

MVLU COLLEGE R PRACTICAL 3

Aim: Exploring data: View() or print() (R).

The screenshot shows the RStudio interface with the following code in the script editor:

```
1 install.packages(c("readr", "psych"))
2
3 library(readr) # For efficient data reading
4 library(psych) # For descriptive statistics
5
6 my_data <- read.csv("Fastest.Fifties.csv")
7
8 # View the first few rows
9 head(Fastest.Fifties)
10
11 # View the last few rows
12 tail(Fastest.Fifties)
13
14 # Get the dimensions (rows and columns)
15 dim(Fastest.Fifties)
16
17 # Get the dimensions (rows and columns)
18 cat("Dimensions (Rows, Columns): ", dim(Fastest.Fifties), "\n")
19
20 # Get the structure (variable types and number of observations)
21 str(Fastest.Fifties)
22
23 # See a summary of the dataset
24 summary(Fastest.Fifties)
25
26 # Get the column names
27 names(Fastest.Fifties)
28 cat("Column Names: ", names(Fastest.Fifties), "\n")
29
30 # Use the 'psych' package for more detailed descriptive statistics
31 # 'describe()' provides: n, mean, sd, median, trimmed mean, mad, min, max, range, skew, kurtosis
32 describe(Fastest.Fifties)
33
34
35
36
```

The Environment pane on the right shows the loaded datasets:

Object	Size	Modified
Fastest.Fifties	106 obs. of 9 variables	
Fastest.Fiftie...	107 obs. of 9 variables	

The screenshot shows the RStudio console with the following output:

```
> Fastest.Fifties
      POS   Player Runs BF  X4s  X6s Against
1      1    KL Rahul  51 14   6   4      DC
2      2  Ishan Kishan  62 17   5   6      KKR
3      3  Sunil Narine  50 17   4   5      RCB
4      4   Jos Buttler  67 18   4   7      DC
5      5   Sam Billings  56 21   2   5      KKR
6      6    KL Rahul  66 22   2   7      KKR
7      7  Chris Gayle  63 22   7   4      CSK
8      8   MS Dhoni  51 22   2   5      DC
9      9  Kieron Pollard  50 22   5   3      PBKS
10     10  Dinesh Karthik  50 22   5   3      PBKS
11     11  Rishabh Pant  69 23   7   5      RR
12     12  AB de Villiers  68 23   2   8      CSK
13     13  AB de Villiers  90 24  10   5      DC
14     14    KL Rahul  60 24   9   2      KKR
15     15  Shane Watson  78 25   4   7      DC
16     16  Dwayne Bravo  68 25   3   7      MI
17     17  Moeen Ali  65 25   2   6      SRH
18     18  Prithvi Shaw  65 25   6   3      SRH
19     19  Karun Nair  54 25   3   5      CSK
20     20   Jos Buttler  95 26  11   2      CSK
21     21  Andre Russell  88 26   1  11      CSK
22     22  Sunil Narine  75 26   9   4      PBKS
23     23  Virat Kohli  70 26   7   3      DC
24     24  Virat Kohli  57 26   7   2      RR
25     25  Jason Roy  91 27   6   6      MI
26     26   Jos Buttler  82 27   9   1      PBKS
27     27  Ambati Rayudu  79 27   9   4      SRH
28     28  Rishabh Pant  61 27   5   4      RCB
29     29  Shane Watson  106 28   9   6      RR
30     30  Kane Williamson  81 28   7   5      RCB
31     31  AB de Villiers  72 28   4   6      DC
32     32  Chris Gayle  62 28   5   6      KKR
33     33  Vijay Shankar  54 28   1   5      CSK
```

Hariprasad vishwakarma
S126
SYCS

MVLU COLLEGE R PRACTICAL 3

The screenshot shows the RStudio environment with a project named 'None'. The console displays a data frame with columns for player names, runs, balls faced, wickets, and matches. The environment pane shows the 'Data' tab with two objects: 'Fastest.Fifties' (106 obs. of 9 variables) and 'Fastest.Fiftie...' (107 obs. of 9 variables). The file explorer shows the 'Home' directory with various files and folders.

Player	Runs	Balls Faced	Wickets	Matches
Hardik Pandya	50	41	5	1
Ajinkya Rahane	65	42	5	1
Chris Lynn	62	42	7	1
Kane Williamson	50	43	4	1
KL Rahul	84	44	7	3
Suresh Raina	61	45	4	2
KL Rahul	95	48	11	2
Manish Pandey	54	48	3	1

The screenshot shows the RStudio environment with the same project. The console displays the output of the following R commands:

```
> describe(Fastest.Fifties)
> names(Fastest.Fifties)
> summary(Fastest.Fifties)
```

The output shows the structure of the data frame, including the number of observations (106) and the number of variables (9). The summary statistics for each variable are displayed.

Variable	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
POS	1.00	27.25	53.50	53.50	79.75	106.00
Player*	1.00	27.25	53.50	53.50	79.75	106.00
Runs	1.00	27.25	53.50	53.50	79.75	106.00
BF	1.00	27.25	53.50	53.50	79.75	106.00
X4s	1.00	27.25	53.50	53.50	79.75	106.00
X6s	1.00	27.25	53.50	53.50	79.75	106.00
Against*	1.00	27.25	53.50	53.50	79.75	106.00
Venue*	1.00	27.25	53.50	53.50	79.75	106.00
Match.Date*	1.00	27.25	53.50	53.50	79.75	106.00

Hariprasad vishwakarma
S126
SYCS

MVLU COLLEGE

R RACTICAL 3

RStudio interface showing the console output for the 'Fastest.Fifties' dataset. The console displays summary statistics and a tail of the dataset. The environment pane shows the loaded data frames.

```
Min. : 0.000 Length:106 Length:106 Length:106
1st Qu.: 2.000 Class :character Class :character Class :character
Median : 3.000 Mode :character Mode :character Mode :character
Mean : 3.755
3rd Qu.: 5.000
Max. : 11.000

> str(Fastest.Fifties)
'data.frame': 106 obs. of 9 variables:
 $ POS : int 1 2 3 4 5 6 7 8 9 10 ...
 $ Player : chr "KL Rahul" "Ishan Kishan" "Sunil Narine" "Jos Buttler" ...
 $ Runs : int 51 62 50 67 56 66 63 51 50 50 ...
 $ BF : int 14 17 17 18 21 22 22 22 22 22 ...
 $ X4s : int 6 5 4 4 2 2 2 2 5 5 ...
 $ X6s : int 4 6 5 7 5 7 4 5 3 3 ...
 $ Against : chr "dc" "KKR" "RCB" "dc" ...
 $ Venue : chr "IS Bindra stadium" "Eden Gardens" "Eden Gardens" "Arun Jaitley Stadium" ...
 $ Match.Date: chr "08 April 2018" "09 May 2018" "08 April 2018" "02 May 2018" ...

> cat('Dimensions (rows, columns): ', dim(Fastest.Fifties), "\n")
Dimensions (rows, columns): 106 9

> dim(Fastest.Fifties)
[1] 106 9

> tail(Fastest.Fifties)
      POS Player Runs BF X4s X6s Against
101 101 Chris Lynn 62 42 7 1 RCB
102 102 Kane Williamson 50 43 4 1 KKR
103 103 KL Rahul 84 44 7 3 RR
104 104 Suresh Raina 61 45 4 2 PBKS
105 105 KL Rahul 95 48 11 2 RR
106 106 Manish Pandey 54 48 3 1 PBKS

      Venue Match.Date
101 M. Chinnaswamy Stadium 29 April 2018
102 Eden Gardens 14 April 2018
103 Holkar Cricket Stadium 06 May 2018
104 Maharashtra Cricket Association's International Stadium 20 May 2018
105 Sawai Mansingh Stadium 08 May 2018
106 Rajiv Gandhi Intl. Cricket Stadium 26 April 2018
```

RStudio interface showing the console output for the 'Fastest.Fifties' dataset. The console displays summary statistics and a head of the dataset. The environment pane shows the loaded data frames.

```
$ POS : int 1 2 3 4 5 6 7 8 9 10 ...
$ Player : chr "KL Rahul" "Ishan Kishan" "Sunil Narine" "Jos Buttler" ...
$ Runs : int 51 62 50 67 56 66 63 51 50 50 ...
$ BF : int 14 17 17 18 21 22 22 22 22 22 ...
$ X4s : int 6 5 4 4 2 2 2 2 5 5 ...
$ X6s : int 4 6 5 7 5 7 4 5 3 3 ...
$ Against : chr "dc" "KKR" "RCB" "dc" ...
$ Venue : chr "IS Bindra stadium" "Eden Gardens" "Eden Gardens" "Arun Jaitley Stadium" ...
$ Match.Date: chr "08 April 2018" "09 May 2018" "08 April 2018" "02 May 2018" ...

> cat('Dimensions (rows, columns): ', dim(Fastest.Fifties), "\n")
Dimensions (rows, columns): 106 9

> dim(Fastest.Fifties)
[1] 106 9

> tail(Fastest.Fifties)
      POS Player Runs BF X4s X6s Against
101 101 Chris Lynn 62 42 7 1 RCB
102 102 Kane Williamson 50 43 4 1 KKR
103 103 KL Rahul 84 44 7 3 RR
104 104 Suresh Raina 61 45 4 2 PBKS
105 105 KL Rahul 95 48 11 2 RR
106 106 Manish Pandey 54 48 3 1 PBKS

      Venue Match.Date
101 M. Chinnaswamy Stadium 29 April 2018
102 Eden Gardens 14 April 2018
103 Holkar Cricket Stadium 06 May 2018
104 Maharashtra Cricket Association's International Stadium 20 May 2018
105 Sawai Mansingh Stadium 08 May 2018
106 Rajiv Gandhi Intl. Cricket Stadium 26 April 2018

> head(Fastest.Fifties)
      POS Player Runs BF X4s X6s Against Venue Match.Date
1 1 KL Rahul 51 14 6 4 DC IS Bindra Stadium 08 April 2018
2 2 Ishan Kishan 62 17 5 6 KKR Eden gardens 09 May 2018
3 3 Sunil Narine 50 17 4 5 RCB Eden gardens 08 April 2018
4 4 Jos Buttler 67 18 4 7 DC Arun Jaitley Stadium 02 May 2018
5 5 Sam Billings 56 21 2 5 KKR chidambaram 10 April 2018
```

Hariprasad vishwakarma

S126

SYCS