

**UNIVERSITY INSTITUTE OF ENGINEERING**

**Department of Computer Science & Engineering**

|  |  |
| --- | --- |
| **Submitted By:                                                                          Submitted To:**  Rahul Kumar (20BCS7081) Er. Nitsha (E13137) | |
| **Subject Name** | Data Mining Lab |
| **Subject Code** | 20CSP\_376 |
| **Branch** | CSE |
| **Semester** | 6th |

**LAB INDEX**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr.No** | **Program** | **Date** | **Evaluation** | | | | **Sign** |
| **LW(12)** | **VV(8)** | **FW(10)** | **Total (30)** |
| 1.1 | Demonstration of preprocessing on .arff file using student data .arff. |  |  |  |  |  |  |
| 1.2 | To perform the statistical analysis of data |  |  |  |  |  |  |
| 1.3 | Demonstration of association rule mining using Aprior algorithm on supermarket data. |  |  |  |  |  |  |
| 1.4 | Demonstration of FP Growth algorithm on supermarket data.  tion of preprocessing on .arff file using student data .arff. |  |  |  |  |  |  |
| 2.1 | To perform the classification by decision tree induction using WEKA tools. |  |  |  |  |  |  |
| 2.2 | To perform classification using Bayesian classification algorithm using R |  |  |  |  |  |  |
| 2.3 | To perform the cluster analysis by k-means method using R. |  |  |  |  |  |  |
| 3.1 | To perform the hierarchical clustering using R programming. |  |  |  |  |  |  |
| 3.2 | Study of Regression Analysis using R programming. |  |  |  |  |  |  |
| 3.3 | Outlier detection using R programming. |  |  |  |  |  |  |

**Experiment 1.1**

**Student Name: Rahul Kumar UID: 20BCS7081**

**Branch: BE-CSE Section/Group:20BCS\_DM-716 B**

**Semester: 6 Date of Performance: 22 Feb 2023**

**Subject Name: DM LAB Subject Code: 20CSP\_376**

**AIM :-**

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

**Code :-**

library(RWeka)

setwd("C:\\Users\\RAHUL KUMAR\\Desktop\\r")

getwd()

Name <- c("Rahul","rohit","Mohan","sohan","Ajay")

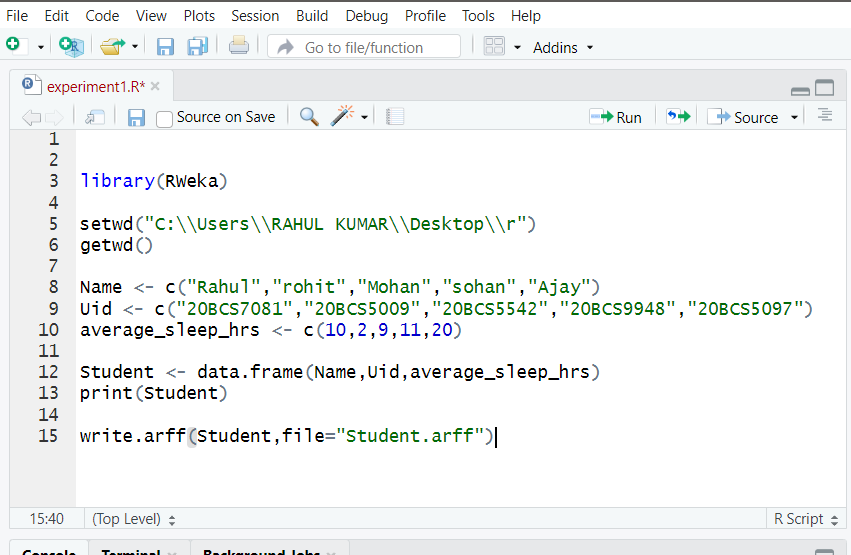
Uid <- c("20BCS7081","20BCS5009","20BCS5542","20BCS9948","20BCS5097")

average\_sleep\_hrs <- c(10,2,9,11,20)

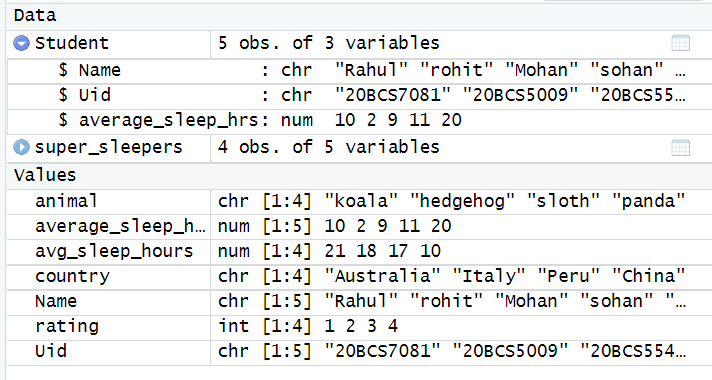
Student <- data.frame(Name,Uid,average\_sleep\_hrs)

print(Student)

write.arff(Student,file="Student.arff")

****

**Data:-**



**Output :-**

