DEPARTMENT OF ACADEMIC AFFAIRS

Discover. Learn. Empower.

NAAC GRADE A+ ACCREDITED UNIVERSITY

UNIVERSITY INSTITUTE OF ENGINEERING

Department of Computer Science & Engineering

Subject Name: Data Mining Lab

Subject Code: 20CSP-376

Submitted to:

Er. Himanshi

Submitted by:

Name: Rithik raj

UID: 20BCS5436

Section: DM-611

Group: B

Experiment-1.2

Student Name:Rithik raj **UID:** 20BCS5436
Branch: BE CSE **Section/Group:** 611/B
Semester: 6th **Date of Performance:** 24/02/2023
Subject Name: DM Lab **Subject Code:** 20CSP-376

1. **Aim/Overview of the practical:**
   To perform the statistical analysis of data.

2. **Apparatus / Simulator Used:**
   - Windows 7 or above
   - R Studio

3. **Objective:**
   - Represent the reading of file using R Studio.
   - Displaying the pattern on Weka Tool.
   - Find mean, median and standard deviation of particular columns.

4. **Script and Output:**

5. ***Code:***

```
library(Rweka)
N=read.arff("Student_frame.arfr')
```

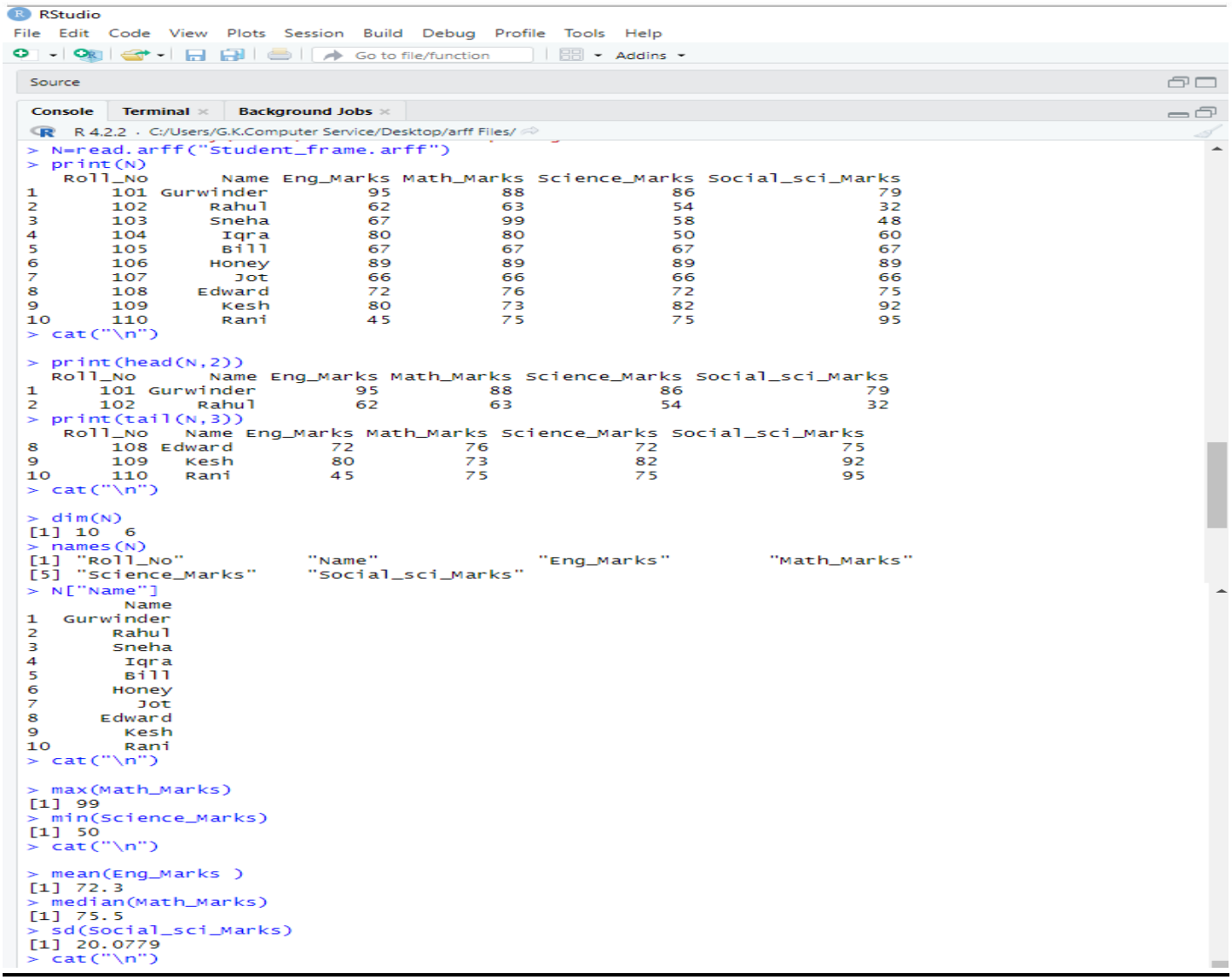
```
print(N)
cat("\n")                          #used to skip lines
 print(head(N,2))                  #used to print first 2 rows
print(tail(N,3))                   #used to print first 3rows
cat("\n")
dim(N)                             #used for finding dimensions
names(N)
N["Name")
cat("\n")
 max(Math_Marks)                   #maximum from column
min(Science_Marks)                 #minimum from column

cat("\n")
mean(Eng_Marks )                           # mean
Median lath_Marks)                         # median
(cial_sci_Marks)                           # standard deviation

cat("\n")

summary(N)
```

Output:


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Source
Console
Terminal
Background Jobs
R 4.2.2 · C:/Users/G.K.Computer Service/Desktop/arff Files/
> N=read.arff("Student_frame.arff")
> print(N)
Roll_No Name Eng_Marks Math_Marks Science_Marks Social_sci_Marks
1 101 Gurwinder 95 88 86 79
2 102 Rahul 62 63 54 32
3 103 Sneha 67 99 58 48
4 104 Iqra 80 80 50 60
5 105 Bill 67 67 67 67
6 106 Honey 89 89 89 89
7 107 Jot 66 66 66 66
8 108 Edward 72 76 72 75
9 109 Kesh 80 73 82 92
10 110 Rani 45 75 75 95
> cat("\n")
> print(head(N,2))
Roll_No Name Eng_Marks Math_Marks Science_Marks Social_sci_Marks
1 101 Gurwinder 95 88 86 79
2 102 Rahul 62 63 54 32
> print(tail(N,3))
Roll_No Name Eng_Marks Math_Marks Science_Marks Social_sci_Marks
8 108 Edward 72 76 72 75
9 109 Kesh 80 73 82 92
10 110 Rani 45 75 75 95
> cat("\n")
> dim(N)
[1] 10 6
> names(N)
[1] "Roll_No" "Name" "Eng_Marks" "Math_Marks"
[5] "Science_Marks" "Social_sci_Marks"
> N["Name"]
Name
1 Gurwinder
2 Rahul
3 Sneha
4 Iqra
5 Bill
6 Honey
7 Jot
8 Edward
9 Kesh
10 Rani
> cat("\n")
> max(Math_Marks)
[1] 99
> min(Science_Marks)
[1] 50
> cat("\n")
> mean(Eng_Marks)
[1] 72.3
> median(Math_Marks)
[1] 75.5
> sd(Social_sci_Marks)
[1] 20.0779
> cat("\n")

```
> summary(N)
    Roll_No          Name              Eng_Marks       Math_Marks    Science_Marks
 Min.   :101.0   Length:10          Min.   :45.00   Min.   :63.0   Min.   :50.00
 1st Qu.:103.2   Class :character   1st Qu.:66.25   1st Qu.:68.5   1st Qu.:60.00
 Median :105.5   Mode  :character   Median :69.50   Median :75.5   Median :69.50
 Mean   :105.5                      Mean   :72.30   Mean   :77.6   Mean   :69.90
 3rd Qu.:107.8                      3rd Qu.:80.00   3rd Qu.:86.0   3rd Qu.:80.25
 Max.   :110.0                      Max.   :95.00   Max.   :99.0   Max.   :89.00
 Social_sci_Marks
 Min.   :32.0
 1st Qu.:61.5
 Median :71.0
 Mean   :70.3
 3rd Qu.:86.5
 Max.   :95.0
>
```

Evaluation Grid (To be created as per the SOP and Assessment guidelines by the faculty):

Sr. No.	Parameters	Marks Obtained	Maximum Marks
1.			
2.			
3.			
4.			

Conclusion:-

I have successfully completed this experiment.