

# Entrega Final



## Data Analytics

*Fabrizio Sgro*

*Facundo Fradeграда*

*Facundo Pereyra*

[ PROFESOR ] Edgardo Osorio

[ TUTOR ] - Joaquín Pablo Luna ;

## Contenido

Historial de versiones .....	3
Introducción.....	3
Descripción .....	4
Alcance .....	4
Hipótesis .....	5
Datasets .....	5
Diagrama de entidad-relación .....	9
Listado de tablas .....	10
Listado de columnas por tablas.....	10
Modelo relacional en Power BI.....	12
Segmentaciones elegidas .....	12
Tooltip .....	15
Medidas calculadas.....	16
Conclusiones.....	17
Líneas futuras .....	17
Bibliografía .....	18

## Historial de versiones

Versión	Fecha
Versión 1.0	17/08/22
Versión 1.1	04/09/22
Versión 1.2	13/10/22
Versión 1.3	31/10/22

## Introducción

En el presente documento trabajaremos con un «dataset» que incluye *“los precios mayoristas de frutas y hortalizas de diferentes mercados concentradores de América del Sur. Los valores están expresados en dólares estadounidenses por kilogramo de producto”*. (Agropecuarios, 2022).

Seleccionamos esta serie de datos porque se encontraban bien ordenados y nos permitiría una mejor interpretación y visualización de la información. Por ser una temática de datos, estrictamente agrícola, puede llegar a tener relevancia para empresas e industrias del rubro, sin embargo, por estar rodeado de un entorno académico, esta base de datos es suficiente para entender mejor la labor de un analista de datos y el proceso completo de obtención, transformación y la carga de datos.

Nuestra hoja de cálculo incluye los precios, productos y productores que reciben los mercados de Brasil, Argentina, Chile y Paraguay. La misma fue relacionada para poder crear un tablero de control con métricas y gráficos relevantes.

## Descripción

La información detallada que se encontrará en el documento es sobre el total de productos y su precio mayorista por mercado centralizador de cada país en los años 2017 y 2018. Es así, que encontrará el precio mayorista para más de 60 frutas y hortalizas que se comercializan en la región. Cobra relevancia este tipo de información para conocer el rendimiento de los mercados, los que manejan mayor cantidad de producto y a un precio menor.

## Alcance

El tablero de control que hemos realizado con la herramienta de visualización de datos, Power BI, está pensado para aquellos encargados de monitorear y tomar decisiones con respecto al Mercado concentrador de frutas y hortalizas de los países Argentina, Brasil, Chile y Uruguay. Es decir, a nivel estratégico, por los altos mandos. El mismo analiza la variación de precios de frutas y hortalizas por cada país, el precio promedio por mercado y cantidad total de productos según el tipo de abastecimiento y la comparativa de producción propia e importación por país. Para ello se utilizó medidas como Cantidad de Productos, Promedio del precio mayorista y Cantidad de países. Por lo tanto, le permitiría al espectador, entender que mercado posee mayor cantidad de productos, en que mercado se ofrecen los productos a mayor y menor costo. También, podrá ver claramente que países producen mayor cantidad de productos y el detalle de los mismos, como así también, los países que obtienen mayor cantidad de productos por importación y producción propia.

## Hipótesis

En un primer momento, teorizamos que el precio mayorista en dólares de las frutas y hortalizas de los mercados centrales de Argentina, Brasil, Chile y Paraguay variaban de igual manera entre el periodo 2017 y 2018. No nos sorprendió ver que estábamos totalmente equivocados, ya que si existe una variación en al menos dos países. Por otro lado, también teníamos la teoría de que el país con mayor cantidad de productos con precios en dólares más bajos es el que tiene mayor producción nacional y por lo tanto menos importación. Esta hipótesis fue descartada ya que el país con el menor promedio de precio era justamente el que más importaba productos (Paraguay).

## Datasets

Nuestro dataset contenía en un principio los valores de los mercados de Argentina, Brasil, Chile y Paraguay. Por este motivo tuvimos que reestructurar y “limpiar” el libro de cálculo.

El resultado que nos quedó fue una tabla llamada mercado con la información de todos los mercados mencionados.

	A	B	C	D	E	F	G	H	I	J
1	ID_MERCADO	Mercado	País_Id	Año	Mes	Producto	ID_PRODUCTOR	Precio Mayorista	Moneda	ID_ABASTECIMIENTO
2	301	Mercado Central de Buenos Aires	32	2017	Enero	Acelga	101	0,45	USD	1000
3	301	Mercado Central de Buenos Aires	32	2017	Febrero	Acelga	101	0,65	USD	1000
4	301	Mercado Central de Buenos Aires	32	2017	Marzo	Acelga	101	0,83	USD	1000
5	301	Mercado Central de Buenos Aires	32	2017	Abril	Acelga	101	0,61	USD	1000
6	301	Mercado Central de Buenos Aires	32	2017	Mayo	Acelga	101	0,38	USD	1000
7	301	Mercado Central de Buenos Aires	32	2017	Junio	Acelga	101	0,23	USD	1000
8	301	Mercado Central de Buenos Aires	32	2017	Julio	Acelga	101	0,46	USD	1000
9	301	Mercado Central de Buenos Aires	32	2017	Agosto	Acelga	101	0,42	USD	1000
10	301	Mercado Central de Buenos Aires	32	2017	Septiembre	Acelga	101	0,34	USD	1000
11	301	Mercado Central de Buenos Aires	32	2017	Octubre	Acelga	101	0,46	USD	1000
12	301	Mercado Central de Buenos Aires	32	2017	Noviembre	Acelga	101	0,41	USD	1000
13	301	Mercado Central de Buenos Aires	32	2017	Diciembre	Acelga	101	0,45	USD	1000
14	301	Mercado Central de Buenos Aires	32	2018	Enero	Acelga	101	0,32	USD	1000
15	301	Mercado Central de Buenos Aires	32	2018	Febrero	Acelga	101	0,34	USD	1000
16	301	Mercado Central de Buenos Aires	32	2018	Marzo	Acelga	101	0,30	USD	1000
17	301	Mercado Central de Buenos Aires	32	2018	Abril	Acelga	101	0,31	USD	1000
18	301	Mercado Central de Buenos Aires	32	2018	Mayo	Acelga	101	0,40	USD	1000
19	301	Mercado Central de Buenos Aires	32	2018	Junio	Acelga	101	0,60	USD	1000
20	301	Mercado Central de Buenos Aires	32	2018	Julio	Acelga	101	0,67	USD	1000
21	301	Mercado Central de Buenos Aires	32	2018	Agosto	Acelga	101	0,47	USD	1000
22	301	Mercado Central de Buenos Aires	32	2018	Septiembre	Acelga	101	0,32	USD	1000
23	301	Mercado Central de Buenos Aires	32	2018	Octubre	Acelga	101	0,34	USD	1000
24	301	Mercado Central de Buenos Aires	32	2018	Noviembre	Acelga	101	0,30	USD	1000
25	301	Mercado Central de Buenos Aires	32	2018	Diciembre	Acelga	101	0,31	USD	1000
26	301	Mercado Central de Buenos Aires	32	2017	Enero	Ajo	101	3,48	USD	1000
27	301	Mercado Central de Buenos Aires	32	2017	Febrero	Ajo	101	3,54	USD	1000
28	301	Mercado Central de Buenos Aires	32	2017	Marzo	Ajo	101	3,35	USD	1000
29	301	Mercado Central de Buenos Aires	32	2017	Abril	Ajo	101	3,33	USD	1000
30	301	Mercado Central de Buenos Aires	32	2017	Mayo	Ajo	101	3,34	USD	1000
31	301	Mercado Central de Buenos Aires	32	2017	Junio	Ajo	101	3,30	USD	1000

Otra tabla de país\_productor que contiene el ID de cada país productor y nombre.

	A	B	C	D	E
1	ID_PRODUCTOR	PRODUCTOR			
2	101	ARGENTINA			
3	102	BRASIL			
4	103	CHILE			
5	104	PARAGUAY			
6	105	ECUADOR			
7	106	IMPORTADO			
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					

Otra tabla denominada País, que tiene el ID e información sobre los países que contienen dicho mercado.

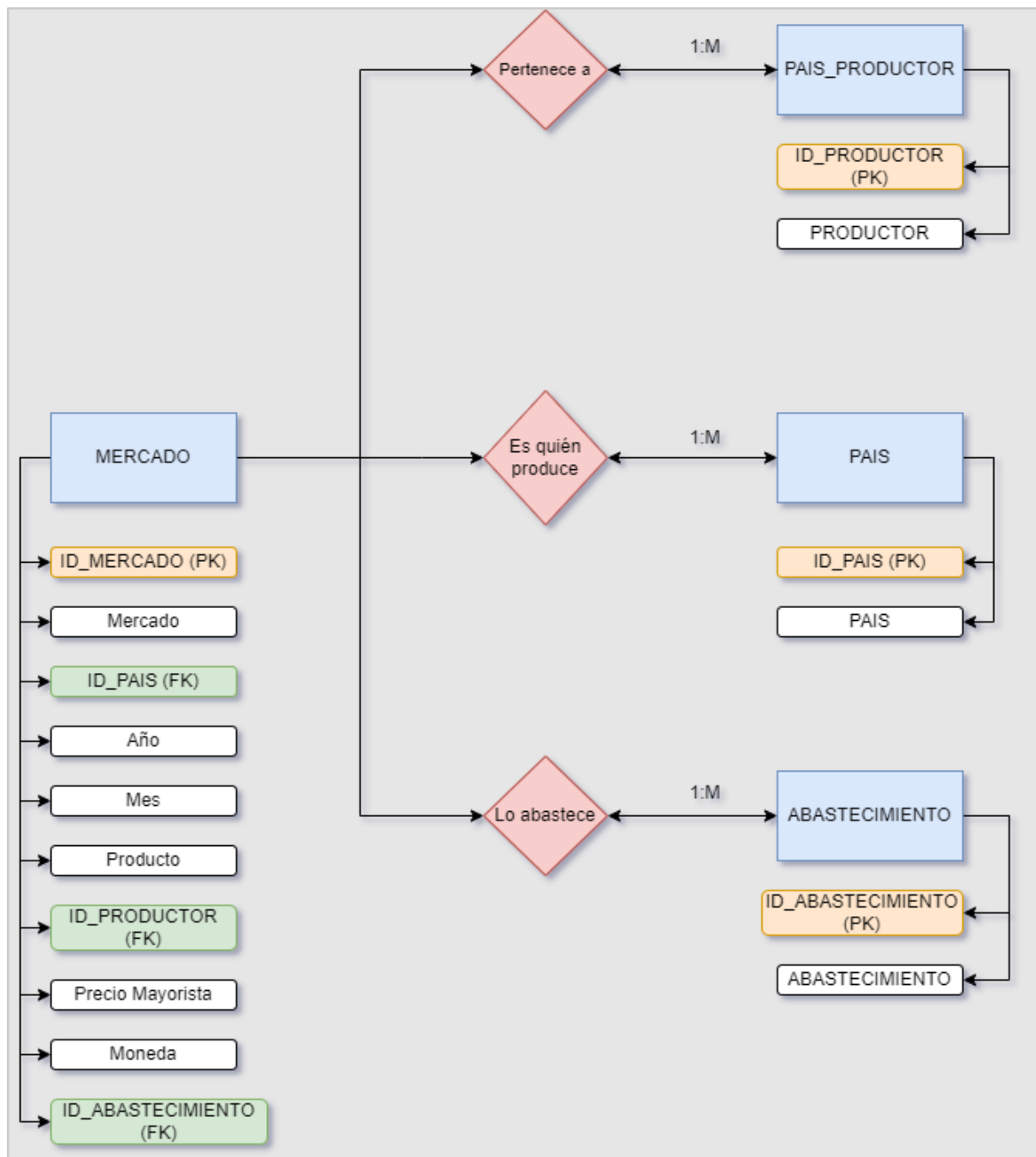
	A	B	C	D	E	F
1	ID_PAIS	PAIS				
2	32	ARGENTINA				
3	76	BRASIL				
4	208	CHILE				
5	221	PARAGUAY				
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						
30						
31						
MERCADO PAIS PAIS_PRODUCTOR ABASTECIMIENTO						

Por último, la tabla abastecimiento, que incluye ID y tipo de abastecimiento (propia, importado). Ya que al relacionarla podemos ver claramente el detalle de cada mercado y producto.

	A	B	C	D	E
1	ID_ABASTECIMIENTO	ABASTECIMIENTO			
2	1000	PRODUCCION PROPIA			
3	2000	IMPORTADO			
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
	MERCADO	PAIS	PAIS_PRODUCTOR	ABASTECIMIENTO	



## Diagrama de entidad-relación



## Listado de tablas

→ **MERCADO:** Contiene la fecha, el precio, tipo de producto de los principales mercados de Brasil, Chile, Paraguay, Ecuador y Argentina. También aclara si dichos mercados obtienen cada producto de manera propia o lo importan de otro país.

◆ **PK:** ID\_MERCADO

◆ **FK:** ID\_PAIS

◆ **FK:** ID\_PRODUCTOR

◆ **FK:** ID\_ABASTECIMIENTO

→ **PAIS:** Contiene la información de cada país y su id.

◆ **PK:** ID\_PAIS

→ **PAIS PRODUCTOR:** Contiene la descripción de cada país productor y su id.

◆ **PK:** ID\_PRODUCTOR

→ **ABASTECIMIENTO:** Contiene la información del tipo de abastecimiento, si es propio o importado. Con su propio identificador.

◆ **FK:** ID\_ABASTECIMIENTO

## Listado de columnas por tablas

<b>MERCADO</b>		
<b>CAMPO</b>	<b>TIPO DE CAMPO</b>	<b>TIPO DE CLAVE</b>
<b>ID_MERCADO</b>	int	<b>PK</b>

Mercado	varchar	-
ID_PAIS	int	FK
Año	int	-
Mes	varchar	-
Producto	varchar	-
ID_PRODUCTOR	int	FK
Precio Mayorista	decimal (5,2)	-
Moneda	varchar	-
ID_ABASTECIMIENTO	int	FK

**PAIS**

CAMPO	TIPO DE CAMPO	TIPO DE CLAVE
ID_PAIS	int	PK
PAIS	varchar	-

**PAIS\_PRODUTOR**

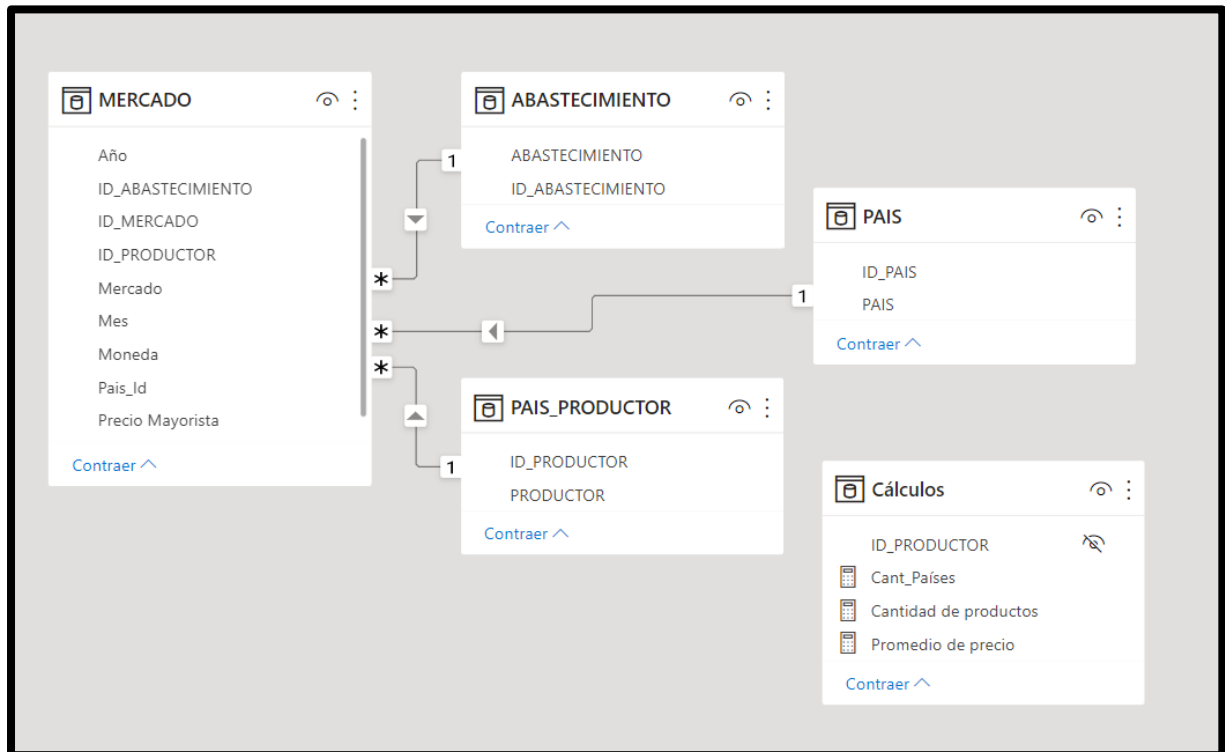
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE
ID_PRODUTOR	int	PK
PRODUTOR	varchar	-

**ABASTECIMIENTO**

CAMPO	TIPO DE CAMPO	TIPO DE CLAVE
ID_ABASTECIMIENTO	int	PK
ABASTECIMIENTO	varchar	-

## Modelo relacional en Power BI

El modelo relacional está conformado por las tablas de Mercado Abastecimiento, País y País Productor. Para las medidas se creó una tabla denominada Cálculos para tener más organizadas las medidas.



Para tener datos homogéneos para el análisis de productos cambiamos el nombre de “palta” por “aguacate”, este cambio se debió a que todos los países, menos Argentina, utilizaba el término aguacate en vez de palta.

## Segmentaciones elegidas

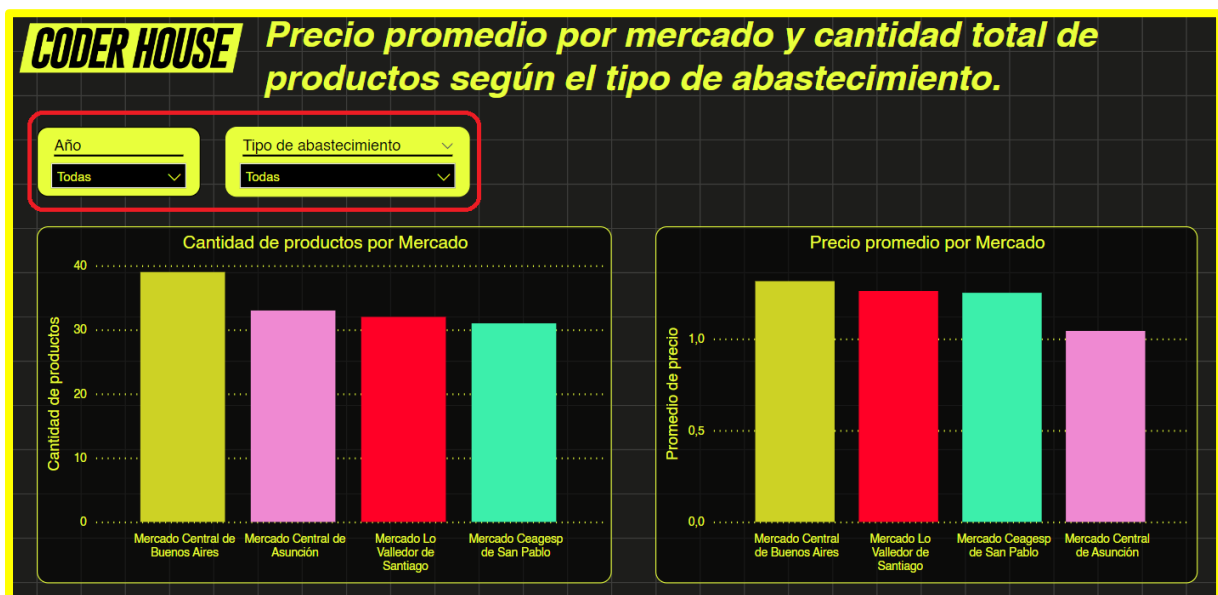
Las segmentaciones que utilizamos en el informe fueron seleccionadas para un análisis integral pero además tener la posibilidad de tener una vista reducida según la segmentación que el analista necesita.

Segmentación en solapa: Variación de precio de frutas y hortalizas por país.



Estas segmentaciones nos permiten analizar el promedio de precio mayorista según los productos, el año y los países que seleccionemos.

Segmentación en solapa: Precio promedio por mercado y cantidad total de productos según el tipo de abastecimiento.



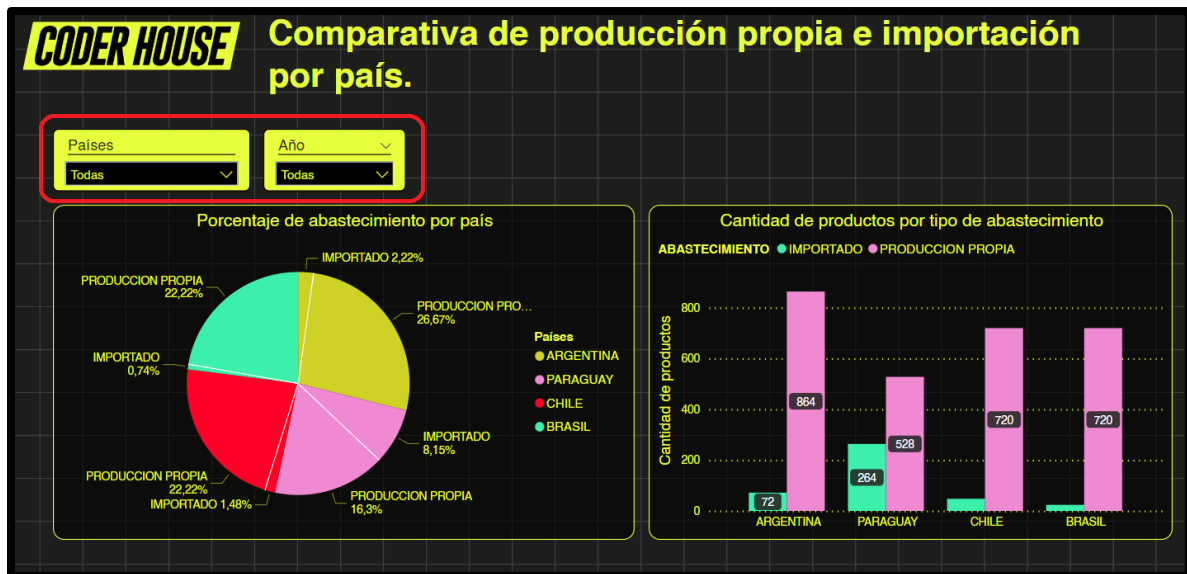
Estas segmentaciones nos permiten analizar el promedio de precio promedio y la cantidad de productos por mercado, pudiendo seleccionar el año y el tipo de abastecimiento según la necesidad del consumidor del informe.

**Segmentación en solapa: TOP 10 de productos (más y menos costosos) por país.**



Estas segmentaciones nos permites ver el top 10 de productos por cada país y cada año. Además, se utilizó la herramienta de marcadores para poder analizar los productos más costosos y los menos costosos superponiendo los dos análisis.

**Segmentación en solapa: Comparativa de producción propia e importación por país.**



Estas segmentaciones nos permiten analizar el porcentaje de abastecimiento por país y la cantidad de productos por tipo de abastecimiento según el año y el país o países que elegimos.

## Tooltip

Para tener un poco más de información se procedió a generar un Tooltip para la solapa “Panorama general”, más precisamente en el mapa, donde se marcan los países donde se encuentran los mercados a analizar. Gracias al Tooltip podemos visualizar la cantidad de productos que hay en ese mercado y su abastecimiento, si es producción propia o bien es importado.



## Medidas calculadas

1. **Cant\_Países** = **DISTINCTCOUNT(PAIS[PAIS])**: Esta medida nos permite tener la cantidad de países que tiene nuestra base de datos. Estos países son donde se localizan los mercados de frutas y hortalizas a analizar.
2. **Cantidad de productos** = **distinctcount (MERCADO[Producto])**: En esta medida tenemos la cantidad de productos diferentes que comercializa cada uno de los mercados concentradores.
3. **Promedio de precio** = **average (MERCADO [Precio Mayorista])**: En esta medida calculamos el promedio de precios mayoristas que tienen los mercados.
4. **TOP 10** = **CALCULATE ([Promedio de precio], TOPN (10, Mercado))**: En esta medida calculamos el top 10 de productos más costosos y menos costosos de cada mercado.



## Conclusiones

Una de las conclusiones que podemos hacer mediante el tablero de control creado, es que el mercado con mayor precio promedio es el que mayor producción propia de productos tiene. Este descubrimiento nos genera un interrogatorio sobre cuál puede ser el motivo. Uno de los caminos de solución que creemos posibles es que debido a que dicho mercado (Argentina) es que el posee la mayor cantidad de productos, aumenta las probabilidades de que el precio promedio total también aumente. También, gracias al tablero de control, podemos ofrecer una solución al tercero interesado en saber qué tipo de fruta y hortaliza es producida en Latinoamérica por alguno de los países de este data set, y quien tiene los menores costos en los productos buscados. Esto podría beneficiar a las empresas y gobiernos a elegir a los proveedores con menores costos.

## Líneas futuras

Este análisis se basa en los mercados concentradores de frutas y hortalizas de Argentina, Brasil, Chile y Paraguay en los años 2017 y 2018. Sería interesante seguir el análisis durante los siguientes años, considerando que en el 2020 se produjo la pandemia por el Coronavirus que afectó a todo el mundo y los países de Latinoamérica no fueron la excepción. Con esto se podría ver si cambió el promedio de precios de cada país, volviéndose más baratos o más caros con respecto a los demás países, como también la importación de productos de cada uno, si aumentó o disminuyó, volviéndose más dependientes o autónomos en la provisión de productos respectivamente.

## Bibliografía

Agropecuarios, D. d.-S. (12 de octubre de 2022). *Datos.gob.ar* . Obtenido de

Datos.gob.ar: <https://datos.gob.ar/dataset/agroindustria-frutas-hortalizas---precios-mayoristas>.