RWorksheet 7a

Krystal Rose M. Rizardo

2022-12-23

```
Student <- seq(1:10)
PreTest <- c(55,54,47,57,51,61,57,54,63,58)
PostTest <- c(61,60,56,63,56,63,59,56,62,61)

tableDF <- data.frame(Student,PreTest,PostTest)
tableDF
```

```
##    Student PreTest PostTest
## 1        1      55       61
## 2        2      54       60
## 3        3      47       56
## 4        4      57       63
## 5        5      51       56
## 6        6      61       63
## 7        7      57       59
## 8        8      54       56
## 9        9      63       62
## 10      10      58       61
```

```
agri_data<- c(10,10,10,20,20,50,10,
           20,10,50,20,50,20,10)
agri_data
```

```
##  [1] 10 10 10 20 20 50 10 20 10 50 20 50 20 10
```

```
data_factor <- factor(agri_data, ordered = TRUE)
data_factor
```

```
##  [1] 10 10 10 20 20 50 10 20 10 50 20 50 20 10
## Levels: 10 < 20 < 50
```

```
abdul_data <- c("l","n","n","i","l","l","n","n","i","l")
abdul_data
```

```
##  [1] "l" "n" "n" "i" "l" "l" "n" "n" "i" "l"
```

```
abdul_dataDF <- data.frame(abdul_data)
abdul_dataDF
```

```
##    abdul_data
## 1           l
## 2           n
## 3           n
## 4           i
## 5           l
## 6           l
## 7           n
## 8           n
## 9           i
## 10          l
```

```
mnemonics_data <- c("tas", "sa", "qld", "nsw", "nsw", "nt", "wa", "wa", "qld",
                "vic", "nsw", "vic", "qld", "qld", "sa", "tas", "sa", "nt",
                "wa", "vic", "qld", "nsw", "nsw", "wa", "sa", "act", "nsw",
                "vic", "vic", "act")
mnemonics_data
```

```
##  [1] "tas" "sa"  "qld" "nsw" "nsw" "nt"  "wa"  "wa"  "qld" "vic" "nsw" "vic"
## [13] "qld" "qld" "sa"  "tas" "sa"  "nt"  "wa"  "vic" "qld" "nsw" "nsw" "wa"
## [25] "sa"  "act" "nsw" "vic" "vic" "act"
```

```
mnemonics_factor <- factor(mnemonics_data)
mnemonics_factor
```

```
##  [1] tas sa  qld nsw nsw nt  wa  wa  qld vic nsw vic qld qld sa  tas sa  nt  wa
## [20] vic qld nsw nsw wa  sa  act nsw vic vic act
## Levels: act nsw nt qld sa tas vic wa
```

```
state_level <- levels(mnemonics_data)
state_level
```

```
## NULL
```

```
accountants_income <- c(60, 49, 40, 61, 64, 60, 59, 54,
            62, 69, 70, 42, 56, 61, 61, 61, 58, 51, 48,
            65, 49, 49, 41, 48, 52, 46, 59, 46, 58, 43)
accountants_income
```

```
##  [1] 60 49 40 61 64 60 59 54 62 69 70 42 56 61 61 61 58 51 48 65 49 49 41 48 52
## [26] 46 59 46 58 43
```

```
income_tapp <- tapply(accountants_income, mnemonics_data, mean)
income_tapp
```

```
##      act      nsw       nt      qld       sa      tas      vic       wa
## 44.50000 57.33333 55.50000 53.60000 55.00000 60.50000 56.00000 52.25000
```

```r
datalength <- length(income_tapp)
datalength
```

```
## [1] 8
```

```r
datastand <- sd(income_tapp)
datastand
```

```
## [1] 4.677966
```

```r
datas <- datastand/sqrt(datalength)
datas
```

```
## [1] 1.653911
```

```r
data(Titanic)
titanic_DF <- data.frame(Titanic)
titanic_DF
```

```
##    Class    Sex   Age Survived Freq
## 1    1st   Male Child       No    0
## 2    2nd   Male Child       No    0
## 3    3rd   Male Child       No   35
## 4   Crew   Male Child       No    0
## 5    1st Female Child       No    0
## 6    2nd Female Child       No    0
## 7    3rd Female Child       No   17
## 8   Crew Female Child       No    0
## 9    1st   Male Adult       No  118
## 10   2nd   Male Adult       No  154
## 11   3rd   Male Adult       No  387
## 12  Crew   Male Adult       No  670
## 13   1st Female Adult       No    4
## 14   2nd Female Adult       No   13
## 15   3rd Female Adult       No   89
## 16  Crew Female Adult       No    3
## 17   1st   Male Child      Yes    5
## 18   2nd   Male Child      Yes   11
## 19   3rd   Male Child      Yes   13
## 20  Crew   Male Child      Yes    0
## 21   1st Female Child      Yes    1
## 22   2nd Female Child      Yes   13
## 23   3rd Female Child      Yes   14
## 24  Crew Female Child      Yes    0
## 25   1st   Male Adult      Yes   57
## 26   2nd   Male Adult      Yes   14
## 27   3rd   Male Adult      Yes   75
## 28  Crew   Male Adult      Yes  192
## 29   1st Female Adult      Yes  140
## 30   2nd Female Adult      Yes   80
## 31   3rd Female Adult      Yes   76
## 32  Crew Female Adult      Yes   20
```

```
survive_data <- subset(titanic_DF, Survived == "Yes")
survive_data
```

```
##    Class    Sex   Age Survived Freq
## 17   1st   Male Child      Yes    5
## 18   2nd   Male Child      Yes   11
## 19   3rd   Male Child      Yes   13
## 20  Crew   Male Child      Yes    0
## 21   1st Female Child      Yes    1
## 22   2nd Female Child      Yes   13
## 23   3rd Female Child      Yes   14
## 24  Crew Female Child      Yes    0
## 25   1st   Male Adult      Yes   57
## 26   2nd   Male Adult      Yes   14
## 27   3rd   Male Adult      Yes   75
## 28  Crew   Male Adult      Yes  192
## 29   1st Female Adult      Yes  140
## 30   2nd Female Adult      Yes   80
## 31   3rd Female Adult      Yes   76
## 32  Crew Female Adult      Yes   20
```

```
didnt_surv_data <- subset(titanic_DF, Survived == "No")
didnt_surv_data
```

```
##    Class    Sex   Age Survived Freq
## 1    1st   Male Child       No    0
## 2    2nd   Male Child       No    0
## 3    3rd   Male Child       No   35
## 4   Crew   Male Child       No    0
## 5    1st Female Child       No    0
## 6    2nd Female Child       No    0
## 7    3rd Female Child       No   17
## 8   Crew Female Child       No    0
## 9    1st   Male Adult       No  118
## 10   2nd   Male Adult       No  154
## 11   3rd   Male Adult       No  387
## 12  Crew   Male Adult       No  670
## 13   1st Female Adult       No    4
## 14   2nd Female Adult       No   13
## 15   3rd Female Adult       No   89
## 16  Crew Female Adult       No    3
```

```
library("readxl")
```

```
## Warning: package 'readxl' was built under R version 4.2.2
```

```
install.packages("readxl")
```

```
## Warning: package 'readxl' is in use and will not be installed
```

```
Read <- read_excel("C:/Users/Acer/Downloads/Breast_Cancer.xlsx")
Read
```

```
## # A tibble: 49 x 11
##         Id CL. thickne~1 Cell ~2 Cell ~3 Marg.~4 Epith~5 Bare.~6 Bl. C~7 Norma~8
##      <dbl>         <dbl>   <dbl>   <dbl>   <dbl>   <dbl> <chr>     <dbl>   <dbl>
##  1 1000025             5       1       1       1       2 1             3       1
##  2 1002945             5       4       4       5       7 10            3       2
##  3 1015425             3       1       1       1       2 2             3       1
##  4 1016277             6       8       8       1       3 4             3       7
##  5 1017023             4       1       1       3       2 1             3       1
##  6 1017122             8      10      10       8       7 10            9       7
##  7 1018099             1       1       1       1       2 10            3       1
##  8 1018561             2       1       2       1       2 1             3       1
##  9 1033078             2       1       1       1       2 1             1       1
## 10 1033078             4       2       1       1       2 1             2       1
## # ... with 39 more rows, 2 more variables: Mitoses <dbl>, Class <chr>, and
## #   abbreviated variable names 1: 'CL. thickness', 2: 'Cell size',
## #   3: 'Cell Shape', 4: 'Marg. Adhesion', 5: 'Epith. C.size',
## #   6: 'Bare. Nuclei', 7: 'Bl. Cromatin', 8: 'Normal nucleoli'
```

```
mean_cell <- mean(Read$`Cell Shape`)
mean_cell
```

```
## [1] 3.163265
```

```
error_lng <- length(Read$`Cell Shape`)
error_lng
```

```
## [1] 49
```

```
standarde_sd <- sd(Read$`Cell Shape`)
standarde_sd
```

```
## [1] 2.910806
```

```
cellshape_standarderror <- standarde_sd/sqrt(error_lng)
cellshape_standarderror
```

```
## [1] 0.4158294
```

```
T_score = 0.05
T_score
```

```
## [1] 0.05
```

```
Ts_standarderror = error_lng - 1
Ts_standarderror
```

```
## [1] 48
```

```
T_score = qt(p=T_score/2, df=Ts_standarderror,lower.tail=F)
T_score
```

```
## [1] 2.010635
```

```
ConfidenceI <- T_score * cellshape_standarderror
ConfidenceI
```

```
## [1] 0.836081
```

```
ConfidenceI_diff <- mean_cell - ConfidenceI
ConfidenceI_diff
```

```
## [1] 2.327184
```

```
sumConfidenceI <- mean_cell + ConfidenceI
sumConfidenceI
```

```
## [1] 3.999346
```

```
Confidence <- c(ConfidenceI_diff, sumConfidenceI)
Confidence
```

```
## [1] 2.327184 3.999346
```

```
percentage_Res <- subset(Read, Class == "malignant")
percentage_Res
```

```
## # A tibble: 18 x 11
##         Id CL. thickne~1 Cell ~2 Cell ~3 Marg.~4 Epith~5 Bare.~6 Bl. C~7 Norma~8
##      <dbl>         <dbl>   <dbl>   <dbl>   <dbl>   <dbl> <chr>     <dbl>   <dbl>
##  1 1017122             8      10      10       8       7 10            9       7
##  2 1041801             5       3       3       3       2 3             4       4
##  3 1044572             8       7       5      10       7 9             5       5
##  4 1047630             7       4       6       4       6 1             4       3
##  5 1050670            10       7       7       6       4 10            4       1
##  6 1054590             7       3       2      10       5 10            5       4
##  7 1054593            10       5       5       3       6 7             7      10
##  8 1057013             8       4       5       1       2 NA            7       3
##  9 1065726             5       2       3       4       2 7             3       6
## 10 1072179            10       7       7       3       8 5             7       4
## 11 1080185            10      10      10       8       6 1             8       9
## 12 1084584             5       4       4       9       2 10            5       6
## 13 1091262             2       5       3       3       6 7             7       5
## 14 1099510            10       4       3       1       3 3             6       5
## 15 1100524             6      10      10       2       8 10            7       3
## 16 1102573             5       6       5       6      10 1             3       1
## 17 1103608            10      10      10       4       8 1             8      10
## 18 1105257             3       7       7       4       4 9             4       8
## # ... with 2 more variables: Mitoses <dbl>, Class <chr>, and abbreviated
## #   variable names 1: 'CL. thickness', 2: 'Cell size', 3: 'Cell Shape',
## #   4: 'Marg. Adhesion', 5: 'Epith. C.size', 6: 'Bare. Nuclei',
## #   7: 'Bl. Cromatin', 8: 'Normal nucleoli'
```