

UNIVERSIDAD DE COSTA RICA

Facultad de Ingeniería

Escuela de Ciencias de la Computación e Informática

Bases de Datos II, CI-1314

Profesora:

Elzbieta Malinowski Gajda

Etapa #2 del Proyecto

Estudiantes:

Ivannia Alvarado González, B10273

Oscar Castro Espinoza, B11616

María José Herrera Barquero, B23258

22 de Noviembre, 2014

1. La especificación y justificación de los cambios realizados al esquema conceptual desarrollado en la primera etapa del proyecto (si aplica).

Pese a que se intentó en la medida de lo posible, adaptar los datos que se tenían para ajustar la base de datos del proyecto al esquema conceptual inicialmente aprobado por la profesora, la dificultad de integración de algunos datos hizo que algunas relaciones de participación total no fueran posibles de implementar como tales. Es por esto que en algunos casos se eliminan del esquema conceptual estas relaciones y debido a esto, se propone un nuevo esquema. El nuevo esquema se aprecia en la Figura 2, y el esquema conceptual inicial es el que se aprecia en la Figura 1.



Figura : esquema conceptual aprobado inicialmente.

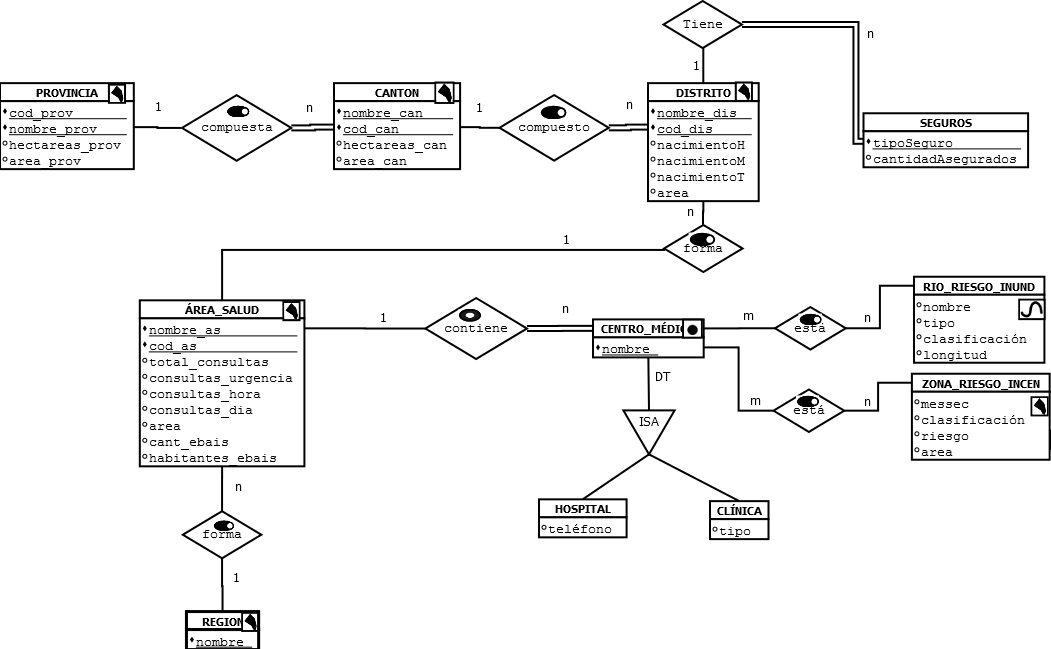


Figura : esquema conceptual modificado para la última etapa.

1. Por cada tabla:
2. El código o pantallazo de la creación de tabla con todas las restricciones indicadas en el esquema lógico.
3. El código del (de los) disparador(es) (*trigger(s)*).
4. Breve descripción de las correspondencias entre atributos de la tabla creada, archivos shape y archivos de datos convencionales y de las transformaciones necesarias para la limpieza y corrección de datos.
5. El código de inserción de datos con las transformaciones especificadas en el punto anterior.
6. Código de creación de índices o especificación de los parámetros usados (entiendo su significado). Los parámetros (especialmente el *bounding box*) tienen que ser adecuados para la extensión espacial de datos usados.
7. Breve referencia a problemas encontrados.

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) |  |
| Descripción de las correspondencias y transformaciones | |  |  |  | | --- | --- | --- | | **Nombre de la tabla:** provincias | **Archivo Shape:** provincias2008crtm05 | Descripción: | | cod\_prov | COD\_PROV | Código de la provincia. | | nombre\_prov | PROVINCIA | Nombre de la provincia. | | area\_prov | NA | Área de la geometría de la provincia. | | geom | SIN NOMBRE | Se aplica las transformaciones para asegurar que las geometrías sean válidas. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | FALTA |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) |  |
| Descripción de las correspondencias y transformaciones | |  |  |  | | --- | --- | --- | | **Nombre de la tabla:** cantones | **Archivo Shape:** cantones2008ctm05 | Descripción: | | cod\_can | CODNUM | Código del cantón. | | cod\_prov | NA | Código de la provincia a la que pertenece el cantón. | | nombre\_can | NCANTON | Nombre del cantón. | | area\_can | NA | Área de la geometría del cantón. | | geom | SIN NOMBRE | Se aplica las transformaciones para asegurar que las geometrías sean válidas. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | FALTA |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) |  |
| Descripción de las correspondencias y transformaciones | |  |  |  |  | | --- | --- | --- | --- | | **Nombre de la tabla:** distritos | **Archivo Shape:** distritos2008crtm05 | **Archivo Excel:** Total de nacimientos por sexo, según provincia, cantón y distrito de residencia de la madre2008 | Descripción: | | cod\_dis | CODNUM | NA | Código del distrito. | | cod\_can | NA | NA | Código del cantón al que pertenece el distrito. | | nombre\_dis | NCANTON | NA | Nombre del distrito. | | area\_dis | NA | NA | Área de la geometría del distrito. | | nacimientoT | NA | Total | Nacimientos totales en el distrito en el año 2008. | | nacimientoH | NA | Hombres | Nacimientos masculinos en el distrito en el año 2008. | | nacimientoM | NA | Mujeres | Nacimientos femeninos en el distrito en el año 2008. | | geom | SIN NOMBRE | NA | Se aplica las transformaciones para asegurar que las geometrías sean válidas. Se aplica la unificación de geometrías. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | FALTA |
| Problemas encontrados | Hubo que hacer unión de geometrías de distintas tuplas y además, solucionar problemas de geometrías con coordenadas equivocadas. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No son necesarios para esta tabla. |
| Descripción de las correspondencias y transformaciones | |  |  |  | | --- | --- | --- | | **Nombre de la tabla:** region | **Archivo Excel:** RelacionAreasSaludRegiones | Descripción: | | id | NA | Identificador de tupla de región. | | nombre\_re | SIN NOMBRE | Nombre de la región. | | geom | NA | Se aplica la unión de geometrías desde la tabla de areas\_salud. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | FALTA |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No pudieron implementarse para esta tabla, por razones de la forma en la que están los datos. |
| Descripción de las correspondencias y transformaciones | |  |  |  |  | | --- | --- | --- | --- | | **Nombre de la tabla:** areas\_salud | **Archivo Shape:** AreaSalud\_crtm05 | **Archivo Excel:** InfoAreasSalud | Descripción: | | id |  | NA | Identificador de tupla de área de salud. | | nombre\_as |  | NA | Nombre del área de salud. | | id\_region | NA |  | Llave foránea a la región a la que pertenece el área de salud. | | total\_consultas | NA | Total consultas en el Area de Salud | Promedio de consultas total en esa área de salud. | | consultas\_urgencia | NA | Consultas en urgencias en el área de salud | Promedio de consultas en urgencia en esa área de salud. | | consultas\_hora | NA | Consultas por hora en el área de salud | Promedio de consultas por hora en esa área de salud. | | consultas\_dia | NA | Consultas por día en el área de salud | Promedio de consultas por día en esa área de salud. | | area | NA | NA | Área de la geometría del área de salud. | | cant\_ebais | NA | Cantidad de EBAIS en el área de salud | Promedio de la cantidad de ebais pertenecientes a esa área de salud. | | habitantes\_ebais | NA | Habitantes por EBAIS | Promedio de la cantidad de habitantes que atienden los ebais en esa área de salud. | | geom |  | NA | Se aplica las transformaciones para asegurar que las geometrías sean válidas. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | FALTA |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) |  |
| Descripción de las correspondencias y transformaciones | |  |  |  | | --- | --- | --- | | **Nombre de la tabla:** centros\_medicos | **Archivo Shape:** hospitales2008crtm05, clinicas2008crtm05 | Descripción: | | id | NA | Identificador de la tupla en centros médicos. | | nombre | NOMBRE (en ambos) | Nombre de la clínica o el hospital. | | id\_as | NA | Llave foránea al área de salud a la que pertenece el centro médico. | | geom | SIN NOMBRE | Se aplica las transformaciones para asegurar que las geometrías sean válidas. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | FALTA |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No son necesarios para esta tabla, pues en la forma de insertar la llave foránea nunca queda nula. |
| Descripción de las correspondencias y transformaciones | |  |  | | --- | --- | | **Nombre de la tabla:** hospitales | Descripción: | | id\_cm | Llave foránea al centro médico. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | No se requieren para esta tabla. |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No son necesarios para esta tabla, pues en la forma de insertar la llave foránea nunca queda nula. |
| Descripción de las correspondencias y transformaciones | |  |  |  | | --- | --- | --- | | **Nombre de la tabla:** provincias | **Archivo Shape:** clinicas2008crtm05 | Descripción: | | id\_cm | NA | Llave foránea al centro médico. | | tipo | TIPO | Tipo de clínica. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | No se requieren para esta tabla. |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No son necesarios para esta tabla. |
| Descripción de las correspondencias y transformaciones | |  |  |  | | --- | --- | --- | | **Nombre de la tabla:** riesgos\_inun | **Archivo Shape:** riesginundacionrtm05 | Descripción: | | id | NA | Indentificador de tupla para riesgos de inundación. | | nombre | NOMBRE | Nombre del río. | | tipo | TIPO | Tipo de fuente de agua. | | clasificacion | CLASIFICAC | Características del tipo de riesgo. | | longitud | NA | Longitud del río. | | geom | SIN NOMBRE | Se aplica las transformaciones para asegurar que las geometrías sean válidas. Se aplica la unión de geometrías de distintas tuplas. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | FALTA |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No son necesarios para esta tabla. |
| Descripción de las correspondencias y transformaciones | |  |  |  | | --- | --- | --- | | **Nombre de la tabla:** riesgos\_incen | **Archivo Shape:** riesgoincendiocrtm05 | Descripción: | | id | NA | Indentificador de tupla para riesgos de incendio. | | messec | MESSEC | Desconocido. | | clasificacion | CLASIFICAC | Características del riesgo. | | riesgo | RIESGO | Tipo de riesgo. | | area | NA | Área de la geometría del riesgo. | | geom | SIN NOMBRE | Se aplica las transformaciones para asegurar que las geometrías sean válidas. Se aplica la unión de geometrías de distintas tuplas. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | FALTA |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No son necesarios para esta tabla, pues en la forma de insertar las llaves foráneas nunca quedan nulas. |
| Descripción de las correspondencias y transformaciones | |  |  | | --- | --- | | **Nombre de la tabla:** centro\_inun | Descripción: | | id\_cm | Llave foránea al centro médico. | | Id\_riesInun | Llave foránea al riesgo de inundación. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | No se requieren para esta tabla. |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No son necesarios para esta tabla, pues en la forma de insertar las llaves foráneas nunca quedan nulas. |
| Descripción de las correspondencias y transformaciones | |  |  | | --- | --- | | **Nombre de la tabla:** centro\_incen | Descripción: | | id\_cm | Llave foránea al centro médico. | | Id\_riesIncen | Llave foránea al riesgo de incendio. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | No se requieren para esta tabla. |
| Problemas encontrados | Ninguno. |

Tabla

|  |  |
| --- | --- |
| Creación de la tabla |  |
| Disparador(es) | No son necesarios para esta tabla, pues en la forma de insertar la llave foránea nunca queda nula. |
| Descripción de las correspondencias y transformaciones | |  |  |  | | --- | --- | --- | | **Nombre de la tabla:** seguro | **Archivo Excel:** Población total por cond. aseguramiento y tipo de seguro social, según provincia, cantón y distrito | Descripción: | | cod\_dis | NA | Llave foránea al distrito. | | tipo\_seguro | TIPO DE SEGURO SOCIAL | Tipo de seguro. | | cantidad\_asegurados | SIN NOMBRE | Número de asegurados de ese tipo en el distrito. | |
| Código de la inserción con transformaciones, limpieza y correcciones de geometrías | Se encuentra el archivo POBLAR\_BASE\_FINAL.sql empezando en la línea #Linea. |
| Código de la creación de índices (incluyendo la verificación de bounding box) o la especificación de parámetros usados | No se requieren para esta tabla. |
| Problemas encontrados | Ninguno. |

Tabla

1. Código y los resultados de cada una de las consultas elaboradas para análisis.

|  |  |  |
| --- | --- | --- |
|  | Código | Resultado |
| Consulta#1 |  |  |
| Consulta#2 |  |  |
| Consulta#3 |  |  |

Tabla : consultas para análisis

1. Una tabla comparativa que debe incluir las consultas y los resultados de tiempo de ejecución, lecturas lógicas, lecturas físicas y los planes de ejecución (para todo es posible usar los pantallazos). Además, deben incluir la **interpretación** (no descripción) de los resultados de cada experimento tomando en cuenta tanto los datos de la tabla comparativa como los planes de ejecución. Se evaluará si la presentación de las conclusiones está expresada en forma clara y concisa.

**Consultas analíticas:**

* Consulta
* Breve explicación de la consulta
* Resultados (o su parte si más que 10 tuplas)

|  |  |  |  |
| --- | --- | --- | --- |
| **Consulta #1** | **Lecturas físicas** | **Lecturas lógicas** | **Tiempo de ejecución** |
| Normal | 11 | 184 | 124 ms |
| Con nested loop join | 6 | 1106 | 207 ms |
| Con nested merge join | 8 | 184 | 245 ms |
| Con nested hash joins | 7 | 184 | 160 ms |
| Modificada con IN | 4 | 36 | 35 ms |
| Modificada con OR | 5 | 180 | 135 ms |
| Plan de ejecución normal |  | | |
| Plan de ejecución Con nested loop join |  | | |
| Plan de ejecución Con nested merge join |  | | |
| Plan de ejecución Con nested hash joins |  | | |
| Plan de ejecución Modificada con IN |  | | |
| Plan de ejecución Modificada con OR |  | | |
| Interpretación | La consulta que sale más barata en cuanto a lecturas físicas es la que fuerza el uso de nested loop join, en contraste con nested hash join, que es la que sale más cara en lecturas. Con respecto a planes de ejecución, podemos ver que el plan de ejecución default elegido por el optimizador es el que utiliza el nested loop join, de hecho vemos que los planes de ejecución de la consulta normal y con hash join son idénticos. En cuanto a costos, observamos que todas las variaciones tienen un costo similar, 100 todas a excepción del nested loop join que tiene 101. Como conclusión con respecto a esta consulta, observamos que definitivamente, la opción elegida por el optimizador no es la mejor, puesto que realiza más lecturas físicas y no tiene un menor costo, la consulta pued optimizarse forzando el uso de otro plan de ejecución mediante hints.  Ahora, para comparar el in versus join, se hizo una pequeña variación a la consulta original para que buscara solo valores en un conjunto. De aquí podemos observar que mientras ambos realizan cantidad similar de lecturas físicas, IN 4 y OR 5, además, el costo es similar también, 99 para IN y 100 para join. Para terminar, observamos que los planes de ejecución lo que varía es la concentración del costo, mientras que en la consulta con IN el costo se centra en los índices sobre distrito y cantones, en la consulta con OR el costo se reparte también en el loop join. Entonces, el IN le gana en eficiencia al OR en algunos puntos, pero son similares ambas consultas. | | |

Tabla : cuadro comparativo de las variaciones de la consulta de análisis #1

|  |  |
| --- | --- |
| **Consulta #1** | **Estadísticas** |
| Normal | Tiempo de análisis y compilación de SQL Server:  Tiempo de CPU = 16 ms, tiempo transcurrido = 28 ms.  (470 row(s) affected)  Tabla 'Worktable'. Recuento de exámenes 0, lecturas lógicas 0, lecturas físicas 0, lecturas anticipadas 0, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tabla 'distritos'. Recuento de exámenes 1, lecturas lógicas 166, lecturas físicas 8, lecturas anticipadas 165, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tabla 'cantones'. Recuento de exámenes 1, lecturas lógicas 18, lecturas físicas 3, lecturas anticipadas 22, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tiempos de ejecución de SQL Server:  Tiempo de CPU = 0 ms, tiempo transcurrido = 124 ms. |
| Con nested loop join | Tiempo de análisis y compilación de SQL Server:  Tiempo de CPU = 0 ms, tiempo transcurrido = 11 ms.  (470 row(s) affected)  Tabla 'cantones'. Recuento de exámenes 0, lecturas lógicas 940, lecturas físicas 4, lecturas anticipadas 0, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tabla 'distritos'. Recuento de exámenes 1, lecturas lógicas 166, lecturas físicas 2, lecturas anticipadas 165, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tiempos de ejecución de SQL Server:  Tiempo de CPU = 0 ms, tiempo transcurrido = 207 ms. |
| Con nested merge join | Tiempo de análisis y compilación de SQL Server:  Tiempo de CPU = 0 ms, tiempo transcurrido = 43 ms.  (470 row(s) affected)  Tabla 'distritos'. Recuento de exámenes 1, lecturas lógicas 166, lecturas físicas 5, lecturas anticipadas 165, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tabla 'cantones'. Recuento de exámenes 1, lecturas lógicas 18, lecturas físicas 3, lecturas anticipadas 22, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tiempos de ejecución de SQL Server:  Tiempo de CPU = 0 ms, tiempo transcurrido = 245 ms. |
| Con nested hash joins | Tiempo de análisis y compilación de SQL Server:  Tiempo de CPU = 16 ms, tiempo transcurrido = 72 ms.  (470 row(s) affected)  Tabla 'Worktable'. Recuento de exámenes 0, lecturas lógicas 0, lecturas físicas 0, lecturas anticipadas 0, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tabla 'distritos'. Recuento de exámenes 1, lecturas lógicas 166, lecturas físicas 4, lecturas anticipadas 165, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tabla 'cantones'. Recuento de exámenes 1, lecturas lógicas 18, lecturas físicas 3, lecturas anticipadas 22, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tiempos de ejecución de SQL Server:  Tiempo de CPU = 15 ms, tiempo transcurrido = 160 ms. |
| Modificada con IN | Tiempo de análisis y compilación de SQL Server:  Tiempo de CPU = 0 ms, tiempo transcurrido = 75 ms.  (9 row(s) affected)  Tabla 'cantones'. Recuento de exámenes 0, lecturas lógicas 18, lecturas físicas 2, lecturas anticipadas 0, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tabla 'distritos'. Recuento de exámenes 9, lecturas lógicas 18, lecturas físicas 2, lecturas anticipadas 0, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tiempos de ejecución de SQL Server:  Tiempo de CPU = 0 ms, tiempo transcurrido = 35 ms. |
| Modificada con OR | Tiempo de análisis y compilación de SQL Server:  Tiempo de CPU = 0 ms, tiempo transcurrido = 74 ms.  (657 row(s) affected)  Tabla 'cantones'. Recuento de exámenes 1, lecturas lógicas 162, lecturas físicas 3, lecturas anticipadas 22, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tabla 'distritos'. Recuento de exámenes 9, lecturas lógicas 18, lecturas físicas 2, lecturas anticipadas 0, lecturas lógicas de LOB 0, lecturas físicas de LOB 0, lecturas anticipadas de LOB 0.  Tiempos de ejecución de SQL Server:  Tiempo de CPU = 15 ms, tiempo transcurrido = 135 ms. |

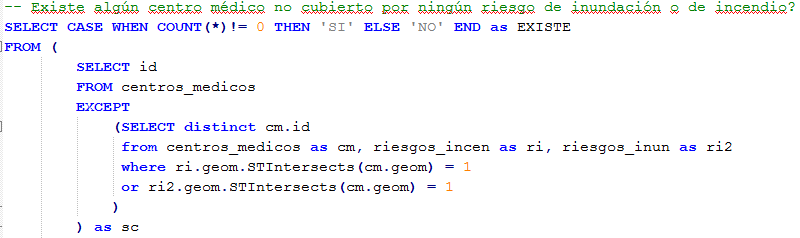
Tabla : estadísticas de la consulta de análisis #1 y sus variaciones

1. Código y los resultados de cada una de las consultas espaciales elaboradas para demostración de uso de la aplicación.

|  |  |  |
| --- | --- | --- |
|  | Código | Resultado |
| Consulta#1 |  |  |
| Consulta#2 |  |  |
| Consulta#3 |  |  |
| Consulta#4 |  |  |
| Consulta#5 |  |  |
| Consulta#6 |  |  |

Tabla : consultas espaciales para demostración

1. Código de algunas consultas extra elaboradas.



Figura