



Trabajo Practico Entregable 1

Data Mining

**Maestría en Explotación de Datos y Descubrimiento del
Conocimiento**

Facultad de Ciencias Exactas y Naturales

Departamento de Computación

Universidad de Buenos Aires

Juan Jose Iguarán Fernández

3 de Junio de 2019

Tabla de Contenido

Introducción.....	3
Obtención de datos.....	3
Análisis exploratorio.....	3
Integración de datos.....	4
Reducción de dimensionalidad.....	4
Tratamiento de faltantes.....	5
Valores extremos.....	5
Análisis.....	6
¿Cuanto cuesta en promedio cada elemento de la canasta básica para cada una de las empresas?.....	6
¿Cual es el precio promedio del total de los productos de la canasta básica para cada una de las empresas?.....	7
¿Cual es el precio medio de los productos de la canasta para las distintas empresas para cada medición?.....	8
¿Cual es la variación entre mediciones del precio promedio de los productos de la canasta por empresa?.....	9
¿Cual es la variación acumulada del precio promedio de los productos de la canasta por empresa?.....	9
.....	10
¿Que tan cerca esta la variación acumulada total del promedio de los productos de la canasta para cada empresa, a la reportada para los bienes en la ciudad de Buenos Aires?.....	10
Conclusiones.....	11
Bibliografía.....	11

Introducción

En la ciudad de Buenos Aires existe una plataforma llamada precios claros, donde se reportan los precios de distintos productos de consumo básico en distintos comercios con el objetivo de brindar información útil para la comparación de precios a los ciudadanos. En general el presente informe tiene como propósito utilizar la información disponible en dicha plataforma para emplear técnicas propias del Data Mining, en cuanto a la generación de conocimiento se refiere. Específicamente, se emplearan técnicas de preprocesamiento y exploración para brindar información previamente desconocida.

Obtención de datos

A través de técnicas de web crawling sobre la plataforma, se realizaron 10 consultas entre noviembre de 2018 hasta finales de febrero de 2019 teniendo cada medición las siguientes fechas: 1:05/11/2018, 2:12/11/2018, 3:26/11/2018, 4:10/12/2018, 5:17/12/2018, 6:31/12/2018, 7:21/01/2019, 8:04/02/2019, 9:11/02/2019, 10:25/02/2019. Estas consultas arrojaron como resultado información sobre tres entidades: los precios, los locales y los productos registrados en ella. Esta información se presentó en forma de tres archivos JSON Lines en tres archivos distintos, uno para cada entidad, cada uno de estos con un conjunto determinados de atributos para cada entrada. El archivo productos contiene información sobre los 1000 productos mas frecuentes; el archivo sucursales, contiene información sobre las 837 sucursales registradas en la plataforma. El archivo precios, contiene los precios para los distintos productos en cada sucursal en cada una de las 10 mediciones realizadas.

Análisis exploratorio

Para poder realizar análisis con los datos disponibles, se hace necesario que estén presentes en un formato adecuado, por ello, la primera etapa consistió en la conversión de estos archivos JSON Lines en CSV. Una vez realizado dicho proceso, se puede apreciar mas claramente la información contenida en cada uno de los archivos mediante el software Rstudio. Allí, se puede apreciar las observaciones y las variables de cada uno de los set de datos. El set sucursales tiene 837 observaciones de los siguientes atributos :X_id(identificación de la consulta), sucursalTipo(clasificación del tipo de sucursal en cuestión), dirección(dirección de la sucursal), provincia(provincia en la que se ubica la sucursal), banderaId(numero que identifica la bandera), localidad(localidad en donde se ubica la sucursal), banderaDescripcion(descripción de la bandera asignada a la sucursal), lat(latitud de la sucursal), comercioRazonSocial(razón social de la empresa propietaria de la sucursal), lng(longitud en donde se ubica la sucursal), sucursalNombre(nombre dado a la sucursal), comercioId(id característico de la empresa propietaria de la sucursal), sucursalId(id propio de la sucursal), id(id que reúne el id de la sucursal, el comercio y la empresa). El set de productos tiene 1000 observaciones de los siguientes atributos: X_id(identificación de la

consulta), nombre(nombre del producto), marca(marca productora del producto), presentación(presentación en la que viene el producto), id(id del producto en cuestión). El set de precios tiene 1584661 de los siguientes atributos: X_id(identificación de la consulta), producto(id del producto en cuestión), sucursal(id de la sucursal de donde se tomo el precio), precio(precio registrado), fecha(fecha en la que se tomo la medición), medición(numero que identifica la medición realizada). Con esta información se puede proceder con las siguientes etapas del análisis.

Integración de datos

En términos de practicidad, resulta mas sencillo realizar los análisis basados en un solo dataset que en tres diferentes, por lo que se resulta coherente avanzar hacia la integración de estos archivos. Para hacerlo es necesario identificar los atributos comunes entre los distintos datasets. Se observa que el dataset precios posee el atributo productos que contiene la id del producto; esta es la misma información contenida en el atributo id del dataset productos, por consiguiente para generar el dataset completo, es menester enlazar ambas conjuntos a través de ese atributo. De igual manera, el atributo sucursal del dataset precios, contiene el atributo sucursal cuyo valor es la id de la sucursal del dataset sucursales, así que lo congruente es proceder a la integración de ambos mediante este atributo. El resultado de esta integración es un dataset con 1585323 observaciones de 23 atributos que contiene toda la información de los datasets originales, es con estos datos que resultara eficiente realizar los análisis.

Al efectuar la integración se observa que existen 662 observaciones con toda la información de productos como datos faltantes. Esto obedece a que existen 662 sucursales para las cuales no se registro información de precios, entre estas, todas las sucursales cuyo tipo de sucursal es autoservicio, así como 5 correspondientes a supermercado y debido a esto, todo análisis realizado sobre estas sucursales es enteramente ficticio, por lo que directamente se procede a su eliminación dejando en total 1584661 observaciones.

Reducción de dimensionalidad

Puesto que el objetivo del presente informe consiste en analizar información relacionada con la variación de los precios de los productos a lo largo del tiempo, existen algunos atributos que no realizan ningún aporte para el análisis y que por consiguiente, serán removidos del dataset. Estos serán el X_id originales de los datasets, la marca de los productos, la presentación, la dirección, la provincia, la id de la bandera, la localidad, la descripción de la bandera, la latitud, la longitud, el nombre de la sucursal, el id del comercio y el id de la sucursal, quedando así un dataset de 1584661 con los atributos sucursal, id del producto, nombre del producto, tipo de sucursal, razón social del comercio, fecha, medición y precio. Para practicidad del tratamiento, la variable original fecha es reemplazada por las fechas asociadas a cada medición. Este dataset presenta un formato mas compacto de la información necesaria.

Tratamiento de faltantes

Debido a que los datos disponibles provienen del resultado de un web crawling, existe la posibilidad de que para determinada medición, la información no se encontrara disponible en los servidores de la plataforma, por consiguiente pueden ocurrir que no se cuente con el registro de algún producto para determinada medición en determinada sucursal. Aun cuando esta situación ocurra, inicialmente no es observable en el dataset generado, pues en el solo estarán presentes las coincidencias producto de la integración de los datasets previos, y al no existir el registro, no se genera el dato faltante. Debido a esto, se hace necesario la introducción de nuevas observaciones para estas mediciones no presentes para aquellos productos que aparecen en otras mediciones en determinada sucursal. Para ello, se genera un dataset en donde aparecen las distintas combinaciones productos-sucursal presentes en el dataset. Una vez obtenida esta lista se replica 10 veces pues es el numero total de mediciones para cada combinación sucursal-producto; el resultado es un dataset de 1673110 observaciones para estas dos variables, el cual se ordena primero por producto y luego por sucursal para facilidad de integración. Así mismo se crea otro data set que contenga los números del 1 al 10 indicando las distintas mediciones, así como la fecha asociada a cada una. Este dataset se replica 167311 veces con el objetivo de obtener la misma cantidad de observaciones que el dataset anterior. Listos estos dataset se unen, y a el dataset resultante se le aplica un `left_Join` para unirlo con el dataset original. El resultado es un dataset de 1673110 con los 8 atributos originales en donde existen 87787 datos faltantes en el atributo precio pues es aquí en donde no existe información disponible como producto del `left_Join`, y estos son precisamente los faltantes deseados.

Una vez generados estos faltantes, es necesario utilizar un método de imputación apropiado para estos datos. Puesto que estos precios están relacionados con la medición, es decir unos valores Y asociados a una variable X resulta lógico pensar en una interpolacion como método adecuado. Esta interpolacion se efectuara con los precios Y de determinado producto para determinada sucursal, en las mediciones X , asignado de esta manera para el precio faltante en determinada medición, el resultado de la interpolacion tomando como X la medición en cuestión . Sin embargo, existen productos que solo aparecen una vez en determinada sucursal; para estos productos, puesto que no es posible una interpolacion, simplemente se asignara este mismo valor para todas las mediciones. De esta manera se completan todos los valores faltantes de los precios obteniendo asi un dataset completo para realizar el análisis.

Valores extremos

El dataset esta compuesto por 1000 productos de distinta naturaleza y por consiguiente, de una variación de precios considerables, por lo que una comparación directa entre los precios de estos no arroja información concluyente sobre un comportamiento atípico de su valor. Así mismo, debido a la naturaleza cambiante de los precios de los productos a lo largo del tiempo, que es en suma la variable objeto de análisis, no resulta prudente clasificar un dato determinado como atípico gracias a su comparación con otros datos del mismo tipo para distintas mediciones, por lo que se opta por no realizar ninguna modificación a los valores de los precios y proseguir con estos inmediatamente a su análisis.

Análisis

Con los datos listos para obtener conocimiento previamente desapercibido, es pertinente cuestionarse que tipo de conocimiento resulta interesante adquirir. Puesto que en la medida en que determinado conocimiento sea de mayor utilidad para un mayor publico resulta mas valioso, es conveniente enfocar la exploración teniendo en cuenta este principio; es por ello que se opta por enfocar el análisis a lo relacionado con el comportamiento de la canasta básica.

Un primer paso consiste en conocer la composición de dicha canasta, para ello, se utiliza información proveniente del INDEC de donde se obtiene que esta está compuesta por los siguientes productos: Pan, Galletitas de agua, Galletitas dulces , Arroz, Harina de trigo, Otras harinas (maíz), Fideos, Papa, Batata, Azúcar, Dulces, Hortalizas, Frutas, Carnes, Menudencias, Fiambres, Huevos, Leche, Queso, Yogur, Manteca, Aceite, Bebidas no alcohólicas, Bebidas alcohólicas, Sal fina, Condimentos, Vinagre, Café y Yerba.

Con el objetivo de conocer cuales de estos productos se encuentran disponibles en el dataset, se procede realizar una búsqueda en la variable que contiene el nombre de los productos por coincidencia de palabras claves, luego de la cual se obtiene que los productos disponibles son los siguientes 21:

Pan, Galletitas, Arroz, Harina de Trigo, Fideos, Azucar, Dulce, Tomate, Leche, Yogur, Manteca, Aceite, Gaseosa, Jugo, Cerveza, Vino, Sal, Ketchup, Mayonesa, Café, Yerba. Puesto que esta información sera de utilidad para el análisis de los precios, se opta por almacenarla.

Así como resulta conveniente conocer los productos de la canasta disponible, es así mismo relevante conocer las distintas empresas registradas en el dataset; para ello se busca los distintos valores contenidos en la variable razón social del comercio, obteniendo así los siguientes nombres: INC S.A., Coto Centro Integral de Comercialización S.A., DIA Argentina S.A, Josimar S.A., Jumbo Retail Argentina S.A., Wal Mart Argentina S.R.L. Esta información es así mismo almacenada.

Ahora que se cuenta con la información sobre los productos de la canasta disponibles, así como de las empresas listadas, es posible extraer nuevo conocimiento al respecto. Dicho conocimiento puede organizarse a través de la formulación y respuesta de distintas preguntas de interés relacionadas con la variación de la canasta familiar a lo largo del tiempo, por lo que en adelante se procede con dicha tarea.

¿Cuanto cuesta en promedio cada elemento de la canasta básica para cada una de las empresas?

Esta información es de resulta de sumo interés para los consumidores y por eso es propio responder la pregunta. Para ello, es inicialmente necesario obtener cuanto cuestan en promedio los productos cuyo nombre contiene cada una de las palabras clave previamente registradas, para cada una de las mediciones, para cada una de las empresas. Esta operación da como resultado un dataset de 210 observaciones pues son 21 palabras claves para 10 observaciones, con 9 atributos, 6 correspondientes a cada una de las empresas, 1 para identificar la medición, otro para identificar la palabra clave y otro en donde se registra la media de los precios para todas las empresas. A partir de

este dataset, se procede a calcular el promedio de todas las mediciones para cada uno de los productos para cada una de las empresas arrojando la siguiente tabla como resultado:

item	media	INC S.A.	Coto Centro Integral de Comercialización S.A.	DIA Argentina S.A	Josimar S.A.	Jumbo Retail Argentina S.A.	Wal Mart Argentina S.R.L.
Pan	50.3	48.6	50.3	49.2	53.8	52.2	50.1
Galletitas	47.6	45.3	47.7	47.4	49.4	49.6	46.8
Arroz	63.2	62.8	62.5	61.7	59.3	65	61.5
Harina de Trigo	33.5	33.5	33.3	33.3	30.7	34	31.1
Fideos	53.5	51.2	53.6	52	49.8	56.1	49.8
Azucar	60.3	58.5	58.3	66.5	49	64	60.8
Dulce	58.8	57.5	58.4	55.5	54.6	61.8	55.5
Tomate	25.8	25.3	25.6	26.4	26.1	26.3	24.7
Leche	51.6	51.2	51.8	48.8	48.3	52.4	51.8
Yogur	45	44	45.1	43.1	42.9	46.3	44.3
Manteca	75.3	75.2	76	71.4	73.4	75.6	71.3
Aceite	122.6	123.4	123.6	103.2	102.5	126.1	118.6
Gaseosa	60.4	58.6	61.4	56.2	61.5	61.7	60.7
Jugo	21.8	21.2	22.1	20.8	23.7	22	24.4
Cerveza	86.2	86.1	87.5	69.4	80.3	88.5	84.4
Vino	136.4	118.6	143.9	114.4	125.8	149.2	128.7
Sal	56.6	52.1	52.7	66.5	47.3	65.5	47.8
Ketchup	53.6	54.5	53.7	42.5	58.5	54.6	51.2
Mayonesa	56.2	53.8	56.6	60	58.1	57.1	52.2
Cafe	201.8	202	202.7	155.7	179	207.8	178
Yerba	94.2	92.9	94.4	91.3	87.3	96	92.8

Figure 1: costo promedio de los productos de la canasta por empresa

A simple vista puede apreciarse que la empresa con el catalogo mas económico es DÍA, el mas costoso es JUMBO, y aquella que ofrece los precios medios es COTO, aun cuando algunos de los productos no obedezcan a esta tendencia. Para tener una idea mas clara obre esta tendencia cabe formularse la siguiente pregunta:

¿Cual es el precio promedio del total de los productos de la canasta básica para cada una de las empresas?

Puesto que ya se cuenta con la suficiente información encapsulada sobre productos y empresas provenientes de la pregunta anterior, para responder esta, solo basta con promediar los precios obtenidos arrojando como resultado la siguiente tabla:

media	INC S.A.	Coto Centro Integral de Comercialización S.A.	DIA Argentina S.A	Josimar S.A.	Jumbo Retail Argentina S.A.	Wal Mart Argentina S.R.L.
69.3	67.4	69.6	63.6	64.8	72	66

Figure 2: Media total para los producto de la canasta básica por empresa

Se puede confirmar la apreciación hecha previamente: DÍA es la empresa mas económica, JUMBO la mas costosa y COTO la promedio. Además se evidencia que los productos de la canasta básica tienen precios que están alrededor de los 69 pesos.

Ahora que se tiene como referencia el precio medio de los productos de la canasta básica, es un buen punto de partida para indagar sobre la variación que presentan los mismos a lo largo de las mediciones; por consiguiente, cabe realizar la siguiente pregunta:

¿Cual es el precio medio de los productos de la canasta para las distintas empresas para cada medición?

Esta información puede obtenerse a partir del dataset que contiene los precios por medición por empresa por producto, pues a partir de este puede calcularse el promedio para todos los productos en cada medición en cada empresa, generando así un dataset con 10 observaciones, uno para cada medición, y 8 atributos: las 6 empresas, el numero de la medición y la media, tal como se aprecia a continuación:

medicion	media	INC S.A.	Coto Centro Integral de Comercialización S.A.	DIA Argentina S.A	Josimar S.A.	Jumbo Retail Argentina S.A.	Wal Mart Argentina S.R.L.
1	66.2	64.7	66.6	60.3	62.3	68.5	65.7
2	66.6	65.2	66.8	61.6	62.7	69	62.4
3	67	65.1	67.4	61.8	62.4	69.5	63.6
4	68	65.7	68.3	62.7	63	71.3	64.6
5	68.6	66.7	69.1	63.1	63.6	71.2	65.6
6	69.4	68.3	69.5	63.8	65.3	71.9	66.7
7	70.7	68.7	71.2	65	67	73.2	67.7
8	71.9	70.3	72	65.6	67.3	74.7	67.7
9	71.9	69.9	72.2	65.7	67.2	74.9	67.9
10	72.3	69.8	72.7	66.1	67.4	75.7	68.2

Figure 3: Precios del promedio de los productos de la canasta por medición por empresa

Es evidente como el precio promedio va cambiando de manera ascendente a lo largo de las mediciones y esta variación es considerable, sin embargo, no es apreciable cuan grande es el cambio entre una medición y otra con respecto al punto de partida, es por eso que resulta pertinente plantearse la siguiente pregunta:

¿Cual es la variación entre mediciones del precio promedio de los productos de la canasta por empresa?

Una vez respondida la pregunta anterior, para poder apreciar esta información se hace necesario realizar el calculo de la variación porcentual entre cada precio de canasta y su dato anterior, obteniendo así la variación. Esta operación se ejecuta para cada uno de los valores en cada una de las columnas de la tabla anterior arrojando como resultado la siguiente tabla en donde la columna

salto quiere decir como su nombre lo indica el salto entre una medición y otra, así el salto 1 es el salto dado entre la medición 1 y la medición 2 y así sucesivamente:

salto	media	INC S.A.	Coto Centro Integral de Comercialización S.A.	DIA Argentina S.A	Josimar S.A.	Jumbo Retail Argentina S.A.	Wal Mart Argentina S.R.L.
1	0.6	0.77	0.3	2.16	0.64	0.73	-5.02
2	0.6	-0.15	0.9	0.32	-0.48	0.72	1.92
3	1.49	0.92	1.34	1.46	0.96	2.59	1.57
4	0.88	1.52	1.17	0.64	0.95	-0.14	1.55
5	1.17	2.4	0.58	1.11	2.67	0.98	1.68
6	1.87	0.59	2.45	1.88	2.6	1.81	1.5
7	1.7	2.33	1.12	0.92	0.45	2.05	0
8	0	-0.57	0.28	0.15	-0.15	0.27	0.3
9	0.56	-0.14	0.69	0.61	0.3	1.07	0.44

Figure 4: Variaciones de los precios medios de los ítems de la canasta por empresa

De esta manera se puede evidenciar como han cambiado los promedios de precios para cada una de las empresas. En primera instancia es apreciable como el salto 6 en promedio, es decir entre la medición 6 y la 7 tuvo el mayor incremento. Cabe resaltar que este salto tuvo lugar entre las fechas 31 de diciembre de 2018 y 21 de enero de 2019, lo que da entender que el mayor aumento de precios se dio en esas fechas. Así mismo, el salto 8 que se dio entre las mediciones 8 y 9 , es decir, entre las fechas 4 de febrero y 11 de febrero de 2019 fue el que reporto la menor variación, dando a entender el menor aumento de precios.

Con esta información es posible profundizar un poco mas en el análisis estableciendo la siguiente pregunta:

¿Cual es la variación acumulada del precio promedio de los productos de la canasta por empresa?

Con la información obtenida en la pregunta anterior, es posible responder esta pregunta a partir de la suma del datos de variación junto con su dato anterior para cada una de las columnas. Esta operación da como resultado el siguiente dataset:

salto	media	INC S.A.	Coto Centro Integral de Comercialización S.A.	DIA Argentina S.A	Josimar S.A.	Jumbo Retail Argentina S.A.	Wal Mart Argentina S.R.L.
1	0.6	0.77	0.3	2.16	0.64	0.73	-5.02
2	1.2	0.62	1.2	2.48	0.16	1.45	-3.1
3	2.69	1.54	2.54	3.94	1.12	4.04	-1.53
4	3.57	3.06	3.71	4.58	2.07	3.9	0.02
5	4.74	5.46	4.29	5.69	4.74	4.88	1.7
6	6.61	6.05	6.74	7.57	7.34	6.69	3.2
7	8.31	8.38	7.86	8.49	7.79	8.74	3.2
8	8.31	7.81	8.14	8.64	7.64	9.01	3.5
9	8.87	7.67	8.83	9.25	7.94	10.08	3.94

Figure 5: Variación acumulada de los precios de los productos de la canasta por empresa

Como puede apreciarse en la tabla, la empresa con mayor variación acumulada fue JUMBO, la de menor fue WAL MART y la que se encuentra mas cerca de la media es COTO. Llama la atención que al igual que para los precios, tanto JUMBO como COTO obedecen a la misma tendencia: COTO es el mas cercano a la media y JUMBO es el mas alejado por arriba. Podría pensarse que el resultado anterior es una consecuencia de este, es decir, COTO posee los precios mas cercanos a la media pues el incremento que hace en cada salto es el mas cercano a la misma; lo mismo aplica para JUMBO pero en cuanto se refiere a lo mas alejado por arriba.

Ahora que se conoce la información de la variación acumulada cabe hacerse una ultima pregunta:

¿Que tan cerca esta la variación acumulada total del promedio de los productos de la canasta para cada empresa, a la reportada para los bienes en la ciudad de Buenos Aires?

En el sitio web de estadísticas de la ciudad se encuentra disponible información sobre la evolución mensual del costo de los bienes en la ciudad, de allí se obtiene que de noviembre a febrero, que es el periodo de estudio del presente análisis, la variación acumulada fue del 7%. A partir de esta información y de la obtenida en la pregunta previamente resuelta se puede proceder a calcular el porcentaje que representa con respecto a este dato, la variación acumulada de cada una de las columnas, arrojando como resultado lo siguiente:

media	INC S.A.	Coto Centro Integral de Comercialización S.A.	DIA Argentina S.A	Josimar S.A.	Jumbo Retail Argentina S.A.	Wal Mart Argentina S.R.L.
126.8	109.5	126.1	132.1	113.6	144	56.2

Figure 6: Porcentaje de la variacion acumulada real de bienes de la ciudad que representa la variación acumulada de cada empresa

Se puede observar que la empresa cuya variación mas cercana fue INC S.A., es decir Carrefour. Por otro lado el mas alejado por arriba fue JUMBO y el mas alejado por abajo fue WAL MART, al igual que se observo con la pregunta anterior, por lo que puede concluirse que el comportamiento del

incremento de los precios de los productos de la canasta de Carrefour es la referencia mas cercana al comportamiento real en la ciudad.

Conclusiones

- La integración de datos no siempre revela los datos faltantes, por lo que técnicas adicionales de generación pueden ser necesarias.
- La interpolación puede ser un método efectivo para imputar datos seriales, sin embargo, solo es posible cuando se tiene al menos dos puntos.
- En ocasiones, la presencia de valores extremos puede revelar información por lo que su remoción debe ser considerada dependiendo del caso
- La empresa con el catalogo mas económico es DÍA, el mas costoso es JUMBO, y aquella que ofrece los precios medios es COTO, aun cuando algunos de los productos no obedezcan a esta tendencia, teniendo como precio medio de los productos de la canasta 69 pesos.
- Es claro el incremento de precios a lo largo del tiempo, lo que denota una perdida de poder adquisitivo de la moneda, es decir inflación.
- Entre las fechas 31 de diciembre de 2018 y 21 de enero de 2019 se presento el incremento mas grande de precios, lo que podría indicar un incremento del consumo y por consiguiente mayor perdida de poder adquisitivo. Así mismo entre las fechas 4 de febrero y 11 de febrero de 2019 se presento menor variación, dando a entender un bajo consumo.
- La empresa con mayor variación acumulada fue JUMBO, la de menor fue WAL MART y la que se encuentra mas cerca de la media es COTO. Al igual que para los precios, COTO es el mas cercano a la media y JUMBO es el mas alejado por arriba. Podría concluirse que COTO posee los precios mas cercanos a la media pues el incremento que hace en cada salto es el mas cercano a la misma; lo mismo aplica para JUMBO pero en cuanto se refiere a lo mas alejado por arriba.
- La empresa cuya variación es mas cercana al incremento real de precios en los bienes en la ciudad de Buenos Aires fue Carrefour. La mas alejada por arriba fue JUMBO y el mas alejado por abajo fue WAL MART, por lo que puede concluirse que el comportamiento del incremento de los precios de los productos de la canasta de Carrefour es la referencia mas cercana al comportamiento real en la ciudad.

Bibliografía

1. <https://www.preciosclaros.gob.ar/>
2. https://www.indec.gob.ar/uploads/informesdeprensa/canasta_11_18.pdf
3. <https://www.estadisticaciudad.gob.ar/eyc/?p=27386>