

# Market Basket Analysis

# author: Marcin Dawidowski

Głównym założeniem projektu jest analiza kupowanych zestawów rzeczy za pomocą napisanego przez siebie algorytmu Apriori do odkrywania reguł asocjacyjnych.

## Realizacja

Projekt został wykonany w języku Java. Aplikacja korzysta z bazy danych SQLite, w której przechowywane są dane. Aplikacja wyposażona została w interfejs graficzny, w którym możliwy jest wybór produktu dla którego szukamy reguł asocjacyjnych.

## Wygląd aplikacji

Wybierz produkt (1 z 17860)      łączna ilość transakcji: 29901

4710265849066 ▼

Produkt ID	Ilość wystąpień	Support ▼	Confidence	Lift	Conviction
4710168705056	13	0,000434768	0,01416122	1,91599381	0,52192235
4710088410610	11	0,000367881	0,01198257	0,773846328	1,292246222
4710685440362	9	0,000300993	0,009803921	1,91599381	0,52192235
4710105002019	5	0,000167218	0,005446623	1,155031681	0,865777075
4710466101130	3	0,000100331	0,003267974	0,280792177	3,56135273
4713089006103	2	0,000066887	0,002178649	1,714310169	0,583325028
0078895121247	2	0,000066887	0,002178649	1,809549809	0,55262363
4711564801656	2	0,000066887	0,002178649	7,238199234	0,138155907
4710043100082	2	0,000066887	0,002178649	0,824604928	1,212701917
4710176123798	1	0,000033444	0,001089325	1,480540633	0,675428927
4901987221210	1	0,000033444	0,001089325	2,326563835	0,429818422
4710032700026	1	0,000033444	0,001089325	1,480540633	0,675428927
4710363196000	1	0,000033444	0,001089325	3,619099617	0,276311815
4711258001508	1	0,000033444	0,001089325	1,252765179	0,798234165
2220122002270	1	0,000033444	0,001089325	4,65212767	0,214000211

Oblicz      Przygotuj dane      Zamknij

Reguły oceniane są pod względem różnych miar:

- support czyli wsparcie (łączne wystąpienia produktów/ilość wszystkich transakcji)
- confidence czyli wiarygodność (support pary produktów/support wybranego produktu)
- lift określa korelację pomiędzy zdarzeniami
  - Lift = 1 – zdarzenia niezależne
  - Lift < 1 – zdarzenia skorelowane negatywnie
  - Lift > 1 – zdarzenia skorelowane pozytywnie
- conviction czyli przekonanie. Jest to odwrotność wartości lift

Dodano także opcję sortowania odkrytych reguł wg każdej z użytych miar.