

21L-6205

Syed Sheraz Hussain

Lab-03

Nominal data:

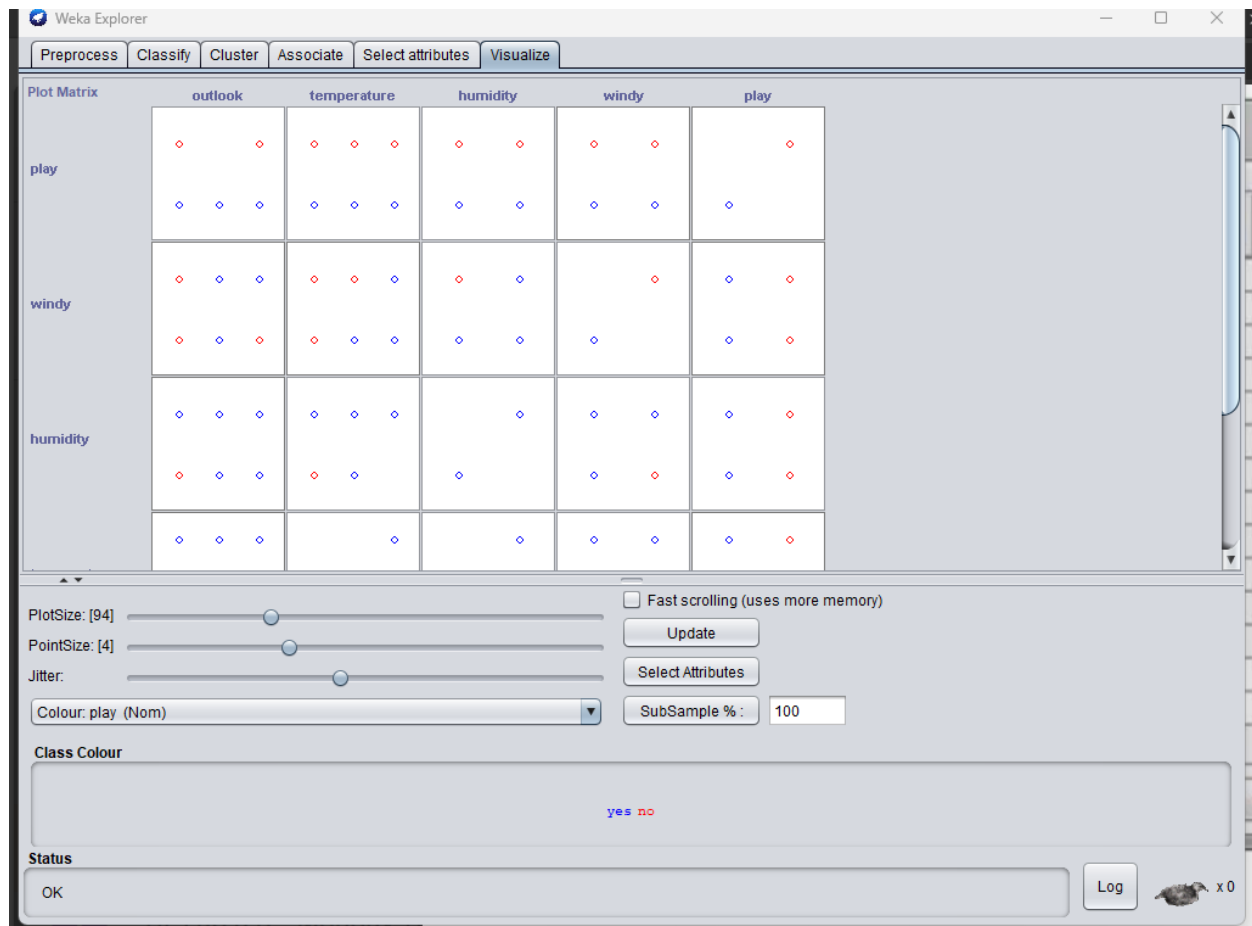
Load the data and click on edit



Relation: weather.symbolic

No.	1: outlook	2: temperature	3: humidity	4: windy	5: play
	Nominal	Nominal	Nominal	Nominal	Nominal
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mild	high	TRUE	no

Visual the data



Numerical Data :

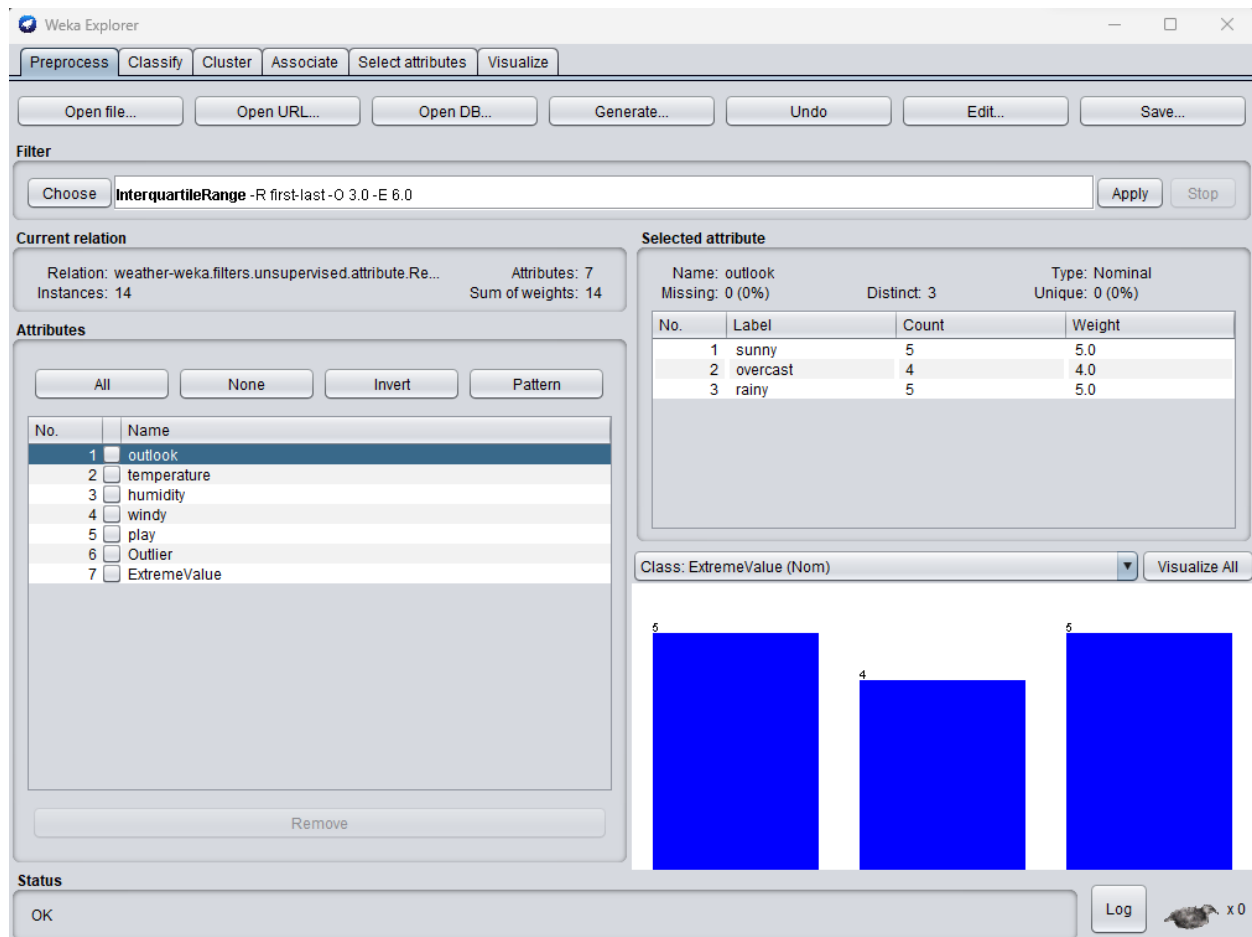
Replace the missing values by apply the filter

Viewer

Relation: weather-weka.filters.unsupervised.attribute.Replace

No.	1: outlook	2: temperature	3: humidity	4: windy	5: play
	Nominal	Numeric	Numeric	Nominal	Nominal
1	sunny	85.0	85.0	FALSE	no
2	sunny	80.0	90.0	TRUE	no
3	overcast	83.0	86.0	FALSE	yes
4	rainy	73.8461538...	96.0	FALSE	yes
5	rainy	68.0	80.0	FALSE	yes
6	rainy	65.0	70.0	TRUE	no
7	overcast	64.0	65.0	TRUE	yes
8	sunny	72.0	95.0	FALSE	no
9	sunny	69.0	82.5384...	FALSE	yes
10	rainy	75.0	80.0	FALSE	yes
11	sunny	75.0	70.0	TRUE	yes
12	overcast	72.0	90.0	TRUE	yes
13	overcast	81.0	75.0	FALSE	yes
14	rainy	71.0	91.0	TRUE	no

Apply the inter quartile range



Remove the attribute:

Remove the 2 attribute by apply the filter

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

Filter

Choose Remove -R 1,2 Apply Stop

Current relation

Relation: weather.symbolic-weka.filters.unsupervised.att... Attributes: 3
Instances: 14 Sum of weights: 14

Attributes

All None Invert Pattern

No.	Name
1	<input checked="" type="checkbox"/> humidity
2	<input type="checkbox"/> windy
3	<input type="checkbox"/> play

Remove

Selected attribute

Name: humidity
Missing: 0 (0%)
Distinct: 2
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	high	7	7.0
2	normal	7	7.0

Class: play (Nom) Visualize All

Status

OK Log x 0

Microsoft Edge FreeCours... 8. Do not charge your personal mobile

Csv:

Load the csv file data and click on the edit option the data is as follows:

weka explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Viewer

Relation: small_telco_labOne

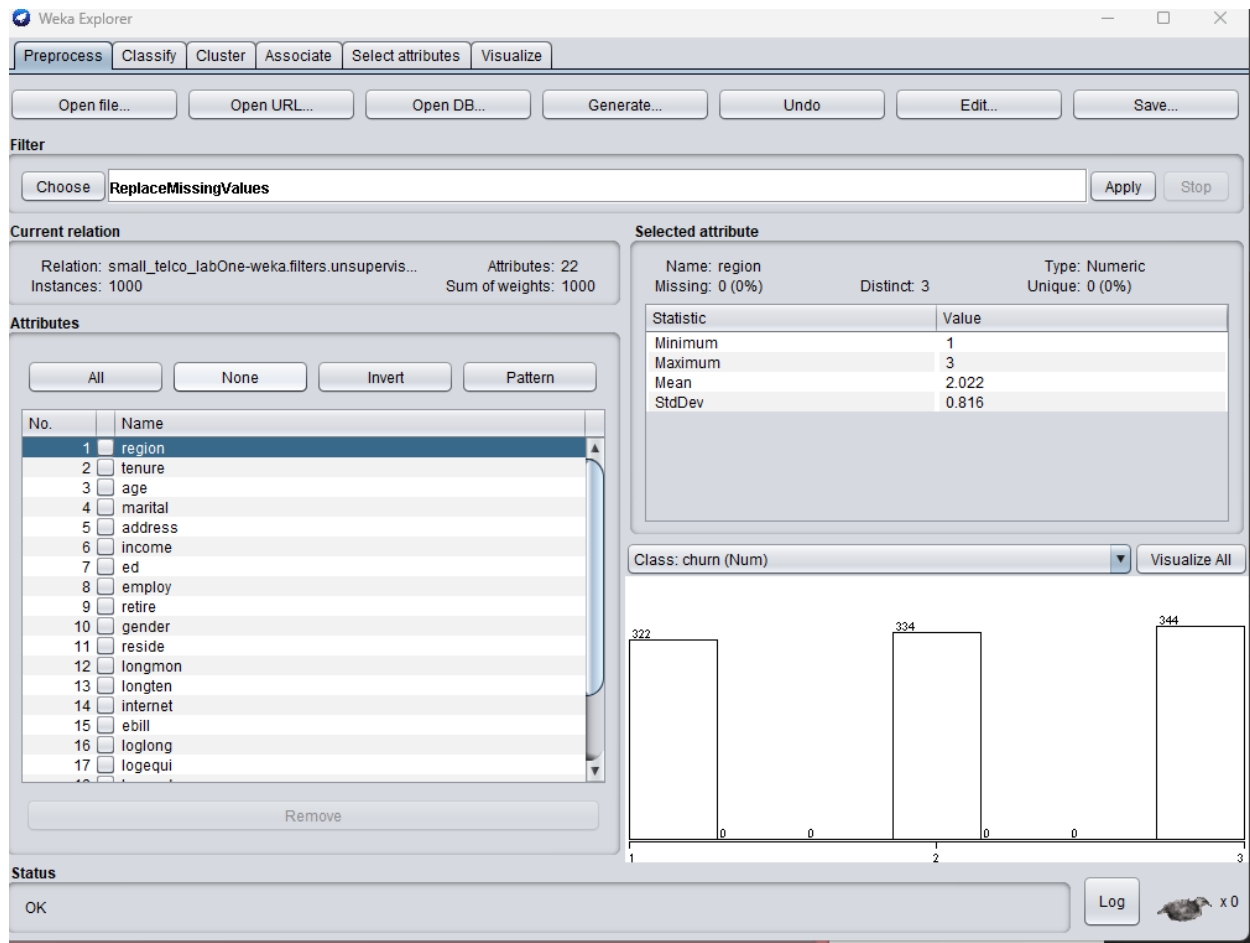
7: ed	8: employ	9: retire	10: gender	11: reside	12: longmon	13: longten	14: internet	15: ebill	16: loglong	17: logequi	18: logcard	19: logwire	20: lninc	21: custcat	22: churn
Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric
4.0	5.0	0.0	0.0	2.0	3.7	37.45	0.0	0.0	1.30833...		2.01490...		4.158...	1.0	1.0
5.0	5.0	0.0	0.0	6.0	4.4	42.0	0.0	0.0	1.48160...		2.72457...	3.57515...	4.912...	4.0	1.0
1.0	29.0	0.0	1.0	2.0	18.15	1300.6	0.0	0.0	2.89867...		3.40949...		4.753...	3.0	0.0
2.0	0.0	0.0	1.0	1.0	9.45	288.8	0.0	0.0	2.24601...				3.496...	1.0	1.0
1.0	2.0	0.0	0.0	4.0	6.3	157.05	0.0	0.0	1.84054...				3.401...	3.0	0.0
2.0	16.0	0.0	1.0	1.0	11.8	487.4	0.0	0.0	2.46809...		2.60268...		4.356...	3.0	0.0
2.0	4.0	0.0	1.0	5.0	10.9	504.5	1.0	1.0	2.38876...		2.16905...		2.944...	2.0	1.0
2.0	10.0	0.0	0.0	3.0	6.05	239.55	1.0	1.0	1.80005...	3.91402...	3.14630...	4.17284...	4.330...	4.0	0.0
4.0	31.0	0.0	0.0	5.0	9.75	449.05	0.0	0.0	2.27726...		2.48490...		5.111...	3.0	0.0
1.0	22.0	0.0	0.0	3.0	24.15	1659.7	0.0	0.0	3.18428...		2.80336...		4.276...	2.0	0.0
4.0	5.0	0.0	1.0	1.0	4.85	17.25	1.0	1.0	1.57897...	3.26384...			4.828...	1.0	1.0
2.0	15.0	0.0	1.0	1.0	7.1	47.45	0.0	0.0	1.96009...		3.16758...		4.382...	3.0	0.0
2.0	9.0	0.0	1.0	3.0	8.55	308.7	0.0	0.0	2.14593...		3.73169...		3.610...	1.0	0.0
4.0	23.0	0.0	1.0	3.0	15.6	825.35	1.0	1.0	2.74727...	3.84374...		4.11169...	4.744...	4.0	1.0
1.0	8.0	0.0	1.0	2.0	4.4	36.8	0.0	0.0	1.48160...				3.218...	1.0	0.0
5.0	1.0	0.0	0.0	4.0	5.1	146.25	1.0	0.0	1.62924...	3.40949...	2.42036...		4.317...	2.0	0.0
2.0	30.0	0.0	0.0	1.0	16.15	946.9	0.0	0.0	2.78192...	3.44361...	3.40119...		5.087...	3.0	0.0
2.0	6.0	0.0	1.0	3.0	6.65	230.8	0.0	0.0	1.89461...				3.891...	3.0	0.0
1.0	3.0	0.0	0.0	1.0	1.05	1.05	0.0	0.0	0.04879...				2.995...	1.0	0.0
4.0	2.0	0.0	0.0	4.0	6.7	140.95	1.0	1.0	1.90210...	3.87328...	3.18841...	3.64544...	4.343...	4.0	1.0
3.0	1.0	0.0	1.0	1.0	3.75	25.65	1.0	0.0	1.32175...	3.52046...		2.92852...	2.772...	2.0	1.0
1.0	24.0	0.0	0.0	2.0	20.7	1391.05	0.0	0.0	3.03013...		3.09104...		4.787...	1.0	0.0
5.0	4.0	0.0	1.0	2.0	5.3	253.35	1.0	1.0	1.66770...	3.90399...	3.28653...	3.93963...	4.615...	4.0	0.0
1.0	25.0	0.0	0.0	3.0	15.05	810.45	0.0	0.0	2.71137...		3.30505...		4.204...	1.0	0.0
1.0	5.0	0.0	0.0	5.0	12.5	153.75	0.0	0.0	2.52572...		2.89037...		3.583...	3.0	0.0
2.0	0.0	0.0	1.0	3.0	2.2	2.2	0.0	0.0	0.78845...		3.70130...		3.496...	3.0	0.0
2.0	8.0	0.0	1.0	1.0	8.25	399.15	1.0	1.0	2.11021...	3.60821...	3.33220...	3.62167...	3.610...	4.0	0.0
1.0	5.0	0.0	0.0	4.0	9.1	234.95	0.0	0.0	2.20827...				3.433...	3.0	0.0
2.0	2.0	0.0	1.0	2.0	2.2	25.25	0.0	0.0	1.08474...				2.933...	1.0	0.0

Add instance Undo OK Cancel

Status

OK Log x 0

Replace the missing values by apply the filter:



Apply the inter quartile range : by the filter

Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Open file...

Open URL...

Open DB...

Generate...

Undo

Edit...

Save...

Filter

Choose

InterquartileRange -R first-last -O 3.0 -E 6.0 -P

Apply

Stop

Current relation

Relation: small_telco_labOne-weka.filters.unsupervis...
Instances: 1000

Attributes: 68
Sum of weights: 1000

Attributes

All

None

Invert

Pattern

No.	Name
1	<input checked="" type="checkbox"/> region
2	<input type="checkbox"/> tenure
3	<input type="checkbox"/> age
4	<input type="checkbox"/> marital
5	<input type="checkbox"/> address
6	<input type="checkbox"/> income
7	<input type="checkbox"/> ed
8	<input type="checkbox"/> employ
9	<input type="checkbox"/> retire
10	<input type="checkbox"/> gender
11	<input type="checkbox"/> reside
12	<input type="checkbox"/> longmon
13	<input type="checkbox"/> longten
14	<input type="checkbox"/> internet
15	<input type="checkbox"/> ebill
16	<input type="checkbox"/> loglong
17	<input type="checkbox"/> logequi

Remove

Selected attribute

Name: region
Missing: 0 (0%)
Distinct: 3
Type: Numeric
Unique: 0 (0%)

Statistic	Value
Minimum	1
Maximum	3
Mean	2.022
StdDev	0.816

Class: churn_ExtremeValue (Nom)

Visualize All

Category	Count
1	322
2	334
3	344

Status

OK

Log

StdDev

0.816

Select the attribute by apply inter-quartile range: by detection true

Weka Explorer

Preprocess | Classify | Cluster | Associate | **Select attributes** | Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

Filter

Choose **NumericCleaner** -min -1.7976931348623157E308 -min-default -1.7976931348623157E308 -max 1.7976931348623157E308 -max-default 1.79 **Apply** **Stop**

Current relation

Relation: 111-weka.filters.unsupervised.attribute.Repl... Attributes: 68
Instances: 1000 Sum of weights: 1000

Attributes

All None Invert Pattern

No.	Name
46	reside_ExtremeValue
47	longmon_Outlier
48	longmon_ExtremeValue
49	longten_Outlier
50	longten_ExtremeValue
51	<input checked="" type="checkbox"/> internet_Outlier
52	<input checked="" type="checkbox"/> internet_ExtremeValue
53	<input checked="" type="checkbox"/> ebill_Outlier
54	<input checked="" type="checkbox"/> ebill_ExtremeValue
55	<input checked="" type="checkbox"/> loglong_Outlier
56	<input checked="" type="checkbox"/> loglong_ExtremeValue
57	<input checked="" type="checkbox"/> logequi_Outlier
58	logequi_ExtremeValue
59	logcard_Outlier
60	logcard_ExtremeValue
61	logwire_Outlier
62	logwire_ExtremeValue

Selected attribute

Name: logequi_Outlier
Missing: 0 (0%) Distinct: 1 Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	no	1000	1000.0
2	yes	0	0.0

Class: logwire_ExtremeValue (Nom) **Visualize All**

Status

OK Log x 0

And the data is as by edit option: given extreme values:

Viewer

relation: small_telco_labOne-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.InterquartileRange-Rfirst-last-O3.0-E6.0

gender	11: reside	12: longmon	13: longten	14: internet	15: ebill	16: loglong	17: logequi	18: logcard	19: logwire	20: lninc	21: custcat	22: churn	23: Outlier	24: ExtremeValue
Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Nominal	Nominal
0.0	2.0	3.7	37.45	0.0	0.0	1.30833...	3.56809...	2.01490...	3.59829...	4.158...	1.0	1.0	no	no
0.0	6.0	4.4	42.0	0.0	0.0	1.48160...	3.56809...	2.72457...	3.57515...	4.912...	4.0	1.0	no	yes
1.0	2.0	18.15	1300.6	0.0	0.0	2.89867...	3.56809...	3.40949...	3.59829...	4.753...	3.0	0.0	no	no
1.0	1.0	9.45	288.8	0.0	0.0	2.24601...	3.56809...	2.85420...	3.59829...	3.496...	1.0	1.0	no	no
0.0	4.0	6.3	157.05	0.0	0.0	1.84054...	3.56809...	2.85420...	3.59829...	3.401...	3.0	0.0	no	no
1.0	1.0	11.8	487.4	0.0	0.0	2.46809...	3.56809...	2.60268...	3.59829...	4.356...	3.0	0.0	no	no
1.0	5.0	10.9	504.5	1.0	1.0	2.38876...	3.56809...	2.16905...	3.59829...	2.944...	2.0	1.0	no	no
0.0	3.0	6.05	239.55	1.0	1.0	1.80005...	3.91402...	3.14630...	4.17284...	4.330...	4.0	0.0	no	yes
0.0	5.0	9.75	449.05	0.0	0.0	2.27726...	3.56809...	2.48490...	3.59829...	5.111...	3.0	0.0	no	no
0.0	3.0	24.15	1659.7	0.0	0.0	3.18428...	3.56809...	2.80336...	3.59829...	4.276...	2.0	0.0	no	no
1.0	1.0	4.85	17.25	1.0	1.0	1.57897...	3.26384...	2.85420...	3.59829...	4.828...	1.0	1.0	no	yes
1.0	1.0	7.1	47.45	0.0	0.0	1.96009...	3.56809...	3.16758...	3.59829...	4.382...	3.0	0.0	no	no
1.0	3.0	8.55	308.7	0.0	0.0	2.14593...	3.56809...	3.73169...	3.59829...	3.610...	1.0	0.0	no	no
1.0	3.0	15.6	825.35	1.0	1.0	2.74727...	3.84374...	2.85420...	4.11169...	4.744...	4.0	1.0	no	yes
1.0	2.0	4.4	36.8	0.0	0.0	1.48160...	3.56809...	2.85420...	3.59829...	3.218...	1.0	0.0	no	no
0.0	4.0	5.1	146.25	1.0	0.0	1.62924...	3.40949...	2.42036...	3.59829...	4.317...	2.0	0.0	no	yes
0.0	1.0	16.15	946.9	0.0	0.0	2.78192...	3.44361...	3.40119...	3.59829...	5.087...	3.0	0.0	no	yes
1.0	3.0	6.65	230.8	0.0	0.0	1.89461...	3.56809...	2.85420...	3.59829...	3.891...	3.0	0.0	no	no
0.0	1.0	1.05	1.05	0.0	0.0	0.04879...	3.56809...	2.85420...	3.59829...	2.995...	1.0	0.0	no	no
0.0	4.0	6.7	140.95	1.0	1.0	1.90210...	3.87328...	3.18841...	3.64544...	4.343...	4.0	1.0	no	yes
1.0	1.0	3.75	25.65	1.0	0.0	1.32175...	3.52046...	2.85420...	2.92852...	2.772...	2.0	1.0	no	yes
0.0	2.0	20.7	1391.05	0.0	0.0	3.03013...	3.56809...	3.09104...	3.59829...	4.787...	1.0	0.0	no	no
1.0	2.0	5.3	253.35	1.0	1.0	1.66770...	3.90399...	3.28653...	3.93963...	4.615...	4.0	0.0	no	yes
0.0	3.0	15.05	810.45	0.0	0.0	2.71137...	3.56809...	3.30505...	3.59829...	4.204...	1.0	0.0	no	no
0.0	5.0	12.5	153.75	0.0	0.0	2.52572...	3.56809...	2.89037...	3.59829...	3.583...	3.0	0.0	no	no
1.0	3.0	2.2	2.2	0.0	0.0	0.78845...	3.56809...	3.70130...	3.59829...	3.496...	3.0	0.0	no	no
1.0	1.0	8.25	399.15	1.0	1.0	2.11021...	3.60821...	3.33220...	3.62167...	3.610...	4.0	0.0	no	yes
0.0	4.0	9.1	234.95	0.0	0.0	2.20827...	3.56809...	2.85420...	3.59829...	3.433...	3.0	0.0	no	no
1.0	3.0	2.0	25.25	0.0	0.0	1.05474...	3.56809...	2.85420...	3.59829...	3.023...	1.0	0.0	no	no

Add instance Undo OK Cancel

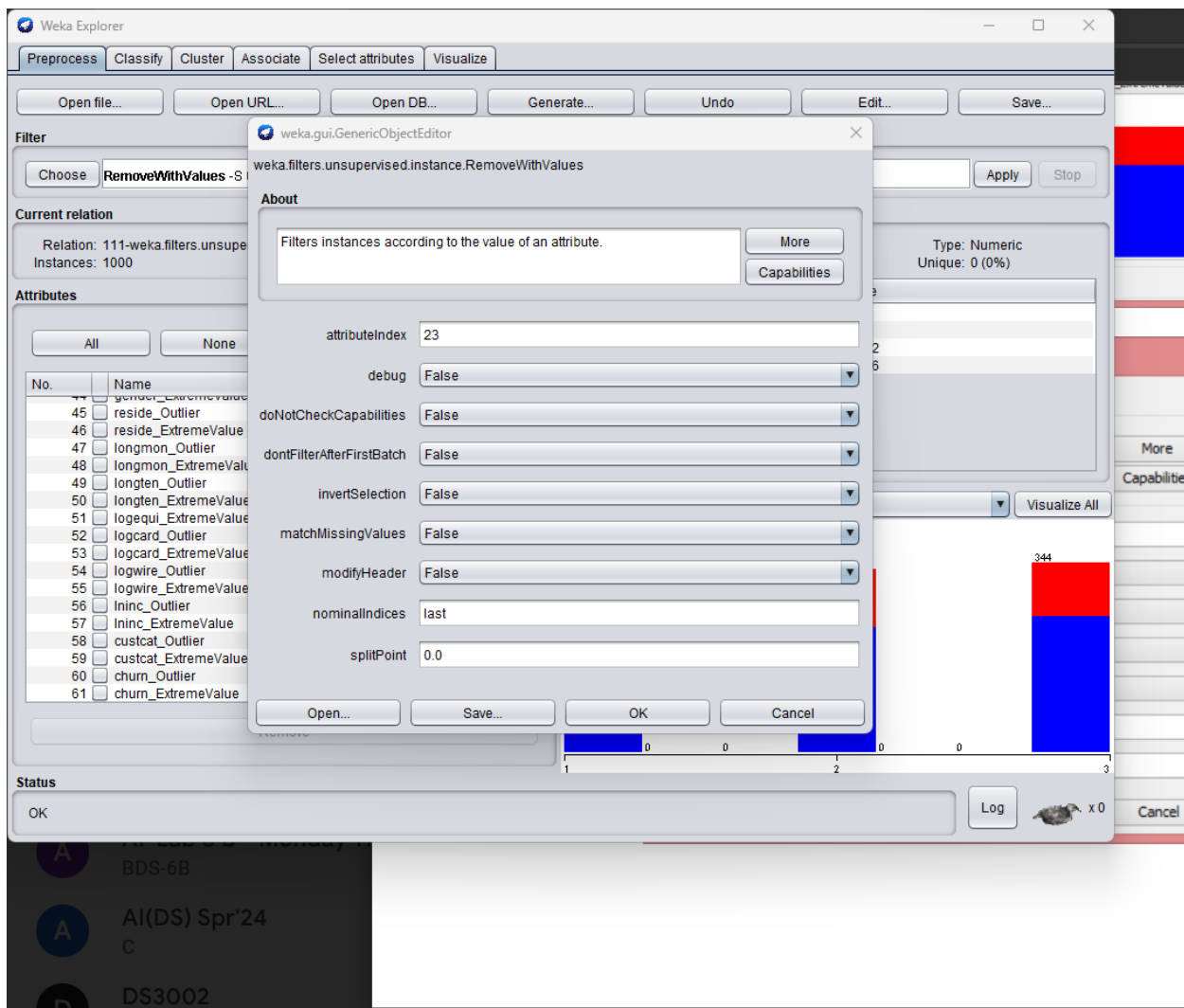
atus

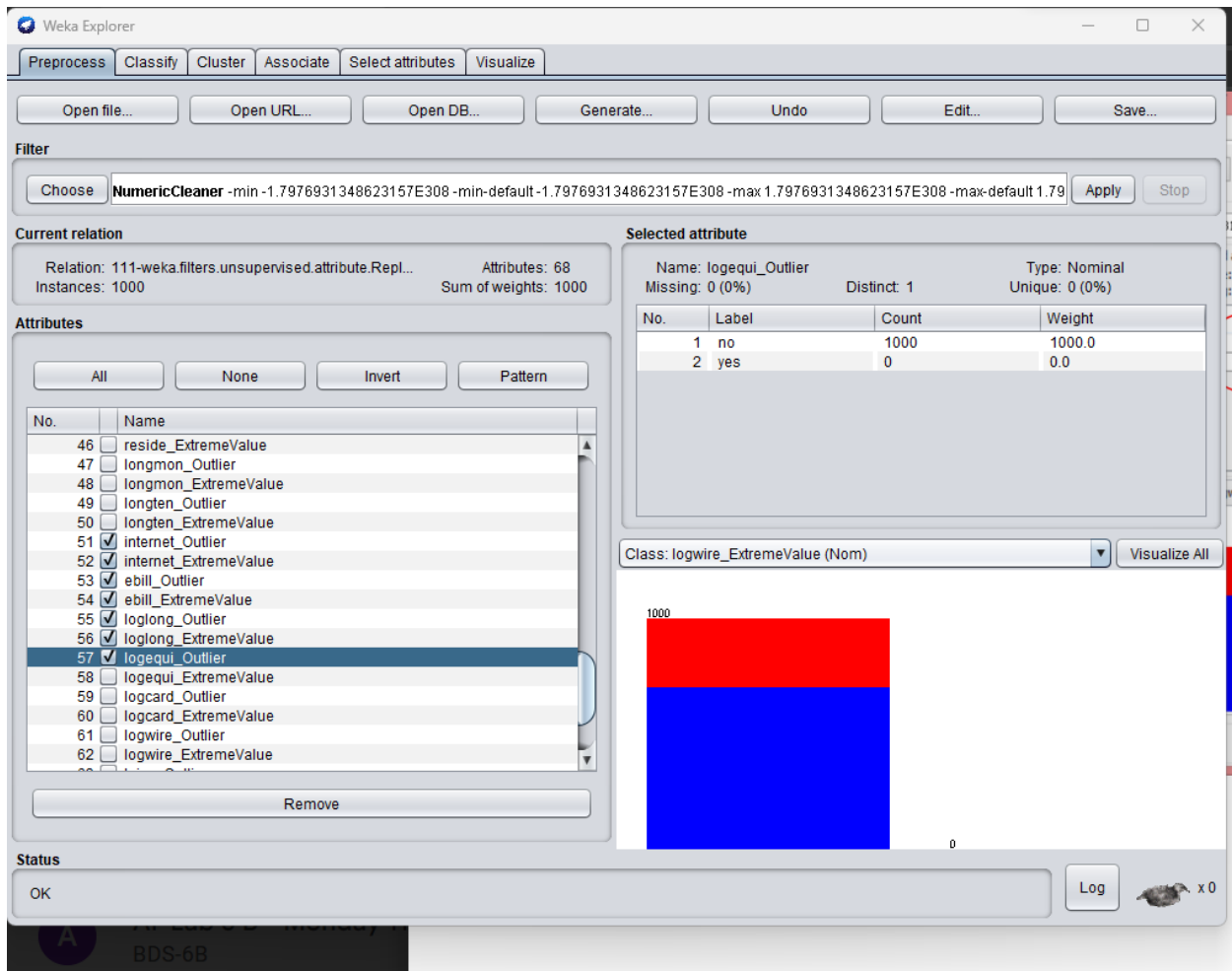
OK Log

And remove the values by just click remove button below:

And apply numeric cleaner:

Apply the numeric cleaner and attribute remove with 23 and last instances:





The attribute is remove

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

Filter

Choose RemoveWithValues -S 0.0 -C 23 -L last Apply Stop

Current relation

Relation: 111-weka.filters.unsupervised.attribute.Repla... Attributes: 61
Instances: 930 Sum of weights: 930

Attributes

All None Invert Pattern

No.	Name
44	gender_ExtremeValue
45	reside_Outlier
46	reside_ExtremeValue
47	longmon_Outlier
48	longmon_ExtremeValue
49	longten_Outlier
50	longten_ExtremeValue
51	logequi_ExtremeValue
52	logcard_Outlier
53	logcard_ExtremeValue
54	logwire_Outlier
55	logwire_ExtremeValue
56	lninc_Outlier
57	lninc_ExtremeValue
58	custcat_Outlier
59	custcat_ExtremeValue
60	churn_Outlier
61	churn_ExtremeValue

Remove

Selected attribute

Name: region Missing: 0 (0%) Distinct: 3 Type: Numeric Unique: 0 (0%)

Statistic	Value
Minimum	1
Maximum	3
Mean	2.015
StdDev	0.817

Class: logwire_ExtremeValue (Nom) Visualize All

303 310 317

1 2 3

Status

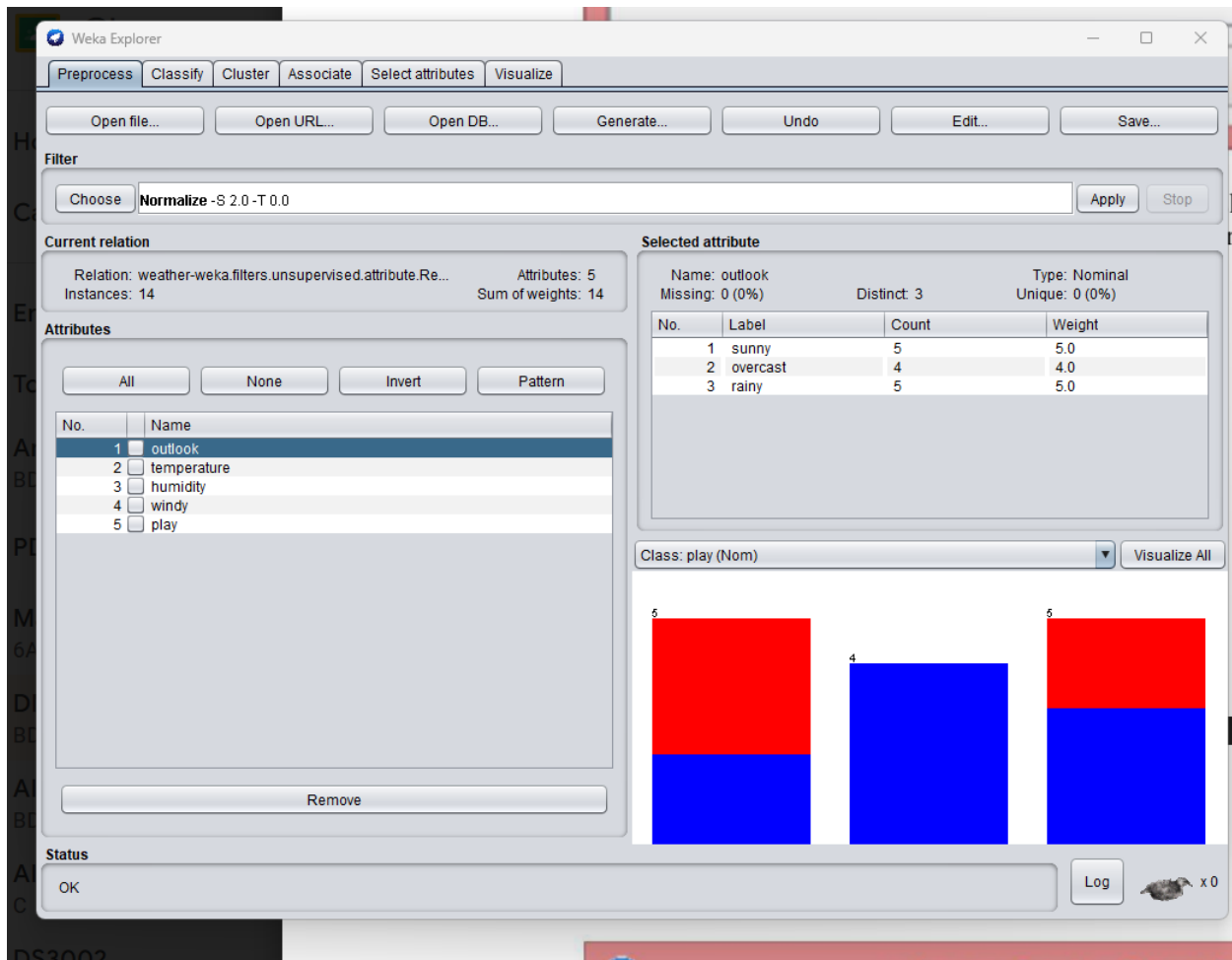
OK Log x 0

Using filters to perform normalization

Unsupervised Attribute Filter – Normalize: This values values in the given dataset for the default ra

Numerical Data Normalization:

Normalize with range (0,1),(-1,1),(2,0) And the normalization results are



Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Open file...

Open URL...

Open DB...

Generate...

Undo

Edit...

Save...

Filter

Choose

Normalize -S 1.0 -T 0.0

Apply

Stop

Current relation

Relation: weather-weka.filters.unsupervised.attribute.Re...

Attributes: 5

Instances: 14

Sum of weights: 14

Attributes

All

None

Invert

Pattern

No.	Name
1	<input checked="" type="checkbox"/> outlook
2	<input type="checkbox"/> temperature
3	<input type="checkbox"/> humidity
4	<input type="checkbox"/> windy
5	<input type="checkbox"/> play

Remove

Selected attribute

Name: outlook

Missing: 0 (0%)

Distinct: 3

Type: Nominal

Unique: 0 (0%)

No.	Label	Count	Weight
1	sunny	5	5.0
2	overcast	4	4.0
3	rainy	5	5.0

Class: play (Nom)

Visualize All

Status

OK

Log

x 0

•
•

Weka Explorer

PreprocessClassifyClusterAssociateSelect attributesVisualize

Open file...Open URL...Open DB...Generate...UndoEdit...Save...

Filter

ChooseNormalize -S 1.0 -T -1.0ApplyStop

Current relation

Relation: weather-weka.filters.unsupervised.attribute.Re...
Instances: 14

Attributes: 5
Sum of weights: 14

Attributes

AllNoneInvertPattern

No.	Name
1	<input checked="" type="checkbox"/> outlook
2	<input type="checkbox"/> temperature
3	<input type="checkbox"/> humidity
4	<input type="checkbox"/> windy
5	<input type="checkbox"/> play

Remove

Selected attribute

Name: outlook
Missing: 0 (0%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	sunny	5	5.0
2	overcast	4	4.0
3	rainy	5	5.0

Class: play (Nom)Visualize All


6

4


6

Status

OK

Log x 0

S3002

 weka.gui.GenericObjectE

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Viewer

Relation: weather-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.Normalize-S1.0-T0.0-weka.filters.unsupervised.attribute.Normalize...

No.	1: outlook	2: temperature	3: humidity	4: windy	5: play
	Nominal	Numeric	Numeric	Nominal	Nominal
1	sunny	0.0	-0.3548...	FALSE	no
2	sunny	-0.2380952...	-0.1935...	TRUE	no
3	overcast	-0.0952380...	-0.3225...	FALSE	yes
4	rainy	-0.5311355...	0.0	FALSE	yes
5	rainy	-0.8095238...	-0.5161...	FALSE	yes
6	rainy	-0.9523809...	-0.8387...	TRUE	no
7	overcast	-1.0	-1.0	TRUE	yes
8	sunny	-0.6190476...	-0.0322...	FALSE	no
9	sunny	-0.7619047...	-0.4342...	FALSE	yes
10	rainy	-0.4761904...	-0.5161...	FALSE	yes
11	sunny	-0.4761904...	-0.8387...	TRUE	yes
12	overcast	-0.6190476...	-0.1935...	TRUE	yes
13	overcast	-0.1904761...	-0.6774...	FALSE	yes
14	rainy	-0.6666666...	-0.1612...	TRUE	no

Add instance Undo OK Cancel

Status

OK Log x 0

Preprocess Classify Cluster Associate Select attributes Visualize

Viewer

Relation: weather-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.Normalize-S1.0-T0.0-weka.filters.unsupervised.attribute.Normalize...

No.	1: outlook	2: temperature	3: humidity	4: windy	5: play
	Nominal	Numeric	Numeric	Nominal	Nominal
1	sunny	2.0	1.29032...	FALSE	no
2	sunny	1.52380952...	1.61290...	TRUE	no
3	overcast	1.80952380...	1.35483...	FALSE	yes
4	rainy	0.93772893...	2.0	FALSE	yes
5	rainy	0.38095238...	0.96774...	FALSE	yes
6	rainy	0.09523809...	0.32258...	TRUE	no
7	overcast	0.0	0.0	TRUE	yes
8	sunny	0.76190476...	1.93548...	FALSE	no
9	sunny	0.47619047...	1.13151...	FALSE	yes
10	rainy	1.04761904...	0.96774...	FALSE	yes
11	sunny	1.04761904...	0.32258...	TRUE	yes
12	overcast	0.76190476...	1.61290...	TRUE	yes
13	overcast	1.61904761...	0.64516...	FALSE	yes
14	rainy	0.66666666...	1.67741...	TRUE	no

Add Instance Undo OK Cancel

Status

OK Log x 0

S3002

Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Open file

Open URL

Open DB

Generate

Undo

Edit

Save

Viewer

Relation: weather-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.Normalize-S1.0-T0.0

No. 1: outlook 2: temperature 3: humidity 4: windy 5: play

	Nominal	Numeric	Numeric	Nominal	Nominal
1	sunny	1.0	0.64516...	FALSE	no
2	sunny	0.76190476...	0.80645...	TRUE	no
3	overcast	0.90476190...	0.67741...	FALSE	yes
4	rainy	0.46886446...	1.0	FALSE	yes
5	rainy	0.19047619...	0.48387...	FALSE	yes
6	rainy	0.04761904...	0.16129...	TRUE	no
7	overcast	0.0	0.0	TRUE	yes
8	sunny	0.38095238...	0.96774...	FALSE	no
9	sunny	0.23809523...	0.56575...	FALSE	yes
10	rainy	0.52380952...	0.48387...	FALSE	yes
11	sunny	0.52380952...	0.16129...	TRUE	yes
12	overcast	0.38095238...	0.80645...	TRUE	yes
13	overcast	0.80952380...	0.32258...	FALSE	yes
14	rainy	0.33333333...	0.83870...	TRUE	no

Add instance

Undo


OK

Cancel

Status

OK

Log

 x 0