

Experiment 1.2

Student Name: Shubham Shrivastava

Branch: CSE

Semester: 6

Subject Name: Data Mining Lab

UID: 20BCS5472

Section/Group: 701 B

Date of Performance: 20/2/23

Subject Code: 20CSP376

- 1. Aim: Statistical analysis of data.**
- 2. Objective: To perform statistical analysis of data.**
- 3. Script and Output:**

```
library(RWeka)
#setting the working directory
setwd("C:\\Users\\dell\\OneDrive\\Desktop\\DataMining")
#checking the working directory
getwd()

#reading arff file
N = read.arff("diabetes.arff")
print(N)

#printing first and last 2 rows
head(N,2)
tail(N,2)

#printing dimension
dim(N)
```



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Discover. Learn. Empower.

#printing column names

```
names(N)
```

#printing maximum and minimum age

```
max(N["age"])
```

```
min(N["age"])
```

#mean age

```
mean(N$age)
```

#median age

```
median(sort(N$age))
```

#standard deviation

```
sd(N$age)
```

#summary

```
summary(N)
```



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Discover. Learn. Empower.

The screenshot displays the RStudio interface on a macOS desktop. The main editor window shows an R script with the following code:

```
15 b <- mean(nameq$age)
16 c <- median(nameq$age)
17 d <- mode(nameq$age)
18 print(b)
19 print(c)
20 print(d)
21 print(head(nameq))
22 print(tail(nameq))
23 print(dim(nameq))
```

The console output shows the results of these operations:

```
R 4.2.2 ~ /
> b <- mean(nameq$age)
> c <- median(nameq$age)
> d <- mode(nameq$age)
> print(b)
[1] 24.8
> print(c)
[1] 21
> print(d)
[1] "numeric"
> print(head(nameq))
  name_id names_sr age_no
1      1    snehil    20
2      2      Anas    21
3      3   Sanjay    21
4      4     Abhas    22
5      5 Roop Lal    40
> print(tail(nameq))
  name_id names_sr age_no
1      1    snehil    20
2      2      Anas    21
3      3   Sanjay    21
4      4     Abhas    22
5      5 Roop Lal    40
> print(dim(nameq))
[1] 5 3
>
```

The Environment pane on the right shows the 'nameq' data frame with 5 observations and 3 variables (b, c, d). The Packages pane shows the installed and available packages, including the System Library and various data analysis packages.