

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

## Aim: Generating basic summaries using str() or summary() (R).

The image displays two screenshots of the RStudio interface, demonstrating the process of loading a dataset and generating summaries using R functions.

**Top Screenshot:** The console shows the following code and output:

```
> wh_df <- read.csv("C:/Users/mvuc/Downloads/weight-height.csv")
> print("--- Data Loaded from CSV ---")
> print(head(wh_df))
  Gender Height Weight
1 Male  73.84702 241.8936
2 Male  68.78190 162.3105
3 Male  74.11011 212.7409
4 Male  71.73098 220.0425
5 Male  69.88180 206.3498
6 Male  67.25302 152.2122
>
> print("--- OUTPUT OF str() ---")
> str(wh_df)
'data.frame':   10000 obs. of  3 variables:
 $ Gender: chr  "Male" "Male" "Male" "Male" ...
 $ Height: num  73.8 68.8 74.1 71.7 69.9 ...
 $ Weight: num  242 162 213 220 206 ...
>
> print("--- OUTPUT OF summary() [Before Factor Conversion] ---")
> summary(wh_df)
  Gender      Height      Weight
Length:10000   Min.   :54.26   Min.   : 64.7
Class :character 1st Qu.:63.51   1st Qu.:135.8
Mode :character  Median :66.32   Median :161.2
                Mean  :66.37   Mean  :161.4
                3rd Qu.:69.17   3rd Qu.:187.2
                Max.   :79.00   Max.   :270.0
>
> # Convert Gender column to Factor (if it exists)
> wh_df$Gender <- as.factor(wh_df$Gender)
>
> print("--- OUTPUT OF summary() [After Factor Conversion] ---")
> summary(wh_df)
  Gender      Height      Weight
Length:10000   Min.   :54.26   Min.   : 64.7
Class :factor   1st Qu.:63.51   1st Qu.:135.8
Mode :factor    Median :66.32   Median :161.2
                Mean  :66.37   Mean  :161.4
                3rd Qu.:69.17   3rd Qu.:187.2
                Max.   :79.00   Max.   :270.0
>
> # Accessing only selected statistics
> avg_height <- mean(wh_df$Height, na.rm = TRUE)
> max_weight <- max(wh_df$Weight, na.rm = TRUE)
>
> print(paste("Average Height:", avg_height))
[1] "Average Height: 66.3675597548212"
> print(paste("Highest Weight:", max_weight))
[1] "Highest Weight: 269.989698505106"
>
> print("Yukta Sonawane S120")
[1] "Yukta Sonawane S120"
```

**Bottom Screenshot:** The console shows the same code as the top screenshot, but with additional code to access specific statistics:

```
> print(paste("Average Height:", avg_height))
[1] "Average Height: 66.3675597548212"
> print(paste("Highest Weight:", max_weight))
[1] "Highest Weight: 269.989698505106"
>
> print("Yukta Sonawane S120")
[1] "Yukta Sonawane S120"
```