



datascience.pe

Tema

**Market Basket y Recomendación de
Productos en Retail: Una revisión de
Casos**



Expositor

Luis Cajachahua – Principal Data Scientist
Center Of Excellence, Teradata Americas



+51 987 654 321



lcajachahua@gmail.com

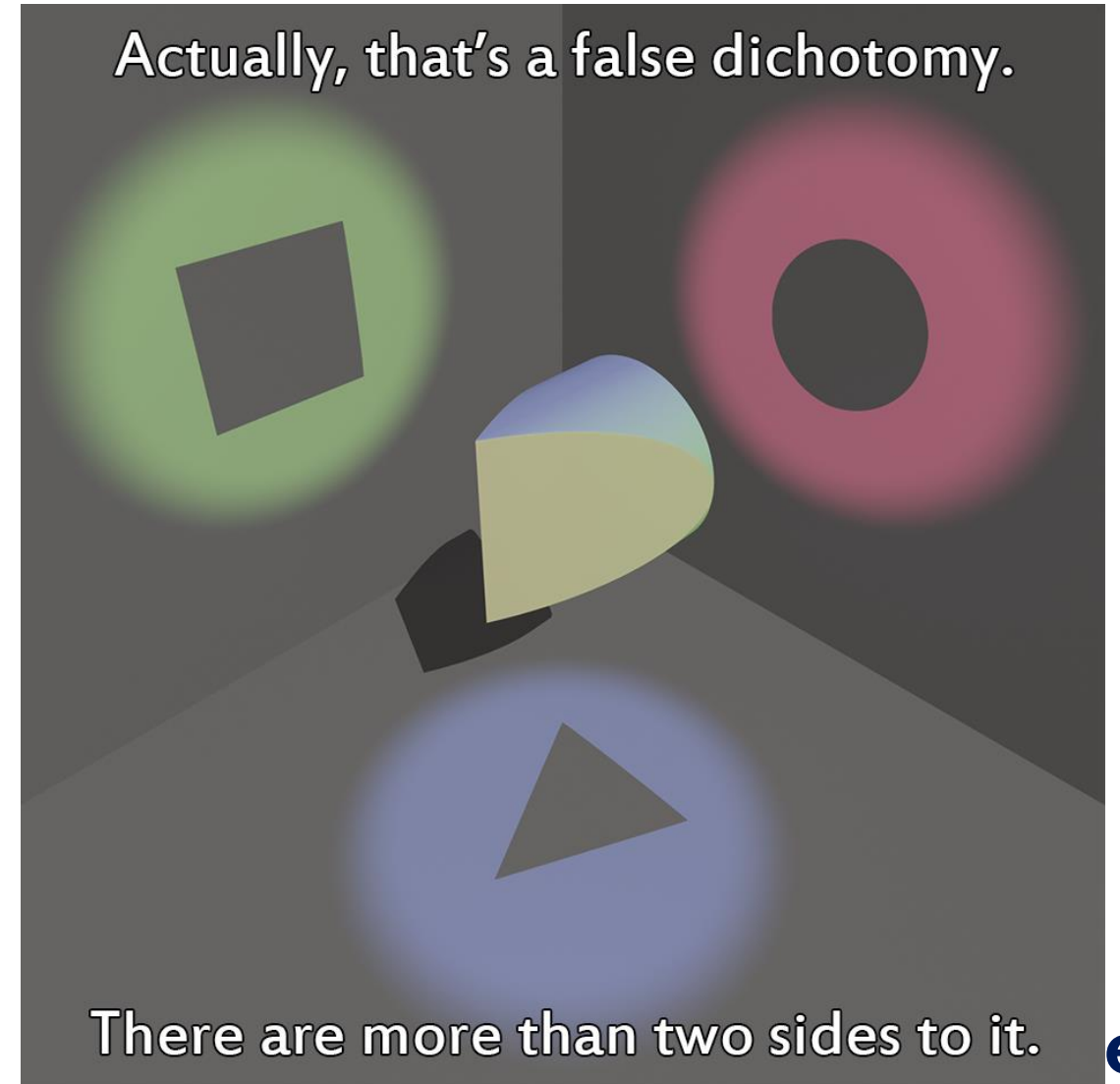
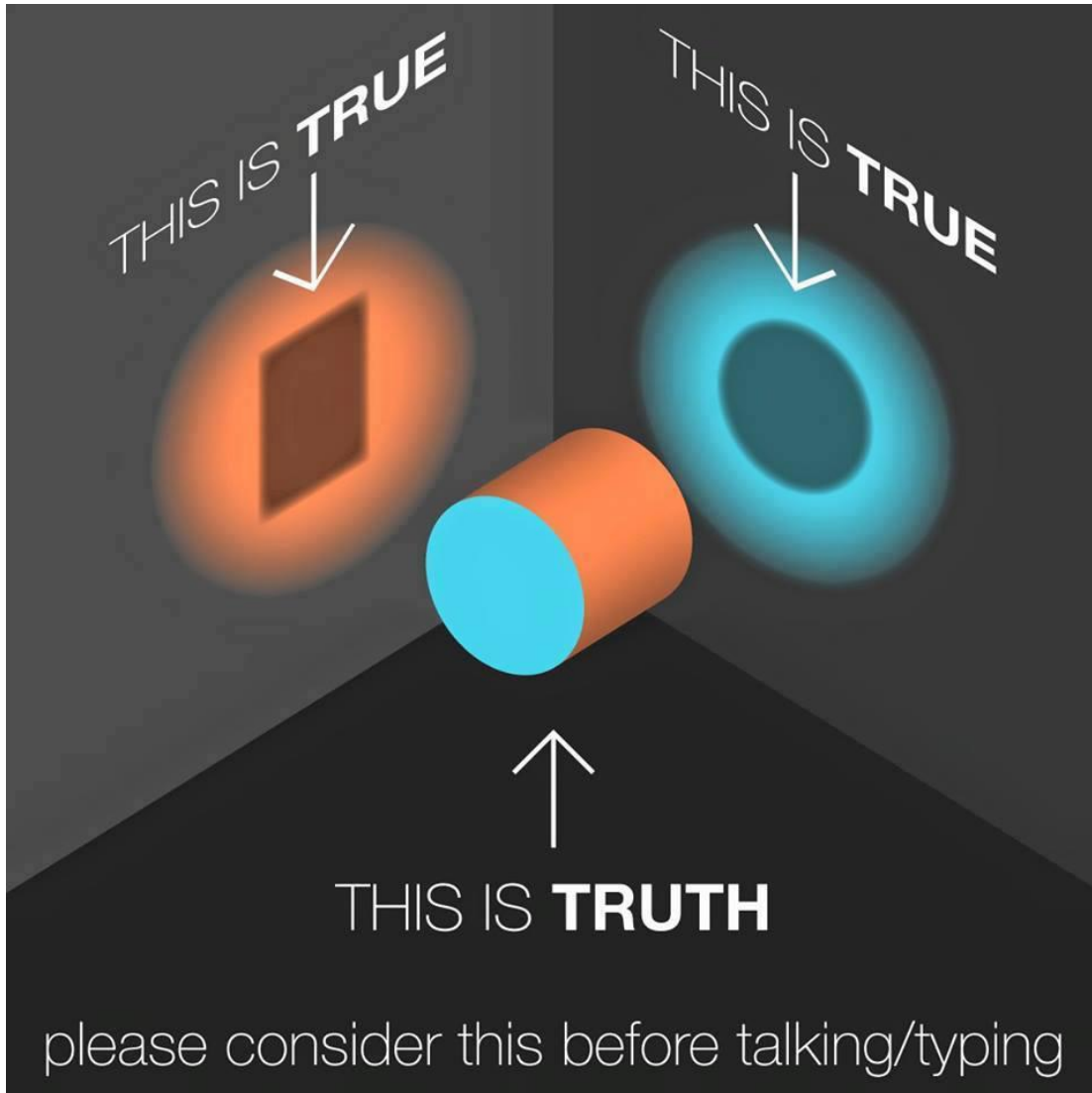


[linkedin.com/in/lcajachahua/](https://www.linkedin.com/in/lcajachahua/)

Agenda

- Market Basket Analysis
- Recomendación de Productos
- Cazando mitos
- Accionamiento de los Insights
- Conclusiones

Antes de empezar...



Market Basket Analysis

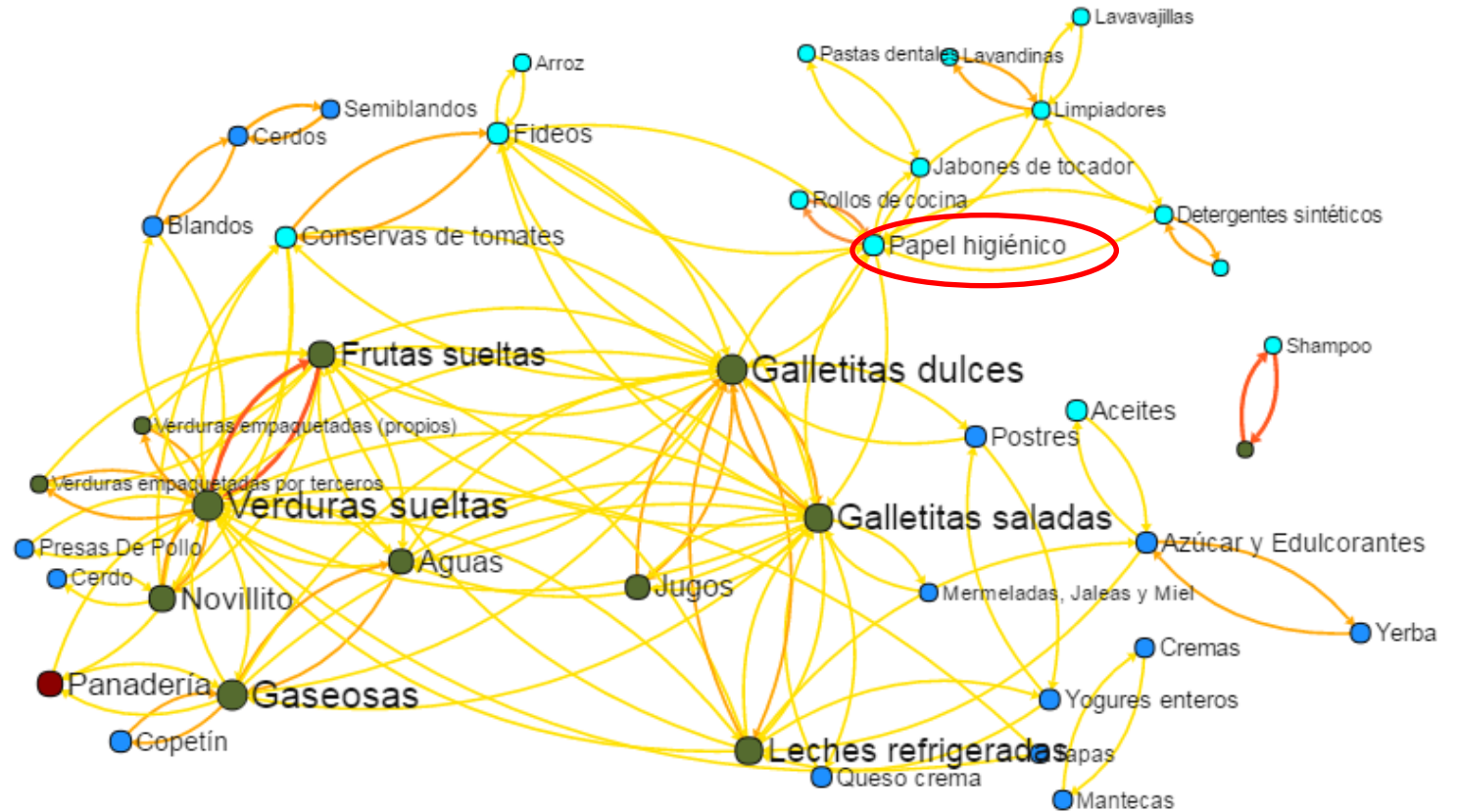
También conocido como:

- Análisis de Afinidad
- Análisis de Asociación
- Reglas de Asociación

“Si compra Frutas sueltas, es altamente probable que compre Verduras sueltas”

Algoritmos:

- APriori
- Frequent Pattern Growth (FPGrowth)
- Direct Hashing Pruning (DHP)



Recomendación de Productos

Accionamiento destinado a:

- Generar más rotación de stocks
- Generar más ventas orgánicas y cruzadas
- Aumentar el ticket promedio por cliente
- Mejorar la Experiencia del usuario
- Personalizar la oferta
- Ganar Relevancia y Engagement

Aspectos importantes en la implementación:

- Canal
- Tipo de Producto/Negocio
- Propuesta de valor
- Oportunidad
- Diferenciación



CASEY JOHNSTON, ARS TECHNICA BUSINESS 04.16.12 08:20 AM

SHARE

SHARE
11

TWEET



COMMENT



EMAIL

NETFLIX NEVER USED ITS \$1 MILLION ALGORITHM DUE TO ENGINEERING COSTS

Netflix awarded a \$1 million prize to a developer team in 2009 for an algorithm that increased the accuracy of the company's recommendation engine by 10 percent. But it doesn't use the million-dollar code, and has no plans to implement it in the future, Netflix [announced](#) on its blog Friday. The post goes on to explain why: a combination of too much engineering effort for the results, and a shift from movie recommendations to the "next level" of personalization caused by the transition of the business from mailed DVDs to video streaming.

Netflix notes that it does still use two algorithms from the team that won the [first Progress Prize](#) for an 8.43 percent improvement to the recommendation engine's root mean squared error (the full \$1 million was [awarded](#) for a 10 percent improvement). But the increase in accuracy on the winning improvements "did not seem to justify the engineering effort needed to bring them into a production environment," the blog post said.

| Rank | Team Name | Best Score | % Improvement | Last Submit Time |
|--|--------------------|------------|---------------|---------------------|
| Grand Prize - \$1M - 0.81563 | | | | |
| 1 | RecommenderTeam | 0.81563 | 9.78 | 2009-09-10 01:54:47 |
| 2 | Netflix.co.uk/Team | 0.81563 | 9.71 | 2009-05-12 08:14:08 |
| 3 | Grand Prize Team | 0.81563 | 9.68 | 2009-05-12 08:20:24 |
| 4 | Clus | 0.81563 | 9.66 | 2009-04-22 06:37:01 |
| 5 | NetflixTeam | 0.81563 | 9.47 | 2009-06-15 18:52:05 |
| Progress Prize - 2008 - \$50K - 0.8424 - Winning Team: Netflix.co.uk/Team | | | | |
| 6 | Netflix | 0.8424 | 9.48 | 2009-09-17 13:47:48 |
| 7 | Clus | 0.8424 | 9.25 | 2009-04-22 18:31:32 |
| 8 | Clus & Netflix | 0.8424 | 9.19 | 2009-09-09 22:24:51 |
| 9 | Netflix | 0.8424 | 9.18 | 2009-09-17 12:47:27 |
| 10 | Netflix.co.uk/Team | 0.8424 | 9.18 | 2009-09-02 17:30:31 |
| 11 | Clus | 0.8424 | 9.17 | 2009-09-12 23:34:29 |
| 12 | mpg2 | 0.8424 | 9.17 | 2009-09-10 01:54:47 |
| 13 | Netflix | 0.8424 | 9.17 | 2009-09-10 01:54:47 |
| 14 | Clus | 0.8424 | 9.11 | 2009-09-10 22:21:19 |
| 15 | Netflix.co.uk/Team | 0.8424 | 9.08 | 2009-09-04 19:52:54 |
| 16 | Team ESP | 0.8424 | 9.05 | 2009-09-10 01:54:47 |
| 17 | Netflix.co.uk/Team | 0.8424 | 9.04 | 2009-09-10 01:54:47 |
| 18 | Netflix.co.uk/Team | 0.8424 | 9.01 | 2009-09-10 01:54:47 |
| 19 | Clus | 0.8424 | 9.00 | 2009-09-10 01:54:47 |
| 20 | Netflix.co.uk/Team | 0.8424 | 9.00 | 2009-09-10 01:54:47 |

MOST POPULAR



GEAR

Why We Should All Be Using the Encrypted Chat App Signal

JORDAN MCMAHON



SECURITY

How to Keep Your Bitcoin Safe and Secure

LILY HAY NEWMAN



SPONSOR CONTENT

Digital transformation: How machine learning could help change business

SIEMENS

Combinación de Criterios de Recomendación

*¿Qué artículos se deben recomendar a un cliente con la intención de aumentar el tamaño de su ticket promedio y el Lifetime Value?
¿Cómo recompensar a los clientes?*



← → ↻ <https://www.enriquedans.com/2006/08/panales-y-cerveza.html> 🔍 📄 ⭐ 📧 📧 📧



Enrique Dans

EN ES



ARCHIVOS TEMAS CONTACTO ACERCA DE...

<https://www.enriquedans.com/2006/08/panales-y-cerveza.html>

20 agosto 2006

En beer, correlation, customer intelligence, diapers, market basket analysis, statistics

¿Pañales y cerveza?

Hasta aquí, todo muy bien. Es, como dice Seth Godin en su fantástico libro, una buena historia, una de esas que aparentemente tiene patas, que se sostiene, que gusta contar. ¿El problema? Muy sencillo: **que es mentira**. Lisa y llanamente falsa. Es una recolección de datos inconexos que comienza con el descubrimiento de una inofensiva correlación entre consumo de pañales y cerveza no adscrita a sexo ni a día de la semana alguno, que continúa con unos

🔍 Buscador por temas

Enrique Dans

Curriculum

Contacto

Mi libro: "Todo va a cambiar"

En Google+

En Facebook

Artículos antiguos

Enlaces en Pinterest

Mis fotos en Flickr

Hasta aquí, todo muy bien. Es, como dice Seth Godin en su fantástico libro, una buena historia, una de esas que aparentemente tiene patas, que se sostiene, que gusta contar. ¿El problema? Muy sencillo: que es mentira. Lisa y llanamente falsa. Es una recolección de datos inconexos que comienza con el descubrimiento de una inofensiva correlación entre consumo de pañales y cerveza no adscrita a sexo ni a día de la semana alguno, que continúa con unos cambios en el *store layout* que las grandes superficies nunca tomaron, y que prosigue con los supuestos buenos resultados que las grandes superficies presuntamente obtuvieron. En realidad, nada. Simplemente, una correlación inofensiva no sustentada que podría servir para ilustrar la diferencia entre la búsqueda orientada de hipótesis y la mera exploración de los datos, y que los vendedores de software de BI suelen usar con profusión para poner los dientes largos a los clientes con eso de "anda que si encuentras algo así, inesperado y rentable..." Pues a lo mejor lo encuentras, pero esperemos que si lo haces sea, al menos, más cierto que la historia de los pañales y la cerveza.

Para más explicaciones, este [artículo en The Register](#) y esta [anotación de Daniel J. Power](#), uno de los implicados en el comienzo de esta leyenda urbana.

teradata.

En la vida real...



| | col1_item1 | col1_item2 | cntb | cnt1 | cnt2 | score | support | confidence | lift |
|---|------------|-------------------|------|--------|--------|-------------|--------------|--------------|-----------|
| 1 | PAÑAL BEBE | CERVEZA IMPORTADA | 2514 | 570328 | 140709 | 7.875605... | 7.5620282... | 0.0044079... | 1.0414... |
| 2 | PAÑAL BEBE | CERVEZA NACIONAL | 8683 | 570328 | 611507 | 2.161789... | 2.6118174... | 0.0152245... | 0.8276... |

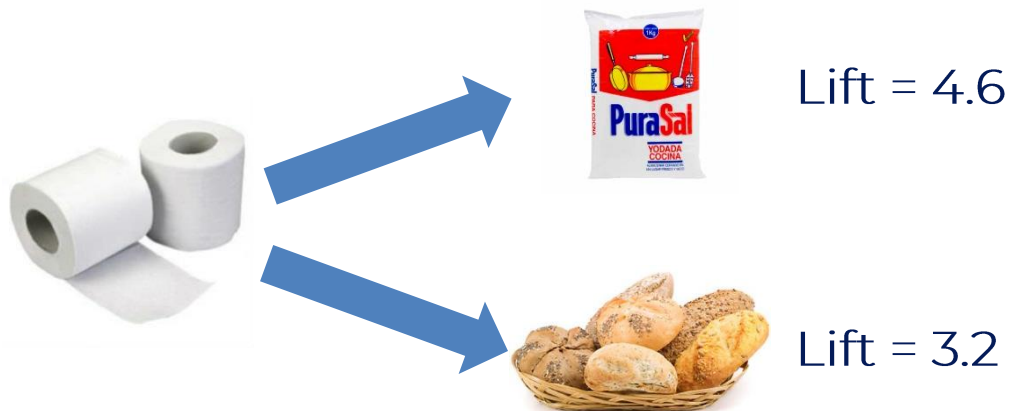
| | col1_item1 | col1_item2 | cntb | cnt1 | cnt2 | score | support | confidence | lift |
|---|-------------------|------------|------|--------|--------|------------|--------------|---------------|-----------|
| 1 | CERVEZA IMPORTADA | PAÑAL BEBE | 2514 | 140709 | 570328 | 7.87560... | 7.5620282... | 0.01786666... | 1.0414... |
| 2 | CERVEZA NACIONAL | PAÑAL BEBE | 8683 | 611507 | 570328 | 2.16178... | 2.6118174... | 0.01419934... | 0.8276... |

Afinidades Significativas



| col1_item1 | col1_item2 | cntb | cnt1 | cnt2 | score | support | confidence | lift |
|------------|-------------------|------|-------|-------|---------------|--------------|--------------|-----------|
| TINTO | VARIOS | 4231 | 77056 | 85624 | 0.00271321... | 0.0011195... | 0.0549081... | 2.4235... |
| TINTO | HIELO | 980 | 77056 | 19845 | 6.28050531... | 2.5930508... | 0.0127180... | 2.4220... |
| TINTO | ENCURTIDOS | 4526 | 77056 | 91718 | 0.00289846... | 0.0011975... | 0.0587365... | 2.4202... |
| TINTO | HONGOS | 3264 | 77056 | 68506 | 0.00201820... | 8.6364468... | 0.0423588... | 2.3368... |
| TINTO | SIN HUESO | 713 | 77056 | 15298 | 4.31258772... | 1.8865767... | 0.0092530... | 2.2859... |
| TINTO | ENJUAGUE BUCAL | 650 | 77056 | 14335 | 3.82492175... | 1.7198806... | 0.0084354... | 2.2239... |
| TINTO | DULCES | 2768 | 77056 | 61277 | 0.00162266... | 7.3240456... | 0.0359219... | 2.2155... |
| TINTO | ORGANIZADORES ... | 36 | 77056 | 799 | 2.10499835... | 9.5254928... | 4.6719269... | 2.2098... |
| TINTO | PAPELERIA | 2 | 77056 | 45 | 1.15356220... | 5.2919404... | 2.5955149... | 2.1798... |
| TINTO | AL HUEVO | 429 | 77056 | 9688 | 2.46532394... | 1.1351212... | 0.0055673... | 2.1718... |

| col1_item1 | col1_item2 | cntb | cnt1 | cnt2 | score | support | confidence | lift |
|------------|---------------------|------|-------|--------|---------------|--------------|--------------|-----------|
| TINTO | PISCO | 2606 | 77056 | 20340 | 0.00433302... | 6.8953984... | 0.0338195... | 6.2839... |
| TINTO | CERVEZA IMPORTA... | 2552 | 77056 | 30430 | 0.00277749... | 6.7525160... | 0.0331187... | 4.1132... |
| TINTO | BOSQUE | 2456 | 77056 | 68409 | 0.00114429... | 6.4985029... | 0.0318729... | 1.7608... |
| TINTO | ESPECIAL | 2410 | 77056 | 170646 | 4.41704182... | 6.3767882... | 0.0312759... | 0.6926... |
| TINTO | JABON | 2409 | 77056 | 81929 | 9.19241208... | 6.3741423... | 0.0312629... | 1.4421... |
| TINTO | ESPECIAS | 2292 | 77056 | 63350 | 0.00107615... | 6.0645637... | 0.0297446... | 1.7745... |
| TINTO | GALLETAS INTEGRA... | 2285 | 77056 | 88294 | 7.67422902... | 6.0460419... | 0.0296537... | 1.2692... |
| TINTO | COMIDAS | 2260 | 77056 | 203881 | 3.25112496... | 5.9798927... | 0.0293293... | 0.5436... |
| TINTO | PREPARADOS | 2245 | 77056 | 59614 | 0.00109718... | 5.9402031... | 0.0291346... | 1.8470... |



| col1_item1 | col1_item2 | cntb | cnt1 | cnt2 | score | support | confidence | lift |
|-----------------|--------------------|-------|-------|--------|-------------|---------------|----------------|-------------|
| PAPEL HIGIENICO | TALCO | 425 | 96338 | 3566 | 5.257737... | 1.12453735... | 0.004411550... | 4.675467... |
| PAPEL HIGIENICO | ACCESORIOS DE BAÑO | 328 | 96338 | 2786 | 4.008380... | 8.67878238... | 0.003404679... | 4.618597... |
| PAPEL HIGIENICO | SAL | 3540 | 96338 | 30089 | 0.004323... | 9.36673465... | 0.036745624... | 4.615438... |
| PAPEL HIGIENICO | VINAGRE | 2250 | 96338 | 19127 | 0.002747... | 5.95343304... | 0.023355269... | 4.614802... |
| PAPEL HIGIENICO | ATUN | 9362 | 96338 | 80181 | 0.011346... | 0.00247715... | 0.097178683... | 4.580517... |
| PAPEL HIGIENICO | CEPILLO DENTAL | 2070 | 96338 | 17833 | 0.002494... | 5.47715839... | 0.021486848... | 4.553688... |
| PAPEL HIGIENICO | CUIDADO CORPORAL | 676 | 96338 | 5833 | 8.132120... | 1.78867588... | 0.007016961... | 4.546447... |
| PAPEL HIGIENICO | ENVASADA | 17196 | 96338 | 151232 | 0.020296... | 0.00455001... | 0.178496543... | 4.460680... |
| PAPEL HIGIENICO | DESODORANTE HOMBRE | 3202 | 96338 | 28347 | 0.003754... | 8.47239670... | 0.033237144... | 4.431304... |

| col1_item1 | col1_item2 | cntb | cnt1 | cnt2 | score | support | confidence | lift |
|-----------------|---------------|-------|-------|--------|-------------|---------------|----------------|-------------|
| PAPEL HIGIENICO | PAN | 19236 | 96338 | 237901 | 0.016144... | 0.00508978... | 0.199671988... | 3.172020... |
| PAPEL HIGIENICO | TROZADO | 18799 | 96338 | 233405 | 0.015716... | 0.00497415... | 0.195135875... | 3.159672... |
| PAPEL HIGIENICO | JAMON COCIDOS | 17608 | 96338 | 313791 | 0.010256... | 0.00465902... | 0.182773152... | 2.201339... |
| PAPEL HIGIENICO | ENVASADA | 17196 | 96338 | 151232 | 0.020296... | 0.00455001... | 0.178496543... | 4.460680... |
| PAPEL HIGIENICO | SIN GAS | 16901 | 96338 | 324172 | 0.009146... | 0.00447195... | 0.175434408... | 2.045287... |
| PAPEL HIGIENICO | SEMI-DUROS | 16321 | 96338 | 263591 | 0.010489... | 0.00431848... | 0.169413938... | 2.429034... |
| PAPEL HIGIENICO | JABON | 16109 | 96338 | 81929 | 0.032877... | 0.00426239... | 0.167213352... | 7.713444... |
| PAPEL HIGIENICO | GRANEL | 15999 | 96338 | 205384 | 0.012936... | 0.00423328... | 0.166071539... | 3.055931... |
| PAPEL HIGIENICO | UHT | 15683 | 96338 | 171845 | 0.014856... | 0.00414967... | 0.162791421... | 3.580219... |

Accionamiento de los Insights



Nº de tarjeta [REDACTED]

10€ de DTO

en la sección de MASCOTAS
de tu hiper habitual

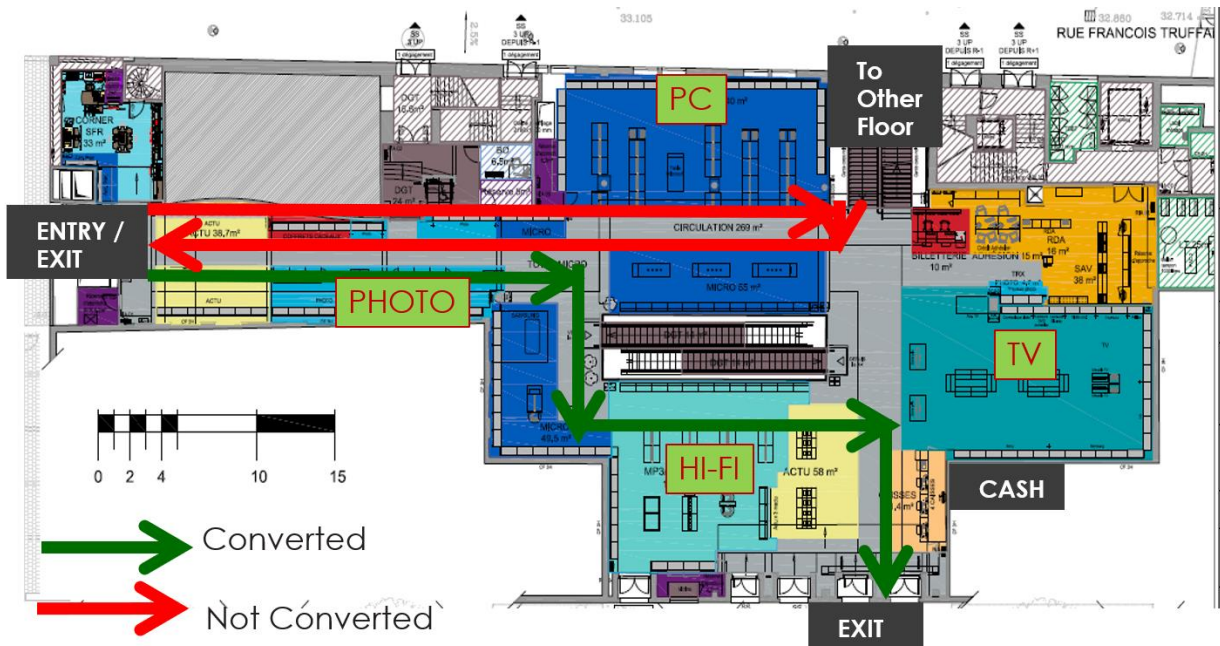
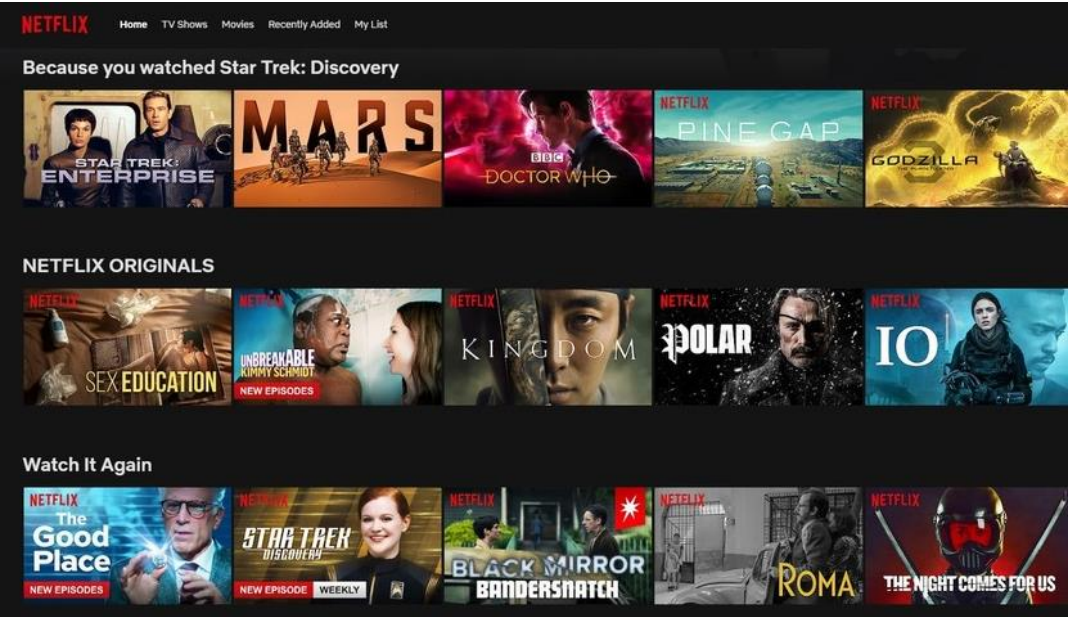
VÁLIDO EN HIPERMERCADOS CARREFOUR

Descuento válido hasta el 30-11-2011

UTILÍCELO EN SU PRÓXIMA COMPRA CON SU
TARJETA CLUB O PASS. NO SE PODRÁN
CANJEAR 2 CUPONES IGUALES EN UN MISMO
TICKET NI USAR EL MISMO CUPÓN MÁS DE UNA
VEZ. CONSULTAR BASES DE EL CLUB

03/09/11 13:12:48 . 0170 018 0049 8793

Accionamiento de los Insights



1 Best sellers in Office & School Supplies

Page 1 of 8

| | | | | |
|---|--|--|---|--|
| <p>Hammermill Paper, Copy Paper, 8.5 x 11 Paper,...</p> <p>2 ★★★★★ 1,757 \$27.18 ✓prime</p> | <p>Sharpie Permanent Markers, Fine Point,...</p> <p>3 ★★★★★ 307 \$21.99 ✓prime</p> | <p>BIC Round Stic Xtra Life Ballpoint Pen, Medium...</p> <p>4 ★★★★★ 177 \$13.71 ✓prime</p> | <p>Pendaflex File Folders, Letter Size, 8-1/2" x 11",...</p> <p>★★★★★ 516 \$9.98 ✓prime</p> | <p>AmazonBasics 92 Bright Multipurpose Copy...</p> <p>★★★★★ 1,927 \$33.98 ✓prime</p> |
|---|--|--|---|--|

Exclusive offers on your mobile

Download our **free** Sainsbury's app for bonus Nectar points, deals and a handy store locator on the go.

[Download Sainsbury's app now](#)

Key Takeaways

1

A la gente le gustan las buenas historias. Si cuentas una Buena historia, podría quedar en la mente de la comunidad por mucho tiempo.

2

“En Dios confiamos, los demás deben traer datos”. Una gran frase de Edwards Deming. Validemos la información disponible.

3

El éxito de un modelo no está en el nivel de acierto, sino en la capacidad de accionamiento. Los resultados son los que hacen la diferencia.

¡Muchas gracias!



#yomequedoencasa