



## **Experiment1.1**

**Student Name:** Mayank Kumar

**Branch:** BE-CSE

**Semester:** 6<sup>th</sup>

**Subject Name:** Data Mining Lab

**UID:** 20BCS1353

**Section/Group:** 20BCS\_DM\_705 A

**Date of Performance:**

**Subject Code:** 20CSP-376

**Aim:** Demonstration of pre-processing on .arff file using student data .arff.

**Objective:** we will be learning about ARFF files and how to create ARFF file

**Script and Output:**

### **Code:**

```
library(RWeka)
setwd("F:\\DataMiningExp1")
getwd()
stdno <- 1:4
stdname <- c("Shubham","Mayank","Ashutosh","Sandeep")
stdclass <- c("20BCSDM705","20BCSNT603","20BCS705","20BCSNT603")
stdmarks <- c(76,88,89,90)
student <- data.frame(stdno,stdname,stdclass,stdmarks,stringsAsFactors = FALSE)
print(student)
print(class(student))
print(str(student))
print(is.factor(stdno))
class(stdno)
```



# DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Discover. Learn. Empower.

```
write.arff(student,file="student.arff")
```

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains the R script code for creating a data frame and writing it to an ARFF file.
- Console:** Displays the output of the executed code, including the data frame structure and the file path.
- Environment Pane:** Shows the objects in the global environment, including 'student' and 'stdno'.

```
1 library(Rweka)
2 setwd("F:\\DataMiningExp1")
3 getwd()
4
5 stdno <- 1:4
6 stdname <- c("Shubham", "Ankur", "Ashutosh", "Sandeep")
7 stdclass <- c("20BCSDM705", "20BCSNT603", "20BCS705", "20BCSNT603")
8 stdmarks <- c(76, 88, 89, 90)
9
10 student <- data.frame(stdno, stdname, stdclass, stdmarks, stringsAsFactors = FALSE)
11
12 print(student)
13 print(class(student))
14 print(str(student))
15 print(is.factor(stdno))
16 class(stdno)
17 write.arff(student, file="student.arff")
18
```

**Console Output:**

```
> library(Rweka)
> setwd("F:\\DataMiningExp1")
> getwd()
[1] "F:/DataMiningExp1"
>
> stdno <- 1:4
> stdname <- c("Shubham", "Ankur", "Ashutosh", "Sandeep")
> stdclass <- c("20BCSDM705", "20BCSNT603", "20BCS705", "20BCSNT603")
> stdmarks <- c(76, 88, 89, 90)
>
> student <- data.frame(stdno, stdname, stdclass, stdmarks, stringsAsFactors = FALSE)
> print(student)
  stdno stdname stdclass stdmarks
1     1 Shubham 20BCSDM705      76
2     2   Ankur 20BCSNT603      88
3     3 Ashutosh 20BCS705      89
4     4  Sandeep 20BCSNT603      90
>
> print(class(student))
[1] "data.frame"
> print(str(student))
'data.frame':   4 obs. of  4 variables:
 stdno: num 1 2 3 4
 stdname: chr "Shubham" "Ankur" "Ashutosh" "Sandeep"
 stdclass: chr "20BCSDM705" "20BCSNT603" "20BCS705" "20BCSNT603"
 stdmarks: num 76 88 89 90
> print(is.factor(stdno))
[1] FALSE
> class(stdno)
[1] "numeric"
>
```

**Environment Pane:**

Object	Class	Attributes
student	data.frame	4 obs. of 4 variables
stdno	numeric	4 obs. of 1 variable
stdname	character	4 obs. of 1 variable
stdclass	character	4 obs. of 1 variable
stdmarks	numeric	4 obs. of 1 variable

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains the R script code for creating a data frame and writing it to an ARFF file.

```
1 library(Rweka)
2 setwd("F:\\DataMiningExp1")
3 getwd()
4
5 stdno <- 1:4
6 stdname <- c("Shubham", "Ankur", "Ashutosh", "Sandeep")
7 stdclass <- c("20BCSDM705", "20BCSNT603", "20BCS705", "20BCSNT603")
8 stdmarks <- c(76, 88, 89, 90)
9
10 student <- data.frame(stdno, stdname, stdclass, stdmarks, stringsAsFactors = FALSE)
11
12 print(student)
13 print(class(student))
14 print(str(student))
15 print(is.factor(stdno))
16 class(stdno)
17 write.arff(student, file="student.arff")
18
```



# DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Discover. Learn. Empower.

Output:

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
+ - [Icons] Go to file/function [Icons] Addins -
Source
Console Terminal x Background Jobs x
R 4.2.2 · F:/DataMiningExp1/ ↗
> library(RWeka)
>
> setwd("F:\\DataMiningExp1")
>
> getwd()
[1] "F:/DataMiningExp1"
>
> stdno <- 1:4
> stdname <- c("Shubham","Ankul","Ashutosh","Sandeep")
> stdclass <- c("20BCSDM705","20BCSNT603","20BCS705","20BCSNT603")
> stdmarks <- c(76,88,89,90)
>
> student <- data.frame(stdno,stdname,stdclass,stdmarks,stringsAsFactors = FALSE)
> print(student)
  stdno stdname  stdclass stdmarks
1     1  Shubham 20BCSDM705      76
2     2   Ankul 20BCSNT603      88
3     3 Ashutosh  20BCS705      89
4     4  Sandeep 20BCSNT603      90
>
> print(class(student))
[1] "data.frame"
>
> print(str(student))
'data.frame':  4 obs. of  4 variables:
 $ stdno  : int  1 2 3 4
 $ stdname : chr  "Shubham" "Ankul" "Ashutosh" "Sandeep"
 $ stdclass: chr  "20BCSDM705" "20BCSNT603" "20BCS705" "20BCSNT603"
 $ stdmarks: num  76 88 89 90
NULL
>
> print(is.factor(stdno))
[1] FALSE
>
> class(stdno)
[1] "integer"
>
> write.arff(student,file="student.arff")
> |
```