Rworksheet_loredo#3b.Rmd

2023-10-31

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

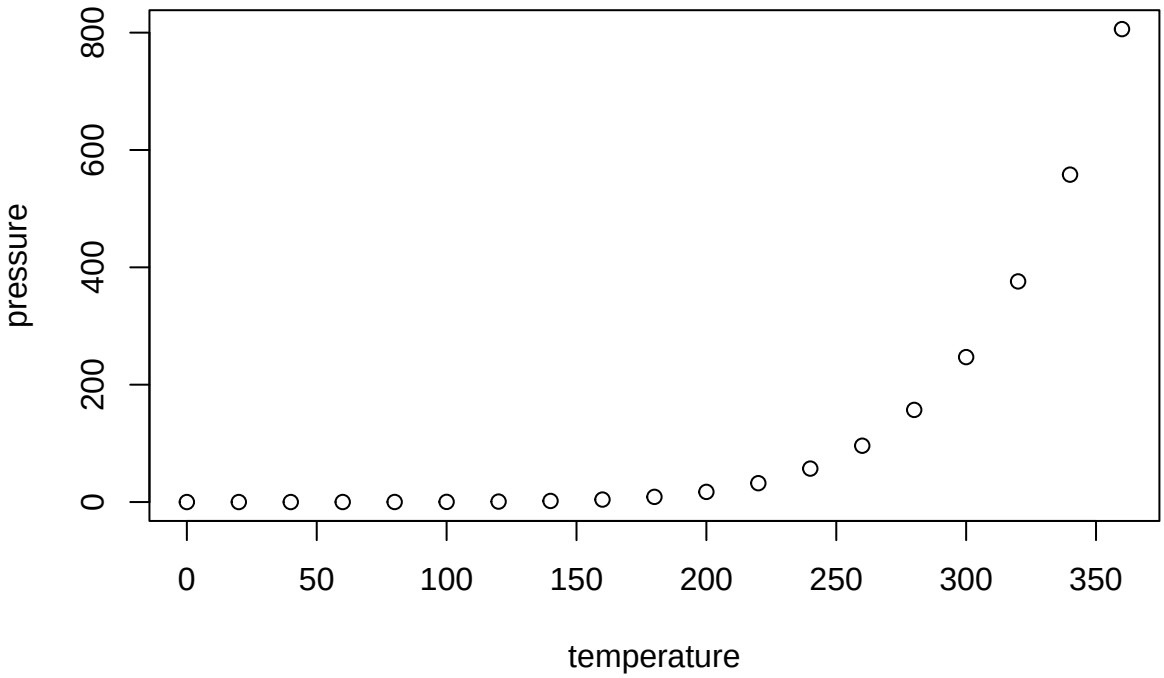
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

#1. The table below shows the data about shoe size and height. Create a data frame.

shoe_size <- c(6.5, 9.0, 8.5, 8.5, 10.5, 7.0, 9.5, 9.0, 13.0, 7.5, 10.5, 8.5, 12.0, 10.5, 13.0, 11.5, 8.5, 5.0, 10.0, 6.5, 7.5, 8.5, 10.5, 8.5 ,10.5, 11.0, 9.0, 13.0) height <- c(66.0, 68.0, 64.5, 65.0, 70.0, 64.0, 70.0, 71.0, 72.0, 64.0, 74.5, 67.0, 71.0, 71.0, 77.0, 72.0, 59.0, 62.0, 72.0, 66.0, 64.0, 67.0, 73.0, 69.0, 72.0, 70.0,69.0,70.0) gender <- c("F","F","F","F","M","F","F","F","M","F","M","F","M","M","M","M","F","F","M","F","F","M","M","F","M","M","M","

#1. Create a data frame data <- data.frame(ShoeSize = shoe_size, Height = height, Gender= gender)

#a. Describe the data. summary(data)

#b. Create a subset by males and females with their corresponding shoe size and height. #What its result? Show the R scripts.

males <- data[data$Gender == "M", c("Gender", "ShoeSize", "Height")]$females < -data[data$Gender == "F", c("Gender", "ShoeSize", "Height")]

males females

#c. Find the mean of shoe size and height of the respondents. Write the R scripts and it result. mean_shoe_size <- mean(data$ShoeSize)$mean_height < -mean(data$Height)

mean_shoe_size mean_height

#d. Is there a relationship between shoe size and height? Why? #No, Both shoe size and height can change over a person's lifetime. correlation <- cor(data$ShoeSize, data$Height) correlation

#2. Create a character vector of months months <- c("March", "April", "January", "November", "January", "September", "October", "September", "November", "August", "January", "November", "November", "February", "May", "August", "July", "December", "August", "August", "September", "November", "February", "April") factor_months_vector <- factor(months) print(factor_months_vector) levels(factor_months_vector) #3 summary # character vector #for numerical analysis summary(months) # the factor # for categorical analysis summary(factor_months_vector)

#4 Create a vector and factor direction_vector <- c("north", "east", "west") frequency_vector <- c(1, 4, 3) factor_data <- factor(direction_vector, levels = c("east", "west", "north"), c(1, 4, 3)) print(factor_data)

#5 A Import the excel library(readr) import_march <- read_csv("import_march.csv")

#5 B.View the dataset head(import_march)

#6 Full Search num <- readline(prompt= "Enter number from 1 to 50:") paste("Your entered number is", num) if(num == 50) { paste("The number you selected is beyond the range of 1 to 50") } else if (num <= 50){ paste("TRUE") } else { paste(num) } #7 Change minimum <- function (price) { bill <- price%/% 50 paste("The minimum number of bills:", bill) } snackprice <- 250 minimum(snackprice)

#8 A. Create a dataframe name <- c("Annie", "Thea", "Steve", "Hanna") grade1 <- c(85, 65, 75, 95) grade2 <- c(65, 75, 55, 75) grade3 <- c(85, 90, 80, 100) grade4 <- c(100, 90, 85, 90) mathgrades <- data.frame(name, grade1, grade2, grade3, grade4) print(mathgrades)

#8 B. the average score of students whose average math score over 90 points during the semester. mathgrades <- data.frame(name = c("Annie", "Thea", "Steve", "Hanna"), grade1 = c(85, 65, 75, 95), grade2 = c(65, 75, 55, 75), grade3 = c(85, 90, 80, 100), grade4 = c(100, 90, 85, 90)) print(mathgrades) mathgrades$average < -(mathgrades$grade1 + mathgrades$grade2 + mathgrades$grade3 + mathgrades$grade4)/4top < -mathgrades[mathgrades$average >= 90,] top if (nrow(top) > 0) { paste(top$name, "'saveragegradethissemesteris", top$average) } else { paste("No students have an average math score over 90.") }

#8 C. the average score was less than 80 out of 4 tests. test1 <- sum(mathgrades$grade1)/nrow(mathgrades)test1test2 < -sum(mathgrades$grade2) / nrow(mathgrades) test2 test3 <- sum(mathgrades$grade3)/nrow(mathgrades)test3test4 < -sum(mathgrades$grade4) / nrow(mathgrades) test4 if (test1 < 80) { paste("The 1st test was difficult") } else if (test2 < 80) { paste("The 2nd test was difficult") } else if (test3 < 80) { paste("The 3rd test was difficult") } else if (test4 < 80) { paste("The 4th test was difficult") } else { paste("No test had an average grade less than 80") }

#8 D.students whose highest score for a semester exceeds 90 points. # annie scores if (mathgrades[1,2] > mathgrades[1,3] && mathgrades[1,2] > mathgrades[1,4] && mathgrades[1,2] > mathgrades[1,5]) { annie <- mathgrades[1,2] } else if (mathgrades[1,3] > mathgrades[1,4] && mathgrades[1,3] > mathgrades[1,5]) { annie <- mathgrades[1,3] } else if (mathgrades[1,4] > mathgrades[1,5] && mathgrades[1,2] > mathgrades[1,5]) { annie <- mathgrades[1,4] } else { annie <- mathgrades[1,5] }

thea scores

if (mathgrades[2,2] > mathgrades[2,3] && mathgrades[2,2] > mathgrades[2,4] && mathgrades[2,2] > mathgrades[2,5]) { thea <- mathgrades[2,2] } else if (mathgrades[2,3] > mathgrades[2,4] && mathgrades[2,3] > mathgrades[2,5]) { thea <- mathgrades[2,3] } else if (mathgrades[2,4] > mathgrades[2,5] && mathgrades[2,2] > mathgrades[2,5]) { thea <- mathgrades[2,4] } else { thea <- mathgrades[2,5] }

steve scores

if (mathgrades[3,2] > mathgrades[3,3] && mathgrades[3,2] > mathgrades[3,4] && mathgrades[3,2] > mathgrades[3,5]) { steve <- mathgrades[3,2] } else if (mathgrades[3,3] > mathgrades[3,4] && mathgrades[3,3] > mathgrades[3,5]) { steve <- mathgrades[3,3] } else if (mathgrades[3,4] > mathgrades[3,5] && mathgrades[3,2] > mathgrades[3,5]) { steve <- mathgrades[3,4] } else { steve <- mathgrades[3,5] }

hanna scores

if (mathgrades[4,2] > mathgrades[4,3] && mathgrades[4,2] > mathgrades[4,4] && mathgrades[4,2] > mathgrades[4,5]) { hanna <- mathgrades[4,2] } else if (mathgrades[4,3] > mathgrades[4,4] && mathgrades[4,3] > mathgrades[4,5]) { hanna <- mathgrades[4,3] } else if (mathgrades[4,4] > mathgrades[4,5] && mathgrades[4,2] > mathgrades[4,5]) { hanna <- mathgrades[4,4] } else { hanna <- mathgrades[4,5] }

mathgrades$highest <- c(annie, thea, steve, hanna)

abovegradeof90 <- mathgrades[mathgrades$highest >= 90,]

if (nrow(abovegradeof90) > 0) { paste(abovegradeof90$name, "'s highest grade this semester is", abovegradeof90$highest) } else { paste("No students have an average math score over 90.") }