

PROJECTE FINAL DE CARRERA

ENGINYERIA D'INFORMÀTICA

Memòria

Disseny i implementació d'una base de dades
relacional per a la gestió d'apostes de futbol

Alumne: Jordi Castilla Miró
Consultor: Juan Martínez Bolaños

Quadrimestre primavera 2013-2014
Universitat Oberta de Catalunya

Dedicatòria

Per na Núria Miró Ribas i en Rafel Castilla Chavarria,
després de 33 anys ja torneu a estar junts.

Agraïments

A Raquel por ser como eres y dejarme ser como soy.

A Joselin i Dayana per escoltar-me encara que no m'entenieu.

A Dudu, Jordi, Jordi i Rafa per demostrar-me la vostra amistat a diari.

Gràcies a totes i a tots.

Resum

La present memòria és el treball final de molts anys de carrera on demostraré tots els coneixements que he adquirit al llarg del temps i les assignatures que he cursat.

L'objectiu d'aquest Projecte de Final de Carrera és el disseny i la implementació d'un sistema de gestió d'apostes de futbol (SGAF) basat en una base de dades Oracle.

He visitat llocs web d'apostes esportives i he investigat al respecte de les temporades, pels històrics, i de les modalitats d'apostes entre d'altres.

Planificaré les tasques, definiré i analitzaré els requeriments, enumeraré els casos d'ús, dissenyaré la base de dades operacional i de data warehouse, implementaré els esquemes i realitzaré les proves unitàries i d'extracció, transformació i càrrega.

L'usuari s'haurà de validar usant la funció *login* i per poder accedir a la informació haurà d'usar els paquets de la base de dades sense poder accedir directament a les dades de les taules.

Índex de continguts

Capítol 1: Introducció	13
1.1 Justificació del PFC i context en el qual es desenvolupa: punt de partida i aportació del PFC	13
1.1.1 Enunciat	13
1.1.2 Justificació	14
1.2 Objectius del PFC	14
1.3 Enfocament i mètode seguit	14
1.4 Planificació del treball	15
1.5 Productes obtinguts.....	15
1.6 Breu descripció dels altres capítols de la memòria.....	16
Capítol 2: Planificació.....	17
2.1 PAC 1: Planificació i Anàlisi (del 28/02 al 16/03).....	17
2.2 PAC 2: Disseny i creació (del 17/03 al 13/04)	17
2.3 PAC 3: Implementació i proves (del 14/04 al 11/05)	17
2.4 Lliurament Final: Memòria i presentació (del 12/05 al 15/06)	18
2.5 Tribunal Virtual.....	18
2.6 Avaluació de riscos i incidències previstes.....	18
2.7 Diagrama de Gantt.....	18
Capítol 3: Requeriments	19
3.1 Requeriments de maquinari.....	19
3.1.1 Punt de treball estàndard de la UOC.....	19
3.2 Requeriments de programari	20
3.2.1 Oracle	20
3.3 Requeriments funcionals bàsics.....	20
3.3.1. Gestió i accés únic a la informació usant procediments de base de dades	20
3.3.2 Gestió de dades esportives	20
3.3.3 Gestió d'apostes esportives	20
3.3.4 Gestió de magatzem de dades (<i>data warehouse</i>)	21
3.3.5 Gestió de problemes	21
3.4 Investigació de les modalitats d'apostes	21
3.4.1 Pàgines visitades.....	22
3.5 Requeriments d'escalabilitat	22
3.6 Altres requeriments.....	22
3.6.1 Interfície d'usuari	22
3.6.2 Requeriments operacionals	23
3.7 Eines i materials a usar	23
3.8 Recursos Humans	23
Capítol 4: Casos d'ús	25
4.1 Diagrama de Casos d'ús.....	25
4.2 Descripció dels casos d'ús.	25
4.2.1 Paquet d'usuaris	26
4.2.2 Paquet de competicions(l·ligues)	26
4.2.3 Paquet de partits.....	27
4.2.4 Paquet de joc	28
4.2.5 Paquet d'apostes.....	28
4.2.6 Paquet de data warehouse.....	29
4.2.7 Paquet de problemes	29

4.3 Actors	30
4.4 Altres restriccions	30
Capítol 5: Disseny	31
5.1 Disseny OLTP	31
5.1.1 Disseny conceptual OLTP	31
5.1.2 Disseny lògic OLTP	35
5.1.3 Disseny físic OLTP	38
5.2 Disseny data warehouse	39
5.2.1 Requeriments	39
5.2.2 Fets	39
5.2.3 Dimensions	39
5.2.4 Indicadors o mètriques	40
5.2.5 Atributs	40
5.2.6 Disseny conceptual data warehouse	41
5.2.7 Disseny lògic data warehouse	42
5.2.8 Disseny físic data warehouse	43
5.2.9 Disseny ETL.....	43
Capítol 6: Implementació	45
6.1 Paquets	45
6.1.1 pkg_competicions.....	46
6.1.2 pkg_usuaris.....	48
6.1.3 pkg_partits.....	49
6.1.4 pkg_joc	50
6.1.5 pkg_apostes	51
6.1.6 pkg_import	52
6.1.7 pkg_util	52
6.2 Registre de log	53
6.3 Dades inicials OLTP	53
6.4 Restriccions implementades.....	54
6.5 Proves unitàries OLTP.....	54
6.6 Disparadors OLTP.....	55
6.7 Paquets data warehouse	55
6.8 ETL: Procés d'importació	56
6.8.1 Neteja i normalització de les dades operacionals	56
6.8.2 Dimensions	56
6.8.3 Fets i indicadors	57
6.8.4 Execució ETL	57
6.9 Magatzem multidimensional	57
6.9.1 Creació del cub	57
6.9.2 Execució a Pentaho.....	58
Capítol 7: Valoració econòmica	61
Capítol 8: Conclusions.....	63
Glossari	65
Bibliografia	67
Annex A: Scripts SQL.....	69
A.1 Creació d'esquemes	69
A.2 Scripts OLTP	69
A.3 Scripts Data Warehouse.....	69
A.4 Scripts paquets OLTP.....	69
A.5 Scripts paquets Data Warehouse	69
A.6 Scripts de proves	69
Annex B: Proves realitzades a la base de dades operacional	71

B.1 Resultat execució del script 15z Execucio TEST.sql.....	71
B.2 Resultat execució de la funció pkg_util.dades: inicials	74
Annex C: Proves realitzades a la base de dades data warehouse	75
C.1 Resultat execució del script 25z Execucio IMPORT.sql.....	75
C.2 Creació del CUB.....	76
C.3 Fitxer XML del cub	79
C.4 Resultat execucions d'informes a Pentaho	80
Annex D: Planificació	91

Índex de figures

Figura 1: Cicle de vida del desenvolupament d'un projecte	14
Figura 2: Fase de seguiment d'un projecte	14
Figura 3: Combinació de DCU i àgil	15
Figura 4: maquinari de què dispoço	19
Figura 5: Casos d'ús	25
Figura 6: Exemple de coeficients 1 X 2	29
Figura 7: Disseny conceptual OLTP	31
Figura 8: Disseny lògic OLTP	35
Figura 9: Disseny conceptual data warehouse.....	41
Figura 10: Disseny lògic data warehouse	42
Figura 11: Procés estàndard ETL de generació d'un data warehouse	55
Figura 12: Procés actual de generació del data warehouse	56
Figura 13: Resultat de la consulta de distribució geogràfica	58
Figura 14: Resultat de la consulta de distribució per edats.....	59
Figura 15: Resultat de la consulta d'equips als que s'aposta més.....	59
Figura 16: Resultat de la consulta de guanys i despeses per sexes	59
Figura 17: Resultat de competicions i temporades amb més apostes.....	59
Figura 18: Resultat de la consulta de sexes i edats amb un gràfic.....	60
Figura 19: Exportació a MSEXcel.....	60
Figura 20: CUB-uoc-pfc (part 1)	76
Figura 21: CUB-uoc-pfc (part 2)	77
Figura 22: Publicació del cub.....	78
Figura 23: Publicació del cub.....	79
Figura 24: Creació d'un nou anàlisi	80
Figura 25: Primera vista de l'anàlisi.....	81
Figura 26: desplegament de l'anàlisi	82
Figura 27: Reorganització de les files, columnes i filtres	83
Figura 28: Anàlisi ja reorganitzat	84
Figura 29: Distribució de zones dels apostadors	85
Figura 30: Distribució per edats, apostes i guanys dels apostadors	86
Figura 31: Distribució per equips apostats.....	87
Figura 32: Distribució per sexe dels apostadors	88
Figura 33: Distribució per competicions i temporades	89
Figura 34: Combinació d'edat i sexe amb gràfic	90
Figura 35: Diagrama de Gantt del projecte.....	91

Capítol 1: Introducció

El primer pas d'aquesta memòria és conèixer l'enunciat, els objectius que cal complir, la metodologia de treball i els productes finals entregables.

1.1 Justificació del PFC i context en el qual es desenvolupa: punt de partida i aportació del PFC

1.1.1 Enunciat

La nostra empresa vol desenvolupant un sistema de gestió d'apostes de futbol.

El nostre equip serà l'encarregat d'analitzar els requeriments, així com la posterior implementació de la base de dades per satisfer-los. Tota la gestió i accés a la informació es farà mitjançant procediments de base de dades, sent aquesta l'única manera d'accedir-hi.

La BD ha de poder emmagatzemar dades esportives, és a dir, dades d'equips i lligues de diferents països, resultat dels partits, nom dels jugadors que han marcat gol i en quin ordre, etc. S'ha de tenir en compte que la BD ha de gestionar més d'una temporada per a oferir als usuaris dades històriques de resultats, historial esportiu de cada jugador, gols marcats, etc.

La BD també ha de permetre la gestió d'apostes: identificació de l'usuari, import de l'aposta, modalitat (victòria o empat, jugador que marcarà primer, resultat al final del partit, resultat a la fi de la primera part, etc.).

La BD haurà de ser escalable per poder anar incorporant progressivament totes aquelles necessitats que sorgeixen durant la seva vigència.

També es voldria definir un magatzem de dades (*data warehouse*) per extreure estadístiques i saber, per exemple, la distribució geogràfica dels usuaris, edats, despeses per intervals d'edat, equips als que s'aposta més, etc.

Finalment, per tal de facilitar el manteniment del sistema, es valorarà molt disposar de mecanismes que permetin resoldre potencials problemes d'integració amb la resta del sistema: un log de les accions fetes amb la BD, mecanismes per testejar la funcionalitat de la BD, etc.

Objectius:

- Detectar quines són les necessitats bàsiques del sistema.
- Detectar possibles funcionalitats addicionals de valor afegit.
- Proposar un disseny que s'ajusti als requeriments exposats.
- Implementar un sistema que encapsuli les funcions d'accés a les dades.

1.1.2 Justificació

Aquesta memòria queda justificada en formar la part més important del Projecte de Final de Carrera i es desenvolupa en un context purament acadèmic tot i que indicaré certs aspectes a tenir en compte si es volgués dur a terme en un entorn de producció.

1.2 Objectius del PFC

L'objectiu d'aquest projecte és el desenvolupament d'un sistema de gestió d'apostes de futbol que desglossaré en aquesta sèrie de subobjectius:

- Creació de l'estructura operacional de la base de dades.
- Creació dels disparadors i procediments emmagatzemats per l'accés i modificació de la informació.
- Creació de l'estructura de magatzem (data warehouse) de la base de dades.
- Creació dels processos ETL de traspàs de l'estructura operacional a la de magatzem.
- Creació de mecanismes de resolució de problemes.
- Redacció de la memòria del projecte.
- Redacció de la presentació del projecte.

1.3 Enfocament i mètode seguit

Usaré la metodologia tradicional de projectes i d'anàlisi descendent per la realització d'aquest projecte, empaquetant els grups de requeriments funcionals per poder realitzar les proves de forma independent i així anar avançant en fases el desenvolupament del projecte (figura 1).

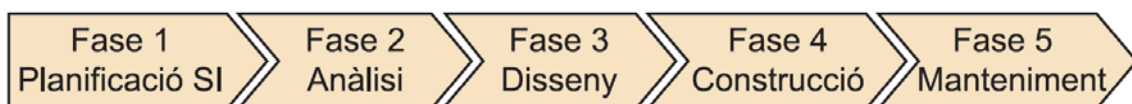


Figura 1: Cicle de vida del desenvolupament d'un projecte

En aquest cas cal modificar el títol de la Fase 5 Manteniment per Tancament i Entrega del projecte, ja que no s'establirà un manteniment d'aquest projecte perquè és acadèmic.

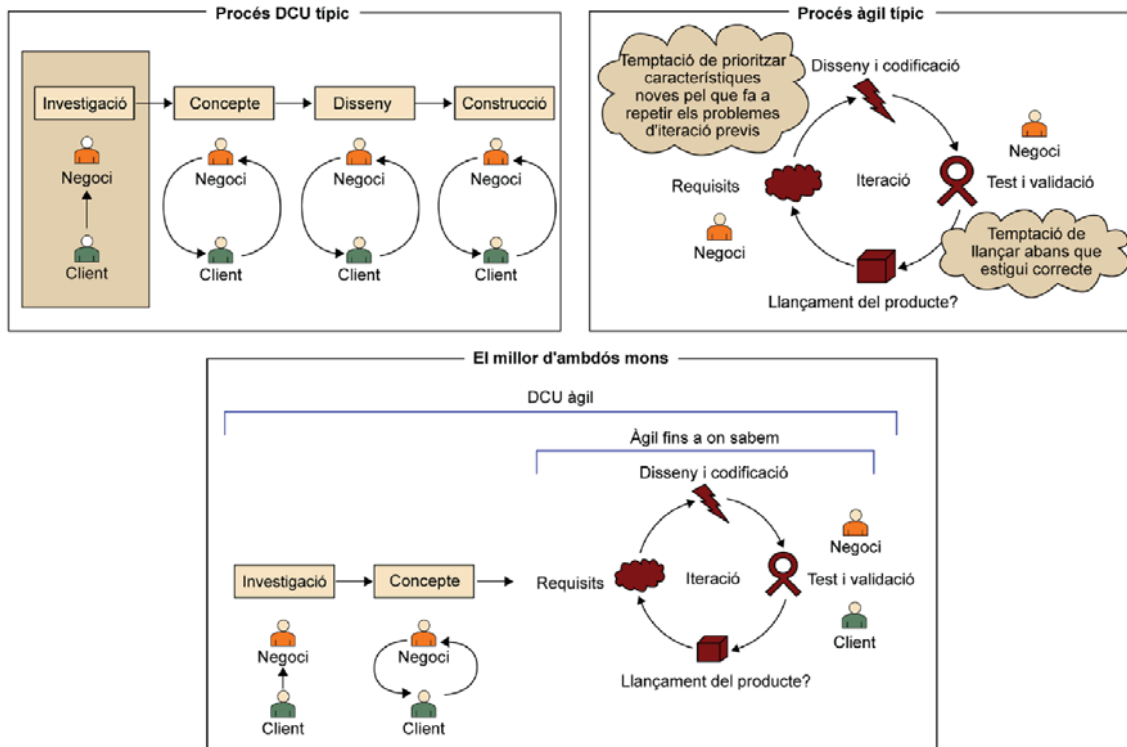
A més hi haurà una fase alternativa i paral·lela al projecte dedicada al seguiment i control del projecte (figura 2)



Figura 2: Fase de seguiment d'un projecte

En el cas de poder realitzar una interfície d'usuari bàsica usaria la combinació de la metodologia àgil i el disseny centrat a l'usuari que crec que serà més útil en aquests casos (figura 3).

Combinació de les aproximacions del desenvolupament àgil i el DCU



Font: adaptat d'Anthony Colfelt (2010). "Bringing User Centered Design to the Agile Environment". *Boxes and Arrows* (abril, 2011).
<<http://www.bboxesandarrows.com/view/bringing-user>>

Figura 3: Combinació de DCU i àgil

1.4 Planificació del treball

Al capítol 2 detallaré la planificació dels treballs i les tasques a realitzar però crec que seguint amb la metodologia de projectes cal associar cada PAC a una fase del projecte i on les dates d'entrega són les següents:

- PAC1 entregada el dia 16/03/2014
- PAC2 entregada el dia 13/04/2014
- PAC3 entregada el dia 11/05/2014
- Entrega final: memòria, presentació i productes el dia 15/06/2014

PAC 1		PAC 2	PAC 3	Lliurament final
Planificació	Anàlisi	Disseny	Implementació	Tancament

1.5 Productes obtinguts

Els productes obtinguts com a resultat del projecte són tota una sèrie de Scripts sql que executats de forma ordenada donen com a resultat les bases de dades operacional i de data warehouse que engloben tant les dades en taules com les funcionalitats en paquets pl/sql.

1.6 Breu descripció dels altres capítols de la memòria

La resta de la memòria està composta pels següents capítols:

Capítol 2: Planificació

Ampliaré la planificació indicada al punt 1.4 indicant totes les tasques a realitzar i el temps necessari per realitzar-les.

Capítol 3: Requeriments

Enumeraré els requeriments tant de maquinari, de programari, funcionals, d'escalabilitat, les eines que usaré i els recursos humans assignats.

Capítol 4: Casos d'ús

Descriuré els casos d'ús extrets de l'anàlisi dels requeriments i de l'enunciat, enumerant-los i agrupant-los per una millor gestió.

Capítol 5: Disseny

Mostraré els diagrames conceptuals, lògics i físics tant de la base de dades operacional com de la data warehouse.

Capítol 6: Implementació

Enumeraré les funcions implementades relacionant-les amb els casos d'ús implementats.

Capítol 7: Valoració econòmica

Realitzaré un resum econòmic indicant els perfils integrants i les hores usades.

Capítol 8: Conclusions

Donaré la meua opinió al respecte d'aquest projecte indicant els aspectes que crec més importants.

Bibliografia

Faré una llista de les referències bibliogràfiques que he usat explícitament pel desenvolupament d'aquest projecte.

Annexos A, B, C i D

Indicaré la llista de scripts sql donats com a resultat final i entregable del projecte, els resultats de les proves, tant en text com en imatges, i el diagrama de Gantt de la planificació.

Capítol 2: Planificació

Hi ha 5 fites ineludibles a aquest projecte i són les entregues de les 3 PACs, la memòria, la presentació i el producte i el tribunal virtual.

2.1 PAC 1: Planificació i Anàlisi (del 28/02 al 16/03)

<i>Temps assignat a aquesta fase:</i>	16d	= 80h
Especificació de requeriments	3d	= 15h
Investigació de les modalitats d'apostes	1d	= 5h
Altres requeriments	0,5d	= 2:30h
Definició de casos d'ús	0,5d	= 2:30h
Llistat de tasques i fites	1d	= 5h
Disseny de la planificació	2d	= 10h
Creació del document PAC 1	4d	= 20h
<i>Farcit</i>	(4d)	

Entrega PAC 1 el dia 16 de març de 2014

2.2 PAC 2: Disseny i creació (del 17/03 al 13/04)

<i>Temps assignat a aquesta fase:</i>	28d	= 140h
<i>Impossibilitat de treballar del 17/03 al 21/03</i>	5d	= 25h
Disseny conceptual	3d	= 15h
Disseny físic	3d	= 15h
Creació de l'esquema	0,4d	= 2h
Creació de les taules	0,8d	= 4h
Creació de les claus foranes	0,8d	= 4h
Creació dels índexs	1d	= 5h
Proves unitàries	1d	= 5h
Disseny de fets i dimensions	2d	= 10h
Disseny de ETL	2d	= 10h
Proves unitàries	1d	= 5h
Creació del document PAC 2	4d	= 20h
<i>Farcit</i>	(4d)	

Entrega PAC 2 el dia 13 d'abril de 2014

2.3 PAC 3: Implementació i proves (del 14/04 al 11/05)

<i>Temps assignat a aquesta fase:</i>	28d	= 140h
Implementació procediments emmagatzemats	7d	= 35h
Implementació dels disparadors de les taules	3d	= 15h
Proves d'integració	4d	= 20h
Implementació del DW	4d	= 20h
Proves del DW	2d	= 10h
Creació del document PAC 3	4d	= 20h
<i>Farcit</i>	(4d)	

Entrega PAC 3 el dia 11 de maig de 2014

2.4 Lliurament Final: Memòria i presentació (del 12/05 al 15/06)

<i>Temps assignat a aquesta fase:</i>	<i>35d</i>	<i>= 175h</i>
Creació de la memòria	20d	= 100h
Creació de la presentació	7d	= 35h
Creació del treball pràctic	3d	= 15h
<i>Farcit</i>	<i>(5d)</i>	
Entrega Lliurament Final el dia 15 de juny de 2014		

2.5 Tribunal Virtual

Tribunal els dies 25 a 27 juny de 2014

2.6 Avaluació de riscos i incidències previstes

El risc més gran que crec que hi ha és el temps. De fet la falta de temps o la impossibilitat d'usar-lo.

El meu calendari laboral m'obliga a usar les tardes com a horari de treball de la realització d'aquest projecte.

He definit aquest horari com:

- 1 dia de treball = 5 hores (16:00h a 21:00h)
- 1 setmana de treball = 7 dies (de DL a DG)

Tinc planificada una incidència dins del pla de projecte:

- La setmana del 17 al 21 de març no estaré disponible degut a un viatge de feina.

Procuraré planificar com a mínim 4 dies de marge (40h de treball) en cadascuna de les PACs com a esponja d'incidències no previstes (aproximadament el 15% del temps de treball de cada PAC) i l'anomenaré *Farcit*.

2.7 Diagrama de Gantt

Veure l'annex D.

Capítol 3: Requeriments

En una primera lectura de l'enunciat es poden deduir aquesta sèrie inicial de requeriments:

- Requeriments de maquinari:
 - Punt de treball estàndard de la UOC.
- Requeriments de programari:
 - Oracle
- Requeriments funcionals bàsics:
 - Gestió i accés únic a la informació mitjançant procediments de base de dades.
 - Gestió de dades esportives.
 - Gestió d'apostes esportives.
 - Gestió de magatzem de dades (*data warehouse*).
 - Gestió de problemes: log d'accions, test de la base de dades, ...
- Requeriment d'escalabilitat

Procedeixo ara a desenvolupar-los amb més profunditat:

3.1 Requeriments de maquinari

3.1.1 Punt de treball estàndard de la UOC

Realitzaré aquest projecte usant el maquinari del qual dispo: (figura 4):

1 CPU DualCore 3 GHZ
3 GB Ram
160 GB Disc

De tota manera si aquest projecte es portés a la pràctica els requeriments de maquinari reals serien diferents:

1 Servidor de Base de Dades
2 CPU DualCore 3 GHz
8 GB Ram
2 TB Disc

1 Servidor d'Aplicacions Web
2 CPU DualCore 3 GHz
4 GB Ram
500 GB Disc



Figura 4: maquinari de què dispo

3.2 Requeriments de programari

3.2.1 Oracle

L'únic requeriment de programari és la utilització de Oracle com a SGBDR.

En aquest projecte utilitzaré la versió Oracle XE 10.2 que no hauré d'instal·lar degut a PACs anteriors.

3.3 Requeriments funcionals bàsics

3.3.1. Gestió i accés únic a la informació usant procediments de base de dades

Aquest requeriment imposa una forma de treballar molt específica al moment de realitzar la implementació.

El requeriment de programari Oracle serà el responsable de donar suport a aquest requeriment fent que l'accés a les taules sigui prohibit a tots els usuaris a excepció del propietari del esquema i caldrà programar tota una sèrie de funcions per proporcionar l'accés CRUD a les dades.

3.3.2 Gestió de dades esportives

Aquest requeriment generarà el primer grup de funcionalitats del sistema permetent la introducció de dades esportives relatives a les lligues, equips, jugadors, temporades, gols marcats, resultats finals, etc.

Cal dissenyar-ho pensant en la necessitat de disposar de dades històriques.

3.3.3 Gestió d'apostes esportives

Aquest requeriment permetrà les funcionalitats necessàries per la gestió dels usuaris que aposten, import de l'aposta i de quina modalitat d'aposta.

He investigat més sobre les modalitats perquè crec que cadascuna de les modalitats requereix un tractament especial i diferent i això ho explico al punt 3.4.

També crec que haurà d'existir un procés encarregat de resoldre cadascuna de les apostes introduïdes per decidir si s'ha encertat o no.

Enlloc es parla tampoc de la forma de pagament de les apostes, tot i que probablement hauria d'usar-se un sistema de pagament via targeta de crèdit o d'altres formes com PayPal.

3.3.4 Gestió de magatzem de dades (*data warehouse*)

Aquest requeriment respon a la necessitat d'anàlisi de les dades històriques del sistema.

Per poder-ho fer amb garanties caldrà dissenyar un nou model de dades amb fets i dimensions per tal de respondre a les estadístiques que se'ns demanen.

Crec que el programari Pentaho serà el més idoni per tal de realitzar aquests anàlisis amb la informació tractada usant la tècnica ETL.

3.3.5 Gestió de problemes

Aquest requeriment proposa una sèrie de procediments per obtenir informació de les tasques realitzades per la base de dades o poder testear les funcionalitats de la base de dades.

3.4 Investigació de les modalitats d'apostes

Es poden apostar a les lligues de tots els països del món, als partits internacionals oficials i amistosos i a les grans competicions com la lliga de campions, l'Europa ligue o el campionat mundial de brasil 2014.

L'aposta principal és saber quin dels dos equips guanya o si finalment hi ha un empat, anomenada 1X2. Per cadascun d'aquests tres resultats hi ha un valor numèric associat que és el valor del guany de l'aposta.

També hi ha dades relatives a la data i l'hora del partit.

Altres modalitats d'apostes són moltíssimes, per exemple:

- quin equip marcarà primer,
- en quin minut es marcarà el primer gol,
- quin serà el jugador que marqui el primer gol,
- quin jugador marcarà més gols,
- resultat a la mitja part,
- resultat final del partit, ...

També es poden escollir apostes combinades amb quin jugador marcarà primer i quin equip guanyarà.

3.4.1 Pàgines visitades

- bWin - Apuestas de futbol <https://sports.bwin.es/es/sports#sportId=4>
- MiApuesta – futbol <http://miapuesta.es/apuestas-de-futbol/>
- William Hill – futbol http://sports.williamhill.es/bet_esp/es

3.5 *Requeriments d'escalabilitat*

Els requeriments de maquinari i de programari són crítics a l'hora de definir i garantir l'escalabilitat del projecte.

Si es definís aquest projecte per un entorn de producció real s'hauria de pensar inicialment en el maquinari que he definit al punt 3.1.1.

Pel que fa al programari, el requeriment d'utilització de Oracle ja ens proporciona una garantia d'escalabilitat i tot un conjunt d'eines de monitorització per fer un seguiment de l'estat de la base i les necessitats de capacitat i escalabilitat.

L'escalabilitat del programari usat es pot traduir en escalabilitats verticals (ampliació del maquinari de la base de dades) o horitzontals (clusterització del maquinari que suposarà un canvi de versió del programari a Oracle RAC).

D'igual manera passaria amb el maquinari servidor d'aplicacions. Es podria ampliar verticalment o horitzontalment usant per exemple un mòdul de proxy d'Apache.

3.6 *Altres requeriments*

Quan he investigat les modalitats d'apostes he visitat unes quantes webs que es dediquen al món de les apostes. I degut a que són webs la interfície d'usuari és molt important: ha de ser clara, fàcil, ràpida, ...

3.6.1 Interfície d'usuari

L'enunciat del projecte no el menciona però queda clar que d'alguna manera l'usuari s'ha de comunicar amb el sistema. El requeriment d'accés és qui marcarà la forma com la interfície d'usuari (crec que hauria de ser web sobre un servidor d'aplicacions Java o PHP, per exemple) haurà d'executar les crides als procediments emmagatzemats de la base de dades.

3.6.2 Requeriments operacionals

Tampoc tindrè en compte tota la sèrie de tasques i processos derivats de la gestió operacional estàndard tant de la base de dades com del possible servidor d'aplicacions, ni de les còpies de seguretat, ni de la disponibilitat, etc.

3.7 Eines i materials a usar

Les eines de treball que usaré sota el sistema operatiu Windows XP SP3 pel disseny i implementació són (en ordre alfabètic):

- Adobe Acrobat Professional 9
- Magic Draw Personal 17.0
- Microsoft Office Project 2003
- Microsoft Office Visio 2003
- Microsoft Office Word 2003
- Oracle Database Express Edition 11.2
- SQLDeveloper 3.0.04

Tot aquest programari ja està instal·lat a l'ordinador de treball.

EL programari d'execució dels informes de data warehouse és Pentaho que s'executa a una màquina virtual Ubuntu Linux 10.11 que ja vaig usar a l'assignatura Models Multidimensionals i Magatzems de Dades i està composta pels programaris:

- Pentaho User Console (Pentaho BI Platform 3.10)
- Pentaho Schemma Workbench 3.3
- Pentaho Data Integrator 4.2

3.8 Recursos Humans

Usaré tres perfils diferents per la resolució d'aquest projecte:

- **Director de projecte:** Encarregat del seguiment i control del projecte i de la revisió i validació dels documents.
- **Analista:** Encarregat de la captura de necessitats, anàlisi de requeriments, definició de casos d'ús i disseny conceptual i físic, disseny de magatzem i dels processos ETL.
- **Programador:** Encarregat de la implementació del disseny de la base de dades i dels disparadors, dels procediments emmagatzemats i de la implementació del magatzem i dels processos ETL.

La valoració econòmica dels recursos humans la faré al capítol 7.

Capítol 4: Casos d'ús

Com a resultat de l'anàlisi dels requeriments funcionals i abans de començar amb el disseny, cal acabar de perfilar i especificar els requeriments funcionals del sistema i quins seran els casos d'ús que finalment s'implementaran i quines seran les dades tractades en cada cas.

4.1 Diagrama de Casos d'ús

Aquest és el diagrama de casos d'ús que desenvoluparé:

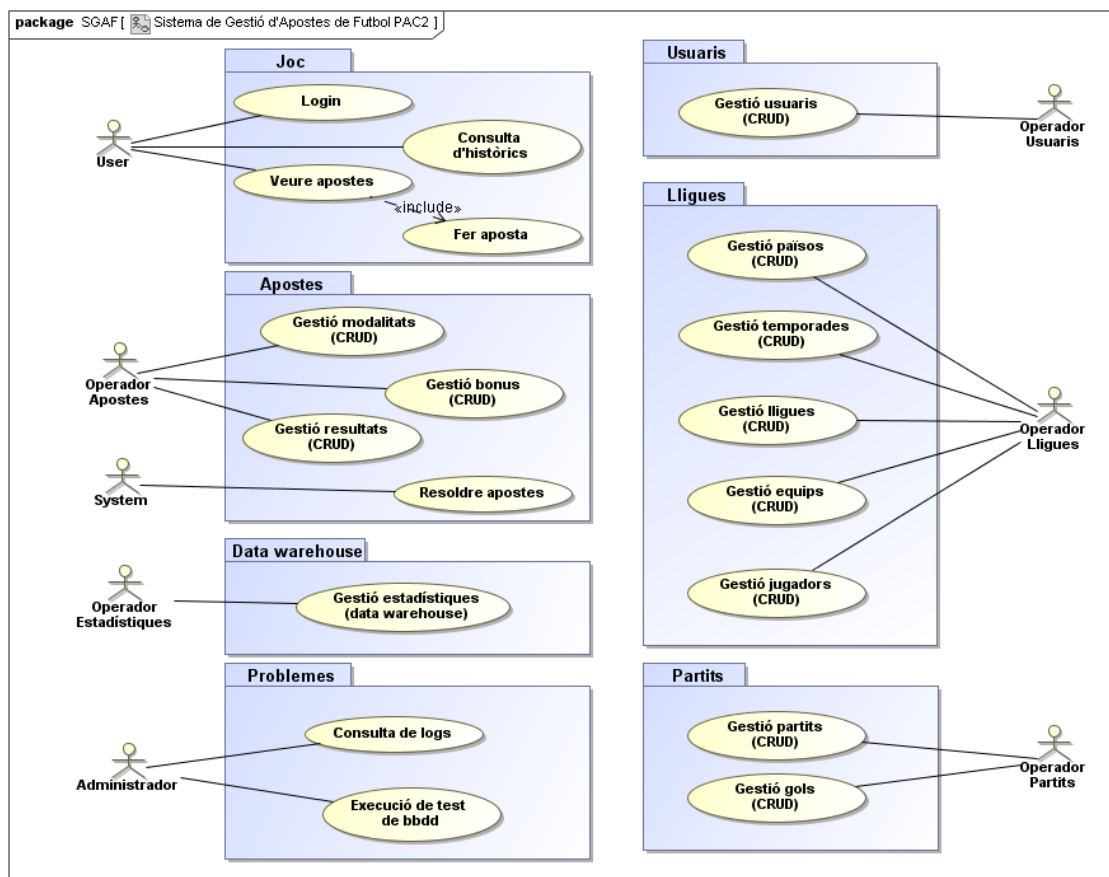


Figura 5: Casos d'ús

Aquest empaquetament està directament relacionat amb el concepte de divisió en fases funcionals que he indicat en la metodologia usada.

4.2 Descripció dels casos d'ús.

He agrupat els casos d'ús en paquets per facilitar tant la comprensió com el desenvolupament i la modularitat de les proves unitàries.

4.2.1 Paquet d'usuaris

L'operador d'usuaris serà l'actor responsable d'aquest paquet.

CU-1: Gestió d'usuaris

Contindrà un conjunt de rutines per realitzar el manteniment CRUD dels usuaris del sistema: altes, baixes, modificacions i consultes.

Les dades que s'hauran d'emmagatzemar d'un usuari són: el seu nom d'usuari i contrasenya, el seu nom i cognoms, el número del document identificatiu (DNI, NIE, Passaport, ...), la data de naixement, el seu gènere (Home, Dona), el país d'origen o de residència i un número de compte bancari o de targeta de crèdit.

Cal assegurar que l'usuari sigui major d'edat.

Per les investigacions que he realitzat crec que fer que el nom d'usuari sigui el seu correu electrònic estalvia possibles problemes de duplicitat i a més identifica unívocament a l'usuari.

4.2.2 Paquet de competicions(lligues)

L'operador de lligues serà l'actor responsable d'aquest paquet.

CU-2: Gestió de zones (països)

Contindrà un conjunt de rutines per realitzar el manteniment CRUD dels països del sistema.

La dada a emmagatzemar dels països és el seu nom. També incloure les estructures supranacionals com els continents i també el mateix món.

Usaré el nom de Zones enlloc de Països per l'agrupament supranacional.

CU-3: Gestió de competicions (lligues)

Contindrà un conjunt de rutines per realitzar el manteniment CRUD de les lligues del sistema.

Les dades a emmagatzemar de les lligues són: el seu nom i el país on es realitzen. En els cas de lligues supranacionals el país correspondrà al continent o al món.

També inclouré les competicions que es realitzin en format de copa de la mateixa manera que les de lliga.

Usaré el nom de Competicions enlloc de Lligues.

CU-4: Gestió de temporades

Contindrà un conjunt de rutines per realitzar el manteniment CRUD de les temporades del sistema.

Les dades a emmagatzemar de les temporades són: l'any de la temporada, la lliga a la que pertany aquella temporada i les dates d'inici i final de la temporada.

CU-5: Gestió d'equips

Contindrà un conjunt de rutines per realitzar el manteniment CRUD dels equips del sistema.

Les dades a emmagatzemar dels equips són: el nom, la data de la seva fundació, i el país al que pertany.

CU-6: Gestió de jugadors

Contindrà un conjunt de rutines per realitzar el manteniment CRUD dels jugadors del sistema.

Les dades a emmagatzemar dels jugadors són: el nom i cognoms, la data de naixement, i la nacionalitat.

Tota la informació del paquet de competicions caldrà que s'hagi introduït prèviament a les dades del paquet de partits.

4.2.3 Paquet de partits

L'operador de partits serà l'actor responsable d'aquest paquet.

CU-7: Gestió de partits

Contindrà un conjunt de rutines per realitzar el manteniment CRUD dels partits del sistema.

Les dades a emmagatzemar dels partits són: la temporada, els equips que jugaran i la data del partit.

L'ordre dels equips és important ja que els partits d'anada i de tornada contenen les mateixes dades però amb l'ordre dels equips a l'inrevés i evidentment amb dates diferents.

CU-8: Gestió de gols

Contindrà un conjunt de rutines per realitzar el manteniment CRUD dels gols del sistema.

Les dades a emmagatzemar dels gols són: el partit, l'equip i el jugador que ha marcat, el minut i de quina forma l'ha marcat (falta, corner, penal,...).

En cas d'haver temps afegit al partit aquest mai passarà de 5 minuts.

Tota la informació del paquet de partits caldrà que s'hagi introduït prèviament a les dades del paquet d'apostes.

4.2.4 Paquet de joc

L'usuari serà l'actor responsable d'aquest.

CU-9: Login

Contindrà un conjunt de rutines per validar l'autenticitat de l'usuari.

Per ampliar la seguretat el sistema haurà de bloquejar l'autenticació de l'usuari en cas d'existir més d'un error de contrasenya al validar-se.

CU-10: Consulta d'històrics

Contindrà un conjunt de rutines per mostrar les dades històriques de resultats, historials de gols marcats dels jugadors,...

S'haurà de poder filtrar per zona, competició, temporada, equip, ...

CU-11: Veure apostes

Contindrà un conjunt de rutines per mostrar les apostes que ha realitzat un usuari.

S'haurà de poder filtrar per modalitat d'aposta, per dates, ...

CU-12. Fer apostes

Contindrà un conjunt de rutines per realitzar les apostes d'un usuari.

Les apostes només es podran realitzar de les modalitats que s'hagin afegit al sistema prèviament.

Tot i que en aquest sistema el pagament de les apostes serà fictici, s'haurà de validar que l'usuari realitzi el pagament de l'aposta usant les dades bancàries o de targeta o també disminuint el saldo del qual disposa l'usuari en el sistema.

4.2.5 Paquet d'apostes

Aquest paquet és nou i conté tres nous casos d'ús: CU-13, CU-14 i CU-15.

L'operador d'apostes serà l'actor responsable d'aquest paquet amb l'excepció de la resolució d'apostes que serà responsabilitat del sistema.

CU-13: Gestió de modalitats

Contindrà un conjunt de rutines per realitzar el manteniment CRUD de les modalitats d'apostes del sistema (1X2, resultat final, etc).

Les dades a emmagatzemar de les modalitats d'apostes és: el nom.

CU-14: Gestió de coeficients

Contindrà un conjunt de rutines per realitzar el manteniment CRUD dels coeficients dels partits del sistema. El concepte de coeficient és el factor multiplicador de l'aposta en cas de resultar guanyadora.

★ Lugo	2.75	X	2.90	★ Alcorcón	2.55
--------	------	---	------	------------	------

Figura 6: Exemple de coeficients 1 X 2

Les dades a emmagatzemar dels coeficients són: la modalitat d'aposta, el partit jugat, el possible resultat i el coeficient resultant.

CU-15: Gestió de resultats

Contindrà un conjunt de rutines per realitzar el manteniment CRUD dels resultats dels partits del sistema.

Les dades a emmagatzemar dels resultats són: la modalitat d'aposta, el partit jugat i el resultat.

CU-16: Resoldre apostes

Contindrà un conjunt de rutines per resoldre les apostes dels partits que ja han acabat.

El sistema usarà les dades dels resultats per resoldre quines apostes han estat guanyadores i afegir els guanys de les apostes als usuaris guanyadors.

4.2.6 Paquet de data warehouse

L'operador d'estadístiques serà l'actor responsable d'aquest paquet.

CU-17: Gestió d'estadístiques

Contindrà un conjunt de rutines destinades a realitzar les tasques ETL per transformar les dades operacionals en dades dissenyades per un entorn de treball data warehouse.

Aquestes dades estadístiques estaran organitzades en dimensions de dades per les que es podran realitzar les recerques dels fets.

4.2.7 Paquet de problemes

L'administrador serà l'actor responsable d'aquest paquet.

CU-18: Consulta de logs

El sistema escriurà el un log totes les accions que se li han demanat d'executar amb el resultat tant sigui correcte com erroni.

L'administrador disposarà d'un conjunt de rutines per poder consultar aquests logs filtrants per data o tipologia.

CU-19: Test de base de dades

Contindrà d'un conjunt de rutines dissenyades per realitzar proves unitàries de totes les dades verificant si els resultats correctes o erronis coincideixen amb els valors esperats de les execucions.

4.3 Actors

He associat un actor diferent a cadascun del paquets de casos d'ús encara que es podrien agrupar els operadors diversos i l'administrador per deixar únicament tres tipologies generals d'actors: Usuaris, Administradors i Sistema.

4.4 Altres restriccions

A més de les restriccions habituals d'unicitat de les dades en aquest cas cal afegir-ne més que només podran ser satisfetes usant la codificació dels procediments emmagatzemats:

- Un Jugador no pot pertànyer a la Plantilla de dos Equips al mateix moment.
- Un Partit ha d'estar compost de dos Equips diferents.
- Un Gol serà marcat per un Jugador d'un dels dos Equips que juguen el Partit.
- Una Aposta es farà sobre un Partit que encara no s'hagi disputat.

Capítol 5: Disseny

Un cop finalitzats l'anàlisi de requeriments i de funcionalitats he arribat a la conclusió que he de realitzar el disseny de dues estructures clarament diferenciades en l'àmbit de les dades: les dades operacionals dins d'una infraestructura OLTP i les dades del magatzem data warehouse.

Les dades del registre de log d'accions de tot el sistema seran una única taula amb les dades de les accions realitzades, l'usuari que les ha executat, la data i hora d'execució i el resultat correcte o incorrecte d'aquella acció.

5.1 Disseny OLTP

El disseny operacional serà l'encarregat d'emmagatzemar i gestionar les dades relacionades amb els processos habituals del sistema d'apostes i manteniment de les dades esportives i d'usuari.

Usa la metodologia d'abstracció de les entitats partint del disseny més genèric que serà el conceptual fins arribar al disseny físic.

5.1.1 Disseny conceptual OLTP

Després de fer-hi moltes voltes i de generar molts esbossos he arribat a aquest esquema conceptual:

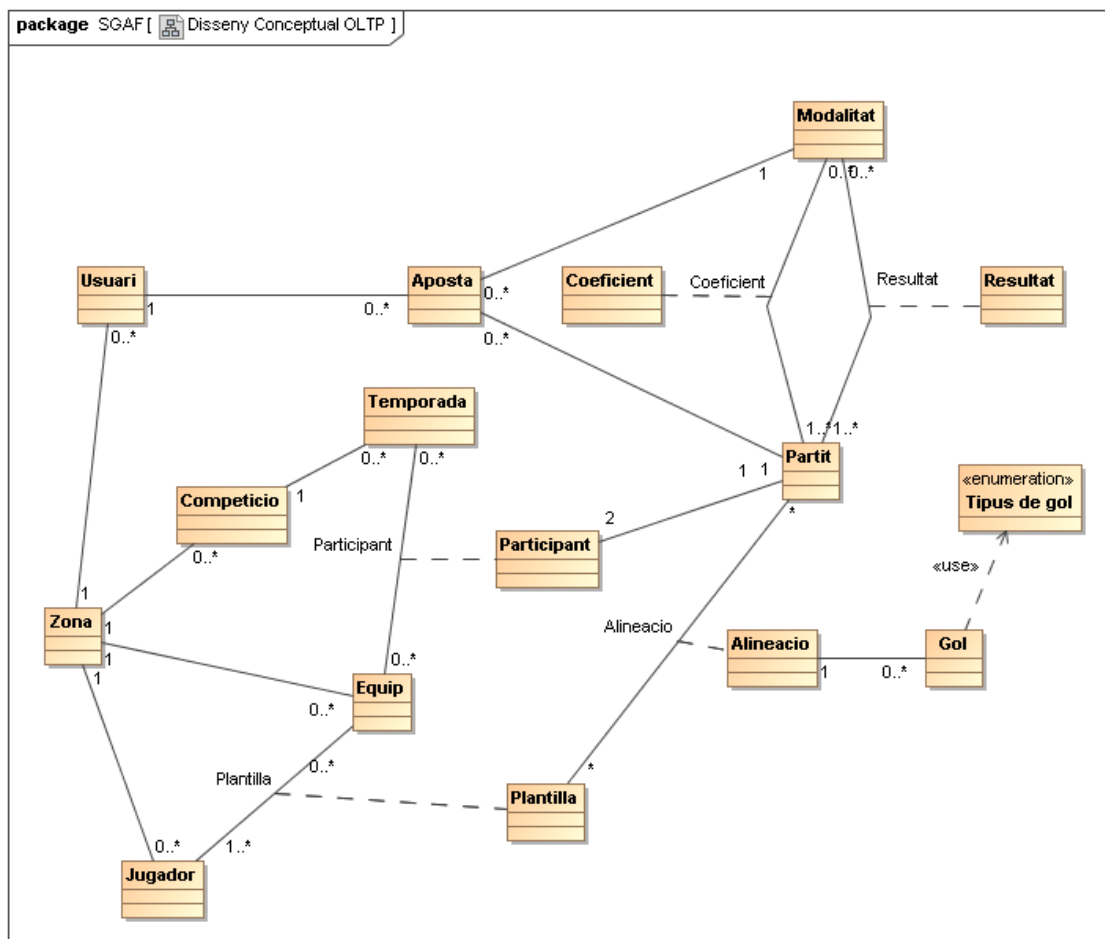


Figura 7: Disseny conceptual OLTP

Aniré explicant totes les entitats, quin és el seu propòsit i amb quines altres entitats estan relacionades.

Zona

Inicialment s'havia anomenat País però degut a que hi ha competicions entre equips de diferents països l'he estès perquè contempli tant països com altres zones supranacionals com Europa, Sud-amèrica o el Món.

Competició

Defineix les competicions que s'organitzen arreu del món com són les lligues nacionals de tots els països: lliga Catalana, lliga BBVA, ..., les competicions europees: lliga de campions, uefa lliga,... i els mundials de futbol o les olimpíades...

Una Competició pertany a una Zona.
Una Zona pot tenir cap o més Competicions.

Temporada

Instancia les competicions futbolístiques al llarg dels anys.

Una Temporada pertany a una Competició.
Una Competició pot tenir cap o més Temporades.

Equip

Seràn els qui participaran als partits de cada temporada.

Un Equip pertany a una Zona.
Una Zona pot tenir cap o més Equips.

Participant

Dins d'una mateixa temporada els equips poden participar en diferents competicions.

Es tracta d'una classe associativa entre Equip i Temporada amb multiplicitat cap o més en ambdós extrems.

Jugador

És una persona contractada per un equip per jugar a futbol.

Un Jugador és nascut a una Zona.
Una Zona pot ser lloc de naixement de cap o més Jugadors.
Un Jugador resideix a una Zona.
Una Zona pot ser lloc de residència de cap o més Jugadors.

Plantilla

És el conjunt de jugadors de què disposa un equip per jugar a futbol en un moment determinat.

Es tracta d'una classe associativa entre Equip i Jugador amb multiplicitat u o més excepte que un jugador pot no pertànyer a cap plantilla.

Partit

És el nucli d'aquest sistema d'apostes i per descomptat del futbol.

Un Partit està compostat per dos equips Participants.

Alineació

Per cada partit els equips preparen les alineacions dels jugadors que hi intervindran, tant els titulars com els suplents.

Es tracta d'una classe associativa entre Partit i Plantilla amb multiplicitat variable que guardarà el dorsal de cada jugador.

Gol

És l'objectiu final dels equips que disputen un partit.

*Un Gol serà marcat per un jugador Alineat.
Un jugador Alineat pot marcar cap o més Gols.*

Tipus de gol

És habitual diferenciar els gols segons la forma com s'han marcat ja sigui de falta, de corner, de penal, ...

Modalitat

Són les tipologies genèriques d'apostes que es permeten fer en el nostre sistema.

Aposta

És la predicció que fa un usuari dels esdeveniments que aquest creu que succeiran en un partit. Només podran fer-se de modalitats existents al sistema.

Poden ser tant el resultat final del partit, a la mitja part o quin jugador marcarà més gols, ...

*Una Aposta serà d'un Partit.
Un Partit pot tenir cap o més Apostes.
Una Aposta és d'una Modalitat.
Una Modalitat pot tenir cap o més Apostes.*

Resultat

Són els fet reals que han succeït en un partit dins les diverses modalitats d'apostes que el sistema admet.

Es tracta d'una classe associativa entre Partit i Modalitat amb multiplicitat u o més excepte que una Modalitat pot no tenir resultats.

Coeficient

És el factor multiplicador dels diners apostats pels usuaris en un partit, per una modalitat i un resultat concret.

Es tracta d'una classe associativa entre Partit i Modalitat amb multiplicitat u o més excepte que una Modalitat pot no tenir coeficient.

Usuari

Són les persones registrades al sistema que poden realitzar apostes en diverses modalitats dels partits que es disputaran en un futur proper.

Un Usuari pot realitzar cap o més Apostes.

Una Aposta li pertany a un Usuari.

Un Usuari és nascut a una Zona.

Una Zona pot ser lloc de naixement de cap o més Usuaris.

5.1.2 Disseny lògic OLTP

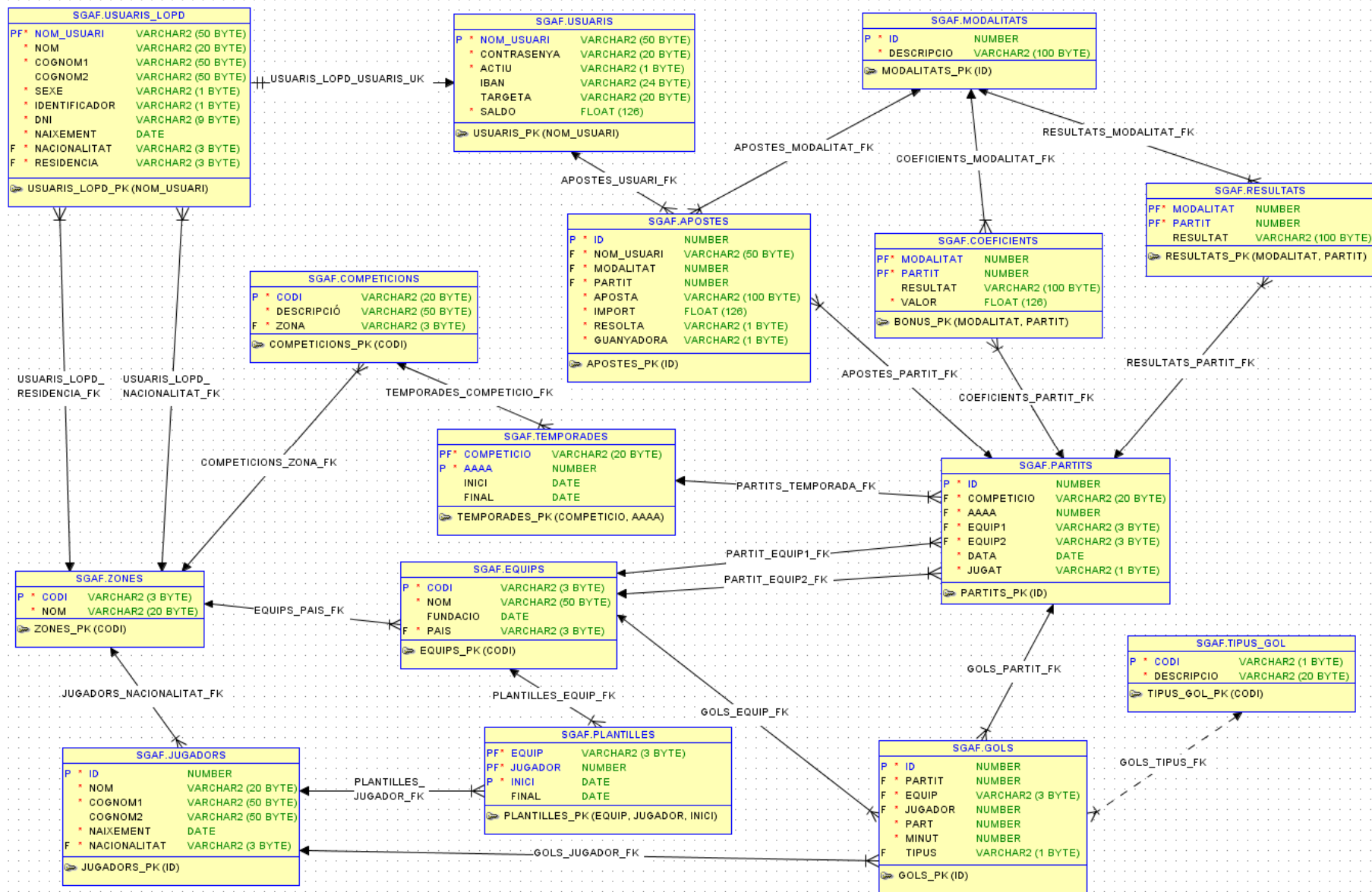


Figura 8: Disseny lògic OLTP

El disseny lògic està adaptat a l'arquitectura requerida del sistema i a la legislació vigent. Per aquest motiu he revisat el disseny conceptual per definir aquest nou disseny lògic de la figura 8.

Aniré explicant els canvis i les motivacions dels mateixos fins a obtenir aquest disseny:

Canvis per disminuir la complexitat sense modificar la funcionalitat

Les entitats Participant, Plantilla i Alineació són les encarregades de gestionar la integritat de la informació envers les Temporades, els Equips, els Jugadors i els Partits.

Però pel sistema d'apostes aquestes entitats no aporten cap valor afegit sinó tot el contrari dificulten la funcionalitat de la operativa del sistema.

El concepte d'equip participant a una temporada ja queda evidenciat per la seva intervenció en un partit (o dos si es compten l'anada i la tornada).

El concepte d'alineació l'he eliminat ja que en cap moment es demana aquesta informació ni el seu tractament històric.

Finalment he relacionat Jugador amb Gol tancant així la simplificació lògica del disseny.

Canvis deguts a l'arquitectura

El requeriment d'usar Oracle com a Gestor de Base de Dades ens obliga a modificar el disseny de les relacions múltiple a múltiple i de les classes associatives generant una nova classe intermèdia on es dipositaran els valors de les relacions.

Aquest és el cas de les classes associatives Plantilla, Coeficient i Resultats que han passat a ser entitats independents.

Canvis deguts a la legislació vigent

La Llei Orgànica de Protecció de Dades de caràcter personal (LOPD 15/1999 del 13 de desembre per adaptar la Directiva 95/46/CE del Parlament Europeu) ens indica que s'han de garantir una sèrie de normes relatives tant a la seguretat de la informació com a l'accés a la mateixa.

Per no entorpir la operativa normal i també per adaptar el disseny a la normativa vigent he decidit disgregar la informació d'usuari en dues classes: Usuari i Usuari LOPD.

La primera no contindrà cap atribut identificatiu de la persona i la segona contindrà totes les dades relatives a la persona que han d'estar sota la tutela de la llei: nom, DNI, sexe, ...

5.1.3 Disseny físic OLTP

Scripts de creació

Un cop realitzats els dissenys conceptuals i lògics cal passar al disseny físic on indicarem tots els atributs que formaran part de les taules i les seves relacions:

(adjunto document SQL de creació de les taules, claus primàries i claus foranes)

Claus primàries

He decidit que en certes taules la clau primària sigui un valor que identifiqui el registre no només per ser clau primària sinó també perquè el seu contingut és entenedor i diferencial.

Aquestes taules són: Zones, Competicions, Temporades, Equips, Usuaris, Usuaris LOPD, i Tipus de gol

Hi ha d'altres taules on he decidit que la clau primària sigui un valor numèric sense cap relació amb les dades degut a que l'única clau candidata és la totalitat dels camps.

Aquestes taules són: Jugadors, Gols, Partits, Modalitats i Apostes.

I per últim tenim les taules derivades de les classes associatives on la seva clau primària està formada per la combinació de les claus primàries de les dues taules relacionades.

Aquestes taules són: Plantilles, Coeficient i Resultats.

Claus foranes i índexs

Totes les claus foranes han estat definides amb nom i són exactament totes les relacions entre les entitats definides.

He generat també tot un conjunt d'índexs començant per indexar tots els camps que formen part de les claus foranes. A mida que vagi desenvolupant segurament n'afegiré algun més per millorar el rendiment de les SELECTs.

(adjunto document SQL de creació dels índexs)

Nul·labilitat

Potser pot semblar que he afegit masses restriccions de no nul·labilitat però les he cregut convenientes totes per indicar no només la seva obligatorietat sinó també perquè les he considerat definidores de seves les entitats.

Proves unitàries

He realitzat proves d'inserció de dades a totes les taules per verificar les integritats referencials però les proves de debò les hauré de fer un cop hagi implementat els procediments emmagatzemats de test.

5.2 Disseny data warehouse

Aquest magatzem de dades ha de respondre a les necessitats indicades a l'enunciat per obtenir estadístiques i resums complexos d'informació.

Encara que s'indiquen certes possibilitats d'estadístiques i consultes, no hi ha peticions específiques a complir, per tant afegiré els requeriments que considero oportuns

5.2.1 Requeriments

Els requeriments que caldrà complir són:

- Distribució geogràfica dels usuaris
- Distribució per edats dels usuaris
- Despeses per intervals d'edats
- Equips als que s'aposta més

A més de:

- Intervals d'edat que produeixen més guanys a l'empresa
- Dies en què s'aposta més
- Competicions a les que s'aposta més
- Temporades a les que s'ha apostat més
- Usuaris que han apostat més
- Equips que produeixen més guanys a l'empresa
- Competicions que produeixen més guanys a l'empresa
- Temporades que ha produït més guanys a l'empresa
- Usuaris que ha produït més guanys a l'empresa

N'he afegit més dels indicats a l'enunciat ja que els he considerat importants per poder aportar informació per prendre decisions a la capa directiva.

5.2.2 Fets

Els fets són les entitats que ens proporcionen la informació nuclear del sistema i dels informes que caldrà dissenyar.

Llavors sembla que els fets haurien de ser Usuaris i Apostes, però degut a la renormalització de la informació del magatzem tota la informació necessària d'usuari ja estarà continguda a les apostes. Per tant només caldrà definir les apostes com a fet del magatzem.

5.2.3 Dimensions

Les dimensions són la informació per la que podem fer recerques, agrupacions o filtres dels fets.

Definiré com a dimensions les entitats: Zona, Data, Sexe, Edat, Usuari, Equip, Competició, Temporada i Modalitat

He afegit Usuari per poder mostrar tota la informació d'usuari que no està continguda en el fet Aposta.

5.2.4 Indicadors o mètriques

Els indicadors o mètriques són les dades sumarem, restarem, agruparem,... dins de les entitats dels fets per poder-les mostrar als informes.

Els indicadors que defineixo al fet Aposta son: Diners apostats (numèric), Aposta guanyadora (booleà), Diners pagats a aqueta aposta(numèric).

El tercer indicador serà zero en cas que el segon indicador sigui fals (no pagarem cap aposta que no hagi estat guanyadora).

5.2.5 Atributs

Els atributs són la informació que finalment emmagatzemaran tant les entitats dels fets com les entitats de les dimensions i seran els valors que s'utilitzaran per filtrar els resultats segons ho demani l'usuari del magatzem.

Primer defineixo els atributs dels fets:

Entitat	Atribut
Fets	
Aposta	diners apostats, aposta guanyadora, diners pagats a aqueta aposta

I posteriorment els atributs de les dimensions:

Entitat	Atribut	Jerarquia
Dimensions		
País	codi, nom	* Codi, nom
Data	any, mes, dia, setmana de l'any, dia de la setmana, hora, minut, segon	* any, setmana de l'any, mes, dia, hora, minut, segon * dia de la setmana
Sexe	sexe	* sexe
Edat	edat	* edat
Equip	codi, nom	* codi, nom
Competició	codi, nom	* codi, nom
Temporada	codi, any, nom	* codi, nom * any
Modalitat	codi, nom	* codi, nom
Usuari	codi usuari	* codi usuari

Cal tenir en compte que a les dimensions hi ha associada la jerarquia dels atributs en el cas de voler agrupar per algun d'aquests atributs.

5.2.6 Disseny conceptual data warehouse

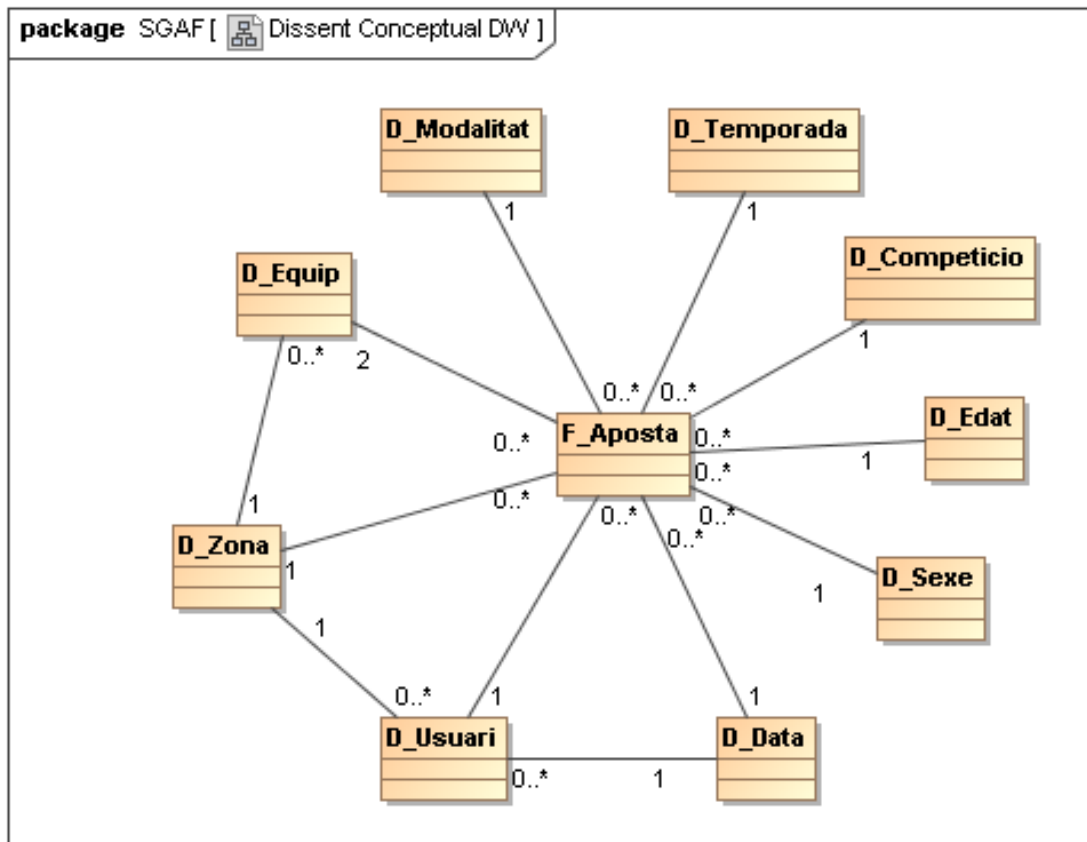


Figura 9: Disseny conceptual data warehouse

He afegit els conceptes de relació entre dimensions per indicar que es podrien filtrar les dades usant valors derivats combinats, com p.e.: Diners apostats en campionats mundials a equips d'un país concret.

Fets

F_Aposta

Dimensions

D_Zona, D_Data, D_Usuari, D_Sexe, D_Edat, D_Equip, D_Competicio, D_Temporada, D_Modalitat

Relacions

1..N entre Fets i Dimensions,
1..N entre D_Data i D_Usuari
1..N entre D_Zona i D_Usuari
1..N entre D_Zona i D_Equip
2..N entre D_Equip i F_Aposta

5.2.7 Disseny lògic data warehouse

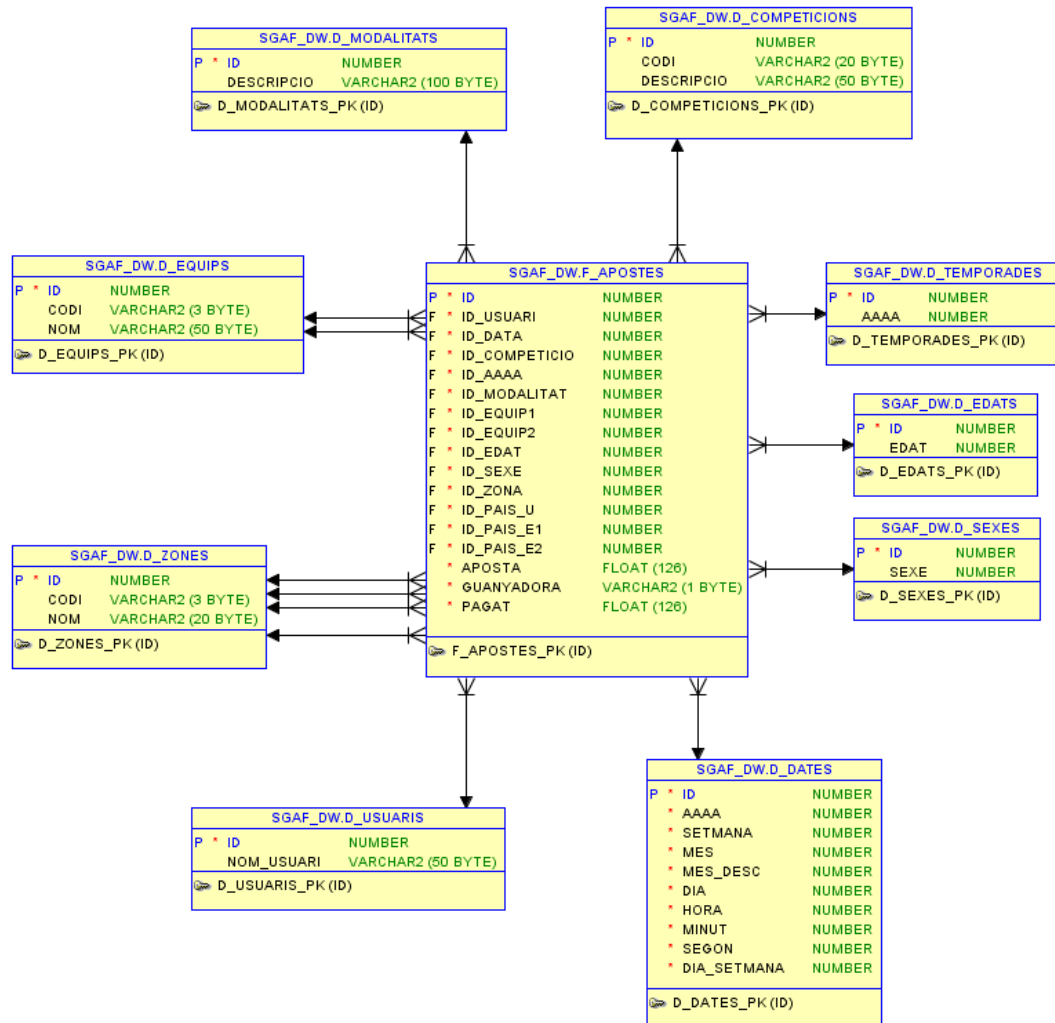


Figura 10: Disseny lògic data warehouse

Canvis respecte del disseny conceptual

He modificat la relació de **D_Zona** amb **F_Aposta** afegint atributs de les zones dels 2 equips i la zona de l'usuari. Per tant queda una relació de 4 zones per cada aposta.

Fets

F_Aposta

Dimensions

D_Zona, **D_Data**, **D_Usuari**, **D_Sexe**, **D_Edat**, **D_Equip**, **D_Competicio**, **D_Temporada**, **D Modalitat**

Relacions

k..N entre Fets i Dimensions

5.2.8 Disseny físic data warehouse

Al disseny físic s'ha hagut d'afegir tota una sèrie d'identificadors arbitraris que seran les claus primàries de les taules per poder-les relacionar.

(adjunto document SQL de creació de les taules, claus primàries i claus foranes)

Proves unitàries

He realitzat proves d'inserció de dades a totes les taules per verificar les integritats referencials però les proves de debò les hauré de fer un cop hagi implementat els procediments emmagatzemats ETL.

5.2.9 Disseny ETL

ETL són el conjunt de processos (extracció, transformació i càrrega) encarregat de preparar el sistema de magatzem a partir del sistema operacional i està profundament lligat a l'arquitectura i a la implementació .

El procés d'extracció està lligat sobre tot a l'arquitectura.

Tal i com he definit l'arquitectura d'aquest sistema crec que la millor manera serà la lectura directe de les dades de l'esquema operacional des de l'esquema de magatzem.

Si es tractés d'un sistema de producció llavors l'extracció seria diferent. Es faria una extracció real de les dades des del servidor operatiu i aquestes dades s'importarien tal qual sobre un nou esquema però en un altre servidor, el servidor de magatzem.

Els processos de transformació i el de càrrega estan lligats sobre tot a la implementació.

Primer transformem les dades operacionals en les dades de les dimensions, usant codis arbitraris com a claus primàries d'aquestes noves taules.

I després es genera la taula dels fets realitzant lectures sobre les taules de dimensions si cal obtenir algun codi d'elles.

Un cop transformada una dada es realitza la càrrega sobre la nova taula de magatzem.

El procés unificat de ETL que construirà el sistema de magatzem estarà implementat sobre el servidor de base de dades Oracle usant el llenguatge de programació PL/SQL.

En cas d'haver de definir una continuïtat d'aquest procés ETL es podria automatitzar fent un script que primer esborrés el magatzem i després realitzés la lectura i la transformació. Tot això ho faria usant OracleJobs per planificar les execucions d'aquesta tasca.

Capítol 6: Implementació

Un cop acabada la fase de disseny i creació cal començar la fase d'implementació del codi associat al nostre sistema.

En aquest cas, al codi s'associen tots els aspectes relacionats amb les dades CRUD, tant la creació, la lectura, la modificació i l'esborrat degut a la restricció de l'enunciat de gestió i accés a la informació.

6.1 Paquets

He implementat els paquets funcionals que havia definit al capítol 4: casos d'ús i els he separat en paquets perquè la gestió d'accessos també sigui coherent amb les definicions dels actors del punt 4.2 i 4.3.

La norma general dels paquets és que tots tenen funcions per realitzar les següents tasques:

- Alta element
- Baixa element
- Modifica element
- Llista elements
- Un element
- Cerca element

Les funcions de recerca només les he creat per les taules on la clau primària és prové de la seqüència, ja que és arbitrària i no té cap valor semàntic tal i com he indicat al punt 5.1.3.

Aquestes taules són: Jugadors, Gols, Partits, Modalitats i Apostes.

Cada funció queda definida per quatre aspectes: paràmetres d'entrada, resultat, precondicions i funcionalitat.

Per no estendre ni complicar les definicions només he indicat aquest aspectes en dos dels casos d'ús diferents: Zones (CU-2) i Jugadors (CU-6), així indicaré a la resta de casos d'ús a quina de les dues s'assembla:

Són similars al CU-2 els casos d'ús: CU-1, CU-3, CU-4, CU-5, CU-14, CU-15.

Són similars al CU-6 els casos d'ús: CU-7, CU-8, CU-13, CU-16.

La resta de casos d'ús tenen funcionalitats diferents i no agrupables.

6.1.1 pkg_competicions

L'operador de lligues serà l'actor responsable d'aquest.

Funcions de zones: (CU-2)

Usaré aquest grup per explicar el funcionament, indicant als altres grups si el seu funcionament és com aquest o com algun altre.

Per no repetir els mateixos conceptes només posaré els paràmetres d'entrada de les funcions d'alta.

- *alta_zona*
Paràmetres d'entrada: codi VARCHAR2, nom VARCHAR2
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: la zona identificada pel codi no ha d'existir.
Funcionalitat: alta d'una nova zona amb el seu codi i el seu nom.
- *baixa_zona*
Paràmetres d'entrada: codi VARCHAR2
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: la zona identificada pel codi ha d'existir.
Funcionalitat: baixa d'una zona identificada pel seu codi.
- *modifica_zona*
Paràmetres d'entrada: codi VARCHAR2, nom VARCHAR2
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: la zona identificada pel codi ha d'existir.
Funcionalitat: modificació del nom d'una zona identificada pel seu codi.
- *llista_zona*
Paràmetres d'entrada: cap
Resultat: CURSOR.
Precondicions: cap.
Funcionalitat: obre un cursor sobre la taula zones sense cap filtre.
- *zona*
Paràmetres d'entrada: codi VARCHAR2
Resultat: ZONES%ROWTYPE.
Precondicions: cap.
Funcionalitat: torna un registre de zones identificat pel seu codi.

Funcions de competicions: (CU-3) (funcionalitat equivalent a CU-2)

- *alta_competicio*
Paràmetres d'entrada: codi VARCHAR2, descripció VARCHAR2, zona VARCHAR2
- *baixa_competicio*
- *modifica_competicio*
- *llista_competicions*
- *competicio*

Funcions de temporades: (CU-4) (funcionalitat equivalent a CU-2)

- *alta_temporada*
Paràmetres d'entrada: competicio VARCHAR2, any NUMBER, inici DATE, final DATE
- *baixa_temporada*
- *modifica_temporada*
- *llista_temporades*
- *temporada*

Funcions de equips: (CU-5) (funcionalitat equivalent a CU-2)

- *alta Equip*
Paràmetres d'entrada: codi VARCHAR2, nom VARCHAR2, fundacio DATE, pais VARCHAR2
- *baixa Equip*
- *modifica Equip*
- *llista Equips*
- *Equip*

Funcions de jugadors: (CU-6)

Usaré aquest grup per explicar el funcionament, indicant als altres grups si el seu funcionament és com aquest o com algun altre.

Degut a que la clau primària aquesta taula és arbitrària usant una seqüència dins del trigger before insert cal afegir una funció extra per fer la recerca del identificador usant els atributs de la clau alternativa.

- *alta_jugador*
Paràmetres d'entrada: nom VARCHAR2, cognom1 VARCHAR2, cognom2 VARCHAR2, naixement DATE, nacionalitat VARCHAR2
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: el jugador no ha d'existir.
Funcionalitat: alta d'un nou jugador amb les seves dades.
- *alta_jugador (sobrecarregada)*
Paràmetres d'entrada: nom VARCHAR2, cognom1 VARCHAR2, cognom2 VARCHAR2, naixement DATE, nacionalitat VARCHAR2, **IN OUT id NUMBER**
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: el jugador no ha d'existir.
Funcionalitat: alta d'un nou jugador amb les seves dades retornant com a paràmetre de sortida el **id** arbitrari que se li ha assignat via trigger.
- *baixa_jugador*
Paràmetres d'entrada: id NUMBER
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: el jugador identificat per l'id ha d'existir
Funcionalitat: baixa d'un jugador identificat pel seu id.

- *modifica_jugador*
Paràmetres d'entrada: id NUMBER, nom VARCHAR2, cognom1 VARCHAR2, cognom2 VARCHAR2, naixement DATE, nacionalitat VARCHAR2
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: el jugador identificat pel id ha d'existir
Funcionalitat: modificació de les dades d'un jugador identificat pel seu id.
- *llista_jugadors*
Paràmetres d'entrada: cap
Resultat: CURSOR
Precondicions: cap
Funcionalitat: obre un cursor sobre la taula jugadors sense cap filtre.
- *jugador*
Paràmetres d'entrada: id NUMBER
Resultat: JUGADORS%ROWTYPE
Precondicions: cap
Funcionalitat: torna un registre de la taula jugadors identificat pel seu id.
- *cerca_jugador*
Paràmetres d'entrada: nom VARCHAR2, cognom1 VARCHAR2, cognom2 VARCHAR2
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: cap
Funcionalitat: torna el id del jugadors identificat per les seves dades.

6.1.2 pkg_usuaris

L'operador d'usuaris serà l'actor responsable d'aquest paquet i se li assignarà el rol d'execució d'aquest paquet.

Funcions: (CU-1) (funcionalitat equivalent a CU-2)

- *alta_usuari*
Paràmetres d'entrada: nom_usuari VARCHAR2, contrasenya VARCHAR2, nom VARCHAR2, cognom1 VARCHAR2, cognom2 VARCHAR2, sexe VARCHAR2, identificador VARCHAR2, dni VARCHAR2, naixement DATE, nacionalitat VARCHAR2, residència VARCHAR2, iban VARCHAR2, targeta VARCHAR2
- *baixa_usuari*
- *modifica_usuari*
- *llista_usuaris*
- *usuari*
- *usuari_lopd*
Ja que les dades dels usuaris estan repartides en dues taules degut a la LOPD, les funcions de retorn de registres també són dues.

6.1.3 pkg_partits

L'operador de partits serà l'actor responsable d'aquest paquet.

Funcions de partits: (CU-7) (funcionalitat equivalent a CU-6)

- *alta_partit*
Paràmetres d'entrada: competicio VARCHAR2, any NUMBER, equip1 VARCHAR2, equip2 VARCHAR2, data DATE, jugar VARCHAR2
- *alta_partit (sobrecarregada)*
Paràmetres d'entrada: competicio VARCHAR2, any NUMBER, equip1 VARCHAR2, equip2 VARCHAR2, data DATE, jugar VARCHAR2, **IN OUT id NUMBER**
- *baixa_partit*
- *modifica_partit*
- *llista_partits*
- *partit*
- *cerca_partit*

Funcions de tipus de gol: (ampliació del CU-8) (funcionalitat equivalent a CU-2)

- *alta_tipus_gol*
Paràmetres d'entrada: codi VARCHAR2, descripció VARCHAR2
- *baixa_tipus_gol*
- *modifica_tipus_gol*
- *llista_tipus_gol*
- *tip_gol*

Funcions de gols: (CU-8) (funcionalitat equivalent a CU-6)

- *alta_gol*
Paràmetres d'entrada: partit NUMBER, equip VARCHAR2, jugador NUMBER, part NUMBER, minut NUMBER, tipus VARCHAR2
- *alta_gol (sobrecarregada)*
Paràmetres d'entrada: partit NUMBER, equip VARCHAR2, jugador NUMBER, part NUMBER, minut NUMBER, tipus VARCHAR2, **IN OUT id NUMBER**
- *baixa_gol*
- *modifica_gol*
- *llista_gols*
- *gol*
- *cerca_gol*

6.1.4 pkg_joc

L'usuari serà l'actor responsable d'aquest.

Funcions:

- **login (CU-9)**
Paràmetres d'entrada: nom_usuari VARCHAR2, contrasenya VARCHAR2
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: cap
Funcionalitat: valida l'usuari.
- **apostar (CU-12)** (funcionalitat equivalent a *alta_aposta* CU-16)
Paràmetres d'entrada: modalitat VARCHAR2, partit NUMBER, aposta VARCHAR2, import FLOAT
Resultat: NUMBER (0 Ok, SqlCode si hi ha hagut algun error)
Precondicions: cap
Funcionalitat: realitza una aposta.
- **historic_resultats (CU-10)**
Paràmetres d'entrada: competicio VARCHAR2, any NUMBER, equip1 VARCHAR2, equip2 VARCHAR2
Resultat: CURSOR
Precondicions: cap
Funcionalitat: obre un cursor sobre la taula de resultats amb els filtres indicats als paràmetres.
- **historic_gols(CU-10)**
Paràmetres d'entrada: jugador NUMBER, competicio VARCHAR2, any NUMBER, equip VARCHAR2
Resultat: CURSOR
Precondicions: cap
Funcionalitat: obre un cursor sobre la taula de gols amb els filtres indicats als paràmetres.
- **historic_apostes (CU-11)**
Paràmetres d'entrada: modalitat VARCHAR2, competicio VARCHAR2, any NUMBER, equip VARCHAR2
Resultat: CURSOR
Precondicions: cap
Funcionalitat: obre un cursor sobre la taula d'apostes amb els filtres indicats als paràmetres i només de l'usuari actual.

Variables globals d'àmbit de sessió:

- **nom_usuari**
S'assigna quan l'usuari executa la funció de logon i s'utilitza per assignar l'alta d'una aposta a un usuari i també per guardar l'usuari dins del registre de log.

6.1.5 pkg_apostes

L'operador d'apostes serà l'actor responsable d'aquest paquet amb l'excepció de la resolució d'apostes que serà responsabilitat del sistema.

Funcions de modalitat: (CU-13) (funcionalitat equivalent a CU-6)

- *alta_modalitat*
Paràmetres d'entrada: descripcio VARCHAR2
- *alta_modalitat (sobrecarregada)*
Paràmetres d'entrada: descripcio VARCHAR2, **IN OUT id NUMBER**
- *baixa_modalitat*
- *modifica_modalitat*
- *llista_modalitats*
- *modalitat*
- *cerca_modalitat*

Funcions de coeficients: (CU-14) (funcionalitat equivalent a CU-2)

- *alta_coeficient*
Paràmetres d'entrada: modalitat VARCHAR2, partit NUMBER, resultat VARCHAR2, valor FLOAT
- *baixa_coeficient*
- *modifica_coeficient*
- *llista_coeficients*
- *coeficient*

Funcions d'apostes: (CU-16) (funcionalitat equivalent a CU-6)

- *alta_aposta*
Paràmetres d'entrada: modalitat VARCHAR2, partit NUMBER, aposta VARCHAR2, import FLOAT
- *alta_aposta (sobrecarregada)*
Paràmetres d'entrada: modalitat VARCHAR2, partit NUMBER, aposta VARCHAR2, import FLOAT, **IN OUT id NUMBER**
- *baixa_aposta*
- *modifica_aposta*
- *llista_apostes*
- *aposta*
- *cerca_aposta*
- *resoldre_apostes*
Resoldrà totes les apostes pendents: apostes amb l'atribut resol a fals de partits amb l'atribut jugat a 1.

Funcions de resultats: (CU-15) (funcionalitat equivalent a CU-2)

- *alta_resultat*
Paràmetres d'entrada: modalitat VARCHAR2, partit NUMBER, resultat VARCHAR2
- *baixa_resultat*
- *modifica_resultat*
- *llista_resultats*
- *resultat*

6.1.6 pkg_import

L'operador d'estadístiques serà l'actor responsable d'aquest paquet.

Funcions d'importació: (CU-17)

- *import_d_zones*
- *import_d_usuaris*
- *import_d_dates*
- *import_d_sexes*
- *import_d_edats*
- *import_d_temporades*
- *import_d_competicions*
- *import_d_modalitats*
- *import_d_equips*

- *import_f_apostes*

- *import_data_warehouse*
Cridarà ordenadament les importacions de les dimensions i els fets llegint les dades de l'esquema operacional i escrivint-les a l'esquema data warehouse.

6.1.7 pkg_util

L'administrador serà l'actor responsable d'aquest paquet.

Funcions de log: (CU-18)

- *alta_log*
PRAGMA AUTONOMOUS_TRANSACTION
Aquesta funció usa una directiva especial per executar-se de forma aïllada a la resta de fils d'aquella sessió d'usuari.

Degut a que aquesta funció s'utilitza dins de tota la resta de tasques cal garantir que s'escriurà la informació al log tant si s'ha executat un COMMIT com si s'ha executat un ROLLBACK dins la tasca.

Aïllant la seva execució en una transacció autònoma no es veurà afectada per la resta de fils d'aquella sessió.

- *llista_logs*

Funcions de test: (CU-19)

- *test_zones*
- *test_competicions*
- *test_temporades*
- *test_equips*
- *test_jugadors*
- *test_plantilles*
- *test_partits*
- *test_tipus_gol*
- *test_gols*
- *test_usuaris*
- *test_modalitats*
- *test_coeficients*
- *test_apostes*
- *test_resultats*

Aquestes funcions són proves unitàries de cadascuna de les taules del sistema.

He definit també un script de proves de test que executa ordenadament totes aquestes proves i en mostra els resultats.

- *dades_inicials*
Afegirà a la base de dades la informació mínima per poder realitzar les proves del sistema.

6.2 Registre de log

S'ha definit una taula de log per enregistrar tots els moviments de les taules tant siguin de lectura com d'escriptura.

Totes les funcions definides anteriorment guarden explícitament al log tant la petició que s'ha fet com el resultat de la mateixa ja sigui correcta o errònia i l'accés de lectura al log també es fa a través d'una funció dissenyada al respecte, complint així la Llei de Protecció de Dades personals a nivell Bàsic.

6.3 Dades inicials OLTP

Per tal de disposar d'un petit conjunt de dades vàlid per poder mostrar les proves funcionals he generat unes funcions per afegir aquest conjunt de dades inicials a totes les taules.

He definit un petit conjunt de zones: Catalunya, España, Europa, Portugal, Argentina, Sud-amèrica, ...

També he definit les competicions més importants, com són: La lliga BBVA, la lliga de campions, el mundial de futbol, les olimpíades, ...

Dins d'aquestes competicions he afegit, només a la temporada 2013-2014, tots els equips i tots els partits de la lliga BBVA de la primera divisió espanyola (dades en brut obtingudes de la pàgina web del diari Marca).

A l'apartat de jugadors només n'he afegit dos: en Leo Messi i en Cristiano Ronaldo, per poder afegir-los també els gols que van fer als respectius partits d'anada i tornada dels seus equips, i he definit quatre tipus diferents de gols: penal, falta, corner o xut.

A les modalitats d'apostes he definit cinc modalitats: 1x2 (la típica de les quínieles), resultat final, equip que marcarà primer, jugador que marcarà primer i minut del primer gol.

Finalment he afegit tots els resultats dels partits de la lliga BBVA de la temporada 2013-2014 que s'han jugat fins el dia 13 d'abril de 2014 però només dins de dues modalitats: 1x2 i resultat final (dades en brut obtingudes de la pàgina web del diari Marca).

6.4 Restriccions implementades

S'han implementat als paquets les restriccions definides al punt 4.4 a més de les restriccions d'integritat definides als scripts de creació de taules:

- Un Usuari ha de ser major d'edat.
- Un Jugador no pot pertànyer a la Plantilla de dos Equips al mateix moment.
- Un Partit ha d'estar compost de dos Equips diferents.
- Un Gol serà marcat per un Jugador d'un dels dos Equips que juguen el Partit.
- Una Aposta es farà sobre un Partit que encara no s'hagi disputat.

6.5 Proves unitàries OLTP

Per garantir que les funcions implementades als paquets són correctes he dissenyat uns processos de proves unitàries per cadascuna de les taules i són les funcions del paquet pkg_util.

L'algorisme usat a totes les funcions de proves segueix la mateixa tècnica que és la següent:

- Alta de dades necessàries degut a les restriccions d'integritat
 - Alta d'un element E1
 - Consulta l'existència d'aquell element E1
 - Esborrat d'aquell element E1
 - Consulta la no existència d'aquell element E1
- Alta de dos elements E2 i E3
 - Consulta l'existència dels elements E2 i E3
 - Esborrat de l'element E2

- Consulta la no existència de l'element E2
 - Consulta l'existència de l'element E3
 - Esborrat de l'element E3
 - Consulta la no existència de l'element E3
-
- Esborrat de les dades necessàries

Les proves realitzades estan desglossades a l'Annex B.

6.6 Disparadors OLTP

He definit els disparadors amb l'únic objectiu de donar valor al codi arbitrari de les taules: jugadors, partits, gols, modalitats i apostes.

Altres restriccions, encara no definides, podrien ser implementades dins d'aquests disparadors.

6.7 Paquets data warehouse

La forma habitual d'obtenir dades per generar un esquema data warehouse és a través del procés ETL:

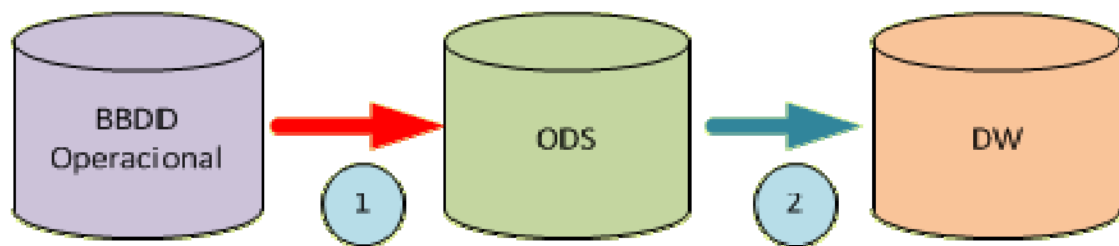


Figura 11: Procés estàndard ETL de generació d'un data warehouse

1. Exportem les dades de la base de dades operacions i les importem a un nou esquema idèntic anomenat ODS (Operational Data Store) deixant l'esquema operacional intacte i sense cap càrrega de treball addicional.
2. Executem les rutines de transformació de dades de l'esquema ODS per obtenir les dimensions i els fets a l'esquema DW.

A un entorn de producció "normal" aquesta seria la millor forma de fer-ho, sent l'esquema operacional un entorn aïllat de l'esquema ODS i del DW per qüestions de rendiment, capacitat i disponibilitat principalment.

Al nostre cas he decidit estalviar-nos l'esquema ODS degut a que aquest no és un entorn de producció real.

Per tant el procés ETL queda reduït només a la transformació i càrrega ja que els requeriments de rendiment, capacitat i disponibilitat no són aplicables en aquest cas.

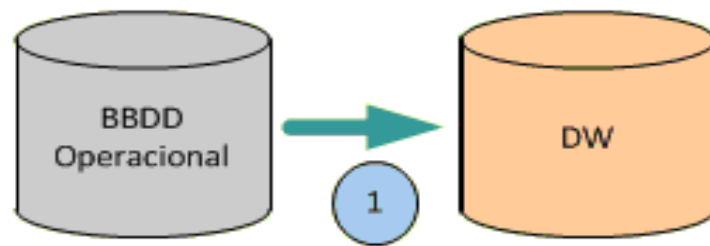


Figura 12: Procés actual de generació del data warehouse

6.8 ETL: Procés d'importació

El procés d'importació està clarament diferenciat en dues parts: la importació de les dimensions i la importació dels fets i cal executar-ho en aquest ordre.

6.8.1 Neteja i normalització de les dades operacionals

A la fase de Transformació del procés ETL cal realitzar tasques de neteja i normalització de les dades operacionals per deixar-les de manera correcta per poder-les importar a la base data warehouse sense tenir problemes.

En aquest cas no cal realitzar aquestes tasques ja que les dades a importar les he dissenyat i afegit correctament per tal de no haver de fer cap feina extra.

En un cas real sí seria necessari fer un estudi previ a la importació per veure com s'haurien de transformar certes dades que per la intervenció humana no serien del tot correctes.

6.8.2 Dimensions

Tal i com vaig indicar al punt 5.2.6 i 5.2.7, les dimensions que cal importar són: D_Zona, D_Data, D_Usuari, D_Sexe, D_Edat, D_Equip, D_Competicio, D_Temporada i D_Modalitat.

Hi haurà dimensions que les obtindrè de les taules anàlogues de la base operacional: D_Zona, D_Usuari, D_Equip, D_Competicio, D_Temporada i D_Modalitat.

La taula d'usuaris també proveirà les dimensions D_Sexe i D_Edat.

Finalment, la D_Data la obtindrè de la taula dels partits.

6.8.3 Fets i indicadors

Els fets són les apostes i les obtindrà de la taula d'apostes de la base operacional i a través de les relacions amb les altres taules obtindrà les dades per poder trobar els identificadors de cada dimensió.

Per poder emmagatzemar correctament els indicadors caldrà usar les taules operacionals d'apostes, de resultats i de coeficients per obtenir els diners apostats, si l'aposta ha estat guanyadora i quants diners s'han pagat (només en cas que hagi estat guanyadora).

Degut a tot això el procés d'importació de F_Aposta serà lent ja que ha de fer recerques en moltes taules abans d'inserir un registre.

6.8.4 Execució ETL

Les dues primeres fases del procés ETL: Extracció i Transformació són buides ja que no cal fer res.

En un entorn real de producció sí existirien aquest dos processos i també seria diferent la Càrrega.

Per aquest sistema la càrrega va precedida de la neteja total de la base de dades data warehouse.

Si l'entorn fos real no es podria fer aquesta neteja total ja que trigaria massa temps per importar totes les dades, per tant seria convenient fer la importació parcial només de les dades que no estan al data warehouse.

Les proves realitzades estan desglossades a l'Annex C.

6.9 Magatzem multidimensional

Posteriorment a la correcta importació de les dades a les dimensions i als fets ja es poden començar a fer proves, però abans cal preparar un cub amb les dimensions i els fets:

6.9.1 Creació del cub

A la creació de cub multidimensional primer indico la taula dels fets i els indicadors que cal mostrar i amb quina funció d'agrupament, en aquest cas la suma:

Taula	Atribut
Fets	
F_Aposta	APOSTA: diners apostats, GUANYADORA: aposta guanyadora, PAGAT: diners pagats a aquesta aposta

I a continuació vaig indicant les dimensions, les taules i els atributs assignats als nivells:

Taula	Atribut	Nivell
Dimensions		
D_Zona	codi, nom	* Codi, nom
D_Data	any, mes, dia, setmana de l'any, dia de la setmana, hora, minut, segon	* any, setmana de l'any, mes, dia, hora, minut, segon * dia de la setmana
D_Sexe	sexe	* sexe
D_Edat	edat	* edat
D-Equip	codi, nom	* codi, nom
D_Competició	codi, nom	* codi, nom
D_Temporada	codi, any, nom	* codi, nom * any
D_Modalitat	codi, nom	* codi, nom
D_Usuari	codi usuari	* codi usuari

Per fer aquest cub he usat la eina Pentaho Schemma Workbench que ja vaig usar a l'assignatura de Models Multidimensionals i Magatzems de Dades.

Les pantalles de les execucions i els detalls de la creació del cub estan desglossades a l'Annex C.2.

6.9.2 Execució a Pentaho

Ara ja podem executar informes usant l'eina d'anàlisi Pentaho indicant-li el cub a tractar i la base de dades data warehouse.

Així doncs un cop he indicat el cub origen de la informació només cal anar mostrant les columnes i les files, posant filtres i mostrant les dades agrupades per obtenir les diferents respostes a les preguntes que es van indicar al punt 5.2.1:

- Distribució geogràfica dels usuaris

	Medidas	
jerarquiaPaisUsuari	● Apostes	● Pagat
<input type="checkbox"/> All Pais_Usuari.jerarquiaPaisUsuaris	92	51,5
<input type="checkbox"/> CAT	92	51,5
Catalunya	92	51,5

Figura 13: Resultat de la consulta de distribució geogràfica

- Distribució per edats dels usuaris i Despeses per intervals d'edats

jerarquiaEdat	Medidas	
	● Apostes	● Pagat
<input type="checkbox"/> All Edat.jerarquiaEdats	92	51,5
20	4,5	5,25
24	50	0
44	10	20
45	27,5	26,25

Figura 14: Resultat de la consulta de distribució per edats

- Equips als que s'aposta més

jerarquiaEquip1	Medidas	
	● Apostes	● Pagat
<input type="checkbox"/> All Equip1.jerarquiaEquip1s	92	51,5
<input type="checkbox"/> FCB	92	51,5
Futbol Club Barcelona	92	51,5

Figura 15: Resultat de la consulta d'equips als que s'aposta més

- Guany i despeses per sexe

jerarquiaSexe	Medidas	
	● Apostes	● Pagat
<input type="checkbox"/> All Sexe.jerarquiaSexes	92	51,5
D	14,5	25,25
H	77,5	26,25

Figura 16: Resultat de la consulta de guany i despeses per sexes

- Competicions a les que s'aposta més i Temporades a les que s'ha apostat més

		Medidas	
jerarquiaCompeticio	jerarquiaTemporada	● Apostes	● Pagat
<input type="checkbox"/> All Competicio.jerarquiaCompeticios	<input type="checkbox"/> All Temporada.jerarquiaTemporadas	92	51,5
<input type="checkbox"/> BBVA	<input type="checkbox"/> All Temporada.jerarquiaTemporadas	92	51,5
Lliga Española de primera divisió	<input type="checkbox"/> All Temporada.jerarquiaTemporadas	92	51,5
	2013	92	51,5

Figura 17: Resultat de competicions i temporades amb més apostes

- Etc.

(Les pantalles de totes les execucions amb Pentaho estan desglossades a l'Annex C.4.)

També es pot mostrar la informació usant combinacions de dimensions i gràfics:

		Medidas	
jerarquiaSexe	jerarquiaEdat	● Apostes	● Pagat
<input type="checkbox"/> All Sexe.jerarquiaSexes	<input type="checkbox"/> All Edat.jerarquiaEdats	92	51,5
D	<input type="checkbox"/> All Edat.jerarquiaEdats	14,5	25,25
	20	4,5	5,25
	44	10	20
H	<input type="checkbox"/> All Edat.jerarquiaEdats	77,5	26,25
	24	50	0
	45	27,5	26,25

Slicer:

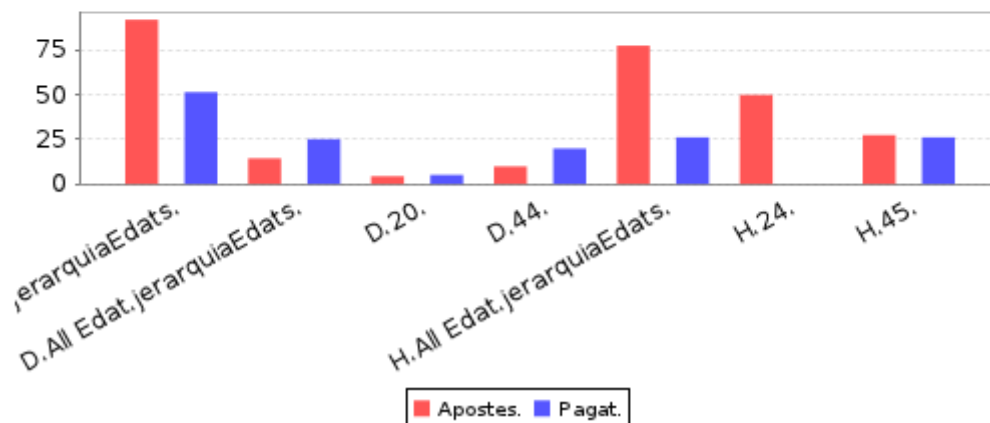


Figura 18: Resultat de la consulta de sexes i edats amb un gràfic

De tota manera cal tenir en compte que aquesta eina també permet fer exportacions a MSExcel i podent usar tota la potència d'aquesta eina:



Figura 19: Exportació a MSExcel

Capítol 7: Valoració econòmica

He assignat totes les tasques a un dels tres perfils de treball del projecte:

Tasca	Perfil	Hores
Especificació de requeriments	Director	15
Investigació de les modalitats d'apostes	Director	5
Altres requeriments	Director	2,5
Definició de casos d'ús	Director	2,5
Llistat de tasques i fites	Director	5
Disseny de la planificació	Analista	10
Creació del document PAC 1	Analista	20
Disseny conceptual	Analista	15
Disseny físic	Analista	15
Creació de l'esquema	Programador	2
Creació de les taules	Programador	4
Creació de les claus foranes	Programador	4
Creació dels índexs	Programador	5
Proves unitàries	Programador	5
Disseny de fets i dimensions	Analista	10
Disseny de ETL	Analista	10
Proves unitàries	Programador	5
Creació del document PAC 2	Analista	20
Creació de la memòria	Analista	100
Creació de la presentació	Analista	35
Creació del treball pràctic	Programador	15
Total hores		305

És possible que hi hagi certes tasques on es pugui discutir si han de ser realitzades per un perfil o per un altre, però jo les he considerat de la manera indicada a la taula anterior.

He agrupant les hores del perfils amb el seu preu d'hora podem trobar el cost final dels recursos humans:

Perfil	Hores	Preu hora	Import
Director de projecte	30	31,2500 €	937,50 €
Analista	235	22,7272 €	5.340,89 €
Programador	40	18,9375 €	757,50 €
Total Projecte	305 h		7.035,89 €

Les dades del preu hora les he calculat de les dades extretes de la pàgina web salaryexplorer.com usant la mensualitat de 22 dies laborables al mes i 8 hores al dia:

Perfil	Mensualitat	Preu hora
Director de projecte	5.500 €	31,2500 €
Analista	4.000 €	22,7272 €
Programador	3.333 €	18,9375 €

I fent una mitjana del cost hora de la totalitat de projecte queda d'aquesta manera:

Total Projecte	305 h	7.035,89 €	23,0685 €/h
-----------------------	-------	------------	--------------------

Capítol 8: Conclusions

Un cop realitzades totes les tasques indicades al Pla de treball només queda fer el tancament del projecte que en aquest cas és la seva presentació i esperar les preguntes del Tribunal.

He pogut comprovar que disposar d'un Pla de treball rigorós i extensiu és molt bo a l'hora d'executar-lo ja que ha estat una càrrega de treball intensa i m'ha anat molt bé només fer el PFC aquest darrer semestre per poder-li així dedicar el temps necessari.

Un cop més he pogut experimentar les diferències creatives, d'abstracció i metòdiques que tenen les tasques associades a cadascun dels punts principals del projecte.

Crec que he complert els objectius fixats al punt 1.2 ajudat per tots els coneixements adquirits durant la carrera i en especial en les assignatures de Gestió de Projectes, Enginyeria del Programari, Bases de Dades i Models Multidimensionals.

Ara entenc, un cop acabat el projecte, que no es demani la realització d'una interfície d'usuari ja que no es disposa del temps per poder-la realitzar amb garanties de qualitat.

Glossari

Apache: servidor web multiplataforma basat en codi obert.

Base de dades: lloc on s'emmagatzema la informació.

Booleà: tipus de dades amb resultat cert o fals.

Cluster: grup de dos o més serveis o servidors que proporcionen tots plegats el mateix servei.

CRUD: sigles de Create, Read, Update i Delete.

Cub: estructura d'informació multidimensional.

Data warehouse: magatzem d'informació.

Diagrama conceptual: disseny gràfic abstracte dels serveis o les dades.

Diagrama de gantt: planificació gràfica de la temporització de les tasques i les seves relacions.

Diagrama físic: implementació real dels serveis o les dades.

Diagrama lògic: especialització del diagrama conceptual usant les restriccions de l'arquitectura.

ETL: sigles de Extract, Transform, Load.

Java: llenguatge de programació orientat a objectes.

Log: registre d'esdeveniments d'un sistema.

Login: validació d'un usuari a un sistema.

LOPD: Llei de Protecció de Dades de caràcter Personal.

OLTP: sigles de OnLine Transaction Process.

Oracle: fabricant de programari famós per la seva base de dades.

PFC: sigles de Projecte de Final de Carrera.

PHP: llenguatge de programació interpretat.

PL/SQL: llenguatge de programació de codi de la base de dades Oracle.

Script: fitxer de text que conté comandes a executar.

SGBDR: sigles de Sistema de Gestió de Bases de Dades Relacionals.

Bibliografia

Reial decret 1720/2007, de 21 de desembre, pel qual s'aprova el Reglament de desenvolupament de la Llei Orgànica 15/1999, de 13 de desembre, de protecció de dades de caràcter personal.

Publicat a: "BOE" núm. 17, de 19/01/2008.

Entrada en vigor: 19/04/2008

Departament: Ministeri de Justícia

<https://www.boe.es/buscar/act.php?id=BOE-A-2008-979>

Dades dels equips de la Lliga BBVA de la temporada 2013-2014

http://ca.wikipedia.org/wiki/Primera_divisi%C3%B3_espanyola_de_futbol

Dades del calendari de la Lliga BBVA de la temporada 2013-2014

<http://www.marca.com/futbol/primera/calendario.html>

Director de projecte: Project Leader

<http://www.salaryexplorer.com/salary-search.php?loc=203&loctype=1&jobtype=2&job=1>

Analista: Software Architect

<http://www.salaryexplorer.com/salary-search.php?loc=203&loctype=1&jobtype=2&job=9>

Programador: Developer / Programmer (general)

<http://www.salaryexplorer.com/salary-search.php?loc=203&loctype=1&jobtype=2&job=3>

Annex A: Scripts SQL

A continuació indico els scripts sql que he creat en aquesta part i l'anterior:
(El número indica l'ordre correcte d'execució dels scripts)

A.1 Creació d'esquemes

00 creacio esquemes.sql

A.2 Scripts OLTP

10 neteja taules OLTP.sql
11 creacio taules OLTP.sql
12 creacio sequence OLTP.sql
13 creacio triggers OLTP.sql
14 creacio indexs OLTP.sql

A.3 Scripts Data Warehouse

20 neteja taules DW.sql
21 creacio taules DW.sql

A.4 Scripts paquets OLTP

15a Package usuaris.sql
15b Package competicions.sql
15c Package partits.sql
15d Package joc.sql
15e Package apostes.sql
15f Package util.sql
16 creacio rols i grants.sql

A.5 Scripts paquets Data Warehouse

22 creacio sequence DW.sql
25 Package import.sql

A.6 Scripts de proves

15z Execucio TEST.sql
25z Execucio IMPORT.sql

Annex B: Proves realitzades a la base de dades operacional

A continuació indico els resultats de les proves realitzades sobre la base de dades operacional.

B.1 Resultat execució del script 15z Execucio TEST.sql

```
TEST_ZONES: ALTA_ZONA Z1 -> CERT
TEST_ZONES: ZONA Z1 -> CERT
TEST_ZONES: BAIXA_ZONA Z1 -> CERT
TEST_ZONES: ZONA Z1 NO DISPONIBLE -> CERT
TEST_ZONES: ALTA_ZONA Z2 -> CERT
TEST_ZONES: ALTA_ZONA Z3 -> CERT
TEST_ZONES: ZONA Z2 -> CERT
TEST_ZONES: ZONA Z3 -> CERT
TEST_ZONES: BAIXA_ZONA Z2 -> CERT
TEST_ZONES: ZONA Z2 NO DISPONIBLE -> CERT
TEST_ZONES: ZONA Z3 -> CERT
TEST_ZONES: BAIXA_ZONA Z3 -> CERT
TEST_ZONES: ZONA Z3 NO DISPONIBLE -> CERT
FINAL TEST_ZONES OK
```

```
TEST_COMPETICIONS: ALTA_ZONA Z1 -> CERT
TEST_COMPETICIONS: ALTA_COMPETICIO C1 -> CERT
TEST_COMPETICIONS: COMPETICIO C1 -> CERT
TEST_COMPETICIONS: BAIXA_COMPETICIO C1 -> CERT
TEST_COMPETICIONS: COMPETICIO C1 NO DISPONIBLE -> CERT
TEST_COMPETICIONS: ALTA_COMPETICIO C2 -> CERT
TEST_COMPETICIONS: ALTA_COMPETICIO C3 -> CERT
TEST_COMPETICIONS: COMPETICIO C2 -> CERT
TEST_COMPETICIONS: COMPETICIO C3 -> CERT
TEST_COMPETICIONS: BAIXA_COMPETICIO C2 -> CERT
TEST_COMPETICIONS: COMPETICIO C2 NO DISPONIBLE -> CERT
TEST_COMPETICIONS: COMPETICIO C3 -> CERT
TEST_COMPETICIONS: BAIXA_COMPETICIO C3 -> CERT
TEST_COMPETICIONS: COMPETICIO C3 NO DISPONIBLE -> CERT
TEST_COMPETICIONS: BAIXA_ZONA Z1 -> CERT
FINAL TEST_COMPETICIONS OK
```

```
TEST_TEMPORADES: ALTA_ZONA Z1 -> CERT
TEST_TEMPORADES: ALTA_COMPETICIO C1 -> CERT
TEST_TEMPORADES: ALTA_TEMPORADA C1-2001 -> CERT
TEST_TEMPORADES: TEMPORADA C1-2001 -> CERT
TEST_TEMPORADES: TEMPORADA C1-2001 -> CERT
TEST_TEMPORADES: TEMPORADA C1-2001 NO DISPONIBLE -> CERT
TEST_TEMPORADES: ALTA_TEMPORADA C1-2002 -> CERT
TEST_TEMPORADES: ALTA_TEMPORADA C1-2003 -> CERT
TEST_TEMPORADES: TEMPORADA C1-2002 -> CERT
TEST_TEMPORADES: TEMPORADA C1-2003 -> CERT
TEST_TEMPORADES: BAIXA_TEMPORADA C1-2002 -> CERT
TEST_TEMPORADES: TEMPORADA C1-2002 NO DISPONIBLE -> CERT
TEST_TEMPORADES: TEMPORADA C1-2003 -> CERT
TEST_TEMPORADES: BAIXA_TEMPORADA C1-2003 -> CERT
TEST_TEMPORADES: TEMPORADA C1-2003 NO DISPONIBLE -> CERT
TEST_TEMPORADES: BAIXA_COMPETICIO C1 -> CERT
TEST_TEMPORADES: BAIXA_ZONA Z1 -> CERT
FINAL TEST_TEMPORADES OK
```

```
TEST_EQUIPS: ALTA_ZONA Z1 -> CERT
TEST_EQUIPS: ALTA_EQUIP EQ1 -> CERT
TEST_EQUIPS: EQUIP E1 -> CERT
TEST_EQUIPS: BAIXA_EQUIP EQ1 -> CERT
TEST_EQUIPS: EQUIP EQ1 NO DISPONIBLE -> CERT
```

TEST_EQUIPS: ALTA_EQUIP EQ2 -> CERT
TEST_EQUIPS: ALTA_EQUIP EQ3 -> CERT
TEST_EQUIPS: EQUIP EQ2 -> CERT
TEST_EQUIPS: EQUIP EQ3 -> CERT
TEST_EQUIPS: BAIXA_EQUIP EQ2 -> CERT
TEST_EQUIPS: EQUIP C2 NO DISPONIBLE -> CERT
TEST_EQUIPS: EQUIP EQ3 -> CERT
TEST_EQUIPS: BAIXA_EQUIP EQ3 -> CERT
TEST_EQUIPS: EQUIP EQ3 NO DISPONIBLE -> CERT
TEST_EQUIPS: BAIXA_ZONA Z1 -> CERT
FINAL TEST_EQUIPS OK

TEST_JUGADORS: ALTA_ZONA Z1 -> CERT
TEST_JUGADORS: ALTA_JUGADOR NOM1 -> CERT
TEST_JUGADORS: JUGADOR NOM1 -> CERT
TEST_JUGADORS: BAIXA_JUGADOR NOM1 -> CERT
TEST_JUGADORS: JUGADOR NOM1 NO DISPONIBLE -> CERT
TEST_JUGADORS: ALTA_JUGADOR NOM2 -> CERT
TEST_JUGADORS: ALTA_JUGADOR NOM3 -> CERT
TEST_JUGADORS: JUGADOR NOM2 -> CERT
TEST_JUGADORS: JUGADOR NOM3 -> CERT
TEST_JUGADORS: BAIXA_JUGADOR NOM2 -> CERT
TEST_JUGADORS: JUGADOR NOM2 NO DISPONIBLE -> CERT
TEST_JUGADORS: JUGADOR NOM3 -> CERT
TEST_JUGADORS: BAIXA_JUGADOR NOM3 -> CERT
TEST_JUGADORS: JUGADOR NOM3 NO DISPONIBLE -> CERT
TEST_JUGADORS: BAIXA_ZONA Z1 -> CERT
FINAL TEST_JUGADORS OK

FINAL TEST_PLANTILLES OK

TEST_PARTITS: ALTA_ZONA Z1 -> CERT
TEST_PARTITS: ALTA_COMPETICIO C1 -> CERT
TEST_PARTITS: ALTA_TEMPORADA C1-2001 -> CERT
TEST_PARTITS: ALTA_EQUIP EQ1 -> CERT
TEST_PARTITS: ALTA_EQUIP EQ2 -> CERT
TEST_PARTITS: ALTA_EQUIP EQ3 -> CERT
TEST_PARTITS: ALTA_PARTIT EQ1-EQ2 -> CERT
TEST_PARTITS: PARTIT EQ1-EQ2 -> CERT
TEST_PARTITS: BAIXA_PARTIT EQ1-EQ2 -> CERT
TEST_PARTITS: PARTIT EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_PARTITS: ALTA_PARTIT RMA-FCB -> CERT
TEST_PARTITS: ALTA_PARTIT ATM-FCB -> CERT
TEST_PARTITS: PARTIT EQ1-EQ1 -> CERT
TEST_PARTITS: PARTIT EQ3-EQ1 -> CERT
TEST_PARTITS: BAIXA_PARTIT EQ2-EQ1 -> CERT
TEST_PARTITS: PARTIT EQ2-EQ1 NO DISPONIBLE -> CERT
TEST_PARTITS: PARTIT EQ3-EQ1 -> CERT
TEST_PARTITS: BAIXA_PARTIT EQ3-EQ1 -> CERT
TEST_PARTITS: PARTIT EQ3-EQ1 NO DISPONIBLE -> CERT
FINAL TEST_PARTITS OK

TEST_TIPUS_GOL: ALTA_TIPUS_GOL 1 -> CERT
TEST_TIPUS_GOL: TIPUS_GOL 1 -> CERT
TEST_TIPUS_GOL: BAIXA_TIPUS_GOL 1 -> CERT
TEST_TIPUS_GOL: TIPUS_GOL 1 NO DISPONIBLE -> CERT
TEST_TIPUS_GOL: ALTA_TIPUS_GOL 2 -> CERT
TEST_TIPUS_GOL: ALTA_TIPUS_GOL 3 -> CERT
TEST_TIPUS_GOL: TIPUS_GOL 2 -> CERT
TEST_TIPUS_GOL: TIPUS_GOL 3 -> CERT
TEST_TIPUS_GOL: BAIXA_TIPUS_GOL 2 -> CERT
TEST_TIPUS_GOL: TIPUS_GOL 2 NO DISPONIBLE -> CERT
TEST_TIPUS_GOL: TIPUS_GOL 3 -> CERT
TEST_TIPUS_GOL: BAIXA_TIPUS_GOL 3 -> CERT
TEST_TIPUS_GOL: TIPUS_GOL 3 NO DISPONIBLE -> CERT
FINAL TEST_TIPUS_GOL OK

TEST_GOLS: ALTA_ZONA Z1 -> CERT
TEST_GOLS: ALTA_COMPETICIO C1 -> CERT
TEST_GOLS: ALTA_TEMPORADA C1-2001 -> CERT
TEST_GOLS: ALTA_EQUIP EQ1 -> CERT
TEST_GOLS: ALTA_EQUIP EQ2 -> CERT

TEST_GOLS: ALTA_JUGADOR Nom1 -> CERT
TEST_GOLS: ALTA_JUGADOR Nom2 -> CERT
TEST_GOLS: ALTA_PARTIT EQ1-EQ2 -> CERT
TEST_GOLS: ALTA_TIPUS_GOL 1 -> CERT
TEST_GOLS: ALTA_TIPUS_GOL 2 -> CERT
TEST_GOLS: ALTA_TIPUS_GOL 3 -> CERT
TEST_GOLS: ALTA_GOL Nom1 Partit EQ1-EQ2 -> CERT
TEST_GOLS: GOL Nom1 Partit EQ1-EQ2 -> CERT
TEST_GOLS: BAIXA_GOL Nom1 Partit EQ1-EQ2 -> CERT
TEST_GOLS: GOL Nom1 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_GOLS: ALTA_GOL Nom2 Partit EQ1-EQ2 -> CERT
TEST_GOLS: ALTA_GOL Nom1 Partit EQ1-EQ2 -> CERT
TEST_GOLS: GOL Nom2 Partit EQ1-EQ2 -> CERT
TEST_GOLS: GOL Nom1 Partit EQ1-EQ2 -> CERT
TEST_GOLS: BAIXA_GOL Nom2 Partit EQ1-EQ2 -> CERT
TEST_GOLS: GOL Nom2 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_GOLS: GOL Nom1 Partit EQ1-EQ2 -> CERT
TEST_GOLS: BAIXA_GOL Nom1 Partit EQ1-EQ2 -> CERT
TEST_GOLS: GOL Nom1 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
FINAL TEST_GOLS OK

TEST_USUARIS: ALTA_ZONA Z1 -> CERT
TEST_USUARIS: ALTA_USUARI USERNAME@UOC.COM -> CERT
TEST_USUARIS: USUARI USERNAME@UOC.COM -> CERT
TEST_USUARIS: BAIXA_USUARI USERNAME@UOC.COM -> CERT
TEST_USUARIS: USUARI USERNAME@UOC.COM NO DISPONIBLE -> CERT
TEST_USUARIS: ALTA_USUARI U2@UOC.COM -> CERT
TEST_USUARIS: ALTA_USUARI U3@UOC.COM -> CERT
TEST_USUARIS: USUARI U2@UOC.COM -> CERT
TEST_USUARIS: USUARI U3@UOC.COM -> CERT
TEST_USUARIS: BAIXA_USUARI U2@UOC.COM -> CERT
TEST_USUARIS: USUARI U2@UOC.COM NO DISPONIBLE -> CERT
TEST_USUARIS: USUARI U3@UOC.COM -> CERT
TEST_USUARIS: BAIXA_USUARI U3@UOC.COM -> CERT
TEST_USUARIS: USUARI U3@UOC.COM NO DISPONIBLE -> CERT
TEST_USUARIS: BAIXA_ZONA Z1 -> CERT
FINAL TEST_USUARIS OK

TEST_MODALITATS: ALTA_MODALITAT PRIMERA_MODALITAT -> CERT
TEST_MODALITATS: MODALITAT PRIMERA_MODALITAT -> CERT
TEST_MODALITATS: BAIXA_MODALITAT PRIMERA_MODALITAT -> CERT
TEST_MODALITATS: ZONA PRIMERA_MODALITAT NO DISPONIBLE -> CERT
TEST_MODALITATS: ALTA_MODALITAT SEGONA_MODALITAT -> CERT
TEST_MODALITATS: ALTA_MODALITAT TERCERA_MODALITAT -> CERT
TEST_MODALITATS: MODALITAT SEGONA_MODALITAT -> CERT
TEST_MODALITATS: MODALITAT TERCERA_MODALITAT -> CERT
TEST_MODALITATS: BAIXA_MODALITAT SEGONA_MODALITAT -> CERT
TEST_MODALITATS: MODALITAT SEGONA_MODALITAT NO DISPONIBLE -> CERT
TEST_MODALITATS: MODALITAT TERCERA_MODALITAT -> CERT
TEST_MODALITATS: BAIXA_MODALITAT TERCERA_MODALITAT -> CERT
TEST_MODALITATS: MODALITAT TERCERA_MODALITAT NO DISPONIBLE -> CERT
FINAL TEST_MODALITATS OK

TEST_COEFICIENTS: ALTA_ZONA Z1 -> CERT
TEST_COEFICIENTS: ALTA_COMPETICIO C1 -> CERT
TEST_COEFICIENTS: ALTA_TEMPORADA C1-2001 -> CERT
TEST_COEFICIENTS: ALTA_EQUIP EQ1 -> CERT
TEST_COEFICIENTS: ALTA_EQUIP EQ2 -> CERT
TEST_COEFICIENTS: ALTA_PARTIT EQ1-EQ2 -> CERT
TEST_COEFICIENTS: ALTA_MODALITAT_MODALITAT1 -> CERT
TEST_COEFICIENTS: ALTA_COEFICIENT Modalitat1 R1 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: COEFICIENT Modalitat1 R1 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: BAIXA_COEFICIENT Modalitat1 R1 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: COEFICIENT Modalitat1 R1 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_COEFICIENTS: ALTA_COEFICIENT Modalitat1 R2 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: ALTA_COEFICIENT Modalitat1 R3 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: COEFICIENT Modalitat1 R2 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: COEFICIENT Modalitat1 R3 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: BAIXA_COEFICIENT Modalitat1 R2 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: COEFICIENT Modalitat1 R2 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_COEFICIENTS: COEFICIENT Modalitat1 R3 Partit EQ1-EQ2 -> CERT
TEST_COEFICIENTS: BAIXA_COEFICIENT Modalitat1 R3 Partit EQ1-EQ2 -> CERT

TEST_COEFICIENTS: COEFICIENT Modalitat1 R3 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
FINAL TEST_COEFICIENTS OK

TEST_APOSTES: ALTA_ZONA Z1 -> CERT
TEST_APOSTES: ALTA_COMPETICIO C1 -> CERT
TEST_APOSTES: ALTA_TEMPORADA C1-2001 -> CERT
TEST_APOSTES: ALTA_EQUIP EQ1 -> CERT
TEST_APOSTES: ALTA_EQUIP EQ2 -> CERT
TEST_APOSTES: ALTA_PARTIT EQ1-EQ2 -> CERT
TEST_APOSTES: ALTA_MODALITAT MODALITAT1 -> CERT
TEST_APOSTES: ALTA_MODALITAT MODALITAT2 -> CERT
TEST_APOSTES: ALTA_MODALITAT MODALITAT3 -> CERT
TEST_APOSTES: ALTA_USUARI USERNAME@UOC.COM -> CERT
TEST_APOSTES: ALTA_USUARI LOGIN USERNAME@UOC.COM -> CERT
TEST_APOSTES: ALTA_APOSTA Modalitat1 A1 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: APOSTA Modalitat1 A1 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: BAIXA_APOSTA Modalitat1 A1 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: APOSTA Modalitat1 A1 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_APOSTES: ALTA_APOSTA Modalitat1 A2 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: ALTA_APOSTA Modalitat1 A3 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: APOSTA Modalitat1 A2 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: APOSTA Modalitat1 A3 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: BAIXA_APOSTA Modalitat1 A2 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: APOSTA Modalitat1 A2 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_APOSTES: APOSTA Modalitat1 A3 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: BAIXA_APOSTA Modalitat1 A3 Partit EQ1-EQ2 -> CERT
TEST_APOSTES: APOSTA Modalitat1 A3 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
FINAL TEST_APOSTES OK

TEST_RESULTATS: ALTA_ZONA Z1 -> CERT
TEST_RESULTATS: ALTA_COMPETICIO C1 -> CERT
TEST_RESULTATS: ALTA_TEMPORADA C1-2001 -> CERT
TEST_RESULTATS: ALTA_EQUIP EQ1 -> CERT
TEST_RESULTATS: ALTA_EQUIP EQ2 -> CERT
TEST_RESULTATS: ALTA_PARTIT EQ1-EQ2 -> CERT
TEST_RESULTATS: ALTA_MODALITAT MODALITAT1 -> CERT
TEST_RESULTATS: ALTA_MODALITAT MODALITAT2 -> CERT
TEST_RESULTATS: ALTA_MODALITAT MODALITAT3 -> CERT
TEST_RESULTATS: ALTA_RESULTAT Modalitat1 R1 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: RESULTAT Modalitat1 R1 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: BAIXA_RESULTAT Modalitat1 R1 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: RESULTAT Modalitat1 R1 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_RESULTATS: ALTA_RESULTAT Modalitat2 R2 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: ALTA_RESULTAT Modalitat3 R3 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: RESULTAT Modalitat2 R2 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: RESULTAT Modalitat3 R3 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: BAIXA_RESULTAT Modalitat2 R2 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: RESULTAT Modalitat2 R2 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
TEST_RESULTATS: RESULTAT Modalitat3 R3 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: BAIXA_RESULTAT Modalitat3 R3 Partit EQ1-EQ2 -> CERT
TEST_RESULTATS: RESULTAT Modalitat3 R3 Partit EQ1-EQ2 NO DISPONIBLE -> CERT
FINAL TEST_RESULTATS OK

B.2 Resultat execució de la funció pkg_util.dades:inicials

INFO: DADES_INICIALS OK

Annex C: Proves realitzades a la base de dades data warehouse

A continuació indico els resultats de les proves realitzades sobre la base de dades data warehouse.

C.1 Resultat execució del script 25z Execucio IMPORT.sql

```
IMPORT_D_ZONES: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_ZONES -> CERT
IMPORT_D_USUARIS: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_USUARIS -> CERT
IMPORT_D_DATES: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_DATES -> CERT
IMPORT_D_SEXES: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_SEXES -> CERT
IMPORT_D_EDATS: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_EDATS -> CERT
IMPORT_D_TEMPORADES: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_TEMPORADES -> CERT
IMPORT_D_COMPETICIONS: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_COMPETICIONS -> CERT
IMPORT_D_MODALITATS: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_MODALITATS -> CERT
IMPORT_D_EQUIPS: IMPORTACIO OK -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_D_EQUIPS -> CERT
IMPORT_DATAWAREHOUSE: IMPORT_F_APOSTES -> CERT
IMPORT DATA WAREHOUSE OK
```

C.2 Creació del CUB

A les figures 4 i 5 es poden observar dins el cub les dimensions i les mesures que he creat usant Pentaho Schema Workbench:

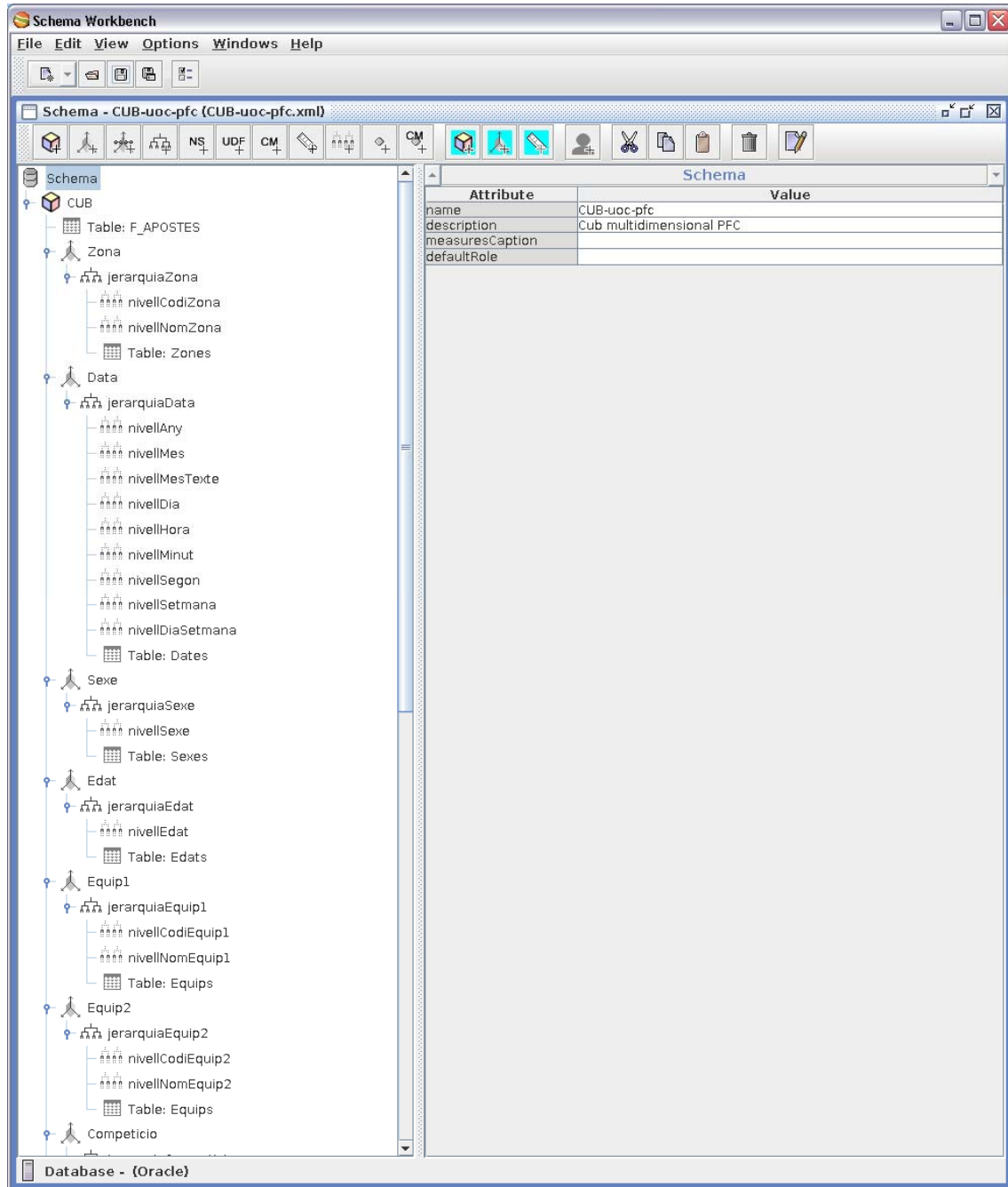


Figura 20: CUB-uoc-pfc (part 1)

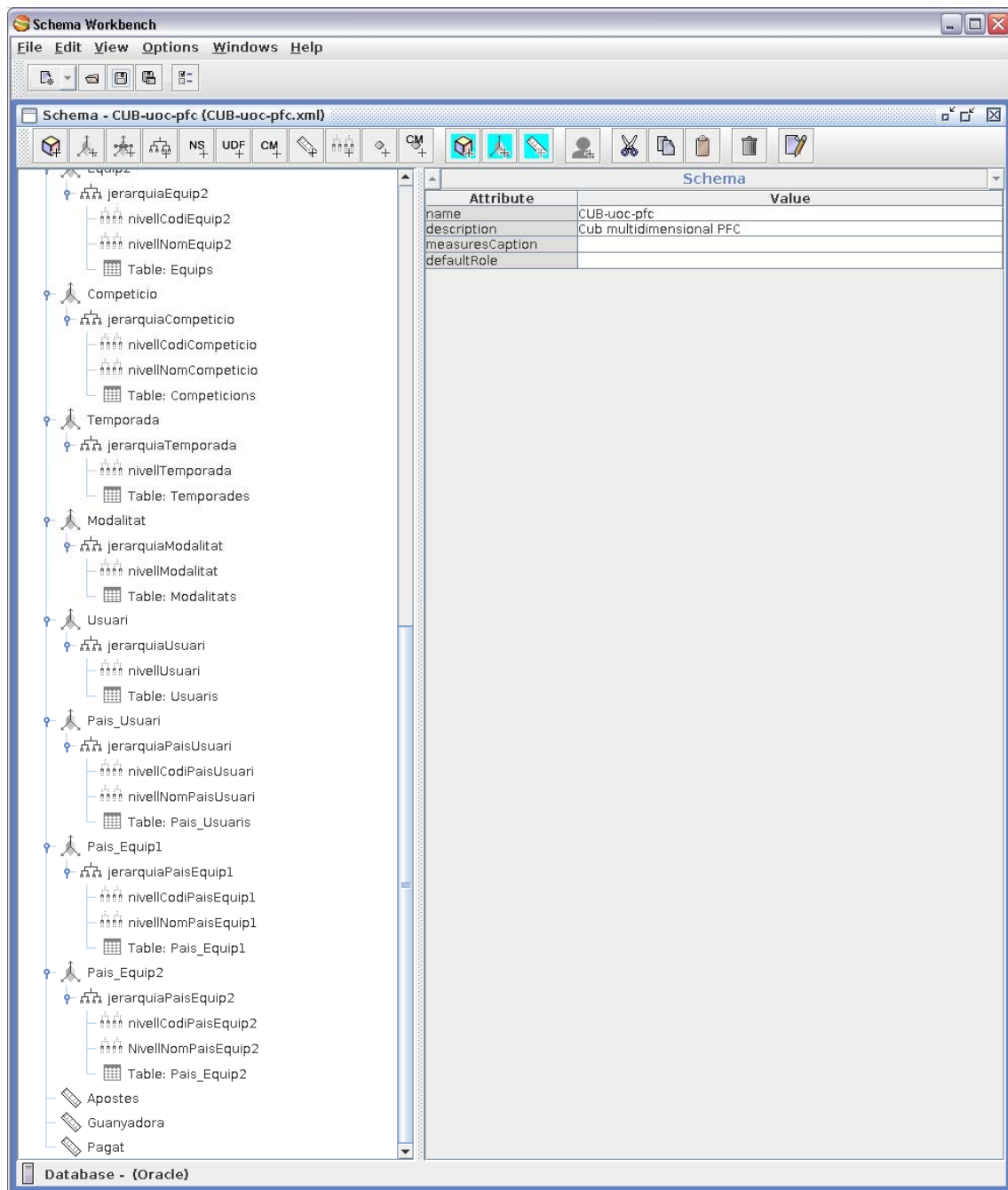


Figura 21: CUB-uoc-pfc (part 2)

Un cop acabat el cub cal publicar-lo, i ho faig a una nova carpeta UOC:

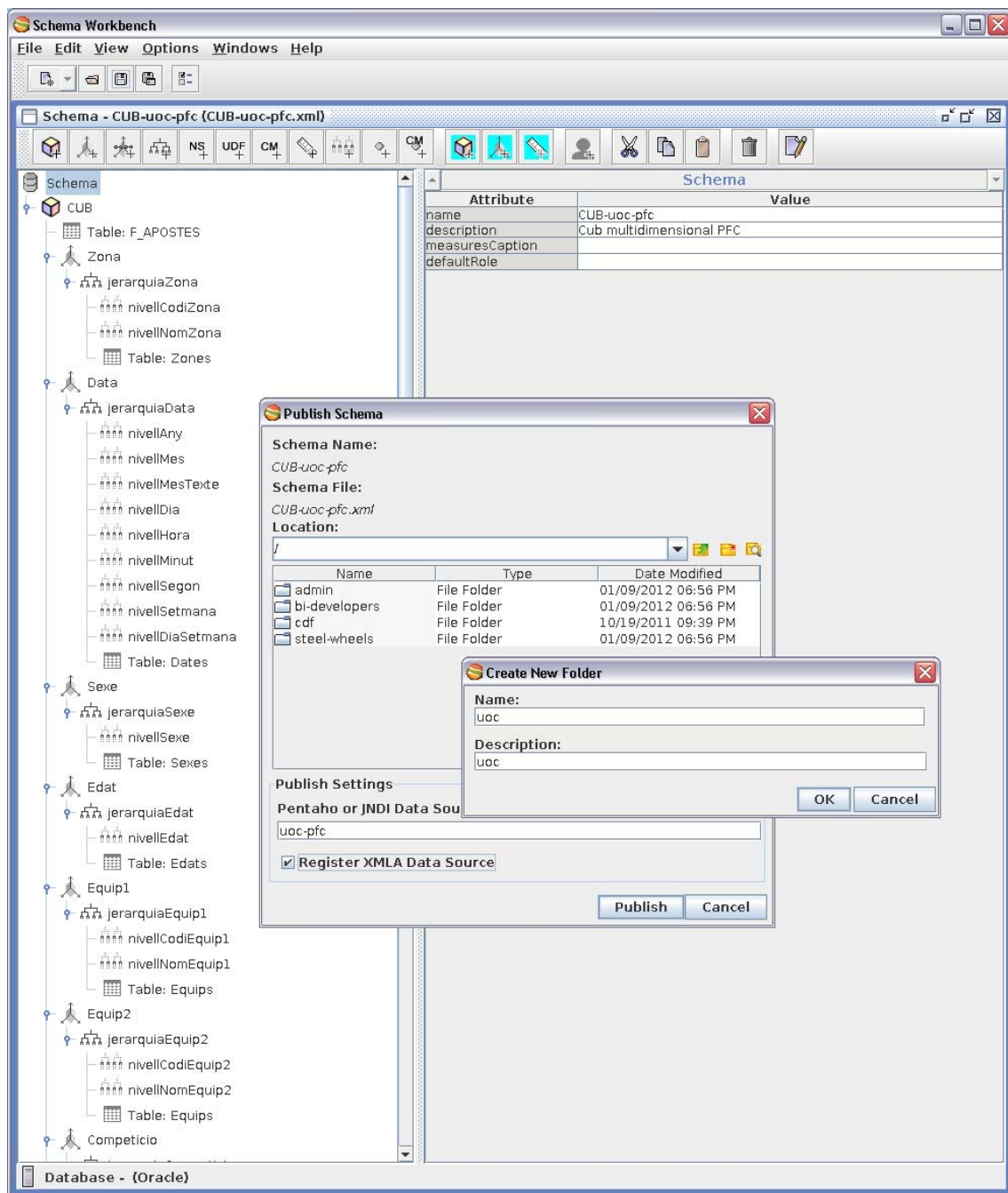


Figura 22: Publicació del cub

Un cop publicat correctament ja podem anar a fer informes:

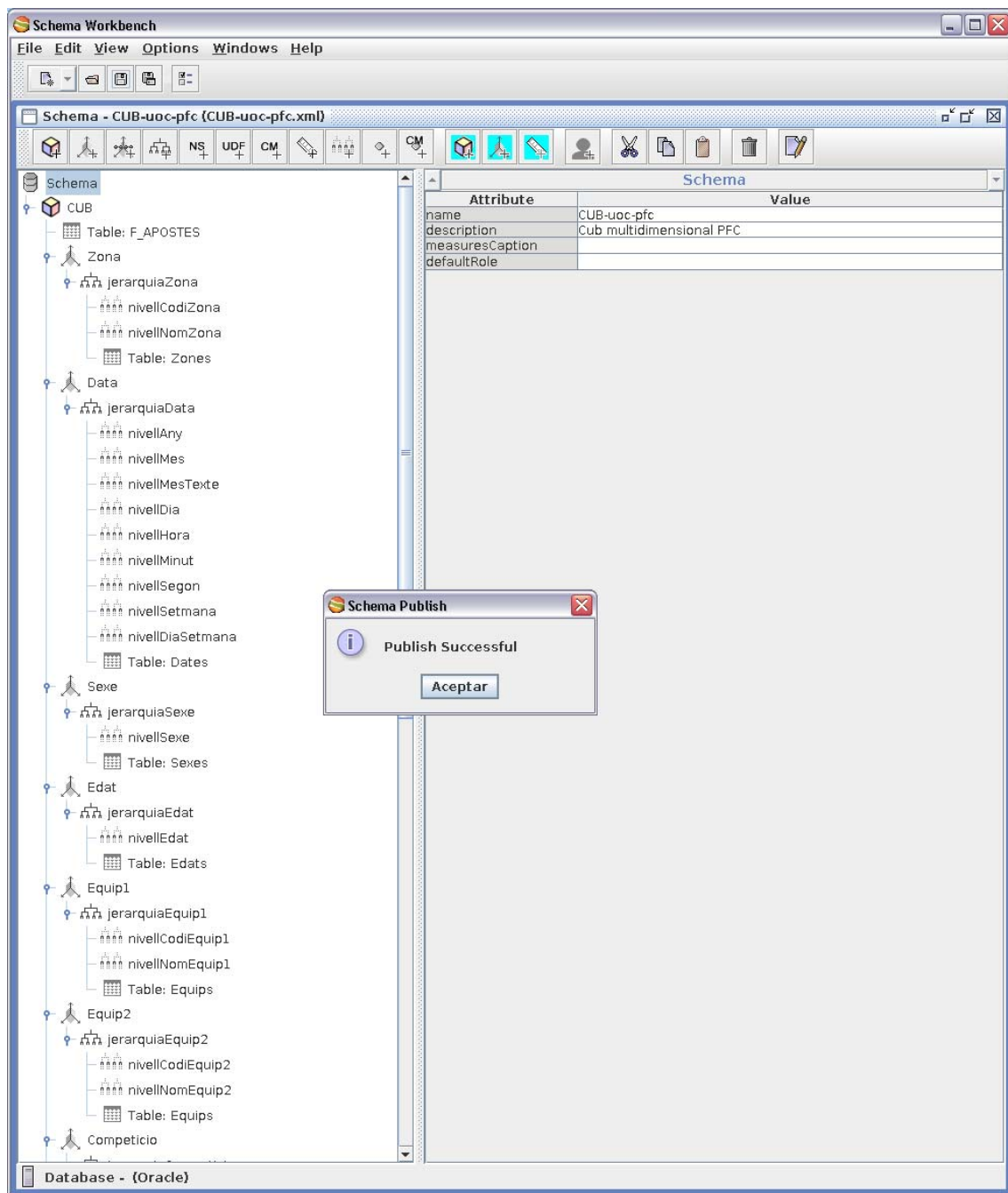


Figura 23: Publicació del cub

C.3 Fitxer XML del cub

Inclouré el fitxer xml del cub que he creat:

CUB-uoc-pfc.xml

C.4 Resultat execucions d'informes a Pentaho

Un cop validats a l'eina web Pentaho podem demanar un nou anàlisi i ens sortirà el cub que acabem de publicar:

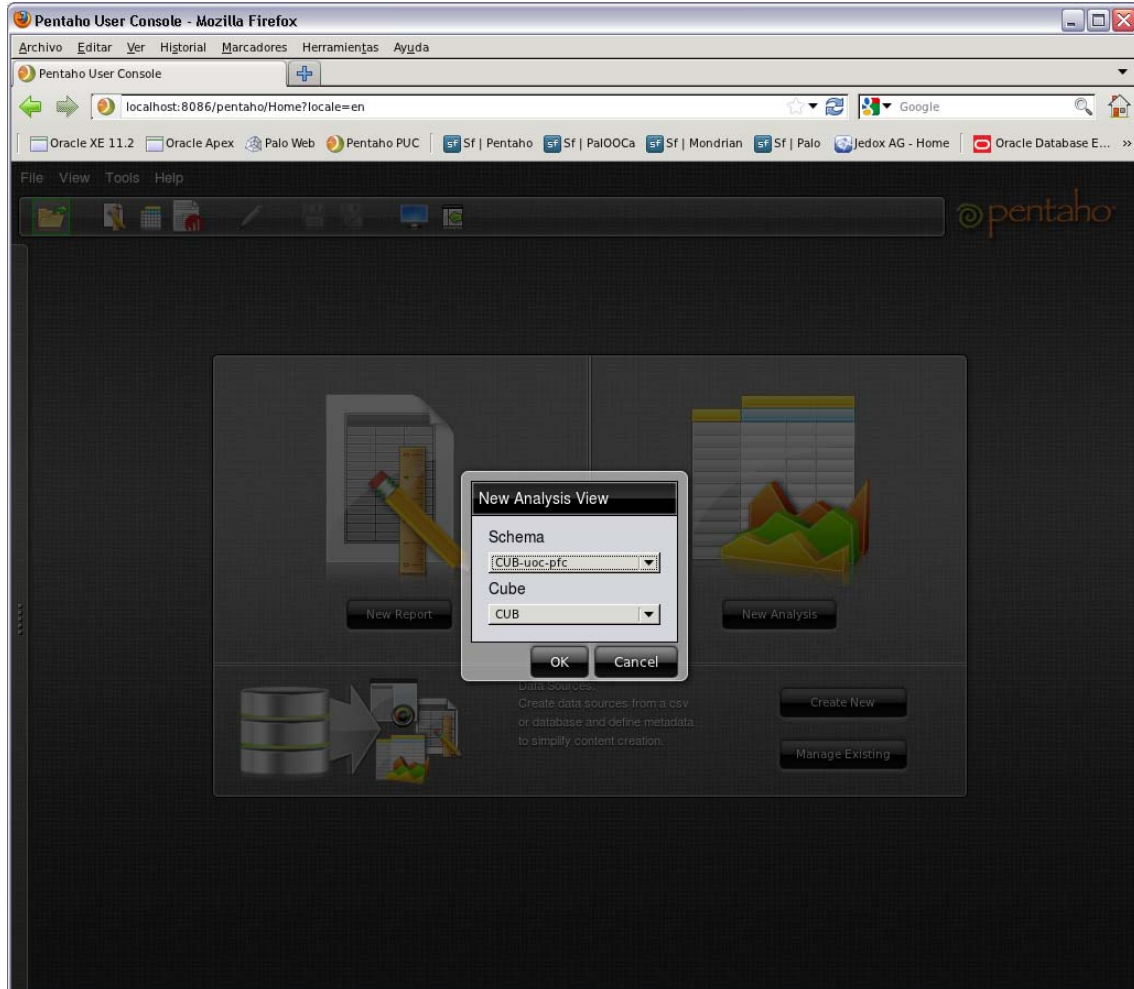


Figura 24: Creació d'un nou anàlisi

La primera vista del nou anàlisi ens mostra totes les dimensions en les columnes per poder-les desplegar o recollir segons convingui:

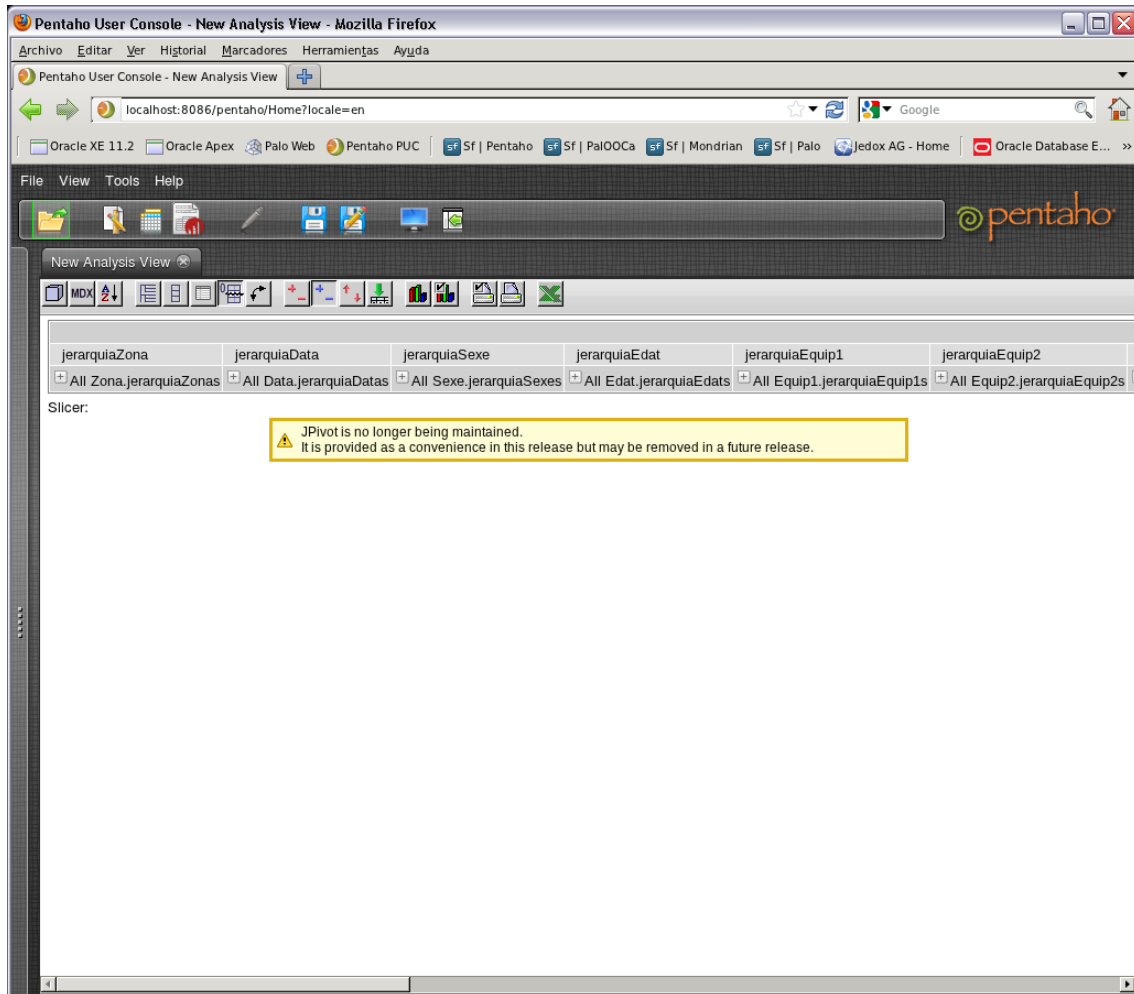


Figura 25: Primera vista de l'anàlisi

Així per cada nivell de jerarquia de les dimensions podem veure les jerarquies de les altres dimensions:

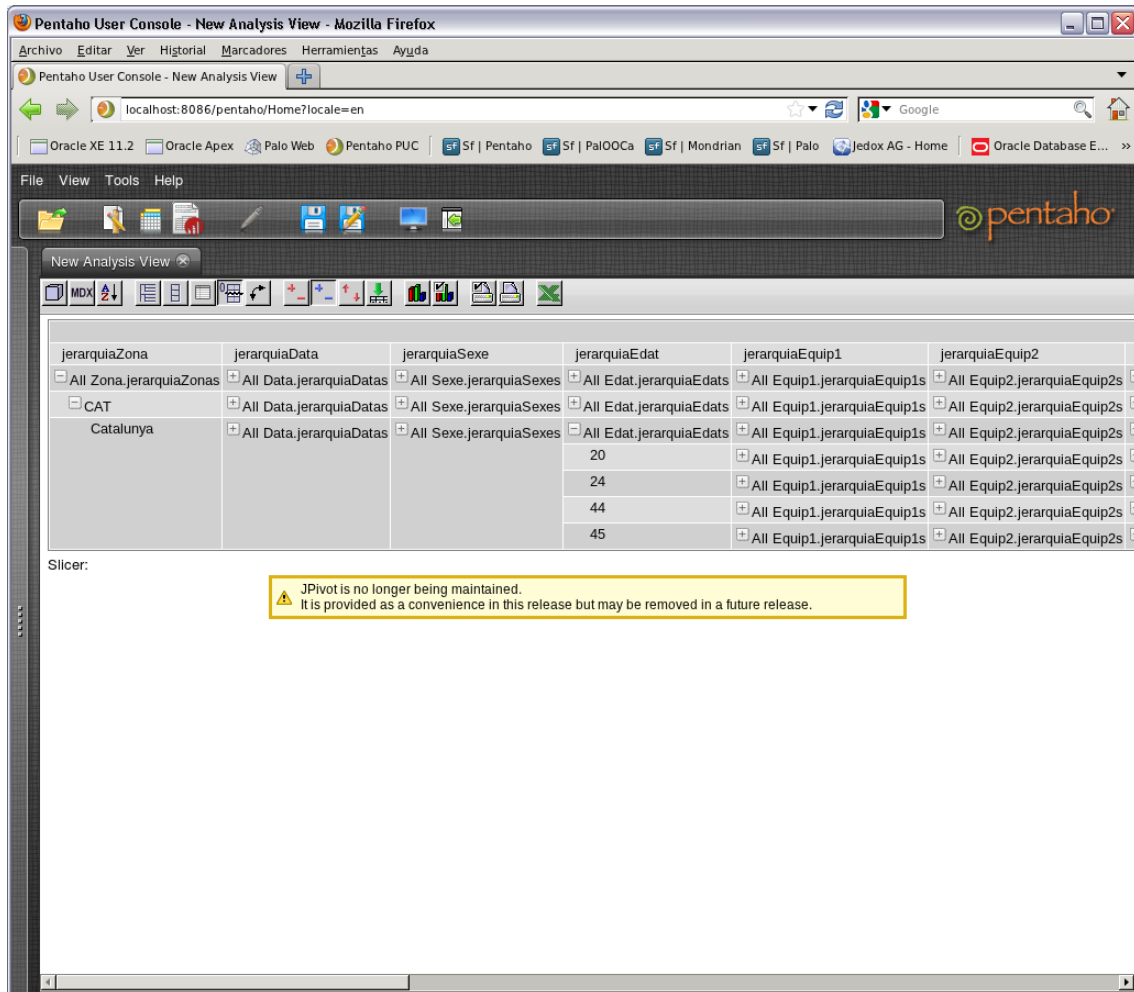


Figura 26: desplegament de l'anàlisi

Si volem simplificar la vista podem reorganitzar-la passant les dimensions de columnes a filtres i deixant només les columnes que ens interessin en cada moment:

The screenshot shows the Pentaho User Console interface in Mozilla Firefox. The browser address bar shows 'localhost:8086/pentaho/Home?locale=en'. The application title is 'Pentaho User Console - New Analysis View'. The interface includes a menu bar (File, View, Tools, Help) and a toolbar with various icons. On the left, there is a sidebar with a tree view showing the hierarchy of dimensions and filters. The main area displays a table with columns for dimensions and a single column for measures.

Columns:

- jerarquiaZona
- jerarquiaSexe
- jerarquiaEdat

Filters:

- jerarquiaCompeticio
- jerarquiaData
- jerarquiaEquip1
- jerarquiaEquip2
- jerarquiaModalitat
- jerarquiaPaisEquip1
- jerarquiaPaisEquip2
- jerarquiaPaisUsuari
- jerarquiaTemporada
- jerarquiaUsuari

Table Data:

jerarquiaZona	jerarquiaSexe	jerarquiaEdat	Medidas
All Zona.jerarquiaZonas	All Sexe.jerarquiaSexes	All Edat.jerarquiaEdats	92
		20	4,5
		24	50
		44	10
		45	27,5
CAT	All Sexe.jerarquiaSexes	All Edat.jerarquiaEdats	92
		20	4,5
		24	50
		44	10
		45	27,5
Catalunya	All Sexe.jerarquiaSexes	All Edat.jerarquiaEdats	92
		20	4,5
		24	50
		44	10
		45	27,5

Slicer:

JPivot is no longer being maintained.
It is provided as a convenience in this release but may be removed in a future release.

Figura 27: Reorganització de les files, columnes i filtres

I de la mateixa manera podem desplegar o recollir les dimensions amb l'anàlisi ja reorganitzat:

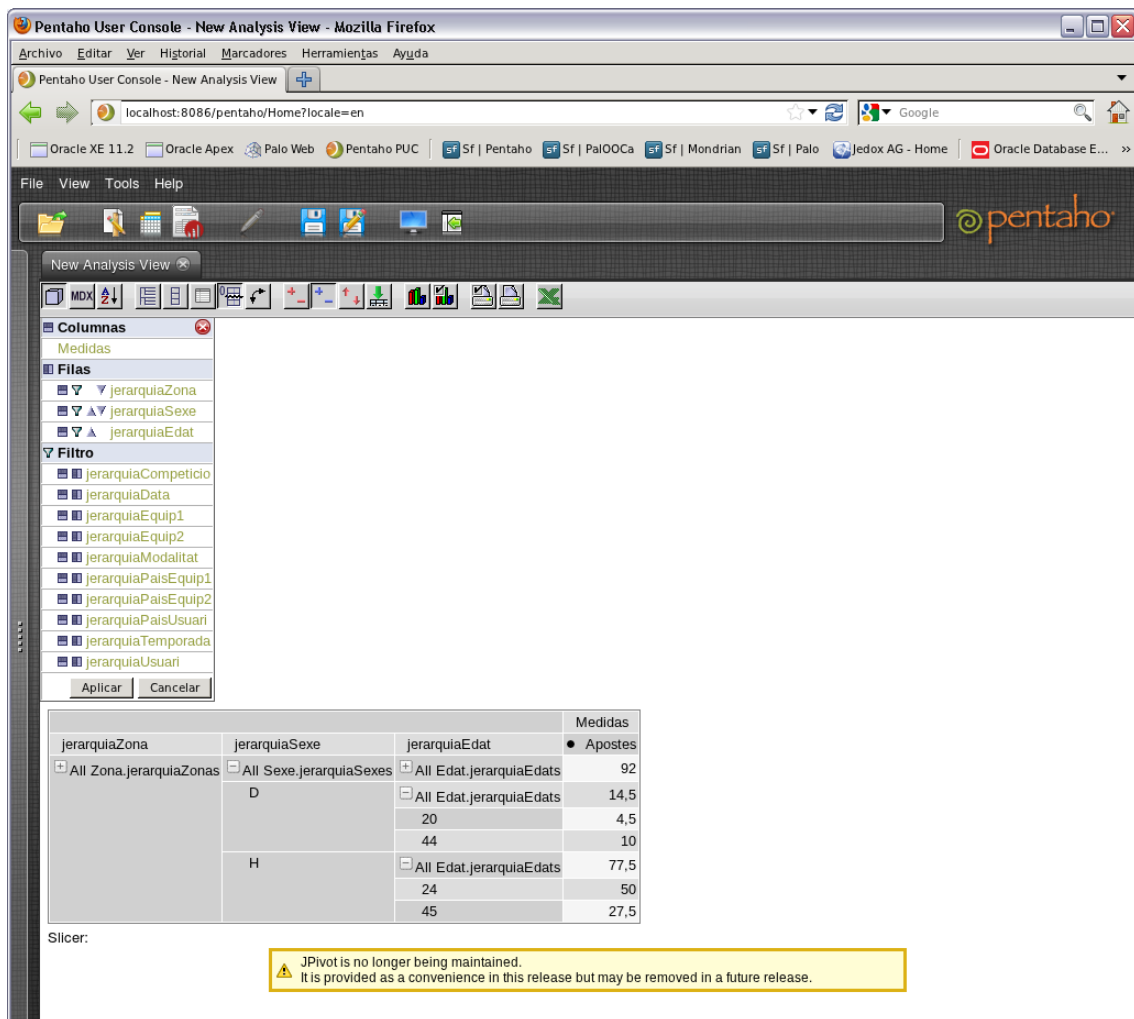


Figura 28: Anàlisi ja reorganitzat

I moltes més respostes a les preguntes formulades al magatzem d'informació:

- Distribució de zones dels apostadors:

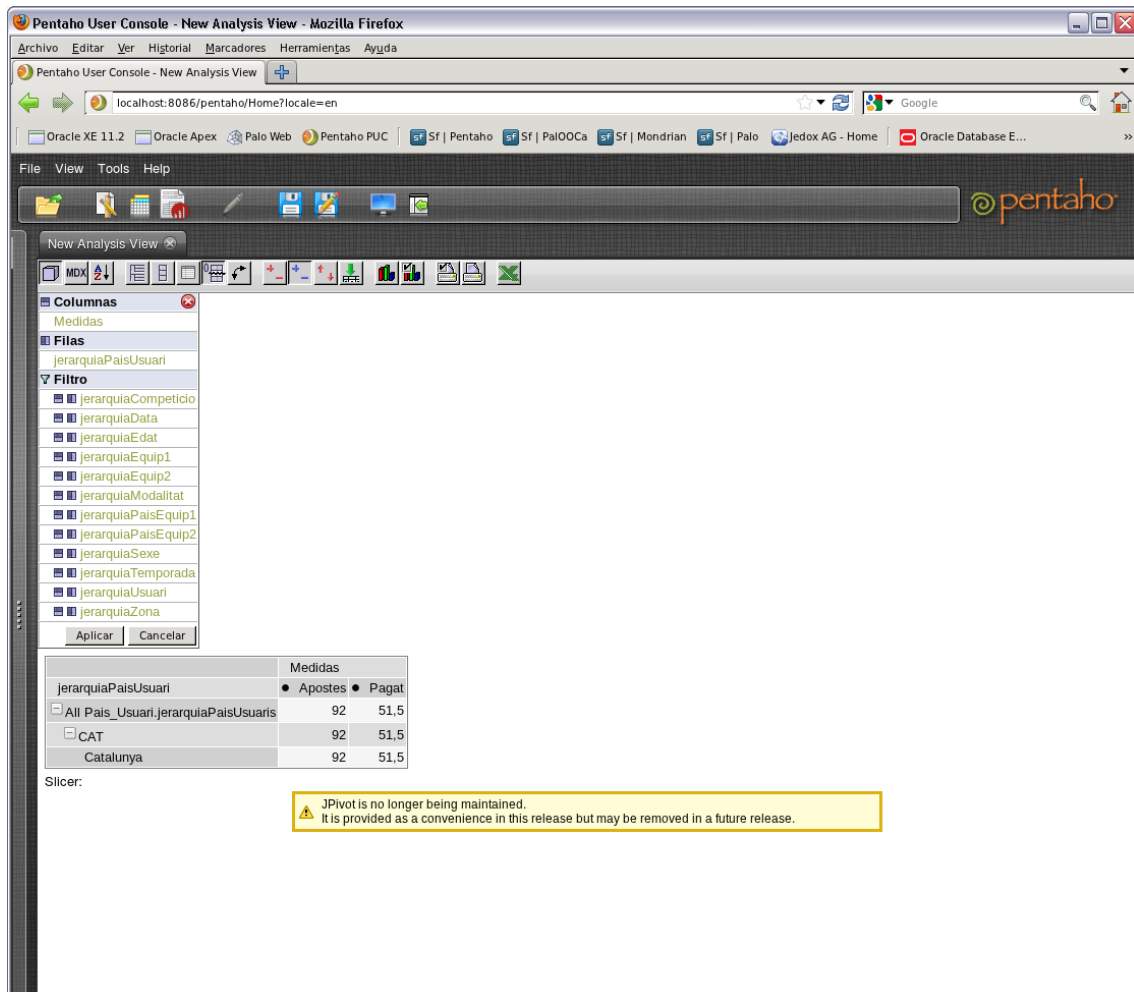


Figura 29: Distribució de zones dels apostadors

- Distribució per edats, apostes i guanyys:

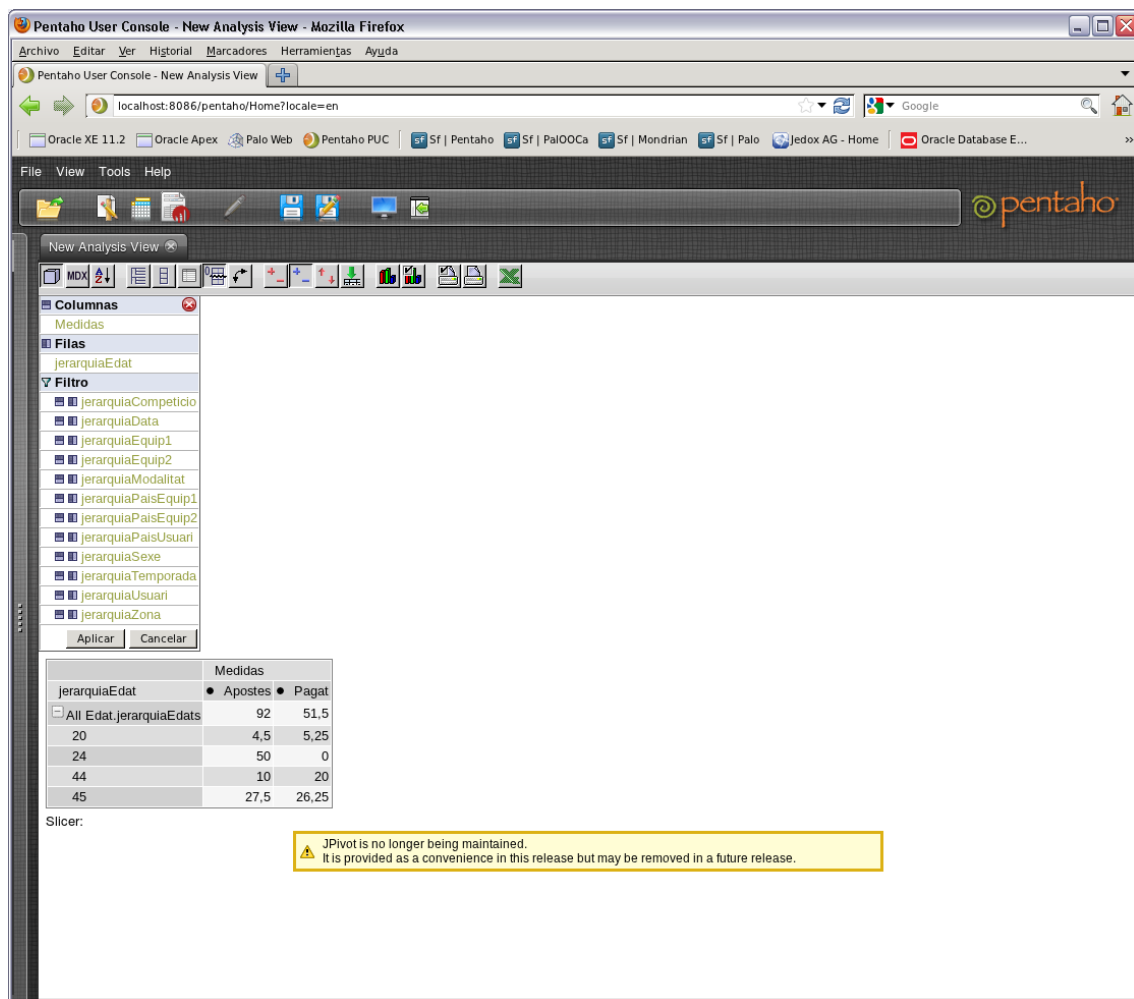


Figura 30: Distribució per edats, apostes i guanyys dels apostadors

- Distribució per equips apostats:

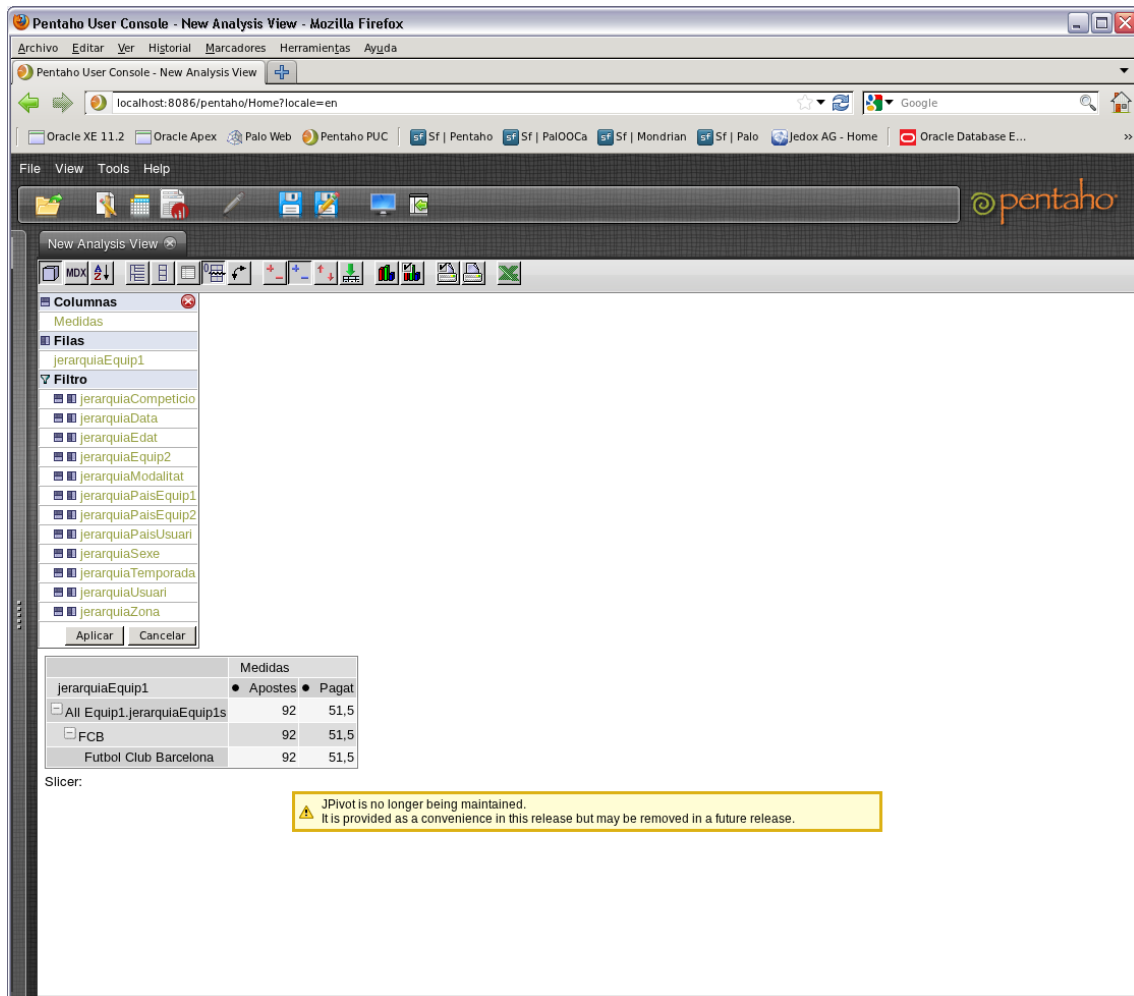


Figura 31: Distribució per equips apostats

- Distribució per sexe dels apostadors:

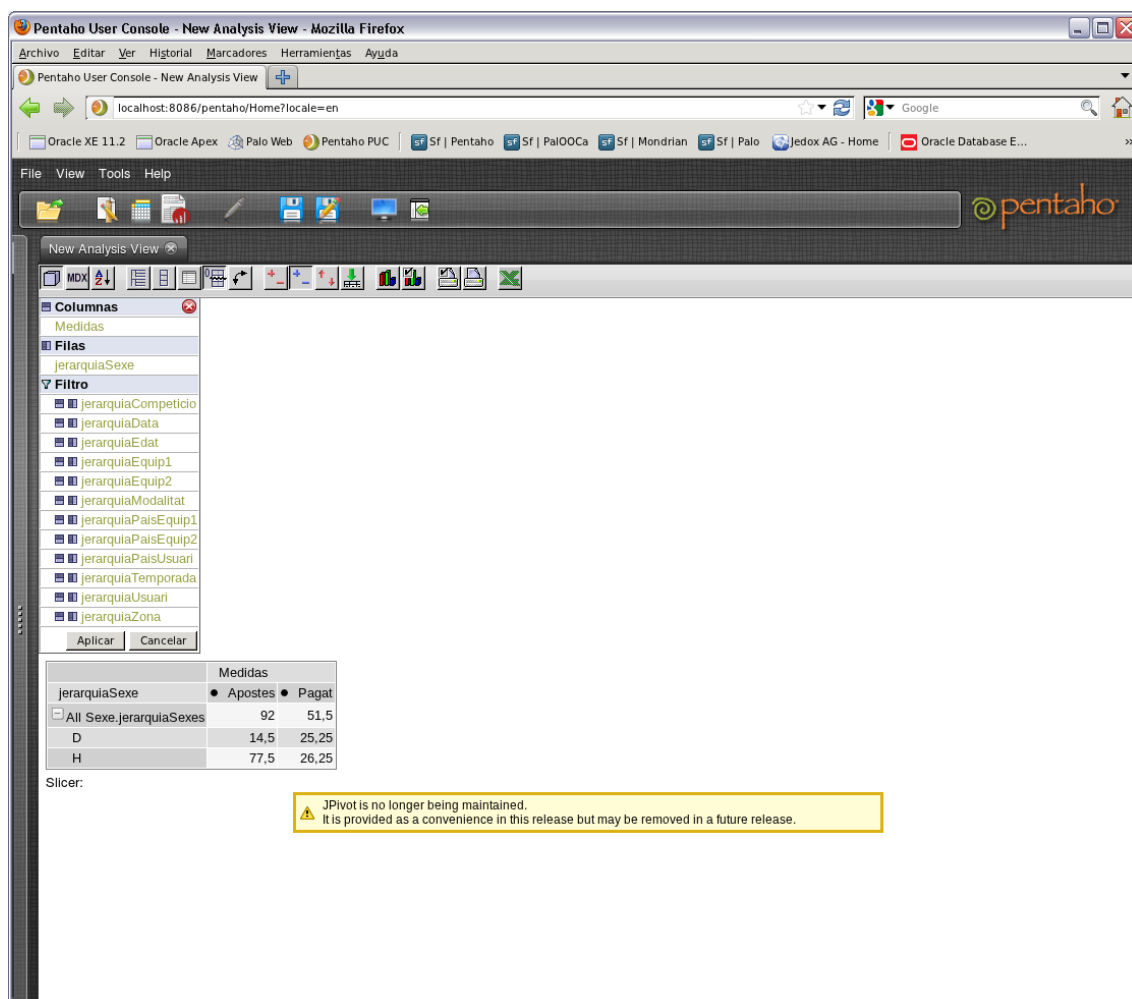


Figura 32: Distribució per sexe dels apostadors

- Distribució per competicions i temporades:

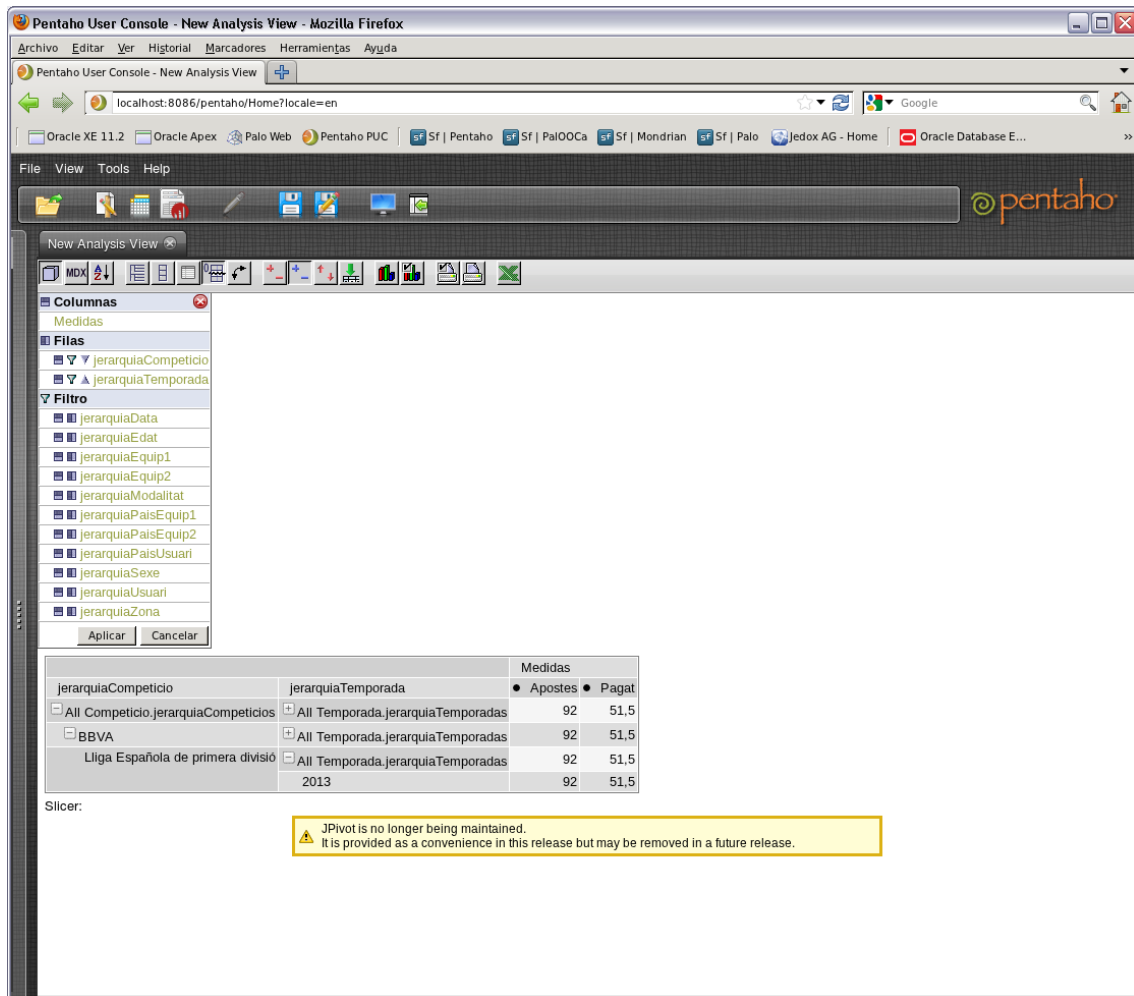


Figura 33: Distribució per competicions i temporades

- Combinació entre edat i sexe amb gràfic:

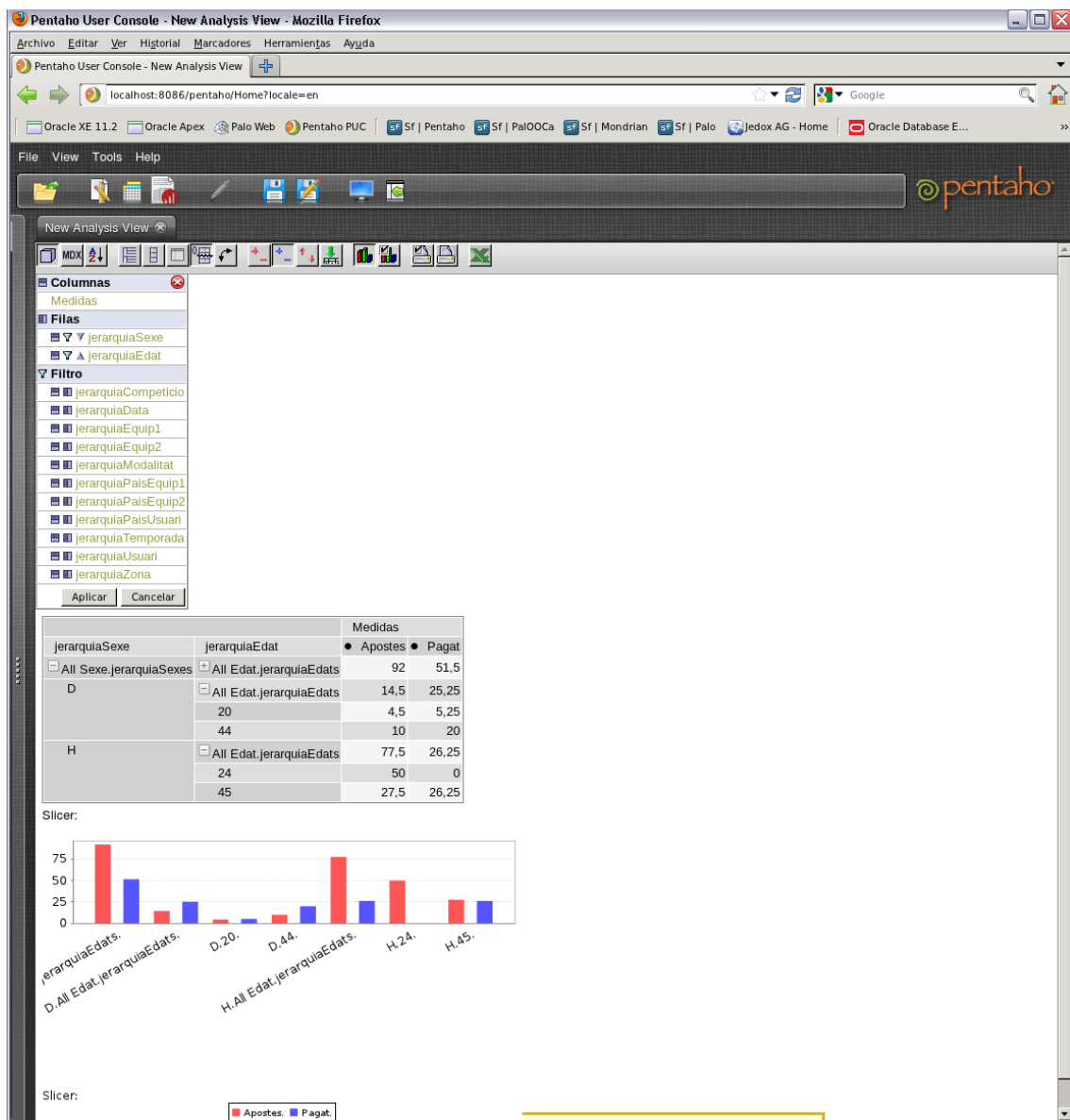


Figura 34: Combinació d'edat i sexe amb gràfic

Annex D: Planificació

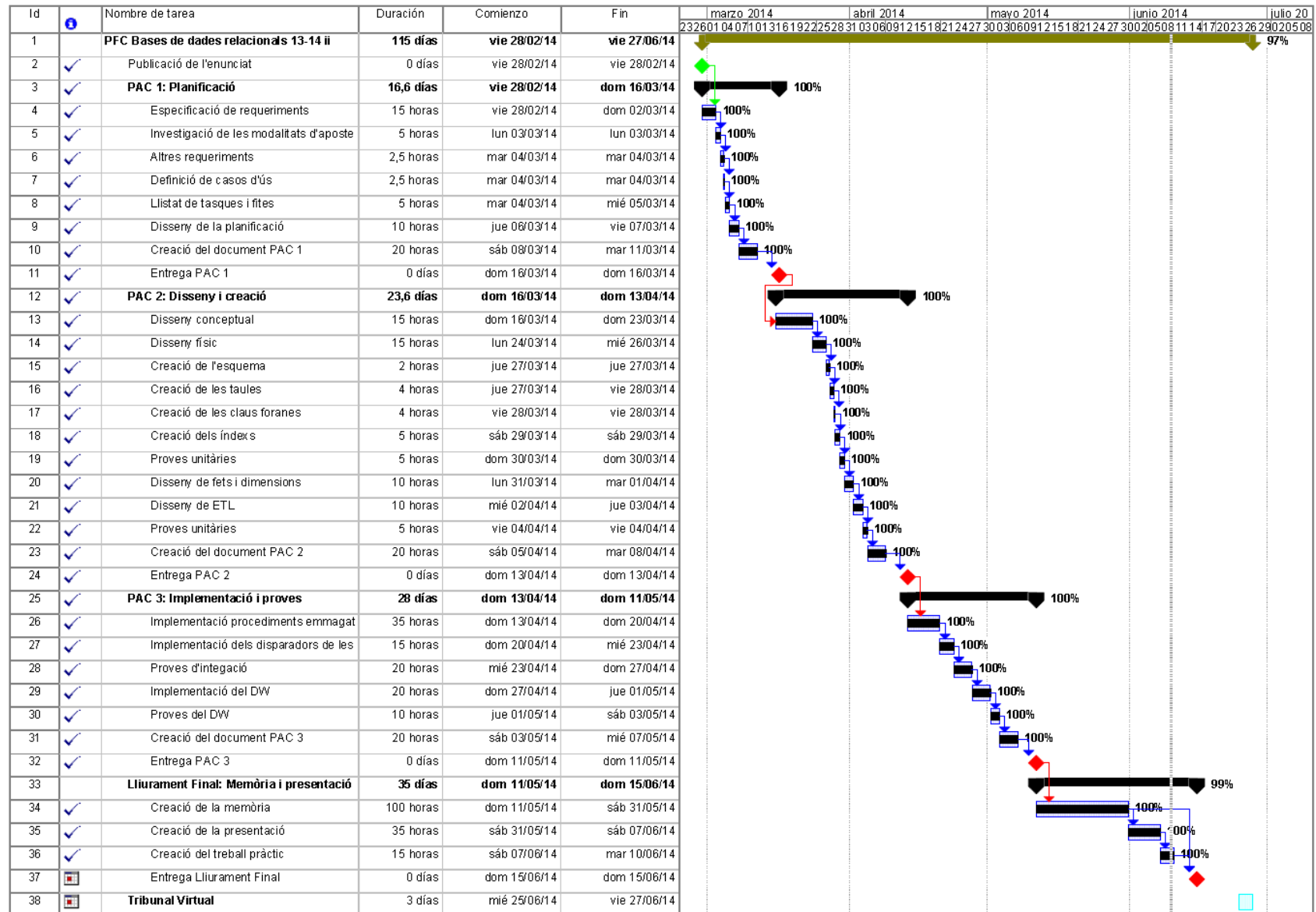


Figura 35: Diagrama de Gantt del projecte