# Modelo relacional

El script para el modelo relacional se encuentra en la carpeta scripts y se llama “asathorRelacional”, este archivo genera todos los elementos necesarios que necesita el data warehouse para funcionar, los elementos son:

Base de datos.

Tablas del modelo.

Inserción de datos de ejemplo (los invente para llenar esta cosa no mas)

Vistas de ejemplo para saber si el modelo responde a las consultas del ejercicio.

Estas vistas te proporcionan un ejemplo para saber si el modelo realmente sirve, para resolver las consultas que solicitan, para esto realizare las consultas necesarias utilizando fechas concretas, esto es debido a que aquí no tengo la dimensión tiempo, así que las consultas que requieran fechas se realizaran explícitamente ya sea con el año que piden o la fecha directamente.

Vistas generadas

vwDimCliente: Vista que emula el comportamiento de la dimensión cliente.

vwDimContrato: Vista que emula el comportamiento de la dimensión contrato.

vwDimDocumento: Vista que emula el comportamiento de la dimensión documento.

vwDimEmpleado: Vista que emula el comportamiento de la dimensión empleado.

vwDimSucursal: Vista que emula el comportamiento de la dimensión sucursal.

vwFactVenta: Vista que emula el comportamiento de la tabla de hechos.

# Diagramas de los modelos

Los diagramas no los he hecho, pero en sql es facilisimo, tan solo debes ir a el apartado diagramas dentro de la base de datos y luego le das a nuevo diagrama, seleccionas las tablas que quieres y este te dibuja solito el monito, pero el acomodar todo ese jaleo me cuesta caleta, asi que queda de tarea para la casa! XD!

# Resultados esperados

Colocare las preguntas del ejercicio y sus resultados aplicando las vistas mencionadas arriba, y podrás visualizar que tal quedo todo este jaleo.

**Pregunta 1: Cuáles son los teléfonos celulares de mayor predilección de los usuarios por región.**

SELECT DISTINCT

    s.nombre\_region,

    c.modelo,

    max(c.MODELOS\_VENDIDOS\_A\_LA\_FECHA) AS Cant

FROM

    vwFactVenta v

    INNER JOIN vwDimSucursal s ON v.id\_sucursal=s.id\_sucursal

    INNER JOIN vwDimContrato c ON v.id\_contrato=c.id\_contrato

GROUP BY s.nombre\_region, c.modelo

HAVING max(c.MODELOS\_VENDIDOS\_A\_LA\_FECHA)=

    (SELECT

        max(con.MODELOS\_VENDIDOS\_A\_LA\_FECHA)

    FROM

        vwFactVenta vta

        INNER JOIN vwDimSucursal suc ON vta.id\_sucursal=suc.id\_sucursal

        INNER JOIN vwDimContrato con ON vta.id\_contrato=con.id\_contrato

    WHERE suc.nombre\_region=s.nombre\_region )

ORDER BY s.nombre\_region

Resultados:

I de Tarapacá   Motorola Moto E6 Plus   17

II de Antofagasta   Samsung Galaxy S10 5G   19

III de Atacama  Alcatel 3l  17

Metropolitana de Santiago   Alcatel 3l  18

Metropolitana de Santiago   Motorola Moto E6 Plus   18

V de Valparaíso Motorola Moto E6 Plus   19

VII del Maule   Samsung Galaxy S10 5G   22

VIII de Concepción  Sony Xperia L3  15

X de Los Lagos  Samsung Galaxy S10 5G   21

XI de Aysén del General Carlos Ibañez del Campo LG K30  14

XIV de Los Ríos Samsung Galaxy S10 5G   15

XV de Arica y Parinacota    Google Plxel 4  10

XV de Arica y Parinacota    iPhone XR   10

**Pregunta 2: Cuáles planes de celulares son los menos escogidos por clientes personas naturales.**

SELECT TOP(1)

    con.nombre\_plan,

    max(con.PLANES\_VENDIDOS\_A\_LA\_FECHA) AS Cant

FROM

    vwFactVenta vta

    INNER JOIN vwDimContrato con ON vta.id\_contrato=con.id\_contrato

WHERE cliente\_coorporativo=0

GROUP BY con.nombre\_plan

ORDER BY cant

Resultado:

Plan 35 GB  74

Los planes de clientes “Persona natural” son:

Plan 35 GB  74

Plan 13 GB  76

Plan 45 GB  83

Plan 25 GB  89

**Pregunta 3: Cuál sucursal es la que mejor desempeño en ventas durante el periodo 2005 – 2012**

SELECT TOP(1)

    suc.sucursal,

    count(suc.sucursal) AS [Cantidad ventas],

    sum(vta.monto\_venta) AS [Total ventas]

FROM

    vwFactVenta vta

    INNER JOIN vwDimSucursal suc ON vta.id\_sucursal=suc.id\_sucursal

WHERE year(CONVERT(DATETIME, vta.id\_fecha)) BETWEEN 2005 AND 2012

GROUP BY suc.sucursal

ORDER BY [Total ventas] DESC

Resultado:

sucursal    Cantidad ventas Total ventas

Suc.Quillota     37   1894130

Pregunta 4: Cuál es el acumulado general de ventas discriminado por región y tipo de cliente.

SELECT

    suc.nombre\_region,

    cli.tipo\_cliente,

    sum(vta.monto\_venta) AS [Monto]

FROM

    vwFactVenta vta

    INNER JOIN vwDimSucursal suc ON vta.id\_sucursal=suc.id\_sucursal

    INNER JOIN vwDimCliente cli ON vta.id\_cliente=cli.id\_cliente

GROUP BY suc.nombre\_region, cli.tipo\_cliente

ORDER BY suc.nombre\_region

Resultado: Aquí difiero en este resultado, personalmente no me gustan tirar así, pero verás la diferencia ya que te tira la lista con los valores correctos pero es una lista no una tabla, después muestro mi forma favorita de mostrar este tipo de resultados.

I de Tarapacá   Cliente corporativo 737900

I de Tarapacá   Persona natural 302330

II de Antofagasta   Cliente corporativo 2969620

II de Antofagasta   Persona natural 700460

III de Atacama  Cliente corporativo 1055870

III de Atacama  Persona natural 283810

…y así hasta pasar por todas las regiones, esto te dara almenos 32 filas, 16 regiones por 2 tipos de personas, la crema no lo crees?, pero aquí esta la magia, pone atención!

SELECT \* FROM (

    SELECT

        suc.nombre\_region,

        cli.tipo\_cliente,

        vta.monto\_venta AS Ventas

    FROM

        vwFactVenta vta

        INNER JOIN vwDimSucursal suc ON vta.id\_sucursal=suc.id\_sucursal

        INNER JOIN vwDimCliente cli ON vta.id\_cliente=cli.id\_cliente) tabla

PIVOT (

    sum(Ventas)

    FOR tipo\_cliente

    IN ([Cliente corporativo],[Persona natural])

) tablaPivot

ORDER BY nombre\_region

…hasta la consulta es mas elegante, esta si que es brutalisima!!! Y el resultado queda:

nombre\_region                               Cliente corporativo Persona natural

I de Tarapacá                                       737900       302330

II de Antofagasta                                   2969620      700460

III de Atacama                                      1055870      283810

Metropolitana de Santiago                           2927650      313680

V de Valparaíso                                     5851270      1203420

VII del Maule                                       3669580      819150

VIII de Concepción                                  1793770      334950

X de Los Lagos                                      2595690      602920

XI de Aysén del General Carlos Ibañez del Campo     603930       64760

XIV de Los Ríos                                     271970       155000

XV de Arica y Parinacota                            663920       153200

…y que tal? Se ve elegante al igual que la consulta, las pivot tables son brutales, hechale un ojo a este tema es super interesante y dominarlas cuesta un poco al principio pero realmente vale la pena usarlas sobre todo en tablas comparativas.

**Pregunta 5: Cuáles son los clientes con mejor comportamiento en el pago de sus facturas durante los años 2009 – 2011.**

SELECT

    cli.nombre\_cliente,

    CONVERT(DATE, CONVERT(DATETIME, id\_fecha)) AS [Fecha de pago],

    vta.monto\_pagado

FROM

    vwFactVenta vta

    INNER JOIN vwDimCliente cli ON vta.id\_cliente=cli.id\_cliente

WHERE vta.es\_factura=1 AND vta.doc\_pagado=1 AND year(CONVERT(DATETIME, id\_fecha)) BETWEEN 2009 AND 2011

ORDER BY vta.id\_fecha

Resultados (extracto):

nombre\_cliente  Fecha de pago       monto\_pagado

Frances Adams       2009-01-15          19790

Miguel Adams        2009-05-10          9990

Mary Adams          2009-06-05          59990

Ben Adams           2009-07-22          13490

**Pregunta 6: Cuál es el monto total de la facturación vencida de clientes corporativos durante el periodo 2005 – 2010**

SELECT

    sum(monto\_venta) AS [Total monto documentos vencidos]

FROM

    vwFactVenta vta

WHERE vta.cliente\_coorporativo=1

    AND vta.es\_factura=1

    AND vta.doc\_vencido=1

    AND year(CONVERT(DATETIME, vta.id\_fecha)) BETWEEN 2005 AND 2010

Resultado:

Total monto documentos vencidos

2913660

**Pregunta 7: En qué porcentaje se ha incrementado o reducido la incorporación de clientes personas naturales y corporativos en los últimos 5 años.**

SELECT

    tabla.tipo\_cliente,

    tabla.periodo\_actual AS [2014-2019],

    tabla.periodo\_anterior AS [2009-2013],

    CASE

        WHEN periodo\_actual<periodo\_anterior THEN 'Disminuye'

        ELSE 'Incremento'

    END AS estado,

    CASE

        WHEN periodo\_actual<periodo\_anterior

            THEN format((CONVERT(FLOAT, (periodo\_actual-periodo\_anterior))/periodo\_anterior)\*100, '000.00')

        ELSE

            format((CONVERT(FLOAT, (periodo\_anterior+periodo\_actual))/periodo\_anterior)\*100, '000.00')

    END AS porcentaje

FROM

    (SELECT

        DISTINCT

        cli.tipo\_cliente,

        (SELECT

            count(v.id\_cliente)

        FROM

            vwFactVenta v

            INNER JOIN vwDimCliente c ON v.id\_cliente=c.id\_cliente

        WHERE year(CONVERT(DATETIME, v.id\_fecha)) BETWEEN 2014 AND 2019 AND v.cliente\_nuevo=1 AND c.tipo\_cliente=cli.tipo\_cliente

    ) AS periodo\_actual,

    (SELECT

        count(v.id\_cliente)

    FROM

        vwFactVenta v

        INNER JOIN vwDimCliente c ON v.id\_cliente=c.id\_cliente

    WHERE year(CONVERT(DATETIME, v.id\_fecha)) BETWEEN 2009 AND 2013 AND v.cliente\_nuevo=1 AND c.tipo\_cliente=cli.tipo\_cliente

) AS periodo\_anterior

FROM

    vwDimCliente cli) tabla

Resultado:

tipo\_cliente        2014-2019   2009-2013   estado      porcentaje

Cliente corporativo     0           12      Disminuye       -100.00

Persona natural         2           6       Disminuye       -66.67

# Conclusión del modelo relacional

Al poder responder las preguntas del ejercicio, puedo visualizar que el modelo responde bien, por lo tanto ahora debería pasar al modelo dimensional, pero si te das cuenta el modelo dimensional siempre lo viste!!!!

Adivina buen adivinador! Las dimensiones y la tabla de hechos son las vistas!!!! O sea con esto comprobamos que el modelo dimensional funciona, tan solo queda escribirlo con los mismos nombre de columnas que te tiras los select a las vistas y listo, prueba los siguiente y lo compruebas, jajajajajajajaj! Vas a quedar loco!!!

SELECT \* INTO dimCliente FROM vwDimCliente

…que hizo? Te crea la dimension!!! Haz esto por todas las vistas y ya tienes tu modelo dimensional hasta con los datos listos para pruebas!!!

Pero ojo, el ejercicio te pide hacer el etl, asi que esto tan solo hazlo para que te tire el modelo y lo compruebes por ti solo.

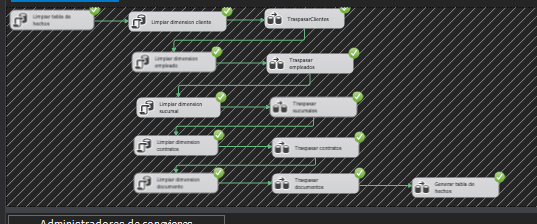
Ahora bien, que hacer con las preguntas?, en las consultas tan solo deberas sacar el prefijo “vw” y listeilors!!! Solo tendras que ajustar las consultas que trabajan con las fechas para que apunten a la tabla tiempo que esa todavia no la tienes, esta se generara en el archivo del script del modelo dimensional que se encuentra en la carpeta “scripts” y el archivo se llama “asathorDimensional.sql”

En dicho script, se encuentra todo lo necesario para generar el modelo dimensional, este lo saque directamente con la ayuda de las vistas y del select que te crea las tablas desde una consulta que te puse mas arriba, el script tiene la consulta para crear fechas desde el año 2004 hasta el 2019 sin necesidad de traer nada desde el etl ya que dicha tabla debe contener las fechas que soporten el ejercicio, asi que esto es todo con respecto a sentencias sql.

Toca el turno de pasar al etl!!!

# ETL.

Aquí ira explicado el proceso para realizar el traspaso ful, que es lo que llevo hasta el momento, este etl se basa en los siguientes flujos

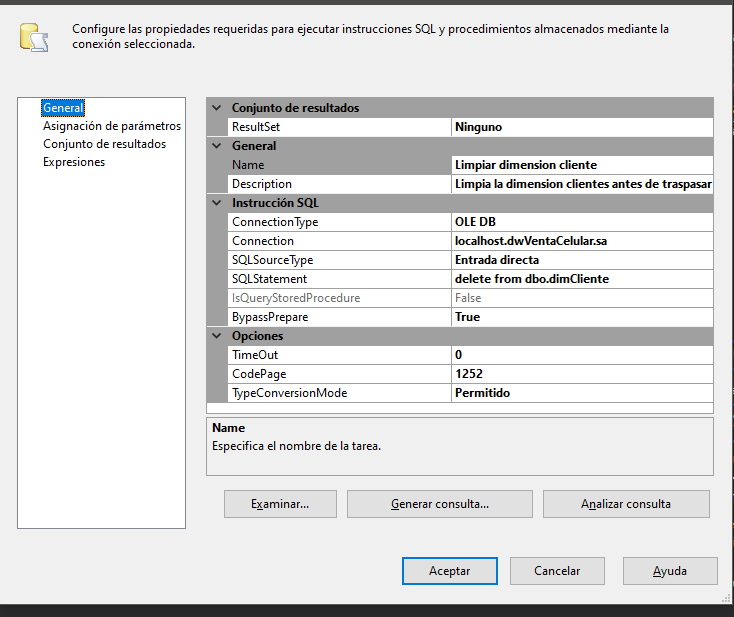


Primero limpio la tabla de hechos para eliminar los registros anteriores, recuerda que este etl es full, así que primeramente se debe eliminar cualquier indicio de información anterior.

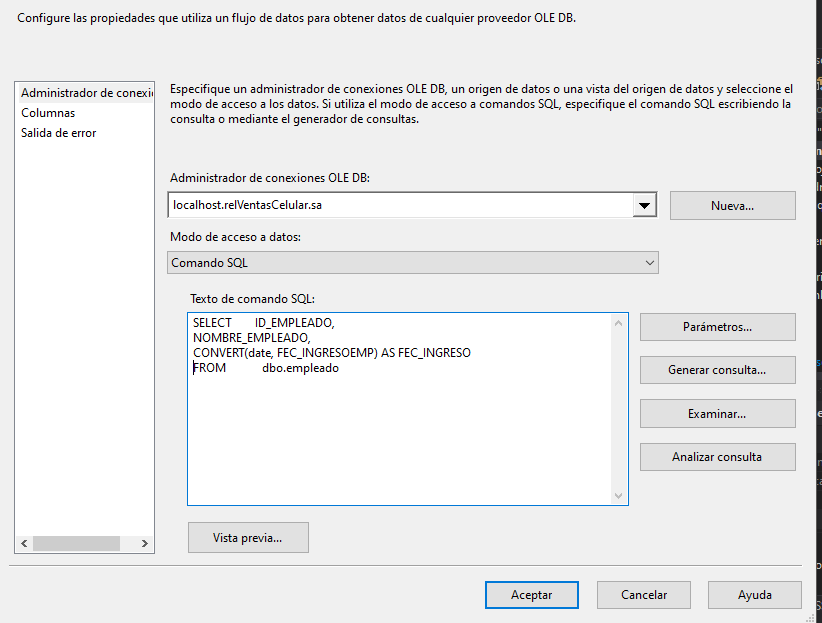
Luego voy limpiando las dimensiones siguientes y en el siguiente flujo realizo el traspaso de la información y así sucesivamente hasta llegar el momento de generar la tabla de hechos que es el ultimo paso. Entonces según esto el proceso es:

1. Limpiar tabla de hechos.
2. Limpiar la dimensión clientes (aquí uso una tarea llamada “ejecutar sql” el cual permite ejecutar una instrucción sql por lo tanto la consulta a ejecutar para este caso ee “delete from dimCliente”, para las otras limpiezas cambiara el nombre de la dimension).
3. Traspasar clientes
4. Limpiar dimensión empleado.
5. Traspasar empleados
6. Limpiar dimensión sucursal
7. Traspasar sucursales
8. Limpiar dimensión contrato
9. Traspasar contratos
10. Limpiar dimensión documento
11. Traspasar documentos
12. Generar tabla de hechos.

Ejemplo de tarea de limpieza.



Para las tareas de traspaso tan solo tiene 1 asistente de origen donde seleccionas la coneccion origen y luego le das a aceptar, das doble click en dicho asistente y donde te pide el metodo colocas consulta sql tan solo ahora deberás colocar los comandos sql necesarios para que cargue la info, pero adivina de donde sale esta consulta, correcto, de las vistas que genera el script del modelo relacional, aquí un ejemplo con la dimension empleado.



Ahora en el asistente de destino tan solo debes seleccionar la coneccion del modelo dimensional y la tabla, cuando vayas al apartado “Asignaciones”, sorpresa!!! Todo esta relacionado, esto se debe a que tu fuente de datos (en este caso el modelo relacional) es quien define por medio de las vistas el funcionar del modelo dimensional, asi que no tendras tanto problema con esas cosas que son las que siempre dan problemas, si todo se define de buena forma al principio estas cosas salen solitas.

Ahora solo falta que hagas los demás enlaces correspondientes y tu etl quedara bien.

Con respecto al etl incremental, este no lo he podido hacer por cuestiones de tiempo, he tenido caleta de pega esta semana, ese te lo debo, en cuante tenga tiempo lo hago y lo comparto si quieres.

# Resolviendo las preguntas desde modelo dimensional.

Ahora que tenemos todo listoco, tan solo nos falta aplicar los cambios a las consultas que responden a lo que el ejercicio te solicita, recuerda que el cambio es super sutil en algunos casos, pondre tan solo las consultas ya que los resultados los puedes comprobar directamente en sql, vale)

## Pregunta 1:

SELECT DISTINCT

    s.nombre\_region,

    c.modelo,

    max(c.MODELOS\_VENDIDOS\_A\_LA\_FECHA) AS Cant

FROM

    factVenta v

    INNER JOIN dimSucursal s ON v.id\_sucursal=s.id\_sucursal

    INNER JOIN dimContrato c ON v.id\_contrato=c.id\_contrato

GROUP BY s.nombre\_region, c.modelo

HAVING max(c.MODELOS\_VENDIDOS\_A\_LA\_FECHA)=

(SELECT

    max(con.MODELOS\_VENDIDOS\_A\_LA\_FECHA)

FROM

    factVenta vta

    INNER JOIN dimSucursal suc ON vta.id\_sucursal=suc.id\_sucursal

    INNER JOIN dimContrato con ON vta.id\_contrato=con.id\_contrato

WHERE suc.nombre\_region=s.nombre\_region )

ORDER BY s.nombre\_region

## Pregunta 2:

SELECT TOP(1)

    con.nombre\_plan,

    max(con.PLANES\_VENDIDOS\_A\_LA\_FECHA) AS Cant

FROM

    factVenta vta

    INNER JOIN dimContrato con ON vta.id\_contrato=con.id\_contrato

WHERE cliente\_coorporativo=0

GROUP BY con.nombre\_plan

ORDER BY cant

## Pregunta 3:

SELECT TOP(1)

    suc.sucursal,

    count(suc.sucursal) AS [Cantidad ventas],

    sum(vta.monto\_venta) AS [Total ventas]

FROM

    factVenta vta

    INNER JOIN dimSucursal suc ON vta.id\_sucursal=suc.id\_sucursal

    inner join dimTiempo tmp on vta.id\_fecha=tmp.id\_fecha

WHERE tmp.anio BETWEEN 2005 AND 2012

GROUP BY suc.sucursal

ORDER BY [Total ventas] DESC

Aquí la diferencia sustancial esta en que ahora tengo que hacer un inner join para enlazar la fecha y el where queda sencillito!!! Brutal o no? Compara el where de arribita con el de la vista que es

where year(convert(datetime, vta.id\_fecha)) between 2005 and 2012

y el de ahora es:

WHERE tmp.anio BETWEEN 2005 AND 2012

Genial!!!

## Pregunta 4:

Consulta feita pero funcional

SELECT

    suc.nombre\_region,

    cli.tipo\_cliente,

    sum(vta.monto\_venta) AS [Monto]

FROM

    factVenta vta

    INNER JOIN dimSucursal suc ON vta.id\_sucursal=suc.id\_sucursal

    INNER JOIN dimCliente cli ON vta.id\_cliente=cli.id\_cliente

GROUP BY suc.nombre\_region, cli.tipo\_cliente

ORDER BY suc.nombre\_region

Consulta brutal y elegante!!! No lo crees?

SELECT \* FROM (

    SELECT

        suc.nombre\_region,

        cli.tipo\_cliente,

        vta.monto\_venta AS Ventas

    FROM

        factVenta vta

        INNER JOIN dimSucursal suc ON vta.id\_sucursal=suc.id\_sucursal

        INNER JOIN dimCliente cli ON vta.id\_cliente=cli.id\_cliente) tabla

PIVOT (

    sum(Ventas)

    FOR tipo\_cliente

    IN ([Cliente corporativo],[Persona natural])

) tablaPivot

ORDER BY nombre\_region

## Pregunta 5:

SELECT

    cli.nombre\_cliente,

    CONVERT(DATE, CONVERT(DATETIME, vta.id\_fecha)) AS [Fecha de pago],

    vta.monto\_pagado

FROM

    factVenta vta

    INNER JOIN dimCliente cli ON vta.id\_cliente=cli.id\_cliente

    inner join dimTiempo tmp on vta.id\_fecha=tmp.id\_fecha

WHERE vta.es\_factura=1

    AND vta.doc\_pagado=1

    AND tmp.anio BETWEEN 2009 AND 2011

ORDER BY vta.id\_fecha

## Pregunta 6:

SELECT

    sum(monto\_venta) AS [Total monto documentos vencidos]

FROM

    factVenta vta

    inner join dimTiempo tmp on vta.id\_fecha=tmp.id\_fecha

WHERE vta.cliente\_coorporativo=1

    AND vta.es\_factura=1

    AND vta.doc\_vencido=1

    AND tmp.anio BETWEEN 2005 AND 2010

## Pregunta 7:

SELECT

    tabla.tipo\_cliente,

    tabla.periodo\_actual AS [2014-2019],

    tabla.periodo\_anterior AS [2009-2013],

    CASE

        WHEN periodo\_actual<periodo\_anterior THEN 'Disminuye'

        ELSE 'Incremento'

    END AS estado,

    CASE

        WHEN periodo\_actual<periodo\_anterior

            THEN format((CONVERT(FLOAT, (periodo\_actual-periodo\_anterior))/periodo\_anterior)\*100, '000.00')

        ELSE

            format((CONVERT(FLOAT, (periodo\_anterior+periodo\_actual))/periodo\_anterior)\*100, '000.00')

    END AS porcentaje

FROM

    (SELECT DISTINCT

        cli.tipo\_cliente,

        (SELECT

            count(v.id\_cliente)

        FROM

            factVenta v

            INNER JOIN dimCliente c ON v.id\_cliente=c.id\_cliente

            inner join dimTiempo t on v.id\_fecha=t.id\_fecha

        WHERE t.anio BETWEEN 2014 AND 2019

            AND v.cliente\_nuevo=1

            AND c.tipo\_cliente=cli.tipo\_cliente

        ) AS periodo\_actual,

        (SELECT

            count(v.id\_cliente)

        FROM

            factVenta v

            INNER JOIN dimCliente c ON v.id\_cliente=c.id\_cliente

            inner join dimTiempo t on v.id\_fecha=t.id\_fecha

        WHERE t.anio BETWEEN 2009 AND 2013

            AND v.cliente\_nuevo=1

            AND c.tipo\_cliente=cli.tipo\_cliente

        ) AS periodo\_anterior

    FROM

        dimCliente cli) tabla

# Conclusión final de todo esto

Aquí con este pequeño ejercicio se puede visualizar como los datos bien estructurados, se pueden lograr virgerias geniales, el modelo dimensional sale solo y los datos que involucra van en relacion al hecho principal, que en este caso es la venta de un telefono, las dimensiones le dan el contexto al hecho, y todo el conjunto permiten realizar un monton de weas con la información que se tiene.

La intención de hacer este ejercicio es que puedas ver en directo lo que una vez hablamos en el patio (lo recuerdas?) y que además puedes apoyarte en varias cosas que te permitan realizar el modelo de mejor manera, puede que haya una mejor pero esta me acomoda (es el metodo Asathor), ya que, no tengo que dibujar modelos dimensionales que posteriormente no me sirvan y tenga que hacer todo nuevamente, si piensas bien el modelo relacional y logras responder las preguntas que te piden en el ejercicio con vistas, estas serán tu guia para el modelo dimensional y posteriormente tambien cuando tengas que hacer el etl, ya que aquí tambien tendras que colocar las consultas para hacer el traspaso, ten encuenta que las asignaciones erroneas al momento de realizar el traspaso son las mas comunes y esta forma de trabajo ni siquiera te preocupas de eso ya que el modelo relacional te proveyo de todos los nombres de columnas que el modelo dimensional necesitaba para su ejecucion.

Bueno esto es todo, espero que te pueda servir de algo este tochaco de weas, si lo pones en practica te daras cuentas de caleta de weas que antes a lo mejor no las comprendias.

Disculpa la falta de ortografia, ya que tan solo me dedique a escribir no mas, lo demas me da paja!!

PD: El lateralus de Tool ha dado como 5 vueltas, por eso la mansa volaita y eso que no me he fumado niuna wea!!!

Ahora si Salud y cuidate!!!