# 1 Introduction

**Mining association rules using ARS, from the SIPINA distribution.**

SIPINA is known for its decision tree induction algorithms. In fact, the distribution includes two other tools that are little known to the public: REGRESS, which is specialized in multiple linear regression, we described it in one of our tutorials[1]; and an association rules extraction tool, called simply Association Rule Software (**ARS**).

At that time (1997-1998), my idea was to build a set of independent executables organized around the same data management grid. Various tools for factor analysis and clustering, in addition to the tools which are included in the current distribution, were developed. But afterward, I realized that a functioning based on visual programming using streams of components was more advantageous in many respects, in terms of ease of use for the users, but also in terms of software evolution. TANAGRA was developed with this in mind.

This does not mean that the tools included in the SIPINA distribution are not interesting. In this tutorial, I describe the use of the ARS tool. Its interactivity with Excel spreadsheet is its main advantage. We launch the software from Excel using the "sipina.xla" add-in[2]. We can easily retrieve the rules in the spreadsheet. Then, we can explore them (the rules) using the Excel data handling capabilities. The ability to filter and sort rules according to different criteria is a great help in detecting interesting rules. This is a very important aspect because the profusion of rules can quickly confuse the data miner.

# 2 Dataset

The **"market_basket.xlsx"** data file describes the contents of **n = 1361** shopping carts (transactions). We have a library of **p = 303** products (items).

In average, each cart contains 9,5 products (min = 0, max = 303). And the products are purchased 42,7 times in average (min = 7, max = 167). The 5 most popular products are: "Eggs", "White bread", "2pct milk", "Potato chips" and "98pct fat free hamburger". The least popular ones are: "Celery", "Oats and Nuts Cereals", "Chicken legs", "Nasal spray" and "Daily Newspaper".

# 3 Mining association rule with ARS

## 3.1 Data importation

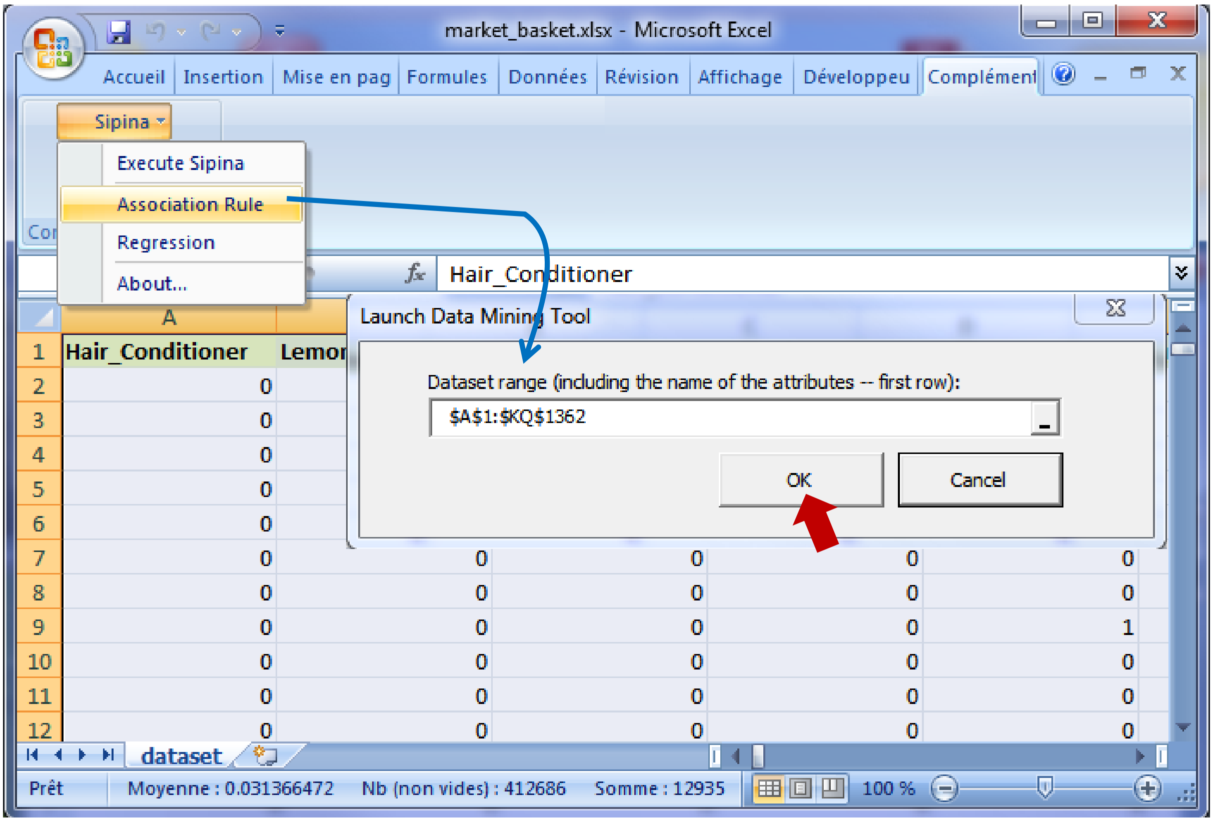
We load the "**market_basket.xlsx**" data file into Excel (we must use Excel 2007 or later version because the number of items [p = 303] exceeds the limitation of 256 columns of Excel 2003 and earlier versions). **Note:** If you do not have the Excel 2007 version, the best solution is to import the "market_basket.txt" data file (tab separated text format) via the "FILE / OPEN / TEXT FILE FORMAT (*.txt)" menu[3]. This file is included in the archive that comes with this document.

[1] http://data-mining-tutorials.blogspot.fr/2011/08/regress-into-sipina-package.html

[2] http://data-mining-tutorials.blogspot.fr/2016/06/sipina-add-in-for-excel-2007-and-2010.html

[3] Sipina – Supported file format - http://data-mining-tutorials.blogspot.fr/2009/11/sipina-supported-file-format.html

After loading the data file into Excel, we select the data range and click the menu SIPINA / ASSOCIATION RULE. A setting dialog box appears. We check the coordinates of the data range and we click the OK button.

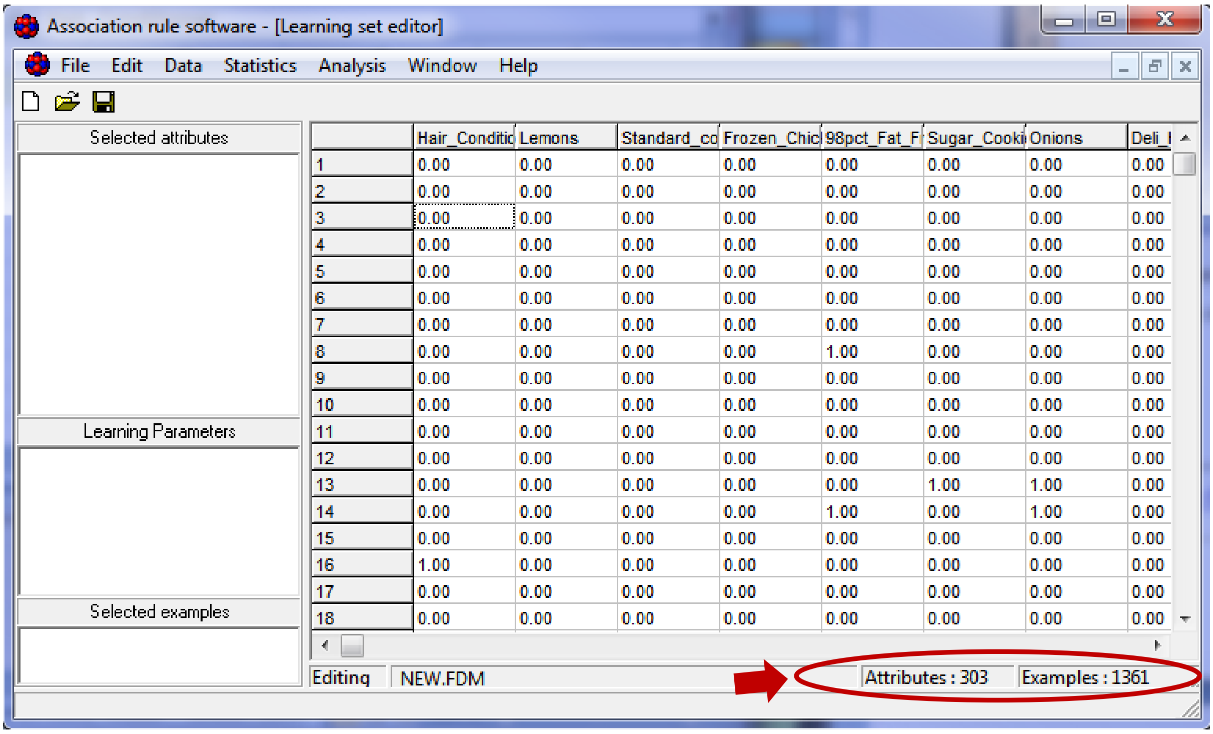


Dataset range (including the name of the attributes -- first row):

$A$1:$KQ$1362

| | A | | | | | |
|---|---|---|---|---|---|---|
| 1 | Hair_Conditioner | Lemor | | | | |
| 2 | 0 | | | | | |
| 3 | 0 | | | | | |
| 4 | 0 | | | | | |
| 5 | 0 | | | | | |
| 6 | 0 | | | | | |
| 7 | 0 | 0 | 0 | 0 | 0 | |
| 8 | 0 | 0 | 0 | 0 | 0 | |
| 9 | 0 | 0 | 0 | 0 | 1 | |
| 10 | 0 | 0 | 0 | 0 | 0 | |
| 11 | 0 | 0 | 0 | 0 | 0 | |
| 12 | 0 | 0 | 0 | 0 | 0 | |

The software is automatically launched and the dataset is imported. We check that we have n = 1361 rows and p = 303 columns.

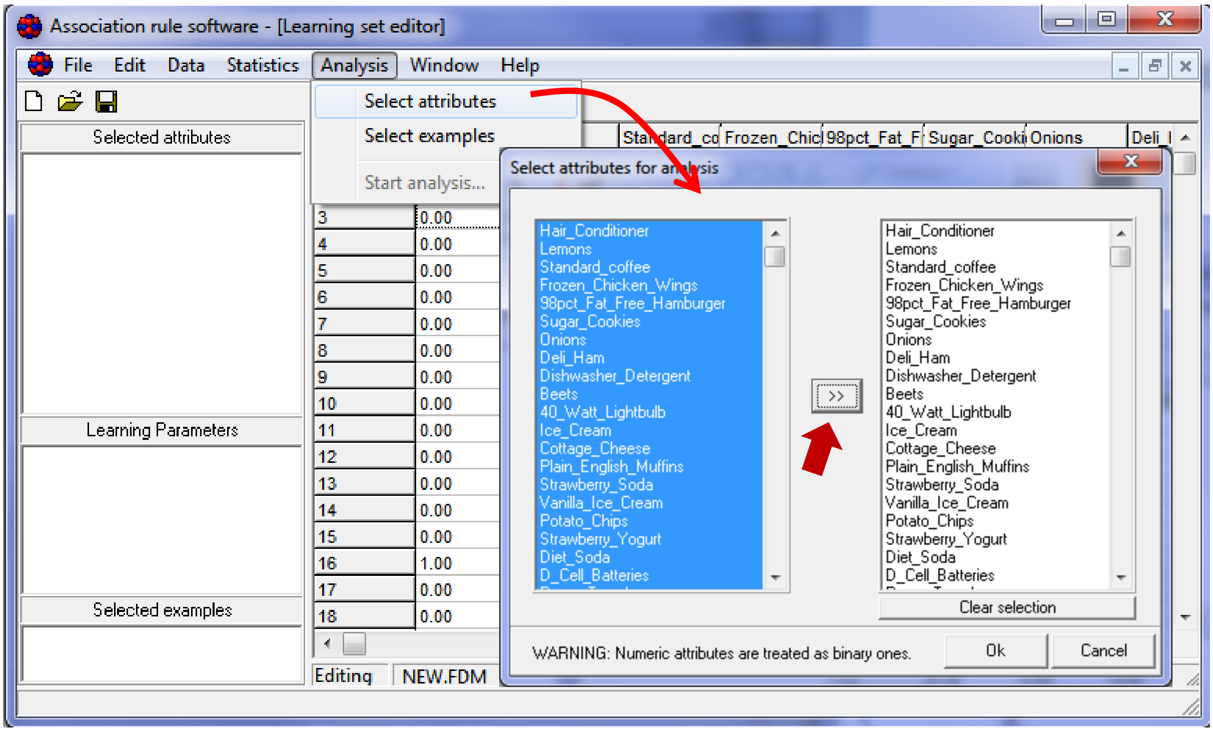


| | Hair_Conditic | Lemons | Standard_co | Frozen_Chic | 98pct_Fat_F | Sugar_Cooki | Onions | Deli_I |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 3 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 4 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 5 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 6 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 7 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 8 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 |
| 9 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 13 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 0.00 |
| 14 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 |
| 15 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 16 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 17 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Editing NEW.FDM Attributes : 303 Examples : 1361

## 3.2 Selecting the items

To select the items to include into the analysis, we click the ANALYSIS / SELECT ATTRIBUTES menu. We select all the columns.

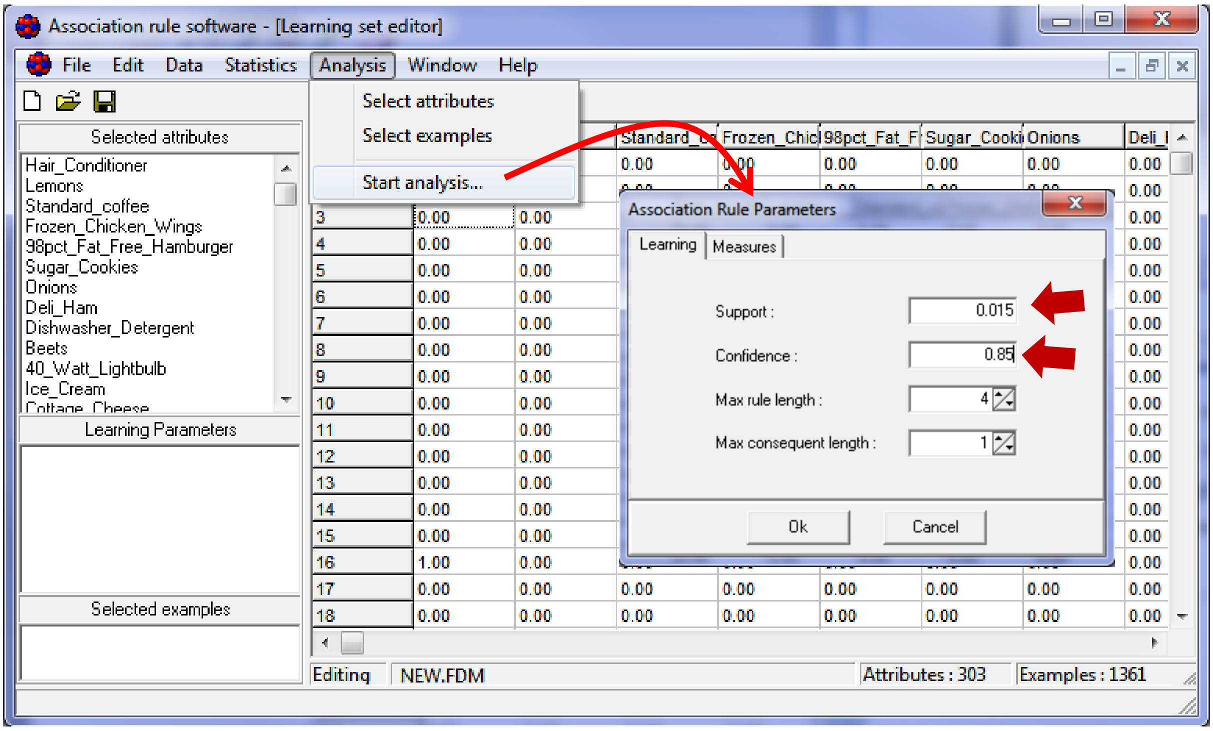


Hair_Conditioner
Lemons
Standard_coffee
Frozen_Chicken_Wings
98pct_Fat_Free_Hamburger
Sugar_Cookies
Onions
Deli_Ham
Dishwasher_Detergent
Beets
40_Watt_Lightbulb
Ice_Cream
Cottage_Cheese
Plain_English_Muffins
Strawberry_Soda
Vanilla_Ice_Cream
Potato_Chips
Strawberry_Yogurt
Diet_Soda
D_Cell_Batteries

WARNING: Numeric attributes are treated as binary ones.

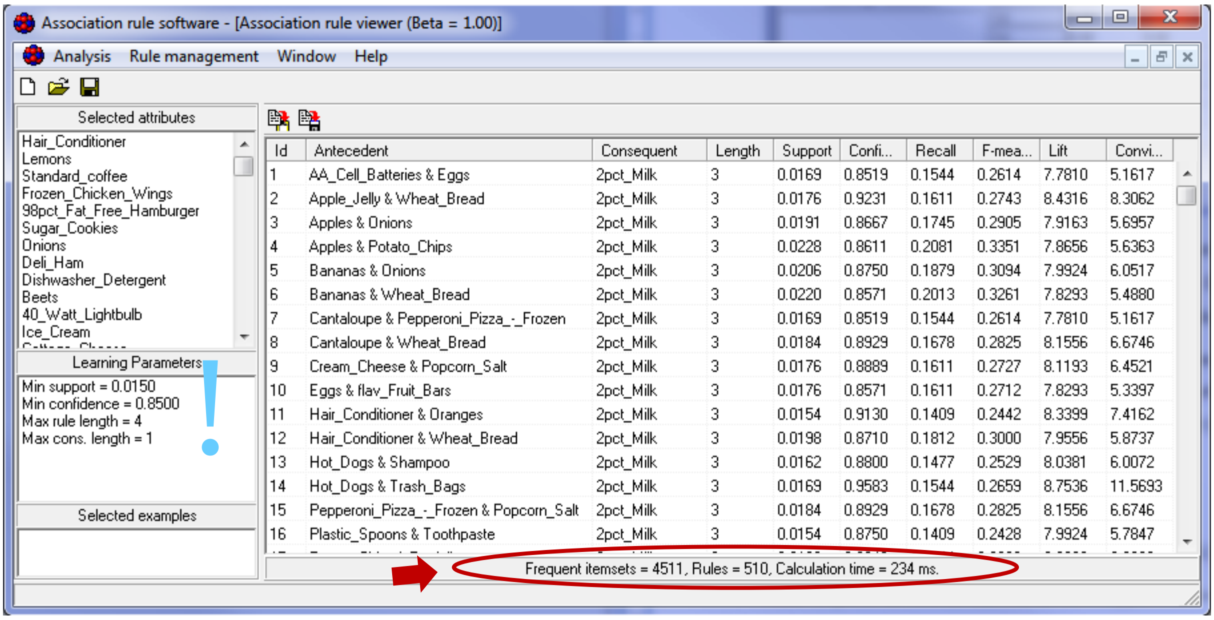
Because our variables are numerical, ARS takes them as dummy variables: 0, absence of the item in the shopping cart; 1 (it can be any value > 0), presence. When we process categorical variables, the tool automatically performs a dummy coding before starting processing.

## 3.3 Settings for the rule extraction process



Support : 0.015
Confidence : 0.85
Max rule length : 4
Max consequent length : 1

To launch the association rule learning, we click the ANALYSIS / START ANALYSIS menu. A setting dialog box appears: we set the minimum support to 0.015 (we accept the rules which occurs at least 1361 x 0.015 ≈ 20 times into the database); the minimum confidence is 0.85; the maximum cardinal of the rule is 4 items; we have at most 1 item into the consequent of the rule.


Association rule software - [Association rule viewer (Beta = 1.00)]
Analysis
Rule management
Window
Help
!

Selected attributes

Hair_Conditioner
Lemons
Standard_coffee
Frozen_Chicken_Wings
98pct_Fat_Free_Hamburger
Sugar_Cookies
Onions
Deli_Ham
Dishwasher_Detergent
Beets
40_Watt_Lightbulb
Ice_Cream

Learning Parameters

Min support = 0.0150
Min confidence = 0.8500
Max rule length = 4
Max cons. length = 1

Selected examples

| Id | Antecedent | Consequent | Length | Support | Confi... | Recall | F-mea... | Lift | Convi... |
|---|---|---|---|---|---|---|---|---|---|
| 1 | AA_Cell_Batteries & Eggs | 2pct_Milk | 3 | 0.0169 | 0.8519 | 0.1544 | 0.2614 | 7.7810 | 5.1617 |
| 2 | Apple_Jelly & Wheat_Bread | 2pct_Milk | 3 | 0.0176 | 0.9231 | 0.1611 | 0.2743 | 8.4316 | 8.3062 |
| 3 | Apples & Onions | 2pct_Milk | 3 | 0.0191 | 0.8667 | 0.1745 | 0.2905 | 7.9163 | 5.6957 |
| 4 | Apples & Potato_Chips | 2pct_Milk | 3 | 0.0228 | 0.8611 | 0.2081 | 0.3351 | 7.8656 | 5.6363 |
| 5 | Bananas & Onions | 2pct_Milk | 3 | 0.0206 | 0.8750 | 0.1879 | 0.3094 | 7.9924 | 6.0517 |
| 6 | Bananas & Wheat_Bread | 2pct_Milk | 3 | 0.0220 | 0.8571 | 0.2013 | 0.3261 | 7.8293 | 5.4880 |
| 7 | Cantaloupe & Pepperoni_Pizza_-_Frozen | 2pct_Milk | 3 | 0.0169 | 0.8519 | 0.1544 | 0.2614 | 7.7810 | 5.1617 |
| 8 | Cantaloupe & Wheat_Bread | 2pct_Milk | 3 | 0.0184 | 0.8929 | 0.1678 | 0.2825 | 8.1556 | 6.6746 |
| 9 | Cream_Cheese & Popcorn_Salt | 2pct_Milk | 3 | 0.0176 | 0.8889 | 0.1611 | 0.2727 | 8.1193 | 6.4521 |
| 10 | Eggs & flav_Fruit_Bars | 2pct_Milk | 3 | 0.0176 | 0.8571 | 0.1611 | 0.2712 | 7.8293 | 5.3397 |
| 11 | Hair_Conditioner & Oranges | 2pct_Milk | 3 | 0.0154 | 0.9130 | 0.1409 | 0.2442 | 8.3399 | 7.4162 |
| 12 | Hair_Conditioner & Wheat_Bread | 2pct_Milk | 3 | 0.0198 | 0.8710 | 0.1812 | 0.3000 | 7.9556 | 5.8737 |
| 13 | Hot_Dogs & Shampoo | 2pct_Milk | 3 | 0.0162 | 0.8800 | 0.1477 | 0.2529 | 8.0381 | 6.0072 |
| 14 | Hot_Dogs & Trash_Bags | 2pct_Milk | 3 | 0.0169 | 0.9583 | 0.1544 | 0.2659 | 8.7536 | 11.5693 |
| 15 | Pepperoni_Pizza_-_Frozen & Popcorn_Salt | 2pct_Milk | 3 | 0.0184 | 0.8929 | 0.1678 | 0.2825 | 8.1556 | 6.6746 |
| 16 | Plastic_Spoons & Toothpaste | 2pct_Milk | 3 | 0.0154 | 0.8750 | 0.1409 | 0.2428 | 7.9924 | 5.7847 |

Frequent itemsets = 4511, Rules = 510, Calculation time = 234 ms.

When we validate these settings by clicking the OK button, we obtain a new window which contains the mined rules. Each rule is characterized by its antecedent, its consequent, and a set of numerical indicators (support, confidence, lift, etc.)[4].

We obtain **510 rules**. The first one (n°1) is:

**IF** purchase (AA_Cell_Batteries & Eggs) **THEN** purchase also (2pct_Milk)

The Support of the rule is:

P(AA_Cell_Batteries & Eggs & 2pct_Milk) = 0.0169

Its Confidence:

P(2pct_Milk / AA_Cell_Batteries & Eggs) = 0.8519

Its Lift:

$$\frac{\mathrm{P(2pct\_milk/AA\_Cell\_Batteries\ \&\ Eggs)}}{\mathrm{P(2pct\_milk)}} = 7.7810$$

Actually, nothing really distinguishes this tool from the association rule extractions components available in TANAGRA. My idea in this tutorial is to highlight in the next section the features... of Excel that allow us to examine the results (mined rules) better.

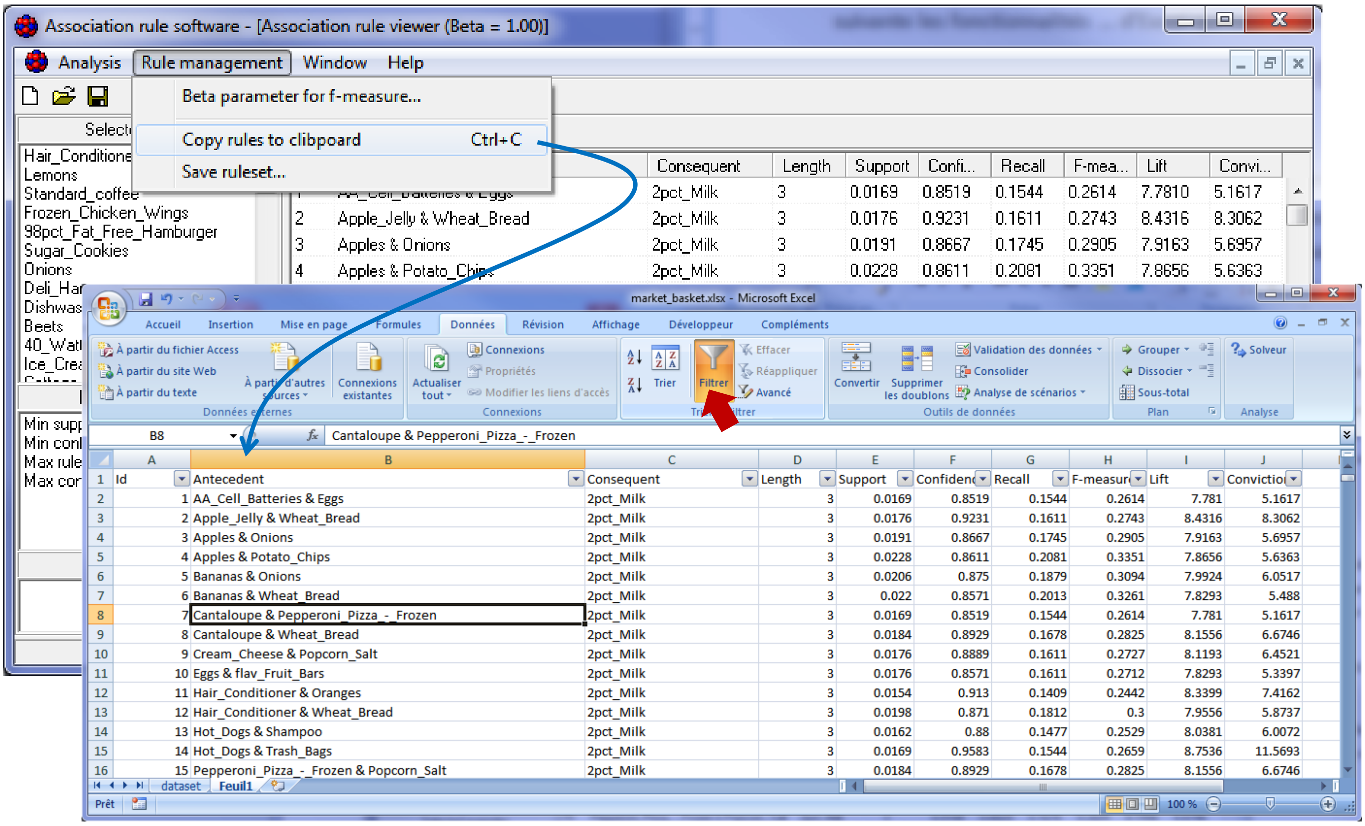
[4] "Interestingness measures for association rules" - http://data-mining-tutorials.blogspot.fr/2009/02/interestingness-measures-for.html

# 4 Exploring the rules under Excel

## 4.1 Retrieving the rules

Association rules learning algorithms often generate a large number of rules, 510 for our study. To be able to examine them, we must be able to organize the rules at our convenience: filter them according to several criteria, sort them according to the measures of interestingness, etc. The Excel's **Sort & Filter** tools are really appropriate for that purpose.

First, we must copy the rules into Excel. From ARS, we click the RULE MANAGEMENT / COPY RULES TO CLIPBOARD menu.


Association rule software - [Association rule viewer (Beta = 1.00)]
Analysis
Rule management
Window
Help
Beta parameter for f-measure...
Copy rules to clipboard Ctrl+C
Save ruleset...
market_basket.xlsx - Microsoft Excel
Données
Filtrer
Trier et filtrer

B8 Cantaloupe & Pepperoni_Pizza_-_Frozen

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Id | Antecedent | Consequent | Length | Support | Confidenc | Recall | F-measur | Lift | Convictio |
| 2 | 1 | AA_Cell_Batteries & Eggs | 2pct_Milk | 3 | 0.0169 | 0.8519 | 0.1544 | 0.2614 | 7.781 | 5.1617 |
| 3 | 2 | Apple_Jelly & Wheat_Bread | 2pct_Milk | 3 | 0.0176 | 0.9231 | 0.1611 | 0.2743 | 8.4316 | 8.3062 |
| 4 | 3 | Apples & Onions | 2pct_Milk | 3 | 0.0191 | 0.8667 | 0.1745 | 0.2905 | 7.9163 | 5.6957 |
| 5 | 4 | Apples & Potato_Chips | 2pct_Milk | 3 | 0.0228 | 0.8611 | 0.2081 | 0.3351 | 7.8656 | 5.6363 |
| 6 | 5 | Bananas & Onions | 2pct_Milk | 3 | 0.0206 | 0.875 | 0.1879 | 0.3094 | 7.9924 | 6.0517 |
| 7 | 6 | Bananas & Wheat_Bread | 2pct_Milk | 3 | 0.022 | 0.8571 | 0.2013 | 0.3261 | 7.8293 | 5.488 |
| 8 | 7 | Cantaloupe & Pepperoni_Pizza_-_Frozen | 2pct_Milk | 3 | 0.0169 | 0.8519 | 0.1544 | 0.2614 | 7.781 | 5.1617 |
| 9 | 8 | Cantaloupe & Wheat_Bread | 2pct_Milk | 3 | 0.0184 | 0.8929 | 0.1678 | 0.2825 | 8.1556 | 6.6746 |
| 10 | 9 | Cream_Cheese & Popcorn_Salt | 2pct_Milk | 3 | 0.0176 | 0.8889 | 0.1611 | 0.2727 | 8.1193 | 6.4521 |
| 11 | 10 | Eggs & flav_Fruit_Bars | 2pct_Milk | 3 | 0.0176 | 0.8571 | 0.1611 | 0.2712 | 7.8293 | 5.3397 |
| 12 | 11 | Hair_Conditioner & Oranges | 2pct_Milk | 3 | 0.0154 | 0.913 | 0.1409 | 0.2442 | 8.3399 | 7.4162 |
| 13 | 12 | Hair_Conditioner & Wheat_Bread | 2pct_Milk | 3 | 0.0198 | 0.871 | 0.1812 | 0.3 | 7.9556 | 5.8737 |
| 14 | 13 | Hot_Dogs & Shampoo | 2pct_Milk | 3 | 0.0162 | 0.88 | 0.1477 | 0.2529 | 8.0381 | 6.0072 |
| 15 | 14 | Hot_Dogs & Trash_Bags | 2pct_Milk | 3 | 0.0169 | 0.9583 | 0.1544 | 0.2659 | 8.7536 | 11.5693 |
| 16 | 15 | Pepperoni_Pizza_-_Frozen & Popcorn_Salt | 2pct_Milk | 3 | 0.0184 | 0.8929 | 0.1678 | 0.2825 | 8.1556 | 6.6746 |

We create a new worksheet in Excel. We paste the rules. Then we activate the **FILTER** button (**Data** tab, **Sort & Filter** group)[5]. Arrows in columns headers enable to select various criteria for filtering the rules. The same functionalities are available in free spreadsheets such as Calc of LibreOffice[6] or Apache OpenOffice[7].

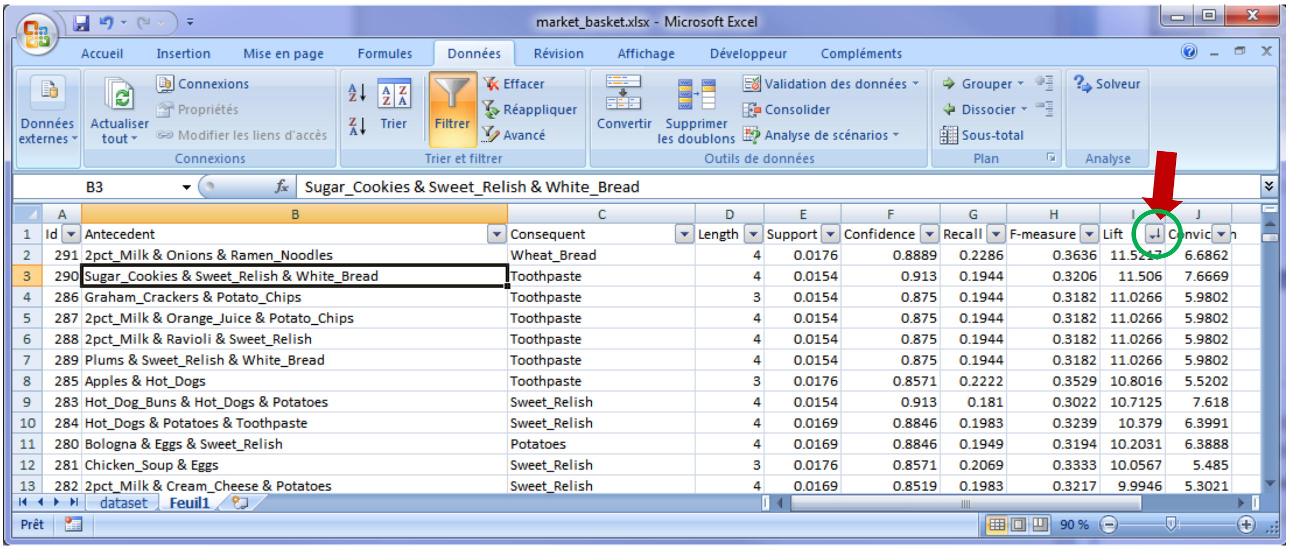
## 4.2 Sorting the rules according to a numerical indicator

We want to sort the rules according to the LIFT criterion. We click on the arrow of the corresponding column, we ask "Sort from the largest to the smallest".

[5] Our screenshots are based on the French version of Excel 2007.

[6] https://fr.libreoffice.org/

[7] https://www.openoffice.org



B3 Sugar_Cookies & Sweet_Relish & White_Bread

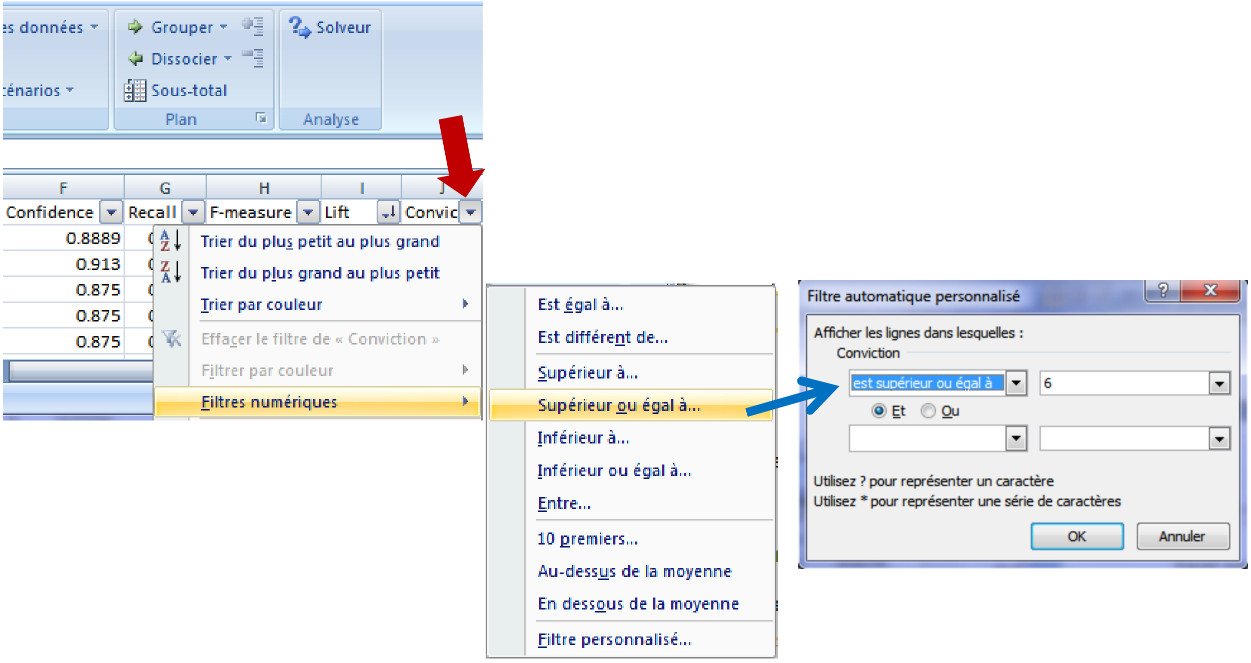
| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Id | Antecedent | Consequent | Length | Support | Confidence | Recall | F-measure | Lift | Convic |
| 2 | 291 | 2pct_Milk & Onions & Ramen_Noodles | Wheat_Bread | 4 | 0.0176 | 0.8889 | 0.2286 | 0.3636 | 11.5217 | 6.6862 |
| 3 | 290 | Sugar_Cookies & Sweet_Relish & White_Bread | Toothpaste | 4 | 0.0154 | 0.913 | 0.1944 | 0.3206 | 11.506 | 7.6669 |
| 4 | 286 | Graham_Crackers & Potato_Chips | Toothpaste | 3 | 0.0154 | 0.875 | 0.1944 | 0.3182 | 11.0266 | 5.9802 |
| 5 | 287 | 2pct_Milk & Orange_Juice & Potato_Chips | Toothpaste | 4 | 0.0154 | 0.875 | 0.1944 | 0.3182 | 11.0266 | 5.9802 |
| 6 | 288 | 2pct_Milk & Ravioli & Sweet_Relish | Toothpaste | 4 | 0.0154 | 0.875 | 0.1944 | 0.3182 | 11.0266 | 5.9802 |
| 7 | 289 | Plums & Sweet_Relish & White_Bread | Toothpaste | 4 | 0.0154 | 0.875 | 0.1944 | 0.3182 | 11.0266 | 5.9802 |
| 8 | 285 | Apples & Hot_Dogs | Toothpaste | 3 | 0.0176 | 0.8571 | 0.2222 | 0.3529 | 10.8016 | 5.5202 |
| 9 | 283 | Hot_Dog_Buns & Hot_Dogs & Potatoes | Sweet_Relish | 4 | 0.0154 | 0.913 | 0.181 | 0.3022 | 10.7125 | 7.618 |
| 10 | 284 | Hot_Dogs & Potatoes & Toothpaste | Sweet_Relish | 4 | 0.0169 | 0.8846 | 0.1983 | 0.3239 | 10.379 | 6.3991 |
| 11 | 280 | Bologna & Eggs & Sweet_Relish | Potatoes | 4 | 0.0169 | 0.8846 | 0.1949 | 0.3194 | 10.2031 | 6.3888 |
| 12 | 281 | Chicken_Soup & Eggs | Sweet_Relish | 3 | 0.0176 | 0.8571 | 0.2069 | 0.3333 | 10.0567 | 5.485 |
| 13 | 282 | 2pct_Milk & Cream_Cheese & Potatoes | Sweet_Relish | 4 | 0.0169 | 0.8519 | 0.1983 | 0.3217 | 9.9946 | 5.3021 |

The rule with the highest LIFT is:

**IF** (2pct_Milk & Onions & Ramen_Noodles) **THEN** (Wheat_Bread) [**LIFT = 11.5217**]

## 4.3 Filtering according to a numerical condition

We want to display only the rules with a conviction > 6.

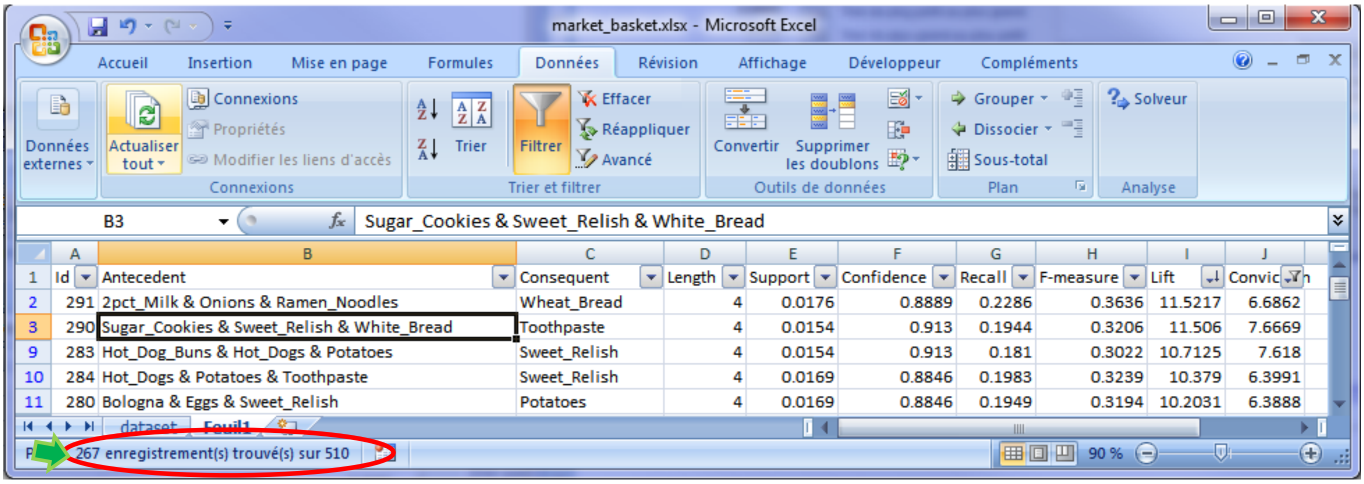


Filtre automatique personnalisé

Afficher les lignes dans lesquelles :
Conviction
est supérieur ou égal à | 6
● Et ○ Ou

Utilisez ? pour représenter un caractère
Utilisez * pour représenter une série de caractères

To do this, we click on the corresponding arrow and we select the option "**Number Filters** / Greater than or Equal to…". We specify the threshold value 6.



B3 — Sugar_Cookies & Sweet_Relish & White_Bread

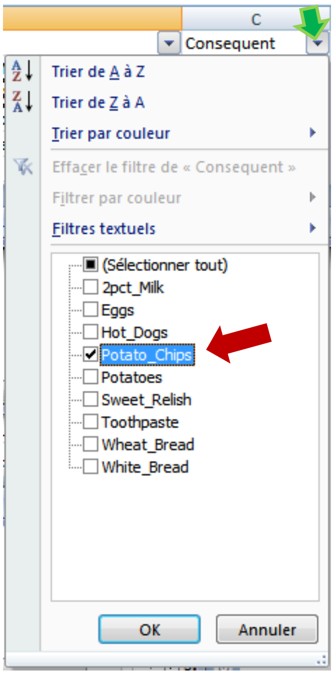
| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Id | Antecedent | Consequent | Length | Support | Confidence | Recall | F-measure | Lift | Convic |
| 2 | 291 | 2pct_Milk & Onions & Ramen_Noodles | Wheat_Bread | 4 | 0.0176 | 0.8889 | 0.2286 | 0.3636 | 11.5217 | 6.6862 |
| 3 | 290 | Sugar_Cookies & Sweet_Relish & White_Bread | Toothpaste | 4 | 0.0154 | 0.913 | 0.1944 | 0.3206 | 11.506 | 7.6669 |
| 9 | 283 | Hot_Dog_Buns & Hot_Dogs & Potatoes | Sweet_Relish | 4 | 0.0154 | 0.913 | 0.181 | 0.3022 | 10.7125 | 7.618 |
| 10 | 284 | Hot_Dogs & Potatoes & Toothpaste | Sweet_Relish | 4 | 0.0169 | 0.8846 | 0.1983 | 0.3239 | 10.379 | 6.3991 |
| 11 | 280 | Bologna & Eggs & Sweet_Relish | Potatoes | 4 | 0.0169 | 0.8846 | 0.1949 | 0.3194 | 10.2031 | 6.3888 |

267 enregistrement(s) trouvé(s) sur 510

267 out of 510 rules meet this condition.

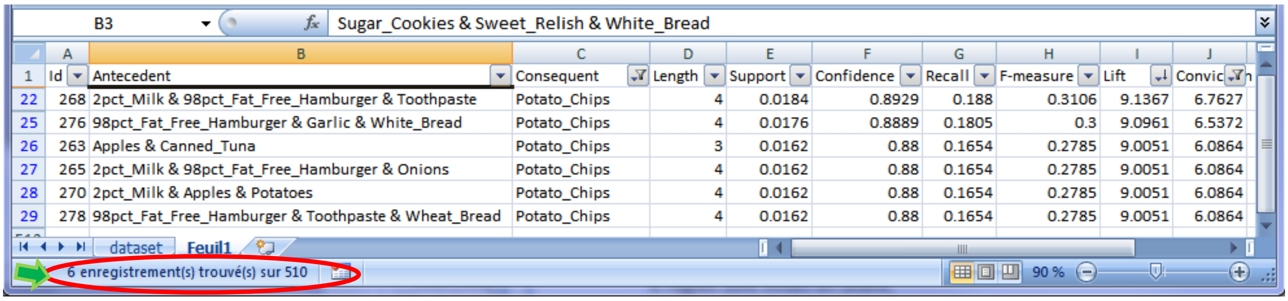
## 4.4 Filtering based on the consequent of the rule

**Among these rules**, we search the ones with the consequent “Purchase of (Potato_Chips)”. We click the arrow in the header of the column and we select the corresponding item.



- ■ (Sélectionner tout)
- ☐ 2pct_Milk
- ☐ Eggs
- ☐ Hot_Dogs
- ☑ Potato_Chips
- ☐ Potatoes
- ☐ Sweet_Relish
- ☐ Toothpaste
- ☐ Wheat_Bread
- ☐ White_Bread

6 rules are highlighted.



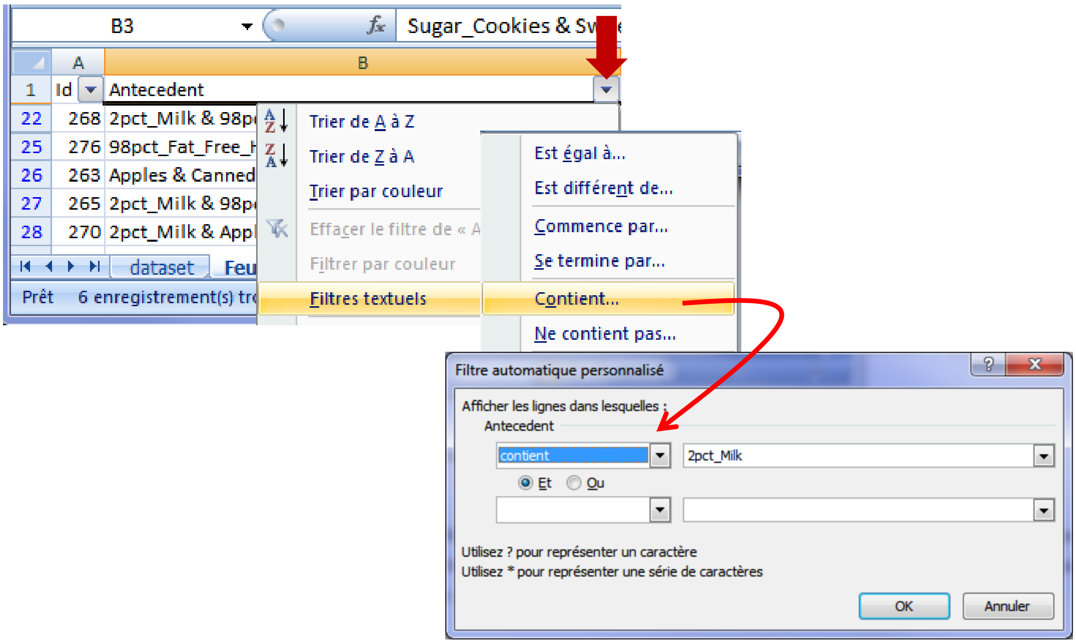
B3 — Sugar_Cookies & Sweet_Relish & White_Bread

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Id | Antecedent | Consequent | Length | Support | Confidence | Recall | F-measure | Lift | Convic |
| 22 | 268 | 2pct_Milk & 98pct_Fat_Free_Hamburger & Toothpaste | Potato_Chips | 4 | 0.0184 | 0.8929 | 0.188 | 0.3106 | 9.1367 | 6.7627 |
| 25 | 276 | 98pct_Fat_Free_Hamburger & Garlic & White_Bread | Potato_Chips | 4 | 0.0176 | 0.8889 | 0.1805 | 0.3 | 9.0961 | 6.5372 |
| 26 | 263 | Apples & Canned_Tuna | Potato_Chips | 3 | 0.0162 | 0.88 | 0.1654 | 0.2785 | 9.0051 | 6.0864 |
| 27 | 265 | 2pct_Milk & 98pct_Fat_Free_Hamburger & Onions | Potato_Chips | 4 | 0.0162 | 0.88 | 0.1654 | 0.2785 | 9.0051 | 6.0864 |
| 28 | 270 | 2pct_Milk & Apples & Potatoes | Potato_Chips | 4 | 0.0162 | 0.88 | 0.1654 | 0.2785 | 9.0051 | 6.0864 |
| 29 | 278 | 98pct_Fat_Free_Hamburger & Toothpaste & Wheat_Bread | Potato_Chips | 4 | 0.0162 | 0.88 | 0.1654 | 0.2785 | 9.0051 | 6.0864 |

6 enregistrement(s) trouvé(s) sur 510

### 4.1 Filtering based on the antecedent of the rule

**Among these rules**, we want to highlight the rules for which the antecedent contains the "2pct_Milk" item. We use the **Text Filters** tool, we set the condition.



Filtre automatique personnalisé

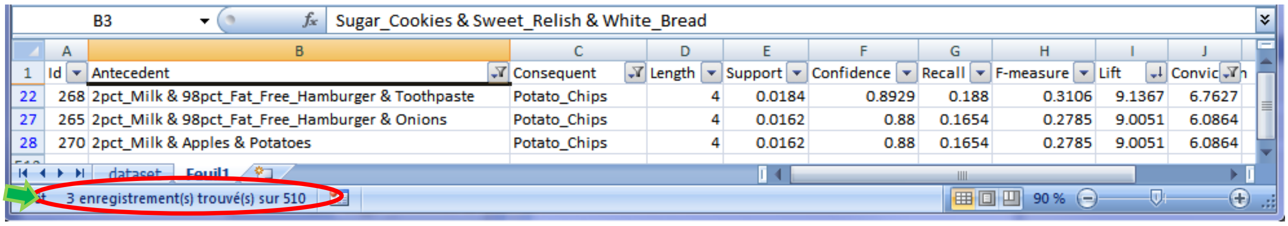
Afficher les lignes dans lesquelles :
Antecedent

contient 2pct_Milk

● Et ○ Ou

Utilisez ? pour représenter un caractère
Utilisez * pour représenter une série de caractères

3 rules are now displayed.


B3 Sugar_Cookies & Sweet_Relish & White_Bread

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Id | Antecedent | Consequent | Length | Support | Confidence | Recall | F-measure | Lift | Convic…n |
| 22 | 268 | 2pct_Milk & 98pct_Fat_Free_Hamburger & Toothpaste | Potato_Chips | 4 | 0.0184 | 0.8929 | 0.188 | 0.3106 | 9.1367 | 6.7627 |
| 27 | 265 | 2pct_Milk & 98pct_Fat_Free_Hamburger & Onions | Potato_Chips | 4 | 0.0162 | 0.88 | 0.1654 | 0.2785 | 9.0051 | 6.0864 |
| 28 | 270 | 2pct_Milk & Apples & Potatoes | Potato_Chips | 4 | 0.0162 | 0.88 | 0.1654 | 0.2785 | 9.0051 | 6.0864 |

3 enregistrement(s) trouvé(s) sur 510

We can thus multiply the combinations to highlight the rules that best meet the specifications of our study. To be honest, this system is only really operational if we deal with a moderate-size rules base. But Excel is easy to use. It is one of the data miner favorite tools for a long time (KDnuggets Polls, Top Analytics / Data Science Tools, May 2017). We note in this tutorial that its functionalities are attractive for the processing of rule bases.

# 5 Conclusion

ARS is an academic software for association rule mining. It is included into the SIPINA distribution. In this tutorial, we describe its combined use with Excel. Among other things, the opportunities about the post-processing of the mined rules is particularly attractive.