Sri Lanka Institute of Information Technology

Lab Submission
Lab sheet 03

IT24101672

Arachchi.K.A.T

Probability & Statistics | IT2120

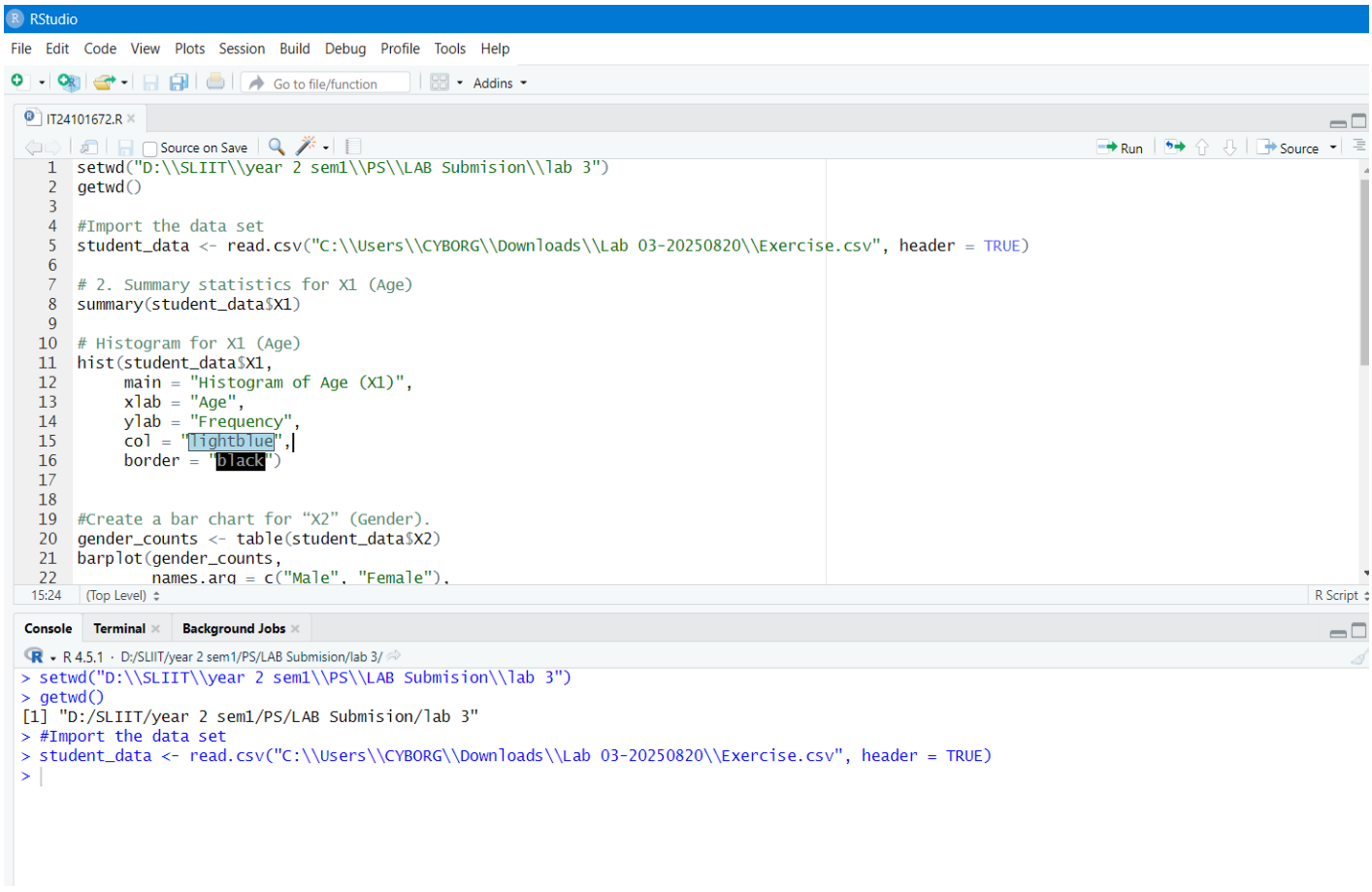
B.Sc. (Hons) in Information Technology

Exercise

Instructions: Create a folder in your desktop with your registration number (Eg: "IT......"). You need to save the R script file and take screenshots of the command prompt with answers and save it in a word document inside the folder. Save both R script file and word document with your registration number (Eg: "IT....."). After you finish the exercise, zip the folder and upload the zip file to the submission link.

1. Import the dataset ('Exercise.csv') into R and store it in a data frame called "student_data".
2. Produce the summary statistics and histogram for the variable "X1" (Age).
3. Create a bar chart and frequency table for "X2" (Gender).
4. How does the age (X1) change according to the accommodation (X3)? Analyze it using a suitable graph and interpret the results. (Note that accommodation has three levels which are type 1, type 2 and type 3)

1)



2)

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | Project: (None)

IT24101672.R

Source on Save | Run | Source

```
1  setwd("D:\\SLIIT\\year 2 sem1\\PS\\LAB Submision\\lab 3")
2  getwd()
3
4  #Import the data set
5  student_data <- read.csv("C:\\Users\\CYBORG\\Downloads\\Lab 03-20250820\\Exercise.csv", header = TRUE
6
7  # 2. Summary statistics for X1 (Age)
8  summary(student_data$X1)
9
10 # Histogram for X1 (Age)
11 hist(student_data$X1,
12      main = "Histogram of Age (X1)",
13      xlab = "Age",
14      ylab = "Frequency",
15      col = "lightblue",
16      border = "black")
17
18
19 #Create a bar chart for "X2" (Gender).
20 gender_counts <- table(student_data$X2)
21 barplot(gender_counts,
22
```

19:1 (Top Level) R Script

Console Terminal Background Jobs

R 4.5.1 · D:/SLIIT/year 2 sem1/PS/LAB Submision/lab 3/

```
> getwd()
[1] "D:/SLIIT/year 2 sem1/PS/LAB Submision/lab 3"
> #Import the data set
> student_data <- read.csv("C:\\Users\\CYBORG\\Downloads\\Lab 03-20250820\\Exercise.csv", header = TRUE)
> # 2. Summary statistics for X1 (Age)
> summary(student_data$X1)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  18.00   19.00   21.00   21.40   23.75   26.00
> # Histogram for X1 (Age)
> hist(student_data$X1,
+      main = "Histogram of Age (X1)",
+      xlab = "Age",
+      ylab = "Frequency",
+      col = "lightblue",
+      border = "black")
>
```

Environment History Connections Tutorial

Import Dataset | 143 MiB | List

R | Global Environment

Data

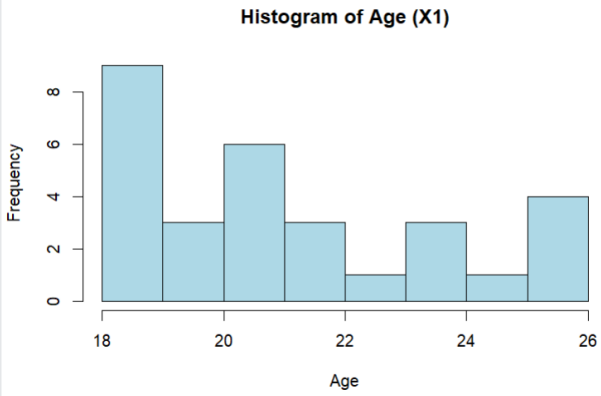
student_data	30 obs. of 3 variables

Values

accommodation_means	num [1:3(1d)] 20.8 21.9 21.5
gender_counts	'table' int [1:2(1d)] 15 15

Files Plots Packages Help Viewer Presentation

Zoom | Export | Publish



3)

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

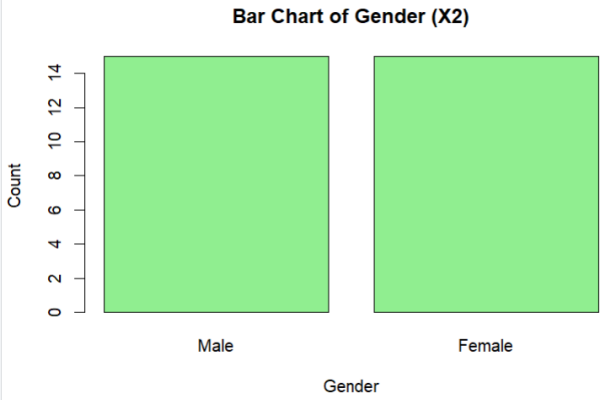
IT24101672.R

```
# 2. Summary statistics for X1 (Age)
summary(student_data$X1)

# Histogram for X1 (Age)
hist(student_data$X1,
     main = "Histogram of Age (X1)",
     xlab = "Age",
     ylab = "Frequency",
     col = "lightblue",
     border = "black")


#Create a bar chart for "X2" (Gender).
gender_counts <- table(student_data$X2)
barplot(gender_counts,
        names.arg = c("Male", "Female"),
        col = "lightgreen",
        main = "Bar Chart of Gender (X2)",
        xlab = "Gender",
        ylab = "Count")
```

Environment | History | Connections | Tutorial

Data	
student_data	30 obs. of 3 variables
Values	
accommodation_means	num [1:3(1d)] 20.8 21.9 21.5
gender_counts	'table' int [1:2(1d)] 15 15

Files | Plots | Packages | Help | Viewer | Presentation



Console | Terminal | Background Jobs

R 4.5.1 · D:/SLIIT/year 2 sem1/PS/LAB Submision/lab 3/

```
> # Histogram for X1 (Age)
> hist(student_data$X1,
+      main = "Histogram of Age (X1)",
+      xlab = "Age",
+      ylab = "Frequency",
+      col = "lightblue",
+      border = "black")
> #Create a bar chart for "X2" (Gender).
> gender_counts <- table(student_data$X2)
> barplot(gender_counts,
+         names.arg = c("Male", "Female"),
+         col = "lightgreen",
+         main = "Bar Chart of Gender (X2)",
+         xlab = "Gender",
+         ylab = "Count")
>
```

4)

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

IT24101672.R

```
27
28 table(student_data$X2)
29
30
31 # Calculate mean age by accommodation type
32 accommodation_means <- tapply(student_data$X1, student_data$X3, mean)
33
34 # Bar chart of average age per accommodation type
35 barplot(accommodation_means,
36         names.arg = c("Type 1", "Type 2", "Type 3"),
37         col = "lightcoral",
38         main = "Average Age by Accommodation Type",
39         xlab = "Accommodation Type",
40         ylab = "Average Age")
41
42 # View numerical results
43 accommodation_means
44
45
```

44:1 (Top Level) R Script

Console Terminal Background Jobs

R 4.5.1 · D:/SLIIT/year 2 sem1/PS/LAB Submision/lab 3/

```
+         ylab = "Count")
> table(student_data$X2)

 1  2
15 15
> # Calculate mean age by accommodation type
> accommodation_means <- tapply(student_data$X1, student_data$X3, mean)
> # Bar chart of average age per accommodation type
> barplot(accommodation_means,
+         names.arg = c("Type 1", "Type 2", "Type 3"),
+         col = "lightcoral",
+         main = "Average Age by Accommodation Type",
+         xlab = "Accommodation Type",
+         ylab = "Average Age")
> # View numerical results
> accommodation_means
       1        2        3
20.77778 21.88889 21.50000
>
```

Environment History Connections Tutorial

Import Dataset 132 MiB List

R Global Environment

Data

student_data	30 obs. of 3 variables

Values

accommodation_means	num [1:3(1d)] 20.8 21.9 21.5
gender_counts	'table' int [1:2(1d)] 15 15

Files Plots Packages Help Viewer Presentation

Zoom Export Publish