

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:



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(psych)
2 df1 <- read.csv("C:/users/mvluc/OneDrive/Desktop/ankita_tiwari/python/sales_data.csv")
3
4
51 (Top Level) :
```

Console Terminal Background Jobs

```
R > R 4.1.2 - ~/ ~
> df1 <- read.csv("C:/users/mvluc/OneDrive/Desktop/ankita_tiwari/python/sales_data.csv")
> library(psych)
> dft <- read.csv("C:/users/mvluc/OneDrive/Desktop/ankita_tiwari/python/sales_data.csv")
> summary(df1$Sales_Amount)
   Min. 1st Qu. Median  Mean 3rd Qu. Max.
100.1  2550.3 5019.3 7507.4 9989.0
> psych::describe(df1$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
>
>
>
> |
```

500312

Search

ENG 11:25

ANKITA TIWARI

SYCS

S122

DATA ANALYSIS PRACTICAL NO.1-6 MODULE 2

MVLU COLLEGE

PRACTICAL NO. 2 OUTPUT:

PRACTICAL NO. 3 OUTPUT:

RStudio

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 Run Source ▾

1 df3 <- read.csv("C:/users/mvluc/OneDrive/Desktop/ankita tiwari/python/Student Mental health.csv", stringsAsFactors = FALSE)

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

(Top Level) 3

Console Terminal Background Jobs

R 4.1.2 - ~

> df3 <- read.csv("C:/users/mvluc/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 >

>

29°C Sunny ENG INR 11:57 15-12-2023

ANKITA TIWARI

SYCS

S122

DATA ANALYSIS PRACTICAL NO.1-6 MODULE 2

MVLU COLLEGE

PRACTICAL NO. 4 OUTPUT:

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source 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
2:27 (Top Level) 2
Console Terminal Background Jobs
R - R 4.1.2 - ~/ ~
> df4 <- read.csv("C:/users/mvluc/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 = 14.80295, 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/mvluc/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/mvluc/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:

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Go to file/function Addins Project: (None)
Console Terminal Background Jobs
R - R 4.1.2 - ~/ ~
> # Load necessary library
> library(dplyr)
>
> # Read the CSV
> df5 <- read.csv("C:/users/mvluc/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 = 0.25915, 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/mvluc/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:
```

ANKITA TIWARI

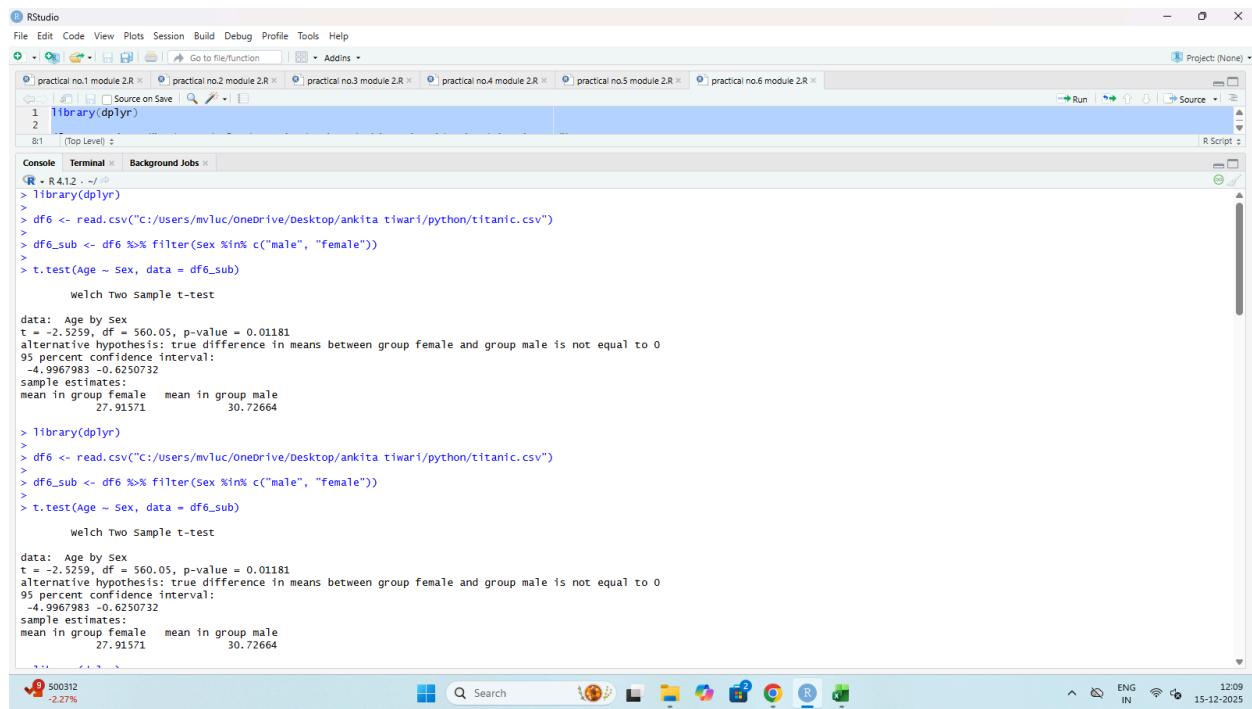
SYCS

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
Go to file/function Addins Project: (None)
1 library(dplyr)
2
81 (Top Level) R Script
Console Terminal Background Jobs
R > R 4.1.2 - ~/practical no.6 module 2.R
> library(dplyr)
> df6 <- read.csv("C:/users/mvluc/onedrive/desktop/ankita tiwari/python/titanic.csv")
> df6_sub <- df6 %>% filter(Sex %in% c("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:
-0.9967983 -0.6250732
sample estimates:
mean in group female mean in group male
27.91571 30.72664

> library(dplyr)
> df6 <- read.csv("C:/users/mvluc/onedrive/Desktop/ankita tiwari/python/titanic.csv")
> df6_sub <- df6 %>% filter(Sex %in% c("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:
-0.9967983 -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