RWorksheet_Camiña7a

Francine Camiña

2022-11-26

```
install.packages("Hmisc")
install.packages("pastecs")
```

```
#1.
  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)

  DF <- data.frame(Student,PreTest,PostTest)
  DF


  #1.a

  library(Hmisc)
  library(pastecs)

    describe(DF)

    stat.desc(DF)
```

```
#2.
    DepartmentofAgriculture <- c(10,10,10,20,20,50,10,
                          20,10,50,20,50,20,10)



  #2.a
    In_Ord <- sort(DepartmentofAgriculture, decreasing = FALSE)
    In_Ord
```

```
#3.
    subjects <- c("l","n","n","i","l","l","n","n","i","l")

  #3.a
    out <- data.frame(subjects)
    out
```

```
#4.
    state <- c("tas", "sa", "qld", "nsw", "nsw", "nt", "wa", "wa", "qld",
```

```r
            "vic", "nsw", "vic", "qld", "qld", "sa", "tas", "sa", "nt",
            "wa", "vic", "qld", "nsw", "nsw", "wa", "sa", "act", "nsw",
            "vic", "vic", "act")
    state
```

  #4.a
```r
    fctor <- function(state)
    fctor
```

#5.

  #5•
```r
    incomes <- 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)
```

  #5.a
```r
    calcu <- tapply(state, incomes, mean)
    calcu
```

  #5.b
```r
    #40 41 42 43 46 48 49 51 52 54 56 58 59 60 61 62 64 65 69 70
    #NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
```

#6.
```r
    calcu_ST.n <- length(calcu)
    calcu_1.sd <- sd(calcu)
    calcu_Final.se <- calcu_1.sd/sqrt(calcu_ST.n)
    calcu_Final.se
```

  #6.a
```r
    #NA
```

  #6.b
```r
    #Results are not available due to some objects are character type, as a result it won't able to get
```

#7.
```r
    data("Titanic")
    head<- data.frame(Titanic)
```

  #7.a
```r
    head_subset <- subset(head, select = "Survived")
    head_subset
```

#8.

  #8.a
```r
    #The dataset s all about Breast Cancer.
```

  #8.b

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

```
   DATA <- read_excel("D:\\Franx_BSIT2A\\RWorksheet_7a//breast_cancer.xlsx")
   DATA
#8.c

#8.c.1
   clump <- length(DATA$'CL. thickness')
   clump_A <- sd(DATA$'CL. thickness')
   clump_B <- clump_A/sqrt(DATA$'CL. thickness')
   clump_B

#8.c.2
   coeff <- sd(DATA$'Marg. Adhesion') / mean(DATA$'Marg. Adhesion')* 100
   coeff

#8.c.3
   null_values <- subset(DATA,'Bare. Nuclei' == "NA")

#8.c.4
   mean(DATA$'Bl. Cromatin')
   sd(DATA$'Bl. Cromatin')

#8.c.5

#Calculate the mean
   calmean <- mean(DATA$'Cell Shape')
   calmean

#Calculate the standard error of the mean
   SE_M <- length(DATA$'Cell Shape')
   SD_B <- sd(DATA$'Cell Shape')
   Ans_1 <- SD_B/sqrt(SE_M)
   Ans_1

#Find the t-score that corresponds to the confidence level
   D = 0.05
   numE = SE_M - 1
   numF = qt(p = D/ 2, df = numE,lower.tail = F)
   numF

#Constructing the confidence interval
   numG <- numF * numE

#Lower
   numH <- calmean - numG

#Upper
   numI <- calmean + numG

   c(numH,numI)


#d. How many attributes?
   attributes(DATA)
```

```
#e. Find the percentage of respondents who are malignant. Interpret the results.
  P_R <- subset(DATA, Class == "maligant")
  P_R
    #There 17 respondents who are malignant.
    #And there are total of 49 respondent.

#Getting the percentage
  17  / 49 * 100
#9. Export the data abalone to the Microsoft excel file. Copy the codes.

install.packages("AppliedPredictiveModeling")

library("AppliedPredictiveModeling")

  data("abalone")
  View(abalone)
  head(abalone)
  summary(abalone)
#Exporting the data abalone to the Microsoft excel file
install.packages("xlsxjars")

library(xlsx)
write.xlsx("abalone","D:\\Franx_BSIT2A\\RWorksheet_7a\\abalone.xlsx")
```