

Worksheet-5 in R

Worksheet for R Programming

Instructions:

- Use RStudio or the RStudio Cloud accomplish this worksheet.
- Save the R script as *RWorksheet_lastname#5.R*.
- Create your own *GitHub repository* and push the R script as well as this pdf worksheet to your own repo. Do not forget to comment your Git repo

Accomplish this worksheet by answering the questions being asked and writing the code manually.

1. The table shows the enrollment of BS in Computer Science, SY 2010-2011.

Course Year	2019 - 2020
1st	80
2nd	75
3rd	70
4th	60

- a. Plot the data using a bar graph. Write the codes and copy the result.

```
year_2019_2020 <- c(80, 75, 70, 60)
barplot(year_2019_2020)
```

- b. Using the same table, label the barchart with
Title = "Enrollment of BS Computer Science"
horizontal axis = "Curriculum Year" and
vertical axis = "number of students"

```
barplot(year_2019_2020,
        main= "Enrollment of BS Computer Science",
        xlab= "Curriculum Year",
        ylab= "number of students", names.arg= c("1st", "2nd", "3rd", "4th"))
```

2. The monthly income of De Jesus family was spent on the following:

60% on Food, 10% on electricity, 5% for savings, and 25% for other miscellaneous expenses.

- a. Create a table for the above scenario.
Write the codes and its result.

```
Monthlyincome_Dejesus <- data.frame(Food= ("60%"), Electricity= ("10%"),  
                                     Savings= ("5%"), Other_miscellaneous_expenses= ("25%"))  
Monthlyincome_Dejesus
```

- b. Plot the data using a pie chart. Add labels, colors and legend.
Write the codes and its result.

```
Monthlyincome_Dejesus <- c(60, 10, 5, 25)  
Monthlyincome_Dejesus  
  
pie(Monthlyincome_Dejesus,  
    main= "cost",  
    col= rainbow(length(Monthlyincome_Dejesus)),  
    labels= c("Food", "Electricity", "Savings", "Other miscellaneous expenses"))  
legend("topright", c("Food", "Electricity", "Savings", "Other miscellaneous expenses"),  
      cex= 0.5, fill= rainbow(length(Monthlyincome_Dejesus)))
```

3. Open the mtcars dataset.

- a. Create a simple histogram specifically for **mpg** (miles per gallon) variable.
Use \$ to select the *mpg* only. Write the codes and its result.

```
simple_histo <- (mtcars$mpg)  
hist(simple_histo,)
```

- b. Colored histogram with different number of bins.
hist(mtcars\$mpg, breaks=12, col="red")
Note: breaks= controls the number of bins

```
hist(simple_histo, breaks=12, col="red")
```

c. Add a Normal Curve

```
x <- mtcars$mpg
h<-hist(x, breaks=10, col="red", xlab="Miles Per Gallon",
        main="Histogram with Normal Curve")
xfit<-seq(min(x),max(x),length=40)
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="blue", lwd=2)
```

Copy the result.

```
histo <-hist(simple_histo, breaks=10, col="red", xlab="Miles Per Gallon",
            main="Histogram with Normal Curve")
xfit<-seq(min(simple_histo),max(simple_histo),length=40)
yfit<-dnorm(xfit,mean=mean(simple_histo),sd=sd(simple_histo))
yfit <- yfit*diff(histo$mids[1:2])*length(simple_histo)
lines(xfit, yfit, col="blue", lwd=2)
```

4. Open the iris dataset. Create a subset for each species.

a. Write the codes and its result.

```
data_set1<- subset(iris, Species == "setosa")
data_set2<- subset(iris, Species == "versicolor")
data_set3<- subset(iris, Species == "virginica")
data_set1
data_set2
data_set3
```

b. Get the mean for every characteristics of each species using colMeans().

Write the codes and its result.

Example: setosa <- colMeans(setosa[sapply(setosaDF,is.numeric)])

```
setosa <- colMeans(data_set1[sapply(data_set1,is.numeric)])
versicolor <- colMeans(data_set2[sapply(data_set2,is.numeric)])
virginica <- colMeans(data_set3[sapply(data_set3,is.numeric)])
setosa
versicolor
virginica
```

c. Combine all species by using rbind()

The table should be look like this:

```
species <- rbind(setosa, versicolor, virginica)
species
```

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
setosa			
versicolor			
virginica			

d. From the data in 4-c: Create the barplot().

Write the codes and its result.

The barplot should be like this.

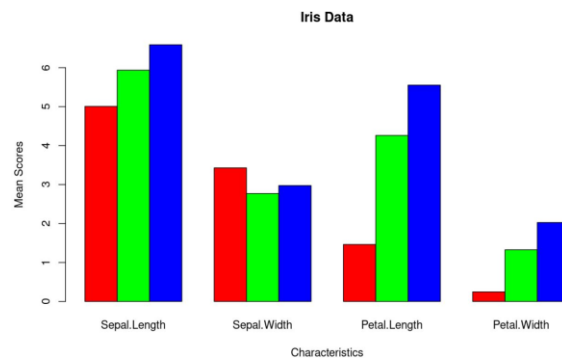


Figure 1: Iris Data using Barplot

```
barplot(species, beside = TRUE, col= c("red", "green", "blue"),
        main= "Iris Data",
        xlab= "Characteristics",
        ylab= "Mean Scores", names.arg= c("Sepal.Length", "Sepal.Width",
                                           "Petal.Length", "Petal.Width"))
```