

# TD1 Data visualization

## Practice with Tableau Public

Ruiwen HE

Computer Science department

Engineering School de VINCI (ESILV)

Objectives of the TD

- Discover [Tableau Public](#)
- Understand the mapping of data onto graphic marks (line, square, etc.) and their properties (size, color, etc.)
- Distinguish between raw and abstract data, as well as between dimensions and measures
- Explore a visualization design space

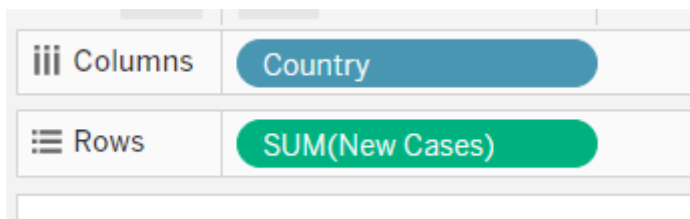
Answer the questionnaire as the practical work progresses on [form](#).

### 1. Load data into Tableau

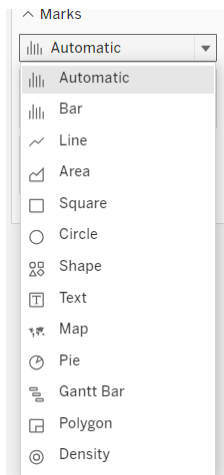
Download the following csv: [Stat\\_covid.csv](#). Launch Tableau and select connect to text file, load the file, and switch to the data view (bottom left of Tableau click on Data Source).

### 2. See the results in graphical form

Now switch to the 1st Sheet (Sheet 1). Start by displaying a graph with the **Country** in column, and **New Case** in line.



Then you should choose a type of chart to visualize the data



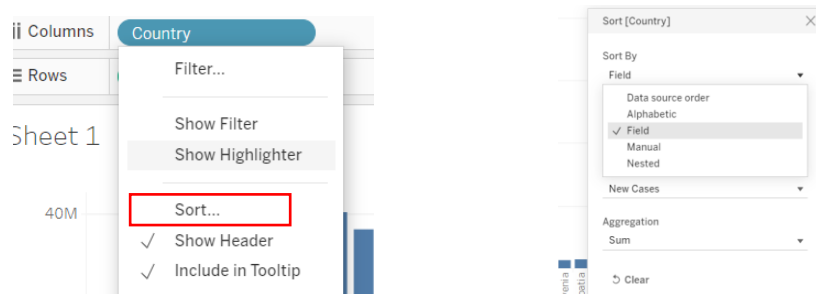
Questions (answer in the form):

2.1 What type of data concerns the **Country** and the **New Case** respectively (Nominal, Ordinal, Interval and Ratio)?

2.2 What type of chart do you think we should use?

2.3 Why do you choose this type of chart?

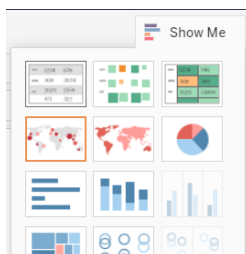
Re-order the Country (in order of most new cases line according to the SUM measure)



Questions (answer in the form):

2.4 After sorting, please indicate the three countries with the most **New Cases**.

Beside the chart chosen, explore other representations in Show Me.



### 3. Filter by region and identity trends

Due to geographical, social, demographic, and other reasons.

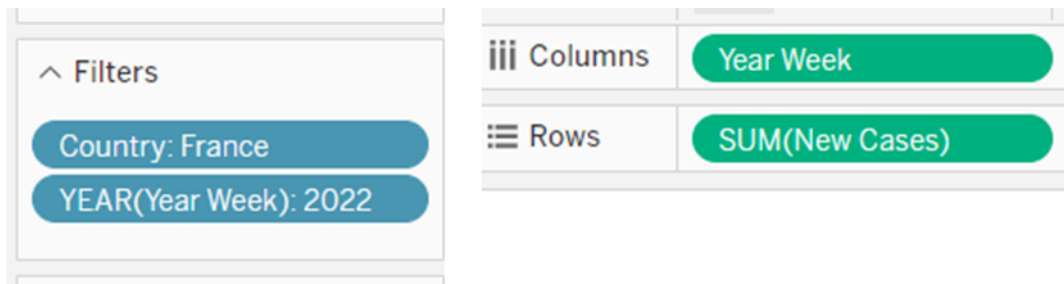
We will filter them to focus on France. At the same time, we focus on the epidemic situation by years individually. Therefore, we set two values respectively: **Country** and **Year Week**, in the filter.

Questions (answer in the form):

3.1 What default data type is the **Year Week** value?

3.2 Could we use it directly? If not, what data type do we need?

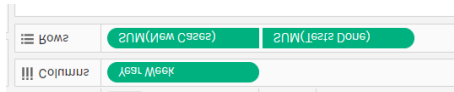
We want to know the trends of covid in France in 2022 by the number of weeks.



3.3 What type of chart do you think we should use?

3.4 Show your result and describe it.

You might find that the state of epidemic seems to have improved starting in May 2022. Is this true? Display again a graph with the new case curve, and those of **Tests Done** in the same figure (Show me > for dual line try).



3.5 What default data type is the **Tests Done** value?

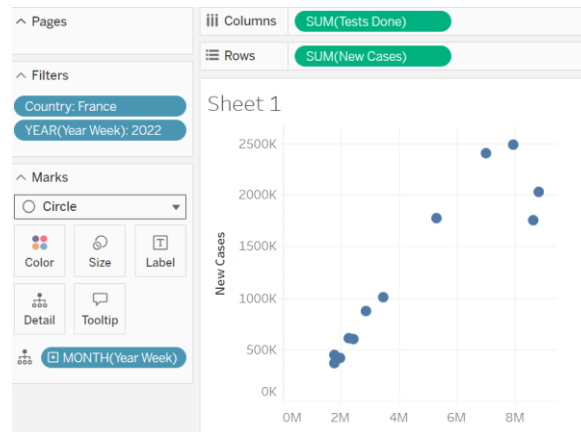
3.6 Could we use it directly? If not, what data type do we need?

3.7 Show your result and describe it.

At the same time, let's take a look at the trends in positive test rates.

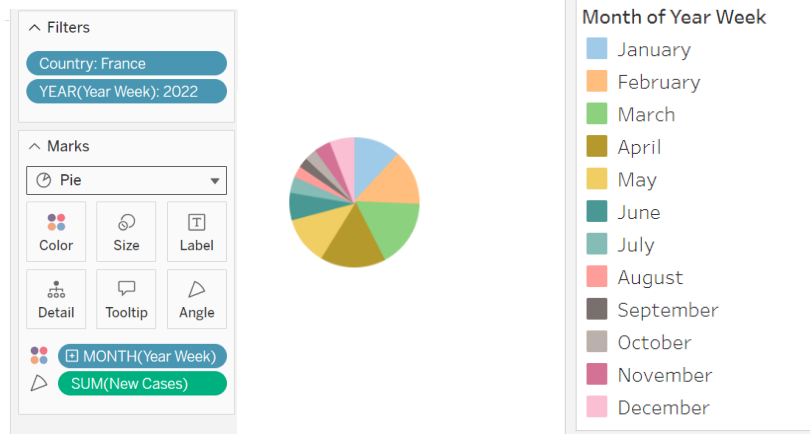
#### 4. Relationship between the **Tests Done** and **New Case**

Create a new sheet, Add the **Year Week** as a detail Mark, and the **New Case** in rows, the **Test Done** in columns.



### 5. Which month has the most *New Case* in 2022?

Create a new sheet, Add the *Year Week* as a color Mark, and the *New Case* as angle.



### 6. Map

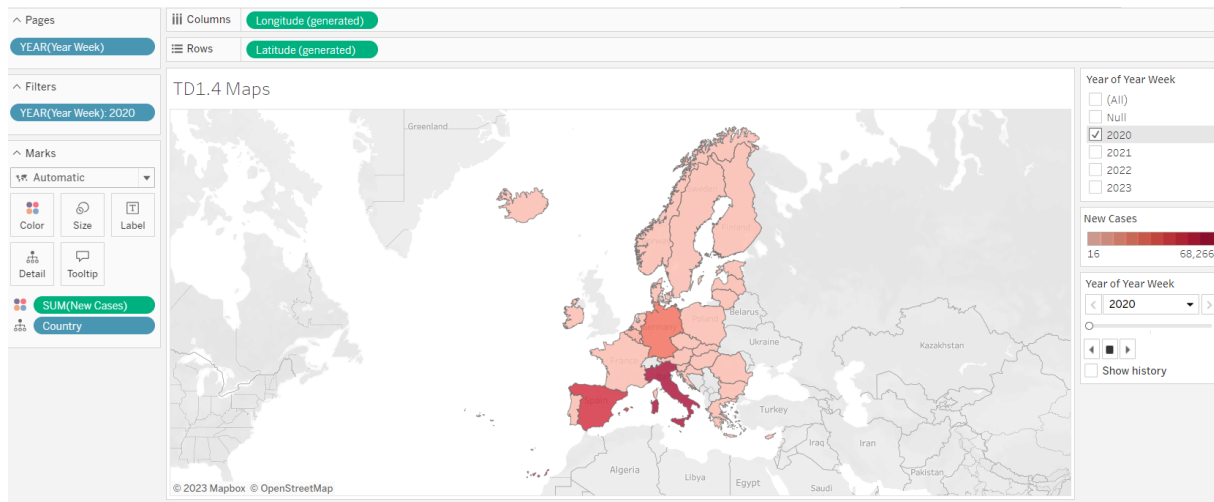
Create a new sheet, add Latitude and Longitude in row and column.

Add the *Country* as a detailed Mark, and the *New Case* in color.

Questions (answer in the form):

Add the time in Pages. Configure it to navigate from one month to the next.

Do the same things for positive test rates.



### Exercise 1:

1. Download data from <http://becomingvisual.com/sales.xls>.
2. Show the age and gender demographics that have the most sales.
3. Select the type of chart that best illustrates your idea.

### Exercise 2 :

1. Download data from [Netflix.xlsx](http://Netflix.xlsx).
2. Give a comparative analysis of films and television shows.
3. Give a graph that can help a producer answer the following question: during which month should new content be broadcast? (Month with the lowest amount of content).
4. Do a movie audience analysis
5. Offer other analyzes