DATA MINING LAB

ASSIGNMENT-1

Ayush Kumar

20214284

CSE-6B1

1) Create a Weather Table with training data set which includes attributes like outlook, temperature, humidity, windy, play with the help of Data Mining Tool WEKA.

weather.arff - Notepad

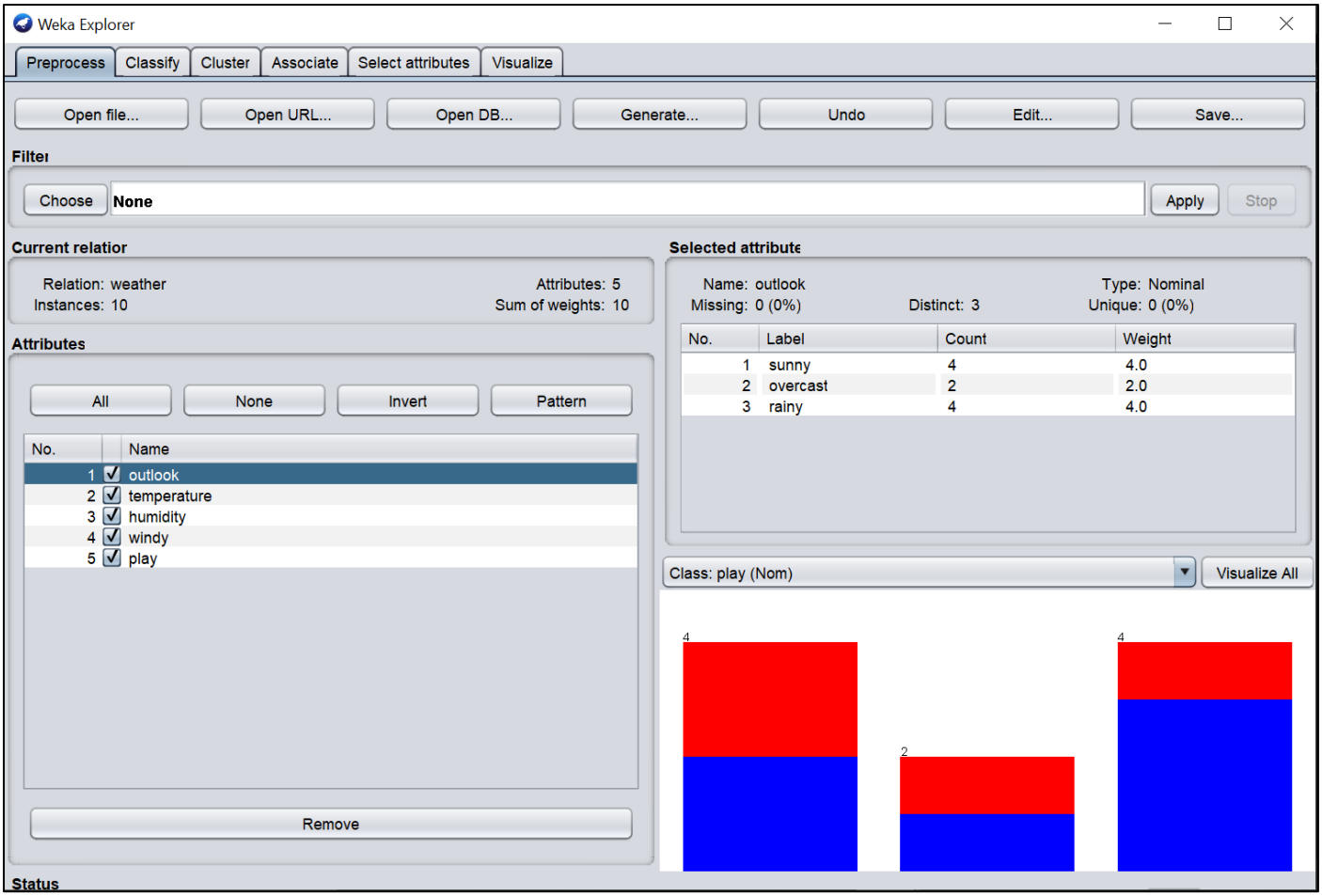
File Edit Format View Help

```
@relation weather

@attribute outlook {sunny, overcast, rainy}
@attribute temperature numeric
@attribute humidity numeric
@attribute windy {false, true}
@attribute play {yes, no}

@data
sunny, 85.0, 85.0, false, no
overcast, 80.0, 90.0, true, no
sunny, 83.0, 86.0, false, yes
rainy, 70.0, 86.0, false, yes
rainy, 68.0, 80.0, false, yes
rainy, 65.0, 70.0, true, no
overcast, 64.0, 65.0, false, yes
sunny, 72.0, 95.0, true, no
sunny, 69.0, 70.0, false, yes
rainy, 75.0, 80.0, false, yes
```

Viewer

Relation: weather

No.	1: outlook Nominal	2: temperature Numeric	3: humidity Numeric	4: windy Nominal	5: **play** Nominal
1	sunny	85.0	85.0	false	no
2	overcast	80.0	90.0	true	no
3	sunny	83.0	86.0	false	yes
4	rainy	70.0	86.0	false	yes
5	rainy	68.0	80.0	false	yes
6	rainy	65.0	70.0	true	no
7	overcast	64.0	65.0	false	yes
8	sunny	72.0	95.0	true	no
9	sunny	69.0	70.0	false	yes
10	rainy	75.0	80.0	false	yes



Q2) Apply Pre-Processing techniques to the training data set of Weather Table using WEKA EXPLORER and KNOWLEDGEFLOW.

i) Add

(Climate Attribute)

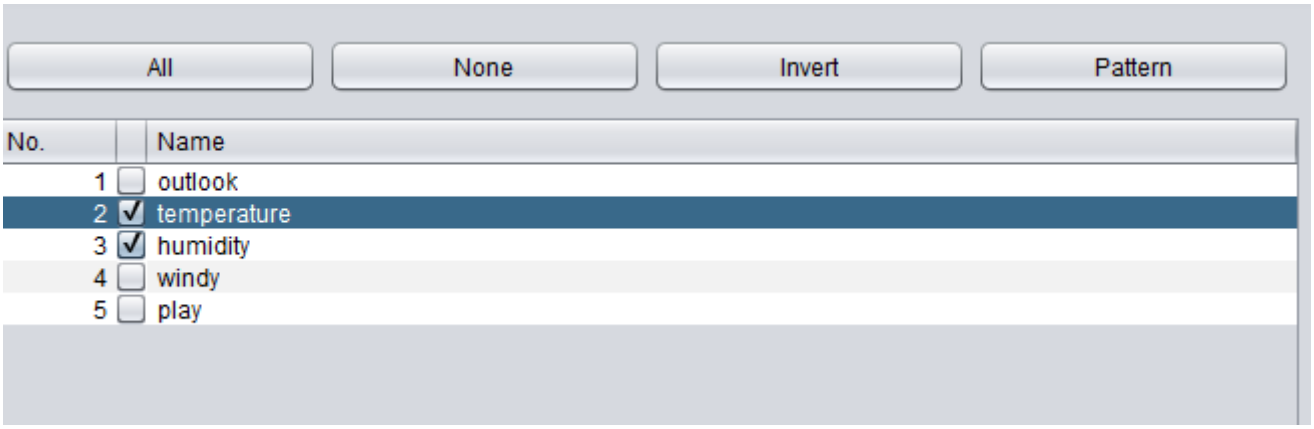
No.	1: outlook Nominal	2: temperature Numeric	3: humidity Numeric	4: windy Nominal	5: play Nominal	6: **Climate** Nominal
1	sunny	85.0	85.0	false	no	
2	overcast	80.0	90.0	true	no	
3	sunny	83.0	86.0	false	yes	
4	rainy	70.0	86.0	false	yes	
5	rainy	68.0	80.0	false	yes	
6	rainy	65.0	70.0	true	no	
7	overcast	64.0	65.0	false	yes	
8	sunny	72.0	95.0	true	no	
9	sunny	69.0	70.0	false	yes	
10	rainy	75.0	80.0	false	yes	

ii) Remove

(Windy and Play attribute)

No.	1: outlook Nominal	2: temperature Numeric	3: humidity Numeric	4: **Climate** Nominal
1	sunny	85.0	85.0	
2	overcast	80.0	90.0	
3	sunny	83.0	86.0	
4	rainy	70.0	86.0	
5	rainy	68.0	80.0	
6	rainy	65.0	70.0	
7	overcast	64.0	65.0	
8	sunny	72.0	95.0	
9	sunny	69.0	70.0	
10	rainy	75.0	80.0	

3) Attribute Selection



5) Normalization

No.	1: outlook Nominal	2: temperature Numeric	3: humidity Numeric	4: windy Nominal	5: **play** Nominal
1	overcast	0.0	0.0	TRUE	yes
2	rainy	0.04761904...	0.16129...	TRUE	no
3	rainy	0.19047619...	0.48387...	FALSE	yes
4	sunny	0.23809523...	0.16129...	FALSE	yes
5	rainy	0.28571428...	1.0	FALSE	yes
6	rainy	0.33333333...	0.83870...	TRUE	no
7	sunny	0.38095238...	0.96774...	FALSE	no
8	overcast	0.38095238...	0.80645...	TRUE	yes
9	rainy	0.52380952...	0.48387...	FALSE	yes
10	sunny	0.52380952...	0.16129...	TRUE	yes
11	sunny	0.76190476...	0.80645...	TRUE	no
12	overcast	0.80952380...	0.32258...	FALSE	yes
13	overcast	0.90476190...	0.67741...	FALSE	yes
14	sunny	1.0	0.64516...	FALSE	no

6) Discretization

No.	1: outlook	2: temperature	3: humidity	4: windy	5: **play**
	Nominal	Nominal	Nominal	Nominal	Nominal
1	sunny	'All'	'All'	FALSE	no
2	sunny	'All'	'All'	TRUE	no
3	overcast	'All'	'All'	FALSE	yes
4	rainy	'All'	'All'	FALSE	yes
5	rainy	'All'	'All'	FALSE	yes
6	rainy	'All'	'All'	TRUE	no
7	overcast	'All'	'All'	TRUE	yes
8	sunny	'All'	'All'	FALSE	no
9	sunny	'All'	'All'	FALSE	yes
10	rainy	'All'	'All'	FALSE	yes
11	sunny	'All'	'All'	TRUE	yes
12	overcast	'All'	'All'	TRUE	yes
13	overcast	'All'	'All'	FALSE	yes
14	rainy	'All'	'All'	TRUE	no