

MVLU COLLEGE.

PRACTICAL NO :- 03

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

CODE :-

```
install.packages(c("readr", "psych"))
```

```
library(readr) # For efficient data reading
```

```
library(psych) # For descriptive statistics
```

```
my_data <- read.csv("C:/Users/itlab/Downloads/S100/sales_data.csv")
```

```
# View the first few rows
```

```
head(my_data)
```

```
tail(my_data)
```

```
dim(my_data)
```

```
cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")
```

```
str(my_data)
```

```
summary(my_data)
```

```
names(my_data)
```

```
cat("Column Names: ", names(my_data), "\n")
```

```
library(psych)
```

```
describe(my_data)
```

MVLU COLLEGE.

The image displays two screenshots of the RStudio interface, showing the process of loading and analyzing data.

Top Screenshot: The R script editor shows the following code:

```
1 install.packages(c("readr", "psych"))
2
3 library(readr) # For efficient data reading
4 library(psych) # For descriptive statistics
5
6 my_data <- read_csv("C:/Users/it1ab/Downloads/S100/sales_data.csv")
7
8 # view the first few rows
9 head(my_data)
10
11 tail(my_data)
12
13 dim(my_data)
14
15 cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")
16
17 str(my_data)
18
19 summary(my_data)
20
21 names(my_data)
22 cat("Column Names: ", names(my_data), "\n")
23
24 library(psych)
25 describe(my_data)
26
```

The Environment pane on the right shows the loaded data:

| Object | Class | Size |
|-----------------------|------------|---------------------------|
| my_data | data.frame | 1000 obs. of 14 variables |
| sales_data | data.frame | 1000 obs. of 14 variables |
| Student.Mental.health | data.frame | 101 obs. of 11 variables |

Bottom Screenshot: The R script editor shows the following code:

```
> view(my_data)
> view(my_data)
> # View the first few rows
> head(my_data)
  Product_ID Sale_Date Sales_Rep Region Sales_Amount Quantity_Sold Product_Category Unit_Cost
1    1052 2023-02-03    Bob North    5033.97         18      Furniture    152.75
2    1093 2023-04-21    Bob West    4384.02         17      Furniture    3816.39
3    1015 2023-09-21   David South    4631.23         30         Food    261.56
4    1072 2023-08-24    Bob South    2167.94         39      Clothing    4330.03
5    1061 2023-02-24   Charlie East    3750.20         13      Electronics    637.37
6    1021 2023-02-11   Charlie West    3761.15         32         Food    900.79

  Unit_Price Customer_Type Discount Payment_Method Sales_Channel Region_and_Sales_Rep
1    267.22      Returning    0.09      Cash      Online      North-Bob
2    4209.44      Returning    0.11      Cash      Retail      West-Bob
3    371.40      Returning    0.20 Bank Transfer      Retail      South-David
4    4467.75        New    0.02 Credit Card      Retail      South-Bob
5    692.71        New    0.08 Credit Card      Online      East-Charlie
6    1106.51        New    0.21      Cash      Online      West-Charlie

> tail(my_data)
  Product_ID Sale_Date Sales_Rep Region Sales_Amount Quantity_Sold Product_Category Unit_Cost
995    1068 2023-04-06    Alice South    9093.50          31      Clothing    3169.37
996    1010 2023-04-15   Charlie North    4733.88           4         Food    4943.03
997    1067 2023-09-07    Bob North    4716.36          37      Clothing    1754.32
998    1018 2023-04-27   David South    7629.70          17      Clothing    355.72
999    1100 2023-12-20   David West    1629.47          39      Electronics    3685.03
1000    1086 2023-08-16    Alice East    4923.93          48         Food    2632.58

  Unit_Price Customer_Type Discount Payment_Method Sales_Channel Region_and_Sales_Rep
995    3304.15      Returning    0.25 Credit Card      Retail      South-Alice
996    5442.15      Returning    0.29      Cash      Online      North-Charlie
997    1856.40        New    0.21 Bank Transfer      Retail      North-Bob
998    438.27      Returning    0.06 Bank Transfer      Online      South-David
999    3743.39        New    0.01 Bank Transfer      Online      West-David
1000    2926.68      Returning    0.14      Cash      Online      East-Alice

> dim(my_data)
[1] 1000 14
> cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")
Dimensions (Rows, Columns): 1000 14
>
> summary(my_data)
  Product_ID Sale_Date Sales_Rep Region Sales_Amount
```

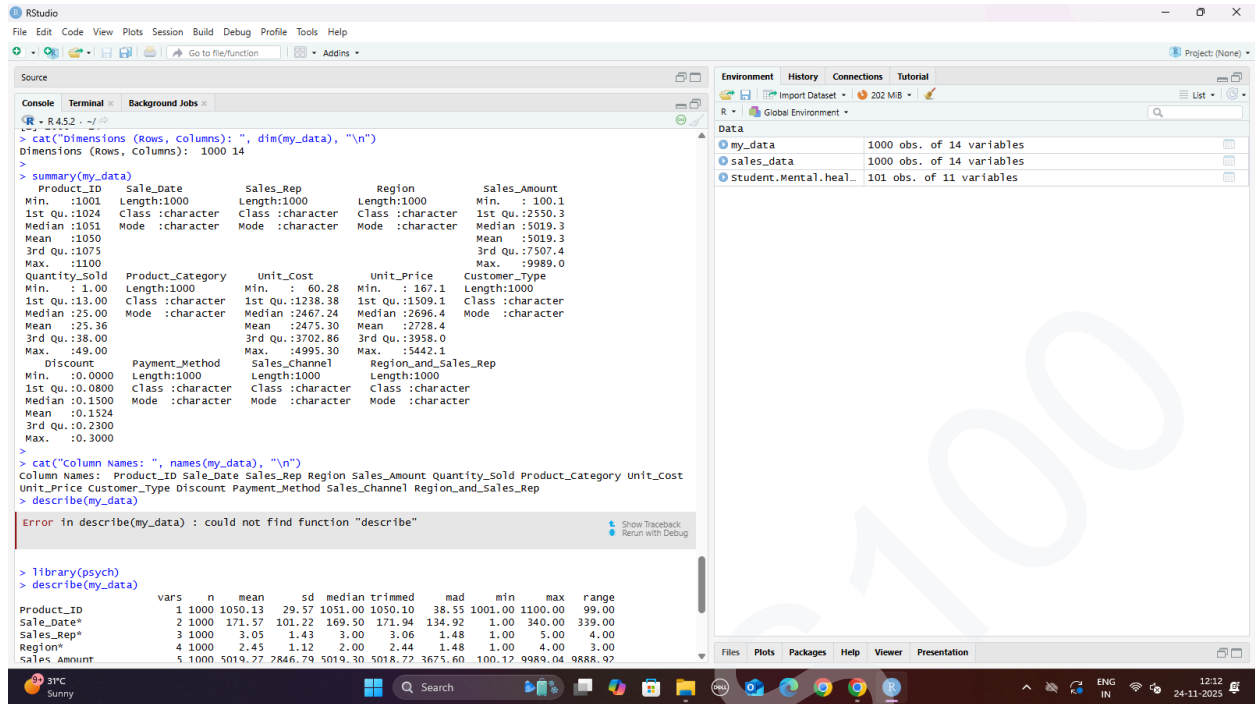
The Environment pane on the right shows the loaded data:

| Object | Class | Size |
|-----------------------|------------|---------------------------|
| my_data | data.frame | 1000 obs. of 14 variables |
| sales_data | data.frame | 1000 obs. of 14 variables |
| Student.Mental.health | data.frame | 101 obs. of 11 variables |

NANDINI PANDIT S100

Data Analysis with SAS / SPSS / R Practical

MVLU COLLEGE.



The screenshot shows the RStudio interface with the following components:

- Source Panel:** Contains R code for data inspection.
- Console Panel:** Displays the output of the R code.
- Environment Panel:** Lists loaded objects: `my_data` (1000 obs. of 14 variables), `sales_data` (1000 obs. of 14 variables), and `Student.Mental.heal...` (101 obs. of 11 variables).
- Terminal Panel:** Shows an error message: `Error in describe(my_data) : could not find function "describe"`.

R Code in Source Panel:

```
> cat("Dimensions (Rows, Columns): ", dim(my_data), "\n")
Dimensions (Rows, Columns): 1000 14

> summary(my_data)
  Product_ID  Sale_Date    Sales_Rep      Region    Sales_Amount
Min.   :1001  Length:1000    Length:1000    Length:1000    Min.   : 100.1
1st Qu.:1024  Class :character  Class :character  Class :character  1st Qu.:2550.3
Median :1051  Mode  :character  Mode  :character  Mode  :character  Median :5019.3
Mean   :1050                                     Mean   :5019.3
3rd Qu.:1075                                     3rd Qu.:7507.4
Max.   :1100                                     Max.   :9989.0
Quantity_Sold Product_Category Unit_Cost    Unit_Price Customer_Type
Min.   : 1.00  Length:1000    Min.   : 60.28  Min.   :167.1  Length:1000
1st Qu.:13.00  Class :character  1st Qu.:1238.38 1st Qu.:1509.1 Class :character
Median :25.00  Mode  :character  Median :2467.24 Median :2696.4 Mode  :character
Mean   :25.36                                     Mean   :2475.30 Mean   :2728.4
3rd Qu.:38.00                                     3rd Qu.:3702.86 3rd Qu.:3958.0
Max.   :49.00                                     Max.   :4995.30 Max.   :5442.1
Discount  Payment_Method Sales_Channel Region_and_Sales_Rep
Min.   :0.0000  Length:1000    Length:1000    Length:1000
1st Qu.:0.0800  Class :character  Class :character  Class :character
Median :0.1500  Mode  :character  Mode  :character  Mode  :character
Mean   :0.1524                                     Mean   :0.1524
3rd Qu.:0.2300                                     3rd Qu.:0.2300
Max.   :0.3000                                     Max.   :0.3000

> cat("Column Names: ", names(my_data), "\n")
Column Names: Product_ID Sale_Date Sales_Rep Region Sales_Amount Quantity_Sold Product_Category Unit_Cost
Unit_Price Customer_Type Discount Payment_Method Sales_Channel Region_and_Sales_Rep
> describe(my_data)
Error in describe(my_data) : could not find function "describe"
```

Output in Console Panel:

```
> library(psych)
> describe(my_data)
      vars   n  mean    sd median trimmed   mad   min   max   range
Product_ID 1 1000 1050.13 29.57 1051.00 1050.10 38.55 1001.00 1100.00  99.00
Sale_Date* 2 1000 171.57 101.22 169.50 171.94 134.92  1.00 340.00 339.00
Sales_Rep*  3 1000  3.05  1.43  3.00  3.06  1.48  1.00  5.00  4.00
Region*    4 1000  2.45  1.12  2.00  2.44  1.48  1.00  4.00  3.00
Sales_Amount 5 1000 5019.27 2846.79 5019.30 5018.72 3675.60 100.12 9989.04 9888.92
```