

SHETH L.U.J. AND SIR M.V. COLLEGE OF ARTS SCIENCE AND COMMERCE

SUBJECT : R Programming

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

Go to file/function Addins

Source

Console Terminal Background Jobs

```
[R] > R4.5.1: ~/R
[1] "R4.5.1: ~/R"
> library(dplyr)
>
> # Import dataset
> marks <- read.csv("data_science_student_marks.csv")
>
> print(" original dataset (First 3 rows) ")
[1] " original Dataset (First 3 rows) "
> print(head(marks, 3))
student_id location age sql_marks excel_marks python_marks power_bi_marks english_marks
1          4 Sydney 24      95       99       87       82        75
2          5 Tokyo   24      99       95       89       86        82
3          6 Berlin  22      72       70       99       79        77
>
> # SELECTING VARIABLES
>
> # select specific columns
> selected_cols <- marks %>%
+   select(student_id, age, python_marks)
>
> print("--- Selected Specific Columns ---")
[1] "--- Selected Specific Columns ---"
> print(head(selected_cols, 3))
student_id age python_marks
1          4    24        87
2          5    24        89
3          6    22        99
>
> # select a range of adjacent columns
> range_cols <- marks %>%
+   select(student_id:python_marks)
>
> print(" Selected Range of Columns ")
[1] " Selected Range of Columns "
> print(head(range_cols, 3))
student_id location age sql_marks excel_marks python_marks
1          4 Sydney 24      95       99       87
2          5 Tokyo   24      99       95       89
3          6 Berlin  22      72       70       99
>
> # select columns starting with 'p'
> starts_with_p <- marks %>%
+   select(starts_with("p"))
```

NAME : SHUBHAM SANJAY KARAPE
ROLL NO : S085

SHETH L.U.J. AND SIR M.V. COLLEGE OF ARTS SCIENCE AND COMMERCE
SUBJECT : R Programming

```
R Studio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins

Source

Console Terminal Background Jobs
R > R4.5.1 : ~
> print(" Selected Range of Columns ")
[1] " Selected Range of Columns "
> print(head(range_cols, 3))
#> #> student_id location age sql_marks excel_marks python_marks
#> #> 1       4   Sydney  24      95        99       87
#> #> 2       5    Tokyo   24      99        95       89
#> #> 3       6   Berlin  22      72        70       99
#> #>
#> #> # select columns starting with 'p'
#> #> starts_with_p <- marks %>%
#> #>   select(starts_with("p"))
#> #>
#> print(" --- columns starting with 'p' ---")
[1] " --- columns starting with 'p' ---"
> print(head(starts_with_p, 3))
#> #> python_marks power_bi_marks
#> #> 1       87        82
#> #> 2       89        86
#> #> 3       99        79
#> #>
#> # DROPPING VARIABLES
#> #>
#> # drop single column
#> drop_one <- marks %>%
#>   select(-location)
#> #
#> print(" After Dropping 'location'")
[1] " After Dropping 'location'"
> print(names(drop_one))
[1] "student_id"      "age"           "sql_marks"      "excel_marks"    "python_marks"
[6] "power.bi_marks"  "english_marks"
#> #
#> # Drop multiple columns
#> drop_multiple <- marks %>%
#>   select(-c(location, english_marks))
#> #
#> print(" After Dropping Multiple Columns ")
[1] " After Dropping Multiple Columns "
> print(names(drop_multiple))
[1] "student_id"      "age"           "sql_marks"      "excel_marks"    "python_marks"
[6] "power.bi_marks"
#>
```

NAME : SHUBHAM SANJAY KARAPE
ROLL NO : S085