



UNIVERSITY INSTITUTE OF ENGINEERING

Department of Computer Science & Engineering

Subject Name: Data Mining Lab

Subject Code: 20CSP 376

Submitted to:

Faculty name:

Anupama Jamwal
(E14665)

Submitted by:

Name: SUMIT KUMAR

UID: 20BCS8226

Section: DM 720 A

EXPERIMENT: 1

Student Name: SUMIT KUMAR
Branch: CSE
Semester: 06
Subject Name: Data Mining Lab

UID: 20BCS8226
Section / Group: -20BCS-DM-720-A
Date of Performance: 20/2/23
Subject Code: 20CSP-376

Question 1:

Demonstrations of preprocessing on .arff file using student data .arff.

Solution / Code:

```
@relation "student"
@attribute id numeric
@attribute f_name string
@attribute l_name string
@attribute contact_num numeric
@attribute dept {CSE,MECHNICAL,BUSINESS,ARTS}
@attribute DOB date dd-mm-yyyy
@attribute city string

@data
1,naman,kumar,1234556678,CSE,02-08-2000,Mohali
2,Brijendra,Tripathi,5555555,CSE,04-05-2001,delhi
3,kishan,kumar,1214556678,ARTS,02-11-2001,kharar
4,Anand,Kumar,5234556678,BUSINESS,03-05-2000,Chandigarh
4,Aman,Kumar,5234556678,MECHNICAL,03-05-2000,Ludhiana
```

Output:

Relation: student							
No.	1: id Numeric	2: f_name String	3: l_name String	4: contact_num Numeric	5: dept Nominal	6: DOB Date	7: city String
1	1.0	naman	kumar	1.234556678E9	CSE	02-08...	Mohali
2	2.0	Brijendra	Tripathi	5555555.0	CSE	04-05...	delhi
3	3.0	kishan	kumar	1.214556678E9	ARTS	02-11...	kharar
4	4.0	Anand	Kumar	5.234556678E9	BUSIN...	03-05...	Chan...
5	4.0	Aman	Kumar	5.234556678E9	MEC...	03-05...	Ludh...

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Open file...

Open URL...

Open DB...

Generate...

Undo

Edit...

Save...

Filter

Choose **None**

Apply

Stop

Current relation

Relation: student

Instances: 5

Attributes: 7

Sum of weights: 5

Attributes

All

None

Invert

Pattern

No.	Name
1	<input checked="" type="checkbox"/> id
2	<input type="checkbox"/> f_name
3	<input type="checkbox"/> l_name
4	<input type="checkbox"/> contact_num
5	<input type="checkbox"/> dept
6	<input type="checkbox"/> DOB
7	<input type="checkbox"/> city

Remove

Selected attribute

Name: id

Missing: 0 (0%)

Distinct: 4

Type: Numeric

Unique: 3 (60%)

Statistic	Value
Minimum	1
Maximum	4
Mean	2.8
StdDev	1.304

Class: city (Str)

Visualize All

\$