

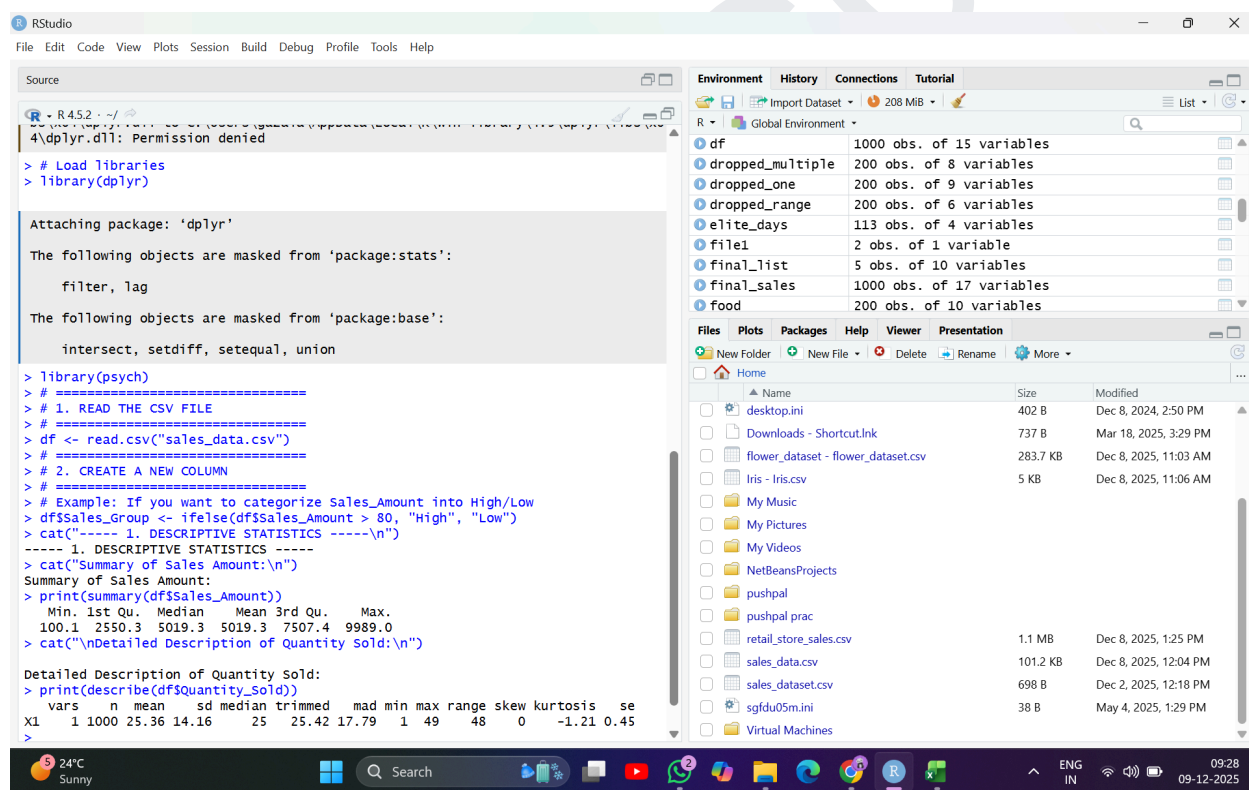
SHETH L.U.J AND SIR M.V COLLEGE

Subject: Data Analysis with SAS / SPSS /R

Practical no. 1

Aim: Generating descriptive statistics using
summary() or describe() (R)

Outputs→



The screenshot displays the RStudio interface. The Source pane on the left contains the following R code:

```
> # Load libraries
> library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':
  filter, lag

The following objects are masked from 'package:base':
  intersect, setdiff, setequal, union

> library(psych)
> # =====
> # 1. READ THE CSV FILE
> # =====
> df <- read.csv("sales_data.csv")
> # =====
> # 2. CREATE A NEW COLUMN
> # =====
> # Example: If you want to categorize Sales_Amount into High/Low
> df$Sales_Group <- ifelse(df$Sales_Amount > 80, "High", "Low")
> cat("----- 1. DESCRIPTIVE STATISTICS -----")
----- 1. DESCRIPTIVE STATISTICS -----
> cat("Summary of Sales Amount:\n")
Summary of Sales Amount:
> print(summary(df$Sales_Amount))
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
100.1  2550.3   5019.3  7507.4 9989.0
> cat("\nDetailed Description of Quantity Sold:\n")

Detailed Description of Quantity Sold:
> print(describe(df$Quantity_Sold))
  vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
X1     1 1000 25.36 14.16    25   25.42 17.79     1  49   48     0   -1.21 0.45
```

The Environment pane on the right shows the following objects:

Object	Obs.	Vars.
df	1000	15
dropped_multiple	200	8
dropped_one	200	9
dropped_range	200	6
elite_days	113	4
file1	2	1
final_list	5	10
final_sales	1000	17
food	200	10

The Files pane at the bottom shows a list of files in the Home directory, including desktop.ini, Downloads - Shortcut.lnk, flower_dataset - flower_dataset.csv, Iris - Iris.csv, My Music, My Pictures, My Videos, NetBeansProjects, pushpal, pushpal prac, retail_store_sales.csv, sales_data.csv, sales_dataset.csv, sgfd05m.ini, and Virtual Machines.