



CURSO DE DATA ANALYTICS: PROYECTO FINAL

AUTOMOBILE SALES DATA



PRESENTADO POR: ALEXANDER GARCIA

FEBRERO 22 DE 2024

Contenido

1.DESCRIPCION DE LA BASE DE DATOS.....	1
2.HIPOTESIS.....	1
3.LISTADO DE TABLAS	1
1. TABLA PRINCIPAL: VENTAS	1
2. TABLA: CLIENTE	2
3. TABLA: LOCALIZACION	2
4. TABLA: PRODUCTO	2
5. TABLA: ENVIO	2
4.DIAGRAMA ENTIDAD-RELACION 1	3
5.LISTADO DE COLUMNAS POR TABLA	4
6.TRANSFORMACIONES REALIZADAS	6
7.DIAGRAMA ENTIDAD RELACION 2	8
8.MEDIDAS CALCULADAS REALIZADAS	9
A.Medidas Calculadas Realizadas	10
B.Implementación de Medidas Calculadas	10
9.COLUMNAS CALCULADAS	17
10. PARAMETROS GENERADADOS.....	18
11.TABLAS REALIZADAS	19
12.CONCLUSIONES	20
13. FUTURAS LINEAS	21

1.DESCRIPCION DE LA BASE DE DATOS

En el presente estudio se puede observar las ventas generadas por la compañía de Autos y motocicletas a nivel mundial. Se pueden evidenciar la fecha de la venta, así como el nombre de los compradores, cantidad de artículos y valor de la misma venta. Todo esto para poder mejorar la productividad de la empresa y poder mejorar la calidad del servicio y ventas para todos los clientes.

2.HIPOTESIS

Identificar los clientes y ciudades donde más se generan ventas, es una información crucial para la compañía en general. Asumimos hipotéticamente que tanto las tiendas de Usa como el mismo país y algunas tiendas de Europa central generar más ventas. Así como dependiendo de la región algunos productos son más vendidos que otros, como pueden ser las motocicletas. Lo importante es tener claro con esta base de datos todas estas variables para mejorar las ventas y servicios de la empresa.

3.LISTADO DE TABLAS

1. TABLA PRINCIPAL: VENTAS

Contiene el Deal size, Quantity Ordered, Price each, Sales, Order Date, Days Since last order, Contact Last name y Contact First name.

* PK: ID_Orden

*FK: ID_Envio

*FK: ID_Producto

*FK: ID_Cliente

*FK: ID_Localizacion

2. TABLA: CLIENTE

Contiene el Customer Name, Phone y Address.

* PK: ID_Cliente

*FK: ID_Localizacion

3. TABLA: LOCALIZACION

Contiene el Pais, Ciudad y Continente.

* PK: ID_Localizacion

4. TABLA: PRODUCTO

Contiene el Product Line.

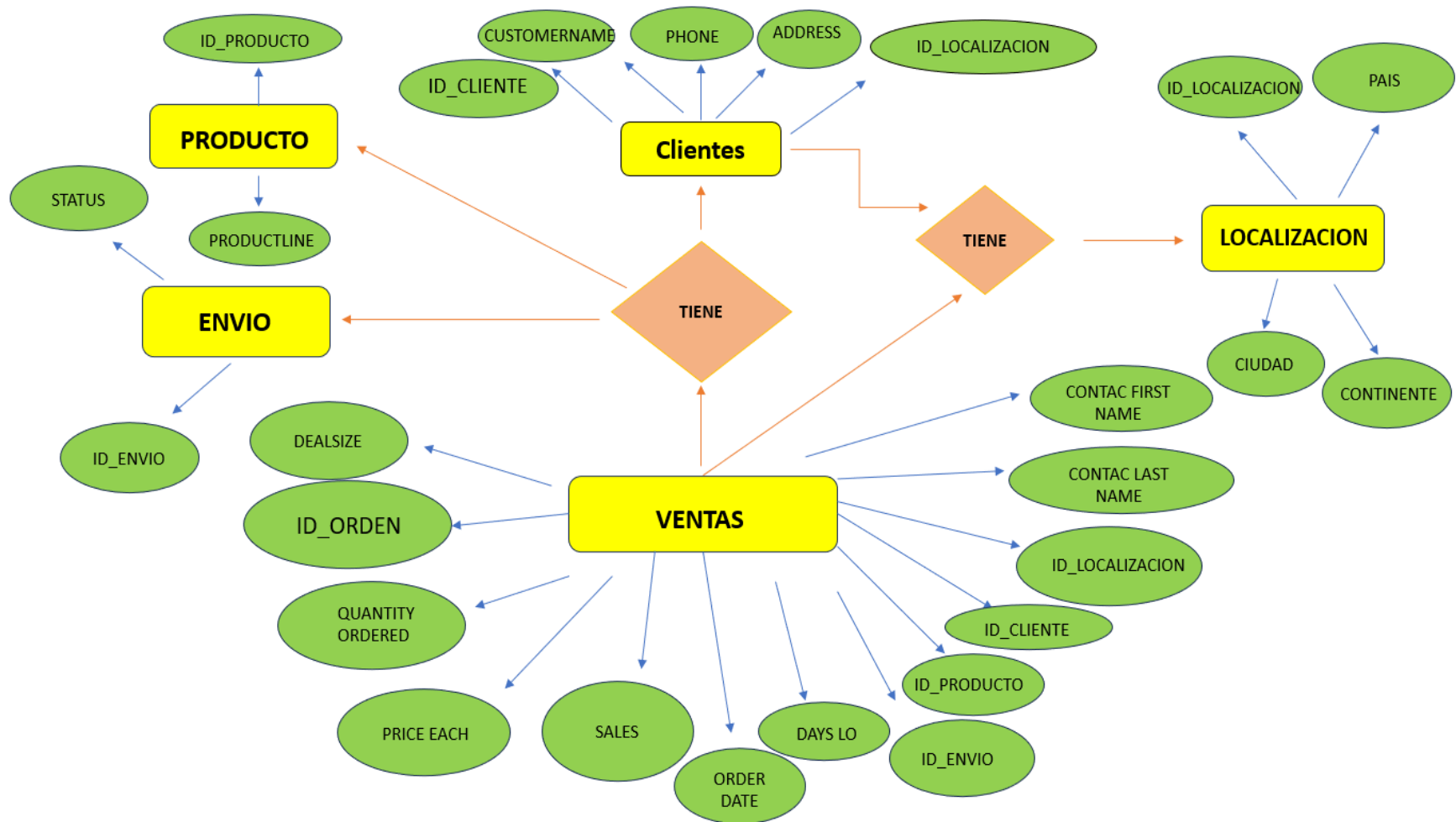
* PK: ID_Producto

5. TABLA: ENVIO

Contiene el Status.

* PK: ID_Envio

4.DIAGRAMA ENTIDAD-RELACION 1



5.LISTADO DE COLUMNAS POR TABLA

VENTAS		
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE
PK: ID_Orden	Int	PK-Index
FK: ID_Envio	Int	FK
FK: ID_Producto	Int	FK
FK: ID_Cliente	Int	FK
FK: ID_Localizacion	Int	FK
Deal size	Text	-
Quantity Ordered	Int	-
Price each	Decimal (18,2)	-
Sales	Decimal (18,2)	-
Order Date	Datetime	-
Days Since last order	Int	-
Contact Last name	Text	-
Contact First name	Text	-

CLIENTE		
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE
PK: ID_Cliente	Int	PK-Index
FK: ID_Localizacion	Int	FK
Customer Name	Text	-
Phone	Int	-
Address	Varchar	-

LOCALIZACION		
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE
PK: ID_Localizacion	Int	PK-Index
Pais	Text	FK
Ciudad	Text	-
Continente	Text	-

PRODUCTO		
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE
PK: ID_Producto	Int	PK-Index
Product Line	Text	-

ENVIO		
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE
PK: ID_Envio	Int	PK-Index
Status	Text	-

6.TRANSFORMACIONES REALIZADAS

TABLA VENTAS:

Let

```
Origen = Excel.Workbook(File.Contents("C:\Alexander\5. Analisis de datos\Data Analytics\Segunda entrega\Correccion2entrega\correccionentrega2.xlsx"), null, true), Ventas_Sheet = Origen{[Item="Ventas",Kind="Sheet"]}[Data],  
#"Encabezados promovidos" = Table.PromoteHeaders(Ventas_Sheet, [PromoteAllScalars=true]), #"Tipo cambiado" =  
Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_ORDEN", Int64.Type}, {"QUANTITY ORDERED", Int64.Type},  
{"PRICE EACH", type number}, {"SALES", type number}, {"ORDERDATE", type date}, {"DAYS_SINCE_LASTORDER", Int64.Type},  
{"ID_ENVIO", Int64.Type}, {"ID_PRODUCTO", Int64.Type}, {"ID_CLIENTE", Int64.Type}, {"ID_LOCALIZACION", Int64.Type},  
{"CONTACT LAST NAME", type text}, {"CONTACT FIRST NAME", type text}, {"DEALSIZE", type text}})
```

In

#"Tipo cambiado"

* Se genero una columna nueva: NOMBRE_COMPLETO_VENDEDOR = Ventas [CONTACT FIRST NAME] & " " & Ventas [CONTACT LAST NAME]

TABLA CLIENTE:

Let

```
Origen = Excel.Workbook(File.Contents("C:\Alexander\5. Analisis de datos\Data Analytics\Segunda entrega\Correccion2entrega\correccionentrega2.xlsx"), null, true), Cliente_Sheet = Origen{[Item="Cliente",Kind="Sheet"]}[Data],  
#"Encabezados promovidos" = Table.PromoteHeaders(Cliente_Sheet, [PromoteAllScalars=true]), #"Tipo cambiado" =  
Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_CLIENTE", Int64.Type}, {"CUSTOMERNAME", type text},  
{"PHONE", type any}, {"ADDRESSLINE1", type text}, {"ID_LOCALIZACION", Int64.Type}})
```

In

#"Tipo cambiado"

TABLA LOCALIZACION:

Let

```
Origen = Excel.Workbook(File.Contents("C:\Alexander\5. Analisis de datos\Data Analytics\Segunda entrega\Correccion2entrega\correccionentrega2.xlsx"), null, true), Localizacion_Sheet = Origen{[Item="Localizacion",Kind="Sheet"]}[Data], #"Encabezados promovidos" = Table.PromoteHeaders(Localizacion_Sheet, [PromoteAllScalars=true]), #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID LOCALIZACION", Int64.Type}, {"PAIS", type text}, {"CIUDAD ", type text}, {"CONTINENTE", type text}}), #"Columnas con nombre cambiado" = Table.RenameColumns(#"Tipo cambiado",{{"CIUDAD ", "CIUDAD"}})
```

In

#"Columnas con nombre cambiado"

TABLA PRODUCTO:

Let

```
Origen =Excel.Workbook(File.Contents("C:\Alexander\5. Analisis de datos\Data Analytics\Segunda entrega\Correccion2entrega\correccionentrega2.xlsx"), null, true), Producto_Sheet = Origen{[Item="Producto", Kind="Sheet"]}[Data], #"Encabezados promovidos" = Table.PromoteHeaders(Producto_Sheet, [PromoteAllScalars=true]), #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_PRODUCT", Int64.Type}, {"PRODUCT LINE", type text}})
```

In

#"Tipo cambiado"

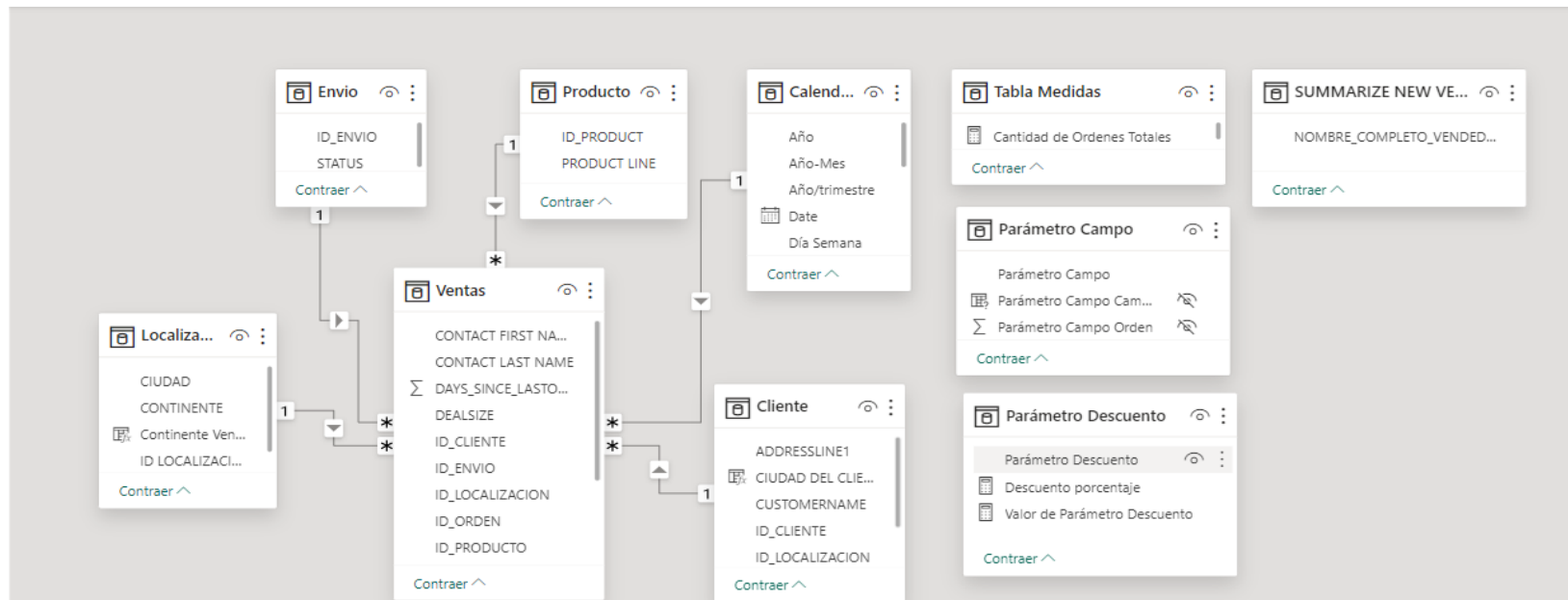
TABLA ENVIO:

Let

```
Origen = Excel.Workbook(File.Contents("C:\Alexander\5. Analisis de datos\Data Analytics\Segunda entrega\Correccion2entrega\correccionentrega2.xlsx"), null, true), Envio_Sheet = Origen[[Item="Envio",Kind="Sheet"]][Data],  
#"Encabezados promovidos" = Table.PromoteHeaders(Envio_Sheet, [PromoteAllScalars=true]), #"Tipo cambiado" =  
Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_ENVIO", Int64.Type}, {"STATUS", type text}})
```

In #"Tipo cambiado"

7.DIAGRAMA ENTIDAD RELACION 2



8.MEDIDAS CALCULADAS REALIZADAS

NUMERO	NOMBRE DE LA MEDIDA	EXPRESION
1.	Cantidad de Ordenes Totales	<code>SUM(Ventas[QUANTITY ORDERED])</code>
2.	MAX Valor de Ventas	<code>MAX(Ventas[SALES])</code>
3.	MIN Valor de Ventas	<code>MIN(Ventas[SALES])</code>
4.	Objetivo	<code>IF(sum(Ventas[QUANTITY ORDERED]) > 3000,</code> <code>"Cumple objetivo de mas de tres mil ordenes", "No cumple")</code>
5.	Promedio de Cantidad de Ordenes	<code>AVERAGE(Ventas[QUANTITY ORDERED])</code>
6.	Valor de Ventas Totales	<code>SUM(Ventas[SALES])</code>
7.	VAR	<code>VAR A = SUM(Ventas[SALES])</code> <code>VAR B = CALCULATE(SUM(Ventas[SALES]), Ventas[DEALSIZE] = "Medium")</code> <code>RETURN B/A</code>
8.	NEW HOY	<code>NOW()</code>
9.	NEW Total ordenes Canceladas	<code>CALCULATE(SUM(Ventas[QUANTITY ORDERED]), Envio[STATUS] = "Cancelled")</code>
10.	NEW Total ordenes Enviadas	<code>CALCULATE(SUM(Ventas[QUANTITY ORDERED]), Envio[STATUS] = "Shipped")</code>
11.	NEW VAR Ventas por las ciudades con mas ventas	<code>VAR Madrid = CALCULATE(SUM(Ventas[SALES]), Localizacion[CIUDAD] = "Madrid")</code> <code>VAR San_Rafael = CALCULATE(SUM(Ventas[SALES]), Localizacion[CIUDAD] = "San Rafael")</code> <code>VAR NYC = CALCULATE(SUM(Ventas[SALES]), Localizacion[CIUDAD] = "NYC")</code> <code>RETURN Madrid + San_Rafael + NYC</code>
12.	NEW Cantidad de Ciudades	<code>DISTINCTCOUNT(Localizacion[CIUDAD])</code>
13.	NEW Cantidad de países	<code>DISTINCTCOUNT(Localizacion[PAIS])</code>

A.Medidas Calculadas Realizadas

- Cantidad de Ordenes Totales
- Max Valor de Ventas
- MIN Valor de ventas
- Objetivo
- Promedio de Cantidad de Ordenes
- Valor de Ventas Totales
- VAR
- NEW HOY
- NEW Total de Ordenes Canceladas
- NEW Total de Ordenes Enviadas
- NEW VAR Ventas por las ciudades con más ventas
- NEW Cantidad de Ciudades
- NEW Cantidad de países

B.Implementación de Medidas Calculadas

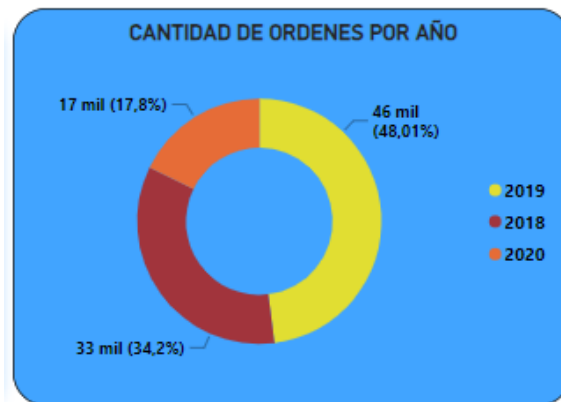
- Cantidad de Ordenes Realizadas

Cantidad de Ordenes Totales = $SUM(Ventas[QUANTITY ORDERED])$

Esta medida se creó para conocer la cantidad de ordenes totales
Generadas durante el estudio, es una de las medidas que más se utilizaron.

Con esta medida se crearon gran cantidad de gráficos, a continuación
mostramos los mas importantes:





- MAX Valor de Ventas

MAX Valor de Ventas = $\text{MAX}(\text{Ventas}[\text{SALES}])$

Esta medida se creó para conocer el valor máximo de las ventas.

Con esta medida se creó un Kpi del valor máximo de las ventas



- MIN Valor de Ventas

MIN Valor de Ventas = $\text{MIN}(\text{Ventas}[\text{SALES}])$

Esta medida se creó para conocer el valor mínimo de las ventas.

Con esta medida se creó un Kpi del valor mínimo de las ventas



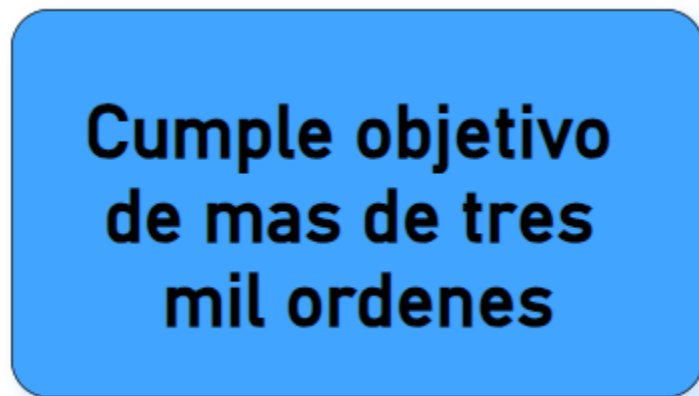
- Objetivo

Objetivo = $IF(\text{sum}(\text{Ventas}[\text{QUANTITY ORDERED}]) > 3000,$

"Cumple objetivo de más de tres mil órdenes", "No cumple")

Esta medida se creó para conocer si se cumplió el objetivo de más de 3000 órdenes generadas

Con esta medida se creó el siguiente Kpi :



- Promedio de Cantidad de Ordenes

Promedio de Cantidad de Ordenes = $AVERAGE(\text{Ventas}[\text{QUANTITY ORDERED}])$

Esta medida se creó para conocer el promedio de la cantidad de ordenes

Con esta medida se creó el Siguiete Kpi:

PROMEDIO DE CANTIDAD DE
ORDENES

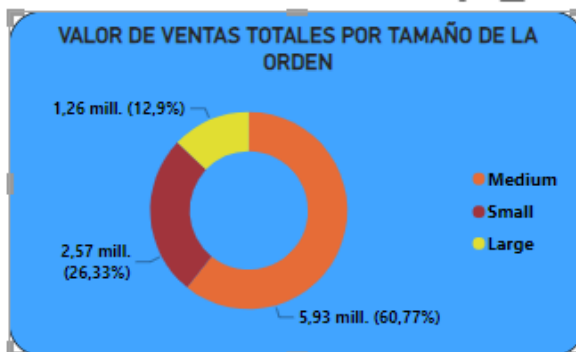
35,10

- Valor de ventas totales

Valor de Ventas Totales = $\text{SUM}(\text{Ventas}[\text{SALES}])$

Esta medida se creó para conocer el valor de las ventas totales
Generadas durante el estudio, es una de las medidas que más se utilizaron.

Con esta medida se crearon gran cantidad de gráficos, a continuación
mostramos los más importantes:



TOP 10 DE VENDEDORES CON MAS VENTAS

NOMBRE_COMPLETO_VENDEDO	OR	Valor de Ventas Totales
Diego Freyre		912.294,11
Valarie Nelson		654.858,06
Peter Ferguson		200.995,41
Jeff Young		197.736,94
Janine Labrune		180.124,90
Eric Natividad		172.989,68
Kwai Yu		164.069,44
Sue Frick		160.010,27
Victoria Ashworth		157.807,81
Anna O'Hara		153.996,13





- VAR

VAR = VAR A = SUM(Ventas[SALES]) VAR B =
CALCULATE(SUM(Ventas[SALES]),
Ventas[DEALSIZE] = "Medium") RETURN B/A

Esta medida se creó para conocer el porcentaje de valor de las ventas de talla media.

Con esta medida se creó el Siguiete Kpi:



- NEW HOY

NEW HOY = NOW()

Esta medida se creó para tener la fecha y hora actualizada

Con esta medida se creó la siguiente tarjeta:



- NEW Total ordenes Canceladas

NEW Total ordenes Canceladas = $\text{CALCULATE}(\text{SUM}(\text{Ventas}[\text{QUANTITY ORDERED}]), \text{Envio}[\text{STATUS}] = \text{"Cancelled"})$

Esta medida se creó para conocer la cantidad de ordenes que tuvieran la característica de “Cancelled” de la columna de estatus de la tabla envío

Con esta medida se creó el siguiente KPI:



- NEW Total ordenes Enviadas

NEW Total ordenes Enviadas = $\text{CALCULATE}(\text{SUM}(\text{Ventas}[\text{QUANTITY ORDERED}]), \text{Envio}[\text{STATUS}] = \text{"Shipped"})$

Esta medida se creó para conocer la cantidad de ordenes que tuvieran la característica de “Enviadas” de la columna de estatus de la tabla envío

Con esta medida se creó el siguiente KPI:



- NEW VAR Ventas por las ciudades con mas ventas

NEW VAR Ventas por las ciudades con mas ventas = $\text{VAR Madrid} = \text{CALCULATE}(\text{SUM}(\text{Ventas}[\text{SALES}]), \text{Localizacion}[\text{CIUDAD}] = \text{"Madrid"})$

```
VAR San_Rafael = CALCULATE(SUM(Ventas[SALES]),
Localizacion[CIUDAD] = "San Rafael") VAR NYC =
CALCULATE(SUM(Ventas[SALES]), Localizacion[CIUDAD] = "NYC")
RETURN Madrid + San_Rafael + NYC
```

Esta medida se creó para conocer la suma del valor de ventas de las 3 ciudades con más valores de ventas .

Con esta medida se creó el siguiente KPI:



- NEW Cantidad de Ciudades

NEW Cantidad de Ciudades = `DISTINCTCOUNT(Localizacion[CIUDAD])`

Esta medida se creó para conocer el numero de ciudades presentes en el estudio.

Con esta medida se creó el siguiente KPI:

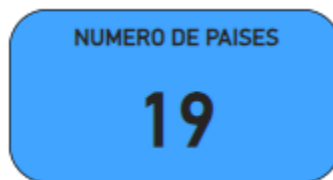


- NEW Cantidad de Paises

NEW Cantidad de Paises = `DISTINCTCOUNT(Localizacion[PAIS])`

Esta medida se creó para conocer el número de ciudades presentes en el estudio.

Con esta medida se creó el siguiente KPI:







9.COLUMNAS CALCULADAS

TABLA	COLUMNA	DESCRIPCION	EXPRESION
Cliente	País del Cliente	En esta columna trajimos la columna de país de la tabla localización	PAIS DEL CLIENTE = LOOKUPVALUE(Localizacion[PAIS],Localizacion[ID LOCALIZACION],Cliente[ID_LOCALIZACION])
Cliente	Ciudad del Cliente	En esta columna trajimos la columna de ciudad de la tabla localización	CIUDAD DEL CLIENTE = LOOKUPVALUE(Localizacion[CIUDAD],Localizacion[ID LOCALIZACION],Cliente[ID_LOCALIZACION])
Localización	Continente Ventas	Usamos un switch para asignar características a los continentes	Continente Ventas = SWITCH(TRUE(), (Localizacion[CONTINENTE] = "Europa"), "Primero en Ventas", (Localizacion[CONTINENTE] = "AMERICA"), "Segundo en Ventas", (Localizacion[CONTINENTE] = "Oceania"), "Tercero en Ventas", "Cuarto en Ventas")
Ventas	Nombre Completo Vendedor	Concatenamos la columna de nombre y apellidos, con espacio.	NOMBRE_COMPLETO_VENDEDOR = Ventas[CONTACT FIRST NAME] & " " & Ventas[CONTACT LAST NAME]

10. PARAMETROS GENERADOS

NUMERO	NOMBRE DEL PARAMETRO	EXPRESION
1.	Parámetro Campo	{ ("CONTINENTE", NAMEOF('Localizacion'[CONTINENTE]), 0),
		("PAIS", NAMEOF('Localizacion'[PAIS]), 1),
		("CIUDAD ", NAMEOF('Localizacion'[CIUDAD]), 2) }
2.	Parámetro Descuento	GENERATESERIES(0, 1, 0.1)
2.1	Descuento porcentaje	[Valor de Ventas Totales] - [Valor de Ventas Totales] * [Valor de Parámetro Descuento]

✓  Parámetro Descuento

- ☐  Descuento porce...
- ☐  Parámetro Descu...
- ☐  Valor de Parámet...

DESCUENTO PARA VENTAS

10 %

- ☐ 0 %
- ☒ 10 %
- ☐ 20 %
- ☐ 30 %
- ☐ 40 %
- ☐ 50 %
- ☐ 60 %

CUSTOMERNAME	DESCU	VENTAS
Alpha Cognac		Descuento porcentaje
Amica Models & Co.		63.439,60
Anna's Decorations, Ltd		84.705,53
Atelier graphique		138.596,52
Australian Collectables, Ltd	64.591,46	21.761,96
Australian Collectors, Co.	200.995,41	58.132,31
Australian Gift Network, Co	59.469,12	180.895,87
Auto Assoc. & Cie.	64.834,32	53.522,21
Auto Canal Petit	93.170,66	58.350,89
Auto-Moto Classics Inc.	26.479,26	83.853,59
AV Stores, Co.	157.807,81	23.831,33
Baane Mini Imports	116.599,19	142.027,03
Bavarian Collectables Imports, Co.	34.993,92	104.939,27
		31.494,53

PARAMETRO POR LUGAR

CONTINENTE

Parámetro Campo

Parámetro Campo



11.TABLAS REALIZADAS

- SUMMARIZE NEW VENDEDORES = SUMMARIZE(Ventas, Ventas[NOMBRE_COMPLETO_VENDEDOR])

Hicimos esta tabla, eliminamos los duplicados de los empleados.

NOMBRE_COMPLETO_VENDEDOR
Jonas Bergulfsen
Maria Hernandez
Francisca Cervantes
Steve Thompson
Kyung Tseng
Leslie Murphy
Adrian Huxley
Valarie Young
Peter Ferguson
Valarie Nelson
Allen Nelson
Jean Fresnisre
Georg Pippis
Ann Brown
Wing C Tam
Matti Karttunen
Annette Roulet
Thomas Hardy
Eduardo Saavedra
Valarie Thompson
Catherine Dewey
Juri Yoshido
Helen Bennett

12.CONCLUSIONES

- Con respecto a las regiones podemos ver que Europa es el continente donde mayor cantidad de ordenes hubo, al igual que el continente que más dinero recaudo. Con respecto a los países Estados Unidos fue el País que mas ventas tuvo, seguido de España y Francia. Y finalmente las Ciudades que mas dinero recaudaron fueron Madrid, San Rafael y Nueva York.
- Con relación a los productos con mayor rotación de ventas, tenemos en primer lugar, a los carros clásicos, seguido de carros antiguos y en tercera posición están las motocicletas.
- También podemos evidenciar que el cliente que mas ordenes facturo fue Euro Shopping Center, en segunda plaza encontramos a mini gifts distributors ltda. Y después esta Australian Collectors.
- En el tamaño de las ordenes, la talla media obtuvo un 60.77%, la talla pequeña un 26,33% y la talla grande un 12,9%.
- El año 2019 fue el año con mas representación de ordenes generadas con un total de 48,01%, a su vez el año 2018 obtuvo un 34,2% y por ultimo el año 2020 con un 17,8%.
- Por el lado de los empleados el señor Diego Freyre de Madrid España, genero el mayor recaudo en el valor de las ventas. Y por otro lado el empleado que menos ventas desarrollo fue la señorita Leslie Young de la ciudad de Glendale Estados Unidos.
- Alrededor de 88764 ordenes fueron enviadas, 2038 fueron canceladas, 1879 en espera, 1660 resueltas, 1490 en proceso y 597 en disputa.

13. FUTURAS LINEAS

- Con base a las conclusiones enunciadas, Sugerimos en primera instancia: Las regiones que quedaron muy rezagadas con respecto a Estados Unidos y Madrid España, Generan promociones, descuentos y mayor publicidad, para poder incentivar mucho más las ventas.
- En los productos menos vendidos, como trenes y barcos, desarrollar descuentos especiales, en dichos vehículos, así como un acompañamiento por parte de mecánicos especializados en el mantenimiento de estos productos.
- Los clientes que no realizaron muchas compras, hay que estimularlos con buenas promociones y con una facilidad de pago para ellos.
- Por último a los mejores empleados, dícese los que mas ventas realizan, darles impulsos económicos, viajes y regalos especiales, para que tengan mucho más entusiasmo de dar buenos resultados en la compañía.