

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

PRACTICAL NO 4

AIM: Performing one-sample t-tests using `t.test()` (R).

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Environment History Connections Tutorial
R > Global Environment
Console Terminal Background Jobs
> library(dplyr)
> library(readr)
> # Load dataset
> df <- readr::read_csv("exams.csv")
Row: 1000 Columns: 8
-- Column specification --
dollar_sign: " "
chr (5): student_id, race/ethnicity, parental level of education, lunch, test preparation course
dbl (3): math score, reading score, writing score
i uses `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
> # Rename math score as Marks (to keep logic similar)
> df <- df %>% rename(Marks = math score)
> # Summary statistics
> print("Summary of Marks:")
[1] "Summary of Marks:"
> summary(df$Marks)
Min. 1st Qu. Median Mean 3rd Qu. Max.
15.00 39.00 66.5 66.4 77.0 100.0
> # Descriptive statistics (base & alternative)
> print("Descriptive statistics of Marks:")
[1] "Descriptive statistics of Marks:"
> mean(df$Marks)
[1] 66.396
> sd(df$Marks)
[1] 15.40287
> nrow(df$Marks)
[1] 13
> max(df$Marks)
[1] 100
> # Frequency table: Test preparation course
> prep_counts <- table(df$test_prep_course)
> print("Frequency Table: Test Preparation Course")
[1] "Frequency Table: test_prep_course"
> print(prep_counts)

completed      none
333          665
> # Data frame format
> prep_df <- df %>% count("test_prep_course")
> print("Test Preparation Course Frequency (Data Frame Format)")
[1] "Test Preparation Course Frequency (Data Frame Format)"
> print(prep_df)

# A tibble: 2 x 2
  test_prep_course     n
  <fct>            <int>
1 completed           333
2 none                665
> # Create marks groups (Low / Medium / High)
> dfMarks_group <- cut(
+   dfMarks,
+   breaks = 3,
+   labels = c("Low", "Medium", "High"))

```

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Environment History Connections Tutorial
R > Global Environment
Console Terminal Background Jobs
> dfMarks_group <- cut(
+   dfMarks,
+   breaks = 3,
+   labels = c("Low", "Medium", "High"))
> print(cross_tabulation("test_prep_course vs Marks Group"))
[1] "Cross Tabulation: Test Prep course vs Marks Group"
> cross_tab <- table(df$test_prep_course, dfMarks_Group)
> print(cross_tab)

Low Medium High
completed 11 166 158
none 52 393 221
> # One-sample t-test: Is mean Marks different from 50?
> print("One-sample t-test: Marks vs mu = 50")
[1] "One-sample t-test: Marks vs mu = 50"
> t_test_one <- t.test(df$Marks, mu = 50)
> print(t_test_one)

One Sample t-test

data: df$Marks
t = 33.662, df = 999, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 50
95 percent confidence interval:
65.4484 67.35182
sample estimates:
mean of x
66.396

> # Independent t-test: Marks by test preparation course
> print("Independent t-test: Marks by Test Preparation Course")
[1] "Independent t-test: Marks by Test Preparation Course"
> t_test_indep <- t.test(
+   Marks ~ test_prep_course,
+   data = df
+ )
> print(t_test_indep)

Welch Two Sample t-test

data: Marks by test_prep_course
t = 4.8511, df = 670.68, p-value = 1.528e-06
alternative hypothesis: true difference in means between group completed and group none is not equal to 0
95 percent confidence interval:
2.945390 6.953053
sample estimates:
mean in group completed      mean in group none
69.68657      64.73835

> # Paired t-test: Math vs Reading scores
> print("Paired t-test: Math vs Reading scores")
[1] "Paired t-test: Math vs Reading Scores"
> t_test_paired <- t.test(

```

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

The screenshot shows the RStudio interface with the following details:

- Console:** Displays R code and its output. The code includes:
 - Creating a table for 'Test Prep Course vs Marks Group'.
 - Performing a Chi-Square test for independence between 'Test Prep Course' and 'Marks Group'.
 - Performing one-sample t-tests for 'Marks' against a mean of 50.
 - Performing independent t-tests for 'Marks' by 'test preparation course'.
 - Performing Welch Two Sample t-test for 'Marks' by 'test preparation course'.
 - Performing paired t-tests for 'Math vs Reading Scores'.
- Environment:** Shows the global environment with objects like df, employee_salary_dataset, exams, prep_df, superstore, t_test_one, t_test_two, courses_counts, courses_freq, cross_tab, marks_freq, and marks_group.
- Data:** Shows the structure of various datasets:
 - df: 1000 obs. of 9 variables
 - employee_salary_dataset: 50 obs. of 9 variables
 - exams: 1000 obs. of 8 variables
 - prep_df: 2 obs. of 2 variables
 - superstore: 51290 obs. of 27 variables
 - t_test_one: List of 10
 - t_test_two: List of 10
 - values: courses_counts, courses_freq, cross_tab, marks_freq, marks_group
- Files:** Shows files in the project directory: employee_salary_dataset.csv, superstore.csv, and exams.csv.
- System:** Shows system status including battery level (27%), network (ENG IN), and date/time (15-12-2025).