

CONSTRUCCIÓ I EXPLOTACIÓ D'UN MAGATZEM DE DADES PER A L'ANÀLISI D'INFORMACIÓ IMMOBILIÀRIA

MEMÒRIA

Tardor 2010



Alumne : *David Caviedes Zambrano (ETIS)*

Consultor : *Pere Juanola Juanola*

AGRAÏMENTS

Voldria agrair especialment a la meva família no només aquest treball, que és la culminació de l'esforç, sinó tots els estudis de l'ETIS, ja que sense el seu ajut i rialles no ho hagués aconseguit. Perdoneu-me de tot cor totes les estones d'estar junts que us he hagut de robar.

M'agradaria agrair també a l'UOC com a institució l'oportunitat que dona a tots els que, com jo, treballen i els hi resulta molt difícil compaginar la seva tasca professional i familiar amb els estudis. Agrair especialment el mètode d'aprenentatge que s'ha triat, perquè considero que estudiar no només és memoritzar i deixar-ho anar a un control, sinó que aprendre significa comprendre el que és essencial d'una matèria, i la tasca de la Universitat és valorar que això s'ha entès, sense artificis ni problemes rebuscats, que per això ja hi ha textos especialitzats i complementaris de gent que porta molts anys investigant i aprofundint en els coneixements corresponents.

Voldria acabar amb una cita dedicada als meus fills, perquè entenguin el perquè continuo dedicant temps a ampliar coneixements:

“Quan ets jove aprens, quan ets gran entens.”

MARIE VON EBNER ESCHENBACH

RESUM

Títol : CONSTRUCCIÓ I EXPLOTACIÓ D'UN MAGATZEM DE DADES PER A L'ANÀLISI D'INFORMACIÓ IMMOBILIÀRIA

Àrea del TFC : MAGATZEMS DE DADES

Paraules clau : Magatzem de Dades, DataWarehouse, DataMart, Anàlisi d'informació immobiliària, Oracle, Business Intelligence, TFC, Sostre per a Tothom

Resum

Aquest TFC (Treball Fi de Carrera) consisteix en crear un projecte de magatzem de dades per a una empresa immobiliària que li permetrà incorporar, modificar i explotar la informació de que disposen d'una forma molt eficient i productiva. La primera etapa de la creació del magatzem consistirà en la definició del projecte, cosa que implicarà determinar els seus objectius i el seu abast, a més a més de detallar la planificació del projecte. La segona fase consistirà en realitzar tot el seu disseny el més acuradament possible, ja que d'això dependrà l'èxit del seu funcionament i aprofitament. En la tercera fase s'implementarà el disseny anterior i es construirà el magatzem de dades, i també es crearan els informes demanats per l'empresa mitjançant eines de Business Intelligence. Com a resultat final, s'haurà construït una base de dades en forma de cub multidimensional (DataWarehouse) amb les dades de tots els immobles que gestiona o ha gestionat l'empresa. Aquesta estructura permet extreure informació sota les diferents perspectives de negoci que puguin interessar a la immobiliària, sense que s'hagin de fer costoses manipulacions de dades que es requeririen en els casos de disposar de SGBD (Sistemes de Gestió de Bases de Dades) convencionals. L'empresa disposarà de diferents informes predissenyats segons els seus requeriments inicials que explotarà de forma dinàmica, transparent i intuïtiva mitjançant l'eina Oracle Discoverer, i podrà crear-ne de nous segons les seves necessitats futures, sense que els usuaris hagin d'aprendre a programar en SQL (Structured Query Language), llenguatge declaratiu d'accés a les bases de dades.

ÍNDEX

AGRAÏMENTS	2
RESUM	3
1 INTRODUCCIÓ	8
1.1 JUSTIFICACIÓ DEL PROJECTE	8
1.2 OBJECTIUS	9
1.3 ENFOCAMENT I METODOLOGIA.....	10
1.4 PLANIFICACIÓ DEL PROJECTE	11
1.5 PRODUCTES OBTINGUTS.....	16
1.6 DESCRIPCIÓ BREU DELS ALTRES CAPÍTOLS DE LA MEMÒRIA.....	16
2. ANÀLISI	18
2.1 DADES D'ORIGEN	18
2.1.1 Fitxer de text del client.....	18
2.1.2 Anàlisi previ i problemes detectats	18
2.1.3 Decisions preses.....	19
2.2 DIAGRAMES DE CASOS D'ÚS	20
2.3 DIAGRAMA DEL MODEL CONCEPTUAL	21
2.3.1 Fet rellevant.....	21
2.3.2 Granularitat	22
2.3.3 Dimensions	22
2.3.4 Atributs de cada dimensió	22
2.3.5 Descriptors i jerarquies d'agregació	23
2.3.6 Mesures a fer servir.....	23
2.3.7 Cel·les	24
2.3.8 Restriccions d'integritat.....	24
2.3.9 Viabilitat.....	24
3. DISSENY	26
3.1 ARQUITECTURA DEL SOFTWARE	26
3.2 ARQUITECTURA DEL HARDWARE	27
3.3 DISSENY LÒGIC.....	27
3.4 DISSENY FÍSIC.....	28

3.4.1 Taula TEMPS	28
3.4.2 Taula LOCALITZACIO	28
3.4.3 Taula PROPIETAT	29
3.4.4 Taula VALOR	29
3.5 TRANSFORMACIÓ DE LES DADES.....	31
3.5.1 Opcions i decisions prèvies	31
3.5.2 Creació del MMDD i càrrega de les dades.....	32
3.6 CREACIÓ D'INFORMES A L'ORACLE DISCOVERER DESKTOP	35
3.7 INFORMES GENERATS	36
3.8 DETERMINACIÓ DEL PIS TIPUS ANDORRÀ I ESPANYOL.....	36
<i>CONCLUSIONS.....</i>	<i>52</i>
<i>LÍNIES D'EVOLUCIÓ FUTURA.....</i>	<i>53</i>
<i>GLOSSARI.....</i>	<i>54</i>
<i>BIBLIOGRAFIA</i>	<i>55</i>

ÍNDIX DE FIGURES

Il·lustració 1 - Taula de la planificació del TFC	12
Il·lustració 2 - Diagrama de Gantt de la planificació del projecte.....	15
Il·lustració 3 - Diagrama d'ús dels usuaris	20
Il·lustració 4 - Diagrama d'ús dels administradors	21
Il·lustració 5 - Model conceptual del MMDD	25
Il·lustració 6 - Diagrama d'arquitectura del software.....	26
Il·lustració 7 - Diagrama d'arquitectura del hardware del MMDD	27
Il·lustració 8 - Disseny lògic del MMDD	28
Il·lustració 9 - Diagrames del disseny físic del MMDD.....	30
Il·lustració 10 - Consultes SQL de la creació de les taules del MMDD.....	33
Il·lustració 11 - Aplicació da càrrega d'Oracle	34
Il·lustració 12 - Blocs d'instruccions del fitxer de control de SQL*Loader.....	35
Il·lustració 13 - Pestanyes de modificació del full de treball	36
Il·lustració 14 - Nombre d'immobles per província - opció de mes i any.....	37
Il·lustració 15 - Nombre d'immobles per província - resum de tots els mesos.....	37
Il·lustració 16 - Nombre d'immobles per tipologia - opció de mes i any.....	38
Il·lustració 17 - Nombre d'immobles per tipologia - resum de tots els mesos.....	38
Il·lustració 18 - Nombre d'immobles per planta - opció de mes i any.....	38
Il·lustració 19 - Nombre d'immobles per superfície - Resum de tots els mesos	39
Il·lustració 20 - Llistat d'immobles que han baixat de preu - per mes i any	39
Il·lustració 21 - Llistat d'immobles que han baixat de preu - sense repeticions	39
Il·lustració 22 - Evolució de preus per província - per mes i any	40
Il·lustració 23 - Gràfica de l'evolució de preus per província - per mesos.....	40
Il·lustració 24 - Evolució de preus per tipologia - per mes i any.....	40
Il·lustració 25 - Evolució de preus per planta - per mes i any.....	41
Il·lustració 26 - Preus màxims, mínims i mitjans per zona ciutat - opció de mes i any	41
Il·lustració 27 - Preus màxims, mínims i mitjans per província - resum de tots els mesos	42

TFC MAGATZEM DE DADES - MEMÒRIA

Il·lustració 28 - Preus màxims, mínims i mitjans per tipologia - opció de mes i any	42
Il·lustració 29 - Preus màxims, mínims i mitjans per tipologia - resum de tots els mesos	42
Il·lustració 30 - Preus màxims, mínims i mitjans per planta - opció de mes i any	43
Il·lustració 31 - Preus màxims, mínims i mitjans per superfície - resum de tots els mesos.....	43
Il·lustració 32 - Superfície màxima, mínima i mitjana per província - resum de tots els mesos	43
Il·lustració 33 - Superfície màxima, mínima i mitjana per tipologia - opció de mes i any	44
Il·lustració 34 - Superfície màxima, mínima i mitjana per tipologia - resum de tots els mesos	44
Il·lustració 35 - Superfície màxima, mínima i mitjana per dormitoris - opció de mes i any	44
Il·lustració 36 - Superfície màxima, mínima i mitjana per planta - resum de tots els mesos	45
Il·lustració 37 - Temps mitjà de venda per província - opció de mes i any.....	45
Il·lustració 38 - Temps mitjà de venda per província - resum de tots els mesos.....	46
Il·lustració 39 - Temps mitjà de venda per tipologia - opció de mes i any	46
Il·lustració 40 - Temps mitjà de venda per tipologia - resum de tots els mesos	46
Il·lustració 41 - Temps mitjà de venda per planta - opció de mes i any	47
Il·lustració 42 - Nombre d'immobles venuts per província - resum de tots els mesos	47
Il·lustració 43 - Nombre d'immobles venuts per tipologia - opció de mes i any	47
Il·lustració 44 - Nombre d'immobles venuts per tipologia - resum de tots els mesos	48
Il·lustració 45 - Nombre d'immobles venuts per planta - opció de mes i any	48
Il·lustració 46 - Consultes del pis tipus Andorrà	49
Il·lustració 47 - Consultes sobre l'immoble tipus d'Espanya	50

1 INTRODUCCIÓ

El projecte d'enginyeria informàtica d'aquest TFC requereix aplicar totes aquelles habilitats, tècniques i coneixements adquirits durant els estudis a l'anàlisi del problema plantejat, per cercar la millor de les solucions possibles. Aquest projecte, al igual que en els que es desenvolupen en el món empresarial, està limitat en el temps i en els recursos disponibles, cosa que fa que la cerca de la solució s'hagi de fer d'una forma àgil i pragmàtica.

1.1 JUSTIFICACIÓ DEL PROJECTE

Tradicionalment, el tractament de la informació de les organitzacions s'han basat en bases de dades de petit tamany, transaccionals i específiques per a cada àrea de negoci, per gestionar les seves operacions quotidianes. Tanmateix, en l'actualitat i cada vegada més, les persones encarregades de la presa de decisions en les empreses han requerit tenir accés a totes les dades del negoci, amb independència de la seva ubicació. Aquesta necessitat ha provocat que es creï un repositori centralitzat o magatzem de dades, alimentat per dades procedents de les diverses fonts d'informació de l'empresa, mantingudes per les seves diferents unitats operatives.

Un aspecte important dels magatzems de dades és la volatilitat de les dades, ja que la informació del repositori no s'ha de modificar ni eliminar, és a dir, no ha de ser volàtil. Per tant, no només s'accedirà als valors actualment emmagatzemats, sinó que són necessaris els seus valors històrics, per facilitar als directius una anàlisi exhaustiva del negoci, de les seves possibilitats i de les possibles tendències futures.

Sobre aquesta base, l'empresa immobiliària Sostre per a Tothom ha encarregat la construcció d'un magatzem de dades per poder extreure informació del seu catàleg d'immobles que li permeti comparar preus, zones i tipologies d'habitatges. D'aquesta forma intentaria ajustar l'oferta i la demanda dels mateixos, obtenint un avantatge competitiu amb respecte altres empreses de la competència.

El Data Warehouse (DW) que dissenyarem s'ajustarà a les característiques generals dels magatzems de dades (Inmon, 1993):

- Es trobarà organitzat pels temes més importants per a l'organització (en aquest cas per productes), en lloc de per l'àrea organitzativa (administració, en el cas de Sostre per a Tothom).
- La informació es trobarà integrada, és a dir, que se li donarà coherència a les dades provinents de diverses fonts, per donar una vista unificada als usuaris.
- Es guardaran les dades històriques, per la qual cosa es donarà una visió de variabilitat en el temps.
- La informació no serà volàtil. Les dades no s'actualitzaran en temps real, sinó que es faran processos periòdics per incorporar dades noves que sempre incrementaran la BBDD actual, sense substituir les anteriors.

Un cop el magatzem de dades reculli la informació que ara mateix Sostre per a Tothom té dispersa en diferents fulls de càlcul, els usuaris de l'empresa podran demanar informes predefinits que l'informin de diferents característiques i preus dels immobles des de diverses perspectives. Per aquesta tasca farem servir eines d'anàlisi sobre grans volums de dades o també conegudes com Business Intelligence (BI). Hi ha dos tipus d'eines de BI, les de processament analític en línia (OLAP) o les de minerias de dades. En el cas que ens ocupa, es farà servir Oracle Discoverer, una eina del primer tipus, perfecta per les necessitats actuals del client.

En resum, integrarem totes les dades (disperses, desestructurades i incoherents) proporcionades pel client, de forma uniforme, en un únic repositori a partir del qual els usuaris puguin executar consultes amb facilitat, generar informes i realitzar diferents anàlisis per a la presa de decisions.

1.2 OBJECTIUS

L'objectiu principal del projecte és el disseny, construcció i explotació d'un magatzem de dades a partir de la informació disponible en una base de dades transaccional.

Una forma més mesurable per assolir aquest objectiu principal és dividir-lo en objectius d'un abast més assequible i immediat:

1. Objectiu 1 : Anàlisi preliminar de les dades i planificació. Realitzar una anàlisi preliminar de les dades facilitades per determinar mancances, errors, inconsistències, dimensions, atributs i indicadors de les dades. A partir d'aquesta anàlisi fer una planificació amb les etapes del treball el més real possible i preveient imprevistos. Es realitzarà una familiarització amb l'entorn de treball previst (eina Oracle 10 versió Express, Developer, Microsoft Visio, Microsoft Project, VirtualBox,...), eines que ens permetran fer el disseny del magatzem, planificar-ho en el temps i la càrrega de les dades.
2. Objectiu 2 : Anàlisi de requeriments i disseny conceptual tècnic. Fer una anàlisi detallada de les dades a partir de la feta preliminarment. Realitzar el disseny amb la descripció del model dimensional que es preveu per donar suport a les necessitats dels usuaris. Realitzar el disseny dels processos d'extracció de les dades a alt nivell, definint els processos i un pseudocodi que ajudi en la construcció del magatzem de dades.
3. Objectiu 3 : Implementació. Construcció del magatzem de dades dissenyat amb l'eina Oracle i realitzar la càrrega de les dades. Instal·lació de l'eina d'explotació de les dades i finalment, construir els informes i analitzar la informació obtinguda per verificar el resultat i detectar possibles errors.

Igualment, podem definir altres objectius més generals relacionats amb el TFC:

- Familiaritzar-se amb la metodologia, vocabulari i eines disponibles per a la construcció d'un magatzem de dades (MMDD).

- Extreure conclusions sobre els avantatges, complexitat i problemes en la creació dels MMDD.
- Analitzar la dificultat d'homogeneïtzar les dades originals per integrar-les en el MMDD, probablement el principal coll d'ampolla del procés.
- Familiaritzar-se amb les eines de BI per a l'explotació de les dades i la creació d'informes.
- Veure els avantatges de dividir un projecte complex i de llarga durada en fites més assumibles econòmicament pel client i generant uns resultats més immediats.
- Familiaritzar-se amb tot el que és inherent al desenvolupament d'un projecte, des de la seva concepció, planificació, disseny i construcció, a la redacció de la memòria (documentació) i la presentació al client.

1.3 ENFOCAMENT I METODOLOGIA

L'enfocament de qualsevol projecte desenvolupat per a un client es basa en la col·laboració i participació activa del mateix tant en la concepció com en el desenvolupament. L'explicació és que una decisió unilateral mal presa per la nostra part pot donar com a resultat un producte poc útil pels usuaris que l'han de fer servir.

Tot i que de forma succinta, s'ha intentat mantenir aquest enfocament en la metodologia seguida i que es detalla a continuació:

1. Definició i assimilació dels requeriments del client, aportants mitjançant documents i ampliats i focalitzats en les reunions mantingudes amb ells.
2. Anàlisi preliminar de les dades d'origen proporcionades pels clients.
3. Planificació del projecte.
4. Preparació del sistema físic. S'haurà de definir els servidors, BBDD, programari i eines de que es disposa i posar-les a punt.
5. Anàlisi exhaustiva de les dades d'origen i homogeneïtzació de les mateixes.
6. Disseny del model del magatzem, proporcionant el model lògic i el disseny tècnic que s'implementarà. Aquesta fase és crítica, ja que un mal disseny pot implicar que el MMDD no pugui complir els requeriments dels clients, proporcionant la informació que necessiten.
7. Construcció del MMDD. Es crearan la taula de fets i les diferents taules de dimensions en el SGBD escollit (Oracle 10g Express) per contenir el repositori.
8. Preparació de les dades per la càrrega al DW i càrrega de les mateixes al MMDD (és el que es coneix com ETL - Extracció, Transformació i Càrrega).
9. Comprovació de la qualitat de la càrrega i de les dades del repositori.
10. Disseny dels informes pel client amb l'eina de BI, Oracle Discoverer, amb la creació prèvia del cub OLAP amb la visió de negoci que interessi al client.

11. Verificació de la qualitat de les dades proporcionades en els informes construïts.
12. Presentació al client, manteniment del producte i actualitzacions. Aspectes que queden fora de l'abast d'aquest projecte però que són molt importants en l'ús futur del MMDD.

1.4 PLANIFICACIÓ DEL PROJECTE

Segons el pla docent, aquestes són les dates clau que es tenen en compte per a la planificació del treball:

Tasca	Data Clau
Inici	22/09/2011
PAC 1	05/10/2011
PAC 2	09/11/2011
PAC 3	21/12/2011
Lliurament Final	10/01/2012
Debat virtual	23/01/2012 al 26/01/2012

En concret, les tasques a realitzar en cada lliurament són les següents:

- PAC 1 : Pla de treball i anàlisi preliminar dels requeriments
- PAC 2 : Anàlisi exhaustiu dels requeriments i disseny conceptual i tècnic (model dimensional, disseny dels procediments d'extracció,...).
- PAC 3 : Implementació (construcció del magatzem de dades, càrregues, explotació de les dades, construcció d'informes i anàlisi de la informació).
- Lliurament Final : Entrega del magatzem de dades construït, de la memòria del Treball Fi de Carrera i de la presentació virtual.
- Debat virtual: Defensa del treball i de la presentació virtual.

TFC MAGATZEM DE DADES - MEMÒRIA

En aquesta taula es resumeix la planificació del treball en les seves diferents fases:

Id	Nom de la tasca	Durada	Començament	Fi
1	1.- Inici TFC	1 dia	jue 22/09/11	jue 22/09/11
2	1.1.- Lectura Pla Docent	1 dia	jue 22/09/11	jue 22/09/11
3	2.- PAC 1	14 días	jue 22/09/11	mié 05/10/11
4	2.1.- Lectura enunciat TFC	1 dia	jue 22/09/11	jue 22/09/11
5	2.2.- Instal·lació del programari Oracle	1 dia	jue 22/09/11	jue 22/09/11
6	2.3.- Importació informació a BBDD relacional	1 dia	vie 23/09/11	vie 23/09/11
7	2.4.- Anàlisi de les dades proporcionades pel client	2 días	sáb 24/09/11	dom 25/09/11
8	2.5.- Consulta TFC anys anteriors	1 dia	dom 25/09/11	dom 25/09/11
9	2.6.- Redacció de l'apartat d'especificacions de la PAC 1	2 días	lun 26/09/11	mar 27/09/11
10	2.7.- Redacció de l'apartat d'anàlisi preliminar de les dades de la PAC 1	2 días	mié 28/09/11	jue 29/09/11
11	2.8.- Redacció de l'apartat de planificació	2 días	vie 30/09/11	sáb 01/10/11
12	2.9.- Revisió del document final de la PAC 1	1 dia	lun 03/10/11	lun 03/10/11
13	3.- PAC 2	35 días	jue 06/10/11	mié 09/11/11
14	3.1.- Instal·lació del programari Vbox + Discoverer	1 dia	jue 06/10/11	jue 06/10/11
15	3.2.- Consulta bibliografia	2 días	vie 07/10/11	sáb 08/10/11
16	3.3.- Depuració de les dades del client	4 días	lun 10/10/11	jue 13/10/11
17	3.4.- Redacció de la portada de la Memòria	1 dia	vie 14/10/11	vie 14/10/11
18	3.5.- Redacció de l'apartat de planificació de la Memòria	1 dia	sáb 15/10/11	sáb 15/10/11
19	3.6.- Redacció de l'apartat d'Anàlisis de Dades de la PAC 2	10 días	dom 16/10/11	mar 25/10/11
20	3.7.- Redacció de l'apartat de Disseny Conceptual i Tècnic de la PAC 2	10 días	mié 26/10/11	vie 04/11/11
21	3.8.- Revisió del document final de la PAC 2	2 días	sáb 05/11/11	dom 06/11/11
22	4.- PAC 3	42 días	jue 10/11/11	mié 21/12/11
23	4.1.- Redacció del Glossari i Bibliografia de la Memòria	4 días	jue 10/11/11	dom 13/11/11
24	4.2.- Configuració del Magatzem de Dades	3 días	lun 14/11/11	mié 16/11/11
25	4.3.- Búsqueda de solucions als possibles problemes de càrrega	4 días	jue 17/11/11	dom 20/11/11
26	4.4.- Càrrega de les dades al Magatzem de Dades	6 días	lun 21/11/11	sáb 26/11/11
27	4.5.- Disseny i anàlisi preliminar dels Informes	10 días	dom 27/11/11	mar 06/12/11
28	4.6.- Construcció dels Informes pel Client	10 días	mié 07/12/11	vie 16/12/11
29	5.- Lliurament Final del Projecte	20 días	jue 22/12/11	mar 10/01/12
30	5.1.- Redacció de la Memòria	15 días	jue 22/12/11	jue 05/01/12
31	5.2.- Revisió Final de la Memòria	4 días	vie 06/01/12	lun 09/01/12
32	5.3.- Elaboració de la Presentació del Projecte	10 días	sáb 31/12/11	lun 09/01/12
33	6.- Debat del Projecte	4 días	lun 23/01/12	jue 26/01/12
34	6.1.- Debat del Projecte	4 días	lun 23/01/12	jue 26/01/12

Il·lustració 1 - Taula de la planificació del TFC

Explicació de detall de les diferents tasques de la planificació:

1. Inici TFC

- 1.1. Es llegirà el pla docent per entendre correctament l'abast, objectius, temporització i resultats requerits del projecte.

2. PAC 1

- 2.1. Lectura de l'enunciat del projecte a desenvolupar. Anàlisi succint dels informes demanats pel client i de les fites de les diferents PAC's que componen el TFC.
- 2.2. S'instal·larà el programari d'Oracle en local per fer proves mentre no es disposa de la màquina virtual proporcionada.
- 2.3. Aquesta tasca es refereix a importar la informació proporcionada pel client en un fitxer pla a un programari com Access o Excel per poder visualitzar més còmodament les diferents problemàtiques de les dades de cara a la seva homogeneïtzació.

- 2.4. Tasca complementària a l'anterior, on analitzarem de forma general les diferents problemàtiques de les dades d'origen per la càrrega posterior.
- 2.5. Tasca de documentació sobre treballs d'anys anteriors per poder assimilar les millors idees de cada treball i aplicar-les en el treball en aquest projecte.
- 2.6. Redacció de l'apartat d'especificacions, que inclou la introducció amb els requeriments dels clients, els objectius, els informes requerits, la descripció de les dades d'origen, el detall dels elements d'anàlisi identificats, i les especificacions tècniques i el termini.
- 2.7. Elaboració de l'apartat de l'anàlisi preliminar de les dades, identificant el format d'origen, els diferents camps i registres, les seves problemàtiques i una valoració i metodologia a fer servir per a la seva homogeneïtzació.
- 2.8. Realitzar l'apartat de planificació d'una forma realista i fixant les dates claus segons les indicacions de l'enunciat del projecte i deixant marge per imprevistos.
- 2.9. Revisió final de la PAC 1, corregint aspectes gramaticals i d'estil de cara a la millor presentació possible.

3. PAC 2

- 3.1. Instal·lació de la màquina virtual proporcionada pels professors i consultors de l'assignatura que conté les eines necessàries pel desenvolupament del TFC i on s'implementarà el producte final.
- 3.2. Tasca de documentació relacionada amb el disseny lògic i físic dels MMDD.
- 3.3. Depuració de les dades d'origen i on prendrem una sèrie de decisions que d'acord amb el client serviran per homogeneïtzar les dades i tractar els errors més evidents.
- 3.4. Es començarà a redactar la memòria final del projecte, començant pels aspectes més bàsics, com la portada.
- 3.5. També redactarem l'apartat de planificació, recollint aquelles recomanacions formulades pel consultor en la seva avaluació de la PAC 1.
- 3.6. Redactarem l'apartat de la introducció i disseny conceptual de PAC 2, destacant el fet rellevant, la granularitat, les dimensions i els seus atributs, els descriptors i agregacions, entre d'altres.
- 3.7. Elaborarem l'apartat del disseny lògic, on es defineixen les diferents connexions entre la taula de fets i les diferents dimensions. També redactarem el disseny tècnic on es detallen les taules i els seus camps per optimitzar les consultes del MMDD. Finalment, també es redacta un apartat de planificació de consultes per poder concloure si el disseny és suficientment adequat per respondre als requeriments d'informació del client.
- 3.8. Revisió final de la PAC 2, corregint aspectes gramaticals i d'estil de cara a la millor presentació possible.

4. PAC 3

- 4.1. En primer lloc prepararem els apartats de la memòria de bibliografia i glossari, mentre esperem l'avaluació de la PAC 2.
- 4.2. Crearem el MMDD en la màquina virtual, segons el nostre disseny anterior, amb l'eina Oracle 10g Express i farem petits ajustos en els dissenys si són necessaris.
- 4.3. Farem les transformacions prèvies necessàries a la càrrega de les dades.
- 4.4. Carregarem les dades al MMDD vigilat els possibles errors i problemàtiques.
- 4.5. Dissenyarem els informes amb Oracle Discoverer i valorarem els resultats obtinguts, vigilat la coherència de les dades.
- 4.6. Acabarem de definir els informes demanats pel client i en cuidarem el format.

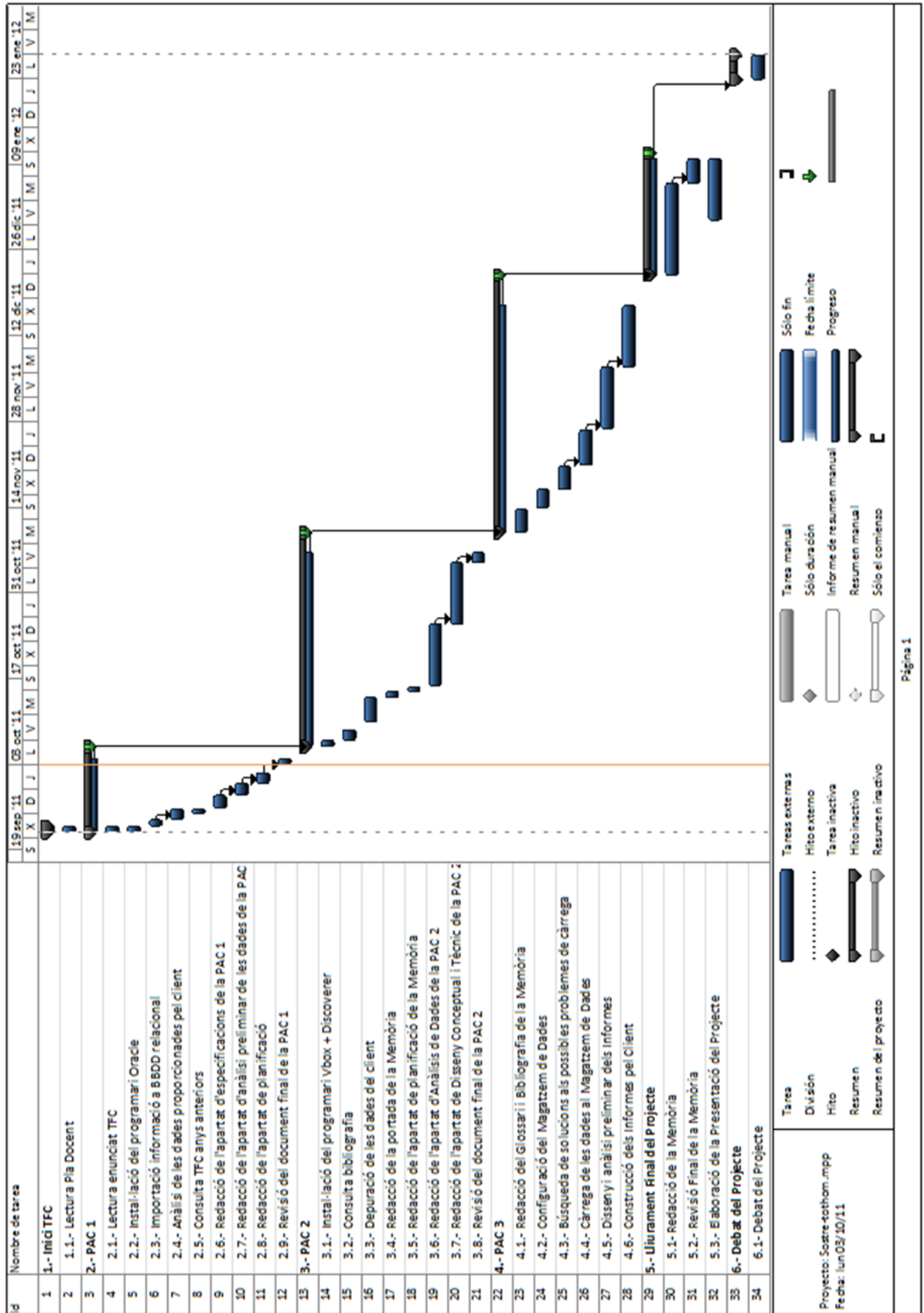
5. Entrega final

- 5.1. Realitzarem la redacció dels apartats de la memòria restants i farem correccions en el producte final si són necessàries.
- 5.2. Revisió final de la memòria, vigilat aspectes formals i gramaticals, i afegint il·lustracions per aclarir conceptes.
- 5.3. Elaboració de la presentació final i entrega de tot el projecte.

6. **Debat del projecte**, on companys i professors podran fer preguntes en relació al projecte presentat (solucions adoptades, decisions preses, etc).

En la plana següent es presenta el diagrama de Gantt amb una visió gràfica de la durada i temporització de les tasques descrites anteriorment:

TFC MAGATZEM DE DADES - MEMÒRIA



Il·lustració 2 - Diagrama de Gantt de la planificació del projecte

1.5 PRODUCTES OBTINGUTS

Durant el projecte s'han obtingut els següents productes:

- PAC 1, que inclou:
 - El pla de treball: planificació temporal de les diferents fases de que consta el projecte.
 - Especificacions del projecte: segons els requeriments del client s'han cercat les solucions més adients per aconseguir l'objectiu de construcció del MMDD.
 - Anàlisi de les dades d'origen: document previ i breu amb els primers problemes detectats sobre les dades d'origen.
- PAC 2, que inclou:
 - El disseny conceptual, lògic i físic: concepció a nivell conceptual i tècnic que ha de tenir el MMDD per poder donar servei a les necessitats del client, preveient possibles desenvolupaments futurs.
- PAC 3, que inclou:
 - La implementació del projecte: construcció del MMDD en base al disseny previ previst.
 - El document annex amb les pantalles del resultat de les consultes: document on es mostren les il·lustracions corresponents als resultats dels diferents informes creats per l'empresa client.
- Entrega final, que inclou:
 - La memòria del projecte: redactat final de la documentació del projecte amb tot el treball d'investigació, disseny i creació del MMDD desenvolupat.
 - La presentació multimèdia: síntesi del treball realitzat explicat davant d'una possible audiència.
 - El producte final, que inclou:
 - Excel amb el tractament previ de les dades
 - BBDD Access amb el tractament previ de les dades
 - Informes demanats pel client
 - Informes "extres" no demanats pel client

1.6 DESCRIPCIÓ BREU DELS ALTRES CAPÍTOLS DE LA MEMÒRIA

En els capítols posteriors es detallaran els processos d'anàlisi i disseny del projecte.

En l'apartat d'anàlisi es detallarà el tractament donat a les dades d'origen, particularment sobre les decisions preses i les homogeneïtzacions portades a terme, es presentaran els diagrames de casos d'ús del projecte i el diagrama del model conceptual.

En l'apartat de disseny es veurà la solució adoptada, amb el seu disseny conceptual, lògic i físic, que ens permetrà crear el MMDD de l'empresa. Complementàriament es proporcionaran els diagrames de l'arquitectura de software, hardware i del model físic que ens permetran entendre visualment l'estructura del DW adoptat. Per últim, es mostrarà com es dissenyen els informes d'Oracle Discoverer i es relacionaran els informes pel client amb una breu descripció.

Finalment, es mostren les conclusions finals del projecte i les línies d'evolució futura que poden derivar-se de la consecució d'aquest.

2. ANÀLISI

En aquesta fase es definiran i analitzaran les dades d'origen provinents de fonts operacionals per poder-les homogeneïtzar i definir el disseny necessari per la construcció del MMDD.

2.1 DADES D'ORIGEN

2.1.1 Fitxer de text del client

Les dades d'inici provenen d'un únic fitxer de text, proporcionat per l'empresa Sostre per a Tothom, en codificació UTF-8, que en el seu origen era informació recollida en diversos fulls de càlcul Excel que han fusionat. Els camps estan delimitats per tabuladors i conté 93.361 línies d'informació.

En aquest fitxer de text, la primera línia conté els noms dels camps:

- Codi i zona : codificació de la territorialitat
- Tipus : tipologia de l'immoble
- Planta : planta de l'immoble (en el cas de ser pisos)
- Dormitoris : nombre de dormitoris
- Metres : metres quadrats de l'immoble
- "Preu mes any" : llista de camps que contenen el preu de l'immoble per al mes i anys detallats.

Hi ha 805 línies que identifiquen la localització dels immobles, mentre que les restants 92.555 línies contenen la informació de les propietats.

Definirem que els camps "codi", "zona", "tipus" i "planta" emmagatzemin valors de tipus text. En canvi els camps "dormitoris", "metres" i tots els "preu mes any" contindran dades de tipus numèric.

Si el camp "Preu mes any" està en blanc significa que l'immoble no està a la venda en aquell període (ha estat venut en l'últim camp Preu mes any informat o bé perquè en aquella data no estava encara en catàleg).

En el fitxer consten unes capçaleres que agrupen el conjunt de línies dels immobles d'una zona i la descriuen textualment. La zona es compon d'una jerarquia de 4 grups territorials que s'agrupen d'esquerre a dreta. En el magatzem de dades es vol poder agregar i desglossar la informació segons aquesta jerarquia territorial de 4 nivells.

2.1.2 Anàlisi previ i problemes detectats

Per poder realitzar l'anàlisi de les dades de forma amigable, s'ha efectuat una importació de les mateixes als programis Excel i Access de Microsoft. En Excel es faran les anàlisis i canvis més evidents, mentre que l'Access es farà servir per comprovacions de dades per mitjà de consultes i filtres de BBDD i transformacions més complexes, com veurem més endavant.

Per facilitar encara més el treball, s'han separat en dos fulls de càlcul diferents els registres que començaven per "#", que corresponen a codificacions de zones, i la resta, que conté el catàleg d'immobles que gestiona el client.

La descripció dels camps que componen el catàleg és la següent:

Codi: identifica el codi de l'immoble i en principi consta de 15 posicions. Com a aspectes a controlar destaquem la comprovació de que no hi hagin codis que no tinguin 15 posicions i la possibilitat que hi hagin codis repetits. No es permeten valors nuls.

Tipus: especifica el tipus d'habitatge. Té com valors: áticos, chalets, dúplex, estudios i pisos. No poden existir valors nuls.

Planta: s'observa algun valor que surt dels paràmetres normals (planta 79 per exemple), i algun sense planta especificada i que no és xalet. No es preveu fer un tractament especial d'aquestes dades perquè no ha d'afectar especialment els informes requerits. És a dir, quan la immobiliària en trobi un error d'aquest tipus es rectificarà manualment, però no en la càrrega de dades al magatzem (per exemple: si consta una planta en un immoble tipus xalet no podem determinar si està malament el tipus o bé la planta que està informada). S'accepten valors nuls.

Dormitoris: es detecten valors extrems com pisos i xalets amb 99 i 64 dormitoris i pisos sense cap. S'hauran de llistar els casos dubtosos i que la immobiliària els estudiï per si cal una rectificació. S'accepten valors a zero, però no nuls.

Metres: correspon a la superfície de l'habitatge. Trobem casos extrems com pisos per sobre dels 3.000 metres quadrats i un xalet amb més de 47.000. També xalets per sota de 50 m² i pisos amb 1 m². Farem un llistat d'aquests casos pel seu estudi detallat. No s'accepten valors nuls.

Preu mes any: els 25 camps existents s'hauran de guardar en un únic camp "preu" amb una visió temporal de mes i any en el magatzem de dades. Pot contenir valors a zero, però no nul

2.1.3 Decisions preses

Amb respecte a les dades, s'han pres les decisions següents:

- S'han unificat els literals de nivell 4 i 5 de localització, concatenant-los amb un guió de separació. Aquesta modificació s'ha realitzat en Excel, concatenant les columnes corresponents als nivells 4 i 5.
- S'ha unificat el codi numèric de les localitzacions en un sol (normalment el més utilitzat). Un cop unificat s'han comprovat les possibles duplicitats i s'han gestionat. Els canvis s'ha fet en Excel, però les consultes per determinar quants registres s'havien de canviar s'han realitzat en Access.
- S'han unificat algunes poblacions que sortien en 2 comarques diferents, optant en aquest cas per traspasar tots els immobles a la comarca actual i eliminant l'antic codi. La detecció d'aquests casos s'ha fet en Access i la correcció en Excel.
- S'han afegit codis nous d'immobles en aquelles localitzacions en que no existien en no haver-hi immobles en cartera, en previsió de futur. La detecció d'aquests casos s'ha fet en Access i la correcció en Excel.
- En aquells immobles amb més de 20 dormitoris, si l'habitatge tenia més de 1.400 m² s'ha canviat el valor a 22, i si en tenia menys de 1.400 m² a 16. Aquestes xifres són de conveniència, ja que no existien i permeten identificar-los posteriorment si el client volgués fer algun tractament especial. La detecció i correcció d'aquests casos s'ha fet en Access.
- Si l'immoble era un pis, xalet o àtic, i el nombre de dormitoris era de 0, s'han canviat els valors a 1. La detecció i correcció d'aquests casos s'ha fet en Access.

- Si la superfície era superior a 3.000 m², si eren pisos s'ha adoptat el valor mig de 125 m² i si eren xalets s'ha dividit el valor actual per 100. La detecció i correcció d'aquests casos s'ha fet en Access.
- Si la superfície era de menys de 15 m², els pisos, xalets i àtics s'han transformat a 100 m² (superfície mitjana, per no distorsionar massa les estadístiques). La detecció i correcció d'aquests casos s'ha fet en Access.

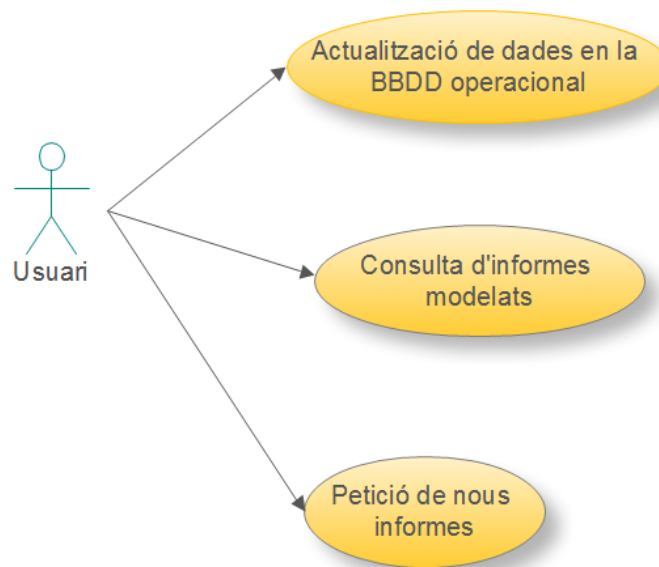
Per evitar pèrdues de les modificacions efectuades entre Excel i Access, primer s'han efectuat tots els canvis pertinents en Excel, després s'ha exportat la informació a Access i tot seguit s'han efectuat els canvis exclusius en Access que hem comentat.

Totes aquestes decisions s'han tractat amb el client i no ha manifestat cap inconvenient.

2.2 DIAGRAMES DE CASOS D'ÚS

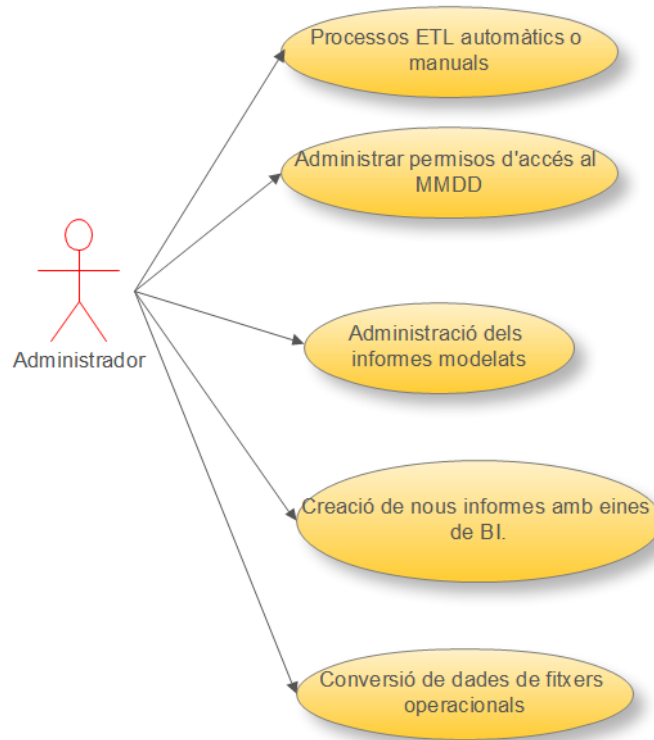
Aquest diagrama definirà els passos que seguirà un actor (usuari o administrador en aquest cas) per a poder fer un procés determinat.

Començarem pel rol d'usuari, però amb una petita llicència, ja que s'ha demostrat que les dades provinents dels fitxers Excel generen molts errors, incongruències i duplicitats. És per això que s'inclou en el diagrama que els usuaris actualitzaran les dades per mitjà d'una BBDD operacional dissenyada ad-hoc per aquest procés. El cost no és elevat i garanteix una molt bona consistència de les dades, aprofitant una estructura de taules semblant a la dissenyada pel MMDD. D'aquesta forma, el procés de d'actualització periòdica del DW es simplifica enormement, sense que "a priori" puguin existir errors i situacions no previstes com les que es podrien generar-se en càrregues de fitxers plans a través d'automatismes creats per SQL*Loader.



II-lustració 3 - Diagrama d'ús dels usuaris

El rol d'administrador haurà de mantenir el MMDD construït, realitzant o supervisant processos de càrrega provinents de les BBDD operacionals, haurà de construir nous informes a petició dels usuaris i en els determinats casos haurà de convertir informació provinent de fonts operacionals a dades llestes per la càrrega al repositori. Veiem el seu diagrama:



Il·lustració 4 - Diagrama d'ús dels administradors

2.3 DIAGRAMA DEL MODEL CONCEPTUAL

Implica expressar en un esquema UML en forma d'estrella tots aquells cubs de dades que ens interessin, triant el fet rellevant, trobant la granularitat adequada al nostre disseny i escollint les dimensions que es faran servir en l'anàlisi de les dades amb:

- els seus atributs
- els atributs que serviran per agrupar (defineixen les jerarquies d'agregació) i els que serviran per seleccionar (anomenats descriptors)
- les mesures que interessin per l'anàlisi de les dades demanat
- l'explicitació de cel·les de mesures que considerem interessants pels analistes
- les restriccions d'integritat del disseny (les dades no inseribles i operacions no permeses)
- l'estudi de la viabilitat (si el volum de dades a generar és adequat pel nostre sistema, si la velocitat de resposta és adequada, si es poden fer les actualitzacions amb la periodicitat i dades necessàries, ...)

2.3.1 Fet rellevant

Un dels primers conceptes a definir en la creació d'un magatzem de dades és la identificació d'aquell element que constitueix el nucli al voltant del qual orbitaran tots els altres elements.

Podem observar que en el projecte que ens ocupa, el centre d'atenció de les dades i consultes demanades són el preu dels immobles. No només es tracta de saber si un immoble està venut i a quin preu, sinó que també ens interessen aquells no venuts, l'evolució dels seus preus i extreure les causes d'aquest fet. En definitiva, el que analitzarem és el valor de l'immoble i si aquest valor fa que l'immoble que gestiona l'empresa es vengui o no.

Aquest fet valor d'un immoble l'analitzarem des de diverses perspectives, ja sigui temporalment (el seu preu evoluciona i es ven o no), per les característiques constructives de la propietat (intentarem esbrinar si el preu és determinant, si ho són les seves característiques més o si és una combinació d'ambdós), per zona i per tipus.

2.3.2 Granularitat

Un bon disseny es caracteritza per l'adequació entre el nivell de precisió que ens proporcionen les dades i el volum de dades del sistema. No és bo un volum excessiu de dades (tant a l'inici com de creixement temporal), ja que provoca malbaratar espai de disc i càrregues excessives al sistema i, conseqüentment, un alentiment de les consultes i processos.

En el cas que del nostre magatzem, el volum de dades inicial no resulta molt elevat i, tot i que es preveu un increment de dades mensuals en actualitzar els preus i incorporar nous immobles, això tampoc significa una càrrega excessiva i és assumible. Per tant, s'opta per la solució de carregar en el sistema la totalitat de les dades proporcionades pel nostre client, previ tractament previ de les mateixes per aconseguir elements d'informació consistents. D'aquesta forma queda garantit l'equilibri entre precisió i volum de les dades.

No es preveuen incorporacions de dades excessives, ja que haurien de venir donades per un augment considerable del tamany de l'empresa client via creixement o compra, perquè el seu dimensionament actual no permet gestionar una cartera d'immobles molt més elevada sense repercussió en la qualitat del servei i/o èxit de vendes. Ara bé, podem preveure que gràcies a aquest projecte i el seu avantatge competitiu, es produirà un increment notable de clients en l'agència immobiliària que voldran que l'empresa gestioni les seves propietats a la venda. Si l'augment de dades a gestionar fos molt superior a l'increment que raonablement hem suposat, s'hauria de revisar el disseny i la seva granularitat, ja que l'entrada d'aquest volum de dades tan gran podria fer baixar el rendiment del sistema en la seva explotació.

2.3.3 Dimensions

A partir de l'anàlisi del fet **Valor**, observem que hi ha dades que contextualitzen el fet i que, per tant, constitueixen una perspectiva d'anàlisi del mateix. Aquestes perspectives diferents poden constituir-se en dimensions :

- Dimensió **Temps** : mes i anys dels diferents preus d'un immoble
- Dimensió **Localització** : zona on està situada la propietat
- Dimensió **Propietat** : característiques físiques de l'immoble

2.3.4 Atributs de cada dimensió

A cada dimensió li corresponen una sèrie d'atributs rellevants pel magatzem de dades, que seran aquells que creiem que representen la informació mínima que s'ha de guardar en aquesta dimensió.

Els atributs de la dimensió "Temps" són el més fàcils i evidents de definir ja que són Mes i Anyo (per no confondre amb la paraula reservada "any" de PL/SQL d'Oracle).

De la dimensió "Localització" hem decidit que siguin els atributs Província, Comarca, Ciutat-Àrea i Zona.

Finalment, per la dimensió "Propietat" tindrem els següents atributs : Tipologia, Planta, Dormitoris i Superfície.

2.3.5 Descriptors i jerarquies d'agregació

En el nostre model, establim les jerarquies d'agregació especificades en la taula següent:

Dimensió	Agregació per:	Comentari
Temps	Anyo	Diferents mesos poden agrupar-se per l'any del fet
Localització	Ciutat-Àrea	Les zones d'una ciutat o àrea es poden agrupar segons una mateixa ciutat o àrea.
	Comarca	Les ciutats o àrees es poden agrupar per les comarques que els corresponen
	Província	Les comarques es poden agrupar per províncies

Els altres atributs de les dimensions els considerarem descriptors i ens serviran únicament per seleccionar dades.

2.3.6 Mesures a fer servir

En aquest projecte la mesura més evident del fet Valor és el "preu", que anirà variant amb el temps. Ens permetrà saber preus mitjos, màxims i mínims sota diferents perspectives (per zones, per tipologies, per característiques, dels immobles més venuts, etc).

Una altra mesura que necessitem en el fet Valor ens ha de servir per calcular si un immoble ha baixat de preu o no. En principi definim que aquesta baixada de preu es produeix amb respecte al seu preu de partida en el moment que es va donar d'alta l'immoble a l'agència. Per tant, incorporarem una mesura que anomenem "preuInicial".

Una tercera mesura que hem d'incorporar per facilitar les consultes posteriors és si l'immoble ha estat venut o no. Si l'immoble tenia un preu en mesos anteriors i un determinat mes ja no en té (ni en els posteriors), interpretem que la propietat s'ha venut i el marcarem com a tal. En aquest cas la mesura serà "venut".

Per últim, haurem d'enregistrar en cada immoble el temps que porta a la venda o que va trigar en vendre's. Per tant, necessitem una mesura que anomenarem "tempsVenda".

Gràcies a aquestes mesures podrem establir diferents agrupacions i a partir d'aquest punt inclús calcular percentatges.

2.3.7 Cel·les

En aquest projecte només hi ha una cel·la a emmagatzemar que mesura un mateix esdeveniment dins l'estructura definida, la que representa el valor atòmic de l'immoble amb les mesures que hem detallat en l'apartat anterior.

2.3.8 Restriccions d'integritat

Es tracta de definir les bases i restriccions del magatzem de dades. La base ens indicarà les dimensions que realment són necessàries per definir de forma unívoca cada cel·la. En el nostre cas tindrem que la taula localització sempre serà fixe per a cadascun dels immobles, és a dir, que quedarà definida per la dimensió Propietat. Per això, només necessitarem la dimensió "Temps" i la dimensió "Propietat" per determinar la cel·la individual.

A aquestes bases que hem configurat són a les que aplicarem les regles d'integritat comunes (en el roll-up i en la transitivitat en l'agregació)

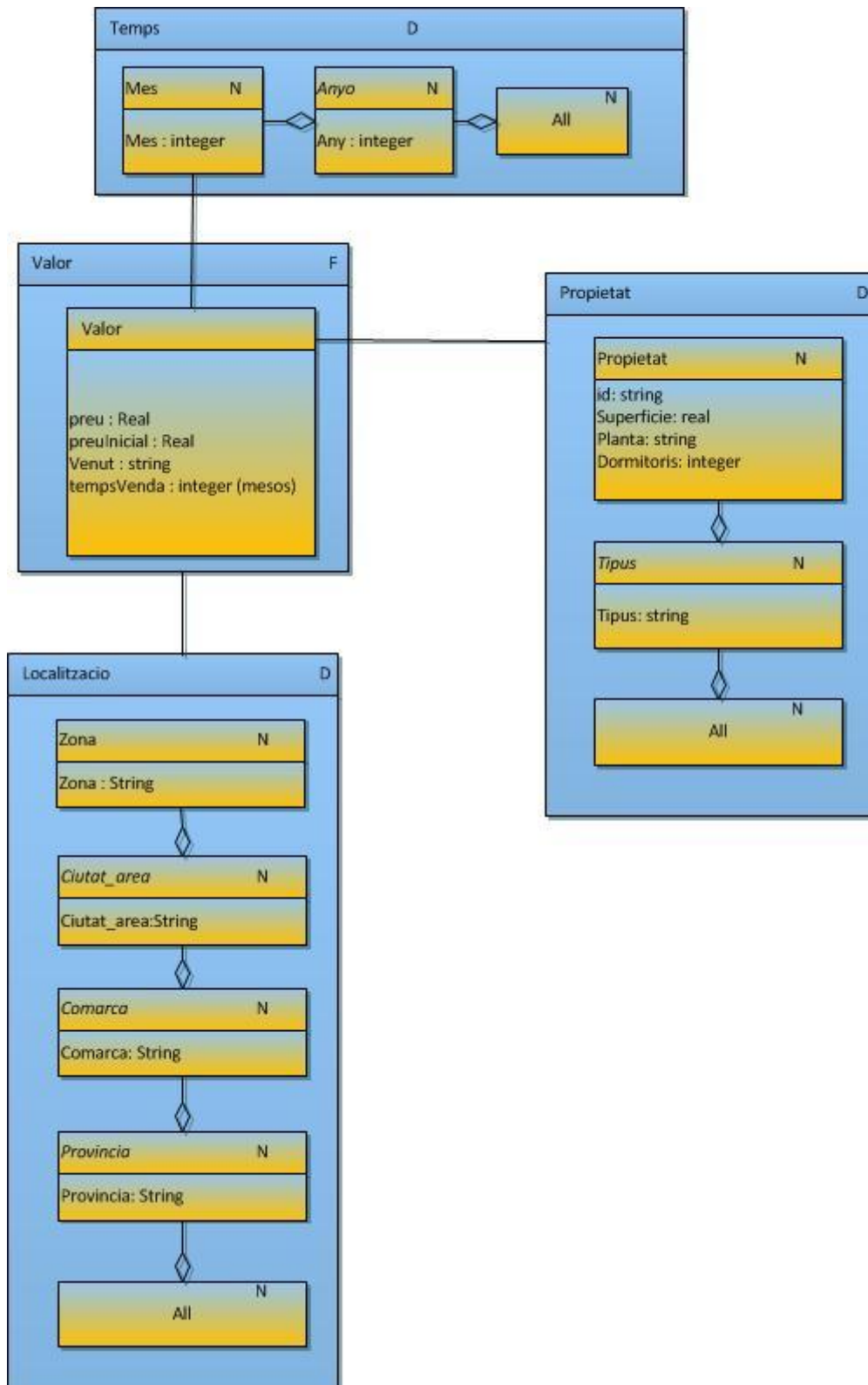
.

2.3.9 Viabilitat

Es tracta de valorar si l'Estrella definida és viable o no, calculant l'espai necessari en el sistema pel seu emmagatzemament.

- Volum de dades : són unes 93.000 línies d'informació aproximadament de les propietats les que incorporarem al magatzem de dades. Considerem que les dades que ocupen les dimensions és un espai negligible amb respecte a les que ocuparà el Fet. Pel Fet tenim , doncs, 93.000 registres amb 25 mesos d'informació històrica, és a dir, 2.325.000 línies d'informació. La cel·la del Fet Valor té 4 atributs que juntament amb els índexs de les dimensions ocupen un màxim de 30 bytes en total. Per tant, tindrem que $2.325.000 \times 30 = 68$ Mb d'ocupació del sistema en dades. Hem de preveure un increment anual de dades per l'entrada de nous immobles i l'evolució de tots en 12 mesos més cada any. Els immobles que s'han venut ja no tenen moviments d'actualització mensual i es mantenen estables en espai ocupat. Inclús uns increments anuals de 68 Mb no haurien de representar un problema seriós pels servidors i computadores actuals, sobretot tenint en compte que el preu per Mb d'emmagatzematge és cada cop més reduït.
- Rendiment : Un volum tan petit de dades no representa un problema de rendiment al sistema en l'explotació de les dades del magatzem que construïm, si tenim en compte les capacitats i rapidesa dels processadors actuals.

Segons aquests passos, el **model conceptual** resultant és el següent:



Il·lustració 5 - Model conceptual del MMDD

3 DISSENY

En aquesta fase es fa l'especificació del disseny lògic, o l'adaptació del model conceptual anteriorment definit a la implementació multidimensional que executarem sobre Oracle.

A continuació es realitza el disseny físic del magatzem a partir del disseny lògic definit i adaptant-lo a les característiques pròpies del SGBD adoptat, que és l'Oracle.

Els objectius prioritaris d'aquesta fase consisteixen en de crear el conjunt de taules relacionals del disseny i les seves restriccions, el d'identificar les estructures d'emmagatzemament específiques i mètodes d'accés per optimitzar el sistema i el de definir les mesures de seguretat del sistema.

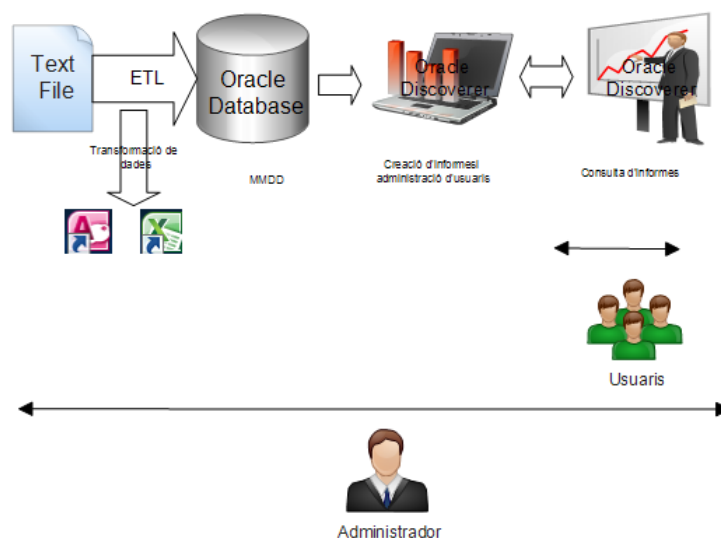
Però abans de tot dibuixarem les arquitectures de software i hardware que es faran servir en aquest projecte.

3.1 ARQUITECTURA DEL SOFTWARE

En quant a software tindrem els següents elements:

- Un fitxer de text pla : és el fitxer font de dades. L'administrador farà els canvis necessaris en el fitxer i generar-ne un de nou en format "csv" per la posterior càrrega de dades.
- Oracle Database 10g XE : Sistema de Gestió de BBDD que contindrà el repositori de la informació.
- Oracle Business Intelligence Discoverer : Eina OLAP de BI que permetrà generar informes i consultar-los sobre les dades del DW d'una forma senzilla.

D'aquesta forma obtenim el següent diagrama d'arquitectura de software:



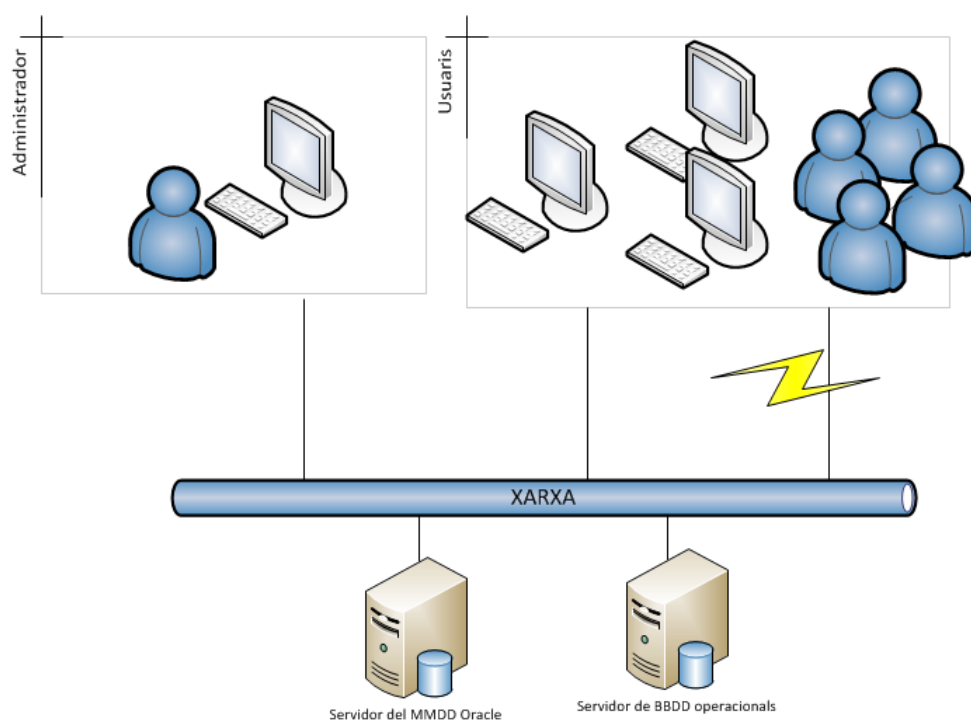
Il·lustració 6 - Diagrama d'arquitectura del software

3.2 ARQUITECTURA DEL HARDWARE

S'identifiquen els següents elements:

- Equips dels administradors del sistema: equips que realitzaran les tasques de manteniment, actualització i disseny de nous informes del MMDD.
- Equips dels usuaris del sistema: equips que es connectaran al DW mitjançant l'Oracle Discoverer per consultar els informes.
- Servidors de les BBDD operacionals: on cada àrea de negoci té la seva BBDD operacional particular
- Servidor Oracle : on es troba el MMDD

Obtenim el següent diagrama d'arquitectura de Hardware:



Il·lustració 7 - Diagrama d'arquitectura del hardware del MMDD

3.3 DISSENY LÒGIC

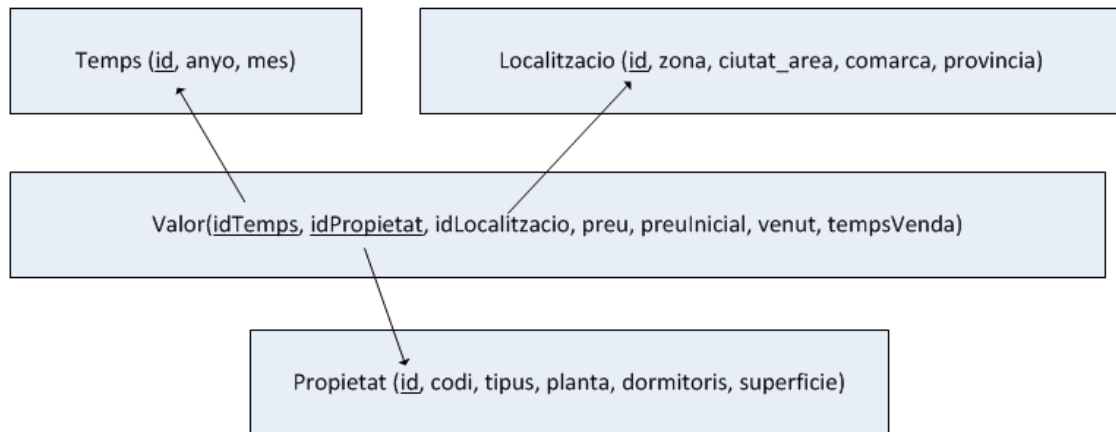
A partir del disseny conceptual i de l'estructura de dades en Estrella que hem vist, farem el disseny lògic del projecte, connexionant les diferents taules que componen el disseny: una pel fet i una per cada una de les dimensions.

- Les connexions les realitzarem per mitjà de les claus primàries de les dimensions que les afegirem com claus foranes a la taula del fet.
- Resulta el següent model lògic:
- Temps (id, anyO, mes)
- Propietat (id, codi, tipus, planta, dormitoris, superfície)

- Localitzacio (id, zona, ciutat_area, comarca, provincia)
- Valor (idTemps, idPropietat, idLocalitzacio, preu, preulnicial, venut, tempsVenda)

Es pot observar que idLocalitzacio no és clau primària pel que hem comentat que no és necessària per determinar una cel·la concreta ja que una propietat sempre està associada a la mateixa zona.

Obtindrem, doncs, el següent diagrama del disseny lògic:



Il·lustració 8 - Disseny lògic del MMDD

3.4 DISSENY FÍSIC

A partir de les relacions definides anteriorment i tenint en consideració les consultes a la BBDD que hem de fer, es realitza el disseny físic amb l'objectiu d'optimitzar els temps de respostes de les diferents consultes demanades pel client.

Es defineixen les diferents taules necessàries que gestionarà el motor de la BBDD.

3.4.1 Taula TEMPS

Amb els tres camps relacionats més avall, tenim els mesos i anys en els quals la propietat ha tingut assignat o no un valor.

ID → Size: 6

MES → Size: 2

ANY → Size: 4

3.4.2 Taula LOCALITZACIO

En aquest cas proposem 5 camps, un de zona de la ciutat o àrea, un de ciutat o àrea on està l'immoble, un de la comarca corresponent, un de la província (interpretarem com a província Andorra, tot i que el tractarem com a país quan el demanin), i un d'identificador de la localització (un codi, en principi numèric).

ID → Size: 20

PROVINCIA → Size: 40

COMARCA → Size: 40

CIUTAT_AREA → Size: 60

ZONA → Size: 60

3.4.3 Taula PROPIETAT

És una taula que conté 6 camps amb les diferents característiques físiques de l'immoble, que ens permetrà establir diferents classificacions.

ID → Size: 15

TIPUS → Size: 20

PLANTA → Size: 10

DORMITORIS → Size: 4

SUPERFICIE → Size: 6, Precision: 2

3.4.4 Taula VALOR

És una taula que conté 7 camps 4 identificadors amb les mesures que hem vist a l'apartat 2.3.6 i els 3 identificadors que ens permetran enllaçar dades amb les 3 dimensions que hem previst.

IDTEMPS → Size: 6

IDPROPIETAT → Size: 15

IDLOCALITZACIO → Size: 20

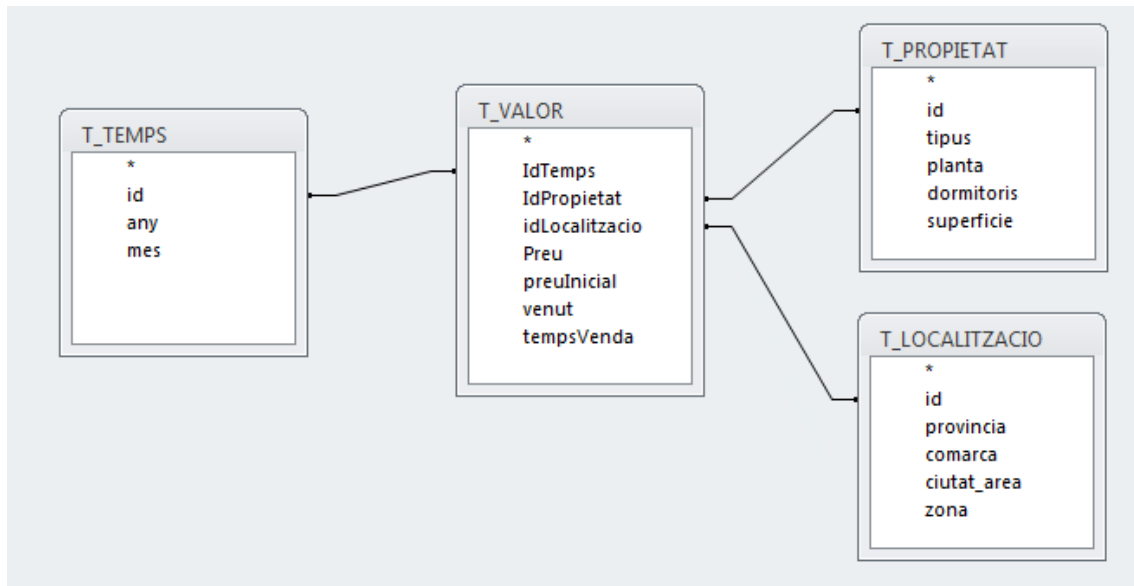
PREU → Size: 12; Precision: 2

PREUINICIAL → Size: 12; Precision: 2

VENUT → Size: 1

TEMPSVENDA → Size: 4

El diagrama del disseny físic obtingut és el següent:



Propietario	Tabla	Columna	Tipo	Nulo
SPAT	LOCALITZACIO	1 ID	Varchar2(20)	
		2 PROVINCIA	Varchar2(40)	
		3 COMARCA	Varchar2(40)	✓
		4 CIUTAT_AREA	Varchar2(60)	✓
		5 ZONA	Varchar2(60)	✓
PROPIETAT	PROPIETAT	1 ID	Varchar2(15)	
		2 TIPUS	Varchar2(20)	
		3 PLANTA	Varchar2(10)	✓
		4 DORMITORIS	Number(4,0)	✓
		5 SUPERFICIE	Number(6,2)	✓
TEMPS	TEMPS	1 ID	Varchar2(6)	
		2 MES	Number(2,0)	
		3 ANYO	Number(4,0)	
VALOR	VALOR	1 IDTEMPS	Varchar2(6)	✓
		2 IDPROPIETAT	Varchar2(15)	✓
		3 IDLOCALITZACIO	Varchar2(20)	✓
		4 PREU	Number(12,2)	✓
		5 PREUINICIAL	Number(12,2)	✓
		6 VENUT	Varchar2(1)	✓
		7 TEMPSVENDA	Number(4,0)	✓
fila(s) 1 - 20 de 20				

II-lustració 9 - Diagrames del disseny físic del MMDD

3.5 TRANSFORMACIÓ DE LES DADES

Donat que les dades proporcionades pel client eren de difícil càrrega immediata, s'ha hagut de fer una verificació prèvia de les mateixes i, com ja s'ha comentat anteriorment, es van detectar errors evidents que feien pensar en altres que poguessin estar ocults i que podrien corrompre el bon funcionament del Magatzem de Dades.

3.5.1 Opcions i decisions prèvies

Una primera opció era la de migrar les dades tal i com estaven en origen i després tractar-les des del sistema Oracle. Es va descartar perquè s'intuïen grans dificultats en la càrrega (el fitxer de control del SQL*Loader hauria de contemplar moltes excepcions de tractament i els errors produïts serien molt probables) i després en el tractament en SQL que s'hauria de fer posteriorment i el temps a emprar per corregir les dades.

La segona opció consistia en un tractament previ limitat, una càrrega menys complicada que l'anterior, però tot i així gens senzilla, i un posterior tractament de les dades en Oracle. Aquesta opció representava la creació de taules temporals que després hauríem d'eliminar del sistema i la probabilitat d'errors encara continuava existint.

La tercera opció, la que s'ha aplicat, ha consistit en el tractament total de les dades prèviament a la càrrega i la generació de fitxers ja preparats per omplir amb dades les taules definitives del SGBD. Pel tractament previ de les dades s'ha fet servir el programari de Microsoft Excel i Microsoft Access, més que res per comoditat, ja que són eines que faig servir amb regularitat i certa destresa. Aquesta opció podria arribar a ser inviable en els casos de gran volums de dades, en els que, probablement, s'hagués optat per la segona opció. A continuació descriuré els passos més importants que es van fer en aquests programes:

- Es va carregar el fitxer pla a Excel i, a continuació, es van filtrar els registres que començaven per “#” per separar la localització de les dades del catàleg.
- Sobre el full d'Excel de localitzacions es va fer una neteja de codis duplicats i es van unificar tots els codis que referenciaven a la mateixa població. El criteri per l'adopció del codi en cas de dubte va ser el que era més utilitzat. Es va unificar els nivells 4 i 5 de localització, concatenant les columnes corresponents i afegint entre mig un caràcter “-” (en aquells que existia nivell 5).
- Amb l'ajuda de consultes en Access sobre els registres del catàleg, es van determinar aquelles poblacions sobre les que no existia un codi de localització únic. La decisió presa, com s'ha dit abans és la de canviar aquelles codificacions que no eren les més habituals, i sempre sota la premissa que tot i que el registre podia tenir un codi d'una altra localització, l'immoble estava ben classificat sota el literal corresponent. Tots aquests canvis també es van fer sobre Excel.
- Un últim pas va ser crear la taula de temps manualment en Excel, ja que el volum de registres a crear era molt petit.

D'aquesta forma, ens garantíem un mínim de qualitat en les localitzacions dels diferents immobles i eliminàvem les possibles duplicitats en la consulta de localitzacions dels informes previstos.

Aquest primer tractament de dades en Excel es va exportar a la BBDD relacional Access, on, per mitjà de filtres i consultes podríem afinar més les transformacions, augmentant encara més la qualitat de les dades.

- La taula temps ja la teníem treballada en Excel i ara la reservàvem per exportar a un fitxer "csv" des d'Access.
- Del catàleg d'immobles es van triar els camps que composaven la taula Propietat i es va crear una taula diferenciada des d'on es van fer certs canvis a nivell de plantes, metres quadrats i plantes, detallats a l'apartat 2.1.3 d'aquesta memòria. Un cop realitzats aquests canvis, la taula Propietat també estava llesta per la seva exportació.
- La taula localització també la teníem acabada des de Excel, per tant aquesta taula també estava preparada per l'exportació.
- En quant a la taula valor, es van afegir els camps venut, preu, idTemps, el preuinicial i mesosVenda.
- Per determinar si estava venut o no, es va realitzar la consulta de tots aquells que a la data 200801 encara tenien preu. Si tenien preu en aquesta data els immobles no estaven venuts, en cas contrari sí. A partir d'aquesta premissa, es va omplir el camp venut amb 'S' o 'N' segons correspongués.
- Per mitjà d'un petit programa en Visual Basic, es van transposar els camps de data que teníem en columnes en un valor del camp idTemps, i el seu preu en el camp preu. Aprofitem també el programa per crear un bucle que ens guardi el primer valor trobat de preu en el camp preuinicial. Tots els demés valors dels registres es mantenen. Amb aquest procés, el nombre de registres s'ha incrementat tantes vegades com s'ha executat el bucle de lectura dels camps de data.
- Ens falta per determinar el valor de mesosVenda. En aquest cas es fa una consulta que filtra sobre els immobles venuts, quants registres hi ha amb una data diferent. Amb el valor donat, només ens cal actualitzar la taula valor amb aquesta dada.
- D'aquesta forma, la taula valor ja està llesta per a l'exportació.

Amb aquest procés, hem aconseguit replicar exactament les taules amb l'estructura requerida pel MMDD i amb totes les dades ja tractades.

En tots els casos tractats, s'ha verificat que el nombre de registres fossin els que per sentit comú corresponien al procés que tractàvem, amb la fi de mantenir la qualitat de les dades.

3.5.2 Creació del MMDD i càrrega de les dades

Un cop connectats com l'usuari SPAT, podem començar el procés de creació de taules en Oracle, i per això, a través de la utilitat SQL d'Oracle escrivim els següents scripts i els executarem:

```
CREATE TABLE temps (  
  id VARCHAR(6 CHAR) CONSTRAINT PK_temps PRIMARY KEY,  
  mes NUMBER(2) NOT NULL,  
  anyo NUMBER(4) NOT NULL  
)
```



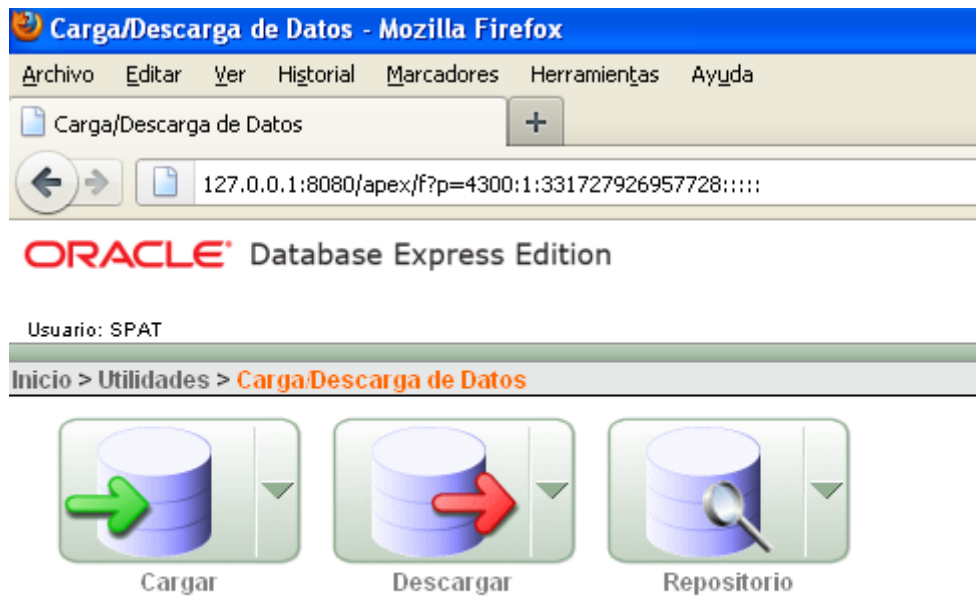
```
CREATE TABLE localitzacio (  
  id VARCHAR(20 CHAR) CONSTRAINT PK_localitzacio PRIMARY KEY,  
  provincia VARCHAR(40 CHAR) NOT NULL,  
  comarca VARCHAR(40 CHAR),  
  ciutat_area VARCHAR(60 CHAR),  
  zona VARCHAR(60 CHAR)  
)  
  
CREATE TABLE propietat (  
  id VARCHAR(15 CHAR) CONSTRAINT PK_propietat PRIMARY KEY,  
  tipus VARCHAR(20 CHAR) NOT NULL,  
  planta VARCHAR(10 CHAR) DEFAULT NULL,  
  dormitoris NUMBER(4) DEFAULT NULL,  
  superficie NUMBER(6,2) DEFAULT NULL  
)  
  
CREATE TABLE valor (  
  idTemps VARCHAR(6 CHAR) CONSTRAINT FK_valor REFERENCES temps(id),  
  idPropietat VARCHAR(15 CHAR) CONSTRAINT FK2_valor REFERENCES propietat(id),  
  idLocalitzacio VARCHAR(20 CHAR) CONSTRAINT FK3_valor REFERENCES localitzacio(id),  
  preu NUMBER(12,2) DEFAULT NULL,  
  preuInicial NUMBER(12,2) DEFAULT NULL,  
  venut VARCHAR(1 CHAR),  
  tempsVenda NUMBER(4) DEFAULT NULL  
)
```

Il·lustració 10 - Consultes SQL de la creació de les taules del MMDD

Per a carregar la informació que tenim en quatre fitxers de text creats des de l'eina Access (T_LOCALITZACIO.txt, T_PROPIETAT.txt, T_TEMPS.txt i T_VALOR.txt) al magatzem de dades, podem triar entre 3 eines que s'utilitzen habitualment:

1. La utilitat SQL*Loader
2. L'aplicació Application Express del propi Oracle, Inicio/Utilidades/Carga Descarga de datos
3. La utilitat data pump import (impdp)

S'ha optat per la segona per la seva senzillesa (menú visual), perquè la càrrega de dades no tenia molts registres i perquè les dades ja les teníem totalment preparades i en els formats requerits per les taules Oracle creades anteriorment.



Il·lustració 11 - Aplicació da càrrega d'Oracle

En el cas que s'haguessin d'haver fet transformacions o filtres en la càrrega, o d'un gran volum de dades a importar, s'hagués optat per la primera utilitat. En el fitxer de control hauríem de posar els següents blocs d'instruccions :

```
LOAD DATA
CHARACTERSET 'WE8MSWIN1252'
BADFILE 'T_TEMPS_BAD.txt'
DISCARDFILE 'T_TEMPS_DISCARD.txt'
INFILE 'T_TEMPS.txt'
FIELDS TERMINATED BY ';'
OPTIONALLY ENCLOSED BY '"'
TRAILING NULLCOLS
(id, any, mes)
```

```
LOAD DATA
CHARACTERSET 'WE8MSWIN1252'
BADFILE 'T_LOCALITZACIO_BAD.txt'
DISCARDFILE 'T_LOCALITZACIO_DISCARD.txt'
INFILE 'T_LOCALITZACIO.txt'
FIELDS TERMINATED BY ';'
OPTIONALLY ENCLOSED BY '"'
TRAILING NULLCOLS
(id, provincia, comarca, ciutat_area, zona)
```

```
LOAD DATA
CHARACTERSET 'WE8MSWIN1252'
BADFILE 'T_PROPIETAT_BAD.txt'
DISCARDFILE 'T_PROPIETAT_DISCARD.txt'
INFILE 'T_PROPIETAT.txt'
FIELDS TERMINATED BY ';'
OPTIONALLY ENCLOSED BY '"'
TRAILING NULLCOLS
(id,tipus,planta,dormitoris,superficie)
```

```
LOAD DATA
CHARACTERSET 'WE8MSWIN1252'
BADFILE 'T_VALOR_BAD.txt'
DISCARDFILE 'T_VALOR_DISCARD.txt'
INFILE 'T_VALOR.txt'
FIELDS TERMINATED BY ';'
OPTIONALLY ENCLOSED BY '"'
TRAILING NULLCOLS
(idTemps,idPropietat,idLocalitzacio,preu,preuInicial,venut,tempsVenda)
```

Il·lustració 12 - Blocs d'instruccions del fitxer de control de SQL*Loader

Per poder executar el SQL*Loader farem servir el següent comanament:

```
sqlldr userid=spat/uoc@XE control=carrega.etl
```

Hem de ser especialment curiosos en les comprovacions de càrregues, per certificar que s'han carregat totes les files i que no hi ha hagut errors. En el cas que n'hi haguessin s'haurien de tractar i fer una càrrega addicional (APPEND) dels registres rebutjats.

3.6 CREACIÓ D'INFORMES A L'ORACLE DISCOVERER DESKTOP

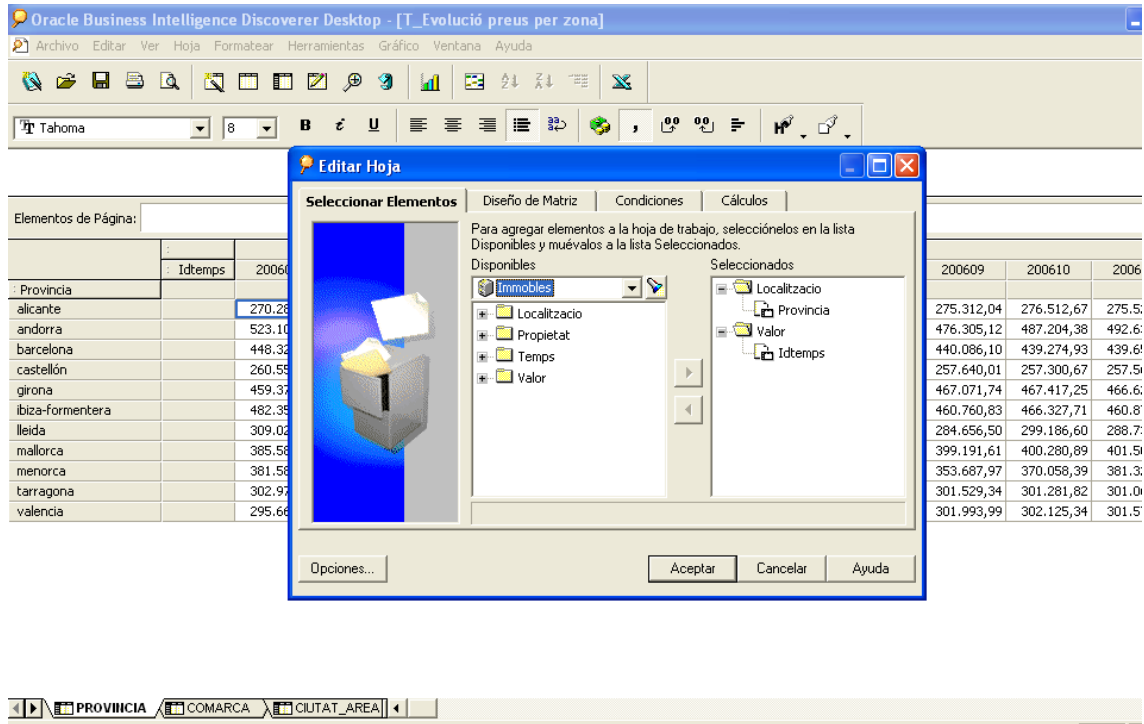
Amb aquesta eina crearem els informes demanats pel client i els posarem a la seva disposició perquè els executin quan creguin convenient.

Primer ens connectarem com a usuari SYSTEM/uoc o SPAT/uoc. Un cop establerta la connexió ja podrem entrar a l'eina de creació d'informes i, per crear-ne un realitzarem els següents passos:

1. Crear llibre de treball
2. Escollir la disposició de sortida de l'informe (com a taula o matriu amb/sense detall)
3. Seleccionem els components de l'àrea de negoci que necessitem per a generar l'informe (full selecció d'elements)
4. Dissenyarem la composició i distribució dels elements del full de treball: files, columnes i valors i/o elements de detall (full disseny matriu/taula)
5. Afegirem els filtres necessaris (full condició)

- Afegirem els càlculs convenients (full càlcul) que seran les dades calculades que obtindrem. Es poden fer servir funcions predeterminades.

De la mateixa forma, podem obrir un full de treball existent i modificar-lo (opció editar full) o executar-lo.



Il·lustració 13 - Pestanyes de modificació del full de treball

3.7 INFORMES GENERATS

A continuació mostrarem algunes captures de pantalles d'informes generats, però abans explicarem l'estructura de directoris que hem creat al disc E:\ per guardar-los.

A la carpeta "Informes amb temporalitat" estan arxivats els informes que ens demanava el client. Cada informe presenta un camp de selecció de mes i any. Un cop feta la tria es mostraran les dades corresponents per a aquell període. S'ha fet així per permetre a l'empresa client una tria sense complicacions i un resultat de pantalla poc carregat. Aquest directori també conté informes d'evolució en què es mostren totes les dades per any i mes, ja que hem considerat que en aquest cas una altra presentació no tenia sentit si el que es volia era veure com variaven les dades.

Hem creat uns altres informes, molt similars als anteriors, que creiem que seran molt útils a l'empresa que donen la visió global de les dades dels immobles sense filtres de temps, per poder veure tota la dimensió de la cartera gestionada. Estan el la carpeta "Informes sense temporalitat".

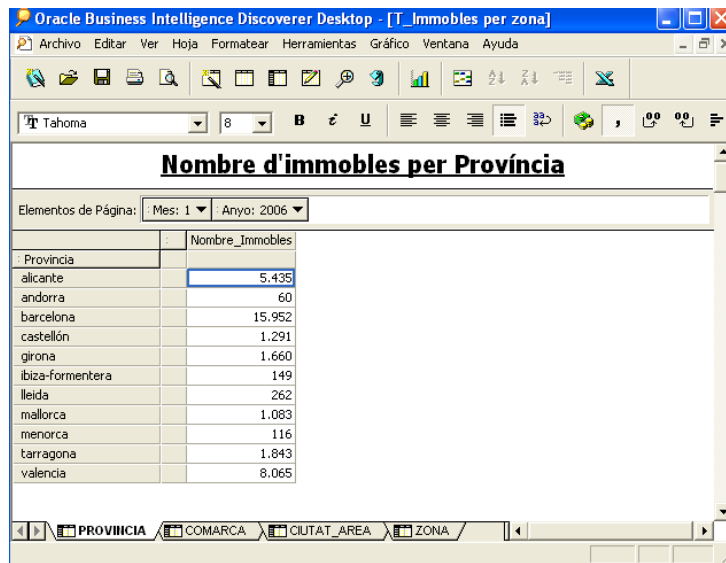
TFC MAGATZEM DE DADES - MEMÒRIA

Altres carpetes en el mateix disc guarden les dades de càrrega i creació de taules, la base de dades Access del tractament previ de les dades, i el fitxer original de dades proporcional pel client.

Finalment, volíem aclarir que per facilitar la tasca als usuaris, alguns informes es guarden en pestanyes diferents si hi ha un agrupament lògic. Simplement s'ha de prémer la pestanya d'interès i executar l'informe.

NOTA : en el cas d'informes agrupats per pestanyes només es mostrarà el primer. Per una consulta més àmplia, existeix el document annex a la PAC 3 on es detallen totes les captures.

- **Nombre d'immobles per zona : càlcul = COUNT**



Oracle Business Intelligence Discoverer Desktop - [T_Immobles per zona]

Archivo Editar Ver Hoja Formatear Herramientas Gráfico Ventana Ayuda

Tahoma 8 B U

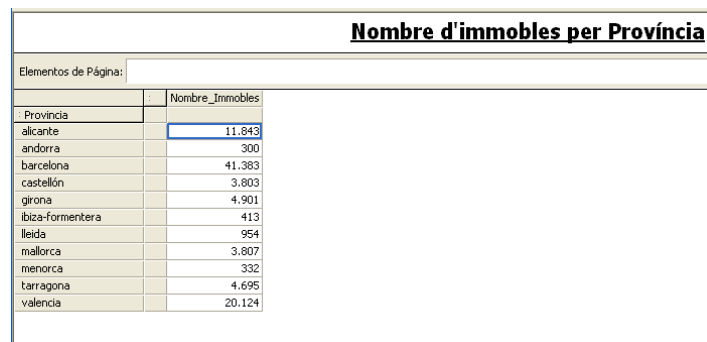
Nombre d'immobles per Província

Elementos de Página: Mes: 1 Año: 2006

Província	Nombre_Immobles
alicante	5.435
andorra	60
barcelona	15.952
castellón	1.291
girona	1.660
ibiza-formentera	149
lleida	262
mallorca	1.083
menorca	116
tarragona	1.843
valencia	8.065

PROVINCIA COMARCA CIUTAT_AREA ZONA

Il·lustració 14 - Nombre d'immobles per província - opció de mes i any



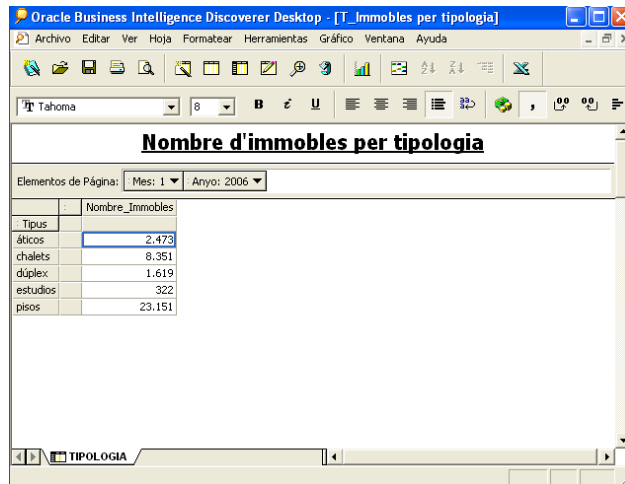
Nombre d'immobles per Província

Elementos de Página:

Província	Nombre_Immobles
alicante	11.843
andorra	300
barcelona	41.383
castellón	3.803
girona	4.901
ibiza-formentera	413
lleida	954
mallorca	3.807
menorca	332
tarragona	4.695
valencia	20.124

Il·lustració 15 - Nombre d'immobles per província - resum de tots els mesos

- **Nombre d'immobles per tipologia : càlcul = COUNT**

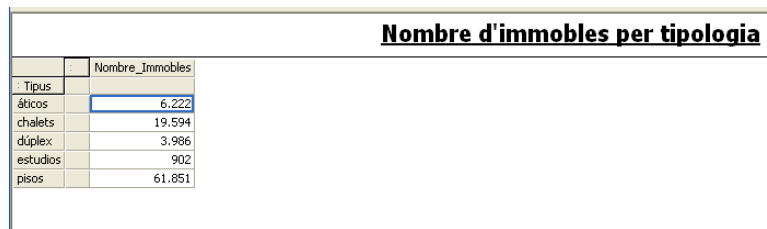


Oracle Business Intelligence Discoverer Desktop - [T_Immobles per tipologia]

Elementos de Página: Mes: 1 Año: 2006

Tipus	Nombre_Inmobles
àtics	2.473
chalets	8.351
dúplex	1.619
estudios	322
pisos	23.151

Il·lustració 16 - Nombre d'immobles per tipologia - opció de mes i any

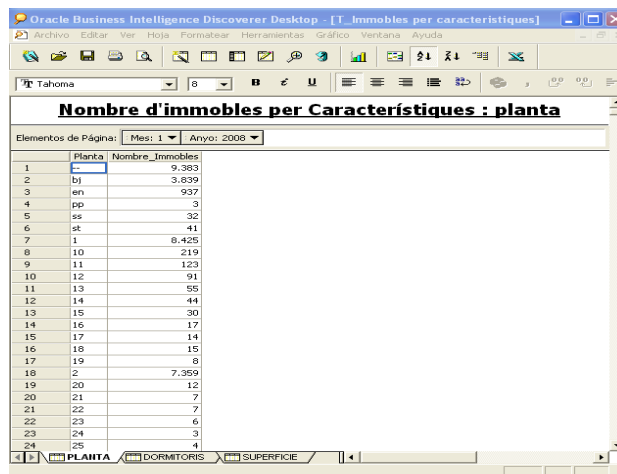


Nombre d'immobles per tipologia

Tipus	Nombre_Inmobles
àtics	6.222
chalets	19.594
dúplex	3.986
estudios	902
pisos	61.851

Il·lustració 17 - Nombre d'immobles per tipologia - resum de tots els mesos

- Nombre d'immobles per característiques: càlcul = COUNT



Oracle Business Intelligence Discoverer Desktop - [T_Immobles per característiques]

Elementos de Página: Mes: 1 Año: 2008

Planta	Nombre_Inmobles
1	9.383
2	3.839
3	937
4	3
5	32
6	41
7	8.425
8	219
9	123
10	91
11	55
12	44
13	30
14	17
15	14
16	15
17	8
18	7.359
19	12
20	7
21	7
22	6
23	3
24	4

Il·lustració 18 - Nombre d'immobles per planta - opció de mes i any

TFC MAGATZEM DE DADES - MEMÒRIA

Nombre d'immobles per Característiques : superfície

	Superfície	Nombre_Inmobles
1	12,00	1
2	14,00	2
3	16,00	5
4	18,00	4
5	19,00	1
6	20,00	24
7	21,00	7
8	22,00	10
9	23,00	7
10	24,00	14
11	25,00	52
12	26,00	28
13	27,00	15
14	28,00	24
15	29,00	14
16	30,00	252
17	31,00	27
18	32,00	52
19	33,00	58
20	34,00	42

II-lustració 19 - Nombre d'immobles per superfície - Resum de tots els mesos

- Llistat d'immobles que han baixat de preu : filtre=preulnicial>preu, càlcul=preulnicial-preu

Llistat d'immobles que han baixat de preu					
	Idpropietat	Idtemp MAX	Preulnicial	Preu MAX	Diferència
1	VC0000000203026	200801	599.000,00	582.000,00	17.000,00
2	VC0000000205233	200707	700.000,00	595.000,00	105.000,00
3	VC0000000205233	200801	700.000,00	550.000,00	150.000,00
4	VC0000000206743	200708	761.000,00	720.000,00	41.000,00
5	VC0000000209364	200801	279.471,00	276.466,00	3.005,00
6	VC0000000209364	200712	279.471,00	276.500,00	2.971,00
7	VC0000000212678	200801	180.000,00	179.000,00	1.000,00
8	VC0000000213635	200801	168.000,00	156.265,00	11.735,00
9	VC0000000215080	200801	192.000,00	186.000,00	6.000,00
10	VC0000000215763	200801	375.000,00	350.000,00	25.000,00
11	VC0000000216927	200801	430.000,00	395.000,00	35.000,00
12	VC0000000216927	200706	430.000,00	415.000,00	15.000,00
13	VC0000000218894	200801	252.425,00	222.374,00	30.051,00
14	VC0000000219222	200702	312.526,00	300.506,00	12.020,00
15	VC0000000219323	200801	258.000,00	240.000,00	18.000,00
16	VC0000000219323	200709	258.000,00	245.000,00	13.000,00
17	VC0000000231303	200801	600.000,00	560.000,00	40.000,00
18	VC0000000234636	200705	510.860,00	498.840,00	12.020,00
19	VC0000000239805	200704	456.770,00	420.000,00	36.770,00
20	VC0000000239829	200801	595.000,00	570.000,00	25.000,00
21	VC0000000255294	200801	360.600,00	306.516,00	54.084,00
22	VC0000000260907	200609	372.630,00	345.582,00	27.048,00
23	VC0000000260913	200801	223.000,00	210.000,00	13.000,00
24	VC0000000272512	200801	367.000,00	360.000,00	7.000,00
25	VC0000000275143	200709	198.000,00	179.500,00	18.500,00
26	VC0000000284778	200801	365.000,00	340.000,00	25.000,00

II-lustració 20 - Llistat d'immobles que han baixat de preu - per mes i any

Llistat d'immobles que han baixat de preu sense duplicats	
	Idpropietat MAX
1	VW0000000814983
2	VW0000000628835
3	VP0000000897406
4	VW0000000480560
5	VW0000000746781
6	VW0000000693037
7	VW0000000688171
8	VW0000000652229
9	VP0000000870293
10	VP0000000848658
11	VW0000000760470
12	VW0000000751182
13	VW0000000770102
14	VW0000000826164
15	VW0000000618534
16	VW0000000855441
17	VW0000000647821
18	VP0000000858517
19	VP0000000899369
20	VW0000000536437
21	VW0000000534965
22	VP0000000834780
23	VP0000000834831
24	VP0000000931719
25	VP0000000762705
26	VP0000000764256

II-lustració 21 - Llistat d'immobles que han baixat de preu - sense repeticions

- Evolució de preus per zona: càlcul=AVG (preu)

TFC MAGATZEM DE DADES - MEMÒRIA

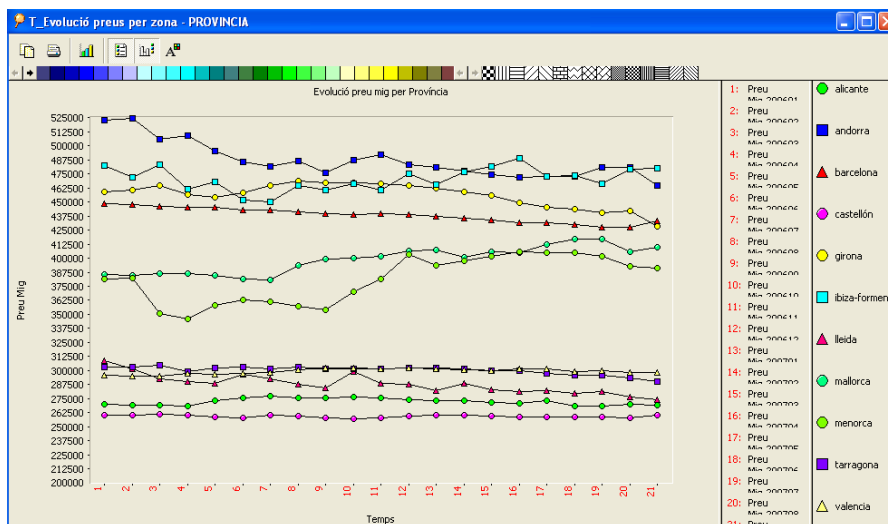
Evolució de preus per Província

Elementos de Página:

	Idtemps	200601	200602	200603	200604	200605	200606	200607	200608
Provincia									
alicante		270.284,23	269.005,23	269.250,62	268.626,19	273.415,75	275.258,52	276.860,83	275.975,27
andorra		523.102,95	524.232,41	505.737,90	508.937,54	495.625,45	485.649,22	481.747,41	486.851,73
barcelona		448.328,79	447.539,84	446.508,16	445.549,75	445.074,98	443.225,56	442.550,32	441.581,63
castellón		260.550,46	260.542,11	260.883,08	259.851,17	258.427,23	257.887,61	260.386,16	259.175,29
girona		459.371,02	460.839,64	464.370,33	456.494,57	454.378,70	458.099,70	464.742,64	469.045,46
ibiza-formentera		482.358,23	472.262,51	483.439,08	461.395,73	468.248,07	451.534,01	449.896,09	464.482,30
lleida		309.027,71	301.467,15	292.314,74	289.972,53	288.253,15	296.240,37	292.292,02	287.889,48
mallorca		385.588,15	385.132,94	386.622,99	386.057,10	384.878,72	381.278,13	380.429,62	393.406,29
menorca		381.586,19	381.956,88	380.465,50	345.631,41	357.896,37	362.899,73	360.889,24	357.460,78
tarragona		302.973,75	302.854,92	304.588,83	299.387,53	302.521,30	303.252,49	301.183,70	303.100,02
valencia		295.668,47	295.066,50	295.203,59	297.775,53	296.870,75	297.545,35	298.078,32	300.310,22

PROVINCIA COMARCA CIUTAT_AREA ZONA

II-lustració 22 - Evolució de preus per província - per mes i any



II-lustració 23 - Gràfica de l'evolució de preus per província - per mesos

- Evolució de preus per tipologia : càlcul=AVG(preu)

Evolució de preus per tipologia

	Idtemps	200601	200602	200603	200604	200605	200606	200607	200608	200609	200610	200611	200612
Tipus													
ático		376.858,05	373.601,64	373.740,10	372.251,40	369.326,22	371.289,05	369.505,99	371.227,70	370.742,31	370.017,47	369.216,56	370.243,76
chalets		557.735,28	556.634,52	556.125,18	555.271,61	555.519,30	554.617,42	558.312,22	562.086,14	563.035,86	563.726,60	566.056,21	566.275,01
dúplex		385.563,27	383.460,33	385.528,96	385.828,07	385.970,79	382.612,43	383.413,42	383.112,94	383.892,65	384.985,77	385.049,99	384.160,53
estudios		157.143,11	154.647,74	154.415,44	154.990,46	156.030,31	155.544,01	156.418,22	158.405,81	158.898,63	159.319,09	160.572,80	160.139,93
pisos		304.273,83	303.957,35	303.375,07	303.912,56	303.519,91	303.344,17	303.440,32	304.313,64	304.028,60	304.361,73	303.785,85	302.969,06

TIPOLOGIA

II-lustració 24 - Evolució de preus per tipologia - per mes i any

TFC MAGATZEM DE DADES - MEMÒRIA

- Evolució de preus per característiques: càlcul=AVG(preu)

Evolució de preus per Característiques : planta											
Elementos de Página:											
	Idtemps	200601	200602	200603	200604	200605	200606	200607	200608	200609	200610
Planta											
--		533.955,41	533.434,15	533.626,85	533.202,77	531.709,18	530.851,75	534.064,97	538.478,29	537.969,44	538.739,26
bj		301.805,68	300.867,43	303.586,51	304.042,34	304.432,50	303.838,47	304.853,82	307.447,58	307.242,71	308.585,55
en		340.256,92	342.825,05	343.688,69	346.981,16	345.562,11	344.816,25	341.325,81	338.187,23	338.273,29	336.401,06
pp		678.304,17	678.304,17	651.068,42	633.384,21	650.238,89	643.627,78	643.627,78	682.620,00	697.200,00	694.271,43
ss		250.029,71	244.821,52	272.817,00	273.130,04	273.421,71	270.204,84	277.966,92	276.451,35	280.256,24	278.252,86
st		249.402,31	246.891,14	246.931,91	249.978,25	256.315,65	250.951,37	250.170,88	252.412,65	255.811,32	256.836,48
1		301.624,48	302.080,29	299.856,11	300.025,76	301.112,04	301.009,23	301.993,86	303.755,84	304.139,80	305.346,28
10		373.604,11	375.707,71	380.645,60	385.973,88	380.743,50	390.433,00	394.550,42	392.214,15	388.098,00	390.839,43
11		429.861,67	433.598,50	432.718,85	431.252,11	419.980,55	423.396,82	422.550,42	421.944,96	426.864,84	424.774,98
12		495.542,69	489.352,35	488.878,24	487.168,58	502.028,13	498.641,33	473.834,77	478.589,49	485.691,65	481.078,73
13		390.050,43	382.286,35	389.656,28	399.113,08	430.060,74	445.818,71	445.271,79	436.178,31	433.387,57	451.937,79
14		291.389,05	288.233,10	283.999,93	296.914,56	289.905,23	294.796,20	287.125,42	285.510,42	287.313,14	282.277,07
15		505.137,57	506.137,57	529.455,85	534.991,20	529.157,50	504.830,89	494.116,61	481.422,93	473.942,17	477.795,18
16		389.763,29	390.787,91	371.320,05	362.787,43	374.323,23	357.804,57	367.233,14	367.233,14	377.394,80	388.873,47
17		495.027,79	467.081,71	391.274,31	391.274,31	368.245,93	372.225,93	383.166,38	389.089,46	404.867,09	404.867,09
18		414.766,11	413.210,56	413.210,56	391.611,88	536.988,33	536.988,33	520.266,50	520.266,50	519.466,50	519.466,50
19		408.179,25	408.179,25	408.179,25	408.179,25	408.179,25	474.095,92	474.095,92	431.179,25	431.179,25	408.857,77
2		298.343,27	297.069,41	295.932,53	297.389,42	297.727,03	297.804,45	297.394,74	298.609,96	297.913,14	297.913,59
20		438.332,00	438.332,00	438.332,00	470.498,80	470.498,80	466.892,80	466.892,80	477.446,36	449.191,00	451.615,45
21		535.785,67	505.345,50	505.345,50	505.145,50	505.145,50	505.145,50	453.713,89	459.154,38	459.154,38	459.154,38
22		588.750,00	588.750,00	588.750,00	588.750,00	657.500,00	759.375,00	814.166,67	814.166,67	674.937,50	674.937,50
23		290.603,00	346.054,75	344.054,75	344.054,75	386.531,67	386.531,67	386.531,67	385.598,57	390.142,86	391.833,33

Il·lustració 25 - Evolució de preus per planta - per mes i any

- Preus màxims, mínims i mitjans per zona: càlcul=Maxim(preu), Mínim (preu) i AVG (preu)

Preus per Zona de la Ciutat							
Elementos de Página: Mes: 1 Año: 2006							
Provincia	Comarca	Ciutat	Area	Zona	Preu màxim	Preu mínim	Preu mitjà
alicante	alt vinalopó		àrea de benejama	centro ciudad	2.625.000,00	75.000,00	473.638,02
				centro ciudad - avenida habaneras	2.625.000,00	75.000,00	473.638,02
				centro ciudad - centro	550.000,00	60.101,00	216.445,04
					550.000,00	60.101,00	216.445,04
					666.000,00	225.000,00	417.000,00
					666.000,00	225.000,00	417.000,00
	bajx segura		àrea de guardamar del segura	centro ciudad	1.200.000,00	70.000,00	184.162,24
				centro ciudad - avenida habaneras	1.200.000,00	70.000,00	184.162,24
				centro ciudad - centro	217.000,00	69.116,00	138.761,80
					217.000,00	69.116,00	138.761,80
					880.000,00	90.000,00	187.795,05
					880.000,00	90.000,00	187.795,05
		àrea de onihuela	centro ciudad	995.000,00	82.500,00	222.192,47	
			centro ciudad - avenida habaneras	995.000,00	82.500,00	222.192,47	
			centro ciudad - centro	840.000,00	80.000,00	145.249,90	
				840.000,00	80.000,00	145.249,90	
				840.000,00	80.000,00	145.249,90	
				840.000,00	80.000,00	145.249,90	

Il·lustració 26 - Preus màxims, mínims i mitjans per zona ciutat - opció de mes i any

Preus per Província			
Elementos de Pàgina:			
Província	Preu màxim	Preu mínim	Preu mitjà
alicante	12.661.095,00	1.000,00	271.134,74
andorra	5.048.501,00	70.000,00	494.094,25
barcelona	9.500.000,00	500,00	436.744,64
castellón	7.362.398,00	800,00	258.846,11
girona	16.000.000,00	2.400,00	448.783,64
ibiza-formentera	4.808.097,00	12.000,00	468.867,28
lleida	5.000.000,00	23.000,00	281.798,89
mallorca	25.000.000,00	55.000,00	403.251,25
menorca	3.790.000,00	110.000,00	384.023,98
tarragona	3.875.000,00	460,00	297.201,95
valencia	4.267.185,00	18.030,00	299.073,24

Il·lustració 27 - Preus màxims, mínims i mitjans per província - resum de tots els mesos

- Preus màxims, mínims i mitjans per tipologia: càlcul=Maxim(preu), Mínim (preu) i AVG (preu)

Preus per tipologia			
Elementos de Pàgina: Mes: 1 Anyo: 2006			
Tipus	Preu màxim	Preu mínim	Preu mitjà
àtics	3.925.000,00	75.000,00	376.858,05
chalets	16.000.000,00	12.000,00	557.735,28
dúplex	3.900.000,00	78.000,00	385.563,27
estudios	1.200.000,00	19.000,00	157.143,11
pisos	3.606.000,00	3.000,00	304.273,83

Il·lustració 28 - Preus màxims, mínims i mitjans per tipologia - opció de mes i any

Preus per tipologia			
Tipus	Preu màxim	Preu mínim	Preu mitjà
àtics	3.925.000,00	750,00	367.072,74
chalets	25.000.000,00	12.000,00	560.645,19
dúplex	4.250.000,00	70.000,00	381.063,59
estudios	1.200.000,00	500,00	157.047,08
pisos	12.661.095,00	460,00	300.474,23

Il·lustració 29 - Preus màxims, mínims i mitjans per tipologia - resum de tots els mesos

- Preus màxims, mínims i mitjans per característiques: càlcul=Maxim(preu), Mínim (preu) i AVG (preu)

Preus per Característiques : planta				
Elementos de Página: Mes: 1 ▾ Anyo: 2006 ▾				
	Planta	Preu màxim	Preu mínim	Preu mitjà
1	--	16.000.000,00	12.000,00	533.955,41
2	bj	3.750.000,00	12.000,00	301.805,68
3	en	1.803.036,00	78.000,00	340.256,92
4	pp	1.795.000,00	285.000,00	678.304,17
5	ss	450.000,00	127.000,00	250.029,71
6	st	1.406.500,00	56.000,00	249.402,31
7	1	1.800.000,00	3.000,00	301.624,48
8	10	1.390.000,00	80.000,00	373.604,11
9	11	1.800.000,00	96.000,00	429.861,67
10	12	2.300.000,00	100.000,00	495.542,69
11	13	1.088.083,00	144.243,00	390.050,43
12	14	950.000,00	75.000,00	291.389,05
13	15	1.950.000,00	97.000,00	505.137,57
14	16	1.180.000,00	109.000,00	389.763,29
15	17	1.680.000,00	134.000,00	495.027,79
16	18	1.700.000,00	168.500,00	414.766,11
17	19	1.100.000,00	116.746,00	408.179,25
18	2	2.910.000,00	15.000,00	298.343,27
19	20	850.000,00	147.250,00	438.332,00
20	21	1.000.000,00	150.000,00	535.785,67
21	22	1.640.000,00	117.500,00	588.750,00
22	23	495.000,00	170.086,00	290.603,00
23	24	811.366,00	189.300,00	423.088,67
24	25	2.500.000,00	227.325,00	754.865,00

PLANTA DORMITORIS SUPERFICIE

Il·lustració 30 - Preus màxims, mínims i mitjans per planta - opció de mes i any

Preus per Característiques : superficie				
	Superficie	Preu màxim	Preu mínim	Preu mitjà
1	12,00	100.000,00	100.000,00	100.000,00
2	14,00	90.000,00	41.640,00	60.227,50
3	16,00	120.250,00	35.900,00	83.819,04
4	18,00	135.000,00	90.000,00	113.688,52
5	19,00	49.000,00	49.000,00	49.000,00
6	20,00	330.000,00	58.000,00	133.772,47
7	21,00	265.046,00	57.000,00	133.428,91
8	22,00	173.212,00	72.000,00	123.149,16
9	23,00	174.000,00	70.787,00	113.264,54
10	24,00	207.000,00	56.000,00	112.555,22
11	25,00	276.465,00	24.600,00	120.921,38
12	26,00	249.340,00	45.000,00	147.281,94
13	27,00	218.000,00	45.000,00	154.109,29
14	28,00	246.400,00	55.000,00	141.647,46

Il·lustració 31 - Preus màxims, mínims i mitjans per superfície - resum de tots els mesos

- Metres quadrats màxims, mínims i mitjans per zona: càlcul=Maxim(superficie), Mínim (superficie) i AVG (superficie)

Superficie per Província				
Elementos de Página:				
Provincia	Superficie màxima	Superficie mínima	Superficie mitjana	
alicante	2.000,00	14,00	121,36	
andorra	2.180,00	25,00	117,27	
barcelona	3.000,00	12,00	113,67	
castellón	1.389,00	26,00	110,78	
girona	2.427,00	16,00	153,94	
ibiza-formentera	1.800,00	28,00	133,37	
lleida	1.000,00	22,00	142,35	
mallorca	2.400,00	20,00	146,09	
menorca	1.000,00	30,00	129,11	
tarragona	2.400,00	20,00	122,55	
valencia	2.300,00	14,00	125,64	

Il·lustració 32 - Superficie màxima, mínima i mitjana per província - resum de tots els mesos

- Metres quadrats màxims, mínims i mitjans per tipologia: càlcul=Maxim(superficie), Mínim (superficie) i AVG (superficie)

Superfície per tipologia

Elementos de Pàgina: Mes: 1 Anyo: 2006

Tipus	Superfície màxima	Superfície mínima	Superfície mitjana
àtics	540,00	26,00	106,58
chalets	2.400,00	20,00	220,59
dúplex	1.120,00	35,00	128,39
estudios	250,00	12,00	42,24
pisos	1.600,00	16,00	89,19

TIPOLOGIA

Il·lustració 33 - Superfície màxima, mínima i mitjana per tipologia - opció de mes i any

Superfície per tipologia

Tipus	Superfície màxima	Superfície mínima	Superfície mitjana
àtics	1.389,00	20,00	106,19
chalets	3.000,00	20,00	220,53
dúplex	1.120,00	25,00	125,58
estudios	285,00	12,00	42,56
pisos	2.300,00	16,00	88,26

Il·lustració 34 - Superfície màxima, mínima i mitjana per tipologia - resum de tots els mesos

- Metres quadrats màxims, mínims i mitjans per característiques: càlcul=Maxim(superficie), Mínim (superficie) i AVG (superficie)

Superfície per Característiques : dormitoris

Elementos de Pàgina: Mes: 1 Anyo: 2006

Dormitoris	Superfície màxima	Superfície mínima	Superfície mitjana	
1	12	2.000,00	250,00	649,68
2	13	250,00	250,00	250,00
3	5	2.300,00	48,00	258,48
4	2	400,00	29,00	78,51
5	20	1.400,00	1.400,00	1.400,00
6	1	1.600,00	16,00	59,29
7	0	250,00	12,00	41,77
8	7	1.600,00	111,00	390,75
9	4	1.900,00	46,00	162,16
10	3	1.200,00	35,00	108,07
11	9	2.400,00	150,00	589,55
12	14	999,00	606,00	751,25
13	6	1.120,00	85,00	325,68
14	15	500,00	340,00	420,00
15	18	1.500,00	1.500,00	1.500,00
16	8	1.200,00	140,00	435,07
17	11	1.800,00	300,00	762,00
18	10	1.953,00	348,00	721,07

PLANTA DORMITORS

Il·lustració 35 - Superfície màxima, mínima i mitjana per dormitoris - opció de mes i any

Superfície per Característiques : planta

	Planta	Superfície màxima	Superfície mínima	Superfície mitjana
1	--	3.000,00	20,00	202,00
2	bj	1.600,00	18,00	93,85
3	en	600,00	12,00	83,91
4	pp	347,00	51,00	131,58
5	ss	285,00	21,00	68,58
6	st	300,00	24,00	70,97
7	1	2.300,00	14,00	89,34
8	10	299,00	25,00	93,62
9	11	385,00	26,00	96,89
10	12	235,00	18,00	97,59
11	13	265,00	30,00	97,96
12	14	330,00	30,00	88,61
13	15	500,00	30,00	113,15

II-lustració 36 - Superfície màxima, mínima i mitjana per planta - resum de tots els mesos

- Temps mitjans de venda per zona: càlcul=COUNT(idPropietat), AVG(mesosVenda)

Temps mitjà de venda per Província		
Elementos de Página: Mes: 1 Anyo: 2006		
Província	Nombre_Inmobles	Temps mijà venda
alicante	2.953	10,55
andorra	10	16,20
barcelona	10.706	10,19
castellón	714	11,49
girona	938	10,91
ibiza-formentera	82	8,99
lleida	141	10,50
mallorca	768	10,41
menorca	66	11,11
tarragona	1.085	10,32
valencia	5.577	9,62

PROVINCIA COMARCA CIUTAT_AREA ZONA

II-lustració 37 - Temps mitjà de venda per província - opció de mes i any

Temps mitjà de venda per Província

Elementos de Página:

Província	Nombre_Inmobles	Temps mitjà venda
alicante	4.961	13,36
andorra	34	13,74
barcelona	20.181	13,07
castellón	1.492	13,58
girona	1.853	13,36
ibiza-formentera	143	11,81
lleida	338	12,37
mallorca	1.798	12,11
menorca	135	13,75
tarragona	2.038	12,89
valencia	10.285	12,54

Il·lustració 38 - Temps mitjà de venda per província - resum de tots els mesos

- Temps mitjans de venda per tipologia: càlcul=COUNT(idPropietat), AVG(mesosVenda)

Temps mitjà de venda per tipologia

Elementos de Página: Mes: 1 Anyo: 2006

Tipus	Nombre_Inmobles	Temps mitjà de venda
àtics	1.565	10,06
chalets	4.790	10,93
dúplex	901	10,63
estudios	200	10,62
pisos	15.584	9,94

Il·lustració 39 - Temps mitjà de venda per tipologia - opció de mes i any

Temps mitjà de venda per tipologia

Tipus	Nombre_Inmobles	Temps mitjà de venda
àtics	2.868	13,08
chalets	8.321	13,68
dúplex	1.628	13,24
estudios	409	12,97
pisos	30.032	12,72

Il·lustració 40 - Temps mitjà de venda per tipologia - resum de tots els mesos

- Temps mitjans de venda per característiques: càlcul=COUNT(idPropietat), AVG(mesosVenda)

Temps mitjà de venda per Característiques : planta

Elementos de Página: Mes: 1 Anyo: 2007

	Planta	Nombre_Inmobles	Temps mitjà de venda
1	--	3.791	15,13
2	bj	1.399	13,72
3	en	405	13,66
4	pp	9	18,22
5	ss	14	12,71
6	st	16	11,56
7	1	3.206	13,80
8	10	74	13,82
9	11	40	16,43
10	12	28	14,89
11	13	13	14,46
12	14	17	14,24
13	15	14	13,50
14	16	7	17,71
15	17	5	13,60
16	18	2	20,50
17	19	5	16,00
18	2	2.902	14,00
19	20	4	15,00
20	21	7	14,14
21	22	2	4,00
22	23	2	20,00
23	24	2	14,00
24	25	1	13,00

PLANTA / DORMITORIS / SUPERFICIE

Il·lustració 41 - Temps mitjà de venda per planta - opció de mes i any

- Nombre d'immobles venuts per zona: càlcul=COUNT(idPropietat)

Nombre d'immobles venuts per Província

Elementos de Página:

	Nombre_Inmobles
Província	
alicante	4.961
andorra	34
barcelona	20.181
castellón	1.492
girona	1.853
ibiza-formentera	143
lleida	338
mallorca	1.798
menorca	135
tarragona	2.038
valencia	10.285

Il·lustració 42 - Nombre d'immobles venuts per província - resum de tots els mesos

- Nombre d'immobles venuts per tipologia: càlcul=COUNT(idPropietat)

Nombre d'immobles venuts per tipologia

Elementos de Página: Mes: 1 Anyo: 2006

	Nombre_Inmobles
Tipus	
áticos	1.565
chalets	4.790
dúplex	901
estudios	200
pisos	15.584

Il·lustració 43 - Nombre d'immobles venuts per tipologia - opció de mes i any

Nombre d'immobles venuts per tipologia	
Tipus	Nombre_Immables
àtics	2.868
chalets	8.321
dúplex	1.628
estudios	409
pisos	30.032

II-lustració 44 - Nombre d'immobles venuts per tipologia - resum de tots els mesos

- **Nombre d'immobles venuts per característiques: càlcul=COUNT(idPropietat)**

Nombre d'immobles venuts per Característiques : planta		
Elementos de Página: Mes: 1 Año: 2006		
Planta	Nombre_Immables	
1	5.049	
2	1.608	
3	477	
4	21	
5	13	
6	21	
7	3.883	
8	87	
9	61	
10	37	
11	21	
12	26	
13	17	
14	17	
15	10	
16	2	
17	6	
18	3.608	
19	5	
20	7	
21	3	
22	6	
23	4	
24	2	
25	2	

II-lustració 45 - Nombre d'immobles venuts per planta - opció de mes i any

3.8 DETERMINACIÓ DEL PIS TIPUS ESPANYOL I ANDORRÀ

Per poder determinar el pis tipus s'hauria de calcular la moda de cada característica i obtindríem el resultat desitjat. En PL/SQL s'hauria de consultar mitjançant STATS.MODE, però en Discoverer és una mica més complex perquè no existeix aquesta funció i no s'ha aconseguit importar-la. Per això, s'ha optat per consultar el nombre d'immobles per les diferents característiques i agafar el valor més gran de tots ells. Així, per Andorra obtenim els següents resultats:

Nombre d'immobles per Característiques : dormitoris		
Elementos de Página: Provincia: andorra		
Dormitoris	Nombre_Immables	
0	14	
1	78	
2	81	
3	87	
4	33	
5	6	
6	1	

Nombre d'immobles existents per Característiques : planta

Elementos de Página: Provincia: andorra ▼		
	Planta	Nombre_Inmobles
1	--	16
2	bj	40
3	en	1
4	st	1
5	1	68
6	12	1
7	16	1
8	2	61
9	3	33
10	35	1
11	4	23
12	5	9
13	6	16
14	7	3
15		26

Nombre d'immobles existents per Característiques : superficie

Elementos de Página: Provincia: andorra ▼		
	Superficie	Nombre_Inmobles
48	87,00	2
49	89,00	1
50	90,00	3
51	91,00	1
52	92,00	4
53	93,00	1
54	95,00	5
55	96,00	1
56	98,00	2
57	99,00	3
58	100,00	15
59	200,00	1

Nombre d'immobles per Tipologia : Andorra

Elementos de Página: Provincia: andorra ▼		
		Nombre Inmobles
Tipus		
àtics		16
chalets		41
dúplex		13
estudios		14
pisos		216

II-lustració 46 - Consultes del pis tipus Andorrà

Per tant, l'immoble tipus d'Andorra seria: un pis, de 100 m² de superfície, amb 3 habitacions i situat a la 1^a planta.

En el cas d'Espanya, no s'ha aconseguit filtrar per Província NOT LIKE 'andorra', per la qual cosa s'ha decidit consultar per totes les províncies i, donat que l'efecte d'Andorra sobre la resta és negligible, podem afirmar que:

Nombre d'immobles per Tipologia : Espanya

Elementos de Página:		
		Nombre Inmobles
Tipus		
àtics		6.222
chalets		19.594
dúplex		3.986
estudios		902
pisos		61.851

TFC MAGATZEM DE DADES - MEMÒRIA

Nombre d'immobles existents per Característiques : planta

Elementos de Página:		
	Planta	Nombre_Inmobles
1	--	17.572
2	bj	7.123
3	en	1.859
4	pp	26
5	ss	62
6	st	74
7	1	16.162
8	10	391
9	11	231
10	12	158
11	13	99
12	14	90
13	15	60
14	16	38
15	17	28
16	18	20
17	19	17
18	2	14.285
19	20	20

Nombre d'immobles per Característiques : dormitorios

Elementos de Página:		
	Dormitorios	Nombre_Inmobles
1	0	827
2	1	6.025
3	2	21.289
4	3	41.214
5	4	17.717
6	5	3.667
7	6	1.091
8	7	361
9	8	169
10	9	124
11	10	29
12	11	5
13	12	19
14	13	1
15	14	7
16	15	3
17	16	2
18	17	1
19	18	1
20	20	1
21	22	2

Nombre d'immobles existents per Característiques : superficie

Elementos de Página:		
	Superficie	Nombre_Inmobles
1	90,00	4.551
2	70,00	4.395
3	80,00	4.169
4	100,00	3.409
5	60,00	3.335
6	75,00	3.206
7	65,00	3.063
8	85,00	2.671
9	110,00	2.503
10	120,00	2.473
11	50,00	1.934
12	95,00	1.714
13	55,00	1.546
14	130,00	1.400
15	150,00	1.386
16	140,00	1.266

Il·lustració 47 - Consultes sobre l'immoble tipus d'Espanya

L'immoble tipus espanyol és un pis, de 90 m², de 3 habitacions i situat en una 1^a planta (com es pot observar, no s'ha tingut en compte els 17 mil registres en què no consta una planta, ja que quasi la totalitat serien xalets, aplicant el sentit comú).

CONCLUSIONS

- S'ha creat un bon disseny del MMDD que ha permès la construcció posterior d'un repositori que acomplirà els requeriments d'informació del client.
- S'ha creat el DW incorporant les dades d'origen amb unes transformacions que assegurin una qualitat bona de les mateixes.
- S'han creat les consultes demanades pel client de forma satisfactòria mitjançant l'eina de BI.
- S'ha aportat la documentació requerida per la realització del projecte.
- S'han assolit els objectius proposats a l'inici del projecte.

LÍNIES D'EVOLUCIÓ FUTURA

Com s'ha pogut constatar, la construcció d'un MMDD és un projecte complex i d'envergadura. Això significa que les empreses podrien ser reticents a invertir la quantitat de diners que es requereixen per la implantació total. Aquesta complexitat fa que, actualment, el projecte marxi en diversos projectes fills més assumibles tant econòmicament com temporalment, ja que els resultats es poden veure reflectits en pocs mesos.

Aquest ha sigut un d'aquests possibles projectes fills, centrat en una unitat operativa determinada i que ha permès a l'empresa Sostre per a Tothom obtenir un avantatge competitiu en aquesta part de negoci gestionada.

Desenvolupant aquest projecte s'han pogut observar certes deficiències que es poden corregir sense fer grans despeses i s'han pogut observar línies d'actuació futura:

- 1) Es necessita una BBDD relacional operativa pel manteniment i actualització de les dades de la unitat de l'empresa sobre les que hem creat el MMDD. D'aquesta forma es garanteix la qualitat de les dades a incorporar al DW, i les càrregues periòdiques serien molt més automàtiques i senzilles. Una bona opció seria crear aquesta BBDD en Oracle, estandarditzant els SGBD de l'empresa.
- 2) Es poden incorporar altres dades per enriquir la informació, com per exemple les dades sobre els venedors que gestionen les propietats i els beneficis de les vendes.
- 3) Es poden incorporar altres àrees de negoci al MMDD que requeriran la realització de nous projectes, com poden ser vendes, comercials, rehabilitacions, lloguers, promotors, etc.
- 4) També sorgiran noves demandes dels usuaris sobre nous informes que s'hauran de crear pels administradors del DW.

GLOSSARI

- **Business Intelligence (BI)** : Estratègies i eines orientades a administrar i crear coneixement per mitjà de l'anàlisi de les dades existents a l'empresa.
- **DataMining** : mineria de dades. Estratègies d'anàlisi de dades per a esbrinar possibles relacions o associacions desconegudes per mitjà d'altres anàlisis tradicionals.
- **Dimensió** : estructura que representa una perspectiva d'anàlisi diferent de les dades d'un repositori. Estableix el context per a realitzar preguntes sobre els fets continguts a la taula de fets.
- **DW (Data Warehouse)**: veure MMDD.
- **ETL (Extract, Transform and Load)**: Es coneix amb aquest nom a tot el procés que s'encarrega de l'extracció, transformació i càrrega de les dades a una BBDD.
- **MMDD (magatzem de dades)**: Sistema d'informació que reuneix en un únic repositori la informació històrica generada per les diferents unitats operatives d'una organització de forma uniforme i sense errors. D'aquesta forma es poden consultar les dades de forma flexible, àgil i fiable.
- **OLAP (processament analític en línia)**: Eines de consulta sobre grans quantitats de dades organitzades en estructures multidimensionals.
- **Oracle DataBase Express Edition**: versió gratuïta del gestor de BBDD relacional creat per Oracle.
- **Oracle Discoverer**: Eina OLAP per a l'anàlisi i consulta de grans volums de dades desenvolupat per la companyia Oracle.
- **PL/SQL** : llenguatge de programació d'Oracle sobre bases de dades, que a més de suportar l'estàndard de SQL afegeix noves característiques com incorporar variables, estructures de control de flux i d'altres.
- **SQL (Structured Query Language)**: Llenguatge de consulta estructurat per crear, modificar i eliminar dades sobre bases de dades relacionals.

BIBLIOGRAFIA

- CONNOLLY Thomas, BEGG Carolyn (2005). *Sistemas de bases de datos. Un enfoque práctico para diseño, implementación y gestión*. Madrid. Pearson Educación.
- GROFF James, WEINBERG Paul (2003). *SQL, manual de referencia*. Madrid. McGraw-Hill.
- GROH Michael i altres (2007). *Access 2007, la Bíblia*. Madrid. Anaya.
- KIMBALL Ralph, ROSS Margy (2002). *The Data Warehouse Toolkit*. USA. John Wiley and Sons.
- RIUS Àngels, SERRA Montse (2003). *Magatzems de dades i models multidimensionals*. Barcelona. Fundació UOC.