

## Proyecto Big Data de salud Comunidad Europea



Diseño e implementación de una base de datos para recogida y análisis de datos de actividad física provenientes de dispositivos 'wearables'

Autor: Mari Cruz López Ramírez

Consultor: Manel Rella Ruiz

Enginyeria Tècnica Informàtica de Gestió (ETIG)

TFC · Bases de datos relacionales

# CONTENIDO

## Pre-análisis

- Descripción del proyecto
- Metodología
- Planificación
- Temporización
- Recursos
- Análisis de riesgos
- Coste

## Análisis

- Requisitos funcionales
  - Flujo de datos de contexto
  - Lista de requisitos
- Requisitos no funcionales
  - Despliegue del sistema
  - Lista de requisitos
- Casos de uso

## Diseño

- Esquema conceptual
- Esquema lógico

## Implementación y control

- Diseño físico
  - Tablas y vistas
    - Índices
    - Secuencias
    - Disparadores
  - Paquetes
    - Procedimientos
    - Funciones
- Optimización
- Pruebas del producto

## Conclusiones

# Pre-análisis – Descripción del proyecto

---

Este proyecto consiste en diseñar e implementar un sistema de información alojado en una base de datos Oracle, con el fin de dar respuesta al proyecto Big Data, cuyo objetivo es cruzar los datos de salud y los datos de actividad física de los ciudadanos europeos.

El resultado es una base de datos que permite almacenar:

- ❖ La información de la actividad física proveniente de los dispositivos móviles.
- ❖ Los datos de salud de los ciudadanos, provenientes de los servicios de salud.
- ❖ Información estadística consultable por los ciudadanos y/o por los especialistas de la salud.

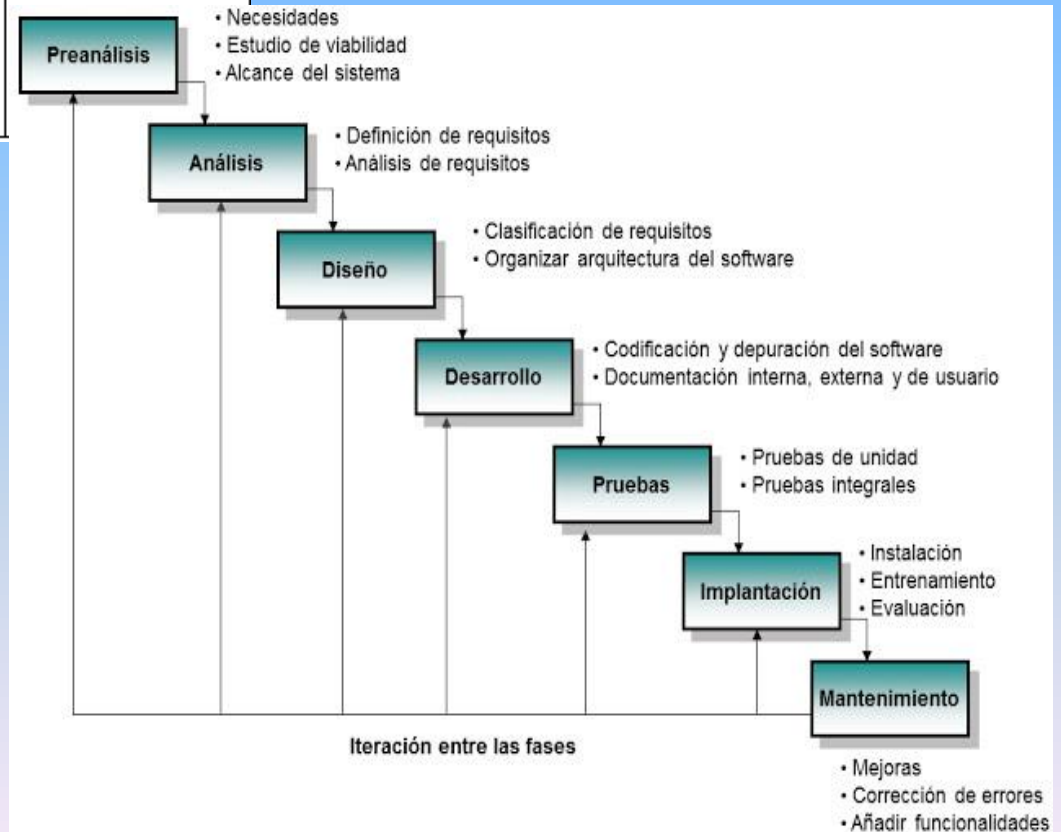
También permite y facilita que la aplicación de alto nivel obtenga la información necesaria del almacén de datos para:

- ❖ Realizar análisis globales del nivel de actividad de la ciudadanía relacionado con su salud y su país.
- ❖ Informar a cada usuario de su nivel de actividad física con recomendaciones para mejorar su salud.
- ❖ Que los médicos conozcan el nivel de actividad física de cada ciudadano.

# Pre-análisis - Metodología

La metodología empleada es el ciclo de vida en cascada, donde las actividades son agrupadas en fases.

FASE 1	FASE 2	FASE 3	FASE 4
PAC1	PAC2	PAC3	Entrega final
<ul style="list-style-type: none"> <li>o Pre_análisis                             <ul style="list-style-type: none"> <li>• Necesidades.</li> <li>• Viabilidad.</li> <li>• Alcance.</li> <li>• Plan de trabajo.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>o Análisis                             <ul style="list-style-type: none"> <li>• Definición de requisitos.</li> <li>• Análisis de requisitos.</li> </ul> </li> <li>o Diseño                             <ul style="list-style-type: none"> <li>• Clasificación requisitos.</li> <li>• Organizar arquitectura.</li> </ul> </li> <li>o Desarrollo objetos BD                             <ul style="list-style-type: none"> <li>• Codificación.</li> <li>• Depuración.</li> <li>• Documentación.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>o Desarrollo consultas                             <ul style="list-style-type: none"> <li>• Codificación.</li> <li>• Depuración.</li> <li>• Documentación.</li> </ul> </li> <li>o Pruebas                             <ul style="list-style-type: none"> <li>• Pruebas unidad.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>o Pruebas                             <ul style="list-style-type: none"> <li>• Pruebas integrales.</li> </ul> </li> <li>o Entrega y evaluación                             <ul style="list-style-type: none"> <li>• Documentación.</li> <li>• Scripts instalación.</li> <li>• Evaluación</li> </ul> </li> </ul>



# Pre-análisis - Planificación

## FASE 1 →

Nombre	Fecha de inicio	Fecha de fin	Duración
• Tarea 1 - Recopilación material	17/09/14	18/09/14	2
• Tarea 2 - Instalación software	19/09/14	19/09/14	1
• Tarea 3 - Lectura documentación	20/09/14	25/09/14	6
• Tarea 4 - Análisis enunciado	26/09/14	28/09/14	3
• Tarea 5 - Identificación tareas	29/09/14	2/10/14	4
• Tarea 6 - Planificación proyecto	3/10/14	3/10/14	1
• Tarea 7 - Estimación de riesgos	4/10/14	4/10/14	1
• Tarea 8 - Estimación de costes	5/10/14	5/10/14	1
• Revisión y entrega	6/10/14	6/10/14	1

## FASE 2 →

Nombre	Fecha de inicio	Fecha de fin	Duración
• Tarea 9 - Análisis de requerimientos	7/10/14	11/10/14	5
• Tarea 10 - Diseño conceptual	12/10/14	16/10/14	5
• Tarea 11 - Diseño lógico de la BD	17/10/14	19/10/14	3
♀ • Tarea 12 - Diseño físico (objetos de BD)	20/10/14	26/10/14	7
• - Crear tablas y secuencias	20/10/14	20/10/14	1
♀ • - Crear disparadores	21/10/14	26/10/14	6
• * Disparadores que alimentan secuencias y columnas de auditoria	21/10/14	21/10/14	1
• * Disp. alimentan módulo estadístico	22/10/14	26/10/14	5
• Tarea 13 - Diseño físico (carga de datos)	27/10/14	28/10/14	2
• Tarea 14 - Diseño físico (métodos de acceso a BD - Actualización)	29/10/14	16/11/14	19
• - Implementación actualización ciudadanos			
• - Implementación actualización centros de salud			
• - Implementación actualización actividad diaria			
• - Implementación actualización historico salud			
• Revisión y entrega	17/11/14	17/11/14	1

## FASE 3 →

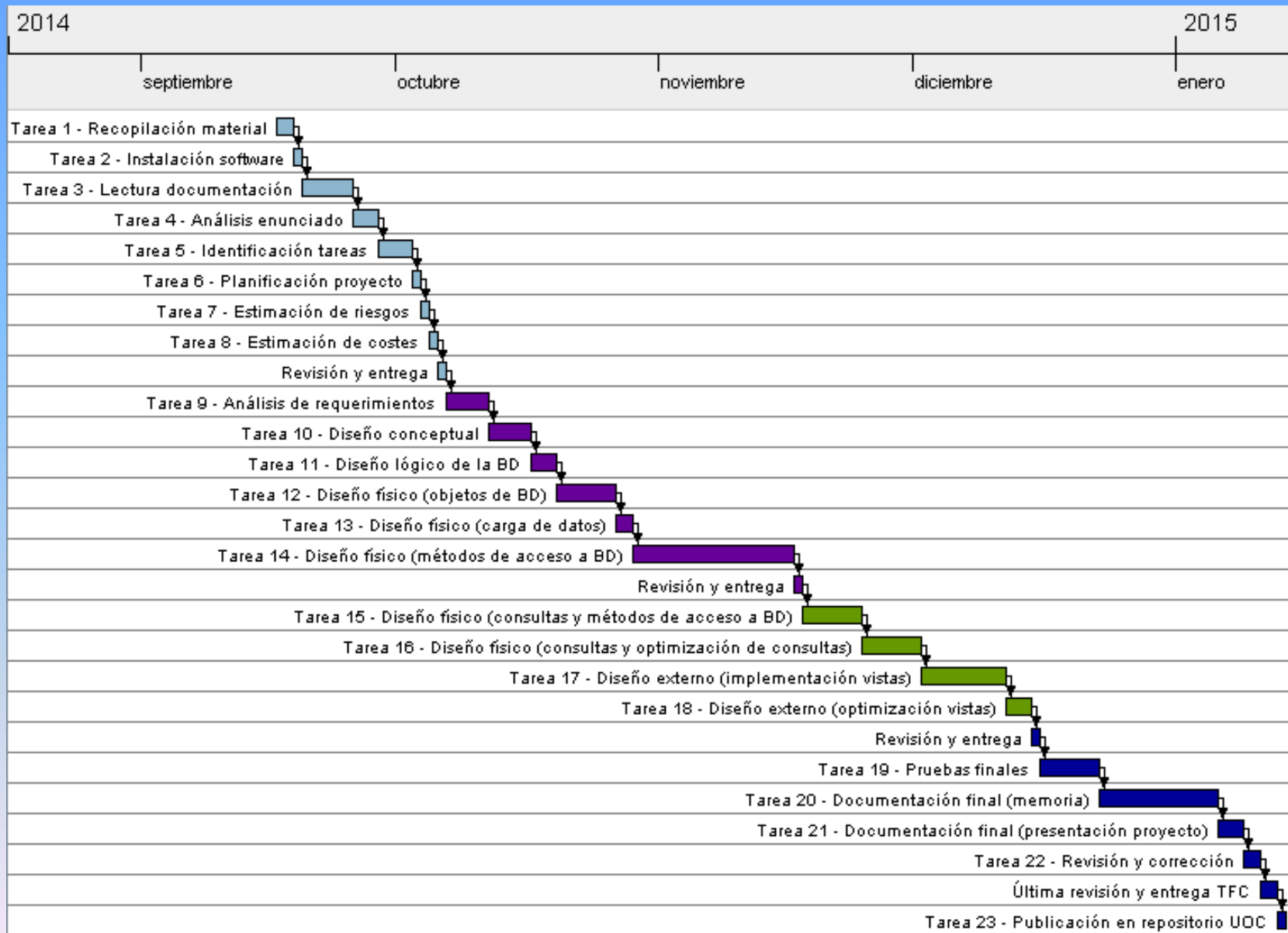
Nombre	Fecha de inicio	Fecha de fin	Duración
• Tarea 15 - Diseño físico (consultas y métodos de acceso a BD)	18/11/14	24/11/14	7
• Tarea 16 - Diseño físico (consultas y optimización de consultas)	25/11/14	1/12/14	7
• Tarea 17 - Diseño externo (implementación vistas)	2/12/14	11/12/14	10
• Tarea 18 - Diseño externo (optimización vistas)	12/12/14	14/12/14	3
• Revisión y entrega	15/12/14	15/12/14	1

## FASE FINAL →

Nombre	Fecha de inicio	Fecha de fin	Duración
• Tarea 19 - Pruebas finales	16/12/14	22/12/14	7
• Tarea 20 - Documentación final (memoria)	23/12/14	5/01/15	14
• Tarea 21 - Documentación final (presentación proyecto)	6/01/15	8/01/15	3
• Tarea 22 - Revisión y corrección	9/01/15	10/01/15	2
• Última revisión y entrega TFC	11/01/15	12/01/15	2
• Tarea 23 - Publicación en repositorio UOC	15/01/15	15/01/15	1

Se planifica realizar tres entregas parciales y una entrega final, donde se entrega el producto obtenido.

# Pre-análisis - Temporización



# Pre-análisis - Recursos

Los recursos de los que se dispone para el desarrollo del proyecto son:

## Software

Software	Función
Oracle Express v11g	Gestor de base de datos
Oracle SQL Developer 4.0.3	Desarrollo software
MagicDraw Personal Edition	Diseño diagrama de clases
Ganttproject 2.6.6	Planificación proyecto
Paquete Microsoft Office 2007	Elaboración documentación

## Recursos humanos

Persona	Rol
Mari Cruz López Ramírez	Desarrollador
Manel Rella Ruiz	Cliente

## Hardware

Componente	Características
Procesador	Intel® Core™2 Duo CPU T7300 @ 2.00GHz 2.00 GHz
Memoria RAM instalada	3 GB
Tipo de sistema	Sistema operativo de 64 bits

## Documentación

Se emplea la documentación de apoyo :

- Material de las asignaturas Base de datos I, Base de datos II y Sistemas de gestión de *BD*.
- Material de la asignatura Ingeniería del Software.
- Material de las asignaturas Gestión de Proyectos Informáticos y Técnicas de desarrollo de software.
- Material de *Presentación de documentos y elaboración de presentaciones*
- Guías y documentación sobre *PL/SQL y base de datos Oracle*

# Pre-análisis – Análisis de riesgos

Descripción de un pequeño plan de contingencias para hacer frente a posibles imprevistos

Descripción	Impacto	Preventiva
Averías: hardware o software	Pérdida del entorno de trabajo	Ordenador de sobremesa de reserva preparado con el mismo entorno que el ordenador habitual.
	Perdida de información	Copia de seguridad periódica del trabajo realizado.
Interrupción de servicio internet	Perdida acceso a parte de fuente de información y perdida de comunicación con la UOC.	Posibilidad de utilizar el servicio de internet de familiares y amigos.
Enfermedad o viajes imprevistos	Disminución de horas de dedicación en la planificación	Se dejan dos días a la semana sin asignación de horas para emplear como extra. Si es necesario se han reservado días de vacaciones en el trabajo.
	Provoca cambio en la planificación	Se pacta con el cliente el cambio de planificación.
Detectar un error en la planificación o cambio de requisitos	Disminución de horas de dedicación en la planificación	Para evitar cambios en la planificación, se dejan dos días a la semana sin asignación de horas para emplear como extra.
	Provoca cambio en la planificación y valoración de coste inicial	Se da la posibilidad al cliente de aumentar coste o eliminar aquellos requisitos que tienen menos prioridad y que se pueden dejar para una segunda versión del proyecto.



# Pre-análisis - Coste

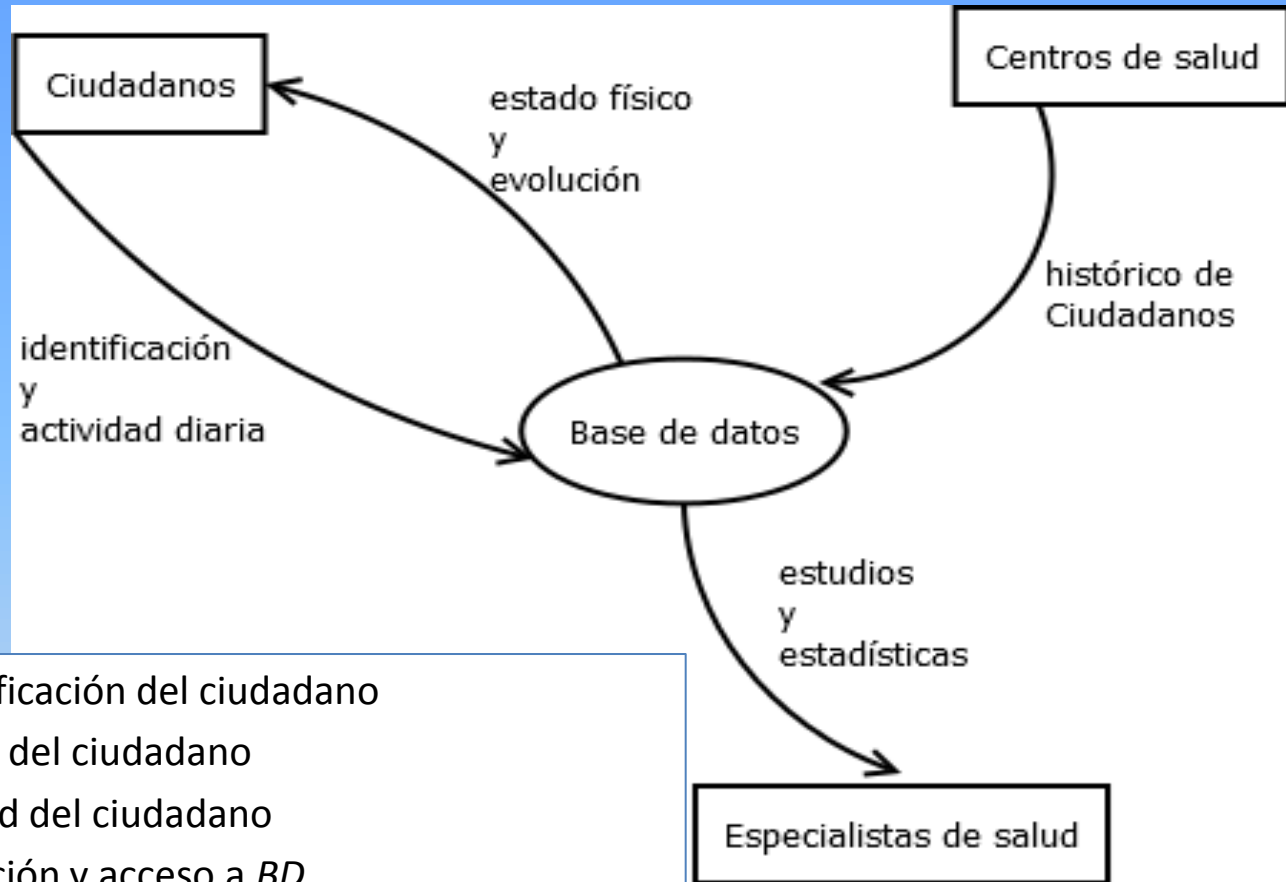
## Valoración económica aproximada

Categoría profesional	Responsabilidades	Precio <sup>31</sup> hora	Nº horas	Coste
Jefe de proyectos	<ul style="list-style-type: none"><li>• Mantiene el contacto con el cliente.</li><li>• Realiza el diseño funcional y técnico del sistema.</li><li>• Documenta el análisis.</li><li>• Define, planifica y realiza las pruebas del sistema y las aplicaciones.</li></ul>	37€	63	2.331€
Analista programador	<ul style="list-style-type: none"><li>• Análisis y la toma de requisitos para su posterior programación.</li><li>• Planifica las tareas de programación a corto plazo.</li><li>• Desarrollo de los elementos de la aplicación.</li></ul>	33€	93	3.069€
			156	5.400€
			16% IVA	864€
			<b>Total</b>	<b>6264€</b>

El número de horas a facturar se obtiene de la planificación propuesta para el desarrollo del proyecto, con los días y horas semanales que el desarrollador puede dedicar al proyecto.

# Análisis – Requisitos funcionales

- **DFD de contexto** >

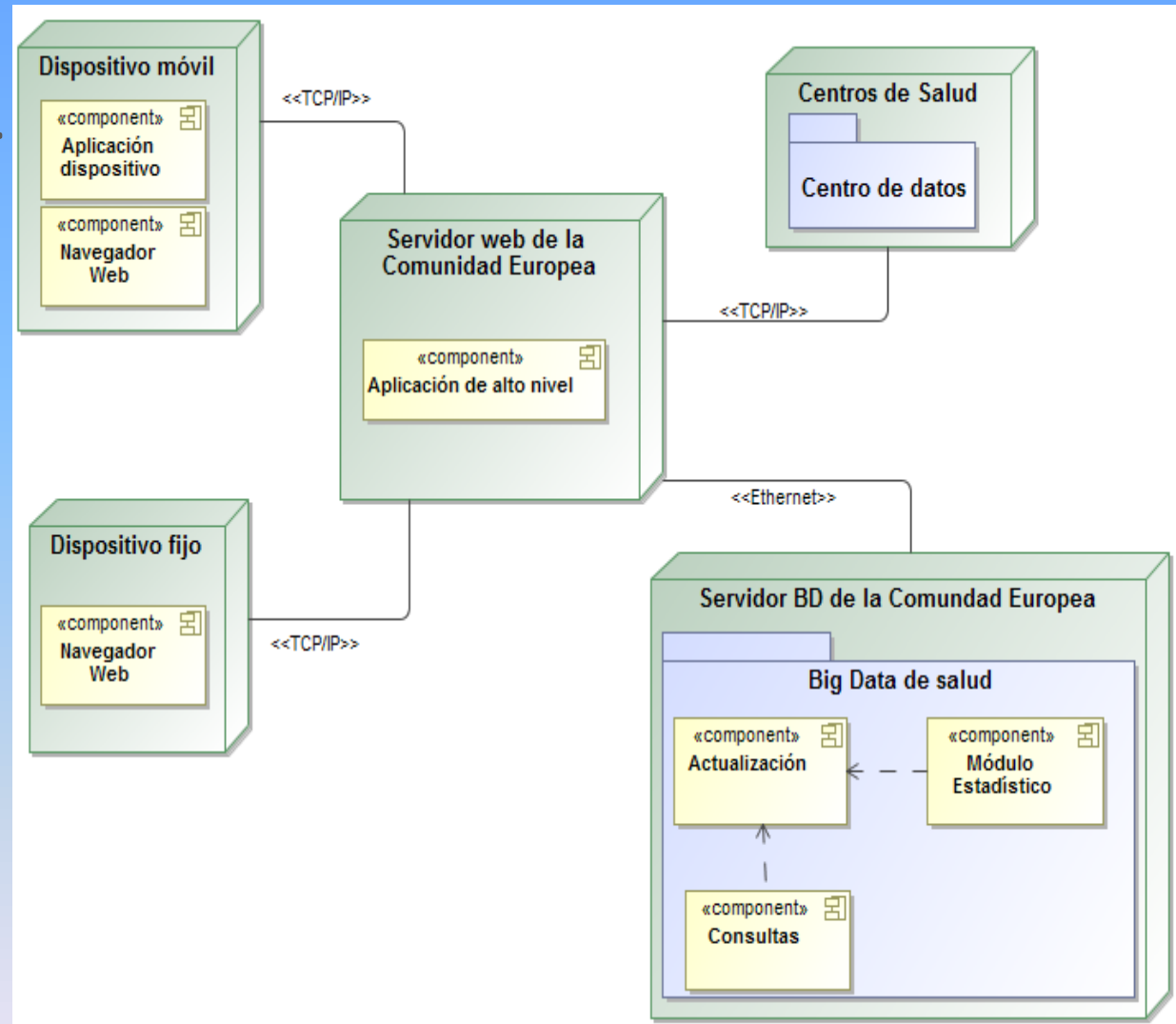


- **Lista de requisitos**

1. Registro de datos de identificación del ciudadano
2. Registro de actividad diaria del ciudadano
3. Registro de historia de salud del ciudadano
4. Funcionalidad de actualización y acceso a *BD*
5. Proporcionar información del estado de salud de ciudadanos
6. Proporcionar información para seguimiento del ciudadano
7. Proporcionar información para estudio de los ciudadanos
8. Módulo estadístico

# Análisis – Requisitos no funcionales

- Diagrama de despliegue del sistema >

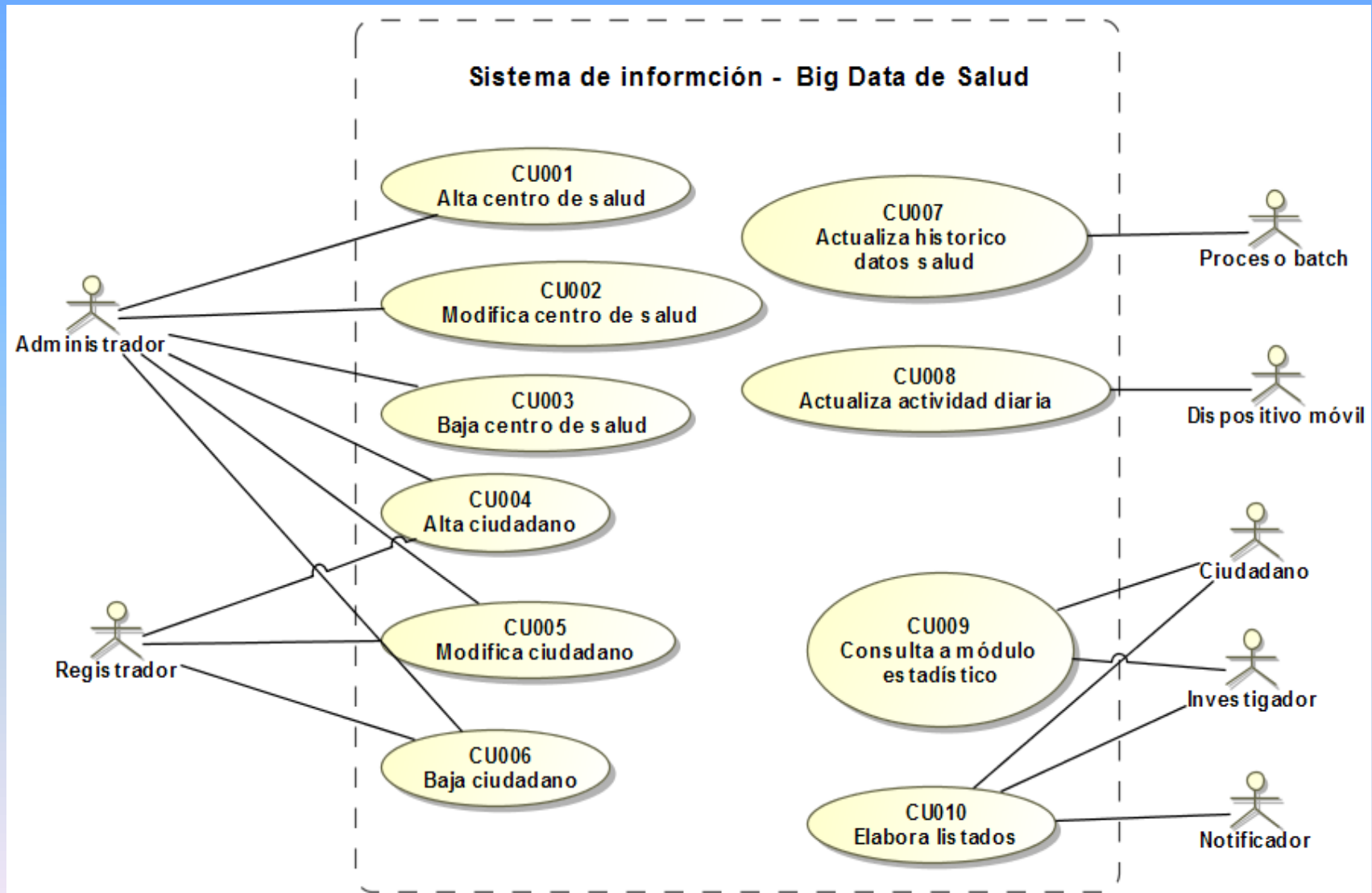


- Lista de requisitos

1. Requisitos de interface
2. Accesibilidad
3. Mantenibilidad
4. Integración
5. Reusabilidad
6. Escalabilidad
7. Rendimiento

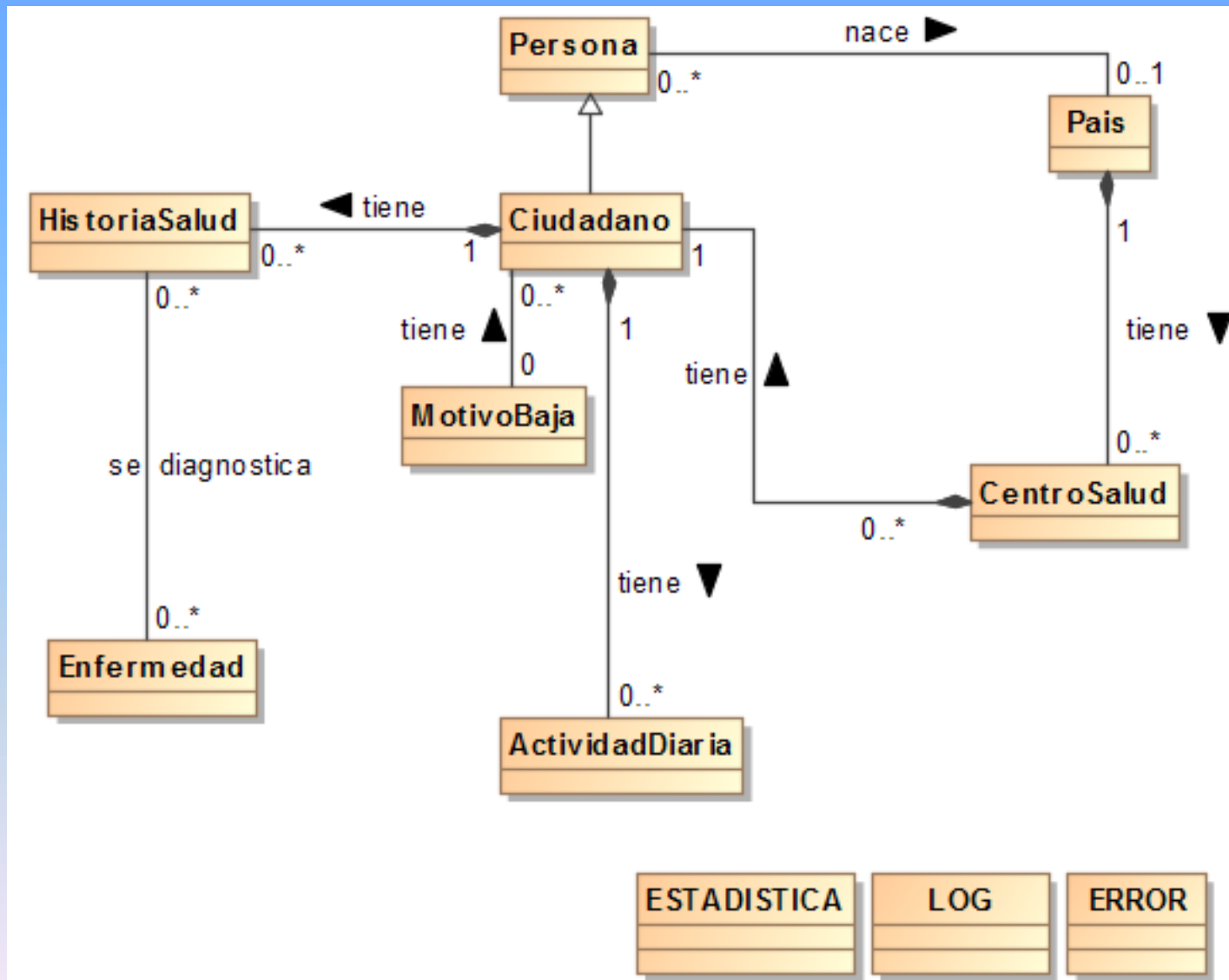
# Análisis – Casos de uso

Escenario de interacción usuario ↔ sistema



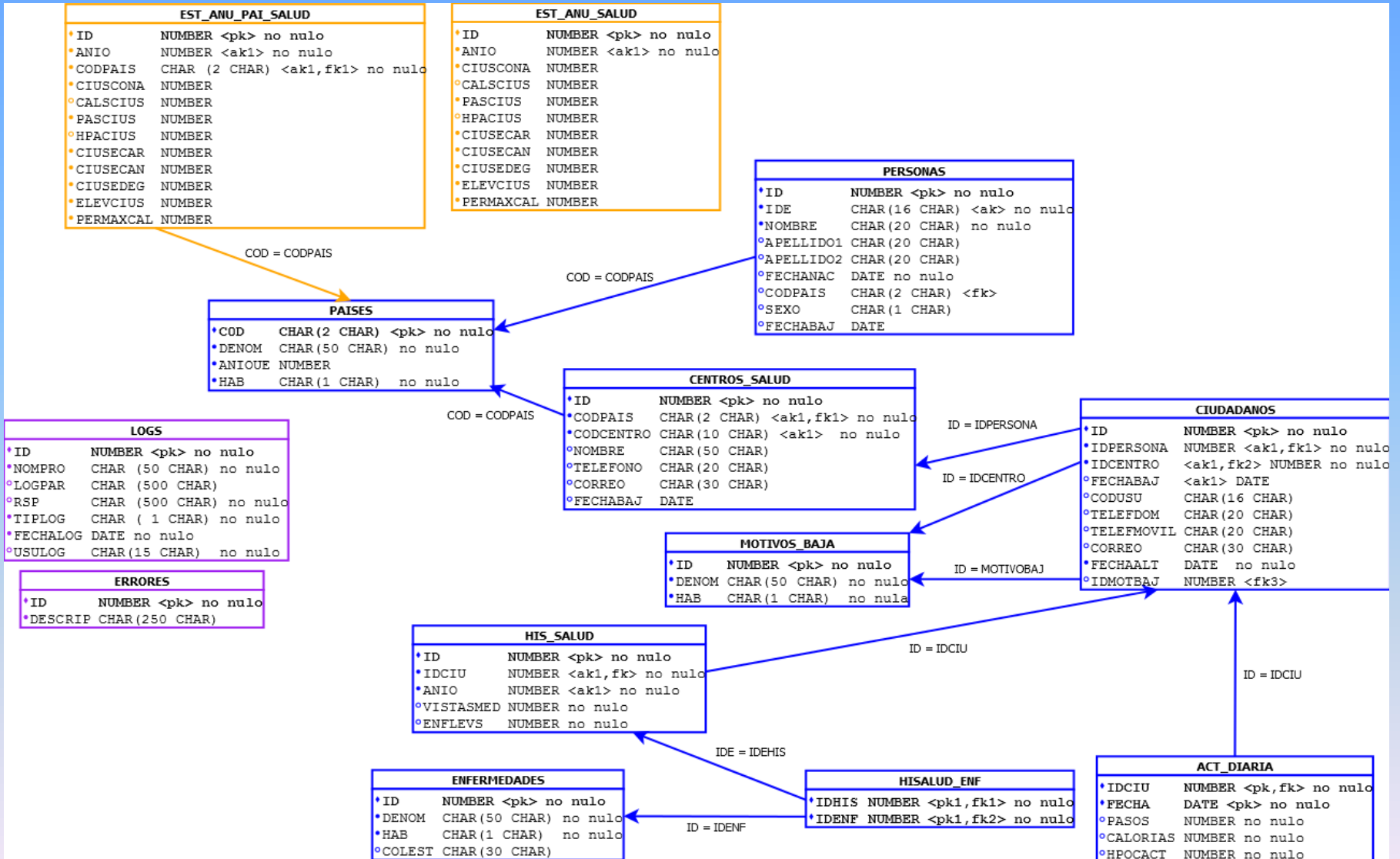
# Diseño – Esquema conceptual

Diagrama entidad/relación (notación UML)



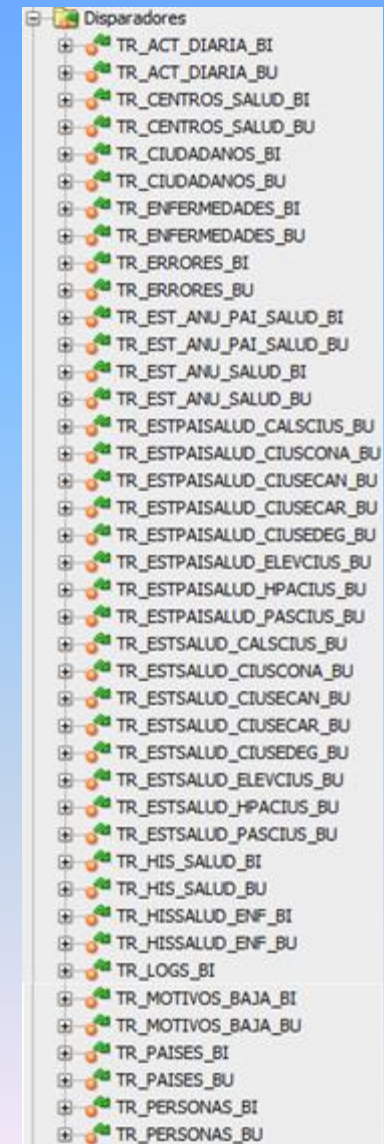
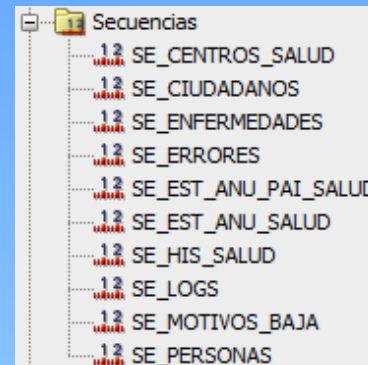
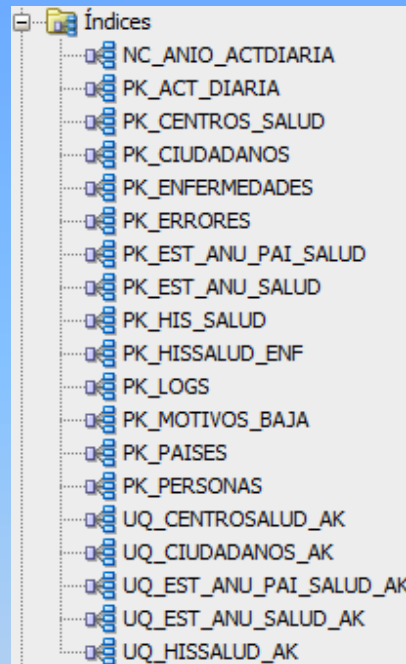
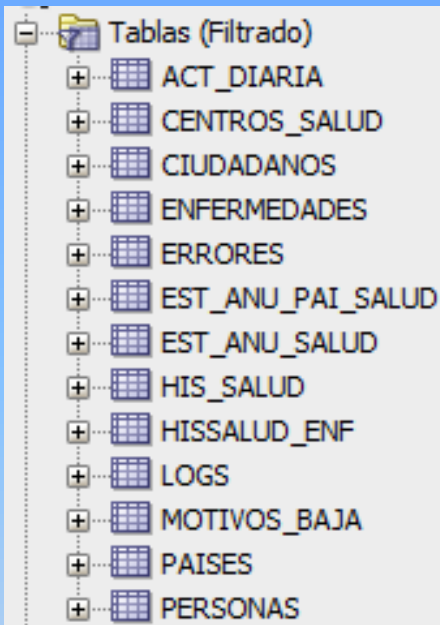
# Diseño – Esquema lógico

## Adaptación del esquema conceptual a la *BD* relacional

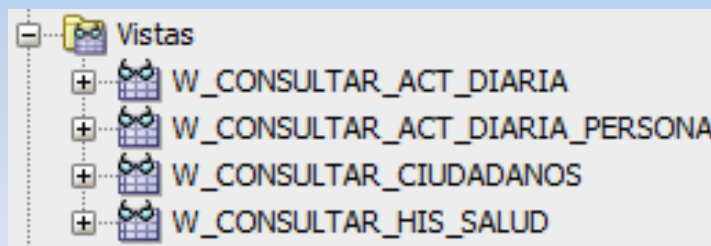


# Implementación – Diseño físico · Tablas y vistas

Tablas que se implementan con sus correspondientes índices, secuencias y disparadores



Vistas que se implementan









# Implementación – Optimización

Para optimizar el acceso a *BD*, se descomponen las consultas más complejas y se analizan aplicando los siguientes criterios:

- ❖ Cuidar el orden de las cláusulas *where* en consultas con varias tablas.
- ❖ Intentar filtrar por valores exactos.
- ❖ Intentar filtrar por columnas indexadas.
- ❖ Valorar si es conveniente añadir algún índice para optimizar los accesos con filtros.
- ❖ Utilizar columnas de longitud corta o parte de columnas para crear índices.

Se aplica la sentencia *EXPLAIN PLAN* para detectar si es necesario crear algún índice que reduzca el coste de los accesos a la *BD*

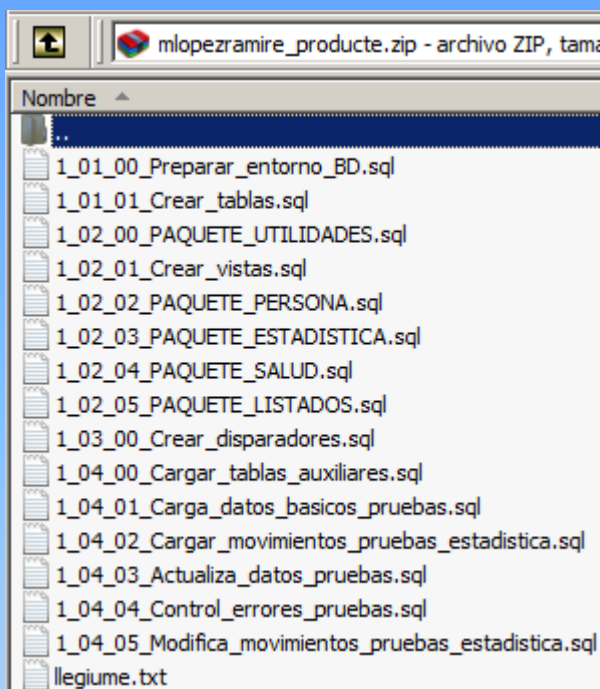
Plan de ejecución antes de crear el índice

OPERATION	OBJECT_NAME	CARDINALITY	COST
SELECT STATEMENT			5 13
SORT (GROUP BY NOSORT)			5 13
MERGE JOIN			5 13
INDEX (FULL SCAN)	PK_PAISES	28	1
SORT (JOIN)			5 11
Access Predicates			
ITEM_1=PA.COD			
Filter Predicates			
ITEM_1=PA.COD			
VIEW	VW_GBC_13		5 10
HASH (GROUP BY)			5 10
HASH JOIN			5 9
Access Predicates			
CI.IDCENTRO=CS.ID			
HASH JOIN			5 6
Access Predicates			
IDCIU=CI.ID			
TABLE ACCESS (FULL)	ACT_DIARIA	5	3
Filter Predicates			
EXTRACT(YEAR FROM INTERNA			
VIEW	index\$_join\$_004		8 3
Filter Predicates			
UTILIDADES.CIUDADANOS_GET			
HASH JOIN			

Plan de ejecución después de crear el índice NC\_ANIO\_ACTDIARIA

OPERATION	OBJECT_NAME	CARDINALITY	COST
SELECT STATEMENT			5 12
SORT (GROUP BY NOSORT)			5 12
MERGE JOIN			5 12
INDEX (FULL SCAN)	PK_PAISES	28	1
SORT (JOIN)			5 10
Access Predicates			
ITEM_1=PA.COD			
Filter Predicates			
ITEM_1=PA.COD			
VIEW	VW_GBC_13		5 9
HASH (GROUP BY)			5 9
HASH JOIN			5 8
Access Predicates			
CI.IDCENTRO=CS.ID			
HASH JOIN			5 5
Access Predicates			
IDCIU=CI.ID			
TABLE ACCESS (BY INDEX ROWID)	ACT_DIARIA	5	2
INDEX (RANGE SCAN)	NC_ANIO_ACTDIARIA	2	1
Access Predicates			
EXTRACT(YEAR FROM INTE			
VIEW	index\$_join\$_004		8 3
Filter Predicates			
UTILIDADES.CIUDADANOS_GET			

# Control – Pruebas del producto



## Scripts para instalación de entorno de pruebas ( para instalar llegime.txt)

1_01_00_Preparar_entorno_BD.sql	1.285	481	Archivo SQL
1_01_01_Crear_tablas.sql	12.994	1.817	Archivo SQL
1_02_00_PAQUETE_UTILIDADES.sql	38.367	3.632	Archivo SQL
1_02_01_Crear_vistas.sql	1.687	523	Archivo SQL
1_02_02_PAQUETE_PERSONA.sql	15.736	1.935	Archivo SQL
1_02_03_PAQUETE_ESTADISTICA.sql	46.897	4.031	Archivo SQL
1_02_04_PAQUETE_SALUD.sql	86.258	6.922	Archivo SQL
1_02_05_PAQUETE_LISTADOS.sql	9.828	2.312	Archivo SQL
1_03_00_Crear_disparadores.sql	14.679	1.601	Archivo SQL
1_04_00_Cargar_tablas_auxiliares.sql	4.159	905	Archivo SQL

## Scripts para las pruebas de los procedimientos de actualización

1_04_01_Carga_datos_basicos_pruebas.sql	60.184	8.379	Archivo SQL
1_04_02_Cargar_movimientos_pruebas_estadistica.sql	82.707	8.537	Archivo SQL
1_04_03_Actualiza_datos_pruebas.sql	1.222	537	Archivo SQL
1_04_04_Control_errores_pruebas.sql	6.257	1.282	Archivo SQL

## Script para las pruebas de módulo estadístico

1_04_05_Modifica_movimientos_pruebas_estadistica.sql	966	427	Archivo SQL
--	-----	-----	-------------

## Demstración utilización de listados

```
SELECT UTILIDADES.CIUDADANOS_GetNomComp (IDCIU) CIUDADANO
, ESTSA.*
FROM TABLE ( LISTADOS.CIUDADANOS_EstadoSalud(2014, 'BE', 35,1500) ) ESTSA
ORDER BY CIUDADANO;
```

Resultado de la Consulta

CIUDADANO	IDCIU	VISITASMED	MEDCALCS	ANIOECAR	ANIOECAN	ANIOEDEG	ENFLEVS
1 JAKOB EDUARD Z	9	2	1391	(null)	(null)	(null)	2
2 JENS ALBES DESC	8	11	1197	(null)	(null)	(null)	4

```
SELECT * FROM TABLE (LISTADOS.ACT_DIARIA_SegMensual(108,2014) );
```

Resultado de la Consulta

MES	TOTPASOS	MEDCALCS	TOTHPOCACT	
1	12	6024	1403	14
2	11	2511	1254	9

# Conclusiones

---

- ❖ **Aplicar la metodología adecuada, definir un estándar de desarrollo y describir los procesos de trabajo permiten que el desarrollo del proyecto informático sea muy ágil y facilita mucho el posterior mantenimiento del producto obtenido.**
- ❖ **Podemos afirmar que el producto obtenido funcionará si se han realizado unas buenas especificaciones, un buen diseño, una buena implementación y un buen control de calidad.**
- ❖ **Un proyecto informático es una misión compartida entre un grupo de personas, en el que es esencial la comunicación.**
- ❖ **Para que el producto obtenido perdure, es necesario mantener formado al personal que se encarga del mantenimiento del producto.**
- ❖ **Un proyecto tiene éxito si es útil, y es útil si su producto es utilizado, y la magnitud de su éxito se basa en el tiempo que el producto obtenido está en uso, cubriendo las necesidades por las que ha sido diseñado.**

# Gracias por la atención prestada



*A person sows a thought and gathers an action  
He sows an action and gathers a habit  
He sows a habit and gathers a character  
He sows a character and gathers a destination  
(Svami Shivananda)*