



UNIVERSIDAD DE GRANADA

Práctica 5. Reglas de asociación.

Tratamiento Inteligente de Datos.

Máster Profesional en Ingeniería Informática

Curso académico 2022/2023

Autor

José Alberto Gómez García

Índice

1. Introducción.....	3
2. Análisis de la base de datos.	3
3. Primer intento de generación de reglas.	4
4. Agrupación de los elementos.	6
5. Segundo intento de generación de reglas.	7
6. Conclusiones.	9

1. Introducción.

En esta práctica utilizaremos reglas de asociación en un intento de descubrir patrones o reglas en las compras que realizan los clientes de un determinado supermercado. Los descubrimientos realizados podrían permitirnos aconsejar a los encargados del supermercado reorganizar sus productos o coordinar ofertas entre ciertos tipos de productos, todo con el objetivo de incrementar las ventas y ganancias.

En la entrega de Prado se adjunta esta memoria y el correspondiente “workflow” de KNIME. Adicionalmente, se adjunta el script Python con el que han sido agrupados los datos (véase sección 4) y la base de datos resultante de la modificación.

2. Análisis de la base de datos.

Tras echar un primer vistazo a la base de datos, de unas 7500 compras aproximadamente, podemos ver que existe una gran disparidad en el número de artículos que compran los diferentes clientes. Las compras de mayor tamaño tienen 19 elementos, mientras que también tenemos numerosas compras de un único artículo, existiendo en medio compras más frecuentes de entre 3 y 8 artículos.

Con un pequeño script Python, hemos calculado cuantas veces se han comprado los diferentes artículos, y cuales han sido. El resultado es el que se muestra a continuación:

```
Número de elementos originales: 120
mineral water: 1788, eggs: 1348, spaghetti: 1306, french fries: 1282, chocolate: 1230, green tea:
991, milk: 972, ground beef: 737, frozen vegetables: 715, pancakes: 713, burgers: 654, cake: 608,
cookies: 603, escalope: 595, low fat yogurt: 574, shrimp: 536, tomatoes: 513, olive oil: 494, frozen
smoothie: 475, turkey: 469, chicken: 450, whole wheat rice: 439, grated cheese: 393, cooking oil:
383, soup: 379, herb & pepper: 371, honey: 356, champagne: 351, fresh bread: 323, salmon: 319,
brownies: 253, avocado: 250, hot dogs: 243, cottage cheese: 239, tomato juice: 228, butter: 226,
whole wheat pasta: 221, red wine: 211, yogurt cake: 205, light mayo: 204, energy bar: 203, ham: 203,
energy drink: 200, pepper: 199, vegetables mix: 193, cereals: 193, muffins: 181, oil: 173, french
wine: 169, fresh tuna: 167, strawberries: 160, meatballs: 157, almonds: 153, parmesan cheese: 149,
mushroom cream sauce: 143, rice: 141, protein bar: 139, mint: 131, white wine: 124, pasta: 118, light
cream: 117, carrots: 115, black tea: 107, tomato sauce: 106, fromage blanc: 102, gums: 101, eggplant:
99, extra dark chocolate: 90, melons: 90, yams: 86, body spray: 86, magazines: 82, barbecue sauce:
81, cider: 79, nonfat milk: 78, candy bars: 73, zucchini: 71, whole wheat flour: 70, salt: 69,
blueberries: 69, green grapes: 68, flax seed: 68, antioxidant juice: 67, bug spray: 65, bacon: 65,
green beans: 65, clothes accessories: 63, toothpaste: 61, shallot: 58, strong cheese: 58, spinach:
53, gluten free bar: 52, pet food: 49, sparkling water: 47, soda: 47, mayonnaise: 46, chili: 46,
pickles: 45, burger sauce: 44, mint green tea: 42, hand protein bar: 39, salad: 37, shampoo: 37,
corn: 36, cauliflower: 36, asparagus: 35, sandwich: 34, babies food: 34, dessert wine: 33, ketchup:
33, oatmeal: 33, chocolate bread: 32, chutney: 31, mashed potato: 31, tea: 29, bramble: 14, cream: 7,
napkins: 5, water spray: 3, asparagus: 1
```

Imagen 1. Elementos únicos y número de apariciones en la base de datos.

Durante este proceso encontramos una errata, y es que “asparagus” aparece en una ocasión con un espacio delante suya. Aprovechamos para corregir esto.

Se puede observar que los elementos que más se han comprado han sido agua mineral, huevos, espaguetis, patatas fritas, chocolate, té verde, etc. Por tanto, presuponemos que tendrán gran presencia e incluso influencia en las reglas de asociación que obtengamos posteriormente.

Otros artículos muchos menos comprados, como los pañales para bebé, las servilletas o la comida para mascotas no creemos que aparezcan en ninguna regla debido a su bajo soporte.

3. Primer intento de generación de reglas.

En un primer momento, intentaremos generar reglas de asociación haciendo uso la base de datos sin modificar (salvo la errata anteriormente mencionada).

Para poder cargar este fichero “.csv” a KNIME haremos uso del nodo “CSV Reader”, en el que deberemos marcar la opción “support short data rows”, dado que muchas filas no tienen tantos elementos como la compra con mayor número de artículos, y KNIME se quejará diciendo que “las filas de datos tienen demasiados pocos elementos”.

A continuación, necesitaremos convertir la base de datos relacional en transaccional, de manera que los algoritmos de reglas de asociación puedan trabajar. Para ello, hacemos uso del nodo “Create Collection Column”. En la configuración de este nodo le diremos que no almacene valores duplicados - puesto que no nos importa cuantas veces se compró un elemento, sino solo el hecho de ser comprado - y que ignore valores nulos, los cuales se generaron en gran medida en el nodo de lectura. Esto genera una nueva columna en nuestra tabla con la agrupación de los items de cada fila, una transacción.

Hecho esto, podremos empezar a trabajar con los algoritmos de reglas de asociación.

Dada la gran cantidad de compras y de artículos distintos que hay, se espera que el soporte de los diferentes conjuntos de elementos frecuentes y reglas de asociación sea bastante bajo (inferior al 5%).

Si hacemos uso del nodo “Association Rule Learner” con un soporte mínimo del 5% y una confianza mínima del 30% sólo obtendremos dos reglas. Estas nos dirán que si se compran espaguetis se compra también agua mineral – lo cual puede tener sentido, pues para cocer los espaguetis necesitas agua – y que si se compra chocolate se compra agua mineral.

Table "default" - Rows: 2 Spec - Columns: 6 Properties Flow Variables						
Row ID	D Support	D Confide...	D Lift	S Conseq...	S implies	[...] Items
rule0	0.053	0.321	1.348	mineral water	<---	[chocolate]
rule1	0.06	0.343	1.439	mineral water	<---	[spaghetti]

Imagen 2. Reglas de asociación. Base de datos original. Soporte 5% y confianza 30%.

Ninguna de estas reglas parece proporcionarnos información útil, por lo que disminuirémos algo más el porcentaje de soporte, hasta el 1%, para obtener reglas adicionales y poder examinar si alguna es útil o novedosa.

En este caso, obtenemos 63 reglas. En la mayoría de ellas, el consecuente es la compra de espaguetis o agua mineral, pero tenemos 6 reglas con consecuentes diferentes. A continuación, se muestran imágenes con los resultados obtenidos.

Row ID	[D] Support	[D] Confide...	[D] Lift	[S] ▼ Con...	[S] implies	[...] Items
rule2	0.01	0.3	1.725	spaghetti	<---	[mineral water, french fries]
rule4	0.01	0.365	2.096	spaghetti	<---	[red wine]
rule6	0.01	0.372	2.136	spaghetti	<---	[mineral water, olive oil]
rule7	0.011	0.317	1.822	spaghetti	<---	[eggs, chocolate]
rule10	0.011	0.34	1.954	spaghetti	<---	[milk, chocolate]
rule16	0.011	0.34	1.952	spaghetti	<---	[mineral water, pancakes]
rule19	0.012	0.336	1.929	spaghetti	<---	[mineral water, frozen vegetables]
rule22	0.013	0.317	1.818	spaghetti	<---	[salmon]
rule29	0.016	0.328	1.883	spaghetti	<---	[mineral water, milk]
rule30	0.016	0.311	1.785	spaghetti	<---	[cooking oil]
rule32	0.016	0.301	1.73	spaghetti	<---	[mineral water, chocolate]
rule34	0.016	0.329	1.889	spaghetti	<---	[herb & pepper]
rule35	0.017	0.316	1.812	spaghetti	<---	[grated cheese]
rule39	0.017	0.417	2.395	spaghetti	<---	[mineral water, ground beef]
rule46	0.021	0.306	1.758	spaghetti	<---	[tomatoes]
rule48	0.023	0.348	2	spaghetti	<---	[olive oil]
rule58	0.039	0.399	2.291	spaghetti	<---	[ground beef]
rule0	0.01	0.507	2.126	mineral water	<---	[eggs, ground beef]
rule1	0.01	0.367	1.54	mineral water	<---	[spaghetti, french fries]
rule3	0.01	0.399	1.674	mineral water	<---	[cereals]
rule5	0.01	0.448	1.878	mineral water	<---	[spaghetti, olive oil]
rule8	0.011	0.389	1.63	mineral water	<---	[red wine]
rule9	0.011	0.474	1.988	mineral water	<---	[ground beef, chocolate]
rule12	0.011	0.469	1.967	mineral water	<---	[frozen vegetables, milk]
rule14	0.011	0.503	2.11	mineral water	<---	[milk, ground beef]
rule15	0.011	0.455	1.909	mineral water	<---	[spaghetti, pancakes]
rule17	0.012	0.348	1.46	mineral water	<---	[avocado]
rule18	0.012	0.431	1.807	mineral water	<---	[spaghetti, frozen vegetables]
rule20	0.013	0.424	1.78	mineral water	<---	[eggs, milk]
rule21	0.013	0.31	1.299	mineral water	<---	[fresh bread]
rule23	0.013	0.406	1.702	mineral water	<---	[eggs, chocolate]
rule24	0.014	0.436	1.828	mineral water	<---	[milk, chocolate]
rule25	0.014	0.391	1.638	mineral water	<---	[eggs, spaghetti]
rule26	0.015	0.317	1.332	mineral water	<---	[honey]
rule28	0.016	0.444	1.861	mineral water	<---	[spaghetti, milk]
rule31	0.016	0.405	1.698	mineral water	<---	[spaghetti, chocolate]
rule36	0.017	0.401	1.683	mineral water	<---	[salmon]
rule37	0.017	0.345	1.447	mineral water	<---	[herb & pepper]
rule38	0.017	0.435	1.826	mineral water	<---	[spaghetti, ground beef]
rule40	0.017	0.333	1.398	mineral water	<---	[grated cheese]
rule41	0.019	0.307	1.288	mineral water	<---	[turkey]
rule43	0.02	0.344	1.443	mineral water	<---	[whole wheat rice]
rule44	0.02	0.394	1.654	mineral water	<---	[cooking oil]
rule45	0.02	0.32	1.342	mineral water	<---	[frozen smoothie]
rule47	0.023	0.38	1.594	mineral water	<---	[chicken]
rule49	0.023	0.456	1.915	mineral water	<---	[soup]
rule50	0.024	0.33	1.385	mineral water	<---	[shrimp]
rule51	0.024	0.314	1.316	mineral water	<---	[low fat yogurt]
rule52	0.024	0.357	1.497	mineral water	<---	[tomatoes]
rule53	0.027	0.339	1.421	mineral water	<---	[cake]
rule54	0.028	0.419	1.758	mineral water	<---	[olive oil]
rule56	0.034	0.355	1.489	mineral water	<---	[pancakes]
rule57	0.036	0.375	1.572	mineral water	<---	[frozen vegetables]
rule59	0.041	0.417	1.748	mineral water	<---	[ground beef]
rule60	0.048	0.37	1.554	mineral water	<---	[milk]
rule61	0.053	0.321	1.348	mineral water	<---	[chocolate]
rule62	0.06	0.343	1.439	mineral water	<---	[spaghetti]
rule13	0.011	0.31	2.39	milk	<---	[mineral water, frozen vegetables]
rule27	0.015	0.301	2.321	milk	<---	[soup]
rule33	0.016	0.323	3.292	ground beef	<---	[herb & pepper]
rule42	0.019	0.311	1.732	eggs	<---	[turkey]
rule55	0.029	0.33	1.838	eggs	<---	[burgers]
rule11	0.011	0.308	1.881	chocolate	<---	[spaghetti, milk]

Imagen 3. Reglas de asociación. Base de datos original. Soporte 1% y confianza 30%.

De estas reglas tampoco parece que podamos extraer demasiada información. El agua mineral es un elemento presente en gran cantidad de compras, y los espaguetis suelen comprarse junto a otros ingredientes de recetas habituales, como pimientos, aceite de oliva, agua, tomate, verduras, distintos tipos de carne, etc. Destaca que haya algunas ocasiones en las que se compre con chocolate, leche o tortitas.

En las últimas seis reglas de asociación podemos ver una con un soporte cercano al 3%, y es que cuando se compran hamburguesas se compran huevos, algo que es relativamente obvio pues combinan bastante bien. Cuando se compra pavo se compran también huevos, con un soporte del 2%, y esto parece algo menos común.

La regla con más empuje (o ratio de mejora) es aquella que nos dice que cuando se compra “hierbas y pimienta”, también se compra carne picada; lo cual es bastante lógico pues se usa de aderezo.

El tener muchos elementos distintos provoca soportes bajos y que las asociaciones sean entre pocos elementos y muy específicas. En un intento de mejorar los resultados y obtener reglas más interesantes, vamos a agrupar los elementos en categorías.

4. Agrupación de los elementos.

A partir de los datos que se mostraban en la figura 1, hemos analizado los artículos vendidos en el supermercado y los hemos agrupado – en su mayoría – en categorías de mayor tamaño. Así pues, realizamos las agrupaciones en las categorías que se muestran en la siguiente imagen; en la cual también aparece cuantas ocasiones se han venido artículos de cada grupo.

```
Número de nuevas categorías: 26  
carne: 3573, postres_desayunos: 2821, lacteos: 2449, verdura: 2321, agua:  
1838, salsas_condimentos: 1729, pasta: 1645, eggs: 1348, chocolate: 1320,  
patatas: 1313, te: 1169, aceites: 1050, pescado: 1022, alcohol: 967, quesos:  
941, fruta: 687, arroz: 580, barritas: 506, otros: 399, panaderia: 393,  
soup: 379, zumos: 295, refrescos: 247, aseo: 184, almonds: 153, sandwich: 34
```

Imagen 4. Nuevas agrupaciones de productos.

Nótese que hay elementos que no se agrupan, como las sopas o huevos, por considerarse artículos con suficiente importancia para ser una categoría por ellos mismos. Además, tampoco se agrupan las almendras (pues no hay más tipos de frutos secos), ni los sándwiches, los cuales pueden ser de muchos tipos. Finalmente, cabe mencionar que se ha creado una categoría “otros” en la que se enmarcan productos de difícil clasificación en el contexto del problema, como comida para bebé, comida para mascotas, revistas o spray para los insectos.

Como podíamos intuir viendo las cantidades de los artículos de forma individualizada, lo que más se compra es carne, productos que se pueden enmarcar en postres y desayunos (tartas, tortitas, cereales, etc.), productos lácteos, verduras y agua. Sorprende la gran

cantidad de salsas y condimentos, entre los que hemos enmarcado la sal y pimienta, mayonesas, ketchup y demás.

5. Segundo intento de generación de reglas.

Con las agrupaciones anteriormente mencionadas hechas, cargamos la nueva base de datos y la volvemos a convertir en transaccional. No tenemos porque preocuparnos por los valores repetidos que puedan haberse generado dentro de una misma compra tras el agrupamiento, pues el nodo “Create Collection Column” tiene una opción para conservar un único elemento de los repetidos en una transacción.

Si utilizamos el nodo “Association Rule Learner” especificando un soporte mínimo del 2% y una confianza mínima del 40%, obtenemos 194 reglas. Dado que no es factible mostrar todas en el documento, analizaremos por partes aquellas con más soporte, con más confianza y con más “lift” (ratio de mejora).

Comencemos con las reglas con más soporte, las cuales mostramos en la siguiente imagen:

Row ID	D ▼ Sup...	D Confide...	D ▼ Lift	S Conseq...	S implies	[...] Items
rule 193	0.121	0.443	1.207	carne	<---	[lacteos]
rule 192	0.119	0.466	1.271	carne	<---	[verdura]
rule 191	0.115	0.471	1.284	carne	<---	[agua]
rule 190	0.111	0.531	1.447	carne	<---	[pasta]
rule 189	0.097	0.478	1.303	carne	<---	[salsas_condimentos]
rule 188	0.085	0.471	1.284	carne	<---	[eggs]
rule 187	0.077	0.437	1.191	carne	<---	[chocolate]
rule 186	0.07	0.463	1.261	carne	<---	[te]
rule 185	0.07	0.401	1.093	carne	<---	[patatas]
rule 184	0.064	0.492	1.34	carne	<---	[aceites]
rule 183	0.061	0.479	1.304	carne	<---	[pescado]
rule 182	0.058	0.456	1.782	verdura	<---	[pescado]
rule 181	0.056	0.469	1.278	carne	<---	[quesos]
rule 180	0.054	0.414	1.516	lacteos	<---	[aceites]
rule 179	0.054	0.418	1.533	lacteos	<---	[pescado]

Imagen 5. Reglas de asociación ordenadas por soporte. Base de datos agrupada. Soporte 2% y confianza 40%.

Como era de esperar, predominan aquellas en los que aparecen la carne, las verduras y la pasta, en tanto que son el tipo de artículos más comprados. Podemos observar algunas reglas curiosas, como que aquellas que nos dicen que las personas que compren té o chocolate compren también carne, con soportes del 7% y confianzas cercanas al 45%; o la última, que nos dice que quien compra pescado también hecha al carrito productos lácteos.

El resto de reglas parecen ser bastante lógicas, en tanto que se compra carne con muchos ingredientes que la complementa; y se compra verdura tanto con carne y pescado (al servir de acompañamiento).

En el caso de ordenar por soporte, obtenemos en las primeras reglas que el consecuente es la carne, la cual es comprada junto a otros productos en diversas combinaciones. En esta ocasión comenzamos a ver antecedentes en los que hay 2 elementos.

De entre estas reglas me llama la atención que se mezcle en una misma compra chocolate y carne, té y carne, o productos del desayuno y carne. Entiendo que estas corresponderán a compras con mayor número de elementos, y no tanto a compras de pequeño tamaño.

Row ID	D Support	D Con...	D Lift	S Conseq...	S implies	[...] Items
rule90	0.026	0.617	1.682	carne	<---	[eggs,pasta]
rule151	0.038	0.617	1.682	carne	<---	[pasta,salsas_condimentos]
rule25	0.022	0.616	1.678	carne	<---	[pasta,patas]
rule13	0.021	0.606	1.652	carne	<---	[lacteos,pasta,verdura]
rule161	0.041	0.602	1.639	carne	<---	[pasta,agua]
rule42	0.022	0.598	1.629	carne	<---	[lacteos,postres_desayunos,agua]
rule63	0.023	0.596	1.623	carne	<---	[te,pasta]
rule122	0.029	0.596	1.623	carne	<---	[pasta,chocolate]
rule165	0.045	0.593	1.616	carne	<---	[lacteos,pasta]
rule118	0.029	0.591	1.609	carne	<---	[aceites,pasta]
rule145	0.036	0.585	1.595	carne	<---	[salsas_condimentos,agua]
rule64	0.023	0.585	1.594	carne	<---	[pescado,salsas_condimentos]
rule160	0.04	0.582	1.586	carne	<---	[pasta,postres_desayunos]
rule100	0.027	0.581	1.584	carne	<---	[pescado,pasta]
rule167	0.045	0.58	1.58	carne	<---	[pasta,verdura]
rule32	0.022	0.577	1.573	carne	<---	[aceites,chocolate]
rule96	0.026	0.571	1.557	carne	<---	[aceites,postres_desayunos]
rule124	0.029	0.571	1.557	carne	<---	[aceites,agua]
rule51	0.022	0.568	1.549	carne	<---	[lacteos,verdura,agua]

Imagen 6. Reglas de asociación ordenadas por confianza. Base de datos agrupada. Soporte 2% y confianza 40%.

Si ordenamos por “lift”, o ratio de mejora, veremos que la carne aparece en menos ocasiones. Las reglas que más cantidad de nueva información nos aportan son las que nos relacionan cosas que pueden no tener demasiado sentido, como la pasta con postres y desayunos (“lift” más alto), pasta y productos lácteos (aunque por recetas como la carbonara podría ser lógico), o verdura, quesos y productos lácteos.

Row ID	D Support	D Conf...	D Lift	S Consequent	S implies	[...] Items
rule33	0.022	0.478	2.298	pasta	<---	[aceites,postres_desayunos]
rule119	0.029	0.449	2.157	pasta	<---	[aceites,carne]
rule15	0.021	0.537	2.101	verdura	<---	[pescado,salsas_condimentos]
rule101	0.027	0.435	2.089	pasta	<---	[pescado,carne]
rule55	0.022	0.434	2.084	pasta	<---	[aceites,agua]
rule18	0.021	0.43	2.065	pasta	<---	[aceites,verdura]
rule111	0.028	0.525	2.052	verdura	<---	[pescado,lacteos]
rule52	0.022	0.517	2.021	verdura	<---	[pescado,postres_desayunos]
rule73	0.024	0.515	2.011	verdura	<---	[pescado,pasta]
rule132	0.031	0.507	1.98	verdura	<---	[pescado,carne]
rule29	0.022	0.407	1.955	pasta	<---	[aceites,lacteos]
rule41	0.022	0.474	1.949	agua	<---	[lacteos,postres_desayunos,carne]
rule62	0.023	0.405	1.947	pasta	<---	[quesos,carne]
rule74	0.024	0.404	1.942	pasta	<---	[pescado,verdura]
rule14	0.021	0.403	1.934	pasta	<---	[lacteos,carne,verdura]
rule66	0.023	0.462	1.898	agua	<---	[soup]
rule56	0.022	0.484	1.892	verdura	<---	[pescado,agua]
rule123	0.029	0.459	1.888	agua	<---	[aceites,carne]
rule54	0.022	0.459	1.886	agua	<---	[aceites,pasta]
rule9	0.021	0.458	1.881	agua	<---	[aceites,postres_desayunos]
rule60	0.023	0.457	1.877	agua	<---	[aceites,verdura]
rule5	0.021	0.477	1.864	verdura	<---	[quesos,lacteos]

Imagen 7. Reglas de asociación ordenadas por “lift”. Base de datos agrupada. Soporte 2% y confianza 40%.

Alguna regla algo más curiosa es aquella que nos dice que si se compran salsas/condimentos y agua, se compran productos del desayuno y postres. En los condimentos se enmarca la miel (aparece en 356 compras), pero la mayoría del

agrupamiento corresponde a artículos como mayonesas, ketchup o salsa barbacoa o salsa de setas.

Si pasamos a generar reglas con un soporte mínimo del 5%, manteniendo la confianza mínima al 40%, el número de resultados se reduce hasta las 19 ocurrencias. Estas se pueden ver en la tabla siguiente:

▼ Row ID	[D] Support	[D] Confide...	[D] ▾ Lift	[S] Conseq...	[S] implies	[...] Items
rule9	0.064	0.492	1.34	carne	<---	[aceites]
rule8	0.061	0.479	1.304	carne	<---	[pescado]
rule7	0.058	0.456	1.782	verdura	<---	[pescado]
rule6	0.056	0.469	1.278	carne	<---	[quesos]
rule5	0.054	0.414	1.516	lacteos	<---	[aceites]
rule4	0.054	0.418	1.533	lacteos	<---	[pescado]
rule3	0.052	0.433	1.178	carne	<---	[alcohol]
rule2	0.052	0.535	1.458	carne	<---	[lacteos,verdura]
rule18	0.121	0.443	1.207	carne	<---	[lacteos]
rule17	0.119	0.466	1.271	carne	<---	[verdura]
rule16	0.115	0.471	1.284	carne	<---	[agua]
rule15	0.111	0.531	1.447	carne	<---	[pasta]
rule14	0.097	0.478	1.303	carne	<---	[salsas_condimentos]
rule13	0.085	0.471	1.284	carne	<---	[eggs]
rule12	0.077	0.437	1.191	carne	<---	[chocolate]
rule11	0.07	0.463	1.261	carne	<---	[te]
rule10	0.07	0.401	1.093	carne	<---	[patatas]
rule1	0.052	0.436	1.597	lacteos	<---	[carne,verdura]
rule0	0.052	0.43	1.681	verdura	<---	[lacteos,carne]

Imagen 8. Reglas de asociación. Base de datos agrupada. Soporte 5% y confianza 40%.

En este caso las reglas son bastante lógicas y no parecen proporcionarnos información relevante (bajos valores de “lift”, pero superiores a 1). Las reglas quizás más interesantes son las que nos dicen que lácteos y aceites se compran juntos, así como que productos lácteos, carne y verduras suelen acabar en un mismo carrito de la compra. Dos reglas que no habíamos mencionado hasta el momento son aquellas que nos dicen que si se compra alguna bebida alcohólica, también se suele comprar carne, lo cual era esperable hasta cierto punto; y que si se compra queso también se compra carne.

6. Conclusiones.

- La generación de reglas de asociación en el marco de una serie de compras en un supermercado es un problema complejo en el cual no es fácil extraer información radicalmente novedosa o que se desconozca a priori, puesto que muchas reglas tienen una lógica muy marcada detrás.
- A la luz de los resultados, se podría recomendar al dueño del supermercado que ubique los productos cárnicos, el pescado, las verduras y la pasta en posiciones algo distantes del supermercado, de manera que en el tránsito entre ellas los clientes se fijen en productos menos vendidos. También, conviene que los aderezos se encuentren entre las carnes y las pastas.

- Dado que el agua mineral es uno de los elementos más comprados, podría ser conveniente colocarla al final del supermercado, de manera que los clientes tuvieran que recorrerlo entero, y se sintieran tentados de divagar en la travesía y comprar algún otro producto que no esperaban echar al carrito en primera instancia.