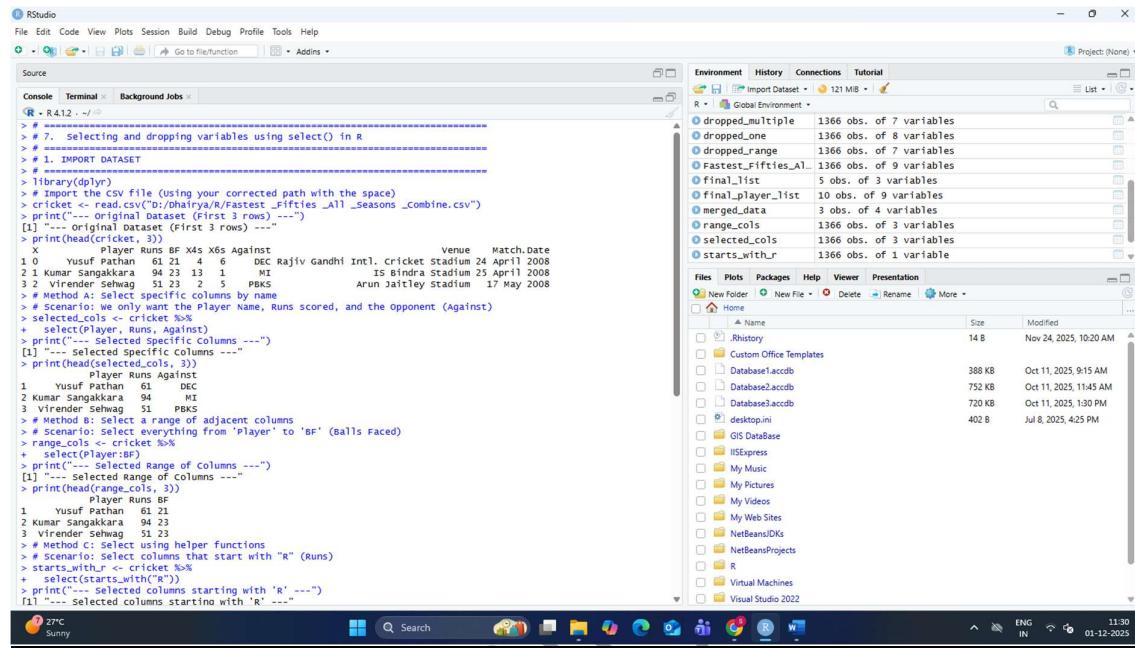


SHETH L.U.J AND SIR M.V. COLLEGE
SUBJECT NAME: Data Analysis with SAS / SPSS /R
Practical No. 7

Aim- Selecting and dropping variables using select () in R. import dataset.

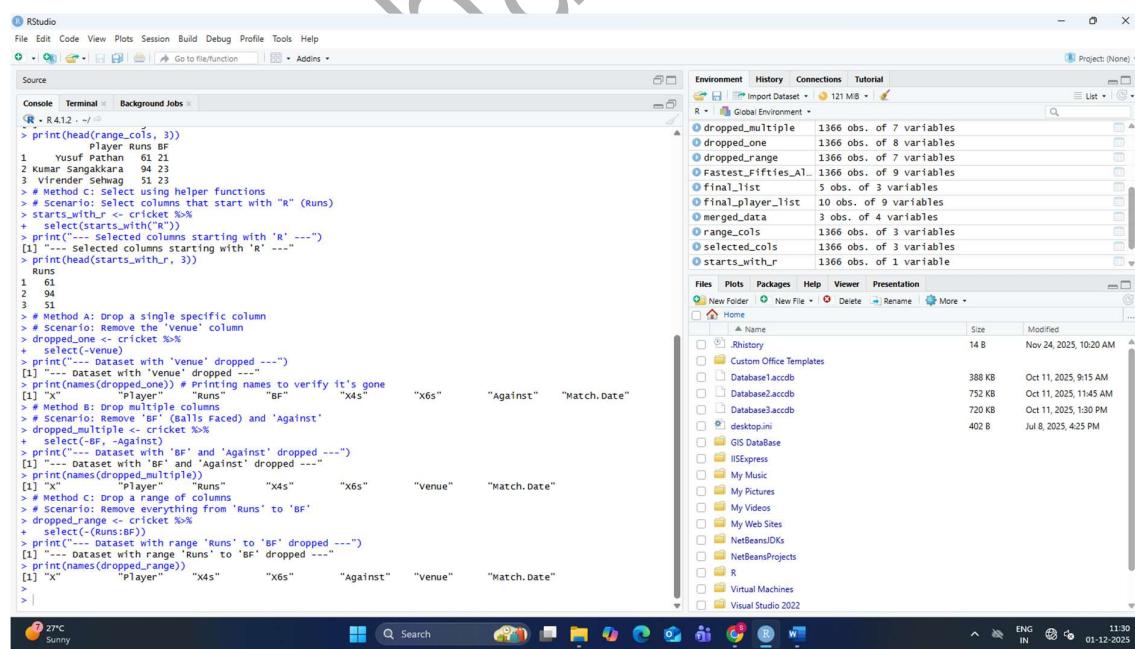
Output-



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
File Project Global Environment
Source
Console Terminal Background Jobs
R - R4.12 - -/-
> # Selecting and dropping variables using select() in R
> # Scenario: Remove the 'venue' column
> # Method A: Select specific columns by name
> # Scenario: we only want the Player Name, Runs scored, and the opponent (Against)
> selected_cols <- cricket %>
> print("---- Selected Specific columns ----")
[1] "---- Selected Specific columns ----"
> print(head(selected_cols, 3))
      Player Runs Against   Venue Match.Date
1  Yusuf Pathan 61 21    DRC
2 Kumar Sangakkara 94 23  MS Dhoni 24 April 2008
3  Virender Sehwag 51 23  5 PBKS
> # Method B: Select a range of adjacent columns
> # Scenario: Select everything from 'player' to 'BF' (Balls Faced)
> range_cols <- cricket %>
> print("---- Selected Range of columns ----")
[1] "---- Selected Range of columns ----"
> print(head(range_cols, 3))
      Player Runs
1  Yusuf Pathan 61 21
2 Kumar Sangakkara 94 23
3  Virender Sehwag 51 23
> # Method C: Select using helper functions
> # Scenario: Select columns that start with "R" (Runs)
> starts_with_r <- cricket %>
+ select(starts_with("R"))
> print("---- Selected columns starting with 'R' ----")
[1] "---- Selected columns starting with 'R' ----"

```



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
File Project Global Environment
Source
Console Terminal Background Jobs
R - R4.12 - -/-
> print(head(range_cols, 3))
      Player Runs BF
1  Yusuf Pathan 61 21
2 Kumar Sangakkara 94 23
3  Virender Sehwag 51 23
> # Method C: Select using helper functions
> # Scenario: Select columns that start with "R" (Runs)
> starts_with_r <- cricket %>
+ select(starts_with("R"))
> print("---- Selected columns starting with 'R' ----")
[1] "---- Selected columns starting with 'R' ----"
> print(head(starts_with_r, 3))
      Runs
1 61
2 94
3 51
> # Method A: Drop a single specific column
> # Scenario: Remove the 'venue' column
> dropped_ones <- cricket %>
+ select(-"venue")
> print("---- Dataset with 'venue' dropped ----")
[1] "---- Dataset with 'venue' dropped ----"
> print(names(dropped_ones)) # Printing names to verify it's gone
[1] "Player"      "Runs"       "BF"        "x4s"       "x6s"       "Against"    "Match.Date"
> # Method B: Drop multiple columns
> # Scenario: Remove 'BF' (Balls Faced) and 'Against'
> dropped_multiple <- cricket %>
+ select(-"BF", -"Against")
> print("---- Dataset with 'BF' and 'Against' dropped ----")
[1] "---- Dataset with 'BF' and 'Against' dropped ----"
> print(names(dropped_multiple))
[1] "Player"      "Runs"       "x4s"      "x6s"      "venue"      "Match.Date"
> # Method C: Drop a range of columns
> # Scenario: Remove everything from 'Runs' to 'BF'
> dropped_range <- cricket %>
+ select(-(Runs:BF))
> print("---- Dataset with range 'Runs' to 'BF' dropped ----")
[1] "---- Dataset with range 'Runs' to 'BF' dropped ----"
> print(names(dropped_range))
[1] "Player"      "x4s"      "x6s"      "Against"    "venue"      "Match.Date"
> |

```