

Carlos Cabello Martin

Treball final de carrera
Enginyeria Tècnica Informàtica de Gestió
UOC

Consultor: Bartomeu Antich Luque
01/2016



TFC – MAGATZEM DE DADES

MOBILITAT D'ESTUDIANTS ERASMUS

Continguts

ERASMUS



Anàlisi preliminar

Elaboració del model

Procés ETL

Explotació de la informació

Conclusions

Introducció:

- **Erasmus** (*EuRopean Community Action Scheme for the Mobility of University Students*)



- **Es vol estudiar la mobilitat d'estudiants entre països participants**



- **Fitxers de Microsoft Excel com a font d'informació**

Estat actual:

- Fitxer de dades mestres d'institucions

A	B	C	D	E	F	G
Country	Charter type	Organisation Name	Erasmus code	Street	Postcode	City
FR	EUCX	IPAC	F ANNECY04	42 chemin de la Prairie	74000	Anncy
FR	EUCX	UNIVERSITE DE REIMS CHAMPAGNE-ARDENNE	F REIMS01	VILLA DOUCE 9 BD DE LA PAIX	51097 CEDEX	REIMS
FR	EUCX	ECOLE NATIONALE SUPERIEURE DE CHIMIE DE LILLE	F LILLE13	BP 90108 - BT C7 CITE SCIENTIFIQUE	59652	Villeneuve d'Ascq cedex

- Fitxer de dades mestres d'assignatures

Full Code	Subject Area Description	Equivalent ISCED97	ISCED97 Description
01.0	Agricultural Sciences	62	Agriculture, forestry and fishery
01.1	Agriculture	620	Agriculture, forestry and fishery (broad programmes)
01.2	Agricultural Economics	629	Agriculture, forestry and fishery (others)
01.3	Food Science and Technology	541	Food processing
01.4	Horticulture	622	Horticulture
01.5	Fisheries	624	Fisheries
01.6	Forestry	623	Forestry

- Fitxer de dades mestres de països

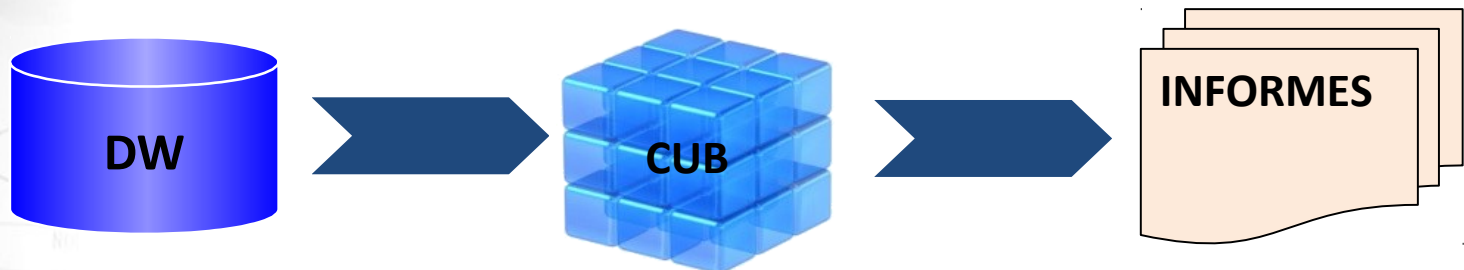
A	B
Code	Country
ad	Andorra
ae	United Arab Emirates
af	Afghanistan

- Fitxers de dades transaccionals Cursos 2011/12 i 2012/13

A	B	C	D	E	F	G	H	I	J	K	L
STUDENT_ID	ID_MOBILITY	HOME_INSTI	HOME_INSTI	STUDENT_AC	STUDENT_GE	STUDENT_NA	STUDENT_SU	STUDENT_ST	NUMB_YRS	IMOBILITY_TY	HOST_INSTIT
47099714S	ES/SM/07/0C	E ALICANTO	01ES	21	F	ES	222	1	3	S	UK BATH01
47397349F	ES/SM/07/0C	E ALICANTO	01ES	22	M	ES	58	1	4	S	P LISBOA04
47397741P	ES/SM/07/0C	E ALICANTO	01ES	20	F	ES	222	1	2	S	F XX
47447533M	ES/SM/07/0C	E ALICANTO	01ES	20	F	ES	223	1	2	S	F PARIS004

Objectius del projecte:

- Fer un magatzem de dades a partir dels fitxers font
- Construir un model multidimensional del tipus *star-schema*
- Crear un conjunt d'informes
- Possibilitat de millores futures
- El programari que utilitzarem serà *BI Pentaho*



Solució proposada:



Continguts

ERASMUS



Anàlisi preliminar

Elaboració del model

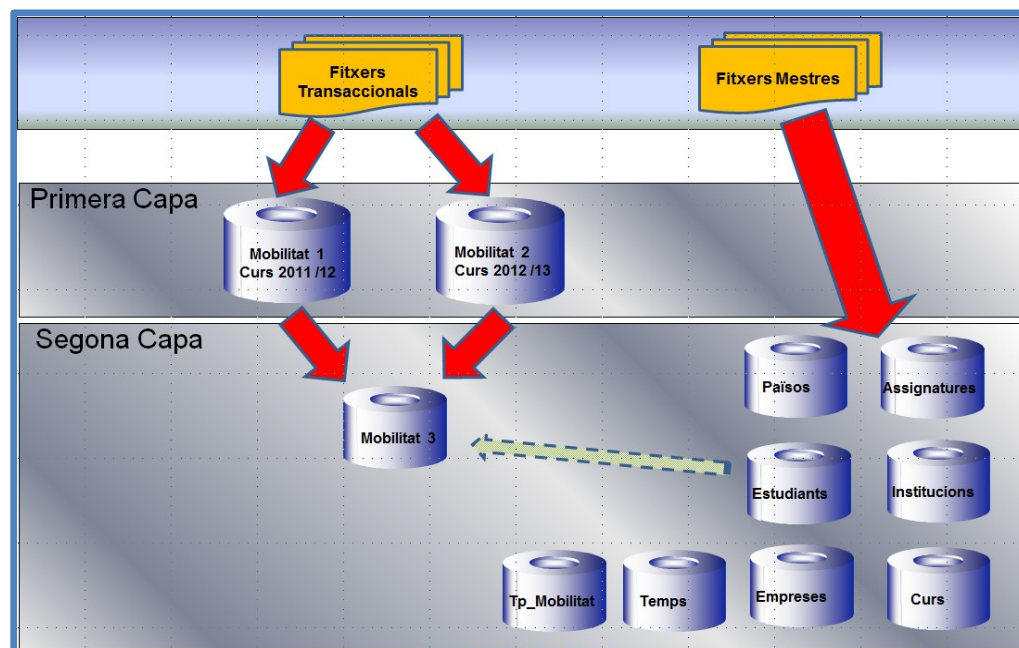
Procés ETL

Explotació de la informació

Conclusions

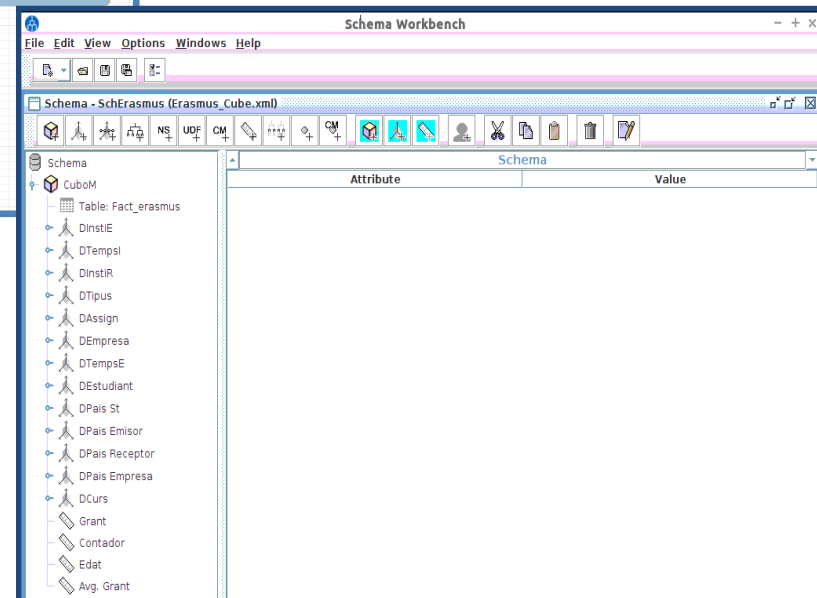
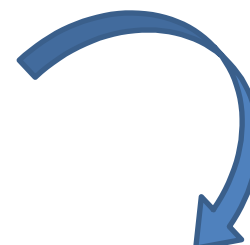
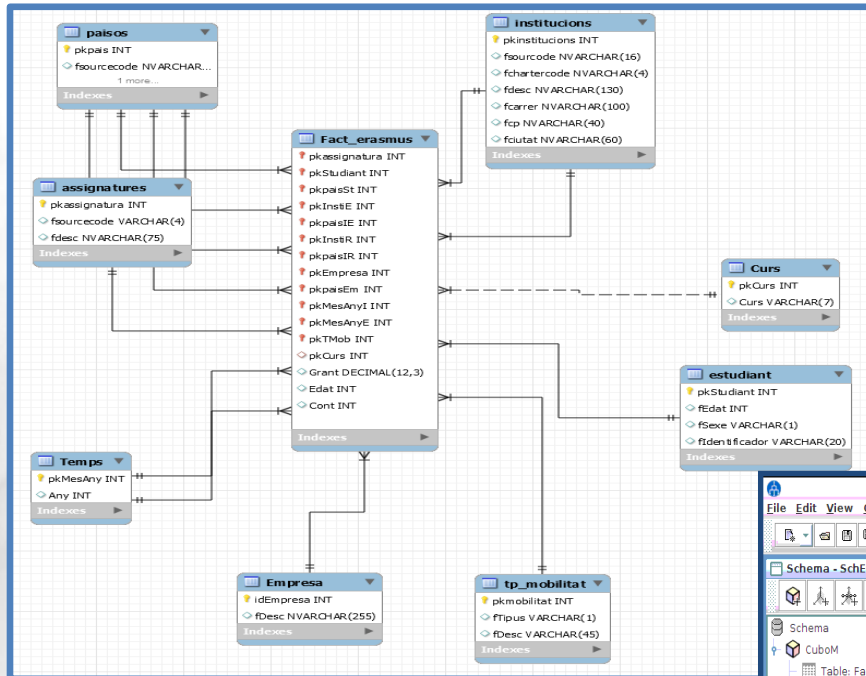
Disseny de la BBDD:

- S'han creat 3 taules de dades transaccionals (Mobilitat1, Mobilitat2 I Mobilitat3).
- S'han creat 8 taules de dades mestres (Països, Assignatures, Estudiants, Institucions, Tp_Mobilitat, Temps, Empreses I Cursos).
- S'ha creat 1 taula de gestió d'errors.



Disseny del cub:

- Relació taula de fets i dimensions



Continguts

ERASMUS



Anàlisi preliminar

Elaboració del model

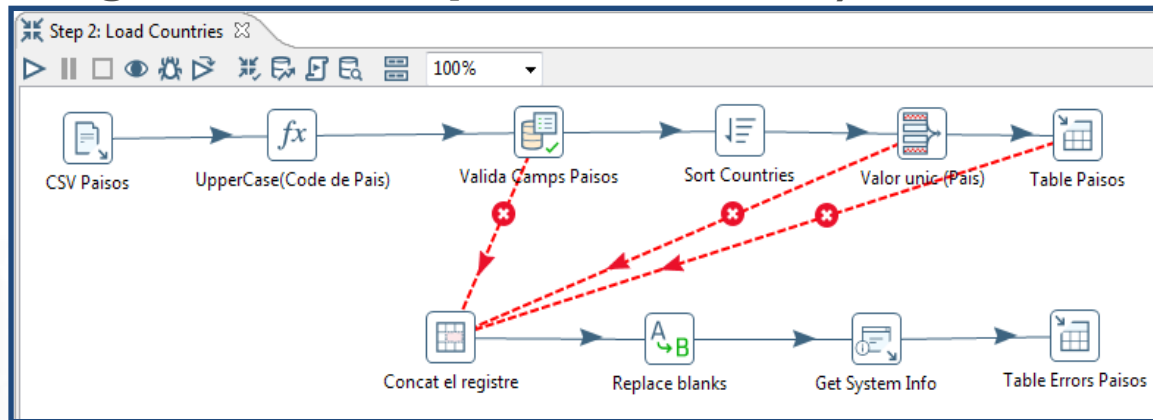
Procés ETL

Explotació de la informació

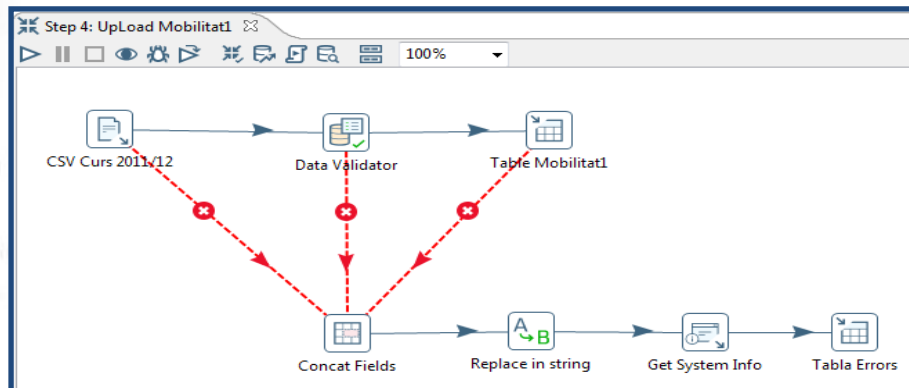
Conclusions

Càrrega de dades (I):

- Fase I: Càrrega de dades mestres amb tractament (Validant integritat dels camps i de les dades)**

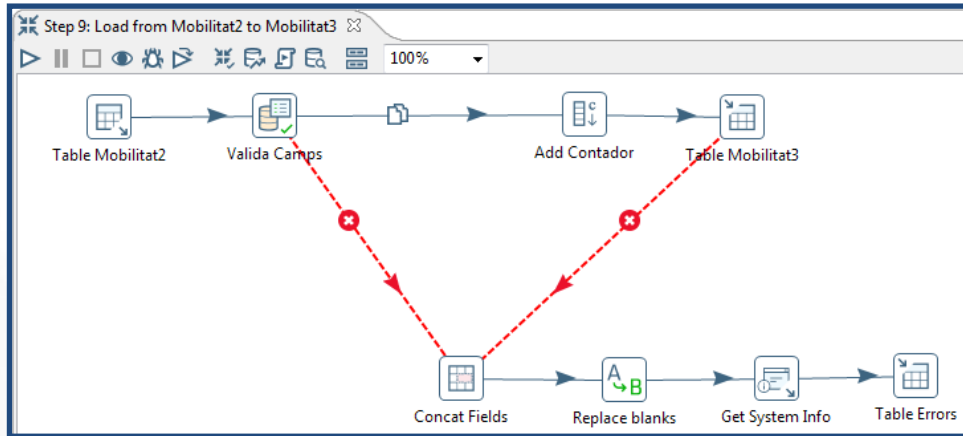


- Fase II: Càrrega de dades transaccionals sense tractament (Validant integritat dels camps)**

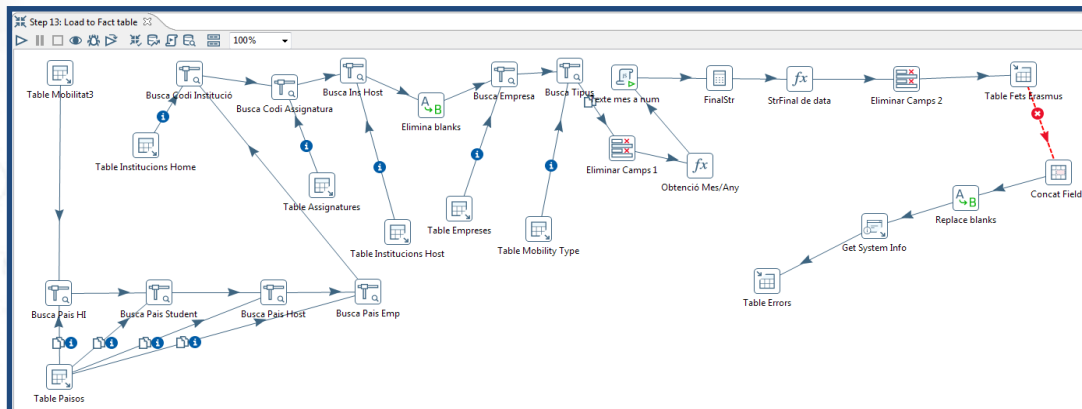


Càrrega de dades (II):

- Fase III: Càrrega de dades transaccionals amb tractament (Validant integritat dels camps i de les dades)**

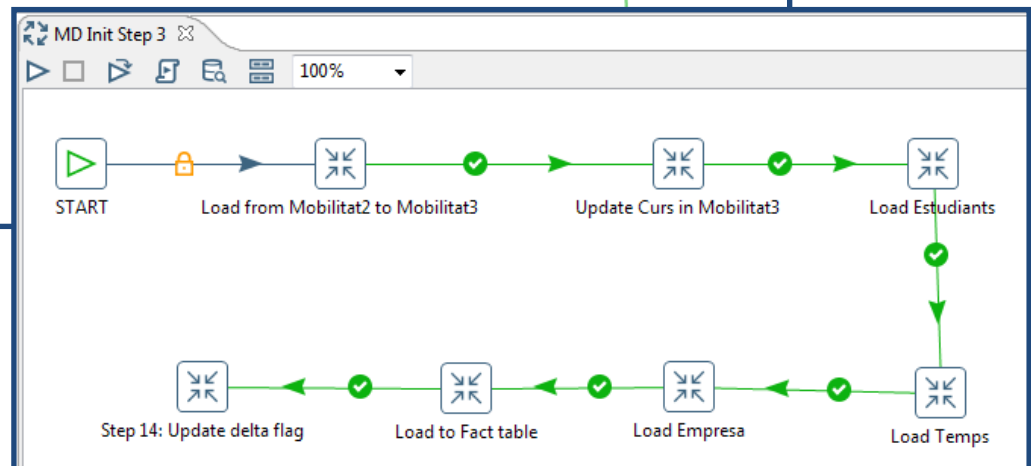
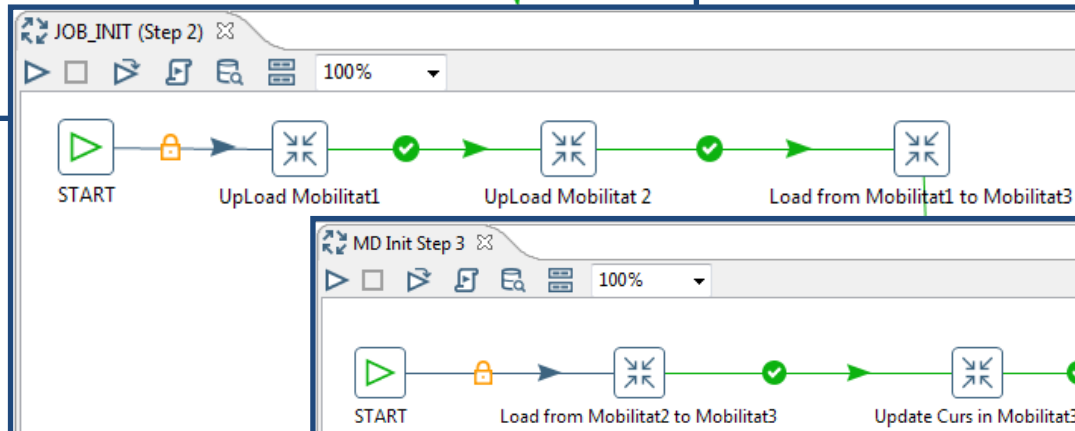
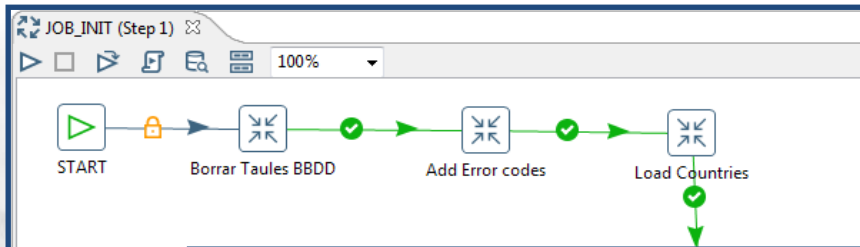


- Fase IV: Càrrega de dades al cub (Substituint, si és necessari valors per codis)**



Càrrega de dades (III):

- Fase V: Automatització de les càrregues per mitjà de *Jobs*, amb possibilitat de fer càrregues programades i periòdiques**



Gestió d'errors:

- **Controlem els errors en la càrrega i transformació del model. Aquests s'emmagatzemen a una taula**

	idErrors	fErrorCode	fErrorDesc	fObjet	fdate
▶	1	Error Long Code Subject	Full;SubjectAreaDescription	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	2	Duplicate Subject code	142;Educationscience	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	3	Duplicate Subject code	144;Trainingforteachersatbasiclevels	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	4	Duplicate Subject code	144;Trainingforteachersatbasiclevels	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	5	Duplicate Subject code	212;Musicandperformingarts	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	6	Duplicate Subject code	214;Design	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	7	Duplicate Subject code	222;Foreignlanguages	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	8	Duplicate Subject code	222;Foreignlanguages	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	9	Duplicate Subject code	222;Foreignlanguages	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	10	Duplicate Subject code	222;Foreignlanguages	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	11	Duplicate Subject code	223;Mothertongue	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	12	Duplicate Subject code	223;Mothertongue	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	13	Duplicate Subject code	223;Mothertongue	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	14	Duplicate Subject code	223;Mothertongue	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	15	Duplicate Subject code	225;Historyandarchaeology	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	16	Duplicate Subject code	225;Historyandarchaeology	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	17	Duplicate Subject code	312;Sociologyandculturalstudies	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	18	Duplicate Subject code	312;Sociologyandculturalstudies	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	19	Duplicate Subject code	313;Politicalscienceandcivics	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39
	20	Duplicate Subject code	321;Journalismandreporting	Step 3: Load Subjects, Institutions & Mobility type	2015-12-12 17:45:39

Continguts

ERASMUS



Anàlisi preliminar

Elaboració del model

Procés ETL

Explotació de la informació

Conclusions

Llistat d'informes:

- **Top 10 Universitats emissores**
- **Top 10 Universitats Receptores**
- **Estudiants per nacionalitat (%)**
- **Estudiants per àrea de coneixement (%)**
- **Evolució comparativa d'estudiants per Curs**
- **Edat mitjana d'estudiants per nacionalitat emissora**
- **Edat mitjana d'estudiants per nacionalitat receptora**
- **Mitjana de beques per nacionalitat emissora**
- **Mitjana de beques per nacionalitat receptora**

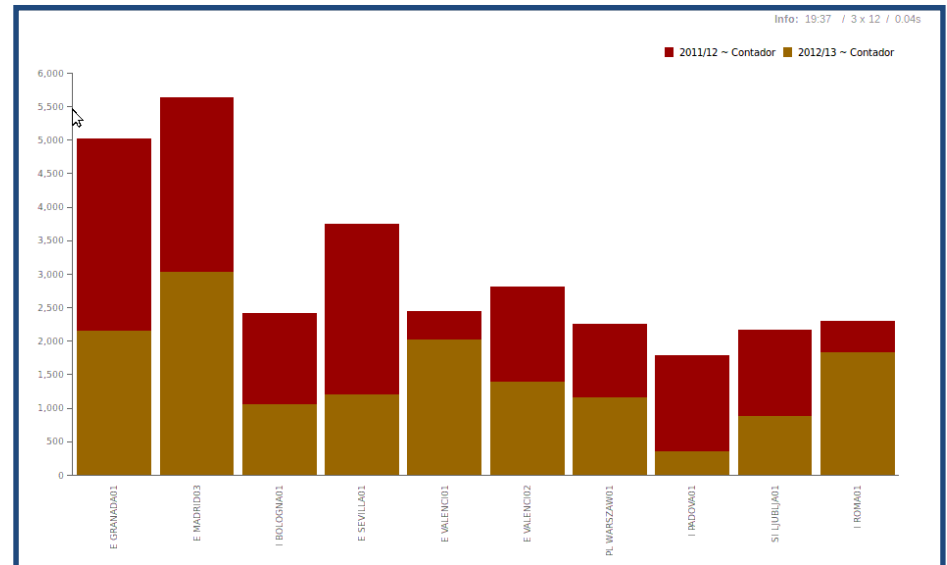
Visualització:

- Els informes es poden veure de dues maneres:

Top 10 Universitats Emisores.saiku

You are using Saiku Community Edition, please consider upgrading to [Saiku Enterprise](#)

Universitat Emisora	2011/12		2012/13	
	Contador	Contador	Contador	Contador
E GRANADA01	1,890	116	1,665	116
E MADRID03	1,831	151	1,674	145
I BOLOGNA01	1,232	328	1,223	424
E SEVILLA01	1,413	99	1,536	58
E VALENCI01	1,335	91	1,326	65
E VALENCI02	1,098	161	950	137
PL WARSZAW01	830	247	820	192
I PADOVA01	755	337	769	397
SI LJUBLJA01	701	340	632	323
I ROMA01	783	395	750	315



Continguts

ERASMUS



Anàlisi preliminar

Elaboració del model

Procés ETL

Explotació de la informació

Conclusions

- **Enunciat amb poca informació**
- **Dades poc consistents. Alta Complexitat**
- **Anàlisi preliminar amb modificacions**
- **Model relacional i multidimensional amb modificacions**
- **Informes poc treballats**
- **Important adaptabilitat, com a la vida real**
- **Gran treball de recerca i desenvolupament**

