DEPARTMENT OF ACADEMIC AFFAIRS
Discover. Learn. Empower.

NAAC GRADE A+ ACCREDITED UNIVERSITY

EXPERIMENT -1

NAME- Jatin	UID- 20BCS5951
BRANCH- BE- CSE	SECTION- DM 605 B
SEMESTER- 6th	DATE- 18/02/2023
SUBJECT- Data Mining Lab	SUBJECT CODE- 20CSP-376

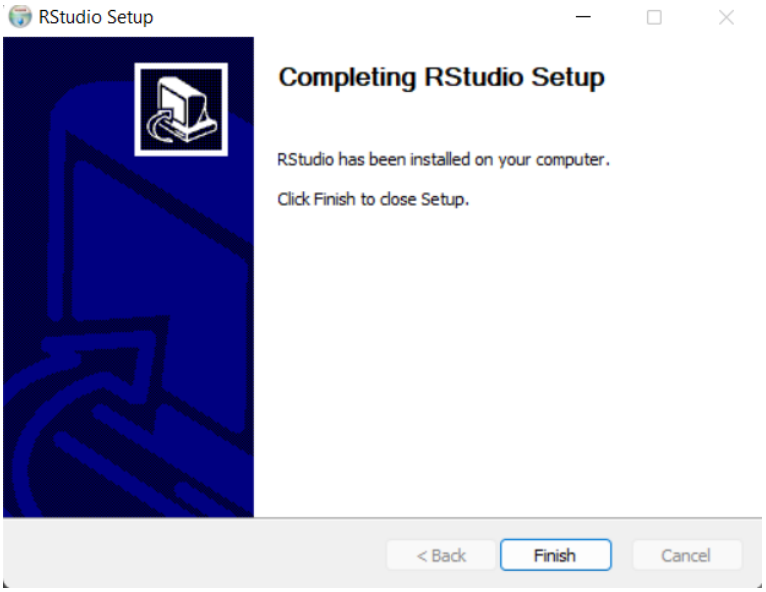
- **Aim** . Demonstration of preprocessing on .arff file using student data .arff.

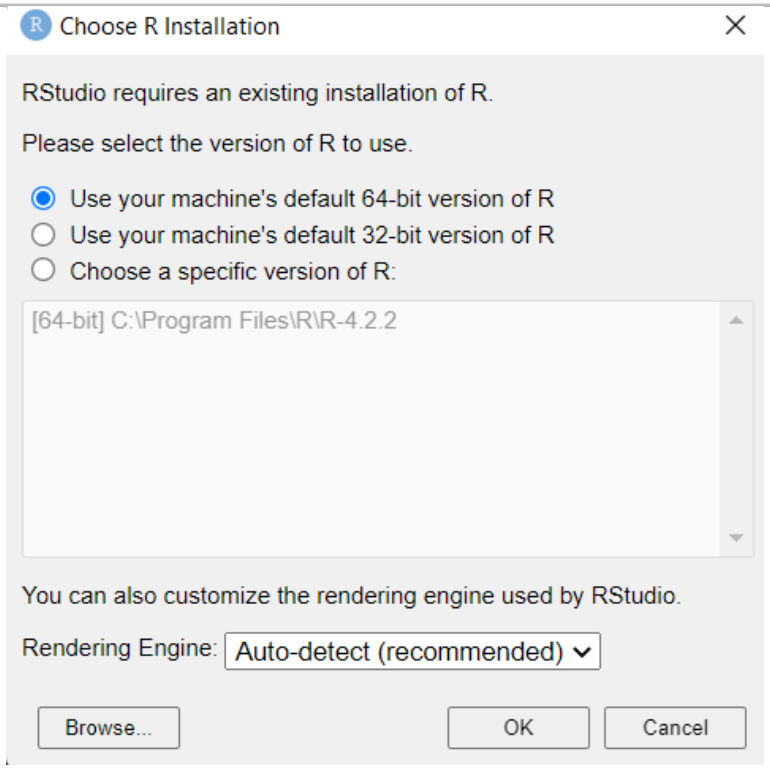
- **Objective:**
  Learning about ARFF files and how to create ARFF File (Attribute relation File Format)

- **Installation:**
  - To install R, go to cran.r-project.org
  - Depending on your operating system, click Download R for (your operating system).
  - Click on install R for the first time.
  - Click Download R for Windows. Open the downloaded file.
  - Select the language you would like to use during the installation. Then click OK.
  - Click Next.
  - Select where you would like R to be installed. It will default to your Program Files on your C Drive. Click Next.
  - You can then choose which installation you would like.
  - (Optional) If your computer is a 64-bit, you can choose the 64-bit User Installation. Then click Next.
  - Then specify if you want to customized your startup or just use the defaults. Then click Next.
  - Then you can choose the folder that you want R to be saved within or the default if the R folder that was created. Once you have finished, click Next.

- You can then select additional shortcuts if you would like. Click Next.
- Click Finish.
- Next, download RStudio. Go to www.rstudio.com
- Click Download RStudio.
- The RStudio installation wizard will pop-up. Click Next and go through the installation steps.
- Congratulations! You have now installed R and RStudio.

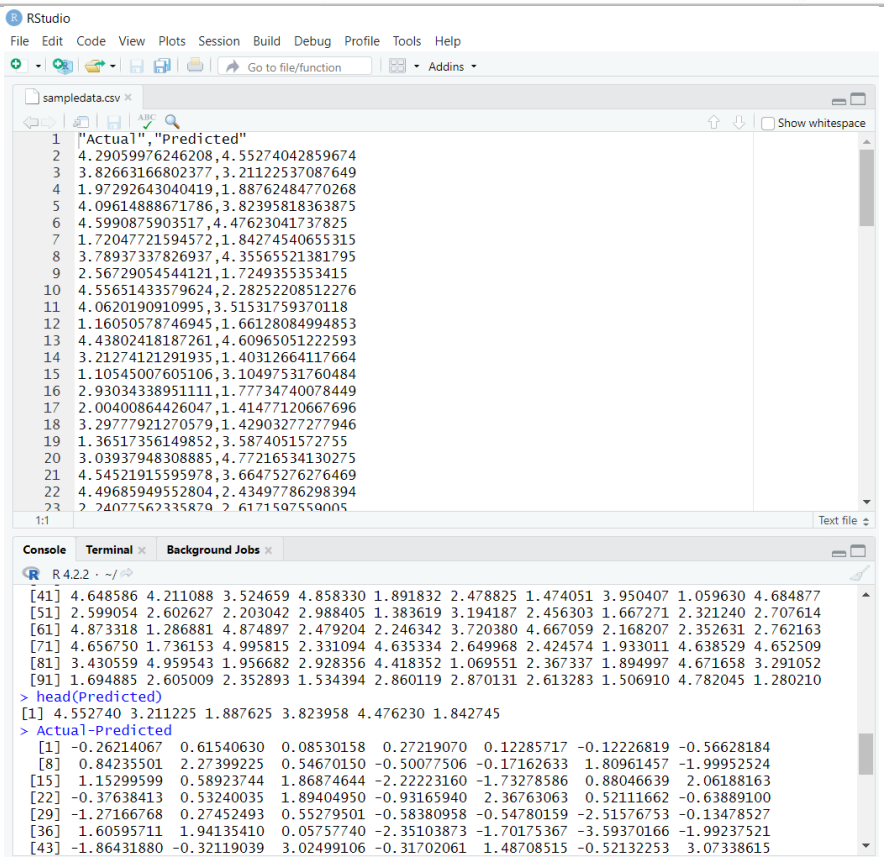
- **Installation Screenshots**





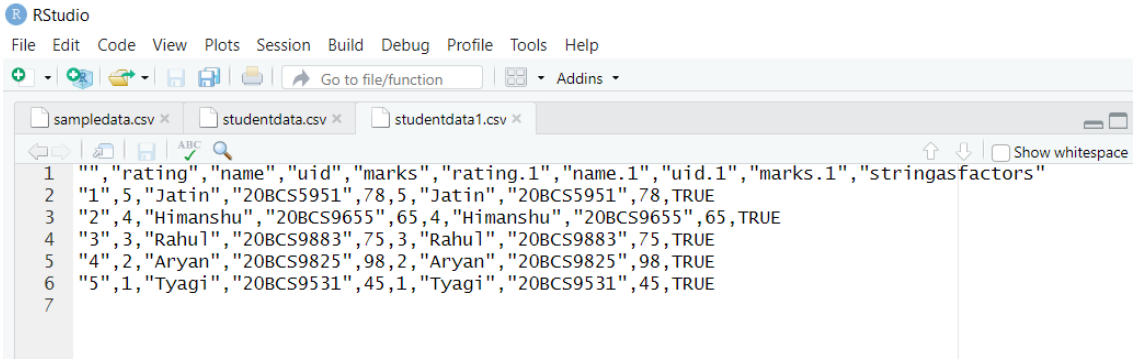
- **Script and Output:**

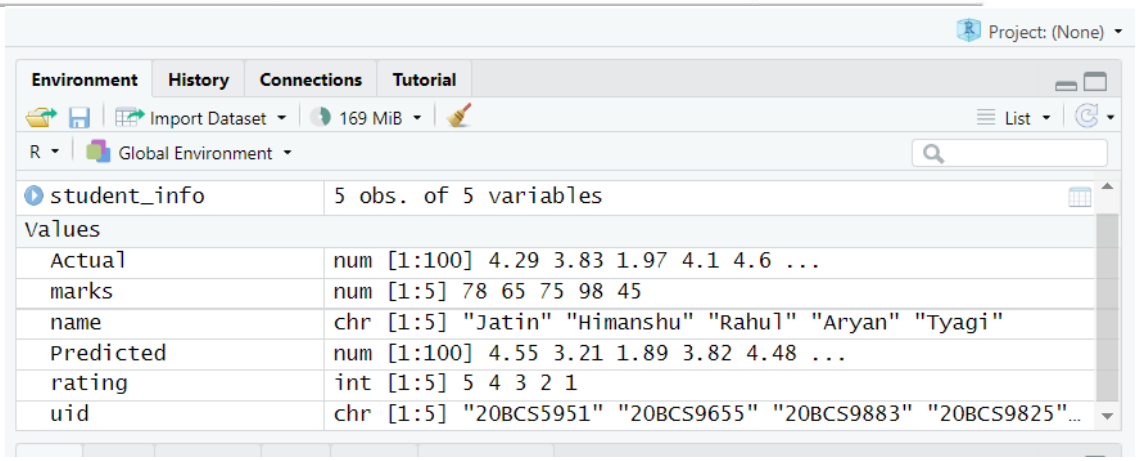
```
exp (-200)
log (100,base-10)
Actual<-runif (100, min-1, max-5)
Actual
head (Actual)
head (Actual, 10)
head (Actual, 20)
Predicted<-runif (100,min-1, max-5)
Predicted
head (predicted)
Actual-Predicted
```



- **Preprocessing on .arff file using student data .arff**

rating<-5:1
name<-c('Jatin','Himanshu','Rahul','Aryan','Tyagi')
uid<-c('20BCS5951','20BCS9655','20BCS9883','20BCS9825','20BCS9531')
marks<-c(78,65,75,98,45)
student_info<-data.frame(rating,name,uid,marks,stringasfactors=TRUE)




Project: (None)
Environment
History
Connections
Tutorial
Import Dataset
169 MiB
List
R
Global Environment
student_info
5 obs. of 5 variables
Values
Actual
num [1:100] 4.29 3.83 1.97 4.1 4.6 ...
marks
num [1:5] 78 65 75 98 45
name
chr [1:5] "Jatin" "Himanshu" "Rahul" "Aryan" "Tyagi"
Predicted
num [1:100] 4.55 3.21 1.89 3.82 4.48 ...
rating
int [1:5] 5 4 3 2 1
uid
chr [1:5] "20BCS5951" "20BCS9655" "20BCS9883" "20BCS9825"...