# Data Analysis

# Modulo 1

Introducción a base de datos. Data-analysis-21-12

# TecnoMark

Fernando Castillo

**Entregable** 



#### **FUENTES DB**

Las bases de datos se obtuvieron de una recoleta de información de una empresa del sector privado.

NOTA: Se modifico la información con datos ficticios en todas las bases de datos.

#### **PROBLEMATICA**

Se crea una nueva reforma al outsourcing en el año 2021 que afecta directamente a las empresas que contratan empleados eventuales. Esta reforma provoca que diversas empresas contraten directamente a sus empleados eventuales o los despidan. Para las empresas con giro de outsourcing, genera términos de contrato con clientes que deciden contratar a su personal.

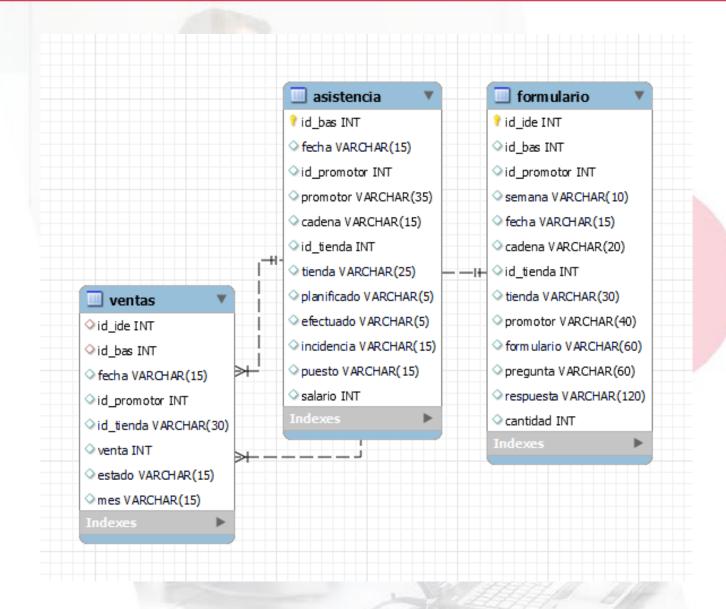
Pandemia global COVID 19 que afecta a diversos giros comerciales. Varias empresas piden a sus empleados no salir a trabajar, lo cual causa un impacto en PDV.

### ESTRUCTURA DE LAS BASES DE DATOS.

Se obtuvieron 3 bases de datos diferentes:

- Asistencia: Contiene la información de la asistencia de todo el personal en PDV.
- Formulario: Contiene la información que el personal recolecta en PDV a través de una aplicación.
- Ventas: Contiene la información de ventas que genera cada empleado en PDV.

# DIAGRAMA RELACIONAL SQL



### **GLOSARIO BASE ASISTENCIA**

ID_BAS	int
FECHA	varchar
ID_PROMOTOR	int
PROMOTOR	varchar
CADENA	varchar
ID_TIENDA	int
TIENDA	varchar
PLANIFICADO	varchar
EFECTUADO	varchar
INCIDENCIA	varchar
PUESTO	varchar
SALARIO	int

ID BAS: Número identificador de cada registro de la base de datos.

FECHA: La fecha registrada en la aplicación del día de la asistencia del personal en PDV.

ID PROMOTOR: ID identificador de cada promotor (empleado).

PROMOTOR: Nombre del empelado que esta registrando la información desde la aplicación.

CADENA: Es el dato identificador de la cadena a la que pertenece la tienda que se esta trabajando. Por ejemplo, la tienda SC ESCOBEDO NORTE pertenece a la cadena SUPERCENTER.

ID TIENDA: ID identificador de cada tienda.

TIENDA: El nombre que tiene asignado cada PDV o tienda. Este dato proviene directamente del cliente, ya que puede ser diferente para cada uno.

PLANIFICADO: Este dato nos indica si la tienda estaba planificado en el plan de trabajo para ser visitado por el promotor y/o empleado en el día especifado.

EFECTUADO: Es el dato que nos indica si el empleado y/o promotor realizo la visita a la tienda que tenía programada el día en cuestión.

INCIDENCIA: En caso de que el empleado y/o promotor no haya visitado la tienda que tenía planificada, debe ingresar el motivo por el cual no asistió. Este dato se representa en la base de datos como Incidencia.

PUESTO: Es el puesto que tiene el empleado dentro de la empresa.

SALARIO: Representa el salario mensual que se le paga al empelado.

### GLOSARIO BASE FORMULARIO

ID_IDE	int
ID_BAS	int
ID_PROMOTOR	int
SEMANA	varchar
FECHA	varchar
CADENA	varchar
ID_TIENDA	int
TIENDA	varchar
PROMOTOR	varchar
FORMUALARIO	varchar
PREGUNTA	varchar
RESPUESTA	varchar
CANTIDAD	int

ID\_IDE: Número identificador de cada registro de la base de datos. Este digito sólo identifica los registros de la base de formularios y ventas.

ID\_BAS: Número identificador de cada registro de la base de datos.

ID\_PROMOTOR: ID identificador de cada promotor (empleado)

SEMANA: Es el dato que indica la semana calendario a la que pertenece el registro.

FECHA: La fecha registrada en la aplicación del día de la asistencia del personal en PDV.

CADENA: Es el dato identificador de la cadena a la que pertenece la tienda que se esta trabajando. Por ejemplo, la tienda SC ESCOBEDO NORTE pertenece a la cadena SUPERCENTER.

ID\_TIENDA: ID identificador de cada tienda.

TIENDA: El nombre que tiene asignado cada PDV o tienda. Este dato proviene directamente del cliente, ya que puede ser diferente para cada uno.

PROMOTOR: Nombre del empelado que esta registrando la información desde la aplicación.

FORMULARIO: Es el nombre identificador del formulario que se esta respondiendo. Este dato nos ayuda a identificar preguntas especificas que pudieran estar repetidas en otros formularios.

PREGUNTA: Es la pregunta que se almacena en el formulario y que esta respondiendo el promotor en PDV.

RESPUESTA: El dato obtenido de la pregunta realizada.

CANTIDAD: Es el número de veces que se respondió la pregunta. Por la estructura de la base, generalmente siempre es 1.

#### **GLOSARIO BASE VENTAS**

ID_IDE	int
ID_BAS	int
FECHA	varchar
ID_PROMOTOR	int
PROMOTOR	varchar
ID_TIENDA	int
TIENDA	varchar
VENTA	int
ESTADO	varchar
MES	varchar

ID\_IDE: Número identificador de cada registro de la base de datos. Este digito sólo identifica los registros de la base de formularios y ventas.

ID\_BAS: Número identificador de cada registro de la base de datos.

FECHA: La fecha registrada en la aplicación del día de la asistencia del personal en PDV.

ID PROMOTOR: ID identificador de cada promotor (empleado)

PROMOTOR: Nombre del empelado que esta registrando la información desde la aplicación.

ID\_TIENDA: ID identificador de cada tienda.

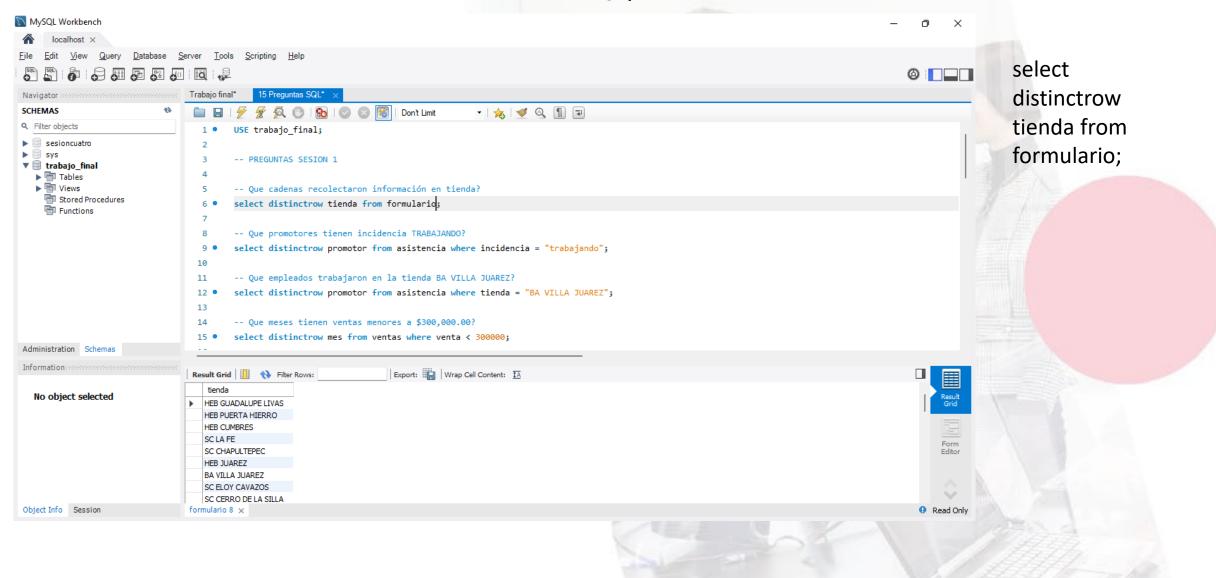
TIENDA: El nombre que tiene asignado cada PDV o tienda. Este dato proviene directamente del cliente, ya que puede ser diferente para cada uno.

VENTA: Es la venta registrada por el promotor en la aplicación de cada PDV de un rango de tiempo específico.

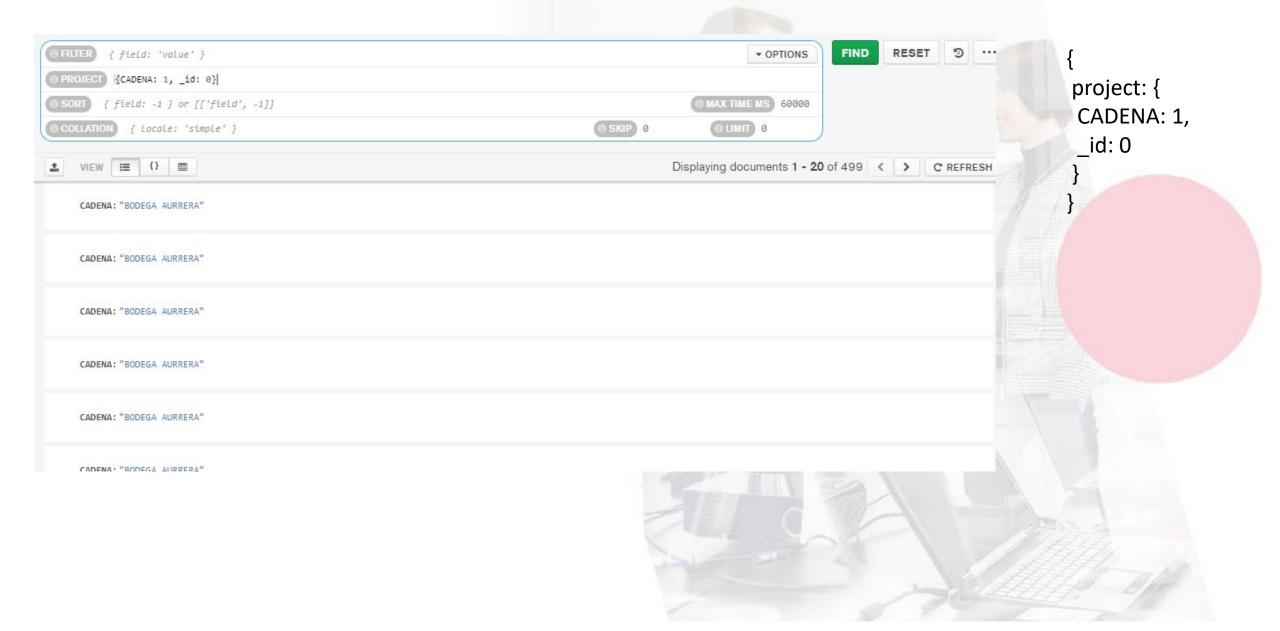
ESTADO: Estado de la republica a la que pertenece la tienda que esta trabajando y donde se recolecto la información.

MES: El mes al que esta asignada la venta registrada en la base.

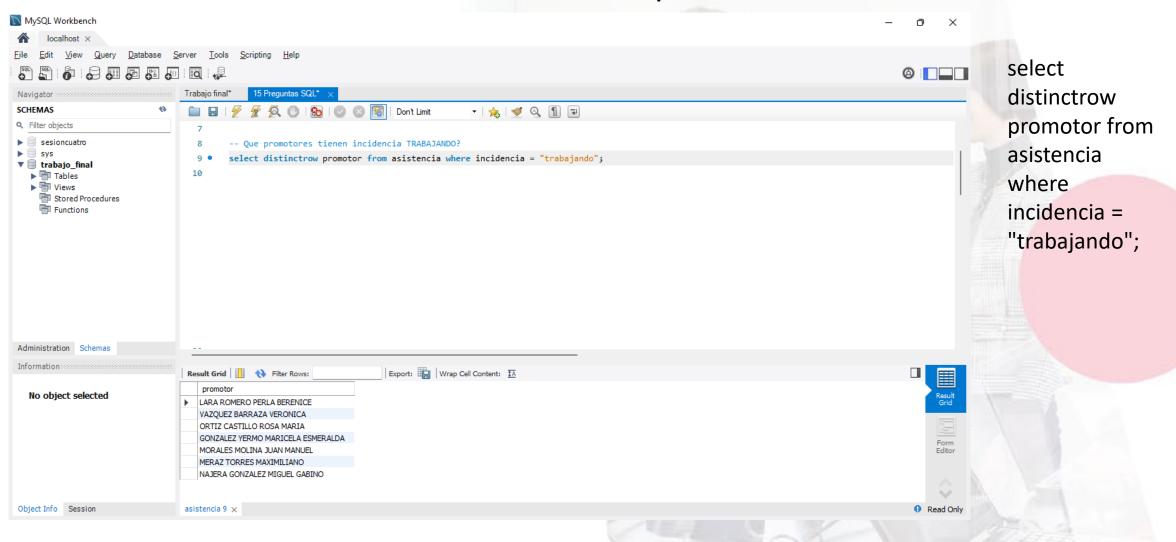
#### ¿Que cadenas recolectaron información en tienda?



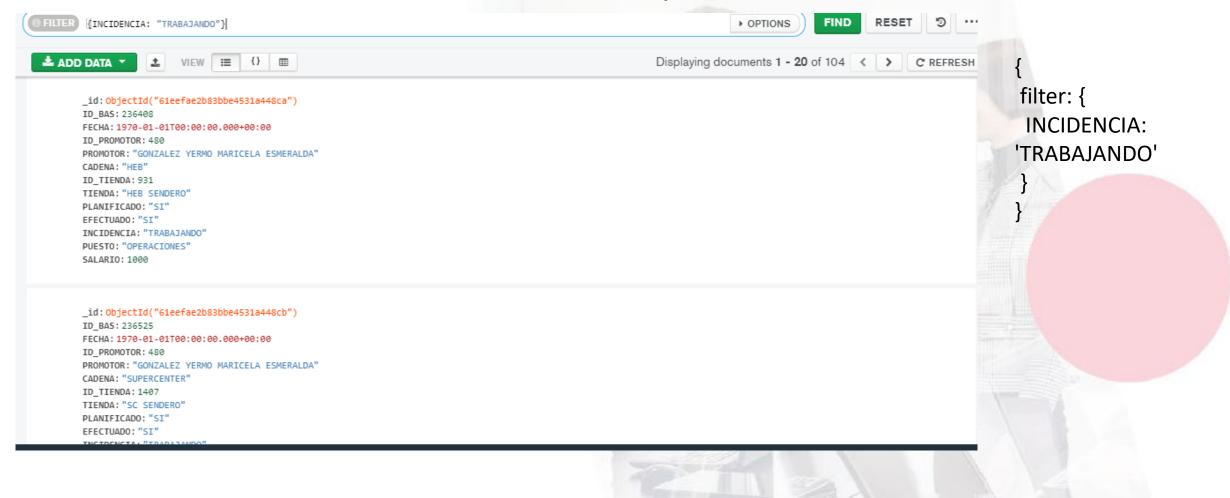
#### ¿Que cadenas recolectaron información en tienda?



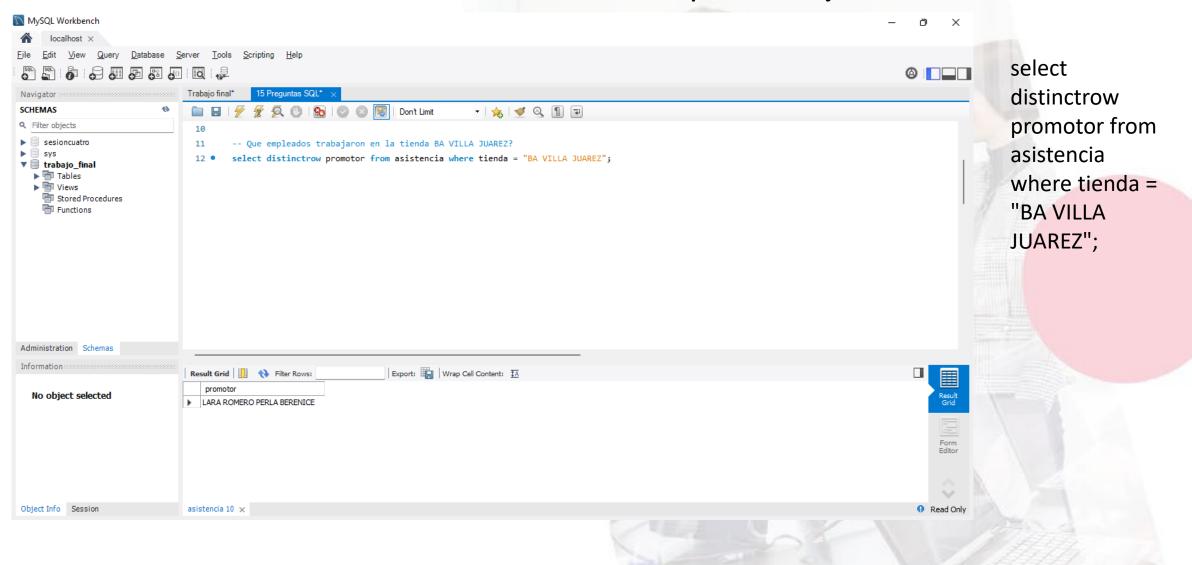
#### ¿Que promotores tienen incidencia TRABAJANDO?



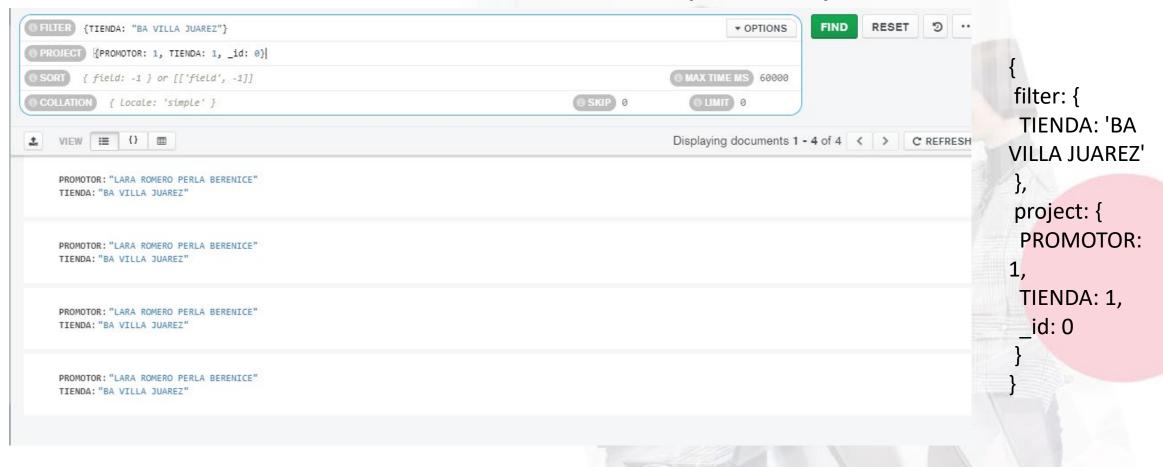
#### ¿Que promotores tienen incidencia TRABAJANDO?



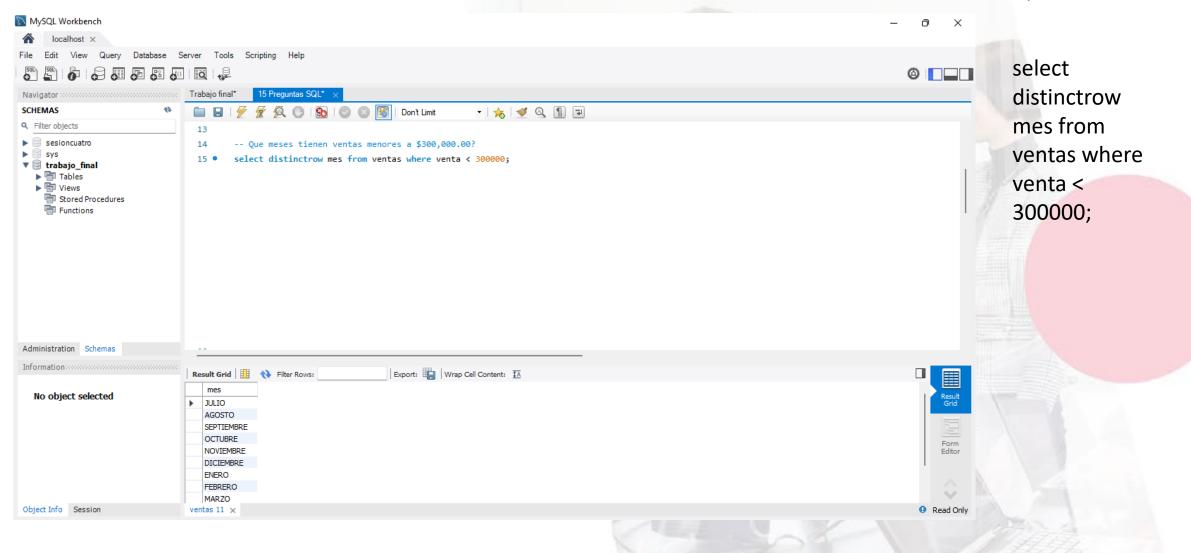
#### ¿Que empleados trabajaron en la tienda BA VILLA JUAREZ?



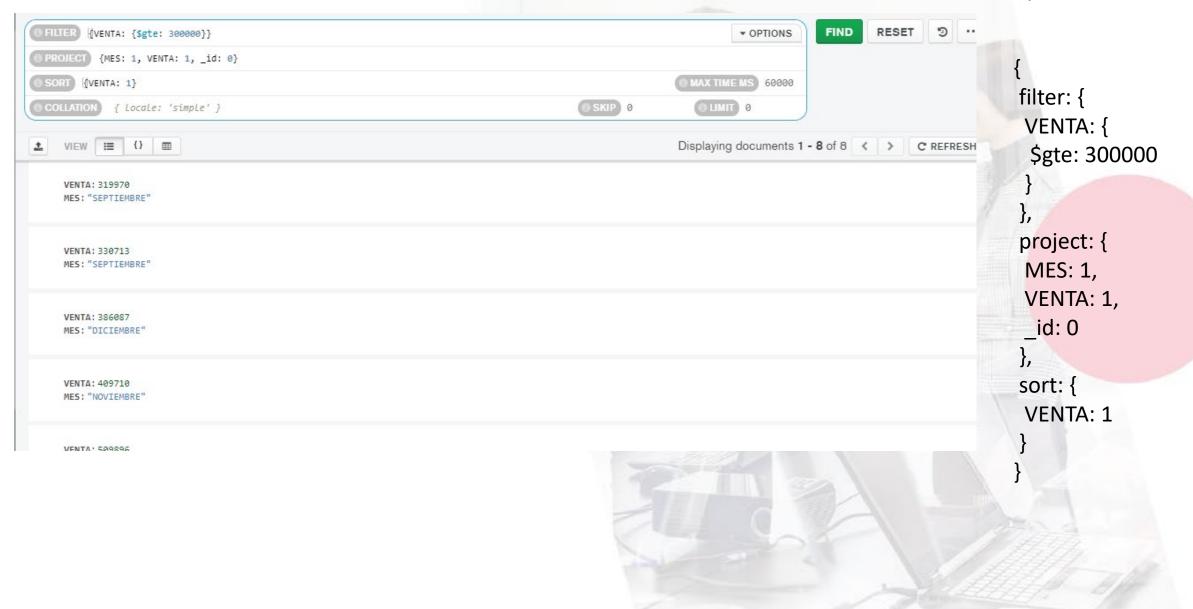
#### ¿Que empleados trabajaron en la tienda BA VILLA JUAREZ?



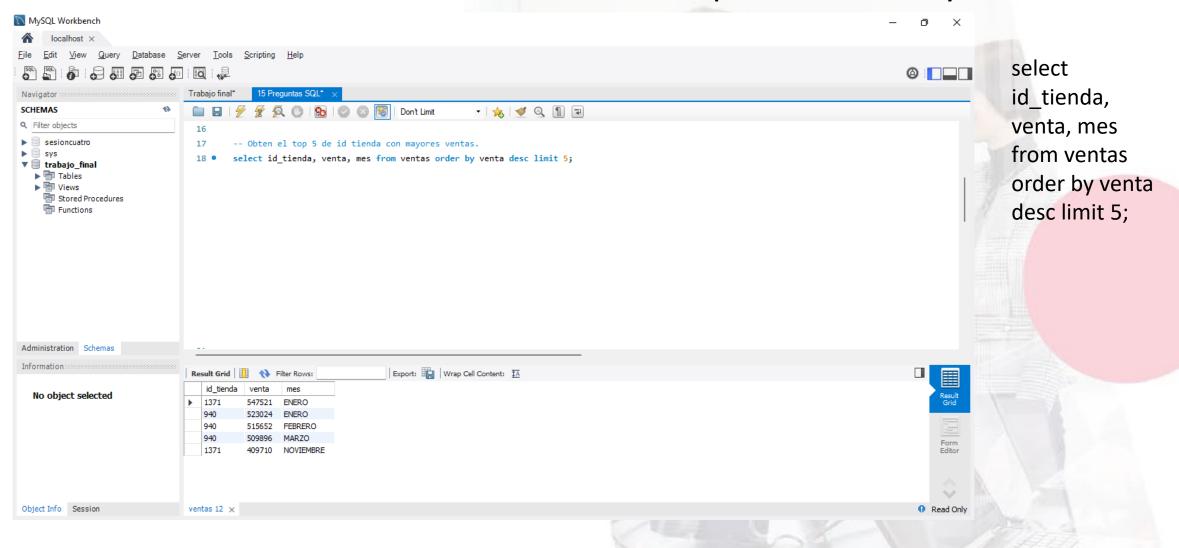
#### ¿Que meses tienen ventas menores a \$300,000.00?



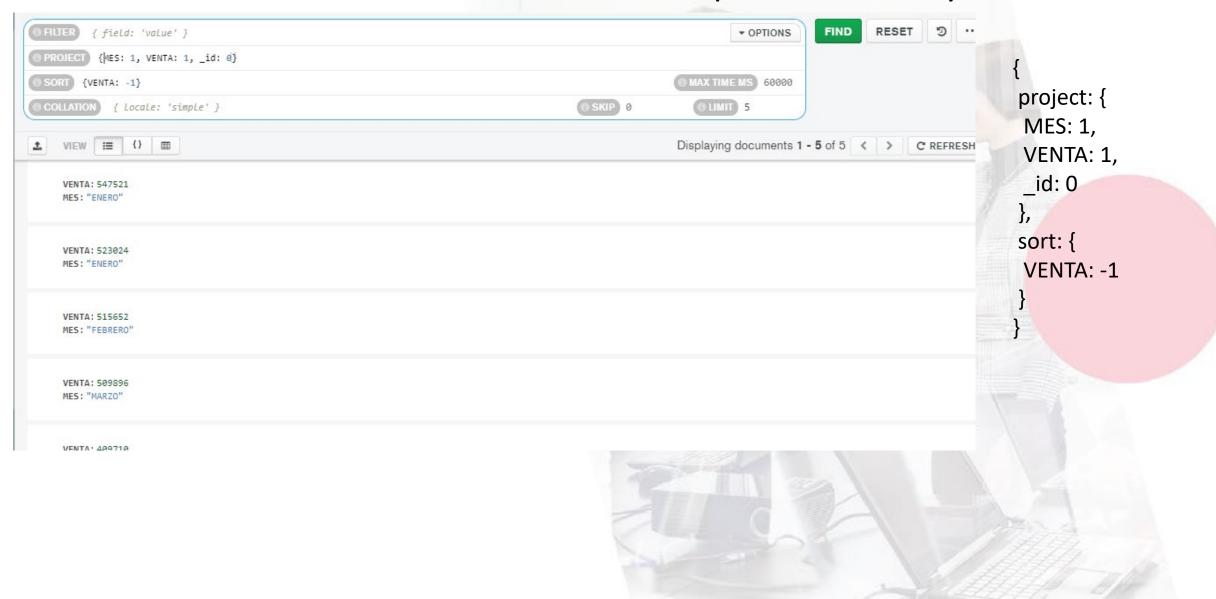
#### ¿Que meses tienen ventas menores a \$300,000.00?



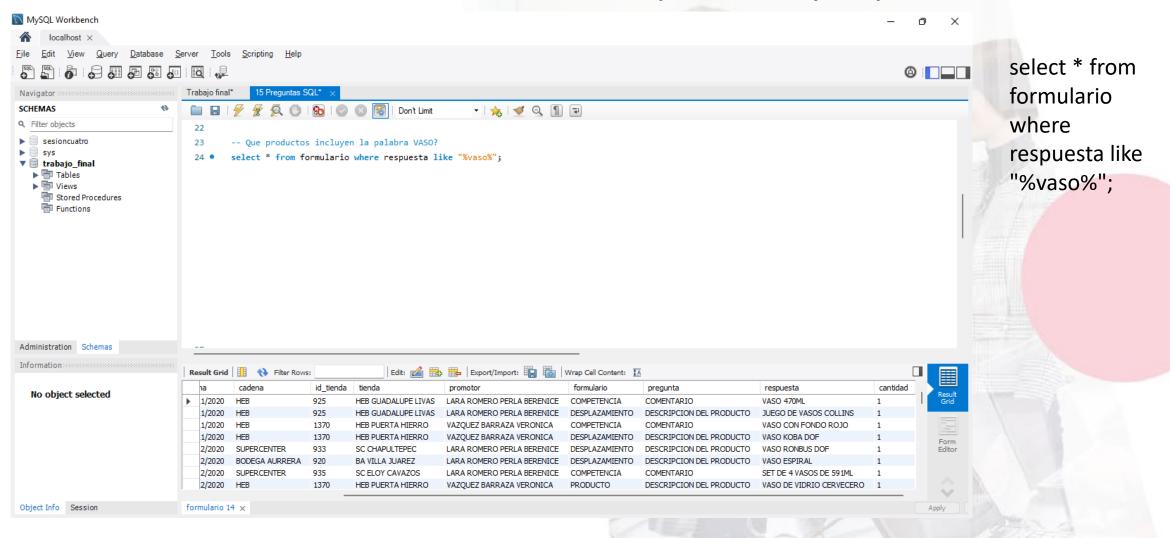
#### Obtén el top 5 de id tienda con mayores ventas



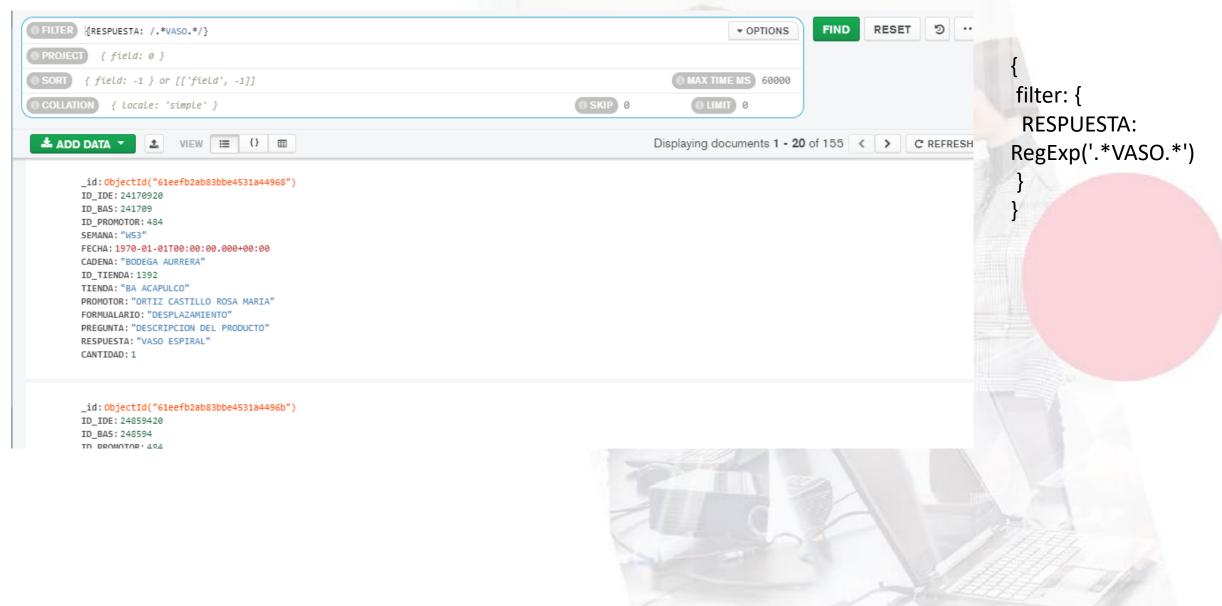
#### Obtén el top 5 de id tienda con mayores ventas



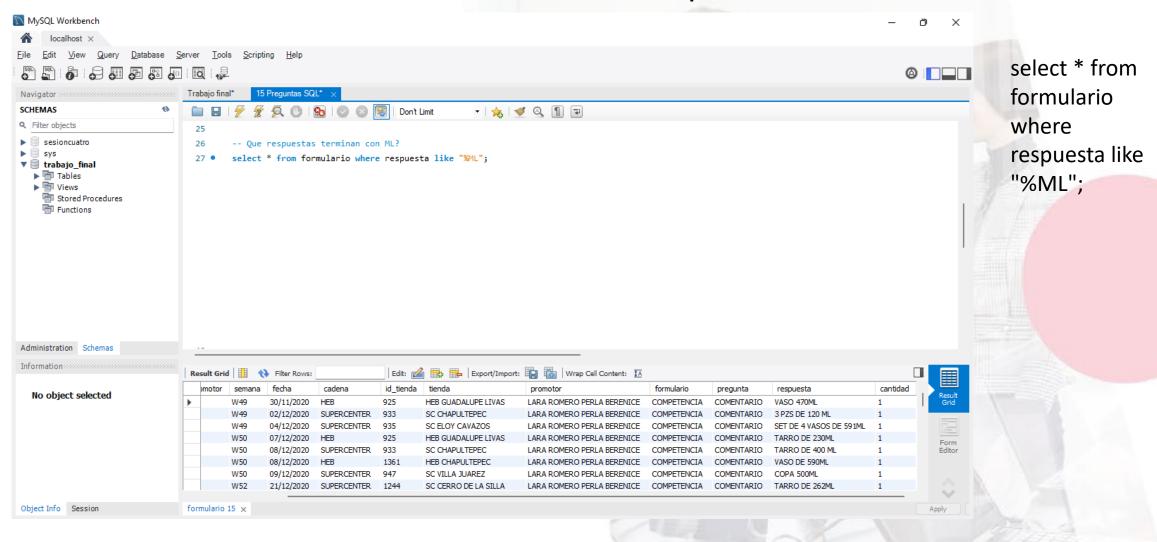
#### ¿Que productos incluyen la palabra VASO?



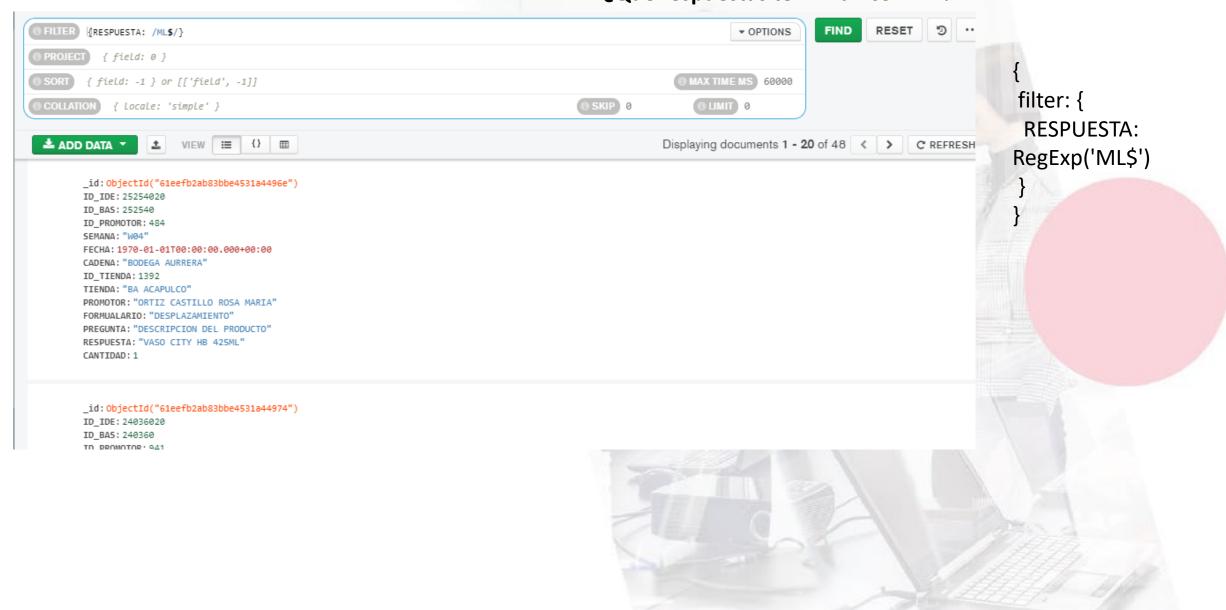
#### ¿Que productos incluyen la palabra VASO?



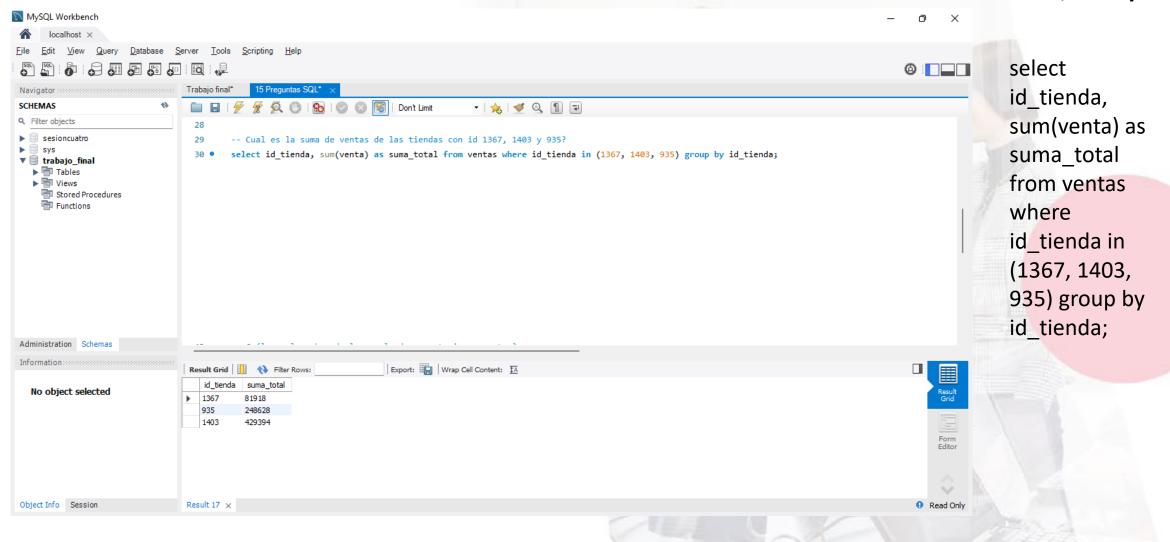
#### ¿Que respuestas terminan con ML?



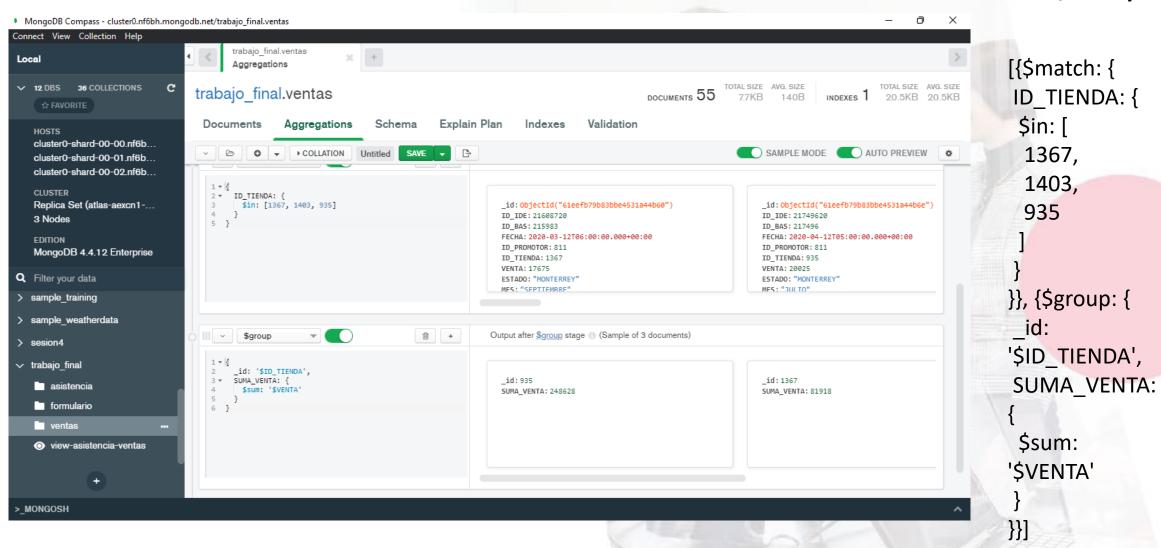
#### ¿Que respuestas terminan con ML?



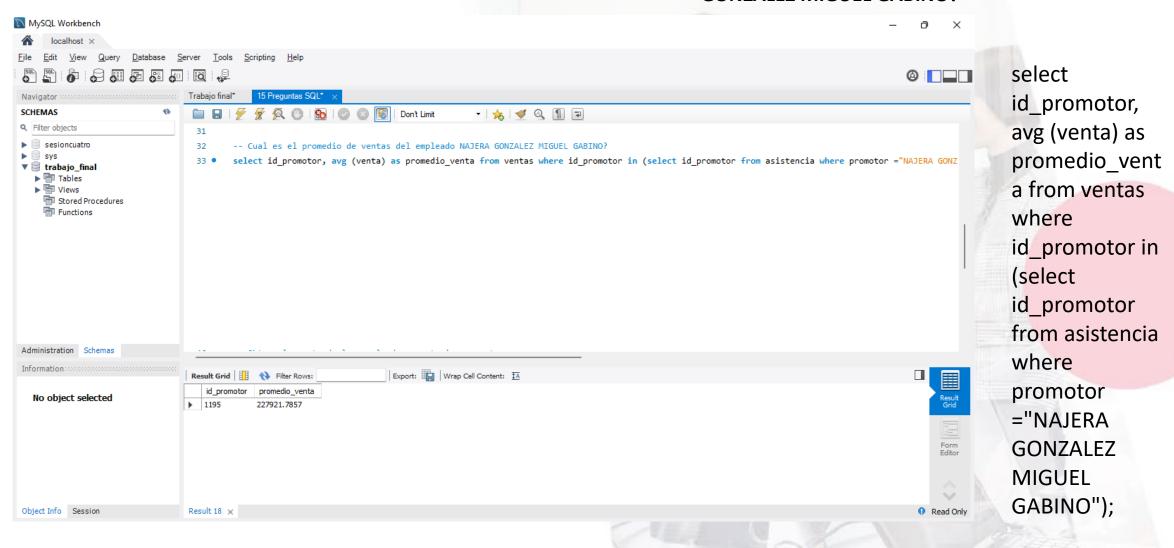
¿Cual es la suma de ventas de las tiendas con id 1367, 1403 y 935?



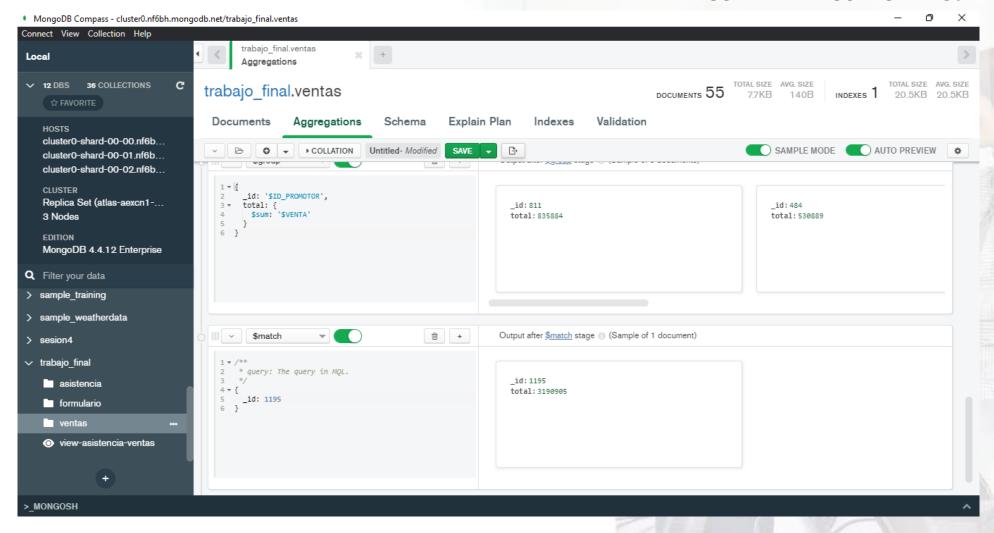
¿Cual es la suma de ventas de las tiendas con id 1367, 1403 y 935?



# ¿Cual es el promedio de ventas del empleado NAJERA GONZALEZ MIGUEL GABINO?

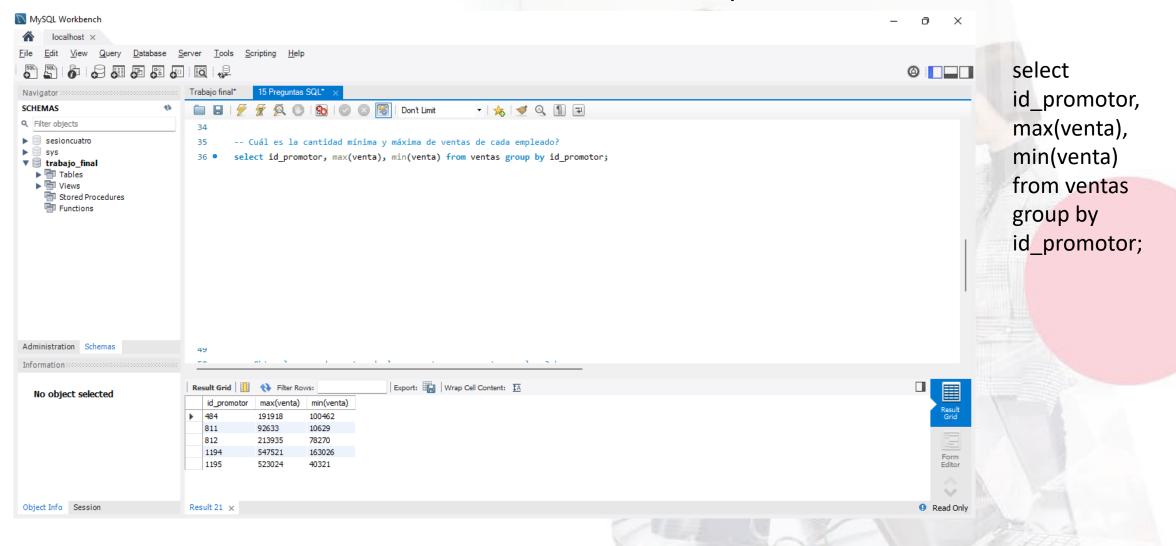


# ¿Cual es el promedio de ventas del empleado NAJERA GONZALEZ MIGUEL GABINO?

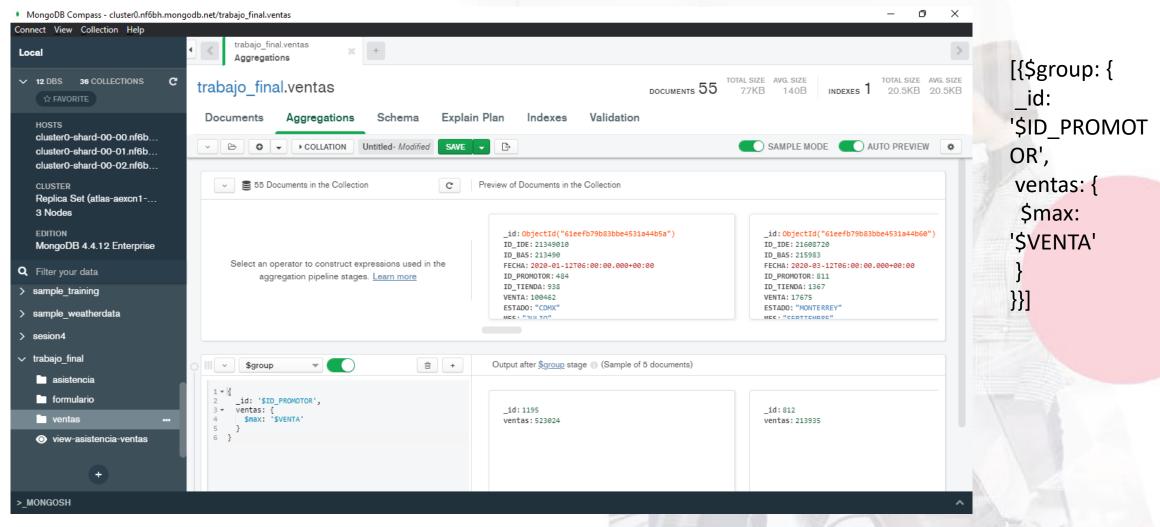


```
[{$lookup: { from:
'asistencia', localField:
'ID BAS', foreignField:
'ID BAS', as:
'nombre promotor'}},
{$addFields: {
promotor: {
$arrayElemAt: [
'$nombre promotor',
0 ] }}}, {$addFields: {
dato buscado:
'$promotor.PROMOTO
R'}}, {$project: { id: 0,
PROMOTOR:
'$dato buscado',
ID PROMOTOR: 1,
VENTA: 1}}, {$group: {
id: '$ID PROMOTOR',
total: { $sum:
'$VENTA' }}}, {$match:
{ id: 1195}}]
```

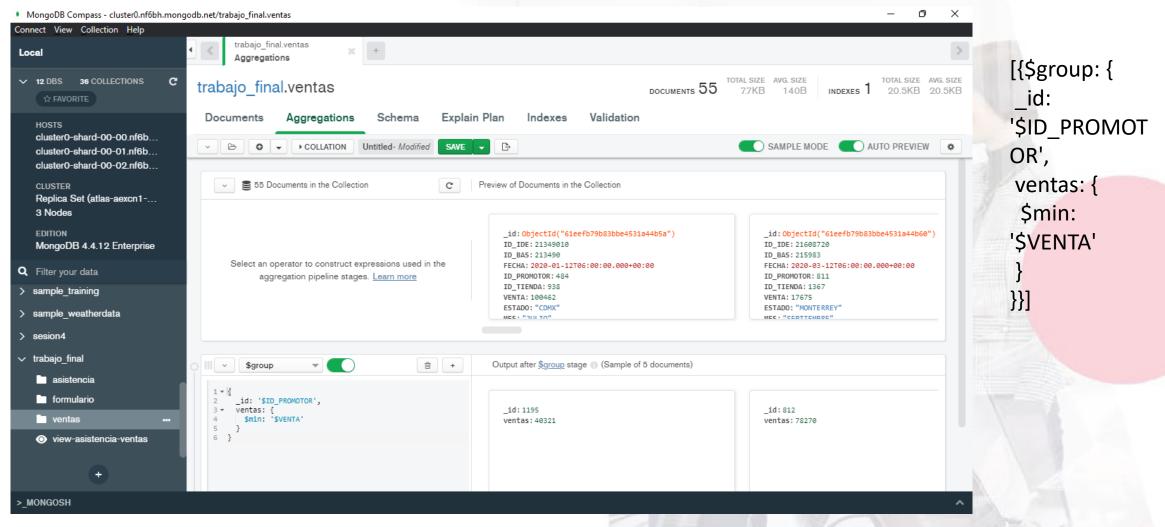
¿Cuál es la cantidad mínima y máxima de ventas de cada empleado?



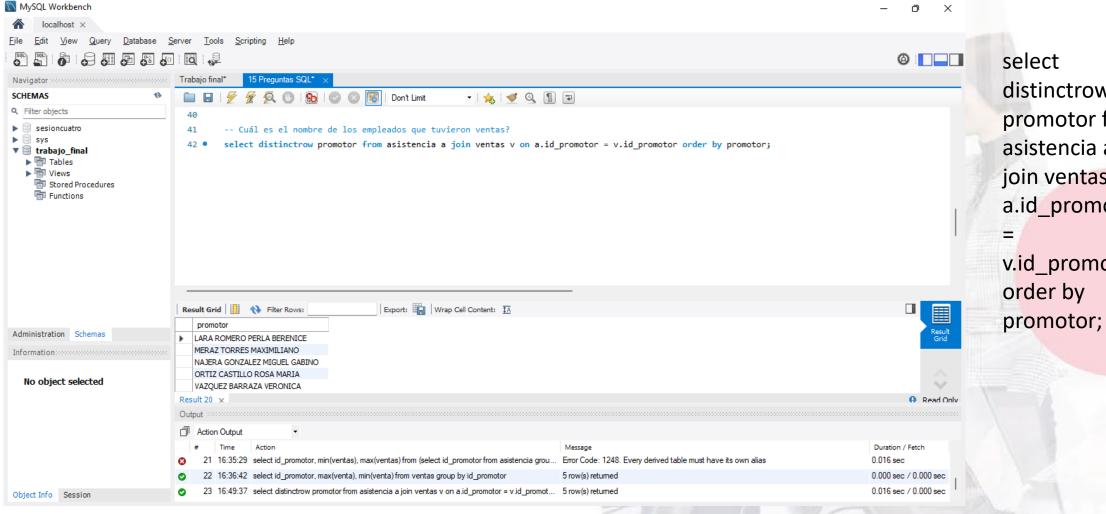
¿Cuál es la cantidad mínima y máxima de ventas de cada empleado?



¿Cuál es la cantidad mínima y máxima de ventas de cada empleado?



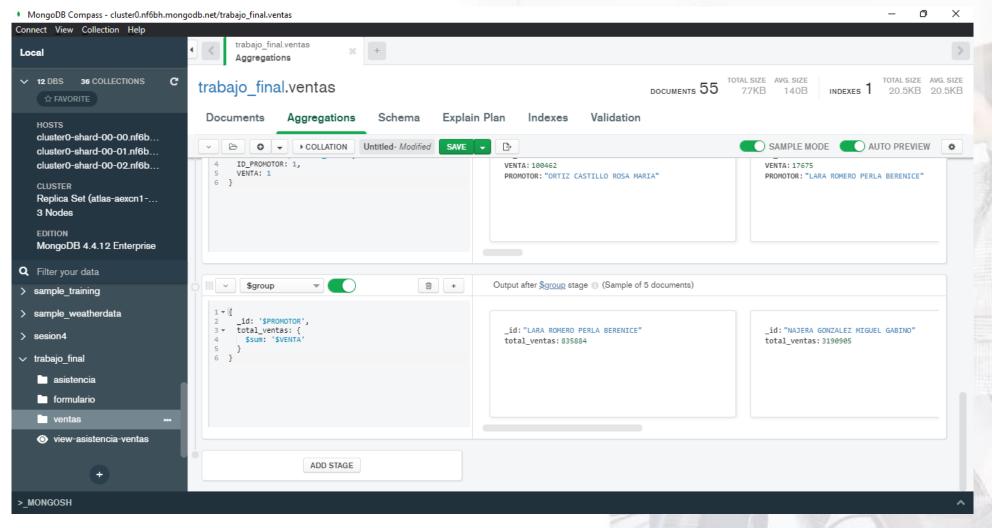
#### ¿Cuál es el nombre de los empleados que tuvieron ventas?



distinctrow promotor from asistencia a join ventas v on a.id\_promotor v.id\_promotor

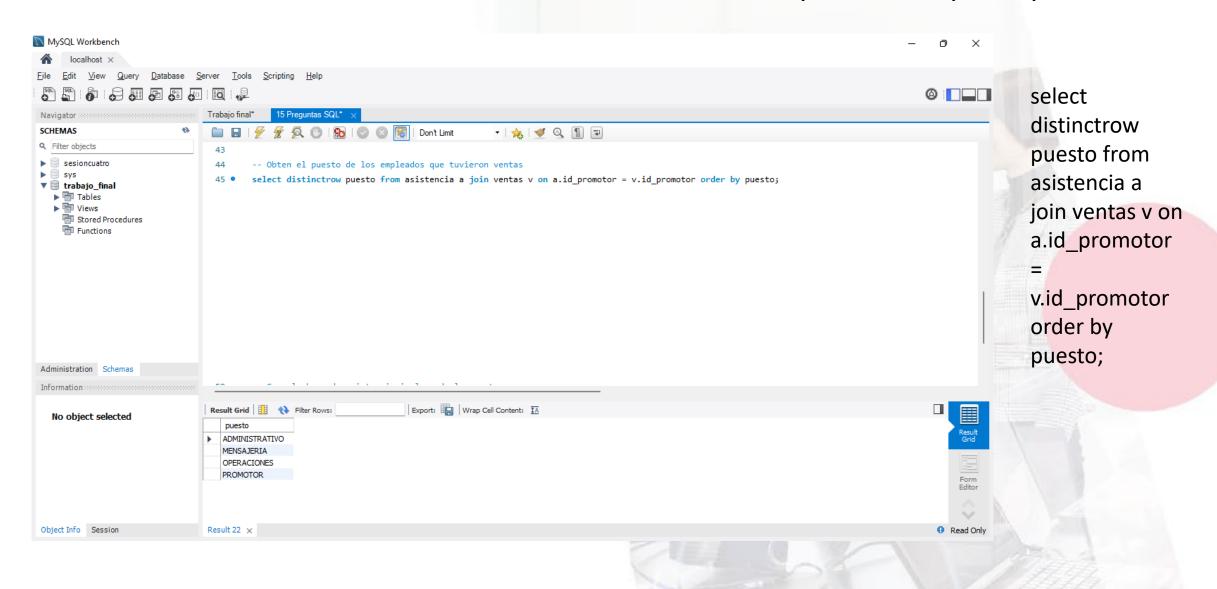
¿Cuál es el nombre de los empleados que tuvieron

ventas?

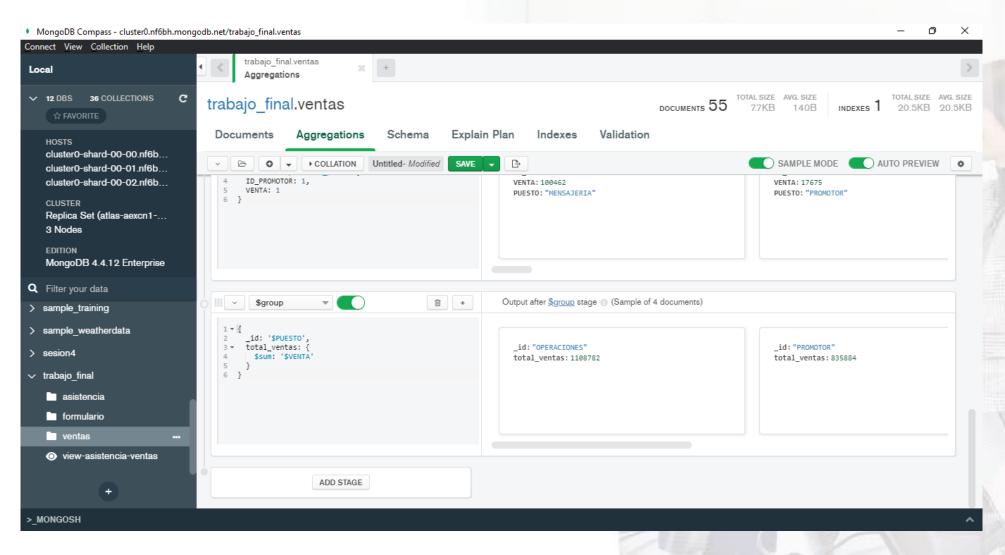


```
[{$lookup: {
from: 'asistencia',
localField: 'ID BAS',
foreignField: 'ID BAS',
as: 'nombre promotor'
}}, {$addFields: {
promotor: {
 $arrayElemAt: [
  '$nombre promotor',
}}, {$addFields: {
promotor venta:
'$promotor.PROMOTOR'
}}, {$project: {
_id: 0,
PROMOTOR:
'$promotor venta',
ID PROMOTOR: 1,
VENTA: 1
}}, {$group: {
 id: '$PROMOTOR',
total ventas: {
 $sum: '$VENTA'
```

#### Obtén el puesto de los empleados que tuvieron ventas

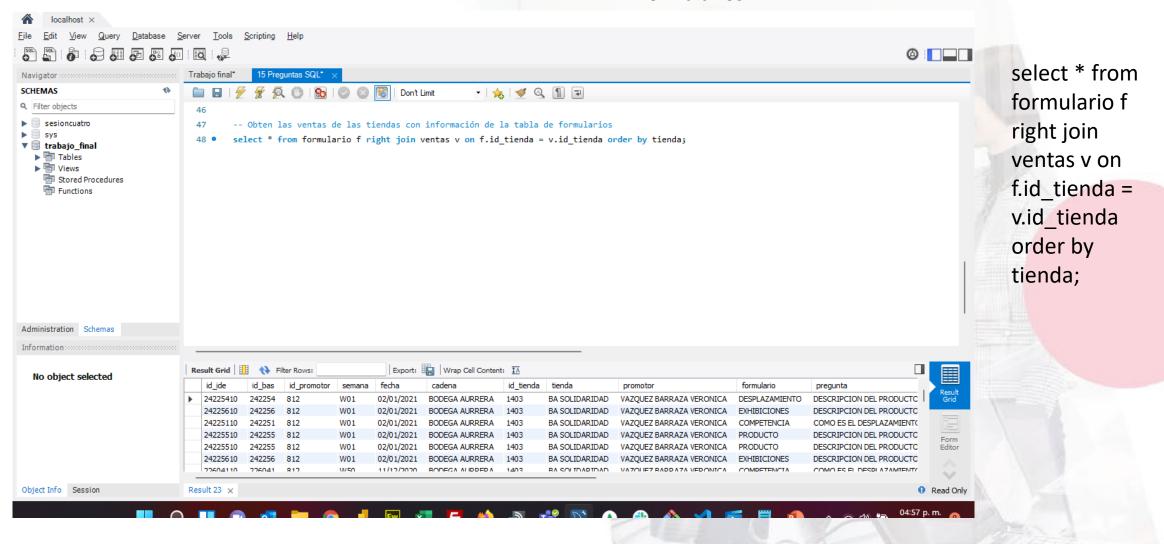


#### Obtén el puesto de los empleados que tuvieron ventas

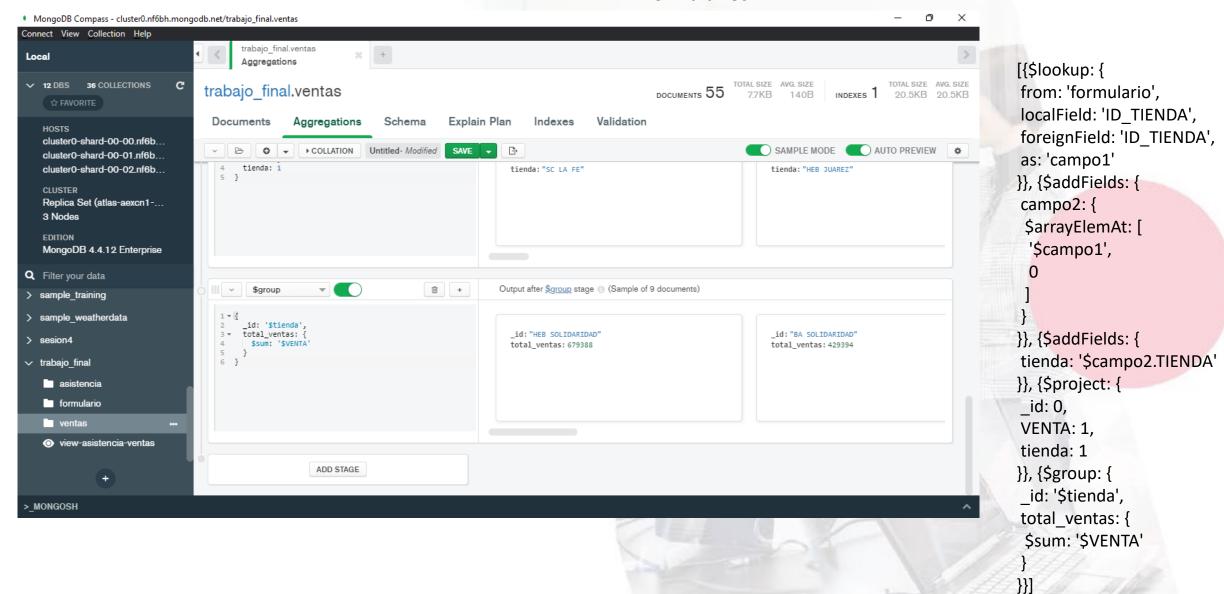


```
[{$lookup: {
from: 'asistencia',
localField: 'ID BAS',
foreignField: 'ID BAS',
as: 'nombre promotor'
}}, {$addFields: {
promotor: {
 $arrayElemAt: [
  '$nombre promotor',
}}, {$addFields: {
promotor puesto:
'$promotor.PUESTO'
}}, {$project: {
id: 0,
PUESTO:
'$promotor puesto',
ID PROMOTOR: 1,
VENTA: 1
}}, {$group: {
_id: '$PUESTO',
total ventas: {
 $sum: '$VENTA'
```

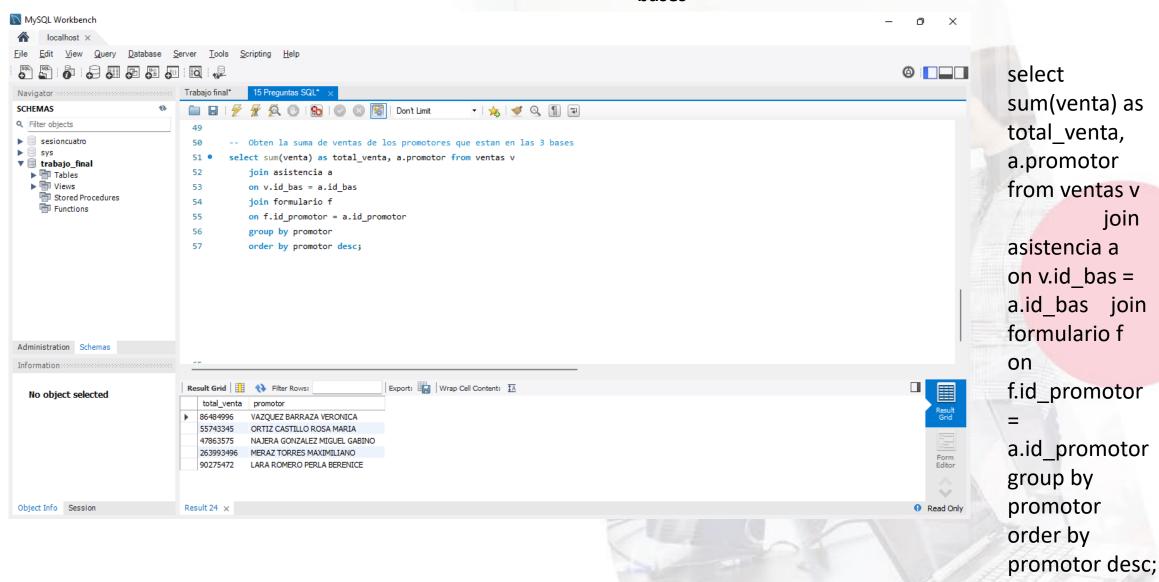
# Obtén las ventas de las tiendas con información de la tabla de formularios



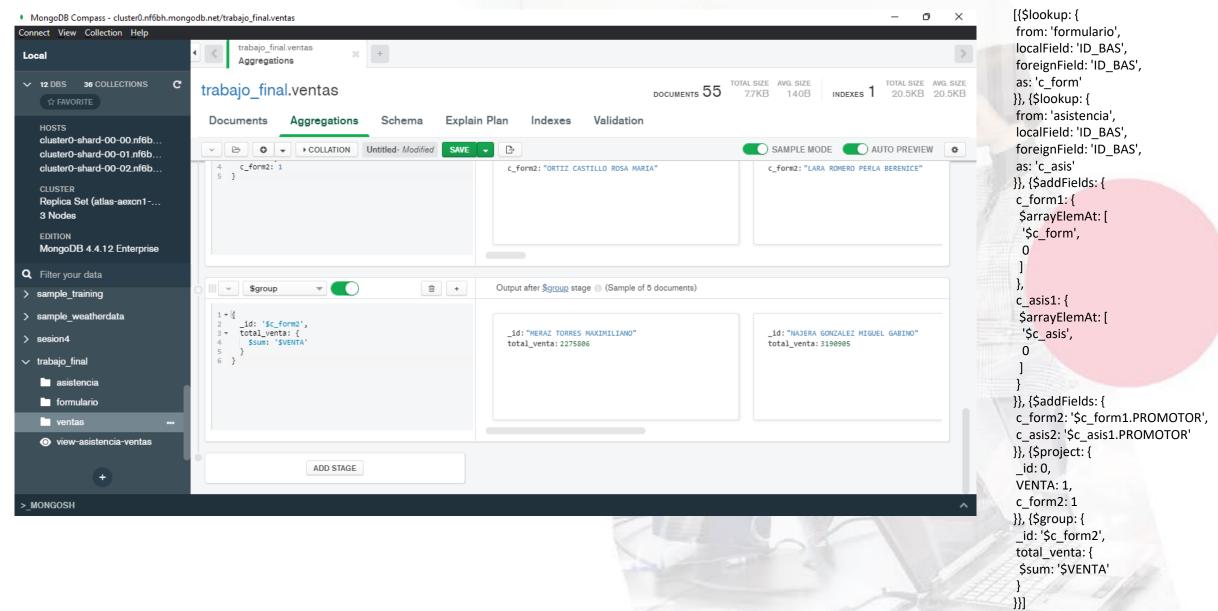
# Obtén las ventas de las tiendas con información de la tabla de formularios



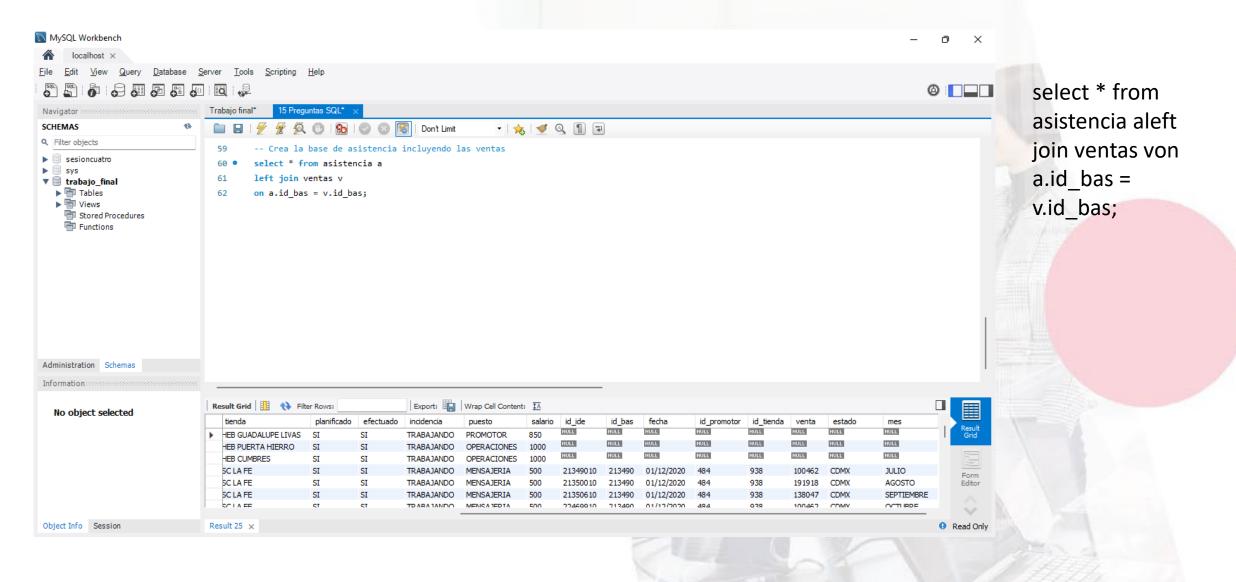
# Obtén la suma de ventas de los promotores que están en las 3 bases



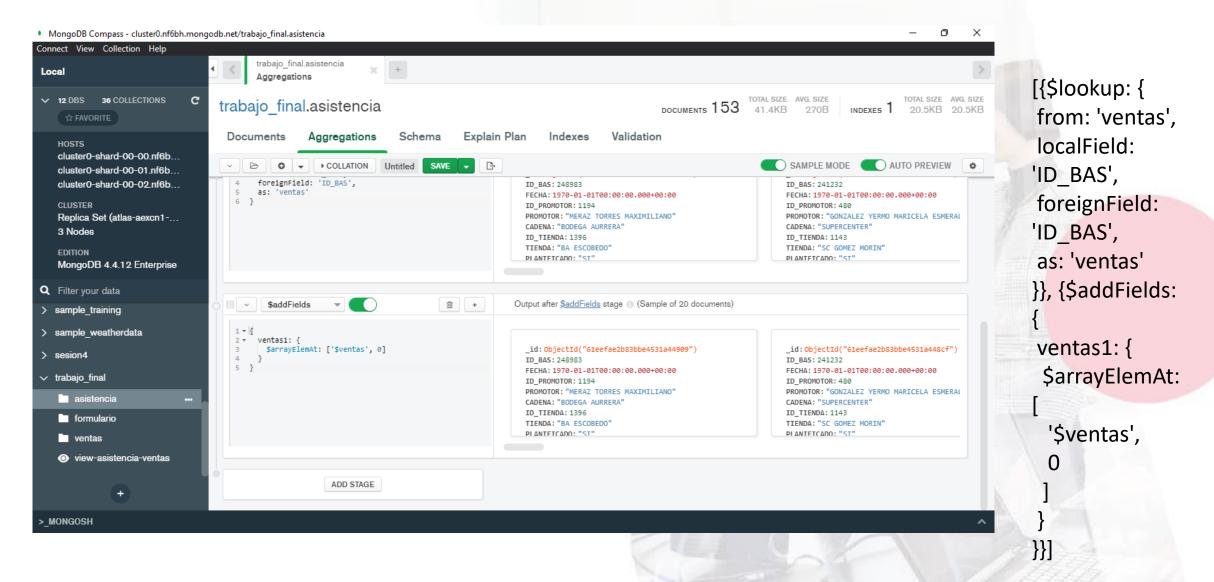
# Obtén la suma de ventas de los promotores que están en las 3 bases



#### Crea la base de asistencia incluyendo las ventas

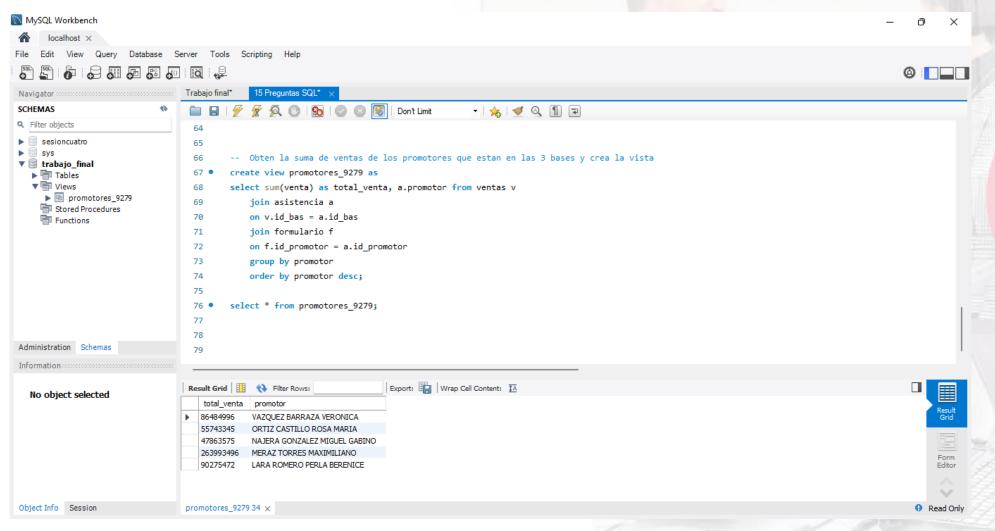


#### Crea la base de asistencia incluyendo las ventas



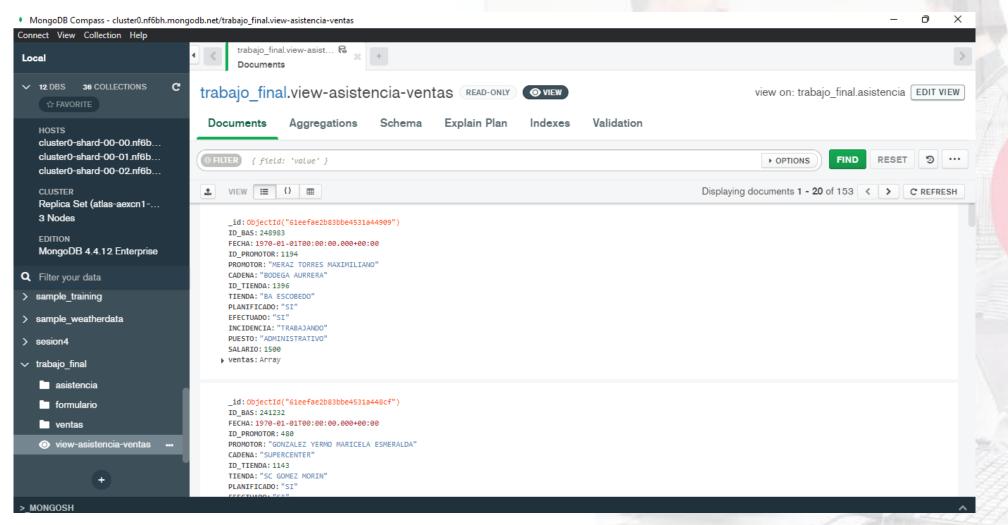
## **VISTA SQL**

Empelados que tuvieron asistencia, recolecta de información y ventas registradas



#### **VISTA MONGODB**

Empelados que tuvieron asistencia, recolecta de información y ventas registradas



### **CONCLUSIONES**

- Haciendo un ejercicio con datos ficticios, observe que se pueden obtener información relevante para mejorar las métricas que se miden y la efectividad en los PDV.
- Haciendo un comparativo entre SQL y MONGO al responder las mismas preguntas, me parece más sencillo obtener los mismos resultados con menor cantidad de pasos en SQL.
- Me gustaría conocer más de MONGO, ya que me parece una herramienta muy potente y que puede funcionar mucho mejor con otro tipo de bases.