Exercise 4.5

Code

```
#########################################################
# Exercise 4.5 - 22th november 2015
#
# Davide Pedranz (362504)
# Jakub Amanowicz (362594)
# Hongmei Liang (292520)
#########################################################

# load needed libraries
library("arulesViz")
data("Groceries")

# a -> mine frequent itemsets
frequent_itemsets <- sort(apriori(Groceries, parameter = list(
    support = 0.05, target="frequent_itemsets")), by="support")
inspect(frequent_itemsets)

# b -> strong association rules
rules <- sort(apriori(Groceries, parameter=list(support=0.004,
    confidence=0.7, minlen = 2, target="rules")), by="support")
inspect(rules)

# c -> scatter plot of association rules
plot(rules, measure=c("support", "confidence"), shading="lift")
```

Output

```
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport support minlen maxlen
        0.8    0.1    1 none FALSE            TRUE    0.05      1     10
          target    ext
 frequent itemsets FALSE

Algorithmic control:
 filter tree heap memopt load sort verbose
    0.1 TRUE TRUE  FALSE TRUE    2    TRUE
```

```
Absolute minimum support count: 491

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[169 item(s), 9835 transaction(s)] done [0.00s].
sorting and recoding items ... [28 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 done [0.00s].
writing ... [31 set(s)] done [0.00s].
creating S4 object  ... done [0.00s].

   items                          support
28 {whole milk}                   0.25551601
27 {other vegetables}             0.19349263
26 {rolls/buns}                   0.18393493
24 {soda}                         0.17437722
25 {yogurt}                       0.13950178
21 {bottled water}               0.11052364
23 {root vegetables}              0.10899847
22 {tropical fruit}               0.10493137
19 {shopping bags}                0.09852567
20 {sausage}                      0.09395018
17 {pastry}                       0.08896797
18 {citrus fruit}                 0.08276563
8  {bottled beer}                 0.08052872
12 {newspapers}                   0.07981698
1  {canned beer}                  0.07768175
16 {pip fruit}                    0.07564820
31 {other vegetables,whole milk} 0.07483477
14 {fruit/vegetable juice}        0.07229283
15 {whipped/sour cream}           0.07168277
9  {brown bread}                  0.06487036
13 {domestic eggs}                0.06344687
7  {frankfurter}                  0.05897306
10 {margarine}                    0.05856634
2  {coffee}                       0.05805796
6  {pork}                         0.05765125
30 {whole milk,rolls/buns}        0.05663447
29 {whole milk,yogurt}            0.05602440
11 {butter}                       0.05541434
4  {curd}                         0.05327911
3  {beef}                         0.05246568
5  {napkins}                      0.05236401

Apriori
```

```
Parameter specification:
 confidence minval smax arem  aval originalSupport support minlen maxlen target
       0.7    0.1    1 none FALSE             TRUE   0.004      2     10  rules
   ext
 FALSE


Algorithmic control:
 filter tree heap memopt load sort verbose
    0.1 TRUE TRUE  FALSE TRUE    2    TRUE


Absolute minimum support count: 39

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[169 item(s), 9835 transaction(s)] done [0.00s].
sorting and recoding items ... [126 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 done [0.00s].
writing ... [5 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
```

	lhs	rhs	support	confidence	lift
1	{tropical fruit, root vegetables, yogurt}	=> {whole milk}	0.005693950	0.7000000	2.739554
2	{butter, curd}	=> {whole milk}	0.004880529	0.7164179	2.803808
3	{curd, domestic eggs}	=> {whole milk}	0.004778851	0.7343750	2.874086
4	{citrus fruit, tropical fruit, root vegetables}	=> {other vegetables}	0.004473818	0.7857143	4.060694
5	{tropical fruit, yogurt, whipped/sour cream}	=> {whole milk}	0.004372140	0.7049180	2.758802

Graph

Scatter plot for 5 rules
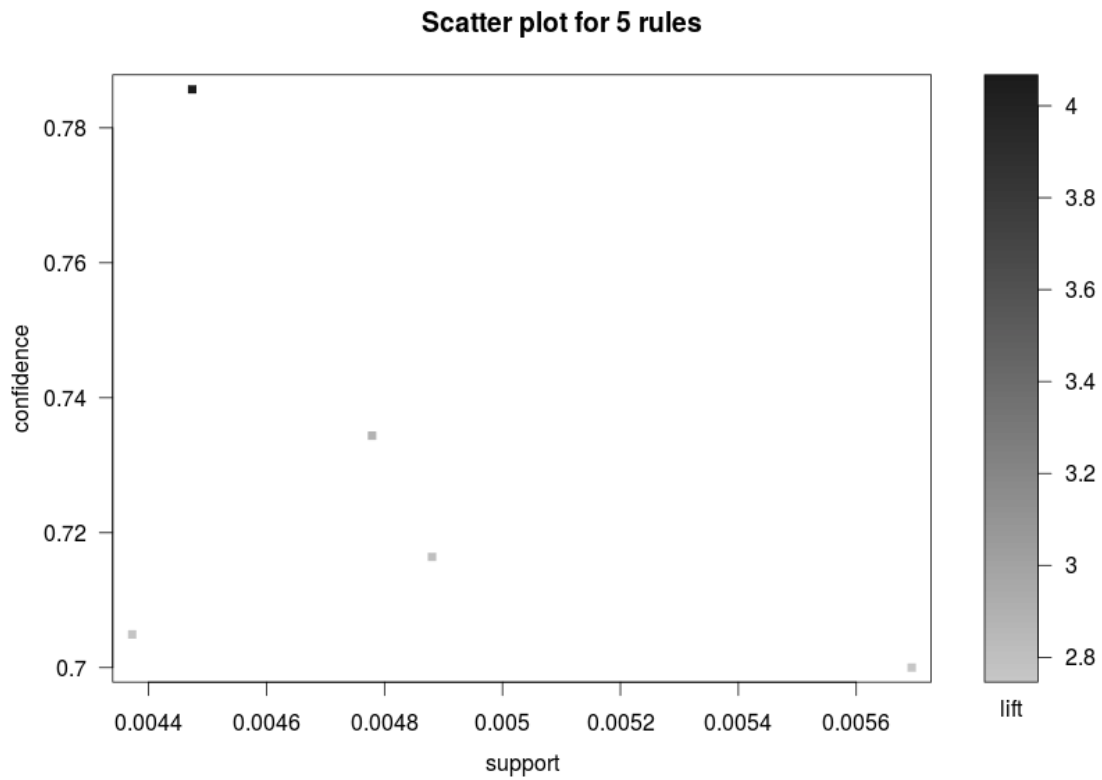
An interesting rule found is the following (with the highest confidence):

$$\{citrus fruit, tropical fruit, root vegetables\} \Rightarrow \{other vegetables\}$$