

Project 1 ‘Part 3-Association’

Sachin Samal (ECU ID 250008)

Oct 21, 2021

DATA MINING WITH R

Here in this part of project, I have decided to work on the association with my previous dataset from **Project 1** and **Project 2**. For this project, I have pulled the data from **English Premier League Results**.

WHAT IS ASSOCIATION?

It is the method to discover relationships between seemingly independent relational databases or other data repositories. It aims to observe frequently occurring patterns, correlations, or associations from datasets found in various kinds of databases such as relational databases, transactional databases, and other forms of repositories.

LETS START OUT ASSOCIATION MINING...

```
library(readxl)
Results <- read_excel("Results.xlsx")
str(Results)
```

LOADING MY EXCEL DATA INTO R ENVIRONMENT

```
## tibble [380 x 6] (S3: tbl_df/tbl/data.frame)
## $ Home_team: chr [1:380] "Arsenal" "Watford" "Chelsea" "Crystal Palace" ...
## $ Away_team: chr [1:380] "Leicester City" "Liverpool" "Burnley" "Huddersfield Town" ...
## $ Home_goal: num [1:380] 4 3 2 0 1 0 1 0 0 4 ...
## $ Away_goal: num [1:380] 3 3 3 3 0 0 0 2 2 0 ...
## $ Result    : chr [1:380] "H" "D" "A" "A" ...
## $ Season    : chr [1:380] "2017-2018" "2017-2018" "2017-2018" "2017-2018" ...

summary(Results$Home_goal)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.000	1.000	1.000	1.532	2.000	7.000

```
library(arules)
```

ASSOCIATION MINING WITH R

```
## Loading required package: Matrix
```

```
##
```

```
## Attaching package: 'arules'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      abbreviate, write
```

```
library(arulesViz)
```

Here, I have added some new libraries. Now, I need to convert the data for association analysis.

Now, I need to look at the columns to see if I can convert them into factors (or Boolean values) for analysis. This is because if not, I have to use different methods like “*Kruskal Wallis test*” or “*Chi-square test*” based on the requirement of my dataset.

```
colnames(Results)[c(1,2,3,4,5)]
```

```
## [1] "Home_team" "Away_team" "Home_goal" "Away_goal" "Result"
```

I don't find anything that needs much attention. So , I will continue with the process as it needs.

CREATE TRANSACTION

I will let R do the default discretization to the rest of the data. This is because I could not come up with better cutoffs for what is left in the dataset.

```
library(ggplot2)
```

```
trans <- transactions(Results)
```

```
## Warning: Column(s) 1, 2, 3, 4, 5, 6 not logical or factor. Applying default
```

```
## discretization (see '? discretizeDF').
```

```
## Warning in discretize(x = c(3, 3, 3, 3, 0, 0, 0, 2, 2, 0, 4, 2, 1, 0, 0, : The calculated breaks are
```

```
## Only unique breaks are used reducing the number of intervals. Look at ? discretize for details.
```

The conversion gives a warning because only discrete features (factor and logical) can be directly translated into items. Continuous features need to be discretized first.

```
summary(Results[5])
```

```
##      Result
```

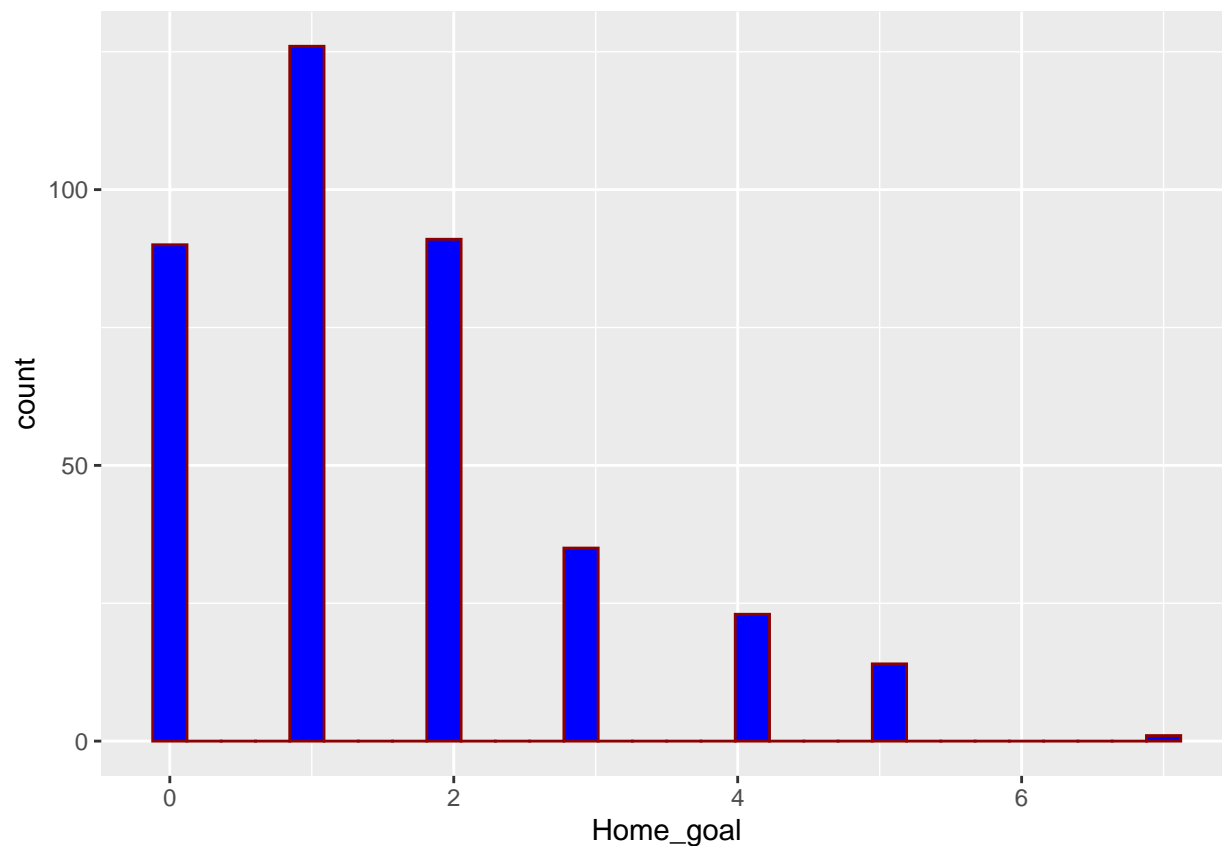
```
## Length:380
```

```
## Class :character
```

```
## Mode  :character
```

```
ggplot(Results, aes(Home_goal)) + geom_histogram(fill='blue', color='darkred')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



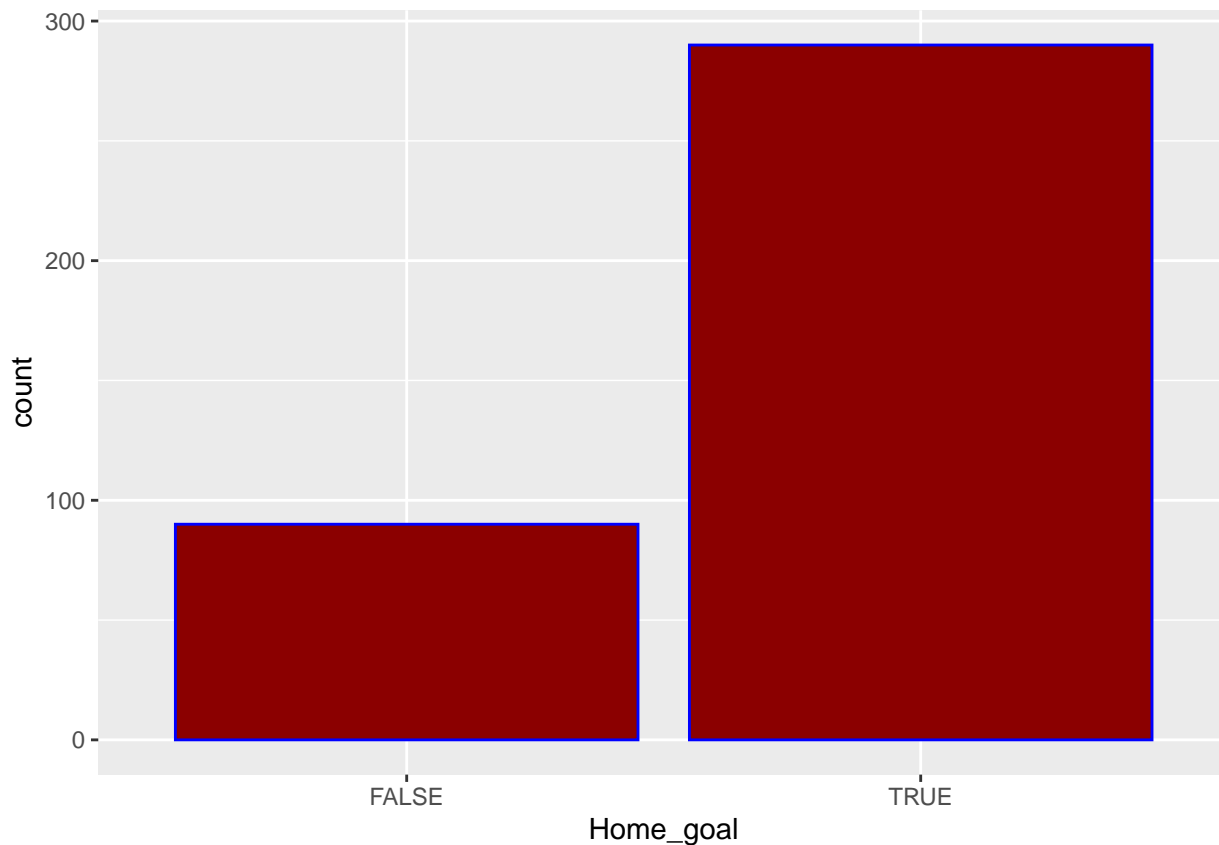
```
table(Results$Home_goal)
```

```
##
##  0  1  2  3  4  5  7
## 90 126 91 35 23 14  1
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:arules':
##
## intersect, recode, setdiff, setequal, union
## The following objects are masked from 'package:stats':
##
## filter, lag
## The following objects are masked from 'package:base':
##
## intersect, setdiff, setequal, union
```

```
Results <- Results %>% mutate(Home_goal = Home_goal > 0)
ggplot(Results, aes(Home_goal)) + geom_bar(fill='darkred', color='blue')
```



```
table(Results$Home_goal)
```

```
##
## FALSE  TRUE
##    90   290
```

The condition we had was if ($\text{Home_goal} > 0$). From the table, we can conclude that home goals are more frequently happening than away goals.

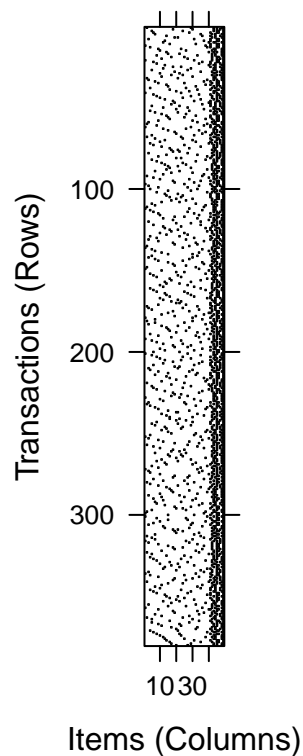
INSPECT TRANSACTION Now, Lets run the transaction and see how the data has cleaned.

```
summary(trans)
```

```
## transactions as itemMatrix in sparse format with
## 380 rows (elements/itemsets/transactions) and
## 49 columns (items) and a density of 0.122449
##
## most frequent items:
## Season=2017-2018 Away_goal=[1,6] Result=H Home_goal=[2,7]
##           380           244           173           164
## Away_goal=[0,1) (Other)
##           136           1183
##
## element (itemset/transaction) length distribution:
## sizes
## 6
```

```
## 380
##
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##         6         6         6         6         6         6
##
## includes extended item information - examples:
##                                labels variables          levels
## 1      Home_team=AFC Bournemouth Home_team      AFC Bournemouth
## 2              Home_team=Arsenal Home_team      Arsenal
## 3 Home_team=Brighton and Hove Albion Home_team Brighton and Hove Albion
##
## includes extended transaction information - examples:
## transactionID
## 1             1
## 2             2
## 3             3
```

```
library(colorRamps)
#plotting image for transaction
image(trans, fill="red")
```



```
frequentItems <- apriori(trans, parameter=list(target = "frequent"))
```

THE MOST FREQUENT ITEMS ON THE DATASET

```
## Apriori
##
## Parameter specification:
```

```

## confidence minval smax arem aval originalSupport maxtime support minlen
##          NA    0.1    1 none FALSE          TRUE      5    0.1    1
## maxlen          target ext
##    10 frequent itemsets TRUE
##
## Algorithmic control:
## filter tree heap memopt load sort verbose
##    0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 38
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[49 item(s), 380 transaction(s)] done [0.00s].
## sorting and recoding items ... [9 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 done [0.00s].
## sorting transactions ... done [0.00s].
## writing ... [53 set(s)] done [0.00s].
## creating S4 object ... done [0.00s].

```

```
inspect(frequentItems)
```

	items	support	transIdenticalToItemsets	count
## [1]	{Home_goal=[0,1]}	0.2368421	0.0000000	90
## [2]	{Result=D}	0.2605263	0.0000000	99
## [3]	{Result=A}	0.2842105	0.0000000	108
## [4]	{Home_goal=[1,2]}	0.3315789	0.0000000	126
## [5]	{Away_goal=[0,1]}	0.3578947	0.0000000	136
## [6]	{Home_goal=[2,7]}	0.4315789	0.0000000	164
## [7]	{Result=H}	0.4552632	0.0000000	173
## [8]	{Away_goal=[1,6]}	0.6421053	0.0000000	244
## [9]	{Season=2017-2018}	1.0000000	0.0000000	380
## [10]	{Home_goal=[0,1], Result=A}	0.1526316	0.0000000	58
## [11]	{Home_goal=[0,1], Away_goal=[1,6]}	0.1526316	0.0000000	58
## [12]	{Home_goal=[0,1], Season=2017-2018}	0.2368421	0.0000000	90
## [13]	{Home_goal=[1,2], Result=D}	0.1184211	0.0000000	45
## [14]	{Away_goal=[1,6], Result=D}	0.1763158	0.0000000	67
## [15]	{Result=D, Season=2017-2018}	0.2605263	0.0000000	99
## [16]	{Away_goal=[1,6], Result=A}	0.2842105	0.0000000	108
## [17]	{Result=A, Season=2017-2018}	0.2842105	0.0000000	108
## [18]	{Home_goal=[1,2], Away_goal=[0,1]}	0.1157895	0.0000000	44
## [19]	{Home_goal=[1,2], Result=H}	0.1157895	0.0000000	44
## [20]	{Home_goal=[1,2], Away_goal=[1,6]}	0.2157895	0.0000000	82
## [21]	{Home_goal=[1,2],			

##	Season=2017-2018}	0.3315789	0.0000000	126
##	[22] {Home_goal=[2,7],			
##	Away_goal=[0,1]}	0.1578947	0.0000000	60
##	[23] {Away_goal=[0,1],			
##	Result=H}	0.2736842	0.0000000	104
##	[24] {Away_goal=[0,1],			
##	Season=2017-2018}	0.3578947	0.0000000	136
##	[25] {Home_goal=[2,7],			
##	Result=H}	0.3394737	0.0000000	129
##	[26] {Home_goal=[2,7],			
##	Away_goal=[1,6]}	0.2736842	0.0000000	104
##	[27] {Home_goal=[2,7],			
##	Season=2017-2018}	0.4315789	0.0000000	164
##	[28] {Away_goal=[1,6],			
##	Result=H}	0.1815789	0.0000000	69
##	[29] {Result=H,			
##	Season=2017-2018}	0.4552632	0.0000000	173
##	[30] {Away_goal=[1,6],			
##	Season=2017-2018}	0.6421053	0.0000000	244
##	[31] {Home_goal=[0,1],			
##	Away_goal=[1,6],			
##	Result=A}	0.1526316	0.0000000	58
##	[32] {Home_goal=[0,1],			
##	Result=A,			
##	Season=2017-2018}	0.1526316	0.0000000	58
##	[33] {Home_goal=[0,1],			
##	Away_goal=[1,6],			
##	Season=2017-2018}	0.1526316	0.0000000	58
##	[34] {Home_goal=[1,2],			
##	Away_goal=[1,6],			
##	Result=D}	0.1184211	0.0000000	45
##	[35] {Home_goal=[1,2],			
##	Result=D,			
##	Season=2017-2018}	0.1184211	0.0000000	45
##	[36] {Away_goal=[1,6],			
##	Result=D,			
##	Season=2017-2018}	0.1763158	0.0000000	67
##	[37] {Away_goal=[1,6],			
##	Result=A,			
##	Season=2017-2018}	0.2842105	0.0000000	108
##	[38] {Home_goal=[1,2],			
##	Away_goal=[0,1],			
##	Result=H}	0.1157895	0.0000000	44
##	[39] {Home_goal=[1,2],			
##	Away_goal=[0,1],			
##	Season=2017-2018}	0.1157895	0.0000000	44
##	[40] {Home_goal=[1,2],			
##	Result=H,			
##	Season=2017-2018}	0.1157895	0.0000000	44
##	[41] {Home_goal=[1,2],			
##	Away_goal=[1,6],			
##	Season=2017-2018}	0.2157895	0.0000000	82
##	[42] {Home_goal=[2,7],			
##	Away_goal=[0,1],			

```

##      Result=H}          0.1578947          0.0000000    60
## [43] {Home_goal=[2,7],
##      Away_goal=[0,1],
##      Season=2017-2018} 0.1578947          0.0000000    60
## [44] {Away_goal=[0,1],
##      Result=H,
##      Season=2017-2018} 0.2736842          0.0000000   104
## [45] {Home_goal=[2,7],
##      Away_goal=[1,6],
##      Result=H}          0.1815789          0.0000000    69
## [46] {Home_goal=[2,7],
##      Result=H,
##      Season=2017-2018} 0.3394737          0.0000000   129
## [47] {Home_goal=[2,7],
##      Away_goal=[1,6],
##      Season=2017-2018} 0.2736842          0.0000000   104
## [48] {Away_goal=[1,6],
##      Result=H,
##      Season=2017-2018} 0.1815789          0.0000000    69
## [49] {Home_goal=[0,1],
##      Away_goal=[1,6],
##      Result=A,
##      Season=2017-2018} 0.1526316          0.1526316    58
## [50] {Home_goal=[1,2],
##      Away_goal=[1,6],
##      Result=D,
##      Season=2017-2018} 0.1184211          0.1184211    45
## [51] {Home_goal=[1,2],
##      Away_goal=[0,1],
##      Result=H,
##      Season=2017-2018} 0.1157895          0.1157895    44
## [52] {Home_goal=[2,7],
##      Away_goal=[0,1],
##      Result=H,
##      Season=2017-2018} 0.1578947          0.1578947    60
## [53] {Home_goal=[2,7],
##      Away_goal=[1,6],
##      Result=H,
##      Season=2017-2018} 0.1815789          0.1815789    69

```

```

#calculating the frequent items
frequentItems

```

```

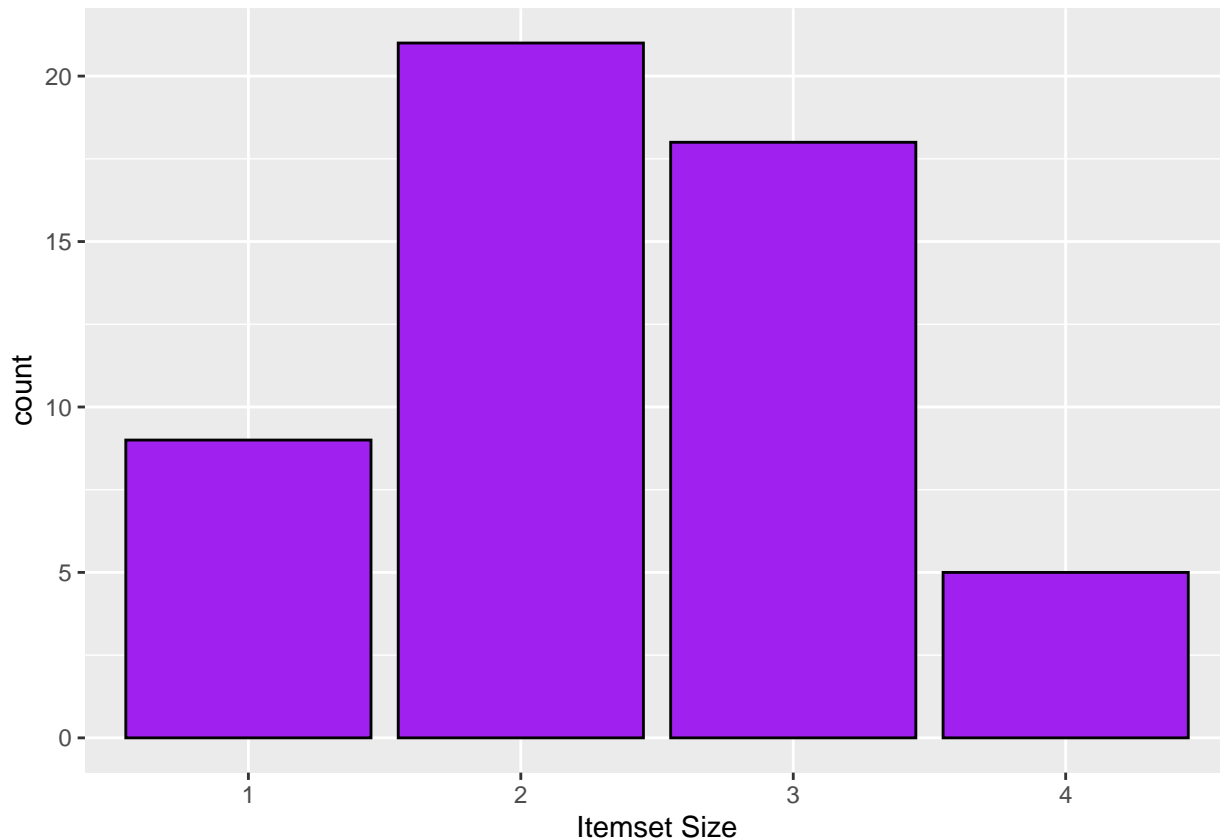
## set of 53 itemsets

```

```

ggplot(tibble(`Itemset Size` = factor(size(frequentItems))), aes(`Itemset Size`)) + geom_bar(fill = "purple")

```

Apriori WITH R - GENERATING RULES We will generate parameters support and confidence for rule mining and lift for interestingness evaluation.

Support indicates how frequently the itemset appears in the dataset.

Confidence is the proportion of the true positive of the rule.

Lets find out the rules using the apriori algorithm.

```
library(arules)
#association rules.
rules <- apriori(Results,
                 parameter = list(supp = 0.05, conf = 0.9,
                                target = "rules"))
```

```
## Warning: Column(s) 1, 2, 4, 5, 6 not logical or factor. Applying default
## discretization (see '? discretizeDF').
```

```
## Warning in discretize(x = c(3, 3, 3, 3, 0, 0, 0, 2, 2, 0, 4, 2, 1, 0, 0, : The calculated breaks are
## Only unique breaks are used reducing the number of intervals. Look at ? discretize for details.
```

```
## Apriori
```

```
##
```

```
## Parameter specification:
```

```
## confidence minval smax arem aval originalSupport maxtime support minlen
##          0.9    0.1    1 none FALSE                TRUE     5    0.05    1
```

```
## maxlen target ext
```

```
##          10 rules TRUE
```

```
##
## Algorithmic control:
## filter tree heap memopt load sort verbose
## 0.1 TRUE TRUE FALSE TRUE 2 TRUE
##
## Absolute minimum support count: 19
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[47 item(s), 380 transaction(s)] done [0.00s].
## sorting and recoding items ... [47 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 done [0.00s].
## writing ... [77 rule(s)] done [0.00s].
## creating S4 object ... done [0.00s].
```

The Apriori algorithm generated 10 rules with the given constraints (parameters). Lets dive into the Parameter Specification section of the output.

minval is the minimum value of the support an itemset should satisfy to be a part of a rule.

smax is the maximum support value for an itemset.

arem is an Additional Rule Evaluation Parameter (similar to lift).

aval is a logical indicating whether to return the additional rule evaluation measure selected with arem.

originalSupport is the traditional support value that consider both LHS and RHS items for calculating support. If you want to use only the LHS items for the calculation then you need to set this to FALSE.

maxtime is the maximum amount of time allowed to check for subsets.

minlen is the minimum number of items required in the rule.

maxlen is the maximum number of items that can be present in the rule.

```
#length
length(rules)
```

```
## [1] 77
```

```
#sorting the rules and printing it
rules.sorted <-sort(rules, by="lift")
inspect(rules.sorted)
```

	lhs	rhs	support	confidence	coverage
## [1]	{Home_goal,				
##	Away_goal=[0,1]}	=> {Result=H}	0.27368421	1	0.27368421 2.1
## [2]	{Home_goal,				
##	Away_goal=[0,1],				
##	Season=2017-2018}	=> {Result=H}	0.27368421	1	0.27368421 2.1
## [3]	{Result=A}	=> {Away_goal=[1,6]}	0.28421053	1	0.28421053 1.5
## [4]	{Home_goal,				
##	Result=D}	=> {Away_goal=[1,6]}	0.17631579	1	0.17631579 1.5
## [5]	{Home_goal,				
##	Result=A}	=> {Away_goal=[1,6]}	0.13157895	1	0.13157895 1.5
## [6]	{Result=A,				
##	Season=2017-2018}	=> {Away_goal=[1,6]}	0.28421053	1	0.28421053 1.5
## [7]	{Home_goal,				
##	Result=D,				

##	Season=2017-2018}	=> {Away_goal=[1,6]}	0.17631579	1	0.17631579	1.5
## [8]	{Home_goal,					
##	Result=A,					
##	Season=2017-2018}	=> {Away_goal=[1,6]}	0.13157895	1	0.13157895	1.5
## [9]	{Result=H}	=> {Home_goal}	0.45526316	1	0.45526316	1.3
## [10]	{Away_goal=[1,6],					
##	Result=D}	=> {Home_goal}	0.17631579	1	0.17631579	1.3
## [11]	{Away_goal=[0,1],					
##	Result=H}	=> {Home_goal}	0.27368421	1	0.27368421	1.3
## [12]	{Away_goal=[1,6],					
##	Result=H}	=> {Home_goal}	0.18157895	1	0.18157895	1.3
## [13]	{Result=H,					
##	Season=2017-2018}	=> {Home_goal}	0.45526316	1	0.45526316	1.3
## [14]	{Away_goal=[1,6],					
##	Result=D,					
##	Season=2017-2018}	=> {Home_goal}	0.17631579	1	0.17631579	1.3
## [15]	{Away_goal=[0,1],					
##	Result=H,					
##	Season=2017-2018}	=> {Home_goal}	0.27368421	1	0.27368421	1.3
## [16]	{Away_goal=[1,6],					
##	Result=H,					
##	Season=2017-2018}	=> {Home_goal}	0.18157895	1	0.18157895	1.3
## [17]	{}	=> {Season=2017-2018}	1.00000000	1	1.00000000	1.0
## [18]	{Away_team=Manchester City}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [19]	{Home_team=Huddersfield Town}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [20]	{Home_team=Swansea City}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [21]	{Home_team=Newcastle United}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [22]	{Away_team=Manchester United}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [23]	{Home_team=Southampton}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [24]	{Away_team=Chelsea}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [25]	{Home_team=Stoke City}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [26]	{Away_team=Tottenham Hotspur}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [27]	{Home_team=Burnley}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [28]	{Home_team=Brighton and Hove Albion}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [29]	{Home_team=AFC Bournemouth}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [30]	{Home_team=Everton}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [31]	{Home_team=Crystal Palace}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [32]	{Away_team=Burnley}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [33]	{Away_team=Brighton and Hove Albion}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [34]	{Away_team=Liverpool}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [35]	{Away_team=Crystal Palace}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [36]	{Away_team=Southampton}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [37]	{Away_team=West Bromwich Albion}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [38]	{Away_team=Watford}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [39]	{Home_team=Leicester City}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [40]	{Away_team=Arsenal}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [41]	{Home_team=West Ham United}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [42]	{Home_team=West Bromwich Albion}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [43]	{Away_team=Swansea City}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [44]	{Home_team=Watford}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [45]	{Home_team=Chelsea}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [46]	{Away_team=AFC Bournemouth}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [47]	{Away_team=Leicester City}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0
## [48]	{Away_team=Everton}	=> {Season=2017-2018}	0.05000000	1	0.05000000	1.0

```

## [49] {Home_team=Liverpool}          => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [50] {Away_team=Huddersfield Town}    => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [51] {Away_team=Stoke City}           => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [52] {Away_team=West Ham United}      => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [53] {Away_team=Newcastle United}     => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [54] {Home_team=Manchester United}    => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [55] {Home_team=Tottenham Hotspur}    => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [56] {Home_team=Manchester City}      => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [57] {Home_team=Arsenal}              => {Season=2017-2018} 0.05000000    1 0.05000000 1.00
## [58] {Result=D}                       => {Season=2017-2018} 0.26052632    1 0.26052632 1.00
## [59] {Result=A}                       => {Season=2017-2018} 0.28421053    1 0.28421053 1.00
## [60] {Away_goal=[0,1]}                => {Season=2017-2018} 0.35789474    1 0.35789474 1.00
## [61] {Result=H}                       => {Season=2017-2018} 0.45526316    1 0.45526316 1.00
## [62] {Away_goal=[1,6]}                => {Season=2017-2018} 0.64210526    1 0.64210526 1.00
## [63] {Home_goal}                      => {Season=2017-2018} 0.76315789    1 0.76315789 1.00
## [64] {Away_goal=[0,1],               => {Season=2017-2018} 0.08421053    1 0.08421053 1.00
##       Result=D}
## [65] {Away_goal=[1,6],               => {Season=2017-2018} 0.17631579    1 0.17631579 1.00
##       Result=D}
## [66] {Home_goal,                     => {Season=2017-2018} 0.17631579    1 0.17631579 1.00
##       Result=D}
## [67] {Away_goal=[1,6],               => {Season=2017-2018} 0.28421053    1 0.28421053 1.00
##       Result=A}
## [68] {Home_goal,                     => {Season=2017-2018} 0.13157895    1 0.13157895 1.00
##       Result=A}
## [69] {Away_goal=[0,1],               => {Season=2017-2018} 0.27368421    1 0.27368421 1.00
##       Result=H}
## [70] {Home_goal,                     => {Season=2017-2018} 0.27368421    1 0.27368421 1.00
##       Away_goal=[0,1]}
## [71] {Away_goal=[1,6],               => {Season=2017-2018} 0.18157895    1 0.18157895 1.00
##       Result=H}
## [72] {Home_goal,                     => {Season=2017-2018} 0.45526316    1 0.45526316 1.00
##       Result=H}
## [73] {Home_goal,                     => {Season=2017-2018} 0.48947368    1 0.48947368 1.00
##       Away_goal=[1,6]}
## [74] {Home_goal,                     => {Season=2017-2018} 0.17631579    1 0.17631579 1.00
##       Away_goal=[1,6],
##       Result=D}
## [75] {Home_goal,                     => {Season=2017-2018} 0.13157895    1 0.13157895 1.00
##       Away_goal=[1,6],
##       Result=A}
## [76] {Home_goal,                     => {Season=2017-2018} 0.27368421    1 0.27368421 1.00
##       Away_goal=[0,1],
##       Result=H}
## [77] {Home_goal,                     => {Season=2017-2018} 0.18157895    1 0.18157895 1.00
##       Away_goal=[1,6],
##       Result=H}

```

```
summary(rules)
```

```

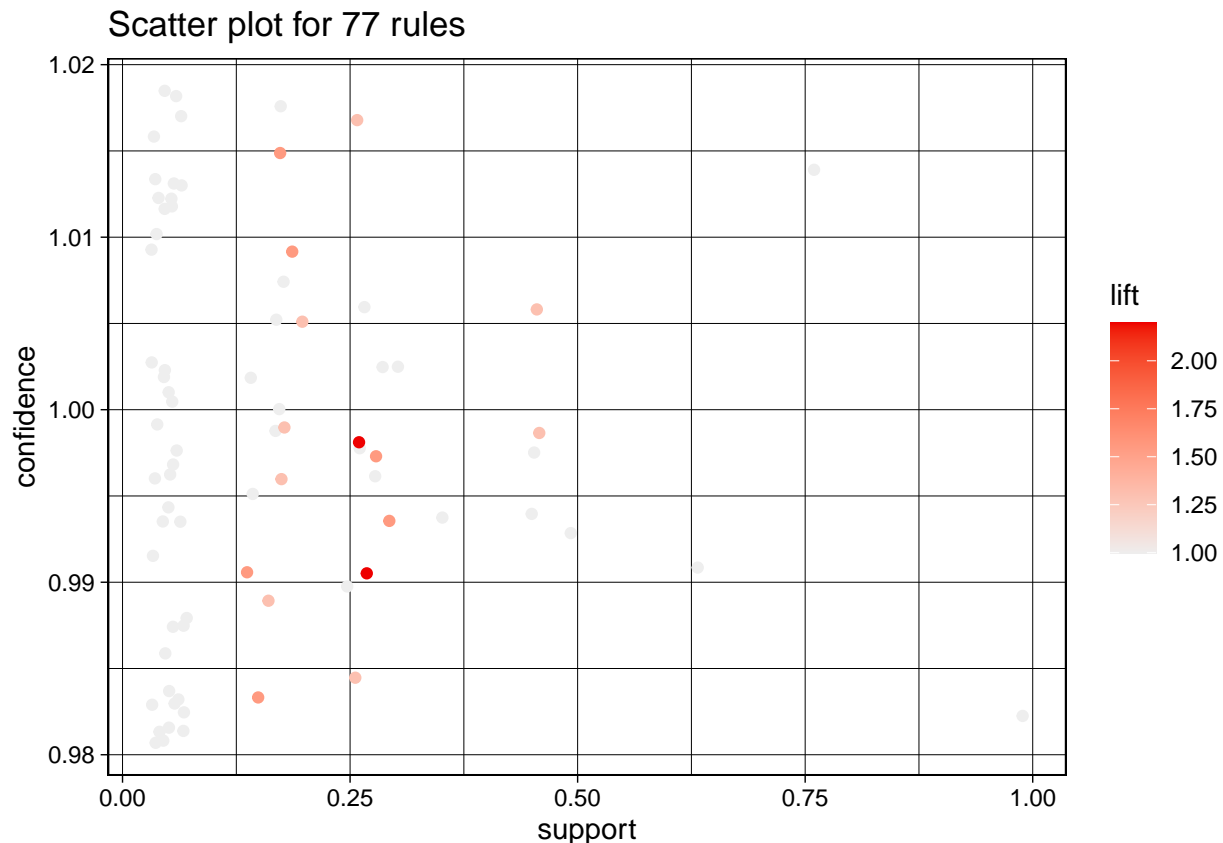
## set of 77 rules
##
## rule length distribution (lhs + rhs):sizes
##  1  2  3  4
##  1 48 18 10

```

```
##
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.000   2.000   2.000   2.481   3.000   4.000
##
## summary of quality measures:
##      support      confidence      coverage      lift
##      Min.    :0.0500   Min.    :1   Min.    :0.0500   Min.    :1.000
##      1st Qu.:0.0500   1st Qu.:1   1st Qu.:0.0500   1st Qu.:1.000
##      Median :0.0500   Median :1   Median :0.0500   Median :1.000
##      Mean   :0.1683   Mean   :1   Mean   :0.1683   Mean   :1.107
##      3rd Qu.:0.2737   3rd Qu.:1   3rd Qu.:0.2737   3rd Qu.:1.000
##      Max.   :1.0000   Max.   :1   Max.   :1.0000   Max.   :2.197
##      count
##      Min.    : 19.00
##      1st Qu.: 19.00
##      Median : 19.00
##      Mean    : 63.95
##      3rd Qu.:104.00
##      Max.    :380.00
##
## mining info:
##      data ntransactions support confidence
##      Results      380      0.05      0.9
```

```
#plot rules.sorted
plot(rules.sorted)
```

To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.



```
plot(rules, method = "graph", measure = "lift", shading = "confidence", engine = "htmlwidget")

## QStandardPaths: XDG_RUNTIME_DIR not set, defaulting to '/tmp/runtime-rstudio-user'
## TypeError: Attempting to change the setter of an unconfigurable property.
## TypeError: Attempting to change the setter of an unconfigurable property.
```

Sort by support

```
rules <- apriori(trans, parameter = list(support = 0.05, confidence = 0.9))

## Apriori
##
## Parameter specification:
## confidence minval smax arem aval originalSupport maxtime support minlen
##           0.9    0.1    1 none FALSE             TRUE      5    0.05      1
## maxlen target  ext
##       10  rules TRUE
##
## Algorithmic control:
## filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE     2    TRUE
##
## Absolute minimum support count: 19
##
```

```
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[49 item(s), 380 transaction(s)] done [0.00s].
## sorting and recoding items ... [49 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 done [0.00s].
## writing ... [101 rule(s)] done [0.00s].
## creating S4 object ... done [0.00s].
```

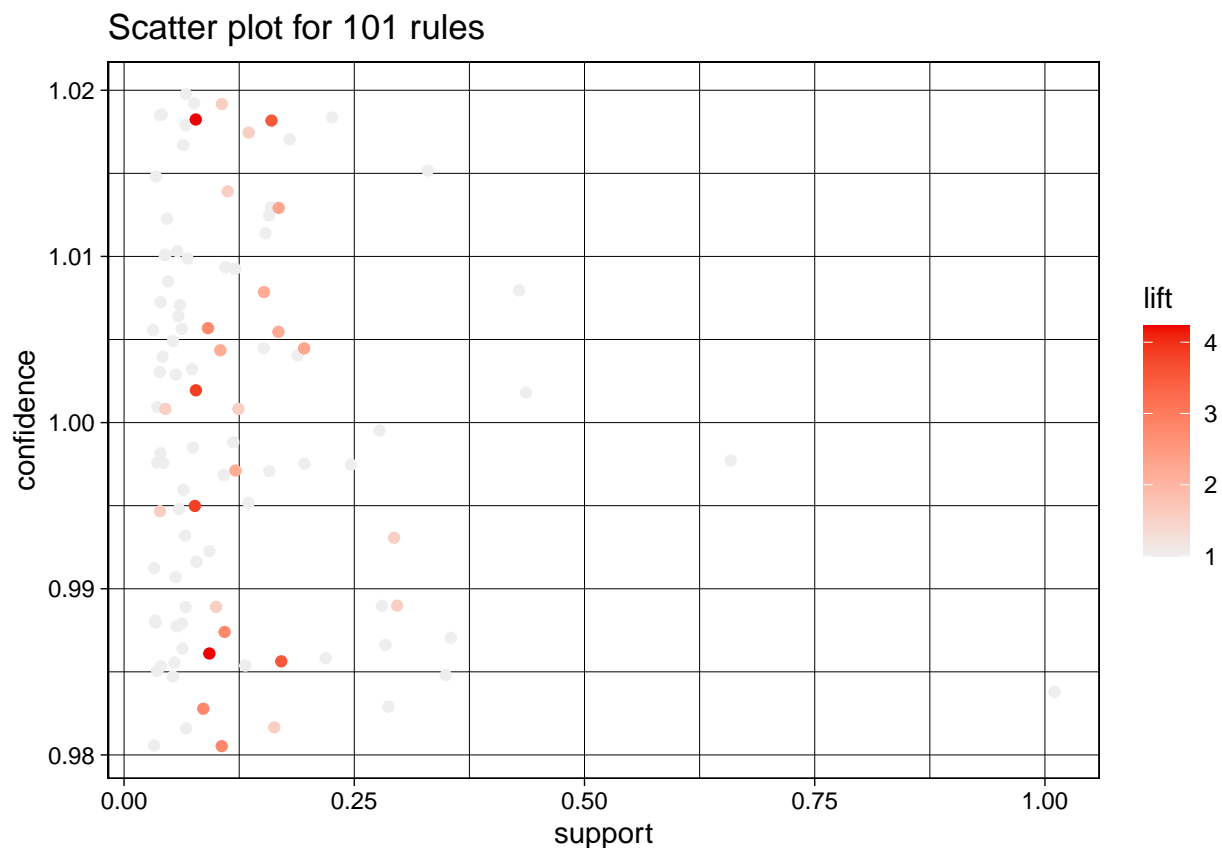
Lets inspect the top three rules sorted by confidence.

```
inspect(head(sort(rules, by = "confidence"), 3))
```

```
##      lhs                                rhs      support confidence
## [1] {}                                => {Season=2017-2018} 1.00      1
## [2] {Home_team=Arsenal}              => {Season=2017-2018} 0.05      1
## [3] {Away_team=Tottenham Hotspur} => {Season=2017-2018} 0.05      1
##      coverage lift count
## [1] 1.00      1    380
## [2] 0.05      1     19
## [3] 0.05      1     19
```

VISUALIZING ASSOCIATION RULES

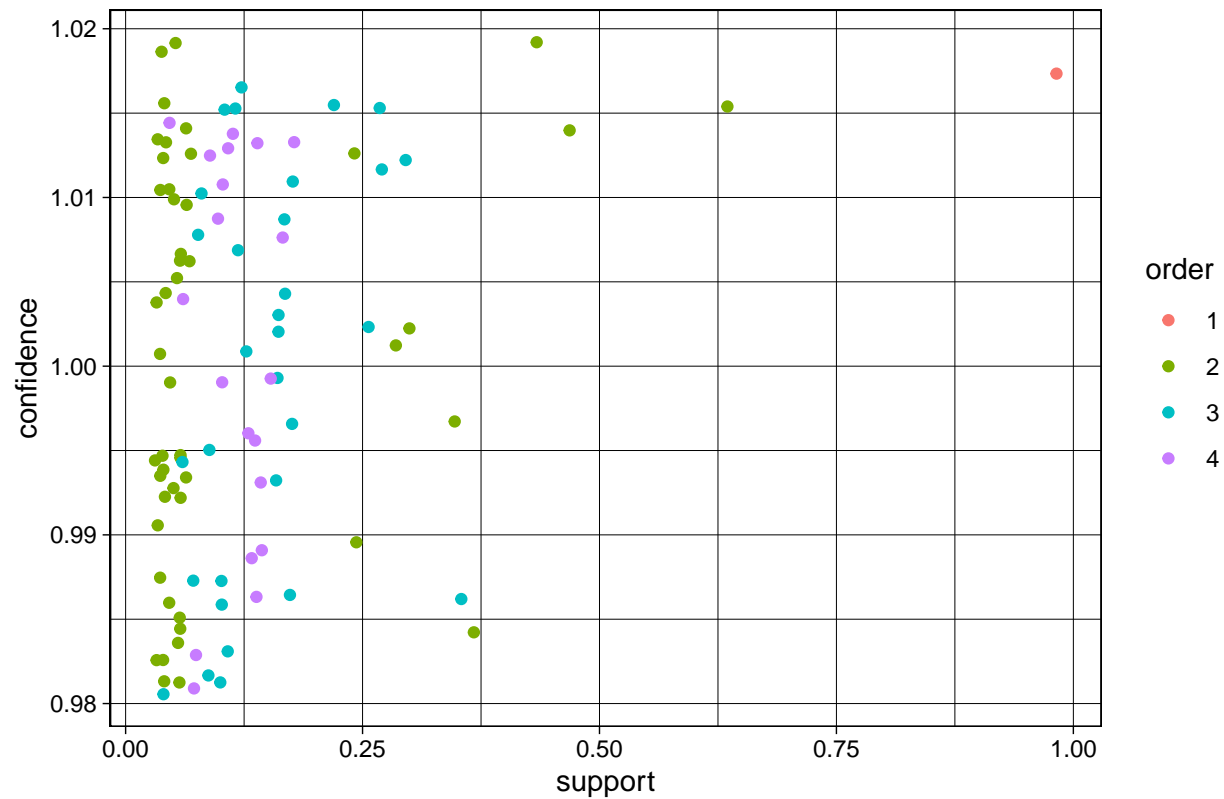
```
plot(rules,jitter = 1)
```



```
plot(rules, shading = "order", color=c("darkred", "purple"))
```

```
## To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
```

Scatter plot for 101 rules

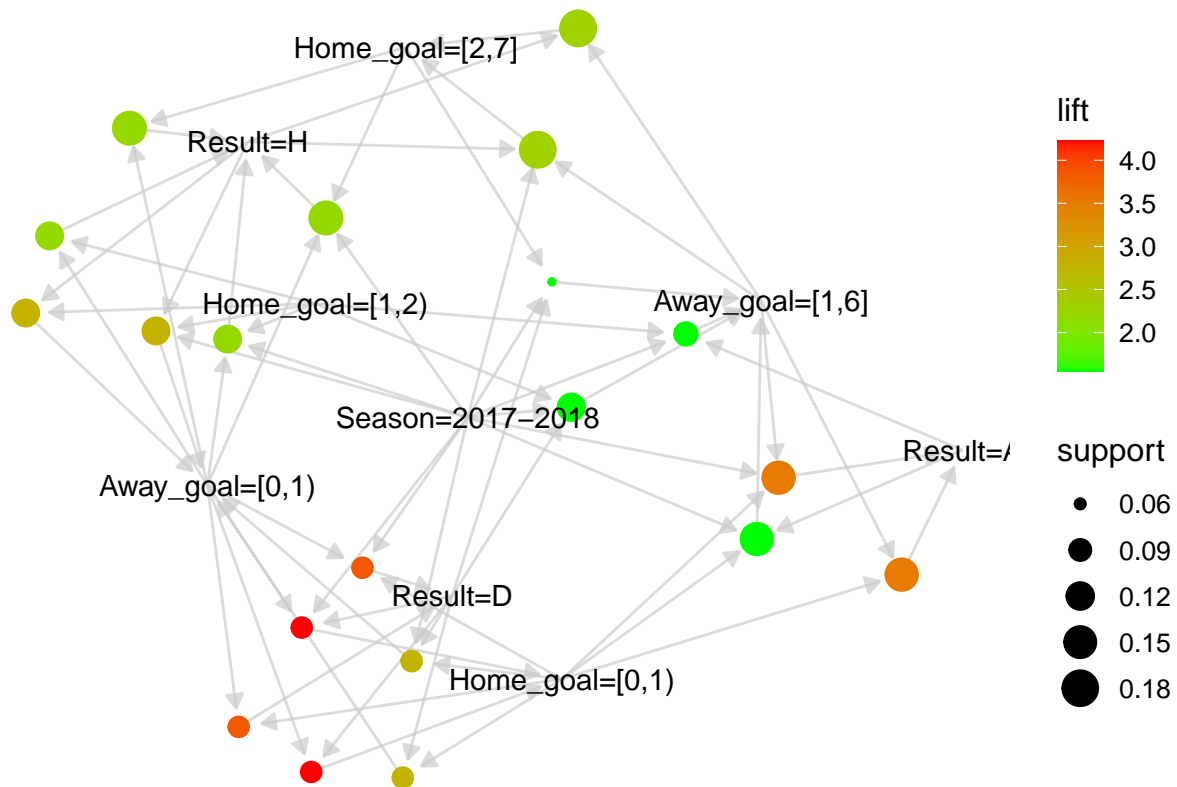


```
#Graph plot for items
```

```
plot(rules, method="graph", max=20, control=list(verbose = FALSE), colors=c("red", "green"))
```

```
## Warning: Too many rules supplied. Only plotting the best 20 rules using lift
```

```
## (change control parameter max if needed)
```

```
plot(rules, method = "matrix", measure=c("support", "confidence"))
```

```
## Itemsets in Antecedent (LHS)
## [1] "{Away_goal=[0,1],Result=D,Season=2017-2018}"
## [2] "{Home_goal=[0,1],Away_goal=[0,1],Season=2017-2018}"
## [3] "{Home_goal=[0,1],Away_goal=[1,6],Season=2017-2018}"
## [4] "{Home_goal=[0,1],Result=D,Season=2017-2018}"
## [5] "{Home_goal=[1,2],Result=H,Season=2017-2018}"
## [6] "{Away_goal=[0,1],Result=D}"
## [7] "{Home_goal=[0,1],Away_goal=[0,1]}"
## [8] "{Away_goal=[1,6],Result=H,Season=2017-2018}"
## [9] "{Home_goal=[0,1],Away_goal=[1,6]}"
## [10] "{Home_goal=[1,2],Away_goal=[0,1],Season=2017-2018}"
## [11] "{Home_goal=[2,7],Away_goal=[0,1],Season=2017-2018}"
## [12] "{Home_goal=[0,1],Result=D}"
## [13] "{Home_goal=[1,2],Result=H}"
## [14] "{Away_goal=[1,6],Result=H}"
## [15] "{Home_goal=[1,2],Away_goal=[0,1]}"
## [16] "{Home_goal=[2,7],Away_goal=[0,1]}"
## [17] "{Result=A,Season=2017-2018}"
## [18] "{Home_goal=[0,1],Result=A,Season=2017-2018}"
## [19] "{Home_goal=[1,2],Result=D,Season=2017-2018}"
## [20] "{Home_goal=[2,7],Result=D,Season=2017-2018}"
## [21] "{Home_goal=[1,2],Result=A,Season=2017-2018}"
## [22] "{Result=A}"
## [23] "{Home_goal=[0,1],Result=A}"
## [24] "{Home_goal=[1,2],Result=D}"
## [25] "{Home_goal=[2,7],Result=D}"
## [26] "{Home_goal=[1,2],Result=A}"
```

```

## [27] "{}"
## [28] "{Home_team=Arsenal}"
## [29] "{Away_team=Tottenham Hotspur}"
## [30] "{Home_team=Manchester United}"
## [31] "{Away_team=West Ham United}"
## [32] "{Home_team=Swansea City}"
## [33] "{Away_team=Manchester United}"
## [34] "{Home_team=AFC Bournemouth}"
## [35] "{Away_team=Watford}"
## [36] "{Home_team=Burnley}"
## [37] "{Away_team=West Bromwich Albion}"
## [38] "{Home_team=Leicester City}"
## [39] "{Away_team=Brighton and Hove Albion}"
## [40] "{Home_team=Liverpool}"
## [41] "{Away_team=Crystal Palace}"
## [42] "{Home_team=Stoke City}"
## [43] "{Away_team=Arsenal}"
## [44] "{Home_team=Huddersfield Town}"
## [45] "{Away_team=Newcastle United}"
## [46] "{Home_team=Tottenham Hotspur}"
## [47] "{Away_team=Chelsea}"
## [48] "{Home_team=Manchester City}"
## [49] "{Away_team=Everton}"
## [50] "{Away_team=Southampton}"
## [51] "{Home_team=Newcastle United}"
## [52] "{Away_team=Manchester City}"
## [53] "{Home_team=Brighton and Hove Albion}"
## [54] "{Away_team=Leicester City}"
## [55] "{Home_team=Watford}"
## [56] "{Away_team=Liverpool}"
## [57] "{Home_team=Chelsea}"
## [58] "{Away_team=Burnley}"
## [59] "{Home_team=Crystal Palace}"
## [60] "{Away_team=Huddersfield Town}"
## [61] "{Home_team=Everton}"
## [62] "{Away_team=Stoke City}"
## [63] "{Home_team=Southampton}"
## [64] "{Away_team=Swansea City}"
## [65] "{Home_team=West Bromwich Albion}"
## [66] "{Away_team=AFC Bournemouth}"
## [67] "{Home_team=West Ham United}"
## [68] "{Home_goal=[0,1]}"
## [69] "{Result=D}"
## [70] "{Home_goal=[1,2]}"
## [71] "{Away_goal=[0,1]}"
## [72] "{Home_goal=[2,7]}"
## [73] "{Result=H}"
## [74] "{Away_goal=[1,6]}"
## [75] "{Away_goal=[1,6],Result=D}"
## [76] "{Away_goal=[1,6],Result=A}"
## [77] "{Home_goal=[1,2],Away_goal=[1,6]}"
## [78] "{Away_goal=[0,1],Result=H}"
## [79] "{Home_goal=[2,7],Result=H}"
## [80] "{Home_goal=[2,7],Away_goal=[1,6]}"

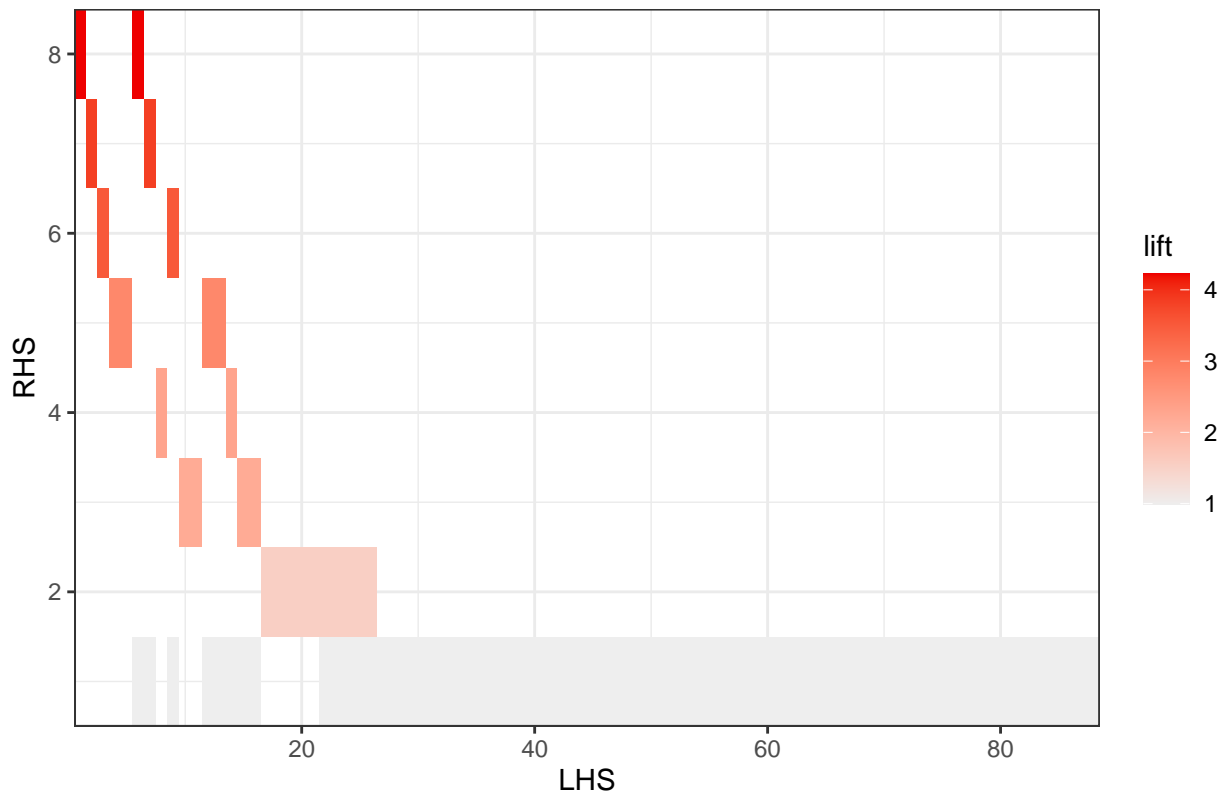
```

```

## [81] "{Home_goal=[0,1),Away_goal=[0,1),Result=D}"
## [82] "{Home_goal=[0,1),Away_goal=[1,6],Result=A}"
## [83] "{Home_goal=[1,2),Away_goal=[1,6],Result=D}"
## [84] "{Home_goal=[2,7),Away_goal=[1,6],Result=D}"
## [85] "{Home_goal=[1,2),Away_goal=[1,6],Result=A}"
## [86] "{Home_goal=[1,2),Away_goal=[0,1),Result=H}"
## [87] "{Home_goal=[2,7),Away_goal=[0,1),Result=H}"
## [88] "{Home_goal=[2,7),Away_goal=[1,6],Result=H}"
## Itemsets in Consequent (RHS)
## [1] "{Season=2017-2018}" "{Away_goal=[1,6]}" "{Result=H}"
## [4] "{Home_goal=[2,7]}" "{Away_goal=[0,1]}" "{Result=A}"
## [7] "{Result=D}" "{Home_goal=[0,1]}"

```

Matrix for 101 rules



ASSOCIATION MINING ANALYSIS

Here, association rules can be used to understand the prediction about which sides either home or away, are going to be more successful in future based on their current performances and result. Understanding the association or co-occurrence will help us plan what promo or recommendation we will need to give to organizer based on their current outcomes. Network analysis help further help us find more insight compared to if only we look at the rules individually.

CONCLUSION

This method can be modified and implemented in different ways, depending on the user's interest. A deeper look into the outcomes can establish additional rules for a more detailed analysis. From this analysis, I have found that the proposed association rule for data mining can be effective to extract football tactics from the team's individual performance.

More than 60% times, the home teams have won when they have scored goals >2 , while the away teams have had draw or lost outcomes on the remaining part. Although the presented technique is not a sophisticated measure for establishing a general recommendation pattern in this dataset, it provides us with an underlying relationships between the teams and their goal differences. Such approach can also be incorporated in many activities, for instance in pitch analysis or a marketing campaign.

END
