

IPL STATISTICS DASHBOARD



Submitted By:

Jitesh Garg
3CS8
102017180

Submitted To:

Mrs. Kashish Goyal


July 2022 – December 2022


OVERVIEW OF THE PROJECT


- The purpose of the project was to create innovative and interactive Tableau dashboards.
- Developed a IPL statistics Dashboard to analysis about the players performance in different seasons of IPL on the basis of Top Runs, Highest Wickets and so on.
- The data to be cleaned using R. The Dataset contained missing values and was cleaned using the R programming language.
- Tableau dashboards were created from the cleaned dataset.

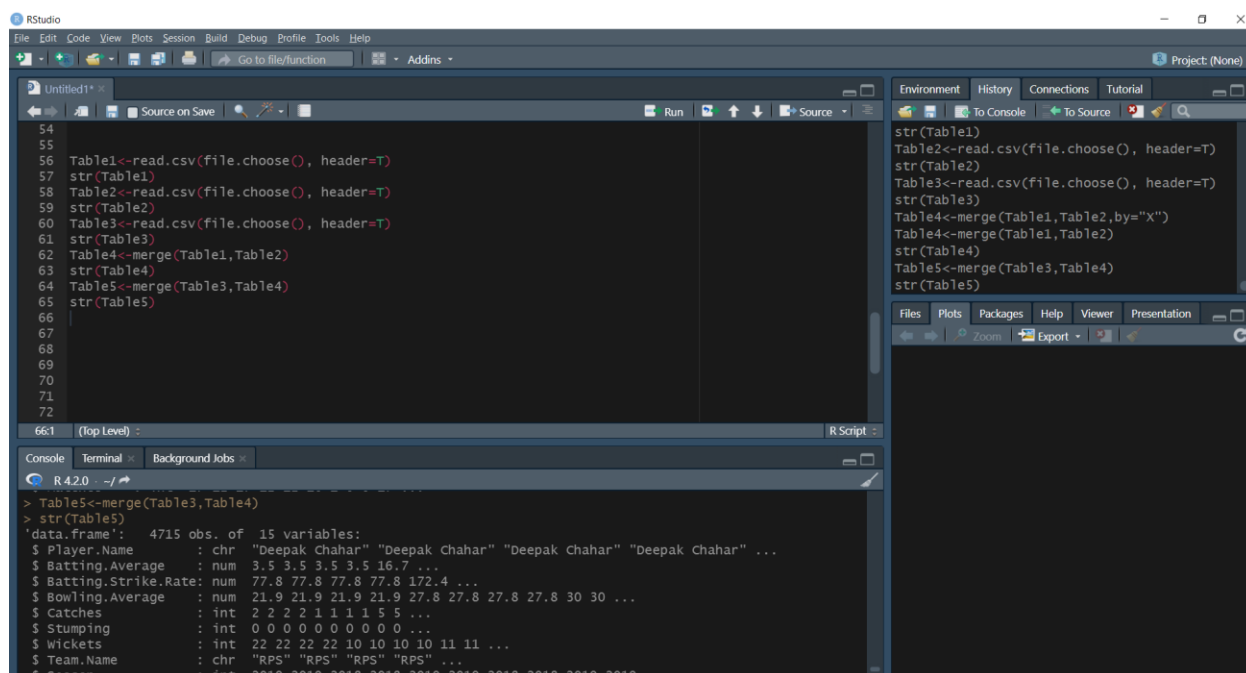
SPLITTING AND MERGING OF DATASET

Name

 Table1

 Table2

 Table3



```

54
55
56 Table1<-read.csv(file.choose(), header=T)
57 str(Table1)
58 Table2<-read.csv(file.choose(), header=T)
59 str(Table2)
60 Table3<-read.csv(file.choose(), header=T)
61 str(Table3)
62 Table4<-merge(Table1,Table2)
63 str(Table4)
64 Table5<-merge(Table3,Table4)
65 str(Table5)
66
67
68
69
70
71
72
66:1 (Top Level) : R Script :
Console Terminal Background Jobs
R 4.2.0 ~/
> Table5<-merge(Table3,Table4)
> str(Table5)
'data.frame': 4715 obs. of 15 variables:
 $ Player.Name : chr "Deepak Chahar" "Deepak Chahar" "Deepak Chahar" "Deepak Chahar" ...
 $ Batting.Average : num 3.5 3.5 3.5 3.5 16.7 ...
 $ Batting.Strike.Rate: num 77.8 77.8 77.8 77.8 172.4 ...
 $ Bowling.Average : num 21.9 21.9 21.9 21.9 27.8 27.8 27.8 27.8 30 30 ...
 $ Catches : int 2 2 2 2 1 1 1 1 5 5 ...
 $ Stumping : int 0 0 0 0 0 0 0 0 0 0 ...
 $ Wickets : int 22 22 22 22 10 10 10 10 11 11 ...
 $ Team.Name : chr "RPS" "RPS" "RPS" "RPS" ...
 $ Season : chr 2018 2018 2018 2018 2018 2018 2018 2018 2018 2018

```

Raw Dataset:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Highest	PlayerNam	TeamNam	Season	BattingAve	BattingStri	BowlingAv	Catches	Fifties	Fours	Matches	Runs	Sixes	Stumping	Wickets								
2	100	Shreyas Iy MI		2018	43	149.75	0	2	3	53	16	250	34	0	0								
3	53	MS Dhoni CSK		2015	25	155	5	90	50	NA	17	400	NA	7	5								
4	57	Shreyas Iy MI		2019	23.5	93.06	0	6	1	20	17	282	7	0	0								
5	100	Shreyas Iy MI		2018	43	149.75	0	2	3	53	16	602	34	0	0								
6	7	Deepak Ch RPS		2015	3.5	85	25	2	ZERO	1	17	NA	0	0	NA								
7	7	Deepak Ch RPS		2019	3.5	77.77	21.9	2	0	1	17	7	0	0	22								
8	39	Deepak Ch RPS		2018	16.66	172.41	27.8	1	0	1	12	50	4	0	10								
9	27	Hardik Pan KKR		2019	16	121.21	30	5	0	6	12	80	3	0	11								
10	68	Hardik Pan KKR		2018	35.25	154.94	38.07	9	1	8	16	141	10	0	14								
11	0	Irfan Pathu SRH		2015	NA	0	0	0	NA	0	0	0	0	NA	0								
12	8	Hardik Pan KKR		2014	0	266.66	0	0	0	0	1	8	1	0	0								
13	6	Hardik Pan KKR		2011	11	183.33	26.16	2	0	1	6	10	1	0	6								
14	6	Hardik Pan KKR		2011	11	183.33	26.16	3	0	1	6	11	1	0	6								
15	62	Hardik Pan KKR		2015	21.66	131.75	16.38	13	1	22	17	195	5	0	26								
16	48	Hardik Pan KKR		2012	46.37	140.53	30.73	9	0	20	19	371	20	0	15								
17	15	Shardul Th CSK		2015	0	300	26.93	4	NA	3	13	15	0	NA	16								
18	23	Hardik Pan KKR		2013	15.12	114.15	15.53	14	0	11	18	121	3	0	32								
19	96	Rohit Shar IOIP		2019	36	123.36	0	12	3	36	12	396	15	0	0								
20	67	Rohit Shar IOIP		2018	32.4	125.58	0	1	1	17	6	162	6	0	0								
21	0	Rohit Shar IOIP		2011	0	0	0	0	0	0	0	0	0	0	0								
22	54	Rohit Shar IOIP		2014	27.54	128.93	0	13	2	26	15	303	7	0	0								
23	55	Rohit Shar IOIP		2015	29.23	125	0	8	1	37	17	380	5	0	0								
24	73	Rohit Shar IOIP		2012	33.16	130.92	0	4	3	29	13	398	17	0	0								
25	31	Ravindra J. DC		2019	NA	120.45	22.86	9	NA	7	16	NA	4	0	NA								
26	0	Rohit Shar IOIP		2013	0	0	0	0	0	0	0	0	0	0	0								
27	0	Irfan Pathu SRH		2015	0	0	0	0	0	0	0	0	0	0	0								

Data Preprocessing:

```

1 ip1_data <- read.csv(file.choose(), header=T)
2 head(ip1_data)
3 dim(ip1_data)
4
5 complete.cases(ip1_data)
6 ip1_data_cleaned <- complete.cases(ip1_data)
7
8 ip1_data <- ip1_data[ip1_data_cleaned,]
9 head(ip1_data)
10 complete.cases(ip1_data)
11 dim(ip1_data)
12
13
14 str(ip1_data)
15
16
17 sum(is.na(ip1_data$Runs))
18
19

```

```

for (i in 1:nrow(ip1_data)) {
  if(!typeof(ip1_data$Runs[i])=='integer')
  {
    print('Value wrong at=')
    print(i)
  }
}
str(ip1_data)
str(ip1_data)
sum(is.na(ip1_data$Runs))

```

```

R 4.2.0 > sum(is.na(ip1_data$Runs))
[1] 2
R 4.2.0 >

```

Before removing NA values:

```

1 ip1_data <- read.csv(file.choose(), header=T)
2 head(ip1_data)
3 dim(ip1_data)
4
5 complete.cases(ip1_data)
6 ip1_data_cleaned<-complete.cases(ip1_data)
7
8 ip1_data<-ip1_data[ip1_data_cleaned,]
9 head(ip1_data)
10 complete.cases(ip1_data)
11 dim(ip1_data)
12

```

```

R 4.2.0 >
6      7 Deepak Chahar      RPS      2019      3.5
   Batting.Strike.Rate Bowling.Average Catches Fifties Fours Matches Runs Sixes Stumping Wickets
1      149.75      0.0      2      3      53      16      250      34      0      0
2      155.00      5.0      90      50      NA      17      400      NA      7      5
3      93.06      0.0      6      1      20      17      282      7      0      0
4      149.75      0.0      2      3      53      16      602      34      0      0
5      85.00      25.0      2      ZERO      1      15      NA      5      4      NA
6      77.77      21.9      2      0      1      17      7      0      0      22
> dim(ip1_data)
[1] 64 15
>

```

After removing NA values:

```

1 ip1_data <- read.csv(file.choose(), header=T)
2 head(ip1_data)
3 dim(ip1_data)
4
5 complete.cases(ip1_data)
6 ip1_data_cleaned<-complete.cases(ip1_data)
7
8 ip1_data<-ip1_data[ip1_data_cleaned,]
9 head(ip1_data)
10 complete.cases(ip1_data)
11 dim(ip1_data)
12

```

```

R 4.2.0 >
6      77.77      21.9      2      0      1      17      7      0      0      22
7      172.41      27.8      1      0      1      12      50      4      0      10
8      121.21      30.0      5      0      6      12      80      3      0      11
> complete.cases(ip1_data)
[1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
[20] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
[39] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
[58] TRUE TRUE
> dim(ip1_data)
[1] 59 15
>

```

Check if all columns have valid values:

```

13
14
15 for (i in 1:nrow(ipl_data)) {
16   if(!typeof(ipl_data$BattingAverage[i])=='double')
17   {
18     print('Value wrong at=')
19     print(i)
20   }
21 }
22
23 for (i in 1:nrow(ipl_data)) {
24   if(!typeof(ipl_data$Fifties[i])=='double')
25   {
26     print('Value wrong at=')
27     print(i)
28   }
29 }
30
31

```

```

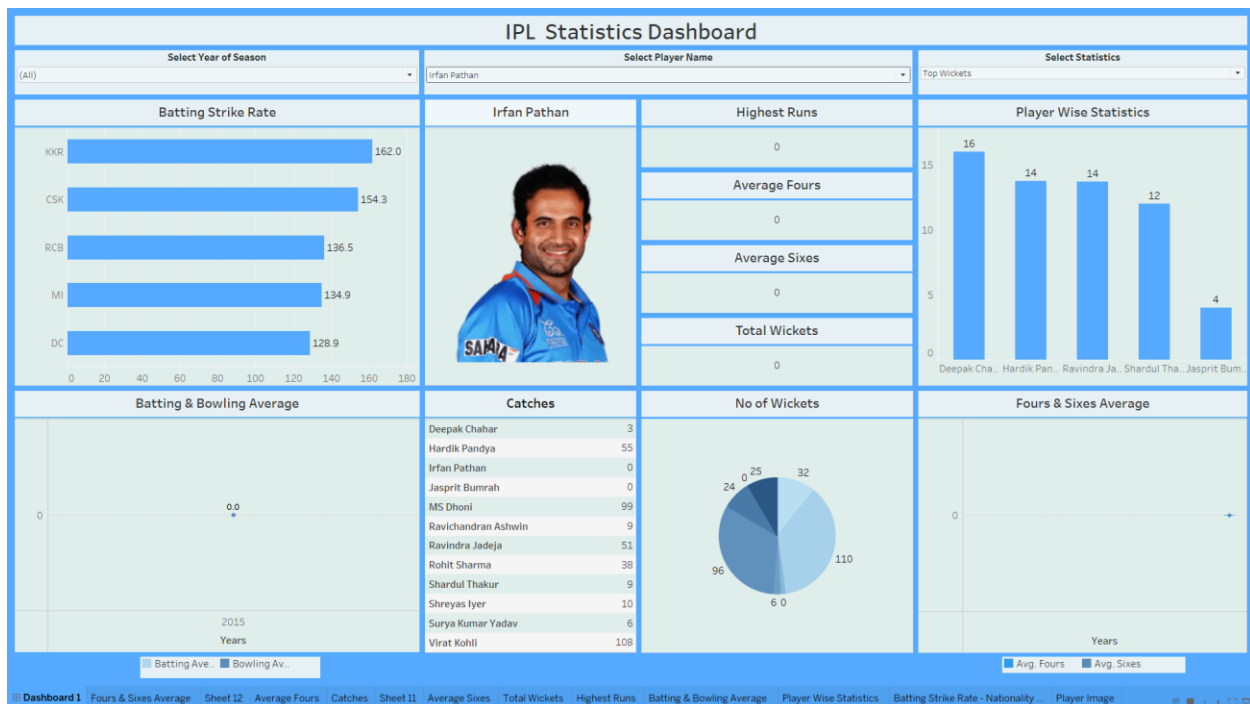
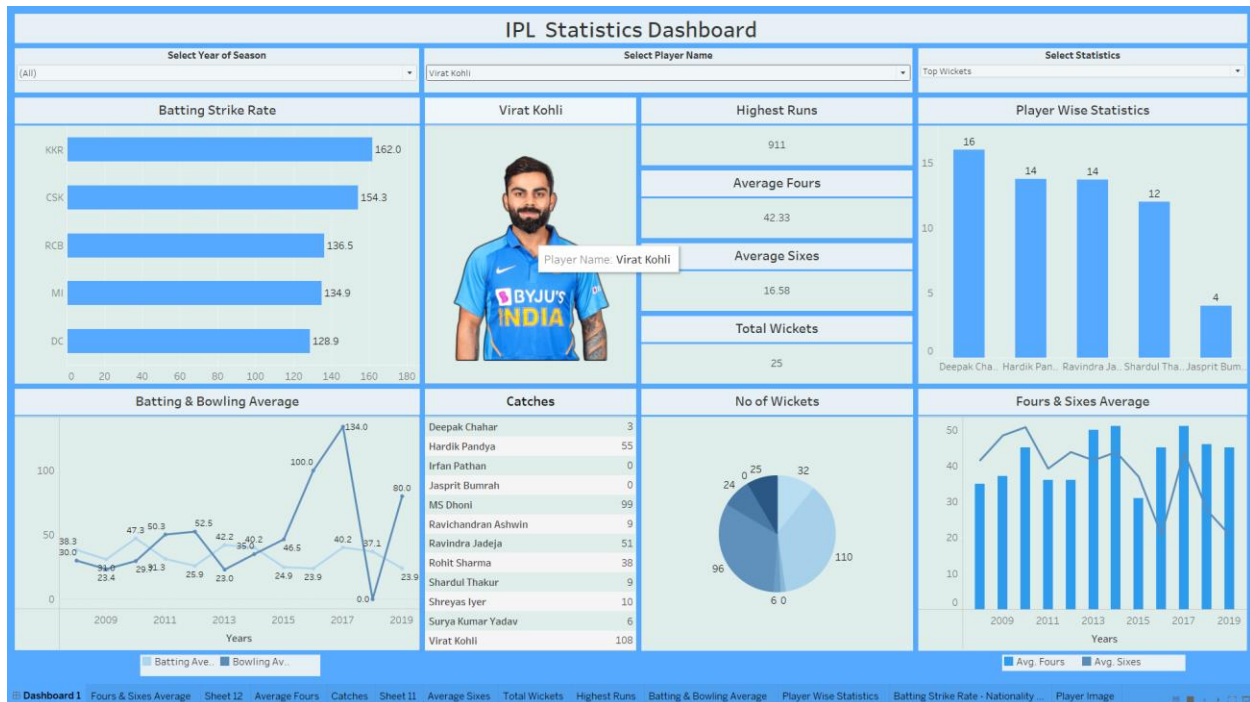
+ print(i)
+ }
+ }
+ }
> for (i in 1:nrow(ipl_data)) {
+   if(!typeof(ipl_data$BattingAverage[i])=='double')
+   {
+     print('Value wrong at=')
+     print(i)
+   }
+ }
+ }
+ }
>

```

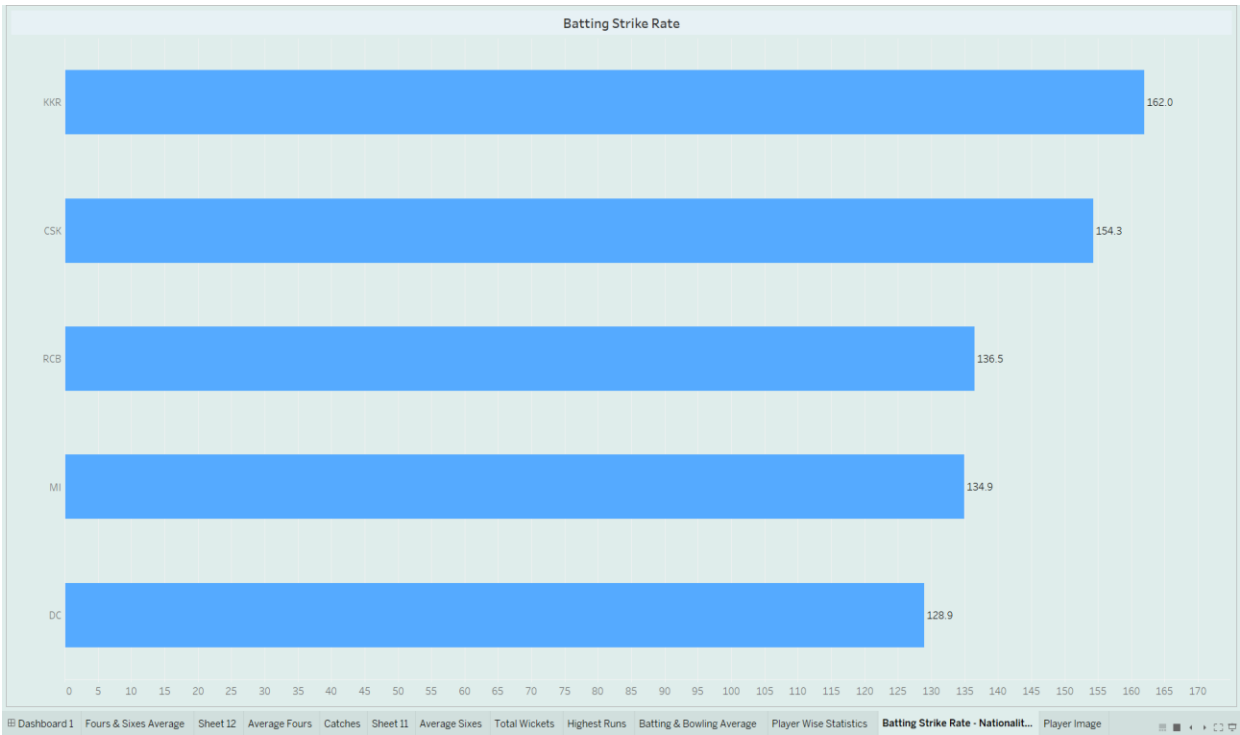
Final Dataset:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Highest	PlayerNam	TeamNam	Season	BattingAve	BattingStri	BowlingAv	Catches	Fifties	Fours	Matches	Runs	Sixes	Stumpings	Wickets								
2	100	Shreyas Iyi	MI	2018	43	149.75	0	2	3	53	16	250	34	0	0								
3	57	Shreyas Iyi	MI	2019	23.5	93.06	0	6	1	20	17	282	7	0	0								
4	100	Shreyas Iyi	MI	2018	43	149.75	0	2	3	53	16	602	34	0	0								
5	7	Deepak Ch RPS	2019	3.5	77.77	21.9	2	0	1	17	7	0	0	22									
6	39	Deepak Ch RPS	2018	16.66	172.41	27.8	1	0	1	12	50	4	0	10									
7	27	Hardik Pani	KKR	2019	16	121.21	30	5	0	6	12	80	3	0	11								
8	68	Hardik Pani	KKR	2018	35.25	154.94	38.07	9	1	8	16	141	10	0	14								
9	8	Hardik Pani	KKR	2014	0	266.66	0	0	0	0	1	8	1	0	0								
10	6	Hardik Pani	KKR	2011	11	183.33	26.16	2	0	1	6	10	1	0	6								
11	6	Hardik Pani	KKR	2011	11	183.33	26.16	3	0	1	6	11	1	0	6								
12	62	Hardik Pani	KKR	2015	21.66	131.75	16.38	13	1	22	17	195	5	0	26								
13	48	Hardik Pani	KKR	2012	46.37	140.53	30.73	9	0	20	19	371	20	0	15								
14	23	Hardik Pani	KKR	2013	15.12	114.15	15.53	14	0	11	18	121	3	0	32								
15	96	Rohit Shar	KKR	2019	36	123.36	0	12	3	36	12	396	15	0	0								
16	67	Rohit Shar	KKR	2018	32.4	125.58	0	1	1	17	6	162	6	0	0								
17	0	Rohit Shar	KKR	2011	0	0	0	0	0	0	0	0	0	0	0								
18	54	Rohit Shar	KKR	2014	27.54	128.93	0	13	2	26	15	303	7	0	0								
19	55	Rohit Shar	KKR	2015	29.23	125	0	8	1	37	17	380	5	0	0								
20	73	Rohit Shar	KKR	2012	33.16	130.92	0	4	3	29	13	398	17	0	0								
21	0	Rohit Shar	KKR	2013	0	0	0	0	0	0	0	0	0	0	0								
22	0	Irfan Pathi	SRH	2015	0	0	0	0	0	0	0	0	0	0	0								
23	84	MS Dhoni	CSK	2019	83.2	134.62	0	11	3	22	15	416	23	5	0								
24	79	MS Dhoni	CSK	2018	75.83	150.66	0	11	3	24	16	455	30	3	0								
25	70	MS Dhoni	CSK	2011	43.55	158.7	0	5	2	25	16	392	23	5	0								
26	84	MS Dhoni	CSK	2016	83.2	134.62	0	11	3	22	15	416	23	5	0								
27	57	MS Dhoni	CSK	2014	74.2	148.4	0	3	1	22	16	371	20	1	0								
28	53	MS Dhoni	CSK	2017	31	121.96	0	8	1	27	17	372	17	3	0								
29	51	MS Dhoni	CSK	2012	29.83	128.77	0	12	1	26	19	358	9	2	0								

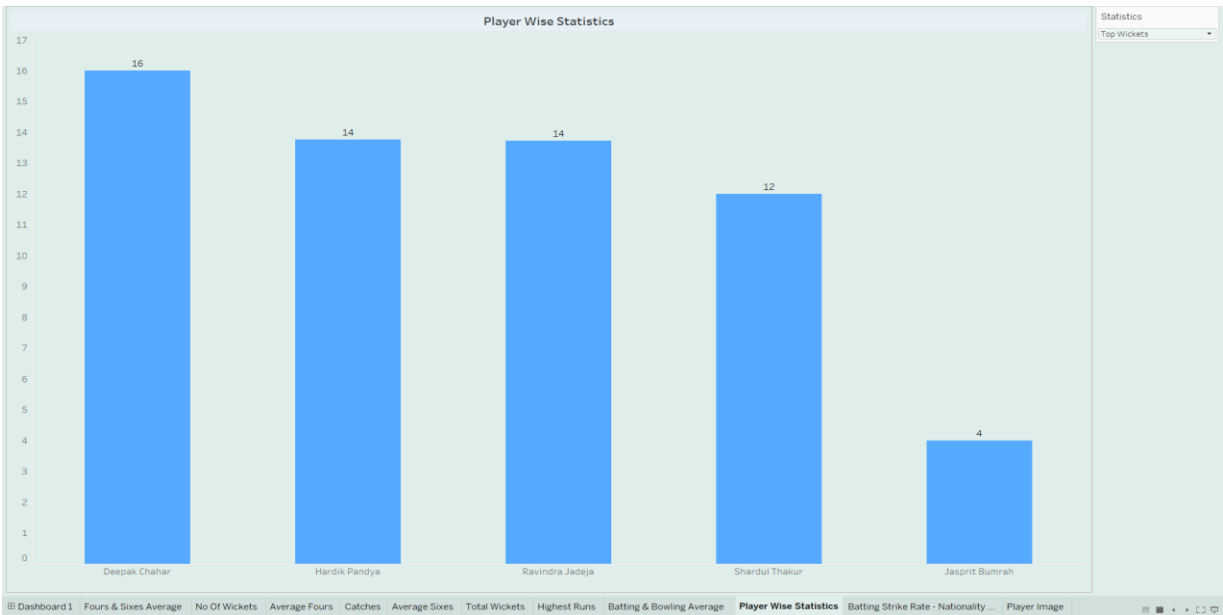
DASHBOARD



Sheet 1: Batting Strike Rate:



Sheet 2: Player Wise Statistics:



Sheet 3: Player Image:

All



Select Player Name

(Multiple values)

Player Name

 Deepak Chahar

 Hardik Pandya

 Irfan Pathan

 Jasprit Bumrah

 MS Dhoni

 Ravichandran Ashwin

 Ravindra Jadeja

 Rohit Sharma

 Shardul Thakur

 Shreyas Iyer

 Surya Kumar Yadav

 Virat Kohli

Dashboard1

Fours & Sixes Average

No Of Wickets

Average Fours

Catches

Average Sixes

Total Wickets

Highest Runs

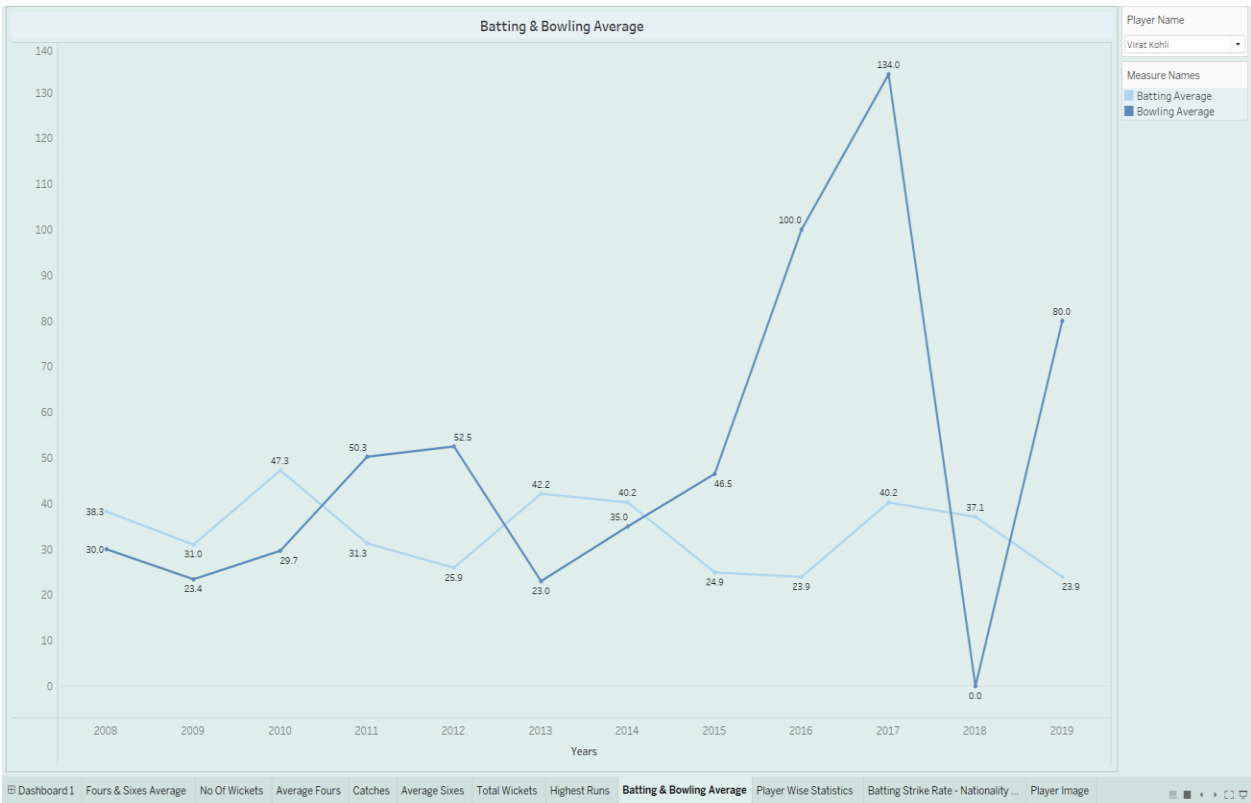
Batting & Bowling Average

Player Wise Statistics

Batting Strike Rate - Nationality ...

Player Image

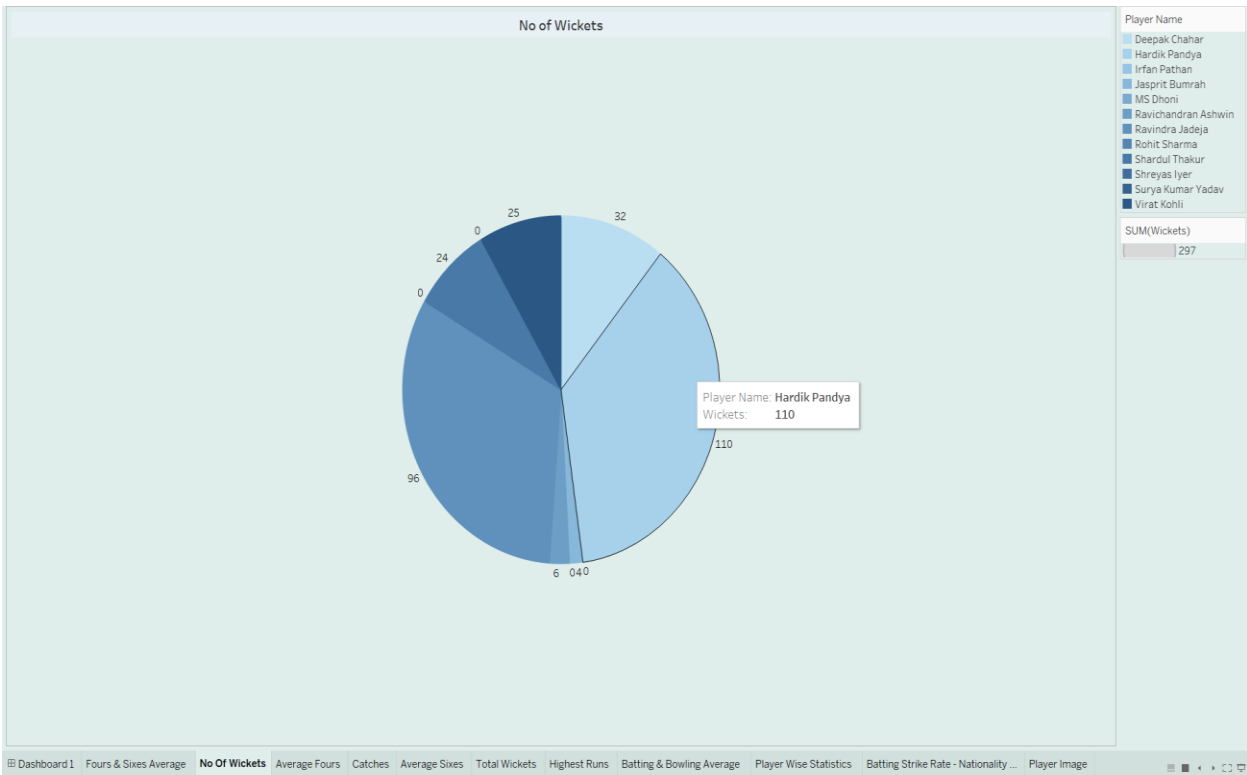
Sheet 8: Batting & Bowling Average:



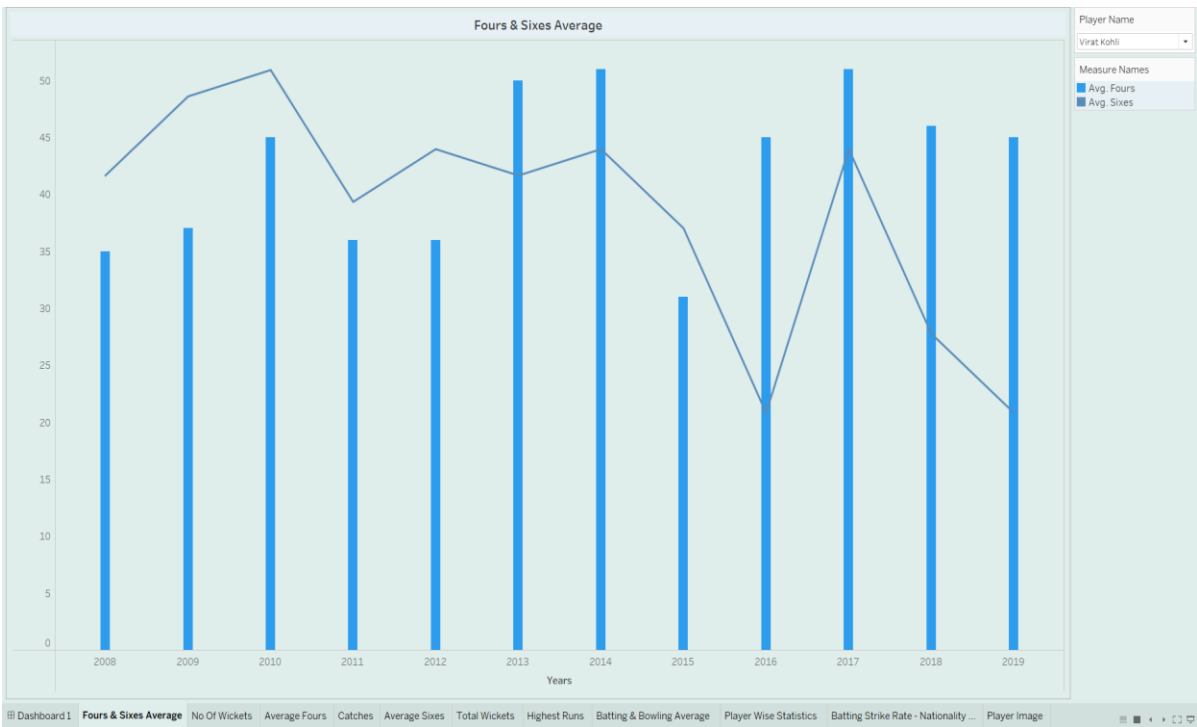
Sheet 9: Catches:

Catches	
Deepak Chahar	3
Hardik Pandya	55
Irfan Pathan	0
Jasprit Bumrah	0
MS Dhoni	99
Ravichandran Ashwin	9
Ravindra Jadeja	51
Rohit Sharma	38
Shardul Thakur	9
Shreyas Iyer	10
Surya Kumar Yadav	6
Virat Kohli	108

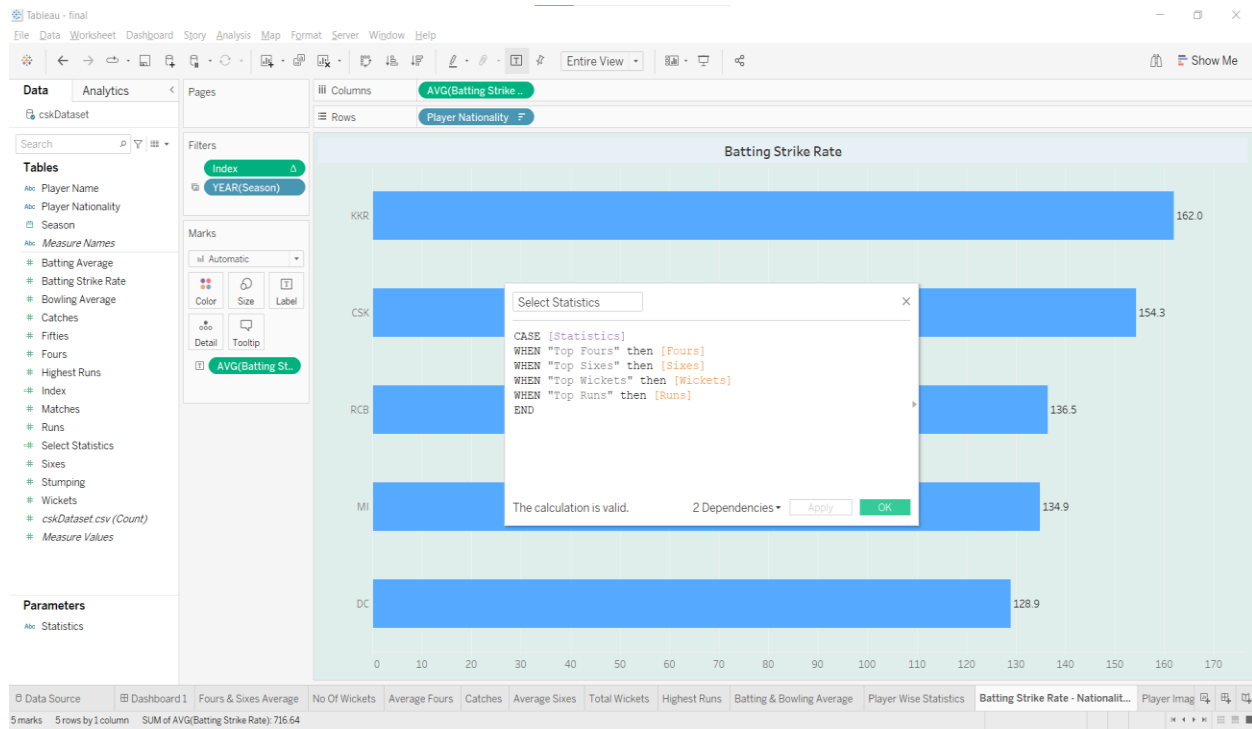
Sheet 10: No of Wickets:



Sheet 11: Fours & Sixes Average:



Filter the Calculations:





Head():

```
> head(df)
  Highest PlayerName TeamName Season BattingAverage BattingStrikeRate BowlingAverage Catches
1      100 Shreyas Iyer      MI    2018          43.00          149.75           0.0         2
2       57 Shreyas Iyer      MI    2019          23.50           93.06           0.0         6
3      100 Shreyas Iyer      MI    2018          43.00          149.75           0.0         2
4        7 Deepak Chahar    RPS    2019           3.50           77.77          21.9         2
5       39 Deepak Chahar    RPS    2018          16.66          172.41          27.8         1
6       27 Hardik Pandya    KKR    2019          16.00          121.21          30.0         5

  Fifties Fours Matches Runs Sixes Stumping Wickets
1        3     53      16  250     34         0         0
2        1     20      17  282      7         0         0
3        3     53      16  602     34         0         0
4        0      1      17    7      0         0        22
5        0      1      12   50      4         0        10
6        0      6      12   80      3         0        11
```

Queries on Dataset:

Q1: Summary of dataset:

```
> summary(df)
  Highest      PlayerName      TeamName      Season      BattingAverage
Min.   : 0.00   Length:59      Length:59      Min.   :2008   Min.   : 0.00
1st Qu.:29.00   Class :character  Class :character  1st Qu.:2012   1st Qu.:16.33
Median :57.00   Mode  :character  Mode  :character  Median :2015   Median :29.23
Mean   :53.08                                     Mean   :2015   Mean   :28.98
3rd Qu.:73.00                                     3rd Qu.:2018   3rd Qu.:39.94
Max.   :117.00                                    Max.   :2019   Max.   :83.20

BattingStrikeRate BowlingAverage      Catches      Fifties      Fours      Matches
Min.   : 0.0   Min.   : 0.00   Min.   : 0.000   Min.   :0.000   Min.   : 0.00   Min.   : 0.00
1st Qu.:122.0   1st Qu.: 0.00   1st Qu.: 3.000   1st Qu.:0.000   1st Qu.: 7.00   1st Qu.:13.00
Median :133.5   Median : 16.38   Median : 6.000   Median :1.000   Median :22.00   Median :16.00
Mean   :130.4   Mean   : 19.64   Mean   : 6.576   Mean   :1.525   Mean   :22.61   Mean   :13.83
3rd Qu.:148.9   3rd Qu.: 29.83   3rd Qu.:10.500   3rd Qu.:3.000   3rd Qu.:36.00   3rd Qu.:17.00
Max.   :300.0   Max.   :134.00   Max.   :15.000   Max.   :5.000   Max.   :53.00   Max.   :19.00

Runs      Sixes      Stumping      Wickets
Min.   : 0.0   Min.   : 0.00   Min.   :0.000   Min.   : 0.000
1st Qu.:107.0   1st Qu.: 4.00   1st Qu.:0.000   1st Qu.: 0.000
Median :332.0   Median : 9.00   Median :0.000   Median : 0.000
Mean   :271.2   Mean   :11.68   Mean   :0.661   Mean   : 5.034
3rd Qu.:415.0   3rd Qu.:19.00   3rd Qu.:0.000   3rd Qu.: 9.000
Max.   :602.0   Max.   :35.00   Max.   :6.000   Max.   :32.000
```

Q2: Seasons played by Players and Teams:

```
> ans<-factor(df$PlayerName)
> levels(ans)
[1] "Deepak Chahar"      "Hardik Pandya"      "Irfan Pathan"      "Jasprit Bumrah"
[5] "MS Dhoni"           "Ravichandran Ashwin" "Ravindra Jadeja"    "Rohit Sharma"
[9] "Shardul Thakur"     "Shreyas Iyer"       "Surya Kumar Yadav"  "Virat Kohli"
> table(ans)
ans
      Deepak Chahar      Hardik Pandya      Irfan Pathan      Jasprit Bumrah      MS Dhoni
              2              8              1              1              12
Ravichandran Ashwin  Ravindra Jadeja  Rohit Sharma  Shardul Thakur  Shreyas Iyer
              2              7              7              2              3
      Surya Kumar Yadav      Virat Kohli
              2              12
> |
```

```
> ans<-factor(df$TeamName)
> levels(ans)
[1] "CSK" "DC" "KKR" "KXIP" "MI" "RCB" "RPS" "RR" "SRH"
> table(ans)
ans
CSK  DC  KKR KXIP  MI  RCB  RPS  RR  SRH
 14   7   8   7   5  12   2   3   1
> |
```


Q3 No. of Runs on the basis of Players:

```
> df %>% group_by(PlayerName) %>% summarise(sum(Runs))
# A tibble: 12 × 2
  PlayerName      sum(Runs)
  <chr>          <int>
1 Deepak Chahar      57
2 Hardik Pandya     937
3 Irfan Pathan        0
4 Jasprit Bumrah      0
5 MS Dhoni         4646
6 Ravichandran Ashwin  953
7 Ravindra Jadeja    1066
8 Rohit Sharma       1639
9 Shardul Thakur      29
10 Shreyas Iyer       1134
11 Surya Kumar Yadav   108
12 Virat Kohli       5433
>
```

Q4 No. of Fours and Sixes on the basis of Teams:

```
> df %>% group_by(TeamName) %>% summarise(sum(Fours))
# A tibble: 9 × 2
  TeamName `sum(Fours)`
  <chr>      <int>
1 CSK        317
2 DC          73
3 KKR         69
4 KXIP       145
5 MI        212
6 RCB        508
7 RPS         2
8 RR          8
9 SRH         0

> df %>% group_by(TeamName) %>% summarise(sum(Sixes))
# A tibble: 9 × 2
  TeamName `sum(Sixes)`
  <chr>      <int>
1 CSK       220
2 DC         37
3 KKR        44
4 KXIP        50
5 MI       130
6 RCB       199
7 RPS         4
8 RR         5
9 SRH         0
>
```

Q5 No. of catches and stumpings on the basis of Seasons:

```
> df %>% group_by(Season) %>% summarise(sum(Catches)+sum(Stumping))
# A tibble: 12 × 2
  Season `sum(Catches) + sum(Stumping)`
  <int>          <int>
1  2008             16
2  2009             15
3  2010             21
4  2011             19
5  2012             40
6  2013             53
7  2014             32
8  2015             56
9  2016             23
10 2017             22
11 2018             61
12 2019             69
> |
```

Q6: Name the player, of which team in which season whose average of Batting, Bowling, Fours and Sixes is Zero:

```
> ans<-subset(df,BattingAverage==0 & BowlingAverage==0 & Fours==0 & Sixes==0,select=c(PlayerName,TeamName,Season))
> ans
```

	PlayerName	TeamName	Season
16	Rohit Sharma	KXIP	2011
20	Rohit Sharma	KXIP	2013
21	Irfan Pathan	SRH	2015
41	Surya Kumar Yadav	RR	2019

```
>
```