

**(Creative Commons)**

Aquest treball està subjecte - excepte que s'indiqui el contrari- en una llicència de Reconeixement-NoComercial-SenseObraDerivada 2.5 Espanya de Creative Commons. Podeu copiar-lo, distribuir-los i transmetre'ls públicament sempre que citeu l'autor i l'obra, no es faci un ús comercial i no es faci còpia derivada. La llicència completa es pot consultar en <http://creativecommons.org/licenses/by-nc-nd/2.5/es/deed.es>

---

## CREACIÓ D'UN MAGATZEM DE DADES

### MEMÒRIA DEL PROJECTE

---

Marta Adell i Roig

Treball Fi de Carrera  
Enginyeria Tècnica d'Informàtica de Gestió  
Curs 2008-2009 – Primer Semestre

**Consultor:** José Ángel Martín Carballo

**Data lliurament:** 7 de gener de 2009

## RESUM

El projecte '**Construcció i explotació d'un magatzem de dades per a l'anàlisi estadístic dels resultats del Campionat de Fórmula 1**' és un treball de disseny i implementació d'una eina de consulta i explotació estadística dels resultats històrics de la Fórmula 1 que, vol donar resposta a les necessitats expressades per l'Institut Català d'Esports de Motor (en endavant, ICEM).

El treball s'inicia amb l'anàlisi de les fonts de dades proporcionades per l'ICEM en un fitxer MS Excel. En funció d'aquest estudi, es realitza el disseny del model de dades i els procediments de transformació i extracció de la informació d'origen.

Posteriorment, es desenvolupa el magatzem de dades i processos associats seguint la solució proposada i, s'implementa l'eina de consulta on l'usuari final, sense coneixements tècnics de bases de dades, podrà realitzar l'explotació de la informació.

Per últim, es realitza un exercici d'anàlisi com a demostració d'utilització de l'eina des del punt de vista de l'usuari final. Aquest anàlisi comença amb la recollida de dades mitjançant l'eina d'explotació del nou sistema i, conclou amb la interpretació d'aquesta informació.

Pel desenvolupament del projecte s'ha utilitzat el Sistema de Gestió de Bases de Dades Oracle 10g Express (en endavant, SGBD Oracle 10g) i el programari de consulta Oracle Discoverer.

## ÍNDIX

<b>1. INTRODUCCIÓ</b> .....	<b>6</b>
1.1. PUNT DE PARTIDA I APORTACIÓ DEL TFC .....	6
1.2. OBJECTIUS DEL TFC.....	6
1.3. ENFOCAMENT I MÈTODE SEGUIT.....	7
1.4. PLANIFICACIÓ DEL PROJECTE .....	8
1.5. PRODUCTES OBTINGUTS .....	11
1.6. BREU DESCRIPCIÓ DELS ALTRES CAPÍTOLS DE LA MEMÒRIA .....	12
<b>2. ANÀLISI FUNCIONAL</b> .....	<b>13</b>
2.1. ESTUDI DE LA FONT D'ORIGEN.....	13
2.2. OBJECTIUS DEL PROJECTE .....	14
2.3. PROPOSTA DE SISTEMA.....	14
2.3.1. <i>Justificació del sistema</i> .....	16
2.3.2. <i>Magatzem de dades</i> .....	16
2.4. DEFINICIÓ DE ROLS .....	17
2.5. DIAGRAMES DE CASOS D'ÚS .....	18
2.6. REQUERIMENTS FUNCIONALS .....	19
2.6.1. <i>Indicadors</i> .....	19
2.6.2. <i>Creació d'informes personalitzats</i> .....	20
2.7. DIAGRAMA DEL MODEL CONCEPTUAL.....	21
2.7.1. <i>Model relacional</i> .....	21
2.7.2. <i>Model dimensional</i> .....	22
<b>3. DISSENY</b> .....	<b>24</b>
3.1. DIAGRAMA DE L'ARQUITECTURA SOFTWARE.....	24
3.2. DIAGRAMA DE L'ARQUITECTURA HARDWARE .....	25
3.3. DISSENY DE LA BASE DE DADES I DIAGRAMA DEL MODEL FÍSIC .....	26
3.3.1. <i>Base de dades relacional</i> .....	26
3.3.2. <i>Magatzem de dades</i> .....	27
3.4. ARQUITECTURA I DISSENY DELS PROCESSOS DE CÀRREGA.....	28
3.4.1. <i>Extracció de les dades d'origen</i> .....	28
3.4.2. <i>Càrrega inicial a la BD Relacional</i> .....	29
3.4.3. <i>Càrrega inicial al Magatzem de Dades</i> .....	30
3.4.4. <i>Manteniment de la informació del magatzem</i> .....	30
3.5. DISSENY I DESCRIPCIÓ DELS INFORMES CREATS .....	32
3.5.1. <i>Rànquing de victòries per Mundial</i> .....	32
3.5.2. <i>Rècords de títols</i> .....	32
3.5.3. <i>Posició mitjana de pilots per Mundial</i> .....	32
3.5.4. <i>Velocitat mitjana de pilots per GP</i> .....	33
3.5.5. <i>Total d'abandonaments per GP</i> .....	33
3.5.6. <i>Rànquing de puntuació per GP</i> .....	33
3.5.7. <i>Pilots de Mundial</i> .....	34
3.5.8. <i>Campionat del món de pilots</i> .....	34
<b>4. ANÀLISI DE LA INFORMACIÓ: RELACIÓ POLE - VICTÒRIA</b> .....	<b>35</b>
<b>5. CAPTURES DE PANTALLA</b> .....	<b>38</b>
5.1. BD RELACIONAL .....	38
5.2. MAGATZEM DE DADES.....	38
5.3. ÀREA DE NEGOCI .....	39
5.4. INFORMES CREATS.....	40
5.4.1. <i>Rànquing de victòries per Mundial</i> .....	41
5.4.2. <i>Rècords de títols</i> .....	43
5.4.3. <i>Posició mitjana de pilots per Mundial</i> .....	45
5.4.4. <i>Velocitat mitjana de pilots per GP</i> .....	45
5.4.5. <i>Total d'abandonaments per GP</i> .....	47
5.4.6. <i>Rànquing de puntuació per GP</i> .....	48
5.4.7. <i>Pilots de Mundial</i> .....	49

5.4.8. <i>Campionat del món de pilots</i> .....	50
<b>6. LÍNIES D'EVOLUCIÓ FUTURA</b> .....	<b>51</b>
6.1. A CURT TERMINI: MILLORES DEL SISTEMA ACTUAL .....	51
6.2. A MIG-LLARG TERMINI: EVOLUCIÓ DEL SISTEMA.....	51
<b>7. CONCLUSIONS I AGRAÏMENTS</b> .....	<b>52</b>
<b>8. GLOSSARI</b> .....	<b>53</b>
<b>9. BIBLIOGRAFIA</b> .....	<b>55</b>
<b>10. ANNEXOS</b> .....	<b>56</b>
10.1. ANNEX 1. ANÀLISI DE LES FONTS DE DADES D'ORIGEN.....	56
10.2. ANNEX 2: CÀRREGA DE DADES D'ORIGEN .....	57
10.3. ANNEX 3: CÀRREGA INICIAL A LA BD RELACIONAL .....	57
10.4. ANNEX 4. TRACTAMENT DADES D'ORIGEN.....	59
10.5. ANNEX 5. REGISTRES PENDENTS DE REVISIÓ .....	59
10.6. ANNEX 6: CÀRREGA INICIAL AL MAGATZEM .....	60
10.7. ANNEX 7: PROCEDIMENTS EMMAGATZEMATS.....	61
10.8. ANNEX 8: AUTOMATITZACIÓ DE LA CÀRREGA .....	63
10.9. ANNEX 9: USUARIS DEL SISTEMA .....	63

**FIGURES**

Figura 1. Cicle de vida clàssic .....	7
Figura 2. Diagrama de Gantt .....	10
Figura 3. Entitats identificades en la font d'origen .....	13
Figura 4. Arquitectura de desenvolupament.....	15
Figura 5. Casos d'ús.....	18
Figura 6. Model conceptual BD relacional.....	21
Figura 7. Diagrama estrella Magatzem de Dades .....	22
Figura 8. Model conceptual dimensions .....	23
Figura 9. Arquitectura final de programari .....	24
Figura 10. Arquitectura de maquinari .....	25
Figura 11. Model físic BD relacional.....	26
Figura 12. Model físic Magatzem de Dades .....	27
Figura 13. Model físic BD temporal .....	28

## 1. INTRODUCCIÓ

### 1.1. Punt de partida i aportació del TFC

L'ICEM és l'entitat encarregada de proporcionar informació oficial als organismes públics i mitjans de comunicació catalans respecte els resultats obtinguts a les diferents competicions mundials d'esports de motor.

Actualment, l'ICEM proporciona aquesta informació de manera molt manual ja que tots els resultats històrics es troben recollits en documents MS Excel. Això implica la realització d'una feina molt feixuga i la necessitat de disposar d'un elevat nombre de recursos humans.

El projecte s'ha posat en marxa amb l'objectiu de millorar el sistema existent, augmentant la fiabilitat de la informació i garantir-ne la integritat i, proporcionant una eina flexible que permeti l'anàlisi dels resultats històrics del Campionat del Món de Fórmula 1 de manera àgil i transparent per l'usuari.

Les tasques a realitzar consisteixen bàsicament en el disseny i creació d'un magatzem de dades i, el desenvolupament d'un entorn d'explotació estadística que inclou un conjunt d'informes predefinits amb la informació bàsica sol·licitada.

Aquest treball representa la primera fase en el procés de construcció d'un sistema molt més ampli que englobarà els resultats obtinguts a les diferents competicions mundials d'esports de motor. És per això, que s'ha de donar especial rellevància a l'escalabilitat de la infraestructura física i operacional d'aquesta nova eina.

Per tal d'avaluar les dades existents, l'ICEM ha lliurat un fitxer Excel amb els resultats de la Fórmula 1 dels últims anys. Aquest és el punt de partida del TFC 'Magatzems de Dades' del primer semestre del curs 2008-2009.

### 1.2. Objectius del TFC

La realització d'aquest treball té com a principal finalitat, familiaritzar-se amb l'arquitectura general d'un magatzem de dades desenvolupant un cas pràctic de caire professional des de la seva planificació inicial fins a l'obtenció del producte final.

Paral·lelament a la consecució d'aquesta fita, s'assoleixen altres objectius importants. Són els següents:

- Aprofundir en el coneixement de les bases de dades i la gestió de la informació
- Introducció en el camp dels magatzems de dades o *Data Warehouse* (DW)
- Anàlisi de tècniques existents per a projectar la base de dades d'un DW
- Adquirir experiència en el disseny, construcció i explotació d'un DW
- Analitzar i comparar les eines de mercat que treballen amb model de dades físic del DW
- Realitzar les anàlisis estadístiques dirigides als usuaris finals del DW
- Aproximar-se al món professional seguint les etapes de desenvolupament d'un sistema segons les necessitats definides pels clients.

### 1.3. Enfocament i mètode seguit

Per a dur a terme el desenvolupament del projecte s'ha seguit l'estratègia del cicle de vida clàssic (també anomenat en cascada) adaptant-lo al pla d'estudis de l'assignatura.

D'aquesta manera, les principals tasques a realitzar s'han fet coincidir amb els lliuraments previstos a l'aula, de la manera següent:

1. PAC1: Pla de Treball i anàlisi preliminar de requeriments
2. PAC2: Anàlisi de requeriments, disseny conceptual i tècnic
3. PAC3: Construcció de l'eina, proves i implantació
4. Lliurament Final: Programari corregit i memòria del projecte

En la **figura 1** es poden veure les diferents etapes del cicle de vida clàssic, on cada fase és el punt de partida de la següent.

Tot i així, en el TFC hi ha etapes que seran parcial o totalment inexistents (p.ex. proves d'usuaris i fase de manteniment) i, al tractar-se d'un treball que promou l'aprenentatge, està obert a modificacions de qualsevol punt al llarg de tot el procés.

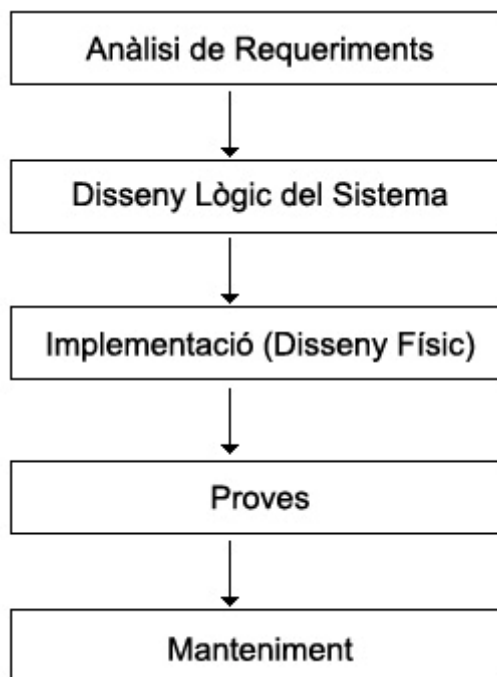


Figura 1. Cicle de vida clàssic

#### 1.4. Planificació del projecte

En aquest apartat es mostra la planificació de les diferent tasques del projecte i les fites de cada una d'elles.

La temporització indicada s'ha ajustat al calendari establert per la Universitat i, a la disponibilitat horària de l'estudiant.

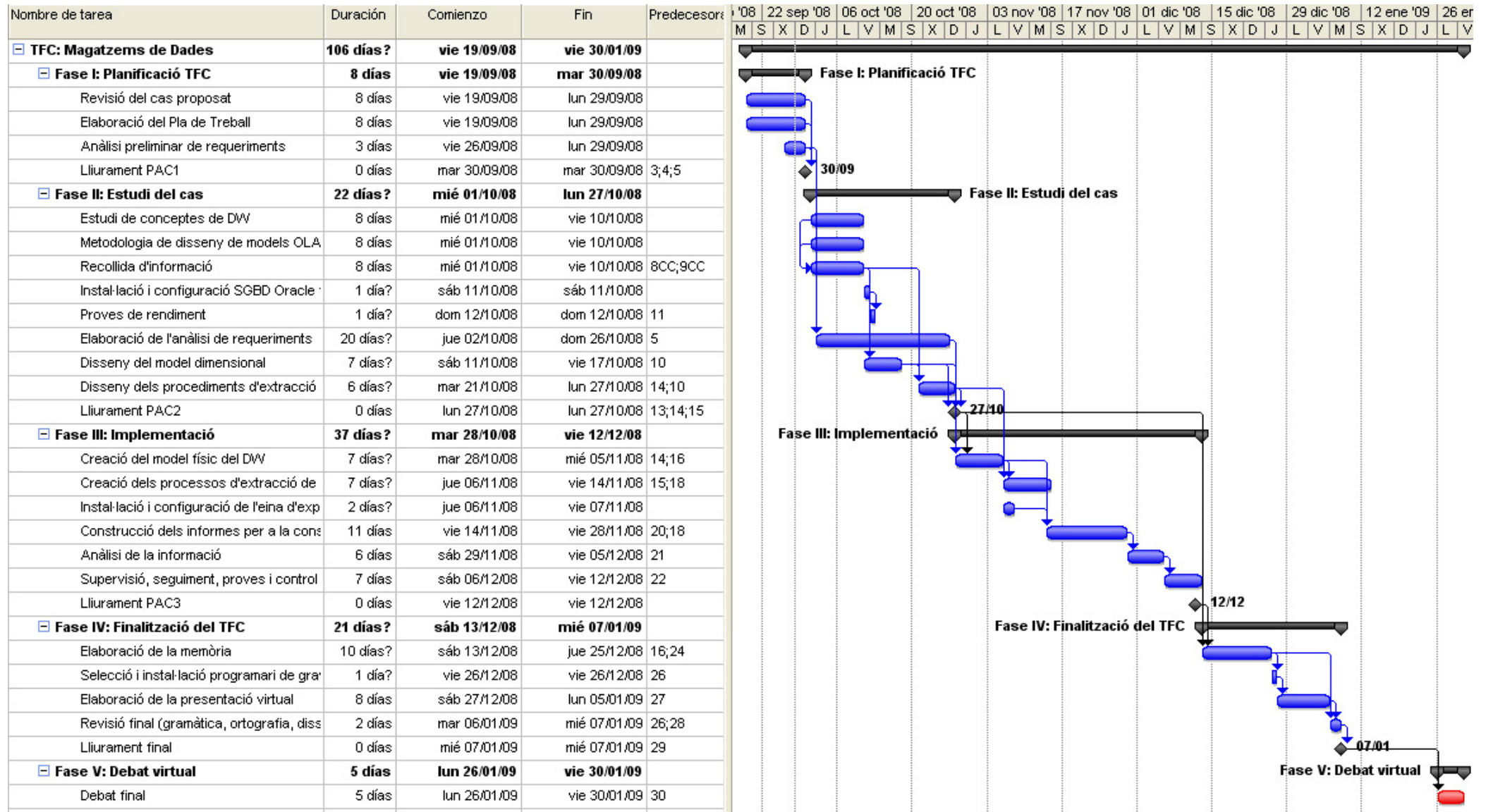
TASCA	Data Inici	Data Fi
<b>Fase I: Planificació del TFC</b>	<b>19/09/08</b>	<b>30/09/08</b>
1. Revisió del cas proposat 2. Elaboració del Pla de Treball 3. Anàlisi preliminar de requeriments	19/09	30/09
<b>FITA: Lliurament PAC1</b>		
<b>Fase II: Estudi del cas</b>	<b>01/10/08</b>	<b>27/10/08</b>
4. Estudi de conceptes de DW 5. Metodologia de disseny de models OLAP a partir de models OLTP 6. Recollida d'informació	01/10	10/10
7. Instal·lació i configuració SGBD Oracle 10g 8. Proves de rendiment	11/10	12/10
9. Elaboració de l'anàlisi de requeriments	01/10	26/10
10. Disseny del model dimensional	12/10	20/10
11. Disseny dels procediments d'extracció de dades d'origen	21/10	27/10
<b>FITA: Lliurament PAC2</b>		



<b>TASCA</b>	<b>Data Inici</b>	<b>Data Fi</b>
<b>Fase III: Implementació</b>	<b>28/10/08</b>	<b>12/12/08</b>
12. Creació del model físic del DW	28/10	5/11
13. Creació dels processos d'extracció de dades origen	6/11	14/11
14. Instal·lació i configuració de l'eina d'explotació de dades	6/11	7/11
15. Construcció dels informes per a la consulta d'informació	14/11	28/11
16. Anàlisi de la informació	29/11	05/12
17. Supervisió, seguiment, proves i control de qualitat	06/12	12/12
<b>FITA: Lliurament PAC3</b>		
<b>Fase IV: Finalització del TFC</b>	<b>13/12/08</b>	<b>07/01/09</b>
18. Elaboració de la memòria	13/12	25/12
19. Selecció i instal·lació programari de gravació d'escriptori	26/12	26/12
20. Elaboració de la presentació virtual	27/12	05/01
21. Revisió final (gramàtica, ortografia, disseny, etc.)	06/01	07/01
<b>FITA: Lliurament final</b>		
<b>Fase V: Debat virtual</b>	<b>26/01/09</b>	<b>30/01/09</b>
22. Debat final	26/01	30/01

La **figura 2** de la pàgina següent presenta el calendari, en mesos, del desenvolupament del TFC desglossat en fases.

Figura 2. Diagrama de Gantt



### 1.5. Productes obtinguts

Al llarg del procés de desenvolupament del projecte s'assoleixen diferents fites. En aquest apartat s'enumeren i es descriuen breument els productes obtinguts en cada una d'elles.

<b>FITA: Lliurament PAC1</b>	<b>30/09/08</b>
<p><u>Pla de Treball</u> Documentació amb la planificació de les tasques a realitzar per a l'elaboració del TFC. Concretament, es determinen les activitats necessàries per assolir els objectius del projecte i, es realitza una estimació de l'esforç i els recursos necessaris per dur-les a terme.</p>	
<b>FITA: Lliurament PAC2</b>	<b>27/10/08</b>
<p><u>Anàlisi de requeriments</u> Estudi que mostra les necessitats funcionals expressades pels usuaris i descriu la situació actual de l'empresa.</p> <p><u>Disseny</u> Document que descriu la solució proposada per satisfer els requeriments dels usuaris i en justifica la seva elecció. Inclou el disseny conceptual i físic del model dimensional i, l'arquitectura i disseny dels procediments d'extracció de dades des de la font d'origen.</p>	
<b>FITA: Lliurament PAC3</b>	<b>12/12/08</b>
<p><u>Fitxers d'implementació</u> Desenvolupament del sistema pròpiament dit. Es realitza a partir del disseny anterior i inclou principalment, el magatzem de dades carregat amb la informació inicial, els processos necessaris pel correcte funcionament i manteniment de les dades i, l'eina de consulta que permetrà als usuaris sense coneixements de bases de dades l'explotació i anàlisi de la informació.</p> <p>El sistema s'ha implementat amb SGBD Oracle 10g i el programari de consulta Oracle Discoverer.</p> <p><u>Manual d'instal·lació</u> Guia que detalla els passos per a realitzar la importació del sistema a partir dels fitxers d'implementació i, descriu la plataforma tecnològica que utilitza.</p> <p><u>Guia de desenvolupament</u> Document que s'adjunta amb la implementació i descriu els passos realitzats durant la construcció del TFC. Inclou detalls tècnics al més baix nivell per tal que el consultor pugui valorar la feina feta.</p> <p><u>Anàlisi de la informació</u> Demostració d'utilització de l'eina mitjançant un exercici d'anàlisi des del punt de vista de l'usuari final. Comença amb la recollida de dades mitjançant l'eina d'explotació del nou sistema implantat i, conclou amb la interpretació d'aquesta informació.</p>	

<b>FITA: Lliurament final</b>	<b>07/01/09</b>
<p><u>Memòria del projecte</u> Síntesis de la feina realitzada durant el procés de desenvolupament del TFC. Conté aquella informació rellevant que permeti comprendre el problema plantejat, la metodologia utilitzada per la seva resolució i mostri el resultat obtingut.</p> <p><u>Presentació virtual del projecte</u> Presentació multimèdia que complementa la memòria proporcionant una visió global del projecte i, en destaca els punts més importants.</p>	

### 1.6. Breu descripció dels altres capítols de la memòria

A continuació es resumeix el contingut dels propers apartats de la memòria.

#### Anàlisi funcional

En aquest capítol es fa un repàs de la situació actual de l'empresa i es detallen els requeriments que ha de complir el sistema. En funció d'aquesta informació es descriu la solució proposada a nivell conceptual i se'n justifica la seva elecció.

#### Disseny

Aquest bloc presenta el disseny físic realitzat seguint l'anàlisi anterior. Inclou els models de bases de dades i l'arquitectura dels processos de càrrega. També es descriu l'arquitectura de programari i maquinari necessari pel correcte funcionament del nou sistema. Finalment es detallen els informes creats responent a la demanda dels usuaris.

#### Anàlisi de la informació

Anàlisi de la informació, des del punt de vista de l'usuari final, segons es demana a l'enunciat del TFC. En concret, s'explica la relació entre el fet d'obtenir la *Pole Position* i guanyar la cursa, detallant els criteris seguits per arribar a la conclusió donada.

#### Captures de pantalla

En aquest apartat es mostren captures de pantalla de les bases de dades del sistema i l'eina d'explotació de la informació. També es pot visualitzar el prototip dels informes creats, juntament amb una petita guia d'utilització.

#### Línies d'evolució futures

Llista de recomanacions que exposen futures accions que poden semblar necessàries, com a resultat directe de les conclusions o d'alguna experiència feta en el curs del treball objecte de la memòria.

#### Conclusions

En aquest apartat s'inclouen les deduccions a les que s'ha arribat com a conseqüència del treball descrit en la memòria.

## 2. ANÀLISI FUNCIONAL

Com s'ha mencionat anteriorment, el sistema que utilitza l'ICEM per informar dels resultats de les diferents competicions mundials d'esports de motor és molt costós, tant pels recursos humans utilitzats com per la tasca laboriosa que representa.

La finalitat d'aquest projecte és la creació d'un entorn que permeti als analistes de l'ICEM l'explotació estadística dels resultats històrics de la Fórmula 1 d'una manera més eficaç.

Tanmateix, aquest treball és en realitat la primera fase en el procés de construcció d'un sistema molt més ampli que englobarà els resultats obtinguts en les diferents competicions d'esports de motor.

### 2.1. Estudi de la font d'origen

Actualment tots els resultats històrics es troben recollits en documents MS Excel. L'ICEM ens fa entrega d'un fitxer que conté les dades dels resultats de la Fórmula 1 des de l'any 1961 fins el 2007.

En la **figura 3** es mostren les entitats identificades en la font d'origen juntament amb els seus atributs. Cada entitat correspon a un full del fitxer Excel.

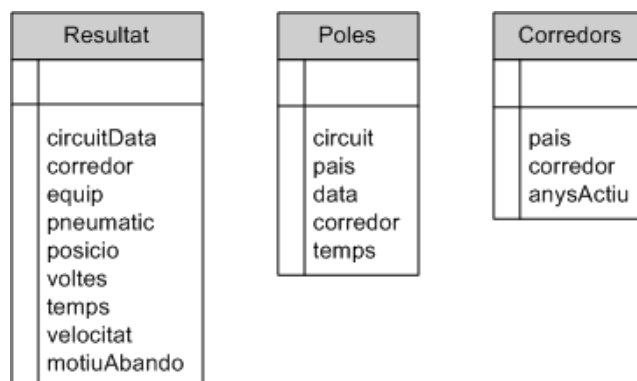


Figura 3. Entitats identificades en la font d'origen

A grans trets, i un cop analitzat aquest fitxer, es detecta l'existència d'entitats no relacionades i sense restriccions d'integritat. Com a conseqüència es produeixen certes deficiències en la informació, com són les següents

- Valors inconsistents
- Duplictat de la informació i redundància
- Manca de coherència
- Formats diferents per mostrar el mateix tipus de dades

**Nota:** En l'annex 1 es detallen les observacions derivades de l'anàlisi inicial de les dades contingudes en el fitxer Excel original.

## 2.2. Objectius del projecte

Els principals objectius que es volen assolir amb aquesta nova eina són els següents:

- Proporcionar a l'ICEM un sistema àgil d'explotació de la informació i anàlisi estadístic
- Millorar la recollida de dades i l'accés als resultats històrics
- Augmentar la fiabilitat de la informació i garantir-ne la integritat
- Millorar el servei proporcionat als organismes públics i mitjans de comunicació
- Alleugerir la càrrega de treball dels tècnics de l'ICEM
- Disminuir el nombre de recursos humans necessaris per a l'explotació de dades
- Permetre l'obtenció d'informació estadística de manera simple i transparent per l'usuari
- Proporcionar un conjunt predefinit d'informes per mostrar la informació definida per l'ICEM i qualsevol altra que pugui resultar útil per l'anàlisi estadístic dels resultats esportius
- Proporcionar una eina flexible que permeti als analistes d'obtenir directament la informació que requereixen per a l'anàlisi i la decisió estratègica de negoci, sense necessitat de realitzar peticions als usuaris tècnics

## 2.3. Proposta de sistema

Per a donar suport a les necessitats d'informació dels analistes, es proposa incorporar el concepte de magatzem de dades a l'ICEM mitjançant la implementació de la **Factoria d'Informació Corporativa** (en endavant, FIC).

La FIC és el conjunt d'accions que cal realitzar i components que s'han de crear per a dur a terme aquest projecte i que té, com a eix central, el magatzem de dades.

Entre aquestes accions, es destaca la necessitat de millorar, adequar i racionalitzar les dades que hi ha en el sistema operacional actual de l'empresa (fitxer Excel). Això és degut a que la poca fiabilitat detectada en algunes d'aquestes dades i la seva redundància podrien generar problemes d'integritat en el magatzem de dades. La solució a aquesta problemàtica passa pel desenvolupament d'una nova base de dades relacional.

En resum, les tasques a realitzar per tal de crear aquest sistema són les següents:

1. Anàlisi de les fonts de dades proporcionades
2. Creació d'una base de dades temporal sobre la que es realitza la importació de les dades contingudes en la font d'origen (fitxer Excel proporcionat per l'ICEM)
3. Creació de la base de dades relacional (en endavant, BD Relacional)
4. Procés de transformació, integració i depuració de les dades
5. Càrrega de dades a la BD relacional

6. Creació del Magatzem de Dades
7. Procés de càrrega inicial de dades provinents de la BD Relacional al magatzem
8. Desenvolupament d'informes predefinitos
9. Creació del procés de càrrega que s'executarà periòdicament per mantenir actualitzada la informació del magatzem a partir de les dades de la BD Relacional

Un cop implementada la FIC i sense requerir coneixements de bases de dades, l'analista podrà accedir a l'explotació estadística de la informació mitjançant els informes predefinitos o la creació d'informes a mida.

Es recomana realitzar el manteniment de la informació directament sobre la nova base de dades operacional. La desaparició del fitxer Excel a favor de la utilització d'aquesta base de dades augmentarà la fiabilitat de la informació i garantirà la integració de les dades.

Tanmateix, queda **fora de l'abast** del projecte la creació d'una funcionalitat específica per introduir (o modificar) informació a la base de dades operacional.

La **figura 4** mostra l'esquema que resumeix el procés de desenvolupament del projecte.

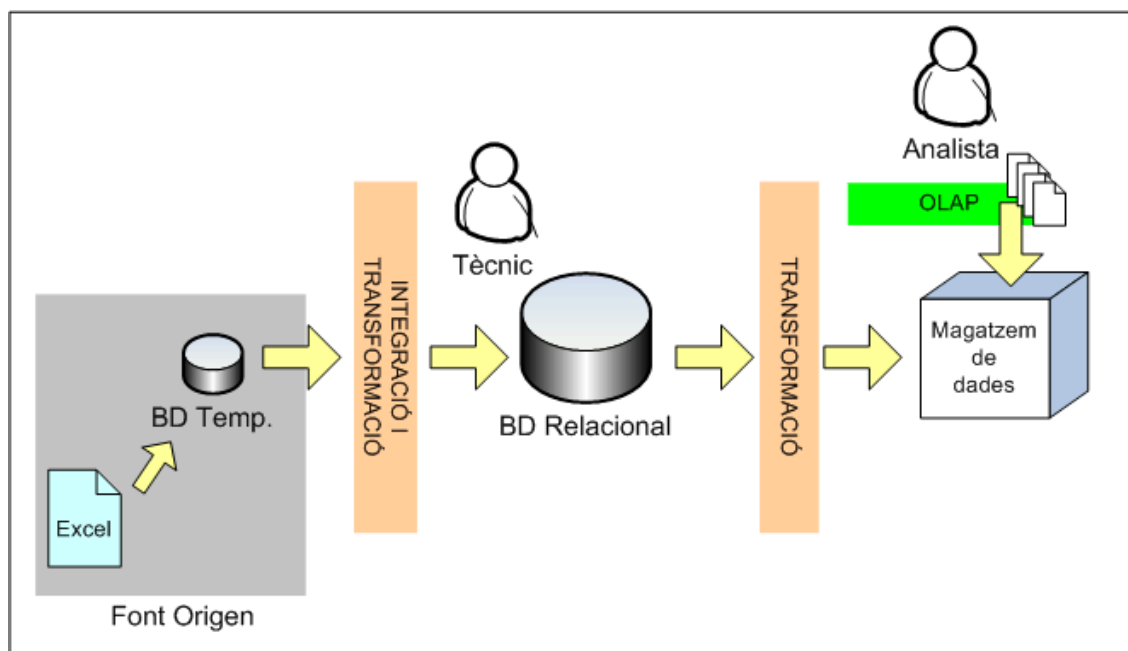


Figura 4. Arquitectura de desenvolupament

### 2.3.1. Justificació del sistema

En relació als resultats de les competicions, l'ICEM s'encarrega d'enregistrar la informació, mantenir les dades històriques i facilitar-les als diferents clients.

Estem doncs, davant una organització que realitza insercions de dades cada cop que finalitza una competició i que, pràcticament, no ha de gestionar un manteniment de la informació donat que els valors introduïts rarament variaran (a no ser que sigui per corregir alguna errada).

Finalment, les dades contingudes en el reservori s'utilitzen per elaborar llistats i estadístiques sobre els resultats de les curses.

Tenint en compte aquesta descripció de les funcions que realitza l'ICEM podem deduir el següent:

- Ens trobem davant un gran volum de dades que volem consultar però sobre el qual no voldrem realitzar operacions de manteniment.
- El factor temps no és crític, és a dir, és més important la optimització de les consultes massives de dades que l'accés immediat a una dada en concret.

És per aquest motiu, que es recomana la creació d'un magatzem de dades.

### 2.3.2. Magatzem de dades

Un magatzem de dades (de l'anglès *data warehouse*) és una col·lecció de dades amb la informació històrica d'una organització dissenyada i estructurada per a realitzar-hi consultes eficientment i donar suport en la presa de decisions de l'organització.

La importància del magatzem de dades està en les dades històriques i no volàtils, sobre les que es realitzen consultes massives de dades orientades a àrees de negoci. Les dades temporals són especialment importants en tasques d'anàlisi; qualsevol dada en el magatzem de dades ha d'anar acompanyada del seu període de validesa. Així, les dades no s'esborren o modifiquen, sinó que s'insereixen les correccions i la data en què s'han fet.

Les principals característiques que defineixen el magatzem de dades són les següents:

- Informació estructurada per temes d'interès (només conté informació d'interès per a l'anàlisi de la informació)
- Capacitat d'emmagatzemar gran volum d'informació de manera homogènia i fiable
- Informació històrica que ajuda a l'estudi de l'evolució del negoci al llarg del temps
- Informació no volàtil que garanteix que no es perd cap dada (ni tant sols les errònies)
- Estructura de dades jeràrquica pensada per facilitar les consultes estratègiques de la direcció de l'organització
- Actualització de manera periòdica a partir de la base de dades operacional
- Optimització de consultes massives de dades



## 2.4. Definició de rols

En relació amb el sistema que es vol desenvolupar s'identifiquen els usuaris següents:

### **Fase de desenvolupament**

#### Interlocutor de l'ICEM

Persona que realitza les tasques de definició de requeriments i funcionalitats de l'aplicació així com l'usuari de contacte en la fase de desenvolupament per aclariments de dubtes que puguin sorgir.

Les tasques principals de l'interlocutor són les següents:

- Definició de requeriments
- Validació i vist-i-plau dels documents lliurats
- Aclariments de dubtes que puguin sorgir durant el desenvolupament.

### **Fase de proves i manteniment**

#### Tècnic de l'ICEM

Usuari tècnic encarregat del manteniment de les dades a nivell de la base de dades operacional.

Podrà realitzar les accions següents:

- Gestionar la informació de la BD relacional
- Introduir nova informació
- Modificar les dades existents

#### Administrador de Bases de Dades (DBA)

Figura encarregada de l'administració de les bases de dades en Oracle 10g, tant de la BD operacional com del magatzem de dades. Podrà executar manualment els processos de càrrega quan sigui necessari. També gestionarà l'àrea de negoci utilitzant Oracle Discoverer Administrator.

Quan es posi en funcionament l'aplicació, aquest perfil pot formar part de la plantilla de l'ICEM o bé tractar-se d'un servei extern.

#### Analista de l'ICEM

Personal que pertany a la cúpula de l'empresa. Té accés al magatzem de dades i, realitzarà l'explotació i anàlisi de la informació mitjançant Oracle Discoverer Desktop. Podrà treballar amb els informes predefinits i crear-ne de nous segons convingui.

Analista i interlocutor solen ser la mateixa persona.

#### Client de l'ICEM

Organismes públics i mitjans de comunicació catalans que reben de l'ICEM la informació relacionada amb els resultats obtinguts a les diferents competicions mundials d'esports de motor.

La interacció d'aquest tipus d'usuari amb el sistema és nul·la (veure apartat 6.2, *línies d'evolució futura – evolució del sistema*).

### 2.5. Diagrames de casos d'ús

A continuació es mostren els diagrames que representen els principals casos d'interacció dels usuaris amb el sistema, descrits en el punt anterior.

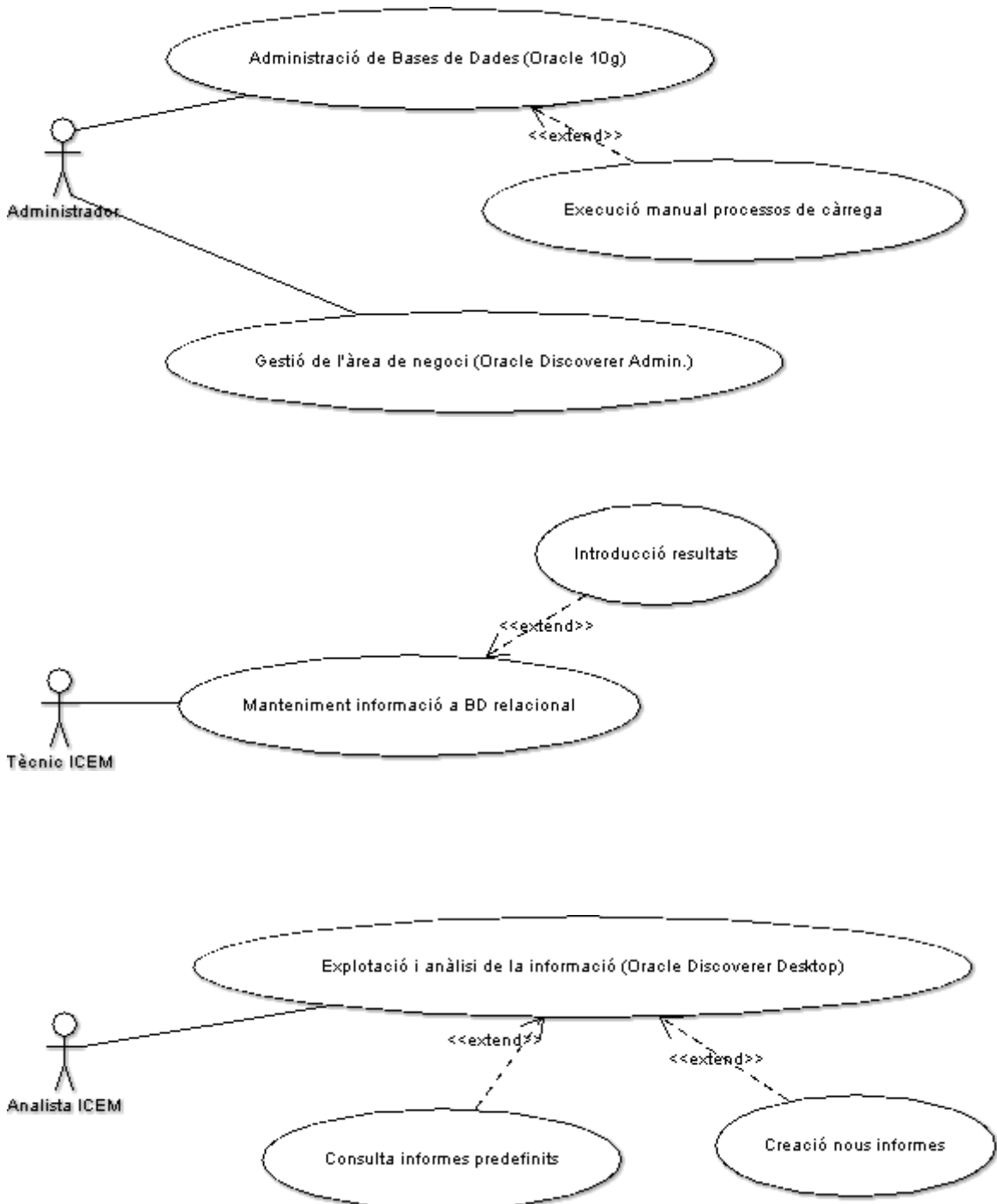


Figura 5. Casos d'ús

## 2.6. Requeriments funcionals

En aquest apartat es descriu el conjunt de prestacions que el sistema ha de subministrar als analistes, per a la realització dels processos de negoci suportats.

### 2.6.1. Indicadors

En funció de les necessitats definides per l'ICEM i tenint en compte la informació proporcionada, es detecten una sèrie d'indicadors. Tots ells es podran observar en funció del temps i són els següents:

- Total de Gran Premis (en endavant, GP) guanyats per corredor
- Total de GP guanyats per tipus de pneumàtic
- Total de GP guanyats per equip (rànkning equips)
- Total Mundials guanyats per equip
- Total Mundials guanyats per corredor
- Total Mundials guanyats per pneumàtic
- Posició mitjana d'un corredor
- Velocitat mitjana per GP
- Velocitat mitjana per corredor
- Total d'abandonaments per GP
- Total d'abandonaments per motiu
- Total punts per equip (rànkning Mundial de Constructors)
- Total punts per corredor (rànkning Mundial de Pilots)
- Total punts de cursa per equip
- Total punts de cursa per corredor

Per tal de visualitzar en detall aquesta informació, s'implementaran una sèrie d'informes predefinits que estaran a disposició de l'analista.

Quan aquest usuari vulgui generar un informe concret, haurà d'accedir al sistema i indicar els valors desitjats, dins de les opcions possibles.

Els informes creats durant el desenvolupament del nou sistema es descriuen en detall a l'apartat 3.5 d'aquest document.

### **2.6.2. Creació d'informes personalitzats**

A part dels informes predefinitos, l'analista podrà fer peticions d'informació a mida amb la utilització d'eines OLAP (On-Line Analytical Processing).

OLAP és una categoria de tecnologia de programari que permet als analistes, gestors i executius de millorar-ne el coneixement de les dades mitjançant l'accés ràpid, consistent i interactiu en una àmplia varietat de possibles vistes d'informació que ha estat transformada des de les dades operacionals per a reflectir la dimensionalitat real de l'empresa com l'entén l'usuari.

Les eines OLAP estan pensades per a ser utilitzades per personal no informàtic. Són senzilles, entenedores i fan èmfasi en la presentació dels resultats. Mitjançant el model multidimensional (molt proper a la manera d'entendre el negoci d'aquest tipus d'usuaris), aconseguen reflectir la complexitat que hi ha en les estructures i relacions de la vida real.

En aquest projecte l'eina utilitzada és Oracle Discoverer.

### 2.7. Diagrama del model conceptual

A continuació es mostra el diagrama del que seran les entitats principals del sistema.

#### 2.7.1. Model relacional

En funció de la informació extreta del fitxer Excel d'origen i dels requeriments especificats pels usuaris, es proposa el model conceptual per a la BD Relacional. És el següent:

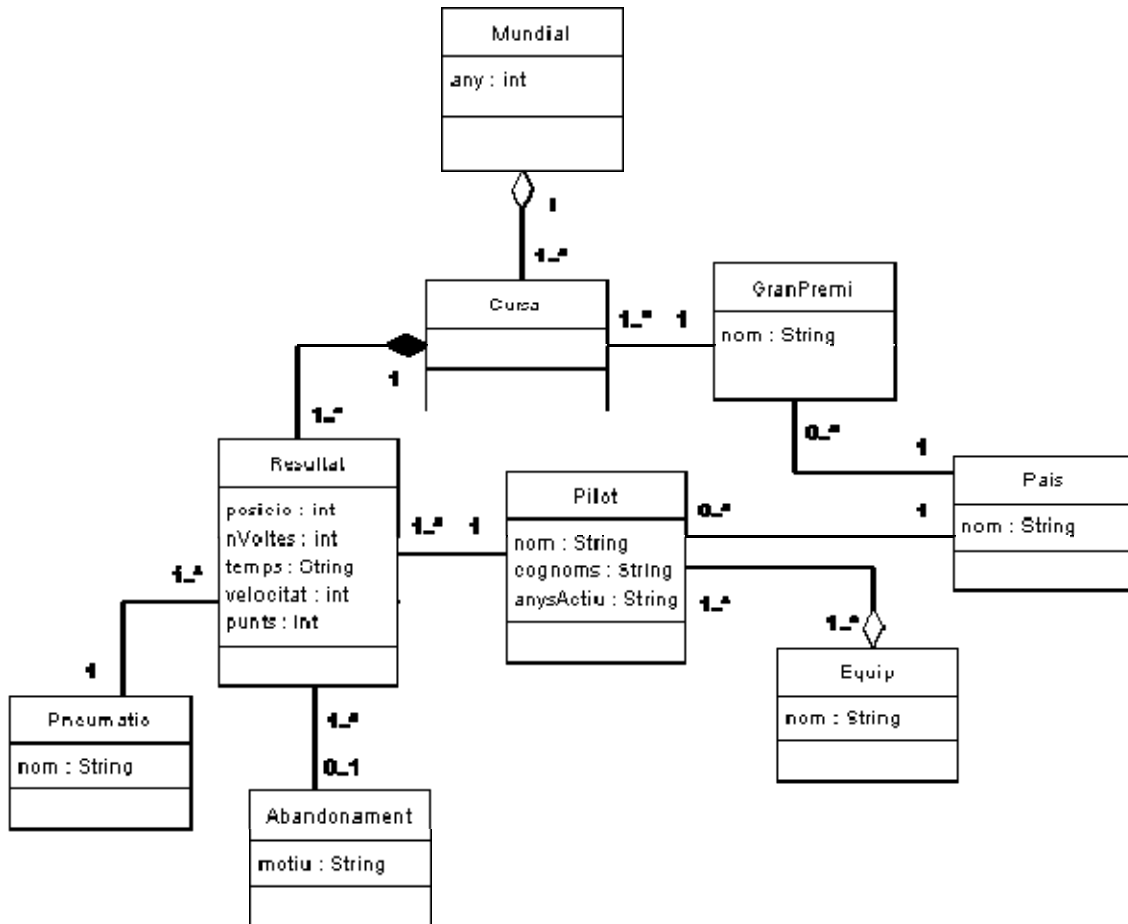
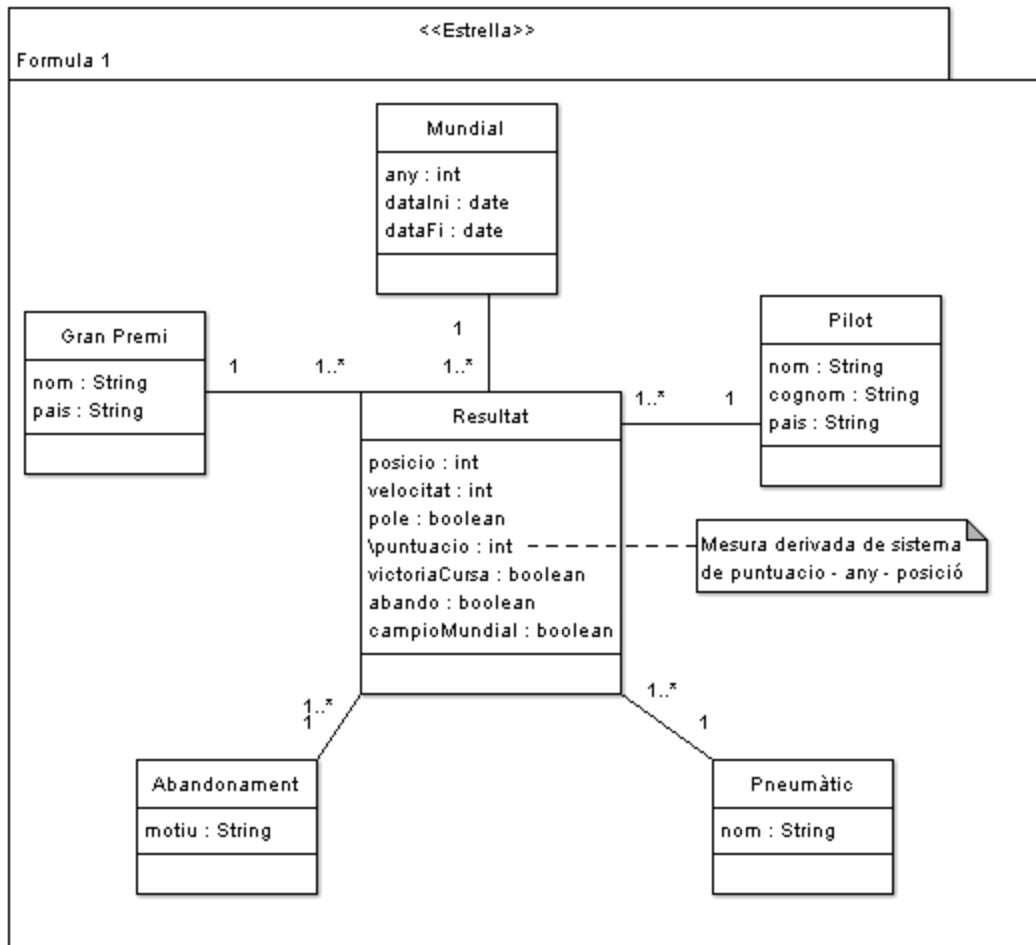


Figura 6. Model conceptual BD relacional

### 2.7.2. Model dimensional

El magatzem de dades inclou només la informació que serà utilitzada pels analistes per a realitzar l'exploració estadística, segons el nivell de detall definit en els requeriments. D'aquesta manera s'optimitza el temps de resposta a les consultes demanades.

El model presentat és un esquema en estrella amb una taula de fets (Resultat) que conté els valors per a l'anàlisi, envoltada de les taules de dimensions.



D:Dimensió N: Nivell F:Fet

Figura 7. Diagrama estrella Magatzem de Dades

Les dimensions Abandonament i Pneumàtic tindran un valor "comodí" per als casos en que no hi hagi informació concreta. Les altres dimensions tindran sempre un valor definit.

Les dimensions representen les perspectives d'anàlisi dels fets que es volen estudiar. Una dimensió pot tenir diferents graus d'agregació, cada un dels quals defineix el nivell de detall de la informació.

Troben les següents:

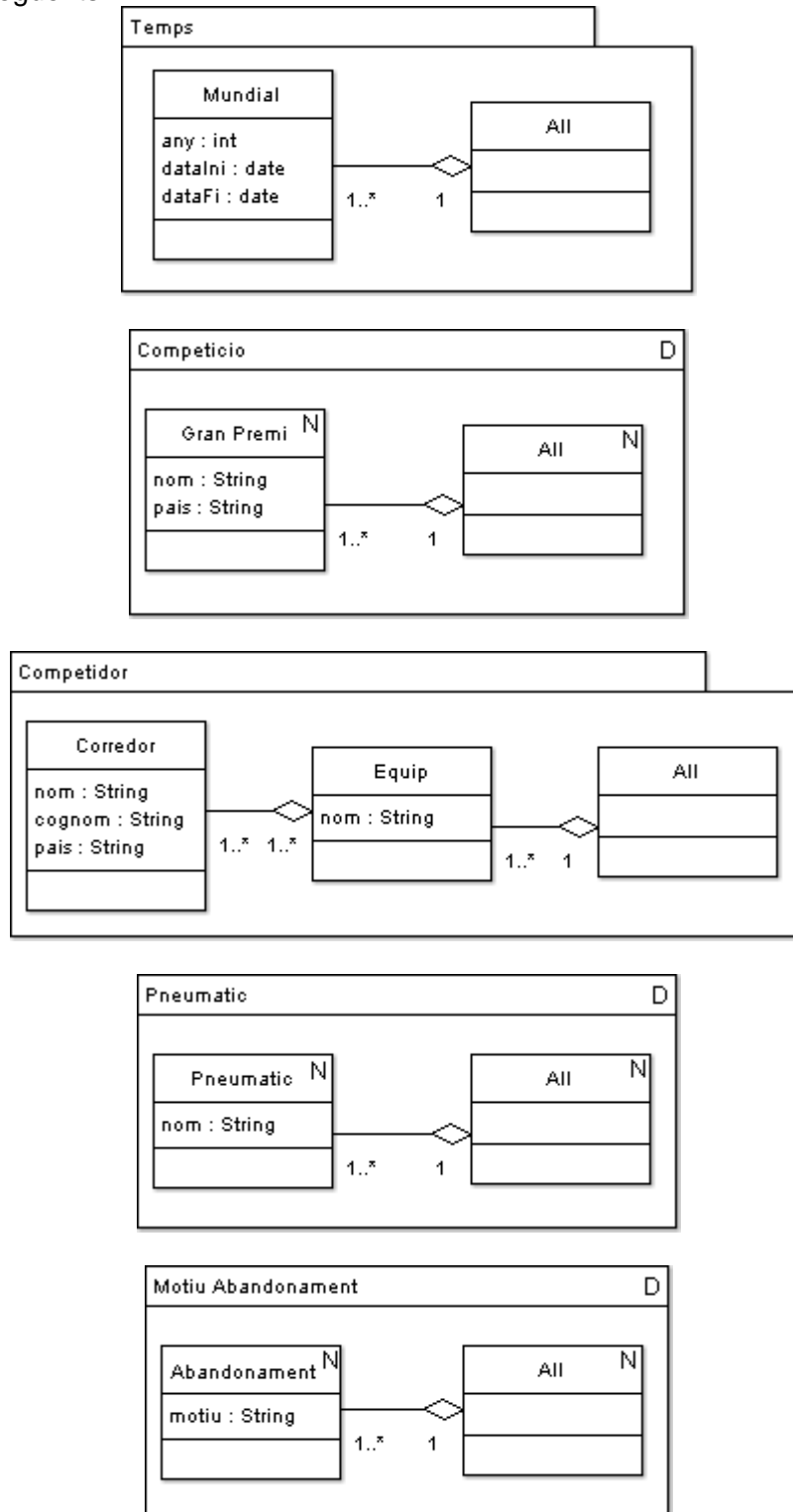


Figura 8. Model conceptual dimensions

### 3. DISSENY

#### 3.1. Diagrama de l'arquitectura software

Com s'ha mencionat anteriorment, un cop implementat el nou sistema, es recomana prescindir dels fitxers Excel i realitzar el manteniment de la informació directament sobre la nova BD Relacional.

L'arquitectura final del sistema seguint aquesta recomanació quedaria com es mostra en la **figura 9**.

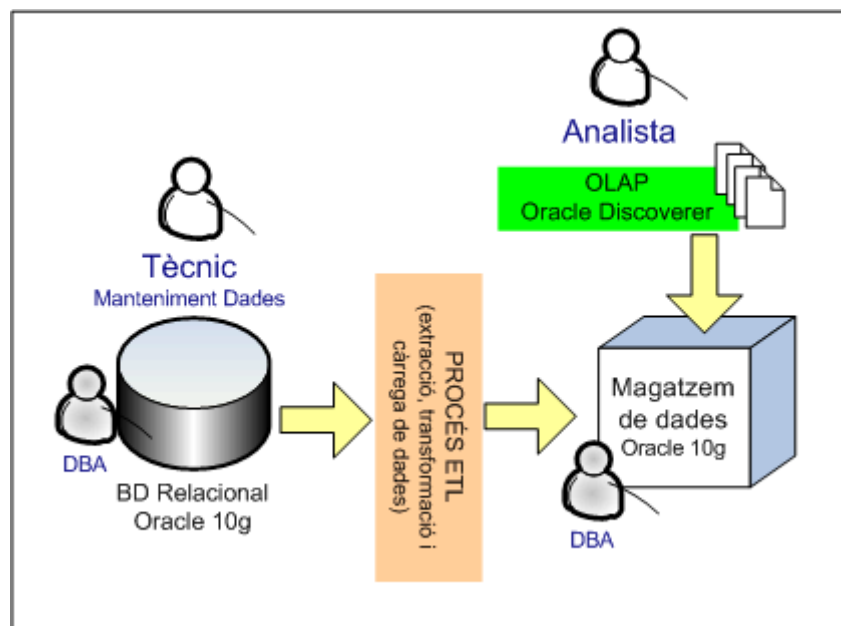


Figura 9. Arquitectura final de programari

D'aquesta manera l'ICEM disposarà d'una base de dades on realitzar l'activitat del dia a dia i que, per tant, estarà sempre actualitzada.

Aquesta base de dades alimentarà periòdicament la informació del magatzem mitjançant processos d'extracció, transformació i càrrega de dades (ETL).

L'usuari analista podrà consultar les dades del magatzem a través de l'eina d'exploració preparada amb aquesta finalitat.



### 3.2. Diagrama de l'arquitectura hardware

Aquest projecte s'inicia sense unes especificacions concretes de maquinari, per tant, s'ha suposat que l'entorn tecnològic de l'ICEM disposa com a mínim d'un servidor de Bases de Dades i d'una Xarxa d'Àrea Local (LAN) que connecta les diferents estacions de treball dels usuaris.

En la **figura 10** es pot veure un esquema d'aquesta arquitectura de maquinari.

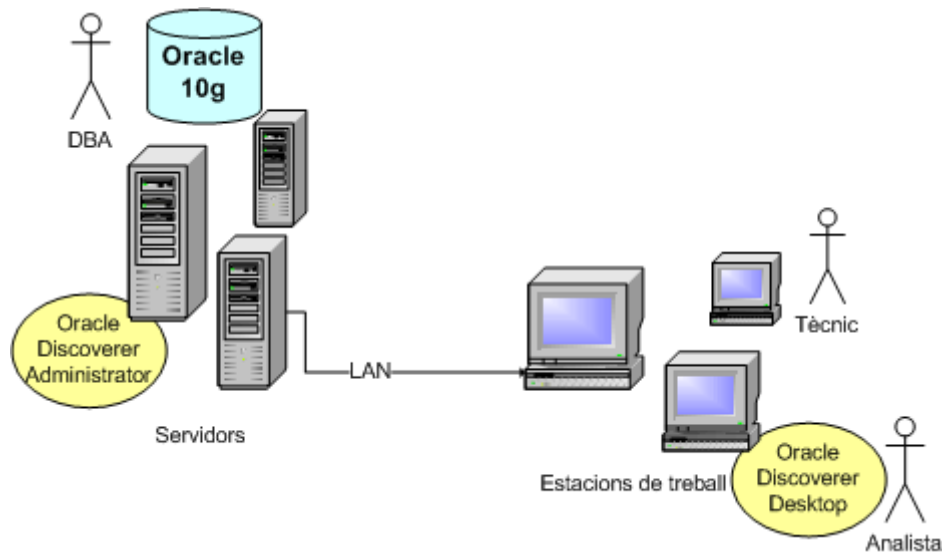


Figura 10. Arquitectura de maquinari

A continuació es defineixen les característiques mínimes que s'han de complir pel correcte funcionament del nou sistema.

#### Servidor de Bases de Dades:

- Sistema Operatiu: Microsoft Windows 2000 Server (o superior)
- Memòria RAM 1GB
- Espai disc dur 5GB
- Còpies de seguretat setmanals
- Sistema de gestió de bases de dades Oracle 10g
- Oracle Discoverer

#### Estacions de treball:

- Pentium II a 266 Mhz amb 64Mb de memòria RAM
- Sistema Operatiu: Windows 98 (o superior)
- Oracle Discoverer Desktop (eina d'anàlisi i *reporting*)

### 3.3. Disseny de la base de dades i diagrama del model físic

#### 3.3.1. Base de dades relacional

El model físic de la BD relacional és el següent:

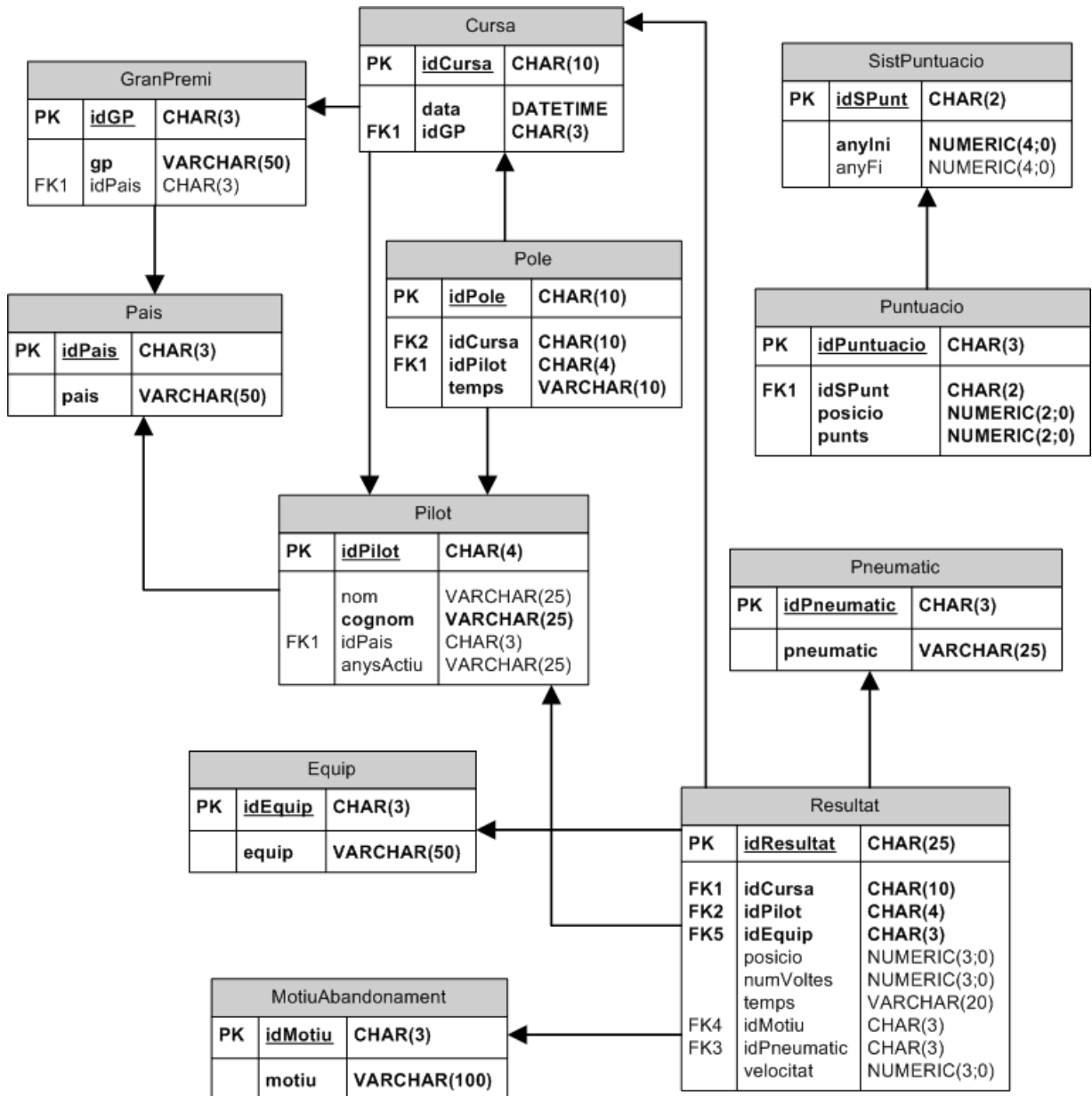


Figura 11. Model físic BD relacional

**PK** indica que és la clau primària de la relació (primary key)  
**FK** indica que l'atribut és clau forana (foreign key)  
 Els atributs en negreta són requerits. La resta accepten valors nuls.

**Informació addicional**

- Un constructor (o equip) està format per dos corredors l'any
- Es disputa un Mundial de Pilots a l'any (guanya el corredor amb més punts totals)
- Es disputa un Mundial de Constructors a l'any (guanya l'equip amb més punts, sumant el total dels seus dos corredors)
- L'identificador de cada registre s'assigna automàticament sense intervenció de l'usuari.
- Es creen les vistes: V\_PilotEquip, V\_Mundial i V\_Resultat per facilitar la càrrega de dades al magatzem.
- Per simplificar el diagrama s'han exclòs les metadades d'entitat: data alta i data modificació. Aquesta informació és bàsica en el procés d'integració i s'actualitzarà automàticament cada cop que es facin modificacions de registres. No requereixen un esforç de manteniment per part de l'usuari.

**3.3.2. Magatzem de dades**

En l'etapa de disseny físic es transforma el model conceptual amb l'objectiu d'aconseguir una major eficiència i, un bon rendiment del magatzem de dades.

El disseny d'esquema en estrella que s'ha utilitzat, permetrà implementar la funcionalitat d'una base de dades multidimensional utilitzant una clàssica base de dades relacional.

Aquest model es caracteritza per la desnormalització de la informació amb la conseqüent aparició de redundància. Aquest fenomen millora la velocitat de resposta a les consultes sobre dades massives.

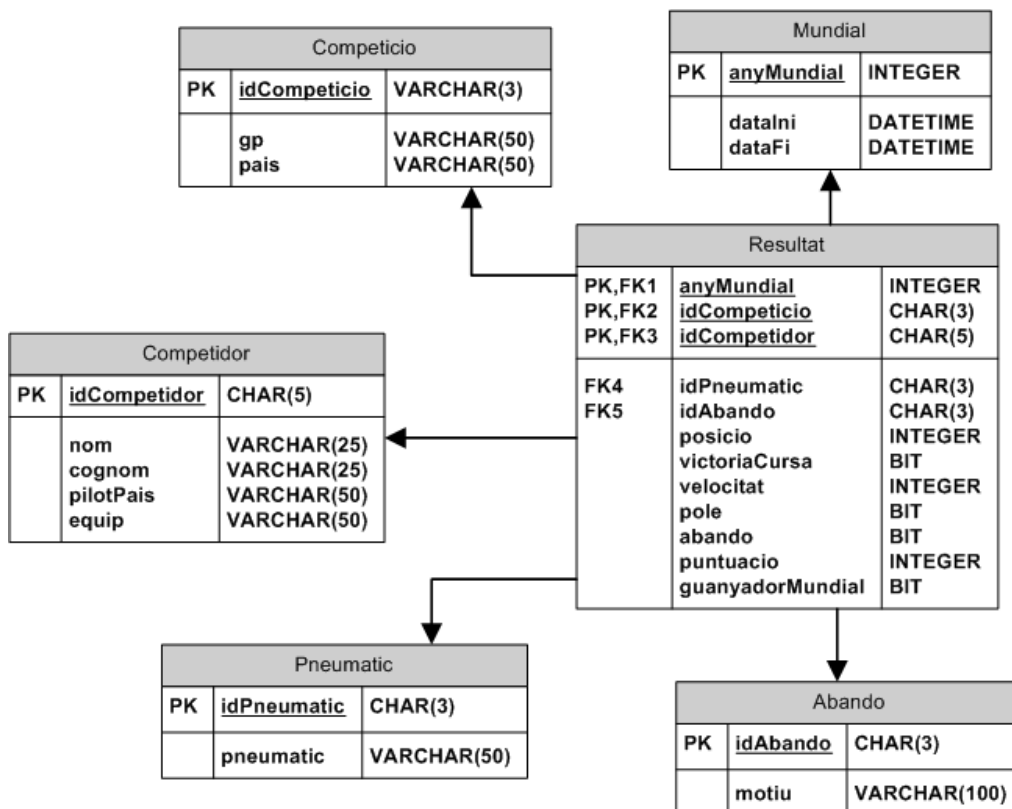


Figura 12. Model físic Magatzem de Dades

### Informació addicional

- Es crea la vista V\_Resultat per facilitar la càrrega de dades al magatzem.
- Per simplificar el diagrama s'ha exclòs del model la metadada d'entitat: data d'importació. Aquest atribut manté actualitzada la data de la darrera modificació de cada registre, incloent la inserció inicial. Aquesta informació és necessària per poder realitzar la càrrega automàtica del magatzem a partir de la BD Relacional.

### 3.4. Arquitectura i disseny dels processos de càrrega

En aquest apartat es descriuen els processos que s'han dut a terme per tal de carregar les dades des del sistema original de l'ICEM fins al magatzem de dades, inicialment buit. I també es descriuen els processos de manteniment periòdic d'aquesta informació.

A grans trets el que fem és crear una base de dades transitòria on s'importen les dades de l'Excel tal qual estan. Aquestes dades es depuren i es transformen convenientment per adaptar la informació a l'estructura normalitzada de la BD relacional. Finalment es creen els processos de càrrega al magatzem a partir de la BD relacional.

#### 3.4.1. Extracció de les dades d'origen

Per poder disposar de les dades d'origen en la plataforma Oracle s'ha creat una base de dades temporal formada per tres taules independents i sense cap mena de restricció d'integritat.

El model de base de dades temporal es pot veure en la **figura 13**.

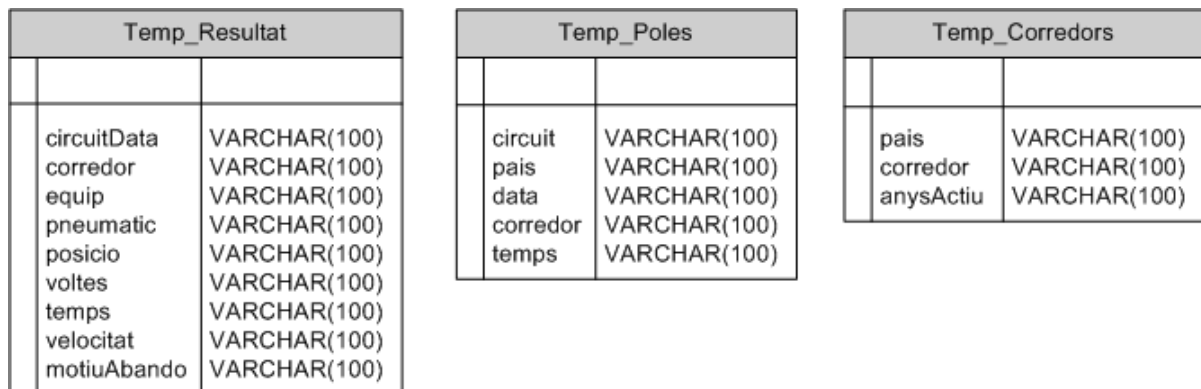


Figura 13. Model físic BD temporal

Aquestes taules equivalen als tres fulls del fitxer Excel proporcionat per l'ICEM i, es carreguen inicialment amb totes les dades contingudes en ell mitjançant la utilitat **SQL\*Loader** d'Oracle (*per a més informació veure annex 2*).

Un cop realitzat el procés per les tres taules, el número de registres que s'han carregat a la base de dades és el següent:

Taula	Núm. Registres
Temp_Resultat	15784
Temp_Poles	728
Temp_Corredors	423

### 3.4.2. Càrrega inicial a la BD Relacional

Un cop creada la base de dades relacional, s'han d'importar les dades contingudes a la BD temporal.

Per poder realitzar aquest pas, primer cal adaptar la informació d'origen a l'estructura normalitzada de la BD relacional. Concretament i, mitjançant instruccions amb llenguatge PL/SQL, es realitzen les tasques següents:

- Transformació de les dades d'origen per adaptar-les a l'estructura de l'esquema de la BD relacional on s'emmagatzemaran.
- Depuració dels errors o conflictes que s'han pogut trobar dins les dades de cada entitat.
- Integració de les dades depurant errors o conflictes localitzats entre dades d'entitats diferents.
- Les taules Puntuació i Sistema de Puntuació s'omplen manualment segons la informació facilitada per l'ICEM.

Finalment, obtenim els resultats següents:

Taula	Núm. registres
Cursa	690
Equip	95
Granpremi	55
Motiuabandonament	132
Pais	36
Pilot	546
Pneumatic	7
Pole	688
Puntuacio	32
Resultat	15690
Sistpuntuacio	5

**Nota:** En l'annex 3 hi ha una mostra d'instruccions PL/SQL de càrrega a la BD. L'annex 4 detalla el tractament realitzat sobre les dades originals i, l'annex 5 informa dels registres que, degut a incoherències en la informació, no s'han pogut importar a la BD Relacional.

### 3.4.3. Càrrega inicial al Magatzem de Dades

La càrrega inicial del magatzem es duu a terme a partir de la BD Relacional que ja disposa de dades depurades i, per tant, no cal fer un tractament previ.

Concretament es realitza mitjançant crides PL/SQL amb les peculiaritats següents:

- S'insereix el valor comodí ('N/A') amb identificador zero a les dimensions pneumàtic i motiu abandonament pels casos que aquesta informació no està disponible.
- Es realitza el càlcul de puntuació obtinguda per cada pilot en cada cursa segons la posició d'arribada i, en funció del sistema de puntuació aplicable per l'any de la competició. Per facilitar la tasca, la vista V\_Resultat de la BD Relacional incorpora aquest valor ja calculat.
- En la taula de fets (Resultat) no hi pot haver valors nuls. En cas de no disposar d'informació sobre posició, velocitat o puntuació, s'inclourà el valor -1. Els camps booleans tindran el valor 'fals' per defecte.
- Es crea un procediment emmagatzemat per omplir els valors 'cert' dels indicadors booleans de la taula de Fets.
- Com s'ha comentat anteriorment, s'informa automàticament la data d'importació de cada registre.

Finalment, obtenim els resultats següents:

Taula	Núm. Registres
Abando	133
Competicio	55
Competidor	1062
Mundial	47
Pneumatic	8
Resultat	15690

**Nota:** En l'annex 6 hi ha una mostra d'instruccions PL/SQL de càrrega inicial al magatzem.

### 3.4.4. Manteniment de la informació del magatzem

Un cop carregat el magatzem inicialment buit, aquest s'haurà d'anar actualitzant periòdicament a partir de la BD relacional.

Aquest procés es realitzarà entre bases de dades en el Servidor de bases de d'Oracle, i s'utilitzaran procediments emmagatzemats per tal de millorar el rendiment i automatitzar la seva execució en moments determinats de poc ús del servidor.

Les metadades de la font d'origen (BD Relacional) emmagatzemen per a cada registre el moment en què es va dur a terme la darrera modificació. Així, obtenir les últimes actualitzacions és immediat, n'hi ha prou d'obtenir els registres marcats amb una empremta de temps posterior a la de l'últim conjunt de canvis ocorregut.

És a dir, per a realitzar el manteniment de la informació, s'executarà automàticament un procés que afegirà els nous registres creats i, aquells que han estat modificats, amb data posterior a la darrera actualització del magatzem.

A continuació s'enumeren algunes característiques dels procediments:

- Les dimensions Abandó, Mundial, Pneumàtic i Competició actualitzen la informació de cada registre cada cop que es realitza la càrrega de dades ja que són taules amb pocs valors i és menys costós refrescar tota la informació que mirar si s'ha modificat. Aquests procediments també incorporen els registres nous.
- En la dimensió Competidor i en la taula de fets, es comprova si hi ha hagut modificacions o insercions. Llavors es modifiquen només els registres que hagin variat i, s'insereixen els nous. Concretament, es compara la data de modificació de la BD Relacional amb la data d'importació del magatzem i, si és més recent, llavors es modifica o s'incorpora la informació al magatzem
- Es crea un procediment emmagatzemat per cada dimensió del magatzem. D'aquesta manera es pot refrescar la informació de cada dimensió independentment de la resta mitjançant la crida manual a un procés determinat.
- Es creen dos procediments per actualitzar la informació de la taula de Fets. El primer actualitza la informació de la taula i el segon s'encarrega de posar els valors 'cert' als atributs booleans guanyadorMundial, victoriaCursa i abando.
- Finalment, es crea un procediment genèric que crida a tots els demés en un ordre determinat. Aquest serà el que s'executi periòdicament per actualitzar tot el magatzem.

La freqüència d'actualització del magatzem serà d'una setmana, tenint en compte que durant el període de competició el temps mínim que pot passar entre dues curses consecutives és de 7 dies. I, tindrà lloc en hores de poca activitat per no alterar el rendiment dels treballs habituals.

Concretament, s'ha creat un procés de refresc de dades que s'inicia automàticament tots els dimarts a la 1am entre els mesos de març i novembre, corresponent al calendari de competició de la Fórmula1.

A més, en cas necessari, es pot executar el procés manualment en qualsevol moment.

**Nota:** En l'annex 7 hi ha una mostra dels processos creats per actualitzar les dades del magatzem. L'annex 8 presenta el codi d'automatització de la càrrega periòdica.

### 3.5. Disseny i descripció del informes creats

A continuació es descriuen els llibres de treball creats amb Oracle Discoverer Desktop a partir de l'àrea de negoci 'Fórmula1' feta amb Oracle Discoverer Administrator.

#### 3.5.1. Rànquing de victòries per Mundial

Informe que genera un llistat amb el número de Gran Premis guanyats per pilot, per tipus de pneumàtic o per constructor, en un Mundial determinat.

Informació prèvia a concretar:

- Mundial (any)
- Concepte (pilot, tipus de pneumàtic o constructor)

Dades que es visualitzaran en el llistat:

- Dades concepte
- Número total de victòries

Ordre de la llista: Descendent per total de grans premis guanyats, el concepte amb més victòries apareix en primer lloc. És a dir, es mostra el rànquing de corredors, pneumàtics o equips d'un Mundial concret.

#### 3.5.2. Rècords de títols

Informe que genera un llistat de número de Mundials guanyats per pilot, per tipus de pneumàtic o per constructor, al llarg de la història de la Fórmula 1.

Informació prèvia a concretar:

- Concepte (pilot, tipus de pneumàtic o constructor)

Dades que es visualitzaran en el llistat:

- Dades concepte
- Número total de victòries

Ordre de la llista: Descendent per total de mundials guanyats, el concepte amb més victòries apareix en primer lloc.

#### 3.5.3. Posició mitjana de pilots per Mundial

Informe que genera un llistat de posició mitjana obtinguda per cada pilot durant el desenvolupament d'un determinat Mundial.

Fórmula: La posició mitjana es calcula com la suma de posicions obtingudes per un pilot en cada un dels Grans Premis del Mundial dividit pel número de curses disputades.

Informació prèvia a concretar:

- Mundial (any)

Dades que es visualitzaran en el llistat generat:

- Dades pilot
- Posició mitjana
- Número de curses disputades

Ordre de la llista: Ascendent per valor de posició, número de curses i nom pilot.



#### 3.5.4. Velocitat mitjana de pilots per GP

Informe que genera un llistat de velocitat mitjana obtinguda per pilot en cada un dels GP disputats durant el desenvolupament d'un determinat Mundial.

Informació prèvia a concretar:

- Mundial (any)

Per cada GP es visualitzaran les dades següents:

- Dades pilot
- Velocitat mitjana

Ordre de la llista: Descendent per valor de velocitat.

#### 3.5.5. Total d'abandonaments per GP

Informe que genera un llistat amb el número d'abandonaments per GP d'un determinat Mundial, classificats per motiu.

Informació prèvia a concretar:

- Mundial (any)

Per cada GP es visualitzaran les dades següents:

- Número d'abandonaments classificats per motiu
- Total d'abandonaments per GP

Ordre de la llista: Descendent per número d'abandonaments.

#### 3.5.6. Rànquing de puntuació per GP

Informe que genera un llistat amb la puntuació obtinguda en cada GP per pilot o per equip, en un Mundial determinat. També es mostrarà la puntuació total.

Escollint el concepte corredor, s'obtindrà la classificació del **Mundial de Pilots**, i pel concepte equip, s'obté la del **Mundial de Constructors**.

Informació prèvia a concretar:

- Mundial (any)
- Concepte (pilot o constructor)

Dades que es visualitzaran en el llistat generat:

- Dades concepte
- Puntuació per GP
- Puntuació total

Ordre de la llista: Descendent per puntuació total.

### 3.5.7. Pilots de Mundial

Informe que genera un llistat dels pilots que han participat en un determinat Mundial, seguit d'una sèrie d'informació relacionada.

Informació prèvia a concretar:

- Mundial (any)

Dades que es visualitzaran en el llistat generat:

- Posició en la classificació del Mundial
- Dades pilot
- Nom Equip
- Número de curses realitzades
- Número de victòries
- Número de poles aconseguides
- Puntuació total

Ordre de la llista: Descendent per puntuació.

### 3.5.8. Campionat del món de pilots

Informe que genera un llistat dels pilots que han guanyat el Mundial de pilots, al llarg de la història de la Fórmula 1.

Dades que es visualitzaran en el llistat generat:

- Any mundial
- Dades pilot
- Nom Equip

Ordre de la llista: Descendent per any mundial.

**Nota:** Es poden veure captures de pantalla dels diferents informes creats a l'apartat 5.4 d'aquest document.

#### 4. ANÀLISI DE LA INFORMACIÓ: RELACIÓ POLE - VICTÒRIA

Com a darrer punt dins del desenvolupament del nou sistema, es demana un exercici d'anàlisi que expliqui la relació existent entre el fet d'obtenir la *Pole Position* i guanyar la cursa i, en justifiqui la conclusió donada.

Per conèixer la relació entre aconseguir la Pole i guanyar la mateixa cursa, realitzem un informe on es visualitza el percentatge de curses guanyades amb Pole aconseguida respecte el % de curses guanyades sense Pole.

Per tal de tenir una visió més global, realitzem el càlcul per anys concrets, per franges d'anys i pel global de dades que tenim.

**Històric (1961-2007):**



**Any 2007:**



Executant l'informe per diferents períodes, obtenim els resultats següents:

Període	% Victòries amb Pole	%Victòries sense Pole
2007	59%	41%
2006	50%	50%
2005	42%	58%
2000-2007	48%	52%
1990-2000	42%	58%
1980-1990	30%	70%
1970-1980	35%	65%
1961-1970	31%	69%
<b>1961-2007</b>	<b>38%</b>	<b>62%</b>

Per les dades obtingudes podem veure que, al llarg de la història de la Fórmula 1, aproximadament un 40% de Poles aconseguides per un pilot van seguides de la seva victòria en la cursa.

És a dir, si una cursa la disputen de mitjana uns 22 corredors, tenim que d'entrada tots tenen un 4,5% (100/22) de probabilitats de guanyar la cursa. Ara bé, si el corredor que aconseguix la Pole estadísticament guanya un 40% dels cops, llavors cada un dels corredors restants té només un 2,85% (60/21) de probabilitats de ser campió. Dit d'una altra manera, **el corredor que aconseguix la Pole té, de mitjana, 14 vegades més opcions de guanyar la cursa que la resta de competidors.**

Això dona una clara avantatge al corredor que aconseguix la Pole i, per tant, es tracta d'una dada molt significativa.

També podem observar que el percentatge va augmentant amb els anys, de manera que actualment hi ha més possibilitats de guanyar una cursa si s'ha aconseguit la Pole, que als inicis de la Fórmula 1.

Finalment, si mirem els resultats en funció del circuit, ens adonem que en algun d'ells, aconseguir la Pole garanteix la victòria o augmenta encara més les possibilitats de guanyar la cursa.

#### Resultats 2000-2007 per GP:

% Poles respecte Victòries aconseguides Inici: '2000', Fi: '2007'			
	Victòria	Pole en Cursa Guanyada	% poles en cursa guanyada
▶ Istanbul	3	3	100,00
▶ Mount Fuji	1	1	100,00
▶ Barcelona	8	7	87,50
▶ Bahrain	4	3	75,00
▶ Suzuka	7	5	71,43
▶ Monza	8	5	62,50
▶ Sepang	8	5	62,50
▶ A1-Ring	4	2	50,00
▶ Hungaroring	8	4	50,00
▶ Melbourne	8	4	50,00
▶ Shanghai	4	2	50,00
▶ Spa-Francorchamps	6	3	50,00
▶ Hockenheim	7	3	42,86
▶ Imola	7	3	42,86
▶ Indianapolis	8	3	37,50
▶ Magny-Cours	8	3	37,50
▶ Monte Carlo	8	3	37,50
▶ Montreal	8	3	37,50
▶ Nurburgring	8	2	25,00
▶ Silverstone	8	2	25,00
▶ Interlagos	8	1	12,50

## Resultats històrics per circuit:

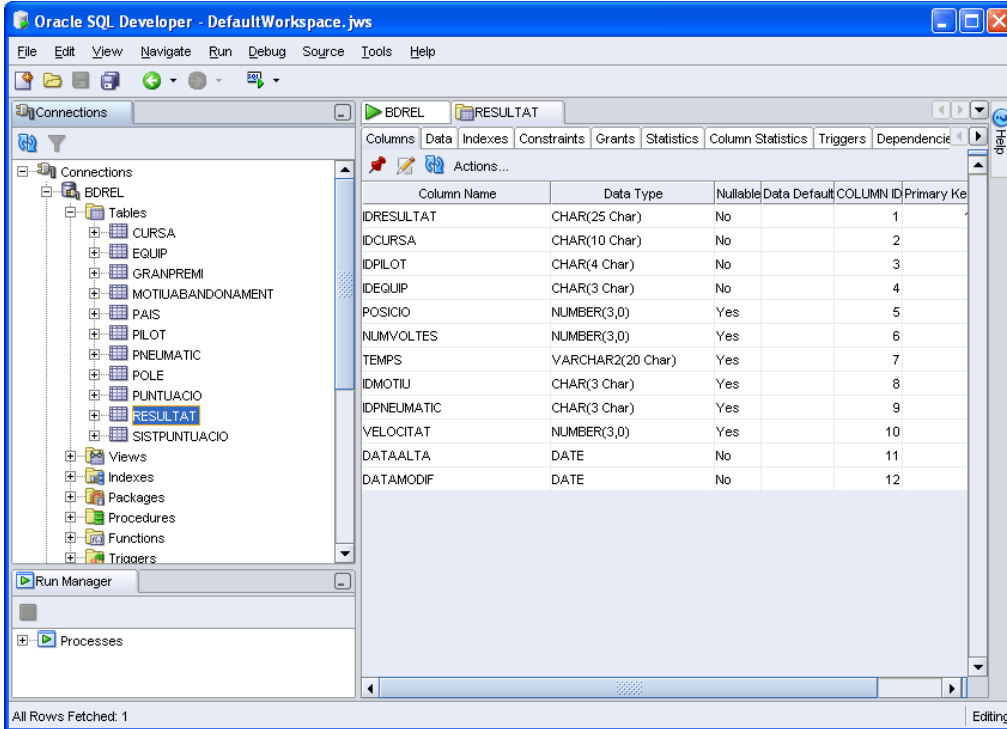
**% Poles respecte Victòries aconseguides per GP  
(en la Història del Campionat)**

	Victòria	Pole en Cursa Guanyada	% poles en cursa guanyada
› Istanbul	3	3	<b>100,00</b>
› Barcelona	17	13	<b>76,47</b>
› Bahrain	4	3	<b>75,00</b>
› East London	3	2	<b>66,67</b>
› Mount Fuji	3	2	<b>66,67</b>
› Paul Ricard	14	8	<b>57,14</b>
› Sepang	9	5	<b>55,56</b>
› Aintree	2	1	<b>50,00</b>
› Clermont-Ferrand	4	2	<b>50,00</b>
› Hockenheim	30	15	<b>50,00</b>
› Long Beach	8	4	<b>50,00</b>
› Mosport Park	8	4	<b>50,00</b>
› Nivelles	2	1	<b>50,00</b>
› Shanghai	4	2	<b>50,00</b>
› Watkins Glen	20	10	<b>50,00</b>
› Zolder	10	5	<b>50,00</b>
› Buenos Aires	13	6	<b>46,15</b>
› Estoril	13	6	<b>46,15</b>
› Hungaroring	22	10	<b>45,45</b>
› Suzuka	20	9	<b>45,00</b>
› Montreal	29	13	<b>44,83</b>
› Jarama	9	4	<b>44,44</b>
› A1-Ring	7	3	<b>42,86</b>
› Detroit	7	3	<b>42,86</b>
› Jerez	7	3	<b>42,86</b>
› Mexico City	14	6	<b>42,86</b>
› Melbourne	12	5	<b>41,67</b>
› Adelaide	10	4	<b>40,00</b>
› Monte Carlo	46	18	<b>39,13</b>
› Indianapolis	8	3	<b>37,50</b>
› Interlagos	25	9	<b>36,00</b>
› Dijon	6	2	<b>33,33</b>
› Imola	27	9	<b>33,33</b>
› Phoenix	3	1	<b>33,33</b>
› Reims	3	1	<b>33,33</b>

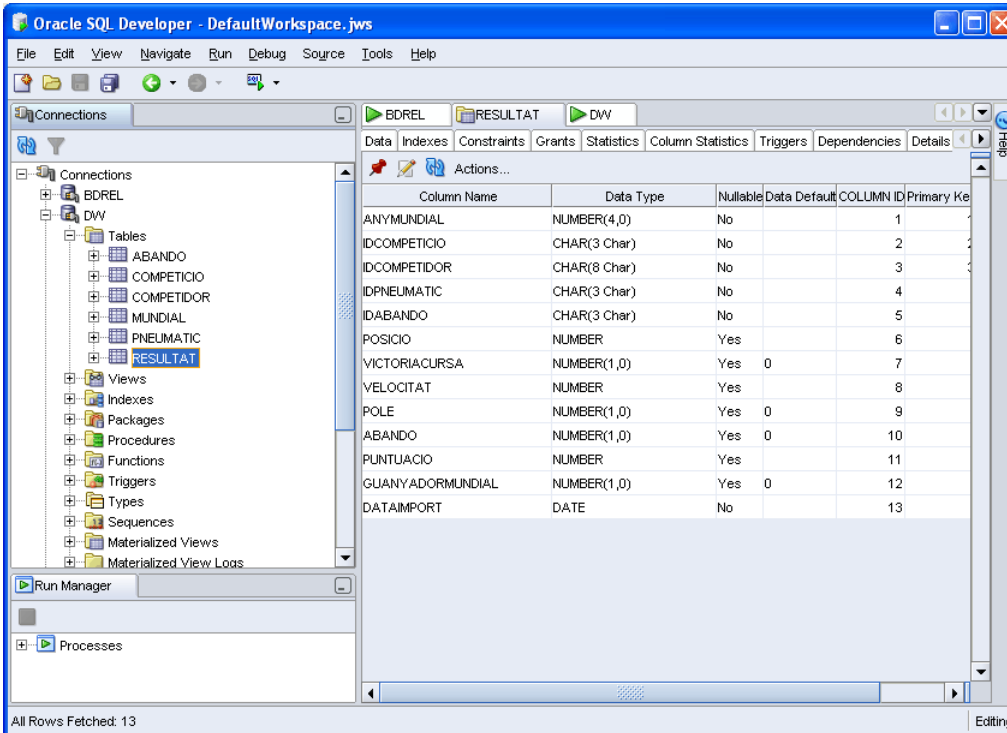
Total
  Per GP
  Anual
  Periode

## 5. CAPTURES DE PANTALLA

### 5.1. BD Relacional

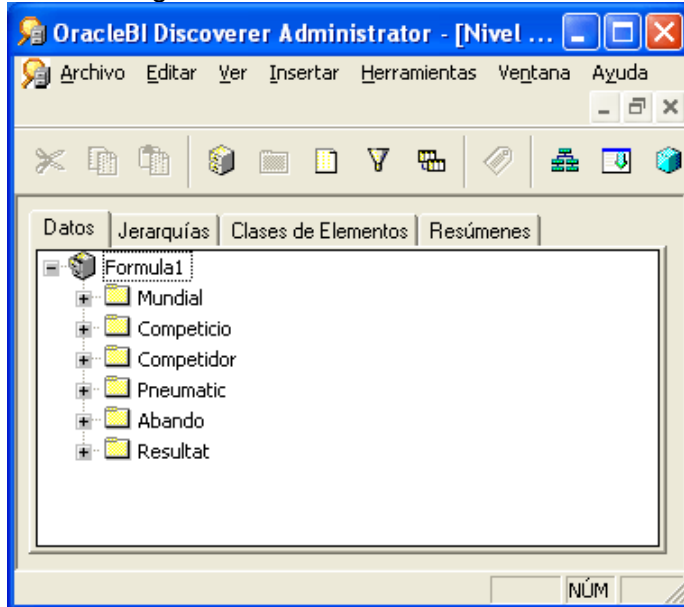


### 5.2. Magatzem de Dades

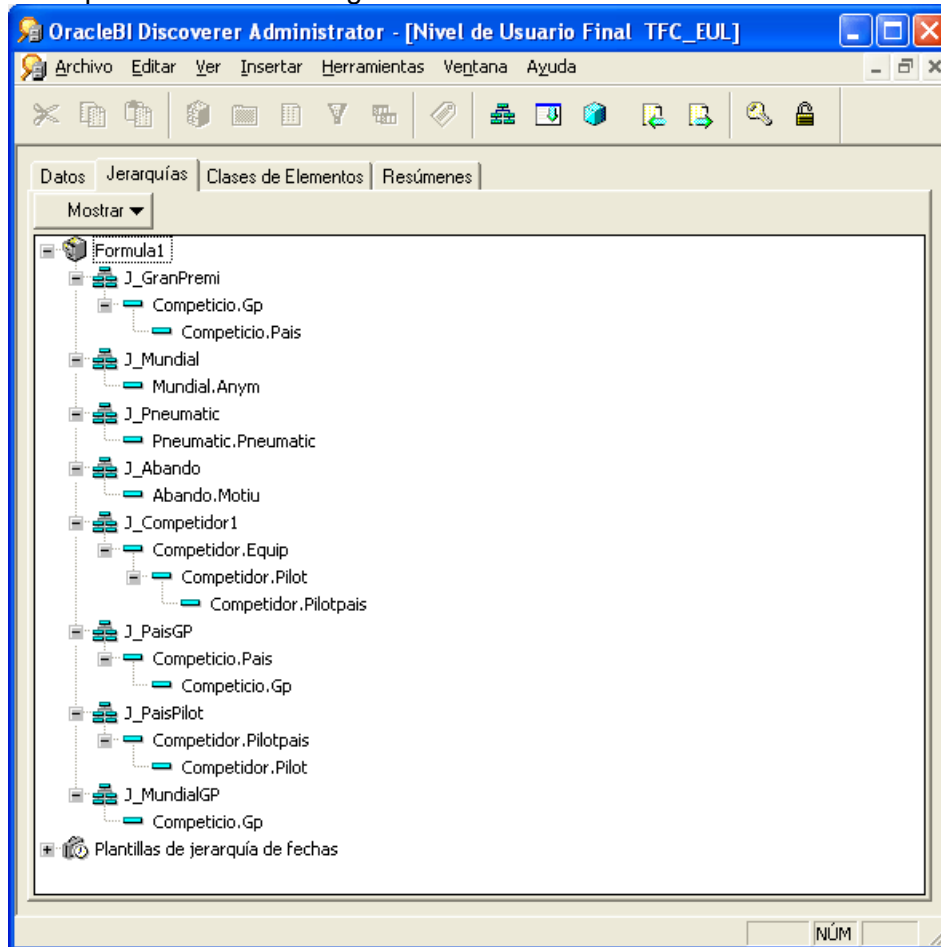


### 5.3. Àrea de negoci

Àrea de negoci 'Fórmula1' creada amb Oracle Discoverer Administrator.



Jerarquies de l'àrea de negoci.

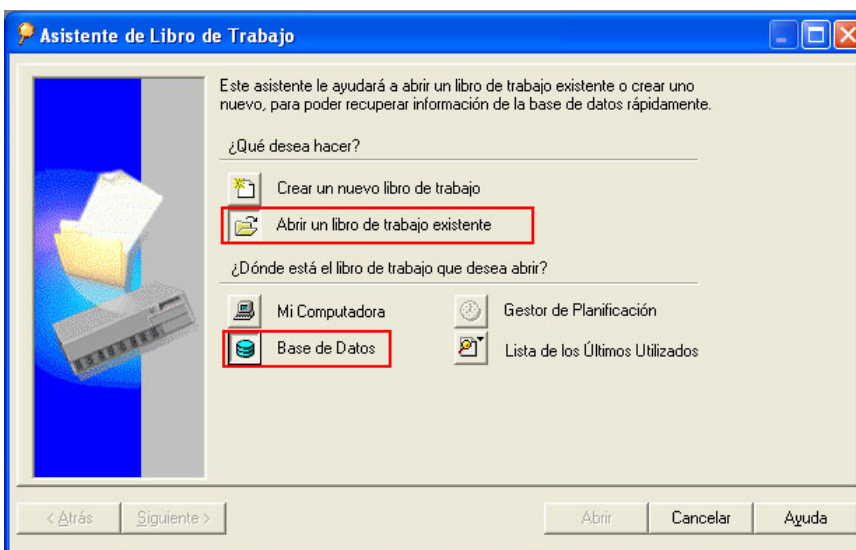


#### 5.4. Informes creats

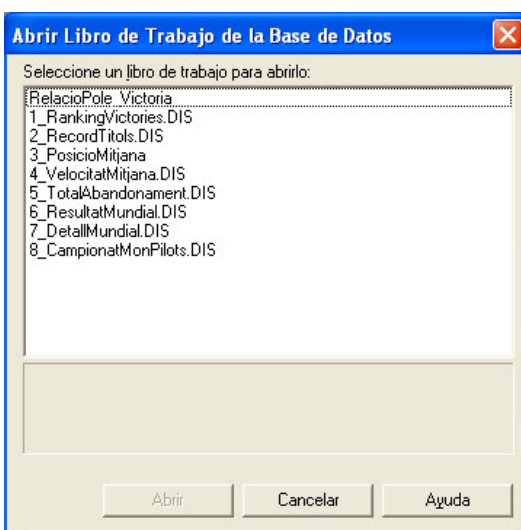
Per a realitzar l'explotació estadística de les dades del magatzem es crea l'usuari específic **TFC\_EUL** (contrasenya: pwdtfc).

L'usuari analista accedirà a la informació del magatzem a través de l'eina **Oracle Discoverer Desktop**, en la qual trobarà diferents llibres de treball desenvolupats segons els requeriments definits.

Un cop executat el programari i introduïdes les dades d'accés, per accedir als informes predefinits s'haurà de seleccionar l'opció 'Abrir un libro de trabajo existente' i després triar 'Base de Datos'.



Aleshores apareix el llistat d'informes existents:



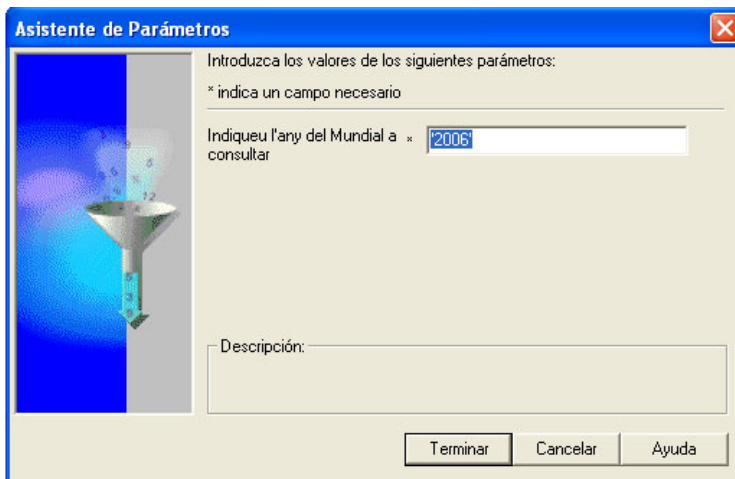


A continuació es mostra el prototip dels informes realitzats.

#### 5.4.1. Rànquing de victòries per Mundial

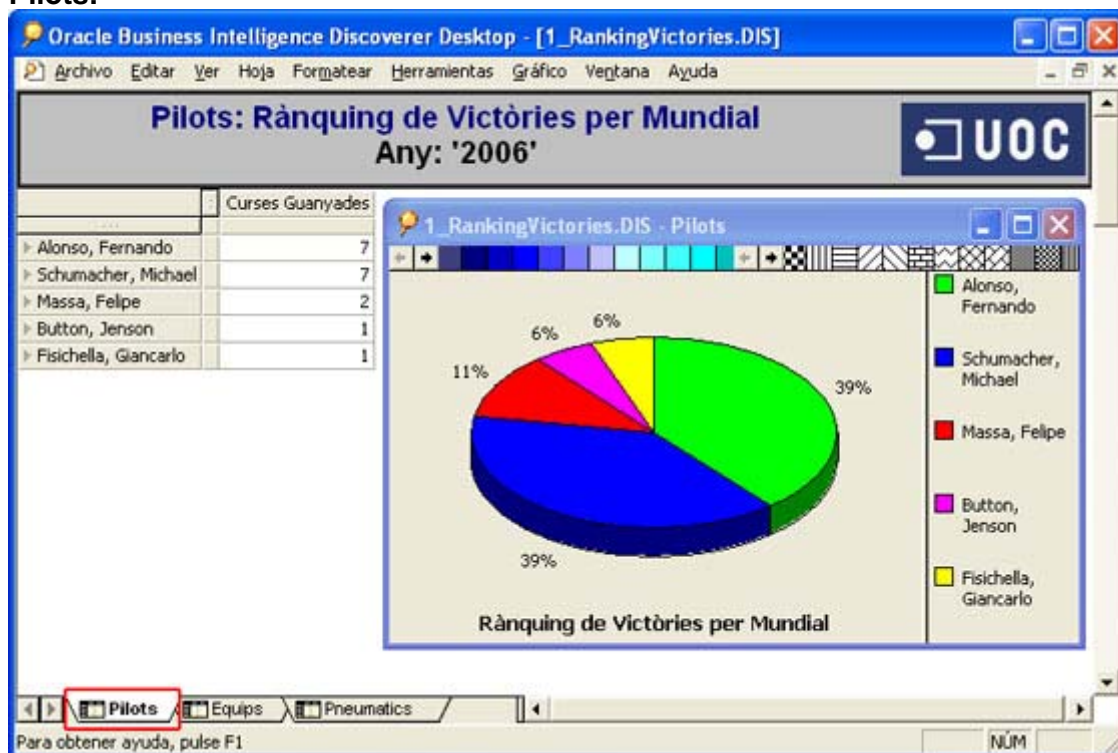
Nom informe: 1\_RankingVictories.DIS

Característiques: En primer lloc s'ha d'indicar l'any del Mundial que es vol consultar.

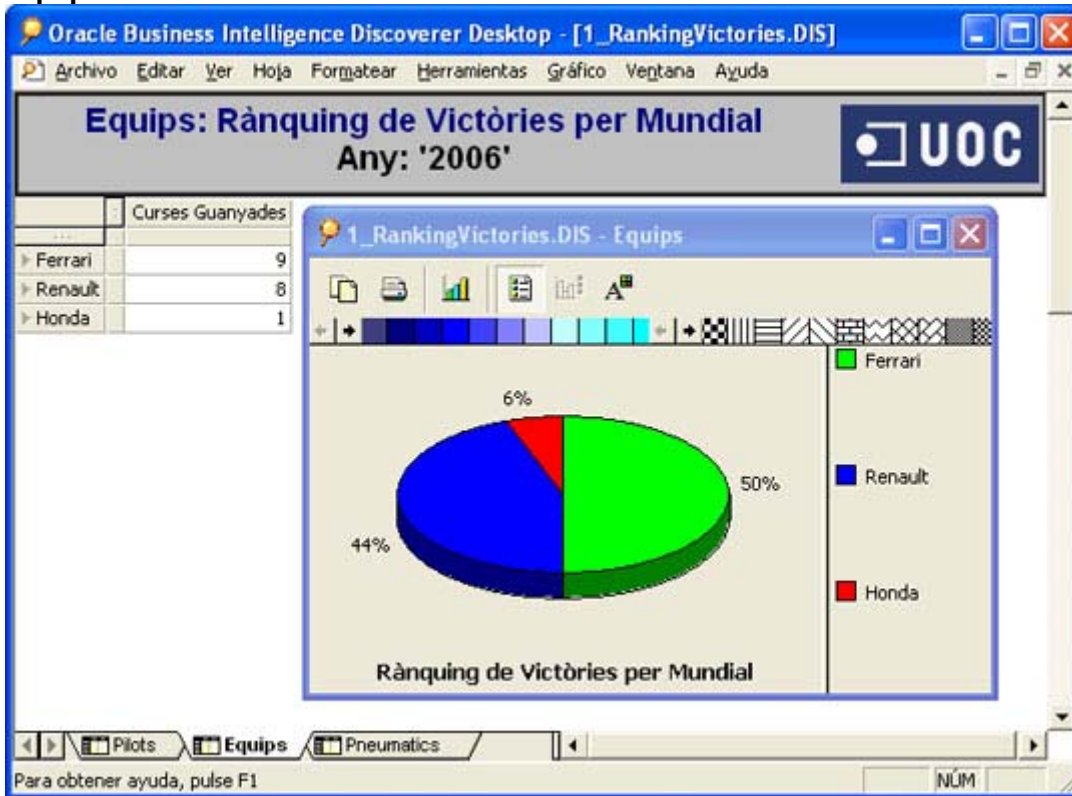


Seguidament, es pot seleccionar una de les tres pestanyes disponibles, corresponents a cada un dels conceptes de l'enunciat.

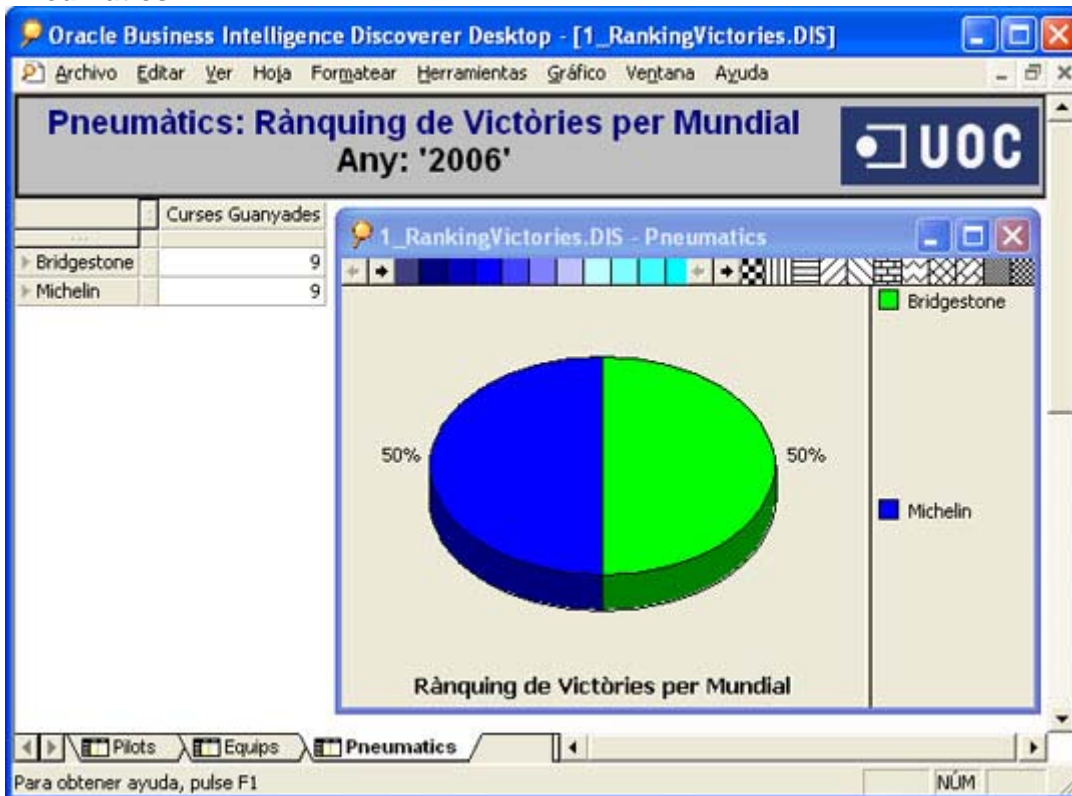
#### Pilots:



**Equips:**



**Pneumàtics:**



### 5.4.2. Rècords de títols

Nom informe: 2\_RecordTitols.DIS

Característiques: En primer lloc s'ha d'indicar l'any del Mundial que es vol consultar i, tot seguit, es pot seleccionar una de les tres pestanyes disponibles, corresponents a cada un dels conceptes de l'enunciat.

#### Pilots:

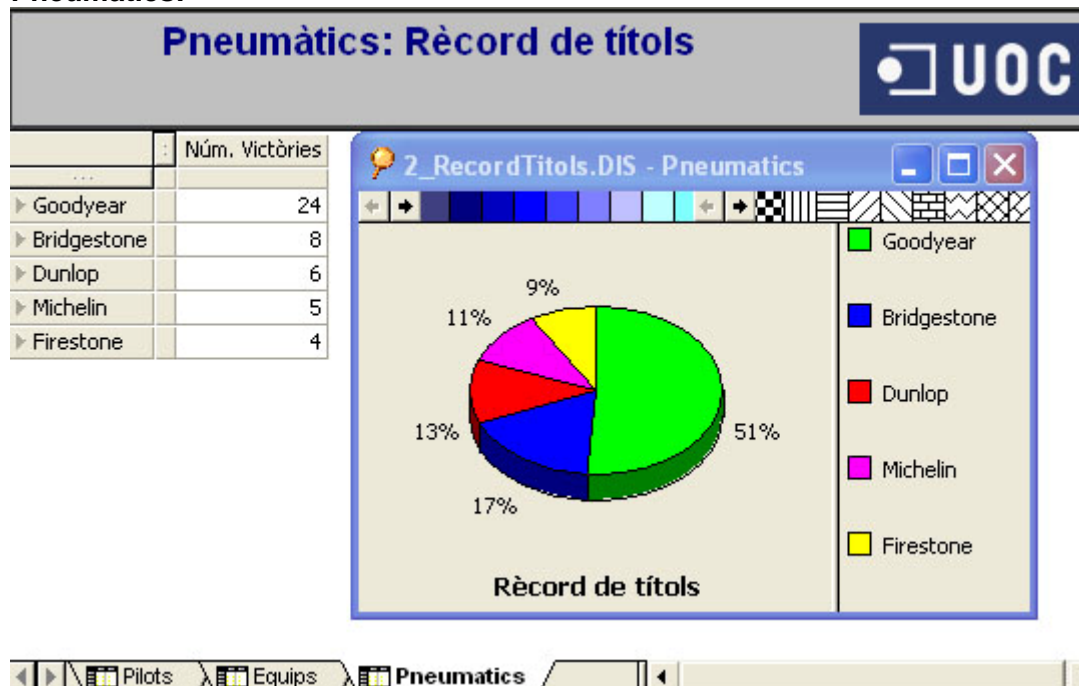
Pilots: Rècord de títols		UOC	
		Núm. Victòries	
▶ Schumacher, Michael		7	
▶ Prost, Alain		6	
▶ Hill, Graham		3	
▶ Piquet, Nelson		3	
▶ Stewart, Jackie		3	
▶ Alonso, Fernando		2	
▶ Clark, Jim		2	
▶ Fittipaldi, Emerson		2	
▶ Hakkinen, Mika		2	
▶ Lauda, Niki		2	
▶ Senna, Ayrton		2	
▶ Andretti, Mario		1	
▶ Brabham, Jack		1	
▶ Hill, Damon		1	
▶ Hill, Phil		1	
▶ Hulme, Denny		1	
▶ Hunt, James		1	
▶ Jones, Alan		1	
▶ Mansell, Nigel		1	
▶ Raikkonen, Kimi		1	
▶ Rindt, Jochen		1	
▶ Rosberg, Keke		1	
▶ Scheckter, Jody		1	
▶ Villeneuve, Jacques		1	

Pilots Equips Pneumatics

**Equips:**



**Pneumàtics:**



### 5.4.3. Posició mitjana de pilots per Mundial

Nom informe: 3\_PosicioMitjana.DIS

Característiques: S'ha d'indicar l'any del Mundial que es vol consultar.

Posició mitjana de pilots per Mundial		
Any: '2007'		
	Posicio Mitjana	Curses disputades
> Raikkonen, Kimi	2,6000	17
> Alonso, Fernando	2,9375	17
> Hamilton, Lewis	3,0000	17
> Massa, Felipe	3,6667	17
> Heidfeld, Nick	5,3333	17
> Kubica, Robert	6,6923	16
> Kovalainen, Heikki	7,8750	17
> Fisichella, Giancarlo	8,1429	17
> Webber, Mark	8,9000	17
> Coulthard, David	9,0000	17
> Rosberg, Nico	9,1429	17
> Wurz, Alexander	9,8182	13
> Nakajima, Kazuki	10,0000	1
> Button, Jenson	10,6364	17
> Trulli, Jarno	10,6667	16
> Schumacher, Ralf	11,1818	17
> Speed, Scott	12,0000	10
> Barrichello, Rubens	12,0667	17

### 5.4.4. Velocitat mitjana de pilots per GP

Nom informe: 4\_VelocitatMitjana.DIS

Característiques: S'ha d'indicar l'any del Mundial que es vol consultar. Hi ha dues pestanyes disponibles, corresponents a velocitat mitjana per pilot en cada GP i velocitat mitjana màxima per circuit.

A més, en la primera pestanya, hi ha un desplegable per seleccionar el circuit que es vol consultar (inclou la opció de veure'ls tots):

Pilots: Velocitat n	
Mundial:	
Elementos de Página:	Gp: Barcelona ▾
	Velocitat
> Hamilton, Lewis	198

## Velocitat mitjana per pilot i circuit:

Pilots: Velocitat mitjana per GP Mundial: '2007'	
Elementos de Pàgina: <b>Gp: Barcelona</b>	
	Velocitat
> Hamilton, Lewis	196
> Massa, Felipe	198
> Alonso, Fernando	197
> Kubica, Robert	197
> Coulthard, David	196
> Kovalainen, Heikki	196
> Rosberg, Nico	196
> Sato, Takuma	195
> Barrichello, Rubens	194
> Davidson, Anthony	194
> Fisichella, Giancarlo	194
> Heidfeld, Nick	192
> Schumacher, Ralf	192

## Velocitat mitjana màxima per circuit:

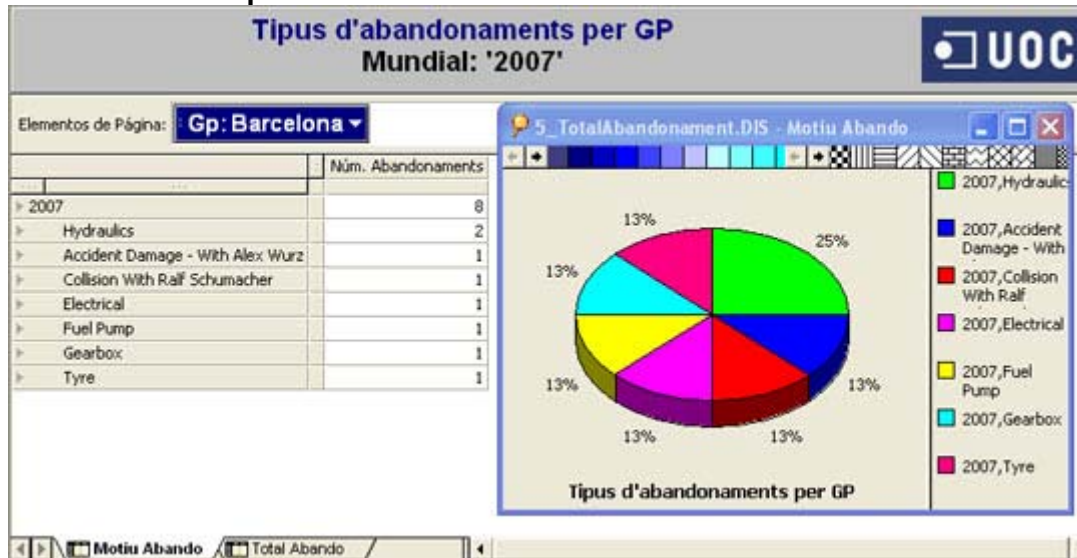
Velocitat mitjana màxima per GP Mundial: '2007'	
max_velocitat	
> Monza	234
> Spa-Francorchamps	229
> Silverstone	223
> Melbourne	216
> Istanbul	214
> Interlagos	208
> Magny-Cours	204
> Sepang	202
> Indianapolis	201
> Bahrain	198
> Barcelona	198
> Montreal	193
> Hungaroring	192
> Shanghai	187
> Monte Carlo	156
> Mount Fuji	152

### 5.4.5. Total d'abandonaments per GP

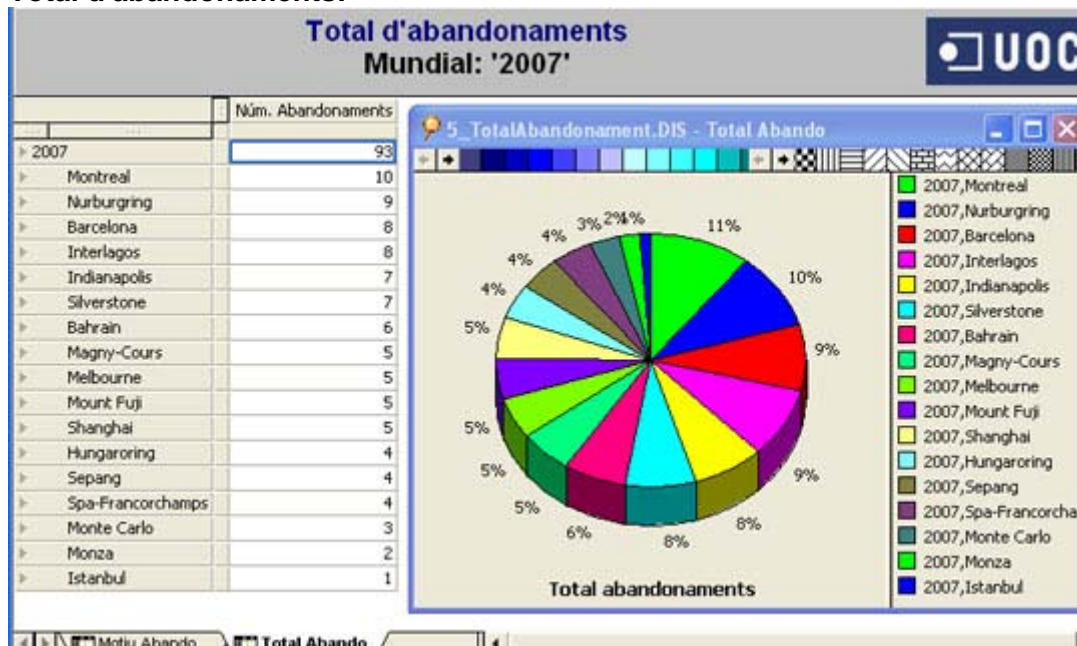
Nom informe: 5\_TotalAbandonament.DIS

Característiques: S'ha d'indicar l'any del Mundial que es vol consultar. Hi ha dues pestanyes disponibles, corresponents a número d'abandonaments classificats per motiu en cada GP i total d'abandonaments per GP.

#### Abandonaments per motiu:



#### Total d'abandonaments:



### 5.4.6. Rànquing de puntuació per GP

Nom informe: 6\_ResultatMundial.DIS

Característiques: S'ha d'indicar l'any del Mundial que es vol consultar. Hi ha dues pestanyes disponibles, corresponents al Mundial de Pilots i al Mundial de Constructors respectivament.

#### Mundial de Pilots:

Mundial de Pilots											UOC
Any: '2006'											
	Magny-Cours	Melbourne	Monte Carlo	Montreal	Monza	Nurburgring	Sepang	Shanghai	Silverstone	Suzuka	Total
	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	
Alonso, Fernando	8	10	10	10		8	8	8	10	10	134
Schumacher, Michael	10		4	8	10	10	3	10	8		121
Massa, Felipe	6			4		6	4		4	8	80
Fisichella, Giancarlo	3	4	3	5	5	3	10	6	5	6	72
Raikkonen, Kimi	4	8		6	8	5			6	4	65
Button, Jenson					4		6	5		5	56
Barrichello, Rubens		2	5		3	4		3			30
Pablo Montoya, Juan			8				5		3		26
Heidfeld, Nick	1	5	2	2	1			2	2	1	23
Schumacher, Ralf	5	6	1				1			2	20
de la Rosa, Pedro	2							4			19
Trull, Jarno				3	2					3	15
Coulthard, David		1	6	1							14
Villeneuve, Jacques		3				1	2		1		7
Webber, Mark								1			7
Kubica, Robert					6						6
Rosberg, Nico						2					4
Klien, Christian											2
Luzzi, Vitantonio											1
Abers, Christijan											

#### Mundial de Constructors:


Mundial de Constructors											UOC
Any: '2006'											
	Magny-Cours	Melbourne	Monte Carlo	Montreal	Monza	Nurburgring	Sepang	Shanghai	Silverstone	Suzuka	Total
	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	Puntuacio	
Renault	11	14	13	15	5	11	18	14	15	16	206
Ferrari	16		4	12	10	16	7	10	12	8	201
McLaren	6	8	8	6	8	5	5	4	9	4	110
Honda		2	5		7	4	6	8		5	86
BMW	1	8	2	2	7	1	2	2	3	1	36
Toyota	5	6	1	3	2		1			5	35
Red Bull		1	6	1							16
Williams						2		1			11
Toro Rosso											1
Midland											
Super Aguri											



## 5.4.7. Pilots de Mundial

Nom informe: 7\_DetallMundial.DIS

Característiques: S'ha d'indicar l'any del Mundial que es vol consultar. Hi ha dues pestanyes disponibles, corresponents al Mundial de Pilots i al Mundial de Constructors respectivament.

Detall Mundial de Pilots							
Any: '2006'							
							
	Pilot	Nacionalitat	Equip	Curses disputades	Curses Guanyades	Poles aconseguides	Puntuacio
▶ 1	<b>Alonso, Fernando</b>	Spain	Renault	18	7	5	<b>134</b>
▶ 2	<b>Schumacher, Michael</b>	Germany	Ferrari	18	7	5	<b>121</b>
▶ 3	<b>Massa, Felipe</b>	Brazil	Ferrari	18	2	3	<b>80</b>
▶ 4	<b>Fisichella, Giancarlo</b>	Italy	Renault	18	1	1	<b>72</b>
▶ 5	<b>Raikkonen, Kimi</b>	Finland	McLaren	18	0	3	<b>65</b>
▶ 6	<b>Button, Jenson</b>	(N/A)	Honda	18	1	1	<b>56</b>
▶ 7	<b>Barrichello, Rubens</b>	Brazil	Honda	18	0	0	<b>30</b>
▶ 8	<b>Pablo Montoya, Juan</b>	Colombia	McLaren	10	0	0	<b>26</b>
▶ 9	<b>Heidfeld, Nick</b>	Germany	BMW	18	0	0	<b>23</b>
▶ 10	<b>Schumacher, Ralf</b>	Germany	Toyota	18	0	0	<b>20</b>
▶ 11	<b>de la Rosa, Pedro</b>	Spain	McLaren	8	0	0	<b>19</b>
▶ 12	<b>Trulli, Jarno</b>	Italy	Toyota	18	0	0	<b>15</b>
▶ 13	<b>Coulthard, David</b>	(N/A)	Red Bull	18	0	0	<b>14</b>
▶ 14	<b>Villeneuve, Jacques</b>	Canada	BMW	12	0	0	<b>7</b>
▶ 15	<b>Webber, Mark</b>	Australa	Williams	18	0	0	<b>7</b>
▶ 16	<b>Kubica, Robert</b>	(N/A)	BMW	6	0	0	<b>6</b>
▶ 17	<b>Rosberg, Nico</b>	Germany	Williams	18	0	0	<b>4</b>
▶ 18	<b>Klien, Christian</b>	Austria	Red Bull	15	0	0	<b>2</b>
▶ 19	<b>Ferrari Vitantonio</b>	Italy	Toro Rosso	18	0	0	<b>1</b>


◀ ▶  Detall Mundial Pilots  Detall Mundial Construct


Detall Mundial de Constructors					
Any: '2006'					
					
	Equip	Recompte Curses	Curses Guanyades	Poles aconseguides	Puntuacio
▶ 1	<b>Renault</b>	18	8	6	<b>206</b>
▶ 2	<b>Ferrari</b>	18	9	8	<b>201</b>
▶ 3	<b>McLaren</b>	18	0	3	<b>110</b>
▶ 4	<b>Honda</b>	18	1	1	<b>86</b>
▶ 5	<b>BMW</b>	18	0	0	<b>36</b>
▶ 6	<b>Toyota</b>	18	0	0	<b>35</b>
▶ 7	<b>Red Bull</b>	18	0	0	<b>16</b>
▶ 8	<b>Williams</b>	18	0	0	<b>11</b>
▶ 9	<b>Toro Rosso</b>	18	0	0	<b>1</b>
▶ 10	<b>Super Aguri</b>	18	0	0	<b>0</b>
▶ 11	<b>Midland</b>	18	0	0	<b>0</b>

◀ ▶  Detall Mundial Pilots  Detall Mundial Const

## 5.4.8. Campionat del món de pilots

Nom informe: 8\_CampionatMonPilots.DIS

Campionat del Món de Pilots				
	▶ Mundial	▶ Pilot	▶ Nacionalitat	▶ Equip
▶ 1	<b>2007</b>	Raikkonen, Kimi	Finland	Ferrari
▶ 2	<b>2006</b>	Alonso, Fernando	Spain	Renault
▶ 3	<b>2005</b>	Alonso, Fernando	Spain	Renault
▶ 4	<b>2004</b>	Schumacher, Michael	Germany	Ferrari
▶ 5	<b>2003</b>	Schumacher, Michael	Germany	Ferrari
▶ 6	<b>2002</b>	Schumacher, Michael	Germany	Ferrari
▶ 7	<b>2001</b>	Schumacher, Michael	Germany	Ferrari
▶ 8	<b>2000</b>	Schumacher, Michael	Germany	Ferrari
▶ 9	<b>1999</b>	Hakkinen, Mika	Finland	McLaren
▶ 10	<b>1998</b>	Hakkinen, Mika	Finland	McLaren
▶ 11	<b>1997</b>	Villeneuve, Jacques	Canada	Williams
▶ 12	<b>1996</b>	Hill, Damon	(N/A)	Williams
▶ 13	<b>1995</b>	Schumacher, Michael	Germany	Benetton
▶ 14	<b>1994</b>	Schumacher, Michael	Germany	Benetton
▶ 15	<b>1993</b>	Prost, Alain	France	Williams
▶ 16	<b>1992</b>	Mansell, Nigel	(N/A)	Williams
▶ 17	<b>1991</b>	Senna, Ayrton	Brazil	McLaren
▶ 18	<b>1990</b>	Senna, Ayrton	Brazil	McLaren
▶ 19	<b>1989</b>	Prost, Alain	France	McLaren
▶ 20	<b>1988</b>	Prost, Alain	France	McLaren
▶ 21	<b>1987</b>	Piquet, Nelson	Brazil	Williams
▶ 22	<b>1986</b>	Prost, Alain	France	McLaren
▶ 23	<b>1985</b>	Prost, Alain	France	McLaren
▶ 24	<b>1984</b>	Prost, Alain	France	McLaren
▶ 25	<b>1983</b>	Piquet, Nelson	Brazil	Brabham
▶ 26	<b>1982</b>	Rosberg, Keke	Finland	Williams
▶ 27	<b>1981</b>	Piquet, Nelson	Brazil	Brabham
▶ 28	<b>1980</b>	Jones, Alan	Australia	Williams
▶ 29	<b>1979</b>	Scheckter, Jody	South Africa	Ferrari
▶ 30	<b>1978</b>	Andretti, Mario	(N/A)	Lotus
▶ 31	<b>1977</b>	Lauda, Niki	Austria	Ferrari
▶ 32	<b>1976</b>	Hunt, James	(N/A)	McLaren
▶ 33	<b>1975</b>	Lauda, Niki	Austria	Ferrari
▶ 34	<b>1974</b>	Fittipaldi, Emerson	Brazil	McLaren
▶ 35	<b>1973</b>	Stewart, Jackie	(N/A)	Torrell

◀▶  Campions de Mundial ◀▶

## 6. LÍNIES D'EVOLUCIÓ FUTURA

### 6.1. A curt termini: millores del sistema actual

Com ja s'ha repetit en aquest document, es recomana aprofitar la feina de depuració feta sobre les dades del fitxer Excel realitzant el manteniment de la informació sobre la nova BD relacional en substitució del sistema actual.

Per facilitar la transició a aquesta nova manera de treballar, s'aconsella la creació d'un aplicatiu que agilitzi les tasques més habituals sobre la BD relacional: inserir nova informació i/o modificar-ne l'existent.

Aquest aplicatiu ha de tenir una interfície amigable i simple, de manera que no requereixi cap esforç d'aprenentatge per part dels usuaris de l'ICEM. En definitiva, la millor manera d'aconseguir canviar el sistema actual és proporcionant una eina que resulti més ràpida i pràctica pel tècnic que manté la informació.

La desaparició del fitxer Excel a favor de la utilització d'aquesta base de dades augmentarà la fiabilitat de la informació i garantirà la integració de les dades. Així mateix, quan finalitzi el procés d'implantació del nou sistema la BD temporal es podrà eliminar.

### 6.2. A mig-llarg termini: evolució del sistema

Tal com s'ha explicat a l'inici d'aquesta memòria, aquest projecte representa la fase pilot en el procés de construcció d'un sistema molt més ampli que englobarà els resultats obtinguts a les diferents competicions mundials d'esports de motor amb què treballa l'ICEM.

Això vol dir que si la implantació d'aquest sistema té èxit, s'iniciarà la següent fase del projecte començant per estudiar les dades de les diferents competicions esportives. Probablement caldrà crear un magatzem de dades per cada tipus d'esport. Paral·lelament s'incorporaran noves àrees de negoci mitjançant Oracle Discoverer Administrator i es dissenyaran els informes a mida amb Discoverer Desktop.

A més, a mida que el projecte creixi poden sorgir noves necessitats que caldrà tenir en compte i, també es poden demanar modificacions o ampliacions en el sistema actual que s'hauran de valorar. Per això, el sistema dissenyat ha de ser dinàmic i estar obert a evolucionar al llarg del temps.

Finalment, quan es disposi d'una FIC que reculli els resultats dels diferents esports amb què treballa l'ICEM, s'aconsella estudiar la viabilitat i oportunitat de desenvolupar una aplicació web d'accés restringit amb usuari i contrasenya que permeti, a cada client, consultar directament determinats informes en funció dels permisos assignats. D'aquesta manera, al ser autosuficients, s'alleujarà la càrrega de treball del personal de l'ICEM.

## 7. CONCLUSIONS I AGRAÏMENTS

L'elaboració d'aquest projecte m'ha donat l'oportunitat d'introduir-me per primer cop en el camp dels magatzems de dades i conèixer les diferències respecte el model relacional que s'estudia durant la carrera.

He conegut els avantatges de tenir dos entorns diferenciats per donar resposta a necessitats diferents segons el perfil d'usuari i, he pogut posar en pràctica la teoria apresada mitjançant el desenvolupament d'un cas concret, completant així el procés d'assimilació dels nous conceptes.

Per a la realització del projecte he aplicat molts dels coneixements adquirits durant aquests anys d'estudis. He pogut aprofundir en el l'estudi d'Oracle i del llenguatge PL/SQL i, he començat a veure el funcionament de l'eina Oracle Discoverer alhora que es feia evident el llarg camí que em falta recórrer, per arribar a aprofitar tot el seu potencial.

En general, considero que la majoria d'objectius definits a l'inici del TFC s'han assolit correctament però, m'hagués agradat poder dedicar més temps i optimitzar el resultat final. De totes maneres, sóc de l'opinió que aquest projecte representa només el primer esglaió en el procés d'aprenentatge i que, cal esforçar-se contínuament per millorar fora de la Universitat.

Per acabar, volia agrair al consultor de l'assignatura, l'Àngel, el seu suport i ajuda durant tot el procés de desenvolupament del projecte i, mencionar que aquesta experiència ha estat molt enriquidora i una bona manera de culminar els estudis d'enginyeria tècnica.

## 8. GLOSSARI

**Base de dades multidimensional:** Base de dades dissenyada per al processament analític en línia (OLAP), estructurada com un hipercub amb un eix per dimensió.

**Base de dades operacional:** Base de dades destinada a gestionar el dia a dia d'una organització, és a dir, emmagatzema la informació referent a l'operativa diària d'una institució.

**Data warehouse** *Veure magatzem de dades.*

**Dimensió:** Punt de vista utilitzat en l'anàlisi d'un cert fet.

**EUL (End User Layer):** Fa referència a l'usuari final d'Oracle Discoverer. EUL emmascara la complexitat i l'estructura del magatzem de dades proporcionant als usuaris finals una interfície fàcil d'entendre i d'utilitzar sense coneixements específics de bases de dades.

**ETL (Extract, Transform and Load):** És el procés que permet a les organitzacions moure dades des de múltiples fonts, transformant-les, depurant-les i carregant-les en una altra base de dades, data mart o magatzem de dades per a analitzar o bé, en un altre sistema operacional per recolzar un procés de negoci.

**Factoria d'informació corporativa:** Conjunt d'elements de programari i maquinari que interactuen per a obtenir informació de suport que ajudi als analistes en processos de presa de decisions en l'organització.

Sigla: FIC

**Fet:** Objecte d'anàlisi.

**FIC:** *Veure factoria d'informació corporativa*

**Granularitat:** Grandària d'un objecte respecte a un altre.

**Jerarquia d'agregació:** Conjunt de relacions entre les instàncies d'una dimensió que indica com n'agrupem unes per a obtenir-ne d'altres.

**Magatzem de dades:** Bases de dades orientades a àrees d'interès de l'empresa que integren dades de diferents fonts amb informació històrica i no volàtil que tenen com a objectiu principal fer de suport en la presa de decisions.

**Metadada:** Dades sobre dades.

**OLAP (On-Line Analytical Processing):** Aplicacions que permeten als usuaris de veure, manipular i analitzar bases de dades multidimensionals i navegar-hi.

**Oracle:** És un sistema de gestió de bases de dades relacional, desenvolupat per Oracle Corporation.

**Oracle Discoverer:** És una eina orientada a temàtiques. S'utilitza per construir consultes ad hoc amb dades d'una àrea específica que són d'interès per un cert grup d'usuaris.

- **Oracle Discoverer Administrator:** Aquest producte serveix per crear, mantenir i administrar dades en la capa d'usuari final (EUL) i per definir com els usuaris interactuen amb les dades.
- **Oracle Discoverer Desktop:** L'usuari final (EUL) utilitza aquest component per executar consultes ad hoc i generar reports.

**PL/SQL:** És el llenguatge de programació estructurada de les bases de dades Oracle. Permet definir procediments mitjançant sentències SQL.

**Procediment** (de l'anglès *procedure*) és un conjunt encadenat de comandes o sentències que realitzen unes tasques determinades.

**Sistema de gestió de bases de dades:** Programari que gestiona i controla bases de dades. Les seves funcions principals són les de facilitar-ne l'ús simultani a molts usuaris de tipus diferents, independitzar l'usuari del món físic i mantenir la integritat de les dades.

**SGBD:** *Veure Sistema de gestió de bases de dades.*

**SQL\*Loader:** El programa SQL\*Loader és una potent utilitat d'importació de dades que possibilita la càrrega automàtica de dades externes (residents en fitxers) en taules de la base de dades. La informació pot carregar-se en una o varies taules prèviament creades i que poden tenir contingut previ. Les noves dades podran substituir a les que ja existien en les taules o bé afegir-se com a noves files.

**Vista** (en base de dades): És el resultat d'una consulta SQL sobre una o varies taules; també es pot considerar com una taula virtual ja que té la mateixa estructura que una taula (files i columnes) amb la diferència que emmagatzemen la definició de la consulta SQL però no les dades.

## 9. BIBLIOGRAFIA

Llibres, material de consulta i referències a pàgines web.

- Inmon, William H. Building the Data Warehouse  
Wiley Publishing, Inc. 2005 (Fourth edition)
- Pla docent. TFC, Magatzems de dades  
UOC, material assignatura
- Pérez Navarro, A., Bataller Díaz, A. [...]. Treball final de carrera  
UOC, material assignatura
- Viquipèdia. Fórmula 1  
[http://ca.wikipedia.org/wiki/F%C3%B3rmula\\_1](http://ca.wikipedia.org/wiki/F%C3%B3rmula_1)
- Oracle® Database SQL Reference  
[http://download.oracle.com/docs/cd/B19306\\_01/server.102/b14200/toc.htm](http://download.oracle.com/docs/cd/B19306_01/server.102/b14200/toc.htm)
- Adrian Billington. Oracle 9i MERGE STATEMENT  
[http://www.quest-pipelines.com/newsletter-v4/0903\\_D.htm](http://www.quest-pipelines.com/newsletter-v4/0903_D.htm)
- Ayuda para Oracle  
<http://ora.u440.com/>
- Wiki Oracle  
<http://www.wikioracle.es/doku.php>
- Oracle® Database Utilities  
[http://download-uk.oracle.com/docs/cd/B19306\\_01/server.102/b14215/toc.htm](http://download-uk.oracle.com/docs/cd/B19306_01/server.102/b14215/toc.htm)

## 10. ANNEXOS

### 10.1. Annex 1: Anàlisi de les fonts de dades d'origen

A continuació es presenta una taula amb les observacions derivades de l'anàlisi de les fonts de dades proporcionades per l'ICEM.

Full	Observació	Tractament proposat
<b>Resultats</b>	Circuit/Data conté "/nomCircuit/dataCursa"	Separar la informació en els dos camps corresponents de la BD relacional. Unificar els diferents formats de data.
	Valors inconsistents (p.ex. a nom corredor trobem tipus de pneumàtic)	Tenint en compte que la resta de camps estan buits, descartarem la informació
	El format del camp temps és variable. (p.e. 17.866, 23 lap(s) o +1:13.348).	El tipus del camp temps es definirà com una cadena de text.
	Apareixen tres resultats amb valor N/A a tots els camps excepte circuit/data	Només es crea la cursa (circuit/data). El valor N/A no s'incorpora.
	Motiu Abandonament té alguns registres amb valor "-" en curses finalitzades	No incloure aquest valor a la taula d'abandonaments.
<b>Poles</b>	Diferents formats de temps	Unificar format de tipus Time.
	Diferents formats de data	Unificar format de tipus Date.
	Apareixen tres resultats amb valor N/A a tots els camps excepte circuit, país i data	Només es crea la cursa (circuit/data). El valor N/A no s'incorpora.
<b>Corredors</b>	Corredor conté "cognom(s), nom"	Separar la informació en els dos camps corresponents de la BD relacional.
	Diferents formats a anys en actiu.	El tipus del camp es definirà com una cadena de text.
	Apareixen resultats amb valor N/A a corredor i anys en actiu.	El valor N/A no s'incorpora.
<b>Altres:</b> - S'hauran d'eliminar espais en blanc davant i darrera dels valors abans d'incorporar les dades a la BD relacional. - Caldrà incloure les metadades: data alta i data modificació.		



## 10.2. Annex 2: Càrrega de dades d'origen

Exemple de càrrega de dades a la BD Temporal d'Oracle a partir del full de resultats del fitxer Excel mitjançant SQL\*Loader

```
-- estructura del fitxer de control RESULTATS
-- cargaresultats.ctl
LOAD DATA
CHARACTERSET 'UTF8'
INFILE 'C:\TFC\resultats.csv'
BADFILE 'resultats.bad'
DISCARDFILE 'resultats.dsc'
INTO TABLE Temp_Resultat
FIELDS TERMINATED BY ";"
OPTIONALLY ENCLOSED BY '"'
TRAILING NULLCOLS
(
    circuitData CHAR "ltrim(rtrim(:circuitData))",
    corredor CHAR "ltrim(rtrim(:corredor))",
    equip CHAR "ltrim(rtrim(:equip))",
    pneumatic CHAR "ltrim(rtrim(:pneumatic))",
    posicio CHAR "ltrim(rtrim(:posicio))",
    voltes CHAR "ltrim(rtrim(:voltes))",
    temps CHAR "ltrim(rtrim(:temps))",
    velocitat CHAR "ltrim(rtrim(:velocitat))",
    motiuAbando CHAR "ltrim(rtrim(:motiuAbando))"
)

-- instrucció de càrrega SQL*Loader
sqlldr RAW_TFC/pwdtfc DATA=resultats.csv CONTROL=cargaresultats.ctl
LOG=resultats.log

-- LOG resultant
Tabla TEMP_RESULTAT:
 15784 Filas se ha cargado correctamente.
   0 Filas no cargada debido a errores de datos.
   0 Filas no cargada porque todas las cláusulas WHEN han fallado.
   0 Filas no cargada porque todos los campos eran nulos.
```

## 10.3. Annex 3: Càrrega inicial a la BD Relacional

Mostra d'instruccions PL/SQL de càrrega inicial a la BD Relacional amb les dades de la BD Temporal.

```
-- Carrega taula Pneumatic
INSERT into pneumatic (pneumatic)
SELECT DISTINCT NLS_INITCAP(pneumatic) FROM raw_tfc.temp_resultat
WHERE pneumatic<>'N/A' AND pneumatic IS NOT NULL AND pneumatic <>'72'
ORDER BY NLS_INITCAP(pneumatic)

7 rows inserted

-- Taula Equip
UPDATE raw_tfc.temp_resultat SET equip =TRIM(BOTH FROM equip);

INSERT into equip (equip)
SELECT DISTINCT equip
FROM raw_tfc.temp_resultat
WHERE equip<>'N/A' AND equip<>'-' AND equip IS NOT NULL AND equip<>'2'
ORDER BY equip

95 rows inserted
```

```
-- Taula Resultat
-- 3 registres duplicats. Inserir només cursa, pilot, equip
INSERT into resultat (idcursa, idpilot, idequip)
SELECT cu.idcursa, pi.idpilot, eq.idequip
FROM (((((raw_tfc.temp_resultat res
INNER JOIN granpremi on TRIM(BOTH FROM SUBSTR(res.circuitdata,1,INSTR(res.circuitdata,'/'))-1)) = granpremi.gp)
INNER JOIN cursa cu on cu.data= TO_DATE(TRIM(BOTH FROM SUBSTR(res.circuitdata,INSTR(res.circuitdata,'/')+2)), 'Day ,Month dd, YYYY', 'NLS_DATE_LANGUAGE = American')
AND cu.idgp=granpremi.idgp)
INNER JOIN pilot pi on TRIM(BOTH FROM SUBSTR(res.corredor,1,instr(res.corredor,' ') -1)) = pi.nom AND TRIM(BOTH FROM SUBSTR(res.corredor,instr(res.corredor,' ') +1))=pi.cognom)
left outer join motiuabandonament mot on NLS_INITCAP(res.motiuabando)=mot.motiu)
left outer join pneumatic pne on NLS_INITCAP(res.pneumatic)=pne.pneumatic)
INNER JOIN equip eq on res.equip = eq.equip
WHERE circuitdata<>'Liria Cespede, Jose Carlos [CAP]' AND res.corredor<>'N/A'
GROUP BY cu.idcursa, pi.idpilot, eq.idequip
having COUNT(*)=2

ORDER BY cu.idcursa, pi.idpilot
```

3 rows inserted

```
-- inserim la resta de resultats
UPDATE temp_resultat
SET velocitat = null
WHERE velocitat='Suspension'

INSERT into resultat (idcursa, idpilot, idequip, idmotiu, idpneumatic, posicio, numvoltes, temps, velocitat)
SELECT cu.idcursa, pi.idpilot, eq.idequip, mot.idmotiu, pne.idpneumatic, res.posicio, SUBSTR(res.voltes,1,3), res.temps, res.velocitat
FROM (((((raw_tfc.temp_resultat res INNER JOIN granpremi on TRIM(BOTH FROM SUBSTR(res.circuitdata,1,INSTR(res.circuitdata,'/'))-1)) = granpremi.gp) INNER JOIN cursa cu on cu.data= TO_DATE(TRIM(BOTH FROM SUBSTR(res.circuitdata,INSTR(res.circuitdata,'/')+2)), 'Day ,Month dd, YYYY', 'NLS_DATE_LANGUAGE = American') AND cu.idgp=granpremi.idgp)
INNER JOIN pilot pi on TRIM(BOTH FROM SUBSTR(res.corredor,1,instr(res.corredor,' ') -1)) = pi.nom AND TRIM(BOTH FROM SUBSTR(res.corredor,instr(res.corredor,' ') +1))=pi.cognom)
INNER JOIN equip eq on res.equip=eq.equip) left outer join motiuabandonament mot on NLS_INITCAP(res.motiuabando)=mot.motiu) left outer join pneumatic pne on NLS_INITCAP(res.pneumatic)=pne.pneumatic
WHERE circuitdata<>'Liria Cespede, Jose Carlos [CAP]' AND res.corredor<>'N/A'
AND not exists (SELECT idcursa, idpilot FROM resultat WHERE idcursa=cu.idcursa AND idpilot=pi.idpilot) AND not exists (SELECT circuitdata, corredor FROM raw_tfc.temp_resultat WHERE corredor not in (SELECT pneumatic FROM pneumatic)
AND circuitdata<>'Liria Cespede, Jose Carlos [CAP]' AND corredor<>'N/A'
AND circuitdata=res.circuitdata AND corredor=res.corredor)
GROUP BY circuitdata, corredor
having COUNT(*)>1)
ORDER BY cu.idcursa, pi.idpilot
```

15685 rows inserted

```
--insetem els resultats de Geki (pilot sense nom)
INSERT into resultat (idcursa, idpilot, idequip, idmotiu, idpneumatic, posicio, numvoltes, temps, velocitat)
SELECT cu.idcursa, pi.idpilot, eq.idequip, mot.idmotiu, pne.idpneumatic, res.posicio, SUBSTR(res.voltes,1,3), res.temps, res.velocitat
FROM (((((raw_tfc.temp_resultat res INNER JOIN granpremi on TRIM(BOTH FROM SUBSTR(res.circuitdata,1,INSTR(res.circuitdata,'/'))-1)) = granpremi.gp)
INNER JOIN cursa cu on cu.data= TO_DATE(TRIM(BOTH FROM SUBSTR(res.circuitdata,INSTR(res.circuitdata,'/')+2)), 'Day ,Month dd, YYYY', 'NLS_DATE_LANGUAGE = American') AND cu.idgp=granpremi.idgp)
INNER JOIN pilot pi on TRIM(BOTH FROM SUBSTR(res.corredor,instr(res.corredor,' ') +1))=pi.cognom) INNER JOIN equip eq on res.equip=eq.equip)
left outer join motiuabandonament mot on NLS_INITCAP(res.motiuabando)=mot.motiu)
left outer join pneumatic pne on NLS_INITCAP(res.pneumatic)=pne.pneumatic
WHERE TRIM(BOTH FROM SUBSTR(res.corredor,instr(res.corredor,' ') +1))='Geki'
ORDER BY cu.idcursa, pi.idpilot
```

2 rows inserted

#### 10.4. Annex 4: Tractament dades d'origen

En el procés de transformació, depuració i càrrega de les dades de la BD temporal a la BD relacional es realitza el tractament de la informació que s'enumera a continuació:

- Neteja d'espais en blanc davant i darrera de les dades
- Es descarta la informació amb valor 'N/A',
- Es descarta el valor 'Suspensión' en el camp de velocitat
- Es descarta el valor '-' en el camp de motiuabando
- Es descarten els valors de temps (%laps%) en el camp motiuabando
- Es descarten els registres on el valor de circuitdata és 'Liria Cespede, Jose Carlos [CAP]'
- Es descarten els registres on el nom de corredor és un tipus de pneumàtic
- Es descarten els registres amb valor null al camp equip
- Es descarta '2' com a nom d'equip
- Es descarta '72' com a pneumàtic
- Es descarta UK i US com a país degut a que no hi ha informació vàlida associada i, en canvi, USA i Great Britain sí en tenen.
- Hi ha 15 registres que es repeteixen dues vegades a temp\_resultats (mateix circuitdata, i corredor). Només 3 d'aquests registres tenen el mateix valor pels camps circuitdata, pilot i equip. Inserim aquests camps dels tres registres i descartem tota la resta
- Es transformen els camps data passant l'any de 2 a 4 dígits per evitar errors d'interpretació

#### 10.5. Annex 5: Registres pendents de revisió

En aquest apartat es detallen els registres que, degut a incoherències en la informació, no s'han pogut importar a la BD Relacional i, també aquells registres que només s'han importat parcialment.

Els tècnics de l'ICEM hauran de revisar i introduir els valors vàlids a la nova BD relacional si volen realitzar l'explotació de les dades.

Dades carregades parcialment:

Circuit/Data	Corredor	Equip
/Watkins Glen/Sunday, October 08, 1961	Masten Gregory	Lotus
/Watkins Glen/Sunday, October 04, 1964	Jim Clark	Lotus
/Watkins Glen/Sunday, October 04, 1964	Mike Spence	Lotus

Dades carregades completament però amb incoherències en el número de posició. S'ha introduït la informació a la BD perquè no afecta la puntuació de la cursa, però caldria fer-ne la correcció pertinent.

Circuit/Data	Corredor	Equip	Neumàtic	Posició
/Watkins Glen/ Sunday, October 03, 1965	Moises Solana	Lotus	Dunlop	12
/Watkins Glen/ Sunday, October 03, 1965	Jackie Stewart	BRM	Dunlop	12

Dades pendents de càrrega:

Circuit/Data	Corredor
/Mount Fuji/Sunday, September 30, 2007	Sebastian Vettel
/Interlagos/Sunday, October 24, 2004	Christian Klien
/Monza/Sunday, September 04, 2005	Ralf Schumacher
/Monza/Sunday, September 04, 2005	Tiago Monteiro
/Suzuka/Sunday, October 10, 2004	Olivier Panis
/Shanghai/Sunday, September 26, 2004	Olivier Panis
/Hungaroring/Sunday, July 31, 2005	Juan Pablo Montoya
/Aintree/Saturday, July 15, 1961	Stirling Moss
/Hockenheim/Sunday, July 24, 2005	Tiago Monteiro
/Montreal/Sunday, June 13, 2004	David Coulthard
/Monza/Sunday, September 09, 2007	Jarno Trulli
/Spa-Francorchamps/Sunday, September 16, 2007	Sakon Yamamoto
Liria Cespede, Jose Carlos [CAP]	Riccardo Patrese
Liria Cespede, Jose Carlos [CAP]	Didier Pironi
Liria Cespede, Jose Carlos [CAP]	Andrea de Cesaris
Liria Cespede, Jose Carlos [CAP]	Nigel Mansell
Liria Cespede, Jose Carlos [CAP]	Elio de Angelis
Liria Cespede, Jose Carlos [CAP]	Derek Daly
Liria Cespede, Jose Carlos [CAP]	Alain Prost
Liria Cespede, Jose Carlos [CAP]	Brian Henton
Liria Cespede, Jose Carlos [CAP]	Marc Surer
Liria Cespede, Jose Carlos [CAP]	Michele Alboreto
Liria Cespede, Jose Carlos [CAP]	Keke Rosberg
Liria Cespede, Jose Carlos [CAP]	Niki Lauda
Liria Cespede, Jose Carlos [CAP]	Nelson Piquet
Liria Cespede, Jose Carlos [CAP]	John Watson

## 10.6. Annex 6: Càrrega inicial al magatzem

Mostra d'instruccions PL/SQL de càrrega inicial al Magatzem.

```
-- Dimensions
-- Taula Mundial
INSERT into mundial (anyM, dataIni, dataFi)
SELECT  EXTRACT(YEAR FROM cu.data) as anyC, min(cu.data) as dataini,
MAX(cu.data) as datafi
FROM bdrel.cursa cu
GROUP BY EXTRACT(YEAR FROM cu.data)
ORDER BY EXTRACT(YEAR FROM cu.data)
47 rows inserted

-- Abandonament
INSERT into abando (idabando, motiu)
SELECT idmotiu, motiu
FROM bdrel.motiuabandonament ab
ORDER BY ab.motiu

INSERT into abando (idabando, motiu) values (0,'(N/A)')
133 rows inserted

-- Competidor
INSERT into competidor (idcompetidor, nom, cognom, pilotpais, equip)
SELECT DISTINCT (TRIM(BOTH FROM idpilot) || 'E' || TRIM(BOTH FROM idequip)),
NVL(nom, '(N/A)'), cognom, NVL(pais, '(N/A)'), EQUIP
FROM BDREL.V_PILOTEQUIP
1062 rows inserted
--Taula de Fets
```

```

INSERT into resultat (anymundial, idcompeticio, idcompetidor, idpneumatic,
idabando, posicio, velocitat, pole, puntuacio)

SELECT DISTINCT mun.anym, competicio.idcompeticio, comp.idcompetidor
,NVL(pne.idpneumatic,(SELECT to_number(idpneumatic) FROM pneumatic WHERE
pneumatic='(N/A)'))
,NVL(aba.idabando,(SELECT to_number(idabando) FROM abando WHERE motiu='(N/A)'))
,res.posicio
,res.velocitat
,CASE WHEN (res.idpilot=pole.idpilot) THEN 1 ELSE 0 END
,puntuacio.punts

FROM ((((((bdrel.v_resultat res
LEFT JOIN mundial mun on EXTRACT(YEAR FROM res.data)=mun.anym)
LEFT JOIN competicio on res.gp=competicio.gp)
LEFT JOIN competidor comp on ((res.nom=comp.nom AND res.cognom=comp.cognom AND
res.equip=comp.equip) OR (res.nom IS NULL AND res.cognom=comp.cognom AND
res.equip=comp.equip)))
LEFT JOIN pneumatic pne on res.pneumatic = pne.pneumatic)
LEFT JOIN abando aba on res.motiu= aba.motiu)
LEFT JOIN bdrel.pole pole on res.idcursa=pole.idcursa)
LEFT JOIN bdrel.sistpuntuacio SP ON ((EXTRACT(YEAR FROM res.data) BETWEEN
SP.ANYINI AND SP.ANYFI)))
LEFT JOIN bdrel.puntuacio puntuacio ON (SP.IDSPUNT=puntuacio.IDSPUNT AND
res.POSICIO=puntuacio.POSICIO)

ORDER BY mun.anym, competicio.idcompeticio, comp.idcompetidor

```

15690 rows inserted

## 10.7. Annex 7: Procediments emmagatzemats

Mostra dels diferents processos creats per al manteniment de les dades del magatzem.

```

-- manteniment taula de fets: resultat
create OR replace PROCEDURE CARREGA_RESULTAT IS

BEGIN

MERGE INTO resultat res
  USING (
    SELECT anym, idcompeticio, idcompetidor, idpneumatic, idabando
    ,posicio, velocitat, pole, puntuacio
    FROM v_resultat vres
    WHERE vres.datamodif > (SELECT MAX(dataImport) FROM resultat)
  ) bdorigen
ON (res.anymundial = bdorigen.anyM
AND res.idcompeticio= bdorigen.idcompeticio
AND res.idcompetidor = bdorigen.idcompetidor)

WHEN MATCHED THEN
  UPDATE
  SET
    res.idpneumatic = bdorigen.idpneumatic,
    res.idabando = bdorigen.idabando,
    res.posicio = bdorigen.posicio,
    res.velocitat = bdorigen.velocitat,
    res.pole = bdorigen.pole,
    res.puntuacio = bdorigen.puntuacio

```

```

    WHEN NOT MATCHED THEN
        INSERT (anymundial, idcompeticio, idcompetidor, idpneumatic,
            idabando, posicio, velocitat, pole, puntuacio)
        VALUES (bdorigen.anyM, bdorigen.idcompeticio, bdorigen.idcompetidor,
            bdorigen.idpneumatic,
            bdorigen.idabando, bdorigen.posicio, bdorigen.velocitat, bdorigen.pole,
            bdorigen.puntuacio);

EXCEPTION
    WHEN OTHERS THEN
        NULL;
END CARREGA_RESULTAT;

-- indicadors de resultat
create OR replace PROCEDURE CARREGA_BOOLEANS IS
    numAny NUMBER:=0;

CURSOR curs_guanyador IS
    SELECT anymundial, idcompetidor, sum(puntuacio)
    FROM resultat
    WHERE puntuacio IS NOT NULL
    GROUP BY anymundial, idcompetidor
    ORDER BY anymundial asc, sum(puntuacio) desc;

BEGIN
    UPDATE resultat
    SET guanyadorMundial=0, victoriaCursa=0, abando=0;

    UPDATE resultat
    SET victoriaCursa=1
    WHERE posicio=1;

    UPDATE resultat
    SET abando=1
    WHERE idabando<>0;

    FOR item IN curs_guanyador LOOP
        EXIT WHEN curs_guanyador%NOTFOUND;
        IF (numAny < item.anymundial) THEN
            UPDATE resultat
            SET guanyadorMundial=1
            WHERE idcompetidor=item.idcompetidor AND anymundial=item.anymundial
            AND rownum=1;
        END IF;
        numAny:=item.anymundial;
    END LOOP;

EXCEPTION
    WHEN OTHERS THEN
        NULL;
END CARREGA_BOOLEANS;

-- PROCEDURE D'ACTUALITZACIÓ DEL MAGATZEM
create OR replace PROCEDURE ACTUALITZACIO_MAGATZEM IS
BEGIN
    CARREGA_MUNDIAL;    CARREGA_COMPETICIO;
    CARREGA_PNEUMATIC; CARREGA_ABANDO;
    CARREGA_COMPETIDOR; CARREGA_RESULTAT;
    CARREGA_BOOLEANS;
EXCEPTION
    WHEN OTHERS THEN
        NULL;
END ACTUALITZACIO_MAGATZEM;

```

## 10.8. Annex 8: Automatització de la càrrega

Detall del procés d'automatització creat per dur a terme el manteniment automàtic de les dades del magatzem a partir de la BD Relacional.

```
--JOB D'AUTOMATITZACIÓ

BEGIN
DBMS_SCHEDULER.CREATE_JOB (
  job_name          => 'DW.Refresca_Magatzem_JOB',
  job_type          => 'STORED_PROCEDURE',
  job_action        => 'DW.ACTUALITZACIO_MAGATZEM',
  start_date        => '02-DIC-08 01.00.00 AM Europe/Warsaw',
  repeat_interval   => 'FREQ=WEEKLY;BYDAY=TUE;BYMONTH=3,4,5,6,7,8,9,10,11');
END;
/

-- PER EXECUTAR EL JOB MANUALMENT
BEGIN
DBMS_SCHEDULER.RUN_JOB (
  job_name =>'DW.Refresca_Magatzem_JOB',
  use_current_session => TRUE);
END;
/
```

## 10.9. Annex 9: Usuaris del sistema

A continuació es detallen les sentències de creació dels diferents usuaris del sistema i els permisos assignats.

```
-- Usuari BD Temporal:
CREATE USER RAW_TFC
IDENTIFIED BY pwdtfc
DEFAULT TABLESPACE Users TEMPORARY TABLESPACE Temp QUOTA UNLIMITED ON Users;

GRANT CREATE SESSION, CREATE TABLE TO RAW_TFC;

-- Usuari BD Relacional:

CREATE USER BDREL
IDENTIFIED BY pwdtfc
DEFAULT TABLESPACE Users TEMPORARY TABLESPACE Temp QUOTA UNLIMITED ON Users;

GRANT CREATE SESSION, CREATE TABLE, CREATE VIEW, CREATE TRIGGER, CREATE SEQUENCE TO BDREL;

    -- Des de RAW_TFC: donar permisos sobre les tres taules.

-- Usuari Magatzem de Dades:
CREATE USER DW
IDENTIFIED BY pwdtfc
DEFAULT TABLESPACE Users TEMPORARY TABLESPACE Temp QUOTA UNLIMITED ON Users;

GRANT CREATE SESSION, CREATE TABLE, CREATE VIEW,
CREATE TRIGGER, CREATE PROCEDURE, CREATE SEQUENCE TO DW;

    -- Des de BDREL: donar permisos sobre cada taula i vista existent.

-- End User Layer:
CREATE USER TFC_EUL
IDENTIFIED BY pwdtfc
DEFAULT TABLESPACE Users TEMPORARY TABLESPACE Temp QUOTA UNLIMITED ON Users;

GRANT CREATE SESSION, CREATE TABLE, CREATE VIEW,
CREATE TRIGGER, CREATE SEQUENCE, CREATE PROCEDURE TO TFC_EUL;

    -- Des de DW: permisos sobre cada taula i vista existent.
```