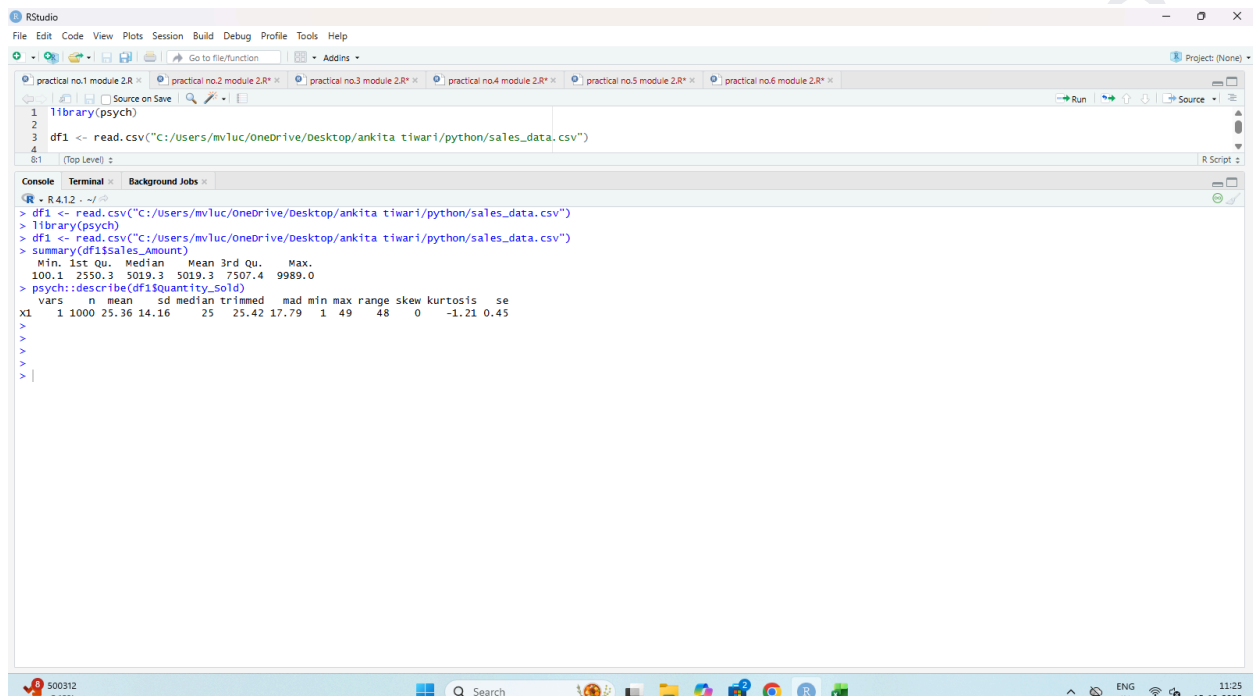


# MVLU COLLEGE

**AIM:** Practical 1 to 6 Generating descriptive statistics using summary() or describe() (R). Generating frequency tables using table() or count() (R). Creating cross-tabulations and two-way tables using table() (R). Performing one-sample t-tests using t.test() (R). Performing independent two-sample t-tests using t.test() with grouping (R). Performing paired t-tests using t.test(paired=TRUE) (R).

## PRACTICAL NO. 1 OUTPUT:



```
1 library(psych)
2
3 df1 <- read.csv("C:/Users/mvluc/OneDrive/Desktop/ankita tiwari/python/sales_data.csv")
4
5 summary(df1$Sales_Amount)
6
7 psych::describe(df1$Quantity_Sold)
8
9 vars      n mean    sd median trimmed  mad min max range skew kurtosis  se
10 X1      1000 25.36 14.16    25  25.42 17.79    1  49  48    0  -1.21 0.45
```

The screenshot shows the RStudio interface. The script editor contains R code for loading a CSV file and performing descriptive statistics. The console shows the output of these commands, including a summary of the 'Sales\_Amount' variable and a detailed psychometric description of the 'Quantity\_Sold' variable.

**ANKITA TIWARI**

**SYCS**

**S122**

**DATA ANALYSIS PRACTICAL NO.1-6 MODULE 2**

# MVLU COLLEGE

## PRACTICAL NO. 2 OUTPUT:

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Background Jobs
R - R 4.1.2 - ~/R
> # Load the dplyr library
> library(dplyr)
>
> # Read the csv file
> df2 <- read.csv("C:/Users/mvlu/OneDrive/Desktop/ankita tiwari/python/sales_data.csv")
>
> # Display the count of each unique value in the 'Region' column
> table(df2$Region)

East North South West
263 267 226 244
>
> # Count the number of occurrences for each 'Product_Category'
> df2 %>%
+   count(Product_Category)
+   count(Product_Category)
1      Clothing 268
2      Electronics 246
3          Food 226
4      Furniture 260
>
> # Load the dplyr library
> library(dplyr)
>
> # Read the csv file
> df2 <- read.csv("C:/Users/mvlu/OneDrive/Desktop/ankita tiwari/python/sales_data.csv")
>
> # Display the count of each unique value in the 'Region' column
> table(df2$Region)

East North South West
263 267 226 244
>
> # Count the number of occurrences for each 'Product_Category'
> df2 %>%
+   count(Product_Category)
+   count(Product_Category)
1      Clothing 268
2      Electronics 246
3          Food 226
4      Furniture 260
> # Load the dplyr library
```

## PRACTICAL NO. 3 OUTPUT:

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
practical no.1 module 2.R practical no.2 module 2.R practical no.3 module 2.R practical no.4 module 2.R practical no.5 module 2.R practical no.6 module 2.R
1 df3 <- read.csv("C:/Users/mvlu/OneDrive/Desktop/ankita tiwari/python/Student Mental health.csv", stringsAsFactors = FALSE)
2
6:1 (Top Level)
R Script
Console Terminal Background Jobs
R - R 4.1.2 - ~/R
> df3 <- read.csv("C:/Users/mvlu/OneDrive/Desktop/ankita tiwari/python/Student Mental health.csv", stringsAsFactors = FALSE)
>
> colnames(df3)
[1] "timestamp" "choose.your.gender" "Age"
[4] "what.is.your.course." "your.current.year.of.study" "what.is.your.CGPA."
[7] "Marital.status" "do.you.have.Depression." "do.you.have.Anxiety."
[10] "do.you.have.Panic.attack." "did.you.seek.any.specialist.for.a.treatment."
>
> table(df3$choose.your.gender, df3$do.you.have.Depression?)
< table of extent 0 x 0 >
>
```

**ANKITA TIWARI**

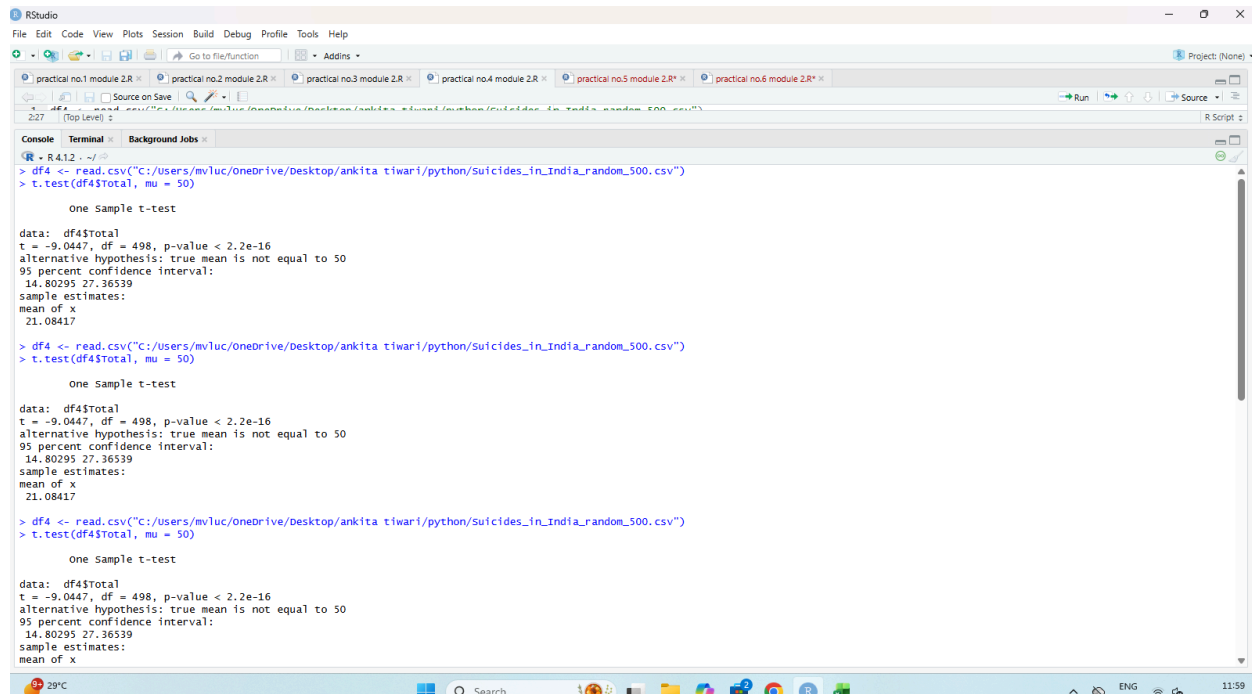
**SYCS**

**S122**

**DATA ANALYSIS PRACTICAL NO.1-6 MODULE 2**

# MVLU COLLEGE

## PRACTICAL NO. 4 OUTPUT:



```
R - R4.1.2 ~ /
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins Project: (None)
practical no.1 module 2.R practical no.2 module 2.R practical no.3 module 2.R practical no.4 module 2.R practical no.5 module 2.R practical no.6 module 2.R
Source on Save
227 (Top Level)
Console Terminal Background Jobs
R - R4.1.2 ~ /
> df4 <- read.csv("C:/Users/mvlu/OneDrive/desktop/ankita tiwari/python/Suicides_in_India_random_500.csv")
> t.test(df4$total, mu = 50)

One Sample t-test

data: df4$total
t = -9.0447, df = 498, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 50
95 percent confidence interval:
 14.80295 27.36539
sample estimates:
mean of x
 21.08417

> df4 <- read.csv("C:/Users/mvlu/OneDrive/desktop/ankita tiwari/python/Suicides_in_India_random_500.csv")
> t.test(df4$total, mu = 50)

One Sample t-test

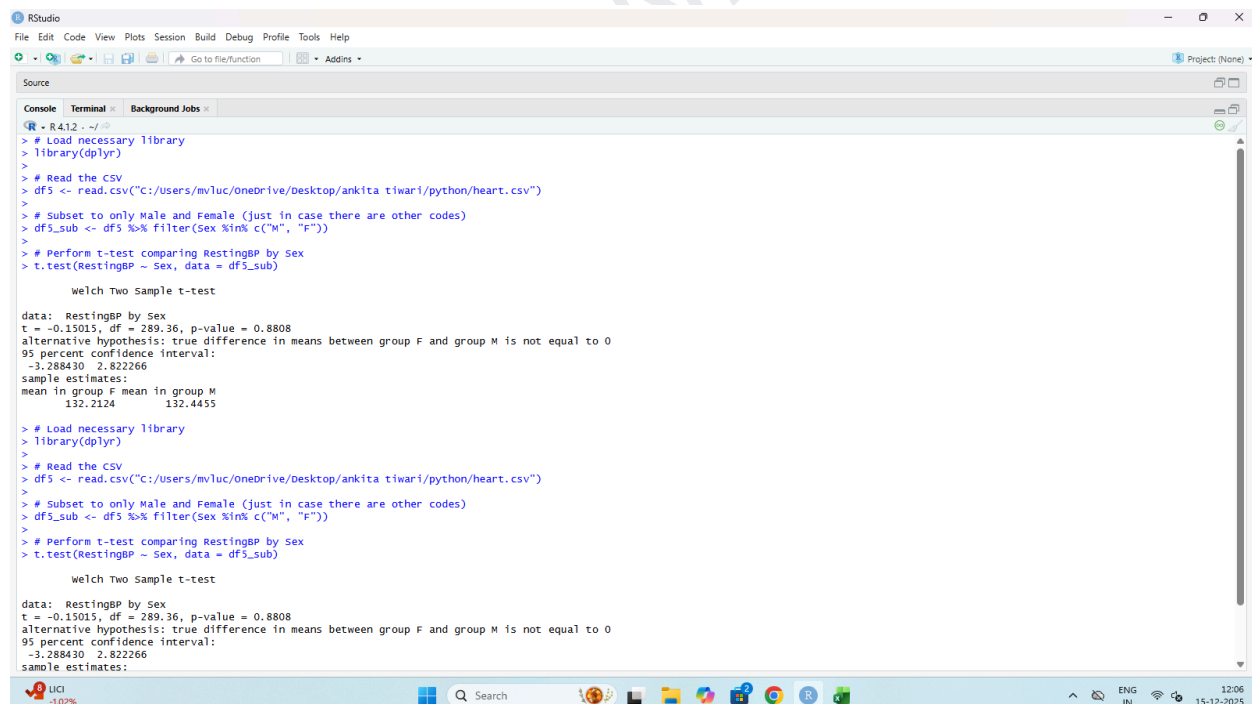
data: df4$total
t = -9.0447, df = 498, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 50
95 percent confidence interval:
 14.80295 27.36539
sample estimates:
mean of x
 21.08417

> df4 <- read.csv("C:/Users/mvlu/OneDrive/desktop/ankita tiwari/python/Suicides_in_India_random_500.csv")
> t.test(df4$total, mu = 50)

One Sample t-test

data: df4$total
t = -9.0447, df = 498, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 50
95 percent confidence interval:
 14.80295 27.36539
sample estimates:
mean of x
 21.08417
```

## PRACTICAL NO. 5 OUTPUT:



```
R - R4.1.2 ~ /
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins Project: (None)
Source
Console Terminal Background Jobs
R - R4.1.2 ~ /
> # Load necessary library
> library(dplyr)
>
> # Read the CSV
> df5 <- read.csv("C:/Users/mvlu/OneDrive/desktop/ankita tiwari/python/heart.csv")
>
> # Subset to only Male and Female (just in case there are other codes)
> df5_sub <- df5 %>% filter(sex %in% c("M", "F"))
>
> # Perform t-test comparing RestingBP by Sex
> t.test(RestingBP ~ Sex, data = df5_sub)

Welch Two Sample t-test

data: RestingBP by Sex
t = -0.15015, df = 289.36, p-value = 0.8808
alternative hypothesis: true difference in means between group F and group M is not equal to 0
95 percent confidence interval:
 -3.288430  2.822266
sample estimates:
mean in group F mean in group M
 132.2124      132.4455

> # Load necessary library
> library(dplyr)
>
> # Read the CSV
> df5 <- read.csv("C:/Users/mvlu/OneDrive/desktop/ankita tiwari/python/heart.csv")
>
> # Subset to only Male and Female (just in case there are other codes)
> df5_sub <- df5 %>% filter(sex %in% c("M", "F"))
>
> # Perform t-test comparing RestingBP by Sex
> t.test(RestingBP ~ Sex, data = df5_sub)

Welch Two Sample t-test

data: RestingBP by Sex
t = -0.15015, df = 289.36, p-value = 0.8808
alternative hypothesis: true difference in means between group F and group M is not equal to 0
95 percent confidence interval:
 -3.288430  2.822266
sample estimates:
mean in group F mean in group M
 132.2124      132.4455
```

ANKITA TIWARI

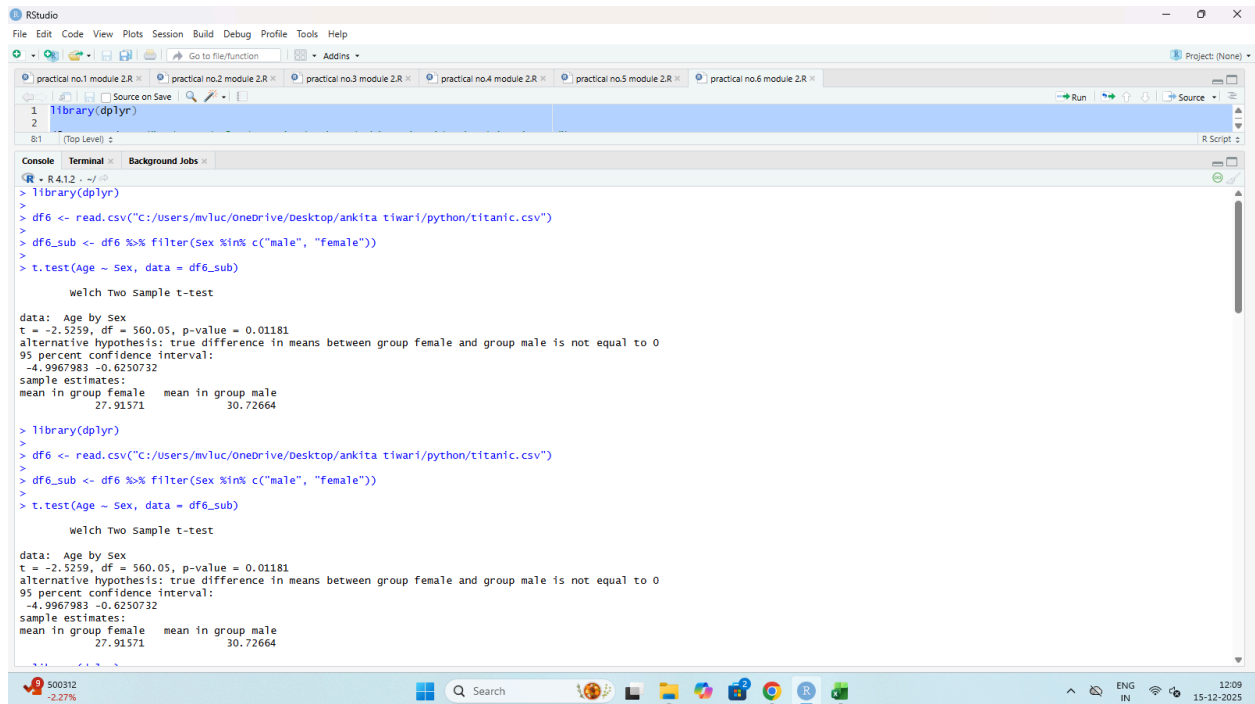
SYCS

S122

DATA ANALYSIS PRACTICAL NO.1-6 MODULE 2

# MVLU COLLEGE

## PRACTICAL NO. 6 OUTPUT:



```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
practical no.1 module 2.R practical no.2 module 2.R practical no.3 module 2.R practical no.4 module 2.R practical no.5 module 2.R practical no.6 module 2.R
1 library(dplyr)
2
8:1 (Top Level)
Console Terminal Background Jobs
R - R 4.1.2 - ~/R
> library(dplyr)
> df6 <- read.csv("C:/Users/mvlu/OneDrive/Desktop/ankita tiwari/python/titanic.csv")
> df6_sub <- df6 %>% filter(Sex != "male", "female")
> t.test(Age ~ Sex, data = df6_sub)

Welch Two Sample t-test

data: Age by Sex
t = -2.5259, df = 560.05, p-value = 0.01181
alternative hypothesis: true difference in means between group female and group male is not equal to 0
95 percent confidence interval:
-4.9967983 -0.6250732
sample estimates:
mean in group female mean in group male
27.91571 30.72664

> library(dplyr)
> df6 <- read.csv("C:/Users/mvlu/OneDrive/Desktop/ankita tiwari/python/titanic.csv")
> df6_sub <- df6 %>% filter(Sex != "male", "female")
> t.test(Age ~ Sex, data = df6_sub)

Welch Two Sample t-test

data: Age by Sex
t = -2.5259, df = 560.05, p-value = 0.01181
alternative hypothesis: true difference in means between group female and group male is not equal to 0
95 percent confidence interval:
-4.9967983 -0.6250732
sample estimates:
mean in group female mean in group male
27.91571 30.72664
```

**ANKITA TIWARI**

**SYCS**

**S122**

**DATA ANALYSIS PRACTICAL NO.1-6 MODULE 2**