

PROYECTO FINAL DE DATA ANALYTICS



CAFETERÍA
**CHEZ
NOUS**
COFFEE ROASTERS

CoderHouse

Docente

Luciano Julián Gómez

Alumno

Federico Varveri





índice

Introduccion

Modificacion Tablas

Descripcion Tematica

Medidas Calculadas

Hipótesis

Dimensiones

Base de Datos

Análisis Funcional
del Tablero

Diagrama E-R

Visualizacion de
Datos

Tablas

Conclusion

Power BI -
Transformacion de Datos

Futuras Lineas

Columnas Calculadas





Introducción

Informe detallado para la identificación de potenciales clientes de nuestro negocio y medidas a tomar para maximizar ganancias de los ya consumidores de nuestros productos.





Alcance y Temática de los Datos

A partir de la información recopilada en forma de encuesta en nuestras distintas cafeterías de especialidad "Chez Nous", la temática a analizar serán los datos biométricos y el tipo de consumo de las personas, con el fin de poder identificar potenciales clientes y/o maximizar las ventas de los ya a activos. Con el fin de tener una base de datos completa y extensa en el tiempo se recopilaron los datos durante 2 años, comenzando el 1º de Enero del 2018 hasta el 31 de Diciembre del 2019. Con esto en mente, nuestro objetivo de este trabajo es reconocer patrones de conducta de los consumidores analizando distintas variables: Género, edad, tipo de empleo, posición económica, consumo de Café semanal, compra de productos del local, entre otras. Para ubicarnos en el tipo de análisis, este informe será realizado desde un diagnóstico, el cual definirá ciertas conclusiones de la forma de actuar en base a los datos históricos arrojados en la base de datos.

Hipótesis

Comprender si la compra de café y productos relacionados está directamente ligado a la posición económica y datos biográficos de la población de la ciudad de Chicago.

Comprobar también si es conveniente la venta de producto final o para su preparación dependiendo del tipo de cliente. Además identificar si es conveniente realizar una campaña de marketing para captar potenciales clientes que no están en el target del producto. Y la última pregunta que nos generamos es corroborar si el crecimiento en las redes sociales genera un impacto en la venta futura.



Herramientas tecnológicas implementadas

Para el presente trabajo se utilizaron los siguientes programas:

- Excel para la lectura y limpieza del Dataset.
- Canva para la creación del diseño de la documentación adicional
- Moqups para la creación del diagrama entidad-relación (<https://app.moqups.com/>)
- Power BI Desktop para la creación del tablero de control





Base de Datos relacionada a Power BI

Para poder hacer el análisis más profundo debemos trabajar con los datos de manera prolja, es por ello que a partir de lo recopilado realizamos una pequeña limpieza en nuestra base de datos que se originó en un archivo Excel. Distinguimos los campos que son llaves primarias y foráneas de manera que sean fácilmente identificables. Modificamos estructuras de las celdas, la tipología y tamaño. Organizamos campos y paginas para mantener un orden de lectura.

A continuación, se adjunta el archivo en formato Excel el cual vincularemos al Programa Power BI para su correspondiente transformación de Datos, relación de tablas y generación de reportes gráficos:



Diagrama E-R

El diagrama Entidad-Relación que se presenta en la siguiente diapositiva busca describir las relaciones entre las distintas entidades, donde se muestra el atributo principal del mismo. En cada Tabla o Entidad se identifica sus relaciones 1:N así como también sus Primary Key y/o su Foreing Key.

A partir de la información que conforma las tablas podemos decir que el PUBLICO ENCUESTADO puede ser de GENEROS diversos, VIVE EN un lugar determinado, tiene distintos tipo de EMPLEO, una BEBIDA preferida y compra diferentes PRODUCTOS. Cada entidad tiene sus diferentes subcategorías y campos.



CAFÉ CHEZ NOUS
COFFEE ROASTERS

Diagrama E-R

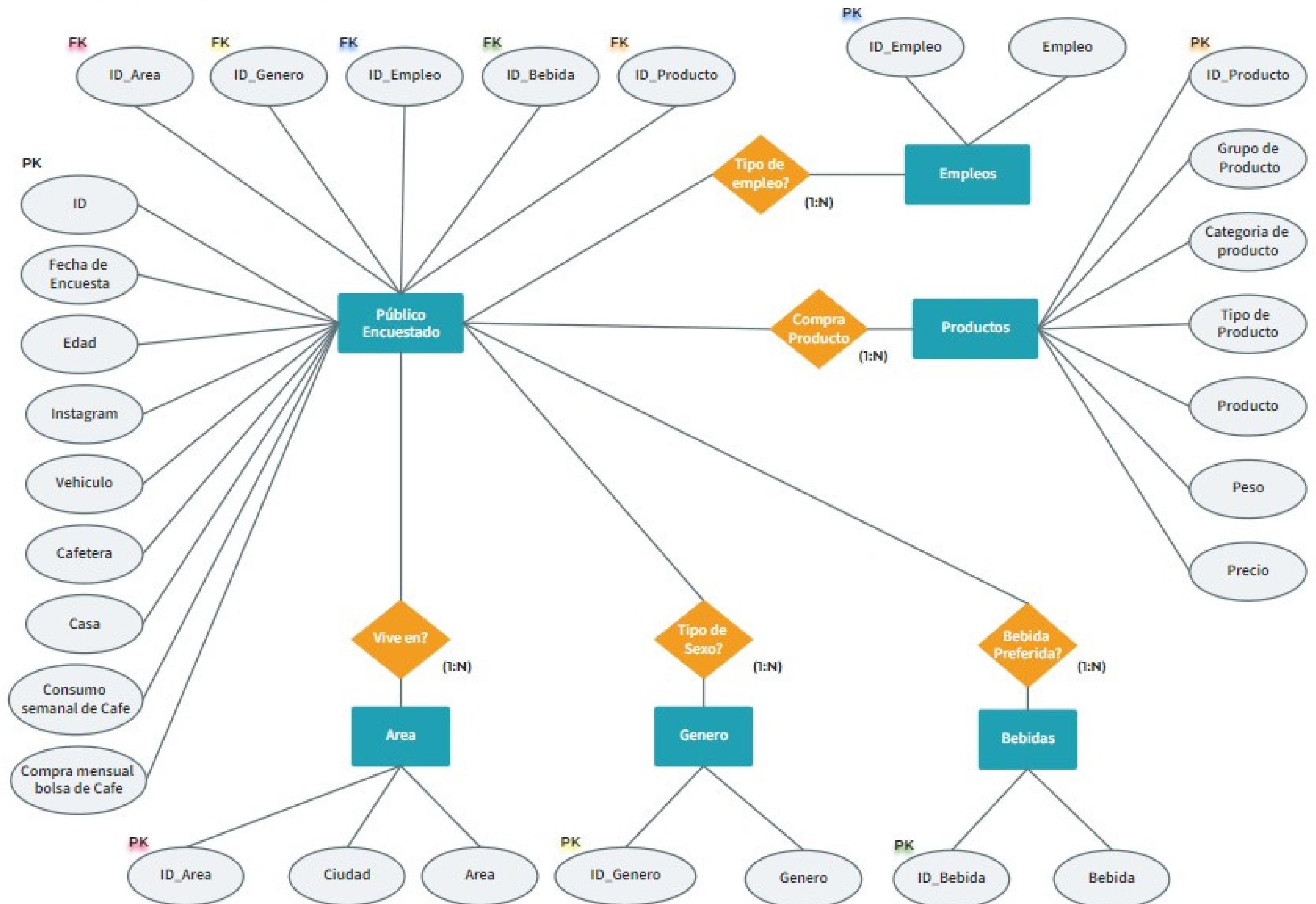


Tabla 1: Público Encuestado

Tabla Principal del Dataset en la cual obtenemos la mayor información para hacer nuestro análisis. Podemos visualizar todos los datos del público encuestado: Edad, Género, Tipo de Empleo, etc; partiendo del ID de cada persona. De esta tabla se relacionaran las tablas dependientes con las correspondientes Foreign Keys

Tipo de Clave	Campo	Tipo de Dato
PK	ID	INT
FK	Area_ID	INT
FK	Gender_ID	INT
FK	Employment_ID	INT
FK	Beverage_ID	INT
FK	Product_id	INT
-	Date_contacted	DATE
-	Age	TINYINT
-	Follow_Instagram	BIT
-	Owns_car	BIT
-	Owns_coffee_machine	BIT
-	Owns_home	BIT
-	Weekly_coffee_consumption	TINYINT
-	Buy_coffeebug_monthly	TINYINT
-	Bought_product	BIT

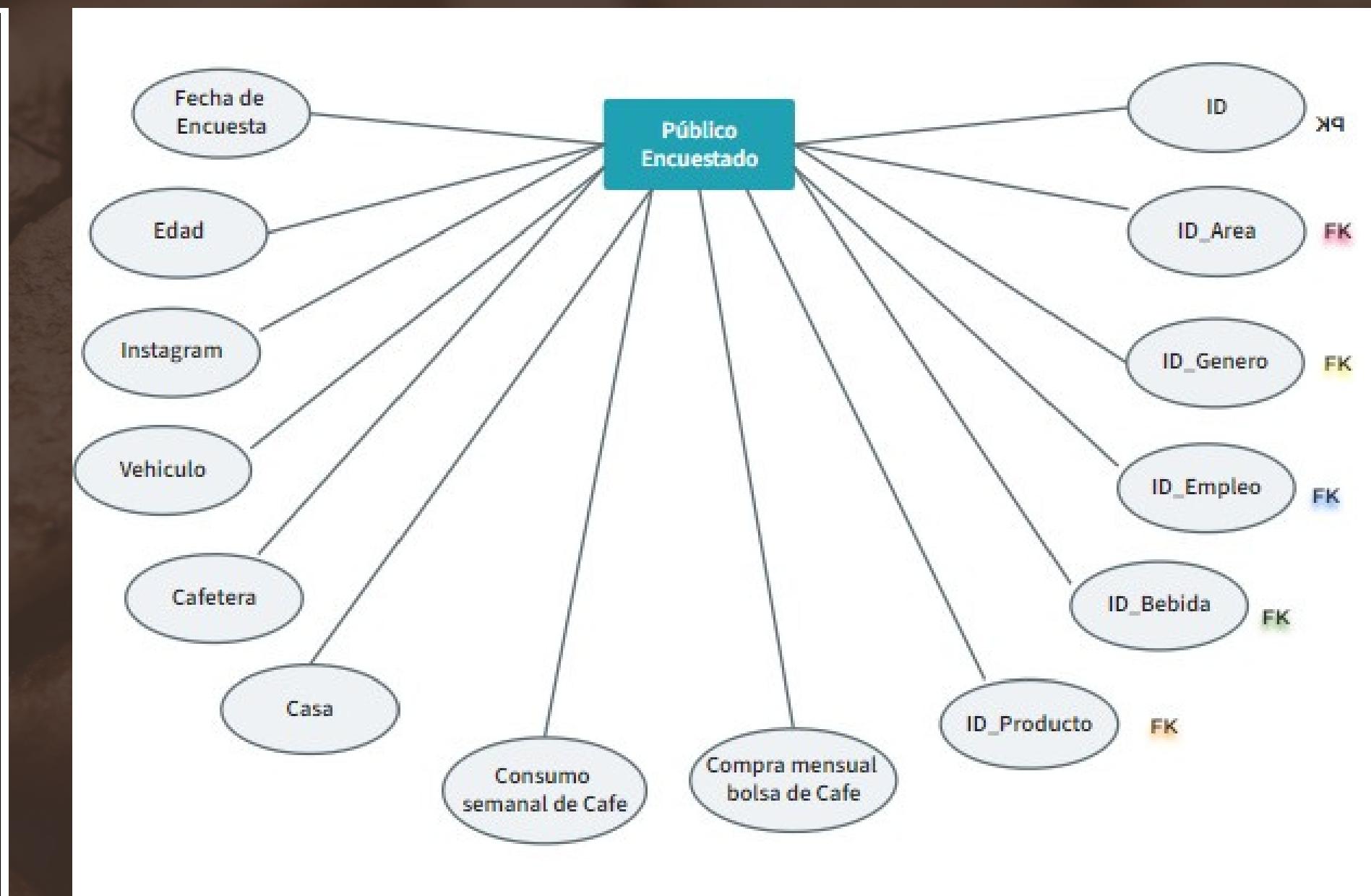


Tabla 2: Género

Tabla correspondiente para identificar el género del público encuestado

Tipo de Clave	Campo	Tipo de Dato
PK	Gender_ID	INT
-	Gender	VARCHAR(50)

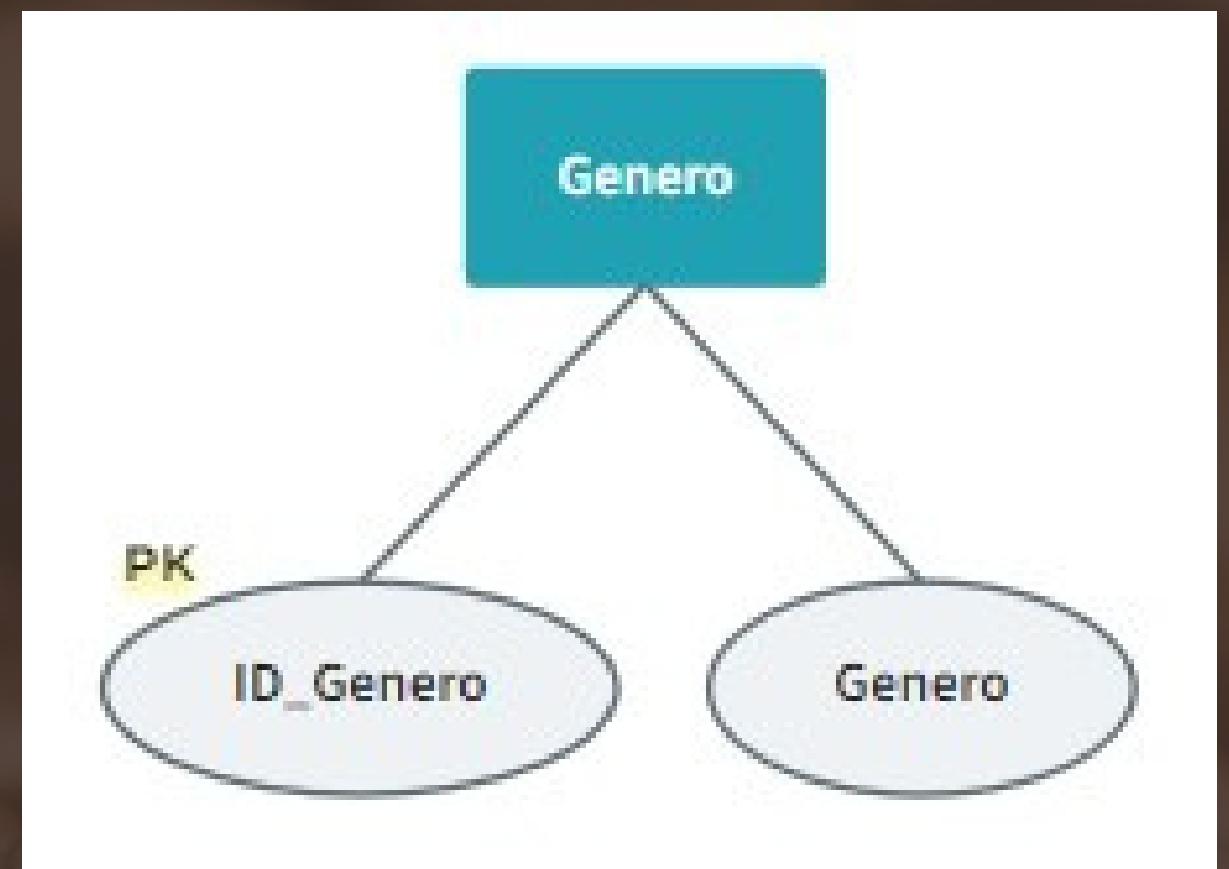


Tabla 3: Tipo de Empleo

En la siguiente tabla se desglosa los tipos de empleo que tienen las personas encuestadas. Si se trabajan part o fulltime, si están retiradas o si son estudiantes.

Tipo de Clave	Campo	Tipo de Dato
PK	Employment_ID	INT
-	Employment	VARCHAR(50)

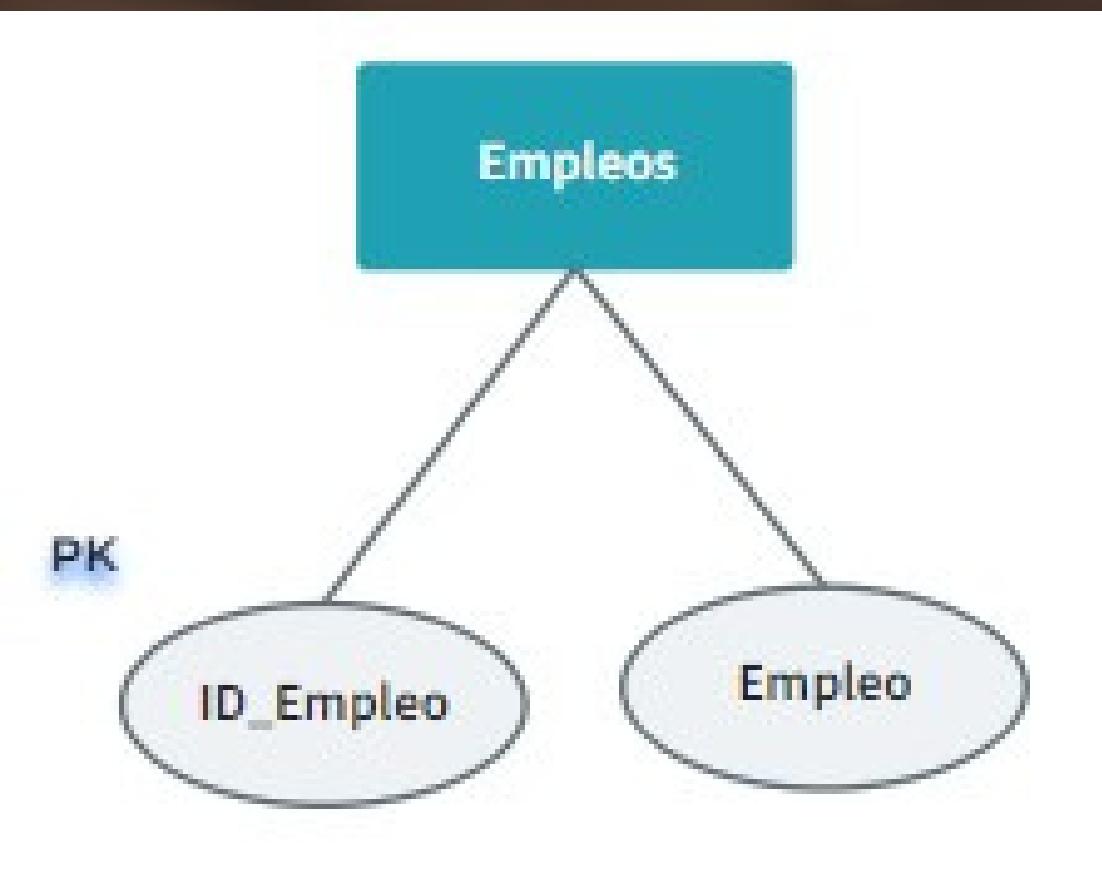


Tabla 4: Área

En nuestra tabla numero 4 podemos observar la distribución de las Áreas y barrios de origen de los masa encuestada En principio agrupado por ciudades.

Tipo de Clave	Campo	Tipo de Dato
PK	Area_ID	INT
-	City	VARCHAR(50)
-	Area	VARCHAR(100)

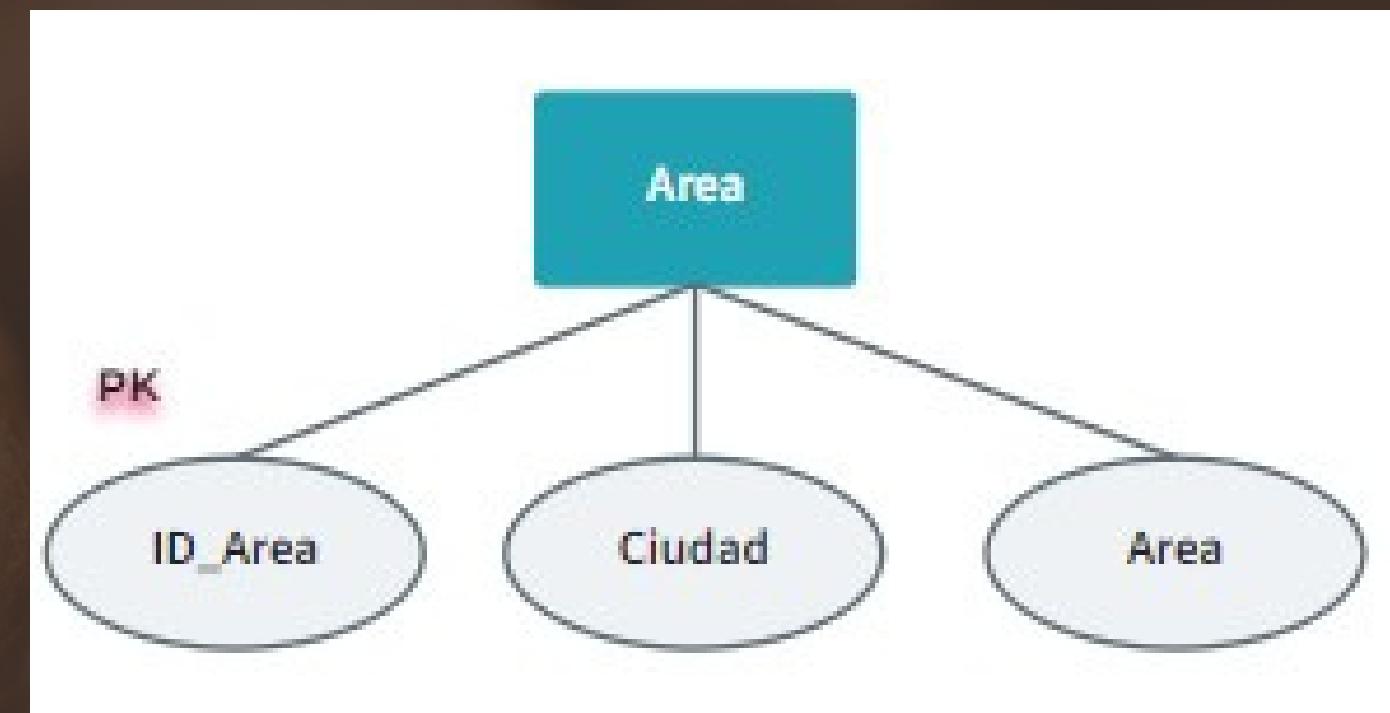


Tabla 5: Bebida de Preferencia

Tabla en la cual podemos observar la bebida preferida de cada persona. La Primary Key va a ser el ID de cada una de las bebidas elegidas.

Tipo de Clave	Campo	Tipo de Dato
PK	Beverage_ID	INT
-	beverage_preference	VARCHAR(50)

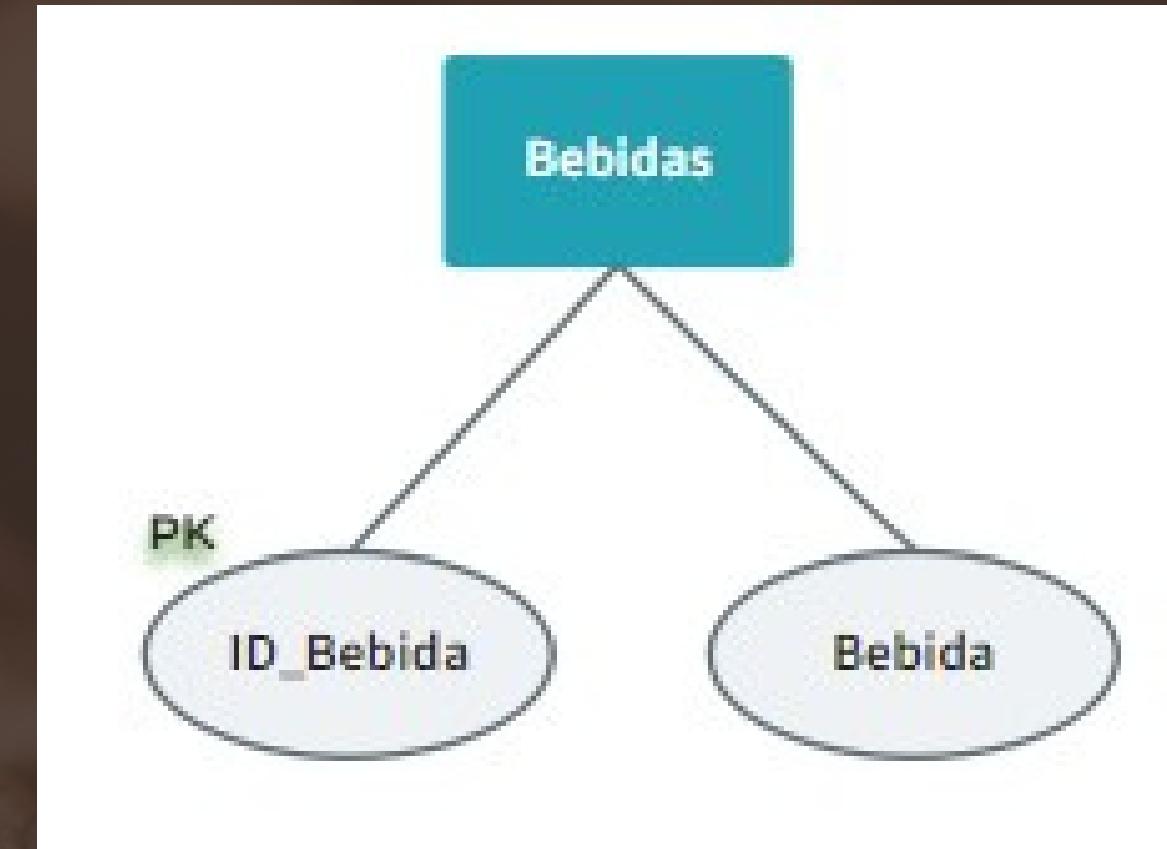
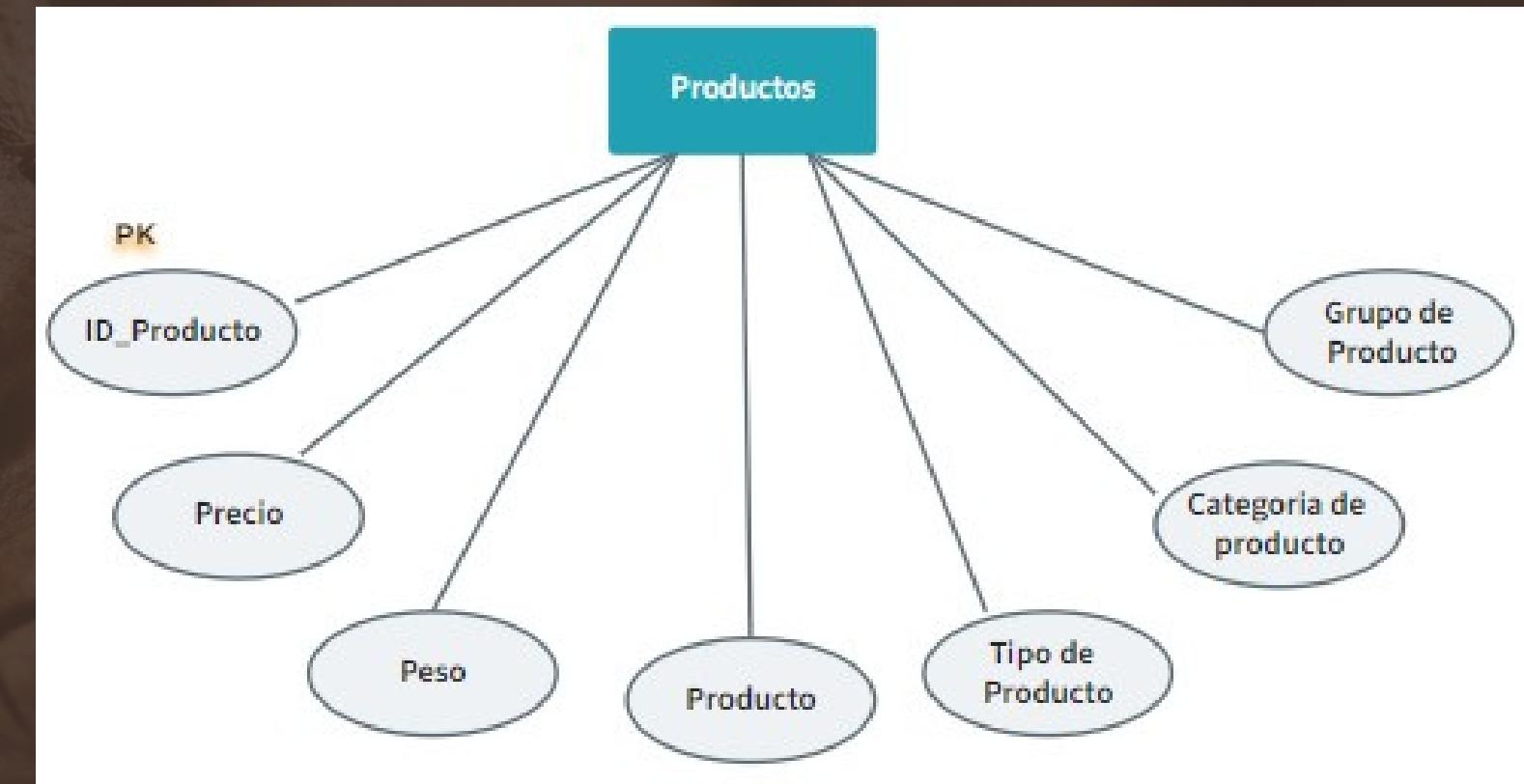


Tabla 6: Productos

Una de las tablas más completa y con más Campos. De la cual podemos determinar los grupos y subgrupos de cada producto para poder hacer un análisis más profundo de las compras por parte de los clientes: que se compra, el tipo, el precio y la unidad de medida de cada producto

Tipo de Clave	Campo	Tipo de Dato
PK	product_id	INT
-	product_group	VARCHAR(50)
-	product_category	VARCHAR(100)
-	product_type	VARCHAR(100)
-	product	VARCHAR(100)
-	unit_of_measure	VARCHAR(10)
-	current_retail_price	MONEY

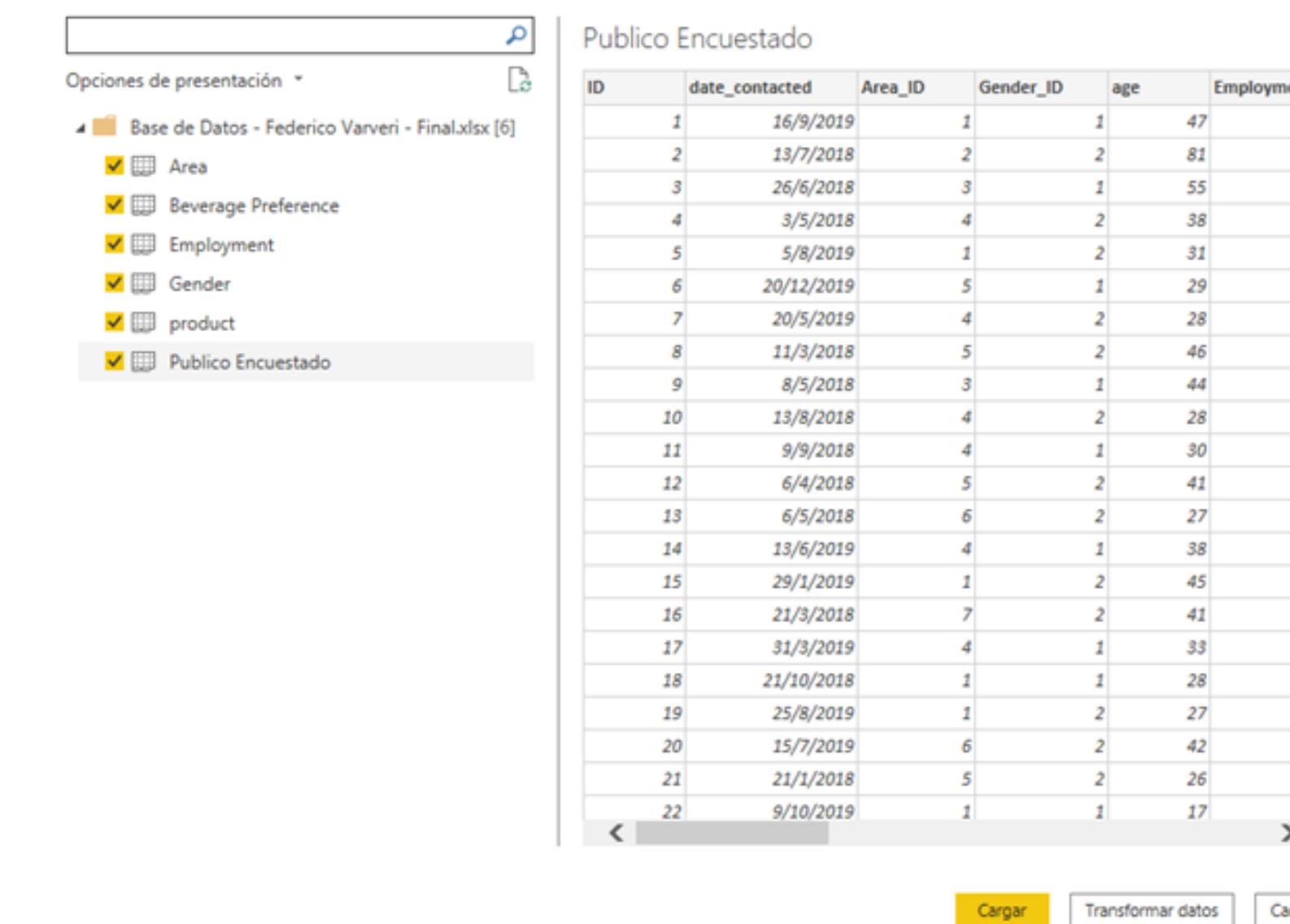




Modelo Relacional Power BI

Transformación de datos

Para poder realizar la transformación de Datos en el programa Power BI tuvimos que cargar, al mismo, los archivos planos de nuestra base de datos (Excel). En nuestro caso disponemos de uno solo: “Base de Datos - Federico Varveri – Final”, el cual cuenta con varias hojas referenciando a cada una de las tablas de hechos y dimensiones. Utilizaremos todas las tablas para realizar el informe:



Opciones de presentación

- Base de Datos - Federico Varveri - Final.xlsx [6]
 - Area
 - Beverage Preference
 - Employment
 - Gender
 - product
 - Publico Encuestado

ID	date_contacted	Area_ID	Gender_ID	age	Employment_ID
1	16/9/2019	1	1	47	
2	13/7/2018	2	2	81	
3	26/6/2018	3	1	55	
4	3/5/2018	4	2	38	
5	5/8/2019	1	2	31	
6	20/12/2019	5	1	29	
7	20/5/2019	4	2	28	
8	11/3/2018	5	2	46	
9	8/5/2018	3	1	44	
10	13/8/2018	4	2	28	
11	9/9/2018	4	1	30	
12	6/4/2018	5	2	41	
13	6/5/2018	6	2	27	
14	13/6/2019	4	1	38	
15	29/1/2019	1	2	45	
16	21/3/2018	7	2	41	
17	31/3/2019	4	1	33	
18	21/10/2018	1	1	28	
19	25/8/2019	1	2	27	
20	15/7/2019	6	2	42	
21	21/1/2018	5	2	26	
22	9/10/2019	1	1	17	

Cargar Transformar datos Cancelar

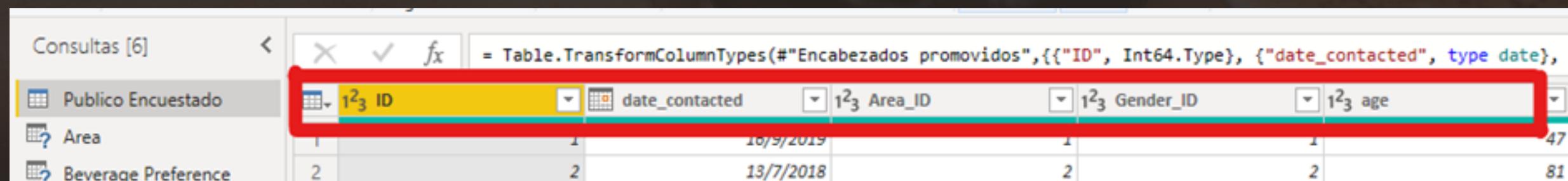
Transformación de datos

Una vez accedido a la opción “Transformar datos” nos abrirá el Power Query donde realizaremos la tan mencionada transformación y limpieza. En principio revisamos si los tipos de datos asignados automáticamente por el programa fueron reconocidos correctamente o debemos modificarlos. Los iconos en el encabezado nos definirán que Tipo de Dato es el campo.

Ejemplo con la Tabla “Publico Encuestado”:

Tabla 1: Publico Encuestado

Tipo de Clave	Campo	Tipo de Dato
PK	ID	INT
FK	Area_ID	INT
FK	Gender_ID	INT
FK	Employment_ID	INT
FK	Beverage_ID	INT
FK	Product_id	INT
-	Date_contacted	DATE
-	Age	TINYINT
-	Follow_Instagram	BIT
-	Owns_car	BIT
-	Owns_coffee_machine	BIT
-	Owns_home	BIT
-	Weekly_coffee_consumption	TINYINT
-	Buy_coffeebug_monthly	TINYINT
-	Bought_product	BIT



Consultas [6]

= Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID", Int64.Type}, {"date_contacted", type date},

ID	date_contacted	Area_ID	Gender_ID	age
1	10/9/2019	1	1	47
2	13/7/2018	2	2	81



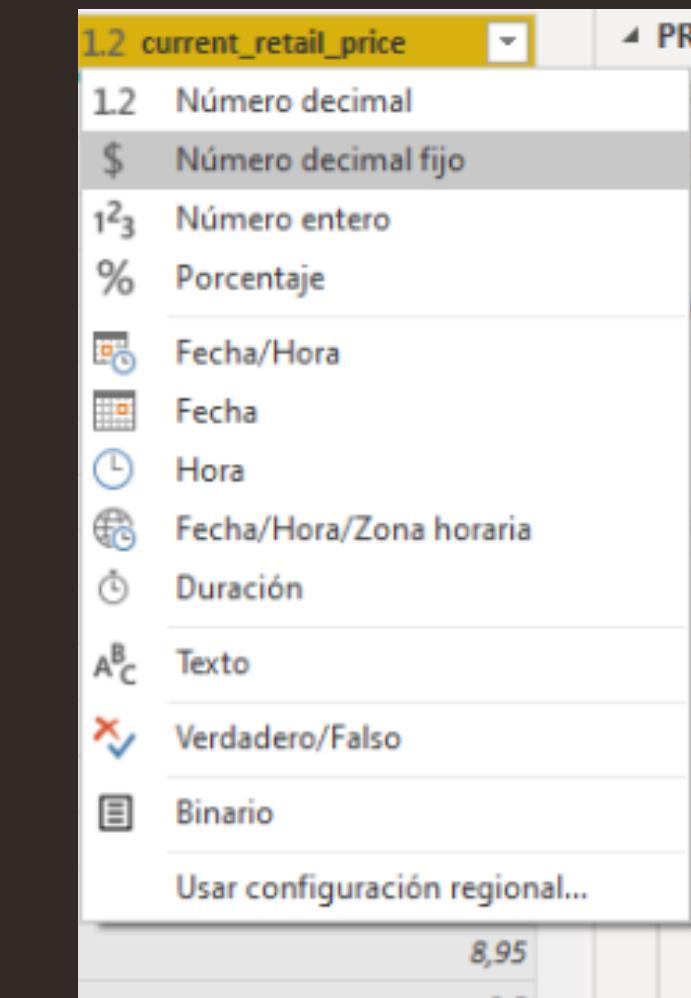
CAFÉ CHEZ NOUS

COFFEE ROASTERS



Transformación de datos

Luego de revisar todas las tablas distinguimos en Productos que el campo “Current_retail_price” tenía asignado el tipo de dato Número decimal (1.2) mientras que nosotros lo habíamos declarado como **Money** que consideramos era el correcto para hablar de precios. Teniendo en cuenta esto decidimos establecer el tipo de dato **Número decimal fijo (\$)** dentro de las opciones.



CAFÉ CHEZ NOUS
COFFEE ROASTERS

Transformación de datos

Eliminamos las columnas de las tablas que consideramos son innecesarias para la realización del informe.

Por ejemplo, decidimos que el campo "Unit_of_measure" (entre otros) no iba a ser utilizada por eso la suprimimos de la tabla Product.

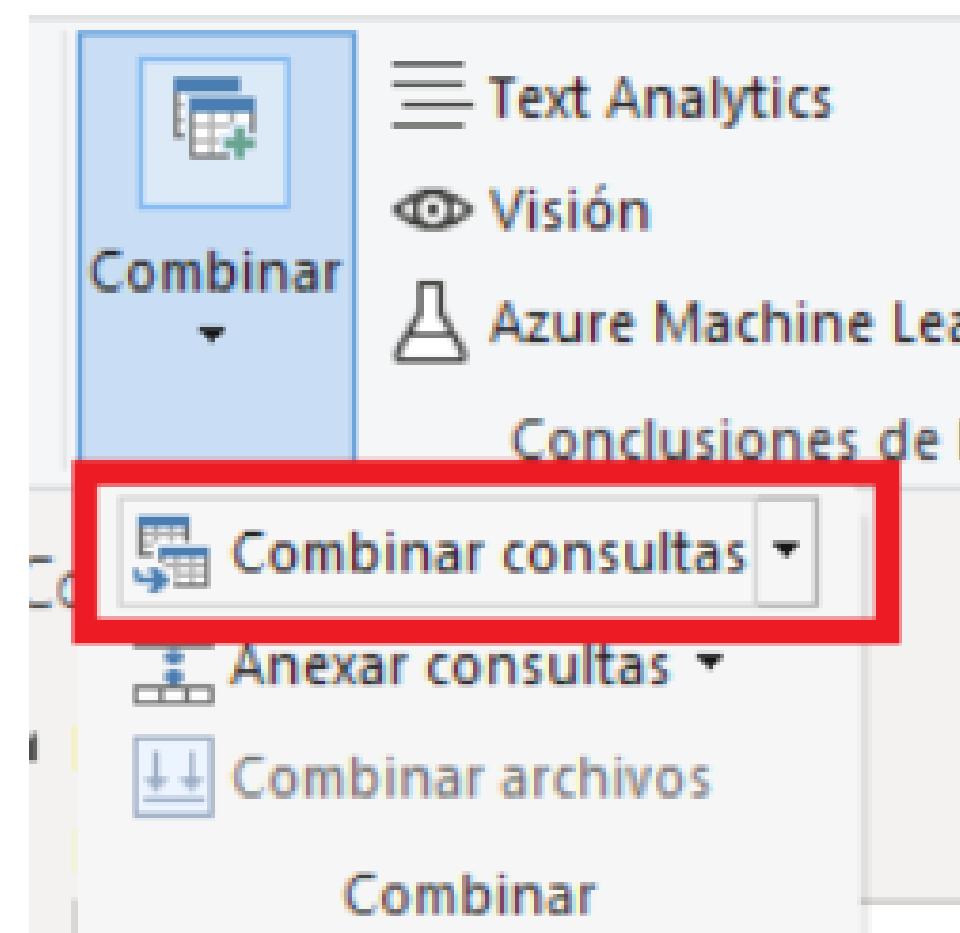
A	B	C
unit_of_measure	current_retail_price	
12 oz		Copiar
12 oz		Quitar
1 lb		Quitar otras columnas
1 lb		Duplicar columna
1 lb		Agregar columna a partir de los ejemplos...



Columnas Calculadas

Columna Gasto x Cliente

Creamos una nueva columna en nuestra tabla **Público Encuestado** donde calculamos el costo (y por ende lo gastado) de cada producto comprado por la persona. El inconveniente radicaba en que el campo del precio de los productos se encontraba en otra de las tablas (**Product**) por lo que tuvimos traer esa columna realizando la Combinación de consultas.



Columnas Calculadas

Para realizar esto había que vincular ambas tablas a partir de un campo en común (FK). Elegimos el campo product_id que comparten ambas tablas.

Combinar

Seleccione una tabla y las columnas coincidentes para crear una tabla combinada.

Publico Encuestado

home	beverage_ID	weekly_coffee_consumption	buy_coffeebug_monthly	bought_product	product_id
FALSE	1	2	0	FALSE	null
FALSE	2	2	1	FALSE	null
TRUE	3	5	0	FALSE	null
FALSE	3	6	1	TRUE	34
...

product

product_id	product_group	product_category	product_type	product	current_retail_price
1	Whole Bean/Teas	Coffee beans	Organic Beans	Brazilian - Organic	18,
2	Whole Bean/Teas	Coffee beans	House blend Beans	Our Old Time Diner Blend	18,
3	Whole Bean/Teas	Coffee beans	Espresso Beans	Espresso Roast	14,
4	Whole Bean/Teas	Coffee beans	Espresso Beans	Primo Espresso Roast	20,

Tipo de combinación

Externa izquierda (todas de la primera, coincidencias...)

Use las coincidencias aproximadas para comparar la combinación.

↳ Opciones de coincidencia aproximada

✓ La selección coincide con 1216 de 28657 filas de la primera tabla.

Aceptar Cancelar





Columnas Calculadas

Una vez que nos trae el campo a la Tabla que queremos combinar, elegimos la columna a expandir. Referenciamos current_retail_price, correspondiente al precio del producto vinculado.

The screenshot shows a software interface for expanding calculated columns. At the top, there are tabs for 'product_id' and 'product'. Below them is a search bar labeled 'Buscar columnas para expandir' and a sorting icon 'AZ'. A radio button 'Expandir' is selected, while 'Agregar' is unselected. A list of columns is displayed, with 'current_retail_price' checked. At the bottom, a checkbox 'Usar el nombre de columna original como prefijo' is checked. There are 'Aceptar' and 'Cancelar' buttons at the bottom right.

Columna	Opción
(Seleccionar todas las columnas)	<input type="checkbox"/>
product_id	<input type="checkbox"/>
product_group	<input type="checkbox"/>
product_category	<input type="checkbox"/>
product_type	<input type="checkbox"/>
product	<input type="checkbox"/>
current_retail_price	<input checked="" type="checkbox"/>

Usar el nombre de columna original como prefijo

Aceptar Cancelar





Columnas Calculadas

El próximo paso fue reemplazar los null asignados automáticamente (porque no había producto otorgado por parte del cliente) por el valor 0, para que nos queden valores numéricos manteniendo el mismo tipo de dato (Número decimal fijo \$):

1 ↵ 2 Reemplazar los valores

Reemplazar los valores

Reemplace un valor con otro de las columnas seleccionadas.

Valor que buscar

null

Reemplazar con

0

Quedando así la Columna Gasto_x_Cliente:

1 ² ₃ product_id	\$ Gasto_x_Cliente
null	0,00
null	0,00
null	0,00
34	2,45
null	0,00
2	18,00
32	3,00
36	3,75

Tabla Calendario

Creamos la tabla calendario para poder gestionar los indicadores y las segmentaciones de manera eficiente, con ella disponemos todos los periodos de tiempo necesarios para nuestro análisis. También permite utilizar funciones de inteligencia de Tiempo.

Para establecer la tabla en Power Query utilizamos la siguiente formula:

```
let
    // configurations start
    Today=Date.From(DateTime.LocalNow()), // today's date
    FromYear = 2018, // set the start year of the date dimension. dates start from 1st of January of this year
    ToYear = 2019, // set the end year of the date dimension. dates end at 31st of December of this year
    StartofFiscalYear=7, // set the month number that is start of the financial year. example; if fiscal year start is July, value is 7
    firstDayofWeek=Day.Monday, // set the week's start day, values: Day.Monday, Day, Sunday....
    // configuration end
   FromDate=#date(FromYear,1,1),
   ToDate=#date(ToYear,12,31),
    Source=List.Dates(
        FromDate,
        Duration.Days(ToDate-FromDate)+1,
        #duration(1,0,0,0)
    ),
    #"Converted to Table" = Table.FromList(Source, Splitter.SplitByNothing(), null, null, ExtraValues.Error),
    #"Renamed Columns" = Table.RenameColumns(#"Converted to Table",{{"Column1", "Fecha"}}),
    #"Changed Type" = Table.TransformColumnTypes(#"Renamed Columns",{{"Fecha", type date}})
in
    #"Changed Type"
```



Tabla Calendario

Debido a que la creación de la tabla nos traerá los datos en un único campo en formato DATE (dd/mm/aaaa) realizamos las siguientes transformaciones para tener separado por días, meses y años para poder trabajar de manera efectiva. La idea es tener una columna por cada intervalo de tiempo (DIAS, MESES, AÑOS) y uno más en donde el mes sea expresado con el nombre del mismo. Para esto primero duplicamos la columna originada por cada intervalo de tiempo que queríamos agregar, en nuestro caso fueron 4.

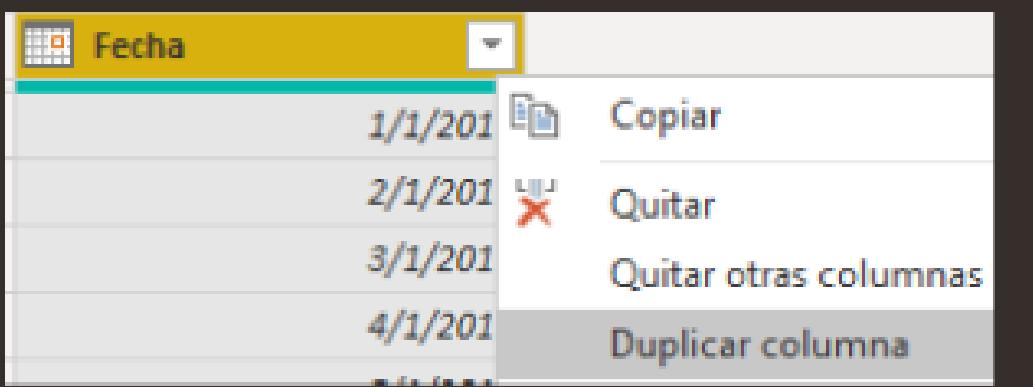




Tabla Calendario

Luego, cada campo lo transformamos en el intervalo correspondiente: Año, Mes, Nombre de mes y Día. También se puede hacer por trimestre o semana.

Cambiar tipo
Transformar
1 2 Reemplazar los valores...
Reemplazar errores...
Agrupar por...
Rellenar
Anulación de dinamización de columnas

Año
Trimestre
Mes
Semana
Día
Transformaciones de texto

Año
Año
Inicio del año
Final del año

Modificamos el nombre del campo para poder identificar más fácil.

Fecha	1 2 Año	1 2 Fecha - Copia.1	A B C Fecha - Copia.2	1 2 Fecha - Copia.3
1/1/2018	2018		1 enero	1
2/1/2018	2018		1 enero	2

Una vez finalizada la tabla lo último que queda por hacer es declararla como nuestra TABLA DE FECHA en Power BI y la relacionamos a la tabla de Hechos a partir de la columna de Fecha que dispongamos.

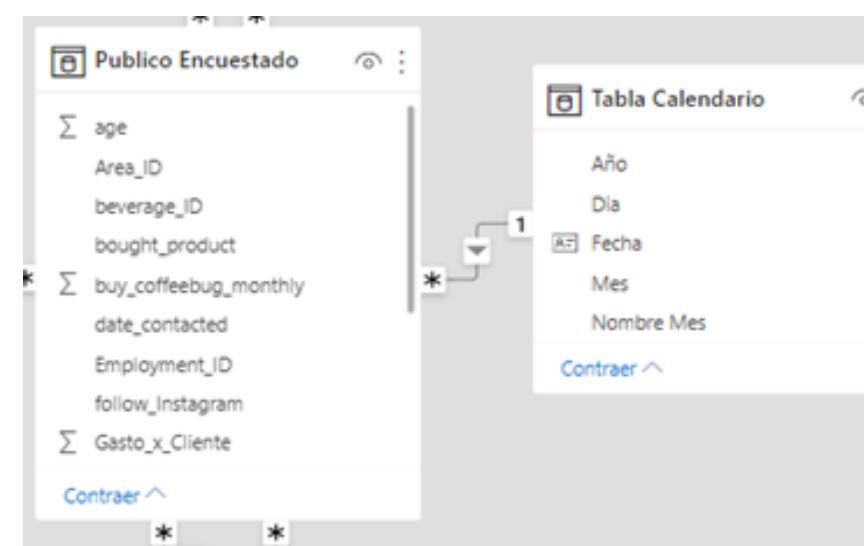


Tabla Beverage Preference

Se decidió agregar una columna con imágenes a la tabla de bebidas preferidas para poder cargarlas a nuestro tablero de control y para que puedan ser plasmadas en visualizaciones.



beverage_ID	beverage_preference	image
1	beer	https://cdn-icons-png.flaticon.com/512/3993/3993988.png
2	wine	https://cdn-icons-png.flaticon.com/512/3995/3995914.png
3	coffee	https://cdn-icons-png.flaticon.com/512/3361/3361122.png
4	water	https://cdn-icons-png.flaticon.com/512/2447/2447774.png
5	tea	https://cdn-icons-png.flaticon.com/512/1764/1764353.png
6	soda	https://cdn-icons-png.flaticon.com/512/735/735842.png

Tabla Audio

Nueva tabla generada para agregar audio al proyecto.

A screenshot of a software interface showing an 'Audio' tab selected. The interface displays code for an audio source, specifically: 1 <audio controls> <source src=

A screenshot of a software interface showing an 'Audio' section. The interface displays a summary: ∑ Audio and a 'Link' button.



Medidas Calculadas



Se creó una tabla **Medidas** para almacenar todas las medidas que generamos para llevar organizado todas las mediciones realizadas en nuestro análisis

- ✓  **Medidas**
 - >  **Cientes**
 - >  **Fecha**
 - >  **Medidas Tendencia Central**
 - >  **Propiedad**
 - >  **Seguidores Instagram**
 - >  **Variables**
 - >  **Variaciones**

Medida Clientes

Medida originada para distinguir a las personas que ya consumen en el local por lo cual son clientes y poder representarlos tanto en cantidad como en porcentaje.

- ✓  **Cientes**
 -  **Cientes**
 -  **Cientes%%**

```
Cientes = CALCULATE(COUNT('Publico Encuestado'[ID]),'Publico Encuestado'[bought_product]=True)
```

```
Cientes%% = CALCULATE(COUNT('Publico Encuestado'[ID]),'Publico Encuestado'[bought_product]=True)/[VAR Personas Encuestadas]
```

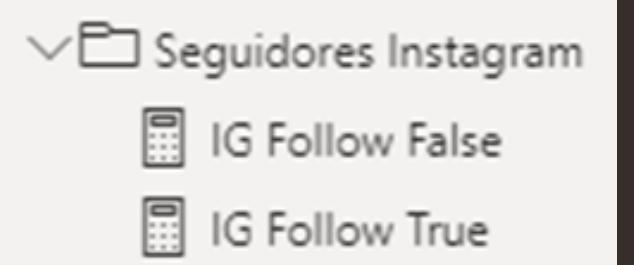


Medida Seguidores Instagram

Se crearon 2 medidas calculadas correspondientes al campo *follow_Instagram* para identificar por medio de un gráfico que porcentaje del publico encuestado es seguidor de nuestra página en las redes sociales.

```
1 IG Follow True =  
2 CALCULATE(Count('Publico Encuestado'[ID]),  
3 'Publico Encuestado'[follow_Instagram]=TRUE())
```

```
1 IG Follow False =  
2 CALCULATE(Count('Publico Encuestado'[ID]),  
3 'Publico Encuestado'[follow_Instagram]=FALSE())
```





Medida de Tendencia Central

- ✓ └ Medidas Tendencia Central
 - └ Moda
 - └ Promedio

Moda

Se estableció una Medida para calcular la Moda (Estadística Descriptiva) de las ventas totales según los filtros aplicados.

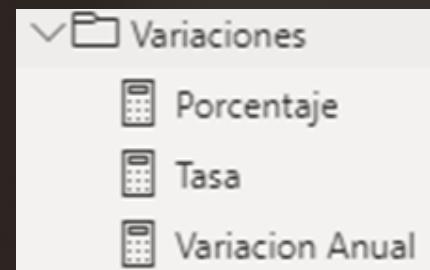
```
1 Moda =  
2 VAR GRUPOS = SUMMARIZE('Publico Encuestado', 'Publico Encuestado'[Gasto_x_Cliente], "CANTIDAD", CALCULATE(COUNT('Publico  
Encuestado'[Gasto_x_Cliente]),NOT(ISBLANK('product'[product_id]))))  
3 VAR MAXIMO = MAXX(GRUPOS, [CANTIDAD])  
4 VAR MODA_FINAL = FILTER(GRUPOS, [CANTIDAD] = MAXIMO)  
5 VAR RESULTADO = CONCATENATEX(MODA_FINAL, FORMAT('Publico Encuestado'[Gasto_x_Cliente], "$ #,###.00"), " - ")  
6 RETURN  
7 "Moda de Venta: " & RESULTADO & UNICHAR(10) & "Se repite "& MAXIMO &" cantidad de veces."
```

Promedio

Otras de las medidas de Tendencia Central para demostrar el Promedio de las ventas según el segmento.

```
1 Promedio = CALCULATE(AVERAGE('Publico Encuestado'[Gasto_x_Cliente]), NOT(ISBLANK('Publico Encuestado'[product_id])))
```

Variaciones



Tasa variacion Mes a Mes

Tasa de variación para comparar las ventas por mes.

```
1 Tasa =
2 VAR __PREV_MONTH =
3     CALCULATE(
4         SUM('Publico Encuestado'[Gasto_x_Cliente]),
5         DATEADD('Tabla Calendario'[Fecha], -1, MONTH)
6     )
7 RETURN
8     DIVIDE(
9         SUM('Publico Encuestado'[Gasto_x_Cliente]) - __PREV_MONTH,
10        __PREV_MONTH
11    )
```

Tasa variacion Año a Año

Medida para poder tener referencia de la variación de compras realizadas por los clientes entre el año 2018 y 2019

```
1 YoY% de Gasto_x_Cliente =
2 VAR __PREV_YEAR =
3     CALCULATE(
4         SUM('Publico Encuestado'[Gasto_x_Cliente]),
5         DATEADD('Tabla Calendario'[Fecha], -1, YEAR)
6     )
7 RETURN
8     DIVIDE(
9         SUM('Publico Encuestado'[Gasto_x_Cliente]) - __PREV_YEAR,
10        __PREV_YEAR
11    )
```

Porcentaje

```
1 Porcentaje = FORMAT(COUNT('Publico Encuestado'[ID])/CALCULATE(COUNT('Publico Encuestado'[ID]), ALL()), "00.00%")
```



CAFÉ CHEZ NOUS

COFFEE ROASTERS



Variables



Narrativa

Narrativa creada con distintas variables con el fin de expresar más información sobre el producto seleccionado.

```
1 Narrativa =
2 VAR Texto1= "Seleccione una categoria de Producto para obtener mas detalle de su consumo."
3 VAR Texto2= "Usted ha seleccionado: "
4 VAR Texto3= ". Las cantidad de productos vendidos fue "
5 VAR Texto4= ", lo cual representa una venta total de "
6 VAR Total= FORMAT(SUM('Publico Encuestado'[Gasto_x_Cliente]), "$ #,###.00")
7 VAR Contar= CALCULATE(COUNT('Publico Encuestado'[ID]),'Publico Encuestado'[product_id])
8 RETURN
9 IF(SELECTEDVALUE('product'[product_category])=BLANK(), Texto1, Texto2 & SELECTEDVALUE('product'[product_category])&Texto3&Contar&Texto4&Total)
```

Variable Total Personas

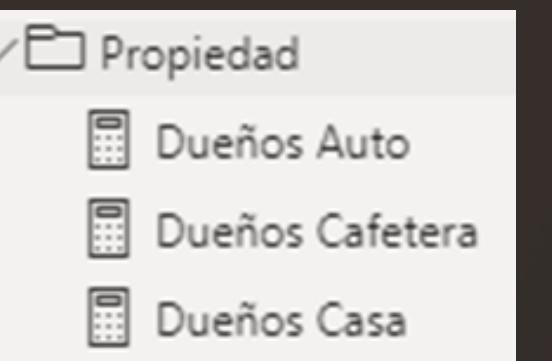
Variable para poder utilizar de una manera más rápida el recuento total de las personas encuestadas.

```
1 VAR Personas Encuestadas = count('Publico Encuestado'[ID])
```



Medida Propiedad (Dueños)

Se crearon 3 variables para poder identificar a las personas según si son dueños de casa, auto o cafetera para luego poder hacer segmentaciones.



```
Dueños Auto = CALCULATE(COUNT('Publico Encuestado'[ID]),'Publico Encuestado'[owns_car]=True)/[VAR Personas Encuestadas]
```

```
Dueños Casa = CALCULATE(COUNT('Publico Encuestado'[ID]),'Publico Encuestado'[owns_home]=True)/[VAR Personas Encuestadas]
```

```
Dueños Cafetera = CALCULATE(COUNT('Publico Encuestado'[ID]),'Publico Encuestado'[owns_coffee_machine]=True)/[VAR Personas Encuestadas]
```

Fecha Hoy

Recuperar el día y hora del momento a partir de la función NOW() para poder mostrarlo en nuestra portada.

```
Fecha Hoy= NOW()
```





Dimensiones

Grupo Etario

Realizaremos una columna calculada o dimensión para agrupar el público encuestado según su edad. A partir de la columna creada (*Age Agrupada*) y la función que le designemos, determinaremos si el público es Joven (16-26), Adulto (27-55) o Adulto Mayor (55-85). Esta dimensión servirá en el futuro para poder filtrar las visualizaciones entre otras cosas.

- 1 *Age Agrupada* = SWITCH(TRUE(),
- 2 'Publico Encuestado'[age]<26,"Joven",
- 3 'Publico Encuestado'[age]<55,"Adulto",
- 4 "Persona Mayor")



Clase Social

Se decidió realizar una dimensión para clasificar la clase social de las personas según si son propietarios de inmuebles, vehículos y máquina de Café para realizar un análisis más exhaustivo.

Consideramos que las personas que disponen de una casa propia pertenecen a una clase social “alta”, mientras que las que tienen únicamente medio de transporte son de una clase social “media” y quienes no poseen ninguna de estas dos cosas corresponden a una clase social “baja”:

```
1 Clase Social = IF(  
2 'Publico Encuestado'[owns_home]=True,"Clase Alta",  
3 | IF(  
4 | | 'Publico Encuestado'[owns_car]=True,"Clase Media", "Clase Baja"))  
5
```



Trabajo

Con esta dimensión se busca poder segmentar y separar a las personas que se encuentran empleadas de las que son desempleadas para realizar un análisis más completo. Se realizó con la función SWITCH tomando el campo *Employement* de la tabla con el mismo nombre.

```
1 Trabajo = SWITCH(True,  
2 Employment[Employment]="part-time","Empleada",  
3 Employment[Employment]="full-time","Empleada",  
4 Employment[Employment]="retired","Desempleado",  
5 Employment[Employment]="unemployed","Desempleado",  
6 Employment[Employment]="student","Desempleado")
```



Consumo Semanal

A partir de la columna “*weekly_coffee_consumption*” de la tabla **Publico Encuestado** realizamos agrupamientos según la cantidad de veces que las personas toman café semanalmente para poder visualizar con un gráfico de embudo como era el consumo.

Registrando: 0-No consume, de 1 a 3 veces – Consumo Bajo, 4 a 6 – Consumo Medio y 7 a 9 – Consumo Alto.

Grupos

Nombre *	Campo
Consumo Semanal	<i>weekly_coffee_consumption</i>

Tipo de grupo

Lista

Valores no agrupados

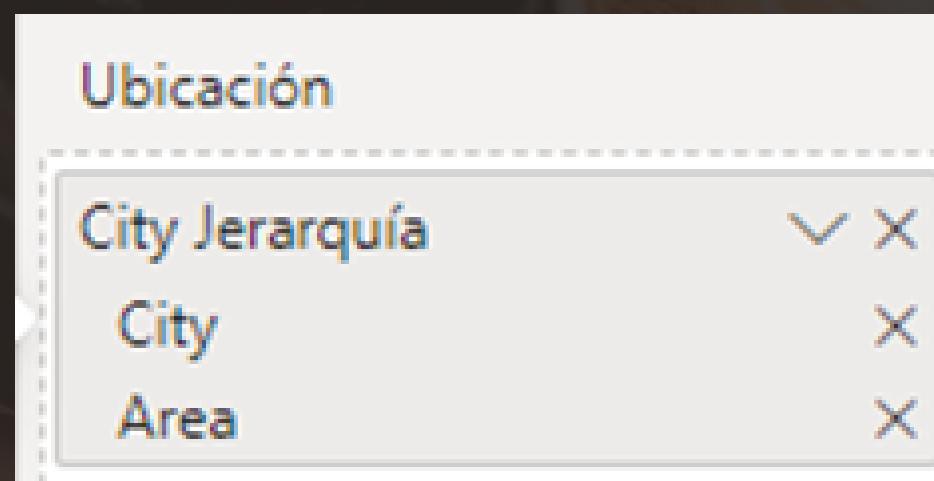
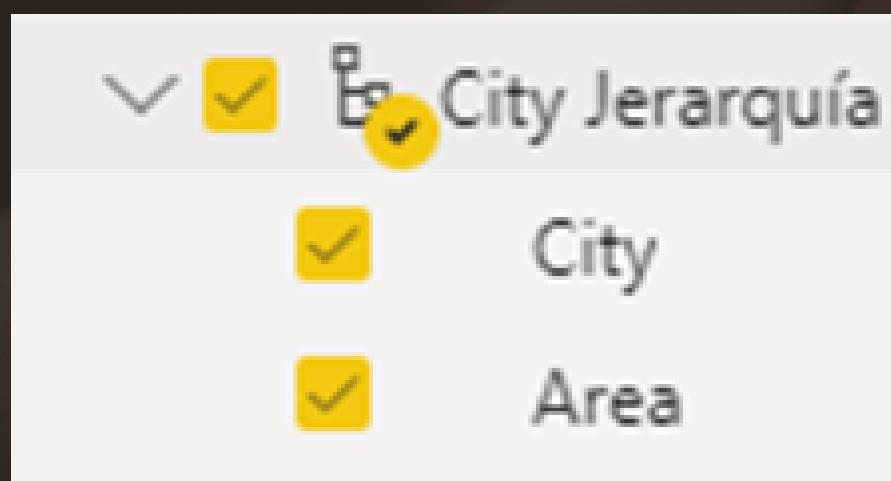
Grupos y miembros

- ▶ Alto (7-9 Veces)
- ▶ Bajo (1-3 Veces)
- ▶ Medio (4-6 Veces)
- ▶ No Consume
- ◀ Otros
 - Contiene todos los valores desagrupados



Grafico Mapa

Decidimos crear una jerarquía debido a que en nuestra tabla Area donde tenemos nuestros campos para graficar el mapa no disponemos de la información completa. Se pueden identificar los barrios y las ciudades donde se realizaron las encuestas pero no agrupados al país que pertenece (Estados Unidos). Con la creación de la Jerarquía haremos que las columnas queden agrupadas por el país, teniendo una mejor referencia del lugar. En el grafico colocaremos la jerarquía como ubicación:



Análisis Funcional del Tablero

Botonera:

Se diseñaron los siguientes botones.

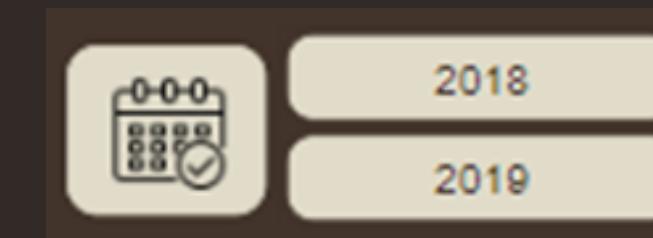
- Solapa Portada: Botones para dirigirse a las redes sociales.
- Solapa Glosario: Marcadores fijado en imágenes para ir directamente a la solapa elegida.
- Ubicados en todas las solapas se generaron botones para poder moverse entre páginas o poder ir directamente a la Portada.
- Botón para quitar todos los filtros que estaban aplicados.
- Ubicado en la solapa “Barrios” este botón permite cambiar la visualización de la cantidad de clientes en porcentaje o valor absoluto.



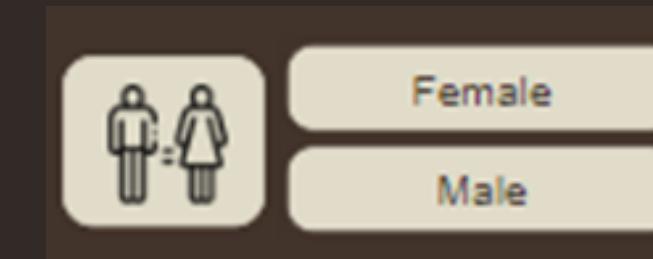


Segmentadores

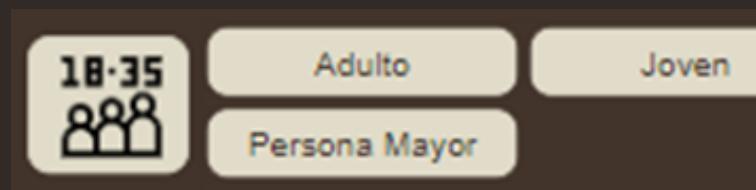
Segmentadores emplazados en todas las páginas donde disponemos gráficos para poder filtrar a elección.



Años: Se muestran los años del período encuestado para filtrar en lista vertical



Género: Elección según genero del publico



Grupo etario: Agrupamos las edades por grupo etario para poder segmentar de una manera más eficiente.

Paleta de Colores

Se tuvo en cuenta para el diseño de la paleta los colores de la empresa y los que más se asocian al café como producto. En caso de necesitar más colores se recurrió a variaciones de esos mismos pero intentando mantener el tablero monocromático.



Visualización de los datos

A continuación expondremos con más detalle cada una de las solapas creadas en nuestro proyecto. En todas las páginas incluimos un lienzo relacionado con el tema del proyecto (Café) con la idea de darle un marco más profesional al trabajo. Se intentó incluir todo tipo de visualizaciones, medidas, formulas, entre otras cosas, vistas en clase con el fin de ponerlas en práctica.

- Portada
- Glosario
- Perfiles
- Barrios
- Consumo
- Preferencias
- Tooltip (Oculta)



Solapa 1: Portada

A continuación expondremos con más detalle cada una de las solapas creadas en nuestro proyecto. En todas las páginas incluimos un lienzo relacionado con el tema del proyecto (Café) con la idea de darle un marco más profesional al trabajo. Se intentó incluir todo tipo de visualizaciones, medidas, formulas, entre otras cosas, vistas en clase con el fin de ponerlas en práctica.

The slide features a dark brown background image of coffee beans. Overlaid on the left side is a smaller image of a person holding a coffee cup, with the text "CAFE CHEZ NOUS COFFEE ROASTERS". The main title "PROYECTO FINAL DE DATA ANALITYCS" is centered in large, bold, white letters. Below the title is the logo for "CAFE CHEZ NOUS COFFEE ROASTERS", which includes a cartoon character of a person wearing a hat and holding a coffee cup. At the bottom left, there is a timestamp "09/02/2023 10:57:50 a.m.". Social media icons for WhatsApp, LinkedIn, and Instagram are at the bottom center, along with the text "CoderHouse Docente Luciano Julián Gómez Alumno Federico Varveri". A small circular arrow icon is in the bottom right corner.

Solapa 2: Glosario

En esta solapa el usuario tendrá la información introductaria del proyecto, el alcance, la temática de los datos y nuestra Hipótesis. Podemos identificar también las imágenes que nos dirigirán a las siguientes páginas y un Scroller, al lado del título, con narrativa para informar el modo de uso de los botones.

GLOSARIO

Por favi

ENCUESTA DE CONSUMO DE CAFE

Introducción:

Informe detallado con el fin de identificar potenciales clientes para nuestro negocio y medidas a tomar para maximizar ganancias de los ya consumidores de nuestros productos.

Hipótesis:

Comprender si la compra de café y productos relacionados está directamente ligado al perfil, tipo de consumo y preferencias de la población de la ciudad de Chicago. Comprobar también si es conveniente la venta de producto final o para su preparación dependiendo del tipo de cliente. Además identificar si es conveniente realizar una campaña de marketing para captar potenciales clientes que no están en el target del producto. Y la última pregunta que intentaremos responder es corroborar si el crecimiento en la redes sociales genera un impacto en la venta futura.

Alcance y Temática de los Datos:

A partir de la información recopilada en forma de encuesta presencial en nuestras distintas cafeterías de especialidad "Chez Nous", la temática a analizar serán los datos biométricos y el tipo de consumo de las personas, con el fin de poder identificar potenciales clientes y/o maximizar las ventas de los ya activos.

Con el fin de tener una base de datos completa y extensa en el tiempo se recopilaron los datos durante 2 años, comenzando el 1º de Enero del 2018 hasta el 31 de Diciembre del 2019. Con esto en mente, nuestro objetivo de este trabajo es reconocer patrones de conducta de los consumidores analizando distintas variables: Género, edad, tipo de empleo, posición económica, consumo de Café semanal, compra de productos del local, entre otras.

Para ubicarnos en el tipo de análisis, este informe será realizado desde un diagnóstico, el cual definirá ciertas conclusiones de la forma de actuar en base a los datos históricos arrojados en la base de datos.

09/02/2023 10:57:50 a.m.

CAFÉ CHEZ NOUS COFFEE ROASTERS

← → ⌂

Solapa 3: Perfiles

Primera página con Visualizaciones en la cual detallaremos rasgos generales de las personas encuestadas: Género, Edad, clasificación de empleo, consumo de café semanal. Diseñamos los Segmentadores en la parte superior de la página para poder filtrar por año, género y grupo etario, estos segmentos se mantendrán a lo largo del proyecto. Incluimos también el botón para quitar todos los filtros y 2 tarjetas sobre el total de ventas y personas encuestadas.



Solapa 4: Barrios

Información detallada sobre los datos geográficos del público encuestado. A partir de esa información lo relacionamos con su condición laboral, clase social y si son clientes del negocio o no. Se mantienen los botones de la página anterior pero se agregó un Tooltip en nuestra tabla de los barrios para obtener más información según el área.



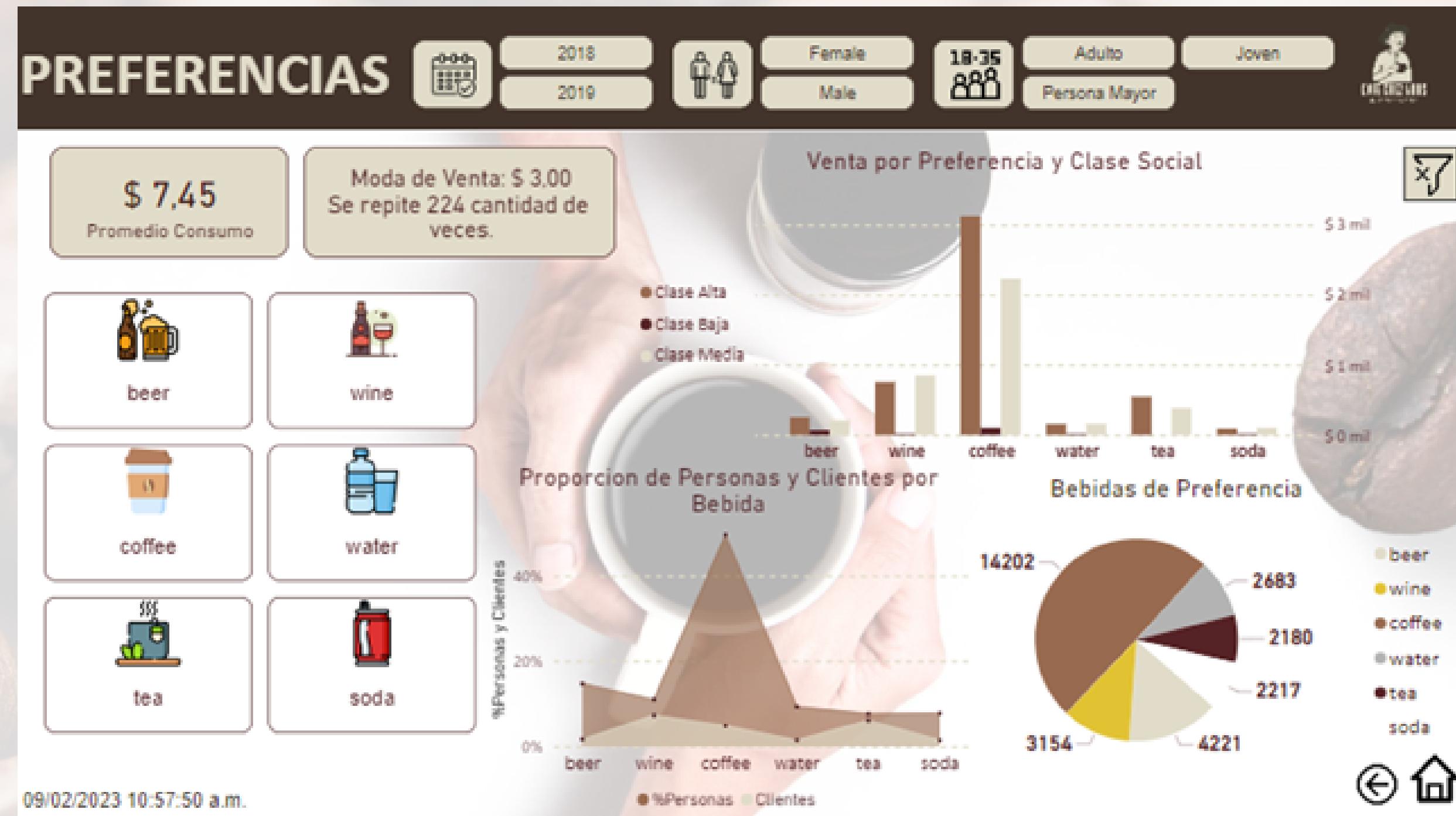
Solapa 5: Consumo

El enfoque dado en esta solapa es el consumo (ventas) de las personas, agrupado por grupo, tipo y producto. Podemos observar también la variación de las ventas anuales y por meses. Se incluyó una narrativa con uso de variables para ver más información de acuerdo al tipo de producto.



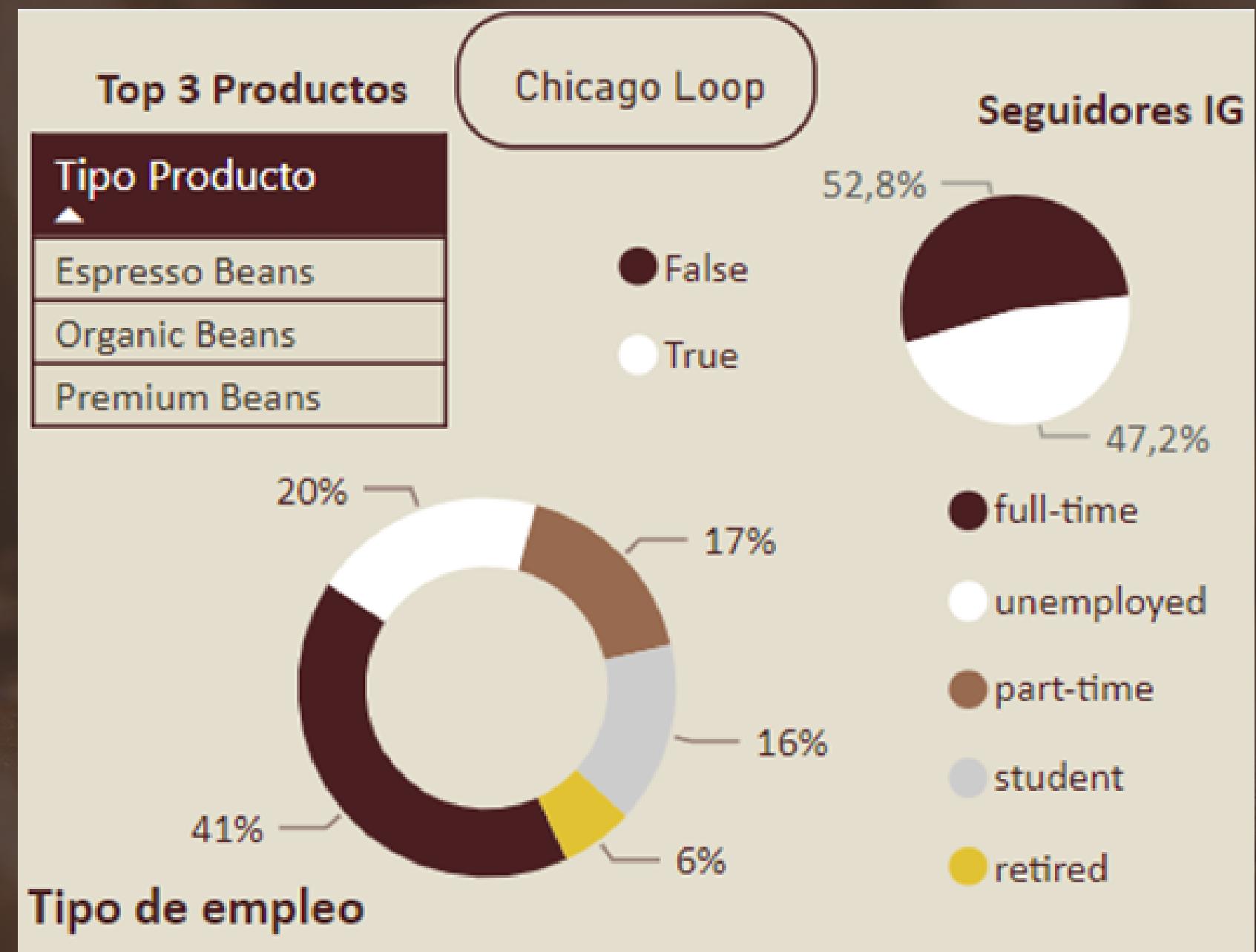
Solapa 6: Preferencias

Según las preferencias de bebidas por parte del público se proyectó esta página para demostrar el impacto en el negocio. Se agregó un Chiclet Slicer con el tipo de bebida (con su imagen correspondiente) y dos tarjetas con Medidas de Tendencia Central.



Solapa 7: Tooltip (Oculta)

Nuestra última solapa es un Tooltip que la mantendremos oculta de cara al usuario. En ella nos encargamos de aportar más información del área en el cual el usuario se posicione (ubicado en la página "barrios"). Top de productos, relación con el tipo de empleo y porcentaje de seguidores en instagram son los gráficos que decidimos incluir.



Conclusion



A nivel general podemos observar a partir de lo plasmado en el proyecto que se puede realizar una campaña de marketing para captar como clientes gran flujo de las personas que fueron encuestadas. Mismo seguir con las iniciativas tomadas durante el 2019 que fueron fructíferas de acuerdo al aumento del margen de venta y cantidad de clientes.



De acuerdo a los datos reflejados el género femenino fue quien más venta generó en el negocio si comparamos con el masculino, teniendo en cuenta que están equilibrado en otros aspectos como el tipo de trabajo y consumo semanal como podemos observar en la primera solapa del proyecto.

Si visualizamos el grafico de los seguidores de instagram la representación nos muestra que están muy equilibrado los porcentajes de cuantos son followers de la cafetería y cuantos no, pero esto se puede entender teniendo en cuenta que abarca grupos etarios muy amplios (como las personas mayores). Si segmentamos por "Jóvenes" los porcentajes se mantienen y acá es donde deberíamos enfatizar para incorporar más de esta juventud que se involucra más con las redes.

Conclusion

Pasando por la segunda solapa, Barrios, identificamos que el porcentaje de clientes es muy bajo en proporción a la cantidad de personas, pero aumenta en los barrios más alejados de la ciudad (donde las condiciones socio económicas son más bajas) y esto está directamente ligado a la clase social. En este ítem, teniendo en cuenta las visualizaciones, el grueso de clientes pertenecen a las clase Media y quienes se encuentran desempleados. Ahí podríamos enfocar como incrementar los clientes de las clases sociales más altas, teniendo en cuenta que son más pudientes. Mismo para abarcar todos los colectivos sociales también se puede incluir un grupo de productos de bajo precio de venta.

La venta de granos representa el porcentaje de venta más grande dentro de todos los tipos de productos, esto se debe a que los precios dentro de este grupo son más altos en comparación a otros y siendo también el producto estrella de nuestro negocio. El té, tanto en hierbas como la bebida, sigue muy de cerca ubicándose en 2º y 4º puesto respectivamente. Teniendo en cuenta ello deberíamos encausar nuestro plan a generar más venta del café bebida (3º) como producto final, es indispensable acaparar más consumidores de la más famosa infusión.

En tanto, el merchandising tuvo su explosión en 2019 alcanzando su pico de venta a finales de ese año, es decir, genera impacto en la clientela. Sería conveniente enfatizar su oferta para que siga siendo adquirido.

Nuestras últimas conclusiones serán basadas a las visualizaciones a partir de las bebidas preferidas de las personas encuestadas. Haciendo un paneo rápido podemos identificar que claramente el café es la bebida preferida del público, pero es muy bajo el porcentaje de clientes (como ya mencionamos). En el mismo grafico podemos ver que la cerveza también es una de las más convocantes (mayor aun en el género masculino) pero en este caso casi no tenemos alcance a estos clientes, una de las medidas a tomar tal vez es incorporar esa consumición a nuestra carta. Mismo el vino, preferido por las personas mayores (>55) puede ser otra bebida a incluir.



Futuras Líneas

Sería conveniente, de cara futuro, enriquecer de datos la encuesta realizada para poder hacer un análisis más exhaustivo. Se podría incluir por ejemplo los siguientes ítems:

- Ofrecer una gama de productos gastronómicos y puedan elegir su preferido para acompañar la bebida adquirida.
- Preguntar sobre el servicio prestado durante la compra y fase postcompra.
- Consultar que considera el público sobre el precio de los productos para por hacer una comparación con otras cafeterías.
- Realizar una gestión de Customer Experience con los ya clientes y ver en qué aspectos se pueden mejorar en la infraestructura del local.
- Intentar que la encuesta no se haga tan extensa en el tiempo porque las respuestas pueden estar condicionadas a factores externos.

