

Seidenberg School of Computer Science and Information Systems  
Pace University, New York

---

# **Predicting the Accuracy in ICC Men's Cricket World Cup**

Data Warehousing, Mining and Visualization Project

---

**Submitted By :-**

Anup Savant

Manali Gaikwad

Shraddha Sawant

Udayan Sawant

**Submitted To:-**

**Prof. Yegin Genc**

Assistant Professor, Pace University

Seidenberg School of Computer Science and Information Systems

Department of Information Systems

December 12<sup>th</sup>, 2019

# Abstract

*Cricket* is one of the most popular team games in the world. Player selection is one the most important tasks for any sport and cricket is no exception. The performance of the players depends on various factors such as the opposition team, the venue, his current form etc. The team management, the coach and the captain select eleven players for each match from a squad of 15 to 20 players. They analyze different characteristics and the statistics of the players to select the best playing eleven for each match. Each batsman contributes by scoring maximum runs possible and each bowler contributes by taking maximum wickets and conceding minimum runs.

With the advent of statistical modelling in sports, predicting the outcome of a game has been established as a fundamental problem. With this article, we embark on predicting the outcome of a *One Day International* (ODI) cricket match using a supervised learning approach from a team composition perspective. Our work suggests that the relative team. Modelling the team strength boils down to modelling individual player's batting and bowling performances, forming the basis of our approach. We use career statistics as well as the recent performances of a player to model consistency, form and fitness.

**Keywords :** *Cricket, One Day International, Modelling Players, Modelling Team.*

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Introduction to Cricket . . . . .	1
1.2	ICC Cricket World Cup . . . . .	1
1.3	Cricket Analytics - Statistics & Actuarial Sciences . . . . .	2
1.4	Improvements in Decision Making . . . . .	2
<b>2</b>	<b>Related Work</b>	<b>4</b>
2.1	Data sets . . . . .	4
2.2	Data Processing . . . . .	4
2.2.1	<i>matches.csv</i> . . . . .	5
2.2.2	<i>batsmen.csv</i> . . . . .	6
2.2.3	<i>bowler.csv</i> . . . . .	8
<b>3</b>	<b>Exploratory Analysis</b>	<b>10</b>
<b>4</b>	<b>Predictive Analysis</b>	<b>17</b>
4.1	Model 1: Bowler's Consistency in World Cups . . . . .	17
4.2	Model 2: Batsman's Form in World Cups . . . . .	19
<b>5</b>	<b>Conclusion</b>	<b>21</b>
<b>6</b>	<b>Future Work</b>	<b>22</b>

## List of Figures

1	Top 10 Batsman with most runs in ICC Cricket World Cup . . . . .	10
2	Exploratory Analysis - Top 10 Batsman with most runs in ICC Cricket World Cup . . . . .	10
3	Top 5 leading wicket takers in ICC Cricket World Cup . . . . .	11
4	Exploratory Analysis - Top 5 leading wicket takers in ICC Cricket World Cup . . . . .	11
5	Closest Margin Victory in terms of runs . . . . .	12
6	Exploratory Analysis - Closest Margin Victory in terms of runs) . . .	12
7	Closest Margin Victory in terms of wickets . . . . .	13
8	Exploratory Analysis - Closest Margin Victory in terms of wickets .	13
9	Matches won by each team in ICC World Cups . . . . .	14
10	Exploratory Analysis - Matches won by each team in ICC World Cups	14
11	Measures on Central Tendency for wickets & runs . . . . .	15
12	Descriptive Statistics for Matches & Wins by Runs . . . . .	15
13	Measures on Central Tendency for wickets & runs . . . . .	16
14	Descriptive Statistics for Matches & Wins by Wickets . . . . .	16
15	Summarizing the dependent & independent variable : Runs & Wickets	17
16	Displaying the accuracy of the predicted model: Bowler's Consistency	17
17	XY Plot : Runs & Wickets . . . . .	18
18	Bowler's Consistency in World Cups . . . . .	18
19	Summarizing the dependent & independent variable : Runs & Mat .	19
20	Displaying the accuracy of the predicted model: Batsman's Form . .	19
21	XY Plot : Runs & Mat . . . . .	20
22	Batsman's Form in World Cups . . . . .	20

# 1 Introduction

## 1.1 Introduction to Cricket

*Cricket* is a sport played by two teams with each side having eleven players. Each team is a right blend of batsmen, bowlers and all rounders. The batsmen's role is to score maximum runs possible and the bowlers have to be maximum wickets and restrict the other team from scoring runs at the same time. All rounders are the players who can both bat and bowl and they contribute by scoring runs and taking wickets. Each player contributes towards the overall performance of the team by giving his best performance in each match.

Each player's performance varies with factors like the team he is playing against and the ground at which the match is being played. It is important to select the right players that can perform the best in each match. The performance of a player also depends on several factors like his current form, fitness, his performance against a particular team, his performance at a particular venue etc. The team management, the coach and the captain analyze each player's characteristics, abilities and past stats to select the best playing XI for a given match. In other words, they try to predict the players' performance for each match.

The game of cricket is played in three formats - Test Matches, ODIs & T20s. We focus our research on International Cricket Council's One Day International Men's Cricket World Cup, the most popular tournament of the game.

## 1.2 ICC Cricket World Cup

The ICC Cricket World Cup is the international championship of One Day International (ODI) cricket. The event is organised by the sport's governing body, the International Cricket Council (ICC), every four years, with first qualification rounds leading up to a semifinals and then finals tournament. The tournament is one of

the world's most viewed sporting events and is considered the "**flagship event of the International Cricket Calendar**" by the ICC.

### 1.3 Cricket Analytics - Statistics & Actuarial Sciences

*Cricket* has over one billion fans globally, with the Indian sub-continent alone constituting more than 90 percent of them, according to the largest ever market research into the sport conducted by the ICC.

As per the results of the research, the average age of fans (in the surveyed age category of 16-69) is 34 with a demographic breakdown of 61% male and 39% female. The research was undertaken to enable the ICC and its members to understand the growth potential of cricket, help shape the development of the growth focused global strategy for the game, drive decision making and to provide a benchmark upon which to measure the outcomes of the strategy.

It also supports the ICC and its members in shaping and implementing their fan engagement strategies and clearly presents cricket as a vibrant and popular sport globally. Close to 70% fans are interested in Test cricket with fans in England / Wales showing the strongest interest in this format (86%), South Africa lead the interest in ODI cricket (91%) and Pakistan in T20I cricket (98%). Globally, T20I is the most popular format with 92% interest while ODIs are close with 88% interest.

*The importance and popularity of ICC Global Events was evident with 95% of fans interested or very interested, with the ICC Cricket World Cup and the ICC World T20 the most popular*

### 1.4 Improvements in Decision Making

Effective *decision-making* forms an important aspect of the game of cricket. In particular, the ability of players to process relevant perceptual information to select an appropriate course of action is crucial. However, although the importance of decision-making in cricket is universally accepted, there is currently little consistency in the approaches adopted to develop this important ability. This article presents a practical framework for decision making in the game of cricket, specifically focusing on batters and bowlers. Building upon this framework, relevant intervention strategies

are presented that seek to specifically target the development of decision-making in the game of cricket.

*Cricket* is a game where decision-making is of paramount importance. For each discrete passage of play the (ball that is bowled) the batter needs to make a decision about the shot that is going to be played, the bowler needs to make a decision about the shot that is going to be played, the bowler needs to make a decision about the type of ball that is going to be bowled, the wicket keeper needs to decide where to stand, and the captain needs to make decisions regarding the position of the fielders. As a result, effective decision-making is a crucial component of performance, and one of the key factors that distinguishes expert compared to novice players

## 2 Related Work

### 2.1 Data sets

To retrieve all the required statistics, the entire data set has been scraped from the *cricinfo.in* website. The data set includes all the matches played in the ICC Men's ODI World Cup till 2019, the latest edition of the tournament. The data set contains the basic match details including the two competing teams, the outcome of the *toss*, the date when it was held, the venue and the inner of the match for all matches. Along with these, the career statistics of the participating players and their performance in every match is also included.

We have restricted our study to only 4 recent editions of World Cup - 2007, 2011, 2014/15 and 2019 held in West Indies, India, Australia/ New Zealand and the recent one in England and Wales respectively.

### 2.2 Data Processing

As mentioned in the previous section, the stats of the players such as average, strike rate etc, are not available directly for each game, we calculated these attributes from the inning by innings list using aggregate functions and mathematical functions. These attributes are generally used to measure a player's performance. We have also distinguished the matches won by wickets and runs.

The following below are the data sets with their description we have used in the research.



### 2.2.1 *matches.csv*

- **Team 1 (Proponent Team):** The proponent team tries to score possible maximum runs to win the game.
- **Team 2 (Opponent Team):** Opponent team tries to defend the proponent team by scoring more runs by them.
- **Winner:** In limited over matches (one innings matches), the team that scores more runs wins. In First-class cricket (two innings match), the team that scores more runs, having dismissed its opponents twice, is the winner.
- **Win By Wickets :** A team can win a match by a certain number of wickets. This means that they were batting last, and reached the winning target with a certain number of batsmen still not dismissed.
- **Win By Runs :** A team can also win a match by a certain number of runs. This means that they were batting first, they set a target score for other team to be chase down. If the other team fails to chase down the score by a run less than the targeted score, it can be termed as win by runs.
- **Venue:** The International Cricket Stadium at which the game is held/ to be played.
- **Match Date:** The day on which the match is scheduled/ to be played.
- **Match Id:** An unique ODI match number designated by International Cricket Council.
- **Edition:** Edition is basically a series of cricket matches which are held after every four years.

- **D/L Applied:** The *Duckworth–Lewis–Stern* method (DLS) is a mathematical formulation designed to calculate the target score for the team batting second in a limited overs cricket match interrupted by weather or other circumstances.
- **Result :** The result of a match is a "win" when one side scores more runs than the opposing side and all the innings of the team that has fewer runs have been completed. The side scoring more runs has "won" the game, and the side scoring fewer has "lost".
- **Player of the match :** A player of the match award is often given to the outstanding player in a particular match. This can be a player from either team, although the player is generally chosen from the winning team.

### 2.2.2 *batsmen.csv*

- **Player:** The demographic information about the player. E.g Sachin Tendulkar represented Indian Cricket Team playing most number of World Cup games as a batsman.
- **Country :** The country for which the batsmen plays for. Currently, 12 countries are full members of the ICC who play World Cups.
- **Matches:** Total number of matches a batsmen represented his country. On a rare side, a player can represent/play for two or more countries respectively as he can acquire citizenship of the country.
- **Innings:** The number of innings in which the batsman has batted till the day of the match. This attribute signifies the experience of the batsman. The more innings the batsman has played, the more experienced the player is.

- **Not Outs:** In cricket, a batsman is not out if they come out to bat in an innings and have not been dismissed by the end of an innings. The batter is also not out while their innings is still in progress.
- **Runs:** Total number of runs the batsman has scored in the World Cups.
- **Highest Score:** The highest runs scored by a batsman in any (single) innings throughout his career. This attribute is used in the formula for calculating the venue attribute. This attribute shows the run scoring capability of the batsman at the venue. If a player has a very high score at a venue in past, he is more likely to score more runs at that venue.
- **Batting Average :** Batting average commonly referred to as average is the average number of runs scored per innings. This attribute indicates the run scoring capability of the player.

$$\text{Average} = \text{Runs Scored} / \text{Number of times dismissed}$$

- **Strike Rate:** Strike rate is the average number of runs scored per 100 balls faced. In limited overs cricket, it is important to score runs at a fast pace. More runs scored at a slow pace is rather harmful to the team as they have a limited number of overs. This attribute indicates how quickly the batsman can score runs.

$$\text{Strike Rate: (Runs Scored / Balls Faced) * 100}$$

- **Hundreds:** Number of innings in which the batsman scored more than 100 runs. This attribute indicates the capability of the player to play longer innings and score more runs.
- **Fifties:** Number of innings in which the batsman scored more than 50 (and less than 100) runs. This attribute indicates the capability of the player to play longer innings and score more runs.

- **Zeros:** Number of innings in which the batsman was dismissed without scoring a single run. This attribute shows how many times the batsman failed to score runs, hence this being a negative factor, it impacts the batsman's prediction negatively.
- **Fours:** Number of four runs the batsman has scored in the matches he represented/ played the World Cup.
- **Sixes:** Number of six runs the batsman has scored in the matches he represented/ played the World Cup.

### 2.2.3 *bowler.csv*

- **Player:** The demographic information about the player. E.g Muttiah Muralidharan represented Sri Lankan Cricket Team with the most number of appearances in the World Cup games as a bowler.
- **Country :** The country for which the bowler plays for. Currently, 12 countries are full members of the ICC who play World Cups.
- **Matches:** Total number of matches a bowler represented his country. On a rare side, a bowler can represent/play two or more countries respectively as he can acquire citizenship of the country.
- **Innings:** The number of innings in which the bowler bowled at least one ball. It represents the bowling experience of a player. The more innings the player has played, the more experienced the player is or the more number of World Cups he has played.
- **Overs Bowled:** The number of overs bowled by a bowler in World Cups. This attribute also indicates the experience of the bowler. The more overs the bowler has bowled, the more experienced the bowler is.

- **Maidens:** A maiden over is one in which no runs are scored. Leg byes and byes scored in the over are not counted against the bowler in a maiden over.
- **Runs Conceded:** The number of runs conceded by a bowler is determined as the total number of runs that the opposing side have scored while the bowler was bowling, excluding any byes, leg byes, or penalty runs.
- **Bowling Average:** Bowling average is the number of runs conceded by a bowler per wicket taken. This attribute indicates the capabilities of the bowler to restrict the batsmen from scoring runs and taking wickets at the same time. Lower values of bowling average indicate more capabilities.

**Bowling Average: Number of runs conceded / Number of wickets taken**

- **Bowling Economy:** Economy rate is the average number of runs conceded for each over bowled in World Cups. A lower economy rate is seen as preferable – it means that the bowler is able to get more batsmen out with fewer balls.
- **Bowling Strike Rate :** Bowling strike rate is the number of balls bowled per wicket taken. This attribute indicates the wicket taking capability of the bowler in World Cups. Lower values mean that the bowler is capable of taking wickets quickly.

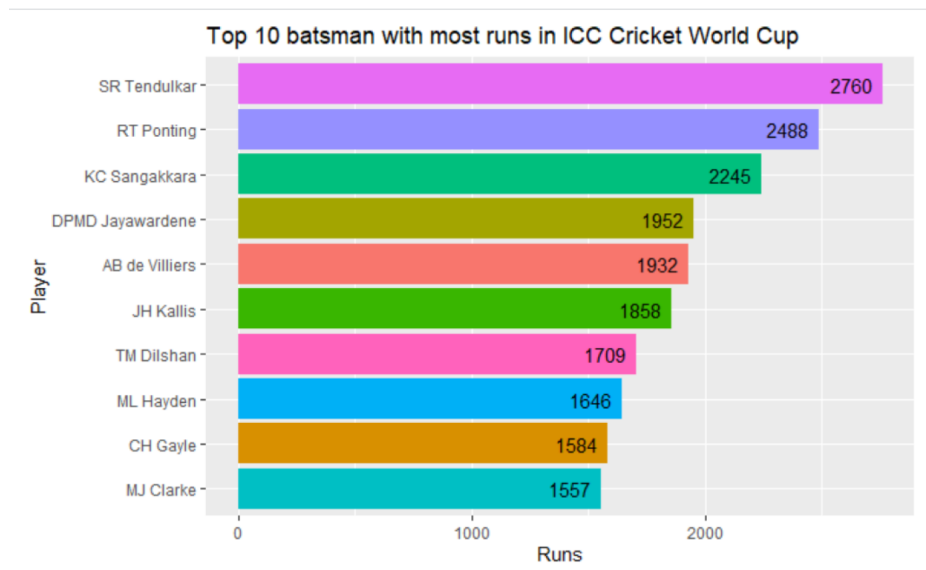
**Strike Rate: Number of balls bowled / Number of wickets taken**

- **Four Wicket Haul:** Number of innings in which the bowler has taken four wickets.
- **Five Wicket Haul:** Number of innings in which the bowler has taken more than four wickets. This attribute indicates the capability of the bowler to take more wickets in an innings. Higher the value, more capable the player

### 3 Exploratory Analysis

```
{r}  
## Top 10 Batsmen who has most runs in ICC Cricket World Cup  
  
library(ggplot2)  
batsmen %>%  
group_by(Player) %>%  
  summarize(Runs = sum(Runs))%>%  
  top_n(n = 10, wt = Runs)%>%  
  ggplot(aes(x = reorder(Player, Runs), y = Runs, fill = Player))+  
  geom_bar(stat = "identity", show.legend = FALSE)+  
  coord_flip()+  
  labs(x = "Player", y = "Runs", title = "Top 10 batsman with most runs in ICC  
Cricket World Cup")+  
  geom_text(aes(label = Runs), hjust = 1.25)
```

**Figure 1:** Top 10 Batsman with most runs in ICC Cricket World Cup

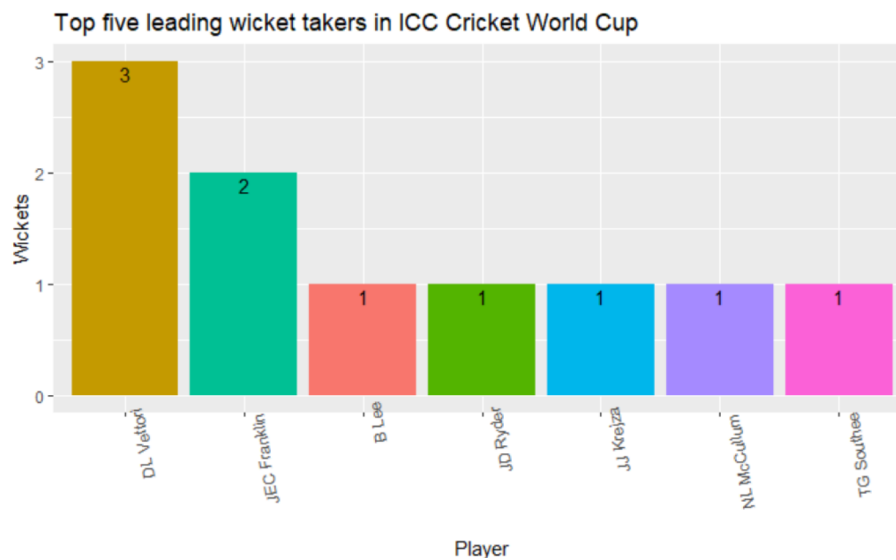


**Figure 2:** Exploratory Analysis - Top 10 Batsman with most runs in ICC Cricket World Cup

```
{r}
## Top 5 leading wicket takers in ICC Cricket World Cup

bowler %>%
group_by(Player)%>%
  summarise(wickets = sum(`Five wicket haul`in% Wickets))%>%
  arrange(desc(wickets))%>%
  top_n(5)%>%
  ggplot(aes(x= reorder(Player, -wickets), y = wickets, fill = Player))+
  geom_bar(stat = "identity",show.legend = FALSE)+
  labs(x = "Player" , y = "wickets", title = "Top five leading wicket takers in ICC
Cricket World Cup")+
  theme(axis.text.x = element_text(angle = 100, hjust = 1.25))+
  geom_text(aes(label = wickets, vjust =1.25))
```

**Figure 3:** Top 5 leading wicket takers in ICC Cricket World Cup



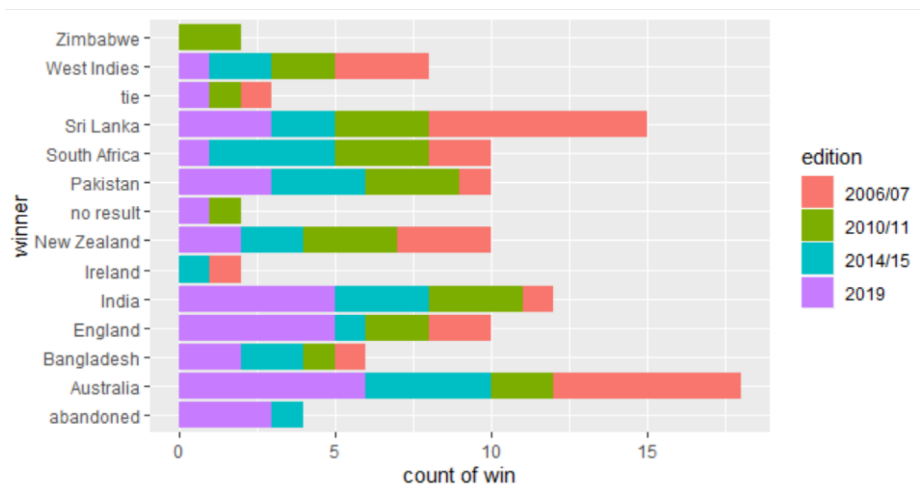
**Figure 4:** Exploratory Analysis - Top 5 leading wicket takers in ICC Cricket World Cup

```

####{r}
#CLOSEST MARGIN VICTORY (minimum runs)
matches%>%
  filter(win_by_runs != 1)%>%
  filter(win_by_runs == min(win_by_runs)) %>%
  select(winner, win_by_runs,edition)%>%
  ggplot(aes(winner,fill = edition))+
  geom_bar()+
  coord_flip()+|
  scale_y_continuous("count of win")
####

```

**Figure 5:** Closest Margin Victory in terms of runs



**Figure 6:** Exploratory Analysis - Closest Margin Victory in terms of runs)

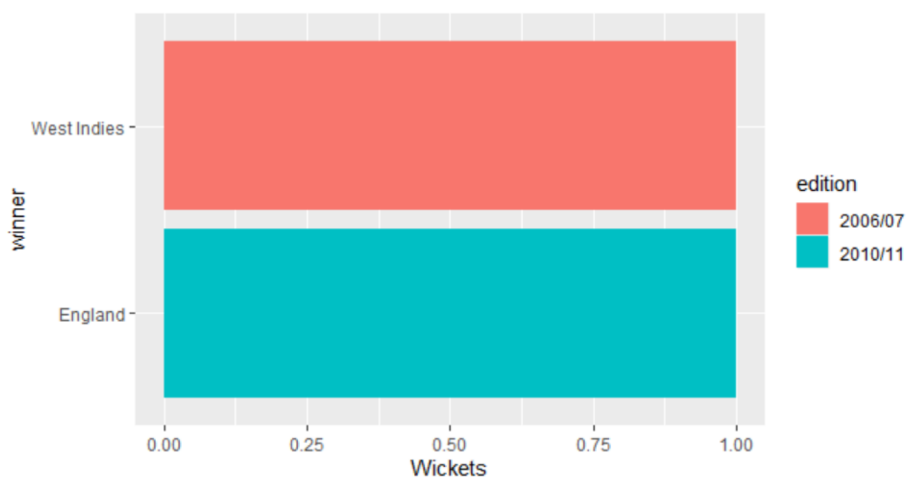


```

{r}
## CLOSEST MARGIN VICTORY by Wickets|
matches%>%
  filter(win_by_wickets >= 6)%>%
  filter(win_by_wickets == min(win_by_wickets))%>%
  select(winner, win_by_runs, edition)%>%
  ggplot(aes(winner, fill = edition))+
  geom_bar()+
  coord_flip()+
  scale_y_continuous("wickets")

```

**Figure 7:** Closest Margin Victory in terms of wickets



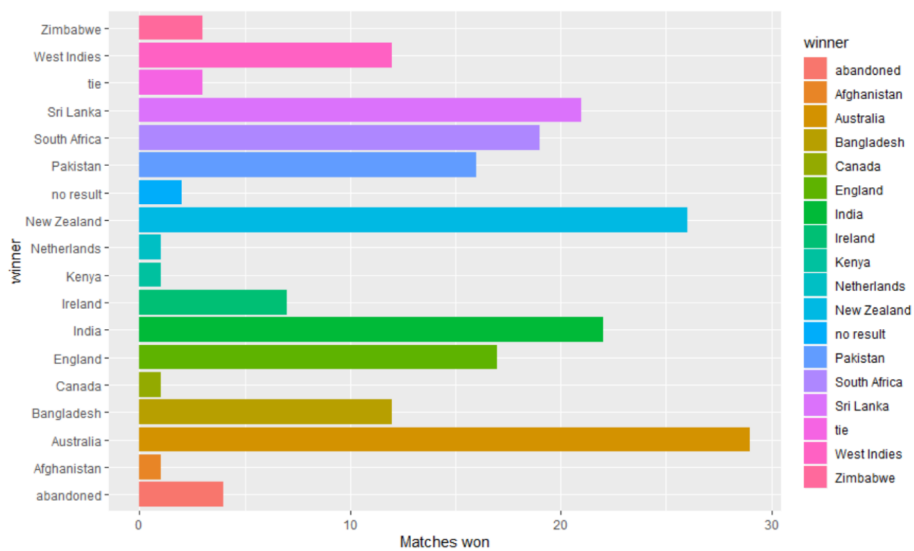
**Figure 8:** Exploratory Analysis - Closest Margin Victory in terms of wickets

```

{r}
## Matches won by each team
matches%>%
  group_by(winner)%>%
  summarize(most_win = n())%>%
  ggplot(aes(x = winner,y = most_win,fill = winner))+
  geom_bar(stat = "identity")+
  coord_flip()+
  scale_y_continuous("Matches won")

```

**Figure 9:** Matches won by each team in ICC World Cups



**Figure 10:** Exploratory Analysis - Matches won by each team in ICC World Cups

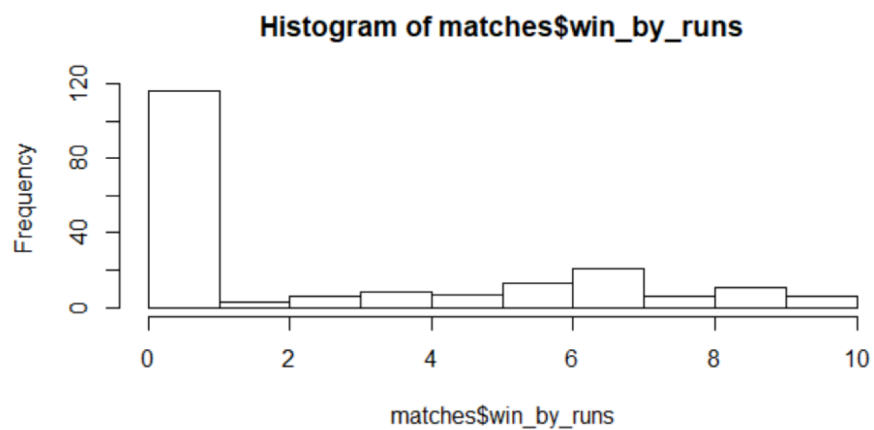
```

####{r}
matches
mean(matches$win_by_runs, na.rm = T)
mean(matches$win_by_wickets, na.rm = T)
median(matches$win_by_wickets, na.rm = T)
hist(matches$win_by_wickets)
hist(matches$win_by_runs)

## Histogram win by wickets and win by runs is rightly skewed and
is positively skewed.
####

```

**Figure 11:** Measures on Central Tendency for wickets & runs



**Figure 12:** Descriptive Statistics for Matches & Wins by Runs

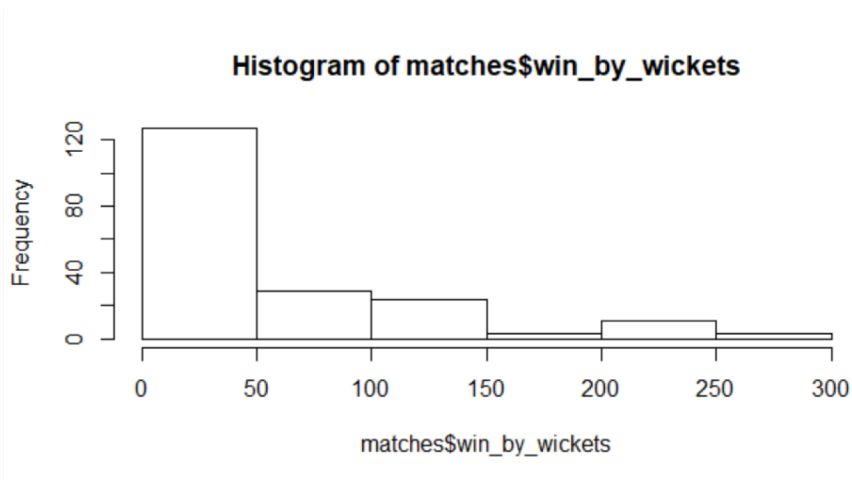
```

## {r}
matches
mean(matches$win_by_runs, na.rm = T)
mean(matches$win_by_wickets, na.rm = T)|
median(matches$win_by_wickets, na.rm = T)
hist(matches$win_by_wickets)

## Histogram win by wickets and win by runs is rightly skewed and
is positively skewed.

```

**Figure 13:** Measures on Central Tendency for wickets & runs



**Figure 14:** Descriptive Statistics for Matches & Wins by Wickets

## 4 Predictive Analysis

### 4.1 Model 1: Bowler's Consistency in World Cups

```
{r}
#Consistent bowler in World Cups
xyplot(Runs ~ Wickets, data = bowler)

{r}
summary(lm(Runs ~ Wickets, data = bowler))

{r}
ggplot(data = bowler, aes(x = Runs, y = Wickets)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE)
```

**Figure 15:** Summarizing the dependent & independent variable : Runs & Wickets

```
Call:
lm(formula = Runs ~ Wickets, data = bowler)

Residuals:
    Min       1Q   Median       3Q      Max
-352.24  -68.40  -24.58   71.23  452.30

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  76.9858    20.9477   3.675 0.000375 ***
Wickets      20.4338     0.8524  23.971 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 133.2 on 106 degrees of freedom
Multiple R-squared:  0.8443,    Adjusted R-squared:  0.8428
F-statistic: 574.6 on 1 and 106 DF,  p-value: < 2.2e-16
```

**Figure 16:** Displaying the accuracy of the predicted model: Bowler's Consistency

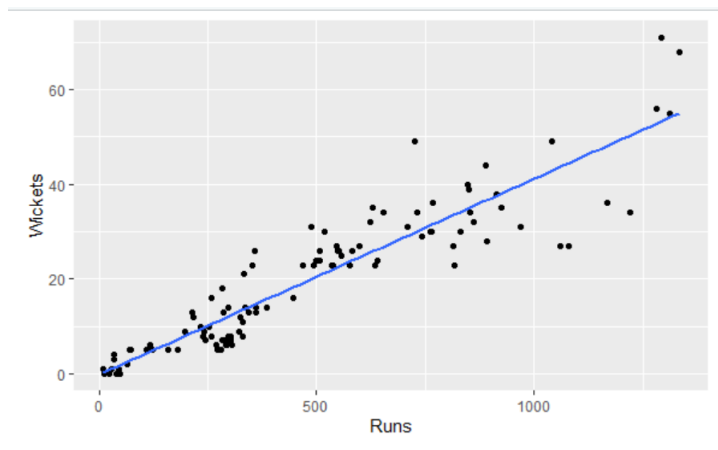


Figure 17: XY Plot : Runs & Wickets

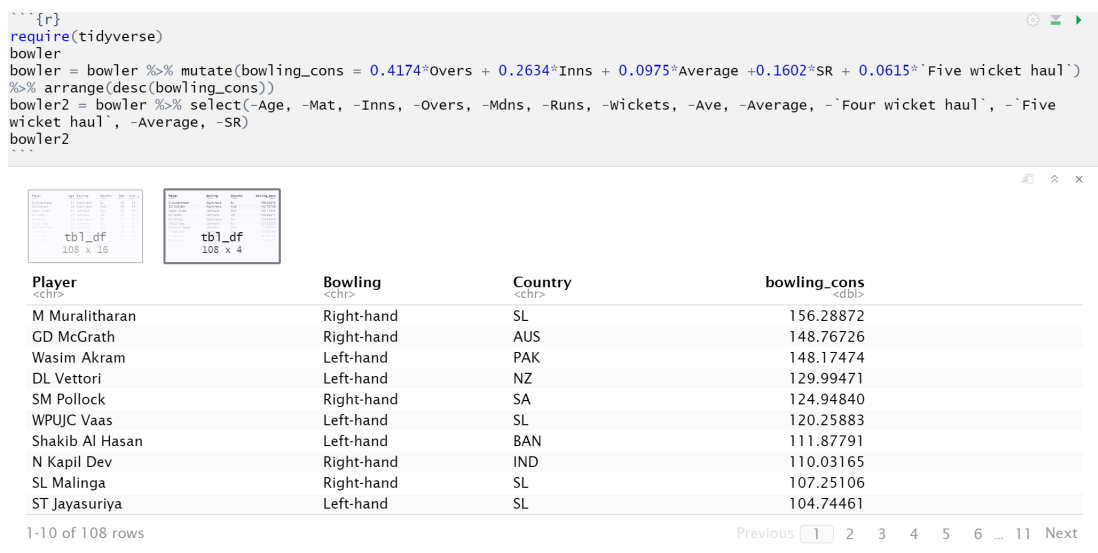


Figure 18: Bowler's Consistency in World Cups

## 4.2 Model 2: Batsman's Form in World Cups

```
## {r}
## Model 1 : Batting Form in world cups
batsmen
xyplot(Runs~., data = batsmen)

## {r}
summary(lm(Runs~Mat , data = batsmen))

## {r}
ggplot(data = batsmen, aes(x = Runs, y = Mat)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE)
```

**Figure 19:** Summarizing the dependent & independent variable : Runs & Mat

```
Call:
lm(formula = Runs ~ Mat, data = batsmen)

Residuals:
    Min       1Q   Median       3Q      Max
-361.15  -91.17  -15.37   74.51  632.88

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -10.668     20.411  -0.523   0.602
Mat           36.795       1.168  31.492 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 142.1 on 147 degrees of freedom
Multiple R-squared:  0.8709,    Adjusted R-squared:  0.87
F-statistic: 991.7 on 1 and 147 DF,  p-value: < 2.2e-16
```

**Figure 20:** Displaying the accuracy of the predicted model: Batsman's Form

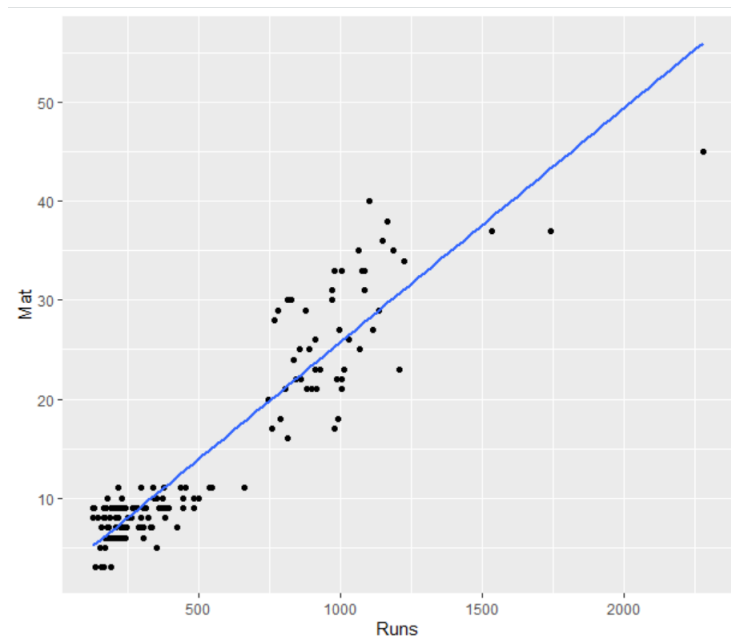


Figure 21: XY Plot : Runs & Mat

```
{r}
#Batsmen Form
batsmen = batsmen %>% mutate(batsmen_form = 0.4262*Ave + 0.2566*Inns + 0.1510*SR + 0.0787*Tons + 0.0556*Fifties -
0.0328*zeros) %>% arrange(desc(batsmen_form))
batsmen1 = batsmen %>% select(-Age, -Mat, -Inns, -NO, -Runs, -HS, -Ave, -BF, -SR, -Tons, -Fifties, -zeros, -fours, -sixes)
batsmen1
```

Player <chr>	Batting <chr>	Country <chr>	batsmen_form <dbl>
SR Watson	Left-hand	AUS	89.15178
AB de Villiers	Right-hand	SA	55.45843
KC Sangakkara	Left-hand	SL	54.58568
MJ Clarke	Right-hand	AUS	54.03842
Yuvraj Singh	Left-hand	IND	53.90689
BJ Hodge	Right-hand	AUS	52.59951
AB de Villiers	Right-hand	SA	50.94541
SB Styris	Right-hand	NZ	50.65273
SR Tendulkar	Right-hand	IND	50.23907
JH Kallis	Right-hand	SA	49.67506

1-10 of 149 rows

Previous  2 3 4 5 6 ... 15 Next

Figure 22: Batsman's Form in World Cups



## 5 Conclusion

Our main goal in this project is to develop a model to predict the outcome of ICC Cricket World Cup matches from the period (2007-2019). We have used the data based on batsmen, bowler and matches being played throughout the four World Cups. Also, we have used linear regression model to establish the relationship between two numerical variables using scatter plots and quantifying the relationship using Correlation to strengthen the relationship between two variables. This knowledge will help us in the future to design a better prediction model.

## 6 Future Work

There is still a lot of space for improvement in our decision making. In the future, we want to design a model that is much more efficient, significant and gives a minor error in predicting the winner of the next edition of the World Cup. Our model cannot handle the result of the matches which are interrupted by rain or other natural calamities. We would like to improve on that with the help of the *Duckworth-Lewis* method. Our future work will be to design a model that will perform prediction on every over.

# Bibliography

- [1] Madam Gopal Jhavar, Vikram Pudi *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2016)*. Conference Center, Riva del Garda, 2016.
- [2] Kalpdrum Passi, Niravkumar Pandey. *IJDKP Vol.8, No.2, March 2018 Increased Prediction Accuracy in the Game of Cricket Using Machine Learning*. Laurentian University, Sudbury, Canada
- [3] Balasaheb, T., Maman, P., Sandhu J.S. (2008) The impact of visual skills training on batting performance in cricketers *Serbian Journal of Sports Sciences*, 2(1), pp. 17-23
- [4] Brooks, R., Bussie're L. F. Jennions M. D & Hunt, J. (2003). Sinister strategies succeed at the cricket World Cup. *Proceedings of the Royal Society*
- [5] Morley, M., & Thomas, D. (2005) An investigation of home advantage and other factors affecting outcomes in English One-Day Cricket matches *Journal of Sports Science*, 23(3):261-168
- [6] J. R. Quinlan, C4.5: Programs for Machine Learning, Elsevier, 2015
- [7] P. Shah "New Performance measure in Cricket" *ISOR Journal of Sports and Physical Education*, Vol.4, No.3, pp. 28-30, 2017