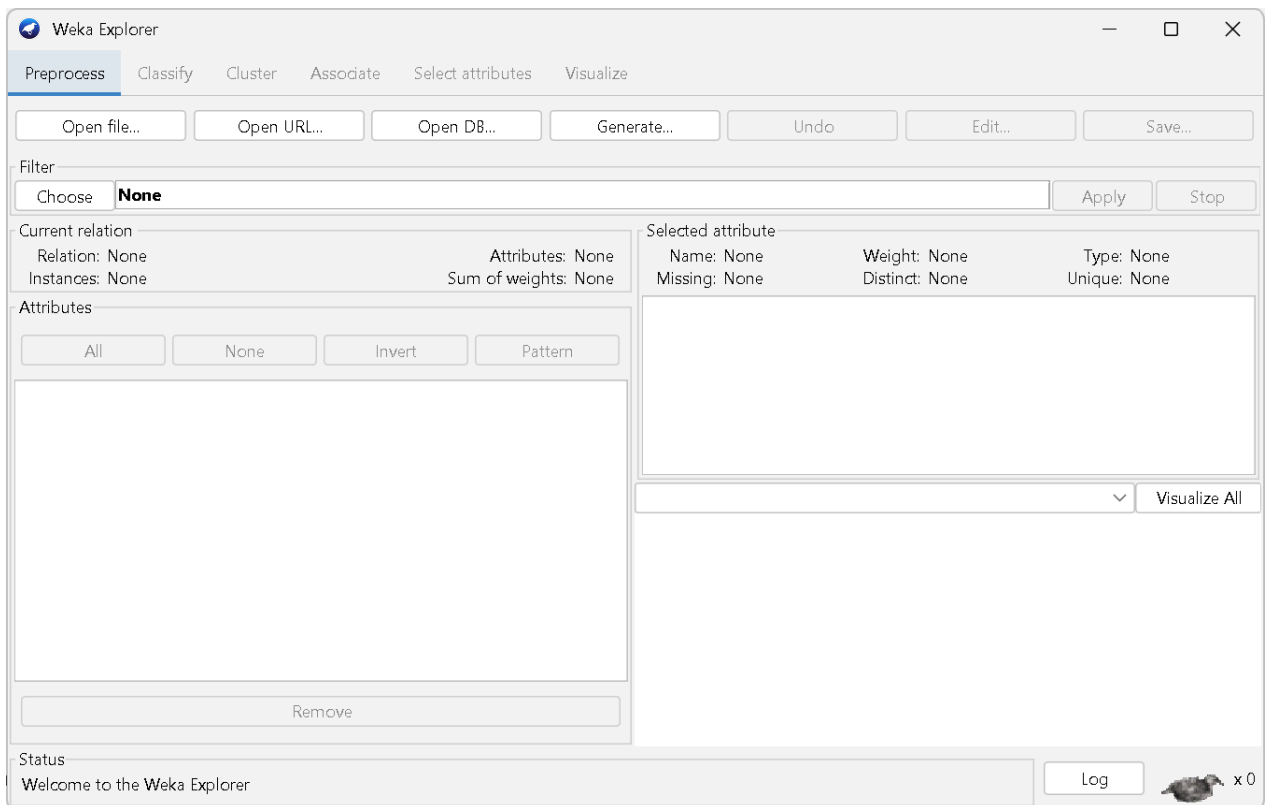
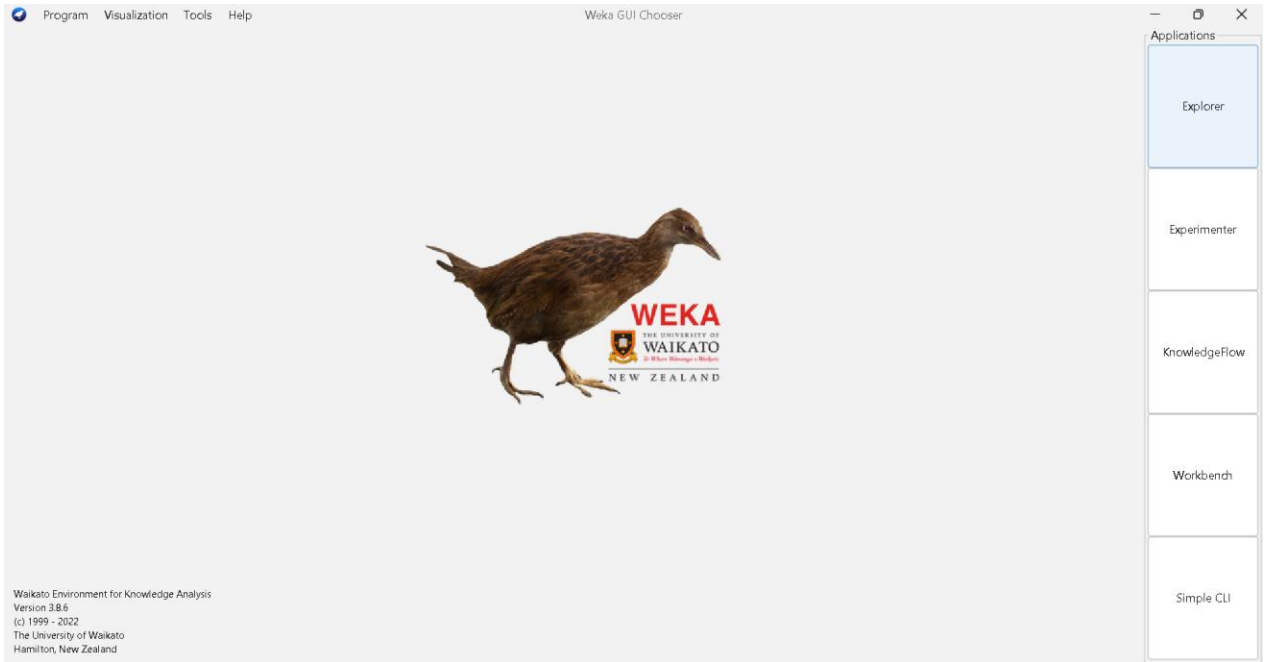
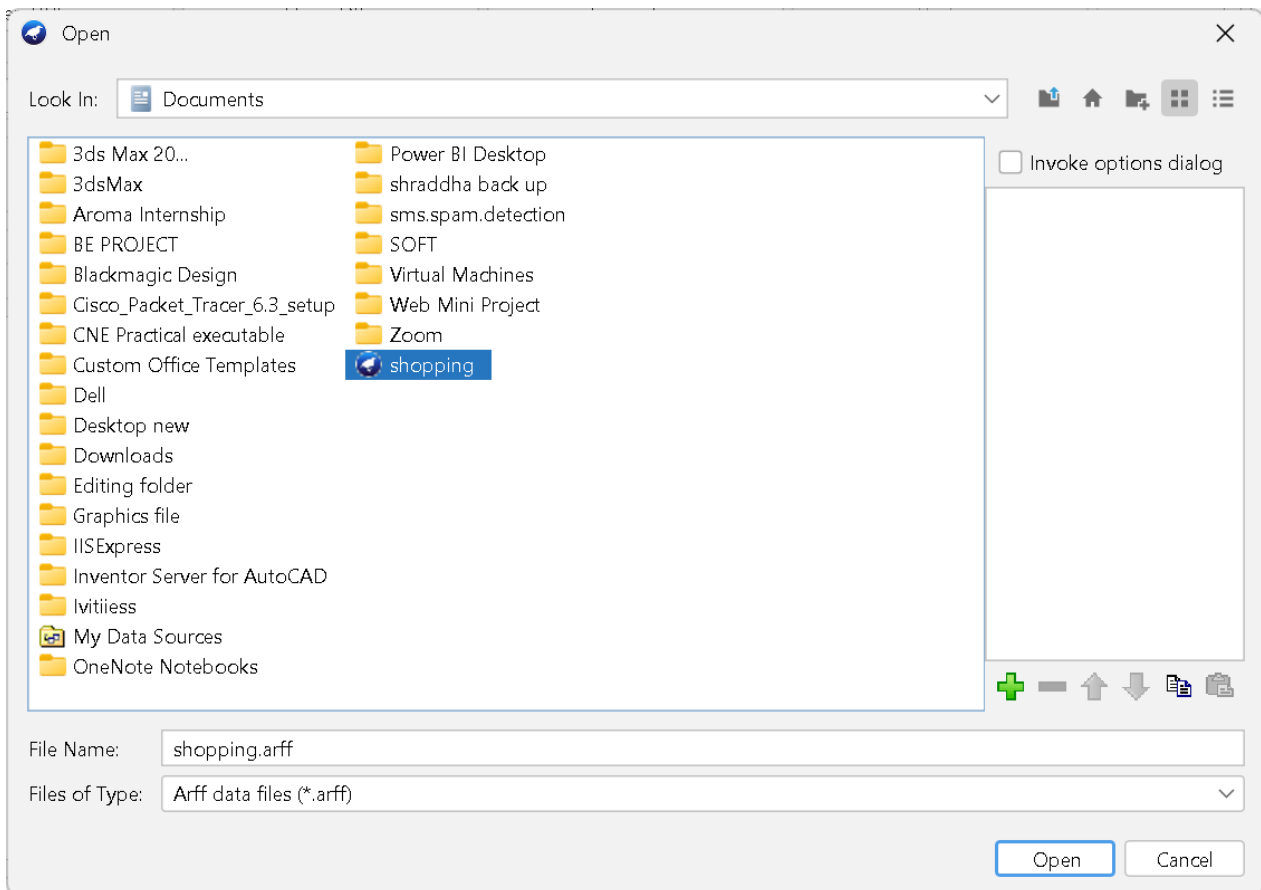


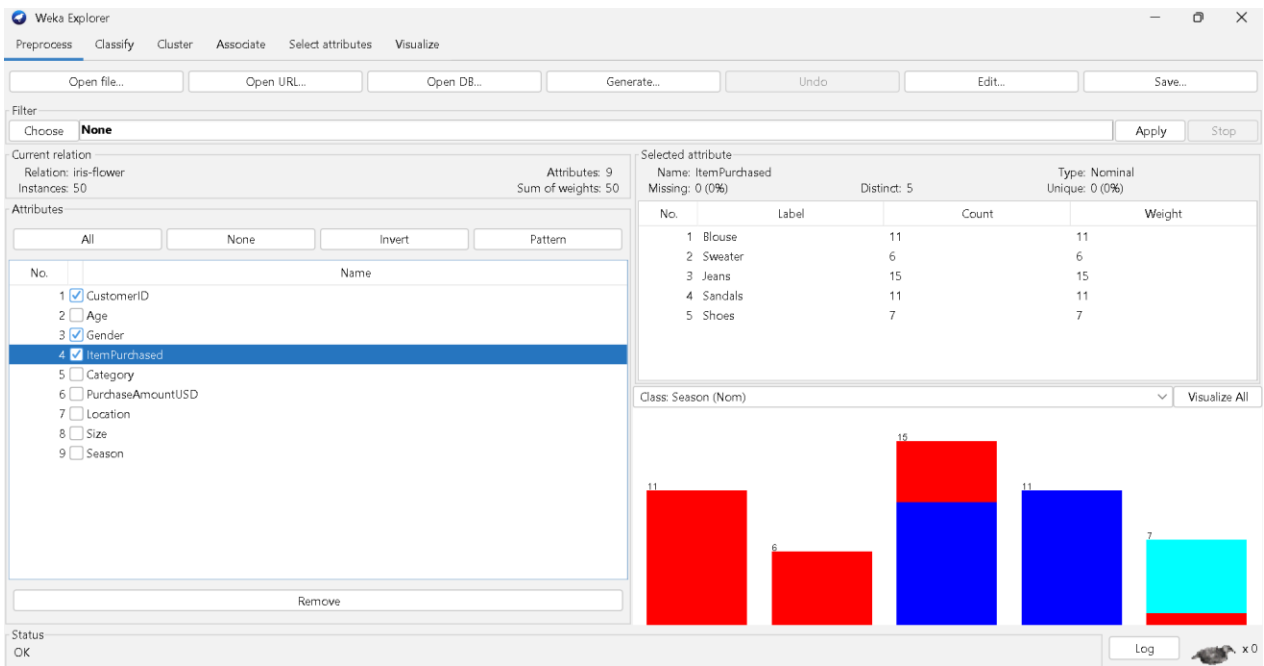
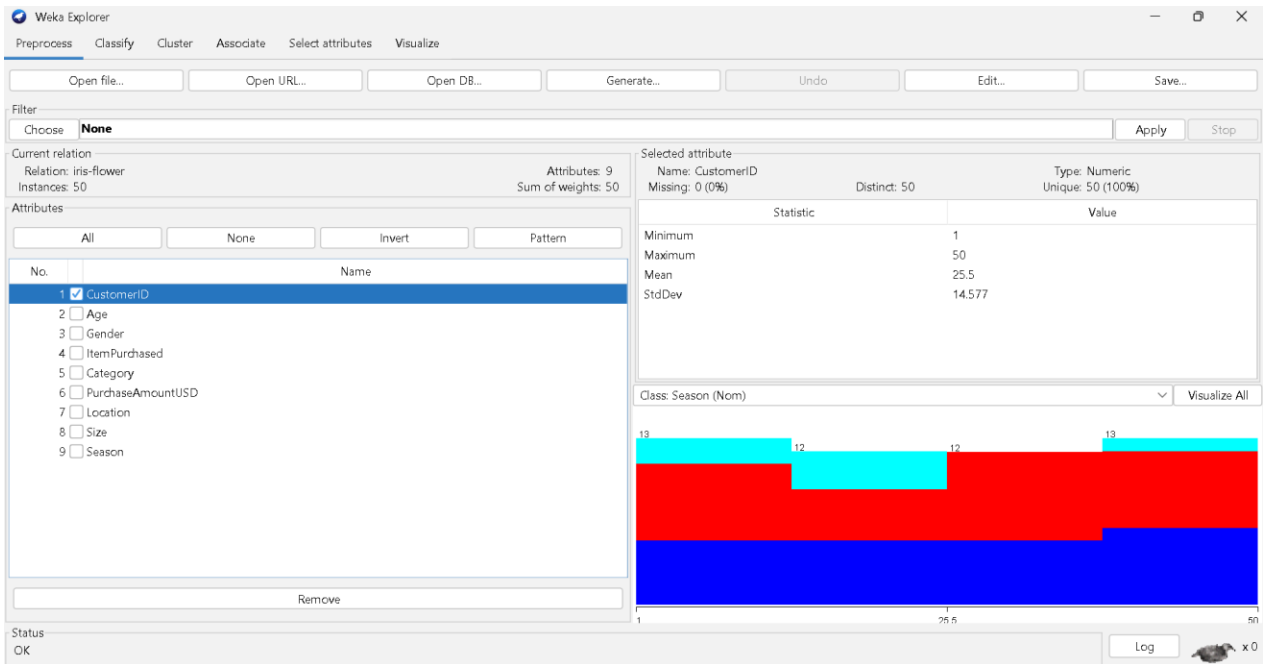
Practical No.1

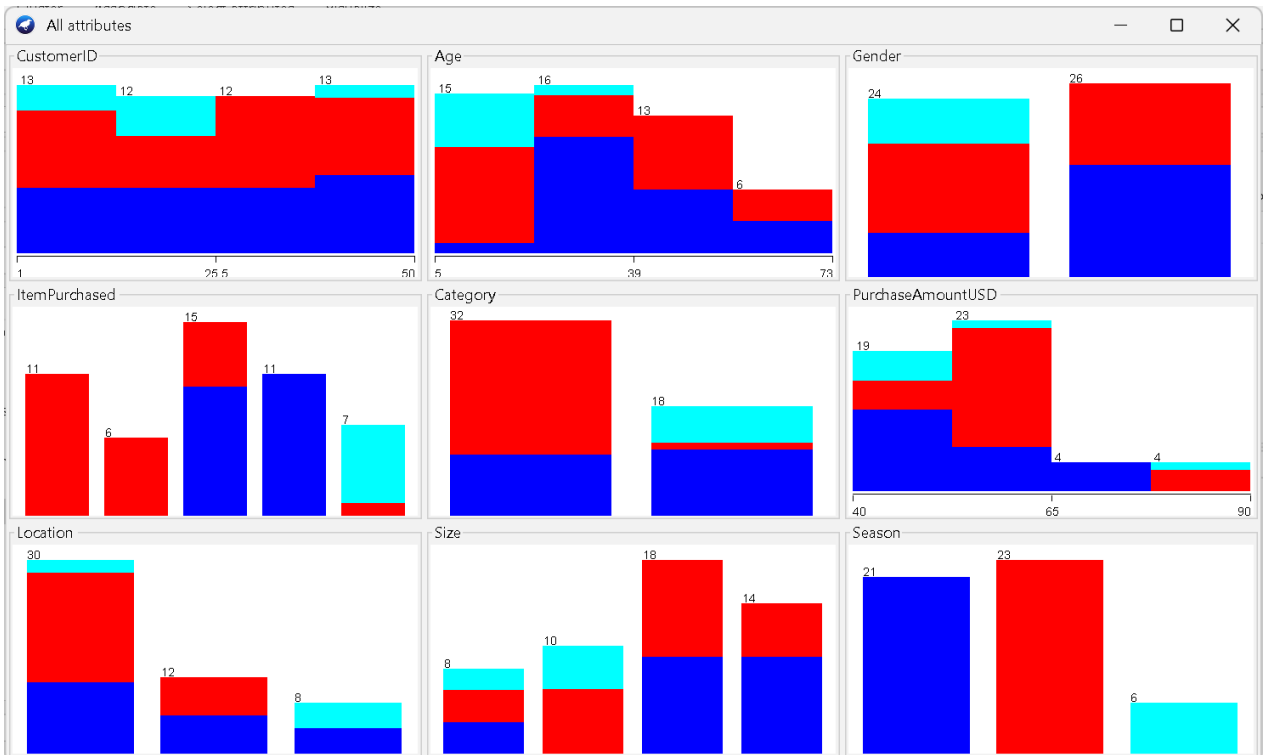
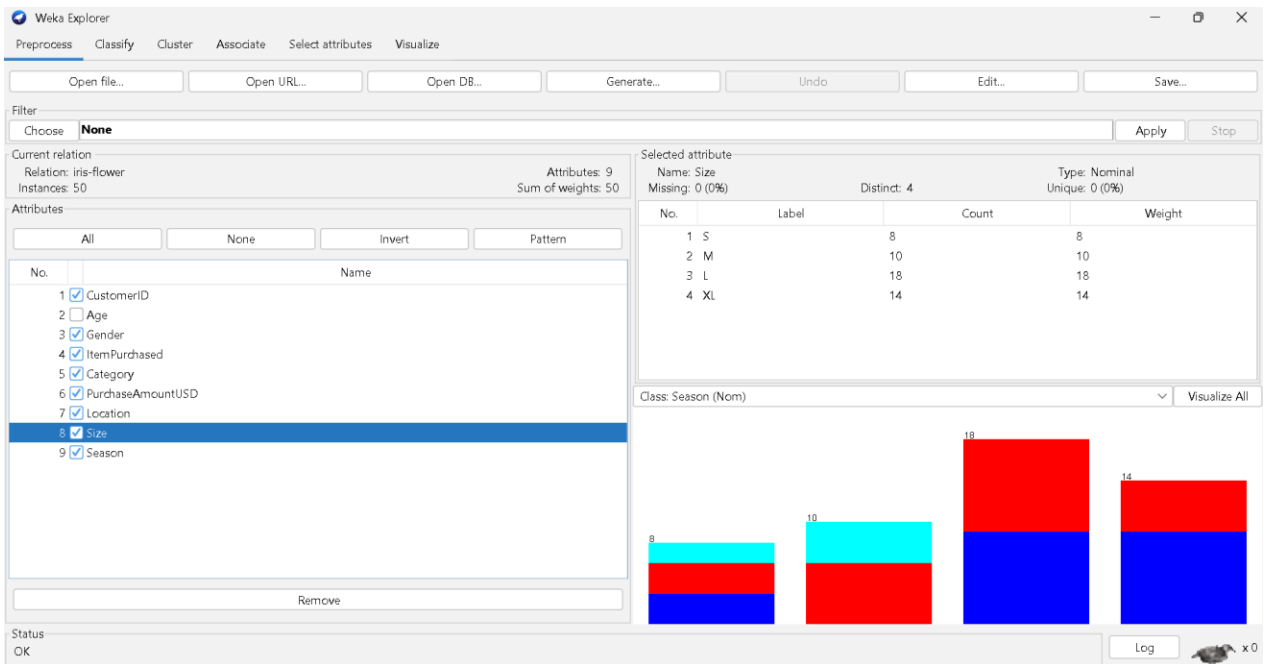


Practical No.2

```
shopping.arff
File Edit View
@attribute Age numeric
@attribute Gender {Male, Female}
@attribute ItemPurchased {Blouse, Sweater, Jeans, Sandals, Shoes}
@attribute Category {Clothing, Footwear}
@attribute PurchaseAmountUSD numeric
@attribute Location {Pune, Satara, Mumbai}
@attribute Size {S,M,L,XL}
@attribute Season {Summer, Winter, Spring}
@data
1, 55, Male, Blouse, Clothing, 53, Pune, L, Winter
2, 73, Female, Sweater, Clothing, 90, Pune, M, Winter
3, 49, Male, Sandals, Footwear, 40, Satara, S, Summer
4, 31, Female, Jeans, Clothing, 65, Pune, XL, Winter
5, 23, Female, Jeans, Clothing, 70, Mumbai, XL, Summer
6, 10, Male, Shoes, Footwear, 80, Pune, S, Spring
7, 51, Male, Blouse, Clothing, 65, Satara, L, Winter
8, 66, Female, Sandals, Footwear, 42, Pune, XL, Summer
9, 17, Male, Shoes, Footwear, 50, Mumbai, M, Spring
Ln 110, Col 1 | 3,113 characters | 100% | Windows (CRLF) | UTF-8
```







Practical No.3

Weka Explorer

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier: Choose **NaiveBayes**

Test options

☐ Use training set
☐ Supplied test set (Set...)
☒ Cross-validation Folds **10**
☐ Percentage split % **66**

More options...

(Nom) Season

Start Stop

Result list (right-click for options)

19:50:47 - bayes NaiveBayes

Classifier output

Scheme: weka.classifiers.bayes.NaiveBayes
Relation: iris-flower
Instances: 50
Attributes: 9

CustomerID
Age
Gender
ItemPurchased
Category
PurchaseAmountUSD
Location
Size
Season

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Naive Bayes Classifier

Attribute	Class		
	Summer (0.42)	Winter (0.45)	Spring (0.13)
CustomerID			
mean	26.5714	26	19.8333
std. dev.	13.7757	15.0159	13.0309
weight sum	21	23	6
precision	1	1	1

Status: OK

Log x0

Weka Explorer

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier: Choose **NaiveBayes**

Test options

☐ Use training set
☐ Supplied test set (Set...)
☒ Cross-validation Folds **10**
☐ Percentage split % **66**

More options...

(Nom) Season

Start Stop

Result list (right-click for options)

19:50:47 - bayes NaiveBayes

Classifier output

Age

	Summer	Winter	Spring
mean	37.418	33.6715	18.2593
std. dev.	14.2741	23.3121	8.8373
weight sum	21	23	6
precision	3.7778	3.7778	3.7778

Gender

	Summer	Winter	Spring
Male	7.0	13.0	7.0
Female	16.0	12.0	1.0
[total]	23.0	25.0	8.0

ItemPurchased

	Summer	Winter	Spring
Blouse	1.0	12.0	1.0
Sweater	1.0	7.0	1.0
Jeans	11.0	6.0	1.0
Sandals	12.0	1.0	1.0
Shoes	1.0	2.0	7.0
[total]	26.0	28.0	11.0

Category

	Summer	Winter	Spring
Clothing	11.0	23.0	1.0
Footwear	12.0	2.0	7.0
[total]	23.0	25.0	8.0

PurchaseAmountUSD

	Summer	Winter	Spring
mean	54.4974	64.2512	55.5556
std. dev.	13.9937	11.564	10.143
weight sum	21	23	6
precision	5.5556	5.5556	5.5556

Status: OK

Log x0

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose NaiveBayes

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) Season

Start Stop

Result list (right-click for options)

19:50:47 - bayes NaiveBayes

Classifier output

Location

Pune	12.0	18.0	3.0
Satara	7.0	7.0	1.0
Mumbai	5.0	1.0	5.0
[total]	24.0	26.0	9.0

Size

S	4.0	4.0	3.0
M	1.0	7.0	5.0
L	10.0	10.0	1.0
XL	10.0	6.0	1.0
[total]	25.0	27.0	10.0

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	35	70	%
Incorrectly Classified Instances	15	30	%
Kappa statistic	0.487		
Mean absolute error	0.1687		
Root mean squared error	0.3327		
Relative absolute error	41.8791 %		
Root relative squared error	74.2377 %		
Total Number of Instances	50		

Status

OK

Log x0

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose NaiveBayes

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) Season

Start Stop

Result list (right-click for options)

19:50:47 - bayes NaiveBayes

Classifier output

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	35	70	%
Incorrectly Classified Instances	15	30	%
Kappa statistic	0.487		
Mean absolute error	0.1687		
Root mean squared error	0.3327		
Relative absolute error	41.8791 %		
Root relative squared error	74.2377 %		
Total Number of Instances	50		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MDC	ROC Area	PRC Area	Class
	0.714	0.241	0.682	0.714	0.698	0.470	0.931	0.921	Summer
	0.739	0.259	0.708	0.739	0.723	0.479	0.894	0.910	Winter
	0.500	0.023	0.750	0.500	0.600	0.572	0.670	0.671	Spring
Weighted Avg.	0.700	0.223	0.702	0.700	0.698	0.486	0.883	0.886	

=== Confusion Matrix ===

a	b	c	<-- classified as
15	6	0	a = Summer
5	17	1	b = Winter
2	1	3	c = Spring

Status

OK

Log x0

Practical No.6

Weka Explorer

Preprocess **Cluster** Associate Select attributes Visualize

Clusterer: Choose **SimpleKMeans** -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Cluster mode

☒ Use training set
☐ Supplied test set Set...
☐ Percentage split % 66
☐ Classes to clusters evaluation (Nom) Season
☒ Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)

19:54:55 - SimpleKMeans

Status: OK

Log

Clusterer output

```
=== Run information ===

Scheme:      weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -
Relation:     iris-flower
Instances:    50
Attributes:   9
CustomerID
Age
Gender
ItemPurchased
Category
PurchaseAmountUSD
Location
Size
Season

Test mode:    evaluate on training data

=== Clustering model (full training set) ===

kMeans
=====

Number of iterations: 7
Within cluster sum of squared errors: 117.13618620589838

Initial starting points (random):
```

Weka Explorer

Preprocess **Cluster** Associate Select attributes Visualize

Clusterer: Choose **SimpleKMeans** -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Cluster mode

☒ Use training set
☐ Supplied test set Set...
☐ Percentage split % 66
☐ Classes to clusters evaluation (Nom) Season
☒ Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)

19:54:55 - SimpleKMeans

Status: OK

Log

Clusterer output

```
Initial starting points (random):

Cluster 0: 14,45,Female,Jeans,Clothing,65,Pune,XL,Winter
Cluster 1: 37,23,Female,Jeans,Clothing,70,Mumbai,XL,Summer

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute      Full Data      Cluster#
                (50.0)         0          1
                (50.0)         (29.0)      (21.0)
=====
CustomerID      25.5          25.4483     25.5714
Age             33.6          35.3793     31.1429
Gender          Female        Female      Male
ItemPurchased   Jeans         Jeans       Sandals
Category        Clothing      Clothing    Footwear
PurchaseAmountUSD 58.1          65.931     47.2857
Location        Pune          Pune        Pune
Size            L             L           S
Season          Winter        Winter      Summer

Time taken to build model (full training data) : 0.01 seconds

=== Model and evaluation on training set ===
```

Weka Explorer

PreprocessClassifyClusterAssociateSelect attributesVisualize

Clusterer

ChooseSimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Cluster mode

Use training set

Supplied test set

Percentage split

Classes to clusters evaluation

Set...

%

66

(Nom) Season

Store clusters for visualization

Ignore attributes

Start

Stop

Result list (right-click for options)

19:54:55 - SimpleKMeans

Clusterer output

Final cluster centroids:

Attribute	Full Data	Cluster#	0	1
	(50.0)	(29.0)	(21.0)	
CustomerID	25.5	25.4483	25.5714	
Age	33.6	35.3793	31.1429	
Gender	Female	Female	Male	
ItemPurchased	Jeans	Jeans	Sandals	
Category	Clothing	Clothing	Footwear	
PurchaseAmountUSD	58.1	65.931	47.2857	
Location	Pune	Pune	Pune	
Size	L	L	S	
Season	Winter	Winter	Summer	

Time taken to build model (full training data) : 0.01 seconds

=== Model and evaluation on training set ===

Clustered Instances

0	29 (58%)
1	21 (42%)

Status

OK

Log

x 0