

ADA_End-Term Project_Group6

Yukti Dewan, Anchal Gupta, Chandni Kaundilya, Rishu Verma, Akshaykumar Mistry

2023-01-01

Project Chosen

The topic we have chosen for our project is **IPL Super 11**. In this project we would be creating our group's dream cricket team by analyzing the data given.



Before starting with our auction process, we would first like to mention the rules and assumptions defined by us for creating the dream team.

Assumptions Taken

- The number of matches played for a player in IPL as given in the data set are less compared to what they have actually played in IPL when checked on internet, So we choose our Super 11 based on the dataset given.
- We have excluded the retired IPL players based on current news.

Rules for Team Formation

- A team is supposed to have 12 players available - 11 on the field, of which at least seven must be Indian, and one substitute - for selection on a matchday.
- Only 4 players in a team can be foreigners.

Auction Process

- In this Auction process, we are selecting an super team for IPL based on a datasets from 2008-2020 containing all ball-by-ball and match-by-match order.

- We have selected players who have played or bowled more than 300 balls only.
- Our team combination will be consisting of strong openers, middle-order hitters, all-rounders, a wicketkeeper, and quick and spin bowlers.
- For this, each player's past IPL performance will be analysed including number of balls played/bowled, runs scored/given, strike rate/economy, 4's and 6's, stumps/catches, etc and finally we will be selecting our Super 11.

Team Combination

Batsmen



- **5 Batsmen**
 - 2 Opener Batsmen
 - 3 Middle Order Batsmen

Bowlers



- 4 Bowlers

- 1 Spinner
- 2 Fast Bowler
- 1 Medium Pacer

All-Rounder



- 2 All-Rounders

- 1 Spinner + Middle-order batsman
- 1 Fast Bowler + Middle-Order batsman

Missing Value Analysis

```
missing_Q1 <- colSums(is.na(Q1))
missing_Q2 <- colSums(is.na(Q2))
missing_Q1
```

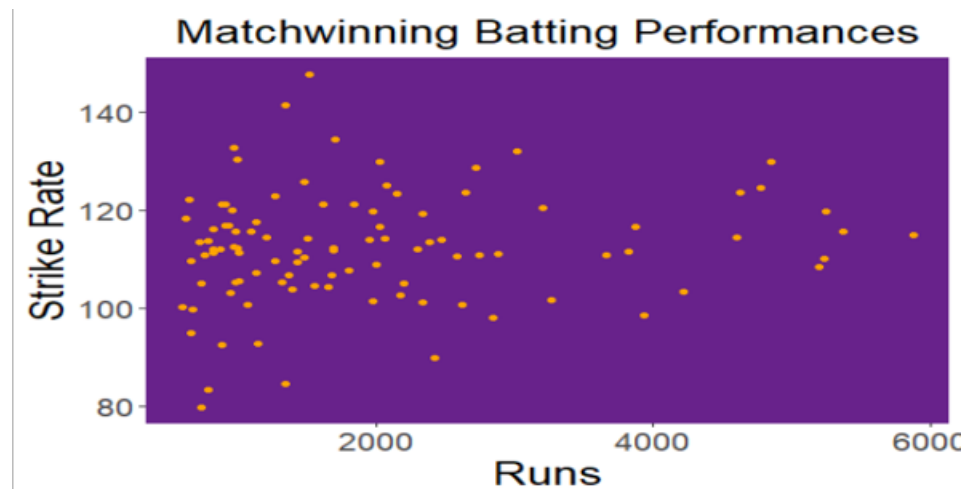
```
      id      inning      over      ball      batsman      non_striker
      0          0          0          0          0          0
    bowler  batsman_runs  extra_runs  total_runs  non_boundary  is_wicket
      0          0          0          0          0          0
dismissal_kind player_dismissed  fielder  extras_type  batting_team  bowling_team
183973      183973      186684      183235          0          191
```

```
missing_Q2
      id      city      date player_of_match      venue  neutral_venue
      0          13          0          4          0          0
    team1      team2  toss_winner  toss_decision      winner      result
      0          0          0          0          4          4
result_margin  eliminator      method      umpire1  umpire2
      17          4          797          0          0
```

- We have not removed the missing values as there be significant loss in data.
- It is also difficult to Impute.

Exploratory Data Analysis

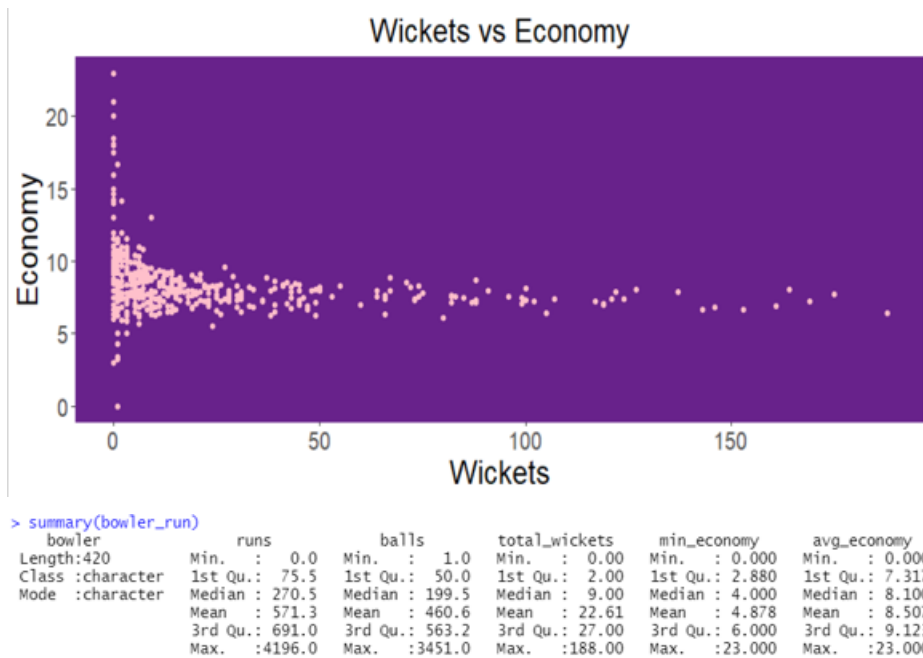
Batsmen - Scatter Plot



```
> summary(batter_run)
      batsman      runs      balls  total_matches  max_strike_rate avg_strike_rate
Length:100   Min.   : 604.0   Min.   : 466   Min.   : 21.00   Min.   :166.7   Min.   : 79.71
Class :character 1st Qu.: 982.2   1st Qu.: 769   1st Qu.: 50.00   1st Qu.:220.0   1st Qu.:105.23
Mode :character  Median :1535.5   Median :1234   Median : 70.00   Median :266.7   Median :112.06
                Mean   :1973.2   Mean   :1552   Mean   : 81.33   Mean   :283.0   Mean   :112.35
                3rd Qu.:2439.5   3rd Qu.:1931   3rd Qu.: 98.00   3rd Qu.:320.9   3rd Qu.:119.39
                Max.   :5878.0   Max.   :4609   Max.   :194.00   Max.   :600.0   Max.   :147.85
```

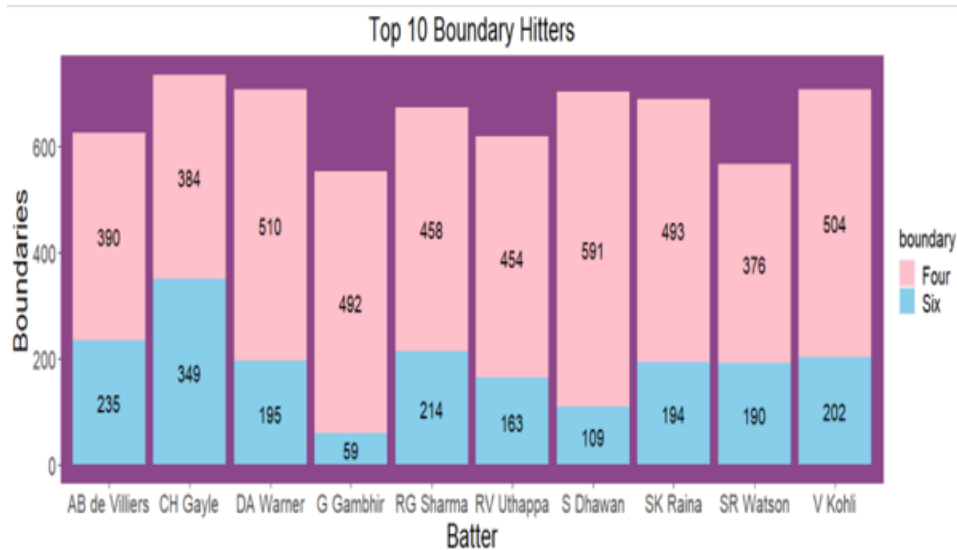
Average Strike Rate should be 112.35

Bowler - Scatter Plot



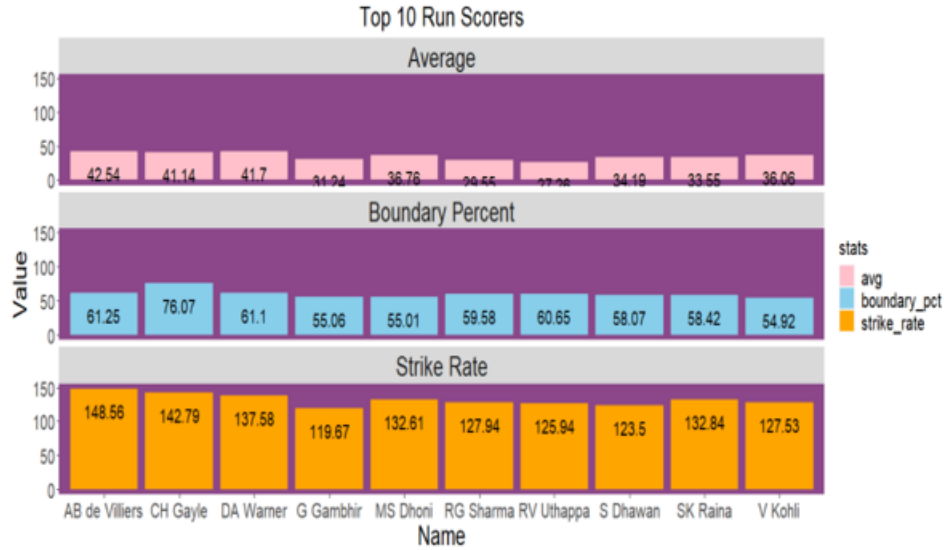
Average economy should be 8.507

Top 10 Hitters



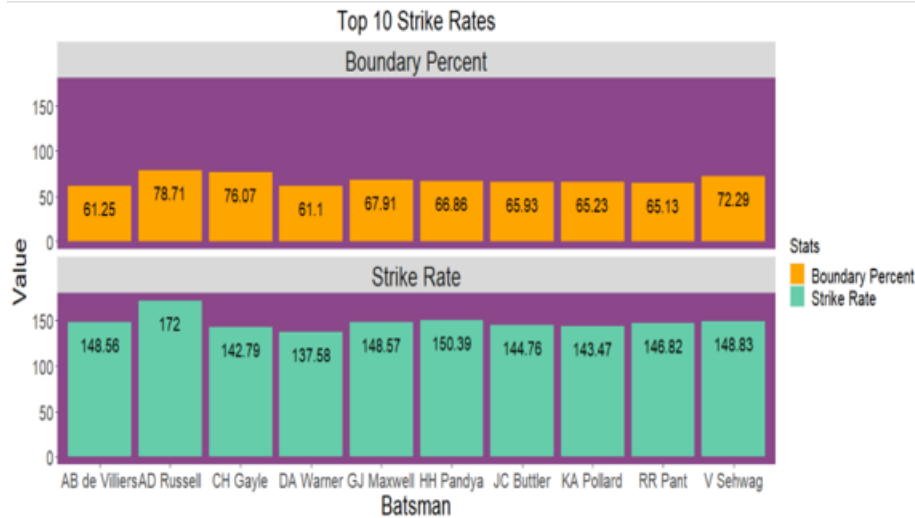
Most players hit way more fours than sixes

Top 10 Run Scorers



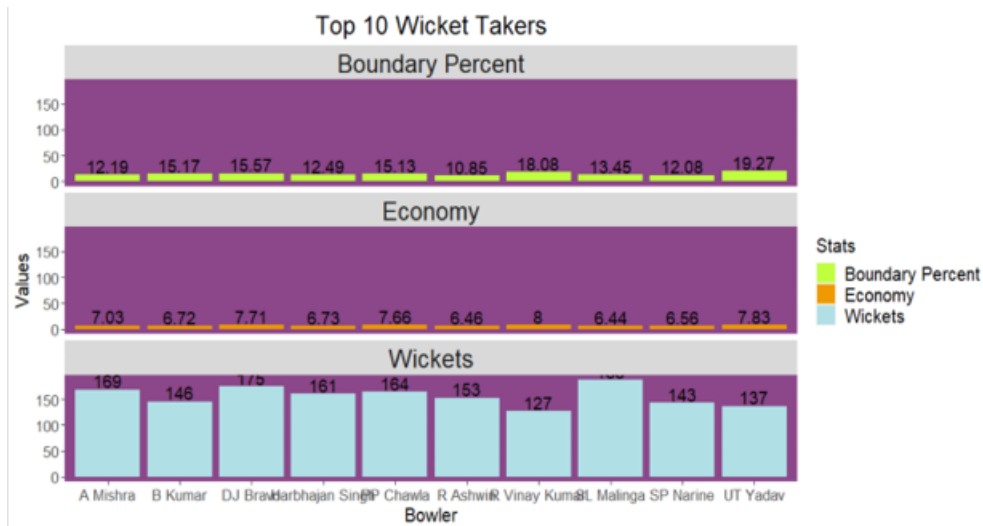
Top 10 scorers with average run, boundary percentage and strike rate

Top 10 Batsmen by Strike Rate



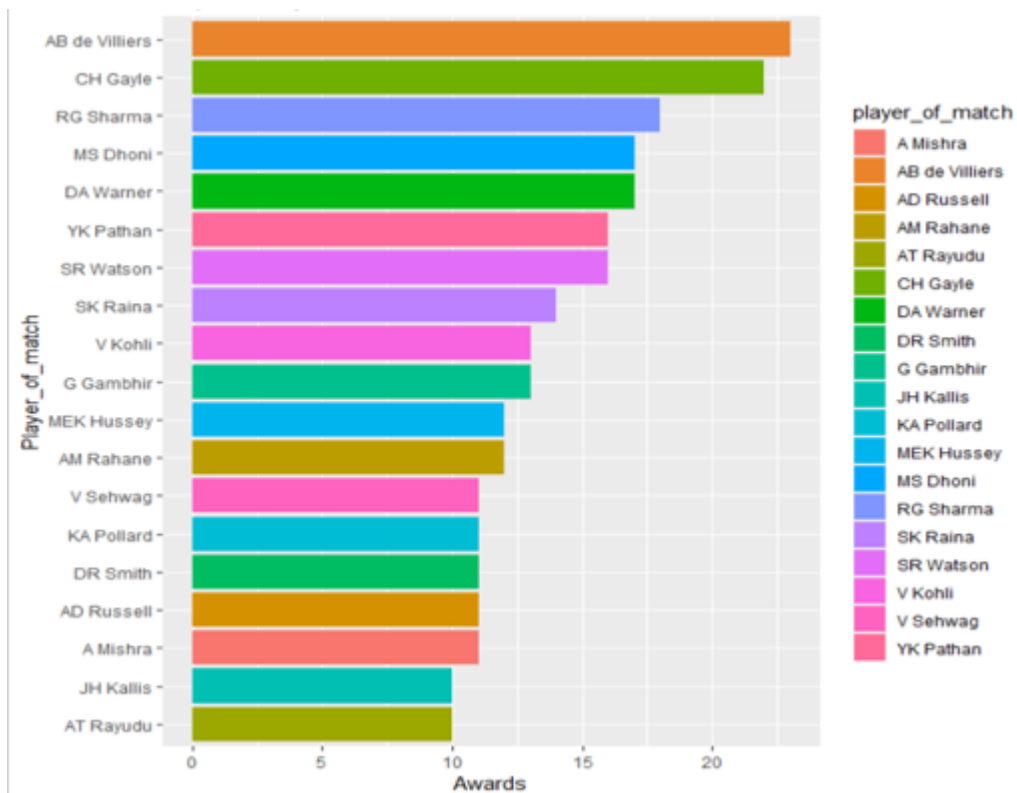
Top 10 batsmen with boundary percentage and strike rate

Top 10 Bowlers by Wickets Taken



Top 10 bowlers with boundary percentage, economy and wickets

Top “Man of the Match” Players



AB de Villiers is the player in IPL to have maximum “Man of the Match” award

Batsmen Selection Criteria

	batsman	runs	balls	total_matches	max_strike_rate	avg_strike_rate
1	V Kohli	5878	4609	184	357.14	114.82
2	SK Raina	5368	4041	189	334.62	115.74
3	DA Warner	5254	3819	142	223.08	119.70
4	RG Sharma	5230	4088	194	350.00	110.11
5	S Dhawan	5197	4208	175	200.00	108.33
6	AB de Villiers	4849	3264	156	344.44	129.90
7	CH Gayle	4772	3342	131	318.18	124.47
8	MS Dhoni	4632	3493	182	281.25	123.67
9	RV Uthappa	4607	3658	182	244.44	114.50
10	G Gambhir	4217	3524	151	220.00	103.44
11	AM Rahane	3933	3325	140	189.47	98.55
12	SR Watson	3874	2888	141	276.92	116.52
13	KD Karthik	3823	3023	176	266.67	111.64
14	AT Rayudu	3659	2970	149	400.00	110.92
15	MK Pandey	3268	2772	134	300.00	101.72
16	YK Pathan	3204	2330	153	327.27	120.50
17	KA Pollard	3023	2107	144	400.00	132.03
18	BB McCullum	2880	2267	109	241.67	110.97

player_of_match	awards
AB de Villiers	23
CH Gayle	22
RG Sharma	18
DA Warner	17
MS Dhoni	17
SR Watson	16
YK Pathan	16
SK Raina	14
G Gambhir	13
V Kohli	13
AM Rahane	12
MEK Hussey	12
A Mishra	11

batsman	Four	Six
CH Gayle	384	349
AB de Villiers	390	235
MS Dhoni	313	216
RG Sharma	458	214
V Kohli	504	202
KA Pollard	196	198
DA Warner	510	195
SK Raina	493	194
SR Watson	376	190
RV Uthappa	454	163
YK Pathan	262	159
Yuvraj Singh	217	149
AT Rayudu	308	132
BB McCullum	293	130
AD Russell	105	129

Batsman/Criteria	Batsman Runs	Balls Played	Total Match	Strike Rate	Player of the match	Batting Type		
AB de Villiers	4849	3264	156	129.9	23	Right Handed	Opener	Wicket Keeper(Stumps:8)
CH Gayle	4772	3342	131	124.47	22	Left Handed		
DA Warner	5254	3819	142	119.7	17	Left Handed		
RG Sharma	5230	4088	194	110.11	18	Right Handed		
V Kohli	5878	4609	184	114.82	13	Right Handed		

- Firstly we took no. of runs scored as the first priority parameter.
- Next we took strike rate as the second priority parameter.
- Last priority criteria was player of match.
- Wicket Keeper selected is based on no. of stumps max out of batsmen.

Bowlers Selection Criteria

	bowler	runs	balls	total_wickets	min_economy	avg_economy
1	SL Malinga	3193	2974	188	2.00	6.42
2	DJ Bravo	3659	2846	175	0.00	7.69
3	A Mishra	3788	3233	169	2.00	7.22
4	PP Chawla	4196	3285	164	3.60	8.09
5	Harbhajan Singh	3868	3451	161	2.25	6.94
6	R Ashwin	3581	3327	153	0.92	6.70
7	B Kumar	3132	2795	146	2.64	6.82
8	SP Narine	3086	2824	143	2.40	6.64
9	UT Yadav	3446	2642	137	2.16	7.91
10	R Vinay Kumar	2906	2179	127	3.43	8.05
11	YS Chahal	2649	2188	124	0.00	7.38
12	RA Jadeja	3408	2759	122	0.00	7.85
13	A Nehra	2422	1974	121	1.50	7.36
14	JJ Bumrah	2513	2159	119	2.33	6.96
15	Z Khan	2691	2276	119	0.00	7.05
16	Sandeep Sharma	2575	2128	117	2.53	7.26
17	SR Watson	2571	2137	107	1.71	7.42
18	DW Steyn	2391	2276	105	1.20	6.43

bowler	maiden_over
P Kumar	12
IK Pathan	9
DW Steyn	8
SL Malinga	8
B Kumar	7
DS Kulkarni	7
DJ Bravo	6
Sandeep Sharma	6
DL Chahar	5
R Ashwin	5
Z Khan	5
A Mishra	4
Harbhajan Singh	4
IC Pandey	4
JJ Bumrah	4

fielder	dismissal_kind	count_dismissal
MS Dhoni	stumped	39
RV Uthappa	stumped	32
KD Karthik	stumped	30
WP Saha	stumped	20
AC Gilchrist	stumped	16
PA Patel	stumped	16
Q de Kock	stumped	12
RR Pant	stumped	11
NV Ojha	stumped	10
KC Sangakkara	stumped	9
AB de Villiers	stumped	8
KM Jadhav	stumped	7
MS Bisla	stumped	7
SP Goswami	stumped	7
BB McCullum	stumped	6

Bowlers/Criteria	Runs Given	Balls Bowled	Wickets taken	Economy	Madein Overs	Bowling Type
A Mishra	3788	3233	169	7.22	4	Leg Spinner
B Kumar	3132	2795	146	6.82	7	Fast Bowler
UT Yadav	3446	2642	137	7.91	3	Fast Bowler
Sandeep Sharma	2575	2128	117	7.26	6	Medium Pacer

- Firstly we took no. of wickets taken(max) as our top priority selection criteria.
- Economy (calculated using formula = runs/no. of balls(6)) was the next priority criteria.

All-Rounders Selection Criteria

	batsman	runs	balls	total_matches	max_strike_rate	avg_strike_rate
1	V Kohli	5878	4609	184	357.14	114.82
2	SK Raina	5368	4041	189	334.62	115.74
3	DA Warner	5254	3819	142	223.08	119.70
4	RG Sharma	5230	4088	194	350.00	110.11
5	S Dhawan	5197	4208	175	200.00	108.33
6	AB de Villiers	4849	3264	156	344.44	129.90
7	CH Gayle	4772	3342	131	318.18	124.47
8	MS Dhoni	4632	3493	182	281.25	123.67
9	RV Uthappa	4607	3658	182	244.44	114.50
10	G Gambhir	4217	3524	151	220.00	103.44
11	AM Rahane	3933	3325	140	189.47	98.55
12	SR Watson	3874	2888	141	276.92	116.52
13	KD Karthik	3823	3023	176	266.67	111.64
14	AT Rayudu	3659	2970	149	400.00	110.92
15	MK Pandey	3268	2772	134	300.00	101.72
33	RA Jadeja	2159	1774	136	400.00	123.37

	bowler	runs	balls	total_wickets	min_economy	avg_economy
1	SL Malinga	3193	2974	188	2.00	6.42
2	DJ Bravo	3659	2846	175	0.00	7.69
3	A Mishra	3788	3233	169	2.00	7.22
4	PP Chawla	4196	3285	164	3.60	8.09
5	Harbhajan Singh	3868	3451	161	2.25	6.94
6	R Ashwin	3581	3327	153	0.92	6.70
7	B Kumar	3132	2795	146	2.64	6.82
8	SP Narine	3086	2824	143	2.40	6.64
9	UT Yadav	3446	2642	137	2.16	7.91
10	R Vinay Kumar	2906	2179	127	3.43	8.05
11	YS Chahal	2649	2188	124	0.00	7.38
12	RA Jadeja	3408	2759	122	0.00	7.85
13	A Nehra	2422	1974	121	1.50	7.36
14	JJ Bumrah	2513	2159	119	2.33	6.96
15	Z Khan	2691	2276	119	0.00	7.05
16	Sandeep Sharma	2575	2128	117	2.53	7.26
17	SR Watson	2571	2137	107	1.71	7.42
18	DIW Steyn	2391	2276	105	1.20	6.43

All Rounder/Criteria	Runs Given	Balls Bowled	Wickets taken	Economy	Made in Overs	Batsman Runs	Balls Played	Total Matches	Strike Rate	Bowling Type
RA Jadeja	3408	2759	122	7.85	1	2159	1774	136	123.37	Spinner
SR Watson	2571	2137	107	7.42	2	3874	2888	141	116.52	Fast Bowler

- The criteria selected was to choose a player which was above average bowler and batsmen both.
- Top bowlers were searched in the top 50 batsmen list.
- Matching top players were chosen as the all-rounders.

Extra Player Selection Criteria

	batsman	runs	balls	total_matches	max_strike_rate	avg_strike_rate
1	V Kohli	5878	4609	184	357.14	114.82
2	SK Raina	5368	4041	189	334.62	115.74
3	DA Warner	5254	3819	142	223.08	119.70
4	RG Sharma	5230	4088	194	350.00	110.11
5	S Dhawan	5197	4208	175	200.00	108.33
6	AB de Villiers	4849	3264	156	344.44	129.90
7	CH Gayle	4772	3342	131	318.18	124.47
8	MS Dhoni	4632	3493	182	281.25	123.67
9	RV Uthappa	4607	3658	182	244.44	114.50
10	G Gambhir	4217	3524	151	220.00	103.44
11	AM Rahane	3933	3325	140	189.47	98.55
12	SR Watson	3874	2888	141	276.92	116.52
13	KD Karthik	3823	3023	176	266.67	111.64
14	AT Rayudu	3659	2970	149	400.00	110.92
15	MK Pandey	3268	2772	134	300.00	101.72
33	RA Jadeja	2159	1774	136	400.00	123.37

bowler	runs	balls	total_wickets	min_economy	avg_economy
S Dhawan	65	49	4	4.00	7.83

Extra Player/Criteria	Runs Given	Balls Bowled	Wickets taken	Economy	Madein Overs	Batsman Runs	Balls Played	Total Matches	Strike Rate	Batting Type
S Dhawan/Extras	65	49	4	7.83	0	5197	4208	175	108.33	Left Handed

- The criteria taken for selecting All-Rounders was taken to select the extra player.

Our Final Team - Uttrakhand Unicorns

UTTRAKHAND UNICORNS

01 AB DE VILLIERS (WC)



AB de Villiers

Role: Wicketkeeper Batter
Nationality: South African
DOB: 17 February 1984
IPL Debut: 2008

02 CHRIS GAYLE



Chris Gayle

Role: All-rounder
Nationality: West Indian
DOB: 21 September 1979
IPL Debut: 2008

03 ROHIT SHARMA



Rohit Sharma

Role: Batter
Nationality: Indian
DOB: 30 April 1987
IPL Debut: 2008

04 VIRAT KOHLI



Virat Kohli

Role: Batter
Nationality: Indian
DOB: 05 November 1988
IPL Debut: 2008

05 DAVID WARNER



David Warner

Role: Batter
Nationality: Australian
DOB: 27 October 1986
IPL Debut: 2009

06 RAVINDRA JADEJA



Ravindra Jadeja

Role: All-rounder
Nationality: Indian
DOB: 06 December 1988
IPL Debut: 2008

07 SHANE WATSON



Shane Watson

Role: All-rounder
Nationality: Australian
DOB: 17 June 1983
IPL Debut: 2008

08 AMIT MISHRA



Amit Mishra

Role: Bowler
Nationality: Indian
DOB: 24 November 1982
IPL Debut: 2008

09 BHUVNESHWAR KUMAR



Bhuvneshwar Kumar

Role: Bowler
Nationality: Indian
DOB: 05 February 1990
IPL Debut: 2012

10 UMESH YADAV



Umesh Yadav

Role: Bowler
Nationality: Indian
DOB: 25 October 1987
IPL Debut: 2010

11 SANDEEP SHARMA



Sandeep Sharma

Role: Bowler
Nationality: Indian
DOB: 12 May 1993
IPL Debut: 2013

12 SHIKHAR DHAWAN EXTRA



Shikhar Dhawan

Role: Batter
Nationality: Indian
DOB: 05 December 1983
IPL Debut: 2008

Logistic Regression

```
Call: glm(formula = winner_toss_winner ~ cat_toss_decision, family = binomial,
  data = toss_decision_obs)
```

Coefficients:

```
(Intercept)  cat_toss_decision
-0.1823      0.3982
```

Degrees of Freedom: 811 Total (i.e. Null); 810 Residual

Null Deviance: 1125

Residual Deviance: 1117 AIC: 1121

- When the team wins the toss, then compared to choosing to 'Bat' if the team chooses to 'Field' then the probability of winning the match increases by 0.39 i.e. 39.82%.

Thank You

