

SHETH L.U.J. AND SIR M.V. COLLEGE

PRACTICAL NO:3

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

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
> # View the first few rows
> head(my_data)
# A tibble: 6 x 7
  age sex   bmi children smoker region  charges
  <dbl> <chr> <dbl>    <dbl> <chr> <chr>    <dbl>
1 19  female 27.9      0 yes   southwest 16885.
2 18  male   33.8      1 no    southeast 1726.
3 28  male   33        3 no    southeast 4449.
4 33  male   22.7      0 no    northwest 21984.
5 32  male   28.9      0 no    northwest 3867.
6 31  female 25.7      0 no    southeast 2757.
```

> # View the last few rows
> tail(my_data)
A tibble: 6 x 7
 age sex bmi children smoker region charges
 <dbl> <chr> <dbl> <dbl> <chr> <chr> <dbl>
1 52 female 44.7 3 no southwest 11412.
2 50 male 31.0 3 no northwest 10601.
3 18 female 31.9 0 no northeast 2206.
4 18 female 36.8 0 no southeast 1630.
5 21 male 25.8 0 no southwest 2008.
6 61 female 29.1 0 yes northwest 29141.

> # Get dataset dimensions
> dim(my_data)
[1] 1338 7

> cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")
Dimensions (Rows, Columns): 1338 7

> # Get structure
> str(my_data)

```
spc_tbl_ [1,338 x 7] (S3: spc_tbl_df/tbl_df/tbl/data.frame)
$ age   : num [1:1338] 19 18 28 33 32 31 46 37 37 60 ...
$ sex   : chr [1:1338] "female" "male" "male" "male" ...
$ bmi   : num [1:1338] 27.9 33.8 33 22.7 28.9 ...
$ children: num [1:1338] 0 1 3 0 0 0 1 3 2 0 ...
$ smoker : chr [1:1338] "yes" "no" "no" "no" ...
```

12:43 24-11-2025

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

```
> # Summary statistics
> summary(my_data)
  age   sex   bmi   children   smoker   region   charges
Min. :18.00  female:1338   Min. :15.96   Min. :0.000
1st Qu.:27.00  Class :character 1st Qu.:26.30   1st Qu.:0.000
Median :39.00  Mode  :character Median :30.40   Median :1.000
Mean   :39.21  Mean   :30.66   Mean   :1.095
3rd Qu.:51.00  3rd Qu.:34.69   3rd Qu.:12.000
Max.   :64.00  Max.   :53.13   Max.   :5.000

  smoker   region   charges
Length:1338  Length:1338  Min. :1122
Class :character Class :character  1st Qu.:4740
Mode  :character Mode  :character  Median :9382
                           Mean   :13270
                           3rd Qu.:16640
                           Max.   :63770

> # Column names
> names(my_data)
[1] "age"     "sex"     "bmi"     "children" "smoker"  "region"  "charges"
> cat("Column Names: ", names(my_data), "\n")
Column Names: age sex bmi children smoker region charges
> # Detailed descriptive statistics using psych
> describe(my_data)
   vars   n   mean     sd median trimmed   mad   min   max
age    1 1338 39.21 14.05 39.00 39.01 17.79 18.00 64.00
sex*   2 1338  1.51  0.50  2.00  1.51  0.00  1.00  2.00
bmi    3 1338 30.66 6.10 30.40 30.50 6.20 15.96 53.13
children 4 1338  1.09  1.21  1.00  0.94  1.48  0.00  5.00
smoker* 5 1338  1.20  0.40  1.00  1.13  0.00  1.00  2.00
region* 6 1338  2.52  1.10  3.00  2.52  1.48  1.00  4.00
charges 7 1338 13270.42 12110.01 9382.03 11076.02 7440.81 1121.87 63770.43
   range   skew kurtosis   se
age   46.00  0.06 -1.25  0.38
sex*  1.00 -0.02 -2.00  0.01
```

12:43 24-11-2025