

# Use Qlik with Apache Kylin

**Date:** September 2016

**Author:** Ramón Portolés, Alberto

[Linkedin](#)

[a.ramonportoles@gmail.com](mailto:a.ramonportoles@gmail.com)

## Intro

It's a new player in the BI tools world, with a high growth from 2013

[On-line demos](#)

Nowadays has connectors for Hadoop Databases: Hive and Impala

But doesn't have Kylin connector

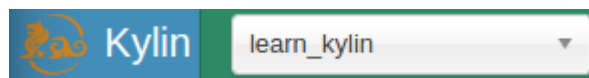
## Used Software

- [Kylin v1.5.2](#)
- [Kylin ODBC v1.5](#)
- [Qlik Descktop v3.0.2.0](#)

## Pre-requisites

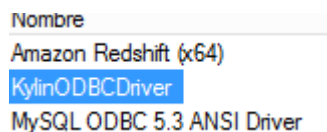
A) We need an instance of Kylin, with a cube: [Quick Start with Sample Cube](#), will be enough

You can check:




B) Download Install and configure Kylin ODBC Driver

1. [Install Kylin ODBC](#)
2. Add DNS System in Windows



3. And configure: User-name /Password: ADMIN / KYLIN

DSN Name	<input type="text" value="kylin"/>
Protocol	<input type="text" value="http"/>
Server Host	<input type="text" value="172.17.0.2"/>
Port	<input type="text" value="7070"/>
Username	<input type="text" value="ADMIN"/>
Password	<input type="password" value="•••••"/>

4. Click Connect and Choose the project

	<input type="button" value="Connect"/>
Project	<input type="text" value="learn_kylin"/>

Click on “Done”

## Install Qlik:

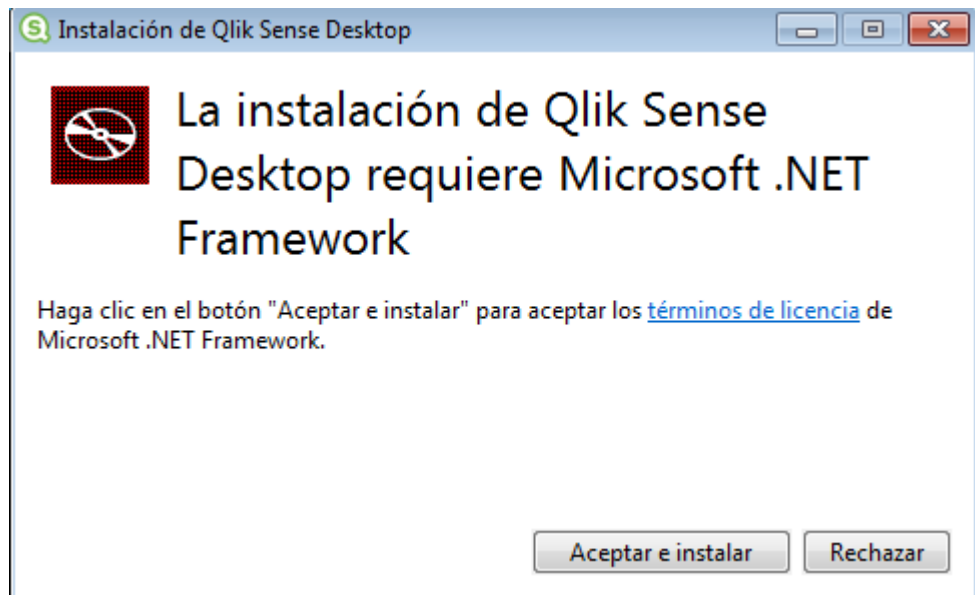
Requires: ([System Requeriments](#))

- Windows 7 SP1 (64 bit) or 8.1 (64 bits)  
Be careful with 64 bits and SP1, its mandatory  
I tested on Windows 2012 and didn't Work
- Framework .Net 4.5.2
- 1.5 GB of free Disk Space

Note: The Language of Qlik Desktop is the same as your S.O. (you can change)

My apologies for some pictures in Spanish language ;)

We launch the process **as Admin:**



It is recommended that you close all other applications before continuing.



#### Install

Installs all program features. Requires up to 1.5 GB of free disk space.

Install



3.0.2.0 has been installed successfully.

## Issue 1: The Burn engine cannot run with an MBA under the .NET 4

(The [Qlik Forum](#) and msg txt wrongly suggests its about .NET problem)

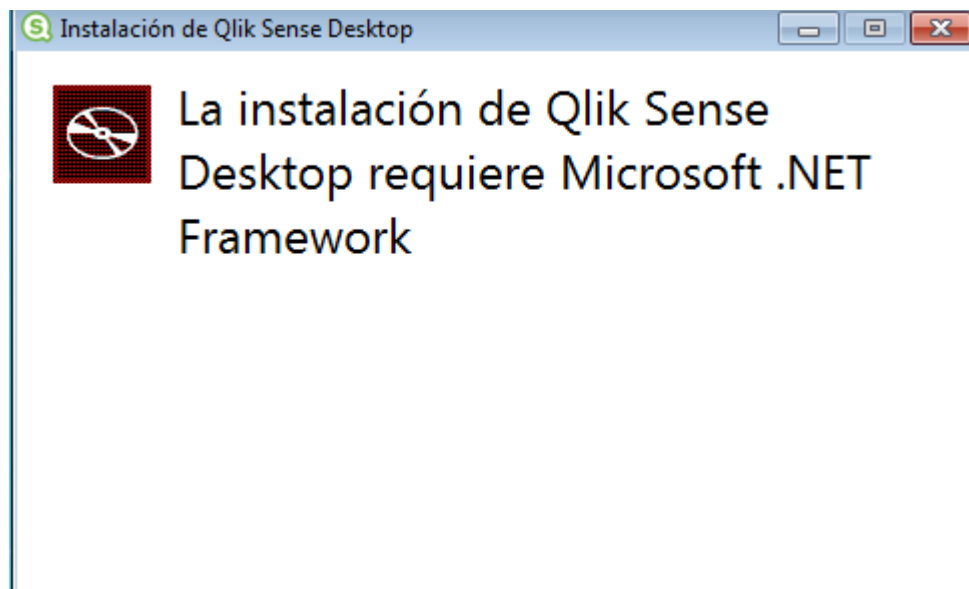
In log you will see:

```
: Error 0x81f403e8: The Burn engine cannot run with an MBA under the .NET 4 CLR on  
Windows 7 RTM with .NET 4.5.2 (or greater) installed.  
[0070:0A8C][2016-09-09T22:34:33]i000: Loading prerequisite bootstrapper application  
because managed host could not be loaded, error: 0x81f403e8.  
[0070:0BC0][2016-09-09T22:34:33]i000: Setting numeric variable 'WixStdBALanguageId' to  
value 3082
```

Solution: Qlik requires SP1 of Windows 7 to install ( this info Isn't in the official [System Requirements](#))

[Download & Install SP1 64 bits](#)

## Issue2: Microsoft .Net framework is required for Qlik Sense desktop setup



And you **don't have** a button "Accept and install"

I tried [Qlik Community 1](#) and [Qlik Community 2](#), but didn't work

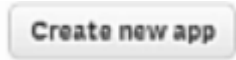
Solution: I deleted in windows registry all references to Qlik

# Create New App

Open Qlik Desktop



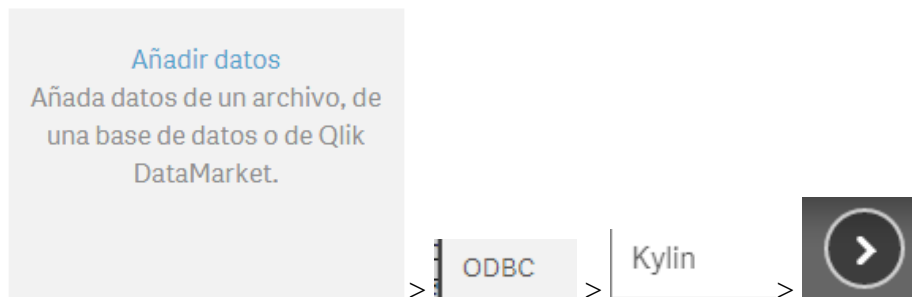
Create New App, With name Kylin



Now, the complex is import your Kylin data into Qlik

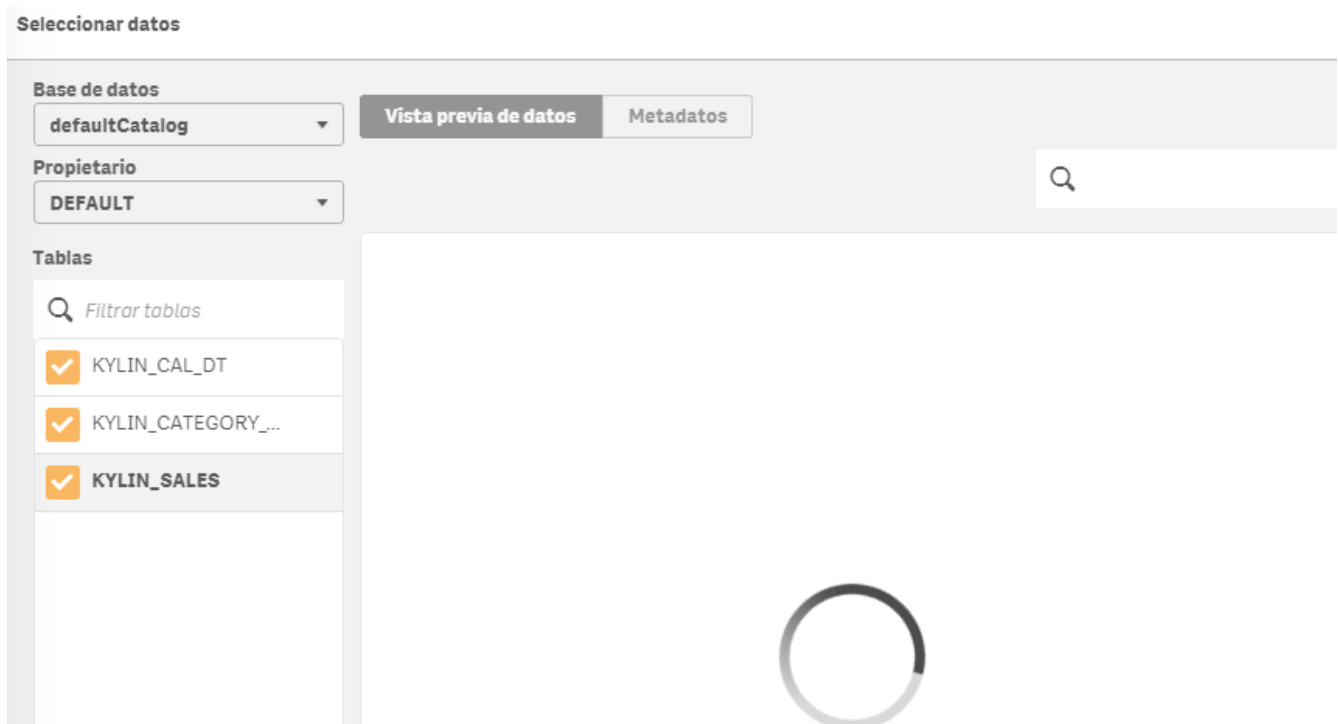
I documented the fails, because I think are also important

## Attempt 1: To add Data (fail)



The login and password can be empty

We can see the Kylin's tables:



Problem: The load **never finish** (I Waited 20 minutes)

See 'Issue 1: Qlik Metadata'

## Attempt 2: Add Data (fail)



The login and password can be empty

An click on: 

Problem: The load **never finish**

See 'Issue 1: Qlik Metadata'


## Attempt 3: Add Data (successful)

We need configure our own data loader using a script

We change some configuration in parameters from connector created in “Attempt 2: Add Data (fail)”

We will need activate debug mode:




And test default script with  , you can see an execution step by step

List of changes:

- Change format to Dates to YYYY-MM-DD

```
6 SET TimeFormat='h:mm:ss';
7 SET DateFormat='YYYY-MM-DD';
```

- Add Connection String: You can add directly with 

```
19 LIB CONNECT TO 'Kylin';
```

- Load columns & data from tables

First, we test query using Kylin UI:

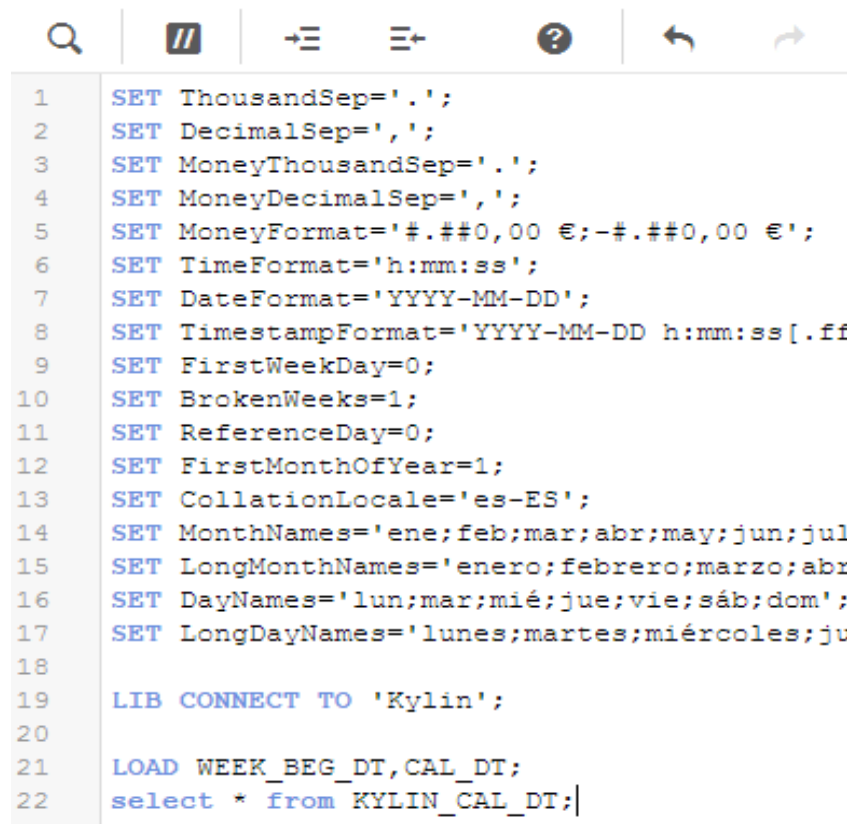
New Query	Saved Queries
1	<code>select * from KYLIN_CAL_DT</code>

The result is: (731 rows)

### Results (731)

WEEK_BEG_..	CAL_DT
2012-03-03	2012-03-09
2013-05-05	2013-05-11
2013-12-08	2013-12-10


The result must be similar to this: (See “script\_data\_1.txt”, on gitHub)




```
1 SET ThousandSep='.';
2 SET DecimalSep=',';
3 SET MoneyThousandSep='.';
4 SET MoneyDecimalSep=',';
5 SET MoneyFormat='#,##0,00 €;-#,##0,00 €';
6 SET TimeFormat='h:mm:ss';
7 SET DateFormat='YYYY-MM-DD';
8 SET TimestampFormat='YYYY-MM-DD h:mm:ss[.ff';
9 SET FirstWeekDay=0;
10 SET BrokenWeeks=1;
11 SET ReferenceDay=0;
12 SET FirstMonthOfYear=1;
13 SET CollationLocale='es-ES';
14 SET MonthNames='ene;feb;mar;abr;may;jun;jul';
15 SET LongMonthNames='enero;febrero;marzo;abr';
16 SET DayNames='lun;mar;mié;jue;vie;sáb;dom';
17 SET LongDayNames='lunes;martes;miércoles;ju';
18
19 LIB CONNECT TO 'Kylin';
20
21 LOAD WEEK_BEG_DT,CAL_DT;
22 select * from KYLIN_CAL_DT;
```

We can test the import:

Save Script: 

Out of debug mode: 

Now, “Load data” button must be enabled: 

The Output must be similar to: (check the row count is the same like Kylin UI)



Tiempo transcurrido 00:00:01

Ya ha comenzado a cargar datos

Conectando a Kylin

Conectado

KYLIN\_CAL\_DT

Líneas recuperadas: 731

App guardada

Completado con éxito

0 error(es) forzado(s)

0 clave(s) sintética(s)

To Import other table (fact table) you only need add

```
LOAD PART_DT,LEAF_CATEG_ID,LSTG_SITE_ID,LSTG_FORMAT_NAME,PRICE,  
SELLER_ID,MIN_PRICE_,MAX_PRICE_,COUNT_, COUNT_DISTINCT_SELLER_ID_,  
COUNT_DISTINCT_LSTG_FORMAT_NAME_;  
select * from KYLIN_SALES;
```

We can put “alias” to data easy:

[CAL\_DT]:

```
LOAD WEEK_BEG_DT,CAL_DT;
```

```
select * from KYLIN_CAL_DT;
```

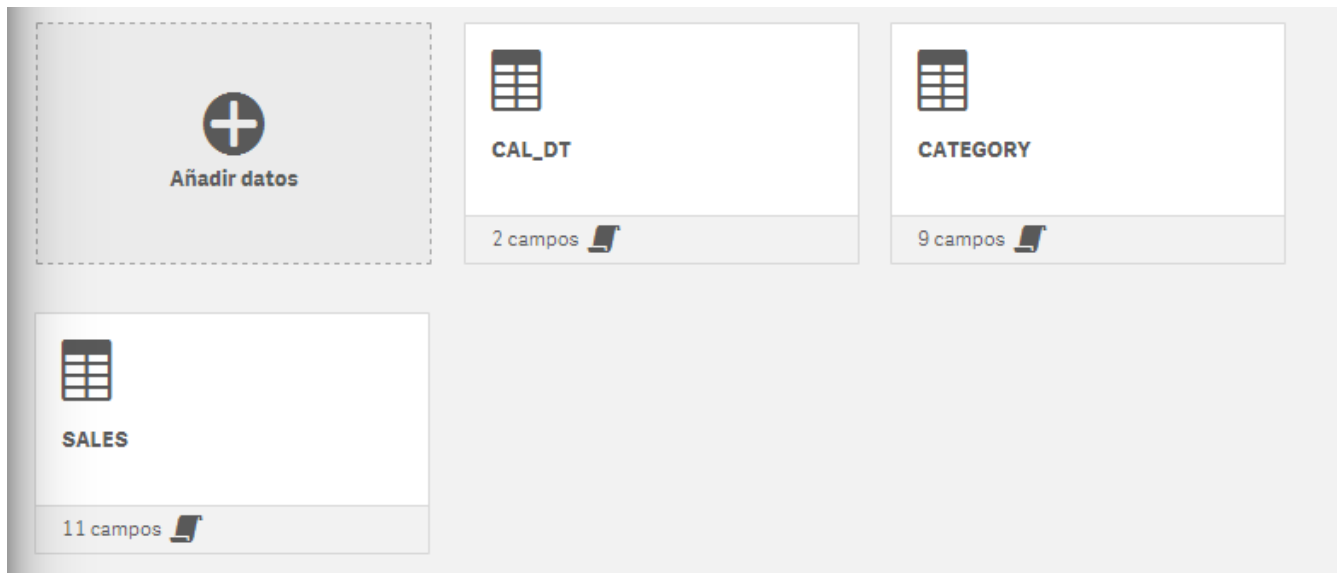
In my Hard: a Laptop with HBase, Kylin, Hive .. In a Docker Container and Qlik in a Virtual Box, the performance to ingest data has been good, only 2 sec for 3 Tables and 10K Rows

Tiempo transcurrido 00:00:02

Ya ha comenzado a cargar datos

```
Conectando a Kylin
Conectado
CAL_DT << KYLIN_CAL_DT
Líneas recuperadas: 731
SALES << KYLIN_SALES
Líneas recuperadas: 9,840
CATEGORY << KYLIN_CATEGORY_GROUPINGS
Líneas recuperadas: 144
```

We can see all ingested tables in “[data manager](#)”:  >  Gestor de datos



NOTE that: If you try to see data of one table: Click in one, **you can't see data** (below)

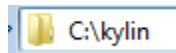
*“Data tables defined in the load script are not managed in Data manager”*



Esta tabla se añade usando el script de carga de datos, y sólo puede visualizarse en el visor del modelo de datos tras haber cargado los datos.

## Issue 1: Qlik Metadata

The log of ODBC kylin are in:



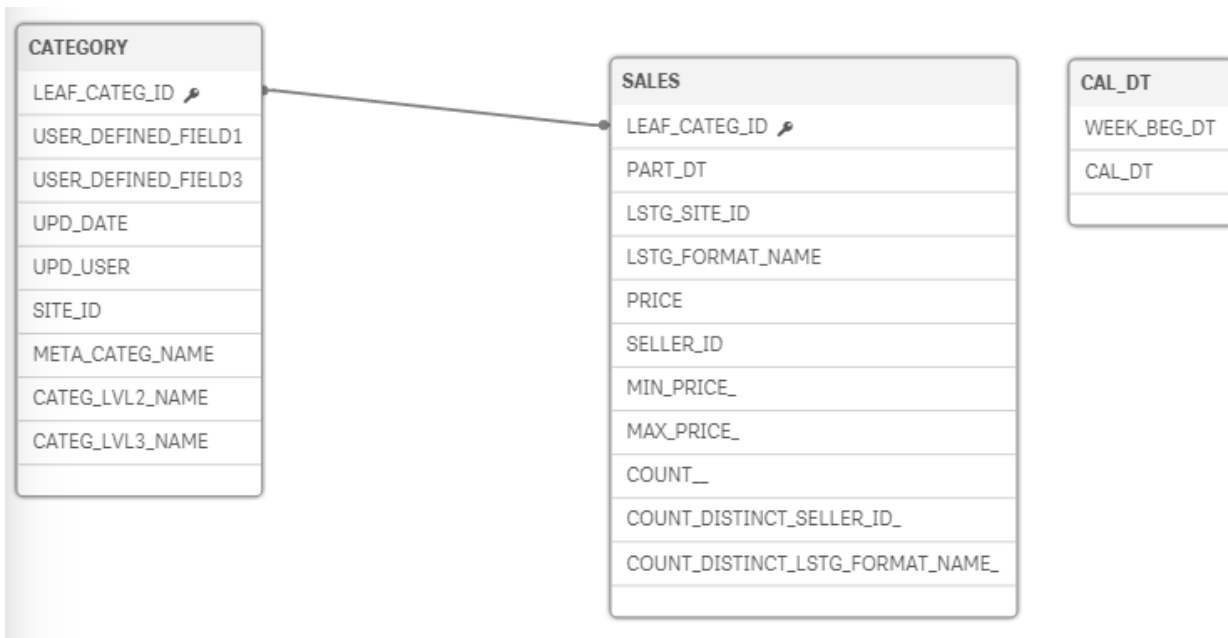
When use Attempt 1 and 2, we can see errors: **SQLGetInfoW**

```
[FATAL] [2016-09-10.14:01:08] _SQLCopywCharDataW - No holder for data size
[FATAL] [2016-09-10.14:01:08] In Shipping mode, Skip popping up window...
[FATAL] [2016-09-10.14:01:09] _SQLCopywCharDataW - No holder for data size
[FATAL] [2016-09-10.14:01:09] In Shipping mode, Skip popping up window...
[FATAL] [2016-09-10.14:01:09] _SQLCopywCharDataW - No holder for data size
[FATAL] [2016-09-10.14:01:09] In Shipping mode, Skip popping up window...
[FATAL] [2016-09-10.14:01:09] _SQLCopywCharDataW - No holder for data size
[FATAL] [2016-09-10.14:01:09] In Shipping mode, Skip popping up window...
[ERROR] [2016-09-10.14:01:09] Function Missing!!! SQLGetInfoW, Field: 89, DataPtr:
```

## Define Table associations

([Managing data associations](#))

You can see you actual table associations in:



But this association isn't correct, you can see the corrects in Kylin UI > Model designer > Data Model:

Fact Table

DEFAULT.KYLIN\_SALES

Lookup Tables				Filter ...
ID	Table Name	Join Type	Join Condition	
1	DEFAULT.KYLIN_CAL_DT	inner	DEFAULT.KYLIN_SALES.PART_DT = DEFAULT.KYLIN_CAL_DT.CAL_DT	
2	DEFAULT.KYLIN_CATEGORY_GROUPINGS	inner	DEFAULT.KYLIN_SALES.LEAF_CATEG_ID = DEFAULT.KYLIN_CATEGORY_GROUPINGS.LEAF_CATEG_ID DEFAULT.KYLIN_SALES.LSTG_SITE_ID = DEFAULT.KYLIN_CATEGORY_GROUPINGS.SITE_ID	

How can I change? :

In "data manager": > Gestor de datos > Associations

In true, you can't change because


*"Data tables defined in the load script are not managed in Data manager"*

## Create dataSheets

There is only way to add data from load script into a sheet:

In "data manager": > Gestor de datos > Load data > Edit Sheet



You can see all Columns in: , you can classifies like Dim or measures with Right click


Also, you can define a Hierarchy in Dim:

**Crear nuevas dimensiones**

☐ Única ☒ Jerárquico

Filtrar por tabla

Todas las tablas

 <i>Buscar</i>
CAL_DT
CATEG_LVL2_NAME
CATEG_LVL3_NAME
COUNT__
COUNT_DISTINCT_LSTG_FO...
COUNT_DISTINCT_SELLER_I...
LEAF_CATEG_ID
LSTG_FORMAT_NAME
LSTG_SITE ID

Campo:

≡ CATEG\_LVL2\_NAME

×

*fx*

≡ =CATEG\_LVL3\_NAME

×

*fx*

*fx*

Nombre:

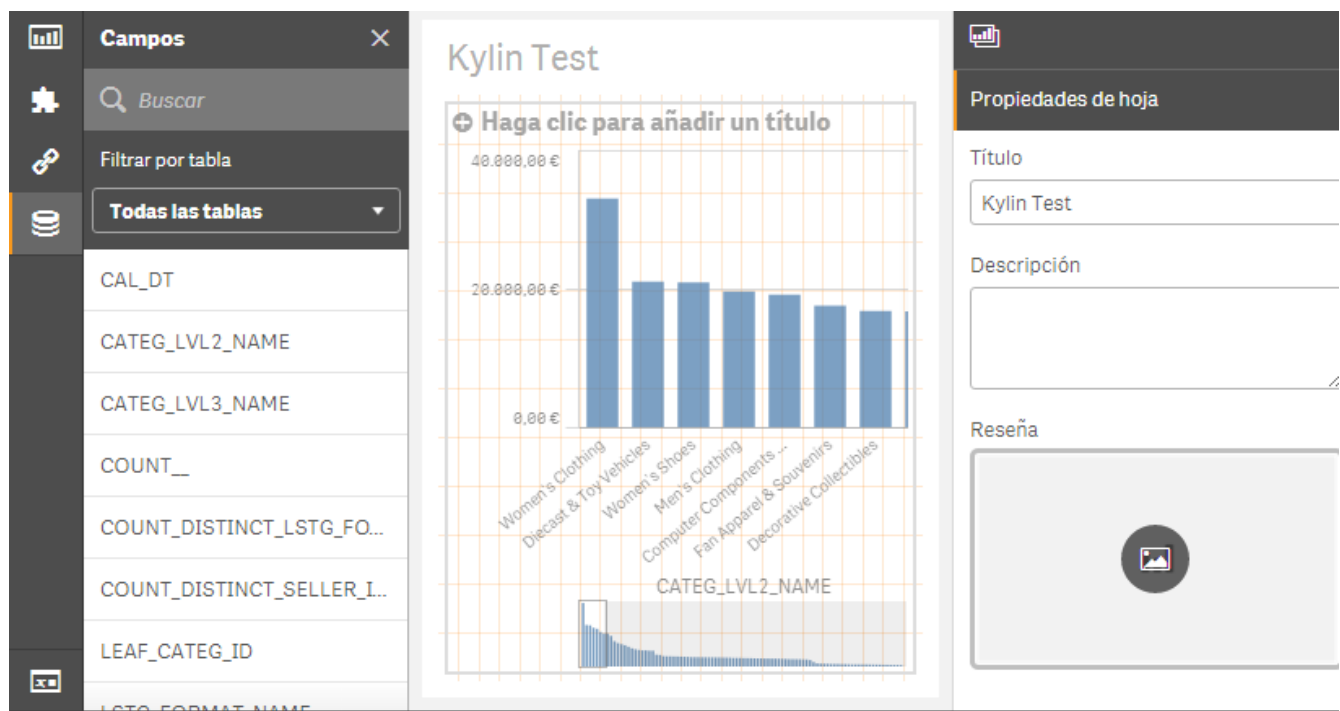
Categ

Descripción:

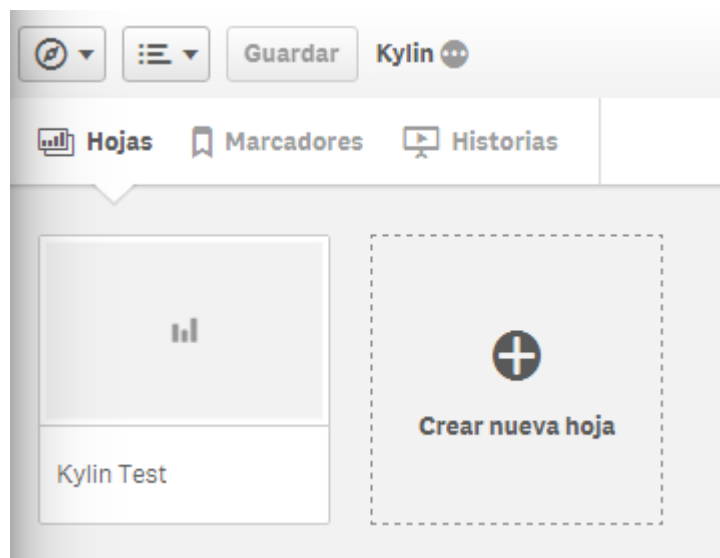
Etiquetas:

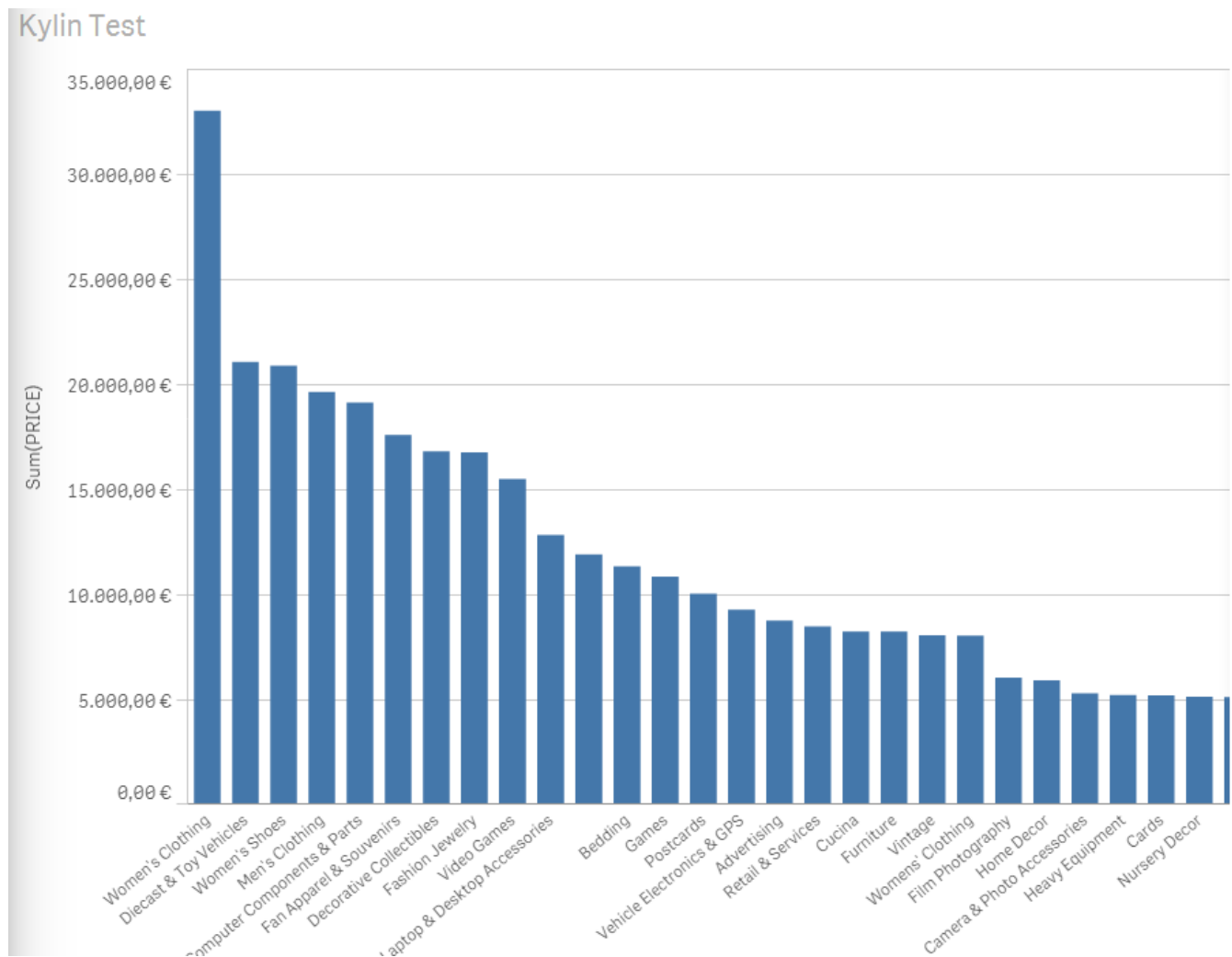
+

Añadir dimensión



You can View the result from main menu:





## Future Work

- Load data using “**load script**” is a workaround to read data from Apache Kylin, but like data isn't controlled by **Data Manager** you can't use the preview of data ([manual reference](#)) ([Qlik Community](#))
- Same problem with “**associations**”, you can't manage/change because this data aren't managed by **Data Manager**

This problems means, if your automatic associations aren't correct, **the data results will be false**

**For any suggestions, feel free to contact me**

**Thanks, Alberto**