streams and these are central and decentral streams. In the central systems, sponsorships, franchises and broadcasting arrangements provide revenue. For 2012, the IPL title was bought by Pepsi for 400 Crores INR while other sponsors such as Vodafone, Yes bank, Star Plus became the sponsors. McDowell's became the Umpire Sponsor while Sony Max, Sony Six and Times Internet purchased the broadcasting rights. The exact amount paid is not clear and while the estimated amount paid is large, some key sponsors such as DLF, the title sponsors for earlier games and Hero MotoCorp have withdrawn. In any case, until 2017, the IPL body would keep 40% of the central revenue stream. The balance 54% is shared among franchises while 5% is given as prize money. In the de-central stream, franchisee owners sell ad space on team uniforms and other ad material, TV displays to other firms. In 2012, Aircel, one of the leading mobile phone service provider paid 85 Crores INR to Chennai Super Kings for a three years deal. Delhi Dare Devils has Muthoot Group as the main sponsor while Mumbai Indians had Hero Honda for three years. Please refer to the following figure that illustrates the revenue model

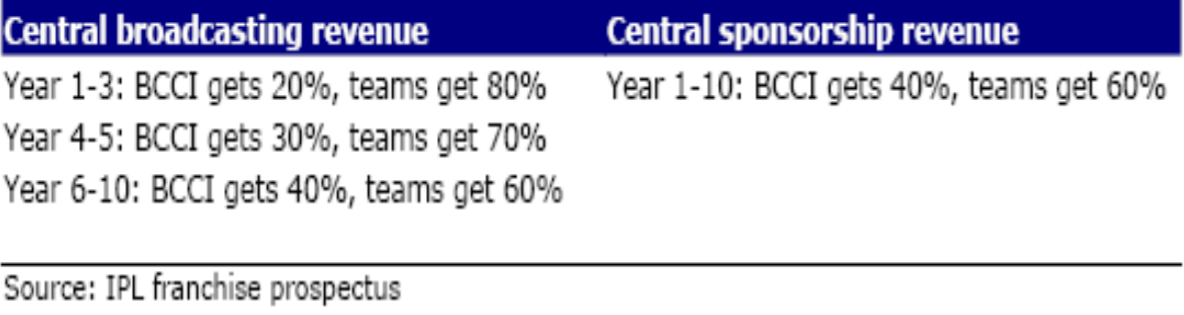
|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

|  |  |
| --- | --- |
|  |  |
|  |  |

**Revenue and Major Cost Heads**



**Figure 2.1: Revenue and Major Cost Heads**



**Figure 2.1:** Broadcasting Revenue

**Franchisee Revenue**

* Media Rights- To be shared equally amongst franchisees after removing IPL’s share. Sponsorship rights (IPL Sponsors)- 60% of the amount collected to be distributed equally amongst the franchisees.
* Local Sponsorships- Team sponsorship revenues
* Gate Receipts- Are anticipated to be a major source of revenues. 20% of tickets are to be allocated to IPL.

## 2.3 Need and Significance of Research

* In order to get winning results, team selection plays an important role. With the help of this study we are to get the best selected team which will result in winning any match. We have worked on various factors in this study so as to get best team selected.
* Selecting team players based on their past performances and their specific opponents and grounds along with consideration of other factors like ground and toss decision so that we can convert result into wining for us.
* Ability of the player as no one benefits from a player playing in a team in which he is either too good or simply not able to contribute.
* Form of the player, it is critical factor but not the only one. A player’s performance will be assessed over a number of matches, not just a single performance and the quality of opposition will also be taken into account.
* Balance in Team both in relation to the spread of batsmen and bowlers in the team as well as the type of batsmen, bowlers, fielders and wicket keeper.
* Batsman and Bowler have different statistics based on their qualities like:-Batsmen’s statistics include batting average, batting strike rate, number of centuries etc. Whereas bowlers’ statistics are measured by bowling average, bowling strike rate, economy rate etc. Other characteristics of batsmen include, batting hand of the batsman, the position at which the batsman bats etc. and those of bowlers include, the type of bowler, bowling hand of the bowler etc. we have considered all these characteristics to find out best bowler and batsman for different categories.

#### Chapter 3

# THE SCOPE OF THE PROJECT

## 3.1 Project objectives

The listed below objectives give an overview of the research:

* To get the list of best batsman with different quality of batting like:-hitting lot of runs in few balls (String rate), keeping the wickets for long term, ability hit runs in most of the matches (Average) and selecting different batsman for different position.
* To get the list of bawlers with different qualities like:-bawling many dot balls, taking many wicket, balling opening spells, middle overs and death overs etc.
* Choosing a wicket keeper who can keep wickets properly and hit some runs when needed.
* Considering fielding along with individual’s skill set for his department as fielding saves some runs for the team and as said catches win matches.

## 3.2 Project questions

The listed below questions we are going to answer with this project:

* Does toss wining matters a lot?
* Does bawling first or batting first is favourable on specific stadium?
* Any specific team getting advantage of winning toss?
* Who is highest run scorer in the tournament?
* Who has faced many balls?
* Who is most destructive batsman?
* Most not out batsman
* Batsman with good average in the tournament
* Top batsman based on the stadium
* Batman with most no. of boundaries and sixes
* Most wicket taking bawlers
* Most Economical bowler
* Bowlers with good strike rate
* Bowlers with good average
* Preferred bowlers for different phases of match like power play, middle overs and death overs.
* Top Bowlers Based on the stadium
* Other information like most dot balls and extra runs
* Wicket keeper analysis with most stumping and catches
* Fielder analysis with most run outs and catches
* Most important player for the match - Man of the match
* Player's performance throughout
* Whole report card for batsman and bowlers
* Importance of having good umpires

## 3.3 Project design

The project design will let us know that how we are going to address above questions.

* As answering the above questions we need to do some analysis and for that analysis we need some data to be there, so I collected the data from some online sources.
* After collection of data the next task would be to answer above questions that will need to be calculated through some tools.
* Here I will be going to use some python language for data processing and visualisation with some important inbuilt functions which helps me to answer these questions. I will be using jupyter notebook for that python code as editor.
* There are some formula’s I will be using to calculate average of batsman and bowlers, Strike rate of batsman and bowlers along with some data science concepts like standard deviations and all.
* Most of the time will be given for data cleaning and extracting out some important relevant data from the bunch of files.
* Design will be focusing on extraction relevant data and visualisation of the same and for that we need to have good understanding of different charts and the way of using those from relevant language which we are going to use.

#### Chapter 4

# IMPLEMENTATION AND METHODOLOGY FOR CATEGORIAL ANALYSIS

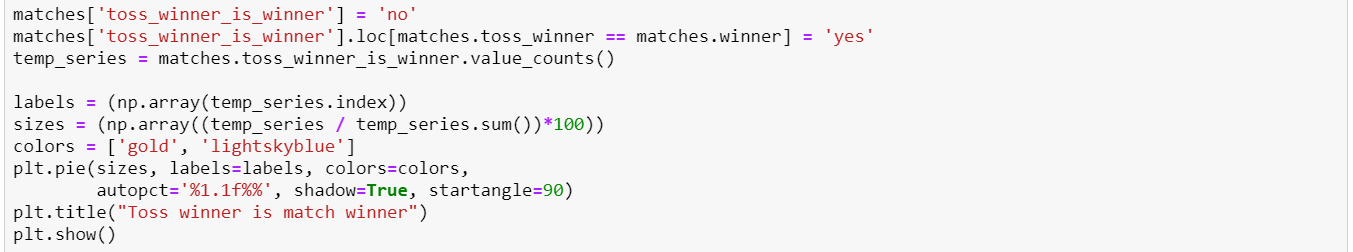
We will divide each analysis in 4 sections, first we will give some introduction of individual analysis, second we will focus on how to get the result through python code and then result of the same and at last we will conclude the analysis.

## 4.1 Toss impact on match result:

**Introduction:**

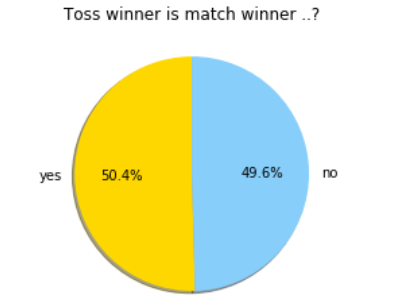
In this analysis, we are trying to find out whether winning a toss matters a lot or has any greater impact on match result.

**Methodology:**



**Figure 4.1.1: Revenue and Major Cost Heads**

**Result:**



**Figure 4.1.2: Revenue and Major Cost Heads**

**Conclusion:**

As stated above, we can deduce that there is no such big difference that we can say Toss winner is match winner. Conclusion is that winning a toss doesn’t matters to win the match.

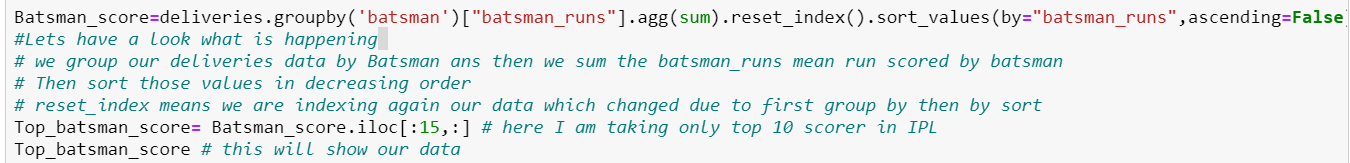
**Batsman Analysis:**

## 4.2 Finding Highest Run Scorer Batman for the team

**Introduction:**

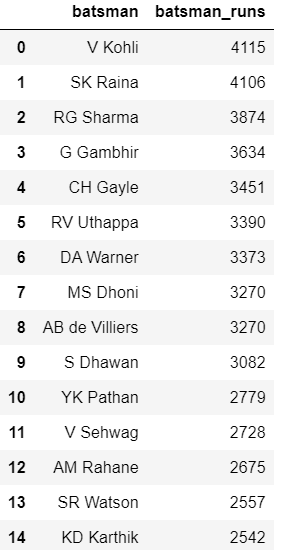
In this analysis, we are trying to find out highest run scorer of the tournament and this way we can get one of the most run getter batsman for our team and there will be mental pressure on other team when such player is part of our team. Also such batsman would cost more so we cannot get all of them so we will try to get at least one of them.

**Methodology:**



**Figure 4.2.1: Revenue and Major Cost Heads**

**Result:**



**Conclusion:**

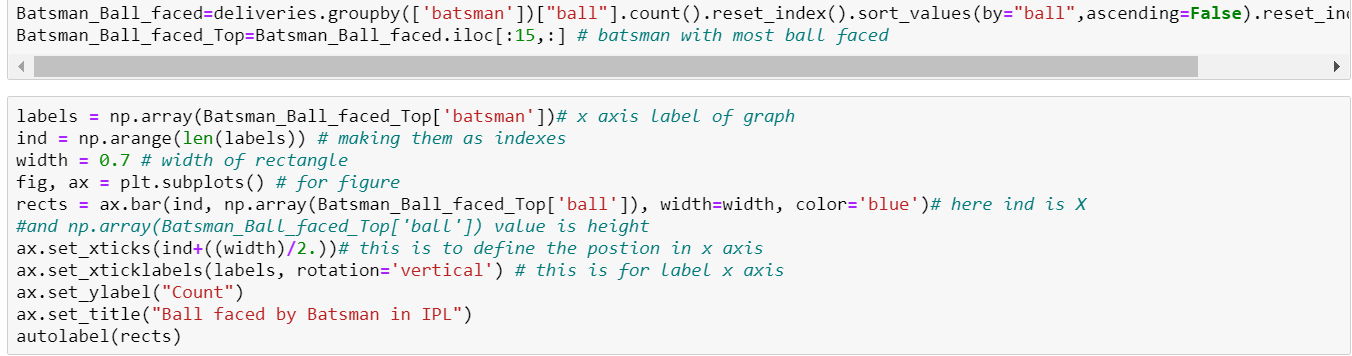
We can see now the top run scorer of the tournament and as clearly seen Virat Kohli, Suresh Raina and Rohit Sharma are the top most run scorer of the tournament.

## 4.3 Selecting one of the reliable batting partner (who faced most balls)

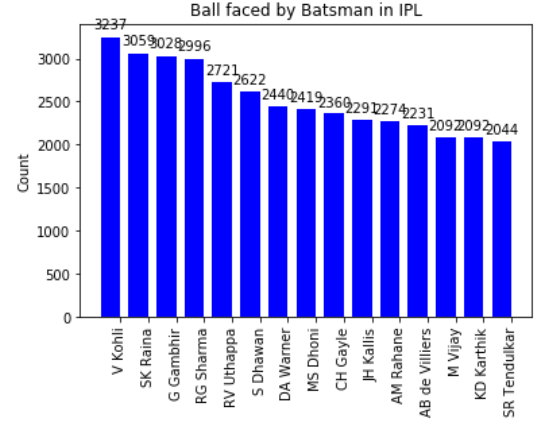
**Introduction:**

In this analysis, we are trying to find out most reliable partner by finding out the batsman who has faced many balls in the tournament and this way the partner can play without being worried of wickets from other side.

**Methodology:**



**Result:**



**Conclusion:**

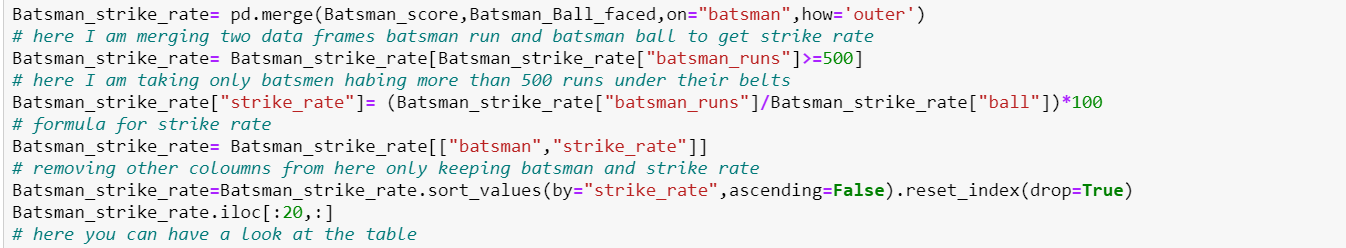
We can now see that players those are in the list of highest run scorer are among the most reliable batsman in the list.

## 4.4 Selecting one of the most destructive batsman’s (More runs in less balls)

**Introduction:**

In this analysis, we are trying to find out most destructive batsman by finding out the batsman who has faced very few balls and hit a lot of runs in the tournament and this way the partner can play without being worried of wickets from other side.

**Methodology:**



**Result:**



**Conclusion:**

We can see now that players those are in the list of highest run scorer are among the most reliable batsman in the list.

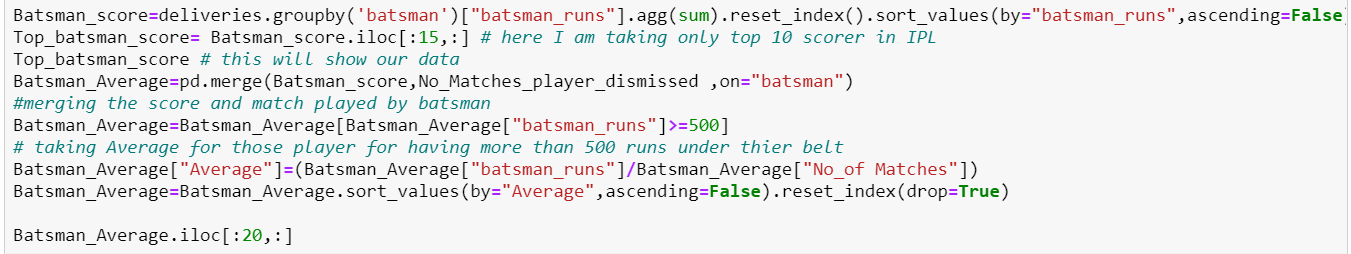
## 4.5 Selecting a batsman with highest average runs (More runs in less balls)

**Introduction:**

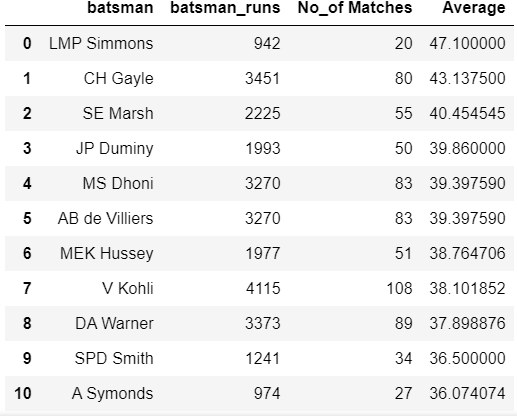
In this analysis, we are trying to find out most destructive batsman by finding out the batsman who has scored well in most of the matches.

**Methodology:**





**Result:**



**Conclusion:**

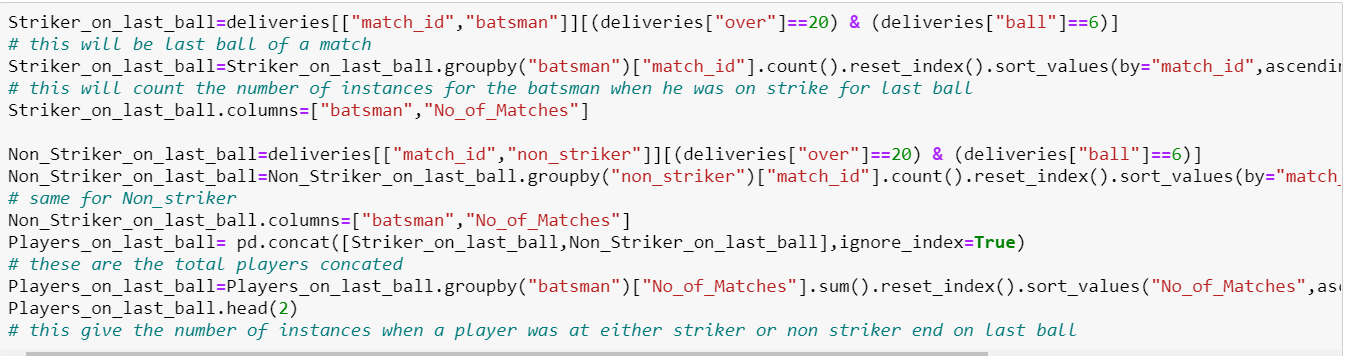
We can see now that Simmons and Gayle has good average runs we can select at least one of the batsman from the list get some assured runs on the board.

## 4.6 Selecting one of the most not out batsman to keep wickets for long

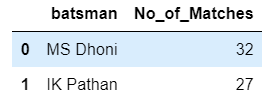
**Introduction:**

In this analysis, we are trying to find out the batsman who has been not out for most of the time and be behave as a wall for the batting team.

**Methodology:**



**Result:**



**Conclusion:**

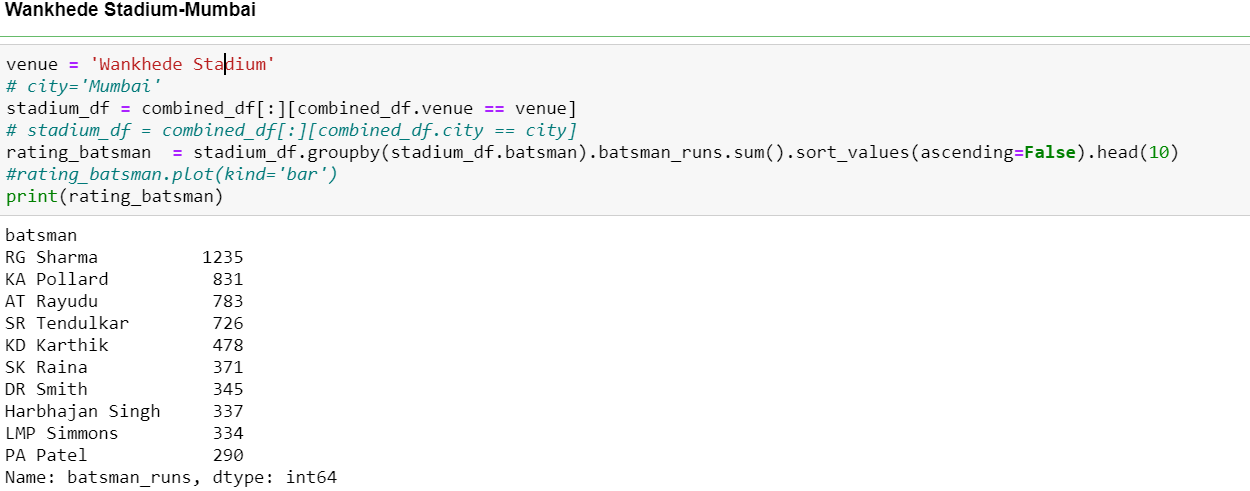
As we have seen Dhoni keeping the wickets for long time and the same thing we are getting through this analysis.

## 4.7 Selecting one of the batsman based on his performance at specific ground

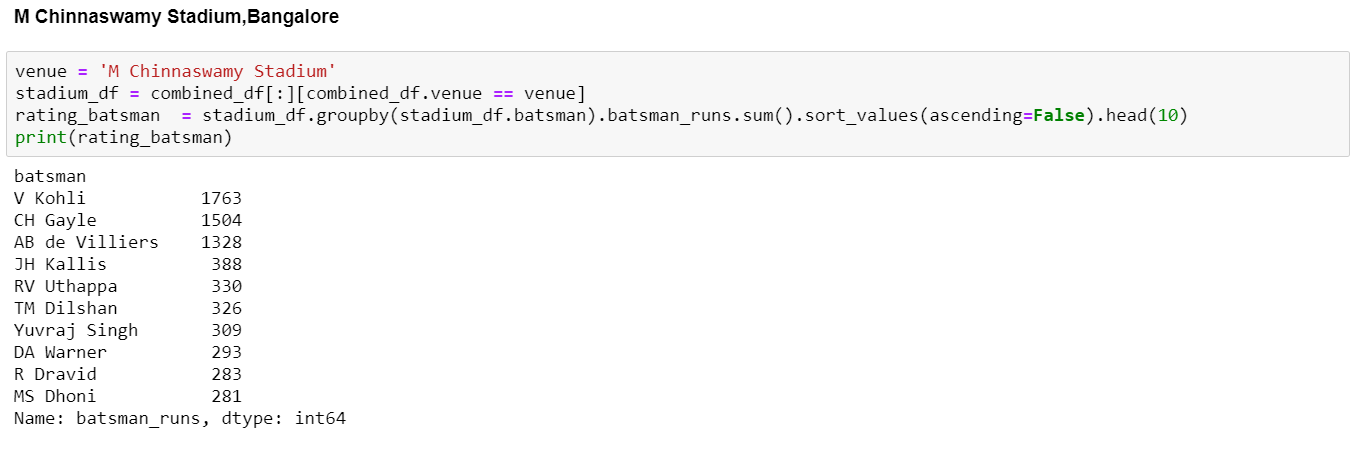
**Introduction:**

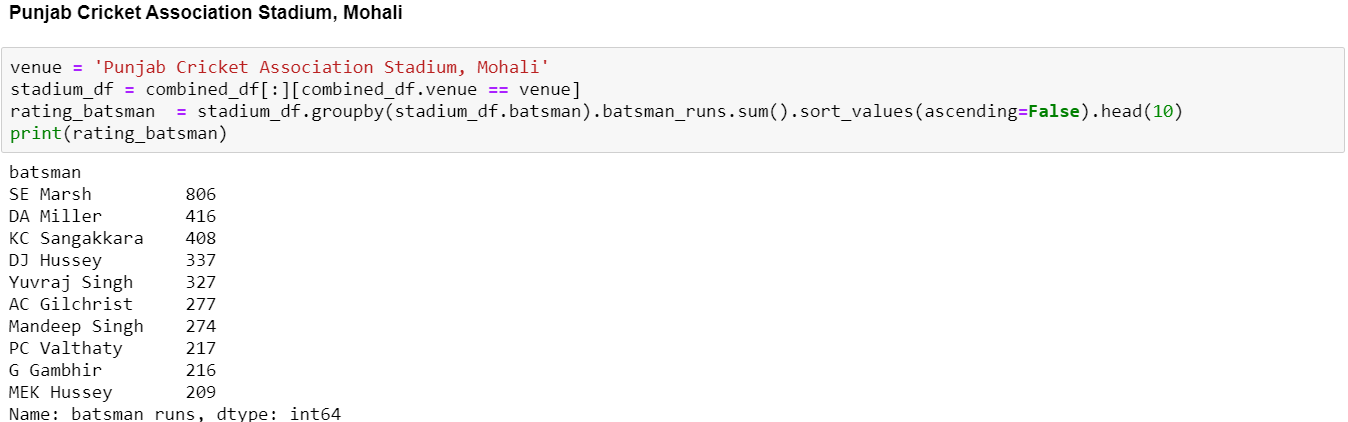
In this analysis, we are trying to select a batsman from available options who has performed well on the specific ground.

**Methodology & Result:**

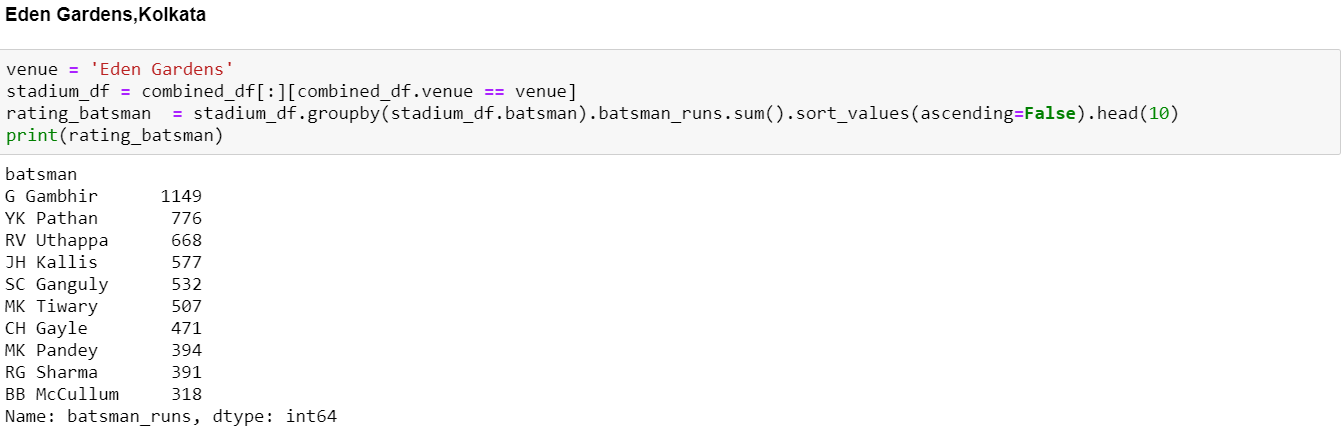


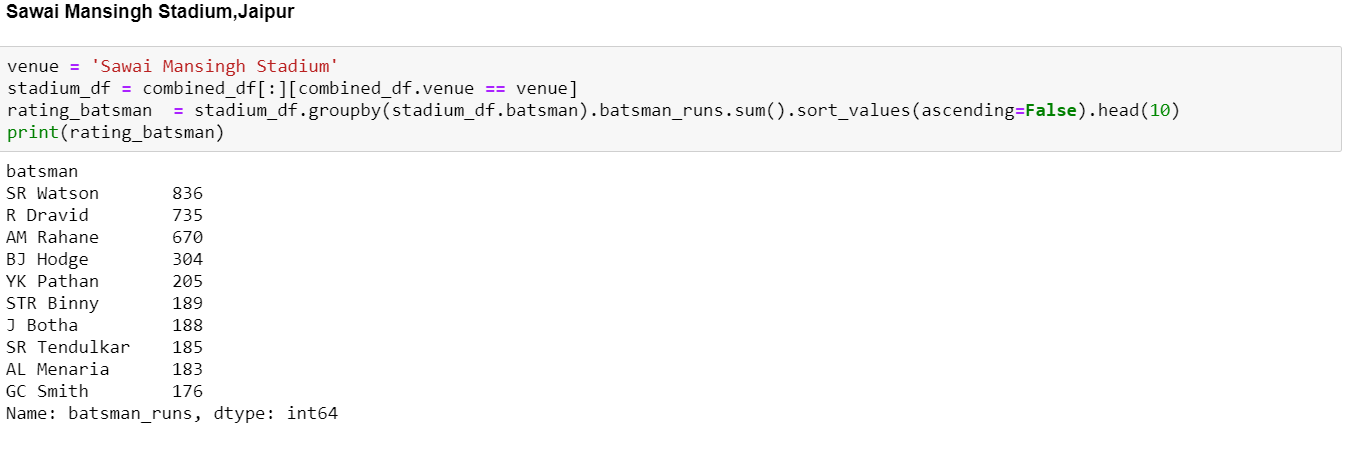










**Conclusion:**

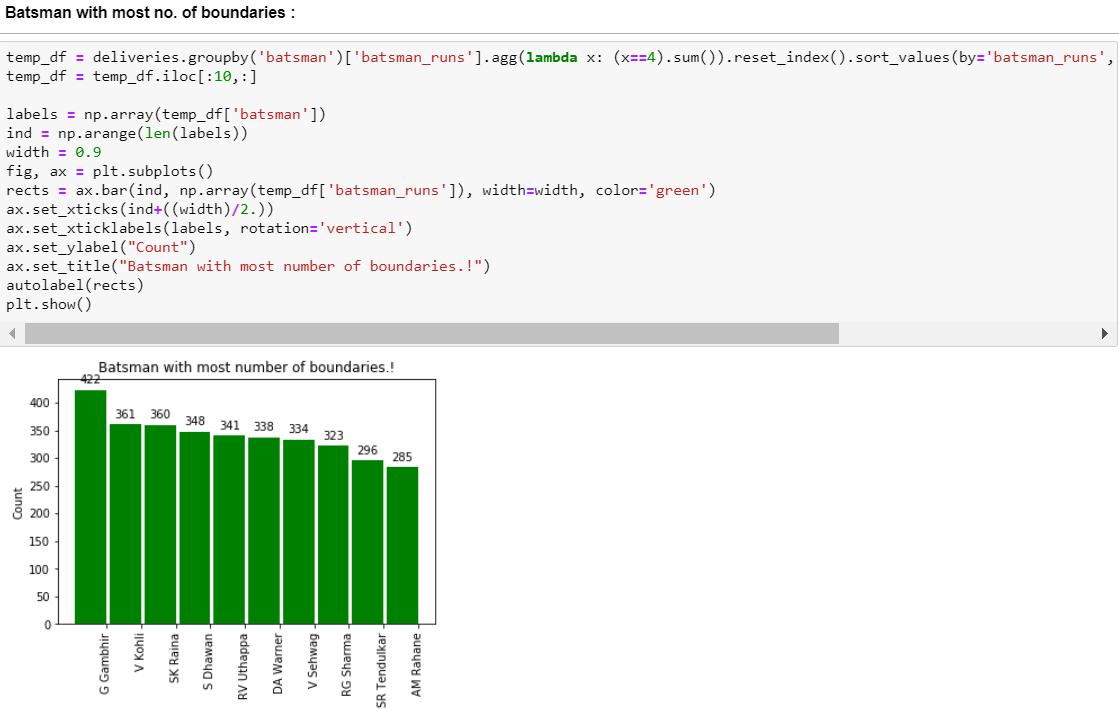
We can see list of players whose performance very from stadium to stadium and we can use the above information to select at least one of the player who is good at batting at specific ground.

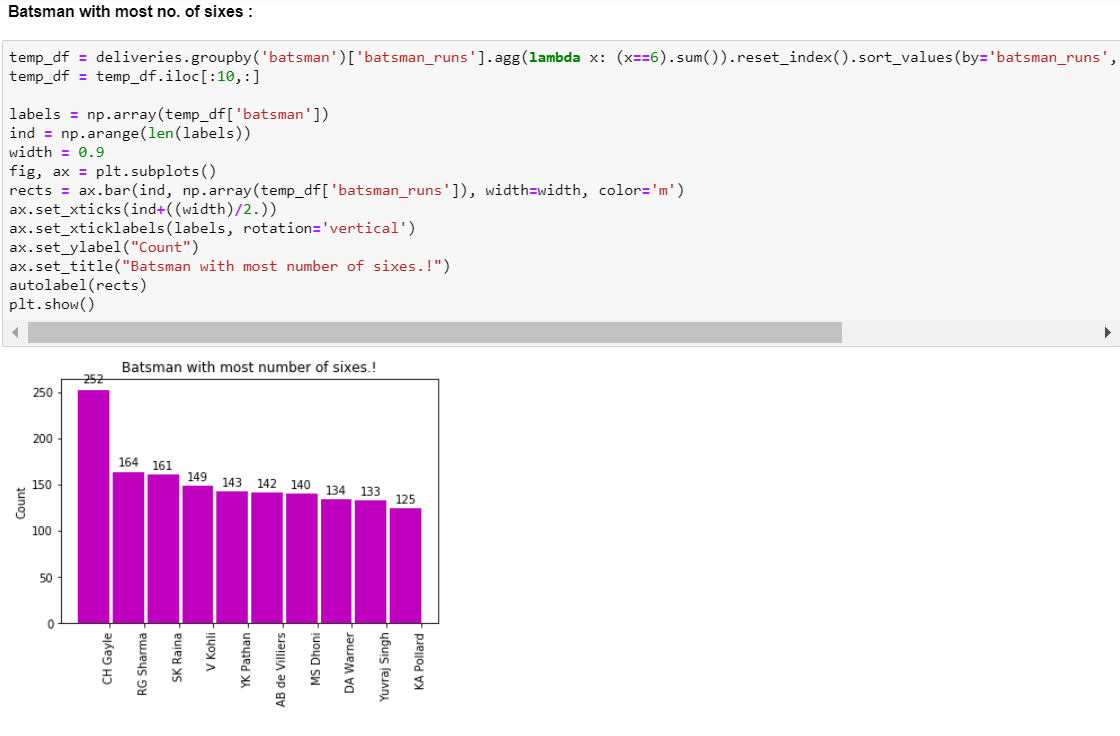
## 4.8 Finding outa batsman who has hit more boundaries and sixes (optional)

**Introduction:**

In this analysis ,we are trying to find out most batsman those have hit lot of boundaries although we have analysed most destructive batsman and this analysis is just optional one to find can include tell enders in the list which can hit run boundary in some situation.

**Methodology & Result:**





**Conclusion:**

We can see the list of players those can be helpful in the situation where target needs to be achieved with only few balls remaining.

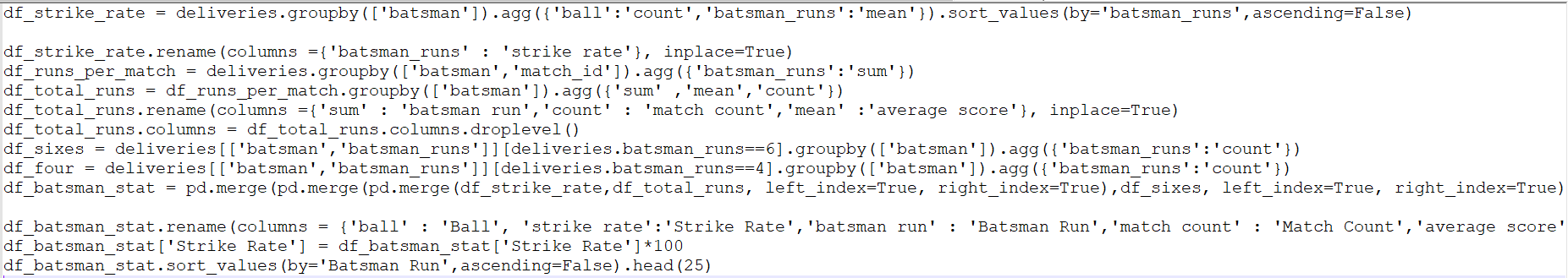
## 4.9 Select an individual batsman by making whole card of their performance

(It is another option where we can include all the information related to all qualities)

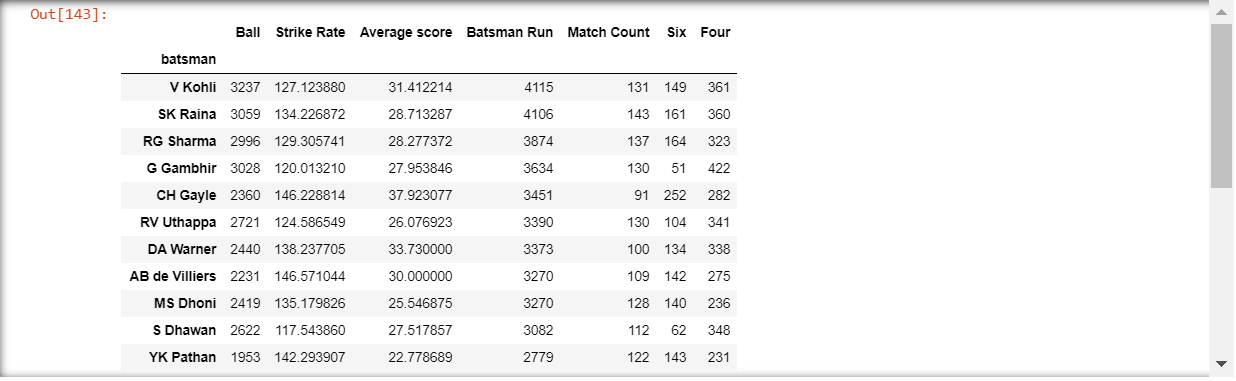
**Introduction:**

In this analysis, we are trying to select batsman based on all of the qualities which we have considered above and finding out which is will be more preferred based on all the qualities we discussed in the sections earlier.

**Methodology:**



**Result:**



**Conclusion:**

We can now see the list of players with their overall batting performances. This information can help us to decide which one to choose from the list with overall qualities.

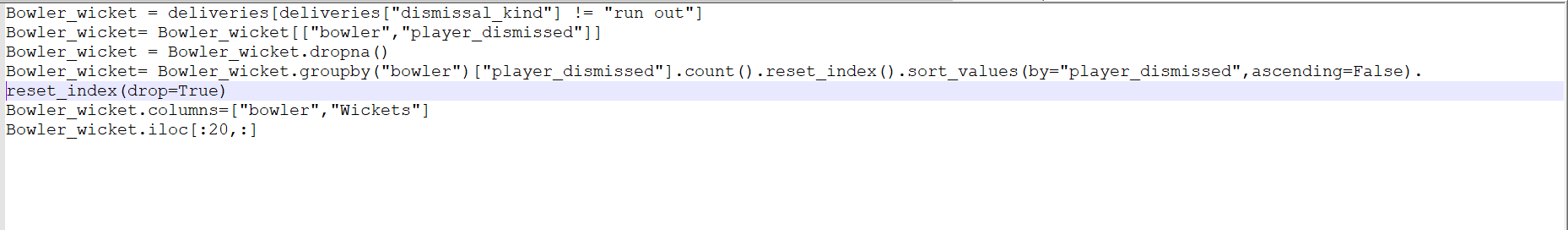
**Bowler Analysis**

## 4.10 Selecting the most wicket taking bowler

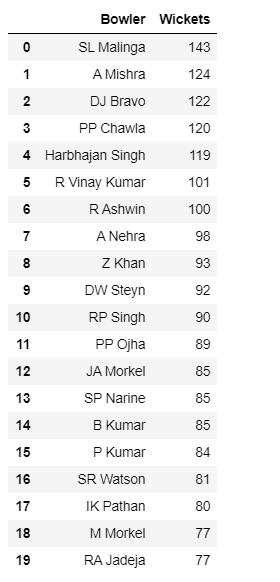
**Introduction:**

In this analysis, we are trying to find out most wicket taking bowlers and select them to get destructive bowling attack.

**Methodology:**



**Result:**



**Conclusion:**

We can now see the list of highest wicket taking batsman and can choose from the list and available options.

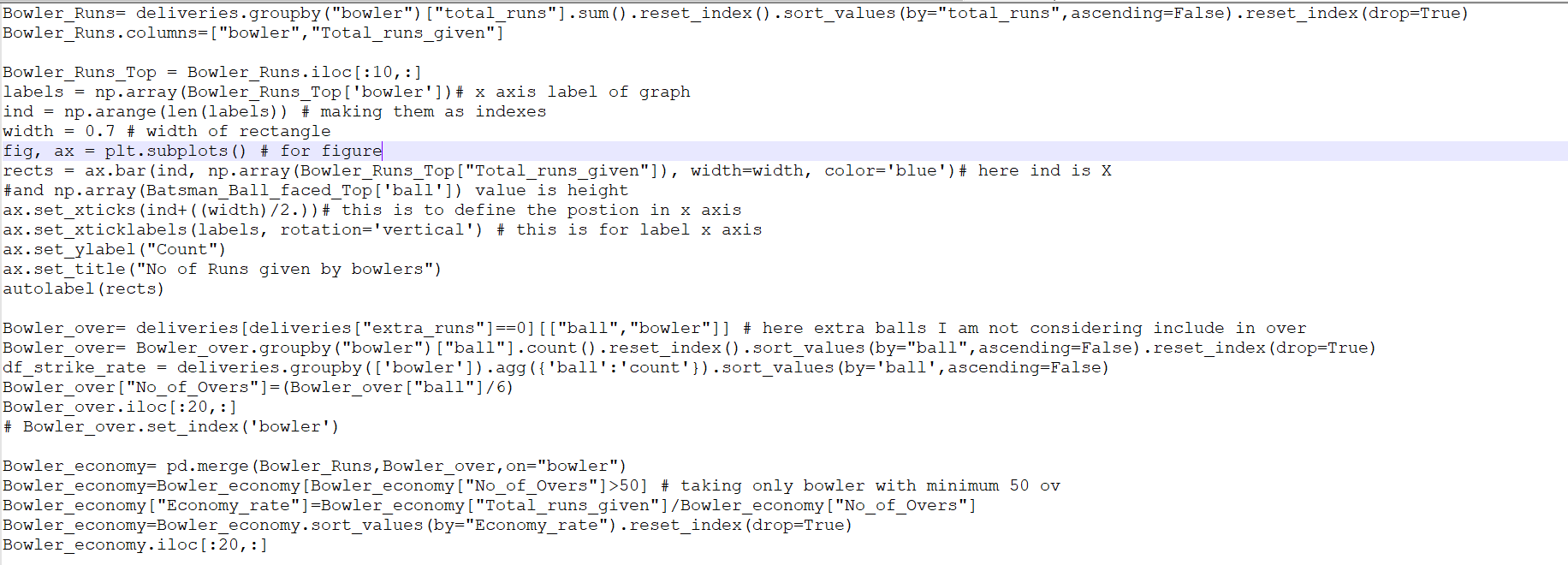
## 4.11 Select the most economical bowler

**Introduction:**

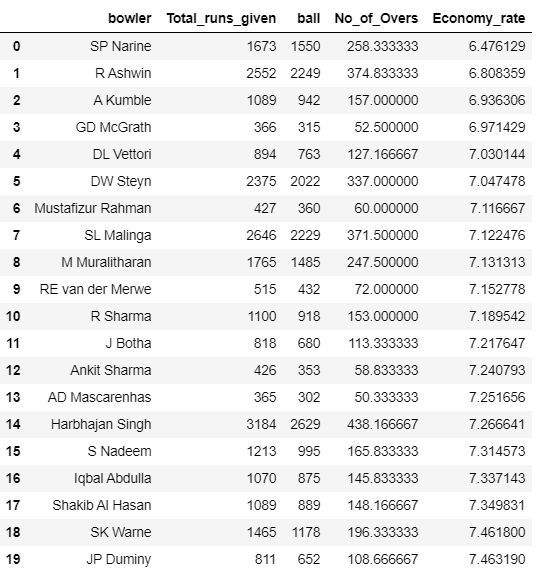
In this analysis, we are trying to find out most economical bowler who can save runs in the short format of cricket where everyone is trying to hit each bad ball hard and get maximum runs on the board.

To get this we need to separate out our analysis into two par, first by getting the overs thrown by each bowler and their given runs and then dividing the total runs by total overs. We have considers at least 50 overs for each bowlers for this analysis.

**Methodology:**



**Result:**



**Conclusion:**

We now got the list of most economical bowler and we can opt them from the list to get as minimum number of run as possible on the opponent’s score board.

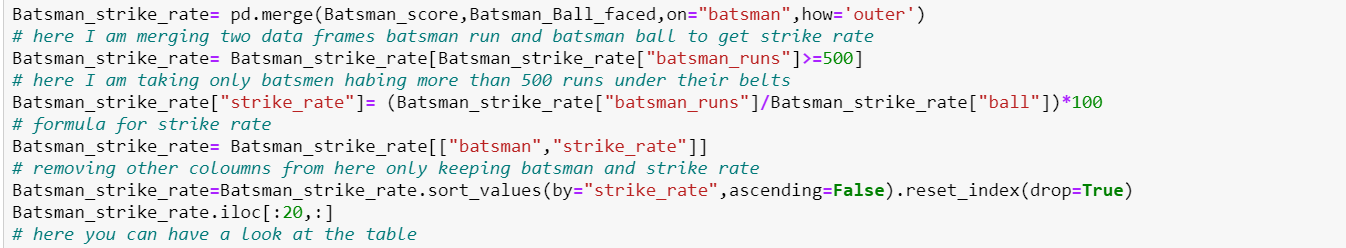
## 4.12 Finding out the best strike rate bowler (Less strike rate will be better)

Strike Rate of a bowler = the average number of balls bowled per wicket taken

**Introduction:**

In this analysis, we are trying to find out most destructive batsman by finding out the batsman who has faced very few balls and hit a lot of runs in the tournament and this way the partner can play without being worried of wickets from other side.

**Methodology:**



**Result:**



**Conclusion:**

We can see now that players those are in the list of highest run scorer are among the most reliable batsman in the list.

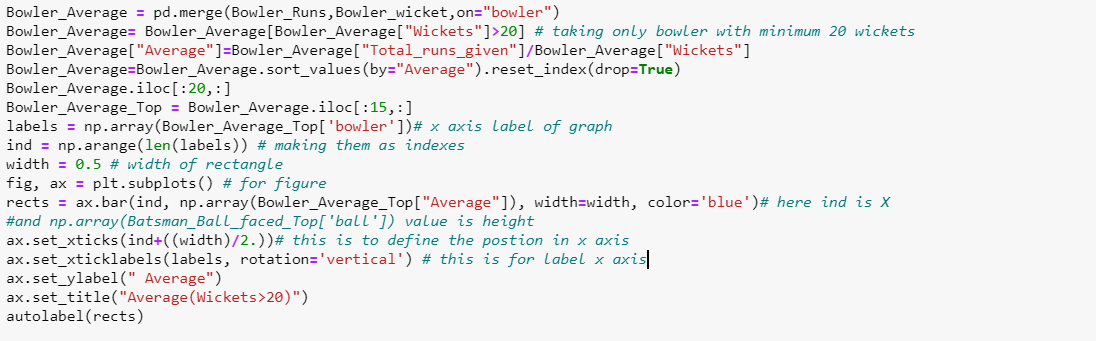
## 4.13 Now calculate average of bowler (Less average will be better)

Bowling Average=The average number of runs conceded per wicket

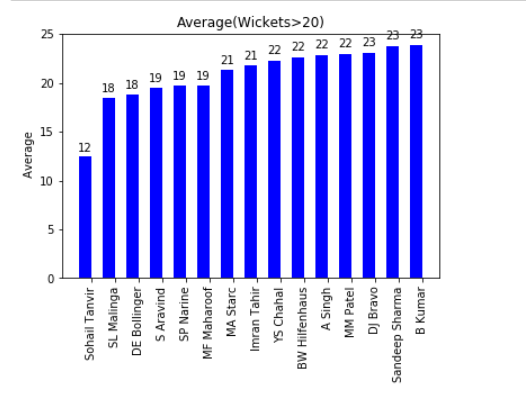
**Introduction:**

In this analysis, we are trying to find out bowlers those have given minimum runs to get a wicket. This way we will be able to identify the bowler who take wicket without conceding a lot of runs, mean that bowler will be better economically and destructive when taking wickets.

**Methodology:**



**Result:**



**Conclusion:**

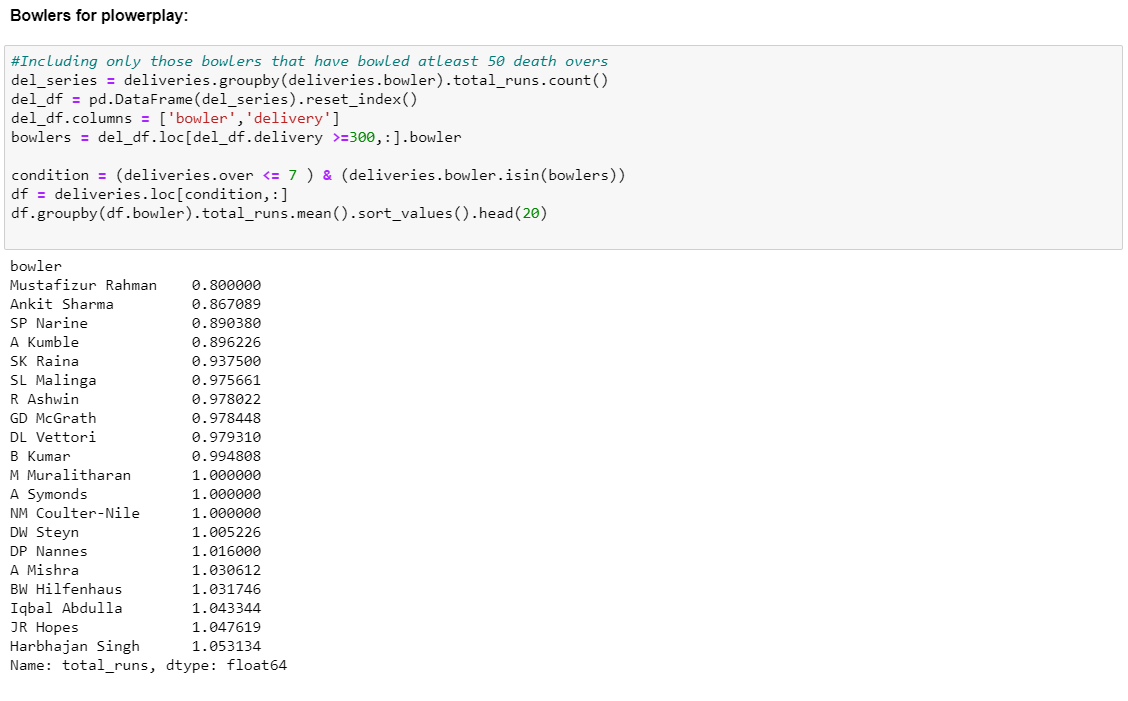
We can now get the list of players those will be better economically and destructive when taking wickets.

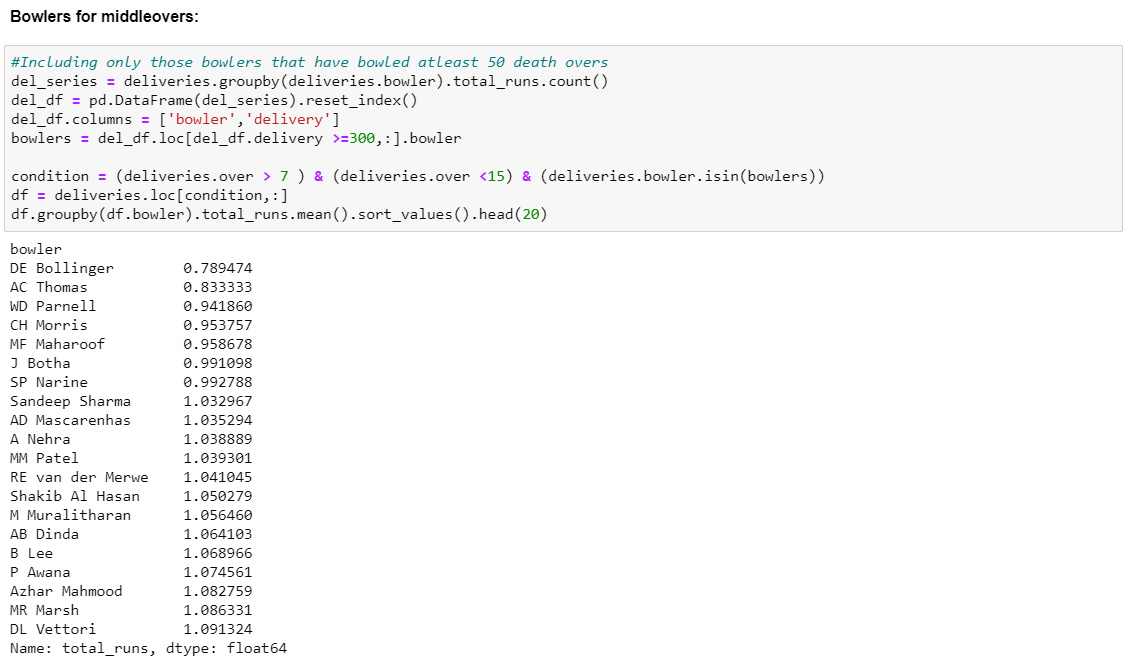
## 4.14 Let's pick the bowlers for different phases of match:

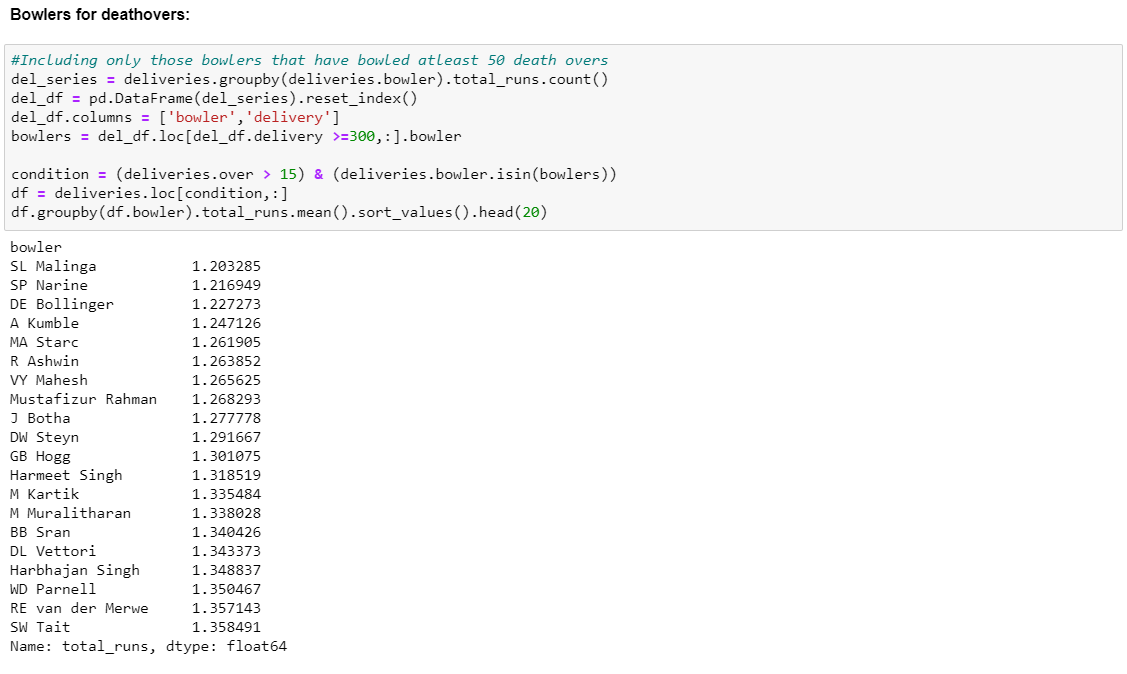
**Introduction:**

We have selected bowlers for our team, let’s organise their spell to get the maximum output from their skill sets. Now we pick bowlers for the different phases of match based on their past performance in different stages of game.

**Methodology & Result:**







**Conclusion:**

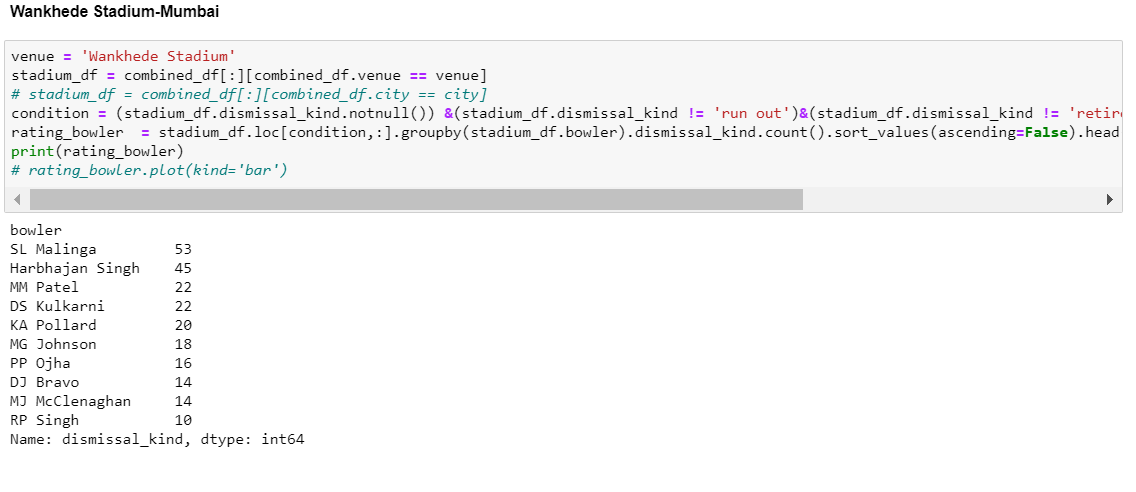
We can now see the list of players those fit into the different phases of game and we can arrange them accordingly those are part of team in the individual matches.

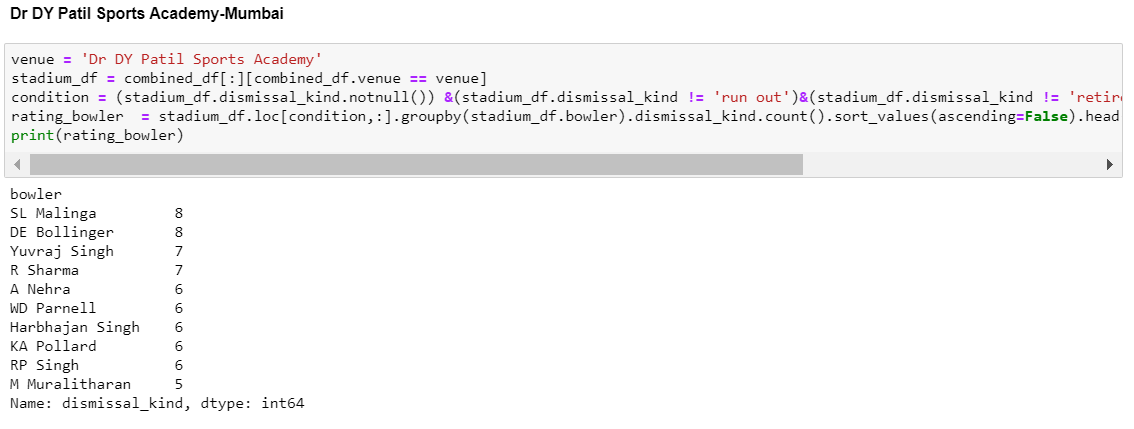
## 4.15 Top Bowlers Based on the stadium

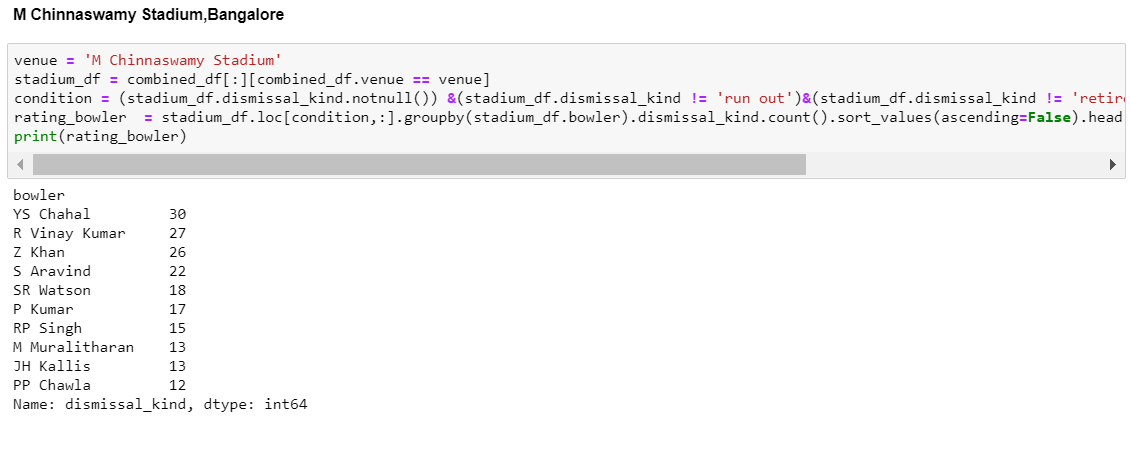
**Introduction:**

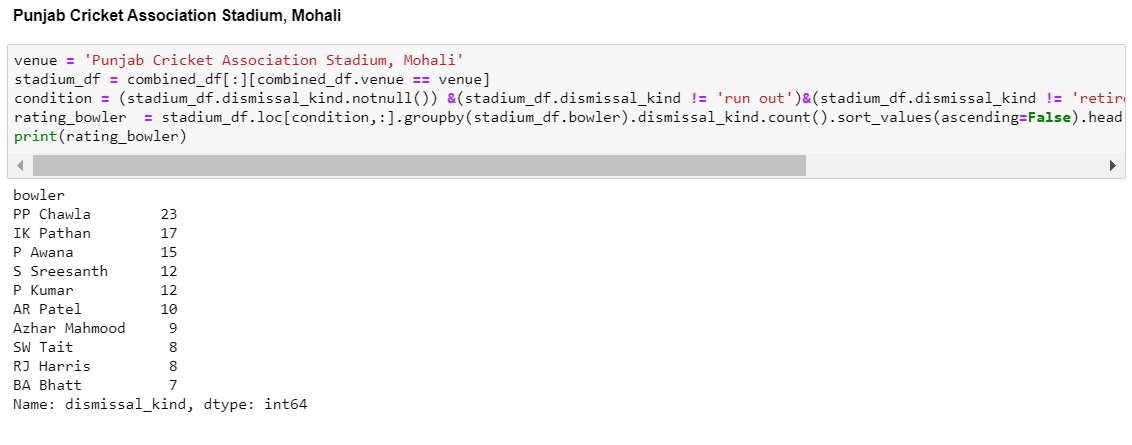
In this analysis, we are trying to select a bowler from available options who has performed well on the specific ground.

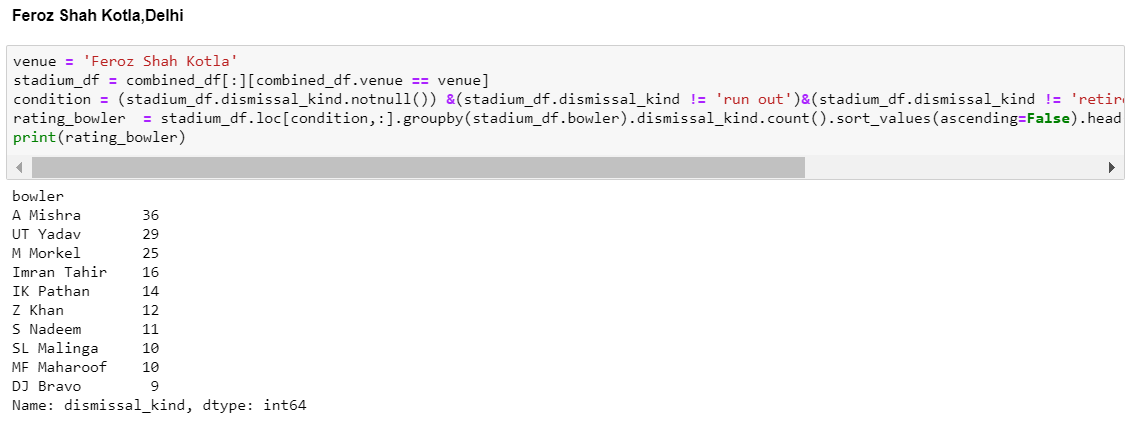
**Methodology & Result:**

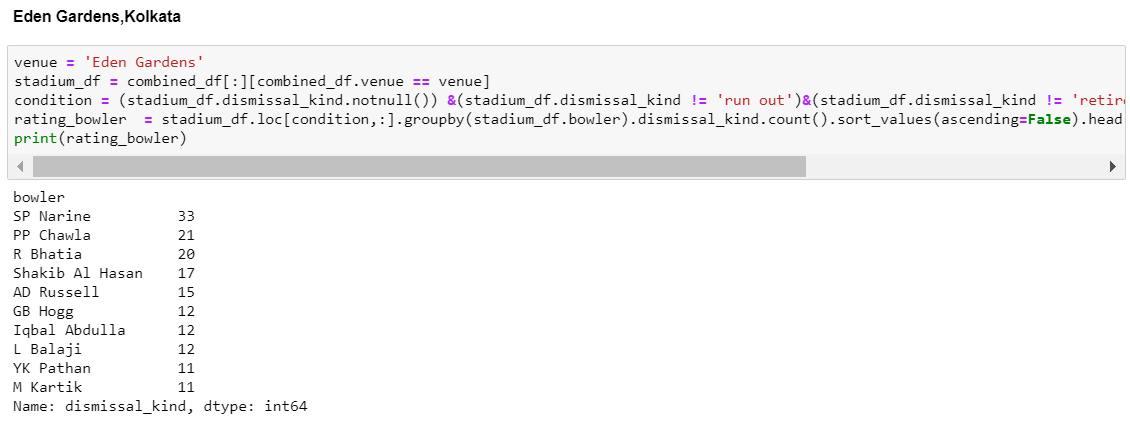


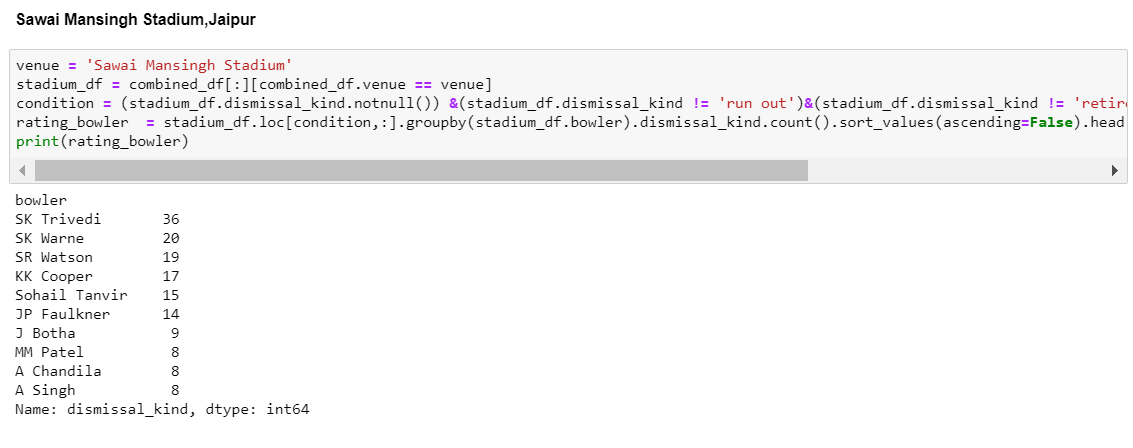


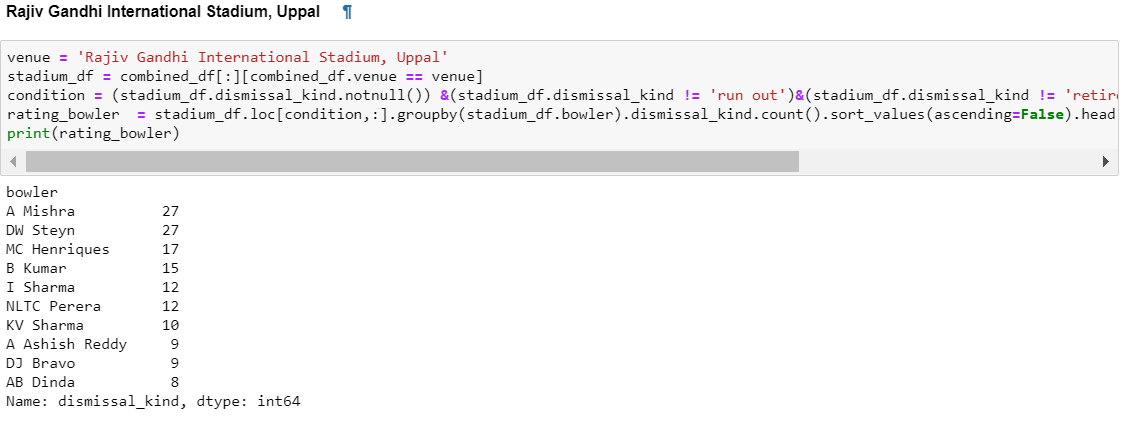


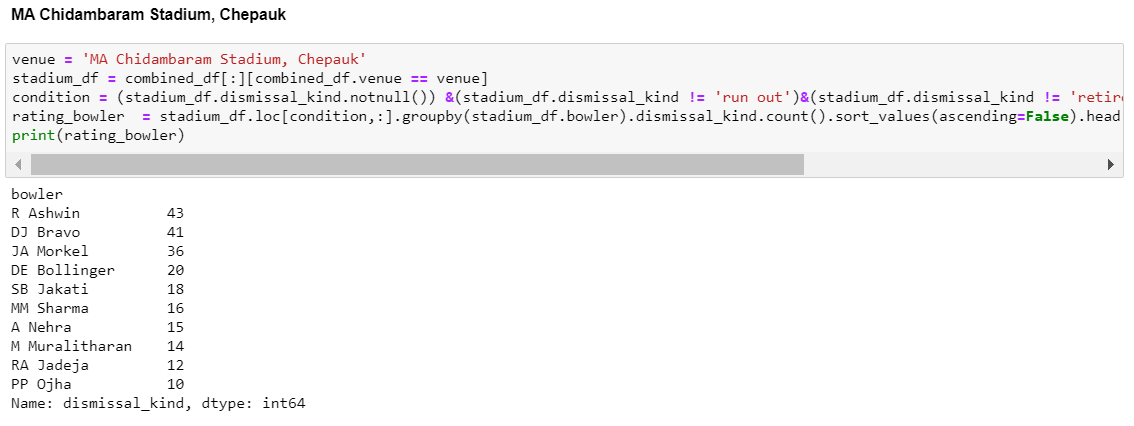












**Conclusion:**

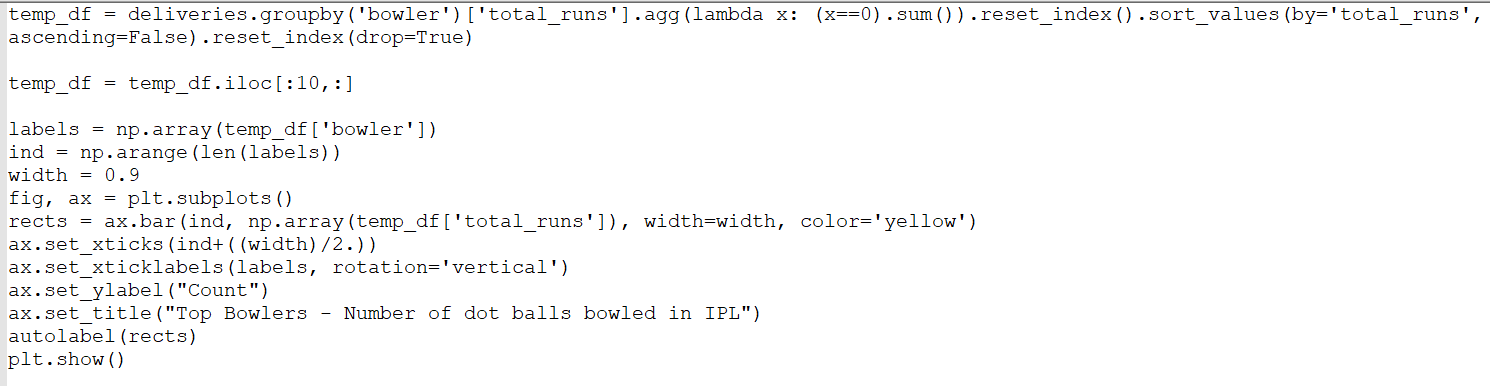
We can see list of players whose performance very from stadium to stadium and we can use the above information to select at least one of the player who is good at bowling at specific ground.

4.16 Finding out bowlers with maximum number of dot balls(Optional, Not useful alone)

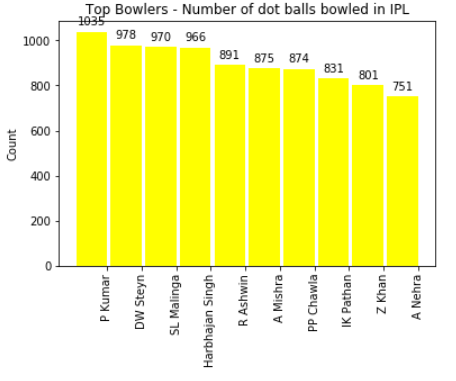
**Introduction:**

In this analysis, we are trying to find out bowler with maximum number of dot balls, it is an optional finding as it does not make much difference alone.

**Methodology:**



**Result:**



**Conclusion:**

We can see now that players those are in the list of bowlers those have ball maximum number of dot balls, it is not that important finding alone but yes we can use these type of bawlers in the death overs where bawling dot balls are similar to taking wickets.

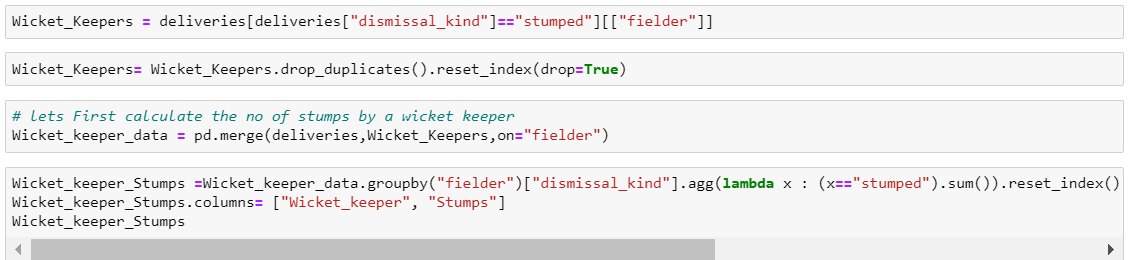
**Wicket Keeper Analysis**

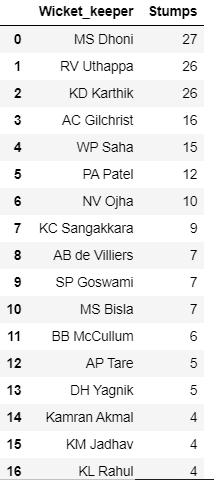
## 4.17 Selecting the best wicket keeper out of available options

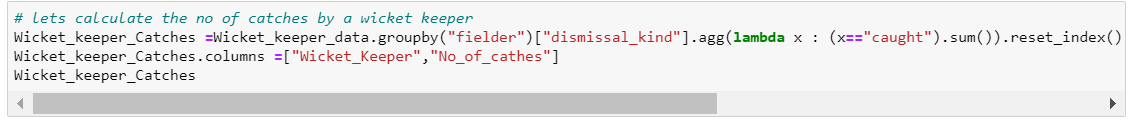
**Introduction:**

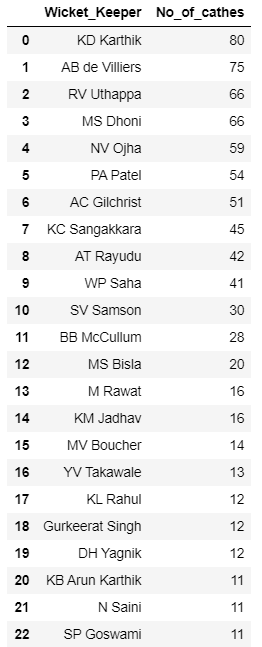
In this analysis, the most important question is, how do we know who is wicket keeper..? To answer this question we need to find out the fielders who has dismissed the batsman with dismissal type equals to stumping. We know that only wicket keeper can out the batsman by stumping. We can also divide this list into two parts as stumping and catches.

**Methodology & Result:**









**Conclusion:**

We can see now see the list of wicket keepers and can opt one of them based on our priority like either we need good wicket keeper who is active in wicket keeping or we can check these wicket keepers with our batsman analysis lists and can take the batter batsman if we give same priority to batting performance as well for wicket keepers.

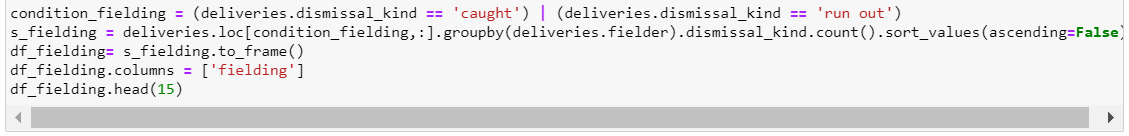
**Fielder Analysis**

## 4.18 Finding out the best fielders on the field

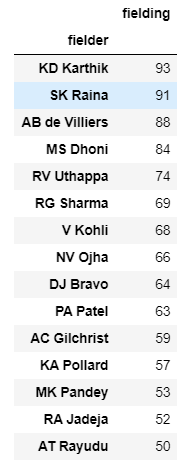
**Introduction:**

In this analysis, we are trying to find out best fielders around and it is important to analysis even though it cannot be considered for selection criteria but there can be some replacement on the field and we can be benefited by replacing one of the players with one who is good at fielding. Fielding is very important aspect as we can save a lot of runs from fielding only and that can be added advantage to us. We have seen in the past that players like Jaunty Rhodes used to get selected in team because of his fielding skill only and he used to save 30-35 run by his fielding only.

**Methodology:**



**Result:**



**Conclusion:**

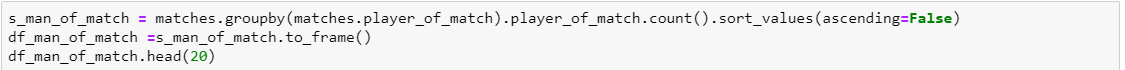
We can now see the players those are good at fielding by checking the number of wickets they have participated in either by catch or run outs.

## 4.19 Finding out the game changes (Man of the match)

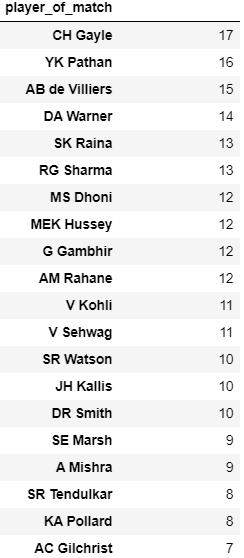
**Introduction:**

In this analysis, we are trying to find out most useful player of the match who has changed by his own and become an important player of the match most of the times and here the analysis is for finding out most Man of match awardees.

**Methodology:**



**Result:**



**Conclusion:**

We can now see the list most important players of the match who have been the game changer most of the times.

**Umpire Analysis**

## 4.20 Umpire Analysis (important when you are thinking to take review)

Last but most important for crucial decisions, reminds me the world cup final 2020

**Introduction:**

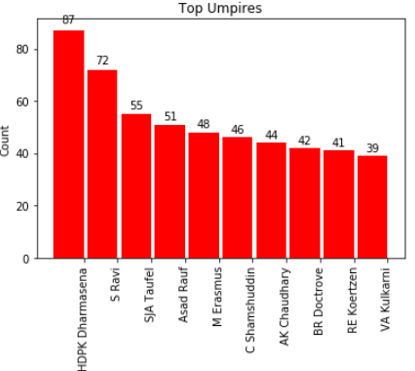
In this analysis, we are trying to find out experienced umpire and this analysis can help us to take review or not, like we have only single review available and if we lose it against a correct umpiring decision then we won’t be able to raised finger at umpire decision.

There are usually two umpires on the field and if one is very experience and given decision against us then before trying to take review we must remember that this umpire is very experience and there is less chance of their giving the wrong decision, so once we are very sure then only we should go for the review system.

**Methodology:**



**Result:**



**Conclusion:**

As clearly seen, Dharmasena is very experienced umpire and we need to think twice before appealing against their umpiring decisions.

#### Chapter 5

# CHALLENGES FOR THE STUDY

The Twenty20 format of cricket is probably the most significant development of the

Game in the twenty first century. In cricket, different measures like batting average and strike rate that are mostly used to measure the performances of batsmen and bowling average, economy rate and bowling strike rate to measure the performance of bowlers.

Despite of Team management’s accurate prediction to select best players for each match for team’s victory. There is no surety for the performance of any player, there are few challenges which may result in losing the game. Some features that affect player’s performance are weather or the nature of wicket, fitness of the player as well as opponent player’s fitness, Performance and form of opponent player. In addition to this following are important factor which are also important for winning any match.

PLAYING CONDITIONS:-

Playing condition of the ground also plays a vital role for the game as sometime weather can fever medium pacers if there is some due on the pitch and can fever batsman is it is very dry and not bouncing much also it can fever the spinner if wicket is getting good turn.

PLAYER’S CURRENT FORM:-

Our study is based on the past records but it doesn’t help much if player is currently out of form, so we need to check his current form as well after selecting his name after this research.

PLAYER’S DAY:-

We cannot help with study if today is not that player’s day, like Sachin gets bowled on first ball or mcgrath can get to much runs in a spell, like wise there is no assurance that the selected player will perform as his past performance.

TEAM BALANCE:-

A captain is entitled to have team balance, both in relation to the spread of batsmen and bowlers in the team as well as the type of batsman and bowlers.

ATTITUDE AND WORK ETHICS:-

Players are expected to work hard during training and support each other at all times. If this happens at training it will also happen in games.

AVAILABILITY:-

Every player will be considered available for selection and ready to play in every game of the season unless they have notified the club accordingly. Players are expected to notify the club of their unavailability immediately they are aware of it. Hence each player have to keep themselves fit unless may not perform well.

INJURY AND FITNESS LEVEL:-

Injury and fitness level are the some issues which we cannot control, all these issues can be happened with anyone at any time and we cannot handle these issues.

BEHAVIOUR:-

All players must adhere to cricket club code of conduct at all times when representing the club both on and off the field. Players bringing the game of cricket into disrepute will jeopardise their chances of future selection. This may affect management procedure and hence the game too.

SUPPORTING THE TEAM-

Cricket is a team sport and it is expected that players will work together at both training sessions and on match day to help set up and pack away equipment, boundary markers, afternoon tea etc.

Hence these all of the above factors may also be kept in mind by the team so as to get the winning performance despite of best players selected for the team.

#### Chapter 6

# SUMMARY OF REQUIREMENT

## 6.1 Prerequisite skills and knowledge

* Need to have good cricketing knowledge
* Should have good command over python language to read/write and understand the queries
* Good to have some basic knowledge of data science/stats functions like standard deviation etc.
* Should have knowledge of basic mathematical functions like average and all.

## 6.2 General hardware requirements

* Modern Operating System:
  + Windows 7 or 10
  + Mac OS X 10.11 or higher, 64-bit
  + Linux: RHEL 6/7, 64-bit (almost all libraries also work in Ubuntu)
* x86 64-bit CPU (Intel / AMD architecture)
* 4 GB RAM
* 5 GB free disk space

## 6.3 General software requirements

* Python or Anaconda with Python distribution/support included (Latest version)

#### Chapter 7

# CONCLUSION

.For any Teams Victory, Selection of the right players for match play an important role. Management Teams need to select the best players for each match for that an accurate prediction of how many runs a batsman is likely to score and how many wickets a bowler is likely to take in a match will help. In this study we created batting and bowling datasets based on player’s statistics and characteristics. In this study, we used Standard Deviation to predict batsmen performances in a match. We used Python functions as well. We targeted both the problems as classification problems where we classified runs and wickets into different ranges. We experimented with four supervised machine learning algorithms and compared their performance. In the study we ﬁnd a way of valuing players based on on-ﬁeld performance which can be used to build a winning team. The most valuable attribute to a team is the number of wins a player can contribute as part of the team.

For any Teams Victory, Selection of the right players for match play an important role. Management Teams need to select the best players for each match for that an accurate prediction of how different players will fit into our required categories of players like batsman with good average, batsman with good strike rate, batsman who can keep wickets for a long and bowlers who take more wickets without giving a lot of run, bowlers who are likely to save a lot of run for the team(economy bowling),bowlers those suitable for specific phases of inning like powerplay, middle overs and death overs along with finding out the best wicket keeper who is good behind the stumps and can hit good runs for the team and also the fielder analysis who can perform good in the field, it is important as fielders can be substituted with other selected players on the field and last but not the least the umpire analysis to identify check whether we should appeal for review or not , if we are not sure about the taking this then we can check umpire’s experience and assume if he is very experience the it is very less likely to give wrong decision. Along with all this analysis, we are also checking the ground wise record for batting first or bowling first results to take batter toss decisions and ground wise individual’s performance.

#### Chapter 8

# REFERENCES

This thesis is aimed primarily at the CATEGORIAL ANALYSIS FOR TEAM SELECTION BASED ON IPL DATA

An attempt is made to keep the statistics involved as simple as possible .There are several key reasons for measuring and evaluating performance in team sport. Organisational Behaviour Theory proves particularly useful in drawing together sport statistics and selection. According to Greenberg and Baron (1997) to build high performance teams appropriate performance measures are required. Tests and measurements are tools that can be used for evaluation of an individual's performance (Franks, B. & Deutsch, H., 1973). For a high performance team, the right team members need to be selected (Greenberg et al, 1973). This means combining all available evidence, quantitative and qualitative, to make correct selection decisions.

**Background-**

**Importance of Cricket analysis**- Finding out Data for Reviewing, Match Coding, Analyzing Team Data, Analysing data from other team, an analysis is being carried out to decide the order of players dynamically, performance data such as( batting average and bowling average. etc)

**SPORTS ANALYTICS**

Sports analytics play a major role in various problems associated with sport. Some of these problems are the ranking of individual players and their specialized skills, the composition of teams with an optimal balance of specialized skills, the ranking of teams, the negotiation of contracts, the evaluation of sports businesses and their potential revenue streams, the planning of both physical and mental training, the development of strategies for winning games and tournaments, assessing the effectiveness of coaches and referees, the medical and actuarial aspects of sports injuries (health and insurance), the analysis of existing rules and the need for improving such rules, the improvement of equipment and technology, the determination of awards, the keeping of historical records and the generation of odds for gambling activities. Related to all of the above is the coherent statistical presentation of both raw data and its inferences to the decision makers to facilitate successful planning and implementation. Furthermore, the media and the public have a great appetite for well visualized statistics.

The sport of [cricket](https://en.wikipedia.org/wiki/Cricket) has a known history beginning in the late 16th century. Having originated in south-east England, it became the country's national sport in the 18th century and has developed globally in the 19th and 20th centuries. International matches have been played since 1844 and [Test cricket](https://en.wikipedia.org/wiki/Test_cricket) began, retrospectively recognised, in 1877. Cricket is the world's second most popular spectator sport after [association football](https://en.wikipedia.org/wiki/Association_football) (soccer). Governance is by the [International Cricket Council](https://en.wikipedia.org/wiki/International_Cricket_Council) (ICC) which has over one hundred countries and territories in membership although only twelve currently play Test cricket.

#### The Indian Premier League

The inaugural India Premier League took place in 2008 with 8 teams competing. Since then, the number of franchises (teams) increased to 11 in 2011 before returning to 8 by the start of the 2014 season. At the original franchise auction, the franchises were sold for $723.59 millions to Indian businesses and celebrities (Cricinfo staff 2008). The 2014 format featured 8 teams competing in a double round-robin tournament and took place in venues across India and the UAE between April and June. Each team play every other team at their respective venues, with some teams sacrificing home games to allow matches to be played in the UAE.

IPL Overview

Indian Premier League (IPL) tournament is the league of the cricket of India. The 2013 season of the Indian Premier League, established by the Board of Control for Cricket in India (BCCI) in 2007. The premier league is the shortest type of Cricket. A T20 match only has 20 over’s per inning. Indian and International Players take part in IPL and contributing there what is the world’s “richest cricket tournament”. IPL tournament is basically a cricket tournament and in this tournament all team plays exact two matches with all other teams’ one on the home town stadium and second on opponent home town. Till now total ……… IPL tournaments have been done.

**AIM-**

Our aim is to ﬁnd a way of valuing players based on on-ﬁeld performance which can be used to build a winning team. A team can only recruit new players through the auction or via trades, so our aim should also be to build the best team possible with budget constraints. The most valuable attribute to a team is the number of wins a player can contribute as part of the team.