Fundamentals of Data Mining Assignment #7 John M. Warlop Due Nov. 7th

1 Describe the Association learning method. How are the frequent item sets created in an efficient way?

Example database with 5 transactions and 5 items

transaction ID	milk	bread	butter	beer	diapers
1	1	1	0	0	0
2	0	0	1	0	0
3	0	0	0	1	1
4	1	1	1	0	0
5	0	1	0	0	0

The association learning method is used to find association between an item/items and other item/items. There is no class per se in this method.

If I is set of binary attributes and D is a
transaction. For example, D1 = 1,1,0,0,0 and I1 =
{milk, bread}

A rule is defined as $X \Rightarrow Y$ given that X & Y are subsets of I.

An example, assocation, that could be made from the table is {bread,butter} => milk or if a shopper buys bread butter, then they will buy milk

The algorithm needs to take into account a minimum support and minimum confidence for each association rule. There are two main steps:

- Find frequent datasets from database given a minimum support, these are the item sets
- The minimum confidence metric is applied to all frequent item sets from step one.

The first step's complexity of performance is $BigO(n^2)$. For large datasets, this could be

prohibitive. Fortunately, there is a proof(downward-closure property(\S ,#), that can be used on a frequent dataset. For, the proof says that if a dataset is frequent, then any subset of that dataset would also be frequent(i.e. if {bread,butter,milk} where frequent, then {bread, butter} would be frequent).

\$

Knowledge Discovery in Databases, AAAI/MIT Press, Cambridge, MA.

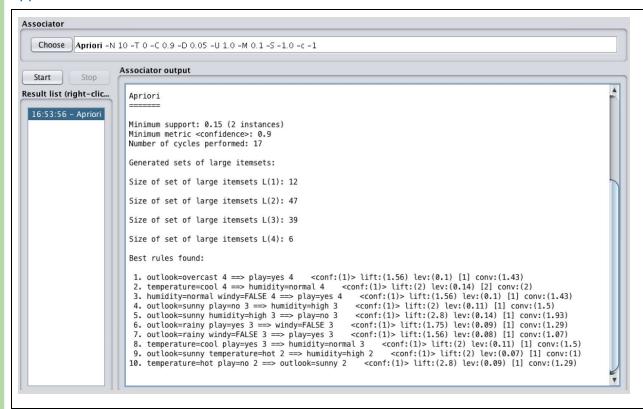
^ Jump up to: a b c d e f Agrawal, R.; Imieliński, T.; Swami, A. (1993). "Mining association rules between sets of items in large databases". *Proceedings of the 1993 ACM SIGMOD international conference on Management of data - SIGMOD '93*. p. 207. doi:10.1145/170035.170072. ISBN 0897915925.

#

Tan, Pang-Ning; Michael, Steinbach; Kumar, Vipin (2005). "Chapter 6. Association Analysis: Basic Concepts and Algorithms" (PDF). Introduction to Data Mining. Addison-Wesley. ISBN 0-321-32136-7.

Use the Association rule learner APRIORI method to find the association rule in the Weather.nominal data set. How many rules did it produce? How large are the itemsets? What was the largest one? What happened when you increased/decreased the confidence level? What about the number of rules? What happens when you increase the confidence parameter to 2? Why?

<u>Appendix</u>

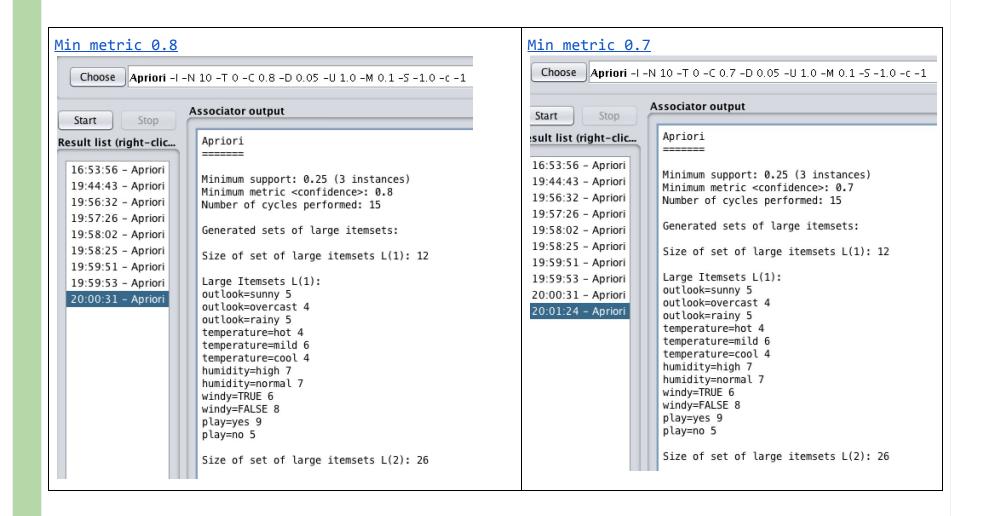


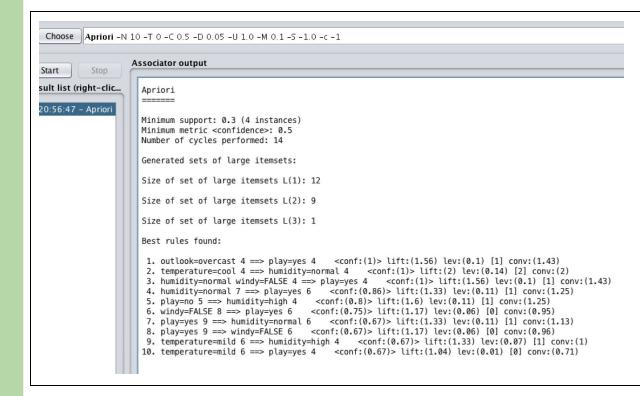
It produced 10 rules.

The item sets are: 12, 47,39 & 6 for attribute count of 1, 2, 3 & 4 respectively.

The largest data set was 47 with 2 attributes.

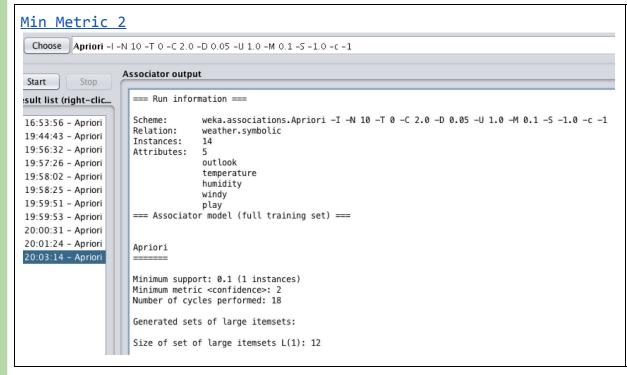
The default confidence level is 0.9. When lower confidence levels are entered, the itemsets tend to get lower.





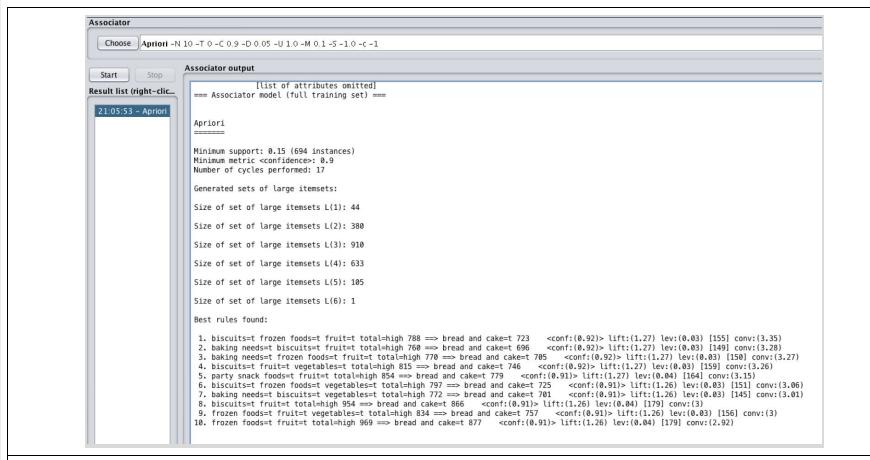
When the confidence level was set to a max of 0.5, the number of rules that were closer to .5 appeared. Of the 10 best rules, 3 were 100%.

When the confidence threshold is set lower, fewer rules with 100% confidence are generated. This makes sense because the APRIORI algorithm can stop sooner(<= .5) rather than (<= .9)



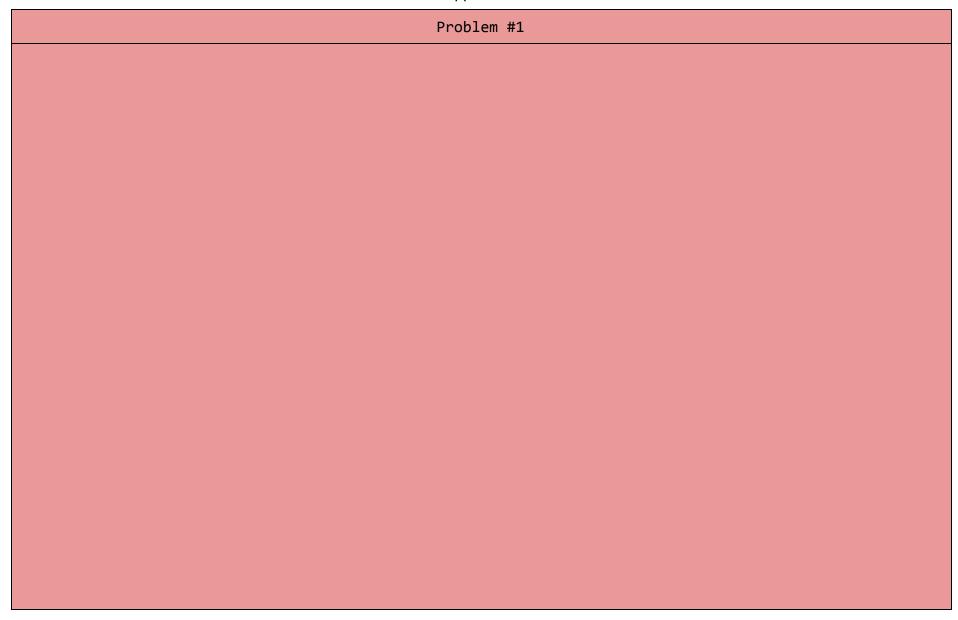
When the confidence parameter is set to 2 or 200%, you get no 10 best rules. A best rule is 100%, having a rule that is better than the best possible, does not make any sense. You also have too many items sets. Using a value of 2 leads to overfitting of data.

Use the Association rule learner A PRIORI method to find the association rule on the supermarket data set. What is the size of the largest item set? What was the highest confidence level produced? How many rules with that confidence? Any interesting rules you found?



The size of the largest data set, using default parameters, is 910. The highest confidence level(when max conf level set @ .9) was .92. There were four rules with confidence level of 0.92. The 10 rules are in HYP=>CONL, I thought it was interesting that all 10 rules had the same conclusion(bread & cake).

Appendix



```
Problem #2
Home
=== Run information ===
              weka.associations.Apriori -I -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Scheme:
Relation:
              weather.symbolic
              14
Instances:
Attributes:
              outlook
              temperature
              humidity
              windy
              play
=== Associator model (full training set) ===
Apriori
======
Minimum support: 0.15 (2 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 17
Generated sets of large itemsets:
Size of set of large itemsets L(1): 12
Large Itemsets L(1):
outlook=sunny 5
outlook=overcast 4
outlook=rainy 5
temperature=hot 4
```

```
temperature=mild 6
temperature=cool 4
humidity=high 7
humidity=normal 7
windy=TRUE 6
windy=FALSE 8
play=yes 9
play=no 5
Size of set of large itemsets L(2): 47
Large Itemsets L(2):
outlook=sunny temperature=hot 2
outlook=sunny temperature=mild 2
outlook=sunny humidity=high 3
outlook=sunny humidity=normal 2
outlook=sunny windy=TRUE 2
outlook=sunny windy=FALSE 3
outlook=sunny play=yes 2
outlook=sunny play=no 3
outlook=overcast temperature=hot 2
outlook=overcast humidity=high 2
outlook=overcast humidity=normal 2
outlook=overcast windy=TRUE 2
outlook=overcast windy=FALSE 2
outlook=overcast play=yes 4
outlook=rainy temperature=mild 3
outlook=rainy temperature=cool 2
outlook=rainy humidity=high 2
outlook=rainy humidity=normal 3
outlook=rainy windy=TRUE 2
outlook=rainy windy=FALSE 3
outlook=rainy play=yes 3
```

```
outlook=rainy play=no 2
temperature=hot humidity=high 3
temperature=hot windy=FALSE 3
temperature=hot play=yes 2
temperature=hot play=no 2
temperature=mild humidity=high 4
temperature=mild humidity=normal 2
temperature=mild windy=TRUE 3
temperature=mild windy=FALSE 3
temperature=mild play=yes 4
temperature=mild play=no 2
temperature=cool humidity=normal 4
temperature=cool windy=TRUE 2
temperature=cool windy=FALSE 2
temperature=cool play=yes 3
humidity=high windy=TRUE 3
humidity=high windy=FALSE 4
humidity=high play=yes 3
humidity=high play=no 4
humidity=normal windy=TRUE 3
humidity=normal windy=FALSE 4
humidity=normal play=yes 6
windy=TRUE play=yes 3
windy=TRUE play=no 3
windy=FALSE play=yes 6
windy=FALSE play=no 2
Size of set of large itemsets L(3): 39
Large Itemsets L(3):
outlook=sunny temperature=hot humidity=high 2
outlook=sunny temperature=hot play=no 2
outlook=sunny humidity=high windy=FALSE 2
```

```
outlook=sunny humidity=high play=no 3
outlook=sunny humidity=normal play=yes 2
outlook=sunny windy=FALSE play=no 2
outlook=overcast temperature=hot windy=FALSE 2
outlook=overcast temperature=hot play=yes 2
outlook=overcast humidity=high play=yes 2
outlook=overcast humidity=normal play=yes 2
outlook=overcast windy=TRUE play=yes 2
outlook=overcast windy=FALSE play=yes 2
outlook=rainy temperature=mild humidity=high 2
outlook=rainy temperature=mild windy=FALSE 2
outlook=rainy temperature=mild play=yes 2
outlook=rainy temperature=cool humidity=normal 2
outlook=rainy humidity=normal windy=FALSE 2
outlook=rainy humidity=normal play=yes 2
outlook=rainy windy=TRUE play=no 2
outlook=rainy windy=FALSE play=yes 3
temperature=hot humidity=high windy=FALSE 2
temperature=hot humidity=high play=no 2
temperature=hot windy=FALSE play=yes 2
temperature=mild humidity=high windy=TRUE 2
temperature=mild humidity=high windy=FALSE 2
temperature=mild humidity=high play=yes 2
temperature=mild humidity=high play=no 2
temperature=mild humidity=normal play=yes 2
temperature=mild windy=TRUE play=yes 2
temperature=mild windy=FALSE play=yes 2
temperature=cool humidity=normal windy=TRUE 2
temperature=cool humidity=normal windy=FALSE 2
temperature=cool humidity=normal play=yes 3
temperature=cool windy=FALSE play=yes 2
humidity=high windy=TRUE play=no 2
humidity=high windy=FALSE play=yes 2
```

```
humidity=high windy=FALSE play=no 2
humidity=normal windy=TRUE play=yes 2
humidity=normal windy=FALSE play=yes 4
Size of set of large itemsets L(4): 6
Large Itemsets L(4):
outlook=sunny temperature=hot humidity=high play=no 2
outlook=sunny humidity=high windy=FALSE play=no 2
outlook=overcast temperature=hot windy=FALSE play=yes 2
outlook=rainy temperature=mild windy=FALSE play=yes 2
outlook=rainy humidity=normal windy=FALSE play=yes 2
temperature=cool humidity=normal windy=FALSE play=yes 2
Best rules found:
 1. outlook=overcast 4 ==> play=yes 4
                                         <conf:(1)> lift:(1.56) lev:(0.1) [1] conv:(1.43)
 2. temperature=cool 4 ==> humidity=normal 4
                                                <conf:(1)> lift:(2) lev:(0.14) [2] conv:(2)
 3. humidity=normal windy=FALSE 4 ==> play=yes 4
                                                    <conf:(1)> lift:(1.56) lev:(0.1) [1] conv:(1.43)
 4. outlook=sunny play=no 3 ==> humidity=high 3
                                                   <conf:(1)> lift:(2) lev:(0.11) [1] conv:(1.5)
 5. outlook=sunny humidity=high 3 ==> play=no 3
                                                   <conf:(1)> lift:(2.8) lev:(0.14) [1] conv:(1.93)
 6. outlook=rainy play=yes 3 ==> windy=FALSE 3
                                                  <conf:(1)> lift:(1.75) lev:(0.09) [1] conv:(1.29)
 7. outlook=rainy windy=FALSE 3 ==> play=yes 3
                                                  <conf:(1)> lift:(1.56) lev:(0.08) [1] conv:(1.07)
 8. temperature=cool play=yes 3 ==> humidity=normal 3
                                                         <conf:(1)> lift:(2) lev:(0.11) [1] conv:(1.5)
 9. outlook=sunny temperature=hot 2 ==> humidity=high 2
                                                           <conf:(1)> lift:(2) lev:(0.07) [1] conv:(1)
10. temperature=hot play=no 2 ==> outlook=sunny 2
                                                     <conf:(1)> lift:(2.8) lev:(0.09) [1] conv:(1.29)
Home
MinMetric .8
=== Run information ===
Scheme:
              weka.associations.Apriori -I -N 10 -T 0 -C 0.8 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
              weather.symbolic
Relation:
```

```
Instances:
              14
Attributes:
              5
              outlook
              temperature
              humidity
              windy
              play
=== Associator model (full training set) ===
Apriori
======
Minimum support: 0.25 (3 instances)
Minimum metric <confidence>: 0.8
Number of cycles performed: 15
Generated sets of large itemsets:
Size of set of large itemsets L(1): 12
Large Itemsets L(1):
outlook=sunny 5
outlook=overcast 4
outlook=rainy 5
temperature=hot 4
temperature=mild 6
temperature=cool 4
humidity=high 7
humidity=normal 7
windy=TRUE 6
windy=FALSE 8
play=yes 9
```

```
play=no 5
Size of set of large itemsets L(2): 26
Large Itemsets L(2):
outlook=sunny humidity=high 3
outlook=sunny windy=FALSE 3
outlook=sunny play=no 3
outlook=overcast play=yes 4
outlook=rainy temperature=mild 3
outlook=rainy humidity=normal 3
outlook=rainy windy=FALSE 3
outlook=rainy play=yes 3
temperature=hot humidity=high 3
temperature=hot windy=FALSE 3
temperature=mild humidity=high 4
temperature=mild windy=TRUE 3
temperature=mild windy=FALSE 3
temperature=mild play=yes 4
temperature=cool humidity=normal 4
temperature=cool play=yes 3
humidity=high windy=TRUE 3
humidity=high windy=FALSE 4
humidity=high play=yes 3
humidity=high play=no 4
humidity=normal windy=TRUE 3
humidity=normal windy=FALSE 4
humidity=normal play=yes 6
windy=TRUE play=yes 3
windy=TRUE play=no 3
windy=FALSE play=yes 6
Size of set of large itemsets L(3): 4
```

```
Large Itemsets L(3):
outlook=sunny humidity=high play=no 3
outlook=rainy windy=FALSE play=yes 3
temperature=cool humidity=normal play=yes 3
humidity=normal windy=FALSE play=yes 4
Best rules found:
 1. outlook=overcast 4 ==> play=yes 4
                                        <conf:(1)> lift:(1.56) lev:(0.1) [1] conv:(1.43)
 2. temperature=cool 4 ==> humidity=normal 4
                                                <conf:(1)> lift:(2) lev:(0.14) [2] conv:(2)
 3. humidity=normal windy=FALSE 4 ==> play=yes 4
                                                    <conf:(1)> lift:(1.56) lev:(0.1) [1] conv:(1.43)
 4. outlook=sunny play=no 3 ==> humidity=high 3
                                                   <conf:(1)> lift:(2) lev:(0.11) [1] conv:(1.5)
 5. outlook=sunny humidity=high 3 ==> play=no 3
                                                   <conf:(1)> lift:(2.8) lev:(0.14) [1] conv:(1.93)
 6. outlook=rainy play=yes 3 ==> windy=FALSE 3
                                                  <conf:(1)> lift:(1.75) lev:(0.09) [1] conv:(1.29)
 7. outlook=rainy windy=FALSE 3 ==> play=yes 3
                                                  <conf:(1)> lift:(1.56) lev:(0.08) [1] conv:(1.07)
 8. temperature=cool play=yes 3 ==> humidity=normal 3
                                                         <conf:(1)> lift:(2) lev:(0.11) [1] conv:(1.5)
 9. humidity=normal 7 ==> play=yes 6 <conf:(0.86)> lift:(1.33) lev:(0.11) [1] conv:(1.25)
10. play=no 5 ==> humidity=high 4 <conf:(0.8)> lift:(1.6) lev:(0.11) [1] conv:(1.25)
MinMetric .7
=== Run information ===
Scheme:
              weka.associations.Apriori -I -N 10 -T 0 -C 0.7 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
              weather.symbolic
Relation:
Instances:
              14
Attributes:
              outlook
             temperature
              humidity
              windy
              play
=== Associator model (full training set) ===
```

```
Apriori
======
Minimum support: 0.25 (3 instances)
Minimum metric <confidence>: 0.7
Number of cycles performed: 15
Generated sets of large itemsets:
Size of set of large itemsets L(1): 12
Large Itemsets L(1):
outlook=sunny 5
outlook=overcast 4
outlook=rainy 5
temperature=hot 4
temperature=mild 6
temperature=cool 4
humidity=high 7
humidity=normal 7
windy=TRUE 6
windy=FALSE 8
play=yes 9
play=no 5
Size of set of large itemsets L(2): 26
Large Itemsets L(2):
outlook=sunny humidity=high 3
outlook=sunny windy=FALSE 3
outlook=sunny play=no 3
```

```
outlook=overcast play=yes 4
outlook=rainy temperature=mild 3
outlook=rainy humidity=normal 3
outlook=rainy windy=FALSE 3
outlook=rainy play=yes 3
temperature=hot humidity=high 3
temperature=hot windy=FALSE 3
temperature=mild humidity=high 4
temperature=mild windy=TRUE 3
temperature=mild windy=FALSE 3
temperature=mild play=yes 4
temperature=cool humidity=normal 4
temperature=cool play=yes 3
humidity=high windy=TRUE 3
humidity=high windy=FALSE 4
humidity=high play=yes 3
humidity=high play=no 4
humidity=normal windy=TRUE 3
humidity=normal windy=FALSE 4
humidity=normal play=yes 6
windy=TRUE play=yes 3
windy=TRUE play=no 3
windy=FALSE play=yes 6
Size of set of large itemsets L(3): 4
Large Itemsets L(3):
outlook=sunny humidity=high play=no 3
outlook=rainy windy=FALSE play=yes 3
temperature=cool humidity=normal play=yes 3
humidity=normal windy=FALSE play=yes 4
Best rules found:
```

```
<conf:(1)> lift:(1.56) lev:(0.1) [1] conv:(1.43)
 1. outlook=overcast 4 ==> play=yes 4
 2. temperature=cool 4 ==> humidity=normal 4
                                                <conf:(1)> lift:(2) lev:(0.14) [2] conv:(2)
 3. humidity=normal windy=FALSE 4 ==> play=yes 4
                                                    <conf:(1)> lift:(1.56) lev:(0.1) [1] conv:(1.43)
 4. outlook=sunny play=no 3 ==> humidity=high 3
                                                   <conf:(1)> lift:(2) lev:(0.11) [1] conv:(1.5)
 5. outlook=sunny humidity=high 3 ==> play=no 3
                                                   <conf:(1)> lift:(2.8) lev:(0.14) [1] conv:(1.93)
                                                  <conf:(1)> lift:(1.75) lev:(0.09) [1] conv:(1.29)
 6. outlook=rainy play=yes 3 ==> windy=FALSE 3
 7. outlook=rainy windy=FALSE 3 ==> play=yes 3
                                                  <conf:(1)> lift:(1.56) lev:(0.08) [1] conv:(1.07)
 8. temperature=cool play=yes 3 ==> humidity=normal 3
                                                         <conf:(1)> lift:(2) lev:(0.11) [1] conv:(1.5)
 9. humidity=normal 7 ==> play=yes 6 <conf:(0.86)> lift:(1.33) lev:(0.11) [1] conv:(1.25)
10. play=no 5 ==> humidity=high 4 <conf:(0.8)> lift:(1.6) lev:(0.11) [1] conv:(1.25)
MinMetric 2
=== Run information ===
              weka.associations.Apriori -I -N 10 -T 0 -C 2.0 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Scheme:
             weather.symbolic
Relation:
Instances:
              14
Attributes:
              outlook
              temperature
              humidity
              windy
              play
=== Associator model (full training set) ===
Apriori
======
```

```
Minimum support: 0.1 (1 instances)
Minimum metric <confidence>: 2
Number of cycles performed: 18
Generated sets of large itemsets:
Size of set of large itemsets L(1): 12
Large Itemsets L(1):
outlook=sunny 5
outlook=overcast 4
outlook=rainy 5
temperature=hot 4
temperature=mild 6
temperature=cool 4
humidity=high 7
humidity=normal 7
windy=TRUE 6
windy=FALSE 8
play=yes 9
play=no 5
Size of set of large itemsets L(2): 54
Large Itemsets L(2):
outlook=sunny temperature=hot 2
outlook=sunny temperature=mild 2
outlook=sunny temperature=cool 1
outlook=sunny humidity=high 3
outlook=sunny humidity=normal 2
outlook=sunny windy=TRUE 2
outlook=sunny windy=FALSE 3
```

```
outlook=sunny play=yes 2
outlook=sunny play=no 3
outlook=overcast temperature=hot 2
outlook=overcast temperature=mild 1
outlook=overcast temperature=cool 1
outlook=overcast humidity=high 2
outlook=overcast humidity=normal 2
outlook=overcast windy=TRUE 2
outlook=overcast windy=FALSE 2
outlook=overcast play=yes 4
outlook=rainy temperature=mild 3
outlook=rainy temperature=cool 2
outlook=rainy humidity=high 2
outlook=rainy humidity=normal 3
outlook=rainy windy=TRUE 2
outlook=rainy windy=FALSE 3
outlook=rainy play=yes 3
outlook=rainy play=no 2
temperature=hot humidity=high 3
temperature=hot humidity=normal 1
temperature=hot windy=TRUE 1
temperature=hot windy=FALSE 3
temperature=hot play=yes 2
temperature=hot play=no 2
temperature=mild humidity=high 4
temperature=mild humidity=normal 2
temperature=mild windy=TRUE 3
temperature=mild windy=FALSE 3
temperature=mild play=yes 4
temperature=mild play=no 2
temperature=cool humidity=normal 4
temperature=cool windy=TRUE 2
temperature=cool windy=FALSE 2
```

```
temperature=cool play=yes 3
temperature=cool play=no 1
humidity=high windy=TRUE 3
humidity=high windy=FALSE 4
humidity=high play=yes 3
humidity=high play=no 4
humidity=normal windy=TRUE 3
humidity=normal windy=FALSE 4
humidity=normal play=yes 6
humidity=normal play=no 1
windy=TRUE play=yes 3
windy=TRUE play=no 3
windy=FALSE play=yes 6
windy=FALSE play=no 2
Size of set of large itemsets L(3): 96
Large Itemsets L(3):
outlook=sunny temperature=hot humidity=high 2
outlook=sunny temperature=hot windy=TRUE 1
outlook=sunny temperature=hot windy=FALSE 1
outlook=sunny temperature=hot play=no 2
outlook=sunny temperature=mild humidity=high 1
outlook=sunny temperature=mild humidity=normal 1
outlook=sunny temperature=mild windy=TRUE 1
outlook=sunny temperature=mild windy=FALSE 1
outlook=sunny temperature=mild play=yes 1
outlook=sunny temperature=mild play=no 1
outlook=sunny temperature=cool humidity=normal 1
outlook=sunny temperature=cool windy=FALSE 1
outlook=sunny temperature=cool play=yes 1
outlook=sunny humidity=high windy=TRUE 1
outlook=sunny humidity=high windy=FALSE 2
```

```
outlook=sunny humidity=high play=no 3
outlook=sunny humidity=normal windy=TRUE 1
outlook=sunny humidity=normal windy=FALSE 1
outlook=sunny humidity=normal play=yes 2
outlook=sunny windy=TRUE play=yes 1
outlook=sunny windy=TRUE play=no 1
outlook=sunny windy=FALSE play=yes 1
outlook=sunny windy=FALSE play=no 2
outlook=overcast temperature=hot humidity=high 1
outlook=overcast temperature=hot humidity=normal 1
outlook=overcast temperature=hot windy=FALSE 2
outlook=overcast temperature=hot play=yes 2
outlook=overcast temperature=mild humidity=high 1
outlook=overcast temperature=mild windy=TRUE 1
outlook=overcast temperature=mild play=yes 1
outlook=overcast temperature=cool humidity=normal 1
outlook=overcast temperature=cool windy=TRUE 1
outlook=overcast temperature=cool play=yes 1
outlook=overcast humidity=high windy=TRUE 1
outlook=overcast humidity=high windy=FALSE 1
outlook=overcast humidity=high play=yes 2
outlook=overcast humidity=normal windy=TRUE 1
outlook=overcast humidity=normal windy=FALSE 1
outlook=overcast humidity=normal play=yes 2
outlook=overcast windy=TRUE play=yes 2
outlook=overcast windy=FALSE play=yes 2
outlook=rainy temperature=mild humidity=high 2
outlook=rainy temperature=mild humidity=normal 1
outlook=rainy temperature=mild windy=TRUE 1
outlook=rainy temperature=mild windy=FALSE 2
outlook=rainy temperature=mild play=yes 2
outlook=rainy temperature=mild play=no 1
outlook=rainy temperature=cool humidity=normal 2
```

```
outlook=rainy temperature=cool windy=TRUE 1
outlook=rainy temperature=cool windy=FALSE 1
outlook=rainy temperature=cool play=yes 1
outlook=rainy temperature=cool play=no 1
outlook=rainy humidity=high windy=TRUE 1
outlook=rainy humidity=high windy=FALSE 1
outlook=rainy humidity=high play=yes 1
outlook=rainy humidity=high play=no 1
outlook=rainy humidity=normal windy=TRUE 1
outlook=rainy humidity=normal windy=FALSE 2
outlook=rainy humidity=normal play=yes 2
outlook=rainy humidity=normal play=no 1
outlook=rainy windy=TRUE play=no 2
outlook=rainy windy=FALSE play=yes 3
temperature=hot humidity=high windy=TRUE 1
temperature=hot humidity=high windy=FALSE 2
temperature=hot humidity=high play=yes 1
temperature=hot humidity=high play=no 2
temperature=hot humidity=normal windy=FALSE 1
temperature=hot humidity=normal play=yes 1
temperature=hot windy=TRUE play=no 1
temperature=hot windy=FALSE play=yes 2
temperature=hot windy=FALSE play=no 1
temperature=mild humidity=high windy=TRUE 2
temperature=mild humidity=high windy=FALSE 2
temperature=mild humidity=high play=yes 2
temperature=mild humidity=high play=no 2
temperature=mild humidity=normal windy=TRUE 1
temperature=mild humidity=normal windy=FALSE 1
temperature=mild humidity=normal play=yes 2
temperature=mild windy=TRUE play=yes 2
temperature=mild windy=TRUE play=no 1
temperature=mild windy=FALSE play=yes 2
```

```
temperature=mild windy=FALSE play=no 1
temperature=cool humidity=normal windy=TRUE 2
temperature=cool humidity=normal windy=FALSE 2
temperature=cool humidity=normal play=yes 3
temperature=cool humidity=normal play=no 1
temperature=cool windy=TRUE play=yes 1
temperature=cool windy=TRUE play=no 1
temperature=cool windy=FALSE play=yes 2
humidity=high windy=TRUE play=yes 1
humidity=high windy=TRUE play=no 2
humidity=high windy=FALSE play=yes 2
humidity=high windy=FALSE play=no 2
humidity=normal windy=TRUE play=yes 2
humidity=normal windy=TRUE play=no 1
humidity=normal windy=FALSE play=yes 4
Size of set of large itemsets L(4): 64
Large Itemsets L(4):
outlook=sunny temperature=hot humidity=high windy=TRUE 1
outlook=sunny temperature=hot humidity=high windy=FALSE 1
outlook=sunny temperature=hot humidity=high play=no 2
outlook=sunny temperature=hot windy=TRUE play=no 1
outlook=sunny temperature=hot windy=FALSE play=no 1
outlook=sunny temperature=mild humidity=high windy=FALSE 1
outlook=sunny temperature=mild humidity=high play=no 1
outlook=sunny temperature=mild humidity=normal windy=TRUE 1
outlook=sunny temperature=mild humidity=normal play=yes 1
outlook=sunny temperature=mild windy=TRUE play=yes 1
outlook=sunny temperature=mild windy=FALSE play=no 1
outlook=sunny temperature=cool humidity=normal windy=FALSE 1
outlook=sunny temperature=cool humidity=normal play=yes 1
outlook=sunny temperature=cool windy=FALSE play=yes 1
```

```
outlook=sunny humidity=high windy=TRUE play=no 1
outlook=sunny humidity=high windy=FALSE play=no 2
outlook=sunny humidity=normal windy=TRUE play=yes 1
outlook=sunny humidity=normal windy=FALSE play=yes 1
outlook=overcast temperature=hot humidity=high windy=FALSE 1
outlook=overcast temperature=hot humidity=high play=yes 1
outlook=overcast temperature=hot humidity=normal windy=FALSE 1
outlook=overcast temperature=hot humidity=normal play=yes 1
outlook=overcast temperature=hot windy=FALSE play=yes 2
outlook=overcast temperature=mild humidity=high windy=TRUE 1
outlook=overcast temperature=mild humidity=high play=yes 1
outlook=overcast temperature=mild windy=TRUE play=yes 1
outlook=overcast temperature=cool humidity=normal windy=TRUE 1
outlook=overcast temperature=cool humidity=normal play=yes 1
outlook=overcast temperature=cool windy=TRUE play=yes 1
outlook=overcast humidity=high windy=TRUE play=yes 1
outlook=overcast humidity=high windy=FALSE play=yes 1
outlook=overcast humidity=normal windy=TRUE play=yes 1
outlook=overcast humidity=normal windy=FALSE play=yes 1
outlook=rainy temperature=mild humidity=high windy=TRUE 1
outlook=rainy temperature=mild humidity=high windy=FALSE 1
outlook=rainy temperature=mild humidity=high play=yes 1
outlook=rainy temperature=mild humidity=high play=no 1
outlook=rainy temperature=mild humidity=normal windy=FALSE 1
outlook=rainy temperature=mild humidity=normal play=yes 1
outlook=rainy temperature=mild windy=TRUE play=no 1
outlook=rainy temperature=mild windy=FALSE play=yes 2
outlook=rainy temperature=cool humidity=normal windy=TRUE 1
outlook=rainy temperature=cool humidity=normal windy=FALSE 1
outlook=rainy temperature=cool humidity=normal play=yes 1
outlook=rainy temperature=cool humidity=normal play=no 1
outlook=rainy temperature=cool windy=TRUE play=no 1
outlook=rainy temperature=cool windy=FALSE play=yes 1
```

```
outlook=rainy humidity=high windy=TRUE play=no 1
outlook=rainy humidity=high windy=FALSE play=yes 1
outlook=rainy humidity=normal windy=TRUE play=no 1
outlook=rainy humidity=normal windy=FALSE play=yes 2
temperature=hot humidity=high windy=TRUE play=no 1
temperature=hot humidity=high windy=FALSE play=yes 1
temperature=hot humidity=high windy=FALSE play=no 1
temperature=hot humidity=normal windy=FALSE play=yes 1
temperature=mild humidity=high windy=TRUE play=yes 1
temperature=mild humidity=high windy=TRUE play=no 1
temperature=mild humidity=high windy=FALSE play=yes 1
temperature=mild humidity=high windy=FALSE play=no 1
temperature=mild humidity=normal windy=TRUE play=yes 1
temperature=mild humidity=normal windy=FALSE play=yes 1
temperature=cool humidity=normal windy=TRUE play=yes 1
temperature=cool humidity=normal windy=TRUE play=no 1
temperature=cool humidity=normal windy=FALSE play=yes 2
Size of set of large itemsets L(5): 14
Large Itemsets L(5):
outlook=sunny temperature=hot humidity=high windy=TRUE play=no 1
outlook=sunny temperature=hot humidity=high windy=FALSE play=no 1
outlook=sunny temperature=mild humidity=high windy=FALSE play=no 1
outlook=sunny temperature=mild humidity=normal windy=TRUE play=yes 1
outlook=sunny temperature=cool humidity=normal windy=FALSE play=yes 1
outlook=overcast temperature=hot humidity=high windy=FALSE play=yes 1
outlook=overcast temperature=hot humidity=normal windy=FALSE play=yes 1
outlook=overcast temperature=mild humidity=high windy=TRUE play=yes 1
outlook=overcast temperature=cool humidity=normal windy=TRUE play=yes 1
outlook=rainy temperature=mild humidity=high windy=TRUE play=no 1
outlook=rainy temperature=mild humidity=high windy=FALSE play=yes 1
outlook=rainy temperature=mild humidity=normal windy=FALSE play=yes 1
```

```
outlook=rainy temperature=cool humidity=normal windy=TRUE play=no 1
outlook=rainy temperature=cool humidity=normal windy=FALSE play=yes 1

Best rules found:

Problem #3
```

Home

```
@relation supermarket
                                                                @attribute 'department109' { t}
@attribute 'department1' { t}
                                                                @attribute 'department110' { t}
@attribute 'department2' { t}
                                                                @attribute 'department111' { t}
@attribute 'department3' { t}
                                                                @attribute 'department112' { t}
@attribute 'department4' { t}
                                                                @attribute 'department113' { t}
@attribute 'department5' { t}
                                                                @attribute 'department114' { t}
                                                                @attribute 'health food bulk' { t}
@attribute 'department6' { t}
@attribute 'department7' { t}
                                                                @attribute 'department116' { t}
@attribute 'department8' { t}
                                                                @attribute 'department117' { t}
@attribute 'department9' { t}
                                                                @attribute 'department118' { t}
@attribute 'grocery misc' { t}
                                                                @attribute 'department119' { t}
@attribute 'department11' { t}
                                                                @attribute 'department120' { t}
                                                                @attribute 'bake off products' { t}
@attribute 'baby needs' { t}
@attribute 'bread and cake' { t}
                                                                @attribute 'department122' { t}
@attribute 'baking needs' { t}
                                                                @attribute 'department123' { t}
@attribute 'coupons' { t}
                                                                @attribute 'department124' { t}
@attribute 'juice-sat-cord-ms' { t}
                                                                @attribute 'department125' { t}
@attribute 'tea' { t}
                                                                @attribute 'department126' { t}
@attribute 'biscuits' { t}
                                                                @attribute 'department127' { t}
@attribute 'canned fish-meat' { t}
                                                                @attribute 'department128' { t}
```

```
@attribute 'canned fruit' { t}
                                                                @attribute 'department129' { t}
@attribute 'canned vegetables' { t}
                                                               @attribute 'department130' { t}
@attribute 'breakfast food' { t}
                                                               @attribute 'small goods2' { t}
                                                               @attribute 'offal' { t}
@attribute 'cigs-tobacco pkts' { t}
@attribute 'cigarette cartons' { t}
                                                               @attribute 'mutton' { t}
                                                               @attribute 'trim pork' { t}
@attribute 'cleaners-polishers' { t}
                                                               @attribute 'trim lamb' { t}
@attribute 'coffee' { t}
@attribute 'sauces-gravy-pkle' { t}
                                                               @attribute 'imported cheese' { t}
@attribute 'confectionary' { t}
                                                               @attribute 'department137' { t}
@attribute 'puddings-deserts' { t}
                                                               @attribute 'department138' { t}
@attribute 'dishcloths-scour' { t}
                                                               @attribute 'department139' { t}
@attribute 'deod-disinfectant' { t}
                                                               @attribute 'department140' { t}
@attribute 'frozen foods' { t}
                                                               @attribute 'department141' { t}
@attribute 'razor blades' { t}
                                                               @attribute 'department142' { t}
@attribute 'fuels-garden aids' { t}
                                                               @attribute 'department143' { t}
@attribute 'spices' { t}
                                                               @attribute 'department144' { t}
@attribute 'jams-spreads' { t}
                                                               @attribute 'department145' { t}
@attribute 'insecticides' { t}
                                                               @attribute 'department146' { t}
@attribute 'pet foods' { t}
                                                               @attribute 'department147' { t}
@attribute 'laundry needs' { t}
                                                               @attribute 'department148' { t}
@attribute 'party snack foods' { t}
                                                               @attribute 'department149' { t}
@attribute 'tissues-paper prd' { t}
                                                               @attribute 'department150' { t}
@attribute 'wrapping' { t}
                                                               @attribute 'department151' { t}
@attribute 'dried vegetables' { t}
                                                               @attribute 'department152' { t}
@attribute 'pkt-canned soup' { t}
                                                               @attribute 'department153' { t}
@attribute 'soft drinks' { t}
                                                               @attribute 'department154' { t}
@attribute 'health food other' { t}
                                                               @attribute 'department155' { t}
@attribute 'beverages hot' { t}
                                                               @attribute 'department156' { t}
@attribute 'health&beauty misc' { t}
                                                               @attribute 'department157' { t}
@attribute 'deodorants-soap' { t}
                                                               @attribute 'department158' { t}
@attribute 'mens toiletries' { t}
                                                               @attribute 'department159' { t}
@attribute 'medicines' { t}
                                                               @attribute 'department160' { t}
@attribute 'haircare' { t}
                                                               @attribute 'department161' { t}
                                                               @attribute 'department162' { t}
@attribute 'dental needs' { t}
@attribute 'lotions-creams' { t}
                                                               @attribute 'department163' { t}
@attribute 'sanitary pads' { t}
                                                               @attribute 'department164' { t}
@attribute 'cough-cold-pain' { t}
                                                               @attribute 'department165' { t}
@attribute 'department57' { t}
                                                               @attribute 'department166' { t}
```

```
@attribute 'meat misc' { t}
                                                                @attribute 'department167' { t}
@attribute 'cheese' { t}
                                                                @attribute 'department168' { t}
@attribute 'chickens' { t}
                                                                @attribute 'department169' { t}
                                                                @attribute 'department170' { t}
@attribute 'milk-cream' { t}
@attribute 'cold-meats' { t}
                                                               @attribute 'department171' { t}
@attribute 'deli gourmet' { t}
                                                               @attribute 'department172' { t}
@attribute 'margarine' { t}
                                                                @attribute 'department173' { t}
@attribute 'salads' { t}
                                                                @attribute 'department174' { t}
@attribute 'small goods' { t}
                                                                @attribute 'department175' { t}
@attribute 'dairy foods' { t}
                                                                @attribute 'department176' { t}
@attribute 'fruit drinks' { t}
                                                                @attribute 'department177' { t}
@attribute 'delicatessen misc' { t}
                                                               @attribute 'department178' { t}
@attribute 'department70' { t}
                                                               @attribute 'department179' { t}
@attribute 'beef' { t}
                                                                @attribute 'casks white wine' { t}
@attribute 'hogget' { t}
                                                                @attribute 'casks red wine' { t}
@attribute 'lamb' { t}
                                                                @attribute '750ml white nz' { t}
@attribute 'pet food' { t}
                                                                @attribute '750ml red nz' { t}
@attribute 'pork' { t}
                                                               @attribute '750ml white imp' { t}
@attribute 'poultry' { t}
                                                               @attribute '750ml red imp' { t}
@attribute 'veal' { t}
                                                               @attribute 'sparkling nz' { t}
@attribute 'gourmet meat' { t}
                                                               @attribute 'sparkling imp' { t}
@attribute 'department79' { t}
                                                               @attribute 'brew kits/accesry' { t}
@attribute 'department80' { t}
                                                               @attribute 'department189' { t}
@attribute 'department81' { t}
                                                               @attribute 'port and sherry' { t}
@attribute 'produce misc' { t}
                                                               @attribute 'ctrled label wine' { t}
@attribute 'fruit' { t}
                                                                @attribute 'department192' { t}
@attribute 'plants' { t}
                                                                @attribute 'department193' { t}
@attribute 'potatoes' { t}
                                                                @attribute 'department194' { t}
@attribute 'vegetables' { t}
                                                               @attribute 'department195' { t}
@attribute 'flowers' { t}
                                                               @attribute 'department196' { t}
@attribute 'department88' { t}
                                                               @attribute 'department197' { t}
@attribute 'department89' { t}
                                                               @attribute 'department198' { t}
@attribute 'variety misc' { t}
                                                               @attribute 'department199' { t}
@attribute 'brushware' { t}
                                                                @attribute 'non host support' { t}
@attribute 'electrical' { t}
                                                                @attribute 'department201' { t}
@attribute 'haberdashery' { t}
                                                               @attribute 'department202' { t}
@attribute 'kitchen' { t}
                                                               @attribute 'department203' { t}
@attribute 'manchester' { t}
                                                               @attribute 'department204' { t}
```

```
@attribute 'pantyhose' { t}
                                                               @attribute 'department205' { t}
@attribute 'plasticware' { t}
                                                               @attribute 'department206' { t}
@attribute 'department98' { t}
                                                               @attribute 'department207' { t}
@attribute 'stationary' { t}
                                                               @attribute 'department208' { t}
@attribute 'department100' { t}
                                                               @attribute 'department209' { t}
@attribute 'department101' { t}
                                                               @attribute 'department210' { t}
@attribute 'department102' { t}
                                                               @attribute 'department211' { t}
@attribute 'prepared meals' { t}
                                                               @attribute 'department212' { t}
@attribute 'preserving needs' { t}
                                                               @attribute 'department213' { t}
@attribute 'condiments' { t}
                                                               @attribute 'department214' { t}
@attribute 'cooking oils' { t}
                                                               @attribute 'department215' { t}
@attribute 'department107' { t}
                                                               @attribute 'department216' { t}
@attribute 'department108' { t}
                                                               @attribute 'total' { low, high} % low < 100
                                                               @data
```

Home

```
=== Run information ===
Scheme:
              weka.associations.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Relation:
              supermarket
Instances:
              4627
Attributes:
              217
              [list of attributes omitted]
=== Associator model (full training set) ===
Apriori
======
Minimum support: 0.15 (694 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 17
Generated sets of large itemsets:
Size of set of large itemsets L(1): 44
```

```
Size of set of large itemsets L(2): 380
Size of set of large itemsets L(3): 910
Size of set of large itemsets L(4): 633
Size of set of large itemsets L(5): 105
Size of set of large itemsets L(6): 1
Best rules found:
 1. biscuits=t frozen foods=t fruit=t total=high 788 ==> bread and cake=t 723
                                                                              <conf:(0.92)> lift:(1.27) lev:(0.03) [155] conv:(3.35)
 2. baking needs=t biscuits=t fruit=t total=high 760 ==> bread and cake=t 696
                                                                              <conf:(0.92)> lift:(1.27) lev:(0.03) [149] conv:(3.28)
 3. baking needs=t frozen foods=t fruit=t total=high 770 ==> bread and cake=t 705
                                                                                  <conf:(0.92)> lift:(1.27) lev:(0.03) [150] conv:(3.27)
 4. biscuits=t fruit=t vegetables=t total=high 815 ==> bread and cake=t 746
                                                                            <conf:(0.92)> lift:(1.27) lev:(0.03) [159] conv:(3.26)
 5. party snack foods=t fruit=t total=high 854 ==> bread and cake=t 779
                                                                         <conf:(0.91)> lift:(1.27) lev:(0.04) [164] conv:(3.15)
 6. biscuits=t frozen foods=t vegetables=t total=high 797 ==> bread and cake=t 725
                                                                                   <conf:(0.91)> lift:(1.26) lev:(0.03) [151] conv:(3.06)
 7. baking needs=t biscuits=t vegetables=t total=high 772 ==> bread and cake=t 701
                                                                                   <conf:(0.91)> lift:(1.26) lev:(0.03) [145] conv:(3.01)
 8. biscuits=t fruit=t total=high 954 ==> bread and cake=t 866
                                                              <conf:(0.91)> lift:(1.26) lev:(0.04) [179] conv:(3)
 9. frozen foods=t fruit=t vegetables=t total=high 834 ==> bread and cake=t 757
                                                                                <conf:(0.91)> lift:(1.26) lev:(0.03) [156] conv:(3)
10. frozen foods=t fruit=t total=high 969 ==> bread and cake=t 877 <conf:(0.91)> lift:(1.26) lev:(0.04) [179] conv:(2.92)
```