

# PROYECTO 1:

## Base de datos.

Bootcamp: Data Analyst 3

Carlos Avendaño

### **Integrantes:**

- Nestor Edilberto Castro Salas
- Numas Enrique Salazar Martinez
- Mónica Alexia Castro Ortega

San Salvador, 16 de septiembre, 2023

## Entidades identificadas

---

De acuerdo a los datos proporcionados por los usuarios, se han definido las siguientes entidades:

**Inventory:** Esta entidad tiene la función de almacenar los datos relacionados al inventario como lo son gestión de existencias, control y estado de inventario.

Inventory	
PK	<u>inventory_id</u>
FK1	product_id product_quantity product_available

**Category:** Surge de la necesidad de clasificar nuestros productos en una categoría, con el objetivo de identificar patrones de compra en nuestros clientes.

Category	
PK	<u>category_id</u>
	category_name description

**Customers:** (Clientes), como en cualquier situación de negocio, donde se implica una venta de un artículo, necesitamos información de quien ha realizado la compra, de ahí la importancia de asignar una entidad para tomar decisiones y planes de marketing.

Customers	
PK	<u>customer_id</u>
	customer_name
	customer_adress
	customer_age
	customer_city
	customer_phone
	customer_email
	customer_gender

**Employees:** (Empleados), relacionado al punto anterior, ya que se necesita saber quién realizó la venta, para tomar decisiones internas de la empresa, como lo son bonos por meta alcanzada.

Employees	
PK	<u>employee_id</u>
FK1	store_id
	employee_name
	employee_gender
	employee_birthdate
	employee_phone

**Shipment:** (Envíos), al ser una empresa que tiene presencia internacional, necesitamos una entidad en la cual incluimos todos los datos relacionados a envíos.

Shipment	
PK	<u>shipment_id</u>
	shipper_name
	shipper_phone

**Products:** (Productos), surge de la necesidad para identificar qué productos se venden más, a qué categoría pertenecen y además con qué proveedor se relacionan.

Products	
PK	<u>product_id</u>
FK1	supplier_id
FK2	category_id
	product_name
	product_price

**Fact\_sales** (Tabla de hechos): Es la tabla de hechos relacionada a todas nuestras ventas, en la cual se encuentra, los clientes, empleados, detalles de venta, envíos y ventas totales

fact_sales	
PK	<u>sales_id</u>
FK1	customer_id
FK2	employee_id
FK3	shipment_id
FK4	campaign_id
	salesdate
	totalsales

**Marketing\_campaigns:** (Campañas de marketing), surge de la necesidad de relacionar las ventas con las campañas y hacer análisis sobre los resultados.

Marketing_campaigns	
PK	<u>campaign_id</u>
	campaign_type
	campaign_mediatype
	startdate
	finishdate

**Suppliers:** (Proveedores), nace de la necesidad de identificar los proveedores de cada producto, para realizar ofertas de acuerdo a la demanda de los productos.

Suppliers	
PK	<u>supplier_id</u>
	supplier_name
	supplier_address
	supplier_country
	supplier_email
	supplier_phone

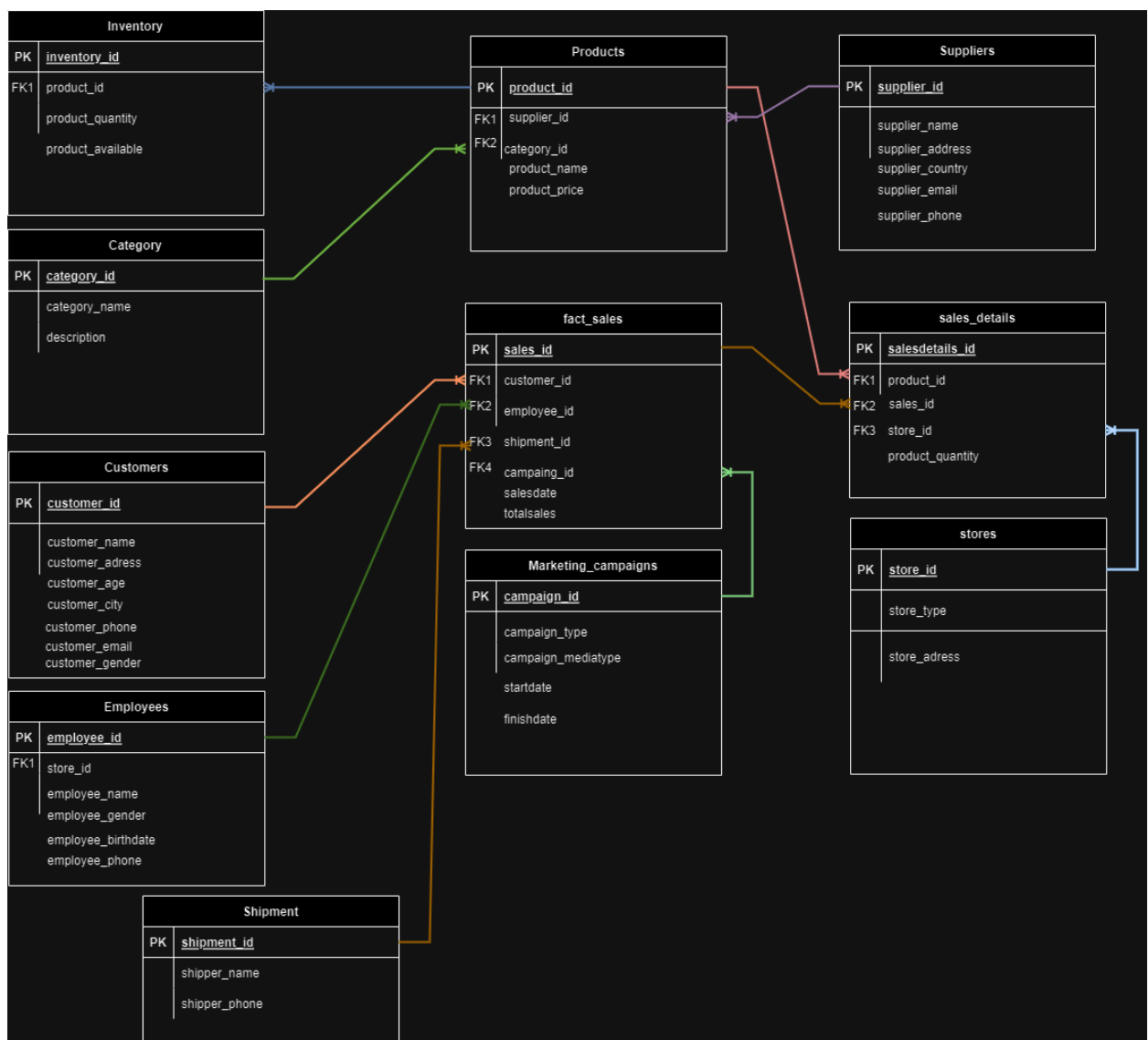
**Sales\_details:** (Detalles de ventas), esta entidad se creó como un apoyo a nuestra tabla de hechos ventas, para dar instrucciones de cada atributo de las ventas, por ejemplo, la cantidad de producto, tipos de productos, tienda, etc.

sales_details	
PK	<u>salesdetails_id</u>
FK1	product_id
FK2	sales_id
FK3	store_id
	product_quantity

**Store:** (Tienda), como la empresa tiene dos tipos de comercio, online y físico, se creó esta entidad para identificar el impacto que genera cada una y tomar decisiones sobre los datos obtenidos.

stores	
PK	<u>store_id</u>
	store_address
	store_type

## Tabla entidad - relación



## Consultas

---

### 1. Primera consulta:

*¿Qué tipo de clientes compran más productos o tienen ventas altas, relacionado edad, género y venta total?*

*Además, se identificaron ventas totales por mes en el periodo estudiado.*

#### **Requerimientos:**

“Nuestro principal objetivo es obtener información sobre el rendimiento de las ventas y el comportamiento del cliente. Necesitamos entender **qué productos se están vendiendo más** y a **qué tipo de clientes**. También queremos identificar tendencias y patrones de compra para mejorar nuestras estrategias de marketing”

#### **Top 10 de clientes con ventas altas:**

Listado de clientes con mayores ventas totales en el periodo, según su edad, género y ciudad.

#### **SELECT**

cu.customer\_name,

sl.customer\_id,

cu.customer\_age,

cu.customer\_gender,

cu.customer\_city,

**SUM** (sl.totalsales) **AS** VENTATOTAL

#### **FROM**

fact\_sales **sl**

#### **INNER JOIN**



customers **cu**

**ON**

sl.customer\_id = cu.customer\_id

**GROUP BY**

cu.customer\_name,sl.customer\_id, cu.customer\_age,  
cu.customer\_gender,cu.customer\_city

**ORDER BY**

VENTATOTAL **DESC**

**LIMIT** 10;

	customer_name character varying (255) 🔒	customer_id integer 🔒	customer_age integer 🔒	customer_gender character varying (255) 🔒	customer_city character varying (255) 🔒	ventatotal numeric 🔒
1	Annette Roulet	41	22	FEMALE	Toulouse	31515.01
2	Frédérique Citeaux	7	27	MALE	Strasbourg	23891.15
3	Maria Anders	1	37	FEMALE	Berlin	22581.21
4	Roland Mendel	20	23	MALE	Graz	20981.36
5	Fran Wilson	48	39	MALE	Portland	14861.52
6	Carlos Hernández	35	41	MALE	San Cristóbal	13790.60
7	Felipe Izquierdo	47	35	MALE	I. de Margarita	13732.05
8	Martín Sommer	8	26	MALE	Madrid	12684.37
9	Thomas Hardy	4	33	MALE	London	11941.4
10	Aria Cruz	21	22	FEMALE	São Paulo	11889.75

### Top 10 productos más vendidos:

Top 10 productos más vendidos, pero no necesariamente representan las mayores ventas.

**SELECT**

p.product\_id,

p.product\_name,

**SUM**(sd.product\_quantity) **AS** total\_vendido,

**SUM**(S.totalsales) **AS** venta\_total

**FROM**

salesdetails **sd**

**INNER JOIN**

products **p**

**ON**

p.product\_id = sd.product\_id

**INNER JOIN**

fact\_sales **s**

**ON**

sd.sales\_id = s.sales\_id

**INNER JOIN**

customers **cu**

**ON**

s.customer\_id = cu.customer\_id

**GROUP BY**

p.product\_id, p.product\_name

**ORDER BY**

total\_vendido **DESC**

**LIMIT** 10;

	product_id [PK] integer	product_name character varying (255)	total_vendido bigint	venta_total numeric
1	31	Gorgonzola Telino	458	30106.20
2	60	Camembert Pierrot	430	43472.45
3	35	Steeleye Stout	369	36779.94
4	59	Raclette Courdavault	346	37540.45
5	2	Chang	341	33787.51
6	16	Pavlova	338	31220.96
7	71	Fløtemysost	336	21599.02
8	17	Alice Mutton	331	27484.45
9	62	Tarte au sucre	325	30639.6
10	33	Geitost	316	13812.86

### Top 10 productos que generan ventas altas:

Productos que generan altos valores monetarios.

**SELECT**

```
p.product_id,  
p.product_name,  
    SUM(sd.product_quantity) AS total_vendido,  
    SUM(S.totalsales) AS venta_total
```

**FROM**

```
salesdetails sd
```

**INNER JOIN**

```
products p
```

**ON**

```
p.product_id = sd.product_id
```

**INNER JOIN**

```
fact_sales s
```

**ON**

```
sd.sales_id = s.sales_id
```

**INNER JOIN**

```
customers cu
```

**ON**

```
s.customer_id = cu.customer_id
```

**GROUP BY**

```
p.product_id, p.product_name
```

**ORDER BY**

```
venta_total DESC
```

**LIMIT 10;**

	product_id [PK] integer	product_name character varying (255)	total_vendido bigint	venta_total numeric
1	38	Côte de Blaye	239	79624.07
2	60	Camembert Pierrot	430	43472.45
3	68	Scottish Longbreads	199	38340.5
4	72	Mozzarella di Giovanni	270	37591.77
5	59	Raclette Courdavault	346	37540.45
6	29	Thüringer Rostbratwurst	168	37107.77
7	35	Steeleye Stout	369	36779.94
8	2	Chang	341	33787.51
9	46	Spegesild	145	31729.15
10	16	Pavlova	338	31220.96

### Ventas totales en los meses durante el periodo:

**SELECT**

**EXTRACT**(YEAR FROM salesdate) AS anio,

**EXTRACT**(MONTH FROM salesdate) AS mes,

**SUM**(totalsales) AS Ventas\_totales

**FROM**

fact\_sales

**WHERE**

salesdate

**BETWEEN** '2022-01-01' **AND** '2023-09-11'

**GROUP BY EXTRACT**(YEAR FROM salesdate), **EXTRACT**(MONTH FROM salesdate)

**ORDER BY** Ventas\_totales **DESC**

	anio numeric 🔒	mes numeric 🔒	ventas_totales numeric 🔒
1	2023	1	79570.88
2	2022	12	63721.23
3	2022	11	62163.99
4	2022	10	51528.69
5	2022	7	37779.85
6	2022	9	34565.60
7	2022	8	33285.49
8	2023	2	19978.91

## 2. Segunda consulta:

*¿Qué relación existe entre las campañas promocionales y las ventas generadas?*

### **Requerimientos:**

“**Juan:** Para el análisis de marketing, necesitamos información sobre las campañas promocionales que realizamos, incluyendo el tipo de campaña, el medio utilizado y las fechas de inicio y fin. También nos gustaría conocer los resultados de cada campaña en términos de ventas generadas.

**Juan:** Estoy de acuerdo. Además, sería útil conocer el impacto de nuestras campañas de marketing en las ventas y determinar si hay alguna correlación entre las acciones promocionales y el aumento de las ventas.”

## Relación entre campañas promocionales y ventas:

### SELECT

s.campaign\_id,  
 SUM (totalsales) AS total, cam.campaign\_mediatype, cam.campaign\_type,  
 cam.startdate AS Fecha\_inico\_campana, cam.finishdate AS  
 Fecha\_final\_campana

### FROM

fact\_sales AS s

### JOIN

campaigns AS cam

### ON

s.campaign\_id = cam.campaign\_id

### WHERE

salesdate

BETWEEN '2022-01-01' AND '2023-09-11'

### GROUP BY

s.campaign\_id, cam.startdate, cam.finishdate, cam.campaign\_mediatype,  
 cam.campaign\_type

### ORDER BY total DESC

campaign_id integer	total numeric	campaign_mediatype character varying (255)	campaign_type character varying (255)	fecha_inico_campana date	fecha_final_campana date
3	50930.61	newspaper	get a free trip all paid for purchases over \$ 3000	2023-01-01	2023-01-15
2	38568.69	TV commercial	Get a free TV 60" for 2000 dollars spent	2022-10-01	2022-10-31
1	30706.95	social media	20% discount on your next sale for 500 dollars spe...	2022-08-01	2022-08-30

Al relacionar la consulta anterior “Relación entre campañas de venta” con la consulta de “Ventas totales en los meses durante el periodo”, se obtiene como resultado lo siguiente:

	anio numeric 🔒	mes numeric 🔒	ventas_totales numeric 🔒
1	2023	1	79570.88
2	2022	12	63721.23
3	2022	11	62163.99
4	2022	10	51528.69
5	2022	7	37779.85
6	2022	9	34565.60
7	2022	8	33285.49
8	2023	2	19978.91

La campaña **número 3**, generó un fuerte impacto de compra en el mes de enero del año 2023 representando un **64% de la venta total** del respectivo mes.

### 3. Tercera consulta:

*Estado de inventario: Costo, proveedores, gestión de existencia, productos que se venden rápido y los que no se venden (rotación).*

#### **Requerimientos:**

“**María:** Desde el punto de vista del inventario, necesitamos hacer un seguimiento de los productos que tenemos en stock, la cantidad vendida y la cantidad restante. Además, sería útil conocer el costo de los productos y los proveedores correspondientes.”

#### **Costos de productos y proveedores:**

**SELECT**

p.product\_name, sp.supplier\_name, p.product\_price

**FROM**

products **p**

**INNER JOIN**

suppliers **sp**

**ON**

p.supplier\_id = sp.supplier\_id

**GROUP BY**

p.product\_name, sp.supplier\_name, p.product\_price;

product_name character varying (255)	supplier_name character varying (255)	product_price numeric
Grandma's Boysenberry Spread	Grandma Kelly's Homestead	25.00
Ipoh Coffee	Leka Trading	46.00
Sasquatch Ale	Bigfoot Breweries	14.00
Nord-Ost Matjeshering	Nord-Ost-Fisch Handelsgesellschaft mbH	25.89
NuNuCa Nuß-Nougat-Creme	Heli Süßwaren GmbH & Co. KG	14.00
Louisiana Fiery Hot Pepper Sauce	New Orleans Cajun Delights	21.05
Ravioli Angelo	Pasta Buttini s.r.l.	19.50
Teatime Chocolate Biscuits	Specialty Biscuits, Ltd.	9.20
Mishi Kobe Niku	Tokyo Traders	97.00
Original Frankfurter grüne Soße	Plutzer Lebensmittelgroßmärkte AG	13.00
Rhönbräu Klosterbier	Plutzer Lebensmittelgroßmärkte AG	7.75
Camembert Pierrot	Gai pâturage	34.00

### **Productos agotados:**

**SELECT**

inv.inventory\_id, inv.product\_id, pro.product\_name, inv.product\_sold,  
inv.product\_available

**FROM**

inventory **AS** inv

**JOIN**



products AS pro

ON

inv.product\_id = pro.product\_id

ORDER BY product\_available ASC;

	inventory_id integer	product_id integer	product_name character varying (255)	product_sold integer	product_available integer
1	1017	17	Alice Mutton	331	0
2	1002	2	Chang	341	0
3	1073	73	Röd Kaviar	45	0
4	1072	72	Mozzarella di Giovanni	270	0
5	1050	50	Valkoinen suklaa	70	0
6	1039	39	Chartreuse verte	266	0
7	1036	36	Inlagd Sill	198	0
8	1011	11	Queso Cabrales	182	0
9	1014	14	Tofu	152	1

### Productos menos vendidos:

SELECT

inv.inventory\_id, inv.product\_id, pro.product\_name, inv.product\_sold,  
inv.product\_available

FROM

inventory AS inv

JOIN

products AS pro

ON

inv.product\_id = pro.product\_id

ORDER BY product\_sold ASC;

	inventory_id integer	product_id integer	product_name character varying (255)	product_sold integer	product_available integer
1	1067	67	Laughing Lumberjack Lager	5	45
2	1045	45	Røgede sild	15	85
3	1022	22	Gustaf's Knäckebröd	18	27
4	1009	9	Mishi Kobe Niku	20	42
5	1007	7	Uncle Bob's Organic Dried Pears	25	30
6	1015	15	Genen Shouyu	25	20
7	1012	12	Queso Manchego La Pastora	27	53
8	1006	6	Grandma's Boysenberry Spread	36	44

### Productos más vendidos:

#### SELECT

inv.inventory\_id, inv.product\_id, pro.product\_name, inv.product\_sold,  
inv.product\_available

#### FROM

inventory AS inv

#### JOIN

products AS pro

#### ON

inv.product\_id = pro.product\_id

ORDER BY product\_sold DESC;

	inventory_id integer	product_id integer	product_name character varying (255)	product_sold integer	product_available integer
1	1031	31	Gorgonzola Telino	458	42
2	1060	60	Camembert Pierrot	430	70
3	1035	35	Steeleye Stout	369	131
4	1059	59	Raclette Courdavault	346	154
5	1002	2	Chang	341	0
6	1016	16	Pavlova	338	62
7	1071	71	Fløtemysost	336	164
8	1017	17	Alice Mutton	331	0
9	1062	62	Tarte au sucre	325	175

#### 4. Cuarta consulta:

*Comparación entre las ventas de la tienda física contra las ventas de la tienda online.*

##### **Requerimiento:**

“Ana: En cuanto a los procesos, realizamos ventas diarias en nuestras tiendas físicas y a través de nuestra tienda en línea. Además, llevamos un registro de los clientes que compran nuestros productos y su información personal, como nombre, edad y ubicación.”

##### **Comparando venta de tienda física o tienda online.**

###### **SELECT**

st.store\_id, st.store\_type,  
**SUM**(s.totalsales) **AS** totalsales

###### **FROM**

stores **st**

###### **INNER JOIN**

salesdetails **sd**

**ON** st.store\_id = sd.store\_id

###### **INNER JOIN**

fact\_sales **s**

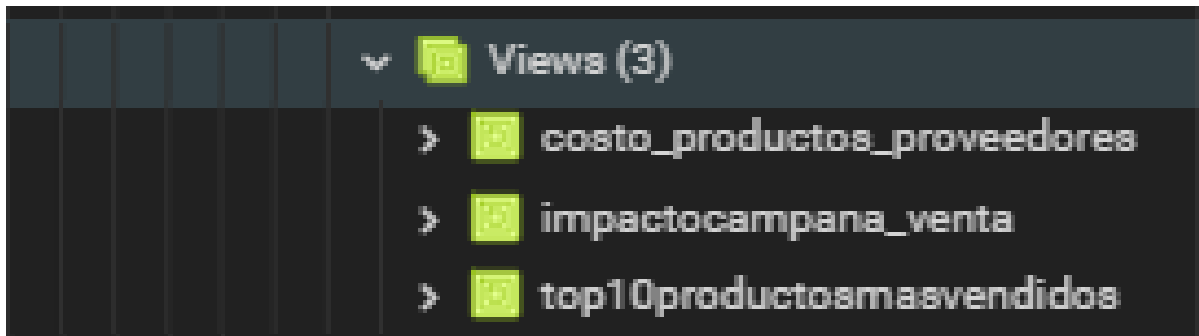
**ON** sd.sales\_id = s.sales\_id

**GROUP BY** st.store\_id, st.store\_type;

	store_id [PK] integer	store_type character varying (255)	totalsales numeric
1	2	online store	267942.85
2	1	physical store	972171.85

## Vistas

---



### VISTA 1: Top 10 productos más vendidos.

**CREATE VIEW** top10productosmasvendidos **AS**

**SELECT**

cu.customer\_name,sl.customer\_id,cu.customer\_age,  
cu.customer\_gender,cu.customer\_city,

**SUM** (sl.totalsales) **AS** VENTATOTAL

**FROM** fact\_sales sl

**INNER JOIN**

customers cu

**ON**

sl.customer\_id = cu.customer\_id

**GROUP BY** cu.customer\_name,sl.customer\_id,cu.customer\_age,  
cu.customer\_gender,cu.customer\_city

**ORDER BY** VENTATOTAL **DESC**

**LIMIT** 10;

**SELECT \* FROM** top10productosmasvendidos;

**VISTA 2: El impacto de las campañas promocionales en las ventas.**

**CREATE VIEW** impactocampana\_venta **AS**

**SELECT**

s.campaign\_id,

**SUM** (totalsales) **AS** total, cam.startdate **AS** Fecha\_inico\_campana,  
cam.finishdate **AS** Fecha\_final\_campana

**FROM**

fact\_sales **AS** s

**JOIN**

campaigns **AS** cam

**ON**

s.campaign\_id = cam.campaign\_id

**WHERE**

salesdate

**BETWEEN** '2022-01-01' AND '2023-09-11'

**GROUP BY**

s.campaign\_id, cam.startdate, cam.finishdate

**ORDER BY** total **DESC**;

**SELECT \* FROM** impactocampana\_venta;

### **VISTA 3: Relación de costo, productos y proveedores.**

```
CREATE VIEW costo_productos_proveedores AS

SELECT
    p.product_name,sp.supplier_name,p.product_price

FROM
    products p

INNER JOIN
    suppliers sp

ON
    p.supplier_id = sp.supplier_id

GROUP BY
    p.product_name,sp.supplier_name,p.product_price;

SELECT * FROM costo_productos_proveedores;
```