

Experiment 1.2

Student Name: Yash Gupta
Branch: BE-CSE
Semester: 6
Subject Name: DM LAB

UID: 20BCS5009
Section/Group: 20BCS_DM-716 B
Date of Performance: 21/02/23
Subject Code: 20CSP_376

AIM :-

Using the pre-functions of r programming languages on the data frame of .arff file.

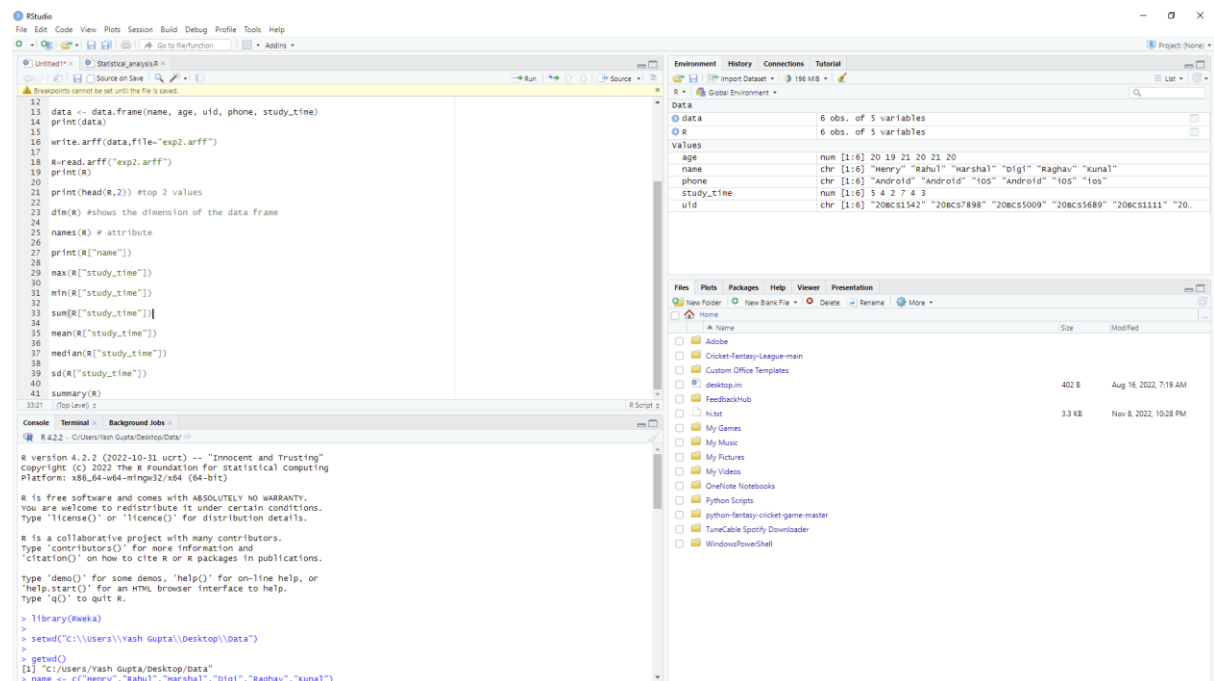
```
library(RWeka)
setwd("C:\\Users\\Yash Gupta\\Desktop\\Data")
getwd()

name <- c("Henry","Rahul","Harshal","Digi","Raghav","Kunal")
age <- c(20,19,21,20,21,20)
phone <- c("Android","Android","iOS","Android","iOS","ios")
uid <-
c("20BCS1542","20BCS7898","20BCS5009","20BCS5689","20BCS1111","20BCS2054")
study_time <- c(5,4,2,7,4,3)
data <- data.frame(name, age, uid, phone, study_time)
print(data)
write.arff(data,file="exp2.arff")

R=read.arff("exp2.arff")
print(R)

print(head(R,2)) #top 2 values
dim(R) #shows the dimension of the data frame
names(R) # attribute
print(R["name"])
max(R["study_time"])
```

```
min(R["study_time"])
sum(R["study_time"])
mean(R["study_time"])
median(R["study_time"])
sd(R["study_time"])
summary(R)
```



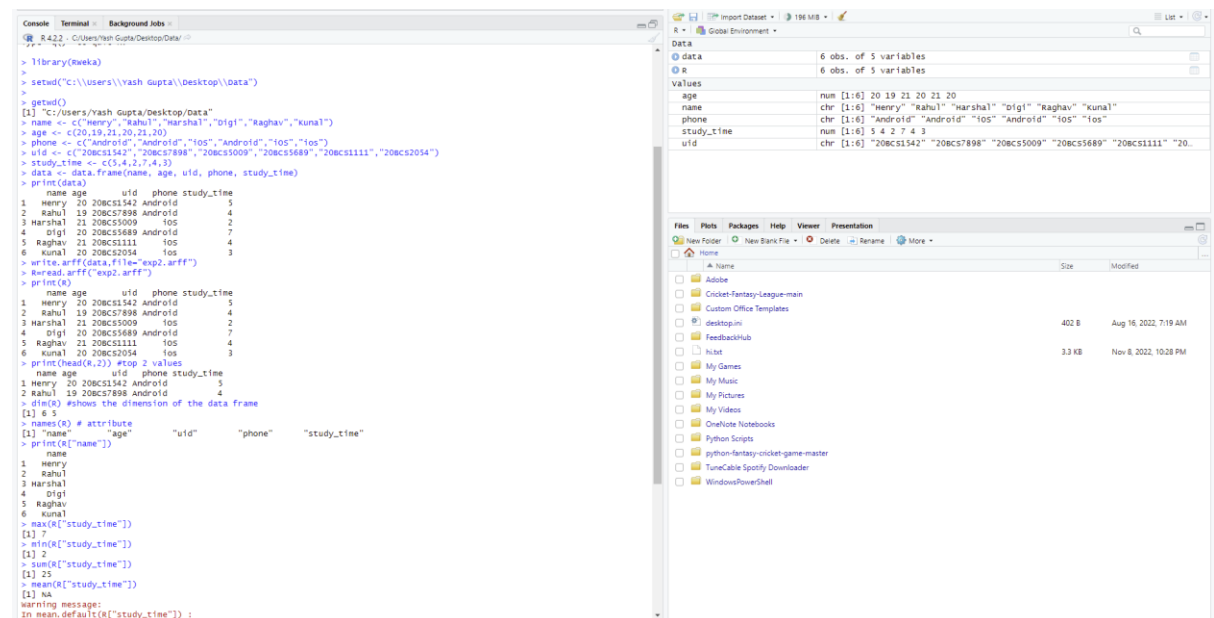
The screenshot shows the RStudio interface with a script editor on the left and the Environment pane on the right. The script editor contains the following R code:

```
12 data <- data.frame(name, age, uid, phone, study_time)
13 print(data)
14
15 write.arff(data,file="exp2.arff")
16
17 R<-read.arff("exp2.arff")
18 print(R)
19
20 print(head(R,2)) #top 2 values
21
22 dim(R) #shows the dimension of the data frame
23
24 names(R) # attribute
25
26 print(R[,"name"])
27
28 max(R["study_time"])
29
30 min(R["study_time"])
31
32 sum(R["study_time"])
33
34 mean(R["study_time"])
35
36 median(R["study_time"])
37
38 sd(R["study_time"])
39
40
41 summary(R)
```

The Environment pane on the right shows the following objects:

Object	Class	Attributes
data	data.frame	6 obs. of 5 variables
R	data.frame	6 obs. of 5 variables

The Files pane at the bottom shows the file explorer with various folders and files.



The screenshot shows the RStudio interface with a script editor on the left and the Environment pane on the right. The script editor contains the following R code:

```
> library(rweka)
> setwd("C:/Users/vash gupta/Desktop/Data")
>
> getwd()
[1] "C:/Users/vash gupta/Desktop/Data"
> name <- c("henry","Rahul","harshal","Digi","Raghav","kunal")
> age <- c(20,19,21,20,21,20)
> phone <- c("Android","Android","ios","Android","ios","ios")
> uid <- c("208cs1542","208cs7898","208cs5009","208cs5689","208cs1111","208cs2054")
> study_time <- c(5,4,2,7,4,3)
> data <- data.frame(name, age, uid, phone, study_time)
> print(data)
  name age uid phone study_time
1 henry 20 208cs1542 Android      5
2 Rahul 19 208cs7898 Android      4
3 Harshal 21 208cs5009 ios        2
4 Digi 20 208cs5689 Android      7
5 Raghav 21 208cs1111 ios         4
6 Kunal 20 208cs2054 ios         3
> write.arff(data,file="exp2.arff")
> R<-read.arff("exp2.arff")
> print(R)
  name age uid phone study_time
1 henry 20 208cs1542 Android      5
2 Rahul 19 208cs7898 Android      4
3 Harshal 21 208cs5009 ios        2
4 Digi 20 208cs5689 Android      7
5 Raghav 21 208cs1111 ios         4
6 Kunal 20 208cs2054 ios         3
> print(head(R,2)) #top 2 values
  name age uid phone study_time
1 henry 20 208cs1542 Android      5
2 Rahul 19 208cs7898 Android      4
> dim(R) #shows the dimension of the data frame
[1] 6 5
> names(R) # attribute
[1] "name" "age" "uid" "phone" "study_time"
> print(R[,"name"])
  name
1 henry
2 Rahul
3 Harshal
4 Digi
5 Raghav
6 Kunal
> max(R["study_time"])
[1] 7
> min(R["study_time"])
[1] 2
> sum(R["study_time"])
[1] 25
> mean(R["study_time"])
[1] NA
Warning message:
In mean.default(R["study_time"]) :
```

The Environment pane on the right shows the following objects:

Object	Class	Attributes
data	data.frame	6 obs. of 5 variables
R	data.frame	6 obs. of 5 variables

The Files pane at the bottom shows the file explorer with various folders and files.

```
exp2 - Notepad
File Edit Format View Help
@relation R_data_frame

@attribute name string
@attribute age numeric
@attribute uid string
@attribute phone string
@attribute study_time numeric

@data
Henry,20,20BCS1542,Android,5
Rahul,19,20BCS7898,Android,4
Harshal,21,20BCS5009,iOS,2
Digi,20,20BCS5689,Android,7
Raghav,21,20BCS1111,iOS,4
Kunal,20,20BCS2054,ios,3
```

Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Open file...

Open URL...

Open DB...

Generate...

Undo

Edit...

Save...

Filter

Choose

None

Apply

Stop

Current relation

Relation: R_data_frame

Instances: 6

Attributes: 5

Sum of weights: 6

Attributes

All

None

Invert

Pattern

No.	Name
1	<input checked="" type="checkbox"/> name
2	<input checked="" type="checkbox"/> age
3	<input checked="" type="checkbox"/> uid
4	<input checked="" type="checkbox"/> phone
5	<input checked="" type="checkbox"/> study_time

Remove

Selected attribute

Name: study_time

Missing: 0 (0%)

Distinct: 5

Type: Numeric

Unique: 4 (67%)

Statistic	Value
Minimum	2
Maximum	7
Mean	4.167
StdDev	1.722

Class: study_time (Num)

Visualize All

4

2

2

4.5

7

Status

OK

Log

x 0