CONSTRUCCIÓN Y EXPLOTACIÓN DE UN ALMACÉN DE DATOS PARA EL ANÁLISIS DE INFORMACIÓN SOBRE TRÁNSITO DE VEHÍCULOS

Trabajo Final de Carrera
ALMACÉN DE DATOS
Memoria del Proyecto

Roberto Sarabia Crespo
Ingeniería Técnica en Informática de Sistemas
Diciembre de 2013
## ÍNDICE DE CONTENIDOS

1 INTRODUCCIÓN .........................................................................................................................4

1.1 JUSTIFICACIÓN ......................................................................................................................4

2 OBJETIVOS .....................................................................................................................................5

2.1 OBJETIVOS GENERALES .......................................................................................................5

2.2 OBJETIVOS ESPECÍFICOS .....................................................................................................5

3 PLANIFICACIÓN ..........................................................................................................................6

3.1 PLANIFICACIÓN GLOBAL ......................................................................................................6

3.1.1 Plan de Trabajo ..................................................................................................................6

3.1.2 Análisis y Diseño ...............................................................................................................6

3.1.3 Implementación ................................................................................................................6

3.1.4 Entrega de la Memoria y Defensa ..................................................................................6

3.2 PLANIFICACIÓN DETALLADA .............................................................................................7

3.3 DIAGRAMA DE GANTT .........................................................................................................8

3.4 INCIDENCIAS Y RIESGOS ....................................................................................................8

3.5 ENTREGABLES .......................................................................................................................9

3.5.1 PEC1 ..................................................................................................................................9

3.5.2 PEC2 ..................................................................................................................................9

3.5.3 PEC3 ..................................................................................................................................9

3.5.4 Memoria ..........................................................................................................................9

3.5.5 Presentación Virtual .........................................................................................................9

4 ANÁLISIS ......................................................................................................................................9

4.1 ANÁLISIS DE LAS FUENTES PROPORCIONADAS ...............................................................9

4.2 ANÁLISIS DE LOS DATOS ....................................................................................................10

4.2.1 Fichero Dades_municipis.xls .........................................................................................10

4.2.2 Fichero Dades_vehicles.xls .........................................................................................11

4.2.3 Fichero Radars_SCT.txt ...............................................................................................12

4.2.4 Ficheros Tipo Dades_conductors_xxxx.txt .................................................................13

4.3 DIAGRAMA DE CASOS DE USO ..........................................................................................14

4.3.1 Perfil Administrador .......................................................................................................14

4.3.2 Perfil Usuario FECRES .................................................................................................15

4.4 MODELO CONCEPTUAL .....................................................................................................15

4.4.1 Granularidad ..................................................................................................................15

4.4.2 Definición de las Dimensiones ......................................................................................16

4.4.3 Identificación del Hecho .................................................................................................16

4.4.4 Identificación de Jerarquías ............................................................................................16

4.4.5 Identificación las Medidas .............................................................................................17

4.5 ESQUEMA CONCEPTUAL Y LÓGICO ................................................................................17

5 DISEÑO .......................................................................................................................................18

5.1 VISIÓN GENERAL DE LA ARQUITECTURA SOFTWARE ................................................18

5.2 DISEÑO FÍSICO .....................................................................................................................19

5.2.1 Dimensiones ...................................................................................................................19
5.2.2 Tabla de Hechos
5.2.3 Diagrama Modelo Físico
5.2.4 Sentencias SQL de la creación de Tablas y Vistas

6 IMPLEMENTACIÓN

6.1 Visión General del proceso ETL
6.2 Creación del Usuario y Tablespaces
6.3 Estructura de Directorios
6.4 Carga de Datos

6.4.1 Preparación de Ficheros
6.4.2 Proceso de Carga en la Base de Datos

7 CONFIGURACIÓN DE MICROSOFT SQL SERVER ANALYSIS SERVICES 2012

8 DEFINICIÓN DE INFORMES

8.1 Informe del Total de Vehículos
8.2 Informe del Total de Conductores
8.3 Informe del % de Vehículos respecto Población
8.4 Informe de la Densidad de Población y Densidad de Vehículos
8.5 Informe del Número de Vehículos / Número de Radares
8.6 Informe del % de Conductores por Radar
8.7 Indicador de conductores vs habitantes por género
8.8 Indicador de Radares vs Vehículos
8.9 Ratio de Vehículos x Conductor
8.10 Cantidad de Vehículos / Superficie

9 CONCLUSIONES

10 LÍNEAS DE EVOLUCIÓN FUTURA

11 REFERENCIAS
1 Introducción

Este proyecto es el final de la carrera de Ingeniería Técnica en Informática de Sistemas de la Universitat Oberta de Catalunya. El proyecto consiste en el desarrollo de un caso práctico en el cual se deben plasmar todos los conocimientos adquiridos a lo largo de estos años.

1.1 Justificación

Como consecuencia del continuo crecimiento del número de desplazamientos en vehículo de motor durante el año 2012, la Fundación de Estudios para la Conducción RESponsable (FECRES), necesita profundizar en la evolución de este tipo de tráfico y analizar las posibles correlaciones entre medios de locomoción, perfiles de conductores y algunas variables de seguridad vial.

FECRES ha solicitado los datos necesarios para su análisis por un lado al IDESCAT, con información sobre municipios y vehículos y por otro lado le ha solicitado a la DGT los censos de conductores de los últimos 5 años. Finalmente, los datos de radares fijos se han solicitado al Servei Catalá de Trànsit.

FECRES nos ha encargado la creación del almacén de datos para poder explotar la información según sus necesidades.
2 Objetivos

Se va a diferenciar entre objetivos generales, que podría aplicar a cualquier tipo de proyecto dentro de los almacenes de datos y objetivos específicos, que son los que se requieren en este proyecto.

2.1 Objetivos Generales

El objetivo general del proyecto es aprender a construir y explotar un almacén de datos a partir de los datos entregados y de las necesidades del enunciado del propio proyecto.

Con el fin de desarrollarlo correctamente, se dividirá en las siguientes fases:

- Plan de trabajo y análisis preliminar de requerimientos
- Análisis de requerimientos y diseño conceptual y técnico
- Implementación

2.2 Objetivos Específicos

Los objetivos específicos son el resultado de la implantación del almacén de datos que FECRES nos ha encargado como consultora externa independiente para que pueda:

- Profundizar en la evolución del número de desplazamientos en vehículo de motor.
- Analizar las posibles correlaciones entre medios de locomoción, perfiles de conductores y algunas variables de seguridad vial.

Del almacén de datos se deberá obtener, como mínimo, la información relativa a:

- Total de vehículos
- Total de conductores
- % de vehículos respecto población
- Densidad de población (habitantes/km$^2$) y densidad de tráfico (vehículos/km$^2$)
- Número de vehículos / Número de radares
- % de conductores por radar
- Indicador de conductores vs habitantes por género
- Indicador de radares vs vehículos
- Ratio de vehículos x conductor
- Cantidad de vehículos / superficie del territorio
3 Planificación

3.1 Planificación Global

A continuación se detallan las fechas clave para la presentación de trabajos:

<table>
<thead>
<tr>
<th>NOMBRE</th>
<th>ENUNCIADO</th>
<th>ENTREGA</th>
<th>CALIFICACIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEC1 - Plan de Trabajo</td>
<td>19/09/2013</td>
<td>01/10/2013</td>
<td>13/10/2013</td>
</tr>
<tr>
<td>PEC2 - Análisis y Diseño</td>
<td>02/10/2013</td>
<td>05/11/2013</td>
<td>17/11/2013</td>
</tr>
<tr>
<td>PEC3 - Implementación</td>
<td>06/11/2013</td>
<td>18/12/2013</td>
<td>29/12/2013</td>
</tr>
<tr>
<td>MEMORIA - Entrega y Defensa</td>
<td>19/12/2013</td>
<td>06/01/2014</td>
<td>28/01/2014</td>
</tr>
</tbody>
</table>

3.1.1 Plan de Trabajo

El Plan de Trabajo, que se entrega con la PEC1, se indica la planificación estimada de las diferentes tareas a realizar para llevar a cabo el proyecto. Se hace un análisis preliminar de dimensiones, atributos, etc. y se hace un primer análisis de las fuentes de datos proporcionados.

3.1.2 Análisis y Diseño

El Análisis y Diseño se entregan con la PEC2 y consiste en un análisis detallado que se basa en el análisis preliminar entregado con el Plan de Trabajo. Se detalla el modelo dimensional que dará soporte a las necesidades de los usuarios. En esta parte del proyecto se produce un aprendizaje y familiarización de las herramientas que se van a usar en la fase de Implementación.

3.1.3 Implementación

La Implementación se detalla y entrega con la PEC3 y consta de varias tareas como la construcción del almacén de datos incluyendo la extracción, transformación y carga de datos (ETL), la configuración de la herramienta de explotación de datos y la construcción de los informes.

3.1.4 Entrega de la Memoria y Defensa

El documento final a entregar es la Memoria que básicamente está compuesta por el conjunto de documentos entregados con anterioridad y que junto una presentación virtual, concluyen el trabajo de fin de carrera.
3.2 Planificación Detallada

Para estimar el fuerzo en horas/día de las tareas a realizar, se ha estimado una media de 3 horas al día, que aunque los fines de semana se podría alargar, muchos días laborales probablemente no se llegue a alcanzar.

<table>
<thead>
<tr>
<th>TAREA</th>
<th>INICIO</th>
<th>FIN</th>
<th>DÍAS</th>
<th>HORAS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PEC1</strong></td>
<td>19/09/2013</td>
<td>01/10/2013</td>
<td>13</td>
<td>39</td>
</tr>
<tr>
<td>Descargar la documentación</td>
<td>19/09/2013</td>
<td>19/09/2013</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Lectura de la documentación</td>
<td>20/09/2013</td>
<td>20/09/2013</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Instalación del software adicional</td>
<td>21/09/2013</td>
<td>21/09/2013</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Prueba del Entorno de Trabajo</td>
<td>22/09/2013</td>
<td>23/09/2013</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Análisis de Requerimientos</td>
<td>24/09/2013</td>
<td>26/09/2013</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Planificación del Trabajo</td>
<td>27/09/2013</td>
<td>27/09/2013</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Elaborar el documento para PEC1</td>
<td>28/09/2013</td>
<td>01/10/2013</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td><strong>PEC2</strong></td>
<td>02/10/2013</td>
<td>05/11/2013</td>
<td>35</td>
<td>105</td>
</tr>
<tr>
<td>Revisión/Corrección PEC1</td>
<td>02/10/2013</td>
<td>03/10/2013</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Estudio de Requerimientos</td>
<td>04/10/2013</td>
<td>09/10/2013</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Modelo conceptual</td>
<td>10/10/2013</td>
<td>13/10/2013</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Diseño del modelo de datos</td>
<td>14/10/2013</td>
<td>27/10/2013</td>
<td>14</td>
<td>42</td>
</tr>
<tr>
<td>Diseño proceso ETL a alto nivel</td>
<td>28/10/2013</td>
<td>30/10/2013</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Manejo del tratamiento de errores</td>
<td>31/10/2013</td>
<td>01/11/2013</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Elaborar el documento para PEC2</td>
<td>02/11/2013</td>
<td>05/11/2013</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td><strong>PEC3</strong></td>
<td>06/11/2013</td>
<td>18/12/2013</td>
<td>43</td>
<td>129</td>
</tr>
<tr>
<td>Revisión/Corrección PEC2</td>
<td>06/11/2013</td>
<td>07/11/2013</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Creación de la estructura de base de datos</td>
<td>08/11/2013</td>
<td>14/11/2013</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>Validación de datos</td>
<td>15/11/2013</td>
<td>19/11/2013</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Creación de Informes</td>
<td>20/11/2013</td>
<td>29/11/2013</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>Pruebas funcionales</td>
<td>30/11/2013</td>
<td>04/12/2013</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Análisis de Resultados</td>
<td>05/12/2013</td>
<td>12/12/2013</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>Elaborar el documento para PEC3</td>
<td>13/12/2013</td>
<td>18/12/2013</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td><strong>MEMORIA</strong></td>
<td>19/12/2013</td>
<td>06/01/2014</td>
<td>19</td>
<td>57</td>
</tr>
<tr>
<td>Revisión/Corrección PEC3</td>
<td>19/12/2013</td>
<td>20/12/2013</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Conclusiones</td>
<td>21/12/2013</td>
<td>23/12/2013</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Elaborar el documento para entrega</td>
<td>24/12/2013</td>
<td>31/12/2013</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>Síntesis</td>
<td>01/01/2014</td>
<td>04/01/2014</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Revisión Final y Entrega</td>
<td>05/01/2014</td>
<td>06/01/2014</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>
3.3 Diagrama de Gantt

A continuación se muestra el diagrama de Gantt con los detalles de los tiempos del proyecto.

![Diagrama de Gantt]

3.4 Incidencias y Riesgos

Es muy importante detectar y prever los posibles riesgos antes de acometer cualquier proyecto para así reducir o anular, en la medida de lo posible, el posible impacto en caso de sufrir una incidencia.

- **Hardware y Software Utilizado.** Se utilizará un portátil de 2 años de antigüedad con sistema operativo Windows 7 64 bits además del resto del software específico para el proyecto. Al utilizar la máquina virtual de Amazon, todos los riesgos derivados de un posible fallo de una máquina virtual en mi portátil desaparecen, salvo el documento de trabajo que hay que entregar. En caso de sufrir algún tipo de fallo hardware, podría seguir trabajando con el equipo de sobremesa instalando el software necesario.

- **Desviaciones de tiempo.** Es posible que durante cada una de las fases del proyecto, se produzcan desviaciones en las tareas por diferentes motivos, ya sean personales, laborales, etc. con lo que será necesario ir ajustando el resto de las tareas para poder ir cumpliendo cada hito.
- **Realización de copias de seguridad.** Se realizará una copia de las carpetas del disco duro principal del portátil a un disco duro USB externo. También se enviará periódicamente las diferentes versiones de los documentos realizados a una cuenta de correo personal.

### 3.5 Entregables

Aunque se ha mencionado anteriormente, a continuación se detallan los documentos a entregar en el proyecto, que son:

#### 3.5.1 PEC1

Este documento detalla el Plan de Trabajo y la planificación de las tareas. Incluye el primer análisis del proyecto.

#### 3.5.2 PEC2

Documento que consiste en un análisis detallado que se basa en el análisis preliminar entregado con la PEC1. Se detalla el modelo dimensional que dará soporte a las necesidades de los usuarios.

#### 3.5.3 PEC3

Documento que consta de varias tareas, como la construcción del almacén de datos la extracción, transformación y carga de datos (ETL), la configuración de la herramienta de explotación de datos y la construcción de los informes.

#### 3.5.4 Memoria

Documento que básicamente está compuesto por el conjunto de documentos entregados con las PECs.

#### 3.5.5 Presentación Virtual

Documento en formato video digital basado en una presentación en formato Microsoft Powerpoint que incluye una demostración del funcionamiento del proyecto.

### 4 Análisis

#### 4.1 Análisis de las fuentes proporcionadas

Los ficheros de datos han sido solicitados por la Fundación de Estudios para la Conducción Responsable (FECRES). Por una parte, al IDESCAT se le ha solicitado información sobre municipios y vehículos. Y por otra parte, a la DGT la información sobre los censos de conductores de los últimos 5 años. Finalmente, se ha solicitado al Servei Català de Trànsit los datos relativos a radares fijos.
FECRES nos advierte que como la información se ha extraído de diferentes sistemas, es posible que los ficheros tengan formatos diferentes (csv, excel y txt).

Será necesario depurar los ficheros para prepararlos para su carga en la base de datos.

A continuación se analizan los datos recibidos:

- Por parte del IDESCAT.
  Los ficheros generados han sido:
  
  **Dades_vehicles.xls**
  **Dades_municipis.xls**
  
- Por parte de la DGT se reciben 5 ficheros:
  
  **Dades_conductors_2007.txt**
  **Dades_conductors_2008.txt**
  **Dades_conductors_2009.txt**
  **Dades_conductors_2010.txt**
  **Dades_conductors_2011.txt**
  
- Por parte del Servei Català de Trànsit se recibe 1 fichero.
  **Radars_SCT.txt**

### 4.2 Análisis de los datos

#### 4.2.1 Fichero **Dades_municipis.xls**

- Fichero que contiene los datos agregados de los municipios, provincias y/o CC.AA. relativos al sumatorio de población por año y extensión en km$^2$

- El formato es de tipo excel

- Los Campos son:
  
  - Nombre de Municipio/Provincia/CC.AA.
  - Código INE
  - Población 2012
  - Población 2011
  - Población 2010
- Población 2009
- Población 2008
- Población 2007
- Extensión (km²)
  - Tiene un volumen de 466 registros

Hay que destacar que hay un municipio (Canonja) que sólo tiene datos del año 2011 y 2012.

4.2.2 Fichero Dades_vehicles.xls
  - Fichero que contiene los datos agregados de los municipios, provincias y/o CC.AA. relativos al sumatorio de los años 2007 al 2012 de:
    - Vehículos de motor
    - Automóviles
    - Camiones y furgonetas
    - Otros vehículos de motor
  - El formato es de tipo excel
  - Los miles aparecen separados por punto “.”
  - Los campos son:
    - Nombre de Municipio/Provincia/CC.AA.
    - Código INE
    - Vehículos de motor 2011
    - Vehículos de motor 2010
    - Vehículos de motor 2009
    - Vehículos de motor 2008
    - Vehículos de motor 2007
    - Automóviles 2012
    - Automóviles 2011
    - Automóviles 2010
    - Automóviles 2009
Hay que destacar que hay un municipio (Canonja) que sólo tiene datos del año 2012.

Faltan los vehículos a Motor de 2012, que hay que calcularlos como la suma de **automóviles, camiones y furgonetas** y **otros** de 2012

Otros vehículos de motor 2012 es la suma de (Motocicletas, Autobuses, Tractores industriales y Resto de vehículos)

### 4.2.3 Fichero Radars_SCT.txt

- Fichero que las vías y los municipios donde se encuentran los radares.

- Es un fichero .txt y los campos están separados por tabulador.
La cabecera del fichero no está preparada para la importación de datos, ya que tiene texto inicial que no sirve.

Los Campos son:

- Vía
- Municipi
- Comarca
- Demarcació

Tiene un volumen de 217 registros

Si hay vías duplicadas, significa que hay varios radares.

### 4.2.4 Ficheros Tipo Dades_conductors_xxxx.txt

- xxxx indica el año de los datos, que van desde el 2007 hasta el 2011
- La cabecera de los ficheros no está preparada para la importación de datos, ya que tiene texto inicial que no sirve.
- Los ficheros contienen datos agregados de número de permisos y licencias por género y municipio.
- Cada fichero consta de 950 registros de datos, salvo en 2010 y 2011, que tienen 951 registros.
- El separador de campos es el tabulador, salvo en 2011, que utiliza “;”
- Los campos son:
  - Municipio
  - Permisos hombre
  - Permisos Mujer
  - Licencias Hombre
  - Licencias Mujer
4.3 Diagrama de casos de uso

Se pueden diferenciar dos tipos de perfiles:

4.3.1 Perfil Administrador

**Administrar la Base de Datos.** El administrador se ocupará de los procesos de gestión de la propia base de datos, como por ejemplo añadir espacio, hacer las copias de seguridad, actualizaciones de software, etc.

**Procesos ETL.** Los procesos de Extracción, Transformación de Carga serán responsabilidad del administrador.

**Gestión de Usuarios.** La creación, modificación y eliminación, así como los permisos y roles las gestionará el administrador.

**Diseño de Informes.** El diseño y la preparación de los informes de usuario son tarea del administrador.
4.3.2 Perfil Usuario FECRES

**Consulta de Informes.** Consultará los informes preparados por el administrador.

**Petición de Nuevos Informes.** El usuario deberá comunicar sus necesidades al administrador para poder visualizar los datos.

**Entregar Datos.** FECRES deberá entregar los datos a analizar.

### 4.4 Modelo Conceptual

En el modelo conceptual se identificarán las dimensiones y la tabla de hechos.

#### 4.4.1 Granularidad

La granularidad va a determinar el grado de detalle en los informes que se van a ejecutar. En el caso de este proyecto, la granularidad temporal viene determinada por la temporalidad de los datos suministrados, que tienen un carácter **anual**.
4.4.2 Definición de las Dimensiones

Las Dimensiones son información complementaria necesaria para la presentación de los datos a los usuarios. Incluye información general que complementa a la Tabla de Hechos.

A continuación se detallan las dimensiones identificadas:

**Fecha.** El atributo de la fecha es el año y su tipo. Disponemos datos desde el año 2007 hasta el año 2011

**Vía.** Tiene los datos de la vía en la que se encuentra un radar. Los atributos son el identificador de la vía, el nombre y el tipo de vía (Autopista, Autovía, Nacional, Comarcal y Local)

**Género.** Define los tipos de género posibles en los permisos. A parte del identificador del propio género, los posibles valores son Hombre y Mujer.

**Municipio.** Los atributos de municipio son el identificador, el código INE, el propio nombre del municipio, la comarca (sin uso), la provincia y la extensión.

**Vehículo.** Sus atributos serán identificador del vehículo y el tipo ('Motor', 'Automóvil', 'Camiones y Furgonetas', 'Otros Vehículos de Motor')

**Permiso.** A parte de un identificador de permiso, contiene los posibles tipos de permiso en sí, que pueden ser Permiso y Licencia.

4.4.3 Identificación del Hecho

Los datos proporcionados que debemos relacionar son los datos de conductores, datos de vehículos, datos de provincias y radares, con lo que el hecho identificado es el propio tránsito de vehículos. Dicho tránsito está determinado por una fecha, un tipo de vehículo y por los propios tipos de conductores en una zona determinada. En el caso de este proyecto, el hecho se identifica con la tabla TRANSITO

4.4.4 Identificación de Jerarquías

Las jerarquías representan una relación lógica entre dos o más atributos, siempre y cuando posean su correspondiente relación padre-hijo. La ventaja principal de las jerarquías es poder analizar los datos desde un nivel más general a otro más detallado y viceversa, al poder desplazarse por los diferentes niveles.

Jerarquía : Provincia → Población
4.4.5 Identificación las Medidas

Las medidas son las propiedades de la tabla de hechos que se usarán posteriormente en el análisis mediante los informes, por ejemplo. Las medidas de la tabla de hechos TRANSITO son:

- Número de Conductores
- Número de Vehículos
- Número de Radares
- Número de Habitantes

4.5 Esquema Conceptual y Lógico

El modelo de datos quedaría como se muestra es el siguiente esquema. Las tablas de dimensiones alrededor de la tabla de hechos. Este es un modelo de **estrella**.
5 Diseño

5.1 Visión General de la Arquitectura Software

A continuación se detalla la arquitectura general del proyecto:

Mediante SQL*Loader, los ficheros originales en sus diferentes formatos (Excel y txt) son cargados y transformados en tablas temporales previo a la carga final para su explotación desde MS SQL Analysis Services.
5.2 Diseño Físico

5.2.1 Dimensiones

Como se ha indicado anteriormente, las Dimensiones son información complementaria necesaria para la presentación de los datos a los usuarios. Incluye información general que complementa a la Tabla de Hechos (TRANSITO). En el siguiente apartado se describen las dimensiones usadas en detalle.

**Dimensión FECHA**

En esta tabla, el identificador ID_ANYO es el propio año en sí. Es la clave primaria de la tabla (PK_ANYO) y crea automáticamente un índice.

El campo BISIESTO nos indica simplemente si es o no bisiesto ese año. En principio no se utiliza y es información complementaria.

<table>
<thead>
<tr>
<th>CAMPO</th>
<th>TIPO</th>
<th>DESCRIPCIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID_ANYO</td>
<td>NUMBER</td>
<td>Identificador del año (AAAA)</td>
</tr>
<tr>
<td>BISIESTO</td>
<td>VARCHAR(1)</td>
<td>S/N</td>
</tr>
</tbody>
</table>

**Dimensión VIA**

El campo ID_VIA toma valores secuenciales de la secuencia de base de datos VIA_SEQ, empezando por el valor 1 e incrementándose en 1 por cada valor:

```
create sequence via_seq start with 1 increment by 1 nocache;
```

Es la clave primaria de la tabla (PK_VIA) y crea automáticamente un índice.

El campo NOMBRE se carga directamente del campo VIA de la tabla temporal TMP_RADARS

El campo TIPO se analiza durante la carga.

<table>
<thead>
<tr>
<th>CAMPO</th>
<th>TIPO</th>
<th>DESCRIPCIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID_VIA</td>
<td>NUMBER(3)</td>
<td>Identificador de la vía</td>
</tr>
<tr>
<td>NOMBRE</td>
<td>VARCHAR(10)</td>
<td>Nombre de la vía</td>
</tr>
<tr>
<td>TIPO</td>
<td>VARCHAR(20)</td>
<td>'Autopista', 'Autovía', 'Nacional',</td>
</tr>
<tr>
<td></td>
<td></td>
<td>'Comarcal', 'Local'</td>
</tr>
</tbody>
</table>
**Dimensión GENERO**

El campo ID_GENERO toma los valores 1 y 2.

Es la clave primaria de la tabla (PK_GENERO) y crea automáticamente un índice.

<table>
<thead>
<tr>
<th>CAMPO</th>
<th>TIPO</th>
<th>DESCRIPCIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID_GENERO</td>
<td>NUMBER(1)</td>
<td>Identificador del Tipo de género</td>
</tr>
<tr>
<td>NOMBRE_GENERO</td>
<td>VARCHAR2(10)</td>
<td>Hombre y Mujer</td>
</tr>
</tbody>
</table>

**Dimensión MUNICIPIO**

El campo ID_MUNICIPIO toma valores secuenciales de la secuencia de base de datos MUNICIPIO_SEQ, empezando por el valor 1 e incrementándose en 1 por cada valor:

create sequence municipio_seq start with 1 increment by 1 nocache;

Es la clave primaria de la tabla (PK_MUNICIPIO) y crea automáticamente un índice.

El campo NOMBRE se carga directamente del campo NOMBRE de la tabla temporal TMP_MUNICIPIES

El campo PROVINCIA se analiza durante la carga.

El campo COMARCA no se usa de momento

<table>
<thead>
<tr>
<th>CAMPO</th>
<th>TIPO</th>
<th>DESCRIPCIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID_MUNICIPIO</td>
<td>NUMBER(3)</td>
<td>Identificador del Municipio</td>
</tr>
<tr>
<td>NOMBRE</td>
<td>VARCHAR2(60)</td>
<td>Nombre del Municipio</td>
</tr>
<tr>
<td>COD_INE</td>
<td>VARCHAR2(5)</td>
<td>Código INE de Municipio</td>
</tr>
<tr>
<td>COMARCA</td>
<td>VARCHAR2(60)</td>
<td>Nombre Comarca (Sin uso)</td>
</tr>
<tr>
<td>PROVINCIA</td>
<td>VARCHAR2(60)</td>
<td>Nombre Provincia</td>
</tr>
<tr>
<td>EXTENSION</td>
<td>NUMBER(10)</td>
<td>Extensión del municipio en km2</td>
</tr>
</tbody>
</table>

**Dimensión VEHICULO**

El campo ID_TIPO_VEHICULO es la clave primaria y toma los valores del 1 al 4.
El campo TIPO es el nombre del tipo y los posibles valores están descritos en la tabla siguiente

<table>
<thead>
<tr>
<th>CAMPO</th>
<th>TIPO</th>
<th>DESCRIPCIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID_TIPO_VEHICULO</td>
<td>NUMBER(2)</td>
<td>Identificador del Tipo vehículo</td>
</tr>
<tr>
<td>TIPO</td>
<td>VARCHAR2(60)</td>
<td>'Motor', 'Automóvil', 'Camiones y Furgonetas', 'Otros Vehículos de Motor'</td>
</tr>
</tbody>
</table>

_Dimensión PERMISO_

El campo ID_PERMISO es la clave primaria y toma los valores 1 y 2.

El campo TIPO_PERMISO es el nombre del tipo y los posibles valores están descritos en la tabla siguiente

<table>
<thead>
<tr>
<th>CAMPO</th>
<th>TIPO</th>
<th>DESCRIPCIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID_PERMISO</td>
<td>NUMBER(1)</td>
<td>Identificador del tipo de permiso</td>
</tr>
<tr>
<td>TIPO_PERMISO</td>
<td>VARCHAR2(10)</td>
<td>Permiso y Licencia</td>
</tr>
</tbody>
</table>
5.2.2 Tabla de Hechos

La tabla de Hechos es la tabla central del diseño y representa los datos principales que van a ser analizados y donde se van a lanzar los informes para el análisis de la información.

Tabla de Hechos : TRANSITO

Los campos ID_VIA, ID_PERMISO, ID_GENERO, ID_ANYO, ID_MUNICIPIO y ID_TIPO_VEHICULO son la clave primaria de la tabla de hechos, que a su vez son claves foráneas de las tablas de dimensiones correspondientes.

Los campos NUM_CONDUCTORES, NUM_VEHICULOS, NUM_RADARES y NUM_HABITANTES son campos calculados y preparados para mejorar el rendimiento en los informes.

<table>
<thead>
<tr>
<th>TRANSITO (TABLA DE HECHOS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAMPO</td>
</tr>
<tr>
<td>ID_VIA</td>
</tr>
<tr>
<td>ID_PERMISO</td>
</tr>
<tr>
<td>ID_GENERO</td>
</tr>
<tr>
<td>ID_ANYO</td>
</tr>
<tr>
<td>ID_MUNICIPIO</td>
</tr>
<tr>
<td>ID_TIPO_VEHICULO</td>
</tr>
<tr>
<td>NUM_CONDUCTORES</td>
</tr>
<tr>
<td>NUM_VEHICULOS</td>
</tr>
<tr>
<td>NUM_RADARES</td>
</tr>
<tr>
<td>NUM_HABITANTES</td>
</tr>
</tbody>
</table>
5.2.3 Diagrama Modelo Físico

En el siguiente esquema, se puede ver que la tabla de hechos TRANSITO es la que enlaza con las tablas de dimensiones (GENERo, PERMISO, FECHA, VIA, MUNICIPIO y VEHICULO) mediante claves foráneas.

5.2.4 Sentencias SQL de la creación de Tablas y Vistas

5.2.4.1 Tablas Temporales

Las tablas temporales se crean y cargan con la información proveniente de ficheros planos en diferentes formatos. Posteriormente, se utilizan para extraer la información necesaria para las tablas finales (dimensiones y hechos).

Tabla para cargar los datos del fichero “Dades_Municipis.xls”

```sql
create table tmp_municipis (    nombre Varchar2(60),    ine Varchar2(5),    p2012 number(10),    p2011 number(10),    p2010 number(10),    p2009 number(10),
```

Roberto Sarabia Crespo

Pág. 23 / 52
Tabla para cargar los datos del fichero “Radars_SCT.txt”

```sql
create table tmp_radars (
    via     Varchar2(10),
    municipio     Varchar2(60),
    comarca     Varchar2(60),
    demarcacion     Varchar2(60)
)
```

Tabla para cargar los datos del fichero “Dades_vehicles.xls”

```sql
create table tmp_vehicles (
    nombre     Varchar2(60),
    ine     Varchar2(5),
    vmotor2012     number(10),
    vmotor2011     number(10),
    vmotor2010     number(10),
    vmotor2009     number(10),
    vmotor2008     number(10),
    vmotor2007     number(10),
    auto2012     number(10),
    auto2011     number(10),
    auto2010     number(10),
    auto2009     number(10),
    auto2008     number(10),
    auto2007     number(10),
    camfur2012     number(10),
    camfur2011     number(10),
    camfur2010     number(10),
    camfur2009     number(10),
    camfur2008     number(10),
    camfur2007     number(10),
    otros2012     number(10),
    otros2011     number(10),
    otros2010     number(10),
    otros2009     number(10),
    otros2008     number(10),
    otros2007     number(10),
    motos     number(10),
    buses     number(10),
    tractores     number(10),
    resto     number(10)
)
```

Tabla para cargar los datos del fichero “Dades_Conductors_2007.txt”

```sql
create table tmp_dades_conductors_2007 (
    municipio     Varchar2(60),
    pmujer     number(10),
    phombre     number(10),
    lmujer     number(10),
    lhombre     number(10),
    demarcacion     Varchar2(60)
)
```
Tablas Finales

Las tablas Finales se crean y cargan con la información proveniente de las tablas y vistas temporales. Posteriormente, se utilizan para lanzar informes de análisis. Al ser un modelo en estrella, las dimensiones proveen de los detalles y la descripción de la información que proporciona la tabla principal o tabla de hechos.

**Tabla de Hechos**

```sql
create table transito (
    id_via   number(3),
    id_permiso   number(1),
)```
Construcción y Explotación de un almacén de datos para el análisis de información sobre tránsito de vehículos

id_genero   number(1),
id_anyo   number,
id_municipio   number(3),
id_tipo_vehiculo   number(2),
um_conductores number(10),
um_vehiculos number(10),
um_radars number(3),
um_habitantes number(10),

constraint pk_transito primary key
(id_via,id_permiso,id_genero,id_anyo,id_municipio,id_tipo_vehiculo)
using index tablespace tfc ,
constraint fk_via foreign key (id_via) references via(id_via),
constraint fk_permiso foreign key (id_permiso) references
permiso(id_permiso),
constraint fk_genero foreign key (id_genero) references
genero(id_genero),
constraint fk_anyo foreign key (id_anyo) references fecha(id_anyo),
constraint fk_municipio foreign key (id_municipio) references
municipio(id_municipio),
constraint fk_tipo_vehiculo foreign key (id_tipo_vehiculo) references
vehiculo(id_tipo_vehiculo)
)

Dimensión VIA

create table via (    id_via   number(3),
nombre   Varchar2(10),
tipo   Varchar2(20),
constraint pk_via primary key (id_via)
)

Dimensión PERMISO

create table permiso (    id_permiso   number(1),
tipo_permiso   Varchar2(10),
constraint pk_permiso primary key (id_permiso)
)

Dimensión GENERO

create table genero (    id_genero   number(1),
nombre_genero   Varchar2(10),
constraint pk_genero primary key (id_genero)
)

Dimensión FECHA

create table fecha (    id_anyo   number,
bisiesto   Varchar2(1),
constraint pk_anyo primary key (id_anyo)
Dimensión MUNICIPIO

```sql
create table municipio
(
  id_municipio number(3),
  nombre Varchar2(60),
  cod_ine Varchar2(5),
  comarca Varchar2(60),
  provincia Varchar2(60),
  extension number(10),
  constraint pk_municipio primary key (id_municipio)
)
```

Dimensión VEHICULO

```sql
create table vehiculo
(
  id_tipo_vehiculo number(2),
  tipo Varchar2(60),
  constraint pk_tipo primary key (id_tipo_vehiculo)
)
```

5.2.4.3 Vistas Temporales

Las vistas temporales se han creado para ayudar a la carga final de la Tabla de Hechos

Vista **V_TMP_DADES_CONDUCTORS**. Une las tablas temporales de datos de conductores de los diferentes años

```sql
create or replace view V_TMP_DADES_CONDUCTORS as
  select 2007 anyo, TMP_DADES_CONDUCTORS_2007.*
  from TMP_DADES_CONDUCTORS_2007
  union
  select 2008 anyo, TMP_DADES_CONDUCTORS_2008.*
  from TMP_DADES_CONDUCTORS_2008
  union
  select 2009 anyo, TMP_DADES_CONDUCTORS_2009.*
  from TMP_DADES_CONDUCTORS_2009
  union
  select 2010 anyo, TMP_DADES_CONDUCTORS_2010.*
  from TMP_DADES_CONDUCTORS_2010
  union
  select 2011 anyo, TMP_DADES_CONDUCTORS_2011.*
  from TMP_DADES_CONDUCTORS_2011
/
```

Vista **V_TMP_MUNICIPIS**. Prepara la tabla temporal TMP_MUNICIPIS para un mejor acceso a los datos por año y número de habitantes

```sql
create or replace view V_TMP_MUNICIPIS as
  select 2007 anyo, nombre, ine, p2007 habitantes
  from TMP_MUNICIPIS
  union
  select 2008 anyo, nombre, ine, p2008 habitantes
  from TMP_MUNICIPIS
  union
  select 2009 anyo, nombre, ine, p2009 habitantes
  from TMP_MUNICIPIS
  union
```
select 2010 anyo, nombre, ine, p2010 habitantes
from TMP_MUNICIPIS
union
select 2011 anyo, nombre, ine, p2011 habitantes
from TMP_MUNICIPIS
/

Vista **V_TMP_VEHICLES**. Prepara la tabla temporal TMP_VEHICLES para un mejor acceso a los datos por año y tipo de vehículo.

```sql
create or replace view V_TMP_VEHICLES as
select 2007 anyo, nombre, ine, vmotor2007 vmotor, auto2007 auto,
camfur2007 camfur, otros2007 otros
from TMP_VEHICLES
union
select 2008 anyo, nombre, ine, vmotor2008 vmotor, auto2008 auto,
camfur2008 camfur, otros2008 otros
from TMP_VEHICLES
union
select 2009 anyo, nombre, ine, vmotor2009 vmotor, auto2009 auto,
camfur2009 camfur, otros2009 otros
from TMP_VEHICLES
union
select 2010 anyo, nombre, ine, vmotor2010 vmotor, auto2010 auto,
camfur2010 camfur, otros2010 otros
from TMP_VEHICLES
union
select 2011 anyo, nombre, ine, vmotor2011 vmotor, auto2011 auto,
camfur2011 camfur, otros2011 otros
from TMP_VEHICLES
/
```

Vista **V_TMP_VEHICLES_AUX**. Prepara la vista temporal V_TMP_VEHICLES para un mejor acceso a los datos por año y tipo de vehículo.

```sql
create or replace view V_TMP_VEHICLES_AUX as
select 'Motor' TIPO, V_TMP_VEHICLES.ANYO, V_TMP_VEHICLES.NOMBRE,
V_TMP_VEHICLES.VMOTOR CANTIDAD
from V_TMP_VEHICLES
union
select 'Automóvil' TIPO, V_TMP_VEHICLES.ANYO, V_TMP_VEHICLES.NOMBRE,
V_TMP_VEHICLES.AUTO CANTIDAD
from V_TMP_VEHICLES
union
select 'Camiones y Furgonetas' TIPO, V_TMP_VEHICLES.ANYO, V_TMP_VEHICLES.NOMBRE,
V_TMP_VEHICLES.CAMFUR CANTIDAD
from V_TMP_VEHICLES
union
select 'Otros Vehículos de Motor' TIPO, V_TMP_VEHICLES.ANYO, V_TMP_VEHICLES.NOMBRE,
V_TMP_VEHICLES.OTROS CANTIDAD
from V_TMP_VEHICLES
/
```
Vista **V_TMP_RADARS**. Prepara los datos relativos a radares, municipios y tipo de vía.

```sql
create or replace view V_TMP_RADARS as
select m.nombre municipio, count(*) cantidad , v.id_via idvia
from municipio m ,tmp_radars r , via v
where m.nombre=r.municipio and r.via=v.nombre
group by m.nombre,v.id_via
union
select nombre municipio, 0 cantidad , 0 idvia
from municipio
where nombre not in (select municipio from tmp_radars)
/
```

Vista **V_TMP_DADES_CONDUCTORS_AUX**. Prepara la vista temporal **V_TMP_DADES_CONDUCTORS** para un mejor acceso a los datos por tipo de permiso y género.

```sql
create or replace view V_TMP_DADES_CONDUCTORS_AUX as
select 'Permiso' TIPO,'Mujer' GENERO,
V_TMP_DADES_CONDUCTORS.ANYO,V_TMP_DADES_CONDUCTORS.MUNICIPIO,
V_TMP_DADES_CONDUCTORS.PMUJER CANTIDAD
from V_TMP_DADES_CONDUCTORS
union
select 'Permiso' TIPO,'Hombre'
GENER0,V_TMP_DADES_CONDUCTORS.ANYO,V_TMP_DADES_CONDUCTORS.MUNICIPIO,
V_TMP_DADES_CONDUCTORS.PHOMBRE CANTIDAD
from V_TMP_DADES_CONDUCTORS
union
select 'Licencia' TIPO,'Mujer' GENERO,
V_TMP_DADES_CONDUCTORS.ANYO,V_TMP_DADES_CONDUCTORS.MUNICIPIO,
V_TMP_DADES_CONDUCTORS.LMUJER CANTIDAD
from V_TMP_DADES_CONDUCTORS
union
select 'Licencia' TIPO,'Hombre'
GENER0,V_TMP_DADES_CONDUCTORS.ANYO,V_TMP_DADES_CONDUCTORS.MUNICIPIO,
V_TMP_DADES_CONDUCTORS.LHOMBRE CANTIDAD
from V_TMP_DADES_CONDUCTORS
/
```
Implementación

6.1 Visión General del proceso ETL

Los ficheros de datos han sido solicitados por la Fundación de Estudios para la Conducción Responsable (FECRES). Por una parte, al IDESCAT se le ha solicitado información sobre municipios y vehículos. Y por otra parte, a la DGT la información sobre los censos de conductores de los últimos 5 años. Finalmente, se ha solicitado al Servei Català de Trànsit los datos relativos a radares fijos.
6.2 Creación del Usuario y Tablespaces

- Para crear el esquema de base de datos, utilizaremos la base de datos XE, donde necesitamos crear previamente un tablespace donde ubicar los objetos que vamos a crear. Se puede crear desde una herramienta gráfica como Oracle SQL Developer o ejecutando desde la línea de comandos:

```sql
sqlplus /nolog
connect / as sysdba
```

- Posteriormente ejecutamos las sentencias de creación del tablespace TFC, usuario TFC y sus quotas y privilegios de base de datos:

```sql
-- TABLESPACE
CREATE TABLESPACE TFC
  DATAFILE 'C:\oraclexe\app\oracle\oradata\XE\TFC.DBF' SIZE 104857600
  AUTOEXTEND ON NEXT 10485760 MAXSIZE 1048576000
  DEFAULT NOCOMPRESS
  ONLINE
  EXTENT MANAGEMENT LOCAL AUTOALLOCATE;

-- USUARIO
CREATE USER tfc IDENTIFIED BY tfc
DEFAULT TABLESPACE "TFC"
TEMPORARY TABLESPACE "TEMP";

-- QUOTAS
ALTER USER TFC QUOTA UNLIMITED ON TFC;

-- ROLES
GRANT "RESOURCE" TO TFC;
GRANT "CONNECT" TO TFC;
GRANT "CREATE VIEW" TO TFC;
```

- La creación de uno o varios tablespaces dedicados a los objetos del esquema TFC se hace por motivos de rendimiento y de organización.
6.3 Estructura de Directorios

Se ha creado una estructura de directorios para facilitar la organización y la ejecución de los distintos pasos. Todos los ficheros y carpetas están ubicados en la carpeta principal \C:TFC y cuya estructura de subdirectorios y contenido de ficheros es la siguiente:

En la carpeta raíz TFC se encuentra el fichero ejecutable principal (coordinador):

Lanza_Creacion.bat

Este fichero será el encargado de lanzar los siguientes ficheros ejecutables en el orden correcto:

CALL C:TFC\BAT\1_Crea_Tablas_Temporales.bat
CALL C:TFC\BAT\2_Carga_Datos_Temporales.bat
CALL C:TFC\BAT\3_Crea_Tablas_Finales.bat
CALL C:TFC\BAT\4_Crea_Vistas.bat
CALL C:TFC\BAT\5_Carga_Datos_Finales.bat
CALL C:TFC\BAT\6_Carga_Tabla_Hechos.bat
6.4 Carga de Datos

6.4.1 Preparación de Ficheros
Antes de la carga de los datos en la base de datos con SQL*Loader, es necesario prepararlos para su correcta carga.

6.4.1.1 Transformaciones manuales


Los ficheros Dades_municipis.xls y Dades_vehicles.xls, se tratarán con OpenOffice Calc para guardarlo en formato .csv:

Dades_municipis.csv y Dades_vehicles.csv
6.4.2 Proceso de Carga en la Base de Datos

El fichero Lanza_Creación.bat ejecutará los siguientes ficheros .bat:

1 _Crea_Tablas_Temporales.bat

```
cd C:\TFC
set NLS_LANG=AMERICAN_AMERICA.AL32UTF8
sqlplus tfc/tfc @"C:\TFC\SQL\1_Crea_Tablas_Temporales.sql"
```

Encargado de ejecutar el script SQL _1_Crea_Tablas_Temporales.sql_, para crear las tablas temporales necesarias para la carga desde SQL*Loader

2_Carga_Datos_Temporales.bat

```
cd C:\TFC
set NLS_LANG=AMERICAN_AMERICA.AL32UTF8
sqlldr userid=tfc/tfc control=CTL\Dades_municipis.ctl log=LOG\Dades_municipis.log
sqlldr userid=tfc/tfc control=CTL\Radars_SCT.ctl log=LOG\Radars_SCT.log
sqlldr userid=tfc/tfc control=CTL\Dades_vehicles.ctl log=LOG\Dades_vehicles.log
sqlldr userid=tfc/tfc control=CTL\Dades_conductors_2007.ctl log=LOG\Dades_conductors_2007.log
sqlldr userid=tfc/tfc control=CTL\Dades_conductors_2008.ctl log=LOG\Dades_conductors_2008.log
sqlldr userid=tfc/tfc control=CTL\Dades_conductors_2009.ctl log=LOG\Dades_conductors_2009.log
sqlldr userid=tfc/tfc control=CTL\Dades_conductors_2010.ctl log=LOG\Dades_conductors_2010.log
sqlldr userid=tfc/tfc control=CTL\Dades_conductors_2011.ctl log=LOG\Dades_conductors_2011.log
```

Es el encargado de ejecutar las diferentes cargas mediante SQL*Loader de los ficheros de texto en la base de datos. Durante la carga se hacen filtrados y transformaciones previos para preparar los datos.

A continuación se detallan los diferentes ficheros de control:

**Dades_municipis.ctl**

```
options (skip=1)
load data
CHARACTERSET WE8MSWIN1252
infile 'DATOS\Dades_municipis.csv'
badfile 'LOG\Dades_municipis.bad'
discardfile 'LOG\Dades_municipis.dsc'
into table tmp_municipis
truncate
FIELDS TERMINATED BY ":" optionally enclosed by "\"
trailing nullcols
(nombre ,
 ine ,
p2012 "decode(:p2012,'n.d.',null,replace(:p2012,'.',''))", 
p2011 "decode(:p2011,'n.d.',null,replace(:p2011,'.',''))", 
p2010 "decode(:p2010,'n.d.',null,replace(:p2010,'.',''))", 
p2009 "decode(:p2009,'n.d.',null,replace(:p2009,'.',''))",
```
Durante la carga de la tabla TMP_MUNICIPIS:
- Se cambia el literal "n.d." por NULL
- Se eliminan los “.” de los valores numéricos
- Se salta la cabecera (primera línea)
- El campo carácter separador es “;”
- Previo a la carga, se trunca la tabla

```sql
options (skip=1)
load data
CHARACTERSET WE8MSWIN1252
infile 'DATOS\Dades_vehicles.csv'
into table tmp_vehicles
truncate
FIELDS TERMINATED BY ';' optionally enclosed by '"'
trailing nullcols
{
  nombre,
  ine,
  vmotor2001
  "decode(:vmotor2001,'n.d.',null,replace(:vmotor2001,'.',''))",
  vmotor2002
  "decode(:vmotor2002,'n.d.',null,replace(:vmotor2002,'.',''))",
  vmotor2003
  "decode(:vmotor2003,'n.d.',null,replace(:vmotor2003,'.',''))",
  vmotor2004
  "decode(:vmotor2004,'n.d.',null,replace(:vmotor2004,'.',''))",
  vmotor2005
  "decode(:vmotor2005,'n.d.',null,replace(:vmotor2005,'.',''))",
  vmotor2006
  "decode(:vmotor2006,'n.d.',null,replace(:vmotor2006,'.',''))",
  vmotor2007
  "decode(:vmotor2007,'n.d.',null,replace(:vmotor2007,'.',''))",
  auto2008
  "decode(:auto2008,'n.d.',null,replace(:auto2008,'.',''))",
  auto2009
  "decode(:auto2009,'n.d.',null,replace(:auto2009,'.',''))",
  auto2010
  "decode(:auto2010,'n.d.',null,replace(:auto2010,'.',''))",
  auto2011
  "decode(:auto2011,'n.d.',null,replace(:auto2011,'.',''))",
  auto2012
  "decode(:auto2012,'n.d.',null,replace(:auto2012,'.',''))",
  camfur2007
  "decode(:camfur2007,'n.d.',null,replace(:camfur2007,'.',''))",
  camfur2008
  "decode(:camfur2008,'n.d.',null,replace(:camfur2008,'.',''))",
  camfur2009
  "decode(:camfur2009,'n.d.',null,replace(:camfur2009,'.',''))",
  camfur2010
  "decode(:camfur2010,'n.d.',null,replace(:camfur2010,'.',''))",
  camfur2011
  "decode(:camfur2011,'n.d.',null,replace(:camfur2011,'.',''))",
  camfur2012
  "decode(:camfur2012,'n.d.',null,replace(:camfur2012,'.',''))",
  camfur2013
  "decode(:camfur2013,'n.d.',null,replace(:camfur2013,'.',''))",
  camfur2014
  "decode(:camfur2014,'n.d.',null,replace(:camfur2014,'.',''))",
  otros2007
  "decode(:otros2007,'n.d.',null,replace(:otros2007,'.',''))",
  otros2008
  "decode(:otros2008,'n.d.',null,replace(:otros2008,'.',''))",
  otros2009
  "decode(:otros2009,'n.d.',null,replace(:otros2009,'.',''))",
  otros2010
  "decode(:otros2010,'n.d.',null,replace(:otros2010,'.',''))",
  otros2011
  "decode(:otros2011,'n.d.',null,replace(:otros2011,'.',''))",
  otros2012
  "decode(:otros2012,'n.d.',null,replace(:otros2012,'.',''))",
  motos
  "decode(:motos,'n.d.',null,replace(:motos,'.',''))",
  km2
}
```
Durante carga de la tabla TMP_VEHICLES:
- Se cambia el literal “n.d.” por NULL
- Se calcula el campo VMOTOR2012 como la suma de:
  - AUTO2012, CAMFUR2012 y OTROS2012
- Se eliminan los “.” de los valores numéricos
- Se salta la cabecera (primera línea)
- El campo carácter separador es “;”
- Previo a la carga, se trunca la tabla

```
options (skip=12)
load data
CHARACTERSET WE8MSWIN1252
infile 'DATOS\Radars_SCT.txt'
into table tmp_radars
truncate
FIELDS TERMINATED BY x'09' optionally enclosed by ""
trailing nullcols
(via ,
 municipio ,
 comarca ,
 demarcacion)
```

Durante carga de la tabla TMP_RADARS:
- Se salta la cabecera (primeras 12 líneas)
- El campo carácter separador es el Tabulador
- Previo a la carga, se trunca la tabla

```
options (skip=7)
load data
CHARACTERSET WE8MSWIN1252
infile 'DATOS\Dades_conductors xxxx.txt'
into table tmp_dades_conductors_2007
WHEN municipio != 'Copyright DGT 2013' AND municipio != 'Fuente:Dirección General de Tráfico'
FIELDS TERMINATED BY x'09' optionally enclosed by ""
trailing nullcols
{
  municipio "trim(substr(:municipio,instr(:municipio,'-',1)+2,length(:municipio)))",
  pmujer ,
  phombre ,
  lmujer ,
  lhombre ,
  demarcacion "substr(:municipio,1,instr(:municipio,'-',1)-2)"
}
```
Durante carga de la tabla TMP_DADES_CONDUCTORS_xxxx:
- Se separan el municipio y la demarcación del fichero de texto y se insertan en los campos MUNICIPIO y DEMARCACIÓN
- Se salta la cabecera (primeras 7 líneas)
- Se salta el pie (las últimas líneas)
- El campo carácter separador es el tabulador
- Previo a la carga, se trunca la tabla

```
options (skip=7)
load data
CHARACTERSET WE8MSWIN1252
infile 'DATOS\Dades_conductors_2011.csv'
truncate into table tmp_dades_conductors_2011
FIELDS TERMINATED BY ';' optionally enclosed by '"'
trailing nullcols
(
  municipio  "trim(substr(:municipio,instr(:municipio,'-',1)+2,length(:municipio)))",
  pmujer ,
  phombre ,
  lmujer ,
  lhombre ,
  demarcacion "substr(:municipio,1,instr(:municipio,'-',1)-2)"
)
```

Durante carga de la tabla TMP_DADES_CONDUCTORS_2011:
- Se separan el municipio y la demarcación del fichero de texto y se insertan en los campos MUNICIPIO y DEMARCACIÓN
- Se salta la cabecera (primeras 7 líneas)
- El campo carácter separador es “;”
- Previo a la carga, se trunca la tabla

3_Crea_Tablas_Finales.bat
```
cd C:\TFC
set NLS_LANG=AMERICAN_AMERICA.AL32UTF8
sqlplus tfc/tfc @"C:\TFC\SQL\2_Crea_Tablas_Finales.sql"
```
Este fichero será el encargado de ejecutar el script SQL “2_Crea_Tablas_Finales.sql” que es el encargado de crear las tablas finales (Dimensiones y tabla de Hechos).

4_Crea_Vistas.bat
```
cd C:\TFC
set NLS_LANG=AMERICAN_AMERICA.AL32UTF8
sqlplus tfc/tfc @"C:\TFC\SQL\4_Crea_Vistas.sql"
```
Este fichero será el encargado de ejecutar el script SQL “4_Crea_Vistas.sql” que crea las vistas auxiliares para las carga de la tabla de hechos.
Construcción y Explotación de un almacén de datos para el análisis de información sobre tránsito de vehículos

5_Carga_Datos_Finales.bat

cd C:\TFC
set NLS_LANG=AMERICAN_AMERICA.AL32UTF8
sqlplus tfc/tfc @"C:\TFC\SQL\5_Carga_Datos_Finales.sql"

Este fichero será el encargado de ejecutar el script SQL “5_Carga_Datos_Finales.sql” que se encarga de crear las secuencia VIA_SEQ de la tabla VIA y la secuencia MUNICIPIO_SEQ , de la tabla MUNICIPIO.

Debido a que la población “Canonja (La)” no existe en los años 2007, 2008 y 2009, se inserta manualmente con valores a 0 y su correspondiente Provincia (Tarragona)

6_Carga_Tabla_Hechos.sql

cd C:\TFC
set NLS_LANG=AMERICAN_AMERICA.AL32UTF8
sqlplus tfc/tfc @"C:\TFC\SQL\6_Carga_Tabla_Hechos.sql"

Finalmente se carga la tabla de Hechos.

Este fichero será el encargado de ejecutar el script SQL “6_Carga_Tabla_Hechos.sql” que se encarga de cargar la tabla TRANSITO con los datos finales para ser analizados.
7 Configuración de Microsoft SQL Server Analysis Services 2012

Para generar nuevos informes, seleccionamos Microsoft Visual Studio 2010 y elegimos “New Project”

Creamos una nueva conexión a la base de datos XE, con el usuario tfc

Finalizamos la creación del Data Source, donde posteriormente se crearán los informes.
Construcción y Explotación de un almacén de datos para el análisis de información sobre tránsito de vehículos

Roberto Sarabia Crespo
8 Definición de Informes

Todos los informes están guardados y definidos con Microsoft Visual Studio 2010 en el proyecto Report_TFC

8.1 Informe del Total de Vehículos

```sql
SELECT     FECHA.ID_ANYO, MUNICIPIO.PROVINCIA, MUNICIPIO.NOMBRE, VEHICULO.TIPO, TRANSITO.NUM_VEHICULOS AS TOTAL
FROM         TRANSITO, MUNICIPIO, FECHA, VEHICULO
WHERE     TRANSITO.ID_MUNICIPIO = MUNICIPIO.ID_MUNICIPIO AND
TRANSITO.ID_ANYO = FECHA.ID_ANYO AND
TRANSITO.ID_TIPO_VEHICULO = VEHICULO.ID_TIPO_VEHICULO
GROUP BY MUNICIPIO.PROVINCIA, MUNICIPIO.NOMBRE, FECHA.ID_ANYO, VEHICULO.TIPO, TRANSITO.NUM_VEHICULOS
ORDER BY FECHA.ID_ANYO, MUNICIPIO.PROVINCIA, MUNICIPIO.NOMBRE, VEHICULO.TIPO
```

En el ejemplo del informe, se muestra el total de vehículos por tipo y año, además, organizado por Provincia y Municipio. Se habilita la posibilidad de “drilldown” por año y provincia.

<table>
<thead>
<tr>
<th>Total Vehículos</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Automóviles</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Barcelona</td>
</tr>
<tr>
<td>Granada</td>
</tr>
<tr>
<td>Valladolid</td>
</tr>
<tr>
<td>Almería</td>
</tr>
<tr>
<td>Albacete</td>
</tr>
<tr>
<td>Albacete</td>
</tr>
<tr>
<td>Alicante</td>
</tr>
<tr>
<td>Almería</td>
</tr>
<tr>
<td>Almería</td>
</tr>
<tr>
<td>Almería</td>
</tr>
<tr>
<td>Almería</td>
</tr>
<tr>
<td>Almería</td>
</tr>
<tr>
<td>Ávila</td>
</tr>
<tr>
<td>Ávila</td>
</tr>
<tr>
<td>Ávila</td>
</tr>
<tr>
<td>Albacete</td>
</tr>
<tr>
<td>Belchite</td>
</tr>
<tr>
<td>Bellacron</td>
</tr>
<tr>
<td>Bellacron</td>
</tr>
<tr>
<td>Bellacron</td>
</tr>
<tr>
<td>Bellacron</td>
</tr>
</tbody>
</table>
8.2 Informe del Total de Conductores

```
SELECT     TRANSITO.NUM_CONDUCTORES AS TOTAL, MUNICIPIO.NOMBRE, MUNICIPIO.PROVINCIA, GENERO.NOMBRE_GENERO, TRANSITO.ID_ANYO, PERMISO.TIPO_PERMISO
FROM         TRANSITO, MUNICIPIO, FECHA, GENERO, PERMISO
WHERE     TRANSITO.ID_MUNICIPIO = MUNICIPIO.ID_MUNICIPIO AND TRANSITO.ID_ANYO = FECHA.ID_ANYO AND TRANSITO.ID_GENERO = GENERO.ID_GENERO AND TRANSITO.ID_PERMISO = PERMISO.ID_PERMISO
GROUP BY MUNICIPIO.NOMBRE, MUNICIPIO.PROVINCIA, GENERO.NOMBRE_GENERO, TRANSITO.ID_ANYO, TRANSITO.NUM_CONDUCTORES, PERMISO.TIPO_PERMISO
```

En el ejemplo del informe, se muestra el total de conductores por género y año y tipo de permiso, además, organizado por Provincia y Municipio. Se habilita la posibilidad de “drilldown” por año y provincia.
### Informe del % de Vehículos respecto Población

```sql
SELECT FECHA.ID_ANYO, MUNICIPIO.PROVINCIA, MUNICIPIO.NOMBRE, VEHICULO.TIPO, TRANSITO.NUM_VEHICULOS AS VEHICULOS, TRUNC(TRANSITO.NUM_VEHICULOS * 100 / TRANSITO.NUM_HABITANTES) AS PORCENTAJE
FROM TRANSITO, MUNICIPIO, FECHA, VEHICULO
WHERE TRANSITO.ID_MUNICIPIO = MUNICIPIO.ID_MUNICIPIO AND TRANSITO.ID_ANYO = FECHA.ID_ANYO AND TRANSITO.ID_TIPO_VEHICULO = VEHICULO.ID_TIPO_VEHICULO
GROUP BY MUNICIPIO.PROVINCIA, MUNICIPIO.NOMBRE, FECHA.ID_ANYO, VEHICULO.TIPO, TRANSITO.NUM_VEHICULOS, TRANSITO.NUM_HABITANTES
```

<table>
<thead>
<tr>
<th>PROVINCIA</th>
<th>NOMBRE</th>
<th>TIPO</th>
<th>VEHICULOS</th>
<th>HABITANTES</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barcelona</td>
<td>Albora</td>
<td>Automóvil</td>
<td>8516</td>
<td>11278</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Motor</td>
<td>9202</td>
<td>11278</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>1441</td>
<td>11278</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Camiones y Furgonetas</td>
<td>1243</td>
<td>11278</td>
<td>11</td>
</tr>
<tr>
<td>Alacantina</td>
<td>Otros Vehículos de Motor</td>
<td>2490</td>
<td>2423</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>1207</td>
<td>2423</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>370</td>
<td>2423</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Automóvil</td>
<td>1247</td>
<td>2423</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Alicante</td>
<td>Motor</td>
<td>8026</td>
<td>9260</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>974</td>
<td>9260</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>2078</td>
<td>9260</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Automóvil</td>
<td>4974</td>
<td>9260</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>Ametilla del Valles (L')</td>
<td>Motor</td>
<td>6700</td>
<td>7796</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Automóvil</td>
<td>4120</td>
<td>7796</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>1096</td>
<td>7796</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>1476</td>
<td>7796</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Arenys de Mar</td>
<td>Automóvil</td>
<td>5399</td>
<td>14449</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>10346</td>
<td>14449</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>2446</td>
<td>14449</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>1301</td>
<td>14449</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Arenys de Munt</td>
<td>Automóvil</td>
<td>5812</td>
<td>8023</td>
<td>47</td>
<td></td>
</tr>
</tbody>
</table>
```

En el ejemplo del informe, se muestra el porcentaje de vehículos por tipo respecto al número de habitantes por año, que se podrá seleccionar al ejecutar el informe, además, organizado por Provincia y Municipio. Al ejecutar el informe se podrá seleccionar uno o varios años.
8.4 Informe de la Densidad de Población y Densidad de Vehículos

```
SELECT NOMBRE, PROVINCIA, ID_ANYO, HABITANTES, EXTENSION, 
      SUM(VEHICULOS) AS VEHICULOS, trunc(HABITANTES / EXTENSION) AS DEN_HAB,  
      trunc(SUM(VEHICULOS) / EXTENSION) AS DEN_TRAF  
FROM (SELECT MUNICIPIO.NOMBRE, MUNICIPIO.PROVINCIA, TRANSITO.ID_ANYO, TRANSITO.NUM_HABITANTES AS HABITANTES, MUNICIPIO.EXTENSION,  
       TRANSITO.NUM_VEHICULOS AS VEHICULOS  
FROM MUNICIPIO, TRANSITO, FECHA  
WHERE MUNICIPIO.ID_MUNICIPIO = TRANSITO.ID_MUNICIPIO AND TRANSITO.ID_ANYO = FECHA.ID_ANYO  
GROUP BY MUNICIPIO.NOMBRE, MUNICIPIO.PROVINCIA, MUNICIPIO.EXTENSION, TRANSITO.ID_ANYO, TRANSITO.NUM_VEHICULOS, TRANSITO.NUM_HABITANTES)  
derivedtbl_1  
GROUP BY NOMBRE, PROVINCIA, ID_ANYO, HABITANTES, EXTENSION
```

En el ejemplo del informe, se muestran los habitantes, vehículos y extensión, así como la densidad de población (habitantes/km²) y la densidad de tráfico (vehículos/km²) por año, además, organizado por Provincia y Municipio.
8.5 Informe del Número de Vehículos / Número de Radares

```
select ID_ANYO,NOMBRE, PROVINCIA,VEHICULOS ,TIPO,sum(RADARES) as radares,decode(sum(radares),0,0,VEHICULOS/sum(RADARES)) as RATIO
from(
    select FECHA.ID_ANYO, MUNICIPIO.NOMBRE,
    MUNICIPIO.PROVINCIA,VEHICULO.TIPO,TRANSITO.NUM_VEHICULOS AS VEHICULOS ,TRANSITO.NUM_RADARES as radares
    FROM TRANSITO, MUNICIPIO, FECHA, VEHICULO
    WHERE TRANSITO.ID_MUNICIPIO = MUNICIPIO.ID_MUNICIPIO AND
    TRANSITO.ID_ANYO = FECHA.ID_ANYO and TRANSITO.ID_TIPO_VEHICULO = VEHICULO.ID_TIPO_VEHICULO
    GROUP BY FECHA.ID_ANYO, MUNICIPIO.NOMBRE,
    MUNICIPIO.PROVINCIA,VEHICULO.TIPO,TRANSITO.NUM_VEHICULOS,
    TRANSITO.NUM_RADARES
)
group by ID_ANYO,NOMBRE, PROVINCIA,VEHICULOS,TIPO
```

Número Vehículos / Número Radares

**Año 2008**

<table>
<thead>
<tr>
<th>PROVINCIA</th>
<th>NOMBRE</th>
<th>VEHICULOS TIPO</th>
<th>RADARES</th>
<th>RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barcelona</td>
<td>Albora</td>
<td>Motor</td>
<td>9302</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Camiones y Forjones</td>
<td>1243</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Automóvil</td>
<td>6618</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>1414</td>
<td>0</td>
</tr>
<tr>
<td>Aquafreda</td>
<td></td>
<td>Motor</td>
<td>1507</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Camiones y Forjones</td>
<td>370</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Automóvil</td>
<td>1247</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>290</td>
<td>0</td>
</tr>
<tr>
<td>Alella</td>
<td></td>
<td>Motor</td>
<td>974</td>
<td>974</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Camiones y Forjones</td>
<td>8026</td>
<td>8026</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Automóvil</td>
<td>4674</td>
<td>4674</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>2078</td>
<td>2078</td>
</tr>
<tr>
<td>Ametella del Vallés (L')</td>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>1475</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Motor</td>
<td>6700</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Camiones y Forjones</td>
<td>1096</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Automóvil</td>
<td>4128</td>
<td>0</td>
</tr>
<tr>
<td>Arenys de Mar</td>
<td>10146</td>
<td>Motor</td>
<td>1</td>
<td>10146</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>2446</td>
<td>2446</td>
</tr>
</tbody>
</table>

En el ejemplo del informe, se muestra el ratio de vehículos y radares por tipo y por año, organizado por Provincia y Municipio. Al ejecutar el informe se podrá seleccionar uno o varios años.
8.6 Informe del % de Conductores por radar

```sql
select ID_ANYO, NOMBRE, PROVINCIA, CONDUCTORES, sum(RADARES), trunc(decode(sum(RADARES), 0, 0, (CONDUCTORES/sum(RADARES))*100/CONDUCTORES)) as PERCENT
from(
  select ID_ANYO, NOMBRE, PROVINCIA, CONDUCTORES, RADARES, id_via
  from(
    select FECHA.ID_ANYO, MUNICIPIO.NOMBRE, MUNICIPIO.PROVINCIA, TRANSITO.NUM_CONDUCTORES AS CONDUCTORES, TRANSITO.id_via, TRANSITO.NUM_RADARES as radares
    from TRANSITO, MUNICIPIO, FECHA
    where TRANSITO.ID_MUNICIPIO = MUNICIPIO.ID_MUNICIPIO AND TRANSITO.ID_ANYO = FECHA.ID_ANYO
    group by FECHA.ID_ANYO, MUNICIPIO.NOMBRE, MUNICIPIO.PROVINCIA, TRANSITO.NUM_RADARES, TRANSITO.NUM_CONDUCTORES, TRANSITO.id_via
  )
  group by ID_ANYO, NOMBRE, PROVINCIA, RADARES, id_via
)
) group by ID_ANYO, NOMBRE, PROVINCIA, CONDUCTORES
```

En el ejemplo del informe, se muestra el porcentaje de conductores respecto al número de radares por año, además, organizado por Provincia y Municipio. Al ejecutar el informe se podrá seleccionar uno o varios años.
### 8.7 Indicador de conductores vs habitantes por género

En el ejemplo del informe, se muestra el indicador de conductores vs habitantes y por género y tipo de Permiso y año, además, organizado por Provincia y Municipio.
8.8 Indicador de Radares vs Vehículos

En el ejemplo del informe, se muestra el indicador de radares vs vehículos por año, además, organizado por Provincia y Municipio.

<table>
<thead>
<tr>
<th>Indicador Radares</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Automóvils</strong></td>
<td><strong>Carrocerías y Fundamentos</strong></td>
</tr>
<tr>
<td><strong>Vehículos</strong></td>
<td><strong>Radares</strong></td>
</tr>
<tr>
<td>Amarilla</td>
<td>6356</td>
</tr>
<tr>
<td>Azul</td>
<td>1107</td>
</tr>
<tr>
<td>Álava</td>
<td>4738</td>
</tr>
<tr>
<td>Aranzadi del Valle (C.)</td>
<td>3950</td>
</tr>
<tr>
<td>Arazuri</td>
<td>6389</td>
</tr>
<tr>
<td>Aranzadi de Mutil</td>
<td>3018</td>
</tr>
<tr>
<td>Argentina</td>
<td>5076</td>
</tr>
<tr>
<td>Artea</td>
<td>2520</td>
</tr>
<tr>
<td>Azua</td>
<td>1123</td>
</tr>
<tr>
<td>Arrazaga</td>
<td>1032</td>
</tr>
<tr>
<td>Arrazospat del Perdido</td>
<td>877</td>
</tr>
<tr>
<td>Badalona</td>
<td>6459</td>
</tr>
<tr>
<td>Badia del Vallès</td>
<td>6088</td>
</tr>
<tr>
<td>Bellaterra</td>
<td>944</td>
</tr>
<tr>
<td>Bellvís</td>
<td>1885</td>
</tr>
<tr>
<td>Bellmunt</td>
<td>1007</td>
</tr>
<tr>
<td>Barberà del Vallès</td>
<td>14044</td>
</tr>
<tr>
<td>Barcelona</td>
<td>600952</td>
</tr>
<tr>
<td>Begues</td>
<td>2706</td>
</tr>
<tr>
<td>Berga</td>
<td>7591</td>
</tr>
<tr>
<td>Begues i Riells</td>
<td>4146</td>
</tr>
<tr>
<td>Bescanó</td>
<td>838</td>
</tr>
<tr>
<td>Cabrera d’Anoia</td>
<td>1</td>
</tr>
<tr>
<td>Ceret</td>
<td>7211</td>
</tr>
</tbody>
</table>
8.9 Ratio de Vehículos x Conductor

<table>
<thead>
<tr>
<th>PROVINCIA NOMBRE</th>
<th>TIPO</th>
<th>VEHICULOS</th>
<th>CONDUCTORES</th>
<th>RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barcelona</td>
<td>Automóvil</td>
<td>6518</td>
<td>6854</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>1243</td>
<td>6854</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>5222</td>
<td>6854</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>1441</td>
<td>6854</td>
<td></td>
</tr>
<tr>
<td>Alguafría</td>
<td>Automóvil</td>
<td>1247</td>
<td>1447</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>370</td>
<td>1447</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>1957</td>
<td>1447</td>
<td>1</td>
</tr>
<tr>
<td>Albora</td>
<td>Automóvil</td>
<td>4974</td>
<td>6343</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>974</td>
<td>6343</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>8026</td>
<td>6343</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td>2078</td>
<td>6343</td>
<td></td>
</tr>
<tr>
<td>Ans. del Vallès (l')</td>
<td>Automóvil</td>
<td>4128</td>
<td>5030</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>1096</td>
<td>5030</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>6700</td>
<td>5030</td>
<td>1</td>
</tr>
<tr>
<td>Anenys de Mar</td>
<td>Automóvil</td>
<td>6395</td>
<td>8236</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td>1301</td>
<td>8236</td>
<td></td>
</tr>
</tbody>
</table>

En el ejemplo del informe, se muestra el ratio de vehículos por conductor, tipo de vehículo y año, además, organizado por Provincia y Municipio. Al ejecutar el informe se podrá seleccionar uno o varios años.
8.10 Cantidad de Vehículos / Superficie

<table>
<thead>
<tr>
<th>PROVINCIA</th>
<th>NOMBRE</th>
<th>TIPO</th>
<th>VEHICULOS</th>
<th>EXTENSION</th>
<th>RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barcelona</td>
<td>Albacete</td>
<td>Automóvil</td>
<td>6559</td>
<td>20</td>
<td>327</td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td></td>
<td>1258</td>
<td>20</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td></td>
<td>9302</td>
<td>20</td>
<td>465</td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td></td>
<td>1485</td>
<td>20</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Aygafreda</td>
<td>Automóvil</td>
<td>1248</td>
<td>8</td>
<td>156</td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td></td>
<td>974</td>
<td>8</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td></td>
<td>1918</td>
<td>8</td>
<td>239</td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td></td>
<td>206</td>
<td>8</td>
<td>57</td>
</tr>
<tr>
<td>Aílla</td>
<td>Automóvil</td>
<td></td>
<td>5019</td>
<td>10</td>
<td>501</td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td></td>
<td>574</td>
<td>10</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td></td>
<td>8140</td>
<td>10</td>
<td>614</td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td></td>
<td>2155</td>
<td>10</td>
<td>215</td>
</tr>
<tr>
<td>Ametlla del Vallés (L')</td>
<td>Automóvil</td>
<td></td>
<td>4231</td>
<td>14</td>
<td>302</td>
</tr>
<tr>
<td></td>
<td>Camiones y Furgonetas</td>
<td></td>
<td>1116</td>
<td>14</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td></td>
<td>6924</td>
<td>14</td>
<td>495</td>
</tr>
<tr>
<td></td>
<td>Otros Vehículos de Motor</td>
<td></td>
<td>1587</td>
<td>14</td>
<td>113</td>
</tr>
<tr>
<td>Arenys de Mar</td>
<td>Automóvil</td>
<td></td>
<td>6426</td>
<td>6</td>
<td>1071</td>
</tr>
<tr>
<td></td>
<td>Camiones y</td>
<td></td>
<td>1320</td>
<td>6</td>
<td>221</td>
</tr>
</tbody>
</table>

En el ejemplo del informe, se muestra la cantidad de vehículos por superficie y año, además, organizado por Provincia y Municipio. Al ejecutar el informe se podrá seleccionar uno o varios años.
9 Conclusiones

A lo largo de las distintas fases de este proyecto y según iba avanzando en el desarrollo del mismo, he ido adquiriendo conocimientos en las diferentes áreas y tecnologías en la construcción y explotación de un almacén de datos.

Desde la planificación del proyecto en su fase inicial, pasando por el análisis y acabando con la implantación del mismo, todas las fases han requerido un gran esfuerzo y tiempo sin los cuales hubiera sido imposible acabar en los plazos indicados.

Lo más importante de la experiencia adquirida es poder aplicarla en cualquier otro tipo de proyecto de explotación de datos.

10 Líneas de evolución futura

En la implantación del proyecto, se ha preparado han preparado las tablas para añadir información sobre comarcas, que no ha sido cubierto en el mismo.

Una vez se estén usando los informes, probablemente sería necesario realizar un ajuste, tanto de las sentencias SQL que se ejecuten, como de la parametrización de la máquina y de la propia base de datos.

Si el volumen de los datos aumentara, también sería necesario revisar la arquitectura de las tablas, pudiendo plantearse añadir tablas particionadas y uso de paralelismo para un acceso más rápido.

Por otra parte, si se quisiera planificar las cargas automáticamente usando datos incrementales, también sería necesario revisar y cambiar el proceso de ETL, ya que está diseñado para partir de cero, es decir, que necesita que no existan datos previamente.
11 Referencias

- Proyectos de ejemplo de Almacenes de Datos UOC
  
  
  
  http://docs.oracle.com/cd/E23943_01/bi.htm
  
  http://es.wikipedia.org/wiki/Almac%C3%A9n_de_datos
  
  http://es.wikipedia.org/wiki/Extract,_transform_and_load
  
  
  http://www.orafaq.com/wiki/SQL*Loader_FAQ
  
  http://es.wikipedia.org/wiki/Clasificaci%C3%B3n_de_carreteras
  
  http://es.wikipedia.org/wiki/Anexo:Red_de_Carreteras_de_Catalu%C3%B1a
  
  http://www.scgenealogia.org/catalunya/catalunya.htm
  
  http://es.wikipedia.org/wiki/Granularidad
  
  http://es.wikipedia.org/wiki/Tabla_de_dimensi%C3%B3n