Data Visualization

Context

Data visualization is viewed by many disciplines as a modern equivalent of visual communication. It involves the creation and study of the visual representation of data, meaning «information that has been abstracted in some schematic form, including attributes or variables for the units of information». It is one of the steps in data analysis or data science.
The primary goal of data visualization is to communicate information clearly and efficiently via statistical graphics, plots and information graphics. Effective visualization helps users analyse and reason about data and evidence. It makes complex data more accessible, understandable and usable.
Data visualization is strategic. All its stake is to help to decision making because the future of company is on a line : data which would not be right could lead to bad decision.

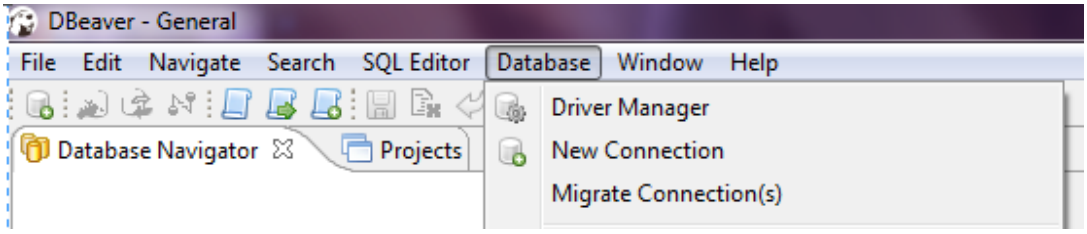
In our projet, the data visualization is very important because it can help clients to visualize their electrical consumption over the year of their home.

DBeaver and Tableau
Why
Tableau is a data visualization tool. Like we said, data visualization is very important nowadays so we decided to choose Tableau which is one of the more famous dataviz's tool.
With Tableau, The user is not dependent anymore. He can create himself his analyses and his displays to publish them then and spread them.

The DBeaver is an SQL client and a database administration tool that we used to connect to vertica.

To connect DBeaver to Vertica, you need to follow these:

In DBeaver, select Database > New Connection

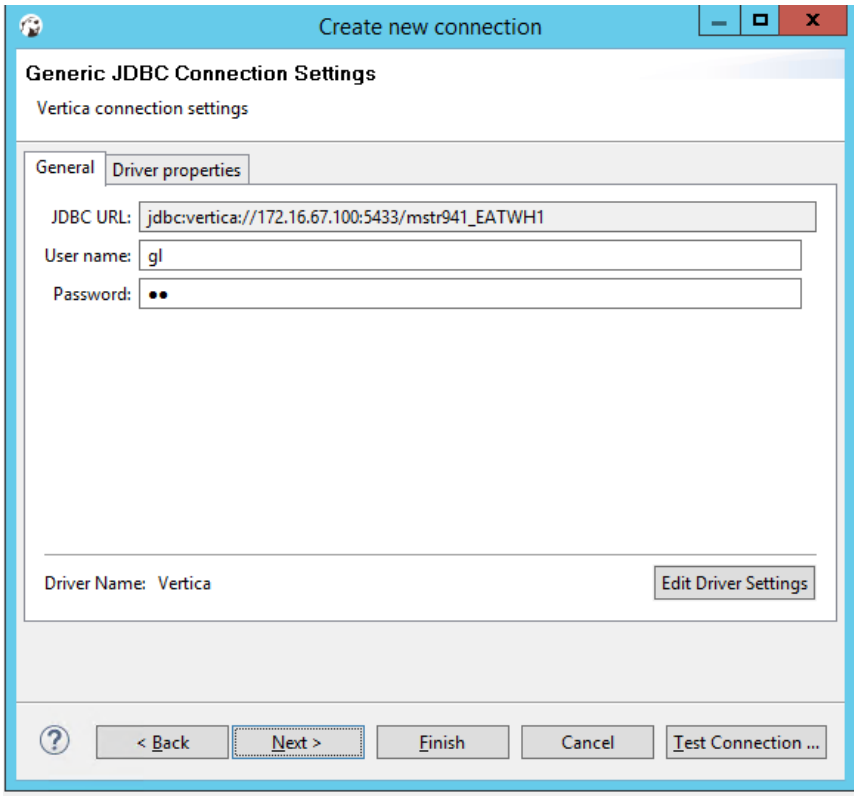


In the Create New Connection window, select Vertica and click Next.

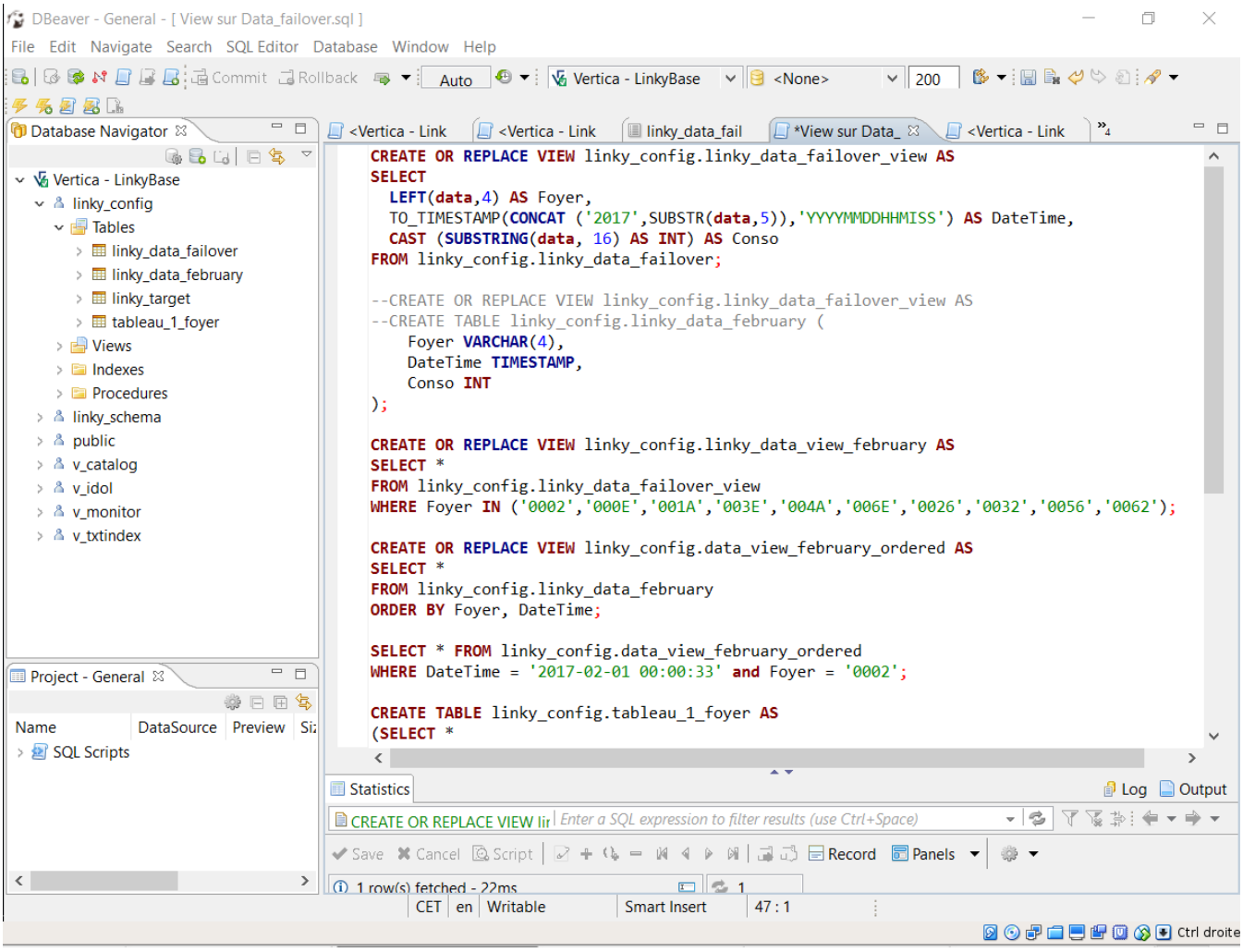
In the general tab, enter you database credentials.

Then, in Edit Driver Settings, select add file and select the Vertica JDBC file you downloaded.

Finally, click Test Connection. If your connection to Vertica us successful, a message appears.



From here, you can run queries and visually explore your Vertica database.



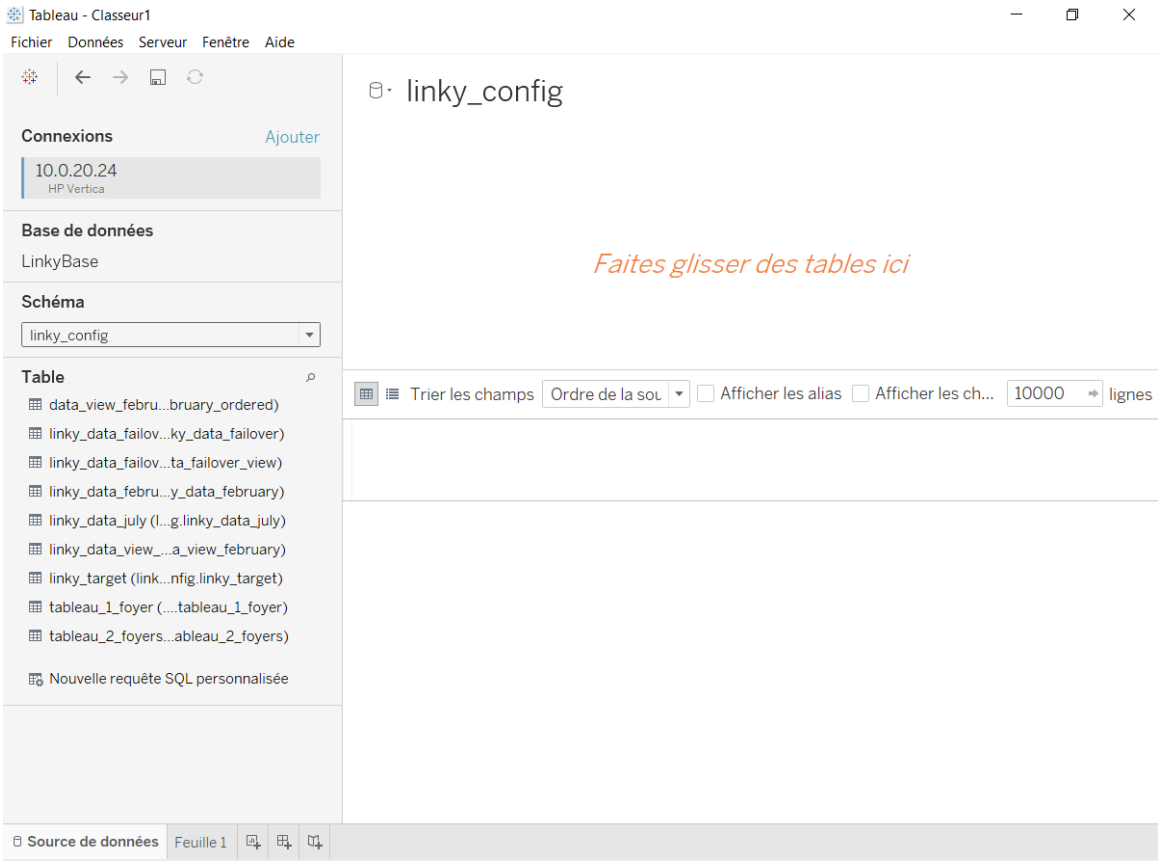
We want to show you how to use Tableau now that we made our queries on DBeaver.

This is the home page when you open Tableau.

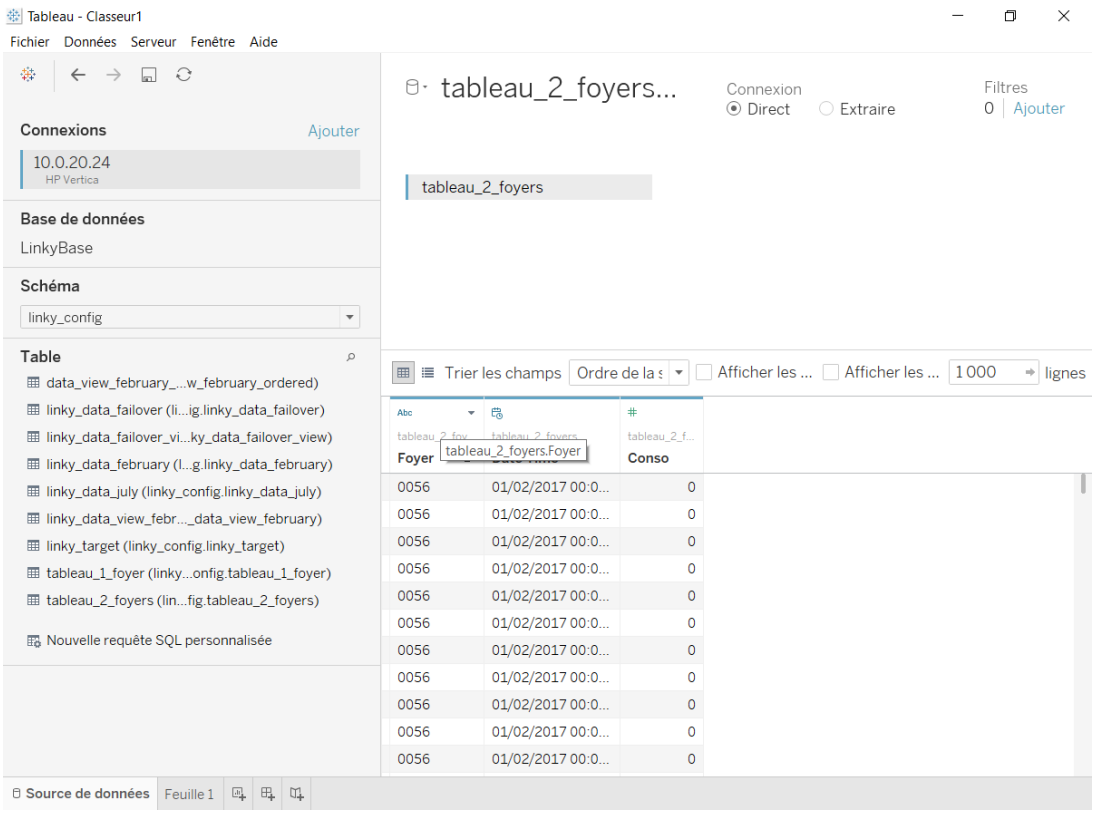
On the left, click on HP Vertica

You will have to put your username and password. Click then on Connexion.

This is the page you see once you are connected. You just have to drag and drop one table at the indicated position



Once you drag and drop your table, you can explore the data as you can see on the next picture. In our case, we wanted to visualize the home's id, the date and the home's consumption



For visualizing your data, click on Feuille 1 at the bottom of Tableau. You can now see the following page.

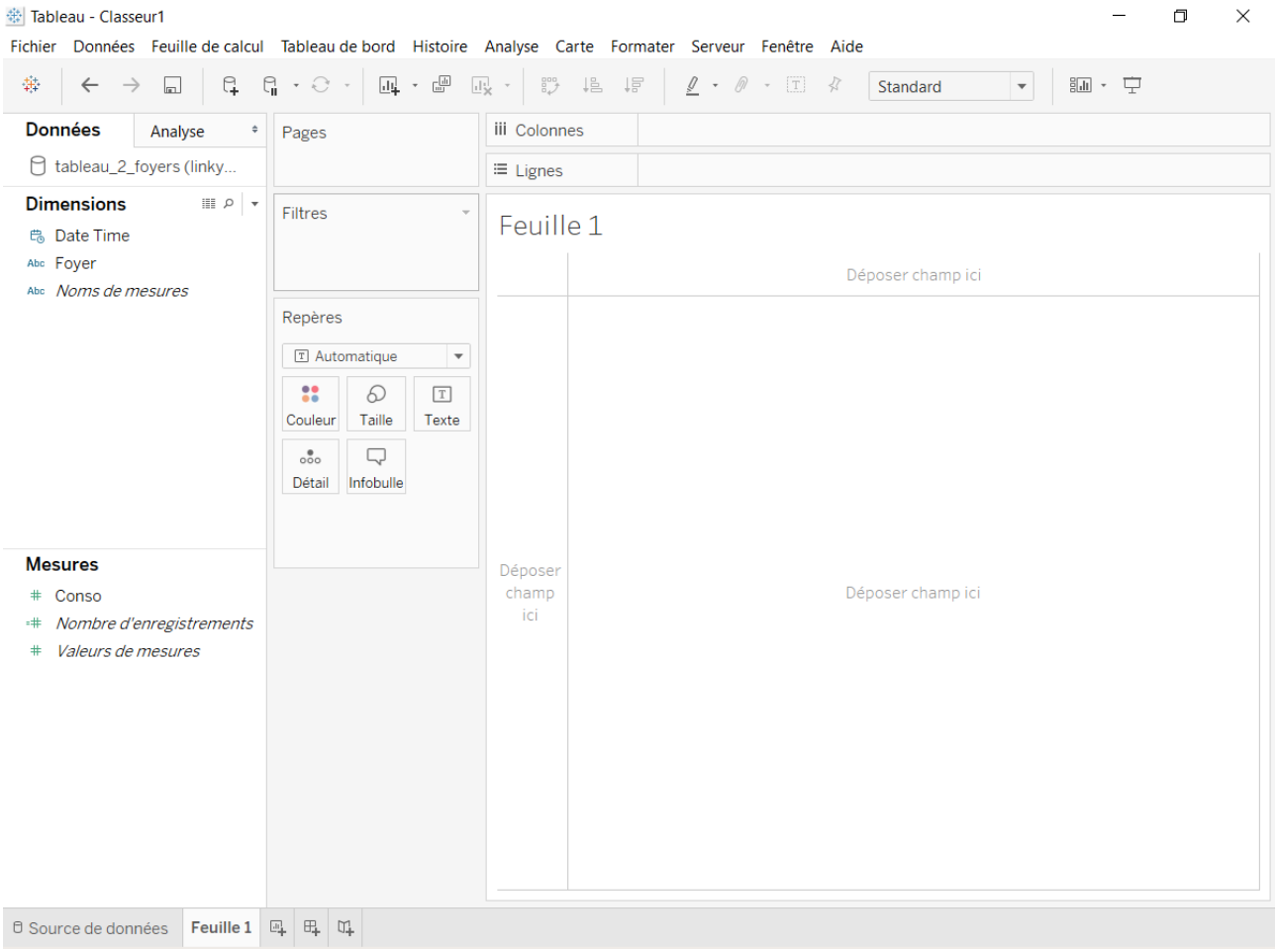
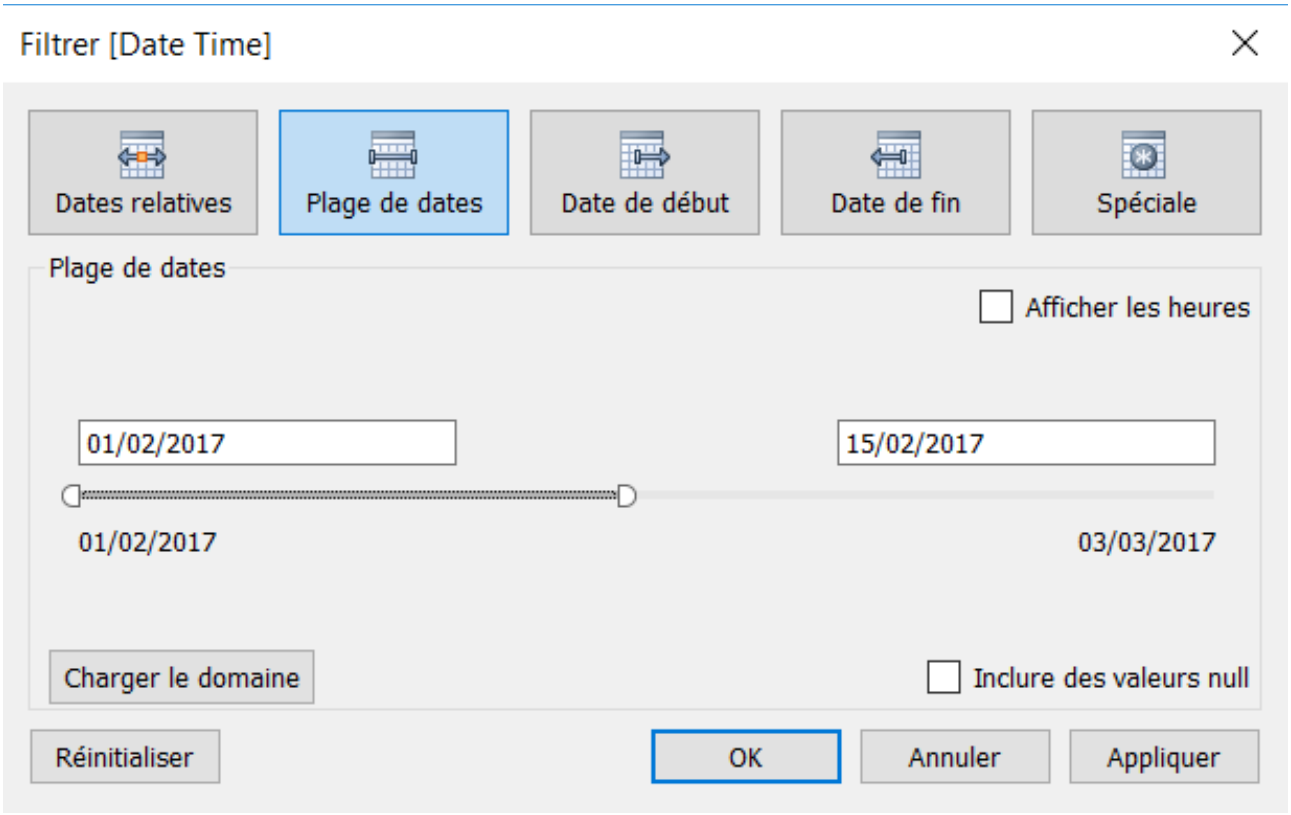
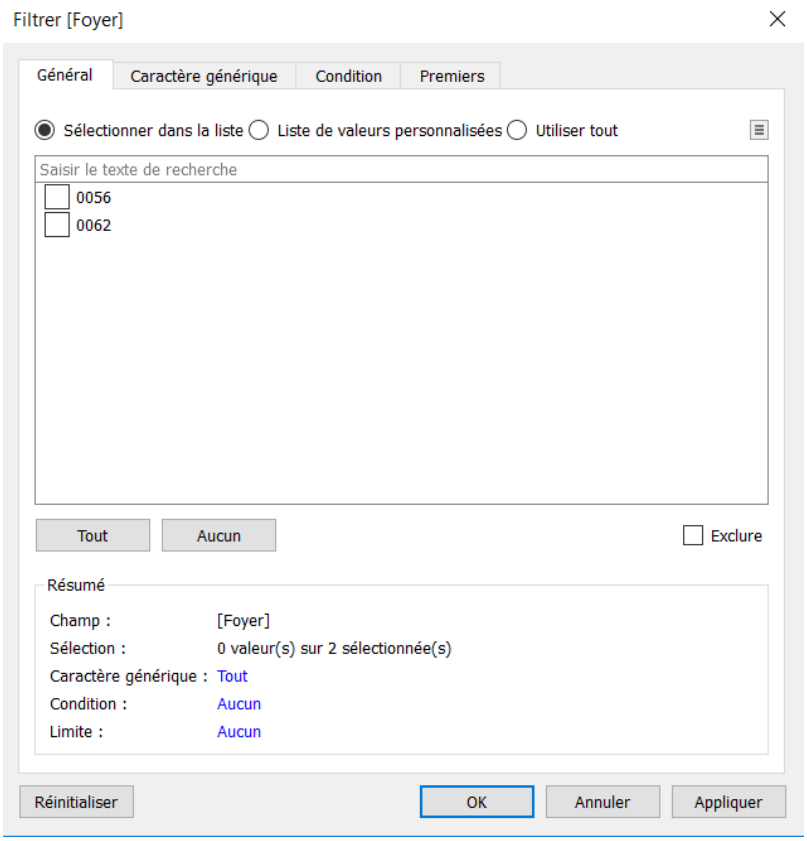


Tableau is really simple and easy to use. You can drag and drop what you need to see into « colonnes » and « Lignes ».
You can also apply filters also by dragging and dropping into the « Filtre » section.

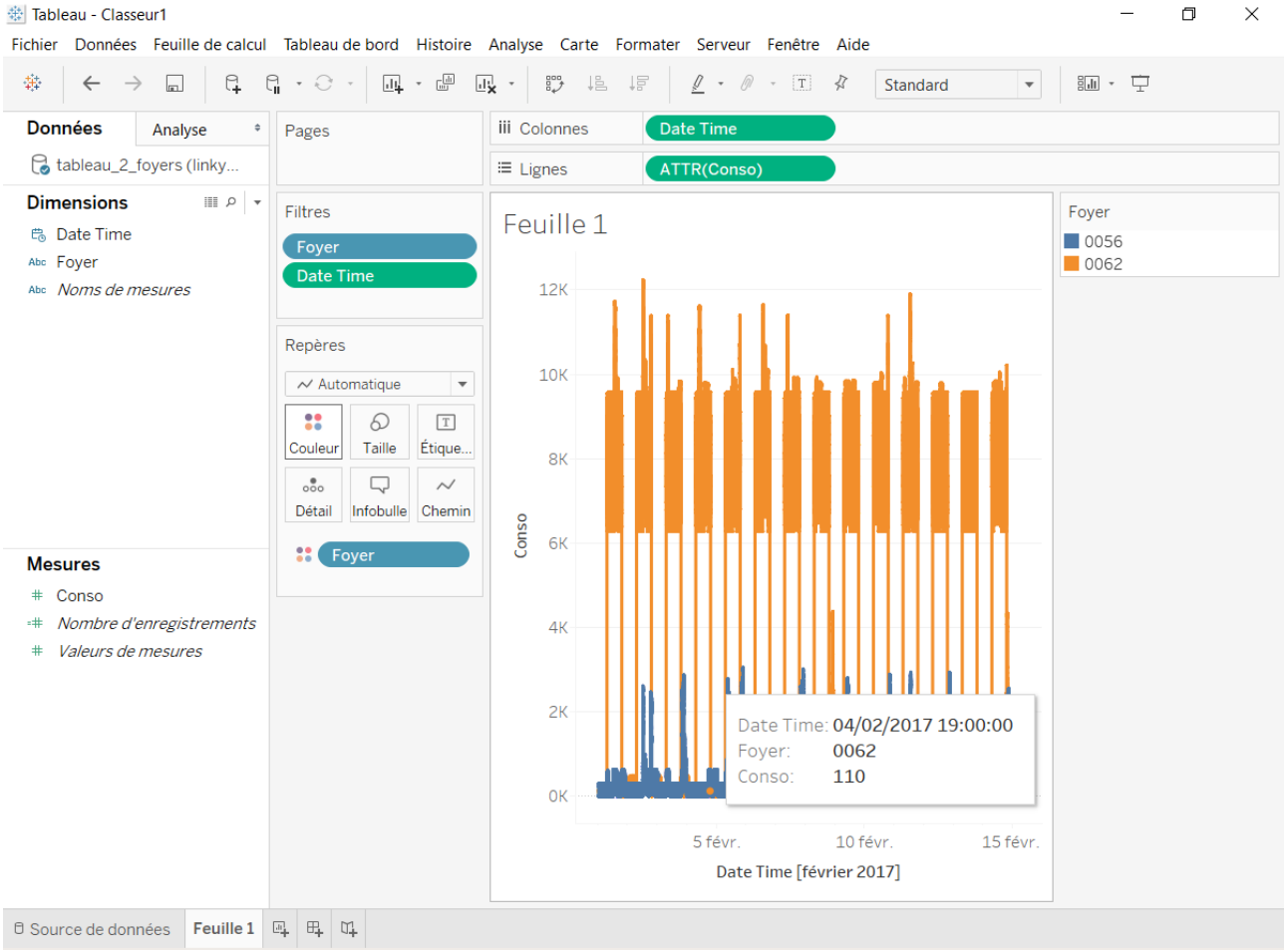
For example, we apply a filter on the date. We only want to select only the two first weeks of February.



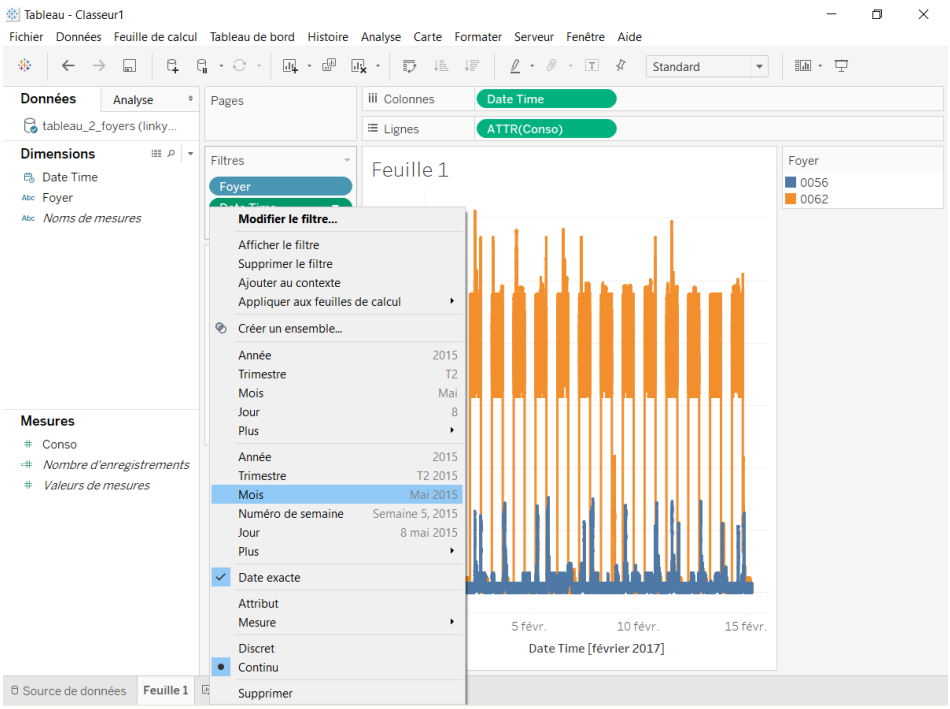
We can also apply a filter on the home to select.



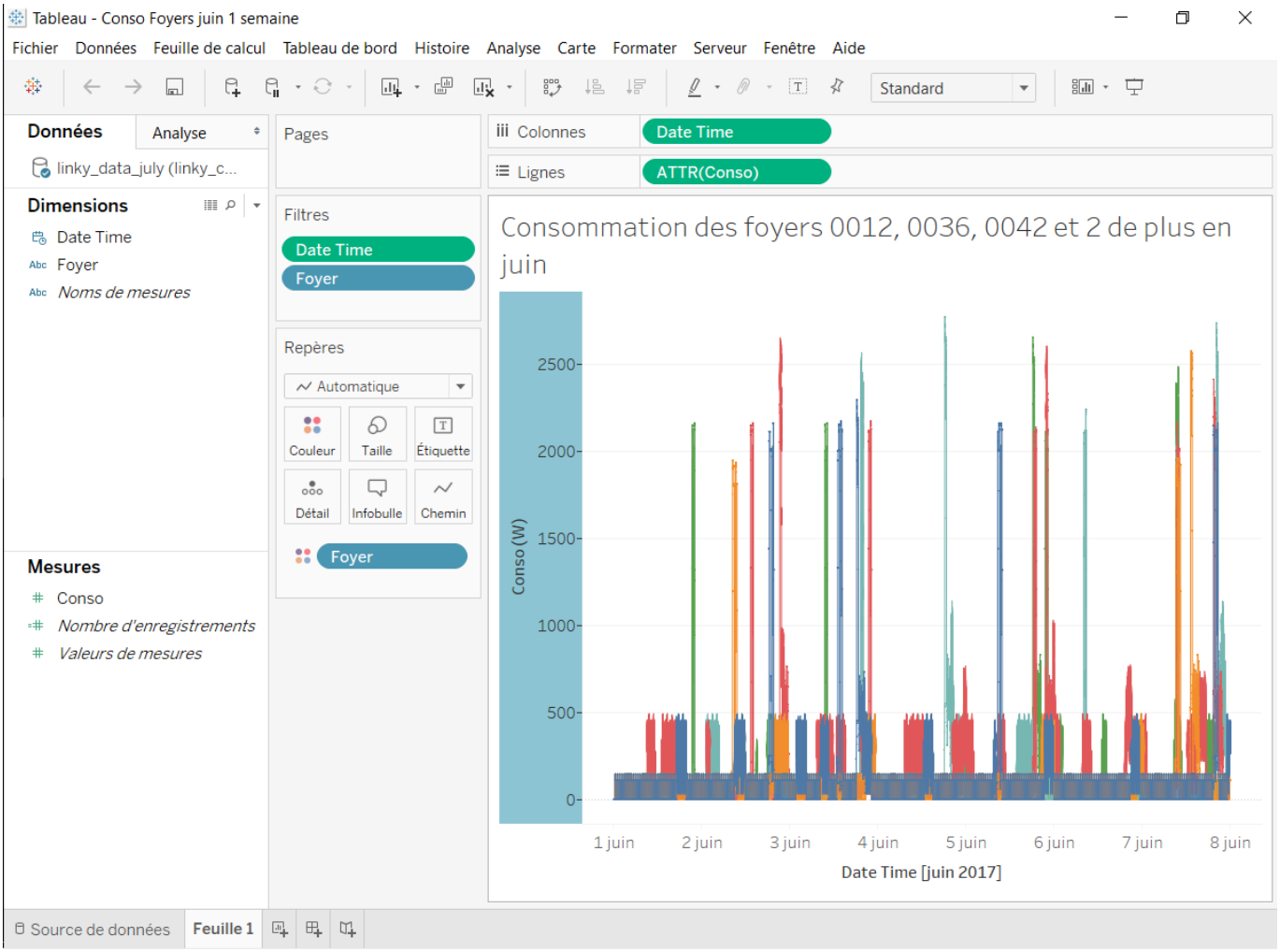
Here is the result of all the filter we apply on our data. you can see the electrical consumption of the two home we selected in our filter at the date we chose.
By default, the colors of the two home is identical but we can also apply a filter on the colors.



You can modify your filters whenever you want by clicking on the filter you want to change.



This is an other example of visualization of our data. You can see the electrical consumption of several home during one week in June.



Amelioration

Spark