Kasu Kiran

1/10/2021

Professor Ji - Young Yun

CRN: ALY6000.71477

Describe the attributes of a given data set using appropriate terms

The given attributes are Student ID, First Name, Last Name, Math, Science, Social Studies in which we can get the Student's Identification number and also their First and Last name and the student scores in Math, Science and Social Studies.

Distinguish between Descriptive statistics and Inferential Statistics

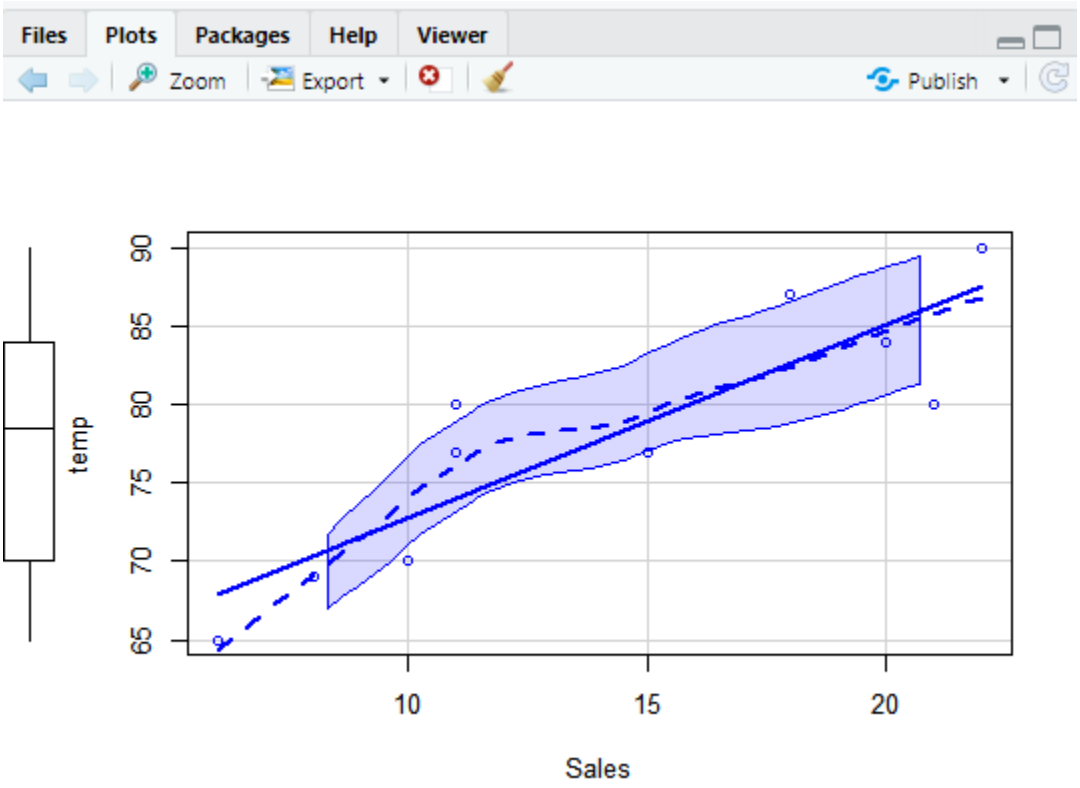
Descriptive Statistics	**Inferential Statistics**
Descriptive Statistics is a branch of statistics that is concerned with describing the population under a study	Inferential Statistics is a type of statistics, that focuses on drawing conclusions about the population, on the basis of sample analysis and observations.
It organizes, analyze and present data in a meaningful way	Compares, test and predicts data
Results are in the form of charts, graphs and Tables	Results are in the form of probability
The use of descriptive statistics is to describe a situation	The use of inferential statistics is to explain the chance of occurrence of an event
Function of descriptive statistics is to explain the data, which is already known, to summarize sample	It attempts to reach the conclusion to learn about the population, that extends beyond the data available.

Key Findings:

(a) A scatter plot of the Sales ~ temp data

Command : Sales <- c (8,11,15,20,21,11,18,10,6,22)
temp <- c (69,80,77,84,80,77,87,70,65,90)
scatterplot (Sales, temp)

Output:



(b) The mean temperature

Command: mean(temp)
Output: 77.9

(c) Display the data after Steps 6 and 7

Command: Sales <- Sales[-3]

Sales_new <- c(Sales[1:4],16,Sales [5:8])

Sales_new

Output: [1] 8 11 20 21 16 11 18 10 6

(d) Display names vector

Command: name <- c ("Tom","Dick","Harry")
name

Output: [1] "Tom" "Dick" "Harry"

(e) Display the 5 rows by 2 columns of 10 integers

Command: y <- matrix (1:10,nrow = 5,ncol = 2)
y

Output:

```
     [,1] [,2]
[1,]  1    6
[2,]   2    7
[3,]   3    8
[4,]   4    9
[5,]   5   10
```

(f) Display the IcSales data frame

Command: icsales <- data.frame(Sales,temp)
icsales

Output: [1] 8 11 20 21 11 18 10 6 22 69 80 77 84 80 77 87 70 65 90

(g) Display the summary icsales data frame

Command: summary(icsales)

Output:

```
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   6.00   14.50   65.00   47.68   78.50   90.00
```

(h) Display the variables only from the Student.csv data set

Command: variable. names(data)
Output: [1] "StudentID" "First" "Last" "Math"
[5] "Science" "Social.Studies

(i) **Summary:**

- Firstly, we have given input values for the Sales and temperature (temp) vectors
- Created a Scatterplot for both Sales and Temp
- Identified the Mean of temp
- Removed the 3rd element from the sales vector and added a different element
- Given input values for names vector
- Created a matrix with 5 rows and 2 columns with 10 integers
- Created and displayed the data frame with sales and temp attributes
- Imported the dataset from Student.csv file
- Displayed the variable names from the student.csv file

Conclusion:

- I learnt how to work on R and R studio
- How to create vector and add input values to vectors
- How to calculate mean value
- How to create matrices and data frames
- Finally, on how to import a data set from a .csv file

Bibliography:

https://keydifferences.com/difference-between-descriptive-and-inferential-statistics.html

S, S., says, R. K., Kausar, R., says, P., Parichita, says, S. sant, sant, S., says, I. B., Ben, I., says, M., Mickeyroni, says, S. S., says, V., Vandana, Says, B., Babatunde, says, E., Efosa, says, S., … Joshua. (2019, December 12). *Difference between descriptive and Inferential Statistics (with comparison chart)*. Key Differences. Retrieved October 1, 2021, from https://keydifferences.com/difference-between-descriptive-and-inferential-statistics.html.

Appendix:

R Script:

```
x <- "Kiran Kasu"

x

install.packages("Hmisc")

install.packages("corrplot")

install.packages("function")

install.packages("vcd")

library("vcd")

library(dplyr)

library(ggplot2)

library(corrplot)

library(RColorBrewer)

library(car)

library(MASS)

library(corrplot)

library(Hmisc)

library(data.table)

Sales <- c  (8,11,15,20,21,11,18,10,6,22)

temp <- c (69,80,77,84,80,77,87,70,65,90)

scatterplot(Sales,temp)
```

```
mean(temp)
Sales <- Sales[-3]
Sales_new <- c(Sales[1:4],16,Sales [5:8])
Sales_new
name <- c("Tom","Dick","Harry")
name
y <- matrix(1:10,nrow = 5,ncol = 2)
y
icsales <- c(Sales,temp)
icsales
str(icsales)
icsales <- data.frame(Sales,temp)
icsales
summary(icsales)
getwd()
data <- read.csv("C:/Users/kasuk/Downloads/Student1.csv",header =TRUE)
data
variable.names(data)
```