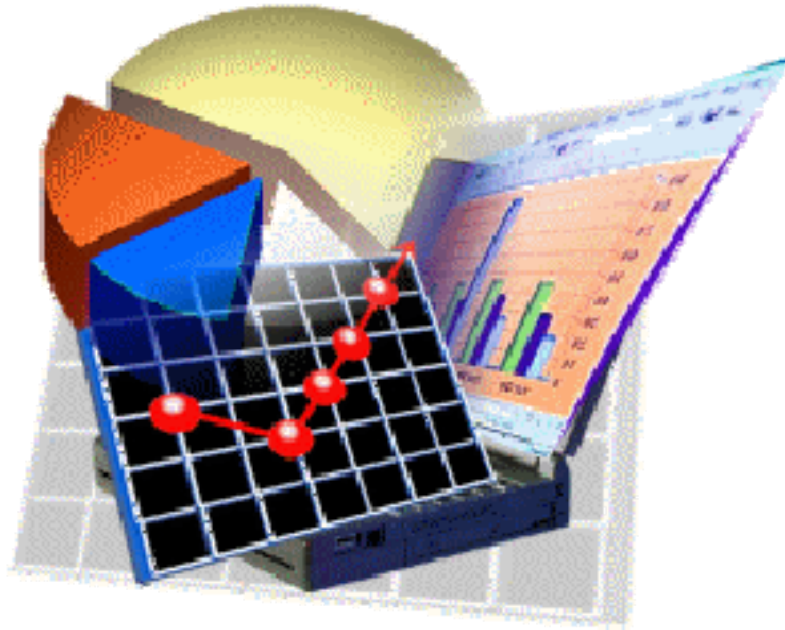


PROYECTO FINAL DE CARRERA

MEMORIA



CONSTRUCCIÓN Y EXPLOTACIÓN DE UN ALMACÉN DE DATOS DE PLANIFICACIÓN HIDROLÓGICA

Pedro Torremocha Molano
I.T.I.G.

Consultor : Pere Juanola Juanola
Enero 2011

Índice

1	DEDICATORIA Y AGRADECIMIENTOS	5
2	RESUMEN Y PALABRAS CLAVE.....	5
2.1	Resumen	5
2.2	Palabras clave.....	5
3	INTRODUCCIÓN	6
3.1	Justificación del TFC y contexto en el que se desarrolla	6
3.2	Objetivos del TFC	6
3.2.1.	<i>Relación con los objetivos de la asignatura.....</i>	<i>7</i>
3.2.2.	<i>Objetivos específicos del proyecto</i>	<i>7</i>
3.2.3.	<i>Informes a realizar en el proyecto</i>	<i>7</i>
3.3	Enfoque y método seguido	8
3.4	Planificación del proyecto.....	9
3.4.1.	<i>Fases.....</i>	<i>9</i>
3.4.2.	<i>Tareas adaptadas al plan docente.....</i>	<i>9</i>
3.4.3.	<i>Hitos y fechas claves.....</i>	<i>10</i>
3.4.4.	<i>Planificación de tareas.....</i>	<i>11</i>
3.4.5.	<i>Diagrama de Gantt.....</i>	<i>12</i>
3.4.6.	<i>Incidencias y riesgos.....</i>	<i>13</i>
3.5	Productos obtenidos	13
3.6	Breve descripción de otros capítulos de la memoria	14
3.6.1	<i>Análisis.....</i>	<i>14</i>
3.6.2	<i>Diseño.....</i>	<i>14</i>
3.6.3	<i>Descripción de las consultas e informes</i>	<i>14</i>
3.6.4	<i>Conclusiones.....</i>	<i>14</i>
4	ANÁLISIS	15
4.1	Datos de partida.....	15
4.2	Diagrama de casos de uso	17
4.3	Diagrama del modelo conceptual	17
4.3.1	<i>Identificar el Hecho.....</i>	<i>17</i>
4.3.2	<i>Encontrar la granularidad adecuada.....</i>	<i>18</i>
4.3.3	<i>Definir las dimensiones a emplear en el análisis.....</i>	<i>18</i>
4.3.4	<i>Establecer las jerarquías y los atributos de cada dimensión.....</i>	<i>18</i>
4.3.5	<i>Identificar las medidas a emplear.....</i>	<i>19</i>
4.3.6	<i>Definir las celdas.....</i>	<i>20</i>
4.3.7	<i>Establecer las restricciones de integridad</i>	<i>20</i>
4.3.8	<i>Estudio de viabilidad.....</i>	<i>20</i>
4.3.9	<i>Resultado</i>	<i>21</i>
5	DISEÑO	21
5.1	Diagrama de la arquitectura de software	21
5.2	Diagrama de la arquitectura de hardware.....	22
5.3	Diseño Lógico.....	22
5.4	Diseño Físico.....	23
5.5	Diseño de la Base de Datos	23
5.5.1	<i>Tabla de hechos TSITUACION.....</i>	<i>24</i>
5.5.2	<i>Tabla de dimensión EMBALSES</i>	<i>25</i>
5.5.3	<i>Tabla de dimensión RIOS.....</i>	<i>25</i>
5.5.4	<i>Tabla de dimensión AUTONOMIAS.....</i>	<i>26</i>
5.5.5	<i>Tabla TNIEVE.....</i>	<i>26</i>
5.6	Proceso E.T.L.	27
5.6.1	<i>Extracción.....</i>	<i>27</i>
5.6.2	<i>Transformación.....</i>	<i>27</i>

5.6.3	<i>Carga</i>	28
5.7	Diseño y descripción de los informes creados.....	31
5.7.1	<i>Creación del modelo de negocio</i>	31
5.7.2	<i>Informes desarrollados</i>	36
6	Capturas de Pantalla	39
6.1	Capacidad total vs Capacidad actual por río y año-mes	39
6.2	Capacidad total vs Capacidad actual por embalse y año-mes	40
6.3	Capacidad total vs Capacidad actual por CCAA y año-mes	42
6.4	Evolución por CCAA de la capacidad agregada de sus embalses.....	45
6.5	Comparación por embalses de cada CCAA entre un año y el anterior.....	46
6.6	Aportación de agua por río, embalse y año-mes	47
6.7	Entrada vs Salida media agregada anual por CCAA	47
6.8	Aportación por río, embalse y mes	48
6.9	Agua diaria en reserva disponible por año y mes	49
6.10	Agua de nieve	50
7	Conclusiones	53
8	Glosario	53
9	Bibliografía.	54
9.1	Publicaciones	54
9.2	Webs	54
10	Anexos	55
10.1	Scripts de creación de las tablas.....	55
10.1.1	<i>Creación de la tabla TSITUACION</i>	55
10.1.2	<i>Creación de la tabla EMBALSES</i>	55
10.1.3	<i>Creación de la tabla RIOS</i>	56
10.1.4	<i>Creación de la tabla AUTONOMIAS</i>	56
10.1.5	<i>Creación de la tabla TNIEVE</i>	57

Índice de figuras

Figura 1. Calendario de fechas claves	10
Figura 2. Planificación de Tareas	11
Figura 3. Diagrama de Gannt.	12
Figura 6. Diagrama de la arquitectura de software	21
Figura 8. Diseño físico.....	23
Figura 10. Tabla de hechos TSITUACION	24
Figura 11. Tabla de dimensión EMBALSES	25
Figura 12. Tabla de dimensión RIOS.....	25
Figura 13. Tabla de dimensión AUTONOMIAS	26
Figura 14. Tabla de previsión de NIEVE	26
Figura 15. Resultado carga tabla de hechos TSITUACION.....	30
Figura 16. Resultado carga tabla previsión de Nieve.	30
Figura 17. Conexión Discoverer Administrator	31
Figura 18. Crear EUL	31
Figura 19. Crear EUL (2 selección usuario)	32
Figura 20. Crear EUL (3 resultado selección usuario)	32
Figura 21. Creando EUL.....	33
Figura 22. Creación Área de negocio (Paso 1)	33
Figura 23. Creación Área de negocio (Paso 2)	33
Figura 24. Creación Área de negocio (Paso 3)	34
Figura 25. Creación Área de negocio (Paso 4).	34
Figura 26. Creación Área de negocio (Paso 5).	34
Figura 27. Área de negocio creada	35
Figura 28. Creación de libro de trabajo	36
Figura 29. Creación Área de negocio (Paso 2)	36
Figura 30. Creación de libro de trabajo (selección de variables).....	37
Figura 31. Creación de libro de trabajo (Elementos de página)	37
Figura 32. Creación de libro de trabajo (Condiciones).....	37
Figura 33. Creación de libro de trabajo (Cálculos).....	38

1 DEDICATORIA Y AGRADECIMIENTOS

A mí mujer y mis hijas por todo su cariño, apoyo y comprensión mostrados durante estos duros años, sin los cuales hubiera sido imposible llegar hasta aquí.

A mis compañeros de trabajo, en especial a Manuel, por sus ánimos y saber tirar de mí cuando lo he necesitado.

Y finalmente agradecer a mi consultor Pere Juanola su dedicación y apoyo ofrecido durante este semestre.

Cuando llegamos al final de una etapa y miramos atrás, nos damos cuenta que sin las buenas personas que están a nuestro lado, hubiera sido imposible alcanzar la meta.

2 RESUMEN Y PALABRAS CLAVE

2.1 Resumen

El objetivo de este proyecto es la construcción y explotación de un almacén de datos de planificación hidrológica para la Confederación Hidrográfica del Norte y Este de España (en adelante CHNE).

Este proyecto permitirá automatizar la recogida de información procedente de ficheros Excel, así como explotar dicha información con herramientas de Business Intelligence. Esto permitirá crear mejores modelos predictivos así como un alto rendimiento en el acceso a los datos.

Para lograr el objetivo del proyecto y tras un estudio pormenorizado de los datos de partida:

- Se diseñan los modelos conceptual, lógico y físico de los datos.
- Se crea un proceso automatizado de Extracción, Transformación (limpieza) y Carga de datos en las tablas que componen el Cubo multidimensional de datos (Cubo M-OLAP).
- Se crea un área de negocio y se diseñan los informes mediante modelos predictivos, en Oracle Discoverer, que ayuden a la CHNE a la toma de decisiones.

2.2 Palabras clave

Almacén de datos, CHNE, Planificación hidrográfica, ETL, Oracle, Discoverer, SQL, Análisis multidimensional, Data Warehouse, OLAP.

3 INTRODUCCIÓN

3.1 Justificación del TFC y contexto en el que se desarrolla

Este proyecto es la parte final de la carrera de Ingeniería Técnica en informática de Gestión, por lo tanto se debe plasmar todo lo aprendido durante estos años en las asignaturas de la carrera, especialmente los conocimientos adquiridos en las asignaturas de bases de datos.

Para este TFC asumimos el reto que supone la realización desde el principio hasta el final de un caso práctico real, para ello debemos definir las tareas a realizar así como una buena planificación del tiempo estimado a dedicar a cada tarea.

El punto de partida del presente proyecto es la petición por parte de la CHNE de un almacén de datos que permita automatizar la recogida de información procedente de ficheros Excel, con los datos correspondientes a los embalses del nordeste de España, y que permita explotar dicha información con herramientas de Business Intelligence. Esto va a permitir crear mejores modelos predictivos así como un alto rendimiento en el acceso a los datos.

Este proyecto se ha diseñado según la funcionalidades y facilidades que nos proporcionan los sistemas multidimensionales OLAP, ha sido creado con Oracle 10g, (que permitirá la carga, el almacenamiento y la posterior explotación del Data Warehouse con las herramientas de business intelligence de Oracle Discoverer), esto unido a la automatización de la carga de los datos acumulativos mensuales va a permitir a la CHNE un análisis rápido, exhaustivo y fiable de la información.

3.2 Objetivos del TFC

El objetivo principal del proyecto es adquirir experiencia en el diseño, construcción y explotación de un almacén de datos a partir de la información disponible hojas de cálculo. Aunque no es objetivo prioritario trabajar a fondo en cuestiones de rendimiento de la base de datos, los tiempos de respuesta de las diferentes consultas deben ser “razonables”.

En el presente TFC, el alumno debe poner en práctica todo lo aprendido en las diferentes asignaturas de que se compone la carrera de Ingeniería Técnica en Informática de Gestión.

Los objetivos principales son:

- Analizar y entender un caso práctico real y plasmarlo en el proyecto.
- Planificar y estructurar el desarrollo del proyecto.
- Sintetizar lo anterior en una solución coherente, fiable y realista.
- Elaborar la memoria del proyecto según lo indicado.

- Realizar una presentación del desarrollo y resultado del proyecto. **Relación con los objetivos de la asignatura**
-

El objetivo principal de la asignatura es la realización de un proyecto real, donde se pueda demostrar los conocimientos adquiridos en planificación, análisis, diseño e implementación de un almacén de datos multidimensional y explotarlo con herramientas de Business Intelligence.

Los objetivos a conseguir son:

- Familiarizarse con la arquitectura general de los almacenes de datos.
- Ampliar conocimientos y experiencia con los Data Warehouse.
- Familiarizarse con los lenguajes de consulta de bases de datos.
- Analizar las técnicas de tratamiento de la información de los almacenes de datos.
- Realizar los análisis, informes y consultas dirigidos a los usuarios finales del almacén de datos.

3.2.2. Objetivos específicos del proyecto

Los objetivos específicos del proyecto son:

- Creación un modelo multidimensional de datos con Oracle para almacenar los datos mensuales sobre las capacidades de los embalses que son objeto de estudio.
- Diseñar un modelo que permita automatizar la carga de los datos mensualmente.
- Realizar los informes estadísticos solicitados por la CHNE y que estos sean intuitivos y fáciles de interpretar por los usuarios.
- Crear las consultas y predicciones pedidas por el cliente.

3.2.3. Informes a realizar en el proyecto

Los informes solicitados por la CHNE son:

- Capacidad total vs. capacidad actual por río y año-mes.
- Capacidad total vs. capacidad actual para embalse y año-mes.
- Capacidad total vs. capacidad actual por comunidad autónoma y año-mes.
- Estadística comparativa anual de la evolución por comunidad autónoma de la capacidad agregada de todos sus embalses.
- Estadística comparativa de la capacidad media agregada de todos los embalses de una comunidad autónoma, de un año con al anterior (valores y % de incremento anual).
- Aportación (Entrada-Salida) por río, para embalse y año-mes; comparación con la media anual de aportación del embalse.

- Entrada vs. Salida, media agregada anual por comunidad autónoma.
- Aportación (Entrada-Salida) por río, embalse y mes; comparativa con la media anual de aportación del embalse.
- Agua diaria en reserva disponible en toda la confederación hidrográfica por año-mes.
- Gráfico de la evolución de agua de nieve equivalente prevista, por año y embalse. Muestra, si los hay, los casos de desbordamiento del embalse.
- Informes sobre los casos de desbordamiento de embalses (si los hubiera).

Todos estos informes deberán ofrecerse en Hm3 (hectómetros cúbicos) o en MI (millones de litros), según selección realizada por el usuario.

3.3 Enfoque y método seguido

Para la realización del presente proyecto se ha seguido el enfoque clásico en la construcción de los almacenes de datos.

- Análisis Previo.
 - Revisión del proyecto.
 - Búsqueda de información adicional.
 - Elaboración del plan de trabajo.
 - Análisis de los datos de entrada.
- Análisis de requerimientos.
 - Análisis detallado de los requerimientos.
 - Análisis de requisitos.
 - Primer borrador del análisis multidimensional.
- Diseño.
 - Diseño conceptual del modelo multidimensional.
 - Diseño físico del modelo multidimensional.
 - Diseño del proceso ETL.
- Implementación
 - Construcción del almacén de datos.
 - Proceso ETL.
 - Instalación de las herramientas de análisis.
 - Generación de informes.
 - Elaboración del manual de usuario.

3.4 Planificación del proyecto

Para la realización de la planificación del proyecto un punto importante ha sido el límite de tiempo disponible para su realización.

Para una buena planificación es necesario conocer las fases y tareas en las que se divide el proyecto, que básicamente serán las descritas en el programa de la asignatura.

3.4.1. Fases.

- **Plan de trabajo y análisis preliminar de requerimientos (Pec 1)**
Análisis preliminar (no detallado) con la enumeración y breve descripción de los elementos de análisis identificados (dimensiones, atributos, indicadores, etc.).
Número de informes a implementar así como el contenido de los mismos.
Análisis de las fuentes de datos operacionales proporcionados por la CHNE.
- **Análisis de requerimientos y diseño técnico (Pec 2)**
Análisis detallado de requerimientos basado en el análisis preliminar realizado.
Diseño del modelo dimensional conceptual y físico.
Diseño de los procedimientos de extracción de datos a alto nivel.
- **Implementación (Pec 3)**
Construcción del almacén de datos: base de datos, cargas, etc.
Instalación de la herramienta de explotación de datos.
Construcción de los informes y análisis de la información.

3.4.2. Tareas adaptadas al plan docente

1. Pec 1: Plan de Trabajo y Análisis preliminar.
 - Revisión del proyecto.
 - Plan de trabajo.
 - Análisis preliminar de requerimientos.
 - Análisis de los datos de entrada.
 - Primer esbozo del análisis multidimensional.
2. Pec 2: Análisis de requerimientos y diseño técnico.
 - Análisis detallado de requerimientos.
 - Diseño conceptual del modelo multidimensional.
 - Diseño físico del modelo multidimensional.
 - Preparación del hardware.
 - Instalación del software necesario.
 - Diseño proceso ETL. (Extracción, transformación y carga).
3. Pec 3: Implementación.

- Creación física del almacén de datos.
 - Finalización proceso ETL.
 - Instalación y configuración del proceso de análisis y explotación de datos.
 - Construcción de los informes.
 - Generación manual de usuario.
4. Entrega Final o Memoria del proyecto.
- Elaboración de la memoria.
 - Elaboración presentación del proyecto.
5. Debate virtual.

3.4.3. Hitos y fechas claves.

El calendario de la asignatura señala las siguientes fechas claves:

Nombre	Fecha de entrega
Pec 1: Plan de Trabajo y Análisis preliminar	05/10/2010
Pec 2: Análisis de requerimientos y diseño técnico	08/11/2010
Pec 3: Implementación	20/12/2010
Entrega Final	10/01/2011
Debate virtual	27/01/2011

Figura 1. Calendario de fechas claves

3.4.4. Planificación de tareas.

Para la realización de la siguiente planificación se ha tenido en cuenta tanto las fechas del calendario de la asignatura como mi disponibilidad horaria (2 horas/día).

ID	Nombre	Fecha de inicio	Fecha de fin	Duración días
0	TFC - Almacenes de Datos	23/09/2010	27/01/2011	126
1	Pec 1: Plan de Trabajo y Análisis preliminar	23/09/2010	05/10/2010	12
2	Lectura y comprensión del proyecto	23/09/2010	24/09/2010	1
3	Elaborar Plan de trabajo	24/09/2010	01/10/2010	7
4	Estudio de requerimientos	24/09/2010	27/09/2010	3
5	Estudio de los datos de partida	24/09/2010	26/09/2010	2
6	Identificación de Dimensiones y medidas	26/09/2010	27/09/2010	1
7	Análisis preliminar de requerimientos	27/09/2010	01/10/2010	4
8	Elaboración documento Pec 1	26/09/2010	04/10/2010	8
9	Entrega Pec 1	04/10/2010	05/10/2010	1
10	Pec 2: Análisis de requerimientos y diseño técnico	06/10/2010	08/11/2010	33
11	Repaso de conceptos	06/10/2010	10/10/2010	4
12	Repaso Análisis de Requerimientos	11/10/2010	12/10/2010	1
13	Diseño conceptual Almacén datos	12/10/2010	15/10/2010	3
14	Diseño lógico	15/10/2010	18/10/2010	3
15	Diseño extracción datos de origen	19/10/2010	22/10/2010	3
16	Transformación datos origen	22/10/2010	26/10/2010	4
17	Elaboración borrador Pec 2	26/10/2010	29/10/2010	3
18	Entrega Borrador	29/10/2010	30/10/2010	1
19	Corrección borrador	01/11/2010	03/11/2010	2
20	Elaboración documento Pec 2	03/11/2010	07/11/2010	4
21	Entrega Pec 2	07/11/2010	08/11/2010	1
22	Pec 3: Implementación	09/11/2010	20/12/2010	41
23	Repaso de conceptos	09/11/2010	11/11/2010	2
24	Instalación y configuración software	12/11/2010	13/11/2010	1
25	Construcción de la base de datos	14/11/2010	17/11/2010	3
26	Elaboración script carga de datos	18/11/2010	19/11/2010	1
27	Carga de datos	19/11/2010	20/11/2010	1
28	Pruebas proceso ETL	21/11/2010	23/11/2010	2
29	Creación informes solicitados	23/11/2010	04/12/2010	11
30	Elaboración borrador	05/12/2010	10/12/2010	5
31	Elaboración documento Pec 3	11/12/2010	19/12/2010	8
32	Entrega Pec 3	19/12/2010	20/12/2010	1
33	Entrega Final	21/12/2010	10/01/2011	20
34	Elaboración Memoria	21/12/2010	31/12/2010	10
35	Elaboración Presentación	31/12/2010	09/01/2011	9
36	Entrega Final	09/01/2011	10/01/2011	1
37	Debate virtual	24/01/2011	27/01/2011	3

Figura 2. Planificación de Tareas.

3.4.5. Diagrama de Gannt.

Teniendo en cuenta la planificación de la tabla anterior, se ha desarrollado la siguiente planificación con la aplicación Gannt Project.

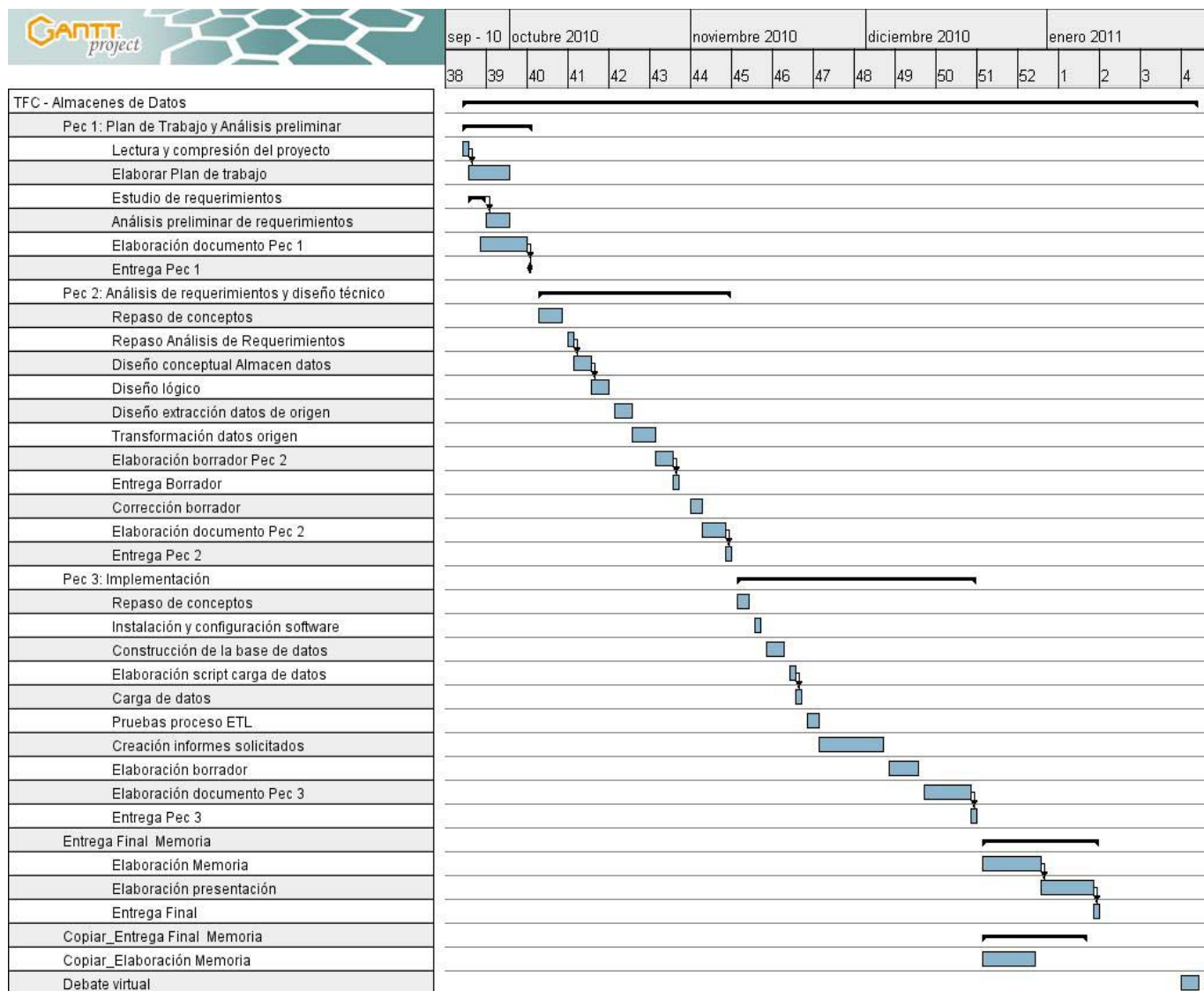


Figura 3. Diagrama de Gannt.

3.4.6. Incidencias y riesgos

A continuación se muestran las posibles incidencias a tener en cuenta en el desarrollo del proyecto y la solución prevista.

Problemas laborales.

Por peticiones puntuales urgentes, así como por la coincidencia en tiempo con mayor carga de trabajo que exijan por mi parte un mayor tiempo en el trabajo.

Para solventar estos problemas se haría ajustes en la programación del TFC y si fuera necesario y se permitiese, en estos casos no se entregaría borrador.

Problemas de salud.

Si se diese algún problema de salud que interfiriera de forma clara en el TFC, se intentaría ponerlo en conocimiento del tutor.

Solapamiento con otra asignatura.

En estos momentos se está cursando al mismo tiempo la asignatura de Sistemas Operativos.

La solución planteada en caso de solapamiento, es que el TFC siempre tendrá prioridad, hasta el punto de que si la interferencia es grande se abandonaría la otra asignatura.

Problemas técnicos.

En estos momentos, se dispone de 2 ordenadores, con lo que la información está replicada.

Si se diera el caso de problemas en los 2 ordenadores a la vez, se buscaría otro ordenador externo, ya que también se dispone de copia de seguridad en una memoria externa usb.

3.5 Productos obtenidos

Los productos obtenidos en el presente proyecto son:

- Documento con el Análisis preliminar.
- Documento con el Plan de Trabajo.
- Documento con el Análisis de requerimientos.
- Diseño conceptual, físico y lógico del modelo de datos.

- Scripts de creación de las tablas (SQL).
- Scripts de creación de índices (SQL).
- Scripts de carga de datos (SQL).
- Construcción del almacén de datos.
- Construcción del área de negocio (Oracle Discoverer Administrator).
- Informes solicitados por la CHNE (Oracle Discoverer Desktop).
- Documento explicativo con el manual para acceso a las bases de datos y a las herramientas de análisis.
- Documento explicativo con el funcionamiento de los procesos de carga.
- Copia de las pantallas de los informes creados.

3.6 Breve descripción de otros capítulos de la memoria

3.6.1 Análisis

En este apartado se incluye todo lo relativo al análisis funcional del proyecto, se parte del análisis de requisitos entregado por la CHNE. Se analizan las necesidades, los datos de partida. Se define la granularidad adecuada y se determinarán los hechos y las dimensiones.

3.6.2 Diseño

Este apartado se compone del diseño de la arquitectura del software así como el diseño conceptual, lógico y físico de los datos.

3.6.3 Descripción de las consultas e informes

En este apartado se describen detalladamente los informes y consultas realizados, mostrando las capturas de pantallas de los mismos.

3.6.4 Conclusiones

En este apartado se presentan las conclusiones finales del presente proyecto así como la valoración del mismo.

4 ANÁLISIS

En este apartado se analizan los datos de partida proporcionados por la CHNE y a partir de aquí se realiza el diseño conceptual del modelo multidimensional, se analizarán las dimensiones, las medidas y los atributos.

4.1 Datos de partida

La Confederación Hidrográfica Norte-Este nos envía 54 ficheros Excel (uno por mes y año), con los datos sobre las capacidades en Hm3 de los embalses y las aportaciones de los diferentes ríos a sus embalses.

- Los datos mensuales de dichos ficheros Excel están agrupados por **embalse**, identificado por su nombre, mostrándonos los siguientes datos:
 - Nombre del Embalse.
 - Nombre del Río que abastece de agua al embalse.
 - Capacidad total del embalse (en Hm3).
 - Cantidad de agua embalsada en la fecha del fichero (en Hm3).
 - Porcentaje del agua embalsada con relación a la capacidad total del embalse.
 - Entrada de agua semanal (en Hm3).
 - Salida de agua semanal (en Hm3).
 - Porcentaje de variación entre la entrada y salida semanal de agua.
 - Cantidad de agua embalsada en el mismo mes del año anterior (en Hm3).
 - Porcentaje de agua embalsada en el mismo mes del año anterior con relación a la capacidad total del embalse.

Los campos **Embalse** y **Río** tienen valores alfanuméricos.

Los campos **Capacidad**, **Contenido**, **Porcentaje**, **Incremento**, **ContenidoAnt** (cantidad año anterior) y **PorcentajeAnt** (% año anterior) tienen valores numéricos enteros.

Los campos **Entrada** y **Salida** tienen valores numéricos decimales.

Valores que pueden presentar los campos:

- **Contenido**: Pueden ir desde 0 hasta el valor del campo Capacidad.
- **Porcentaje**: Desde 0 hasta 100.
- **Entrada**: Cualquier valor numérico, puede haber valores nulos.
- **Salida**: Cualquier valor numérico menor o igual que el valor del campo Contenido, pueden aparecer nulos.
- **Incremento**: Cualquier valor de 0 a 100.

- **ContenidoAnt:** Desde 0 hasta el valor del campo Capacidad
- **PorcentajeAnt:** Valores numéricos desde 0 hasta 100.

Cada uno de estos ficheros tiene 10 filas de cabecera al principio del fichero y 5 filas de resumen al final del mismo. Estas filas no se tendrán en cuenta al hacer la carga de datos en la tabla de hechos (fact table) definitiva.

- Se dispone también de otro fichero Excel, agrupado por **Comunidad autónoma**, con los embalses de cada comunidad autónoma y los ríos que abastecen a cada uno de los embalses.

Los campos que tiene este fichero son: **Comunidad, Provincia, Embalse y Río**. Todos los campos son alfabéticos.

Este fichero tiene una única fila de cabecera.

- Por último, disponemos de una tabla con las estimaciones de agua procedente de nieve en Hm3 que recibirá cada embalse entre 2001 y 2005, los datos de esta tabla nos permitirán hacer unos cálculos estadísticos para mostrar la evolución de los hm3 de agua de nieve prevista y los casos en los que haya desbordamiento.

Los campos que tenemos en esta tabla son:

- **Embalse.**
- **Estimación2001.**
- **Estimación2002.**
- **Estimación2003.**
- **Estimación2004.**
- **Estimación2005.**

El campo **Embalse** tiene valores alfabéticos (nombre de los embalses) y los campos de **estimación** de nieve de los años 2001 a 2005 tienen valores numéricos enteros.

Tras un primer análisis a los datos de entrada, podemos afirmar que habrá que hacer una primera criba, codificación y recálculo de algunos datos para un mejor análisis posterior de los mismos.

Entre los errores encontrados se pueden encontrar:

- Números negativos y caracteres especiales en campos en los que los datos solamente pueden ser números enteros positivos.
- Campos nulos cuando deberían tener valor.
- Datos pocos significativos, en algunos casos sería mejor almacenar la información redondeando a la décima en lugar de a la unidad.

4.2 Diagrama de casos de uso

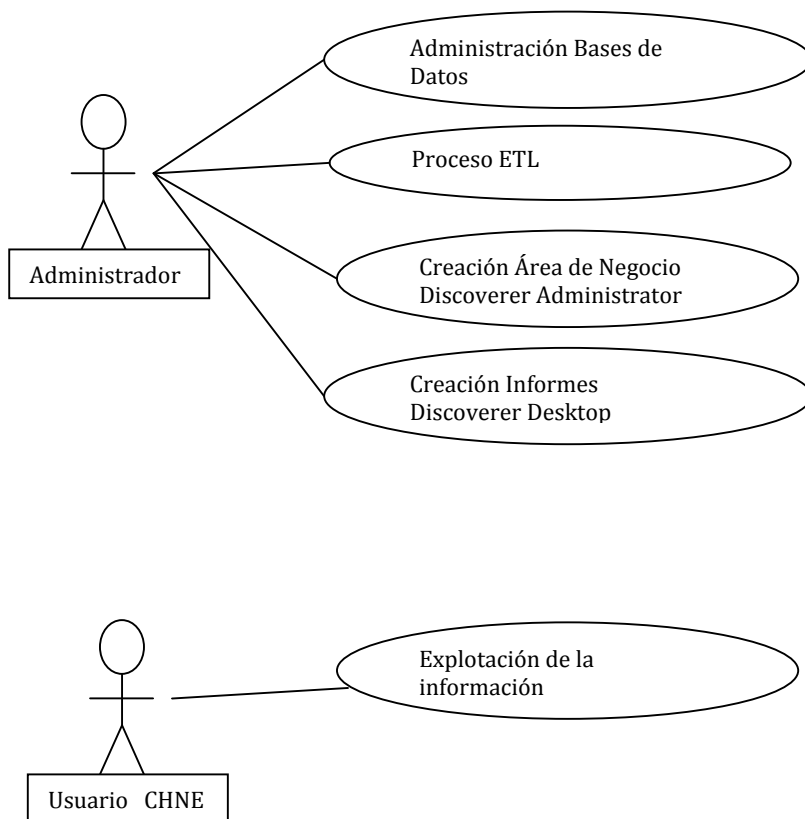


Figura 4. Diagrama de Casos de usos

4.3 Diagrama del modelo conceptual

El diseño conceptual del modelo multidimensional más apropiado al proyecto requiere de la ejecución de unos pasos previos que servirán para identificar el hecho, encontrar la granularidad adecuada, identificar las medidas y definir las dimensiones con sus atributos y jerarquías.

El diseño conceptual se realizará en 8 fases.

4.3.1 Identificar el Hecho

En esta primera fase se tiene que identificar correctamente la actividad que va a ser objeto de análisis (Hecho), que es el núcleo del modelo a realizar.

Del análisis de los datos partida suministrados por la CHNE y de los informes solicitados, se puede deducir que el hecho principal de este proyecto es el **volumen de agua** contenida en los embalses gestionados por la CHNE, **en un período determinado de tiempo**.

4.3.2 Encontrar la granularidad adecuada

En esta fase se debe decidir el nivel de granularidad o nivel de detalle en el que se quiere representar la información, se debe encontrar la granularidad más baja posible sin llegar a perder funcionalidad, ya que hay ocasiones en las que por intentar conseguir un grano demasiado fino se llega a ralentizar tanto la velocidad de cálculo como el acceso a los datos.

Según el estudio de los requerimientos del proyecto y tras analizar los datos de partida, se estima que la mejor opción es llegar al nivel inferior de granularidad (gránulo fino), que sería tener los datos mensuales de un embalse perteneciente a una determinada comunidad autónoma.

4.3.3 Definir las dimensiones a emplear en el análisis

Las dimensiones se emplean para organizar los datos a modo de ejes, y así poder desglosar la información según los criterios determinados en el análisis.

Dicho de otra forma las dimensiones son las variables por las cuales se analizarán los datos (Tiempo, Comunidad, embalse,..).

Después de analizar los datos y haber definido la granularidad adecuada, se pueden definir las siguientes dimensiones:

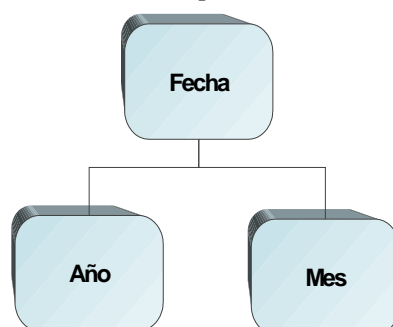
- Fecha.
 - Jerarquía: Año→Mes
- Autonomía.
- Río.
- Embalse.

4.3.4 Establecer las jerarquías y los atributos de cada dimensión

Una vez definidas las dimensiones del diseño del proyecto, debemos establecer para cada una de ellas los atributos (característica descriptiva) que permitirá filtrar los valores de una dimensión.

- Fecha.

Se dispone de información, con periodicidad mensual, desde Julio de 2001 hasta diciembre de 2005, se decide que sea una dimensión temporal de **Año** y **mes**.



- **Autonomía.**
No se establece ninguna jerarquía y los atributos son el código identificador y el nombre de la Autonomía.



- **Río.**
No se establece ninguna jerarquía y los atributos son el código identificador y el nombre del río.



- **Embalse.**
No se establece ninguna jerarquía y los atributos son el código identificador, el nombre del embalse.



4.3.5 Identificar las medidas a emplear

Las medidas o indicadores son los datos relevantes que se quieren estudiar.

Las medidas son los valores contenidos dentro de cada celda del cubo. Sus coordenadas son las dimensiones que sirven para identificarla.

Según el estudio de los informes y consultas que nos solicita la CHNE, las medidas que van a estar en la tabla de hechos serán:

- **Capacidad** (Volumen total de agua que admite el embalse).
- **VolumenAct** (Volumen de agua embalsada en un momento determinado).
- **Entradas** (Aportación de agua semanal).
- **Salidas** (Salida de agua semanal).
- **VolumenAnt** (Volumen de agua embalsada, en el mismo periodo del año anterior).

El **Incremento** no sería una medida a incluir en la tabla de hechos, ya que se puede considerar como una variable calculada ($\text{Incremento} = \text{Entradas} - \text{Salidas}$).

Los **porcentajes**, tanto de volumen como de incremento tampoco se incluirán en la tabla de hechos ya que se pueden calcular fácilmente.

Se incluirá otra tabla que contendrá las estimaciones previstas de nieve por embalse y año, esta tabla tendrá como campos clave, el identificador del embalse y el año y como medida de estudio la estimación en hm³ de nieve prevista para el embalse.

4.3.6 Definir las celdas

En esta fase se definen las celdas que se pueden considerar interesantes para el modelo multidimensional que se está desarrollando y por tanto han de ser almacenadas y cuáles de ellas se pueden considerar derivadas de otras y por tanto no será necesario almacenarlas.

Para este proyecto se estima que habrá una única celda que representará los datos de un embalse en una fecha determinada.

4.3.7 Establecer las restricciones de integridad

De las cuatro dimensiones que se han definido anteriormente, (Fecha, Embalse, Autonomía y Río), las dimensiones que definen de forma única una celda, son Fecha y Embalse.

Las dimensiones Autonomía y Río se pueden determinar por la dimensión Embalse, ya que un Embalse estará asociado al mismo río y autonomía durante toda su vida, así pues, para determinar una celda individual sería suficiente con las dimensiones Fecha y Embalse.

En este caso en concreto sería mucho más eficiente evitar la normalización ya que el ahorro de espacio no es nada significado y durante las consultas se evitaría la multiplicidad de los join.

4.3.8 Estudio de viabilidad

Ahora se debe se debe establecer el espacio que ocuparán los datos del modelo multidimensional para determinar si el modelo se puede implementar.

Se parte de unos datos con información sobre **57** embalses, esta información es mensual y va desde julio de 2001 hasta diciembre de 2005, con lo que se dispone de información de **54** meses, esto hace un total de **3.078** instancias en la tabla de hechos.

Cada celda del cubo tiene la capacidad total, el volumen actual, las entradas, salidas y el volumen del año anterior. Haciendo un cálculo aproximado, el número de bits necesarios que se emplearán, será: 4 bytes para cada uno de los atributos y 3 para cada uno de los 4 índices, $20 + 12 = 32$ bytes. Así pues los datos ocuparán aproximadamente $3.078 * 32 = 98.496$ bytes.

El volumen de datos resultante es muy bajo (a nivel de Kbytes), por lo que el modelo se puede usar en cualquier servidor y no constituye ningún problema para el motor de bases de datos.

4.3.9 Resultado

El modelo del diagrama conceptual queda de la forma siguiente:

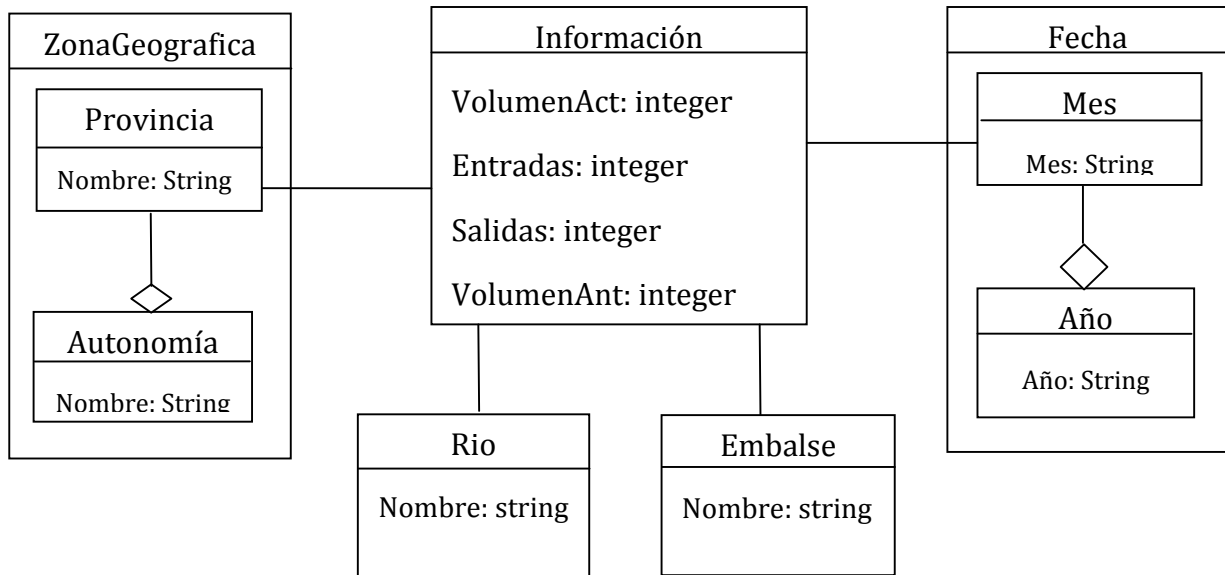


Figura 5. Diagrama conceptual

5 DISEÑO

5.1 Diagrama de la arquitectura de software

El software que se utilizará en el desarrollo del proyecto será:

- **Oracle Database Express Edition v10.2:** para la gestión de la Base de Datos.
- **SQL*loader:** para cargar los datos en la tabla desde los ficheros de origen.
- **Oracle SQL Developer, v1.0.15.27:** para la construcción de la tabla y los procesos ETL.
- **Oracle Discoverer:** como herramienta de Business Intelligent, para la construcción del área de negocio.

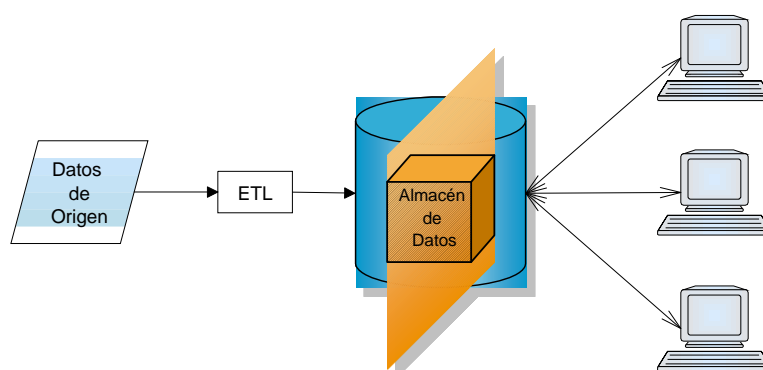


Figura 6. Diagrama de la arquitectura de software

5.4 Diseño Físico

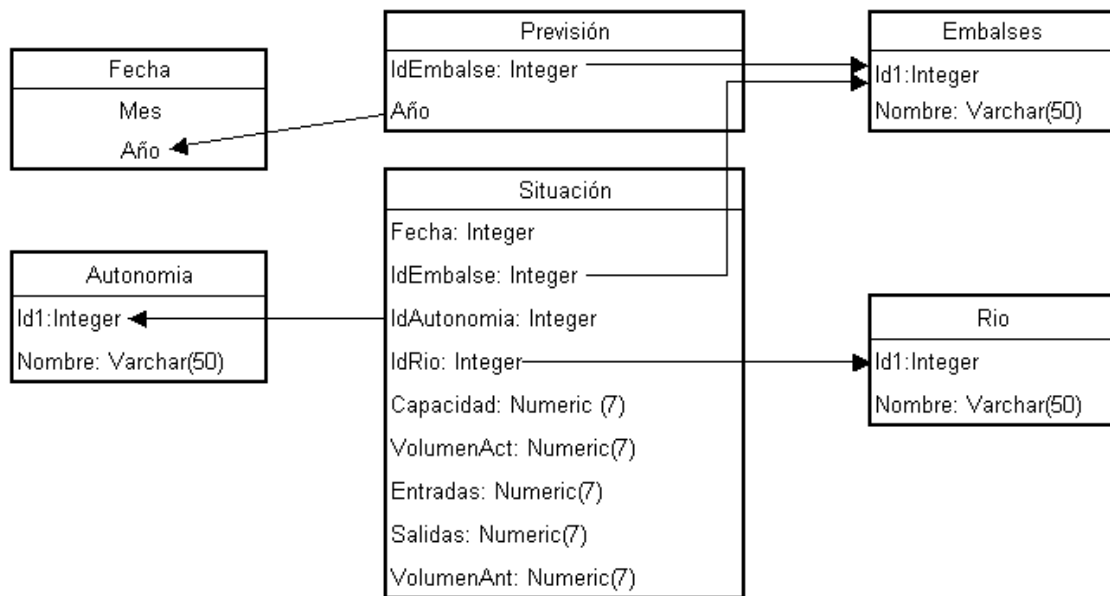


Figura 8. Diseño físico.

5.5 Diseño de la Base de Datos

Para realizar el diseño de la base de datos se ha tomado como base el esquema en estrella definido en documentos anteriores. La tabla de hechos (TSITUACION) enlaza con las tablas de dimensiones (EMBALSES, RIOS y AUTONOMIAS) mediante las claves foráneas.

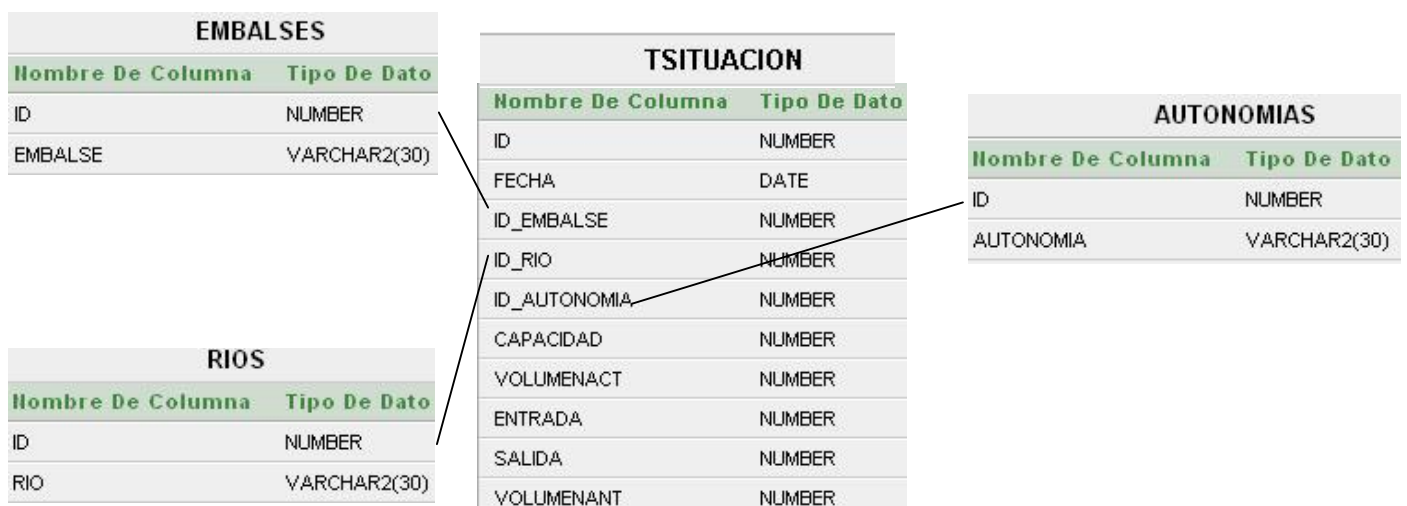


Figura 9. Esquema de la base de datos

5.5.1 Tabla de hechos TSITUACION

Es la tabla central del esquema que representa los datos en el contexto de las entidades de negocio.

Además de contener las claves foráneas por las que se conectará con las tablas de dimensiones, contendrá los campos que identifican a las medidas que se quieren analizar.

TSITUACION				
Nombre De Columna	Tipo De Dato	Nulo	Valor Por Defecto	Clave Primaria
ID	NUMBER	No	-	1
FECHA	DATE	No	-	-
ID_EMBALSE	NUMBER	Yes	-	-
ID_RIO	NUMBER	Yes	-	-
ID_AUTONOMIA	NUMBER	Yes	-	-
CAPACIDAD	NUMBER	Yes	-	-
VOLUMENACT	NUMBER	Yes	-	-
ENTRADA	NUMBER	Yes	-	-
SALIDA	NUMBER	Yes	-	-
VOLUMENANT	NUMBER	Yes	-	-

Figura 10. Tabla de hechos TSITUACION

Los campos principales son:

- **Fecha:** Este campo, se podría considerar como una dimensión temporal, que para este almacén de datos, se podría interpretar como año/mes.
- **Id_Embalse:** Es una clave foránea que permite conectar a la tabla de hechos con la tabla de dimensión EMBALSES, a través de su clave primaria ID.
- **Id_Rio:** Es una clave foránea que permite conectar a la tabla de hechos con la tabla de dimensión RIOS, a través de su clave primaria ID.
- **Id_Autonomia:** Es una clave foránea que permite conectar a la tabla de hechos con la tabla de dimensión AUTONOMIAS, a través de su clave primaria ID.
- **Capacidad:** Este campo contiene la capacidad máxima de un embalse (Id_Embalse) en hectómetros cúbicos. Podrá contener valores NULL.
- **VolumenAct:** Este campo contiene la cantidad de agua embalsada en un embalse (Id_Embalse) a una determinada fecha (Fecha), el volumen está en hectómetros cúbicos. Podrá contener valores NULL.
- **Entrada:** Este campo contiene el volumen de agua que ha suministrado un río (Id_Rio) en una fecha (Fecha), el valor está almacenado en hectómetros cúbicos. Podrá contener valores NULL.
- **Salida:** Este campo contiene el volumen de agua que ha vertido un embalse (Id_Embalse) a un río (Id_Rio) en una fecha (Fecha), el valor está almacenado en hectómetros cúbicos. Podrá contener valores NULL.

- **VolumenAnt:** Este campo contiene la cantidad de agua embalsada en un embalse (Id_Embalse) en el mismo mes del año anterior a la fecha (Fecha), el volumen está en hectómetros cúbicos. Podrá contener valores NULL.

5.5.2 Tabla de dimensión EMBALSES

En esta tabla se almacenarán todos los embalses que son objeto de estudio. Se relaciona con la tabla de hechos mediante la clave ID que se corresponde con la clave foránea ID_Embalse de la tabla de hechos.

Nombre De Columna	Tipo De Dato	Null	Valor Por Defecto	Clave Primaria
ID	NUMBER	No	-	1
EMBALSE	VARCHAR2(30)	Yes	-	-
				1 - 2

Figura 11. Tabla de dimensión EMBALSES

- ID: Es el índice, identificador único de los embalses, generado automáticamente de forma autoincremental.
- Embalse: Es el nombre de cada embalse.

5.5.3 Tabla de dimensión RIOS

En esta tabla se almacenarán todos los ríos que abastecen de agua a los embalses que son objeto de estudio. Se relaciona con la tabla de hechos mediante la clave ID que se corresponde con la clave foránea ID_Rio de la tabla de hechos.

Nombre De Columna	Tipo De Dato	Null	Valor Por Defecto	Clave Primaria
ID	NUMBER	No	-	1
RIO	VARCHAR2(30)	Yes	-	-
				1 - 2

Figura 12. Tabla de dimensión RIOS

- ID: Es el índice, identificador único de los ríos, generado automáticamente de forma autoincremental.
- Rio: Es el nombre del río.

5.5.4 Tabla de dimensión AUTONOMIAS

En esta tabla se almacenarán todas los Autonomías en las que se encuentran los embalses que son objeto de estudio. Se relaciona con la tabla de hechos mediante la clave ID que se corresponde con la clave foránea ID_Autonomia de la tabla de hechos.

Nombre De Columna	Tipo De Dato	Null	Valor Por Defecto	Clave Primaria
ID	NUMBER	No	-	1
AUTONOMIA	VARCHAR2(30)	Yes	-	-
				1 - 2

Figura 13. Tabla de dimensión AUTONOMIAS

- ID: Es el índice, identificador único de los embalses, generado automáticamente de forma autoincremental.
- Autonomia: Es el nombre de cada Autonomía.

5.5.5 Tabla TNIEVE

Esta tabla además de contener la clave foránea EMBALSE mediante la que se conectará con la tabla de dimensión EMBALSES, contendrá los campos que permitirán calcular el agua de nieve para cada embalse y año, que posteriormente nos servirá para pronosticar desbordamientos y periodos de sequía.

Nombre De Columna	Tipo De Dato	Null	Valor Por Defecto	Clave Primaria
ID	NUMBER	No	-	1
ANYO	DATE	No	-	2
ID_EMBALSE	NUMBER	No	-	3
PREVISION	NUMBER	Yes	-	-
				1 - 4

Figura 14. Tabla de previsión de NIEVE

- ID: Es el índice, identificador único de la tabla de previsión de NIEVE, generado automáticamente de forma autoincremental.
- ANYO: Campo de tipo fecha que contendrá el año para el cual se está calculando la previsión de agua de nieve.
- ID_EMBALSE: Es la clave foránea por la cual se conectará con la tabla de dimensión EMBALSES.
- PREVISION: Contiene el dato correspondiente a la previsión de agua de nieve para el embalse y el año correspondiente.

5.6 Proceso E.T.L.

Se llama proceso E.T.L. (Extraction, Transformation, Load), al proceso mediante el cual se realizan las funciones de:

- Extracción de los datos, ya sea de fuentes transaccionales o externas.
- Transformación de los datos (limpieza, consolidación, etc.).
- Carga en el almacén de datos, en este punto estaría la carga inicial de los datos, así como las agregaciones y el refresco periódico de los cambios.

5.6.1 Extracción

En este caso la empresa CHNE, nos ha suministrado 54 ficheros Excel con la información de la situación de los embalses desde julio de 2001 hasta diciembre de 2005.

Además de estos ficheros nos ha suministrado otro fichero Excel con información agrupada por embalse sobre su localización (Provincia y Autonomía), y el río que le suministra agua.

Así mismo la CHNE nos ha suministrado una tabla con la previsión de nieve prevista para cada año y embalse.

5.6.2 Transformación

Consiste en la limpieza, transformación e integración de los datos extraídos de las fuentes de origen, así como el cálculo de datos derivados, creación de claves, etc.

En los datos de origen siempre suelen existir anomalías y estas se eliminan mediante:

- Limpieza de datos: eliminar, corregir y completar datos, eliminar duplicados, etc.
- Estandarización: Unificar codificación, formatos, unidades de medida, tiempo, etc.

A los datos proporcionados por la CHNE se les ha sometido a una transformación previa debido a los siguientes errores:

- Embalses y ríos de los que tenemos información mensual y de los que no se suministra ningún tipo de información adicional. Ha habido un proceso de investigación sobre dichos embalses y ríos.
- Embalses y ríos con nomenclatura diferente en los ficheros de situación con respecto al fichero de identificación. Unificación de los nombres.
- Caracteres alfabéticos en campos numéricos. Se han transformado en valores nulos o 0. (Si (Valor_Campo) = “---“ entonces Valor_Campo = 0).

- Mediante una macro se crea una columna al principio de cada hoja con la fecha correspondiente.
- Mediante una macro se eliminan las filas y columnas innecesarias.
- Se han agrupado todos los ficheros mensuales en un único fichero de texto.

5.6.3 Carga

Esta fase consiste en pasar los datos desde el almacenamiento intermedio hasta el almacén de datos en sus correspondientes estructuras.

Durante la carga inicial se mueven grandes volúmenes de datos, pero en cambio durante los procesos de mantenimiento periódicos se mueven pequeños volúmenes de datos.

En este proyecto, una vez que los datos extraídos se han transformado, se procede a la carga de los ficheros resultantes de la ETL, aprovechando la facilidad que tiene el gestor de base de datos de Oracle para reconocer datos de ficheros de texto (csv), se cargarán directamente a tablas externas de Oracle.

No obstante se han creado los siguientes scripts para la carga posterior si el cliente (CHNE) lo considerase necesario:

Tabla temporal situación_tmp.

```
LOAD DATA
CHARACTERSET 'UTF8'
INFILE 'C:\UOC\TFC\situacion.csv'
BADFILE 'situacion.bad'
DISCARDFILE 'situacion.dsc'
APPEND
INTO TABLE SITUACION_TMP
FIELDS TERMINATED BY ";"
OPTIONALLY ENCLOSED BY ""
TRAILING NULLCOLS
(
  fecha      DATE "DD-Month-YY",
  embalse    CHAR "ltrim(rtrim(:embalse))",
  rio        CHAR "ltrim(rtrim(:rio))",
  capacidad  INTEGER,
  volumenact INTEGER,
  porcenact  INTEGER,
  entrada    INTEGER,
  salida     INTEGER,
  incremento INTEGER,
  volumenant INTEGER,
  porcentant INTEGER
)
```

Tabla temporal embalseaut.

```
LOAD DATA
CHARACTERSET 'UTF8'
INFILE 'C:\UOC\TFC\embalseaut.csv'
BADFILE 'embalseaut.bad'
DISCARDFILE 'embalseaut.dsc'
APPEND
INTO TABLE EMBALSEAUT
FIELDS TERMINATED BY ";"
OPTIONALLY ENCLOSED BY ""
TRAILING NULLCOLS
(
  embalse CHAR "ltrim(rtrim(:embalse))",
  rio CHAR "ltrim(rtrim(:rio))",
  autonomia CHAR "ltrim(rtrim(:autonomia))",
  provincia CHAR "ltrim(rtrim(:provincia))"
)
```

Tabla temporal nieve.

```
LOAD DATA
CHARACTERSET 'UTF8'
INFILE 'C:\UOC\TFC\nieve.csv'
BADFILE 'nieve.bad'
DISCARDFILE 'nieve.dsc'
APPEND
INTO TABLE NIEVE
FIELDS TERMINATED BY ";"
OPTIONALLY ENCLOSED BY ""
TRAILING NULLCOLS
(
  embalse CHAR "ltrim(rtrim(:embalse))",
  año INTEGER,
  prevision INTEGER
)
```

En una primera fase se cargarán las tablas de dimensiones: Embalses, Ríos y Autonomía, así como una tabla temporal con los datos de la situación de los embalses (Situación_tmp) y otra tabla temporal con la relación entre Embalse, Río y Autonomía.

La tabla de hechos definitiva se cargará mediante el cruce de la tabla temporal con las tablas de dimensiones.

Inserción en la tabla de hechos TSITUACION.

```

INSERT INTO "TSITUACION"
select  0,
        "SITUACION_TMP"."FECHA",
        "EMBALSES"."ID",
        "RIOS"."ID",
        "AUTONOMIAS"."ID",
        "SITUACION_TMP"."CAPACIDAD",
        "SITUACION_TMP"."VOLUMENACT",
        "SITUACION_TMP"."ENTRADA",
        "SITUACION_TMP"."SALIDA",
        "SITUACION_TMP"."VOLUMENANT"
FROM    "SITUACION_TMP",
        "EMBRIOAUT",
        "EMBALSES",
        "RIOS",
        "AUTONOMIAS"
where   "SITUACION_TMP"."EMBALSES"="EMBRIOAUT"."EMBALSE"
and     "SITUACION_TMP"."RIOS"="EMBRIOAUT"."RIO"
and     "AUTONOMIAS"."AUTONOMIA"="EMBRIOAUT"."AUTONOMIA"
and     "EMBALSES"."EMBALSE"="EMBRIOAUT"."EMBALSE"
and     "RIOS"."RIO"="EMBRIOAUT"."RIO"

```

Resultados [Explicar](#) [Describir](#) [SQL Guardado](#) [Historial](#)

3078 fila(s) insertada(s).

0,50 segundos

Figura 15. Resultado carga tabla de hechos TSITUACION

Inserción en la tabla de TNIEVE

```

INSERT INTO "TNIEVE"
select  0,
        "PREVISION"."AÑO" as "AÑO",
        "EMBALSES"."ID" as "ID_EMBALSE",
        "PREVISION"."PREVISION" as "PREVISION"
from    "EMBALSES" "EMBALSES",
        "PREVISION" "PREVISION"
where   "PREVISION"."EMBALSE"="EMBALSES"."EMBALSE"

```

Resultados [Explicar](#) [Describir](#) [SQL Guardado](#) [Historial](#)

350 fila(s) insertada(s).

0,08 segundos

Figura 16. Resultado carga tabla previsión de Nieve.

5.7 Diseño y descripción de los informes creados

Para realizar los informes solicitados, primero se debe crear el área de negocio en Oracle Discoverer Administrator, para posteriormente proceder a desarrollar los informes con la herramienta de Business Intelligent Oracle Discoverer Desktop.

5.7.1 Creación del modelo de negocio

Con Oracle Discoverer Administrator, definiremos el escenario para cada usuario.

Creamos un EUL (End User Layer), que es el metadato del modelo de datos, contendrá la información de las tablas, se definirán las relaciones, las jerarquías necesarias, etc.

El usuario y la clave de conexión serán los mismos que los utilizados anteriormente en Oracle Express (usuario: **SYSTEM**, clave: **uoc**).



Figura 17. Conexión Discoverer Administrator

Lo primero que tenemos que hacer es crear un EUL para nuestro usuario, en este caso se utilizará la información del usuario SYSTEM, se hará como muestran las siguientes imágenes:

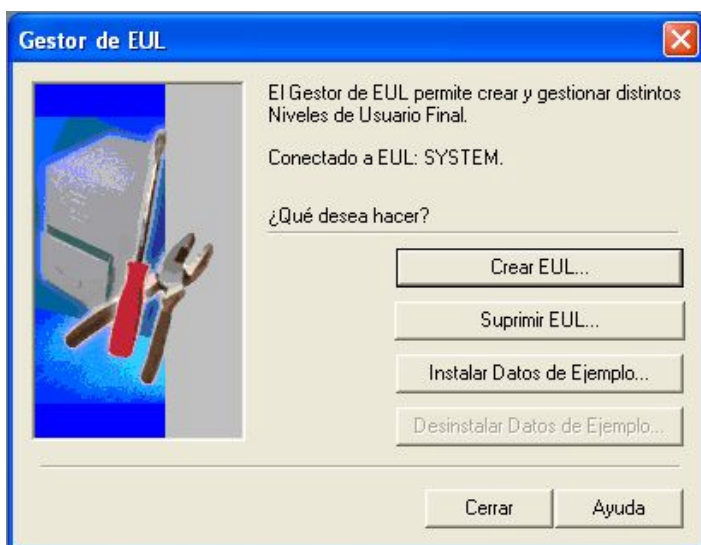


Figura 18. Crear EUL

Se marca *Seleccionar usuario existente* y *Otorgar acceso público*. Debemos seleccionar el usuario SYSTEM.

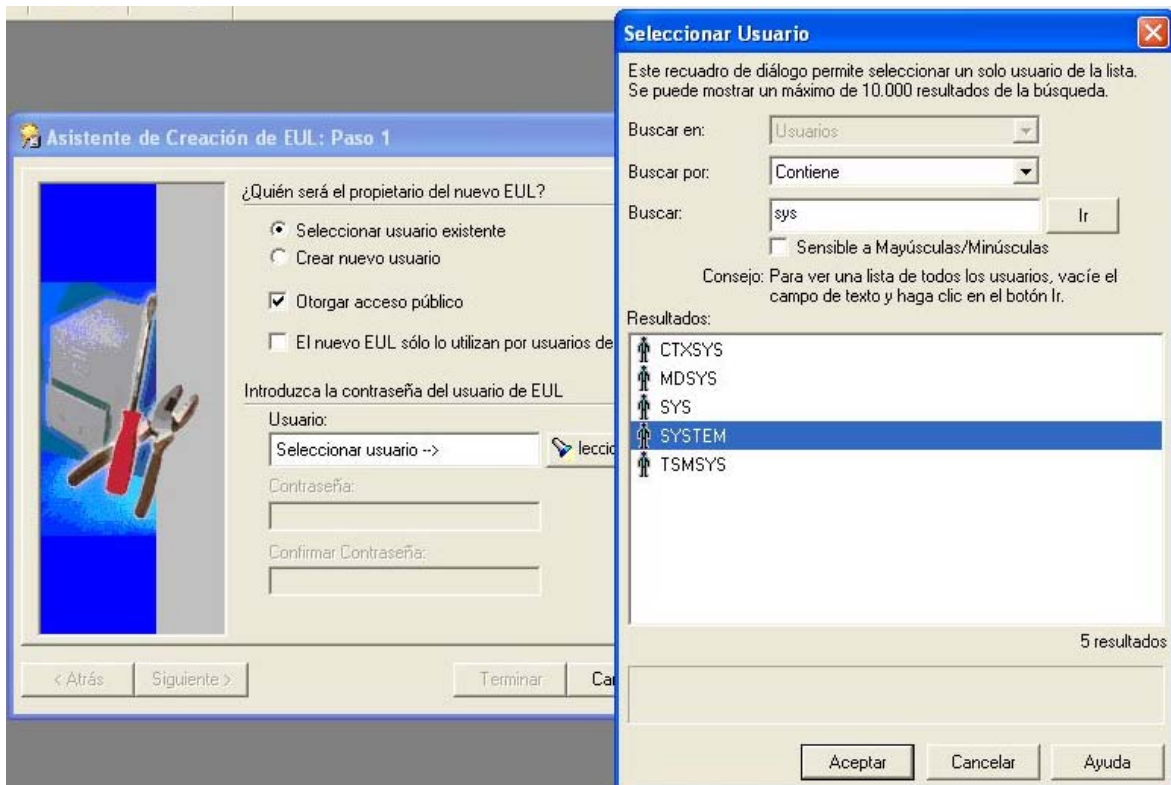


Figura 19. Crear EUL (2 selección usuario)



Figura 20. Crear EUL (3 resultado selección usuario)

Se pulsa *Terminar* y se empieza a crear el EUL correspondiente.



Figura 21. Creando EUL

Ahora se debe crear el área de negocio para el usuario y EUL creado anteriormente.



Figura 22. Creación Área de negocio (Paso 1)

Se debe definir el usuario que tendrá acceso al área de negocio a crear.

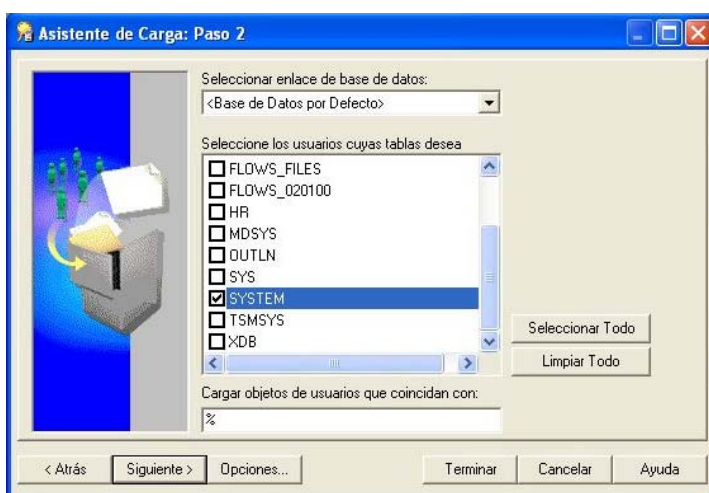


Figura 23. Creación Área de negocio (Paso 2)

Se seleccionan las tablas que compondrán el área de negocio.



Figura 24. Creación Área de negocio (Paso 3)



Figura 25. Creación Área de negocio (Paso 4).



Figura 26. Creación Área de negocio (Paso 5).

El área de negocio quedará de la forma siguiente:

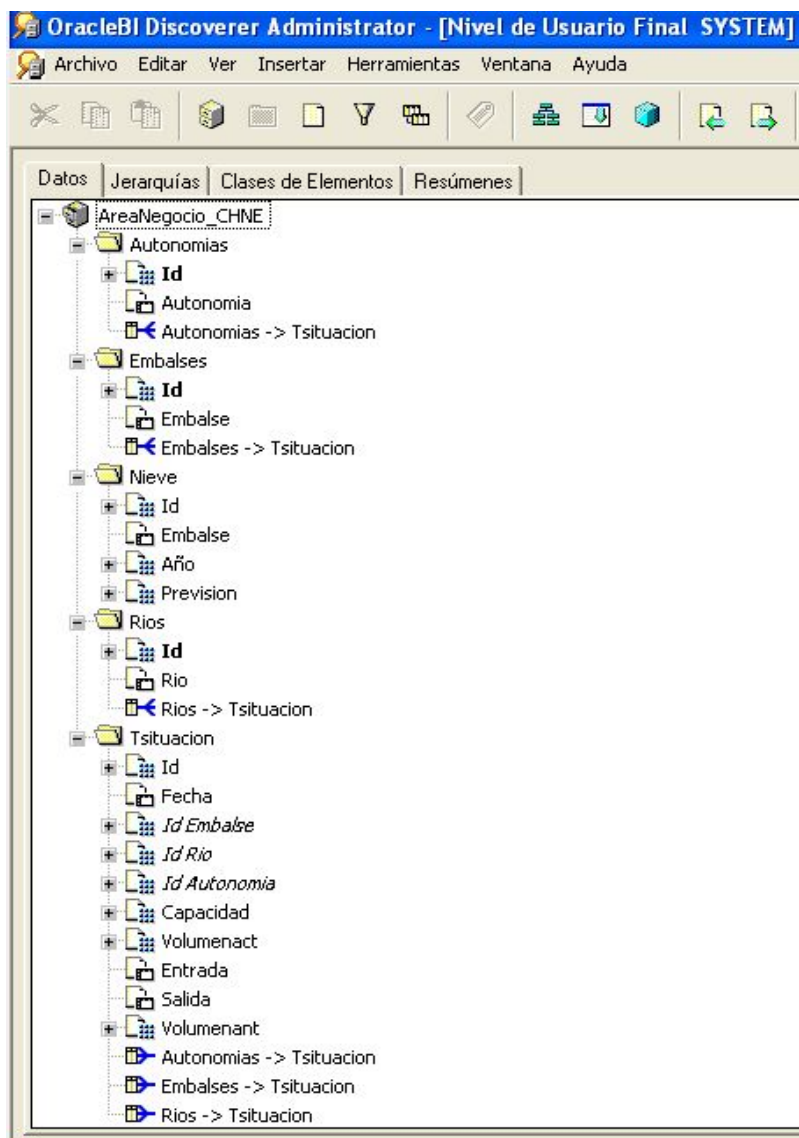


Figura 27. Área de negocio creada

5.7.2 Informes desarrollados

A partir del modelo conceptual de datos creado en Oracle Discoverer Administrador, se procede a desarrollar los informes solicitados en la herramienta de Business Intelligent Oracle Discoverer Desktop.

Los libros de trabajo que contendrán los informes solicitados se han ido creando con el asistente para creación de libros de trabajo que contiene Discoverer Desktop.



Figura 28. Creación de libro de trabajo

Se ha elegido las opciones de visualización en *Matriz* y *Matriz página-detalle*, ya que estas opciones mostrarán los datos multidimensionales y permitirá rotar las dimensiones horizontal o verticalmente.



Figura 29. Creación Área de negocio (Paso 2)

Se seleccionan las dimensiones y medidas, que son objeto de estudio en el informe a desarrollar, del área de negocio creada para ello.



Figura 30. Creación de libro de trabajo (selección de variables)

Ahora se debe elegir el criterio por el cual se agrupa la información a mostrar en la página:



Figura 31. Creación de libro de trabajo (Elementos de página)

Se podrán definir condiciones para limitar los resultados del análisis a realizar.



Figura 32. Creación de libro de trabajo (Condiciones)

Así mismo se podrán definir variables o medidas calculadas manualmente.

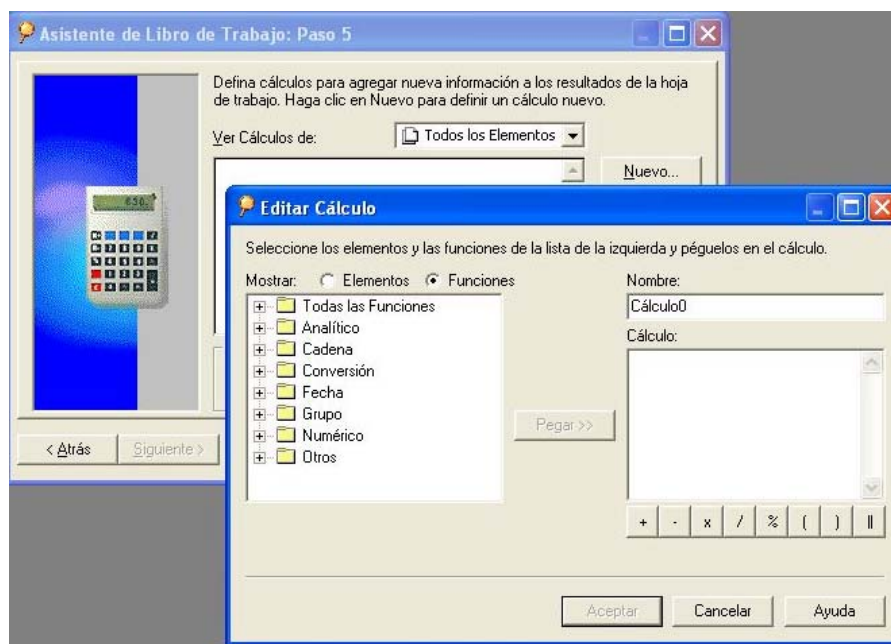


Figura 33. Creación de libro de trabajo (Cálculos)

Según los requerimientos realizados por la CHNE se han creado los siguientes libros:

- Total vs Actual por río y año-mes.
- Total vs Actual para embalse y año-mes.
- Total vs Actual por comunidad autónoma y año-mes.
- Evolución anual por Autonomía de la capacidad agregada.
- Evolución anual por Autonomía comparación año anterior.
- Aportación por río, para embalse y año-mes.
- Entrada vs Salida media agregada anual por Autonomía.
- Agua diaria en reserva.
- Nieve (agua de nieve):
 - Gráfico de la evolución del agua de nieve equivalente prevista por año y embalse.
 - Casos de desbordamiento del embalse.
 - Predicción de períodos de sequía.

Todos los informes se han desarrollado indistintamente en Hm3 y en millones de litros.

Todos los libros que contienen los informes, para un rápido y fácil acceso, se han guardado en la base de datos.

Dentro de cada libro hay una serie de informes con vistas distintas (hm3., Millones de litros, por un determinado embalse y año, por año y la relación de los embalses, etc.), el cliente podrá elegir el informe deseado seleccionando la pestaña adecuada a sus necesidades.

A continuación se muestran una las capturas de pantalla de los informes creados.

6 Capturas de Pantalla

6.1 Capacidad total vs Capacidad actual por río y año-mes

CHNE							
Capacidad total vs capacidad actual por Río y año-mes							
Elementos de Página: Año: 01 ▼							
	Jul	Ago	Sep	Oct	Nov	Dic	
AGUAS LIMPIAS							
Capacidad MAX	18	18	18	18	18	18	
Volumenact MAX	16	16	12	9	9	18	
Millones de litros	16.000	16.000	12.000	9.000	9.000	18.000	
%	100	100	100	100	100	100	
AGUAS VIVAS							
Capacidad MAX	8	8	8	8	8	8	
Volumenact MAX	1	1	0	0	0	0	
Millones de litros	1.000	1.000	0	0	0	0	
%	100	100					
ALBERCOS							
Capacidad MAX	33	33	33	33	33	33	
Volumenact MAX	18	13	8	7	5	4	
Millones de litros	18.000	13.000	8.000	7.000	5.000	4.000	
%	100	100	100	100	100	100	
ALBINA							
Capacidad MAX	6	6	6	6	6	6	
Volumenact MAX	4	4	3	3	3	4	
Millones de litros	4.000	4.000	3.000	3.000	3.000	4.000	
%	100	100	100	100	100	100	
ARAGÓN							
Capacidad MAX	470	470	470	470	470	470	
Volumenact MAX	223	119	50	76	95	74	

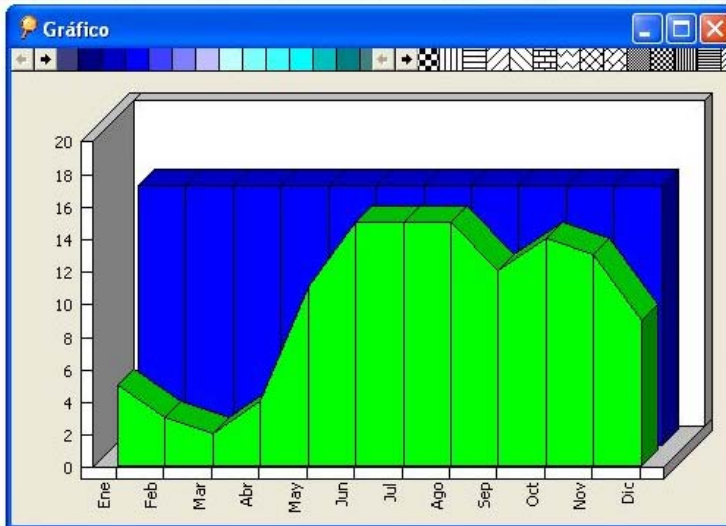
6.2 Capacidad total vs Capacidad actual por embalse y año-mes

CHNE

Capacidad total vs capacidad actual por embalse y año-mes
(en Hectómetros cúbicos)

Elementos de Página: **Embalse: CAVALLERS** **Año: 2005**

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Volumen Embalsado	5	3	2	4	11	15	15	15	12	14	13	9
Capacidad MAX	16	16	16	16	16	16	16	16	16	16	16	16



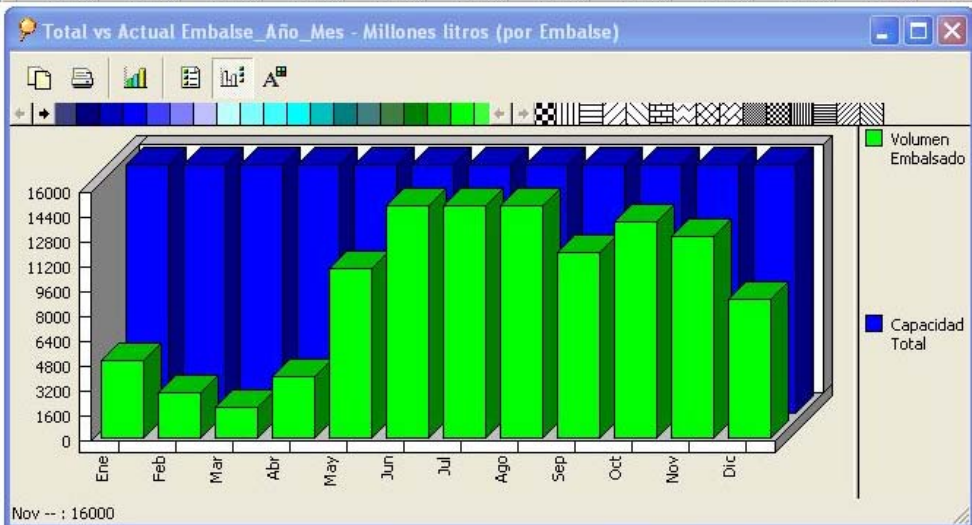
Hm3 (por Embalse) Millones litros (por Embalse)

CHNE

Capacidad total vs capacidad actual por embalse y año-mes
(en Millones de litros)

Elementos de Página: **Embalse: CAVALLERS** **Año: 2005**

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Volumen Embalsado	5.000	3.000	2.000	4.000	11.000	15.000	15.000	15.000	12.000	14.000	13.000	9.000
Capacidad Total	16.000	16.000	16.000	16.000	16.000	16.000	16.000	16.000	16.000	16.000	16.000	16.000



Hm3 (por Embalse) Millones litros (por Embalse) Hm3 (listado de embalses) Millones litros (listado de embalses)

CHNE

Capacidad total vs capacidad actual por embalse y año-mes

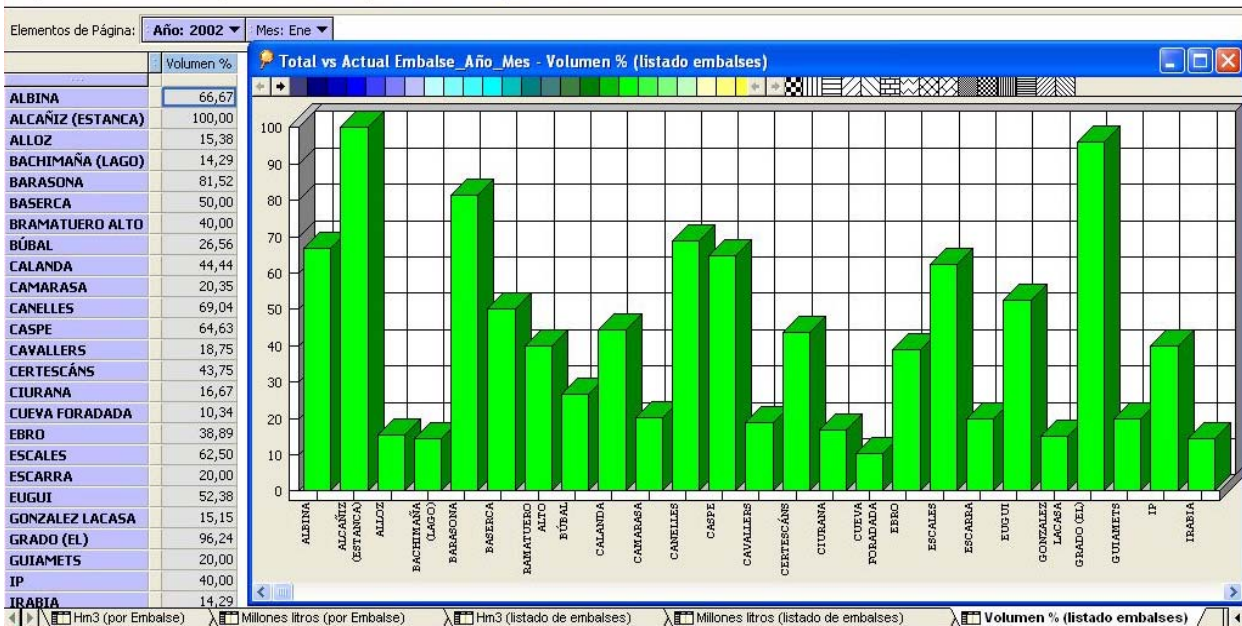
Elementos de Página: Año: 2002 Mes: Ene

	Volumen	Capacidad
ALBINA	4	6
ALCAÑIZ (ESTANCA)	7	7
ALLOZ	10	65
BACHIMAÑA (LAGO)	1	7
BARASONA	75	92
BASERCA	11	22
BRAMATUERO ALTO	2	5
BÚBAL	17	64
CALANDA	24	54
CAMARASA	23	113
CANELLES	475	688
CASPE	53	82
CAVALLERS	3	16
CERTESCÁNS	7	16
CIURANA	2	12
CUEVA FORADADA	3	29
EBRO	210	540
ESCALES	95	152
ESCARRA	1	5
EUGUI	11	21
GONZALEZ LACASA	5	33
GRADO (EL)	384	399
GUIAMETS	2	10
IP	2	5
IRABIA	2	14

Hm3 (por Embalse)
 Millones litros (por Embalse)
 Hm3 (listado de emb.)

CHNE

Capacidad total vs capacidad actual por embalse y año-mes



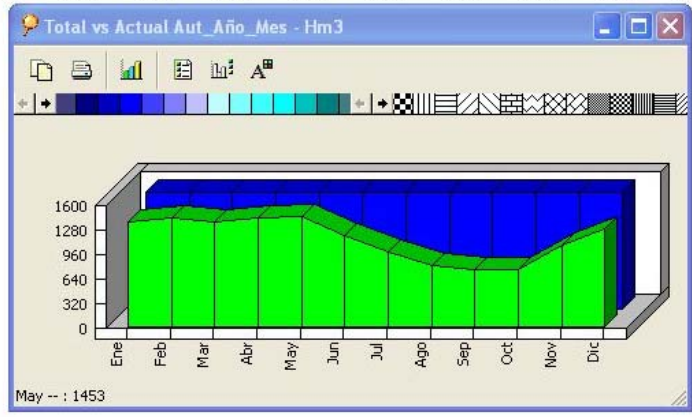
6.3 Capacidad total vs Capacidad actual por CCAA y año-mes

CHNE

Capacidad total vs capacidad actual por Autonomía y año-mes
(Volumen en Hectómetros cúbicos)

Elementos de Página: Autonomía: ARAGÓN Año: 2005

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Volumen Actual	1.372	1.427	1.373	1.429	1.453	1.200	978	806	750	750	1.067	1.292
Capacidad Total	1.534	1.534	1.534	1.534	1.534	1.534	1.534	1.534	1.534	1.534	1.534	1.534



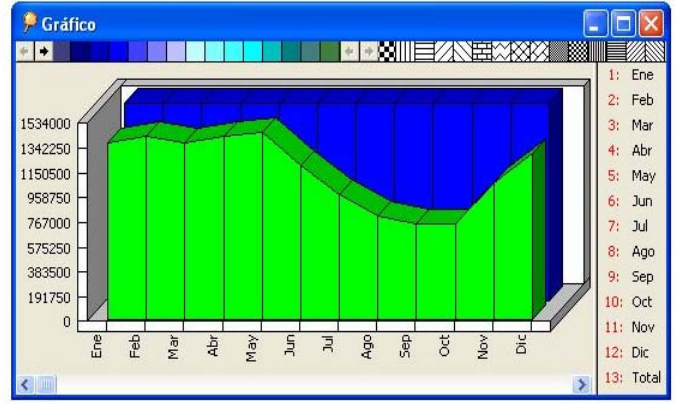
Hm3 Millones litros Porcentaje Hm3 (Relación Autonomías) Millones litros (Relación Autonomías) Porce

CHNE

Capacidad total vs capacidad actual por Autonomía y año-mes
(Volumen en Millones de litros)

Elementos de Página: Autonomía: ARAGÓN Año: 2005

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Total
Volumen Actual	1.372.000	1.427.000	1.373.000	1.429.000	1.453.000	1.200.000	978.000	806.000	750.000	750.000	1.067.000	1.292.000	1.453.000
Capacidad Total	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000	1.534.000

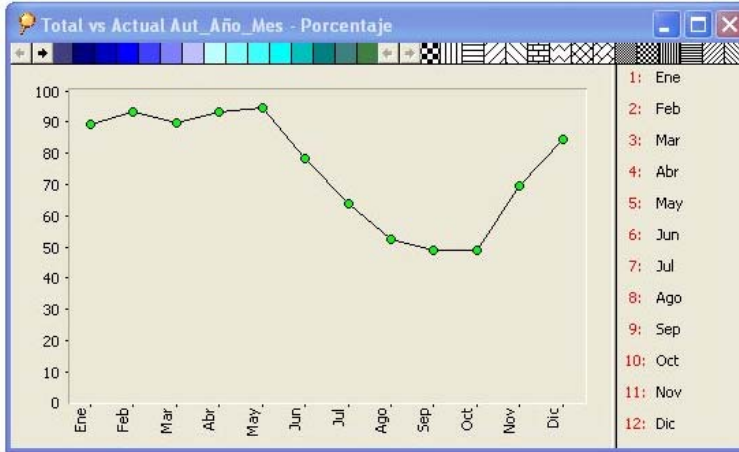


CHNE

Capacidad total vs capacidad actual por Autonomía y año-mes (Porcentaje)

Elementos de Página: Autonomía: ARAGÓN Año: 2005

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
%	89,44	93,02	89,50	93,16	94,72	78,23	63,75	52,54	48,89	48,89	69,56	84,22

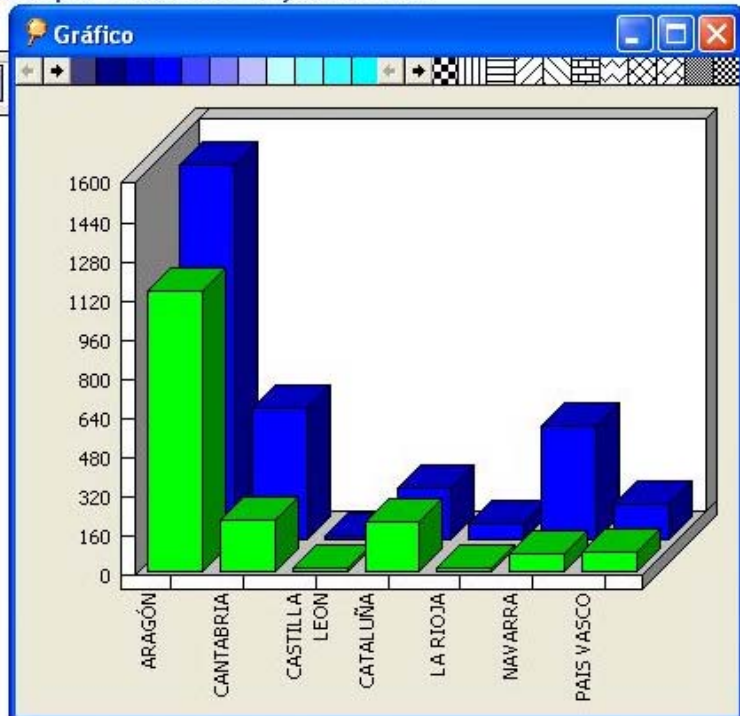


CHNE

Capacidad total vs capacidad actual por Autonomía y año-mes (Volumen en Hectómetros cúbicos)

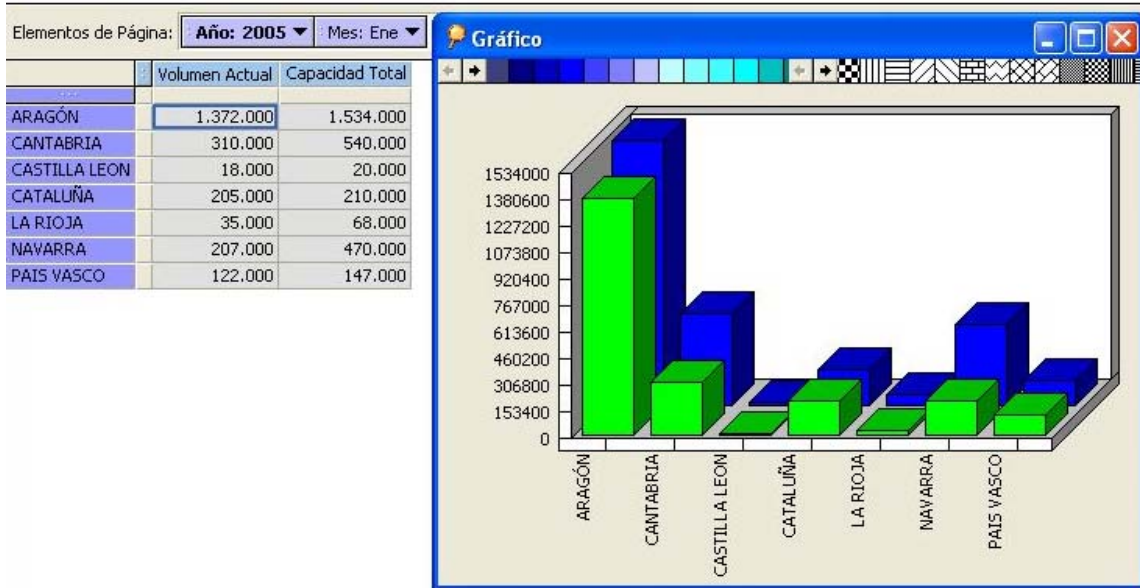
Elementos de Página: Año: 2002 Mes: Ene

	Volumen Actual	Capacidad Total
ARAGÓN	1.144	1.534
CANTABRIA	210	540
CASTILLA LEON	18	20
CATALUÑA	204	210
LA RIOJA	18	68
NAVARRA	73	470
PAIS VASCO	81	147



CHNE

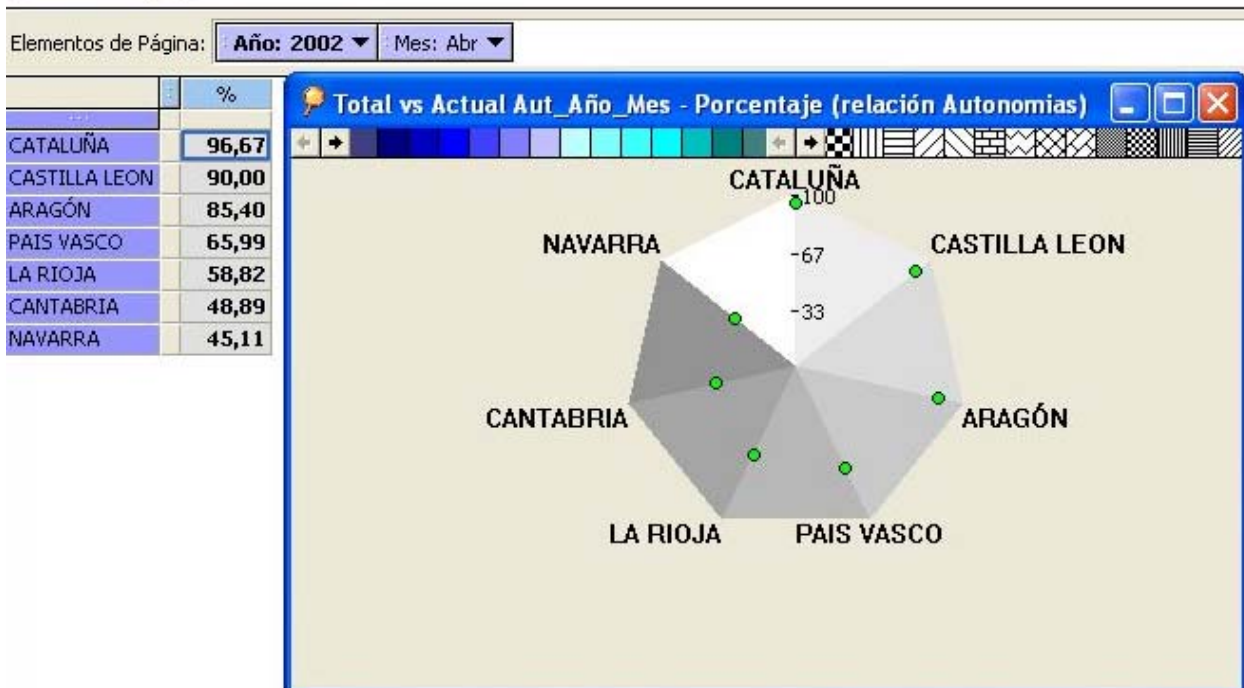
Capacidad total vs capacidad actual por Autonomía y año-mes
(Volumen en Millones de litros)



Hm3 Millones litros Porcentaje Hm3 (Relación Autonomías) Millones litros (Relación Autonomías)

CHNE

Capacidad total vs capacidad actual por Autonomía y año-mes
(Porcentaje)

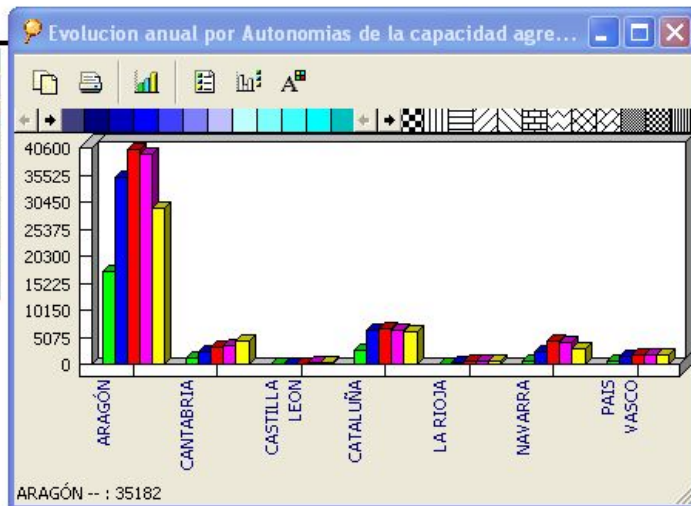


6.4 Evolución por CCAA de la capacidad agregada de sus embalses

CHNE

Evolución anual de la capacidad agregada en por Autonomías (en hectómetros cúbicos)

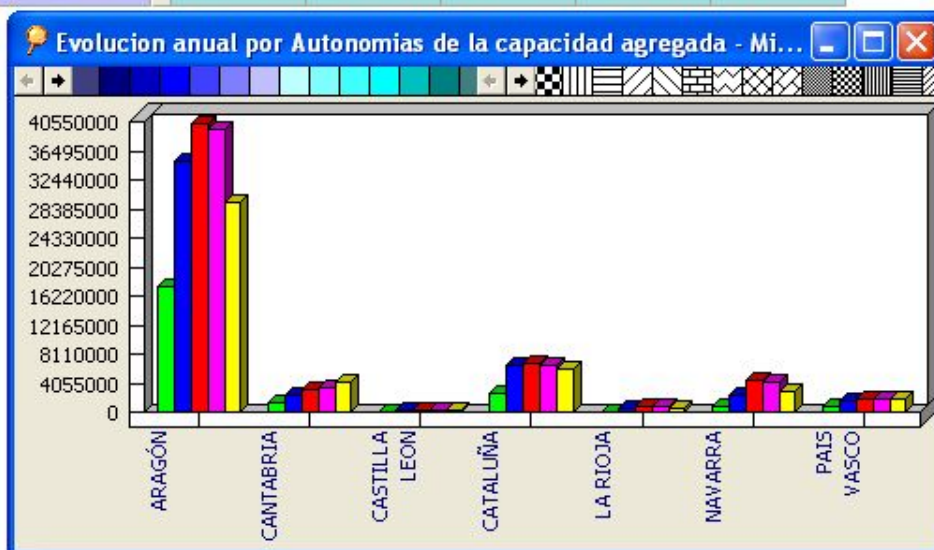
	Hm3 de Agua embalsada				
	2001	2002	2003	2004	2005
ARAGÓN	17.615	35.182	40.546	39.714	29.449
CANTABRIA	1.348	2.381	3.367	3.502	4.425
CASTILLA LEON	141	276	276	301	299
CATALUÑA	2.762	6.612	6.808	6.574	6.148
LA RIOJA	172	567	828	861	716
NAVARRA	819	2.517	4.544	4.244	3.119
PAIS VASCO	825	1.688	1.872	1.987	1.975



CHNE

Evolución anual de la capacidad agregada por Autonomías (en millones de litros)

	Millones de litros				
	2001	2002	2003	2004	2005
ARAGÓN	17.615.000	35.182.000	40.546.000	39.714.000	29.449.000
CANTABRIA	1.348.000	2.381.000	3.367.000	3.502.000	4.425.000
CASTILLA LEON	141.000	276.000	276.000	301.000	299.000
CATALUÑA	2.762.000	6.612.000	6.808.000	6.574.000	6.148.000
LA RIOJA	172.000	567.000	828.000	861.000	716.000
NAVARRA	819.000	2.517.000	4.544.000	4.244.000	3.119.000
PAIS VASCO	825.000	1.688.000	1.872.000	1.987.000	1.975.000



6.5 Comparación por embalses de cada CCAA entre un año y el anterior

CHNE

Evolución de la capacidad media por Autonomía en comparación con Año anterior (en Hectómetros cúbicos)

Elementos de Página: Año : 2003

	Volumen Año Anterior	Volumen Año Actual	Incremento %
ARAGÓN	91,6	105,6	15,27
CANTABRIA	199,9	280,6	40,35
CASTILLA LEON	11,8	11,5	-2,47
CATALUÑA	42,3	43,6	3,14
LA RIOJA	23,5	34,5	46,55
NAVARRA	52,1	94,7	81,69
PAIS VASCO	47,1	52,0	10,31

CHNE

Evolución de la capacidad media por Autonomía en comparación con Año anterior (en Millones de litros)

Elementos de Página: Año : 2003

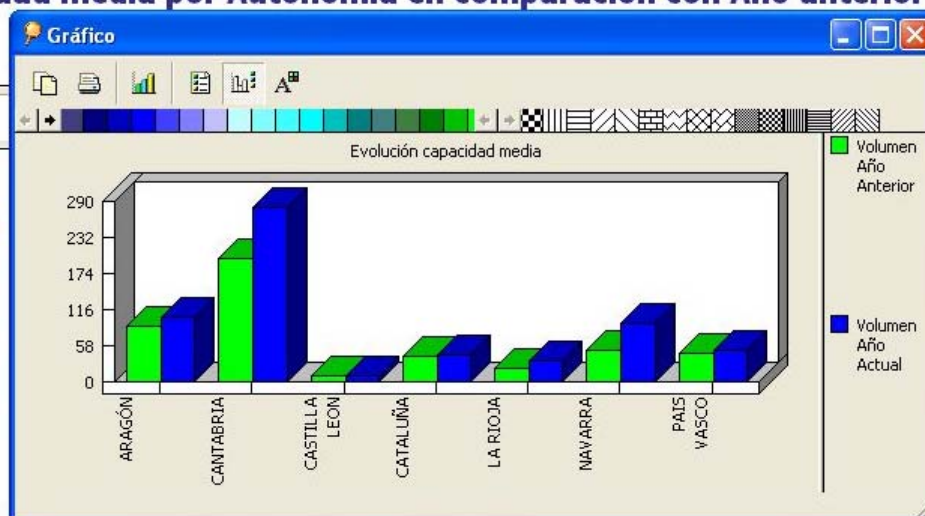
	Volumen Año Anterior	Volumen Año Actual	Incremento %
ARAGÓN	91.604,2	105.588,5	15,27
CANTABRIA	199.916,7	280.583,3	40,35
CASTILLA LEON	11.791,7	11.500,0	-2,47
CATALUÑA	42.314,1	43.641,0	3,14
LA RIOJA	23.541,7	34.500,0	46,55
NAVARRA	52.104,2	94.666,7	81,69
PAIS VASCO	47.138,9	52.000,0	10,31

CHNE

Evolución de la capacidad media por Autonomía en comparación con Año anterior (en Hectómetros cúbicos)

Elementos de Página: Año : 2003

	Volumen Año Anterior	Volumen Año Actual
ARAGÓN	92	106
CANTABRIA	200	281
CASTILLA LEON	12	12
CATALUÑA	42	44
LA RIOJA	24	35
NAVARRA	52	95
PAIS VASCO	47	52



6.6 Aportación de agua por río, embalse y año-mes

Aportación de agua por Río, Embalse, Año y Mes (en Hectómetros cúbicos)

Elementos de Página: Río: EBRO ▼ Año: 2001 ▼

	Mes	Jul	Ago	Sep	Oct	Nov	Dic	Media Anual
		Aportación	Aportación	Aportación	Aportación	Aportación	Aportación	
Embalse								
EBRO		-23,90	-22,40	-12,50	-4,60	5,10	-0,40	-9,78
MEQUINENZA		-19,10	-19,20	-8,00	-1,80	41,80	-3,50	-1,63
RIBARROJA		11,30	-2,00	3,20	2,90	-4,20	3,90	2,52
SOBRÓN		0,00	0,00	0,00	0,00	0,00	0,00	0,00

Aportación de agua por Río, Embalse, Año y Mes (en Millones de litros)

Elementos de Página: Río: EBRO ▼ Año: 2001 ▼

	Jul	Ago	Sep	Oct	Nov	Dic	Media Anual
	Aportación	Aportación	Aportación	Aportación	Aportación	Aportación	
EBRO	-23.900	-22.400	-12.500	-4.600	5.100	-400	-9.783
MEQUINENZA	-19.100	-19.200	-8.000	-1.800	41.800	-3.500	-1.633
RIBARROJA	11.300	-2.000	3.200	2.900	-4.200	3.900	2.517
SOBRÓN	0	0	0	0	0	0	0

6.7 Entrada vs Salida media agregada anual por CCAA

Entrada vs Salida media agregada anual por Autonomía (en Hectómetros cúbicos)

	Año	2001		2002		2003		2004		2005	
		Entrada Media	Salida Media	Entrada Media	Salida Media	Entrada Media	Salida Media	Entrada Media	Salida Media	Entrada Media	Salida Media
Autonomía											
ARAGÓN		4,69	5,07	5,78	5,38	9,15	9,28	9,49	10,19	5,98	6,76
CANTABRIA		2,17	11,95	5,07	5,11	4,49	5,73	4,88	5,47	7,28	6,51
CASTILLA LEON		0,01	0,02	0,08	0,10	0,02	0,02	0,04	0,00	0,25	0,25
CATALUÑA		8,44	9,01	12,83	12,45	23,63	23,80	19,64	19,85	15,13	15,05
LA RIOJA		0,67	1,65	2,25	1,26	2,32	2,62	1,93	2,06	1,55	1,38
NAVARRA		2,23	4,26	5,59	3,86	6,10	7,09	5,57	6,61	2,97	3,70
PAIS VASCO		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Entrada vs Salida media agregada anual por Autonomía (en Millones de litros)

	Año	2001		2002		2003		2004		2005	
		Entrada Media	Salida Media	Entrada Media	Salida Media	Entrada Media	Salida Media	Entrada Media	Salida Media	Entrada Media	Salida Media
Autonomía											
ARAGÓN		4.689,58	5.067,19	5.778,65	5.381,77	9.154,17	9.275,26	9.492,45	10.193,23	5.981,77	6.764,06
CANTABRIA		2.166,67	11.950,00	5.066,67	5.108,33	4.491,67	5.725,00	4.875,00	5.466,67	7.275,00	6.508,33
CASTILLA LEON		8,33	16,67	79,17	104,17	20,83	20,83	37,50	4,17	250,00	254,17
CATALUÑA		8.441,03	9.007,69	12.832,05	12.448,08	23.631,41	23.799,36	19.639,74	19.853,21	15.127,56	15.048,72
LA RIOJA		666,67	1.650,00	2.245,83	1.258,33	2.320,83	2.616,67	1.933,33	2.058,33	1.550,00	1.379,17
NAVARRA		2.225,00	4.258,33	5.591,67	3.858,33	6.102,08	7.091,67	5.566,67	6.606,25	2.966,67	3.702,08
PAIS VASCO		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

6.8 Aportación por río, embalse y mes

Aportación de agua por Río, Embalse, Año y Mes (en Hectómetros cúbicos)

Elementos de Página: Río: EBRO ▼ Año: 2001 ▼

	Mes	Jul	Ago	Sep	Oct	Nov	Dic	Media Anual
		Aportación	Aportación	Aportación	Aportación	Aportación	Aportación	
Embalse								
EBRO		-23,90	-22,40	-12,50	-4,60	5,10	-0,40	-9,78
MEQUINENZA		-19,10	-19,20	-8,00	-1,80	41,80	-3,50	-1,63
RIBARROJA		11,30	-2,00	3,20	2,90	-4,20	3,90	2,52
SOBRÓN		0,00	0,00	0,00	0,00	0,00	0,00	0,00

Aportación de agua por Río, Embalse, Año y Mes (en Millones de litros)

Elementos de Página: Río: EBRO ▼ Año: 2001 ▼

	Jul	Ago	Sep	Oct	Nov	Dic	Media Anual
	Aportación	Aportación	Aportación	Aportación	Aportación	Aportación	
EBRO	-23.900	-22.400	-12.500	-4.600	5.100	-400	-9.783
MEQUINENZA	-19.100	-19.200	-8.000	-1.800	41.800	-3.500	-1.633
RIBARROJA	11.300	-2.000	3.200	2.900	-4.200	3.900	2.517
SOBRÓN	0	0	0	0	0	0	0

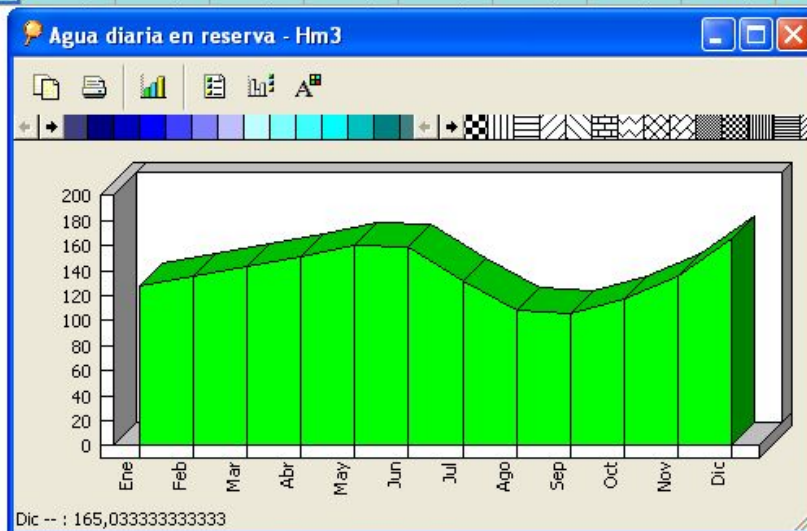
6.9 Agua diaria en reserva disponible por año y mes

Agua diaria en reserva por Año y Mes (en Hectómetros cúbicos)

Elementos de Página:

Año: 2002

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Reserva	127,40	135,20	143,00	150,93	160,93	158,53	131,47	109,00	106,07	117,27	135,93	165,03



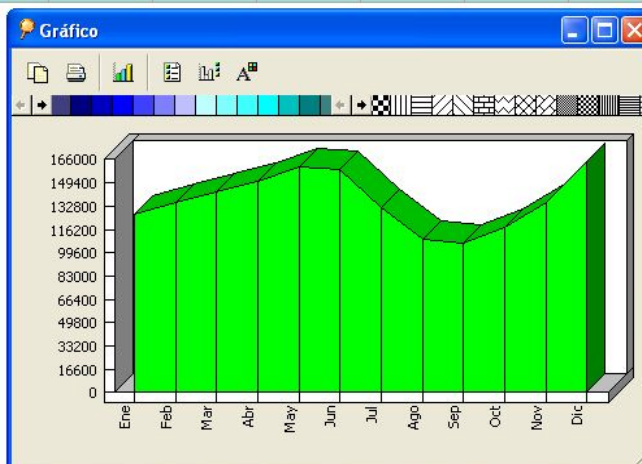
Hm3 Millones de litros

Agua diaria en reserva por Año y Mes (en Millones de litros)

Elementos de Página:

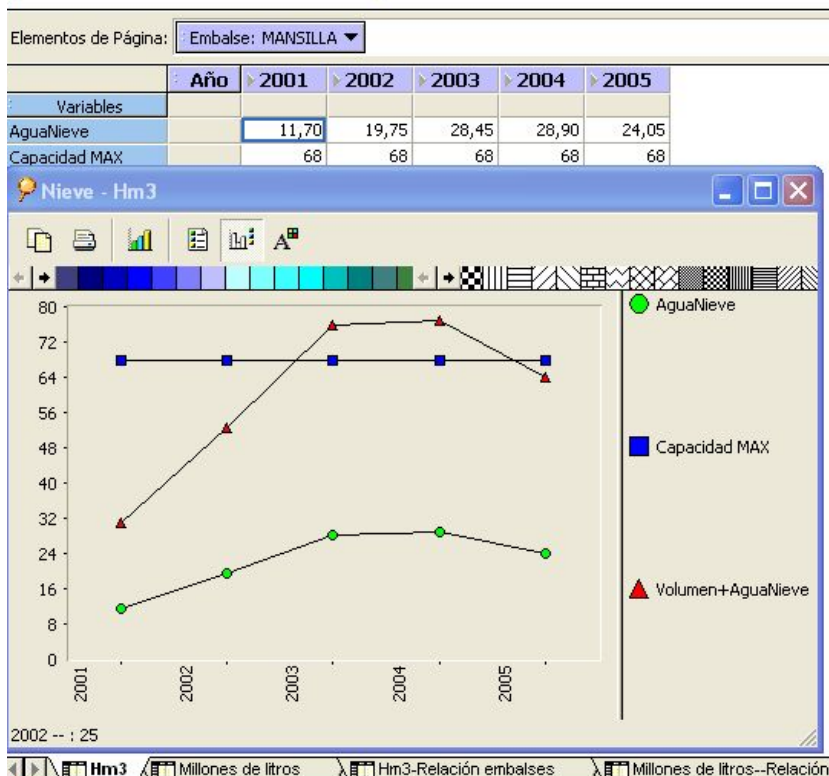
Año: 2002

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Reserva	127.400,00	135.200,00	143.000,00	150.933,33	160.933,33	158.533,33	131.466,67	109.000,00	106.066,67	117.266,67	135.933,33	165.033,33

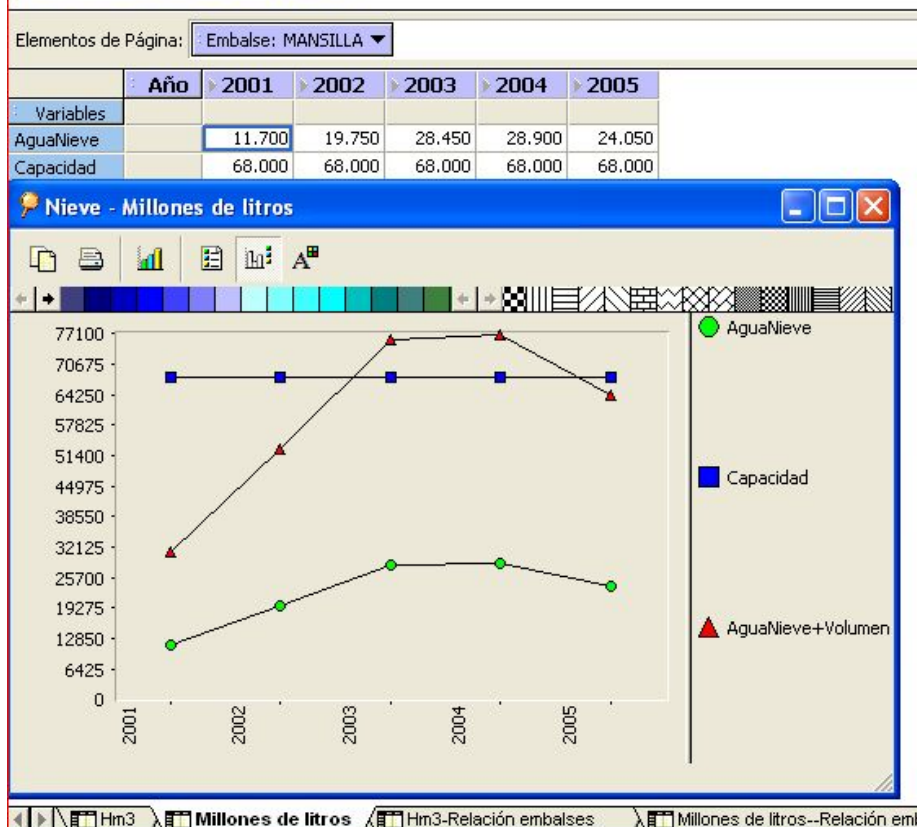


6.10 Agua de nieve

Previsión de agua de nieve por Año y Embalse (en Hectómetros cúbicos)



Previsión de agua de nieve por Año y Embalse (en Millones de litros)



**Previsión de agua de nieve por Año y Embalse
(en Hectómetros cúbicos)**

Embalse	Año	2001		2002		2003		2004		2005	
		Capacidad MAX	AguaNieve	Capacidad MAX	AguaNieve	Capacidad MAX	AguaNieve	Capacidad MAX	AguaNieve	Capacidad MAX	AguaNieve
ALBINA		6	0,70	6	0,80	6	0,75	6	0,80	6	0,72
ALCAÑIZ (ESTANCA)		7	1,30	7	1,33	7	1,33	7	1,23	7	1,17
ALLOZ		65	3,27	65	3,18	65	8,05	65	7,42	65	7,83
BACHIMAÑA (LAGO)		7	0,70	7	0,25	7	0,62	7	0,55	7	0,42
BARASONA		92	9,43	92	12,83	92	13,68	92	11,80	92	9,48
BASERCA		22	3,77	22	2,60	22	3,10	22	2,62	22	2,73
BRAMATUERO ALTO		5	0,90	5	0,10	5	0,30	5	0,37	5	0,00
BÚBAL		64	17,33	64	25,00	64	31,87	64	30,40	64	19,40
CALANDA		54	16,30	54	21,05	54	21,60	54	20,45	54	10,35
CAMARASA		113	8,53	113	12,75	113	15,67	113	15,90	113	12,05
CANELLES		688	101,30	688	86,35	688	101,67	688	107,50	688	56,52
CASPE		82	11,23	82	10,72	82	13,90	82	11,85	82	8,43
CAVALLERS		16	2,03	16	1,82	16	2,35	16	2,05	16	1,97
CERTESCÁNS		16	1,77	16	1,97	16	2,22	16	2,02	16	2,52
CIURANA		12	0,53	12	0,68	12	1,78	12	2,23	12	1,48
CUEVA FORADADA		29	0,43	29	0,75	29	3,00	29	4,27	29	3,43
EBRO		540	44,93	540	39,68	540	56,12	540	58,37	540	73,75
ESCALES		152	23,53	152	24,57	152	24,13	152	21,25	152	19,77
ESCARRA		5	0,37	5	0,50	5	0,75	5	0,27	5	0,40
EUGUI		21	2,53	21	3,32	21	3,23	21	3,15	21	3,18
GONZALEZ LACASA		33	1,83	33	2,87	33	4,32	33	4,72	33	3,92
GRADO (EL)		399	73,97	399	66,98	399	75,92	399	70,87	399	55,85
GUIAMETS		10	0,43	10	0,73	10	1,38	10	1,68	10	1,07
IP		5	0,63	5	0,83	5	0,50	5	0,70	5	0,60
IRABIA		14	0,27	14	1,78	14	1,70	14	1,90	14	1,93
LLANUZA		17	2,17	17	2,73	17	2,72	17	2,40	17	2,23

Hay desbordamiento cuando: (Agua embalsada + Agua de nieve) > Capacidad del embalse.

Casos de desbordamiento de embalses

Embalse	Año	2001	2002	2003	2004	2005
		Desbordamiento	Desbordamiento	Desbordamiento	Desbordamiento	Desbordamiento
LLAuset						
MAIDEVERA						
MANSILLA						
MAR (LAGO)						
MEDIANO						
MEQUINENZA						
MONEVA						
MONTEAGUDO						
NEGRO						
OLIANA						
PENA						
PEÑA (LA)						
RESPOMUSO						
RIBARROJA						
SABURO (LAGO)						
SAN BARTOLOMÉ						
SANTA ANA						
SANTOLEA						
S. MARÍA DE BELSUÉ						
SOBRÓN						
SOTONERA						
TERRADETS						
TORCAS (LAS)						
TORT (LAGO)						
TRANQUERA (LA)						
TREMP O TALÁRN						
ULLIVARRI						

La fórmula que se ha estimado para la predicción de sequías es:

(Agua embalsada + Agua de nieve) sea inferior al 50% de la capacidad del embalse.

Periodos de Sequía

	Año	2001	2002	2003	2004	2005
		Sequía	Sequía	Sequía	Sequía	Sequía
Embalse						
LLAUSSET						
MAIDEVERA						
MANSILLA						
MAR (LAGO)						
MEDIANO						
MEQUINENZA						
MONEVA						
MONTEAGUDO						
NEGRO						
OLIANA						
PENA						
PEÑA (LA)						
RESPOMUSO						
RIBARROJA						
SABURO (LAGO)						
SAN BARTOLOMÉ						
SANTA ANA						
SANTOLEA						
S. MARÍA DE BELSUÉ						
SOBRÓN						
SOTONERA						
TERRADETS						
TORCAS (LAS)						
TORT (LAGO)						
TRANQUERA (LA)						
TREMP O TALÁRN						
ULLIVARRI						

7 Conclusiones

En mi opinión el proyecto es de buena calidad y se ajusta a las especificaciones pedidas por la CHNE.

El sistema permite un acceso óptimo a la información y sirve de ayuda a la toma de decisiones y obtener fácilmente mejores modelos predictivos.

Este proyecto me ha servido para adentrarme aún más en el mundo del Data Warehouse, al realizar un proyecto completo desde el inicio hasta la puesta en marcha del mismo.

8 Glosario

Almacén de datos: Colección de datos no volátiles, orientados al usuario, integrados, variables en el tiempo, principalmente usados como ayuda a la toma de decisiones.

Atributo: Es un tipo básico de información descriptiva de una dimensión.

Hecho: Actividad de la organización que es objeto de análisis.

Data Warehouse: Almacén de datos.

Dimensión: Una estructura que representa una de las caras de un cubo.

Drill: Operación de agregación que se realiza sobre una dimensión.

Roll: Operación de desagregación que se realiza sobre una dimensión.

Medida: Es la información relevante sobre el hecho. Uno de los atributos de la celda del cubo.

Gránulo: Nivel de detalle al que se desea almacenar la información sobre la actividad a modelar.

Data-Mart: Subconjunto de información del DataWarehouse específica y orientada a una línea de negocio de la actividad.

Modelo dimensional: Presenta una visión sencilla y orientada al negocio de datos organizado mediante diferentes perspectivas o “dimensiones” de análisis ordenadas jerárquicamente.

OLAP: Del inglés On-Line Analytical Processing (procesamiento analítico en línea). Permite agilizar la consulta de grandes cantidades de datos.

Tabla de dimensiones: Tablas que contienen los valores por los que se quiere analizar la información. Se emplean para organizar los datos a modo de ejes.

Tabla de hechos: Tablas que contienen los valores a consultar.

Jerarquía: Criterio de agrupación de elementos de una dimensión

Estrella: Esquema de organización clásico del modelo multidimensional.

ETL: Proceso de extracción, transformación y carga (Extraction, Transformation, and Loading).

9 Bibliografía.

9.1 Publicaciones

Data Modeling Techniques for Data Warehousing. Chuck Ballard y otros.

Business Intelligence & Warehouse. Guillermina Moreno.

Building the Data Warehouse (3rd Edition). W. H. Inmon.

The Data Warehouse Toolkit (2 Edition). Ralph Kimball & Margy Ross.

The Data Warehouse ETL Toolkit. Ralph Kimball & Joe Caserta.

An overview of data warehousing and OLAP technology. Chaudhuri S., Dayal U.

9.2 Webs

<http://www.redbooks.ibm.com>

<http://www.dwreview.com>

<http://www.oracle.com/technetwork/developer-tools/warehouse/documentation/index.html>

<http://www.oracle.com/technetwork/indexes/documentation/index.html>

