R Textbook Companion for Statistics for Psychology by Arthur Aron, Elliot J. Coups and Elaine N. Aron[1]

Created by
Shivam Agrawal
B.Tech.
Computer Science and Engineering
Techno India NJR Institute of Technology
Cross-Checked by
R TBC Team

May 23, 2021

[1]Funded by a grant from the National Mission on Education through ICT - http://spoken-tutorial.org/NMEICT-Intro. This Textbook Companion and R codes written in it can be downloaded from the "Textbook Companion Project" section at the website - https://r.fossee.in.

Book Description

Title: Statistics for Psychology

Author: Arthur Aron, Elliot J. Coups and Elaine N. Aron

Publisher: Pearson, USA

Edition: 6

Year: 2013

ISBN: 9780205923922

R numbering policy used in this document and the relation to the above book.

Exa Example (Solved example)

Eqn Equation (Particular equation of the above book)

For example, Exa 3.51 means solved example 3.51 of this book. Sec 2.3 means an R code whose theory is explained in Section 2.3 of the book.

Contents

List of R Codes

Chapter 1

Displaying the Order in a Group of Numbers Using Tables and Graphs

R code Exa 1.1a Frequency Tables an Example

```
# Page no. : 7

stress_rating <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3,
    7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7,
    10, 8, 8)

frequency_table <- data.frame(table(stress_rating))

percent <- round((prop.table(frequency_table$Freq) *
    100), 1)

frequency_table <- cbind(frequency_table, percent)

View(frequency_table)
```

R code Exa 1.1c Frequency Tables Another Example

```
# Page no. : 8 - 9

social_interaction <- c(48, 15, 33, 3, 21, 19, 17,
    16, 44, 25, 30, 3, 5, 9, 35, 32, 26, 13, 14,
                        14, 47, 47, 18, 11, 5, 19,
                           24, 17, 6, 25, 8, 18, 29,
                            1, 18, 22, 3, 22,
                        29, 2, 6, 10, 29, 10, 29,
                           21, 38, 41, 16, 17, 8,
                           40, 8, 10, 18, 7, 4, 4,
                        8, 11, 3, 23, 10, 19, 21,
                           13, 12, 10, 4, 17, 11,
                           21, 9, 8, 7, 5, 3, 22,
                        14, 25, 4, 11, 10, 18, 1,
                           28, 27, 19, 24, 35, 9,
                           30, 8, 26)

breaks <- seq(0, 49, by = 1)

social.interaction <- cut(social_interaction, breaks
    , right = F)

frequency_table <- data.frame(table(social.
    interaction))

View(frequency_table)
```

R code Exa 1.1d Grouped Frequency Tables

```
# Page no. : 9

stress_rating <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3,
     7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7,
```

```
10, 8, 8)

breaks <- seq(-1, 11, by = 2)

stress.rating <- cut(stress_rating, breaks, right = T)

grouped_frequency_table <- data.frame(table(stress.rating))

percent <- round((prop.table(grouped_frequency_table$Freq) * 100), 1)

grouped_frequency_table <- cbind(grouped_frequency_table, percent)

View(grouped_frequency_table)
```

R code Exa 1.1e Grouped Frequency Tables Another Example

```
# Page no. : 10

social_interaction <- c(48, 15, 33, 3, 21, 19, 17, 16, 44, 25, 30, 3, 5, 9, 35, 32, 26, 13, 14,
                        14, 47, 47, 18, 11, 5, 19, 24, 17, 6, 25, 8, 18, 29, 1, 18, 22, 3, 22,
                        29, 2, 6, 10, 29, 10, 29, 21, 38, 41, 16, 17, 8, 40, 8, 10, 18, 7, 4, 4,
                        8, 11, 3, 23, 10, 19, 21, 13, 12, 10, 4, 17, 11, 21, 9, 8, 7, 5, 3, 22,
                        14, 25, 4, 11, 10, 18, 1, 28, 27, 19, 24, 35, 9,
```

```
30, 8, 26)

breaks <- seq(0, 50, by = 5)

social.interaction <- cut(social_interaction, breaks, right = F)

grouped_frequency_table <- data.frame(table(social.interaction))

percent <- round((prop.table(grouped_frequency_table$Freq) * 100), 1)

grouped_frequency_table <- cbind(grouped_frequency_table, percent)

View(grouped_frequency_table)
```

R code Exa 1.2a Histogram for Frequency Table

```
# Page no. : 11

stress_rating <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3, 7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7, 10, 8, 8)

frequency_table <- data.frame(table(stress_rating))

library(ggplot2)

ggplot(data = frequency_table, aes(x = stress_rating, y = Freq)) +
  geom_bar(stat = "identity", fill = "violet") +
  labs(title = "Histogram for frequency table", x = "Stress Rating", y = "Frequency") +
```

```
  theme_bw()
```

R code Exa 1.2b Histogram for Grouped Frequency Table

```
# Page no. : 11

stress_rating <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3,
    7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7,
    10, 8, 8)

frequency_table <- data.frame(table(stress_rating))

View(frequency_table)

library(ggplot2)

ggplot(data = frequency_table, aes(x = stress_rating
   , y = Freq)) +
  geom_bar(stat = "identity", fill = "violet") +
  labs(title = "Histogram for grouped frequency
     table", x = "Stress Rating", y = "Frequency") +
  theme_bw()
```

R code Exa 1.2c Histogram for Grouped Frequency Table Example 2

```
# Page no. : 12

social_interaction <- c(48, 15, 33, 3, 21, 19, 17,
   16, 44, 25, 30, 3, 5, 9, 35, 32, 26, 13, 14,
                        14, 47, 47, 18, 11, 5, 19,
                            24, 17, 6, 25, 8, 18, 29,
                             1, 18, 22, 3, 22,
```

```
                        29, 2, 6, 10, 29, 10, 29, 21, 38, 41, 16, 17, 8, 40, 8, 10, 18, 7, 4, 4,
                        8, 11, 3, 23, 10, 19, 21, 13, 12, 10, 4, 17, 11, 21, 9, 8, 7, 5, 3, 22,
                        14, 25, 4, 11, 10, 18, 1, 28, 27, 19, 24, 35, 9, 30, 8, 26)

breaks <- seq(0, 50, by = 5)

mid_values <- seq(0+5/2,50-5/2, 5)

social.interaction <- cut(social_interaction, breaks, right = F)

grouped_frequency_table <- data.frame(table(social.interaction))
grouped_frequency_table$social_interacton_mid <- mid_values

View(grouped_frequency_table)

library(ggplot2)

ggplot(data = grouped_frequency_table, aes(x = factor(social_interacton_mid), y = Freq)) +
  geom_bar(stat = "identity", fill = "violet") +
  labs(title = "Histogram for grouped frequency table", x = "Social Interaction Mid Values",
       y = "Frequency") +
  theme_bw()
```

R code Exa 1.2d Histogram for Nominal Variables

```
# Page no. : 14

closest_person <- c("Family member", "Nonromantic friend", "Romantic friend", "Other")

frequency <- c(33, 76, 92, 7)

frequency_table <- data.frame(closest_person, frequency)

library(ggplot2)

ggplot(data = frequency_table, aes(x = closest_person, y = frequency)) +
  geom_bar(stat = "identity", fill = "violet") +
  labs(title = "Histogram for nominal variables", x = "Closest Person", y = "Frequency") +
  theme_bw()
```

R code Exa 1.2e How are you doing

```
# Page no. : 15

value <- c(1, 2, 3, 4, 5)

frequency <- c(3, 4, 8, 5, 2)

frequency_table <- data.frame(value, frequency)

library(ggplot2)

ggplot(data = frequency_table, aes(x = value, y = frequency)) +
  geom_bar(stat = "identity", fill = "violet") +
  labs(title = "Histogram", x = "Value", y = "
```

```
    Frequency") +
  theme_bw()
```

R code Exa 1.4a Worked Out Examples

```
# Page no. : 25

Interest <- c(2, 4, 5, 5, 1, 3, 6, 3, 6, 6)

frequency_table <- data.frame(table(Interest))

Percent <- round((prop.table(frequency_table$Freq) *
    100), 0)

frequency_table <- cbind(frequency_table, Percent)

View(frequency_table)

library(ggplot2)

ggplot(data = frequency_table, aes(x = Interest, y =
    Freq)) +
  geom_bar(stat = "identity", fill = "violet") +
  labs(title = "Histogram", x = "Interest", y = "
    Frequency")
```

Chapter 2

Central Tendency and Variability

R code Exa 2.1a Mean Example 1

```
# Page no. : 36 - 37

x <- c(7, 8, 8, 7, 3, 1, 6, 9, 3, 8)

value <- mean(x)

cat("Mean of x is", value)
```

R code Exa 2.1b Mean Example 2

```
# Page no. : 37

x <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3, 7, 6, 5, 0,
    9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7, 10, 8, 8)

value <- mean(x)
```

```
cat("Mean of x is", round(value,2))

library(ggplot2)

ggplot(data = data.frame(x), aes(x = x)) +
  geom_bar(stat = "count", fill = "violet") +
  geom_vline(xintercept = mean(x), col = "red", lwd
    = 0.8) +
  labs(title = "Histogram with mean value", x = "
    Stress Rating", y = "Frequency") +
  theme_bw()
```

R code Exa 2.1c Mean Example 3

```
# Page no. : 38

x <- c(48, 15, 33, 3, 21, 19, 17, 16, 44, 25, 30, 3,
    5, 9, 35, 32, 26, 13, 14, 14, 47, 47, 18,
       11, 5, 19, 24, 17, 6, 25, 8, 18, 29, 1, 18,
          22, 3, 22, 29, 2, 6, 10, 29, 10, 29, 21,
          38,
       41, 16, 17, 8, 40, 8, 10, 18, 7, 4, 4, 8, 11,
          3, 23, 10, 19, 21, 13, 12, 10, 4, 17, 11,
       21, 9, 8, 7, 5, 3, 22, 14, 25, 4, 11, 10, 18,
          1, 28, 27, 19, 24, 35, 9, 30, 8, 26)

value <- mean(x)

cat("Mean of x is", round(value, 2))

library(ggplot2)

ggplot(data = data.frame(table(x)), aes(x = x, y =
   Freq)) +
```

```
  geom_bar(stat = "identity", fill = "violet") +
  geom_vline(xintercept = mean(x), col = "red", lwd = 0.8) +
  labs(title = "Histogram with mean value", x = "Social Interaction", y = "Frequency") +
  theme_bw()
```

R code Exa 2.1d How are you doing

```
# Page no. : 43

scores <- c(2, 3, 3, 6, 6)

value <- mean(scores)

cat("Mean value of scores is", value)

scores <- c(5, 3, 2, 13, 2)

value <- mean(scores)

cat("\nMean value of scores is", value)

mode <- function(v)
{
  x <- unique(v)
  x[which.max(tabulate(match(v, x)))]
}

value <- mode(scores)

cat("\nMode of scores is", value)

value <- median(scores)

```

```
cat("\nMedian of scores is", value)
```

R code Exa 2.2a Variance and Standard Deviation Example 1

```
# Page no. : 48

scores <- c(7, 8, 8, 7, 3, 1, 6, 9, 3, 8)

library(rafalib)

variance <- popvar(scores)

standard_deviation <- popsd(scores)

cat("Variance is", round(variance, 2), "and Standard Deviation is", round(standard_deviation, 2))
```

R code Exa 2.2b Variance and Standard Deviation Example 2

```
# Page no. : 49

interactions <- c(48, 15, 33, 3, 21, 19, 17, 16, 44, 25, 30, 3, 5, 9, 35, 32, 26, 13, 14,
                  14, 47, 47, 18, 11, 5, 19, 24, 17, 6, 25, 8, 18, 29, 1, 18, 22, 3, 22,
                  29, 2, 6, 10, 29, 10, 29, 21, 38, 41, 16, 17, 8, 40, 8, 10, 18, 7, 4, 4,
                  8, 11, 3, 23, 10, 19, 21, 13, 12, 10, 4, 17, 11, 21, 9, 8, 7, 5, 3, 22,
```

```
                  14, 25, 4, 11, 10, 18, 1, 28, 27,
                     19, 24, 35, 9, 30, 8, 26)

library(rafalib)

variance <- popvar(interactions)

standard_deviation <- popsd(interactions)

cat("Variance is", round(variance, 2), "and Standard
    Deviation is", round(standard_deviation, 2))
```

R code Exa 2.2c How are you doing

```
# Page no. : 51 - 52

scores <- c(2, 4, 3, 7)

library(rafalib)

variance <- popvar(scores)

standard_deviation <- popsd(scores)

cat("Variance is", round(variance, 2), "and Standard
    Deviation is", round(standard_deviation, 2))
```

R code Exa 2.3a Worked Out Examples

```
# Page no. : 58 - 60

scores <- c(8, 6, 6, 9, 6, 5, 6, 2)

```

```
value <- mean(scores)

cat("Mean is", value)

scores <- c(1, 7, 4, 2, 3, 6, 2, 9, 7)

value <- median(scores)

cat("\nMedian is", value)

scores <- c(8, 6, 6, 9, 6, 5, 6, 2)

value <- sum((scores - mean(scores))**2)

cat("\nSum of Squares is", value)

library(rafalib)

variance <- popvar(scores)

cat("\nVariance is", variance)

standard_deviation <- popsd(scores)

cat("\nStandard Deviation is", round(standard_
   deviation, 2))
```

Chapter 3

Some Key Ingredients for Inferential Statistics

R code Exa 3.1a Formula to change Raw Score to Z Score and Vice Versa

```
# Page no. : 71

x <- 8
m <- 12
sd <- 4

Z <- (x - m) / sd

cat("Z Score is", Z)

Z <- 1.5
sd <- 4
m <- 12

X <- (Z * sd) + m

cat("\nRaw Score is", X)
```

R code Exa 3.1b Additional Example 1 to change Raw Score to Z Score and Vice Versa

```
# Page no. : 72

m <- 3.40
sd <- 1.47
x <- 6

Z <- (x - m) / sd

cat("Z Score is", round(Z, 2))

Z <- -1.63

X <- (Z * sd) + m

cat("\nRaw Score is", round(X, 2))
```

R code Exa 3.1c Additional Example 2 to change Raw Score to Z Score and Vice Versa

```
# Page no. : 72 - 73

m <- 6.43
sd <- 2.56
x <- 10

Z <- (x - m) / sd

cat("Z Score is", round(Z, 2))

```

```
Z <- -1.73

X <- (Z * sd) + m

cat("\nRaw Score is", round(X, 2))
```

R code Exa 3.1d How are you doing

```
# Page no. : 73

m <- 20
sd <- 5

x1 <- 30

z1 <- (x1 - m) / sd

cat("Z Score is", round(z1, 2))

x2 <- 15

z2 <- (x2 - m) / sd

cat("\nZ Score is", round(z2, 2))

x3 <- 20

z3 <- (x3 - m) / sd

cat("\nZ Score is", round(z3, 2))

x4 <- 22.5

z4 <- (x4 - m) / sd

```

```
cat("\nZ Score is", round(z4, 2))
```

R code Exa 3.1e How are you doing

```
# Page no. : 73

m <- 10
sd <- 2

z1 <- 2

x1 <- (z1 * sd) + m

cat("Raw Score is", round(x1, 2))

z2 <- 0.5

x2 <- (z2 * sd) + m

cat("\nRaw Score is", round(x2, 2))

z3 <- 0

x3 <- (z3 * sd) + m

cat("\nRaw Score is", round(x3, 2))

z4 <- -3

x4 <- (z4 * sd) + m

cat("\nRaw Score is", round(x4, 2))
```

R code Exa 3.2a Examples for finding Percentage from Z Scores and Raw Scores using Normal Curve Table

```
# Page no. : 79 - 81

m <- 100
sd <- 15

IQ1 <- 125

Z1 <- (IQ1 - m) / sd

cat("Z Score is", round(Z1, 2))

percent <- pnorm(Z1, lower.tail = F) * 100

cat("\nPercentage of IQ level is", round(percent, 2)
    )

IQ2 <- 95

Z2 <- (IQ2 - m) / sd

cat("\nZ Score is", round(Z2, 2))

percent <- pnorm(Z2, lower.tail = F) * 100

cat("\nPercentage of IQ level is", round(percent, 2)
    )
```

R code Exa 3.2b Examples for finding Z Scores and Raw Scores from Percentage using Normal Curve Table

```
# Page no. : 81 - 84

```

```r
m <- 100
sd <- 15

percent <- 0.05

Z1 <- qnorm(percent, lower.tail = F)

cat("Z Score is", round(Z1, 2))

IQ1 <- (sd * Z1) + m

cat("\nRaw Score is", round(IQ1, 2))

percent <- 0.55

Z2 <- qnorm(percent, lower.tail = F)

cat("\nZ Score is", round(Z2, 2))

IQ2 <- (sd * Z2) + m

cat("\nRaw Score is", round(IQ2, 2))

percent <- 0.95

Z3 <- qnorm(0.975, lower.tail = T)

cat("\nZ Score is", round(Z3, 2))

IQ3 <- (sd * Z3) + m

cat("\nRaw Score is", round(IQ3, 2))
```

R code Exa 3.3a Worked Out Examples 1

```r
# Page no. : 100

m <- 80
sd <- 20
x <- 65

Z <- (x - m) / sd

cat("Z Score is", Z)

m <- 200
sd <- 50
z <- 1.26

X <- (z * sd) + m

cat("\nRaw Score is", X)
```

R code Exa 3.3b Worked Out Examples 2

```r
# Page no. : 100 - 102

m <- 20
sd <- 3
x <- 24

Z <- (x - m) / sd

cat("Z Score is", round(Z,2))

percent <- pnorm(Z, lower.tail = F) * 100

cat("\nPercentage of people have scores above 24 is"
    , round(percent, 2))

```

```r
m <- 20
sd <- 3
percent <- 0.75

Z <- qnorm(percent, lower.tail = F)

cat("\nZ Score is", round(Z, 2))

X <- (sd * Z) + m

cat("\nRaw Score is", round(X, 2))
```

R code Exa 3.3c Worked Out Examples 3

```r
# Page no. : 102 - 103

apple <- 20
strawberry <- 20
cherry <- 5
grape <- 5

total <- apple + strawberry + cherry + grape

case <- cherry + grape

probability <- case / total

cat("The required probability is", probability)
```

Chapter 4

Introduction to Hypothesis Testing

R code Exa 4.1a The Hypothesis Testing Process

```
# Page no. : 111 - 115

m <- 14
sd <- 3
x <- 6
alpha <- 0.01

cutoff <- qnorm(alpha)
cutoff <- round(cutoff, 2)

z <- (x - m) / sd

cat("\nCutoff Sample Z Score is ", cutoff)
cat("\nActual Sample Z Score is", round(z, 2))

if(z < cutoff)
{
  cat("\nWe reject null hypothesis.")
}else{
```

```
  cat("\nWe cannot reject null hypothesis.")
}
```

R code Exa 4.1b The Second Example

```
# Page no. : 116 – 118

m <- 70
sd <- 10
x <- 80
alpha <- 0.05

cutoff <- qnorm(1 - alpha)
cutoff <- round(cutoff, 2)

z <- (x - m) / sd

if(z > cutoff)
{
  cat("We reject null hypothesis.")
}else{
  cat("We cannot reject null hypothesis.")
}
```

R code Exa 4.1c How are you doing

```
# Page no. : 119 – 120

m <- 30
sd <- 4
x <- 40
alpha <- 0.05

```

```
cutoff <- qnorm(1 - alpha)
cutoff <- round(cutoff, 2)

z <- (x - m) / sd

cat("Cutoff Sample Z Score is ", cutoff)
cat("\nActual Sample Z Score is", z)

if(z > cutoff)
{
  cat("\nWe reject null hypothesis.")
}else{
  cat("\nWe cannot reject null hypothesis.")
}
```

R code Exa 4.2a Example of Hypothesis Testing with a Two Tailed Test

```
# Page no. : 123 - 124

m <- 69.5
sd <- 14.1
x <- 41
alpha <- 0.05

cutoff1 <- qnorm(0.025)
cutoff1 <- round(cutoff1, 2)

cutoff2 <- qnorm(0.975)
cutoff2 <- round(cutoff2, 2)

z <- (x - m) / sd

cat("Cutoff Z scores are", cutoff1, "and", cutoff2)
cat("\nActual Z score is", round(z, 2))

```

```
if(z > cutoff1 & z < cutoff2)
{
 cat("\nWe cannot reject null hypothesis.")
}else{
  cat("\nWe can reject null hypothesis.")
}
```

R code Exa 4.2b How are you doing

```
# Page no. : 125

m <- 500
sd <- 40
x <- 400
alpha <- 0.01

cutoff1 <- qnorm(0.005)
cutoff1 <- round(cutoff1, 2)

cutoff2 <- qnorm(0.995)
cutoff2 <- round(cutoff2, 2)

z <- (x - m) / sd

cat("Cutoff Z scores are", cutoff1, "and", cutoff2)
cat("\nActual Z score is", round(z, 2))

if(z > cutoff1 & z < cutoff2)
{
  cat("\nWe cannot reject null hypothesis.")
}else{
  cat("\nWe can reject null hypothesis..")
}
```

R code Exa 4.3a Worked Out Examples

```
# Page no. : 131 - 132

m <- 19
sd <- 4
x <- 27
alpha <- 0.05

cutoff1 <- qnorm(0.025)
cutoff1 <- round(cutoff1, 2)

cutoff2 <- qnorm(0.975)
cutoff2 <- round(cutoff2, 2)

z <- (x - m) / sd

cat("Cutoff Z scores are", cutoff1, "and", cutoff2)
cat("\nActual Z score is", round(z, 2))

if(z > cutoff1 & z < cutoff2)
{
  cat("\nWe cannot reject null hypothesis.")
}else{
  cat("\nWe can reject null hypothesis.")
}
```

Chapter 5

Hypothesis Tests with Means of Samples

R code Exa 5.1a Determining the Characteristics of a Distribution of Means

```
# Page no. : 144

grade <- 6.67
N <- 2

ans <- grade / N

cat("Answer is", round(ans,2))

pv <- 400
N <- 25

ans <- pv / N

cat("\nAnswer is", round(ans, 2))
```

R code Exa 5.1b Example of Determining the Characteristics of a Distribution of Means

```
# Page no. : 146 - 147

mean <- 200
sd <- 48
N <- 64

var <- sd ** 2
ans <- var / N

cat("Answer is",ans)

sd2 <- sqrt(ans)

cat("\nAnswer is",sd2)
```

R code Exa 5.1c How are you doing

```
# Page no. : 148

mean <- 60
sd <- 10
N <- 4

var <- sd ** 2
ans <- var / N

cat("Answer is",ans)

var2 <- sqrt(ans)

cat("\nAnswer is",var2)
```

R code Exa 5.2a Hypothesis testing with a distribution of means The Z Test

```
# Page no. : 148 - 149

mean <- 18
d_mean <- 10
sd <- 4

z <- (mean - d_mean) / sd

cat("Z score is", z)
```

R code Exa 5.2b Example 1 for Hypothesis testing with a distribution of means The Z Test

```
# Page no. : 150 - 152

n <- 64
mean <- 220
d_mean <- 200
sd <- 48
var <- sd ** 2
d_var <- var / n
d_sd <- sqrt(d_var)
alpha <- 0.05
zval <- qnorm(alpha, lower.tail = F)

z <- (mean - d_mean) / d_sd

cat("The value of z-score is", z)

```

```
if(z > zval)
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nWe cannot reject null hypothesis")
}
```

R code Exa 5.2c Example 2 for Hypothesis testing with a distribution of means The Z Test

```
# Page no. : 152 - 154

n <- 25
mean <- 48
d_mean <- 53
sd <- 7
var <- sd ** 2
d_var <- var / n
d_sd <- sqrt(d_var)
alpha <- 0.01

zval <- qnorm(1 - alpha/2)
zval <- round(c(zval, -zval),2)

z <- (mean - d_mean) / d_sd

cat("The value of z-score is", z)

if(z > zval[1] || z < zval[2])
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nWe cannot reject null hypothesis")
```

```
}
```

R code Exa 5.2d How are you doing

```
# Page no. : 154 - 155

n <- 36
mean <- 70
d_mean <- 75
sd <- 12
var <- sd ** 2
d_var <- var / n
d_sd <- sqrt(d_var)
alpha <- 0.05

zval <- qnorm(1 - alpha/2)
zval <- round(c(zval, -zval),2)

z <- (mean - d_mean) / d_sd

cat("The value of z-score is", z)

if(z > zval[1] || z < zval[2])
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nWe cannot reject null hypothesis")
}
```

R code Exa 5.3a Confidence interval

```
# Page no. : 161
```

```
c_i <- 0.995
n <- 25
d_mean <- 53
mean <- 48
sd <- 7
var <- sd ** 2

se <- sqrt(var/n)

cat("Standard error is", se)

r_s <- round(qnorm(c_i), 2)

x <- se * r_s

y <- mean + c(-x, x)

cat("\nThe margin of error is given by", x)
cat("\nThe 99% confidence interval is given by", y)
```

R code Exa 5.3b How are you doing

```
# Page no. : 163

n <- 36
mean <- 70
d_mean <- 75
sd <- 12
var <- sd ** 2
c_i <- 0.975

se <- sqrt(var/n)

cat("Standard error is", se)
```

```
r_s <- round(qnorm(c_i), 2)

x <- se * r_s

y <- mean + c(-x, x)

cat("\nThe margin of error is given by", x)
cat("\nThe 99% confidence interval is given by", y)
```

R code Exa 5.4a Worked out examples 1

```
# Page no. : 167

sd <- 13
var <- sd ** 2
n <- 20

d_sd <- sqrt(var/n)

cat("Standard deviation for distribution of means is ", round(d_sd, 2))
```

R code Exa 5.4b Worked out examples 2

```
# Page no. : 167 - 168

n <- 75
mean <- 16
d_mean <- 15
sd <- 5
var <- sd ** 2
d_var <- var / n
```

```
d_sd <- sqrt(d_var)
alpha <- 0.05

zval <- qnorm(1 - alpha/2)
zval <- round(c(zval, -zval),2)

z <- (mean - d_mean) / d_sd

cat("Z Score is", z)

if(z > zval[1] || z < zval[2])
{
  cat("\nCan reject null hypothesis")
} else
{
  cat("\nCannot reject null hypothesis")
}
```

R code Exa 5.4c Worked out examples 3

```
# Page no. : 169

n <- 75
mean <- 16
d_mean <- 15
sd <- 5
var <- sd ** 2
d_var <- var / n
d_sd <- sqrt(d_var)
c_i <- 0.995

se <- d_sd

cat("Standard error is", se)

```

```r
r_s <- round(qnorm(c_i), 2)

x <- se * r_s

y <- mean + c(-x, x)

cat("\nThe margin of error is given by", x)
cat("\nThe 99% confidence interval is given by", y)
```

Chapter 6

Making Sense of Statistical Significance

R code Exa 6.1a Figuring effect size

```
# Page no. : 184 - 185

m1 <- 220
m2 <- 200
sd <- 48

d1 <- (m1 - m2) / sd

cat("Estimated effect size is", round(d1, 2))

m1 <- 210

d2 <- (m1 - m2) / sd

cat("\nEstimated effect size is", round(d2, 2))

m1 <- 170

d3 <- (m1 - m2) / sd
```

```r

cat("\nEstimated effect size is", round(d3, 2))
```

R code Exa 6.1b How are you doing

```r
# Page no. : 188 - 189

m1 <- 540
m2 <- 500
sd <- 100

d1 <- (m1 - m2) / sd

cat("Estimated effect size is", round(d1, 2))
```

R code Exa 6.2a Determining power from predicted effect size

```r
# Page no. : 197

m2 <- 200
d <- .20
sd <- 48

m <- m2 + (d * sd)

cat("Predicted value is", round(m, 2))

d <- .50

m <- m2 + (d * sd)

cat("\nPredicted value is", round(m, 2))
```

R code Exa 6.2b How are you doing

```
# Page no. : 204

m2 <- 500
d <- .80
sd <- 100

m <- m2 + (d * sd)

cat("Predicted value is", round(m, 2))
```

R code Exa 6.3a Advanced topic Figuring Statistical Power

```
# Page no. : 214 - 215

m <- 200
sd <- 48
var <- sd ** 2
n <- 64

d_sd <- sqrt(var/n)

m2 <- 208

alpha <- 0.05
zval <- qnorm(alpha, lower.tail = F)

r_s <- round((zval * d_sd) + m, 2)

z <- (r_s - m2) / d_sd

```

```
cat("Z is", round(z, 2))
```

R code Exa 6.3b How are you doing

```
# Page no. : 216 – 217

m <- 500
sd <- 100
var <- sd ** 2
n <- 60

d_sd <- sqrt(var/n)

m2 <- 540

alpha <- 0.05
zval <- qnorm(alpha, lower.tail = F)

r_s <- round((zval * d_sd) + m, 2)

z <- (r_s - m2) / d_sd

cat("Z is", round(z, 2))
```

R code Exa 6.4a Worked out examples 1

```
# Page no. : 218

m1 <- 37
m2 <- 40
sd <- 10

d1 <- (m1 - m2) / sd
```

```
cat("Estimated effect size is", round(d1, 2))
```

R code Exa 6.4b Worked out examples 2

```
# Page no. : 218 - 219

m2 <- 40
d <- -0.20
sd <- 10

m <- m2 + (d * sd)

cat("Predicted effect size is", round(m, 2))
```

R code Exa 6.4c Worked out examples 3

```
# Page no. : 219

m <- 40
sd <- 10
var <- sd ** 2
n <- 25

d_sd <- sqrt(var/n)

m2 <- 49

alpha <- 0.01
zval <- qnorm(alpha, lower.tail = F)

r_s <- round((zval * d_sd) + m, 2)

```

```
z <- (r_s - m2) / d_sd

cat("Z is", round(z, 2))
```

Chapter 7

Introduction to t Tests

R code Exa 7.1a Basic Principle of the t Test Estimating the Population Variance from the Sample Scores

```
# Page no. : 229 - 230

n <- 16
ss <- 694

var <- (ss)/(n - 1)

cat("Estimated population variance is", round(var,
   2))
```

R code Exa 7.1b The Standard Deviation of the Distribution of Means

```
# Page no. : 231

s <- 16
var <- 46.27

```

```
d_var <- var / s

cat("Variance is", round(d_var, 2))

d_sd <- sqrt(d_var)

cat("\nStandard deviation is", round(d_sd, 2))
```

R code Exa 7.1c The Sample means score on the comparison distribution the t score

```
# Page no. : 234

m1 <- 21
m2 <- 17
sd <- 1.70

t <- (m1 - m2) / sd

cat("The t score is", round(t, 2))
```

R code Exa 7.1d Another example of a t test for a single sample

```
# Page no. : 235 - 236

rating <- c(5, 3, 6, 2, 7, 6, 7, 4, 2, 5)
n <- 10

mean <- mean(rating)

diff <- rating - mean

sq_diff <- diff ** 2
```

```
data_frame <- data.frame(rating, diff, sq_diff)

View(data_frame)

df <- n - 1

mu <- 4

var <- sum(data_frame$sq_diff) / df

d_var <- var / n

d_sd <- sqrt(d_var)

alpha <- 0.01

tval <- round(qt(1 - alpha/2, df), 2)
tval <- c(tval, -tval)

t <- (mean - mu) / d_sd

cat("\nValue of t is", t)

if(t > tval[1] || t < tval[2])
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nCannot reject null hypothesis")
}
```

R code Exa 7.1e How are you doing

```
# Page no. : 238 - 240
```

```

n <- 4
df <- n - 1

m1 <- 23

scores <- c(20, 22, 22, 20)

mean <- mean(scores)

diff <- scores - mean

sq_diff <- diff ** 2

alpha <- 0.05

var <- sum(sq_diff) / df

d_var <- var / n

d_sd <- sqrt(d_var)

tval <- round(qt(alpha, df, lower.tail = T), 2)

t <- (mean - m1) / d_sd

cat("Value of t is", t)

if(t < tval)
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nCannot reject null hypothesis")
}
```

R code Exa 7.2a Example of a t test for dependent means

```
# Page no. : 241 - 242

Husband <- LETTERS[1:19]

n <- 19

df <- n - 1

before <- c(126, 133, 126, 115, 108, 109, 124, 98,
    95, 120, 118, 126, 121, 116, 94, 105, 123, 125,
    128)
after <- c(115, 125, 96, 115, 119, 82, 93, 109, 72,
    104, 107, 118, 102, 115, 83, 87, 121, 100, 118)

diff <- after - before
mean <- mean(diff)

deviation <- round(diff - mean, 2)

sq_dev <- round(deviation ** 2, 2)

data_frame <- data.frame(Husband, before, after,
    diff, deviation, sq_dev)

View(data_frame)

mu <- 0

var <- sum(data_frame$sq_dev) / df

d_var <- var / n

```

```
d_sd <- sqrt(d_var)

alpha <- 0.05

tval <- round(qt(1 - alpha/2, df), 2)
tval <- c(tval, -tval)

t <- (mean - mu) / d_sd

cat("\nValue of t is", t)

if(t > tval[1] || t < tval[2])
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nCannot reject null hypothesis")
}
```

R code Exa 7.2b A second example of a t test for dependent means

```
# Page no. : 245 - 246

student <- c(1:10)

n <- 10

df <- n - 1

b_p <- c(1487.8, 1329.4, 1407.9, 1236.1, 1299.8,
   1447.2, 1354.1, 1204.6, 1322.3, 1388.5)
c_p <- c(1487.2, 1328.1, 1405.9, 1234.0, 1298.2,
   1444.7, 1354.3, 1203.7, 1320.8, 1386.8)

diff <- b_p - c_p
```

```
mean <- mean(diff)

deviation <- round(diff - mean, 2)

sq_dev <- round(deviation ** 2, 2)

data_frame <- data.frame(student, b_p, c_p, diff,
    deviation, sq_dev)

View(data_frame)

mu <- 0

var <- sum(data_frame$sq_dev) / df

d_var <- var / n

d_sd <- sqrt(d_var)

alpha <- 0.05

tval <- round(qt(alpha, df, lower.tail = F), 2)

t <- (mean - mu) / d_sd

cat("\nValue of t is", t)

if(t > tval)
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nCannot reject null hypothesis")
}
```

R code Exa 7.2c How are you doing

```
# Page no. : 249 - 250

person <- c(1:5)
n <- 5
df <- n - 1

before <- c(20, 30, 20, 40, 30)
after <- c(30, 50, 10, 30, 40)

diff <- after - before
mean <- mean(diff)

deviation <- round(diff - mean, 2)

sq_dev <- round(deviation ** 2, 2)

data_frame <- data.frame(person, before, after, diff
    , deviation, sq_dev)

View(data_frame)

mu <- 0

var <- sum(data_frame$sq_dev) / df

d_var <- var / n

d_sd <- sqrt(d_var)

alpha <- 0.05

tval <- round(qt(1 - alpha/2, df), 2)
tval <- c(tval, -tval)

t <- (mean - mu) / d_sd

```

```
cat("\nValue of t is", t)

if(t > tval[1] || t < tval[2])
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nCannot reject null hypothesis")
}
```

R code Exa 7.3a Effect size for the t test for dependent means

```
# Page no. : 252

mu1 <- 4
mu2 <- 0
sd <- 8

d <- (mu1 - mu2) / sd

cat("Effect size is", d)

m_d <- -12.05
var <- 153.49
sd <- round(sqrt(var), 2)

d <- round((m_d - mu2) / sd, 2)

cat("\nEffect size is", d)
```

R code Exa 7.3b How are you doing

```
# Page no. : 254
```

```
mu1 <- 40
mu2 <- 0
sd <- 80
alpha <- 0.05

d <- round((mu1 - mu2) / sd, 2)

cat("Effect size is", d)
```

R code Exa 7.4a Worked out examples 1

```
# Page no. : 259

scores <- c(14, 8, 6, 5, 13, 10, 10, 6)
n <- length(scores)
df <- n - 1
mean <- 6
alpha <- 0.05

m <- mean(scores)

diff <- scores - m

sq_diff <- diff ** 2

var <- sum(sq_diff) / df

d_var <- var / n

d_sd <- sqrt(d_var)

tval <- round(qt(1 - alpha/2, df), 2)
tval <- c(tval, -tval)

```

```
t <- (m - mean) / d_sd

cat("Value of t is", t)

if(t > tval[1] || t < tval[2])
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nCannot reject null hypothesis")
}
```

R code Exa 7.4b Worked out examples 2

```
# Page no. : 259 - 261

participant <- c(1:10)
n <- 10
df <- n - 1

before <- c(10.4, 12.6, 11.2, 10.9, 14.3, 13.2, 9.7,
    11.5, 10.8, 13.1)
after <- c(10.8, 12.1, 12.1, 11.4, 13.9, 13.5, 10.9,
    11.5, 10.4, 12.5)

diff <- after - before
mean <- mean(diff)

deviation <- round(diff - mean, 2)

sq_dev <- round(deviation ** 2, 2)

data_frame <- data.frame(participant, before, after,
    diff, deviation, sq_dev)

```

```
View(data_frame)

mu <- 0

var <- sum(data_frame$sq_dev) / df

d_var <- var / n

d_sd <- sqrt(d_var)

alpha <- 0.05

tval <- round(qt(1 - alpha/2, df), 2)
tval <- c(tval, -tval)

t <- (mean - mu) / d_sd

cat("\nValue of t is", t)

if(t > tval[1] || t < tval[2])
{
  cat("\nReject null hypothesis")
} else
{
  cat("\nCannot reject null hypothesis")
}
```

Chapter 8

The t Test for Independent Means

R code Exa 8.1a Estimating the population variance

```
                                    # Page no. : 278
                                        - 279

# Estimating the population variance

n1 <- 11   # Group 1 participants
df1 <- n1 - 1   # Degree of freedom for group 1
var1 <- 60   # Variance

n2 <- 31   # Group 2 participants
df2 <- n2 - 1   # Degree of freedom for group 2
var2 <- 80   # Variance

df <- df1 + df2   # Total degrees of freedom

p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)
    # Pooled variance

cat("Pooled variance is", p_var)
```

R code Exa 8.1b Figuring the variance of each of the two distributions of means

```
                                    # Page no. : 279 -
                                        280

# Figuring the variance of each of the two
    distributions of means

n1 <- 11   # Group 1 participants
n2 <- 31   # Group 2 participants
p_var <- 75   # Pooled variance

var1 <- p_var / n1   # Variance for group 1

cat("Variance for group 1 is", round(var1, 2))

var2 <- p_var / n2   # Variance for group 1

cat("Variance for group 2 is", round(var2, 2))
```

R code Exa 8.1c The variance and standard deviation of the distribution of differences between means

```
                                          # Page no. :
                                              280

# The variance and standard deviation of the
    distribution of differences between means

var1 <- 6.82   # Variance of the distribution of
    means of group 1
```

```
var2 <- 2.42   # Variance of the distribution of means of group 2

diff_var <- var1 + var2   # Variance difference

cat("Variance difference is", diff_var)

diff_sd <- sqrt(diff_var)   # Standard deviation difference

cat("Standard deviation difference is", round(diff_sd, 2))
```

R code Exa 8.1d The t score for the difference between the two actual means

```
                                    # Page no. : 281 - 282

# The t score for the difference between the two actual means

m1 <- 198   # Mean of the first sample
m2 <- 190   # Mean of the second sample

diff_sd <- 3.04   # Standard deviation difference

t <- (m1 - m2) / diff_sd   # t score

cat("t score is", round(t, 2))
```

R code Exa 8.1e How are you doing

```
# Page no. : 282 - 283

# How are you doing?

# 5th Question (a) part

n1 <- 21   # Sample 1 participants
df1 <- n1 - 1   # Degree of  freedom 1
var1 <- 100   # Variance

n2 <- 31   # Sample 1 participants
df2 <- n2 - 1   # Degree of  freedom 1
var2 <- 200   # Variance

df <- df1 + df2   # Total degrees of freedom

p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)   # Pooled variance

cat("Pooled variance is", p_var)

d_var1 <- p_var / n1   # Variance of the distribution of means of sample 1

cat("Variance of the distribution of means of sample 1 is", round(d_var1, 2))

d_var2 <- p_var / n2   # Variance of the distribution of means of sample 2

cat("Variance of the distribution of means of sample 2 is", round(d_var2, 2))

diff_var <- d_var1 + d_var2   # Variance difference

cat("Variance difference is", round(diff_var, 2))

```

```
diff_sd <- sqrt(diff_var)   # Standard deviation
    difference

cat("Standard deviation difference is", round(diff_
    sd, 2))
```

R code Exa 8.2a Example of a T test for independent means

```
#                                     # Page no. :
    284
#
# # Example of a T test for independent means

# Expressive Writing Group

score1 <- c(77, 88, 77, 90, 68, 74, 62, 93, 82, 79)
n1 <- length(score1)   # Sample size 1
df1 <- n1 - 1   # Degree of freedom
m1 <- mean(score1)   # Mean
deviation1 <- score1 - m1
sq_dev1 <- deviation1 ** 2   # Squared deviation

data_frame1 <- data.frame(score1, deviation1, sq_
    dev1)
View(data_frame1)

# Control Writing Group

score2 <- c(87, 77, 71, 70, 63, 50, 58, 63, 76, 65)
n2 <- length(score2)   # Sample size 2
df2 <- n2 - 1   # Degree of freedom
m2 <- mean(score2)   # Mean
deviation2 <- score2 - m2
sq_dev2 <- deviation2 ** 2   # Squared deviation

```

```
data_frame2 <- data.frame(score2, deviation2, sq_dev2)
View(data_frame2)

var1 <- sum(data_frame1$sq_dev1) / df1   # Variance 1
var2 <- sum(data_frame2$sq_dev2) / df2   # Variance 2

df <- df1 + df2   # Total degrees of freedom

p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)   # Pooled variance

cat("Pooled variance is", round(p_var, 2))

d_var1 <- p_var / n1   # Variance of the distribution of means of sample 1

cat("# Variance of the distribution of means of sample 1 is", round(d_var1, 2))

d_var2 <- p_var / n2   # Variance of the distribution of means of sample 2

cat("# Variance of the distribution of means of sample 2 is", round(d_var2, 2))

diff_var <- d_var1 + d_var2   # Variance difference

cat("Variance difference is", diff_var)

diff_sd <- sqrt(diff_var)   # Standard deviation difference

cat("Standard deviation difference is", round(diff_sd, 2))

```

```
# Two - Tailed

alpha <- 0.05   # 5% significance level

tval <- qt(1-alpha/2, df)
tval <- c(tval, -tval)

t <- round((m1 - m2) / diff_sd, 2)   # t score

cat("t score is", t)

if(t > tval[1] || t < tval[2])
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 8.2b A second example of a T test for independent means

```
                                   # Page no. : 286
                                       - 287

# A second example of a T test for independent means

# Experimental Group (Receiving Special Program)

score1 <- c(6, 4, 9, 7, 7, 3)
n1 <- length(score1)   # Sample size 1
df1 <- n1 - 1   # Degree of freedom
m1 <- mean(score1)   # Mean
deviation1 <- score1 - m1
sq_dev1 <- deviation1 ** 2   # Squared deviation

```

```
data_frame1 <- data.frame(score1, deviation1, sq_dev1)
View(data_frame1)

# Control Group (Receiving Ordinary Program)

score2 <- c(6, 1, 5, 3, 1, 1, 4, 3)
n2 <- length(score2)   # Sample size 2
df2 <- n2 - 1   # Degree of freedom
m2 <- mean(score2)   # Mean
deviation2 <- score2 - m2
sq_dev2 <- deviation2 ** 2   # Squared deviation

data_frame2 <- data.frame(score2, deviation2, sq_dev2)
View(data_frame2)

var1 <- sum(data_frame1$sq_dev1) / df1   # Variance 1
var2 <- sum(data_frame2$sq_dev2) / df2   # Variance 2

df <- df1 + df2   # Total degrees of freedom

p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)   # Pooled variance

cat("Pooled variance is", round(p_var, 2))

d_var1 <- p_var / n1   # Variance of the distribution of means of sample 1

cat("# Variance of the distribution of means of sample 1 is", round(d_var1, 2))

d_var2 <- p_var / n2   # Variance of the distribution of means of sample 2

```

```
cat("# Variance of the distribution of means of sample 2 is", round(d_var2, 2))

diff_var <- d_var1 + d_var2   # Variance difference

cat("Variance difference is", diff_var)

diff_sd <- sqrt(diff_var)   # Standard deviation difference

cat("Standard deviation difference is", round(diff_sd, 2))

# Two - Tailed

alpha <- 0.05   # 5% significance level

tval <- qt(1-alpha/2, df)
tval <- c(tval, -tval)

t <- round((m1 - m2) / diff_sd, 2)   # t score

cat("t score is", t)

if(t > tval[1] || t < tval[2])
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 8.2c How are you doing

```r
# Page no. : 289 - 291

# How are you doing?

# 2nd Question (a) part

alpha <- 0.05   # 5% significance level

n1 <- 26   # Sample size 1
df1 <- n1 - 1   # Degree of freedom
m1 <- 5   # Mean 1
var1 <- 10   # Variance 1

n2 <- 36   # Sample size 2
df2 <- n2 - 1   # Degree of freedom
m2 <- 8   # Mean 2
var2 <- 12   # Variance 2

df <- df1 + df2   # Total degrees of freedom

p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)   # Pooled variance

cat("Pooled variance is \n", round(p_var, 2))

d_var1 <- p_var / n1   # Variance of the distribution of means of sample 1

cat("Variance of the distribution of means of sample 1 is \n", round(d_var1, 2))

d_var2 <- p_var / n2   # Variance of the distribution of means of sample 2

cat("Variance of the distribution of means of sample 2 is \n", round(d_var2, 2))

```

```r
diff_var <- d_var1 + d_var2   # Variance difference

cat("Variance difference is \n", round(diff_var, 2))

diff_sd <- sqrt(diff_var)   # Standard deviation difference

cat("Standard deviation difference is \n", round(diff_sd, 2))

# One − Tailed

tval <- qt(alpha, df)

t <- round((m1 - m2) / diff_sd, 2)   # t score

cat("t score is \n", t)

if(t < tval)
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 8.3a Effect size for the t test for independent means

```r
                                                  # Page no. : 293 − 294

# Effect size for the t test for independent means

mu1 <- 29   # Mean 1
mu2 <- 21   # Mean 2
```

```
sd <- 10   # Standard deviation

d <- (mu1 - mu2) / sd   # Effect size

cat("Effect size is \n", d)

m1 <- 33.10   # Mean 1
m2 <- 27.00   # Mean 2
p_sd <- 12.99   # Pooled standard deviation

e_d <- round((m1 - m2) / p_sd, 2)   # Estimated effect size

cat("Estimated effect size is \n", e_d)
```

R code Exa 8.4a Harmonic mean

```
                                    # Page no. : 300

# Harmonic mean

n1 <- 11   # Sample 1
n2 <- 31   # Sample 2

total <- n1 + n2

h_m <- round((2 * n1 * n2) / total, 2)   # Harmonic mean

cat("Harmonic mean is \n", h_m)

# How are you doing?

# 1st Question (Only harmonic mean)

```

```
n1 <- 6   # Sample 1
n2 <- 34   # Sample 2

total <- n1 + n2

h_m <- round((2 * n1 * n2) / total, 2)   # Harmonic mean

cat("Harmonic mean is \n", h_m)
```

R code Exa 8.5a Worked out examples 1

```
                                        # Page no. : 301 - 302

# Worked out examples 1

n1 <- 40   # Sample 1
df1 <- n1 - 1   # Degree of freedom
var1 <- 15   # Variance 1

n2 <- 60   # Sample 2
df2 <- n2 - 1   # Degree of freedom
var2 <- 12   # Variance 2

df <- df1 + df2   # Total degrees of freedom

p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)   # Pooled variance

cat("Pooled variance is \n", round(p_var, 2))

d_var1 <- p_var / n1   # Variance of the distribution of means of sample 1

```

```
cat("Variance of the distribution of means of sample 1 is \n", round(d_var1, 2))

d_var2 <- p_var / n2  # Variance of the distribution of means of sample 2

cat("Variance of the distribution of means of sample 2 is \n", round(d_var2, 2))

diff_var <- d_var1 + d_var2  # Variance difference

cat("Variance difference is \n", round(diff_var, 2))

diff_sd <- sqrt(diff_var)  # Standard deviation difference

cat("Standard deviation difference is \n", round(diff_sd, 2))
```

R code Exa 8.5b Worked out examples 2

```
                                        # Page no. : 302 - 303

# Worked out examples 2

score1 <- c(7, 6, 9, 7, 6)
n1 <- length(score1)  # Sample size 1
df1 <- n1 - 1  # Degree of freedom
m1 <- mean(score1)  # Mean
deviation1 <- score1 - m1
sq_dev1 <- deviation1 ** 2  # Squared deviation

data_frame1 <- data.frame(score1, deviation1, sq_dev1)
```

```
View(data_frame1)

score2 <- c(5, 2, 4, 3, 6)
n2 <- length(score2)   # Sample size 2
df2 <- n2 - 1   # Degree of freedom
m2 <- mean(score2)   # Mean
deviation2 <- score2 - m2
sq_dev2 <- deviation2 ** 2   # Squared deviation

data_frame2 <- data.frame(score2, deviation2, sq_dev2)
View(data_frame2)

var1 <- sum(data_frame1$sq_dev1) / df1   # Variance 1
var2 <- sum(data_frame2$sq_dev2) / df2   # Variance 2

df <- df1 + df2   # Total degrees of freedom

p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)   # Pooled variance

cat("Pooled variance is \n", round(p_var, 2))

d_var1 <- p_var / n1   # Variance of the distribution of means of sample 1

cat("Variance of the distribution of means of sample 1 is \n", round(d_var1, 2))

d_var2 <- p_var / n2   # Variance of the distribution of means of sample 2

cat("Variance of the distribution of means of sample 2 is \n", round(d_var2, 2))

diff_var <- d_var1 + d_var2   # Variance difference
```

```

cat("Variance difference is \n", diff_var)

diff_sd <- sqrt(diff_var)   # Standard deviation difference

cat("Standard deviation difference is \n", round(diff_sd, 2))

# Two - Tailed

alpha <- 0.05   # 5% significance level

tval <- qt(1-alpha/2, df)
tval <- c(tval, -tval)

t <- round((m1 - m2) / diff_sd, 2)   # t score

cat("t score is \n", t)

if(t > tval[1] || t < tval[2])
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 8.5c Worked out examples 3

```
                                                    # Page no. : 304

# Worked out examples 3

```

```r
n1 <- 22    # Sample 1
n2 <- 51    # Sample 2

total <- n1 + n2

h_m <- round((2 * n1 * n2) / total, 2)    # Harmonic mean

cat("Harmonic mean is", h_m)
```

Chapter 9

Introduction to the Analysis of Variance

R code Exa 9.1a Carrying out an Analysis of Variance

```
                                    # Page no. : 325
                                       - 331

# Carrying out an Analysis of Variance

# Page no. : 325 - 326

# Criminal Record Group

rating1 <- c(10, 7, 5, 10, 8)
n1 <- length(rating1)   # Sample 1 size
df1 <- n1 - 1   # Degree of freedom
m1 <- mean(rating1)   # Mean of rating 1
deviation1 <- rating1 - m1   # Deviation from mean 1
sq_dev1 <- deviation1 ** 2   # Squared deviation
   from mean 1

data_frame1 <- data.frame(rating1, deviation1, sq_
   dev1)
```

```r
View(data_frame1)

# Clean Record Group

rating2 <- c(5, 1, 3, 7, 4)
n2 <- length(rating2)   # Sample 2 size
df2 <- n2 - 1   # Degree of freedom
m2 <- mean(rating2)   # Mean of rating 2
deviation2 <- rating2 - m2   # Deviation from mean 2
sq_dev2 <- deviation2 ** 2   # Squared deviation from mean 2

data_frame2 <- data.frame(rating2, deviation2, sq_dev2)
View(data_frame2)

# No Information Group

rating3 <- c(4, 6, 9, 3, 3)
n3 <- length(rating1)   # Sample 3 size
df3 <- n3 - 1   # Degree of freedom
m3 <- mean(rating3)   # Mean of rating 3
deviation3 <- rating3 - m3   # Deviation from mean 3
sq_dev3 <- deviation3 ** 2   # Squared deviation from mean 3

data_frame3 <- data.frame(rating3, deviation3, sq_dev3)
View(data_frame3)

var1 <- var(data_frame1$rating1)   # Variance for rating 1
var2 <- var(data_frame2$rating2)   # Variance for rating 2
var3 <- var(data_frame3$rating3)   # Variance for rating 3

cat("Variance for rating 1 is", var1)
```

```
cat("Variance for rating 2 is", var2)
cat("Variance for rating 3 is", var3)

# Figuring the within - groups estimate of the population variance

# Page no. : 326 - 327

n_g <- 3   # Number of groups
df_within <- (n1 + n2 + n3) - n_g  # Degree of freedom within-groups

var_within <- round((var1 + var2 + var3) / n_g, 2)   # Variance within-groups

cat("Variance within-groups is", var_within)

# Figuring the between - groups estimate of the population variance

# Page no. : 327 - 329

df_between <- n_g - 1   # Degree of freedom between-groups
gm <- round((m1 + m2 + m3) / n_g, 2)   # Grand mean

gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_var)

N <- 5   # 5 elements in each rating
```

```
var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)

# Figuring the F ratio

# Page no. : 329

fvalue <- round(var_between / var_within, 2)   # F ratio
cat("F ratio is", fvalue)

# Direct method to find F ratio

DF <- data.frame(rating1, rating2, rating3)
x <- c(t(as.matrix(DF)))
f <- c("r_1", "r_2", "r_3")
tm <- gl(3, 1, 5*3, factor(f))
result <- anova(lm(x ~ tm))   # Similar to aov(x ~ tm)

cat("F ratio is", round(result$`F value`[1], 2))

# F value

# Page no. : 331

alpha <- 0.05   # 5% Sgnificance level

fval <- qf(0.95,df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
fval <- round(fval, 2)

if(fvalue >= fval)
{
  cat("Reject Null Hypothesis")
} else {
  cat("Cannot reject Null Hypothesis")
```

```
}
```

R code Exa 9.1b How are you doing

```
# Page no. : 331 - 333

# How are you doing?

# Data given

scores_A <- c(5, 7)
m1 <- 6   # Mean 1
n1 <- 2   # Sample size

scores_B <- c(6, 9)
m2 <- 7.5   # Mean 2
n2 <- 2   # Sample size

scores_c <- c(8, 9)
m3 <- 8.5   # Mean 3
n3 <- 2   # Sample size

# 1st Question (c) part

n_s <- 3   # Number of samples
df_within <- (n1 + n2 + n3) - n_s  # Degree of freedom within-groups

var_within <- round((var(scores_A) + var(scores_B) + var(scores_c)) / n_s, 2)
# Variance within-groups


cat("Variance within-groups is", var_within)

```

```r
# 2nd and 3rd Question (c) parts

df_between <- n_s - 1   # Degree of freedom between-groups
gm <- round((m1 + m2 + m3) / n_s, 2)   # Grand mean

gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_var)

N <- 2   # 2 elements in each rating
var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)

# 4th Question (c) part

fvalue <- round(var_between / var_within, 2)   # F ratio
cat("F ratio is", fvalue)

# 5th and 6th Question (c) parts

cat("Between-groups degree of freedom is", df_between)
cat("Within-groups degree of freedom is", df_within)

alpha <- 0.05   # 5% Sgnificance level

fval <- qf(0.95,df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
```

```
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is",
   fval)
```

R code Exa 9.1c Another example

```
                                          # Page no. :
                                            334 - 335

# Another example

# Attachment Styl - Secure

n1 <- 10   # Sample size
df1 <- n1 - 1   # Degree of freedom
m1 <- 2.10   # Mean
sd1 <- 1.66   # Standard deviation
var1 <- 2.76   # Variance

# Attachment Styl - Avoidant

n2 <- 10   # Sample size
df2 <- n2 - 1   # Degree of freedom
m2 <- 3.70   # Mean
sd2 <- 1.89   # Standard deviation
var2 <- 3.57   # Variance

# Attachment Styl - Anxious-Ambivalent

n3 <- 10   # Sample size
df3 <- n3 - 1   # Degree of freedom
m3 <- 4.20   # Mean
sd3 <- 1.93   # Standard deviation
var3 <- 3.72   # Variance
```

```
n_s <- 3   # Number of samples
df_within <- (n1 + n2 + n3) - n_s  # Degree of freedom within-groups
df_between <- n_s - 1   # Degree of freedom between-groups

alpha <- 0.05   # 5% Sgnificance level

fval <- qf(0.95,df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is", fval)

var_within <- round((var1 + var2 + var3) / n_s, 2)   # Variance within-groups

cat("Variance within-groups is", var_within)

gm <- round((m1 + m2 + m3) / n_s, 2)   # Grand mean

gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_var)

N <- 10   # 10 elements in each rating
var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)
```

```

fvalue <- round(var_between / var_within, 2)    # F ratio
cat("F ratio is", fvalue)

if(fvalue >= fval)
{
  cat("Reject Null Hypothesis")
} else {
  cat("Cannot reject Null Hypothesis")
}
```

R code Exa 9.1d How are you doing

```
                                                    # Page no. : 338 - 339

# How are you doing?

# 1st Question (a) part

N <- 16    # Participants in each group
df <- N - 1    # Degree of freedom

# Group 1

m1 <- 20    # Mean
var1 <- 8    # Variance

# Group 2

m2 <- 22    # Mean
var2 <- 9    # Variance

# Group 3
```

```r

m3 <- 18   # Mean
var3 <- 7   # Variance

alpha <- 0.01   # 1% Sgnificance level

n_s <- 3   # Number of samples
df_within <- (3 * N) - n_s  # Degree of freedom within-groups
df_between <- n_s - 1   # Degree of freedom between-groups


fval <- qf(0.99,df_between, df_within)   # 1 - alpha = 1 - 0.01 = 0.99
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is", fval)

var_within <- round((var1 + var2 + var3) / n_s, 2)   # Variance within-groups

cat("Variance within-groups is", var_within)

gm <- round((m1 + m2 + m3) / n_s, 2)   # Grand mean

gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_var)

```

```r
var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)

fvalue <- round(var_between / var_within, 2)   # F ratio
cat("F ratio is", fvalue)

if(fvalue >= fval)
{
  cat("Reject Null Hypothesis")
} else {
  cat("Cannot reject Null Hypothesis")
}
```

R code Exa 9.2a Planned contrast an example

```r
                                                # Page no. : 340 - 341

# Planned contrast an example

m1 <- 8   # Mean 1
m2 <- 5   # Mean 2
n_s <- 2   # Number of samples for contrast
N <- 5   # 5 elements in each sample

var_within <- 5.33   # Variance within-groups

gm <- (m1 + m2) / n_s   # Grand mean
gm_dev <- c((m1 - gm), (m2 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean
df_between <- n_s - 1   # Degree of freedom between-
```

```
groups (2 samples)
df_within <- (3 * N) - (n_s + 1)  # Degree of freedom within-groups (3 samples)

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_var)

var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)

fvalue <- round(var_between / var_within, 2)   # F ratio
cat("F ratio is", fvalue)

alpha <- 0.05   # 5% Sgnificance level

fval <- qf(0.95,df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is", fval)

if(fvalue >= fval)
{
  cat("Reject null hypothesis (Significant)")
} else {
  cat("Cannot reject null hypothesis (not significant)")
}
```

R code Exa 9.2b Planned contrast a second example

```
                                                    # Page no. : 341

# Planned contrast a second example

m1 <- 8   # Mean 1
m2 <- 4   # Mean 2
n_s <- 2   # Number of samples for contrast
N <- 5   # 5 elements in each sample

var_within <- 5.33   # Variance within-groups

gm <- (m1 + m2) / n_s   # Grand mean
gm_dev <- c((m1 - gm), (m2 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean
df_between <- n_s - 1   # Degree of freedom between-groups (2 samples)
df_within <- (3 * N) - (n_s + 1)   # Degree of freedom within-groups (3 samples)

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_var)

var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)

fvalue <- round(var_between / var_within, 2)   # F ratio
cat("F ratio is", fvalue)

```

```
alpha <- 0.05   # 5% Sgnificance level

fval <- qf(0.95,df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is", fval)

if(fvalue >= fval)
{
  cat("Reject null hypothesis (Significant)")
} else {
  cat("Cannot reject null hypothesis (not significant)")
}
```

R code Exa 9.2c How are you doing

```
                                    # Page no. : 342 - 343

# How are you doing?

# 3rd Question

N <- 25   # 25 elements in each sample
var_within <- 100   # Variance within-groups

m1 <- 10   # Mean 1
m2 <- 16   # Mean 2

n_s <- 2   # Number of samples for contrast

gm <- (m1 + m2) / n_s   # Grand mean
```

```
gm_dev <- c((m1 - gm), (m2 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean
df_between <- n_s - 1   # Degree of freedom between-groups (2 samples)
df_within <- (3 * N) - (n_s + 1)  # Degree of freedom within-groups (3 samples)

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_var)

var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)

fvalue <- round(var_between / var_within, 2)   # F ratio
cat("F ratio is", fvalue)

alpha <- 0.05   # 5% Sgnificance level

fval <- qf(0.95,df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is", fval)

if(fvalue >= fval)
{
  cat("Reject null hypothesis (Significant)")
} else {
  cat("Cannot reject null hypothesis (not significant)")
```

```
}
```

R code Exa 9.3a Scheffe Test

```
                                              # Page no. :
                                                 344 − 345

# Scheffe Test

N <- 5   # 5 elements in each sample
f <- 4.22   # F value (overall)

df_between <- 2   # Degree of freedom between−groups
df_within <- (3 * N) - (df_between + 1)   # Degree
    of freedom within−groups (3 samples)
fvalue <- f / df_between   # F value (conrast)

cat("F value (contrast) is", fvalue)

alpha <- 0.05   # 5% Sgnificance level

fval <- qf(0.95,df_between, df_within)   # 1 − alpha
     = 1 − 0.05 = 0.95
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is",
    fval)

if(fvalue >= fval)
{
  cat("Reject null hypothesis (Significant)")
} else {
  cat("Cannot reject null hypothesis (not
      significant)")
}
```

```
# How are you doing?

# 5th Question

N <- 50   # 50 elements in each sample
f <- 12.60   # F value (overall)

df_between <- 3   # Degree of freedom between-groups
df_within <- (4 * N) - (df_between + 1)   # Degree
    of freedom within-groups (4 samples)
fvalue <- f / df_between   # F value (conrast)

cat("F value (contrast) is", fvalue)

alpha <- 0.05   # 5% Sgnificance level

fval <- qf(0.95,df_between, df_within)   # 1 - alpha
     = 1 - 0.05 = 0.95
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is",
    fval)

if(fvalue >= fval)
{
  cat("Reject null hypothesis (Significant)")
} else {
  cat("Cannot reject null hypothesis (not
      significant)")
}
```

R code Exa 9.4a Effect size for the analysis of variance

```
                                    # Page no. : 345
                                        - 346
```

```

# Effect size for the analysis of variance

var_within <- 5.33   # Variance within-groups
var_between <- 21.70   # Variance within-groups

df_between <- 2   # Degree of freedom between-groups
df_within <- 12  # Degree of freedom within-groups

r_sq <- round((var_between * df_between)/((var_between * df_between)+(var_within * df_within)),2)
                                                          # Proportion of variance accounted for

cat("Proportion of variance accounted for is", r_sq)

# Another approach

f <- 4.07   # F value (Ratio)

r_sq <- round((f * df_between)/((f * df_between) + (df_within)),2)
                                              # Proportion of variance accounted for
```

```
cat("Proportion of variance accounted for is", r_sq)
```

R code Exa 9.4b How are you doing

```
                                                    # Page no. :
                                                       348

# How are you doing?

# 2nd Question (d) part

var_within <- 7.20   # Variance within-groups
var_between <- 12.22   # Variance within-groups

df_between <- 2   # Degree of freedom between-groups
df_within <- 8  # Degree of freedom within-groups

r_sq <- round((var_between * df_between)/((var_
   between * df_between)+(var_within * df_within))
   ,2)
                                                          #
                                                             Proportion
                                                              of
                                                             variance

                                                             accounted
                                                              for

cat("Proportion of variance accounted for is", r_sq)

# 3rd Question (c) part

N <- 18   # 18 participants in each group
t_g <- 3   # Total groups
```

```
f <- 4.50   # F value (Ratio)

df_between <- t_g - 1   # Degree of freedom between-groups
df_within <- (t_g * N) - t_g  # Degree of freedom within-groups

r_sq <- round((f * df_between)/((f * df_between) + (df_within)),2)
                                        # Proportion of variance accounted for

cat("Proportion of variance accounted for is", r_sq)
```

R code Exa 9.5a Advanced Topic The Structural Model in the Analysis of Variance an example

```
                                    # Page no. : 351 - 354

# Advanced Topic: The Structural Model in the Analysis of Variance an example

# Criminal Record Group

rating1 <- c(10, 7, 5, 10, 8)

# Clean Record Group

rating2 <- c(5, 1, 3, 7, 4)

# No Information Group
```

```r
rating3 <- c(4, 6, 9, 3, 3)
DF <- data.frame(rating1, rating2, rating3)
# Direct Method
k <- ncol(DF)   # Number of Treatments
n <- nrow(DF)   # Number of Observations for each Treatment
N <- n * k   # Total Observations
x <- c(t(as.matrix(DF)))
f <- c("r_1", "r_2", "r_3")
tm <- gl(k, 1, n*k, factor(f))
result <- anova(lm(x ~ tm))   # Similar to aov(x ~ tm)
result
ss_total <- round(sum(result$`Sum Sq`), 2)   # Total sum of squared deviation
cat("Total sum of squared deviation is", ss_total)
ss_within <- round(result$`Sum Sq`[2], 2)   # Sum of squared deviation within-groups
cat("Sum of squared deviation within-groups is", ss_within)
ss_between <- round(result$`Sum Sq`[1], 2)   # Sum of squared deviation between-groups
cat("Sum of squared deviation between-groups is", ss_between)
df_total <- sum(result$Df)   # Total degrees of
```

```
freedom

cat("Total degrees of freedom is", df_total)

df_within <- result$Df[2]   # Degree of freedom within-groups

cat("Degree of freedom within-groups is", df_within)

df_between <- result$Df[1]   # Degree of freedom between-groups

cat("Degree of freedom between-groups is", df_between)

var_within <- round(result$`Mean Sq`[2], 2)   # Variance within-groups

cat("Variance within-groups is", var_within)

var_between <- round(result$`Mean Sq`[1], 2)   # Variance between-groups

cat("Variance between-groups is", var_between)

f <- round(result$`F value`[1], 2)   # F value (Ratio)

cat("F value (Ratio) is", f)
```

R code Exa 9.6a Worked out examples 1

```
                                   # Page no. : 359 - 360

```

```
# Worked out examples 1

N <- 20   # 20 participants in each group
n_g <- 4   # Numbr of groups

m1 <- 15   # Mean 1
var1 <- 20   # Variance 1

m2 <- 12   # Mean 2
var2 <- 25   # Variance 2

m3 <- 18   # Mean 3
var3 <- 14   # Variance 3

m4 <- 15   # Mean 4
var4 <- 27   # Variance 4

alpha <- 0.05   # 5% significant level

df_between <- n_g - 1   # Degree of freedom between-groups
df_within <- (4 * N) - (n_g)  # Degree of freedom within-groups

var_within <- round((var1 + var2 + var3 + var4) / n_g, 2)   # Variance within-groups
cat("Variance within-groups is", var_within)

gm <- (m1 + m2 + m3 + m4) / n_g   # Grand mean
gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm), (m4 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_
```

```
var)

var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)

fvalue <- round(var_between / var_within, 2)   # F ratio
cat("F ratio is", fvalue)

fval <- qf(0.95,df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
fval <- round(fval, 2)

if(fvalue >= fval)
{
  cat("Reject Null Hypothesis")
} else {
  cat("Cannot reject Null Hypothesis")
}
```

R code Exa 9.6b Worked out examples 2

```
                                          # Page no. : 360 - 361

# Worked out examples 2

m1 <- 12   # Mean 1
var1 <- 25   # Variance 1

m2 <- 18   # Mean 2
var2 <- 14   # Variance 2

alpha <- 0.01   # 1% significant level
```

```r
n_s <- 2   # Number of samples for contrast
N <- 20   # 20 elements in each sample

var_within <- 21.5   # Variance within-groups (overall)

df_between <- n_s - 1    # Degree of freedom between-groups (2 samples)
df_within <- (4 * N) - (n_s + 2)  # Degree of freedom within-groups (4 samples)

gm <- (m1 + m2) / n_s   # Grand mean
gm_dev <- c((m1 - gm), (m2 - gm))   # Deviation from the grand mean
sq_gm_dev <- round(gm_dev ** 2, 2)   # Squared deviation from the grand mean

d_var <- round(sum(sq_gm_dev) / df_between, 2)   # Variance for the distribution of means

cat("Variance for the distribution of means is", d_var)

var_between <- d_var * N   # Variance between-groups

cat("Variance between-groups is", var_between)

fvalue <- round(var_between / var_within, 2)   # F ratio
cat("F ratio is", fvalue)

fval <- qf(0.95,df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is", fval)

```

```
if(fvalue >= fval)
{
  cat("Reject null hypothesis (Significant)")
} else {
  cat("Cannot reject null hypothesis (not
     significant)")
}
```

R code Exa 9.6c Worked out examples 3

```
                                        # Page no. :
                                            361

# Worked out examples 3

n <- 6   # Six planned contrasts
alpha <- 0.05   # 5% significance level

b_c_s <- round(alpha / n, 4)   # Bonferroni
   corrected signficance level

cat("Bonferroni corrected signficance level is", b_c
   _s)

                                       # Page no. :
                                          342

# The Bonferroni Procedure

n <- 2   # Two planned contrasts
alpha <- 0.05   # 5% significance level

b_c_s <- round(alpha / n, 4)   # Bonferroni
   corrected signficance level

```

```
cat("Bonferroni corrected signficance level is", b_c_s)

n <- 3   # Two planned contrasts
alpha <- 0.05   # 5% significance level

b_c_s <- round(alpha / n, 4)   # Bonferroni corrected signficance level

cat("Bonferroni corrected signficance level is", b_c_s)

n <- 3   # Two planned contrasts
alpha <- 0.01   # 1% significance level

b_c_s <- round(alpha / n, 4)  # Bonferroni corrected signficance level

cat("Bonferroni corrected signficance level is", b_c_s)
```

R code Exa 9.6d Worked out examples 4

```
                                        # Page no. : 361

# Worked out examples 4

N <- 10   # 10 elements in each sample
f <- 11.21   # F value (overall)
n_g <- 5   # Number of groups

df_between <- n_g - 1   # Degree of freedom between-groups
df_within <- (5 * N) - (n_g)   # Degree of freedom
```

```
within-groups (5 samples)
fvalue <- f / df_between   # F value (conrast)

cat("F value (contrast) is", fvalue)

alpha <- 0.01   # 1% Sgnificance level

fval <- qf(0.99,df_between, df_within)   # 1 - alpha = 1 - 0.01 = 0.99
fval <- round(fval, 2)

cat("Cutoff F for the 0.01 significance level is", fval)

if(fvalue >= fval)
{
  cat("Reject null hypothesis (Significant)")
} else {
  cat("Cannot reject null hypothesis (not significant)")
}
```

R code Exa 9.6e Worked out examples 5

```
                                        # Page no. : 361

# Worked out examples 5

N <- 20   # 20 participants in each group
n_g <- 4   # Number of groups

df_between <- n_g - 1   # Degree of freedom between-groups
df_within <- (4 * N) - (n_g)  # Degree of freedom
```

```
within-groups

var_between <- 120   # Variance between-groups
var_within <- 21.5   # Variance within-groups

r_sq <- round((var_between * df_between)/((var_between * df_between)+(var_within * df_within)),2)
                                                    # Propo
                                                      of
                                                      varia
                                                      accou
                                                      for

cat("Proportion of variance accounted for is", r_sq)
```

R code Exa 9.6f Worked out examples 6

```
                                        # Page no. :
                                          362

# Worked out examples 6

# Treatment A

x1 <- c(8, 13, 10, 9)

# Treatment B
```

```r

x2 <- c(7, 3, 8)

# Treatment C

x3 <- c(6, 4, 2)

DF <- data.frame( Y=c(x1, x2, x3), Treatment =factor(rep(c("x1", "x2", "x3"),
             times=c(length(x1), length(x2), length(x3)))))

View(DF)

res <- aov(Y~Treatment, data=DF)
result <- anova(res)
result

df_total <- sum(result$Df)   # Total degree of freeodom

cat("Total degrees of freedom is", df_total)

df_within <- result$Df[2]   # Degree of freedom within-groups

cat("Degree of freedom within-groups is", df_within)

df_between <- result$Df[1]   # Degree of freedom between-groups

cat("Degree of freedom between-groups is", df_between)

alpha <- 0.05   # 5% significance level

f <- round(result$`F value`[1], 2)   # F value (Ratio)
```

```
cat("F value (Ratio) is", f)

ss_total <- round(sum(result$`Sum Sq`), 2)   # Total sum of squared deviation

cat("Total sum of squared deviation is", ss_total)

ss_within <- round(result$`Sum Sq`[2], 2)   # Sum of squared deviation within-groups

cat("Sum of squared deviation within-groups is", ss_within)

ss_between <- round(result$`Sum Sq`[1], 2)   # Sum of squared deviation between-groups

cat("Sum of squared deviation between-groups is", ss_between)

fval <- qf(0.95, df_between, df_within)   # 1 - alpha = 1 - 0.05 = 0.95
fval <- round(fval, 2)

cat("Cutoff F for the 0.05 significance level is", fval)

if(f >= fval)
{
  cat("Reject null hypothesis")
} else {
  cat("Cannot reject null hypothesis")
}
```

Chapter 10

Factorial Analysis of Variance

R code Exa 10.1a Table for a Two way Analysis of Variance

```
                                                # Page no. : 406 - 407

# Table for a Two-way Analysis of Variance

DF <- data.frame(Sensitivity = c(rep("Not High", 10), rep("High", 10)),
                 Test_Difficulty = c(rep(c("Easy", "Hard"), 10)),
                 Value = c(2.63, 2.69, 2.53, 2.31, 2.25, 2.45, 2.22, 2.80, 2.52, 2.55, 2.06,
                           3.21, 2.32, 3.21, 2.04, 2.77, 2.31, 2.83, 2.22, 3.03)
                 )

View(DF)     # Table no. 10-10   (Table is different as compared to book)

res <- aov(Value ~ Test_Difficulty*Sensitivity, data
```

```
= DF)
result <- round(anova(res), 2)

# Table no. 10-11

result  # Values differ with book's values (table values)

# Below results are in table no. 10-10

ss_total <- round(sum(result$`Sum Sq`), 2)  # Total sum of squared deviation

cat("Total sum of squared deviation is \n", ss_total)

ss_within <- round(result$`Sum Sq`[4], 2)  # Sum of squared deviation within-groups

cat("Sum of squared deviation within-groups is \n", ss_within)

ss_interaction <- round(result$`Sum Sq`[3], 2)  # Sum of squared deviation of interaction

cat("Sum of squared deviation of interaction is \n", ss_interaction)

ss_rows <- round(result$`Sum Sq`[2], 2)  # Sum of squared deviation of rows

cat("Sum of squared deviation of rows is \n", ss_rows)

ss_columns <- round(result$`Sum Sq`[1], 2)  # Sum of squared deviation of columns

cat("Sum of squared deviation of columns is \n", ss_
```

```
    columns)

alpha <- 0.05   # 5% significance level
df_numerator <- 1   # 2 - 1 = 1   (2 groups)
df_denomenator <- result$Df[4]

fval <- qf(0.95, df_numerator, df_denomenator)
fval <- round(fval, 2)

# Note that fval is same for sensitivity, test difficulty and for interaction effect

cat("F value is \n", fval)

f1 <- result$`F value`[1]   # F value for test difficulty
f2 <- result$`F value`[2]   # F value for sensitivity
f3 <- result$`F value`[3]   # F value for test difficulty : sensitivity

if(f1 >= fval)
{
  cat("Reject null hypothesis for test difficulty")
} else {
  cat("Cannot reject null hypothesis for test difficulty")
}

if(f2 >= fval)
{
  cat("Reject null hypothesis for sensitivity")
} else {
  cat("Cannot reject null hypothesis for sensitivity")
}

if(f3 >= fval)
```

```
{
  cat("Reject null hypothesis for test difficulty : sensitivity")
} else {
  cat("Cannot reject null hypothesis for test difficulty : sensitivity")
}
```

R code Exa 10.1b How are you doing

```
                                              # Page no. : 411 - 413

# How are you doing?

# 5th Question

DF <- data.frame(Participant = LETTERS[1:8],
                 Length_condition = c(rep("Short", 4), rep("Long", 4)),
                 Vividness_condition = c(rep(c("Low", "High"), 4)),
                 Number_recalled = c(6, 9, 4, 5, 2, 1, 4, 1)
                )

View(DF)   # Book's table is little bit different

res <- aov(Number_recalled ~ Vividness_condition*Length_condition, data = DF)
result <- anova(res)
result

ss_total <- round(sum(result$`Sum Sq`), 2)   # Total sum of squared deviation
```

```r

cat("Total sum of squared deviation is", ss_total)

ss_within <- round(result$`Sum Sq`[4], 2)   # Sum of squared deviation within-groups

cat("Sum of squared deviation within-groups is", ss_within)

ss_interaction <- round(result$`Sum Sq`[3], 2)   # Sum of squared deviation of interaction

cat("Sum of squared deviation of interaction is", ss_interaction)

ss_rows <- round(result$`Sum Sq`[2], 2)   # Sum of squared deviation of rows

cat("Sum of squared deviation of rows is", ss_rows)

ss_columns <- round(result$`Sum Sq`[1], 2)   # Sum of squared deviation of columns

cat("Sum of squared deviation of columns is", ss_columns)

alpha <- 0.05   # 5% significance level
df_numerator <- 1   # 2 - 1 = 1   (2 groups)
df_denomenator <- result$Df[4]

fval <- qf(0.95, df_numerator, df_denomenator)
fval <- round(fval, 2)

f1 <- result$`F value`[1]   # F value for vividness condition
f2 <- result$`F value`[2]   # F value for length condition
f3 <- result$`F value`[3]   # F value for vividness
```

```
    condition : length condition

if(f1 >= fval)
{
  cat("Reject null hypothesis")
} else {
  cat("Cannot reject null hypothesis")
}

if(f2 >= fval)
{
  cat("Reject null hypothesis")
} else {
  cat("Cannot reject null hypothesis")
}

if(f3 >= fval)
{
  cat("Reject null hypothesis")
} else {
  cat("Cannot reject null hypothesis")
}
```

R code Exa 10.2a Advanced Topic Effect Size in the Factorial Analysis of Variance

```
                                    # Page no. : 413 -
                                        414

# Advanced Topic : Effect Size in the Factorial
    Analysis of Variance

var_col <- 1.20   # Variance (columns)
df_col <- 1   # Degree of freedom (columns)

```

```
var_within <- 0.03   # Variance within-groups
df_within <- 16   # Degree of freedom within-groups

r_sq1 <- round((var_col * df_col) / ((var_col * df_col) + (var_within * df_within)), 2)
                              # The proportion of Variance announced for (test difficulty)

cat("The proportion of Variance announced for (test difficulty) is", r_sq1)

var_rows <- 0.00   # Variance (rows)
df_rows <- 1   # Degree of freedom (rows)

r_sq2 <- round((var_rows * df_rows) / ((var_rows * df_rows) + (var_within * df_within)), 2)
# The proportion of Variance announced for (sensitivity)

cat("The proportion of Variance announced for (sensitivity) is", r_sq2)

var_int <- 0.60   # Variance (iteraction)
df_int <- 1   # Degree of freedom (interaction)

var_within <- 0.03   # Variance within-groups
df_within <- 16   # Degree of freedom within-groups

r_sq3 <- round((var_int * df_int) / ((var_int * df_int) + (var_within * df_within)), 2)
# The proportion of Variance announced for (test difficulty)

cat("The proportion of Variance announced for (test difficulty) is", r_sq3)
```

R code Exa 10.2b How are you doing

```
                                        # Page no. :
                                           416 - 417

# How are you doing?

# 1st Question (c) part

var_rows <- 9.45   # Variance (rows)
df_rows <- 1   # Degree of freedom (rows)

var_within <- 3.67   # Variance within-groups
df_within <- 36   # Degree of freedom within-groups

r_sq2 <- round((var_rows * df_rows) / ((var_rows *
   df_rows) + (var_within * df_within)), 2)
# The proportion of Variance announced for

cat("The proportion of Variance announced for is", r
   _sq2)
```

R code Exa 10.3a worked out examples 1

```
                                        # Page no. :
                                           419 - 421

# worked out examples 1

# Only (b) and (c) part

```

```r
DF <- data.frame(Variable_A = c(rep("Level1", 8), rep("Level2", 8), rep("Level3", 8)),
                 Variable_B = c(rep(c("Level1", "Level2"), 12)),
                 Values = c(25, 19, 20, 24, 23, 21, 24, 20, 22, 24, 19, 18, 22, 22, 21, 20,
                            16, 18, 19, 21, 13, 16, 16, 17)
)

View(DF)   # Book's table is little bit different

# install.packages(ggplot2)
library(ggplot2)

ggplot(DF, aes(x = Variable_A, y = Values, fill = Variable_B)) + geom_bar(position = "dodge",

# Litte bit different barplot as compare with book's barplot

res <- aov(Values ~ Variable_B*Variable_A, data = DF)
result <- anova(res)
result

ss_total <- round(sum(result$`Sum Sq`), 2)   # Total sum of squared deviation

```

```
cat("Total sum of squared deviation is", ss_total)

ss_within <- round(result$`Sum Sq`[4], 2)   # Sum of squared deviation within-groups

cat("Sum of squared deviation within-groups is", ss_within)

ss_interaction <- round(result$`Sum Sq`[3], 2)   # Sum of squared deviation of interaction

cat("Sum of squared deviation of interaction is", ss_interaction)

ss_rows <- round(result$`Sum Sq`[2], 2)   # Sum of squared deviation of rows

cat("Sum of squared deviation of rows is", ss_rows)

ss_columns <- round(result$`Sum Sq`[1], 2)   # Sum of squared deviation of columns

cat("Sum of squared deviation of columns is", ss_columns)

alpha <- 0.05   # 5% significance level
df_numerator <- 2   # 3 - 1 = 1   (2 groups)
df_denomenator <- result$Df[4]

fval <- qf(0.95, df_numerator, df_denomenator)
fval <- round(fval, 2)

f1 <- result$`F value`[1]   # F value for variable B
f2 <- result$`F value`[2]   # F value for variable A
f3 <- result$`F value`[3]   # F value for variable B : variable A

if(f1 >= fval)
```

```
{
  cat("Reject null hypothesis")
} else {
  cat("Cannot reject null hypothesis")
}

if(f2 >= fval)
{
  cat("Reject null hypothesis")
} else {
  cat("Cannot reject null hypothesis")
}

if(f3 >= fval)
{
  cat("Reject null hypothesis")
} else {
  cat("Cannot reject null hypothesis")
}
```

Chapter 11

Correlation

R code Exa 11.1a Graphing Correlations The Scatter Diagram An example

```
                                    # Page no. : 441 -
                                        442

# Graphing Correlations : The Scatter Diagram An
   example

Hours_slept <- c(5, 7, 8, 6, 6, 10)
Happy_mood <- c(2, 4, 7, 2, 3, 6)

DF <- data.frame(Hours_slept, Happy_mood)
View(DF)

# Install Library if not install

# install.packages("ggplot2")

# Import Library

library(ggplot2)

```

```
ggplot(DF, aes(x = Hours_slept, y = Happy_mood)) + geom_point() + labs(title =
        "Scatter plot", x = "Hours slept last night", y = "Happy mood") + theme_bw() +
  scale_x_continuous(limits=c(1, 12), breaks = c(1:12)) +
  scale_y_continuous(limits=c(1, 8), breaks = c(1:8))
```

R code Exa 11.1b How are you doing

```
                                                  # Page no. : 443

# How are you doing?

# 3rd Question

person <- LETTERS[1:4]
x <- c(3, 6, 1, 4)
y <- c(4, 7, 2, 6)

DF <- data.frame(person, x, y)
View(DF)

# Install Library if not install

# install.packages("ggplot2")

# Import Library

library(ggplot2)

ggplot(DF, aes(x = x, y = y)) + geom_point() + labs(title = "Scatter plot", x = "X", y = "Y") +
```

```
  theme_bw() + scale_x_continuous(limits=c(0, 6), breaks = c(0:6)) +
  scale_y_continuous(limits=c(0, 7), breaks = c(0:7))
```

R code Exa 11.2a Figuring the correlation coefficient an example

```
                                            # Page no. : 455 - 456

# Figuring the correlation coefficient an example

Hours_slept <- c(5, 7, 8, 6, 6, 10)
Happy_mood <- c(2, 4, 7, 2, 3, 6)

# Book Method

m1 <- mean(Hours_slept)   # Mean
m2 <- mean(Happy_mood)   # Mean

deviation1 <- Hours_slept - m1
sq_dev1 <- deviation1 ** 2   # Squared deviation
sd1 <- round(sqrt(sum(sq_dev1) / 6), 2)   # Standard deviation

z_score1 <- round(deviation1 / sd1, 2)

deviation2 <- Happy_mood - m2
sq_dev2 <- deviation2 ** 2   # Squared deviation
sd2 <- round(sqrt(sum(sq_dev2) / 6), 2)   # Standard deviation

z_score2 <- round(deviation2 / sd2, 2)

c_p <- round(z_score1 * z_score2, 2)   # Cross
```

```
Product

DF <- data.frame(Hours_slept, deviation1, sq_dev1, z_score1, Happy_mood, deviation2, sq_dev2,
                 z_score2, c_p)
View(DF)

r <- round(sum(DF$c_p) / nrow(DF), 2)   # Correlation Coefficient

cat("Correlation coefficient is", r)

# Direct method

r2 <- round(cor(Happy_mood,Hours_slept), 2)
cat("Correlation coefficient is", r2)
```

R code Exa 11.2b Figuring the correlation coefficient a second example

```
                                                  # Page no. : 456 - 457

# Figuring the correlation coefficient a second example

No_of_exposures <- c(1:8)
No_of_words_recalled <- c(3, 2, 6, 4, 5, 5, 6, 9)

# Direct method

r <- round(cor(No_of_exposures,No_of_words_recalled), 2)   # Correlation Coefficient
cat("Correlation coefficient is", r)
```

R code Exa 11.2c How are you doing

```
                                        # Page no. : 457
                                            - 458

# How are you doing?

# 4th Question

person <- c("K", "L", "M")
z_score1 <- c(0.5, -1.4, 0.9)
z_score2 <- c(-0.7, -0.8, 1.5)

c_p <- round(z_score1 * z_score2, 2)   # Cross
   product

r <- round(sum(c_p) / 3, 2)   # Correlation
   coefficien

cat("Correlation coefficient is", r)
```

R code Exa 11.3a Significance of a correlation coefficient

```
                                        # Page no. : 458 -
                                           459

# Significance of a correlation coefficient

N <- 3   # No. of perons
r <- 0.24   # Correlation coefficient

```

```
t <- round(r / sqrt((1 - (r ** 2)) / (N - 2)), 2)
    # t value

cat("t value is", t)
```

R code Exa 11.3b Significance of a correlation coefficient an example

```
                                    # Page no. : 459

# Significance of a correlation coefficient an
    example

N <- 6   # Sample size
r <- 0.85   # Correlation coefficient
alpha <- 0.05   # 5% significance level

df <- N - 2   # Degree of freedom

# Two - tailed

tval <- qt(1-alpha/2, df)
tval <- c(tval, -tval)

t <- round(r / sqrt((1 - (r ** 2)) / (N - 2)), 2)
    # t value

cat("t value is", t)

if(t > tval[1] || t < tval[2])
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 11.3c How are you doing

```r
# Page no. : 461 - 462

# How are you doing?

# 3rd Question

N <- 60   # Sample size
r <- -0.31   # Correlation coefficient
alpha <- 0.05   # 5% significance level

df <- N - 2   # Degree of freedom

# Two - tailed

tval <- qt(1-alpha/2, df)
tval <- c(tval, -tval)

t <- round(r / sqrt((1 - (r ** 2)) / (N - 2)), 2)   # t value

cat("t value is", t)

if(t > tval[1] || t < tval[2])
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 11.4a Worked out examples 1

```
                                    # Page no. : 477
                                        - 478

# Worked out examples 1

Elementry_School <- c("Main Street", "Casat", "
   Lakeland", "Shady Grove", "Jefferson")
Class_size <- c(25, 14, 33, 28, 20)
Achievement <- c(80, 98, 50, 82, 90)

DF <- data.frame(Elementry_School, Class_size,
   Achievement)
View(DF)

# Install Library if not install

# install.packages("ggplot2")

# Import Library

library(ggplot2)

ggplot(DF, aes(x = Class_size, y = Achievement)) +
   geom_point() +
  labs(title = "Scatter plot", x = "Class size", y =
      "Achievement test score") +
  theme_bw()
```

R code Exa 11.4b Worked out examples 2

```
# Page no. : 478 - 479

# Worked out examples 2

Elementry_School <- c("Main Street", "Casat", "Lakeland", "Shady Grove", "Jefferson")
Class_size <- c(25, 14, 33, 28, 20)
Achievement <- c(80, 98, 50, 82, 90)

# Direct method

r <- round(cor(Class_size,Achievement), 2)   # Correlation Coefficient
cat("Correlation coefficient is", r)
```

R code Exa 11.4c Worked out examples 3

```
# Page no. : 479 - 480

# Worked out examples 3

N <- 5   # Sample size
r <- -0.90   # Correlation coefficient
alpha <- 0.05   # 5% significance level

df <- N - 2   # Degree of freedom

# Two - tailed

tval <- qt(1-alpha/2, df)
tval <- c(tval, -tval)

t <- round(r / sqrt((1 - (r ** 2)) / (N - 2)), 2)
```

```
   # t value

cat("t value is", t)

if(t > tval[1] || t < tval[2])
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

Chapter 12

Prediction

R code Exa 12.1a The linear prediction rule an example

```
                                        # Page no. : 494 - 496

# The linear prediction rule an example

a <- 0.3   # Regression constant
b <- 0.004   # Regression coefficient
x <- 700   # SAT score

y_cap <- a + (b * x)   # Predicted GPA (Linear predictor)

cat("Predicted GPA (Linear predictor) is", y_cap)

# Another example

a <- -3   # Regression constant
b <- 1   # Regression coefficient
x <- 9   # Hours of sleep

y_cap <- a + (b * x)   # Predicted mood (Linear
```

```
predictor)

cat("Predicted mood (Linear predictor) is", y_cap)
```

R code Exa 12.1b How are you doing

```
                                        # Page no. : 497

# How are you doing?

# 4th Question

a <- -1.23   # Regression constant
b <- 6.11   # Regression coefficient

# (a)

x <- 2.00   # Predictor variable

y_cap <- a + (b * x)   # Predicted score (Linear predictor)

cat("Predicted score (Linear predictor) is", y_cap)

# (b)

x <- 4.87   # Predictor variable

y_cap <- a + (b * x)   # Predicted score (Linear predictor)

cat("Predicted score (Linear predictor) is", round(y_cap, 2))

```

```
# (a)

x <- -1.92   # Predictor variable

y_cap <- a + (b * x)   # Predicted score (Linear
    predictor)

cat("Predicted score (Linear predictor) is", round(y
    _cap, 2))
```

R code Exa 12.2a Another example of drawing the regression line

```
# Page no. : 501

# Another example of drawing the regression line

# Data on page no. : 442

Hours_slept <- c(5, 7, 8, 6, 6, 10)
Happy_mood <- c(2, 4, 7, 2, 3, 6)

DF <- data.frame(Hours_slept, Happy_mood)
View(DF)

# Import Library

library(ggplot2)

ggplot(DF, aes(x = Hours_slept, y = Happy_mood)) +
    geom_point() +
  labs(title =  "Scatter plot", x = "Hours slept
     last night", y = "Happy mood") + theme_bw() +
  scale_x_continuous(limits=c(1, 12), breaks = c
     (1:12)) +
  scale_y_continuous(limits=c(1, 8), breaks = c(1:8)
```

```
    ) +
  geom_smooth(method='lm', se = F)
```

R code Exa 12.3a Finding the best linear prediction rule

```
# Page no. : 502 - 504

# Finding the best linear prediction rule

# Rule 1 ==> y_cap <- 8 - (.18)*X
# Rule 2 ==> y_cap <- 4 + (0)*X
# Rule 3 ==> y_cap <- -2.5 + (1)*X
# Rule 4 ==> y_cap <- -3 + (1)*X

Hours_slept <- c(5, 6, 6, 7, 8, 10)
Actual_mood <- c(2, 2, 3, 4, 7, 6)

# Rule 1

y_cap1 <- 8 - (.18) * Hours_slept

# Rule 2

y_cap2 <- 4 + (0) * Hours_slept

# Rule 3

y_cap3 <- -2.5 + (1) * Hours_slept

# Rule 4

y_cap4 <- -3 + (1) * Hours_slept

DF <- data.frame(Hours_slept, Actual_mood, y_cap1, y
    _cap2, y_cap3, y_cap4)
```

```
View(DF)
```

R code Exa 12.3b The least squared error principle

```
# Page no. : 504 - 505

# The least squared error principle

# Rule 1 ==>  y_cap <- 8 - (.18)*X
# Rule 2 ==>  y_cap <- 4 + (0)*X
# Rule 3 ==>  y_cap <- -2.5 + (1)*X
# Rule 4 ==>  y_cap <- -3 + (1)*X

Hours_slept <- c(5, 6, 6, 7, 8, 10)
Actual_mood <- c(2, 2, 3, 4, 7, 6)

# Rule 1

y_cap1 <- 8 - (.18) * Hours_slept

error1 <- Actual_mood - y_cap1

error1_sq <- round(error1 ** 2, 2)

# Rule 2

y_cap2 <- 4 + (0) * Hours_slept

error2 <- Actual_mood - y_cap2

error2_sq <- error2 ** 2

# Rule 3

y_cap3 <- -2.5 + (1) * Hours_slept
```

```
error3 <- Actual_mood - y_cap3
error3_sq <- error3 ** 2
# Rule 4
y_cap4 <- -3 + (1) * Hours_slept
error4 <- Actual_mood - y_cap4
error4_sq <- error4 ** 2
DF <- data.frame(Hours_slept, Actual_mood, y_cap1, error1, error1_sq, y_cap4, error4, error4_sq)
View(DF)
s1 <- sum(DF$error1_sq)
cat("Rule 1 sum of squared errors is", s1)
s2 <- sum(error2_sq)
cat("Rule 2 sum of squared errors is", s2)
s3 <- sum(error3_sq)
cat("Rule 3 sum of squared errors is", s3)
s4 <- sum(DF$error4_sq)
cat("Rule 4 sum of squared errors is", s4)
```

R code Exa 12.3c Finding a and b for the least squares linear prediction rule

```
                                                # Page no. :
                                                   505 - 507

# Finding a and b for the least squares linear
    prediction rule

Hours_slept <- c(5, 7, 8, 6, 6, 10)
Happy_mood <- c(2, 4, 7, 2, 3, 6)

DF <- data.frame(Hours_slept, Happy_mood)
View(DF)

# Direct method

regressor <- lm(Happy_mood ~ Hours_slept, data = DF)
res <- summary(regressor)

res

b <- res$coefficients[[2]]

a <- res$coefficients[[1]]

cat("Linear Prediction Rule is y_cap =",a,"+",b,"x")
```

R code Exa 12.3d How are you doing

```
# Page no. : 507 - 508

# How are you doing?

# 4th Question

x <- c(4, 6, 7, 3)
y <- c(6, 8, 3, 7)
```

```
DF <- data.frame(x, y)
View(DF)

# (a) part

# Direct method

regressor <- lm(y ~ x, data = DF)
res <- summary(regressor)

res

b <- res$coefficients[[2]]

a <- res$coefficients[[1]]

cat("Linear Prediction Rule is y_cap =",a,"+",b,"x")

# (b) part

y_cap1 <- a + b * x

error1 <- y - y_cap1

error1_sq <- round(error1 ** 2, 2)

s1 <- sum(error1_sq)

cat("Sum of squared errors is", s1)

# (c) part

y_cap2 <- 9 - (0.7) * x

error2 <- y - y_cap2

error2_sq <- round(error2 ** 2, 2)
```

```
s2 <- sum(error2_sq)

cat("Sum of squared errors is", s2)
```

R code Exa 12.4a The standardized regression coefficient

```
# Page no. : 509 - 510

# The standardized regression coefficient

ss_x <- 16
ss_y <- 22
b <- 1

beta <- round(b * (sqrt(ss_x) / sqrt(ss_y)), 2)

cat("Standardized regression coefficient is", beta)
```

R code Exa 12.4b How are you doing

```
# Page no. : 511 - 512

# How are you doing?

# 2nd Question (b) part

ss_x <- 2.57
ss_y <- 7.21
b <- -1.21

beta <- round(b * (sqrt(ss_x) / sqrt(ss_y)), 2)

```

```r
cat("Standardized regression coefficient is", beta)
```

R code Exa 12.5a Multiple regression

```r
                                        # Page no. : 512 - 513

# Multiple regression

a <- -3.78

b1 <- 0.87
x1 <- 7

b2 <- 0.33
x2 <- 3

b3 <- 0.20
x3 <- 1

y_cap <- round(a + (b1 * x1) + (b2 * x2) + (b3 * x3), 2)   # Predicted mood (multiple regression)

cat("Predicted mood (multiple regression) is", y_cap)
```

R code Exa 12.5b How are you doing

```r
                                        # Page no. : 514 - 515

# How are you doing?

```

```
# 3rd Question

a <- 2.19

b1 <- -3.16

b2 <- 0.99

# (a) part

x1 <- 0.40
x2 <- 10.50

y_cap1 <- round(a + (b1 * x1) + (b2 * x2), 2)    # 
    Predicted score (multiple regression)

cat("Predicted score (multiple regression) is", y_
    cap1)

# (b) part

x1 <- 0.15
x2 <- 5.50

y_cap2 <- round(a + (b1 * x1) + (b2 * x2), 2)    # 
    Predicted score (multiple regression)

cat("Predicted score (multiple regression) is", y_
    cap2)
```

R code Exa 12.6a Proportionate reduction in error

```
                                        # Page no. :
                                           520 - 522

```

```
# Proportionate reduction in error

Hours_slept <- c(5, 6, 6, 7, 8, 10)
Actual_mood <- c(2, 2, 3, 4, 7, 6)

DF <- data.frame(Hours_slept, Actual_mood)
View(DF)

# Direct method

regressor <- lm(Actual_mood ~ Hours_slept, data = DF)
res <- summary(regressor)

res

p_r <- round(res$r.squared, 2) # Proportionate reduction in error

cat("Proportionate reduction in error is", p_r)
```

R code Exa 12.6b How are you doing

```
                                        # Page no. : 523 - 524

# How are you doing?

# 4th Question

score <- c(6, 4, 2, 8)
predicted_score <- c(5.7, 4.3, 2.9, 7.1)

m <- mean(score)   # mean

```

```
error <- score - m
error_sq <- error ** 2   # Error square

error2 <- score - predicted_score
error2_sq <- error2 ** 2   # Error square

DF <- data.frame(score, error, error_sq, predicted_
    score, error2, error2_sq)
View(DF)

ss_total <- sum(DF$error_sq)

ss_error <- sum(DF$error2_sq)

p_r <- round((ss_total - ss_error) / ss_total, 2) #
    Proportionate reduction in error

cat("Proportionate reduction in error is", p_r)

r <- round(sqrt(p_r), 2)   # Correlation coefficient

cat("Correlation coefficient is", r)
```

R code Exa 12.7a Worked out examples 1

```
                                        # Page no. :
                                            526

# Worked out examples 1

# This example includes solution of worked out
    examples 3

Elementry_School <- c("Main Street", "Casat", "
    Lakeland", "Shady Grove", "Jefferson")
```

```
Class_size <- c(25, 14, 33, 28, 20)
Achievement <- c(80, 98, 50, 82, 90)

DF <- data.frame(Elementry_School, Class_size, Achievement)
View(DF)

# Direct method

regressor <- lm(Achievement ~ Class_size, data = DF)
res <- summary(regressor)

res

b <- round(res$coefficients[[2]], 2)

a <- round(res$coefficients[[1]], 2)

cat("Linear Prediction Rule is y_cap =",a,"+",b,"x")

# (a) part

x <- 23   # Class size

y_cap <- a + (b * x)   # Predicted achievement

cat("Predicted achievement is", y_cap)

# (b) part

x <- 14   # Class size

y_cap <- a + (b * x)   # Predicted achievement

cat("Predicted achievement is", y_cap)
```

R code Exa 12.7b Worked out examples 2

```
                                        # Page no. :
                                            526

# Worked out examples 2

Elementry_School <- c("Main Street", "Casat", "
    Lakeland", "Shady Grove", "Jefferson")
Class_size <- c(25, 14, 33, 28, 20)
Achievement <- c(80, 98, 50, 82, 90)

DF <- data.frame(Elementry_School, Class_size,
    Achievement)
View(DF)

# Install Library if not install

# install.packages("ggplot2")

# Import Library

library(ggplot2)

ggplot(DF, aes(x = Class_size, y = Achievement)) +
    geom_point() +
  labs(title = "Scatter plot", x = "Class size", y =
      "Achievement test score") +
  theme_bw() + geom_smooth(method = "lm", se = F)

# Book's figure is in page 527
```

R code Exa 12.7c Worked out examples 4

```
                                        # Page no. :
                                           528

# Worked out examples 4

ss_x <- 214
ss_y <- 1328
b <- -2.25

beta <- round(b * (sqrt(ss_x) / sqrt(ss_y)), 2)    #
   Standardized regression coefficient

cat("Standardized regression coefficient is", beta)
```

R code Exa 12.7d Worked out examples 5

```
                                           # Page no. :
                                              528 -
                                              529

# Worked out examples 5

a <- 2.13

b1 <- 1.32
x1 <- 4

b2 <- 1.21
x2 <- 5

b3 <- 1.41
x3 <- 3

```

```
y_cap <- round(a + (b1 * x1) + (b2 * x2) + (b3 * x3)
    , 2)   # Predicted talkativeness

cat("Predicted talkativeness is", y_cap)
```

R code Exa 12.7e Worked out examples 6

```
                                          # Page no. : 529
                                              - 530

# Worked out examples 6

Elementry_School <- c("Main Street", "Casat", "
    Lakeland", "Shady Grove", "Jefferson")
Class_size <- c(25, 14, 33, 28, 20)
Achievement <- c(80, 98, 50, 82, 90)

DF <- data.frame(Elementry_School, Class_size,
    Achievement)
View(DF)

# Direct method

regressor <- lm(Achievement ~ Class_size, data = DF)
res <- summary(regressor)

res

p_r <- round(res$r.squared, 2) # Proportionate
    reduction in error

cat("Proportionate reduction in error is", p_r)
```

Chapter 13

Chi Square Tests

R code Exa 13.1a An example

```
# Page no. : 543 -544

# An example

gender <- c("Male", "Female")
o <- c(996, 390)   # Observed frequency
e <- c(693, 693)   # Expected frequency

diff <- o - e
diff_sq <- diff ** 2

ans <- round(diff_sq / e, 2)   # Difference squared
    weighted by expected frequency

DF <- data.frame(gender, o, e, diff, diff_sq, ans)
View(DF)
```

R code Exa 13.1b The chi square statistic and the chi square test for goodness of fit

```r
# Page no. : 545 -546

# The chi - square statistic and the chi - square
  test for goodness of fit

o <- c(996, 390)   # Observed frequency
e <- c(693, 693)   # Expected frequency

diff <- o - e
diff_sq <- diff ** 2
ans <- round(diff_sq / e, 2)

chi_sq <- sum(ans)

cat("Value of chi-square is", chi_sq)
```

R code Exa 13.1c Another example

```r
# Page no. : 549 -550

# Another example

condition <- c("Anxiety disorder", "Alcohol or drug
  abuse", "Mood disorder",
               "Impulse - control disorder", "None
                  of these conditions")

observed <- c(138, 99, 123, 111, 529)

expected <- c(146, 80, 110, 128, 536)

DF <- data.frame(condition, observed, expected)
View(DF)

alpha <- 0.05
```

```
df <- nrow(DF) - 1

chi_sq_val <- round(qchisq(alpha,df,lower.tail = F),
    2)

diff <- DF$observed - DF$expected
diff_sq <- diff ** 2
ans <- round(diff_sq / DF$expected, 2)

chi_sq <- sum(ans)

cat("Value of chi-square is", chi_sq)

if(chi_sq > chi_sq_val)
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 13.1d How are you doing

```
# Page no. : 551 - 553

# How are you doing?

# 5th Question (a) part

alpha <- 0.05
categories <- 2

observed <- c(15, 35)
N <- sum(observed)
expected <- c(0.6 * N, 0.4 * N)
```

```
df <- categories - 1

chi_sq_val <- round(qchisq(alpha,df,lower.tail = F), 2)

diff <- observed - expected
diff_sq <- diff ** 2
ans <- round(diff_sq / expected, 2)

chi_sq <- sum(ans)

cat("Value of chi-square is", chi_sq)

if(chi_sq > chi_sq_val)
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 13.2a The chi square test for independence

```
# Page no. : 553 - 557

# The chi - square test for independence

# Contingency Table

gender <- c(rep("Male", 28), rep("Female", 30), rep("Male", 125), rep("Female", 39))
age <- c(rep("Child", 58), rep("Adult", 164))

c_t <- table(age, gender)
```

```r
c_t

# Direct method

result <- chisq.test(c_t)
result

# Expected values

expected <- round(result$expected, 2)
expected

# Figuring chi - square

chi_sq <- round(result$statistic, 2)

cat("Chi-square value is", chi_sq)
# Degrees of freedom

df <- result$parameter

cat("Degrees of freedom is", df)

# Hypothesis testing

alpha <- 0.05

chi_sq_val <- round(qchisq(alpha,df,lower.tail = F), 2)

if(chi_sq > chi_sq_val)
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 13.2b Another example

```
# Page no. : 558 - 559

# Another example

# Contingency Table

generation <- c(rep("First", 73), rep("Other", 89),
   rep("First", 657), rep("Other", 1226))
drop <- c(rep("Dropped out", 162), rep("Did not drop
    out", 1883))

c_t <- table(drop, generation)

c_t

# Direct method

result <- chisq.test(c_t)
result

# Expected values

expected <- round(result$expected, 2)
expected

# Figuring chi - square

chi_sq <- round(result$statistic, 2)

cat("Chi-square value is", chi_sq)

# Degrees of freedom
```

```r

df <- result$parameter

cat("Degrees of freedom is", df)

# Hypothesis testing

alpha <- 0.01

chi_sq_val <- round(qchisq(alpha,df,lower.tail = F), 2)

if(chi_sq > chi_sq_val)
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
# Book's answer differ with our answer
```

R code Exa 13.2c How are you doing

```r
# Page no. : 560 - 561

# How are you doing?

# 4th Question (a) part

nominal_variable_A <- c(rep("Category1", 10), rep("Category2", 10), rep("Category1", 50),
                         rep("Category2", 10), rep("Category1", 10), rep("Category2", 10))
nominal_variable_B <- c(rep("Category1", 20), rep("
```

```
    Category2", 60), rep("Category3", 20))

c_t <- table(nominal_variable_B, nominal_variable_A)

c_t

# Direct method

result <- chisq.test(c_t)
result

# Expected values

expected <- round(result$expected, 2)
expected

# Figuring chi - square

chi_sq <- round(result$statistic, 2)

cat("Chi-square value is", chi_sq)

# Degrees of freedom

df <- result$parameter

cat("Degrees of freedom is", df)

# Hypothesis testing

alpha <- 0.10

chi_sq_val <- round(qchisq(alpha,df,lower.tail = F),
    2)

if(chi_sq > chi_sq_val)
{
  cat("Reject null hypothesis")
```

```
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 13.3a Effect Size for Chi Square Tests for Independence

```
# Page no. : 562 - 563

# Effect Size for Chi-Square Tests for Independence

chi_sq <- 6.73
N <- 2045

phi_coeff <- round(sqrt(chi_sq / N), 2)

cat("Phi coefficient is", phi_coeff)


chi_sq <- 12.70
N <- 100
df_smaller <- 1

c_phi_coeff <- round(sqrt(chi_sq / (N * df_smaller))
    , 2)

cat("Cramer's phi coefficient is", c_phi_coeff)   #
    Book's answer is wrong
```

R code Exa 13.3b How are you doing

```
# Page no. : 565 - 566

```

```r
# How are you doing?

# 2nd Question (d) part

chi_sq <- 12
N <- 100

phi_coeff <- round(sqrt(chi_sq / N), 2)

cat("Phi coefficient is", phi_coeff)

# 3rd Question (d) part

chi_sq <- 20
N <- 200
df_smaller <- 4 - 1

c_phi_coeff <- round(sqrt(chi_sq / (N * df_smaller)), 2)

cat("Cramer's phi coefficient is", c_phi_coeff)
```

R code Exa 13.4a Worked out examples 1

```r
# Page no. : 569 - 570

# Worked out examples 1

grade <- LETTERS[1:5]
observed <- c(10, 34, 140, 10, 6)
expected <- c(5, 28, 134, 28, 5)

DF <- data.frame(grade, observed, expected)
View(DF)

```

```
alpha <- 0.01
df <- nrow(DF) - 1

chi_sq_val <- round(qchisq(alpha,df,lower.tail = F), 2)

diff <- DF$observed - DF$expected
diff_sq <- diff ** 2
ans <- round(diff_sq / DF$expected, 2)

chi_sq <- sum(ans)

cat("Value of chi-square is", chi_sq)

if(chi_sq > chi_sq_val)
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 13.4b Worked out examples 2

```
# Page no. : 571 - 572

# Worked out examples 2

# Contingency Table

participant_gender <- c(rep("Men", 29), rep("Women", 17), rep("Men", 4), rep("Women", 14),
                        rep("Men", 26), rep("Women", 28))
comparision <- c(rep("Same sex", 46), rep("Opposite
```

```
    sex", 18), rep("Both sexes", 54))

c_t <- table(comparision, participant_gender)

c_t

# Direct method

result <- chisq.test(c_t)
result

# Expected values

expected <- round(result$expected, 2)
expected

# Figuring chi - square

chi_sq <- round(result$statistic, 2)

cat("Chi-square value is", chi_sq)

# Degrees of freedom

df <- result$parameter

cat("Degrees of freedom is", df)

# Hypothesis testing

alpha <- 0.05

chi_sq_val <- round(qchisq(alpha,df,lower.tail = F),
    2)

if(chi_sq > chi_sq_val)
{
  cat("Reject null hypothesis")
```

```
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 13.4c Worked out examples 3 and 4

```
                                          # Page no. : 572

# Worked out examples 3 and 4

chi_sq <- 14.41   # Chi-square value
N <- 85   # No. of people

phi_coeff <- round(sqrt(chi_sq / N), 2)   # Phi coefficient

cat("Phi coefficient is", phi_coeff)

chi_sq <- 8.78   # Chi-square value
N <- 118   # No. of people
df_smaller <- 2 - 1   # Degree of freedom on smaller side (2 x 2 contingency table)

c_phi_coeff <- round(sqrt(chi_sq / (N * df_smaller)), 2)   # Cramer's phi coefficient

cat("Cramer's phi coefficient is", c_phi_coeff)   # Book's answer is wrong
```

Chapter 14

Strategies When Population Distributions Are Not Normal

R code Exa 14.1a An example of a data transformation

```
# Page no. : 592 − 593

# An example of a data transformation

highly_sensitive_no <- c(0, 3, 10, 22)
highly_sensitive_yes <- c(17, 36, 45, 75)

m1 <- mean(highly_sensitive_no)
var1 <- round(var(highly_sensitive_no), 2)

m2 <- mean(highly_sensitive_yes)
var2 <- round(var(highly_sensitive_yes), 2)

sq_root1 <- round(sqrt(highly_sensitive_no), 2)
sq_root2 <- round(sqrt(highly_sensitive_yes), 2)

DF <- data.frame(highly_sensitive_no, sq_root1,
    highly_sensitive_yes, sq_root2)
View(DF)
```

```
alpha <- 0.05

df1 <- length(highly_sensitive_no) - 1
df2 <- length(highly_sensitive_yes) - 1

df_total <- df1 + df2

# One - tailed

tval <- round(qt(alpha, df_total, lower.tail = T), 2)

result <-t.test(DF$sq_root1, DF$sq_root2)
result

mean1 <- round(result$estimate[1], 2)
var_1 <- round(var(DF$sq_root1), 2)

mean2 <- round(result$estimate[2], 2)
var_2 <- round(var(DF$sq_root2), 2)

p_var <- round(((df1 / df_total) * var_1) + ((df2 / df_total) * var_2), 2)

cat("Pooled variance is", p_var)

t <- round(result$statistic[[1]], 2)   # t value

cat("t value is", t)

if(t < tval)
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 14.2a An example of a rank order test

```r
                                                # Page no. : 595 - 596

# An example of a rank - order test

highly_sensitive_no <- c(0, 3, 10, 22)
highly_sensitive_yes <- c(17, 36, 45, 75)

DF <- data.frame( highly_sensitive = rep(c("No", "Yes"), each = 4),
                  Values = c(highly_sensitive_no, highly_sensitive_yes)
                 )

Rank <- rank(DF$Values)
DF <- cbind(DF, Rank)

View(DF)

alpha <- 0.05   # 5% significance level

result <- wilcox.test(highly_sensitive_no, highly_sensitive_yes, paired = T)

result

p_value <- result$p.value

if(p_value > alpha)
{
  cat("Reject null hypothesis")
} else
```

```
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 14.2b Using parametric tests with rank transformed data

```
# Page no. : 596 - 597

# Using parametric tests with rank - transformed
   data

highly_sensitive_no <- c(1, 2, 3, 5)
highly_sensitive_yes <- c(4, 6, 7, 8)

m1 <- mean(highly_sensitive_no)
var1 <- round(var(highly_sensitive_no), 2)

m2 <- mean(highly_sensitive_yes)
var2 <- round(var(highly_sensitive_yes), 2)

df1 <- length(highly_sensitive_no) - 1
df2 <- length(highly_sensitive_yes) - 1
df_total <- df1 + df2
# One - tailed

alpha <- 0.05   # 5% significance level

tval <- round(qt(alpha, df_total, lower.tail = T),
   2)

p_var <- round(((df1 / df_total) * var1) + ((df2 /
   df_total) * var2), 2)   # Pooled variance

cat("Pooled variance is", p_var)

```

```
result <-t.test(highly_sensitive_no, highly_sensitive_yes)
result

t <- round(result$statistic[[1]], 2)   # t value

cat("t value is", t)

if(t < tval)
{
  cat("Reject null hypothesis")
} else
{
  cat("Cannot reject null hypothesis")
}
```

R code Exa 14.2c How are you doing

```
# Page no. : 598 - 599

# How are you doing?

# 2nd Question

scores <- c(5, 18, 3, 9, 2)

rank <- rank(scores)

DF <- data.frame(scores, rank)

View(DF)
```

R code Exa 14.3a Worked out examples 1 and 2

```r
                                              # Page no. : 605

# Worked out examples 1 and 2

group_A <- c(15, 4, 12, 14)
group_B <- c(21, 16, 49, 17)
group_C <- c(18, 19, 11, 22)

sq_root1 <- round(sqrt(group_A), 2)   # Square root of group A
sq_root2 <- round(sqrt(group_B), 2)   # Square root of group B
sq_root3 <- round(sqrt(group_C), 2)   # Square root of group C

DF <- data.frame(sq_root1, sq_root2, sq_root3)
View(DF)

DF2 <- data.frame( Groups = rep(c("Group A", "Group B", "Group C"), each = 4),
                   Values = c(group_A, group_B, group_C)
                   )

Rank <- rank(DF2$Values)
DF2 <- cbind(DF2, Rank)

View(DF2)
```