

Datawarehouse y Minería de Datos

DMD941 G01T

DESAFÍO PRACTICO 2

CICLO 01 - 2025

Docente: Karen Medrano

Estudiante: Alberto Alkindi Ramírez Rivas

Carné: RR160517

El porcentaje del desafío que se alcanzó fue el 100%



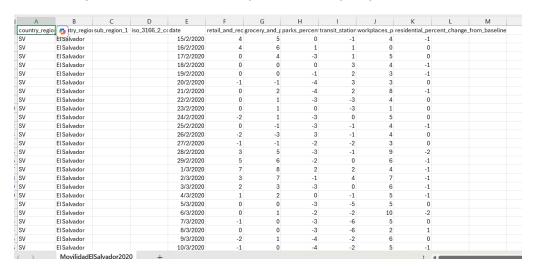
ÍNDICE

INDICE.	2
MODELO	Y CARGA3
DE DATO)\$ 3
Impor	tación de archivo a SQL Server3
1.	Creación de base de datos
2.	Importar archivo .csv4
3.	Creación de tablas de dimensiones y hechos6
DISEÑO	DE 11
CUBO O	LAP11
Defini	ción de Data Source View11
Pas	o 1 11
Pas	o 2 12
Pas	o 3 13
Pas	o 4 13
Pas	o 5 15
Pas	o 6 15
Pas	o 7 16
Pas	o 8 17
Pas	o 9 18
Pas	o 10 18
Pas	o 11 20
ANÁLIS	S DE DATOS
Anális	is de datos en tablas dinámicas21
Carga	de datos en archivo de EXCEL23



MODELO Y CARGA DE DATOS

Los datos se extrajeron a un archivo .csv y se filtraron y depuraron los datos



Importación de archivo a SQL Server

1. Creación de base de datos

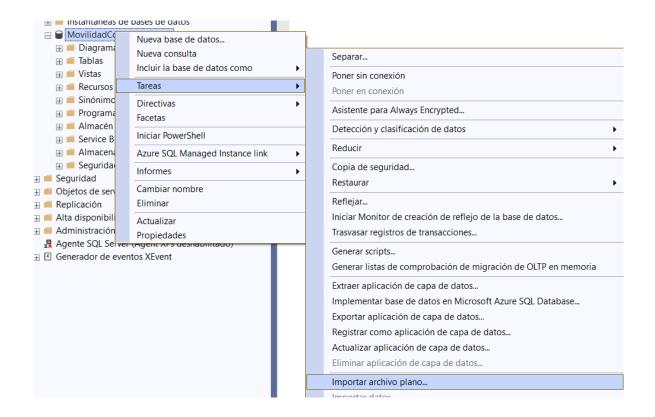
```
SQLQuery1.sq...or (sa (60))* 

CREATE DATABASE MovilidadCovidElSalvador;

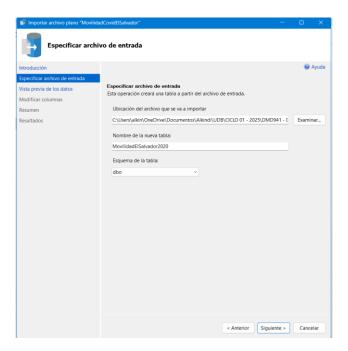
GO
USE MovilidadCovidElSalvador;
```



2. Importar archivo .csv

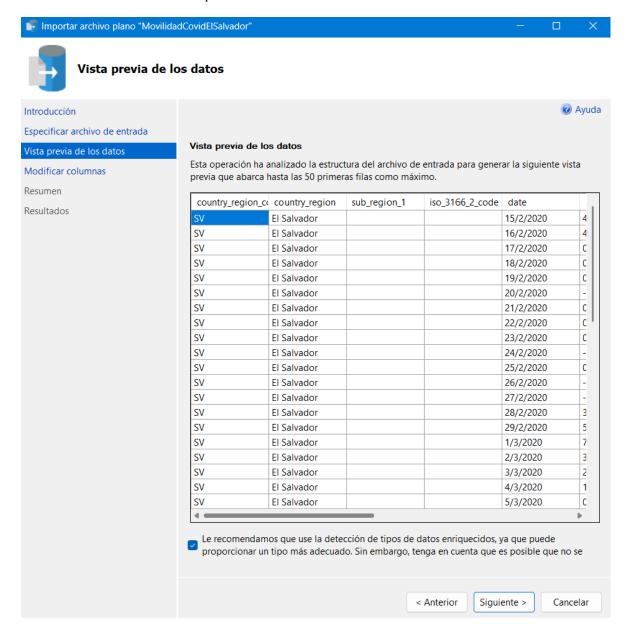


Se eligió el archivo plano .csv





Corroboración de datos importados:





3. Creación de tablas de dimensiones y hechos

Creación de tabla dimensión de tiempo

```
/*
Creación de tabla dimensional de Tiempo
*/
CREATE TABLE dimTiempo (
    idTiempo INT IDENTITY(1,1) PRIMARY KEY,
    fecha DATE,
    anio INT,
    mes INT,
    dia INT,
    nombre_dia NVARCHAR(20)
);
INSERT INTO dimTiempo (fecha, anio, mes, dia, nombre_dia)
SELECT DISTINCT
    date.
    YEAR(date),
    MONTH(date),
    DAY(date),
    DATENAME(WEEKDAY, date)
FROM MovilidadElSalvador2020;
```

Creación de tabla dimensional de ubicación

```
/*
Creación de tabla dimensional de Ubiçación
*/

CREATE TABLE dimUbicacion (
    idUbicacion INT IDENTITY(1,1) PRIMARY KEY,
    pais NVARCHAR(100),
    departamento NVARCHAR(100)
);

INSERT INTO dimUbicacion (pais, departamento)
SELECT DISTINCT
    country_region,
    sub_region_1
FROM MovilidadElSalvador2020;
```

Página 6



Creación de tabla dimensional de categorías

```
/*
Creación de tabla dimensional de Categorías
*/
CREATE TABLE dimCategoria (
    idCategoria INT PRIMARY KEY,
    nombre_categoria NVARCHAR(100)
);

INSERT INTO dimCategoria (idCategoria, nombre_categoria)
VALUES
(1, 'Comercio y recreacion'),
(2, 'Supermercados y farmacias'),
(3, 'Parques y espacios publicos'),
(4, 'Estaciones de transporte'),
(5, 'Lugares de trabajo'),
(6, 'Residencias');
```

Creación de tabla de hechos Movilidad

```
/*
Creación de tabla de hechos
*/

CREATE TABLE HechosMovilidad (
    idHecho INT IDENTITY(1,1) PRIMARY KEY,
    idTiempo INT,
    idUbicacion INT,
    idCategoria INT,
    valorCambio INT,
    FOREIGN KEY (idTiempo) REFERENCES dimTiempo(idTiempo),
    FOREIGN KEY (idUbicacion) REFERENCES dimUbicacion(idUbicacion),
    FOREIGN KEY (idCategoria) REFERENCES dimCategoria(idCategoria)
);

ALTER TABLE HechosMovilidad
ALTER COLUMN valorCambio FLOAT;
```



A continuación, se insertarían los datos a las tablas pertinentes por medio de consultas SQL a cada categoría:

```
Insetar datos en la tabla Hechos
1) Comercio y recreación
INSERT INTO HechosMovilidad (idTiempo, idUbicacion, idCategoria, valorCambio)
SELECT
    t.idTiempo
    u.idUbicacion,
    m.retail_and_recreation_percent_change_from_baseline
FROM MovilidadElSalvador2020 m
JOIN dimTiempo t ON m.date = t.fecha
JOIN dimUbicacion u
    ON ISNULL(LTRIM(RTRIM(LOWER(m.country_region))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.pais))), '')
    AND ISNULL(LTRIM(RTRIM(LOWER(m.sub_region_1))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.departamento))), '');
Insetar datos en la tabla Hechos
2) Supermercados y farmacias
INSERT INTO HechosMovilidad (idTiempo, idUbicacion, idCategoria, valorCambio)
SELECT
   t.idTiempo,
   u.idUbicacion,
   2, /* Supermercados y farmacias */
    m.grocery_and_pharmacy_percent_change_from_baseline
FROM MovilidadElSalvador2020 m
JOIN dimTiempo t ON m.date = t.fecha
JOIN dimUbicacion u
   ON ISNULL(LTRIM(RTRIM(LOWER(m.country_region))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.pais))), '')
   AND ISNULL(LTRIM(RTRIM(LOWER(m.sub_region_1))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.departamento))), '');
 /*
 Insetar datos en la tabla Hechos
 Parques y espacios publicos
 INSERT INTO HechosMovilidad (idTiempo, idUbicacion, idCategoria, valorCambio)
 SELECT
     t.idTiempo,
     u.idUbicacion,
     3, /* Parques y espacios publicos */
     m.parks_percent_change_from_baseline
 FROM MovilidadElSalvador2020 m
 JOIN dimTiempo t ON m.date = t.fecha
 JOIN dimUbicacion u
     ON ISNULL(LTRIM(RTRIM(LOWER(m.country_region))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.pais))), '')
     AND ISNULL(LTRIM(RTRIM(LOWER(m.sub_region_1))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.departamento))), '');
```



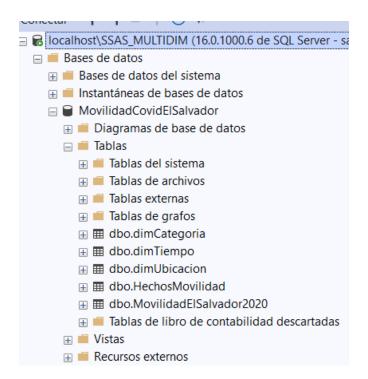
```
Insetar datos en la tabla Hechos
4) Estaciones de transporte
INSERT INTO HechosMovilidad (idTiempo, idUbicacion, idCategoria, valorCambio)
SELECT
    t.idTiempo.
    u.idUbicacion,
    4, /* Estaciones de transporte */
    m.transit_stations_percent_change_from_baseline
FROM MovilidadElSalvador2020 m
JOIN dimTiempo t ON m.date = t.fecha
JOIN dimUbicacion u
    ON ISNULL(LTRIM(RTRIM(LOWER(m.country_region))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.pais))), '')
    AND ISNULL(LTRIM(RTRIM(LOWER(m.sub_region_1))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.departamento))), '');
 Insetar datos en la tabla Hechos
 5) Lugares de trabajo
 INSERT INTO HechosMovilidad (idTiempo, idUbicacion, idCategoria, valorCambio)
 SELECT
    t.idTiempo,
    u.idUbicacion,
    5, /* Lugares de trabajo */
    m.workplaces_percent_change_from_baseline
 FROM MovilidadElSalvador2020 m
 JOIN dimTiempo t ON m.date = t.fecha
 JOIN dimUbicacion u
    ON ISNULL(LTRIM(RTRIM(LOWER(m.country_region))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.pais))), '')
    AND ISNULL(LTRIM(RTRIM(LOWER(m.sub_region_1))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.departamento))), '');
/*
Insetar datos en la tabla Hechos
6) Residencias
INSERT INTO HechosMovilidad (idTiempo, idUbicacion, idCategoria, valorCambio)
SELECT
    t.idTiempo,
    u.idUbicacion,
    6, /* Residencias */
    m.residential_percent_change_from_baseline
FROM MovilidadElSalvador2020 m
JOIN dimTiempo t ON m.date = t.fecha
JOIN dimUbicacion u
    ON ISNULL(LTRIM(RTRIM(LOWER(m.country_region))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.pais))), '')
    AND ISNULL(LTRIM(RTRIM(LOWER(m.sub_region_1))), '') = ISNULL(LTRIM(RTRIM(LOWER(u.departamento))), '');
```

Y por último corroboraremos que los datos estén cargados correctamente en las tablas:



	176 177	SELECT	TOP 100 *	FROM Hec	hosMovilidad;		
90 %	-	_	ncontraron p		4		
⊞ Resultados							
	idHecho	idTiempo	idUbicacion	idCategoria	valorCambio		
1	1	302	2	1	4		
2	2	318	2	1	4		
3	3	161	2	1	0		
4	4	227	2	1	0		
5	5	179	2	1	0		
6	6	135	2	1	-1		
7	7	181	2	1	0		
8	8	298	2	1	0		
9	9	186	2	1	0		
10	10	307	2	1	-2		
11	11	58	2	1	0		
12	12	313	2	1	-2		
13	13	214	2	1	-1		
14	14	267	2	1	3		

Así quedaría exportado los datos y creadas las tablas de dimensiones y hechos correctamente:



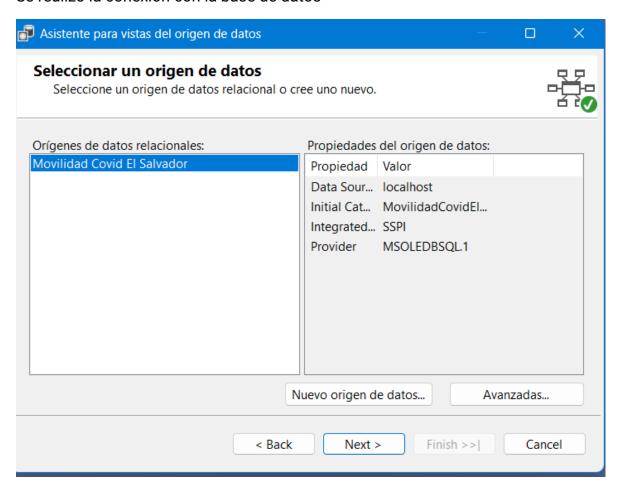


DISEÑO DECUBO OLAP

Primeramente, se creará un nuevo proyecto *Analydis Services Multidimensional* con el nombre *Desafio_2_RR160517*.

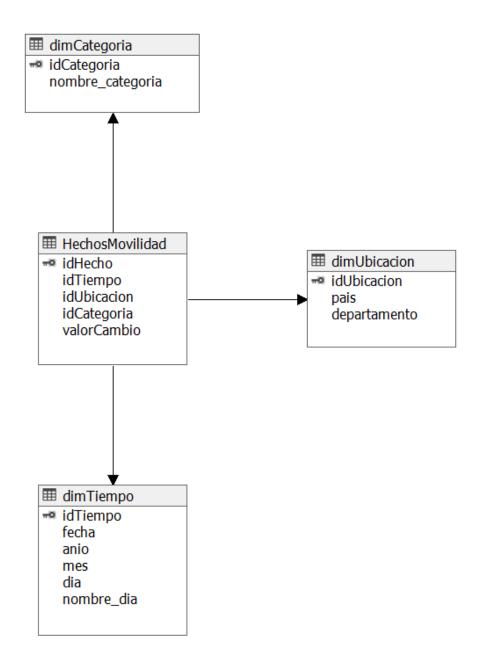
Definición de Data Source View

Paso 1 Se realizó la conexión con la base de datos



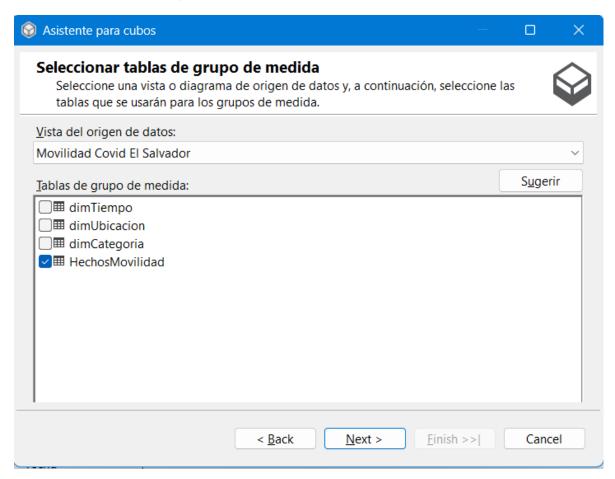


Una vez seleccionado la base de datos y realizada la conexión, podemos visualizar el diagrama:





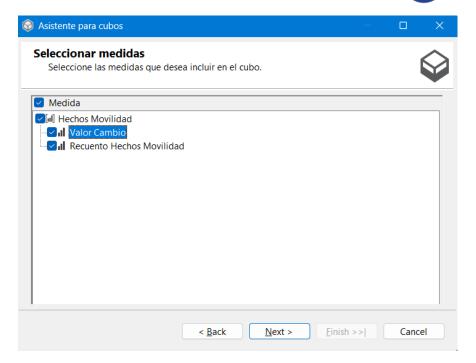
Se configurará por medio del asistente para cubos la creación del cubo, tomando en cuenta la tabla de medida, en nuestro caso será "HechosMovilidad"



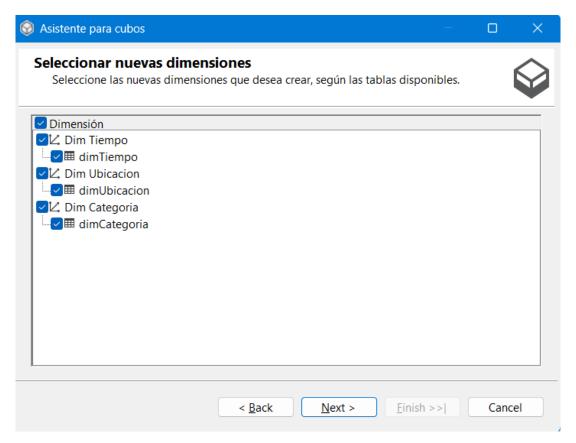
Paso 4

Seleccionaremos las medidas a configurar



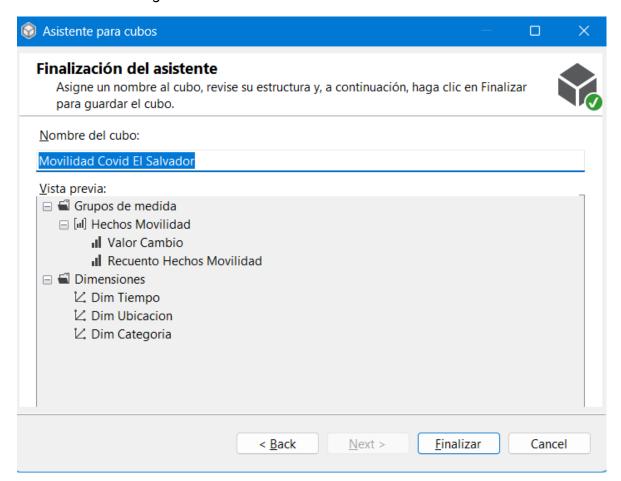


Y luego seleccionaremos las dimensiones creadas:





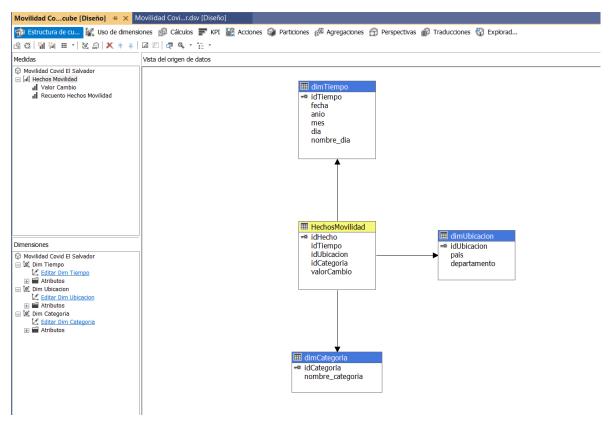
Una vez confirmamos que las medidas y dimensiones están seleccionadas, finalizaremos la configuración del cubo



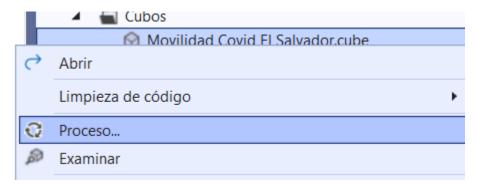
Paso 6

Visualizaremos la venta del diseño del Cubo, con nuestras medidas y dimensiones cargadas correctamente en el panel izquierdo:

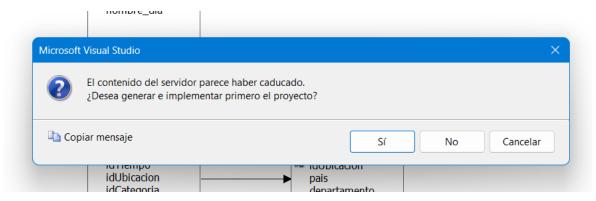




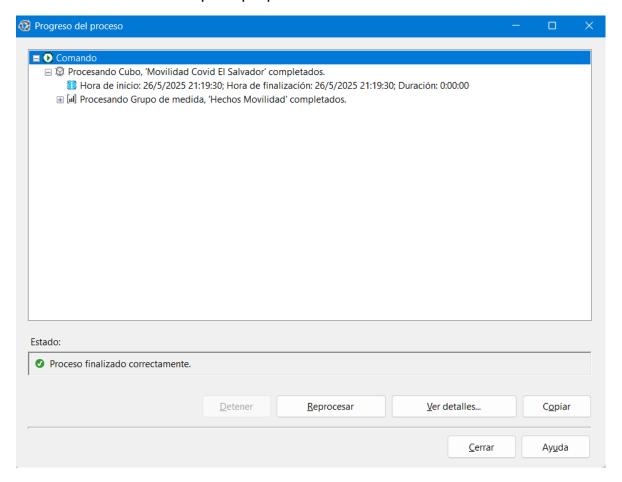
Daremos clic derecho sobre el cubo creado y luego daremos clic en la opción "Proceso". Esto nos permitirá "llenar" el cubo con los datos cargados.





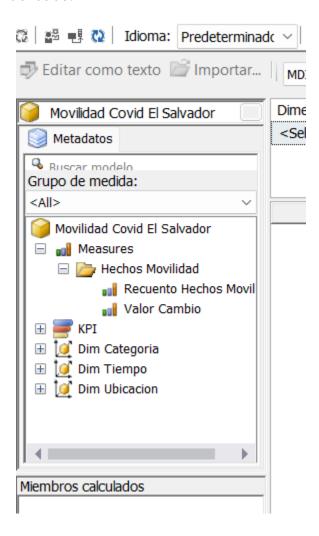


Realizaremos la configuración con el asistente. En este paso se tuvo que corroborar la conexión a la base de datos para que procesen los datos correctamente:





Una vez finalizado este proceso, nos permitirá ver nuestras tablas de dimensiones y hecho en el diseñador del cubo.



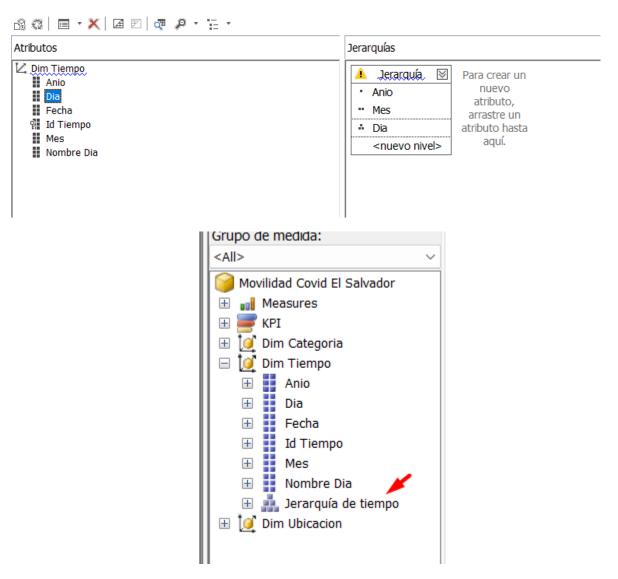
Paso 10

Antes de pasar el explorador y manipular los datos, realizaremos una configuración de jerarquías con los datos cargados.

Si en dado caso no aparecen los datos de las tablas, en la configuración del cubo podemos arrastrar los campos de la tabla al apartado de atributos.

En este caso realizaremos una jerarquía para mostrar la fecha de la dimensión tiempo.





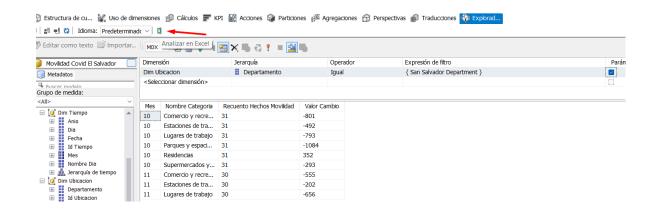
Además, se creará una jerarquía con la dimensión Ubicación





Una vez realizado los cambios, daremos clic derecho en el nombre de nuestro cubo y realizaremos nuevamente el proceso para que actualicen los cambios realizados.

Luego nos iremos a la pestaña de *explorador* y veremos que ya podemos manipular los datos compilados.





ANÁLISISDE DATOS

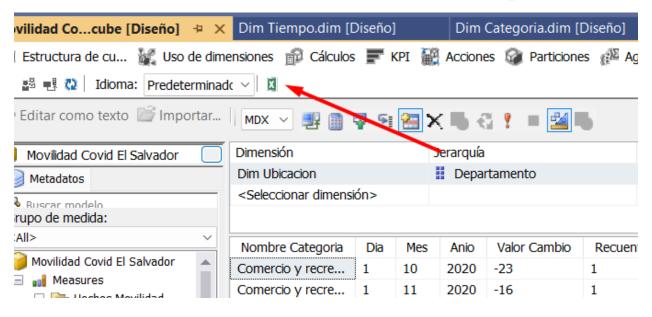
En este apartado corroboramos los datos con un filtro en donde vemos todas las estadísticas de las categorías en el departamento de San Salvador:

Dimensión			Jerarquía		Operador	Expresión de filtro
Dim Ubicacion			■ Departamento		Igual	{ San Salvador Department }
<seleccionar dimens<="" td=""><td>ión></td><td></td><td></td><td></td><td></td><td></td></seleccionar>	ión>					
Nombre Categoria	Dia	Mes	Anio	Valor Cambio	Recuento Hechos Movilidad	
Comercio y recre	1	10	2020	-23	1	
Comercio y recre	1	11	2020	-16	1	
Comercio y recre	1	12	2020	-8	1	
Comercio y recre	1	3	2020	8	1	
Comercio y recre	1	4	2020	-73	1	
Comercio y recre	1	5	2020	-78	1	
Comercio y recre	1	6	2020	-76	1	
Comercio y recre	1	7	2020	-52	1	
Comercio y recre	1	8	2020	-47	1	
Comercio y recre	1	9	2020	-32	1	
Comercio y recre	10	10	2020	-27	1	
Comercio y recre	10	11	2020	-20	1	
Comercio y recre	10	12	2020	-9	1	
Comercio y recre	10	3	2020	0	1	
Comercio y recre	10	4	2020	-87	1	
Comercio y recre	10	5	2020	-79	1	
Comercio y recre	10	6	2020	-70	1	
Comercio y recre	10	7	2020	-59	1	
Comercio v recre	10	8	2020	-49	1	

Análisis de datos en tablas dinámicas

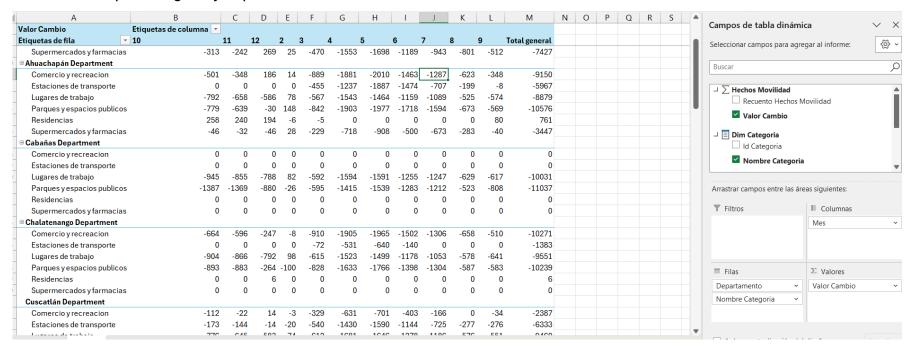
El visor anterior nos permite analizar diversos datos en base a las necesidades que se posean, sin embargo; también podemos realizar el análisis más a profundidad en una tabla dinámica de Excel, dando clic en la función del editor:





Carga de datos en archivo de EXCEL

En este ejemplo podemos visualizar los datos por categorías y departamentos. Esto nos permite evaluar el análisis de movilidad por categoría y departamento:



Cabe destacar que, en el apartado de valores, nos permite realizar una configuración de porcentajes de valores dependiendo de nuestra necesidad:



