Executive Summary Report 1

Introduction to statistics and language R

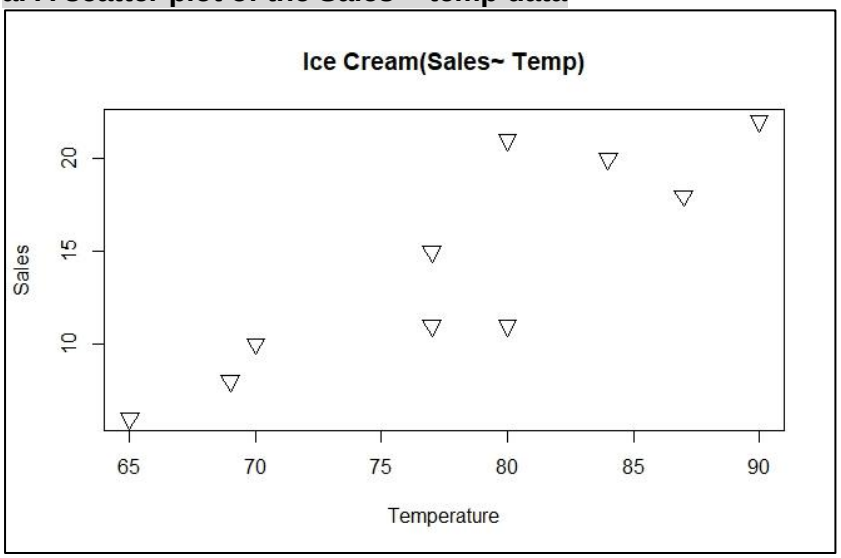
ALY6000: Introduction to Analytics

Prepared by: Heejae Roh
Presented to: Professor Behzad Ahmadi

Date: Sep 26th, 2022

Key findings about the data based on the Dataset Instruction document.

a. A scatter plot of the Sales ~ temp data



Comprehension: The graph above indicates a positive correlation of Temp & sales. If temperature goes up, sales of ic(ice cream) will go up. If I'm the manager of ice cream store, I can draw a conclusion that 'prepare more ice cream. when temperature is high'.

b. The mean temperature

```
> mean(temp)
[1] 77.9
```

c. Display the data after steps 6 and 7

```
> sales=sales[sales!=sales[3]]
> sales <- append(sales, 16, after=2)
> sales
 [1]  8 11 16 20 21 11 18 10  6 22
```

Comprehension: 'append()' function (x, values, after=length(x)) can add values.
If I operate 'sales[3] <- 16', It will just change 3rd vector to 16.

d. Display the names vector

```
> names <- c("Tom", "Dick", "Harry")
> names
[1] "Tom"   "Dick"  "Harry"
```

e. Display the 5 row by 2 column of 10 integers

```
> integers <- matrix(1:10, nrow=5, ncol=2)
> integers
     [,1] [,2]
[1,]    1    6
[2,]    2    7
[3,]    3    8
[4,]    4    9
[5,]    5   10
```

Comprehension: Operating 'integers <- matrix(1:10, nrow=5)', the result will be same.

f. Display the icSales data frame

```
> icSales <- data.frame(sales, temp)
> icSales
   sales temp
1      8   69
2     11   80
3     16   77
4     20   84
5     21   80
6     11   77
7     18   87
8     10   70
9      6   65
10    22   90
```

g. Display the summary of the icSales data frame

```
> summary(icSales)
     sales            temp
 Min.   : 6.00   Min.   :65.00
 1st Qu.:10.25   1st Qu.:71.75
 Median :13.50   Median :78.50
 Mean   :14.30   Mean   :77.90
 3rd Qu.:19.50   3rd Qu.:83.00
 Max.   :22.00   Max.   :90.00
```

Comprehension: The analysis of descriptive characteristics of data conducted by each variable.

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

```
> colnames(data.frame1)
[1] "StudentID"      "First"          "Last"           "Math"
[5] "Science"        "Social.Studies"
```

i. Summary

Data set is a collection of data which shows a characteristic or attribute that can assume values. In R, a researcher could find the correlation of data by drawing a plot. In other words, R helps a researcher convert complex data into visualized form. A researcher assumes the relation of variables, so a researcher could draw a conclusion of the population through the sample. It's inferential.

A researcher could analyze the descriptive data like Mean, Median, Min, Max, etc. from a sample. In our work, we used 'summary()' and discover the sales mean, the temp mean, etc.

Importing datasets into R is very important to analyze big data valuables. The data types of R are numeric (1, 2), character ("Tom") or logical (TRUE). We can know about vector information by 'str()' function. If we use levels and labels options, we can code factors as numeric vector. It helps imported data processed.

Discussion: Which setting or preparation reduce errors occurring in data import?

Bibliography

Kabacoff. Robert. (2015). R In Action: Data Analysis and Graphics with R. Manning.

Bluman, A. (2017). *Elementary Statistics: A Step By Step Approach* (10th ed.). McGraw-Hill Higher Education (US). https://reader2.yuzu.com/books/9781260042054

narasingha, (2018, October 18). How to remove value from a vector in R. Analytics Vidhya. Retrieved from https://discuss.analyticsvidhya.com/t/how-to-remove-value-from-a-vector-in-r/2975

Jitender_1998, (2020, May 10). Adding elements in a vector in R programming – append() method. GeeksforGeeks. Retrieved from https://www.geeksforgeeks.org/adding-elements-in-a-vector-in-r-programming-append-method/

Zach, (2020, October 27). How to Import CSV Files into R (Step-by-Step). STATOLOGY. Retrieved from https://www.statology.org/import-csv-into-r/

sudhanshublaze, (2021, May 27). How to add Header to Dataframe in R ?. GeeksforGeeks. Retrieved from https://www.geeksforgeeks.org/how-to-add-header-to-dataframe-in-r/

Andrie de Vries and Joris Meys, (2016, March 26). How to Add Variables to a Data Frame in R. dummies. Retrieved from How to Add Variables to a Data Frame in R - dummies

aL3xa, (2016, March 26). Control the size of points in an R scatterplot?. stackoverflow. Retrieved from https://stackoverflow.com/questions/2579995/control-the-size-of-points-in-an-r-scatterplot

Ki-young kwahk, (n.d.). R programming / R Basic. YouTube. Retrieved from https://www.youtube.com/watch?v=MbUnU-D01Gg&list=PLY0OaF78qqGA42e0rXqBWCcF_C9Ib-T1x&index=10

Appendix: The R Script

```
#Heejae Roh#


install.packages("vcd")

library(vcd)

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

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

plot(temp,sales, main="Ice Cream(Sales ~ Temp)", xlab="Temperature", ylab="Sales",
pch=25, cex=1.5)

mean(temp)

sales=sales[sales!=sales[3]]

sales <- append(sales, 16, after=2)

names <- c("Tom", "Dick", "Harry")

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

icSales <- data.frame(sales, temp)

str(icSales)

summary(icSales)

data.frame1 <-read.csv("C:\\Users\\14083\\Desktop\\Student.csv", header=TRUE)

data.frame1

colnames(data.frame1)
```