

UNIVERSIDAD CATÓLICA DE CORDOBA

# PROYECTO PARCIAL 2

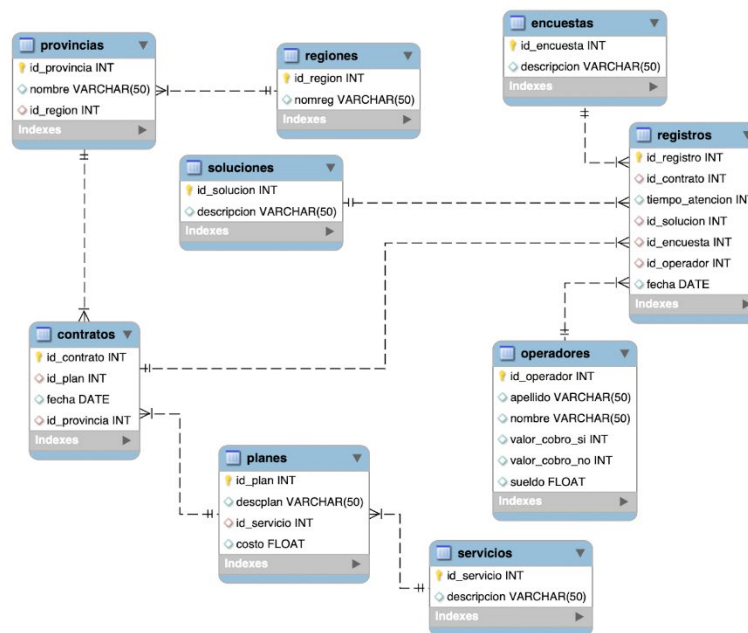
*Materia:* **Base de Datos II.**

*Profesor:* Julio Marcelo Gutierrez

*Alumnos:* Julián Gergolet Albornoz, Felipe Arenillas, Leandro Garza Panelli, Agustín Zanini.

## INTRODUCCIÓN

Para este proyecto utilizamos una base de datos previamente creada llamada Servicios. A continuación, adjuntamos el diagrama de clases de la misma.



Una vez creada utilizaremos la metodología HEFESTO sacada de la página Dataprix. Esta metodología nos hace seguir una serie de pasos y nos guía para crear nuestro Data Warehouse, partiendo de un análisis de requerimientos y terminando con una integración de los datos. El objetivo del armado de nuestro Data Warehouse es analizar los reclamos

## DESARROLLO

### Paso 1 - Análisis de Requerimientos:

Lo primero a realizar en este paso es identificar nuestros requerimientos, a través de preguntas que explicitan los objetivos de este trabajo.

#### a) – *Identificar preguntas:*

Para darle sentido a la información de la base de datos realizamos las siguientes preguntas:

¿Cuántos reclamos hay en cada provincia en un tiempo determinado?

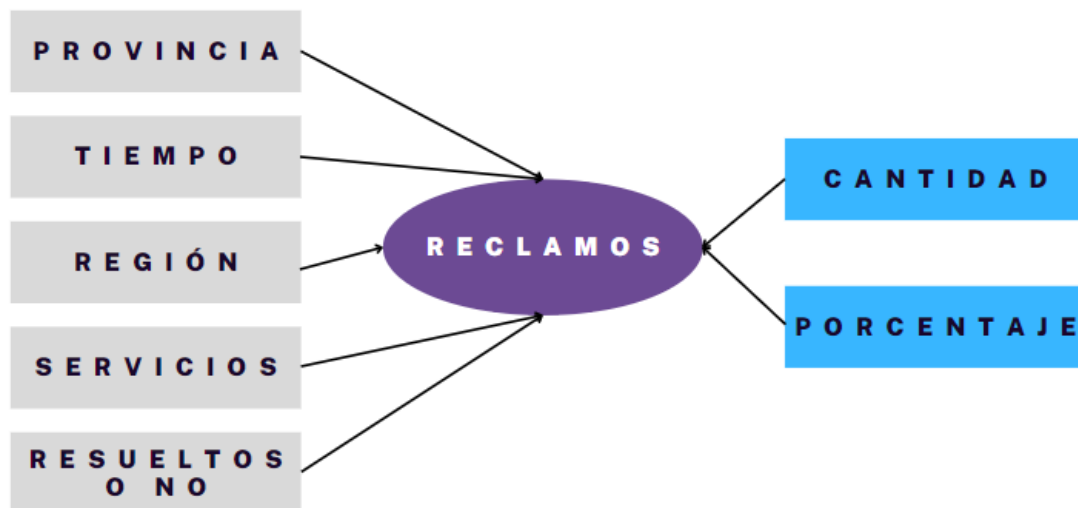
¿Cuántos reclamos hay en cada región en un tiempo determinado?

¿Cuántos reclamos hay por servicio en un tiempo determinado?  
 ¿Qué porcentaje, del total de reclamos, fueron resueltos y que porcentaje no?

b) – Identificar **indicadores** y **perspectivas**:

¿**Cuántos reclamos** hay en cada **provincia** en un **tiempo** determinado?  
 ¿**Cuántos reclamos** hay en cada **región** en un **tiempo** determinado?  
 ¿**Cuántos reclamos** hay por **servicio** en un **tiempo** determinado?  
 ¿**Qué porcentaje**, del total de **reclamos**, fueron **resueltos** y que porcentaje **no** lo fueron, en un **tiempo** determinado?

c) – *Modelo Conceptual*



## Paso 2 - Análisis de los OLTP:

El siguiente paso en la metodología es el de analizar los OLTP, donde se analizan como serán calculados los indicadores y como se configuran las perspectivas.

a) – *Conformar indicadores:*

Como vimos en el paso anterior, nuestros indicadores son 2: cantidad y porcentaje.

- “Cantidad”

Hechos: Cantidad de reclamos

Función: COUNT.

Aclaración: el indicador “Cantidad” representa la cantidad de reclamos efectuados, que luego será observado desde distintas perspectivas.

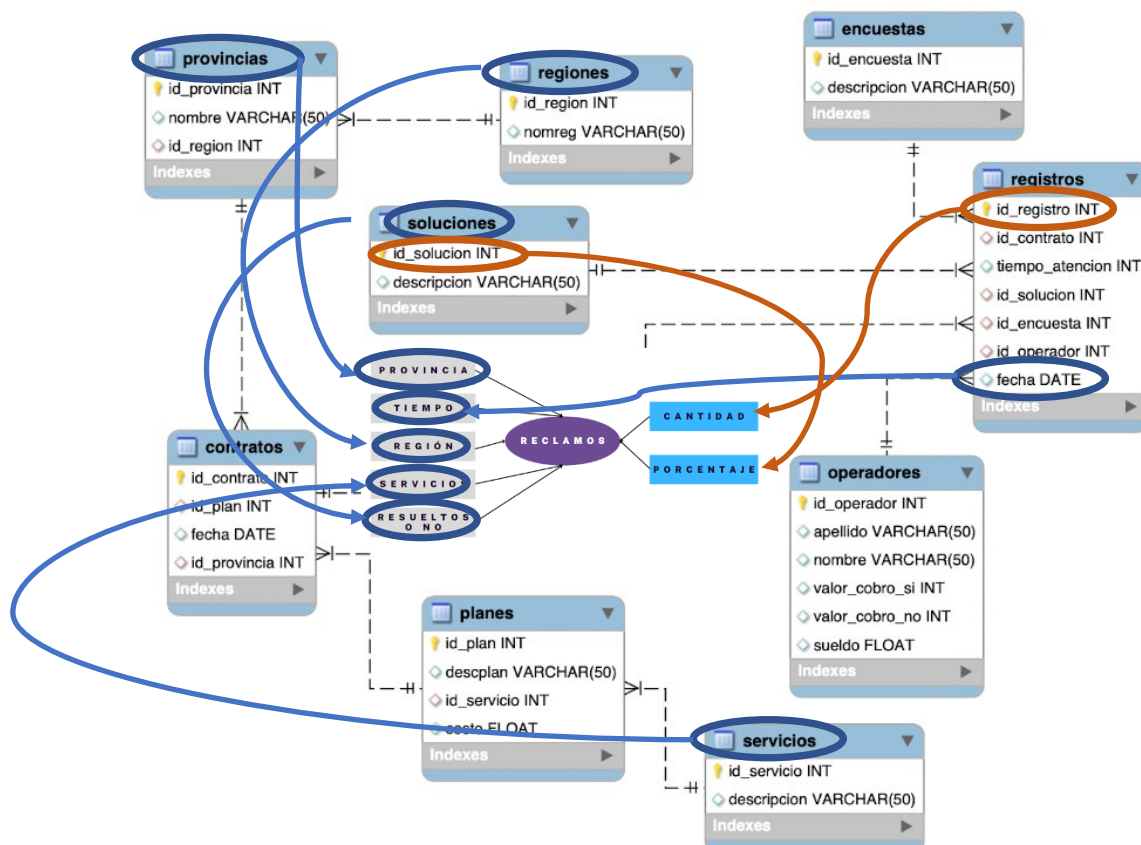
- “Porcentaje”

Hechos: porcentaje de reclamos

Función: COUNT

Aclaración: el indicador “Porcentaje” lo usamos para saber que porcentaje de reclamos fueron resueltos y cuáles no.

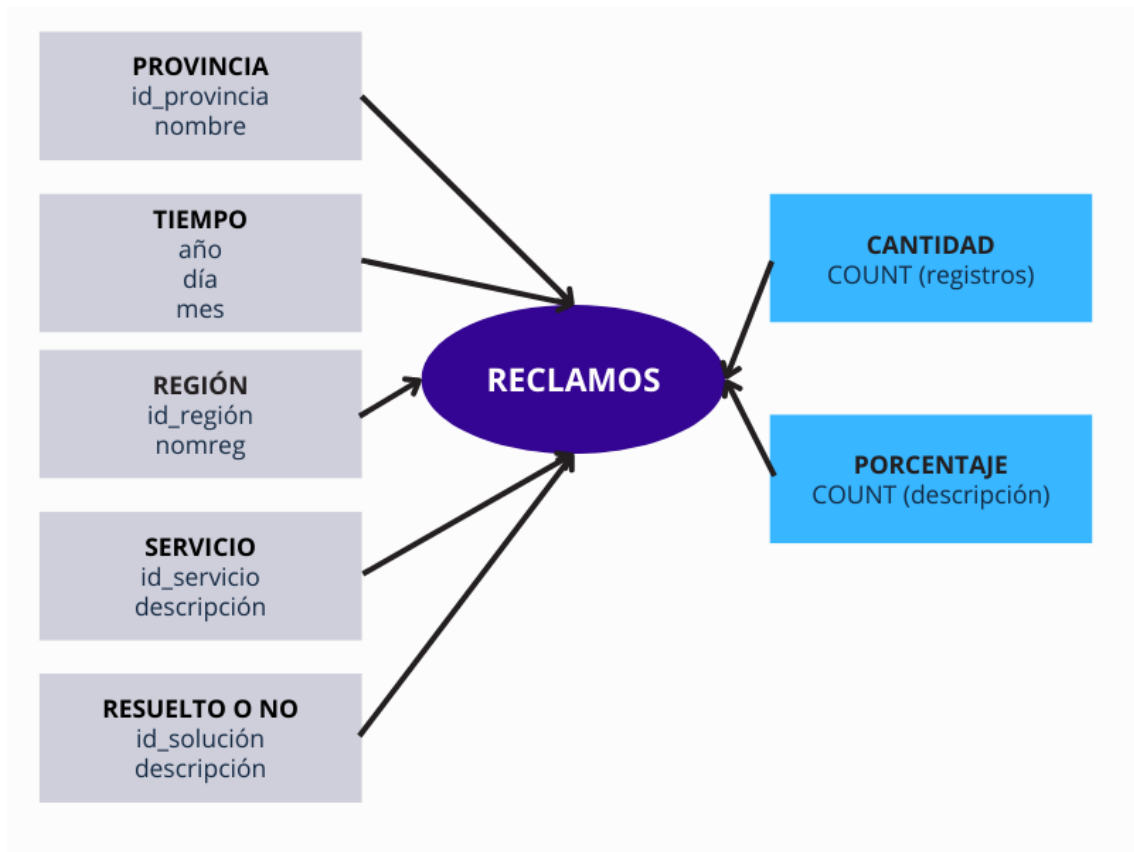
b) – *Establecer Correspondencias:*



c) – *Nivel de granularidad:*

- Perspectiva: **“Provincia”**:  
“id\_provincia” de la tabla “provincias”. Que nos servirá para agrupar por provincias.  
“nombre” de la tabla “provincias”. Nos sirve para, más adelante, poder mostrar los reclamos por nombre de provincia.
- Perspectiva: **“Tiempo”**:  
“año” del atributo “fecha” de la tabla “registros”.  
“mes” del atributo “fecha” de la tabla “registros”.  
“día” del atributo “fecha” de la tabla “registros”.
- Perspectiva: **“Región”**:  
“id\_region” de la tabla “región”. Es el id de cada region, usado para luego agrupar.  
“nomreg” de la tabla “región”. Es el nombre de la region, utilizado para darle sentido a los id.
- Perspectiva: **“Servicios”**:  
“id\_servicio” de la tabla “servicios”. Es el id de cada region, usado para luego agrupar.  
“descripcion” de la tabla “servicios”. Es el nombre de la region, utilizado para darle sentido a los id.
- Perspectiva: **“Resuelto o no”**:  
“id\_solucion” de la tabla “soluciones”. Es el id de cada solucion, usado para luego agrupar.  
“descripcion” de la tabla “soluciones”. En esta descripción se establece si un reclamo fue resuelto o no, con un “sí” o un “no”, respectivamente.

d) – *Modelo conceptual ampliado:*



### Paso 3 – Modelo lógico del DW:

a) – *Tipo del modelo lógico del Data Warehouse:* utilizamos un esquema estrella, es decir, podemos observar un hecho central desde distintas perspectivas.

b) – *Tablas de dimensiones:*

 dwservicios dimservicio

descr  
id\_servicio

 dwservicios dimregion

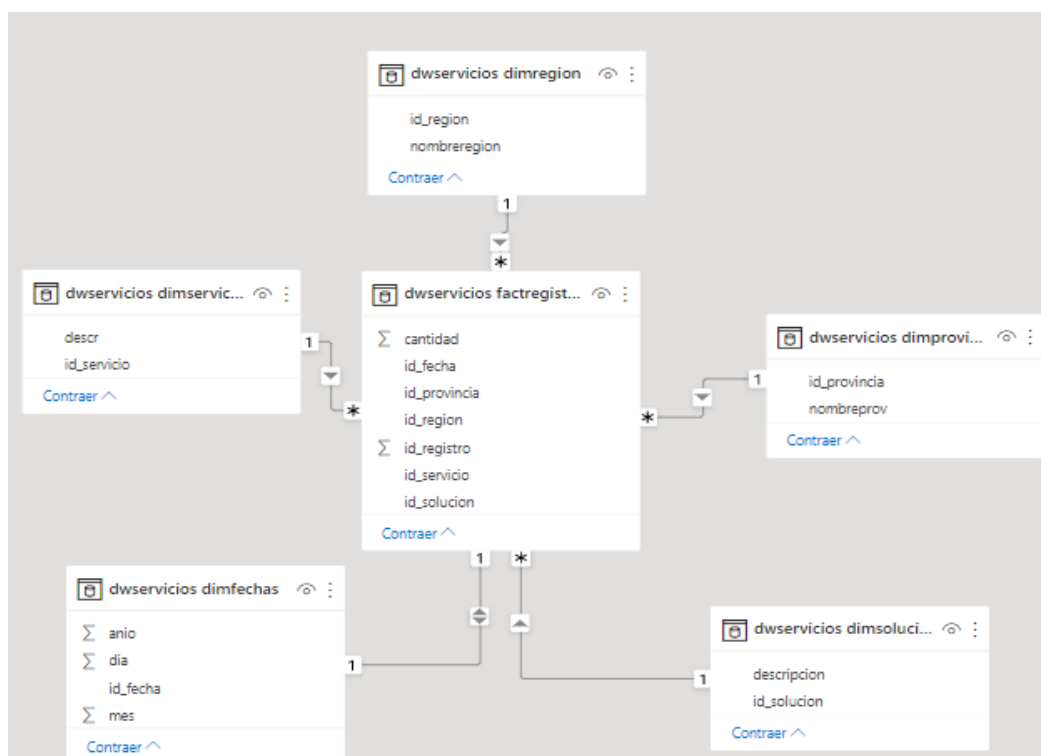
id\_region  
nombreregion

<div><div></div><div>dwservicios dimprovincias</div></div> <div><div>id_provincia</div><div>nombrepov</div></div>	<div><div></div><div>dwservicios dimfechas</div></div> <div><div><div>Σ</div><div>anio</div></div><div><div>Σ</div><div>dia</div></div><div>id_fecha</div><div><div>Σ</div><div>mes</div></div></div>
<div><div></div><div>dwservicios dimsoluciones</div></div> <div><div>descripcion</div><div>id_solucion</div></div>	

c) – Tabla de hechos:

<div><div></div><div>dwservicios factregist...</div><div><div></div><div></div></div></div>
<div><div><div>Σ</div><div>cantidad</div></div><div>id_fecha</div><div>id_provincia</div><div>id_region</div><div><div>Σ</div><div>id_registro</div></div><div>id_servicio</div><div>id_solucion</div></div>
<div>Contraer ^</div>

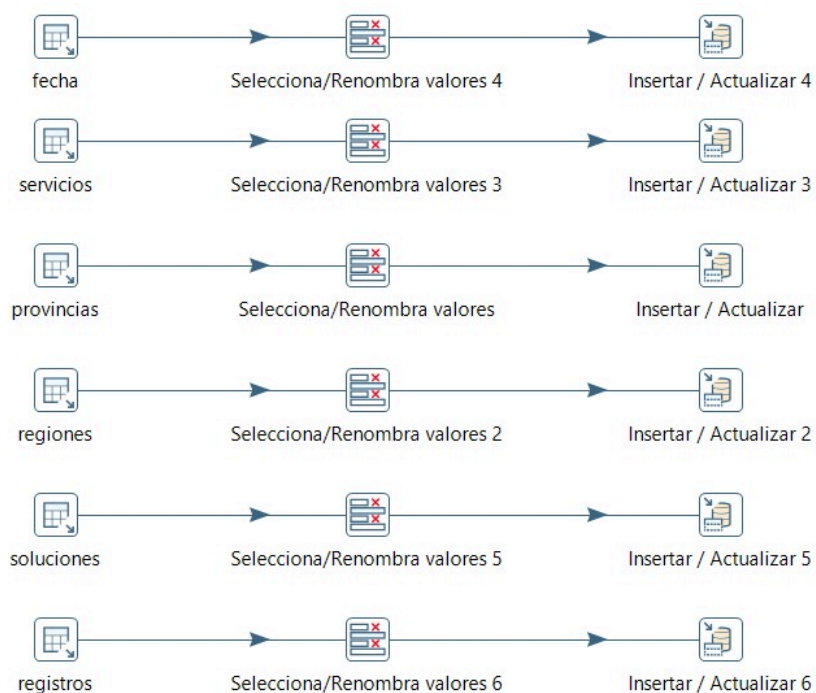
d) – Uniones:



#### **Paso 4 – Integración de datos:**

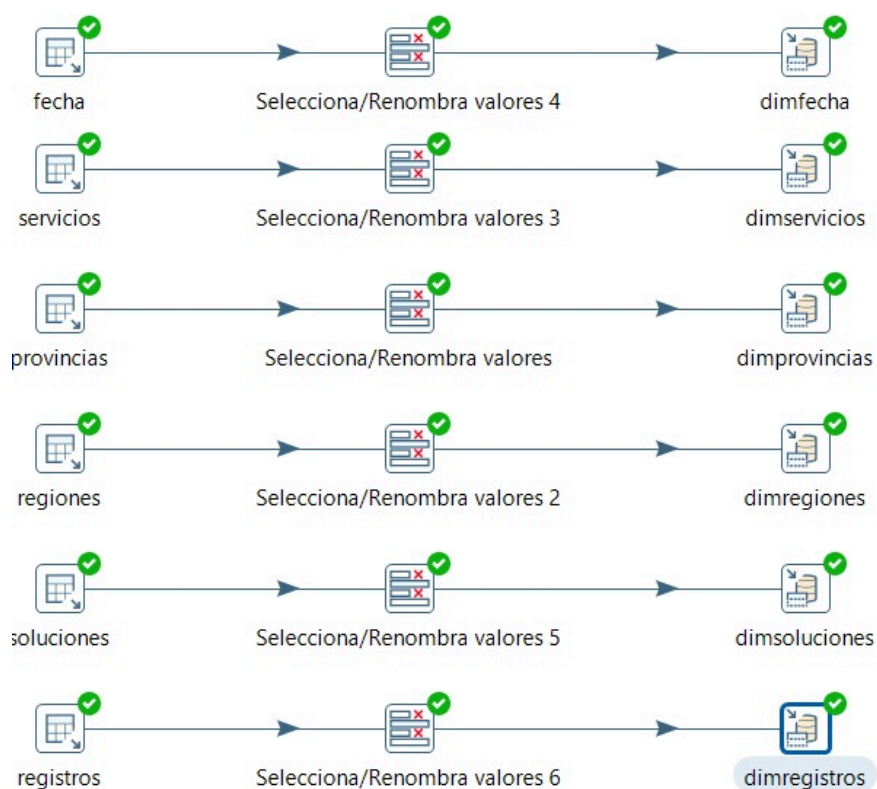
Como primera medida debemos crear el DW en la base de datos.

Una vez construido el DW vacío, se procede a poblarlo con datos utilizando procesos ETL (Extract, Transform and Load) utilizando la herramienta PENTAHO KETTLE.





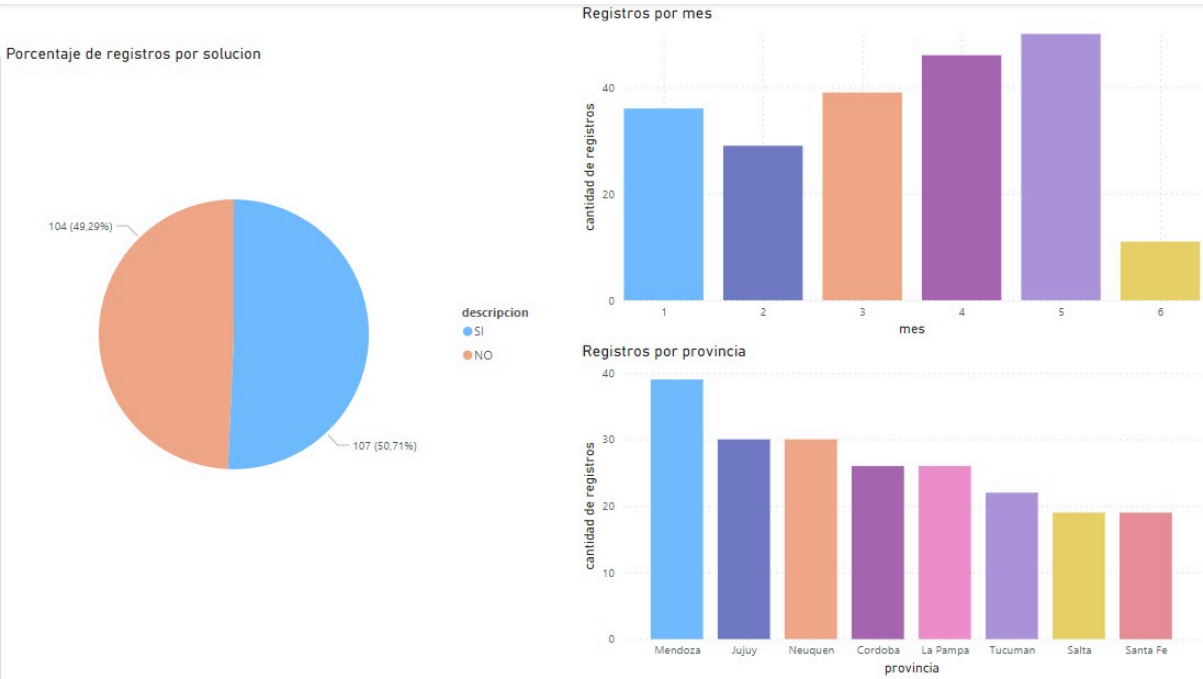
Luego probamos cada conexión para verificar que todo funcione correctamente.



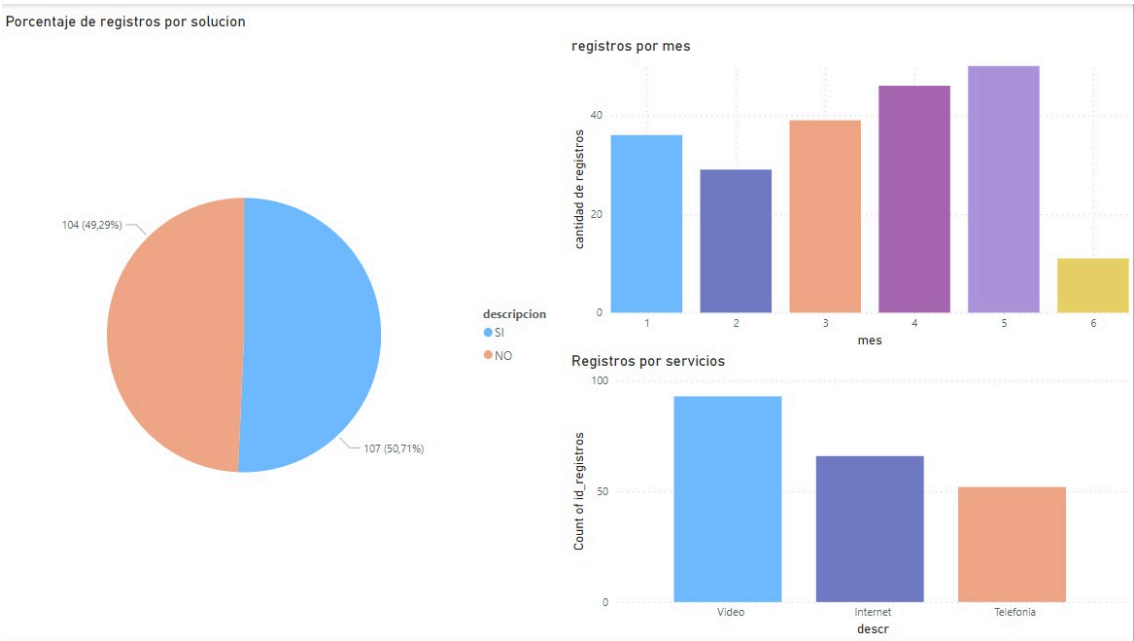
Luego de esto el Data Warehouse queda poblado con los datos obtenidos de la base de datos de Servicios.

Con respecto a la actualización de la base de datos es necesario actualizarla manualmente cada vez que se desee. No programamos ninguna actualización temporal.

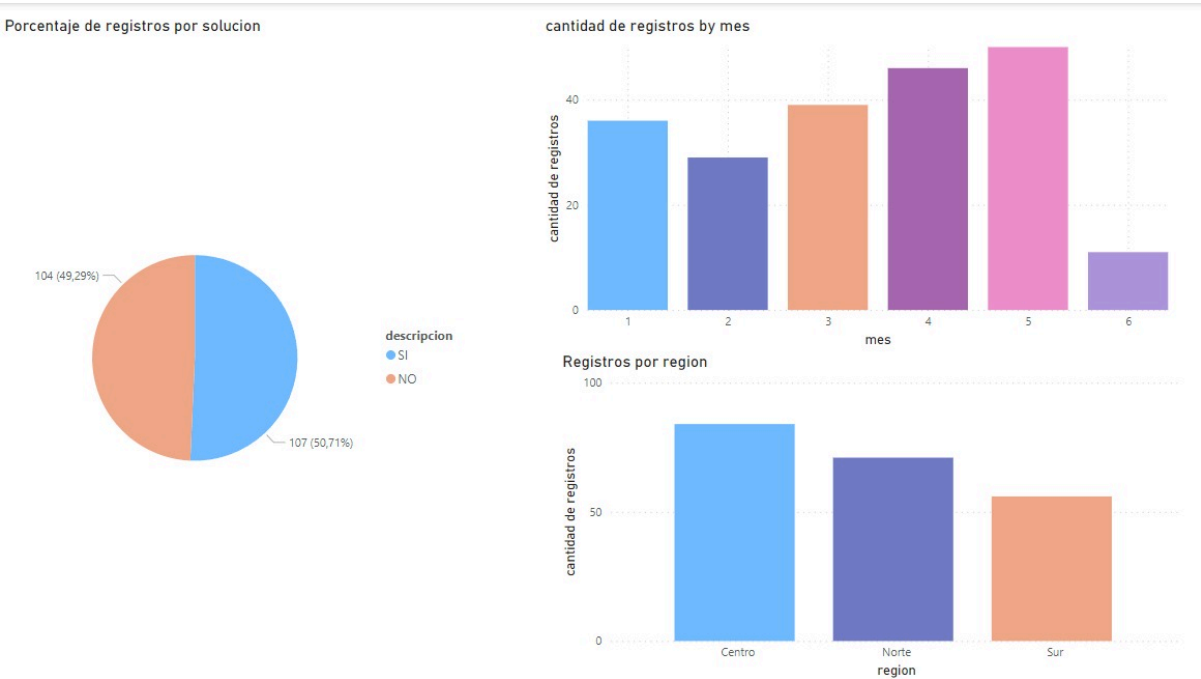
### Paso 5 – Visualización de los datos:



Se puede observar el porcentaje de soluciones efectivas y no efectivas en un gráfico de torta. A su vez podemos ver los registros divididos por provincias. En este caso la empresa podría identificar en que provincias existen más reclamos para realizar determinados cambios en esa provincia.



Luego hicimos un gráfico de los registros por servicios para que en dado caso la empresa vea que servicios funcionan mejor y cuales falta mejorar.



Como uno de los últimos pasos tenemos el grafico de la cantidad de registros dividida por región donde la empresa podría tomar medidas a partir de esto.

Una de las cosas más interesantes que tiene el PowerBI es que podemos clicar en cualquier parte de los gráficos y nos mostrara como se distribuye el resto de información con respecto a eso. Por ejemplo, si en la parte de cantidad de registros por mes hacemos click en cada uno de los meses nos mostrara estos gráficos expuestos anteriormente pero solo para el mes clickeado. También por ejemplo podemos hacer click en la parte azul del grafico de torta y nos mostrara todos los gráficos solo teniendo en cuenta la parte que clickeamos.

