



Almacenes de datos espaciales y SOLAP utilizando software libre

Leonardo Pandolfi González



• Agenda

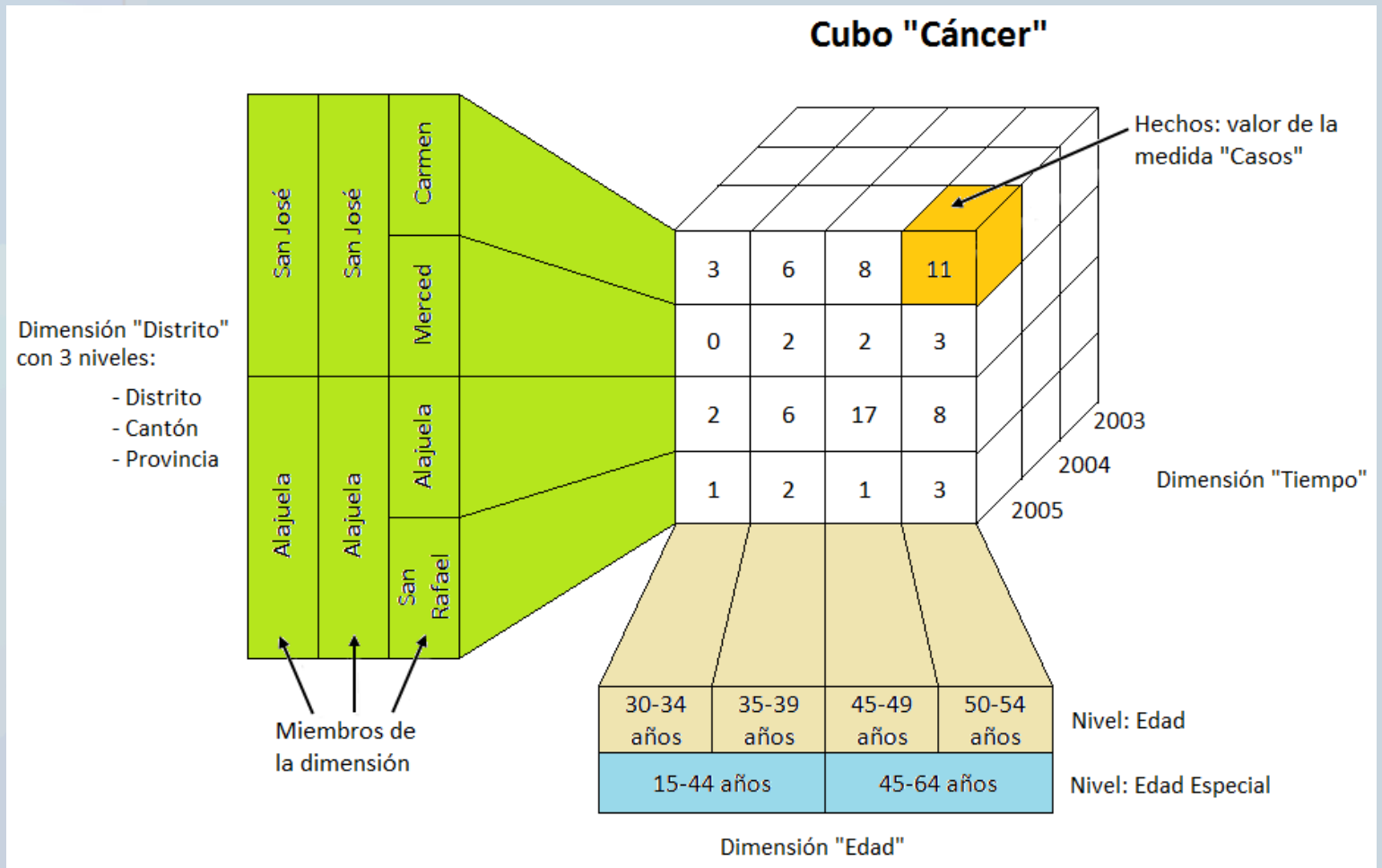
- ¿Por qué SOLAP?
- Modelo Multidimensional
- Operaciones OLAP
- Arquitectura
- Esquema conceptual (caso práctico)
- Definición de los cubos (esquema XML)
- Despliegue de operaciones
- Trabajo futuro
- Conclusiones



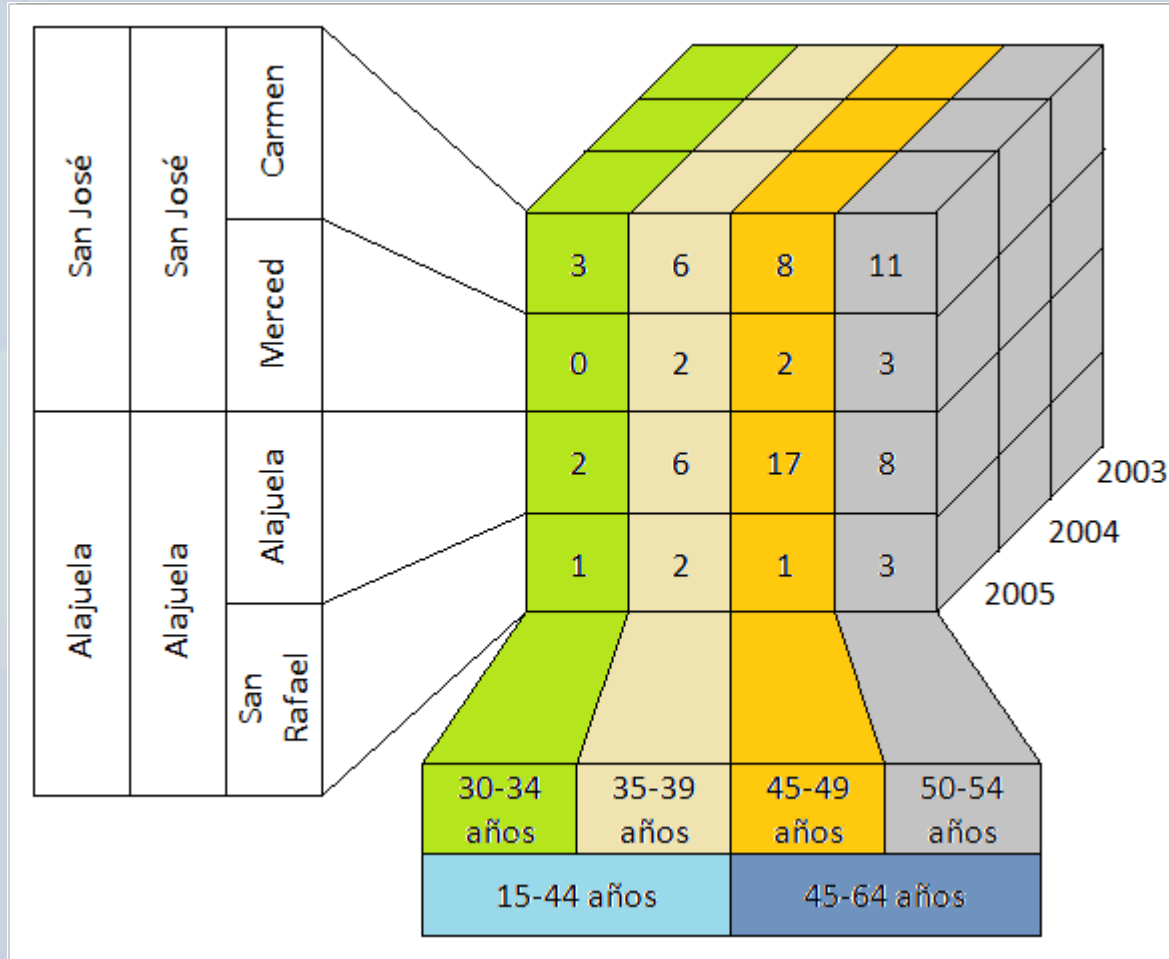
• ¿Por qué SOLAP?

- Spatial On-line Analytical Processing
- Despliegue simultáneo y manipulación dinámica de datos espaciales
 - Utilización de mapas, tablas y gráficos en forma sincronizada
- Integración de datos espaciales y convencionales
 - ¿Cuáles provincias registraron la mayor incidencia de cáncer de próstata entre el año 2000 y el 2005?
- Carácter interdisciplinario
- Interfaz gráfica intuitiva

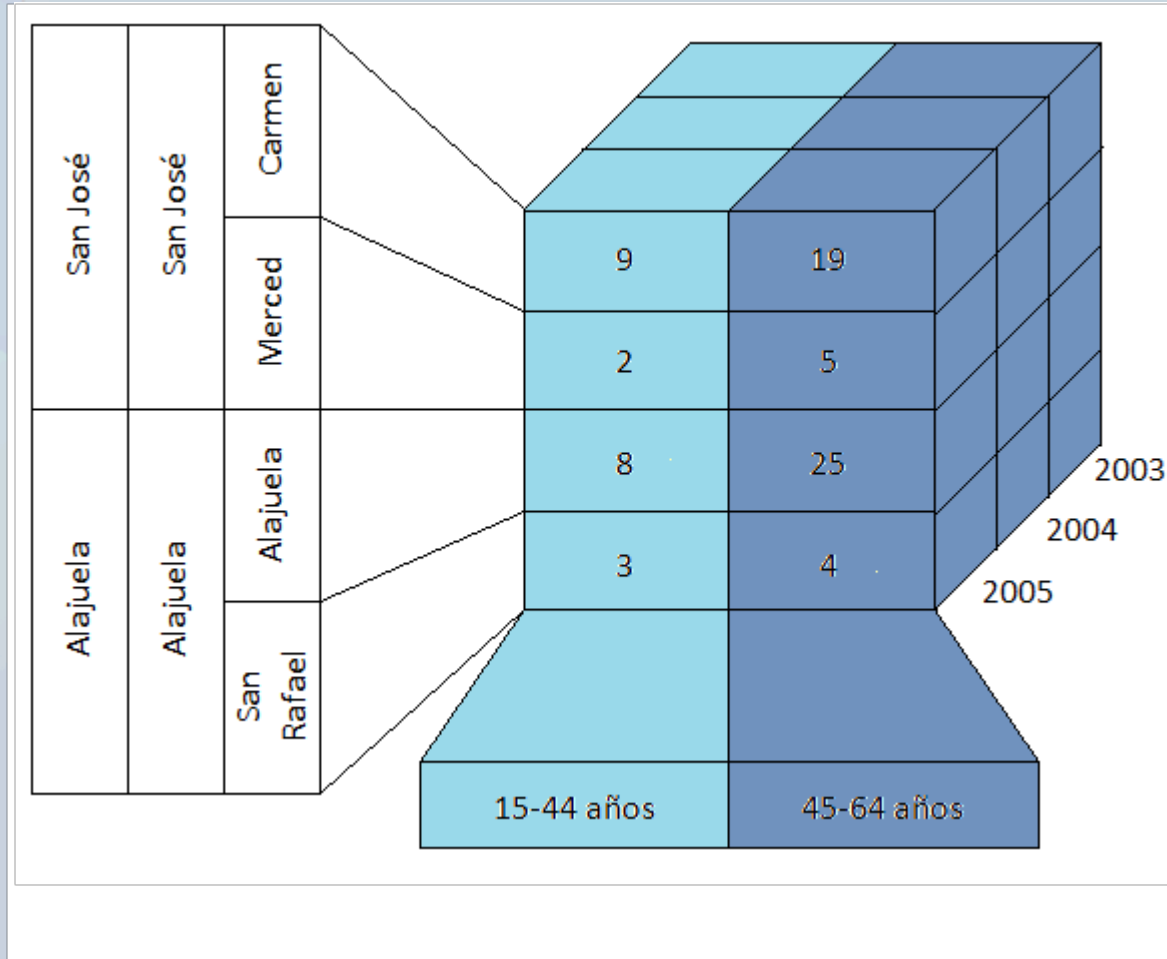
• Modelo Multidimensional



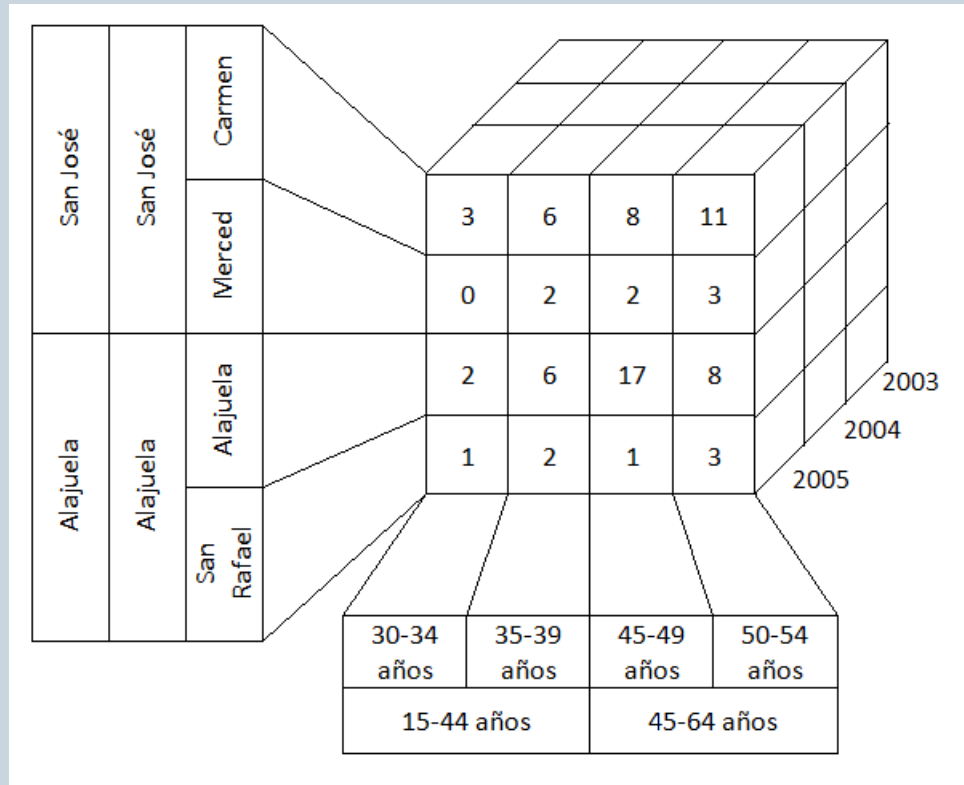
Operaciones OLAP: Drill-down



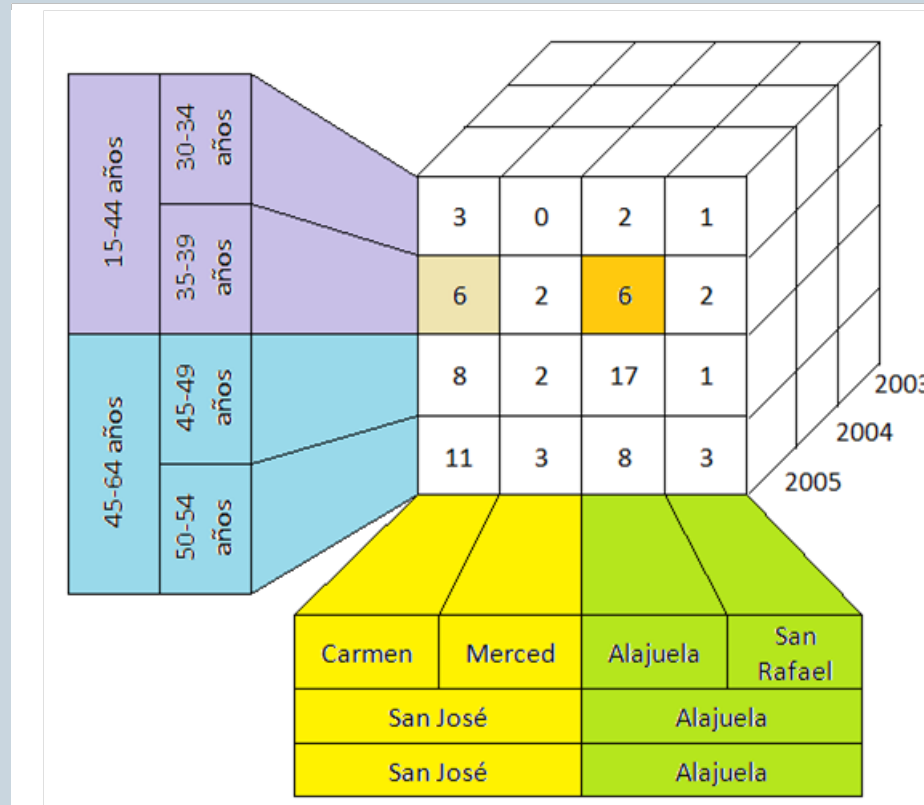
Operaciones OLAP: Roll-up



• Operaciones OLAP: Slice-and-dice



• Operaciones OLAP: Pivot



•Arquitectura

Capa de presentación

GeoOLAP
(Despliegue de tablas, gráficos y mapas)

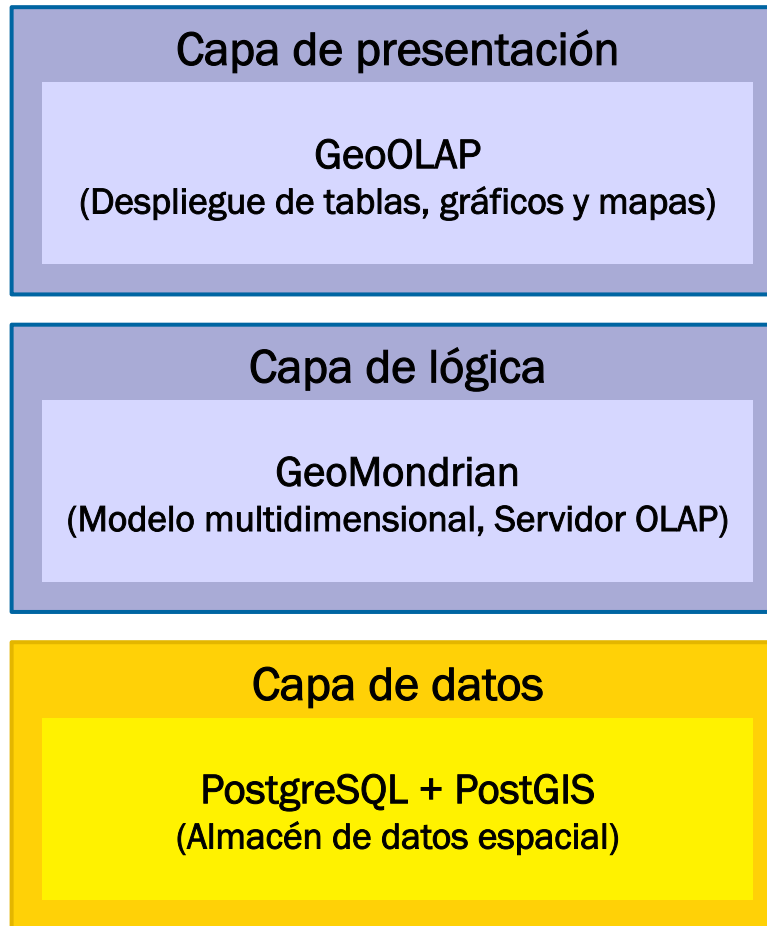
Capa de lógica

GeoMondrian
(Modelo multidimensional, Servidor OLAP)

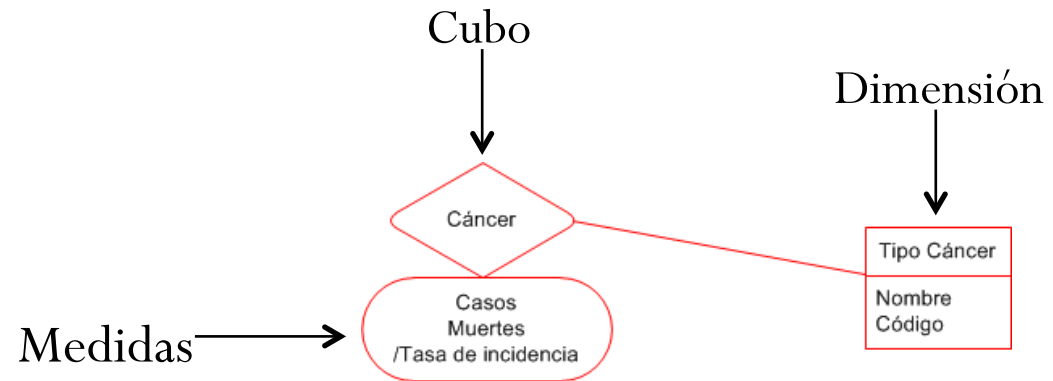
Capa de datos

PostgreSQL + PostGIS
(Almacén de datos espacial)

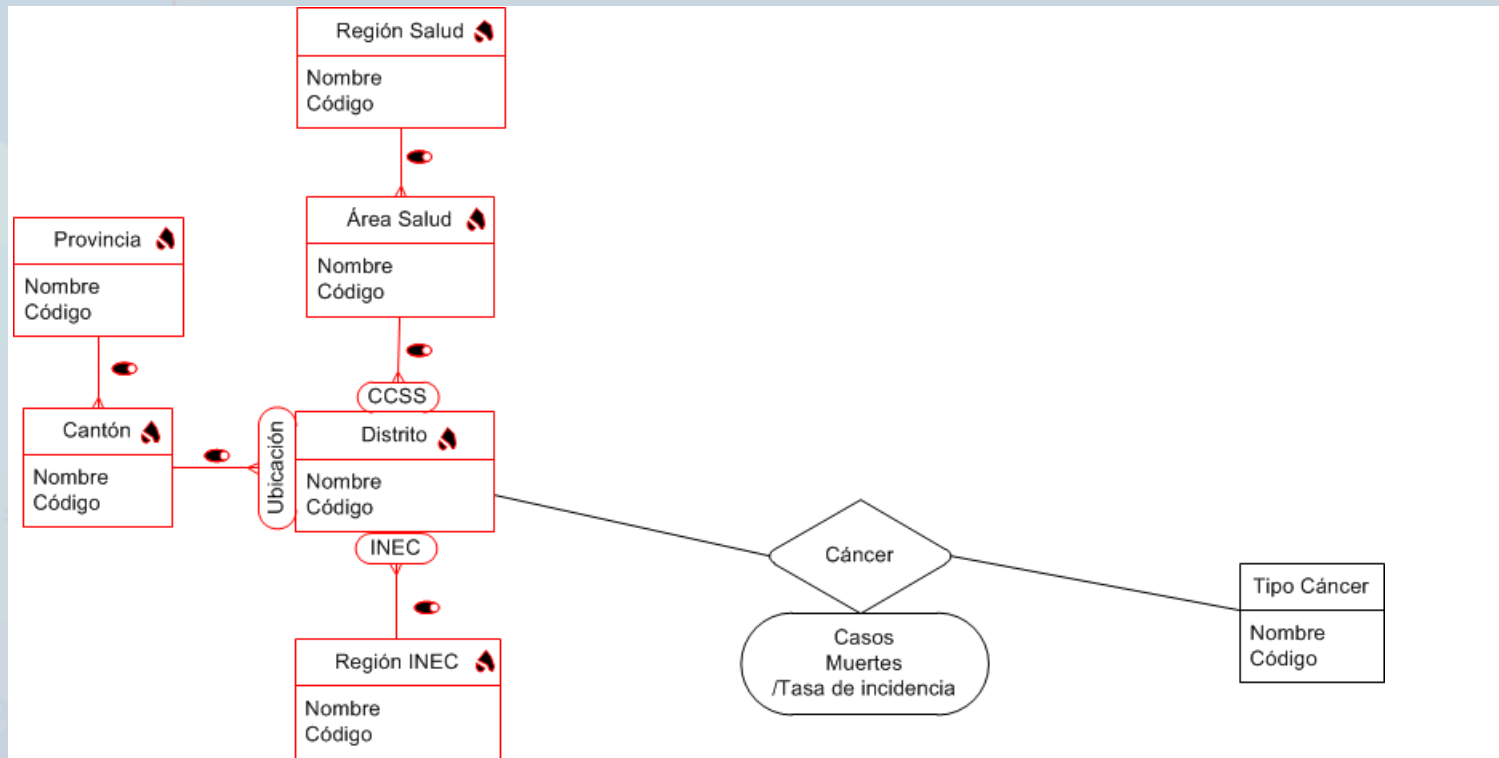
•Capa de datos



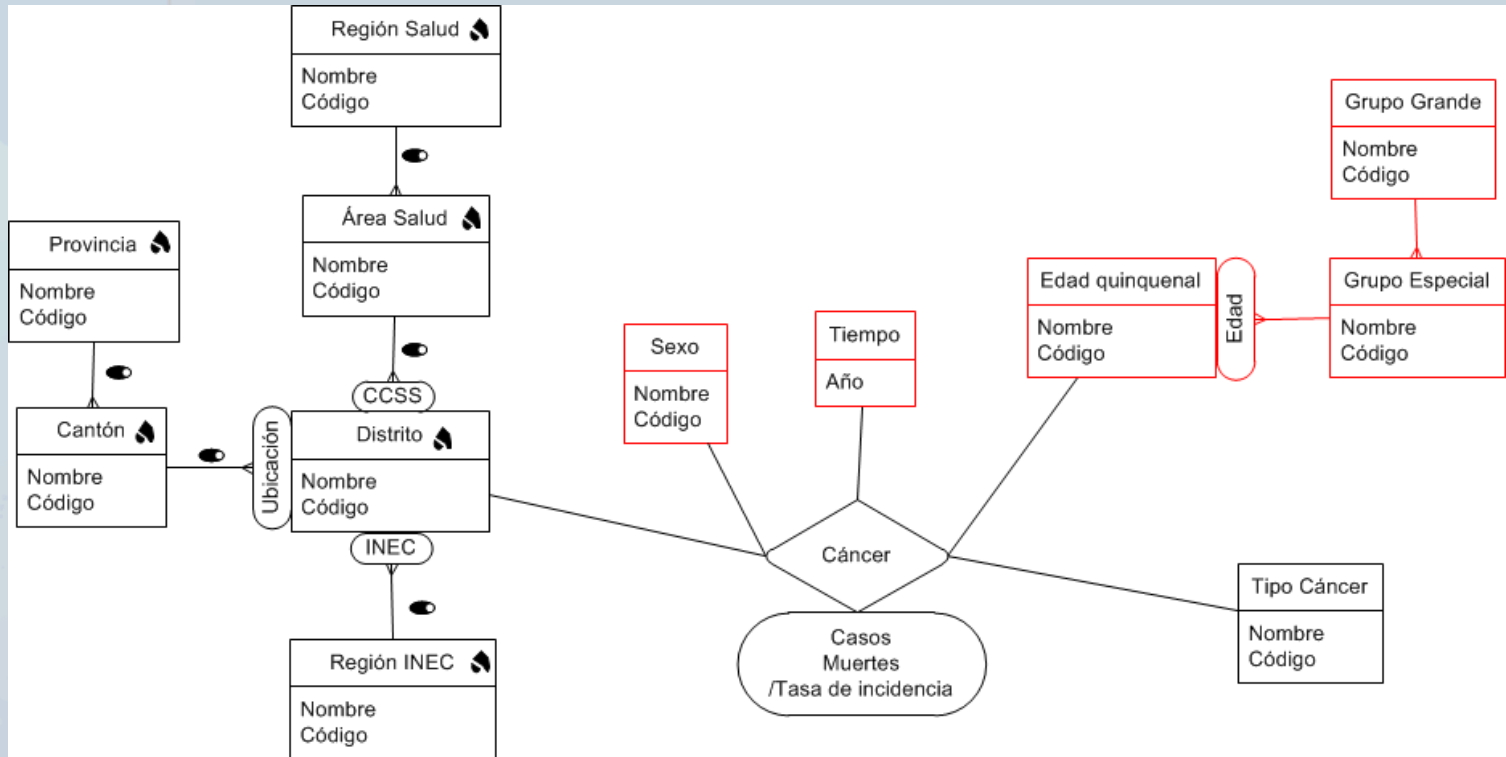
• Esquema conceptual (caso práctico)



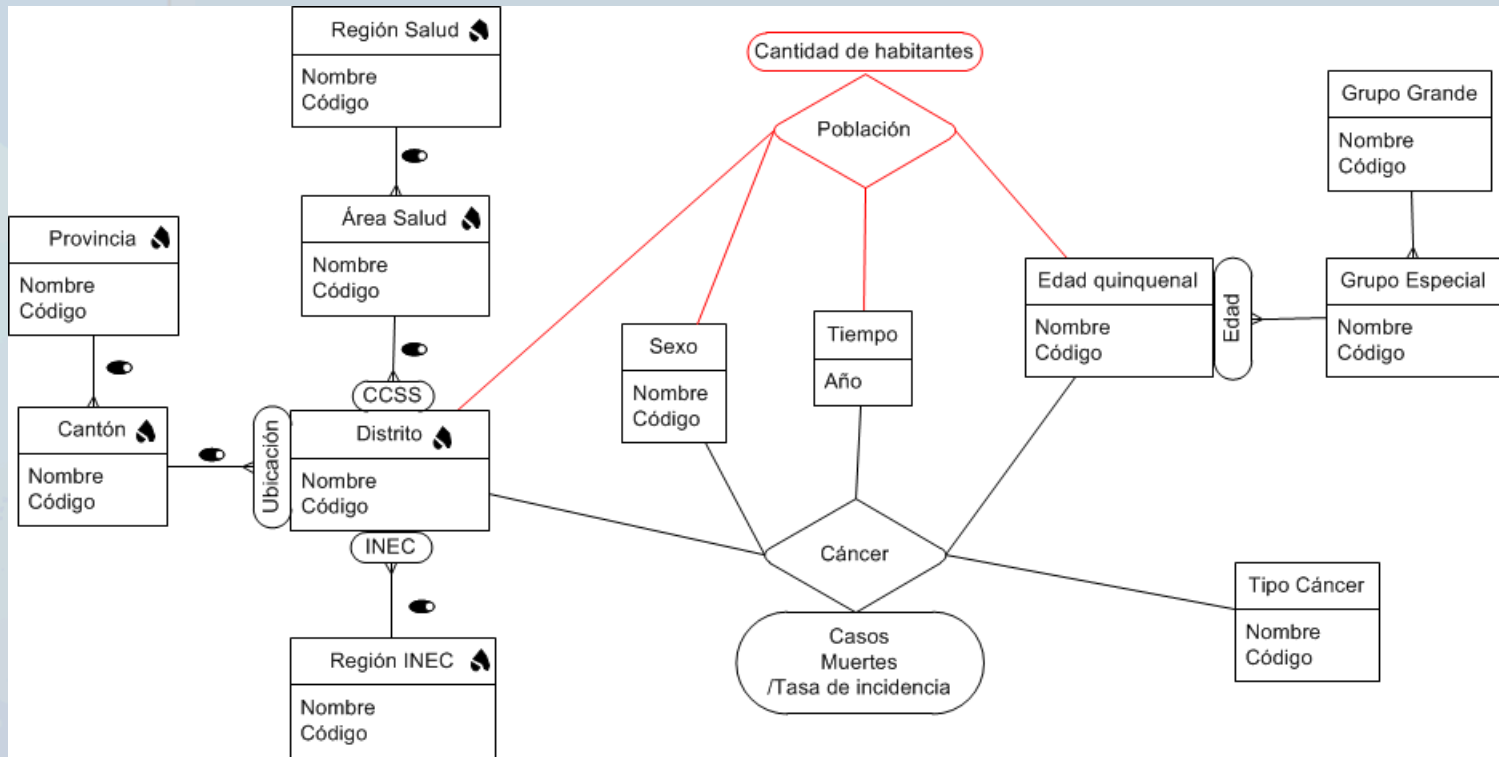
• Esquema conceptual (caso práctico)



• Esquema conceptual (caso práctico)



• Esquema conceptual (caso práctico)



•Capa de lógica

Capa de presentación

GeoOLAP
(Despliegue de tablas, gráficos y mapas)

Capa de lógica

GeoMondrian
(Modelo multidimensional, Servidor OLAP)

Capa de datos

PostgreSQL + PostGIS
(Almacén de datos espacial)

•Características de GeoMondrian

Desarrollador	Spatialytics
Base	Mondrian (Pentaho)
Lenguaje de consulta	MDX
Definición de cubos	Esquema XML
Almacenamiento de resultados	Memoria RAM

Definición de los cubos (esquema XML)

```
<Schema>
  <Cube name="Cancer">
    <Table name="cancer_fact"/>
    <Dimension name="Distritos" foreignKey="codigo_distrito">
      <Hierarchy hasAll="false" primaryKey="codigo_distrito">
        <Table name="geografia"/>
        <Level name="Provincia" column="nombre_provincia" type="String" />
        <Property name="geom" column="geometria_provincia" type="Geometry" />
        <Level name="Cantón" column="nombre_canton" type="String" />
        <Property name="geom" column="geometria_canton" type="Geometry" />
        <Level name="Distrito" column="nombre_distrito" type="String" />
        <Property name="geom" column="geometria_distrito" type="Geometry" />
      </Hierarchy>
    </Dimension>
    <Measure name="Casos" column="casos" aggregator="sum"/>
    <Measure name="Muertes" column="muertes" aggregator="sum"/>
  </Cube>
</Schema>
```

Asociación de dimensiones
Una jerarquía de dimensiones
Física de las dimensiones
Propiedades de las dimensiones
Geometría de las dimensiones
Indicadores de las dimensiones
Definición de medidas y su función de agregación



•Capa de presentación

Capa de presentación

GeoOLAP

(Despliegue de tablas, gráficos y mapas)

Capa de lógica

GeoMondrian

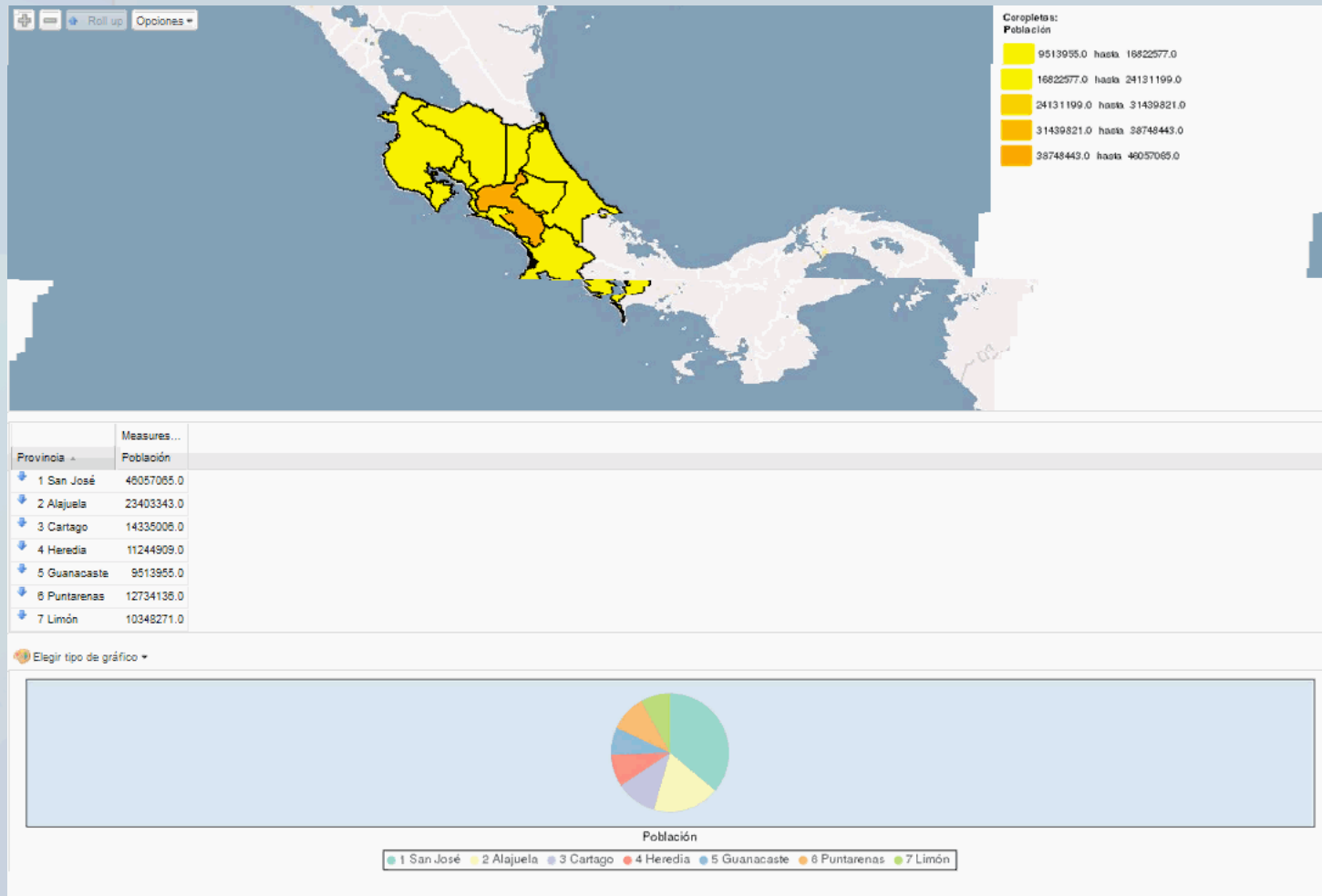
(Modelo multidimensional, Servidor OLAP)

Capa de datos

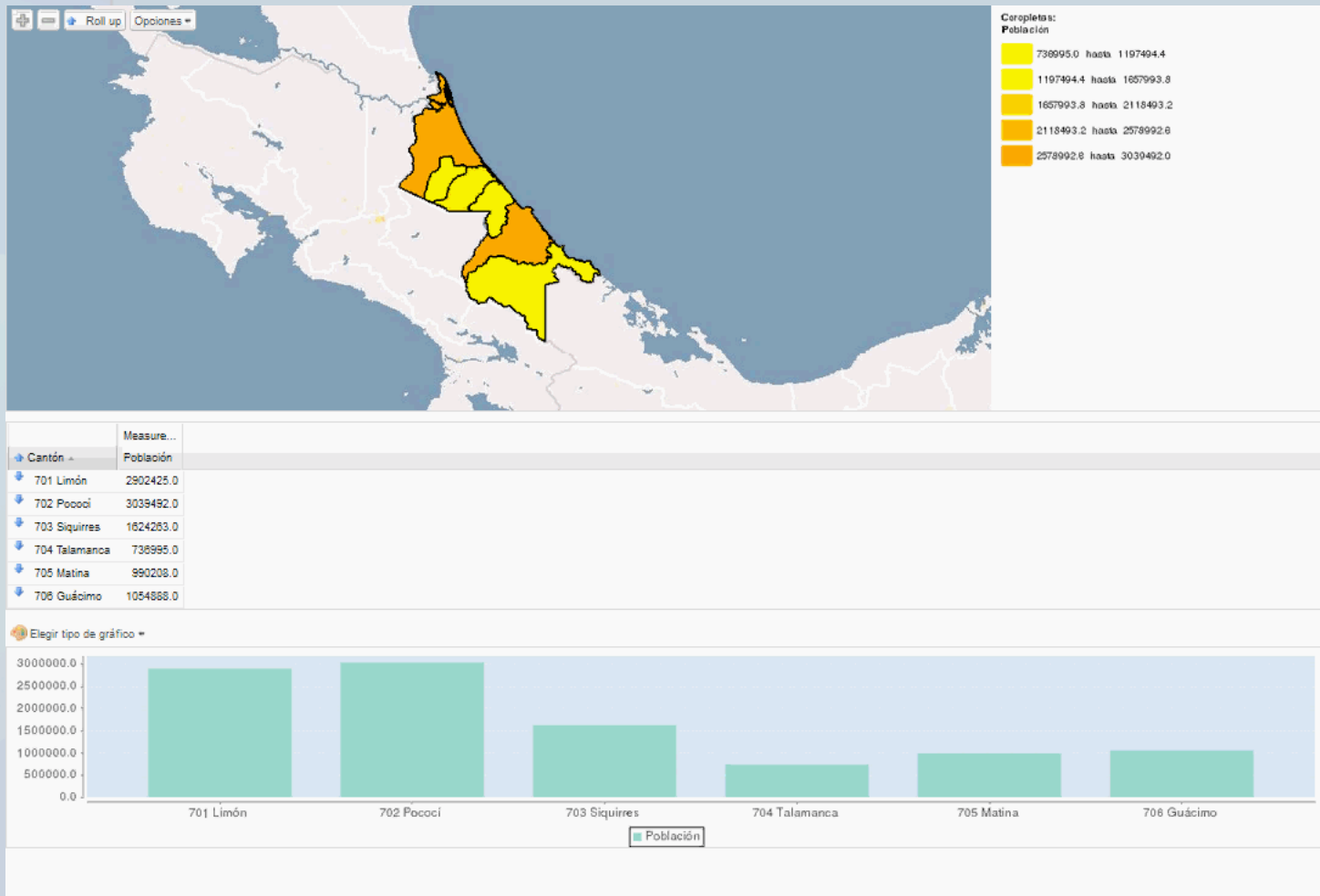
PostgreSQL + PostGIS

(Almacén de datos espacial)

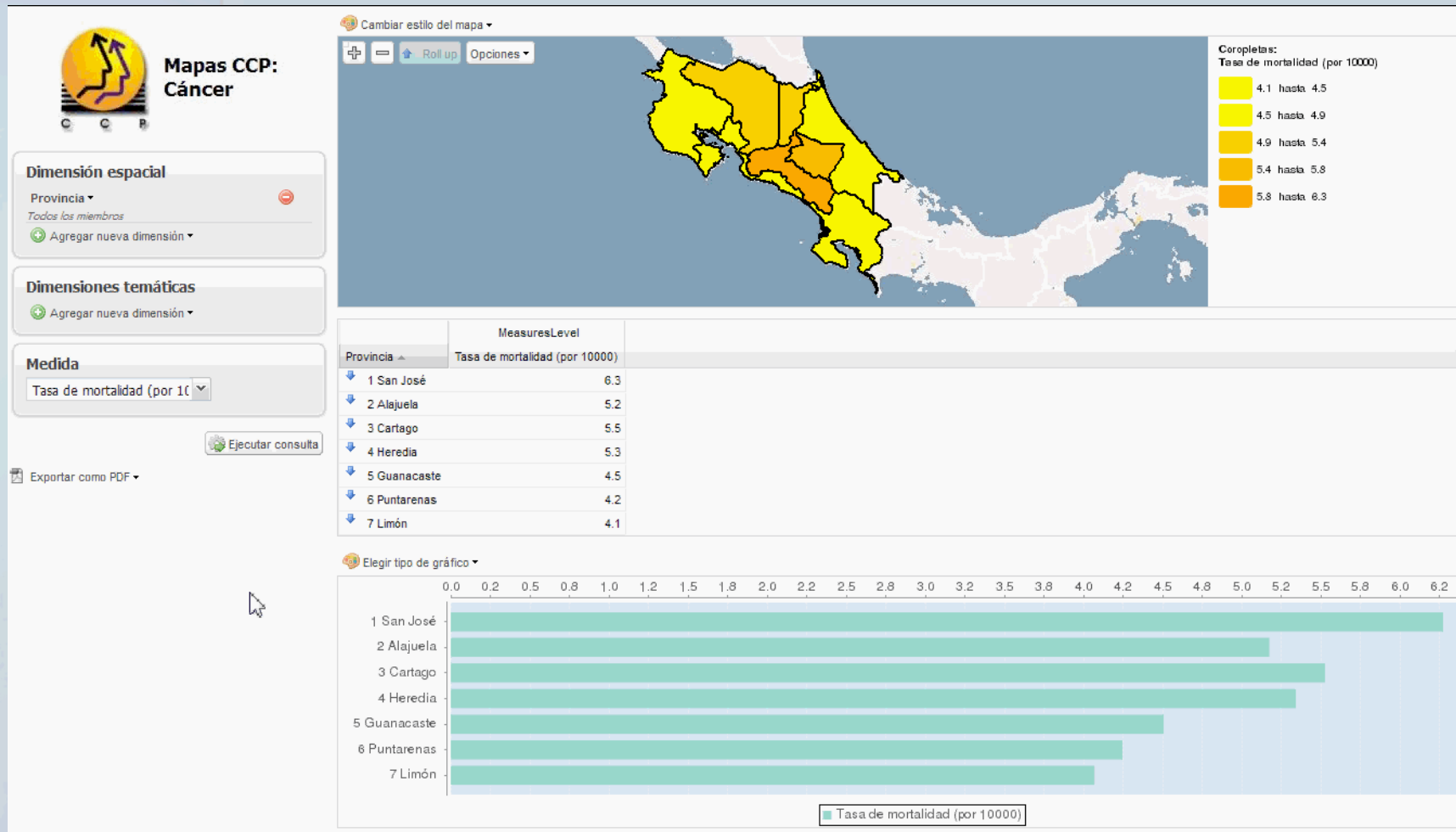
• Despliegue de operaciones: drill-down



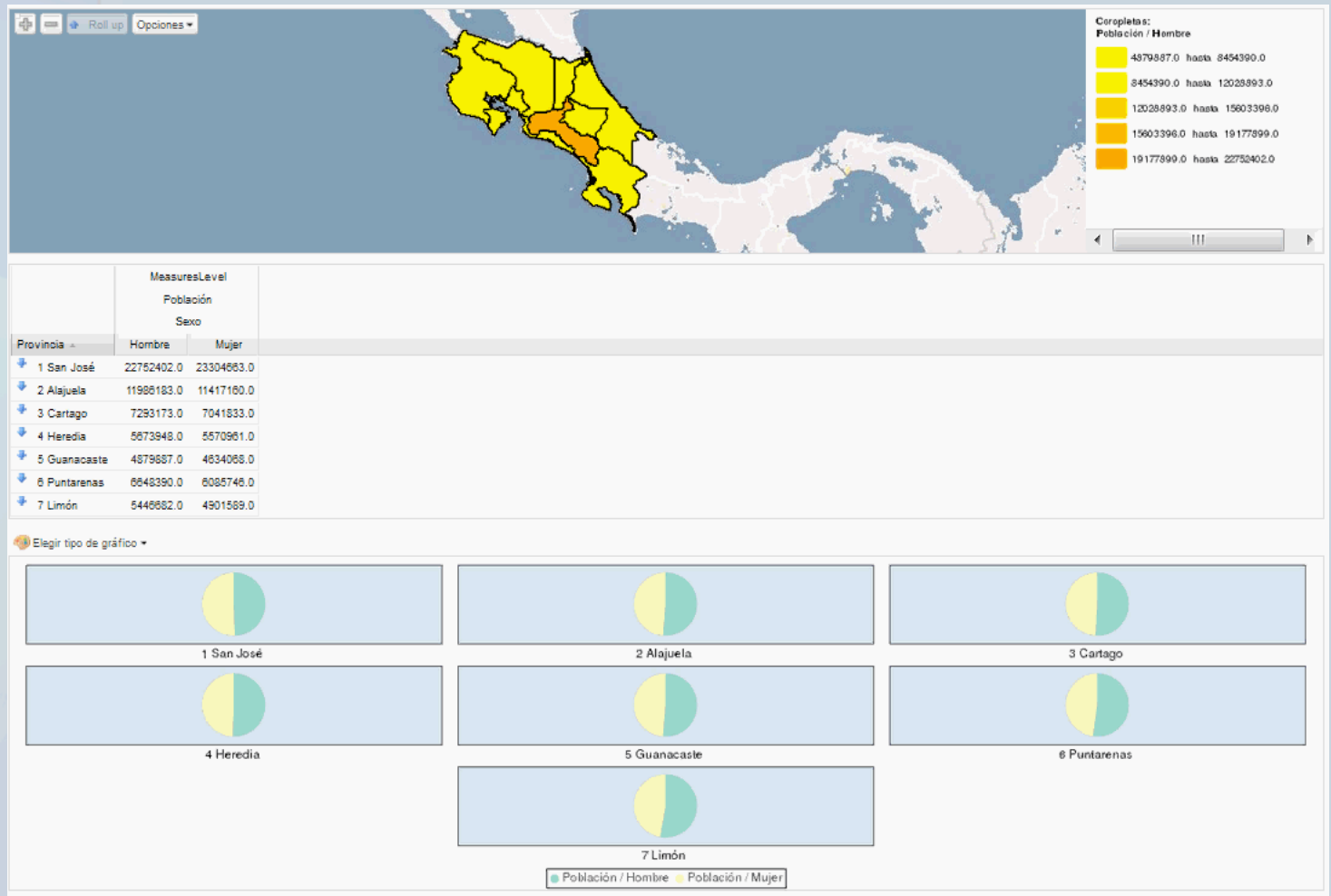
• Despliegue de operaciones: roll-up



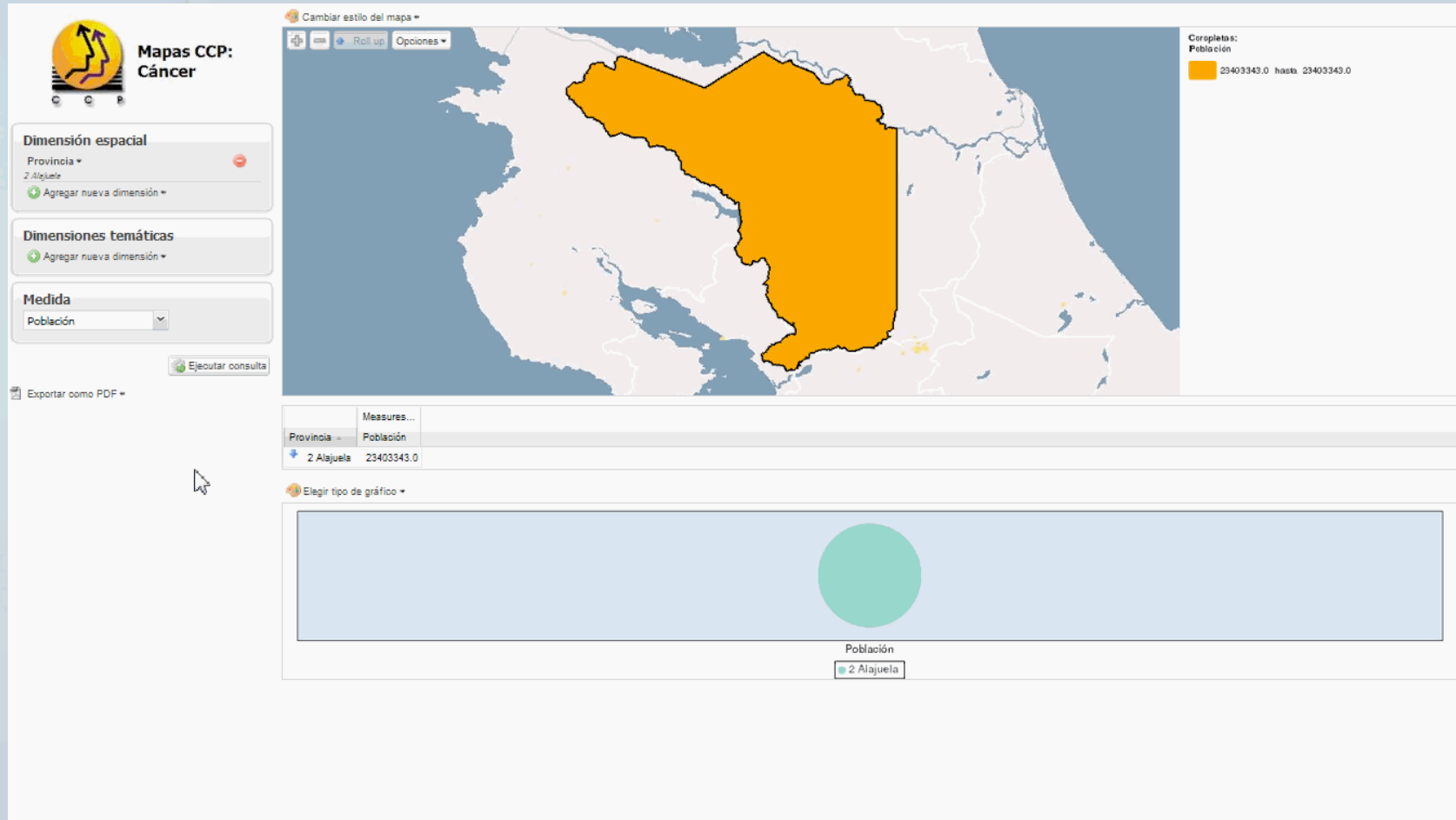
• Despliegue de operaciones: slice-and-dice



• Despliegue de operaciones: pivot



•Despliegue de operaciones: combinación de dimensiones espaciales



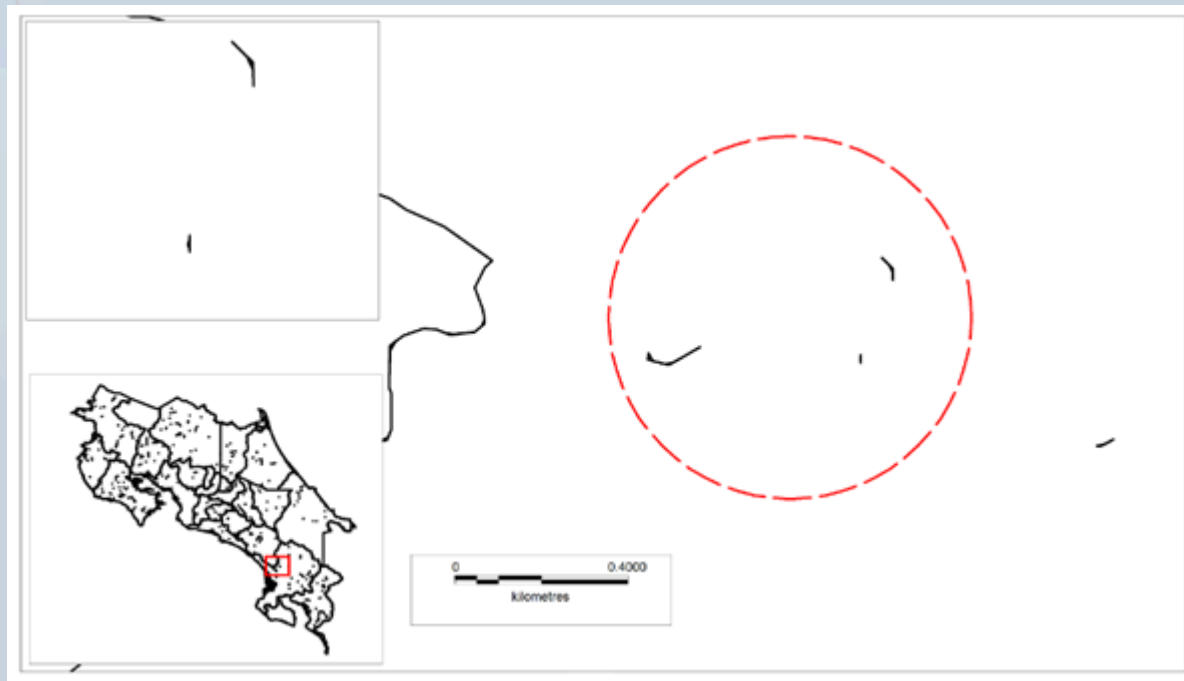


• **Características adicionales de GeoOLAP**

- Cubos virtuales
- Medidas calculadas
 - Simples: cambio de moneda
 - Complejas: tasa de mortalidad por tipo de cáncer
- Medidas por aditividad
 - Aditiva
 - Semiaditiva
 - No aditiva

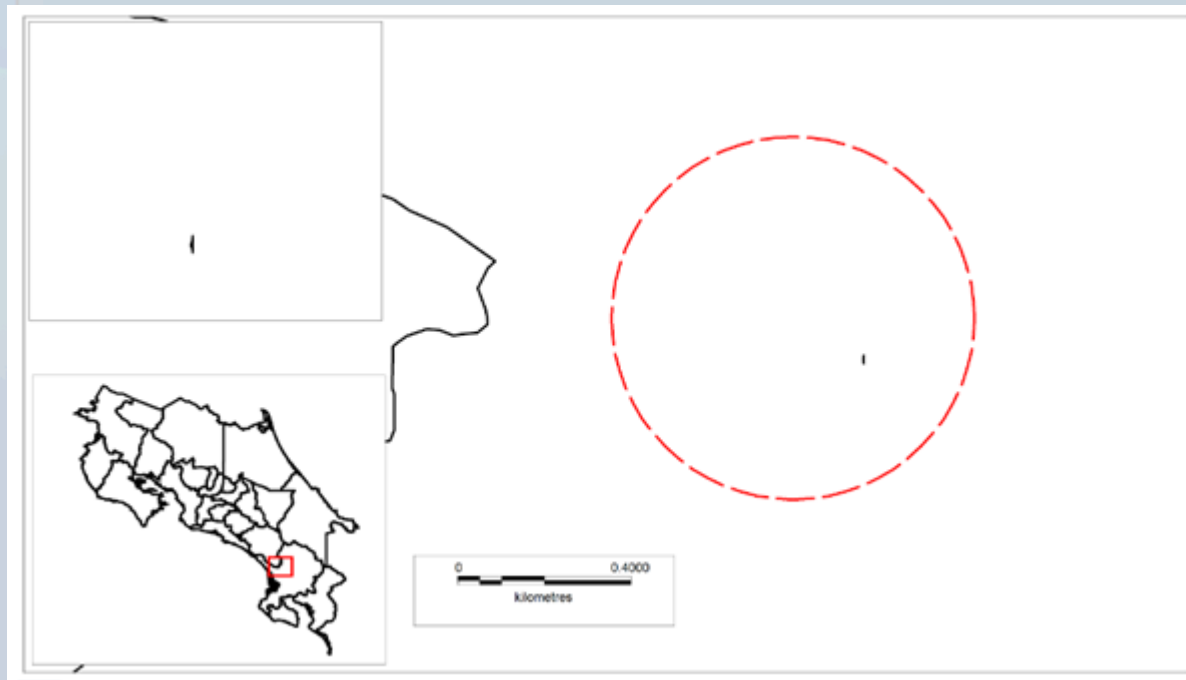
• Trabajo futuro

- Simplificación de geometrías
 - Combinación de datos espaciales
- Verificación de orientación y relaciones topológicas



• Trabajo futuro

- Simplificación de geometrías
 - Combinación de datos espaciales
- Verificación de orientación y relaciones topológicas





• **Trabajo futuro (2)**

- Medidas espaciales
- Vistas materializadas
- Nueva definición de los cubos en Mondrian 4
 - Reescribir o transformar el esquema
- Mejoramiento de la interfaz
 - Sencilla y funcional: amigable con el usuario
 - Múltiples medidas en una misma consulta
 - Utilización de funciones espaciales



• Conclusiones

- Exploración de datos espaciales para ampliar posibilidades de análisis en la toma de decisiones
- Nuevas opciones a los usuarios no expertos en OLAP y SOLAP
- Interfaz es fundamental
- Actualmente, opciones SOLAP *open-source* son limitadas
- Espacio para el desarrollo de una herramienta SOLAP más funcional y gratuita
- Más detalles:
 - D. Bogantes y L. Pandolfi, “Spatial Data Warehouses and SOLAP Using Open-Source Tools”. Simposio de Sistemas Innovadores de Datos, CLEI 2013, Vargas, Venezuela.