Experiment1.2

Student Name: Himanshu
UID: 20BCS7944
Branch: CSE
Section: 905/A
Semester: 6
Date of Performance: 22/02/2023
Subject Name: Data Mining Lab
Subject Code: 20CSP-376

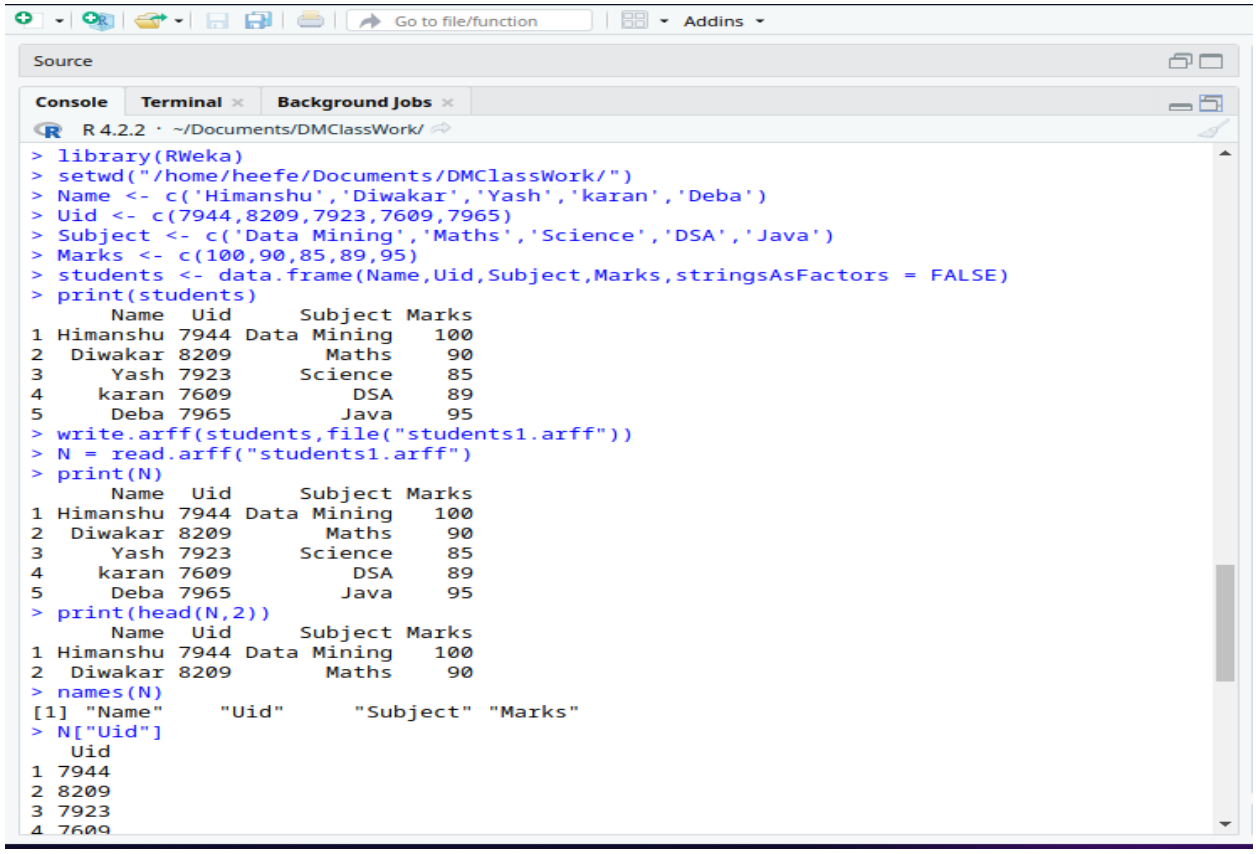
1) Aim:

To perform the statistical analysis of data.

2) Objective:

To analyze the statistical data and perform various mathematical operations.

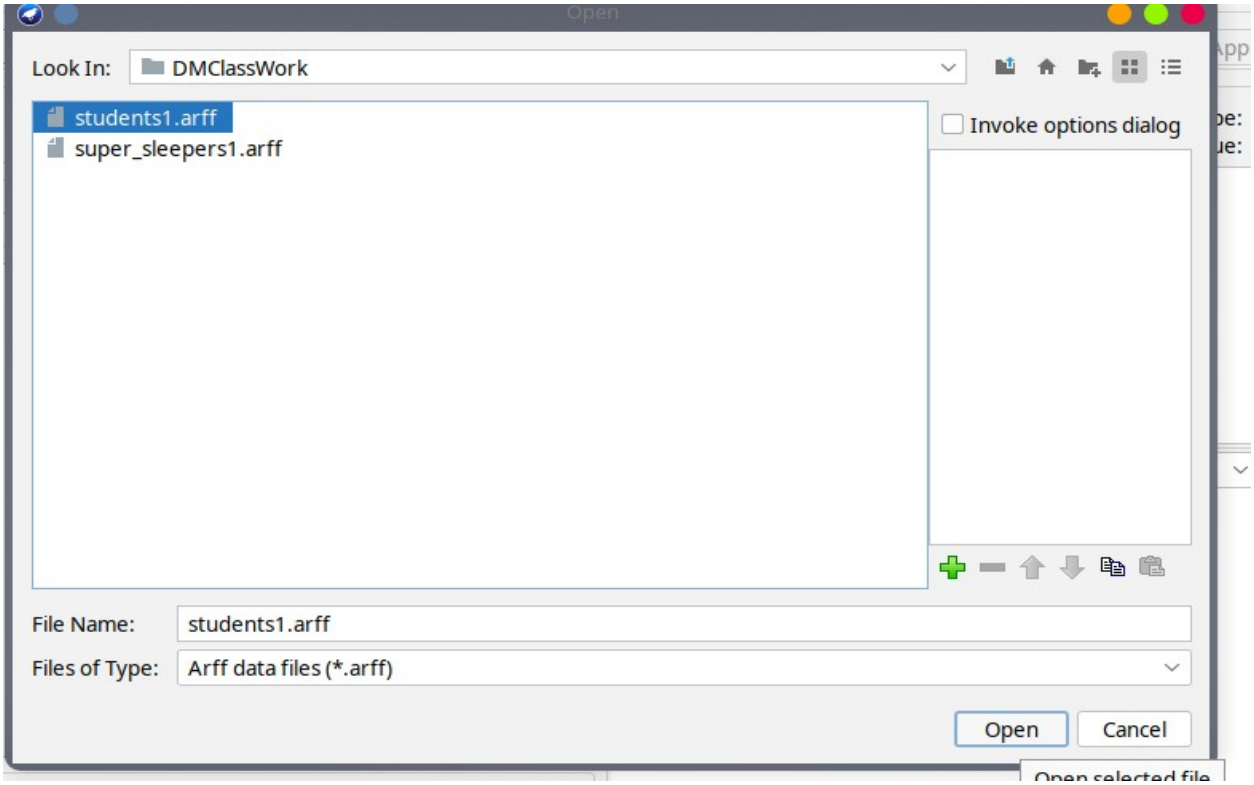
3) Code and Output:

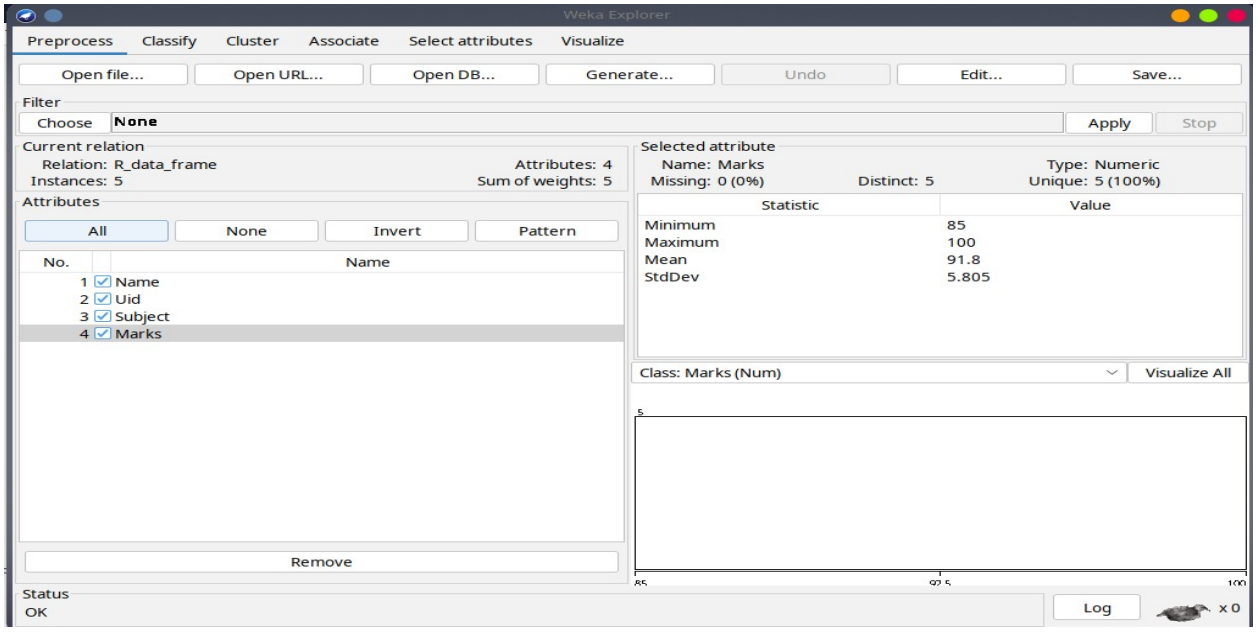


Console | Terminal | Background Jobs

R 4.2.2 · ~/Documents/DMClassWork/

```
> names(N)
[1] "Name"    "Uid"     "Subject" "Marks"
> N["Uid"]
   Uid
1 7944
2 8209
3 7923
4 7609
5 7965
> min(Marks)
[1] 85
> sum(Marks)
[1] 459
> mean(Marks)
[1] 91.8
> median(sort(Marks))
[1] 90
> sd(Marks)
[1] 5.80517
> summary(students)
     Name                Uid          Subject             Marks
 Length:5           Min.   :7609   Length:5           Min.   : 85.0
 Class :character   1st Qu.:7923   Class :character   1st Qu.: 89.0
 Mode  :character   Median :7944   Mode  :character   Median : 90.0
                    Mean   :7930                      Mean   : 91.8
                    3rd Qu.:7965                      3rd Qu.: 95.0
                    Max.   :8209                      Max.   :100.0
> summary(N)
     Name                Uid          Subject             Marks
 Length:5           Min.   :7609   Length:5           Min.   : 85.0
 Class :character   1st Qu.:7923   Class :character   1st Qu.: 89.0
 Mode  :character   Median :7944   Mode  :character   Median : 90.0
                    Mean   :7930                      Mean   : 91.8
                    3rd Qu.:7965                      3rd Qu.: 95.0
                    Max.   :8209                      Max.   :100.0
```


Open
Look In:
DMClassWork
students1.arff
super_sleepers1.arff
Invoke options dialog
File Name:
students1.arff
Files of Type:
Arff data files (*.arff)
Open
Cancel


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: 5
Attributes: 4
Sum of weights: 5
Selected attribute
Name: Marks
Missing: 0 (0%)
Distinct: 5
Type: Numeric
Unique: 5 (100%)
Attributes
All
None
Invert
Pattern
No.
Name
1 Name
2 Uid
3 Subject
4 Marks
Statistic
Value
Minimum 85
Maximum 100
Mean 91.8
StdDev 5.805
Class: Marks (Num)
Visualize All
Remove
Status
OK
Log
x 0

Code -

```
library(RWeka)

setwd("/home/heefe/Documents/DMClassWork/")

Name <- c('Himanshu','Diwakar','Yash','karan','Deba')

Uid <- c(7944,8209,7923,7609,7965)

Subject <- c('Data Mining','Maths','Science','DSA','Java')

Marks <- c(100,90,85,89,95)

students <- data.frame(Name,Uid,Subject,Marks,stringsAsFactors = FALSE)

print(students)

write.arff(students,file("students1.arff"))

N = read.arff("students1.arff")

print(N)

names(N)

N["Uid"]

min(Marks)

sum(Marks)

mean(Marks)

median(sort(Marks))

sd(Marks)

summary(N)
```