



# **Análisis de precios de Capital Federal**

## *Informe TP2 - Data Mining*

Ing. Christian Jorge Marcusa

*Maestría en Data Mining, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires*  
Presentado el 6 de Agosto de 2019

---

### **Abstract**

La fluctuación de precios en los mercados porteños es un factor fundamental en la vida de los ciudadanos. Con variables y factores externos influyentes en ello, como la inflación, la oferta y demanda, el mercado cambiario, e incluso la elaboración nacional y extranjera de los mismos, se hace muy difícil encontrar un precio acorde a cada producto. En este desarrollo se intenta encontrar asociaciones entre las subas de los precios en los puntos de venta al consumidor.

*Keywords:* Data Mining, Precios, Asociacion, Reglas, Arules, KDD

### **Introducción**

Al hablar de precios, y su análisis, debemos siempre recurrir a datos objetivos e históricos de los mismos. Por lo tanto se ha buscado información confiable provista por el gobierno de la ciudad para realizar este análisis. Se ha planteado una estructura de datos sencilla y durante su desarrollo se irá simplificando aún más para análisis específicos.

### **Preparación de datos de precios**

Si bien el mayor análisis ya fue realizado en etapas previas, se opta por hacer una preparación de cero de los datos.

En primer lugar el dataframe de precios contiene muchos datos faltantes y está en formato vertical. Se ha decidido pasar a horizontal la colección de datos y agregar las distintas mediciones como columnas de cada subconjunto "producto - sucursal". Para ello se ha utilizado los paquetes dplyr y tidyr. (En particular las funciones select y spread de los mismos respectivamente).

	producto	sucursal	1	2	3	4	5	6	7	8	9	10
1	0000040084107	10-1-214	36.00	36.00	36.00	36.00	36.00	36.00	43.50	43.50	43.50	55.00
2	0000040084107	10-1-219	36.00	36.00	36.00	36.00	36.00	36.00	43.50	43.50	46.00	55.00
3	0000040084107	10-1-220	36.00	36.00	36.00	36.00	36.00	36.00	NA	43.50	43.50	55.00
4	0000040084107	10-1-26	36.00	36.00	36.00	36.00	36.00	NA	43.50	43.50	43.50	55.00
5	0000040084107	10-1-271	36.00	36.00	36.00	36.00	36.00	36.00	43.50	43.50	46.00	55.00
6	0000040084107	10-1-2	36.00	36.00	36.00	36.00	36.00	36.00	43.50	43.50	43.50	55.00

Luego para limpiar y discretizar las mediciones se han creado una serie de funciones que iteran sobre los datos. Para el desarrollo de las mismas se ha partido en un conjunto de test (Usando la función head) para agilizar su desarrollo, y luego ser aplicadas para el conjunto total (Debido al gran volumen de mediciones)

Las funciones desarrolladas fueron

- **completarFaltantes(dataframe)** : Recibe un dataframe y para cada valor NA dentro de sus mediciones, se le completa con la media entre sus adyacentes, o su adyacente contiguo si es una medición extrema (1 o 10), en el caso que sus adyacentes contengan NAs, ignora la operación.
- **eliminarFaltantes(dataframe)** : Utiliza la función drop\_na de tidyr para descartar los registros que no pudieron ser llenados por la función anterior.
- **agregarColumnasPeriodos(dataframe)** : Recibe un dataframe y para cada fila, separa las mediciones en 4 periodos (1 a 3, 4 a 5, 6 a 7 y 8 a 10), luego calcula el precioPromedio para todos los periodos.
- **agregarVariaciones(dataframe)** : Recibe un dataframe y para cada fila, calcula y agrega las variaciones para cada periodo. Es decir  $[p(x) - p(x-1)] / p(x-1)$ . Luego también calcula y agrega la variación total, del periodo 1 al 4.
- **discretizacionDeVariaciones(dataframe)** : Discretiza las variaciones según la siguiente tabla.

Categorías	Rango
Disminución Fuerte	$[-\infty; -0.05)$
Disminución Media	$[-0.05; -0.02)$
Disminución Leve	$[-0.02; -0.005)$
Mantiene	$[-0.005; 0.005)$
Aumento Leve	$[0.005; 0.05)$
Aumento Medio	$[0.05; 0.1)$
Aumento Fuerte	$[0.1; \infty]$

- **mediasPorProducto(dataframe)** : Recibe un dataframe y utilizando dplyr, calcula las medias de precios para cada producto (Agrupa por el mismo, dejando fuera las sucursales). Retorna un nuevo dataframe con las medias para cada periodo, y la media total del precio.
- **preciosRelativos(dataframe, promedios)** : Recibe un dataframe, y otro dataframe con las medias calculadas por la función antes explicada. Para cada elemento del primer dataframe, calcula su diferencia con las medias calculadas en el segundo.

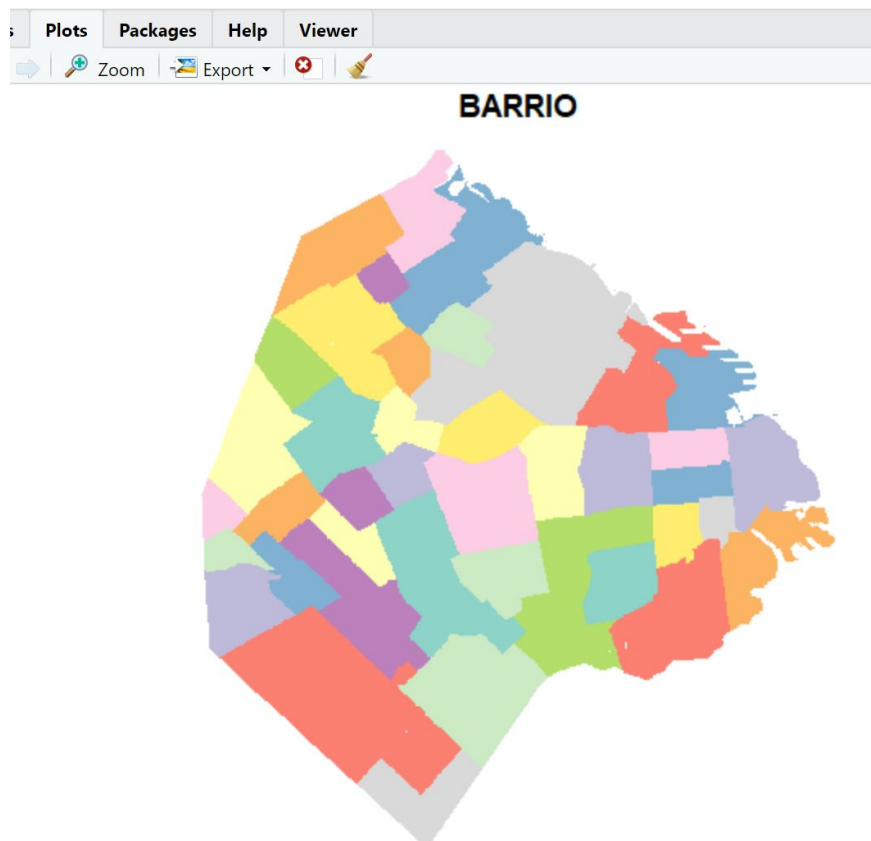
- **discretizacionesDePrecio(dataframe)** : Recibe un dataframe y discretiza las mediciones de la función explicada en el apartado anterior, según la tabla que se detalla a continuación.

Categorías	Rango
Muy caro	$[\infty; 0.1)$
Medio caro	$[0.1; 0.05)$
Levemente caro	$[0.05; 0.01)$
Medio	$[-0.01; 0.01)$
Levemente barato	$[-0.05; -0.01)$
Medianamente barato	$[-0.1; -0.05)$
Muy barato	$[-\infty; 0.1)$

Con este procedimiento ya se disponen los datos de precios para trabajar con las reglas, detallado en los siguientes apartados.

## Preparación de datos de sucursales

Para los datos de sucursales se han utilizado datos geográficos obtenidos de distintas APIs en la red. La primera que se ha incorporado es el dataset "[https://bitsandbricks.github.io/data/CABA\\_rc.geojson](https://bitsandbricks.github.io/data/CABA_rc.geojson)" el cual contiene objetos Simple Feature (sf a partir de ahora) para plotear información geográfica. El dataset contiene mucha información no pertinente a este análisis pero permite graficar las sucursales según los barrios.



A su vez también se necesitó obtener información pertinente a la geolocalización de las sucursales, ya que solo se dispone de la LATITUD y la LONGITUD.

Para esto se ha utilizado la API Open Cage Data (<https://opencagedata.com>), la cual expone una sencilla API HTTPS que con una Key (por login), una coordenada de latitud y una de longitud, brinda información pertinente a nuestro análisis.

El problema que se tiene es que parsear la respuesta desde R fue complicado porque está en formato JSON con mucha información no importante, de la cual solo interesaba lo expuesto en el siguiente snippet:

```
"components": {
  "ISO_3166-1_alpha-2": "AR",
  "ISO_3166-1_alpha-3": "ARG",
  "_type": "building",
  "city": "Buenos Aires",
  "continent": "South America",
  "country": "Argentina",
  "country_code": "ar",
  "house_number": "4877",
  "neighbourhood": "Parque Cornelio Saavedra",
  "postcode": "1603",
  "road": "Avenida Doctor Ricardo Balbu00edn",
  "state": "Ciudad Aut\u00f3noma de Buenos Aires",
  "state_code": "C",
  "suburb": "Saavedra"
}
```

Y más en particularmente el campo suburb (Barrio), ya que el resto de la información ya se dispone (Ciudad, país) o es información más específica no útil (como la dirección exacta) o mismo información redundante (como el campo “neighbourhood”).

Para extraer esta información se desarrolló un pequeño módulo en la tecnología NodeJS (Debido a su facilidad para consumir APIs y trabajar con JSON) el cual obtiene el dataset guardado desde R, lo parsea en JSON, consume las APIs y para cada elemento del dataset le agrega el barrio y lo vuelve a guardar, y esta información ya está disponible para ser usada por el código de R.

Luego el dataset de sucursales ya puede ser agrupado y filtrado por barrio.

	field1	id	lat	lng	barrio
1	46	3-1-3	-34.546397	-58.451884	Belgrano
2	95	15-1-2	-34.5785767	-58.4871575	Villa Urquiza
3	23	10-3-373	-34.543233	-58.463148	Núñez
4	6	10-3-300	-34.540539	-58.472052	Núñez
5	7	15-1-498	-34.5512734	-58.4871771	Saavedra

## Preparación de datos de productos

Los datos de productos tienen 3 campos de texto que brindan información redundante, los mismos son nombre, marca y descripción.

Se han utilizado los paquetes `tm`, `stringi`, `stringr` y `arules` para darle forma al dataset. En particular las funciones `removeNumbers`, `removePunctuation`, `stri_trans_general`, `stripWhitespace`, `unique` y la función `removeWords` con `stopwords` (la cual recibe un parametro `kind='sp'` y permite obtener las palabras no útiles al análisis, como las preposiciones y los artículos). Para guardar el dataset generado y no hacer de nuevo el procesamiento, se ha utilizado `write.csv`, y se ha guardado el dataset limpio, ejemplificado a continuación:

	id	nombre	presentacion	marca
1	0000040084107	huevo chocolate sorpresa	gr	kinder
2	0000075027513	desodorante original	gr	dove
3	0000075032715	cerveza rubia	cc	corona
4	0000077903518	galletitas obleas rellena	gr	opera
5	0000077940131	turrón	gr	arcor
6	0000077940704	desodorante crema pote odorono	gr	rexona
7	0000077953063	postre dulce leche pack	gr	danette

## Metodologia

Ya disponiendo de los dataset limpios, se procede a hacer el análisis correspondiente. En el caso de los productos, que ya se venía trabajando con `arules`, se generó un vocabulario de palabras utilizando una frecuencia mínima de 20 ocurrencias, con la función `findFreqTerms` de `tm`.

Lo mismo arrojó un vocabulario de 32 palabras importantes:

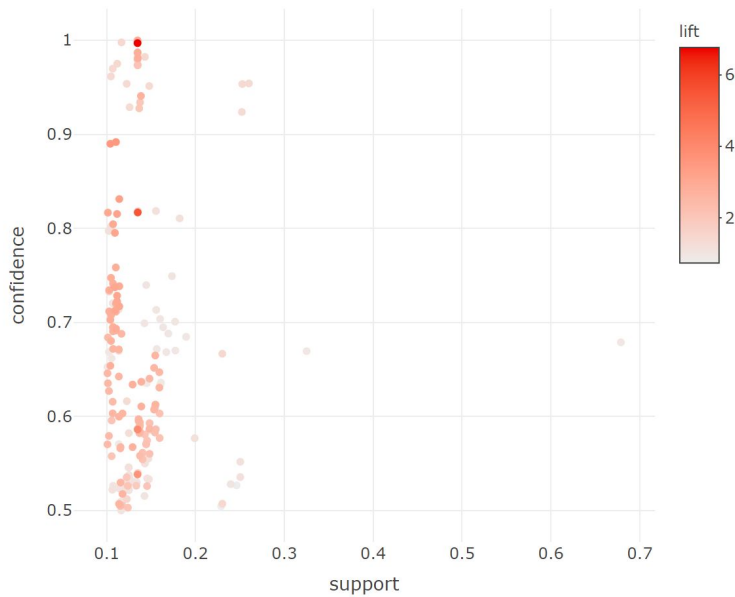
```
[1] "chocolate" "desodorante" "galletitas" "crema" "dulce" "leche"
[7] "pack" "aerosol" "blanco" "agua" "jabon" "liquido"
[13] "doypack" "limon" "cafe" "manzana" "polvo" "jugo"
[19] "naranja" "vainilla" "light" "frutilla" "queso" "yogur"
[25] "fideos" "mate" "vino" "tinto" "malbec" "gas"
[31] "saborizada" "gaseosa"
```

Una aclaración importante, es que si se aumenta la frecuencia del vocabulario a 30, el mismo se reduce a 15 elementos, y ya subiendolo a 100, da cero. Esto quiere decir que hay cero términos que estén en 100 productos simultáneamente. Se probó con varias opciones y se eligió el número 20 porque 32 palabras parece simbólico en un total de 1000 productos.

Luego utilizando la función `merge`, se obtuvo un dataset de trabajo general de 44 columnas, a saber: `variacionesDisc(periodo 1a2, periodo 2a3, periodo 3a4, total)`, `preciosRelativosDisc(periodo 1, periodo 2, periodo 3, periodo 4, total)`, `marca`, `presentación`, `barrio`, y el vocabulario de 32 palabras expuesto en el apartado anterior.

## Análisis descriptivo

Primera iteración, generación de reglas con `support=0.1` y `confidence=0.1`, arroja 359 reglas.



Se observa que la mayoría de las reglas están entre 0.1 y 0.2 de soporte, y las que están por encima de ello tienen un LIFT bajo.

Buscando reglas con mayor LIFT, se encuentran:

	All	All	All	All	2.0882	All
[188]	{variacion1=Mantiene,variacion2=Mantiene,variacion3=Mantiene}	{variacionTotalDiscreta=Mantiene}	0.135	0.997	6.764	21,247.000
[108]	{variacion1=Mantiene,variacion3=Mantiene}	{variacionTotalDiscreta=Mantiene}	0.135	0.817	5.543	21,283.000
[105]	{variacion2=Mantiene,variacion3=Mantiene}	{variacionTotalDiscreta=Mantiene}	0.135	0.586	3.977	21,268.000
[111]	{variacion1=Mantiene,variacion2=Mantiene}	{variacionTotalDiscreta=Mantiene}	0.135	0.538	3.653	21,251.000
[130]	{precio2=Levemente barato,precio4=Levemente barato}	{precioRelativoTotal=Levemente barato}	0.111	0.892	3.529	17,417.000
[115]	{precio1=Levemente barato,precio4=Levemente barato}	{precioRelativoTotal=Levemente barato}	0.104	0.890	3.522	16,426.000
[124]	{precio3=Levemente barato,precio4=Levemente barato}	{precioRelativoTotal=Levemente barato}	0.114	0.831	3.289	18,013.000
[121]	{precio2=Levemente barato,precio3=Levemente barato}	{precioRelativoTotal=Levemente barato}	0.112	0.815	3.227	17,618.000
[112]	{precio1=Levemente barato,precio2=Levemente barato}	{precioRelativoTotal=Levemente barato}	0.109	0.795	3.147	17,233.000
[113]	{precio1=Levemente barato,precioRelativoTotal=Levemente barato}	{precio2=Levemente barato}	0.109	0.737	3.131	17,233.000

Que demuestra lo que se supuso en el apartado anterior. Un producto no cambia su precio relativo en general (No deja de ser barato, o medianamente barato, etc, en estos periodos), y un producto cuyas variaciones se mantienen, mantiene su precio total (lo cual es trivial).

A esta altura, explorando el dataset, se decide extraer algunas palabras que generan reglas redundantes, por ejemplo “vino, tinto, malbec” podrían ser una sola, y se puede bajar el Support sin generar reglas redundantes.

**Se extraen las palabras *tinto, malbec, blanco, liquido, doypack, gas, polvo*.**

Sabiendo esto, se decide hilar fino sobre las reglas, considerando LIFT más bajos para encontrar asociaciones significativas:

[160] {presentacion= gr} => {variacion1=Aumento Fuerte} 0.1555100 0.3201782 1.2078743 24509

Los productos cuya presentación es en gramo, sufrieron un aumento fuerte entre el periodo 1 y el 2 con un 32% de confianza.

[162] {variacion2=Aumento Medio} => {variacionTotalDiscreta=Aumento Fuerte} 0.1555481 0.8185036 1.2057261 24515

Con un 81% de confianza, los productos que sufrieron un aumento medio en entre el periodo 2 y el 3, tuvieron un aumento fuerte en el total.

[164] {variacion1=Aumento Fuerte, variacionTotalDiscreta=Aumento Fuerte} => {presentacion= gr} 0.1479404 0.5853585 1.2051894 23316

Los productos que tuvieron un aumento fuerte entre el 1 y el 2, y en el total, son presentados en gramos. 58% confianza.

[165] {precio1=Medio} => {variacion1=Mantiene} 0.1401868 0.5603632 1.1984893 22094

Los productos con un precio relativo medio en el primer periodo, se mantuvieron así hasta el segundo. 56% confianza.

[168] {variacionTotalDiscreta=Aumento Fuerte} => {variacion3=Aumento Medio} 0.1822035 0.2684014 1.1941041 28716

Los productos que tuvieron un aumento fuerte en el total, en general tuvieron un aumento medio entre los periodos 3 y 4. 26%

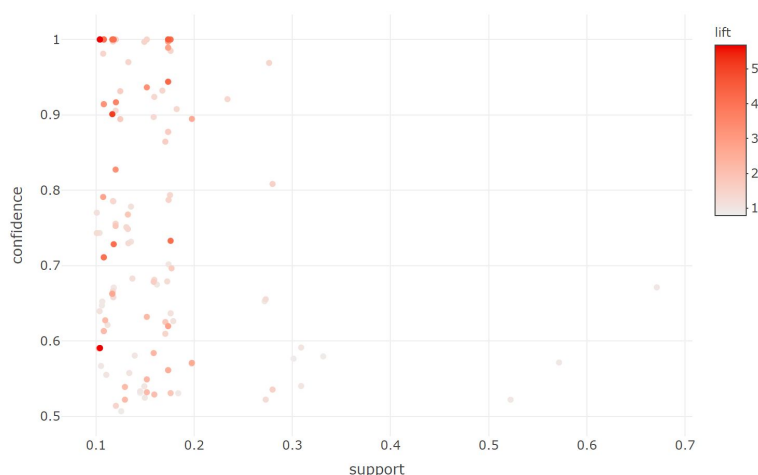
En este punto se sube la confianza a 0.5 para obtener reglas más fuertes. Se obtienen 188 reglas e hilando fino se pueden destacar algunas significativas.

[141] {presentacion= lt} => {variacionTotalDiscreta=Aumento Fuerte} 0.1023324 0.7972712 1.174449 16128

Los productos en litros, tuvieron un aumento fuerte general con un 79% de confianza.

En este punto se observó que debido al soporte, la presencia de las variaciones y precios discretizados es mucho mayor que los campos del vocabulario, y el barrio. Por lo tanto se decide usar el paquete dplyr para obtener subsets de la información y conseguir reglas con ellos.

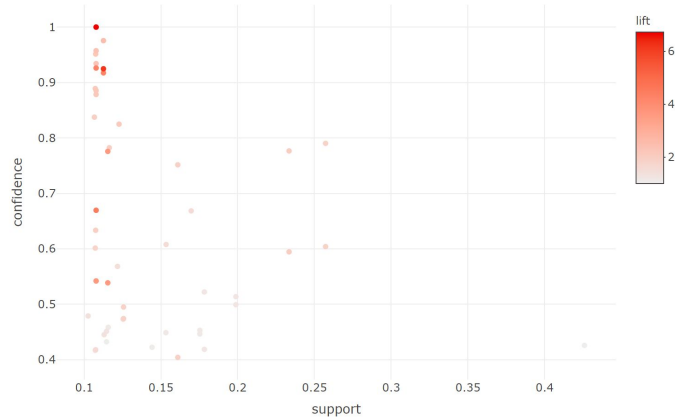
En primer lugar se trabaja con los productos muy caros. Usando filter y select, se extraen las columnas de precios relativos y se vuelven a generar reglas (132 reglas).



Luego se puede observar que hay algunos patrones significativos por LIFT alto, se detallan a continuación:

[63]	{variacion3=Aumento Fuerte, variacionTotalDiscreta=Aumento Fuerte}	{vino}	0.152	0.549	2.288	959.000
[22]	{vino}	{variacion3=Aumento Fuerte}	0.152	0.632	2.216	959.000
[23]	{variacion3=Aumento Fuerte}	{vino}	0.152	0.532	2.216	959.000
[7]	{presentacion= ml}	{variacion3=Aumento Fuerte}	0.109	0.628	2.200	691.000

Se hizo el mismo análisis para productos muy baratos. Con Supp=0.1 y Conf=0.4 se obtuvieron 85 reglas. Se tuvo que bajar la confianza a 0.4 porque las reglas con 0.5 no aportan información.



[48]	{variacion3=Aumento Fuerte, variacionTotalDiscreta=Aumento Fuerte}	{presentacion= gr}	0.107	0.417	1.221	762.000
[24]	{presentacion= gr}	{variacion1=Mantiene}	0.153	0.449	1.160	1,089.000
[30]	{variacion1=Mantiene}	{vino}	0.175	0.453	1.154	1,246.000
[31]	{vino}	{variacion1=Mantiene}	0.175	0.447	1.154	1,246.000
[11]	{presentacion= ml}	{variacion2=Mantiene}	0.116	0.458	1.151	822.000

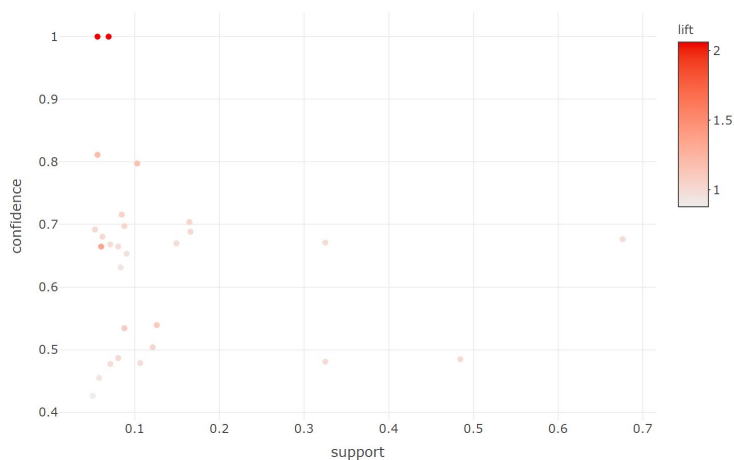
A diferencia de los productos caros, el vino se mantuvo mejor en los productos baratos.

Luego se opta **subdividir por barrios opuestos de la ciudad**, se toman Recoleta y Villa Lugano.

En recoleta se encuentran 144 reglas con 0.1 y 0.5.

Se decide refinar el dataset, no incluyendo la información interperiodo, para obtener reglas más específicas de los barrios. Se destacan las siguientes reglas:

En Recoleta:

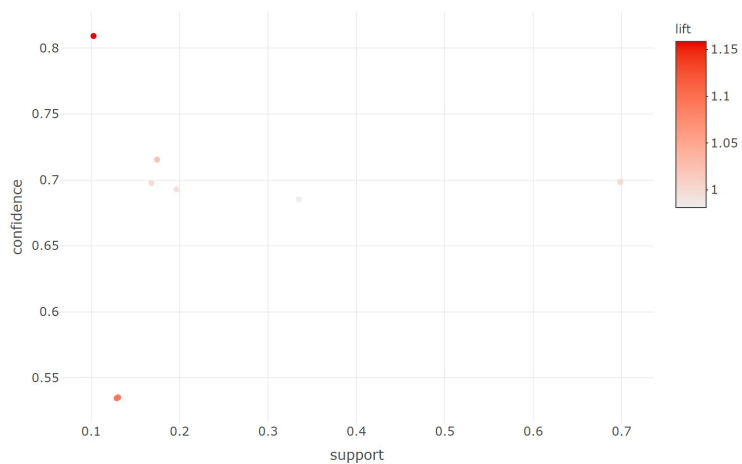




[5]	{galletitas}	{variacionTotalDiscreta=Aumento Fuerte}	0.056	0.811	1.199	821.000
[6]	{pack}	{presentacion= gr}	0.061	0.665	1.371	884.000
[7]	{pack}	{variacionTotalDiscreta=Aumento Fuerte}	0.062	0.680	1.006	905.000
[8]	{precioRelativoTotal=Medianamente caro}	{presentacion= gr}	0.051	0.426	0.879	738.000
[9]	{precioRelativoTotal=Medianamente caro}	{variacionTotalDiscreta=Aumento Fuerte}	0.085	0.716	1.058	1,239.000

Las reglas que más llaman la atención son que los packs, las galletitas, y los productos medianamente caros aumentaron fuertemente.

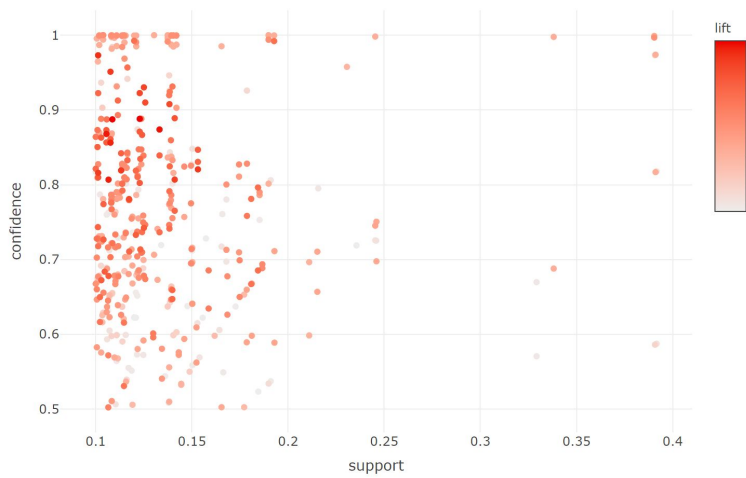
En Villa Lugano:



[2]	{presentacion= lt}	{variacionTotalDiscreta=Aumento Fuerte}	0.103	0.809	1.158	420.000
[3]	{precioRelativoTotal=Levemente barato}	{presentacion= gr}	0.131	0.535	1.095	534.000
[4]	{precioRelativoTotal=Levemente barato}	{variacionTotalDiscreta=Aumento Fuerte}	0.175	0.715	1.024	714.000
[5]	{precioRelativoTotal=Levemente caro}	{presentacion= gr}	0.129	0.534	1.094	527.000

Las reglas que más llaman la atención son que los productos en litro aumentaron fuertemente, y los productos levemente baratos también.

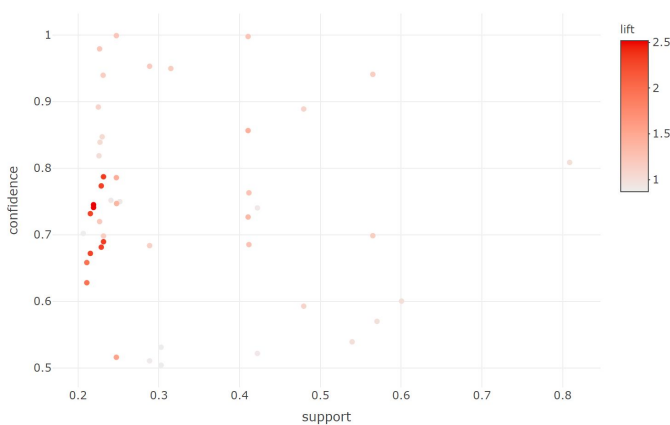
Para analizar la desaceleración de precios en el último periodo, se genera un set de reglas y se filtra por cuya variación en el último periodo sea “Mantiene” o cualquiera de las tres “disminuciones”. Se filtran por el RHS (Siendo uno de los valores del último periodo, mantiene, o disminución). Con un supp de 0.1 y una confianza de 0.5 se obtuvieron las siguientes reglas (filtradas según la utilidad del informe):



```
[15] {v1d=Aumento Fuerte,
      v2d=Mantiene} => {variacion3=Mantiene}    0.06658460 0.5179153 1.499945 10494
[16] {v1d=Aumento Fuerte,
      presentacion= gr} => {variacion3=Mantiene}    0.08487729 0.5457995 1.580701 13377
[32] {variacion1=Aumento Fuerte,
      variacion2=Mantiene,
      variacionTotalDiscreta=Aumento Fuerte} => {variacion=Mantiene}    0.06656557 0.5429281 1.572385 10491
[33] {v1d=Aumento Fuerte,
      variacionTotalDiscreta=Aumento Fuerte,
      presentacion= gr} => {variacion3=Mantiene}    0.08472501 0.5726969 1.658599 13353
```

Como se puede observar, hay reglas con confianza de más del 50% (Y un LIFT elevado) en las que se produjo un aumento fuerte, tanto en períodos individuales como en el total, y en el último periodo se mantuvo.

Trabajando sobre un **lote particular** de productos, se eligen las galletitas, y se generan reglas sobre ellas con un supp de 0.2 y una confianza de 0.5. Se eligen las reglas detalladas a continuación (De un total de 47):



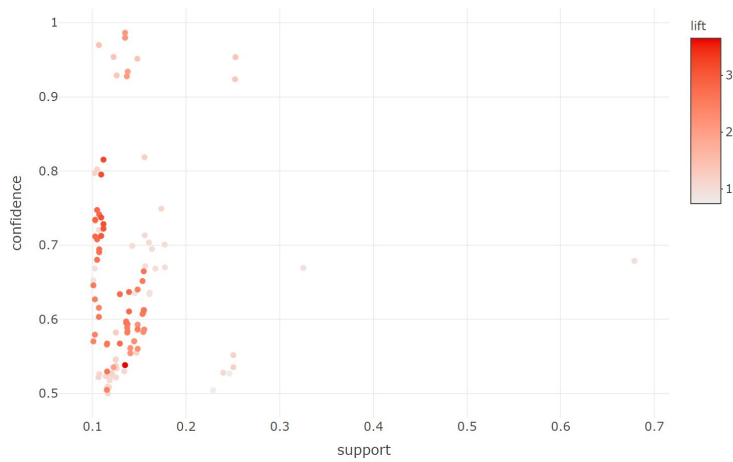
```
{variacion2=Aumento Fuerte,variacionTotalDiscreta=Aumento Fuerte} => {variacion3=Mantiene}    0.2474858 0.7857765
{variacion3=Mantiene,variacionTotalDiscreta=Aumento Fuerte} => {variacion1=Aumento Fuerte}    0.4105869 0.8564073
{variacion2=Aumento Fuerte} => {variacion3=Mantiene}    0.2477601 0.7469680
{variacion1=Aumento Fuerte,variacionTotalDiscreta=Aumento Fuerte} => {variacion3=Mantiene}    0.4105869 0.7266990
{variacion1=Aumento Fuerte} => {variacion3=Mantiene}    0.2477601 0.7469680
{precioTotalRelativo=Levemente barato} => {variacionTotalDiscreta=Aumento Fuerte}    0.2517828 0.7497958
```

Estas reglas marcan que el aumento fuerte de las galletitas (0.80), estuvo dado más que nada en la primer variación, y en general se mantuvo en la última

## Analisis predictivo

Para abordar este análisis se genera un conjunto de reglas utilizando el dataset sin las mediciones de variación y precio relativo totales, y sin el precio y la variación del último período, luego se eligen algunas reglas sobre la variación del segundo periodo, y se contrastan con las métricas del tercero.

Se generan reglas omitiendo el precio final y la ultima variacion:



Se seleccionan las reglas que tengan sentido respecto de los periodos, ordenadas por LIFT (Entre 2.5 y 1.2 de LIFT se utilizó):

[58]	{precio2=Medio}	{variacion1=Mantiene}	0.145	0.571	
[65]	{variacion1=Aumento Fuerte}	{presentacion= gr}	0.156	0.587	
[6]	{variacion2=Aumento Medio}	{variacionTotalDiscreta=Aumento Fuerte}	0.156	0.819	
[52]	{precio1=Medio}	{variacion1=Mantiene}	0.140	0.560	
[63]	{precioRelativoTotal=Medio}	{variacion1=Mantiene}	0.147	0.555	
[2]	{presentacion= lt}	{variacionTotalDiscreta=Aumento Fuerte}	0.102	0.797	
[38]	{precio1=Levemente caro}	{variacion2=Mantiene}	0.121	0.528	
[69]	{variacion1=Mantiene}	{variacion2=Mantiene}	0.250	0.536	
[107]	{variacion2=Mantiene, presentacion= gr}	{variacion1=Mantiene}	0.125	0.522	
[39]	{precio1=Levemente caro}	{variacion1=Mantiene}	0.125	0.546	

Luego se generan reglas que convaliden el análisis en el último periodo (Removiendo del análisis las columnas de los periodos anteriores, y volviendo a utilizar los valores del último periodo). El mismo arroja muy pocas reglas con LIFT lejano a 1, se eligen algunas:

{variacion3=Aumento Medio}	{variacionTotalDiscreta=Aumento Fuerte}	0.182	0.811
----------------------------	---	-------	-------

La primer regla que llama la atención. Aumentos fuertes con variación 3 en aumento medio. Es acorde a la regla que dice que variacion2 = Aumento Medio -> variacionTotal = Aumento Fuerte

*{presentacion= lt} {variacionTotalDiscreta=Aumento Fuerte} 0.102 0.797*

Esta regla está presente en los dos sets.

*{precio4=Levemente caro} {variacionTotalDiscreta=Aumento Fuerte} 0.145 0.740*

En el conjunto anterior se convalidaban reglas que los productos Levemente caros se mantenían, no obstante aquellos que en el periodo 4 resultaron ser levemente caros, tuvieron un fuerte aumento.

*{precio4=Levemente barato,precioRelativoTotal=Levemente barato} {variacionTotalDiscreta=Aumento Fuerte}*

Los productos levemente baratos (En general, y en el último periodo) resultaron tener un aumento fuerte. Regla que no se puede deducir del set anterior.

*{precioRelativoTotal=Levemente caro} {variacionTotalDiscreta=Aumento Fuerte} 0.160 0.704*

En el set anterior se observó que los productos levemente caros se mantuvieron en los primeros periodos, no obstante tomando el total, aumentaron más significativamente.

Se pueden validar algunas reglas, y otras no utilizando la información del último periodo.

## Conclusiones

Los resultados de este informe demuestran varias hipótesis sobre el comportamiento de los precios.

- Primero que los precios en general se desaceleraron en el último periodo.
- Luego que en los barrios más pudientes los productos caros sufrieron un aumento fuerte, mientras que en los demás no hay evidencia para afirmarlo.
- Las predicciones de los primeros periodos no se cumplieron en el último. Se observa una desaceleración, pero en los productos donde se predecía que se mantenían con una fuerte confianza, no se pudo sustentar en el último.
- Los productos baratos aumentaron fuertemente en su mayoría.
- Los productos que se venden por litro sufrieron aumentos considerables.

- Las galletitas sufrieron su mayor aumento en el primer periodo, y luego desaceleraron.
- Al igual que los datos arrojados por el TP1, se valida que el vino es de los grandes afectados por la suba, en especial el de mediana y alta gama (Medianamente caro +).

## Referencias

- Código fuente utilizado para este trabajo está disponible en GitHub, solicitar acceso. (<https://github.com/Chaitooler>).
- "[https://bitsandbricks.github.io/data/CABA\\_rc.geojson](https://bitsandbricks.github.io/data/CABA_rc.geojson)"
- <http://opencagedata.com>
- Dplyr-Documentation (<https://dplyr.tidyverse.org/>)
- R-Documentation (<https://www.rdocumentation.org/>)
- <https://dmuba.github.io/>
- MongoDB-Documentation (<https://docs.mongodb.com/>)