Tableau Class

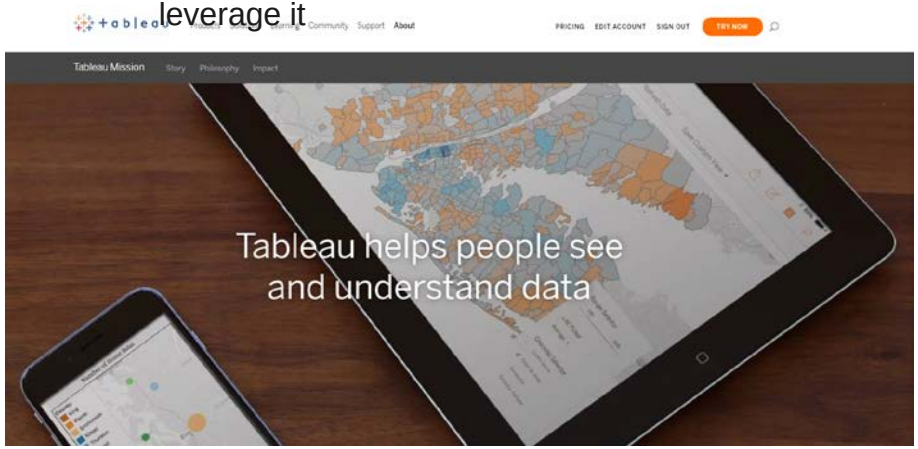
(IE481, Lab #3)

Hansoo Lee

2020. 4. 02

Introduction of Tableau

- Tableau Software is an American interactive data visualization software company founded in January 2003 by Christian Chabot, Pat Hanrahan and Chris Stolte, in Mountain View, California.
- Chabot, Hanrahan and Stolte were researchers at the Department of Computer Science at Stanford University who specialized in visualization techniques for exploring and analyzing relational databases and data cubes
- Tableau helps people see and understand data. With Tableau, you can quickly and easily analyze your data and leverage it



We're focused on your success

1,053	109	550K	1100+
LEARNING RESOURCES	TRAINING VIDEOS	COMMUNITY MEMBERS	PARTNERS

1. Prepare your data for analysis with Prep.

Tableau Prep

Tableau Prep empowers more people to get to analysis faster by helping them quickly and confidently combine, shape, clean and operationalize their data flows.

2. Use powerful drag-and drop analytics in Desktop

Tableau Desktop

Called "the gold standard" in visual analytics, Tableau Desktop upended the business intelligence industry and ushered in a new paradigm of self-service insight.

3. Share data and insights with Server or Online.

Tableau Online

Want the sharing and collaboration of Server, but without having to actually manage a server? Then you want Tableau Online. Secure. Scalable. And Look Ma—No hardware to maintain!

4. Watch data-driven decisions spread.

Tableau Server

Share your data and dashboards to multiply your impact. Whether you keep your Server deployment on-prem or deploy to the public cloud you can keep the management of your server in your hands.

How is Tableau being used in Industry and Research field?

Pepsi: analysis time 90% reduced

PepsiCo cuts analysis time by up to 90% with Tableau + Trifacta

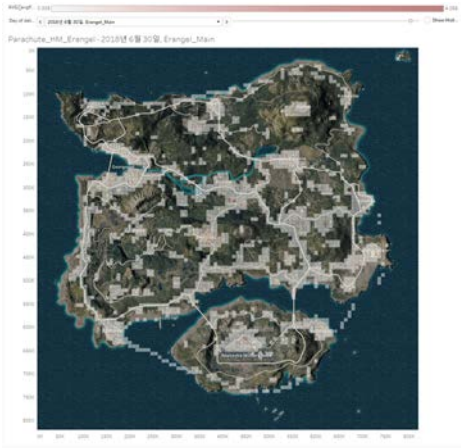
Up to 90% reduction in report production time | Wrangle big data for analytics at scale | "Best-in-class-service" with faster customer analysis

PUBG: increasing game data analysis efficiency, improving accessibility, faster problem identification and resolution

펍지(PUBG) 주식회사, 태블로로 실시간 데이터 기반 의사결정 실현

100개 이상의 지표로 게임 데이터 분석 효율성 향상 | 전세계 직원들의 인사이트에 대한 접근성 개선 | 문제 파악 및 해결 속도 향상

"Battlegrounds maps" with Tableau



What Can We Do with Tableau?

https://www.tableau.com/

The Tableau platform

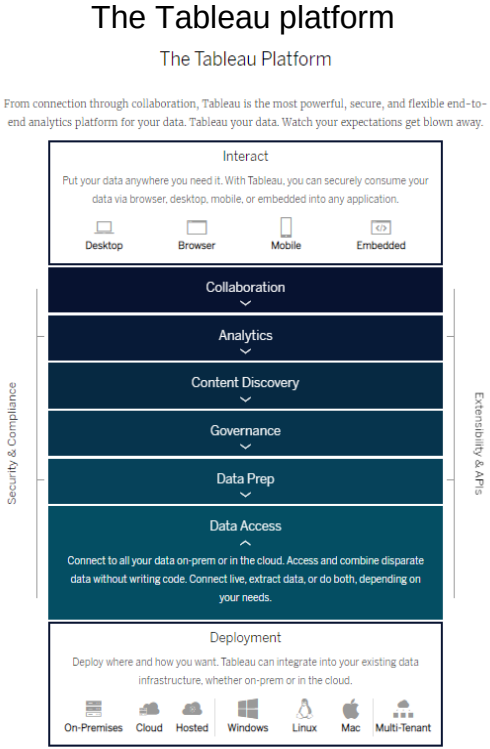


Tableau Prep

Tableau Desktop

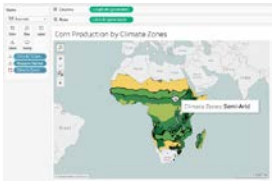


Tableau Developer Tools

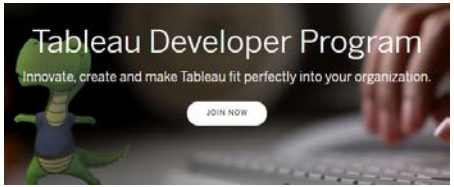
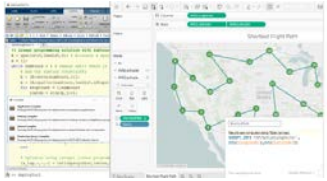


Tableau data/ server management

Tableau Mobile

Data science integration

Integrate and visualize data with coding python, R, matlab



Embedded analytics

Make your Tableau content available anywhere.

Tableau Server

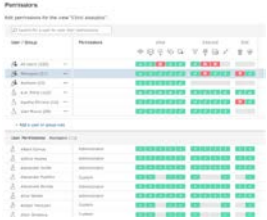


Tableau Online

Automation

Management tasks maintain and update workbooks, data sources, and users on Tableau Server.

Data connectivity

Create connectors to data sources that are not currently supported by Tableau.

What Can We Do with Tableau?

Challenge this later !!

https://www.tableau.com/

The Tableau platform

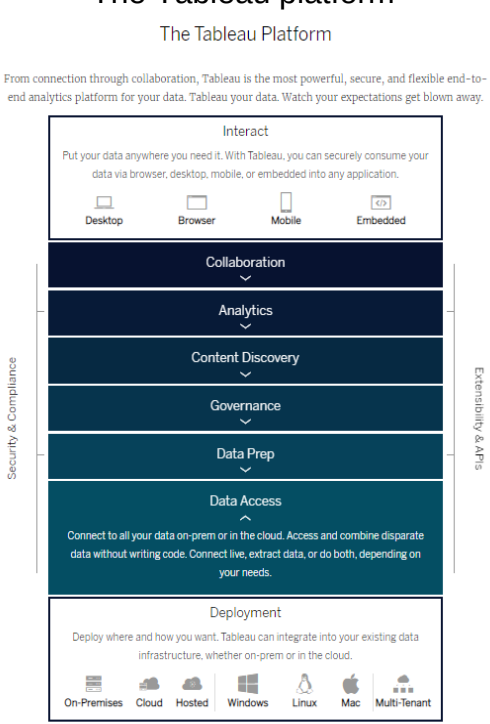


Tableau Prep

Tableau Desktop

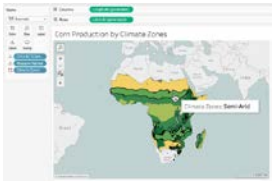


Tableau data/ server management

Tableau Mobile

Tableau Server

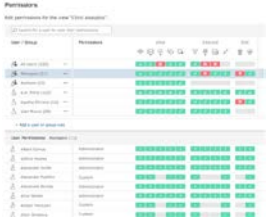
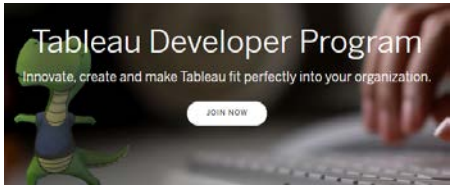
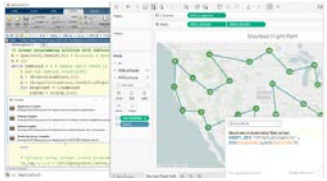


Tableau Online

Tableau Developer Tools



Data science integration
Integrate and visualize data with coding python, R, matlab



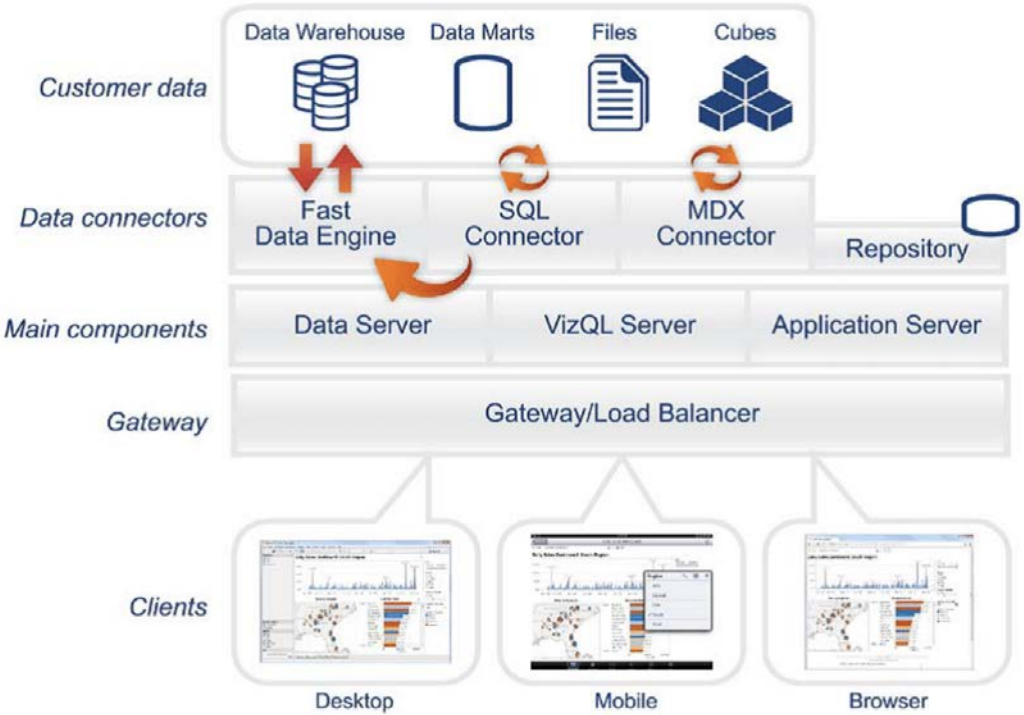
Embedded analytics
Make your Tableau content available anywhere.

Automation
Management tasks maintain and update workbooks, data sources, and users on Tableau Server.

Data connectivity
Create connectors to data sources that are not currently supported by Tableau.

What Can We Do with Tableau?

https://intellipaat.com/blog/tutorial/tableau-tutorial/tableau-server-components/



Let’s Learn the Basic of Tableau

- Sign and Install link
- You will learn basic tutorial for Tableau in this lab
- You will exercise the tableau with COVID - 19 dataset (making a simple corona map) after class

Basic Tutorial of Tableau

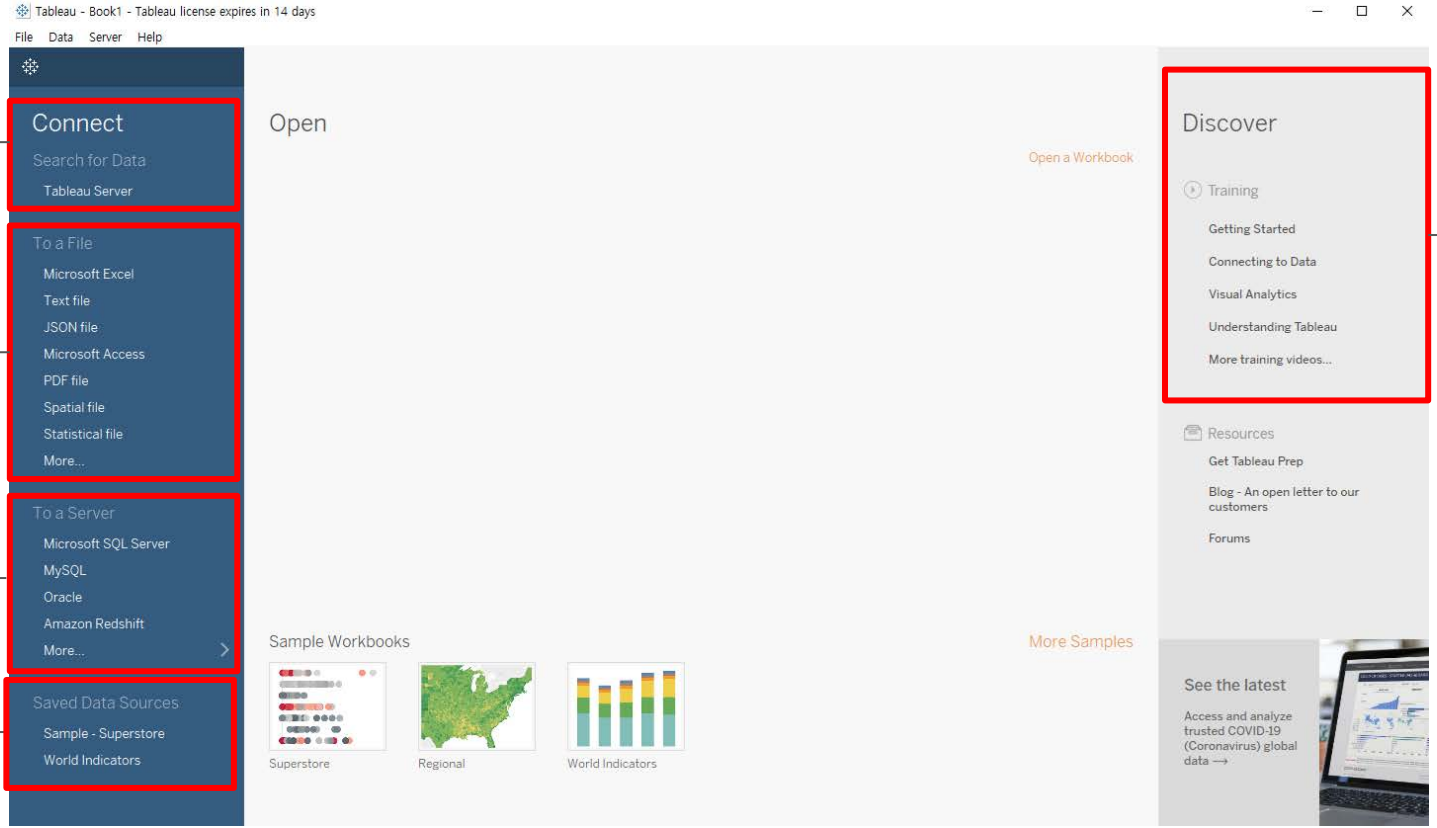
Description of Tableau Start Page

You can search the data in Tableau server

You can load data file from here

You can load data from server

You can load the original saved data file



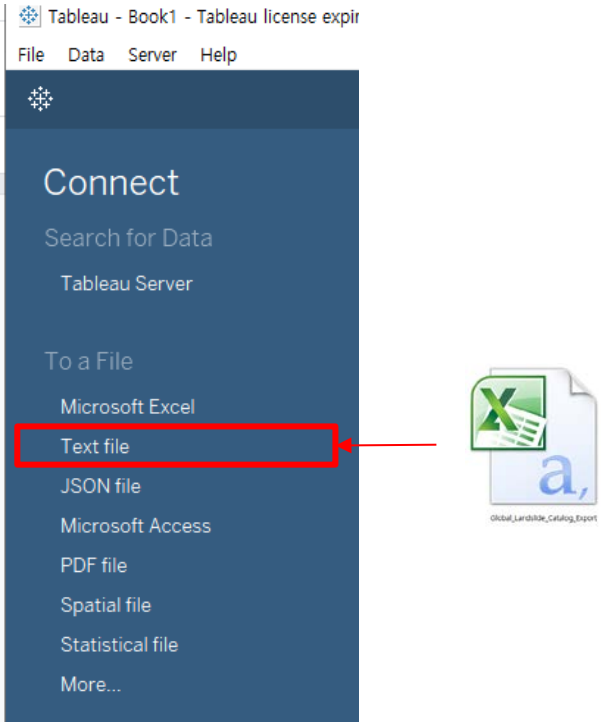
You can watch the basic tutorial education materia for free

Dataset

https://data.nasa.gov/Earth-Science/Global-Landslide-Catalog-Export/dd9e-wu2v

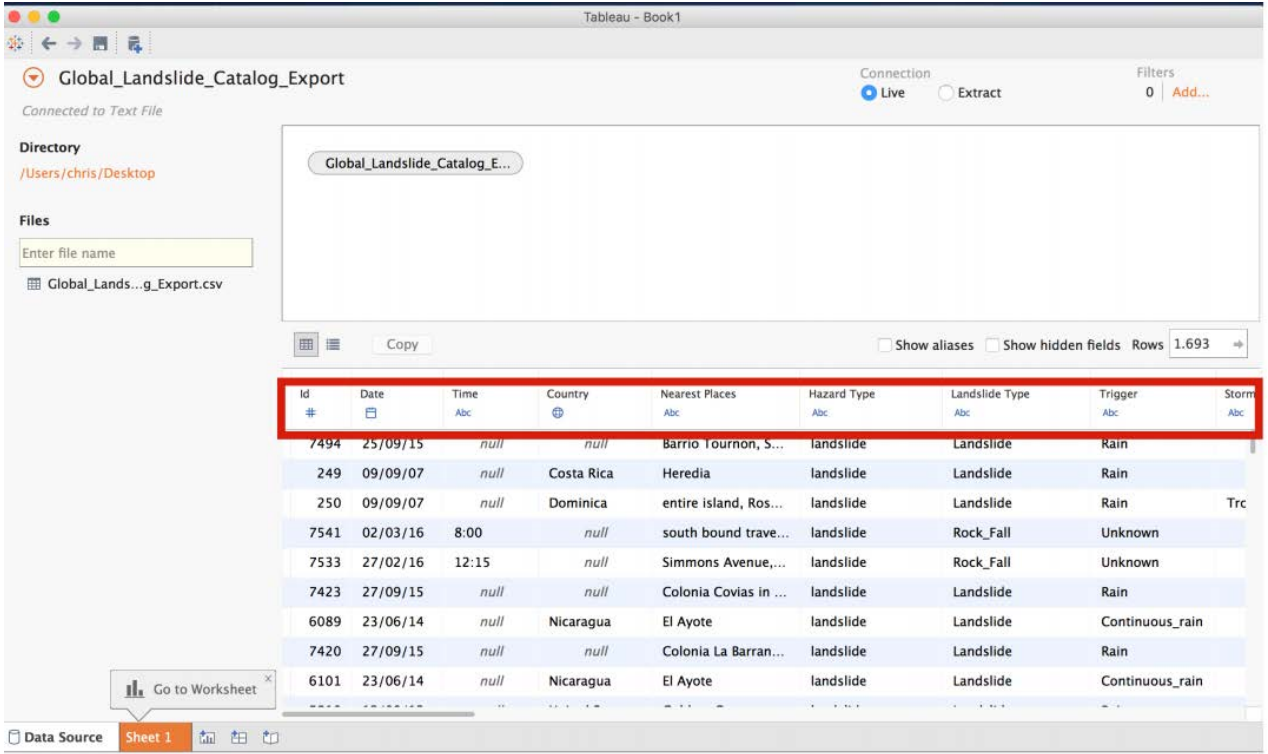
Loading Data

- Download data link: http://vda.univie.ac.at/Teaching/Vis/17w/Tutorials/data/ Global_Landslide_Catalog_Export.csv
- Open the file in Tableau

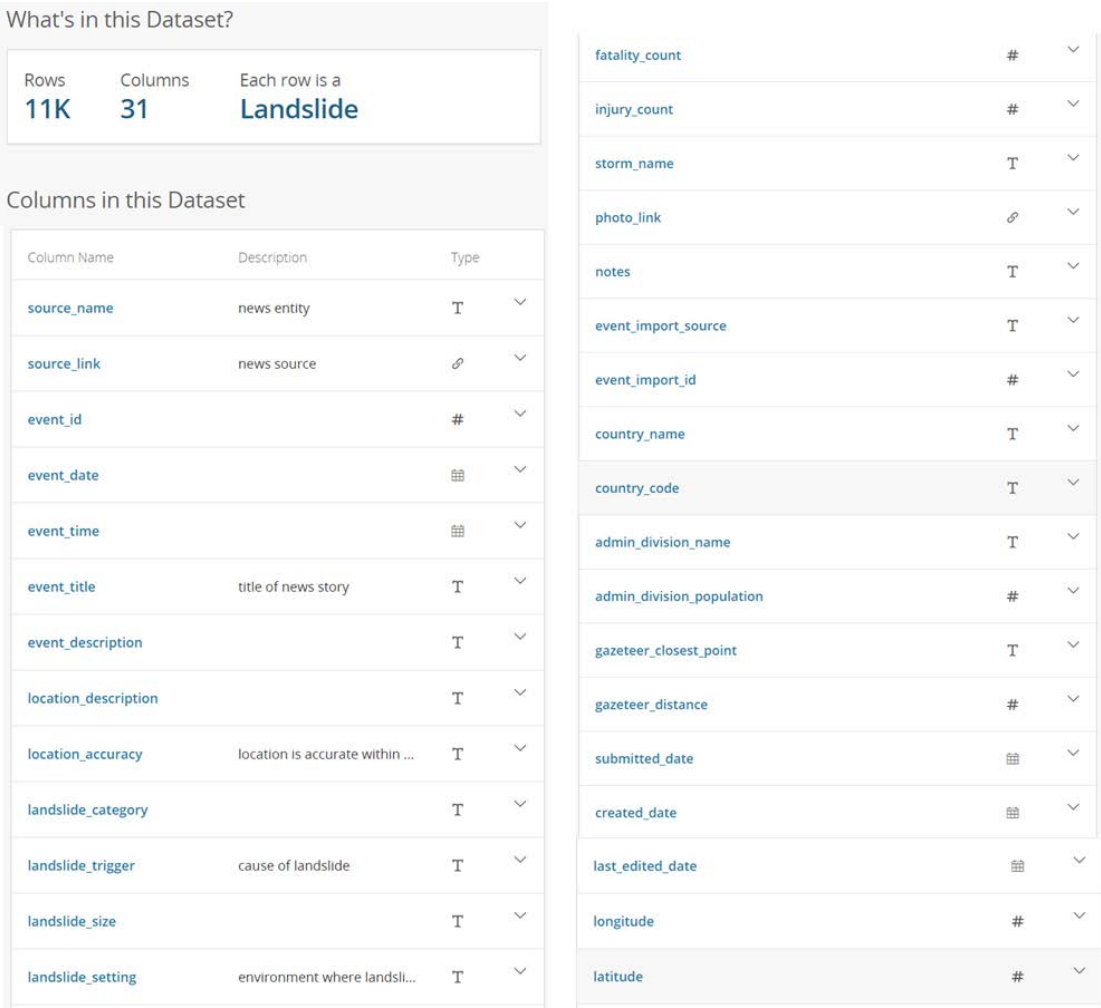


Loading Data

- Always check if the automatic data types are correct by controlling the column headers marked in the figure



Fields



What's in this Dataset?

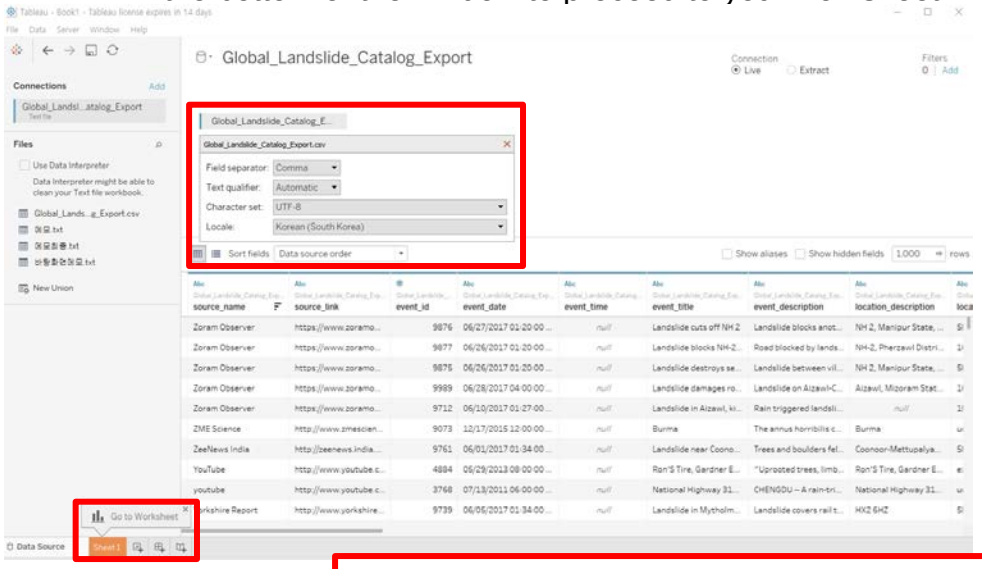
Rows	Columns	Each row is a
11K	31	Landslide

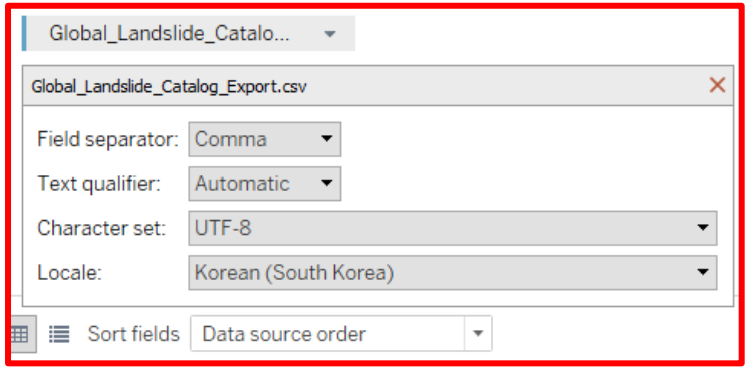
Columns in this Dataset

Column Name	Description	Type
source_name	news entity	T
source_link	news source	🔗
event_id		#
event_date		📅
event_time		📅
event_title	title of news story	T
event_description		T
location_description		T
location_accuracy	location is accurate within ...	T
landslide_category		T
landslide_trigger	cause of landslide	T
landslide_size		T
landslide_setting	environment where landsli...	T
fatality_count		#
injury_count		#
storm_name		T
photo_link		🔗
notes		T
event_import_source		T
event_import_id		#
country_name		T
country_code		T
admin_division_name		T
admin_division_population		#
gazeteer_closest_point		T
gazeteer_distance		#
submitted_date		📅
created_date		📅
last_edited_date		📅
longitude		#
latitude		#

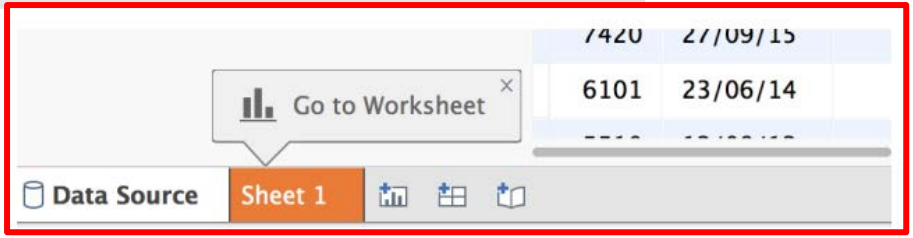
Loading Data

- Mouseover the dataset name and click the little gear to open the settings for the dataset. Choose the correct settings:
- After you have set the correct format and double checked the column headers, the table is useable. Click Sheet 1 at the bottom of the window to proceed to your worksheet.





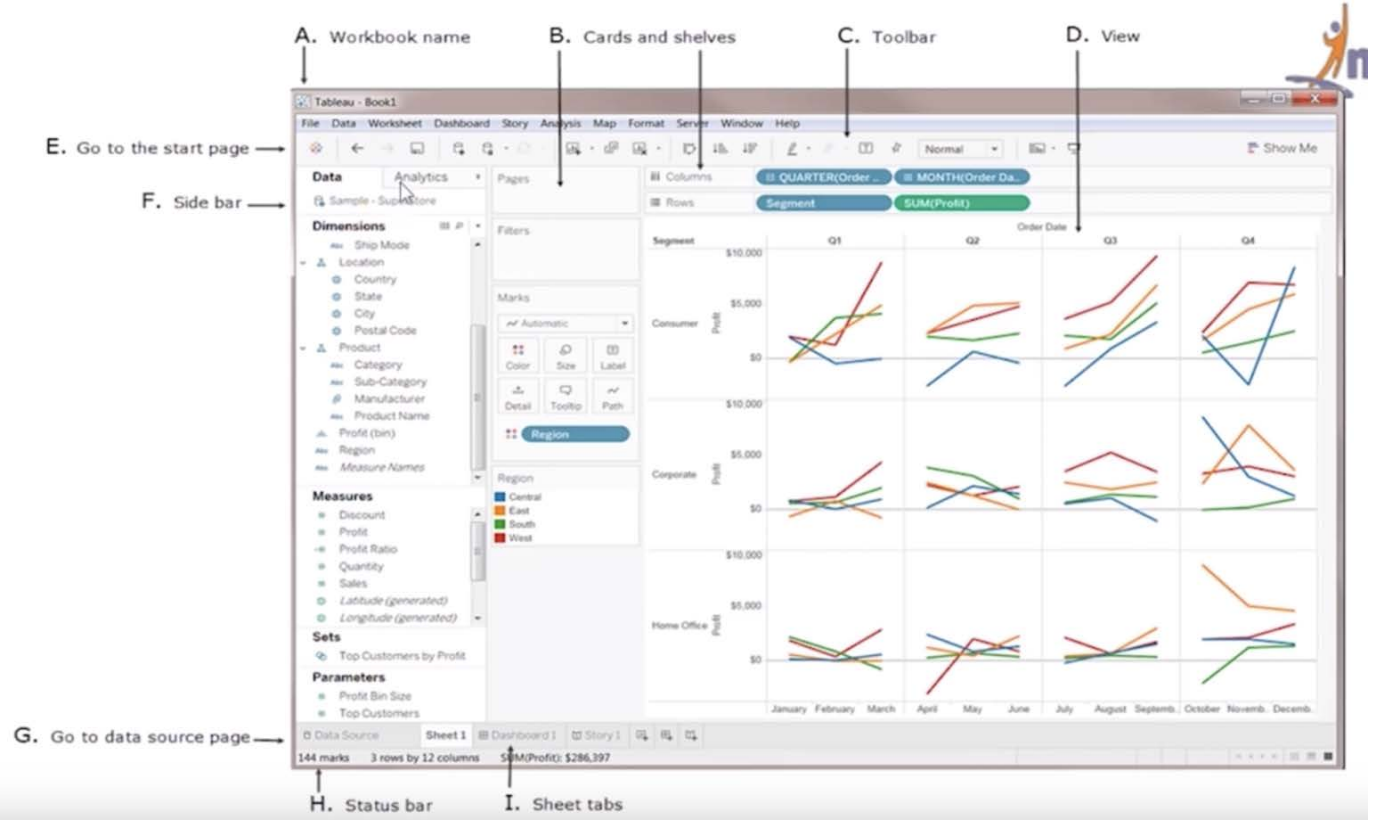
Field Separator: Comma, tab or other
Text Qualifier: How can strings be detected
Character Set: Encoding
Locale: Where is the data from



Description of Tableau Workspace

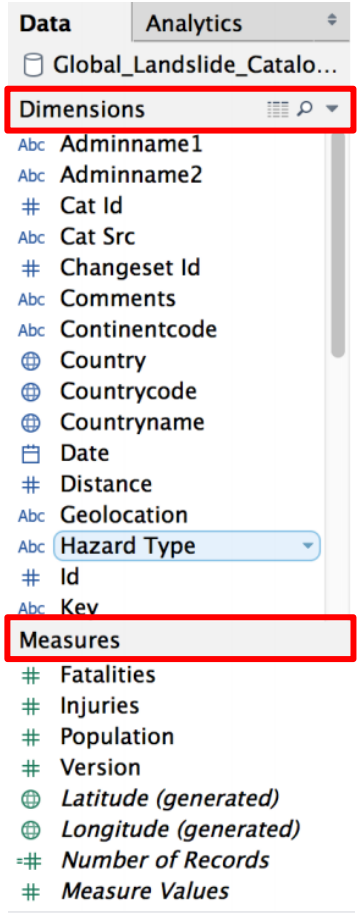
The Tableau workspace consists of menus, a toolbar, the Data plan, cards and shelves, and one or more sheets. Sheets can be worksheets, dashboards, or stories

Worksheets contain shelves and cards that you can drag and drop data fields on to build views



Reference: Free Training Videos Link, Link

Dimension & Measures



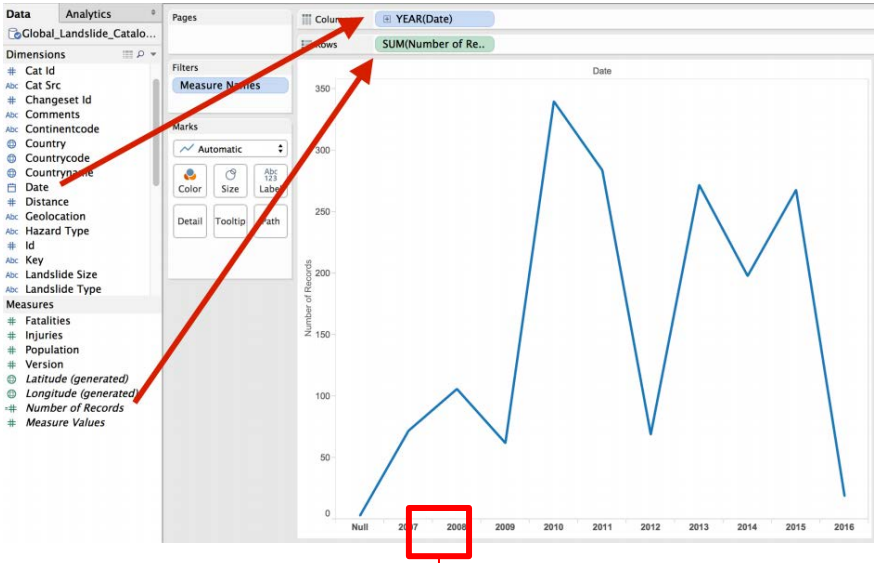
- In the leftmost panel you will see the columns of your table as either dimensions or measures
- You can drag and drop both measures and dimensions to the central panel to create plots
- **Dimensions** contain qualitative values (such as names, dates, or geographical data)

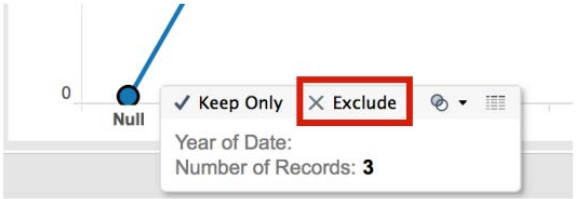
- **Measures** contain numeric, quantitative values that you can measure

https://help.tableau.com/current/pro/desktop/en-us/datafields_typesandroles.htm

Building Charts

- Let’s suppose we want to view the development of the number of landslides over time. Simply **drag** the dimension **Date** to the columns and the measure Number of Records to rows.
- You will see that the records are automatically summed up and the date is set to YEAR. Hence the plot shows the number of landslides per year

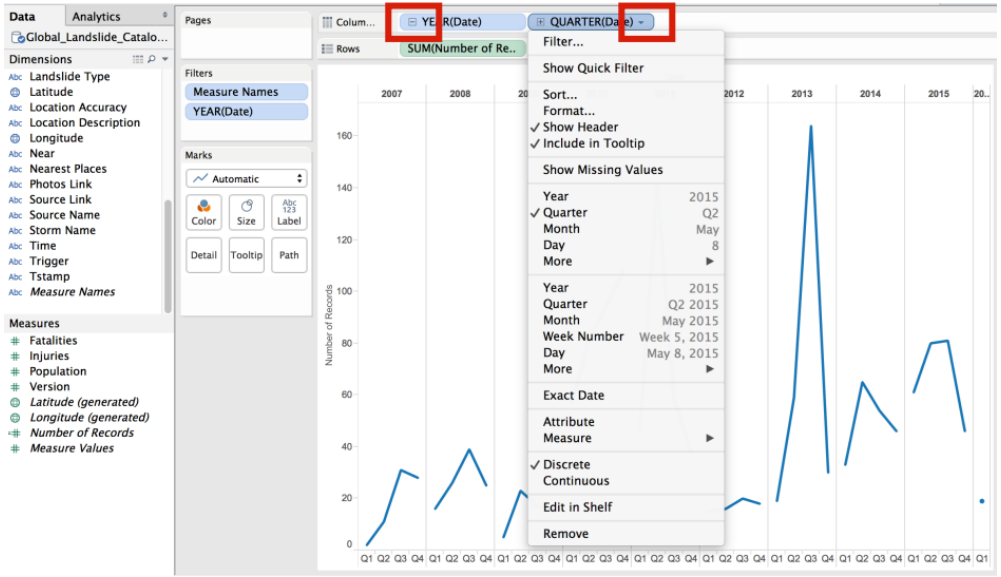


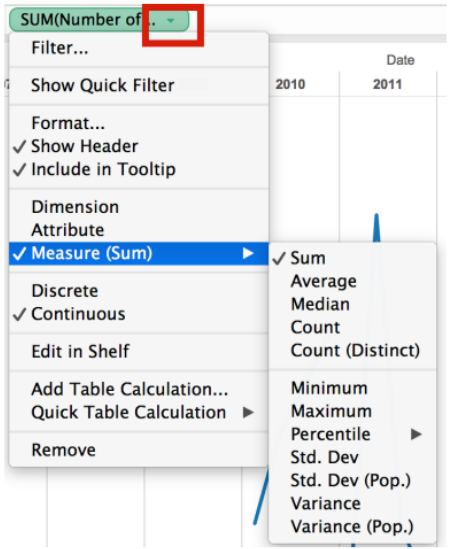


We can see here first data quality problems like **null values**. We can remove them by selecting the point and clicking Exclude.

Building Charts

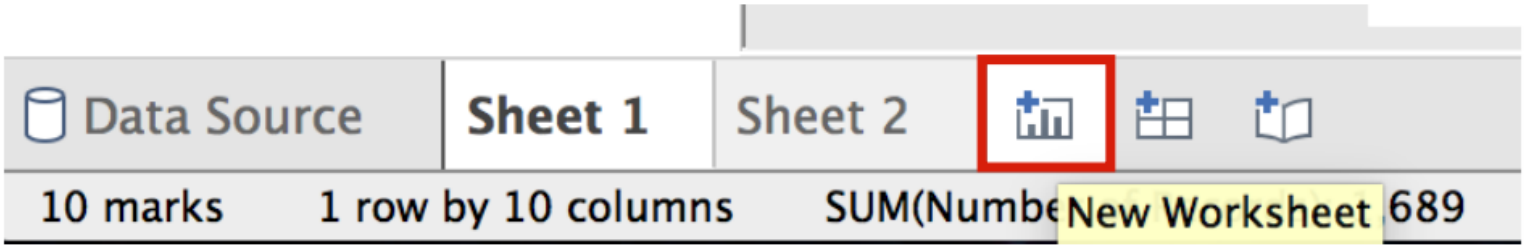
- Click at the + next to **Year(Date)** to get the next smaller dimension.
- By clicking on the small triangle next to a Dimension you can change the **properties of the dimension**.
- You could also **drag** another Date and change it per hand to month. This allows you to stack an arbitrary amount of dimensions.
- The small triangle is available for all datatypes. The measure uses per default **Sum** and you can change it to **Average** or **Variance.**
- This depends on your dataset and the questions you want to answer/things you want to show.





Hands-On

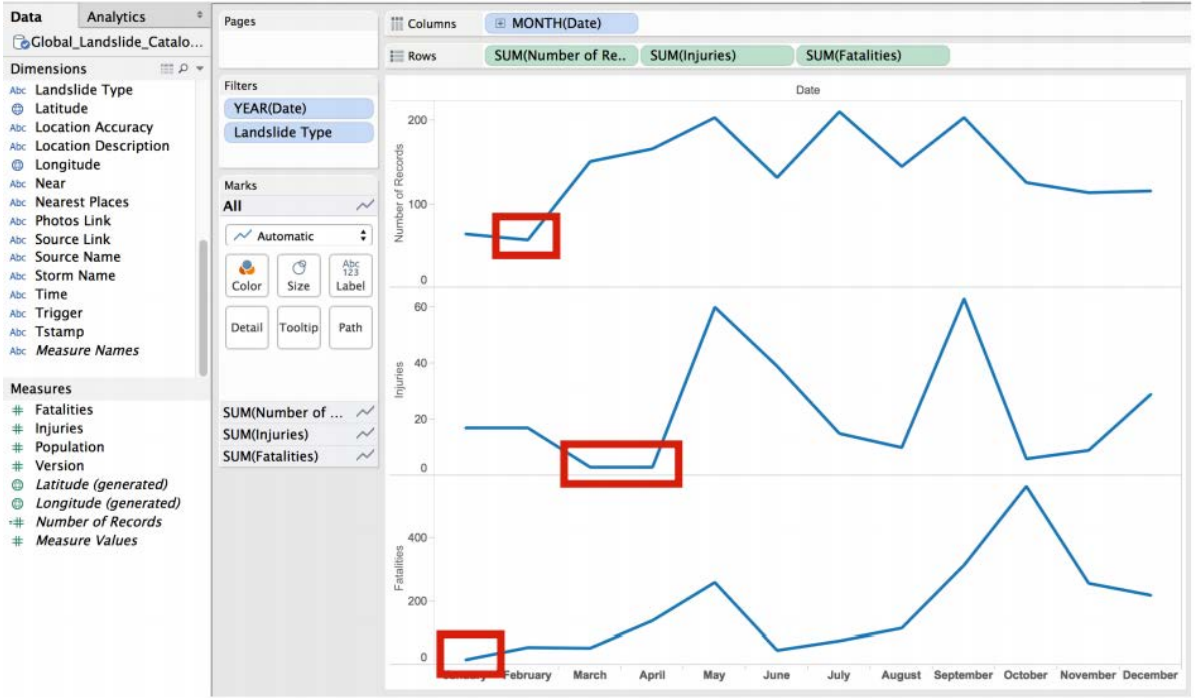
- Question: Which month is the **safest month?** Show (1) the total number of landslides, (2) the number of injured persons and (3) the number of fatalities.
- Create a new sheet by clicking on the **New Worksheet** button and answer the question with help of Tableau:



Hands-On

- One possible solution !

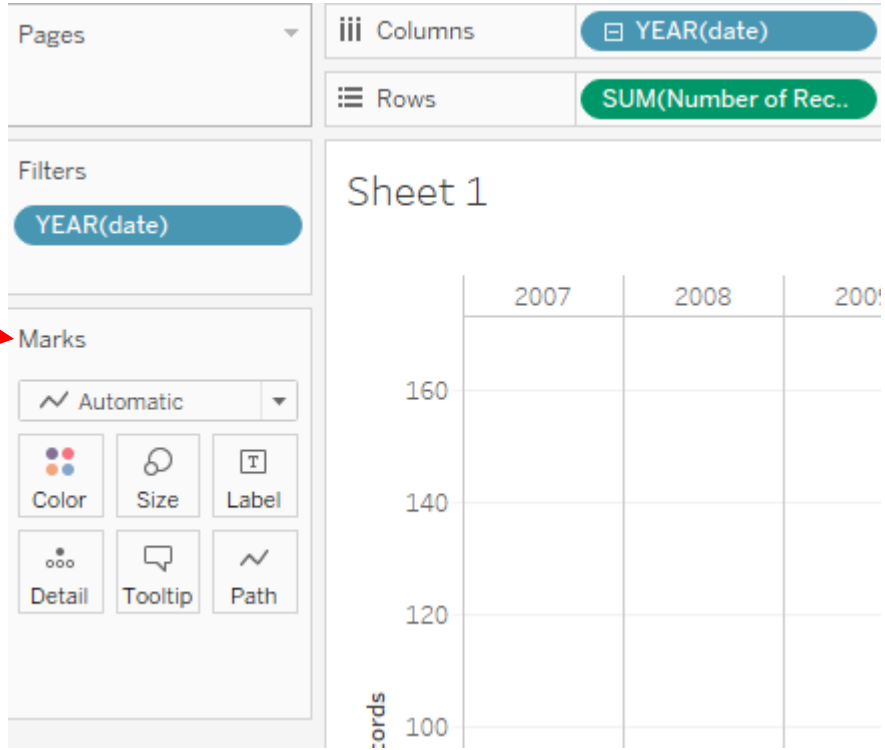
1. From **a total number of events points of view**: February is the safest month.
2. From **a number of injured persons points of view:** March and April are the safest months.
3. From **a number of fatalities point of view**: January is the safest month.



Filters & Marks

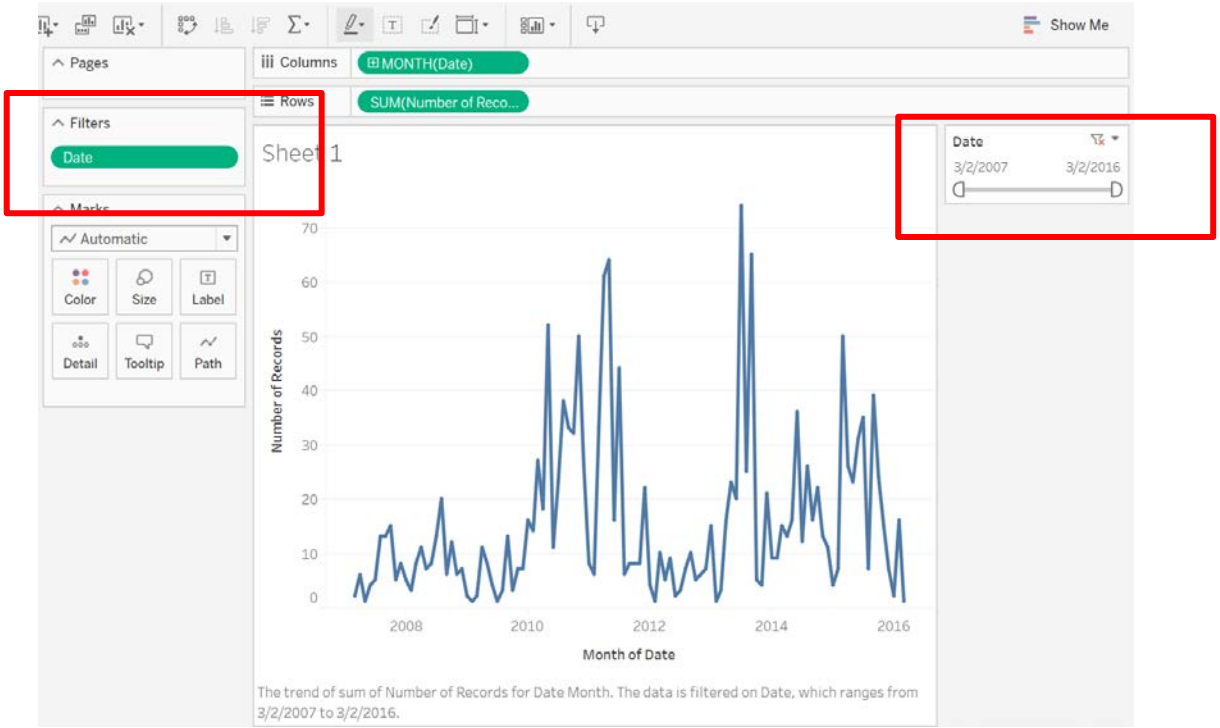
Back to **Sheet 1**.The excluded data is not lost, it is just filtered away. You can always see what filters have been set in the **Filters** panel.

Next let's take a look at the **Marks** panel below it. Here you can modify the appearance of your plot. We want to use marks to encode more information in view.



Filter

- Filter setting: “date” - You can show data within the range specified in “date”



Marks & Appearance

- Let's find out how many small to very large Landslides happened each year:
- Drag **Landslide Size** onto **Color** in the Marks box. We have no multiple lines with weird colors.

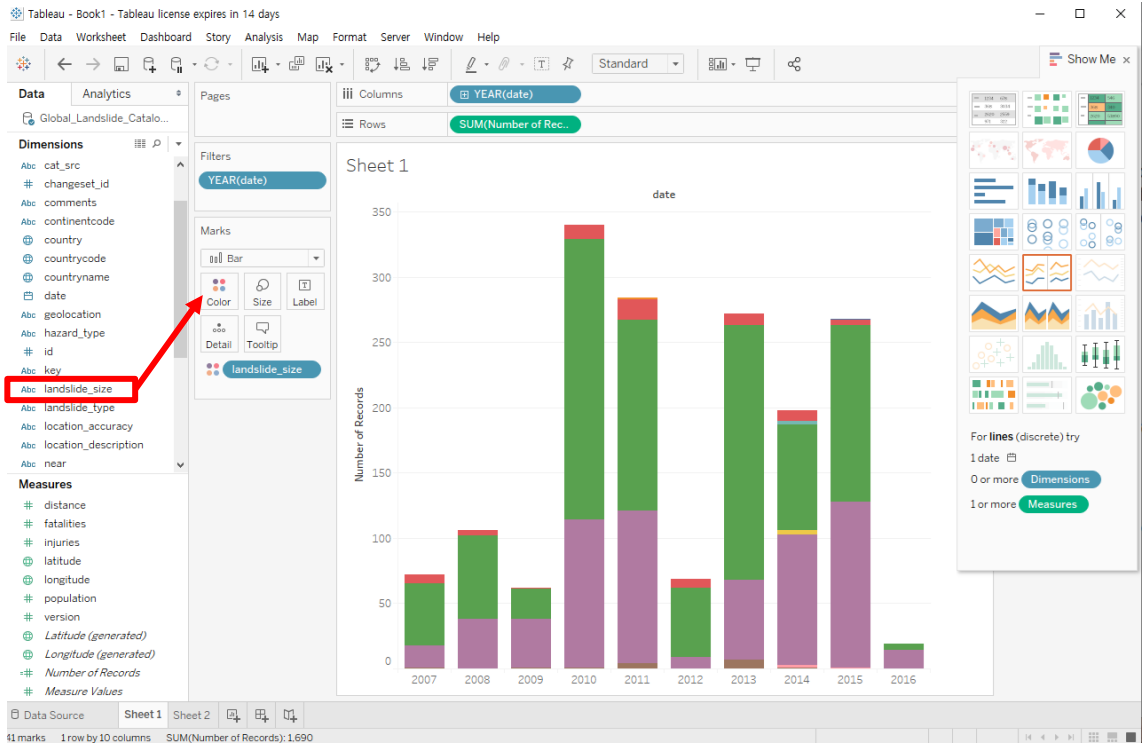
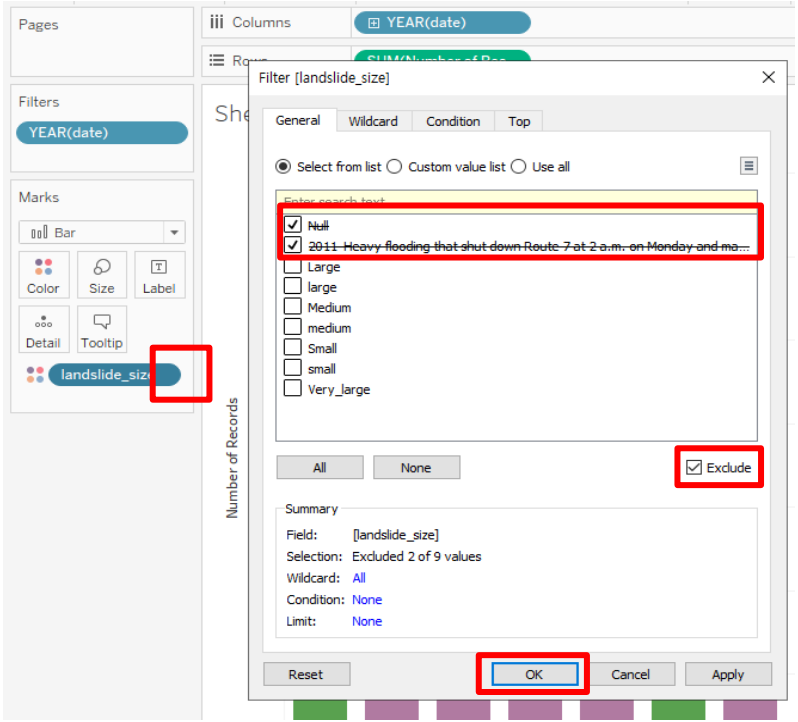


Tableau provides different chart types and by clicking on Show Me in the right upper corner you can choose another one. **Select Stacked Bars.** Your view should look something like this:

Marks & Appearance

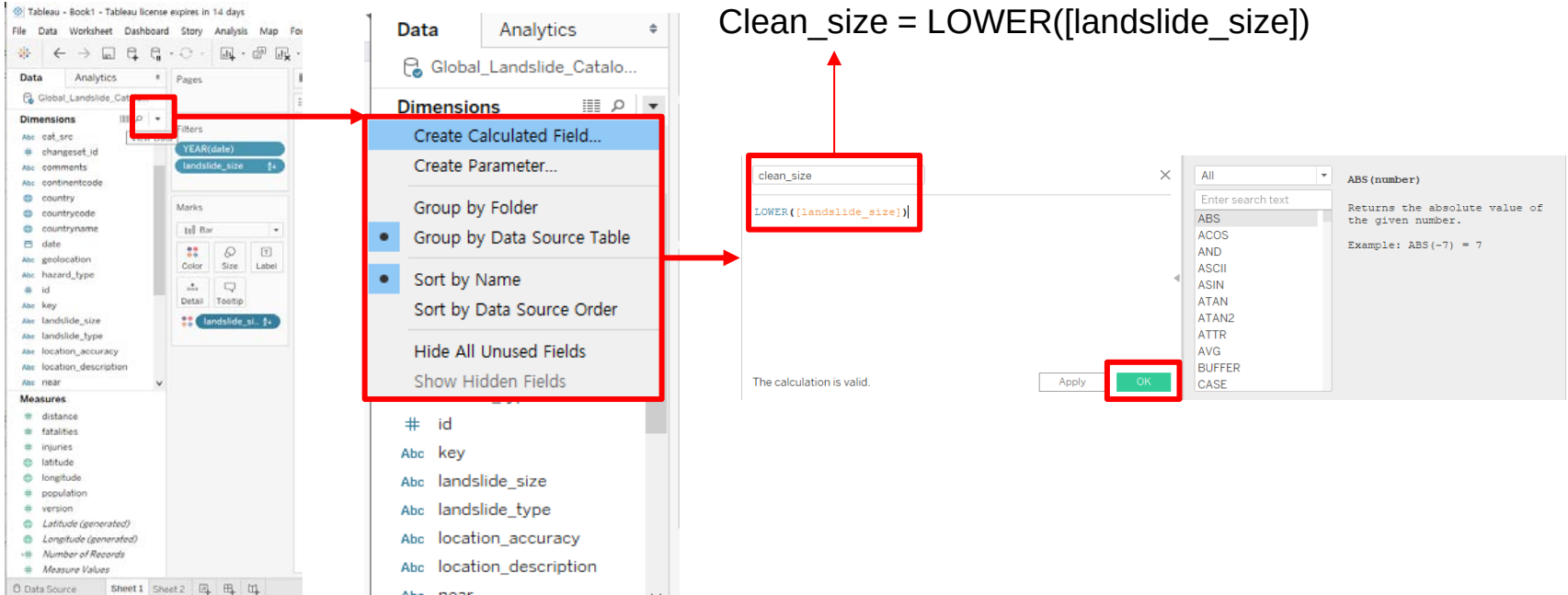
- The default color encoding is actually not that good. Let’s change it!
- Exclude **Null** and the **2011 Heavy flooding** by right clicking them and select **Exclude.**
- **Reorder** the entries by **drag and drop.**

- We can see here **another data quality issue**: upper and lower case answers (e.g. 'Large' and 'large') should not be different groupings



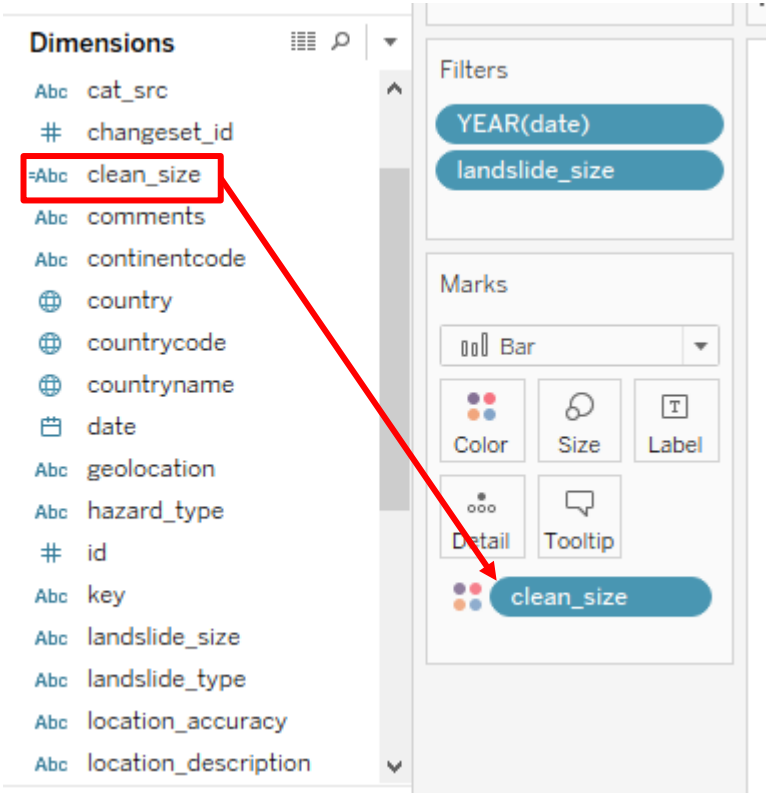
Marks & Appearance

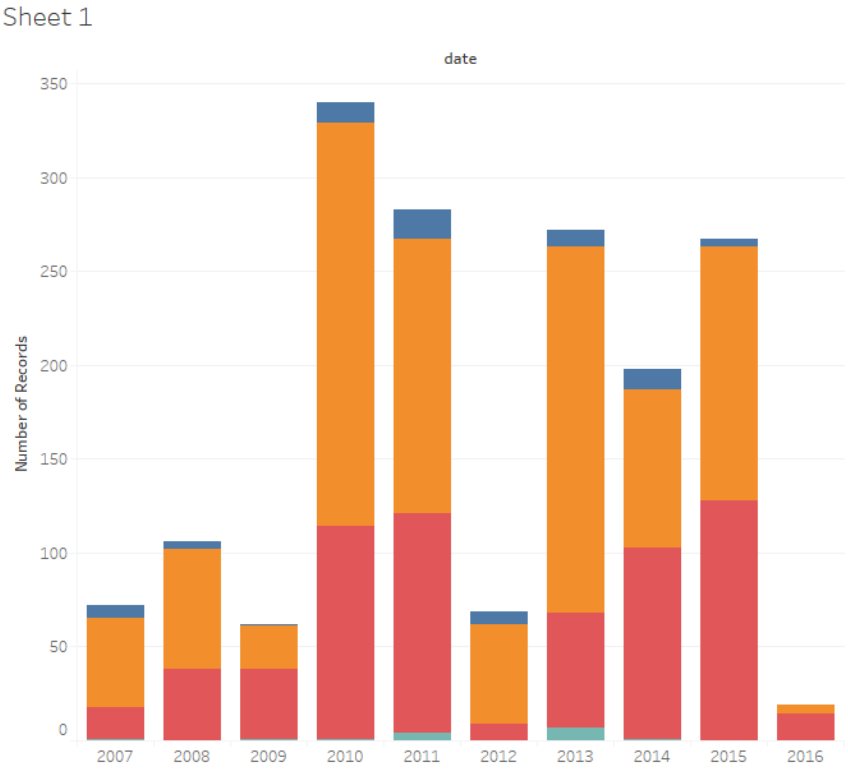
- We can fix this by creating a calculated field
- To do so, in the Data Panel,
  - Click the triangle right next to **Dimensions**
  - Click **Create Calculated Field**
  - Create a new measure “**clean_size**” using the following calculation: LOWER([landslide_size])



Marks & Appearance

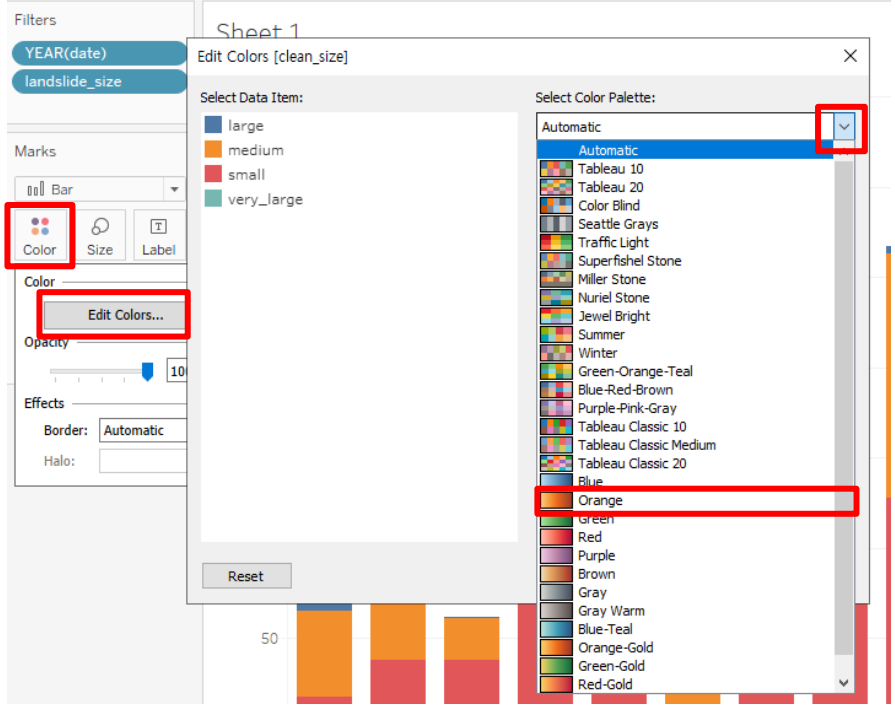
- Drag **clean_size** new dimension onto **Color.** You should see now something like this:

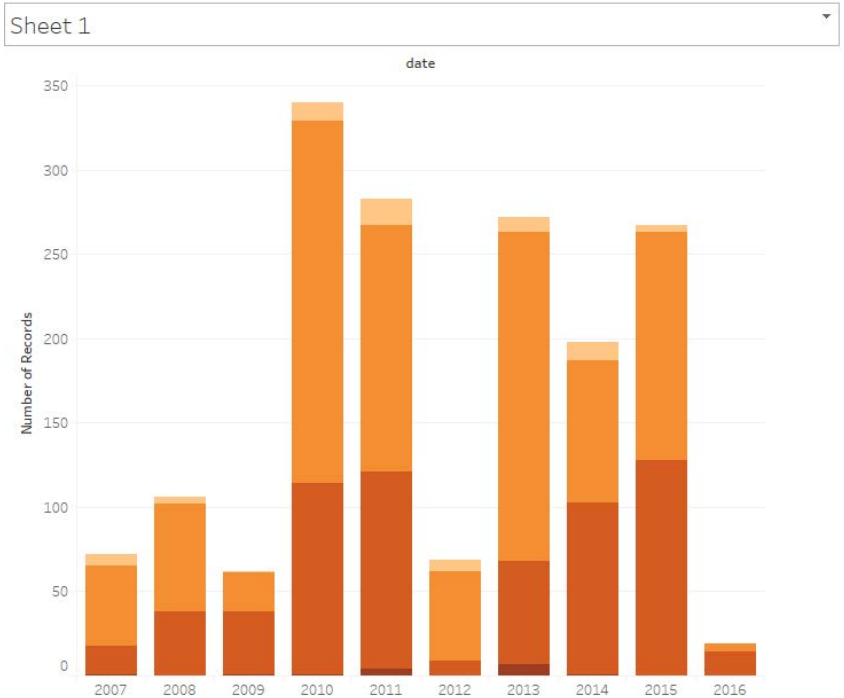




Marks & Appearance

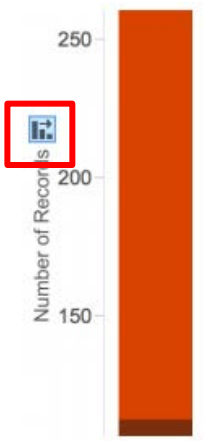
- Click the **Color or Edit Colors** and the small triangle which is visibly when you mouse over and select **Edit Colors**
- Select **Orange -> Assign Palette -> Apply -> OK**
- It should look like this now! We can see that most of the Landslides are small or medium.

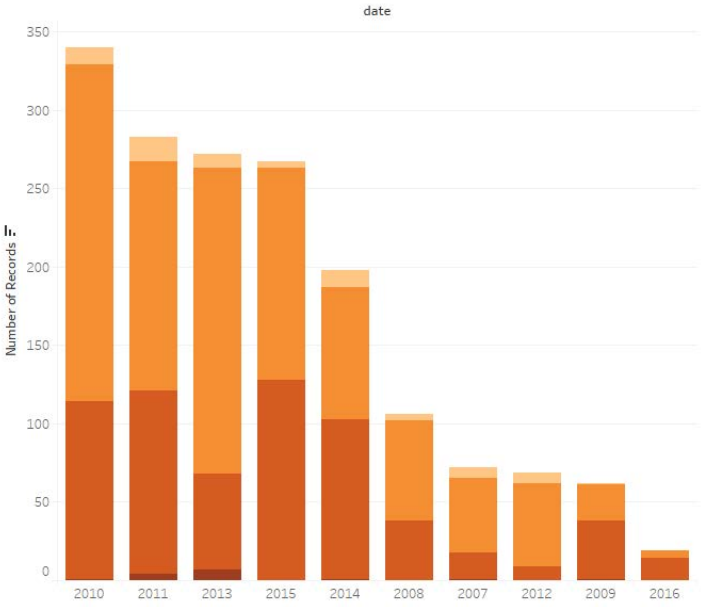




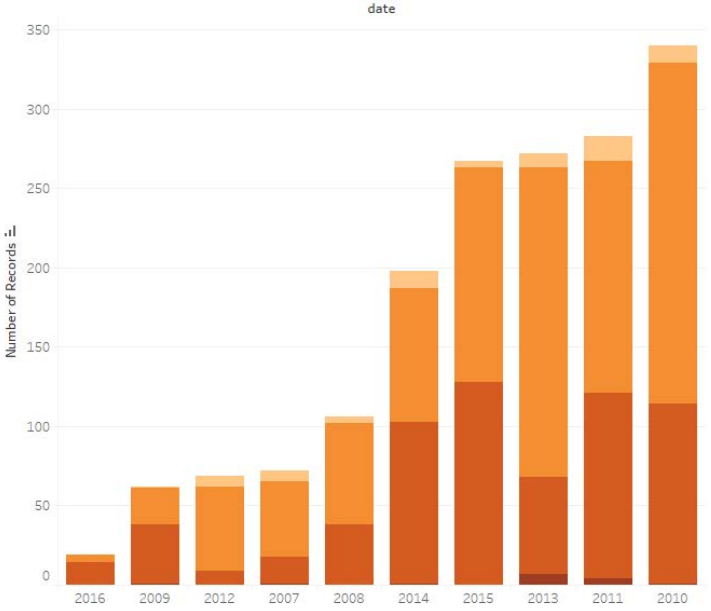
Marks & Appearance

- You can **sort** them now by clicking onto the small sort symbol right above the axis description.





Sort descending



Sort ascending

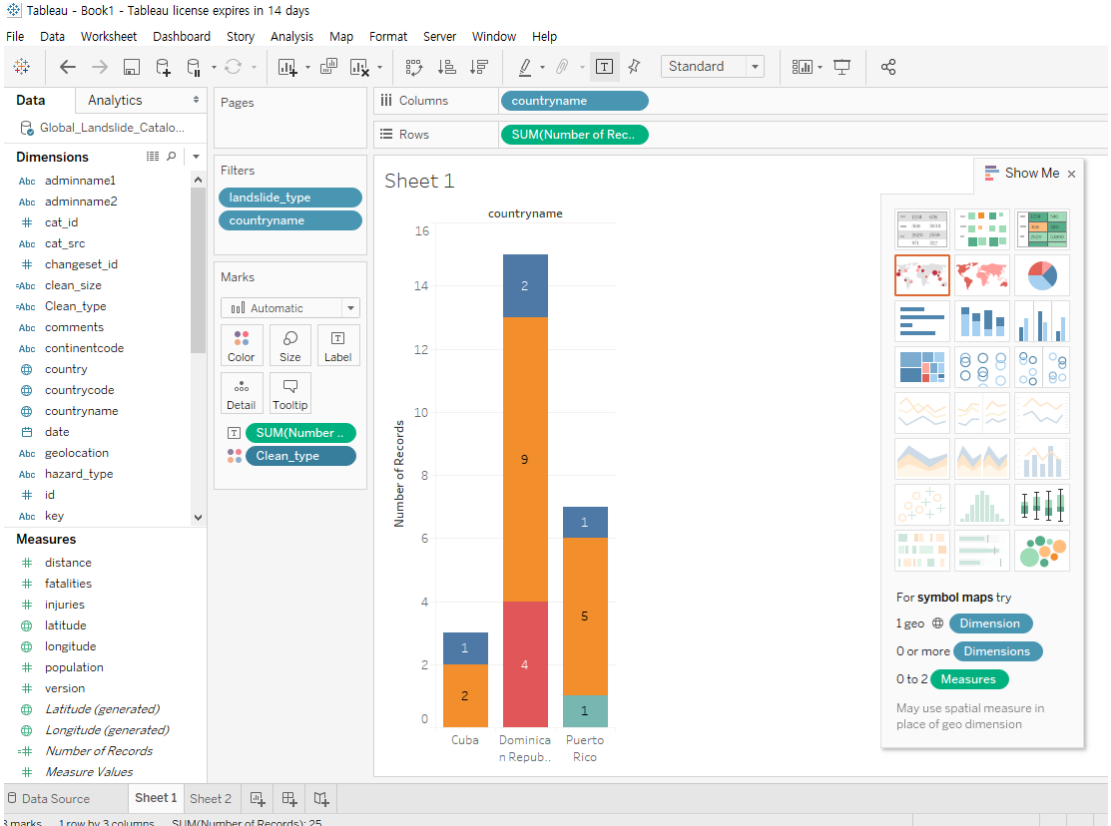
Hands-on

- Question: Which landslide types and how many of them happen in the Dominican Republic, Cuba and Puerto Rico?

  Tips: Use Filters (hint - Quickfilter) and check the dimensions data quality

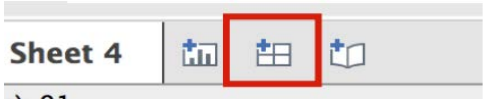
Hands-on

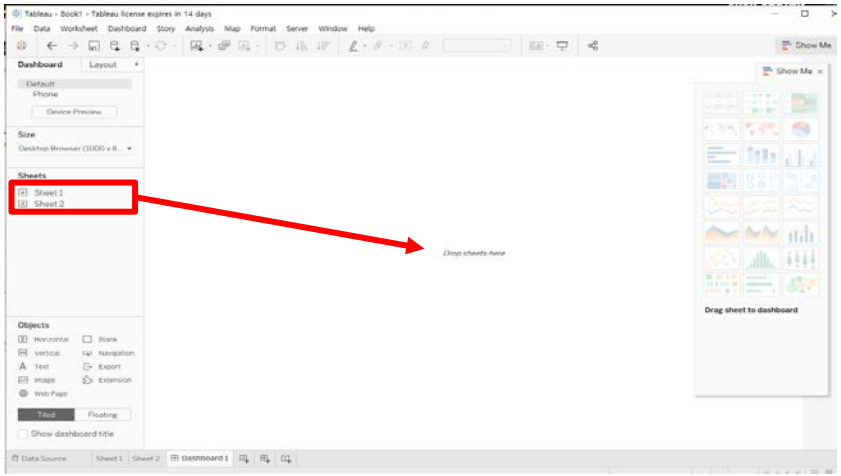
- One possible solution:

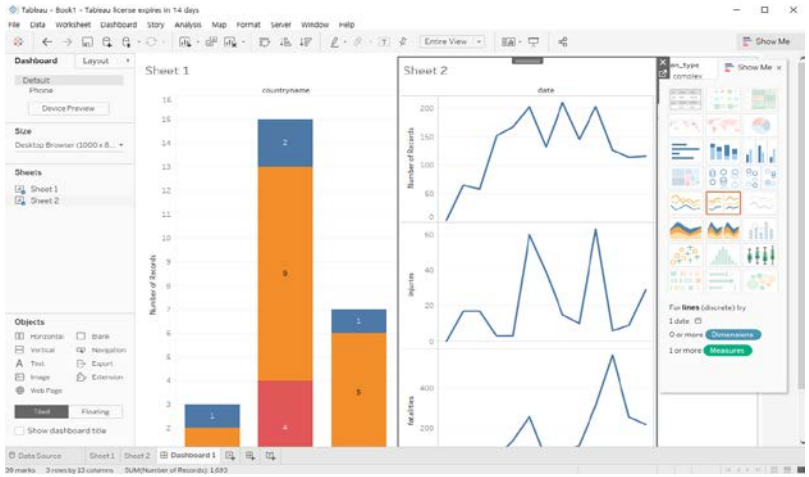


Dashboard

- A Dashboard connects multiple single views with each other.
- This allows for more complex questions to be analysed.
- To create a dashboard click the **new Dashboard button** at the bottom.
- Drag Sheets from the left side into the middle area:

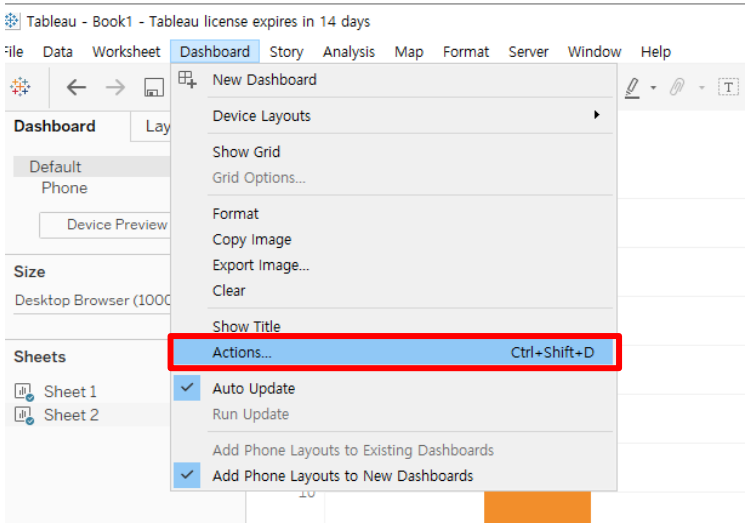


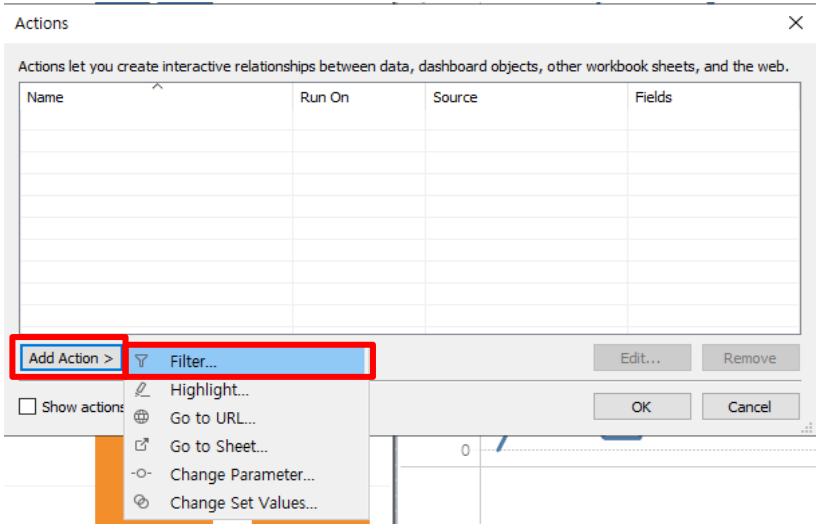




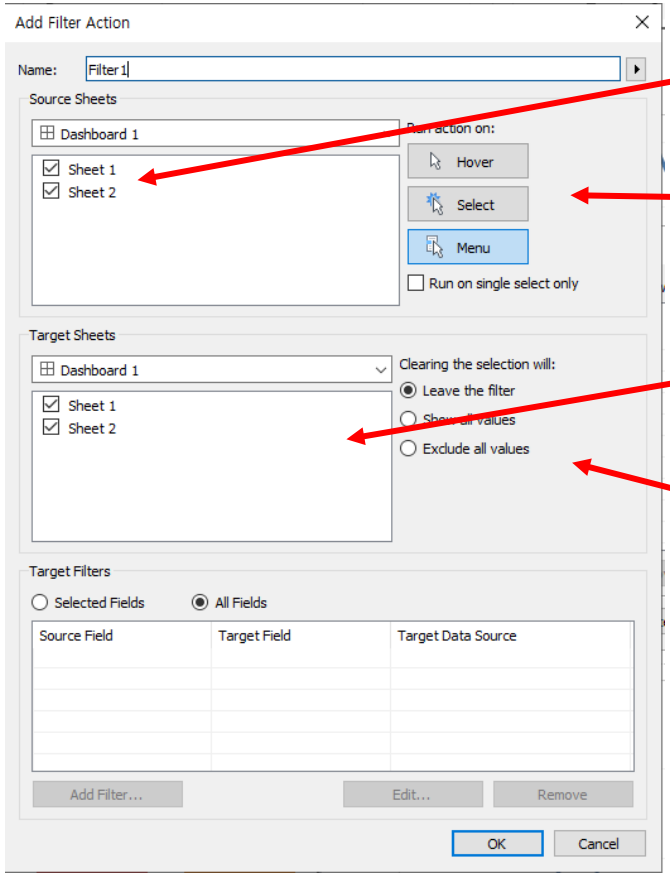
Dashboard

- You can now define Dashboard wide actions like Filters and Highlights. We will add a filter so that can select in one view data and the second view will be updated accordingly. In the Dashboard, add a new Action by selecting Dashboard -> **Actions…**
- Add a new action with **Add Action** > and select **Filter…**





Dashboard



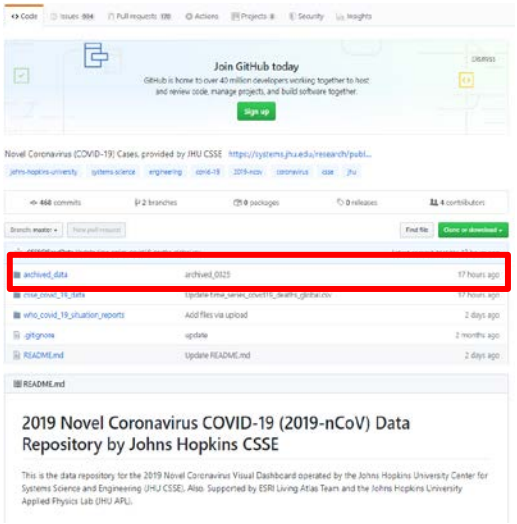
- **Source Sheets** define the sheet where the filter action is recognised. In our case is Sheet 1 the source.
- **Run action on** let you select how the filtering should happen. In our case by selection.
- **Target Sheets** define the sheets where the filter will be applied
- **Clearing the selection** will provides option what happens when you clear the selection.

After Class Practice: Make the Conoronamap with Tableau

- **If you finish and upload (file name: “Covid.twb”) this exercise in the google classroom**
- **You will get extra participant score!!**
- **This task is not mandatory.**

Description of COVID-19 Dataset

- As the COVID-19 dataset, the dataset that was preprocessed (aggregation and cleaning) of the Johns hopkins dataset that provides latitude, longitud number of confirmed, death, and recover information for the world's state, province, and nationality by date is used. Dataset link
- Data preprocessing process will be practiced later in data visualization with python lecture
- Original Johns Hopkins dataset reference: Johns Hopkins dataset download link -> dataset(lasted update 03/29)



We will use this data for the practice

Confirmed

Death

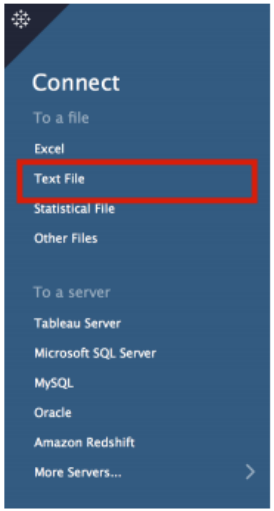
Recovered

Data preprocessing (aggregation and cleaning)

Aggregated data (Confirmed, Death, Recovered)

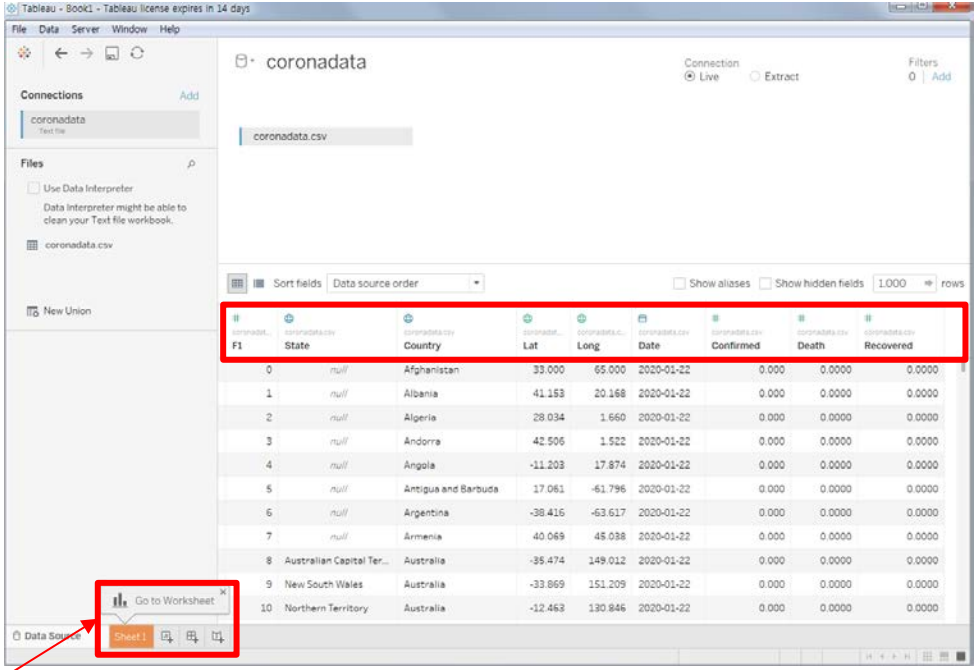
Exercise the Tableau with COVID-19 Dataset

Loading Data



The first step to the visualization is to load the data. On the left side of the screen in the section **To a File** choose **Text File**. This way you can open the .csv file.

After you have selected the data file, you will see an initial table representation of your data. For the data entries to appear, you might have to click **Update Now**.

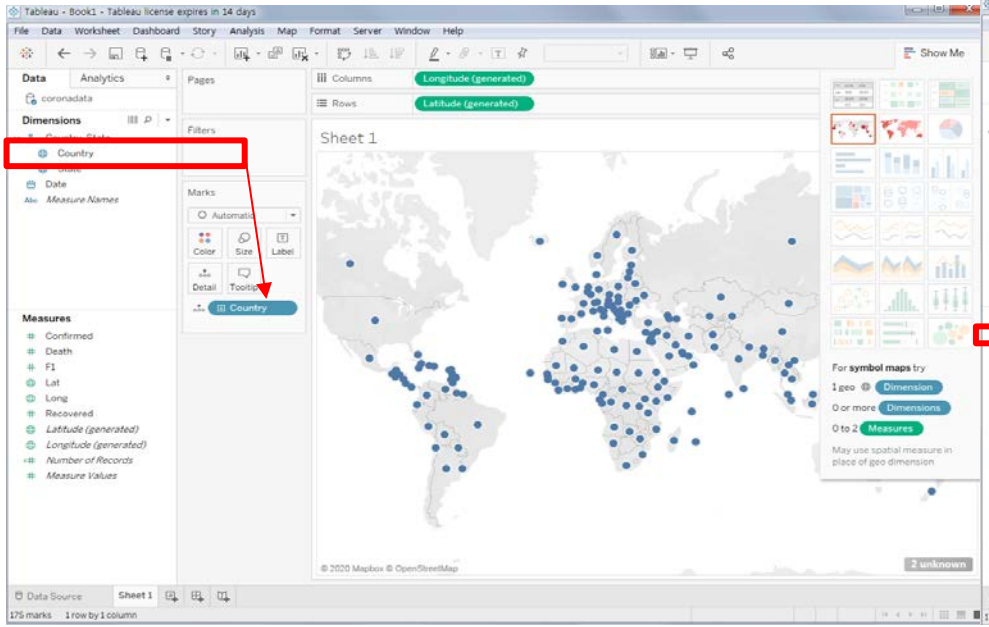


Click the worksheet

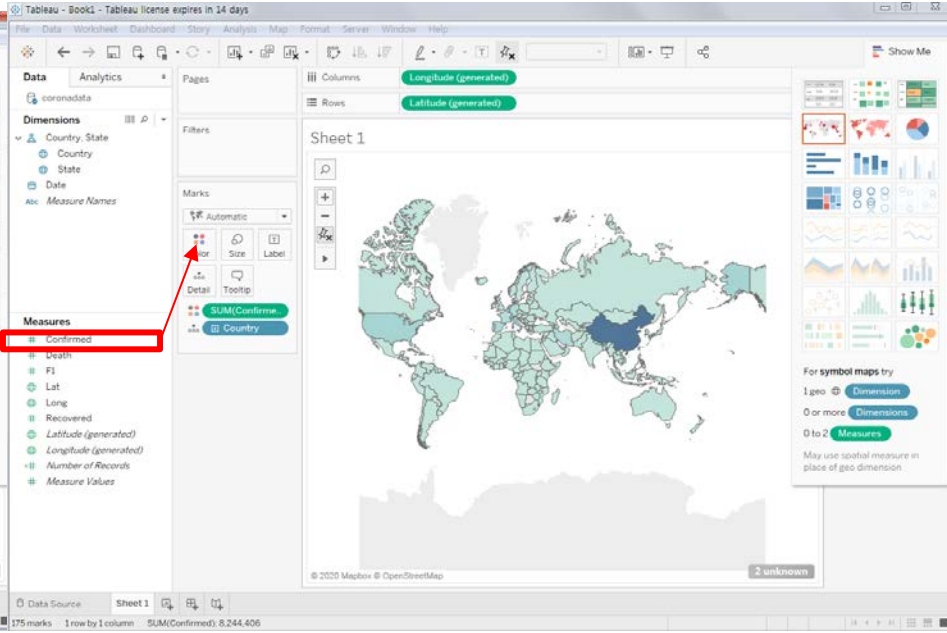
Tableau evaluates automatically what data type is in each column. Always check if the automatic data types are correct by controlling the column headers marked in the figure above.

Exercise the Tableau with COVID-19 Dataset

Drag the country in the “Dimensions” of the sidebar and drop it on the “Marks” to mark each country

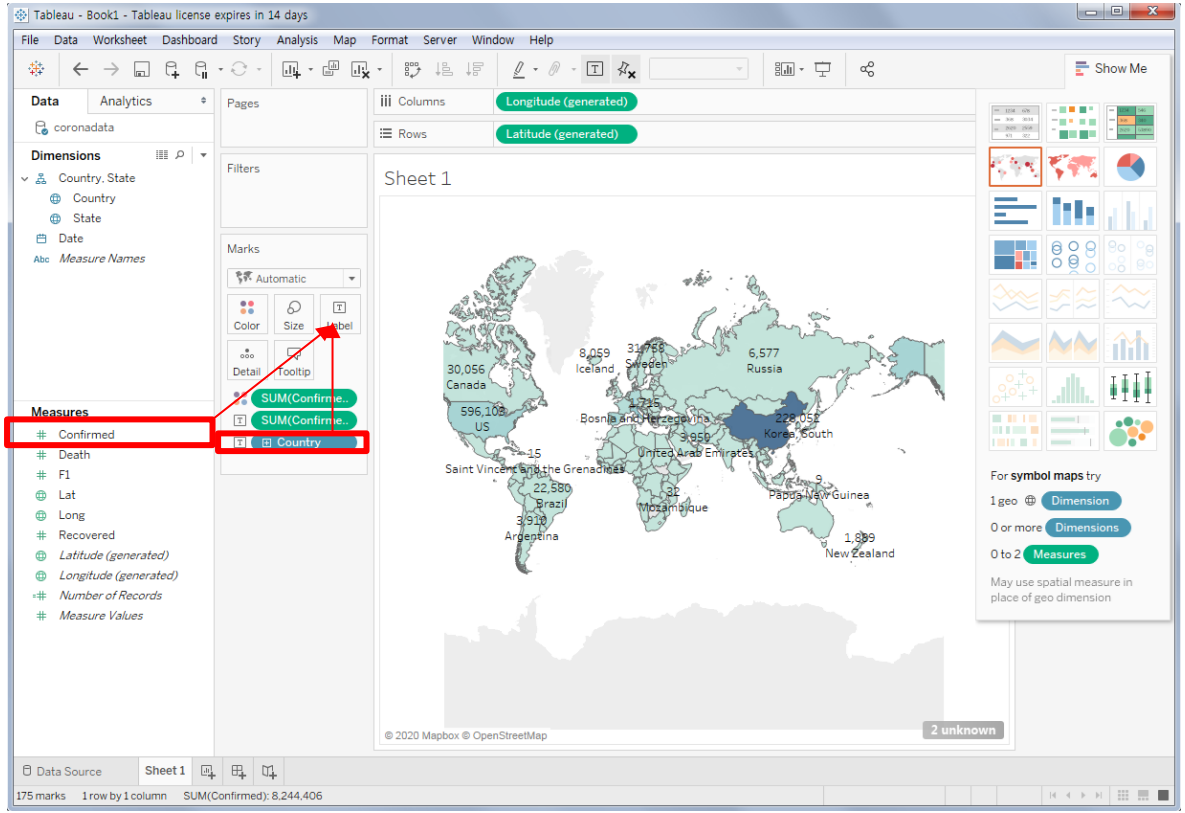


Drag Sum of Confirmed and drop it on the “Color of Marks”. Depending on the number of total confirmed people, the intensity of the color (light to dark) is different for each country.



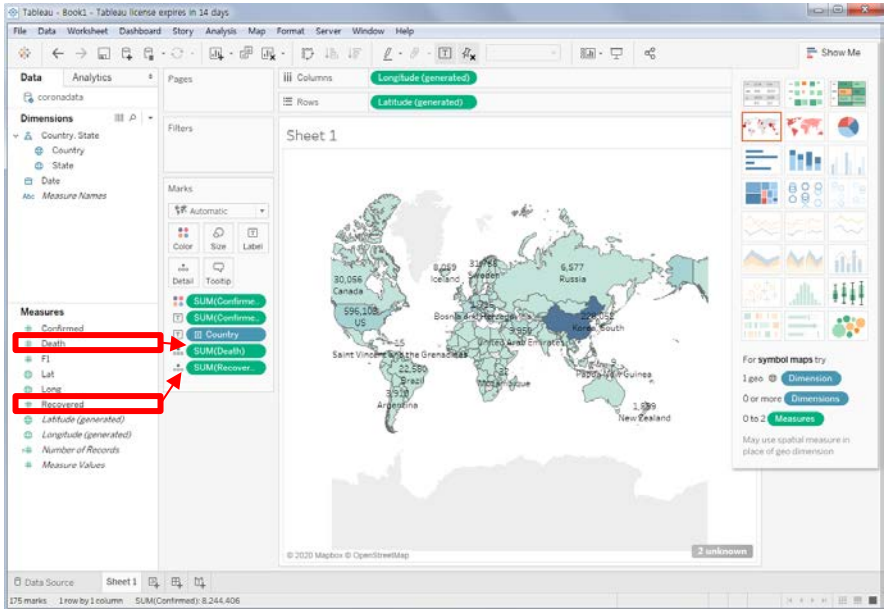
Exercise the Tableau with COVID-19 Dataset

Drag and drop “Country in Measure” and “Sum of Confirmed in Marks” to the “Label in Marks”, then we can label the country name and number of confirmed in corona map.

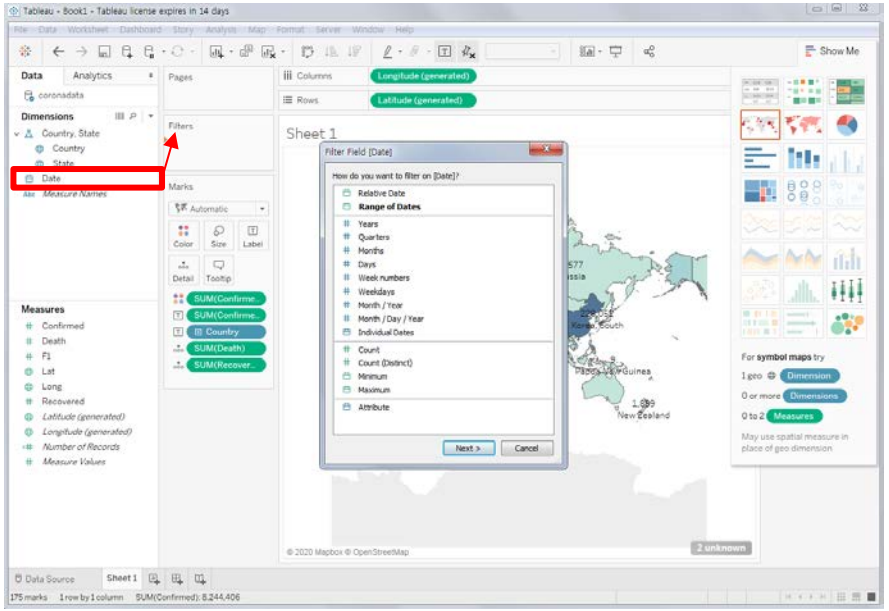


Exercise the Tableau with COVID-19 Dataset

We also add “sum of Death” and “sum of Recovered” in Marks



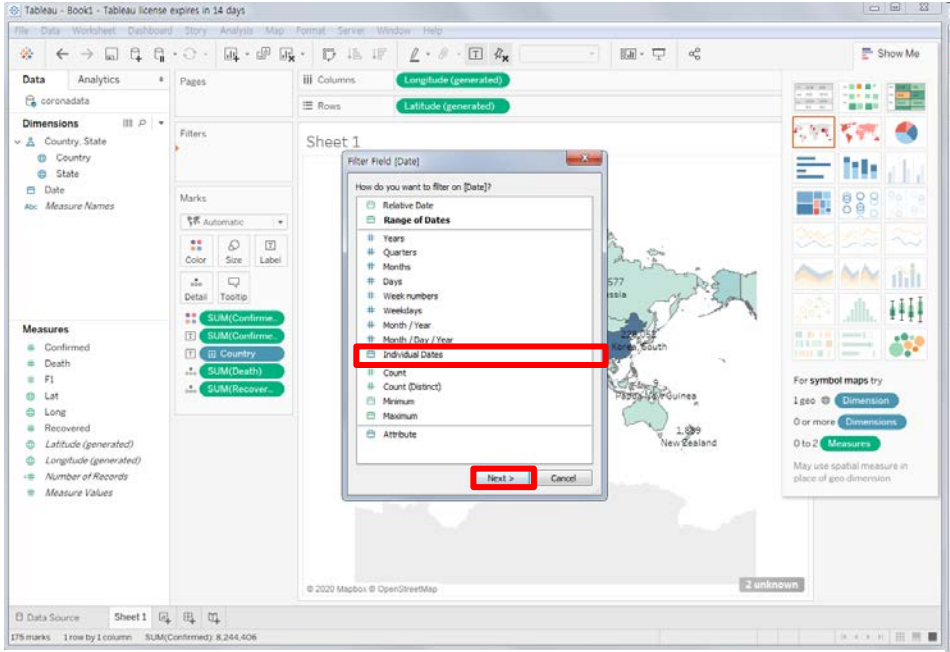
The current map shows the total number of confirmed for the entire date. If you want to see the number of confirmed, death, and recover for a specific date period or country, you can use the filter function.



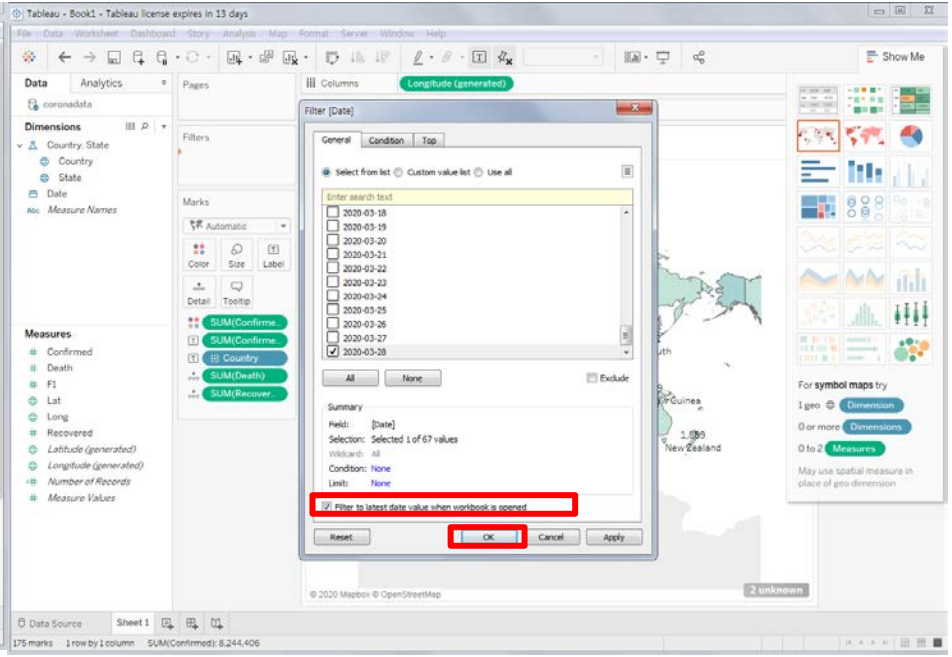
As an example, let's look at the number of confirmers by country for a specific date. Just put the “date in Dimensions” into the “Filters.”

Exercise the Tableau with COVID-19 Dataset

Click “Individual Dates” in the pop-up window of Filter Field [Date] to see the number of confirmers according to the date.

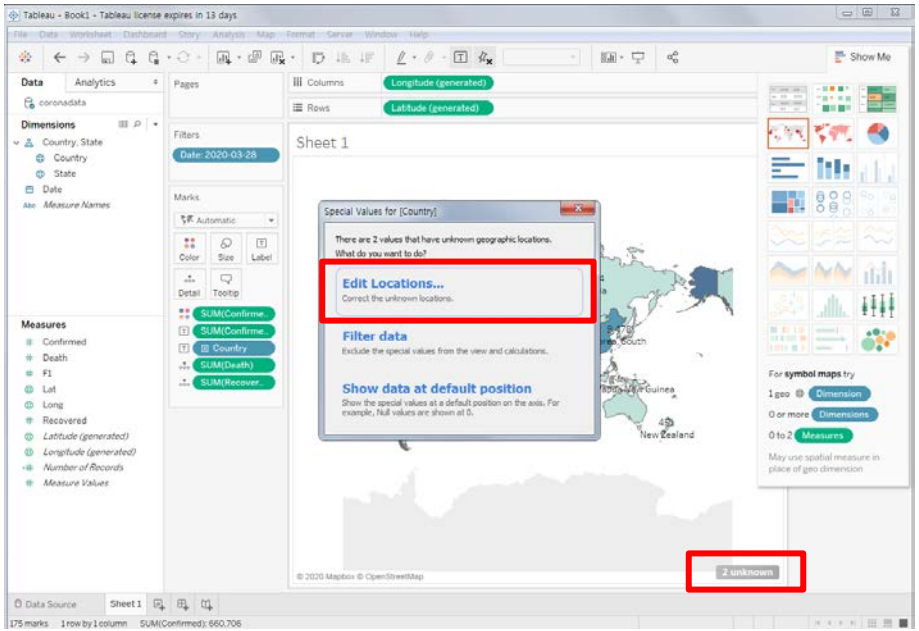


Click “Filter to latest date value when workbook is opened” at the bottom to see the number of confirmed for the most recent date of data, the last date of March 28, 2020.

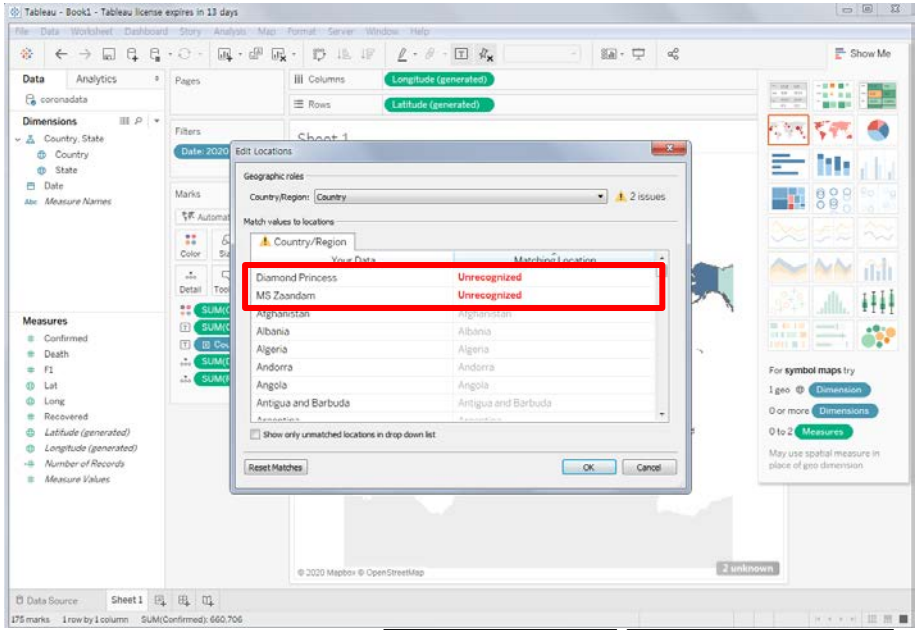


Exercise the Tableau with COVID-19 Dataset

There are two values that have unknown geographic locations, so click “2 unknown” and click the “edit locations” to fix the name.



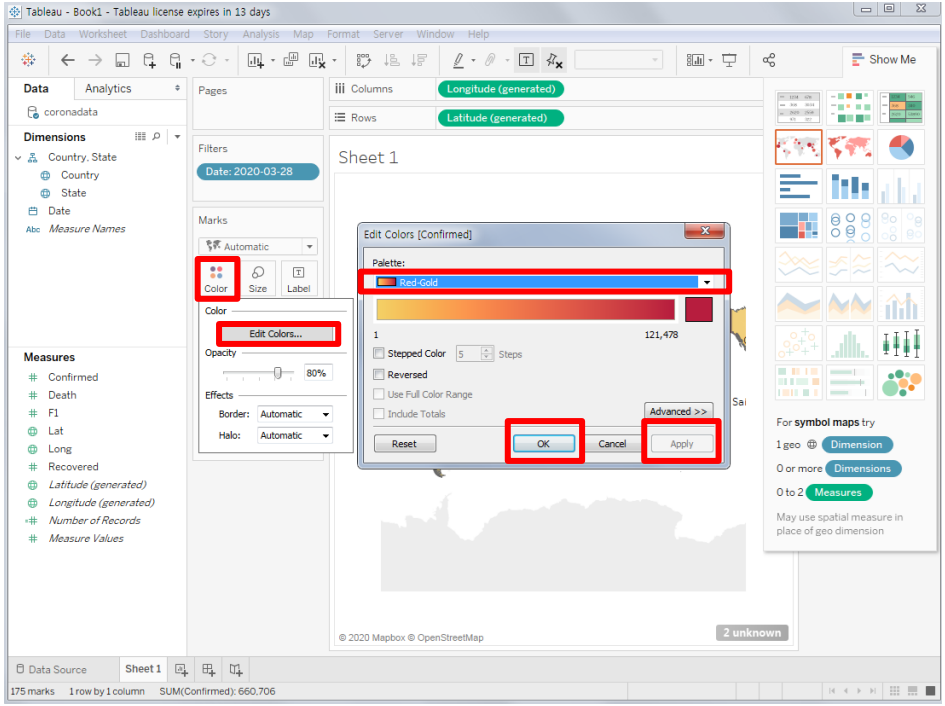
As a result, we can see that diamond princess and MS zanndam are unrecognized.



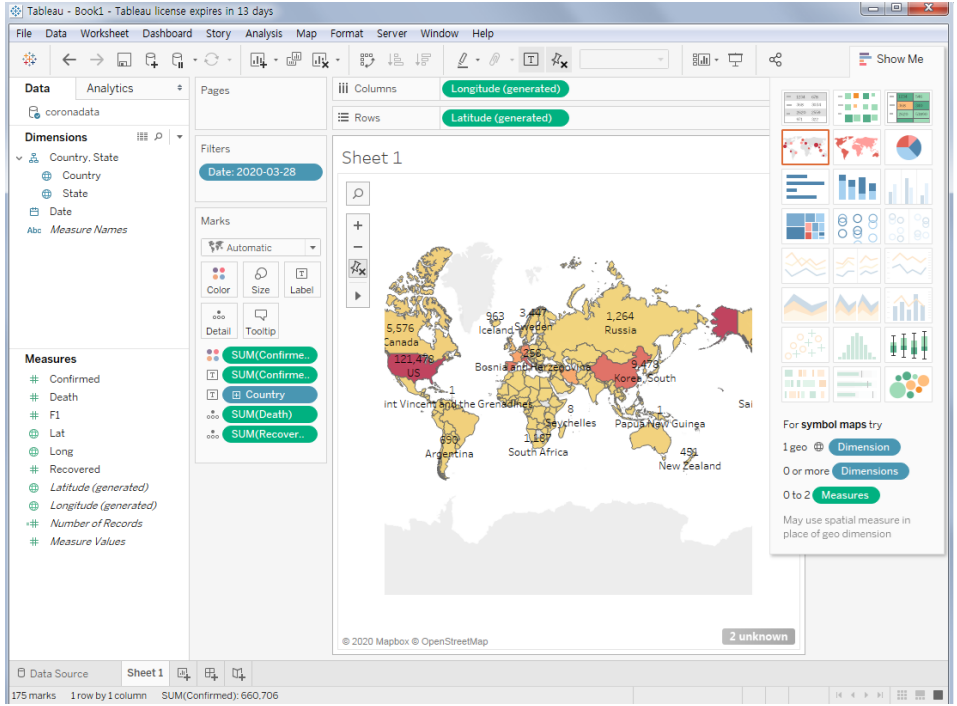
Diamond princess and MS zaandam are cruise ships, so leave the name of territory unchanged.

Exercise the Tableau with COVID-19 Dataset

If you want to change the color of the map, click “Color in Marks”, select Edit Colors, select the color you want from “Palette”, click “Apply”, then click “Ok”.

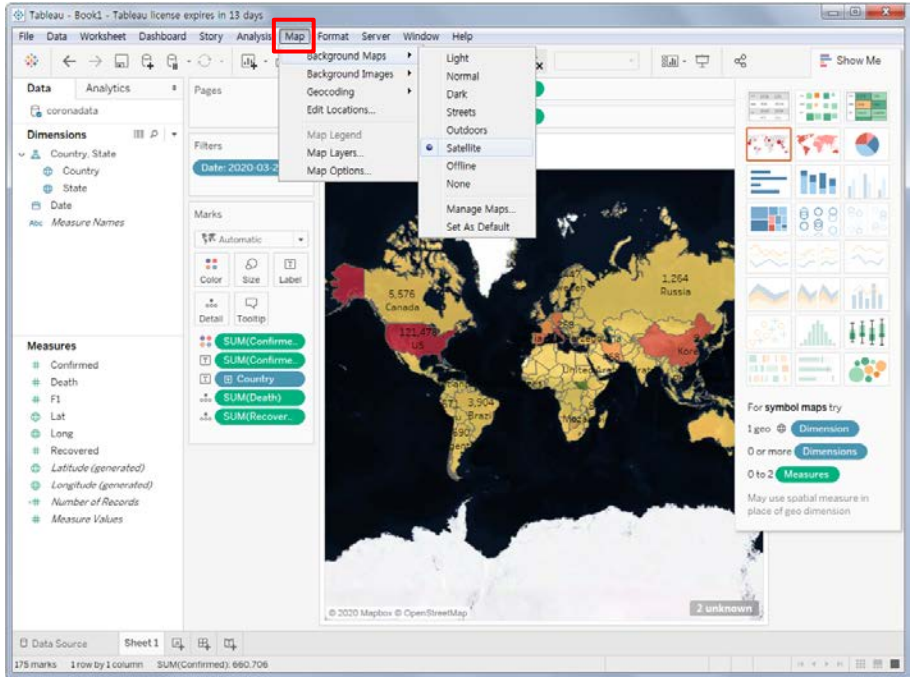


You can see that the color has changed to Red-Gold as shown below.

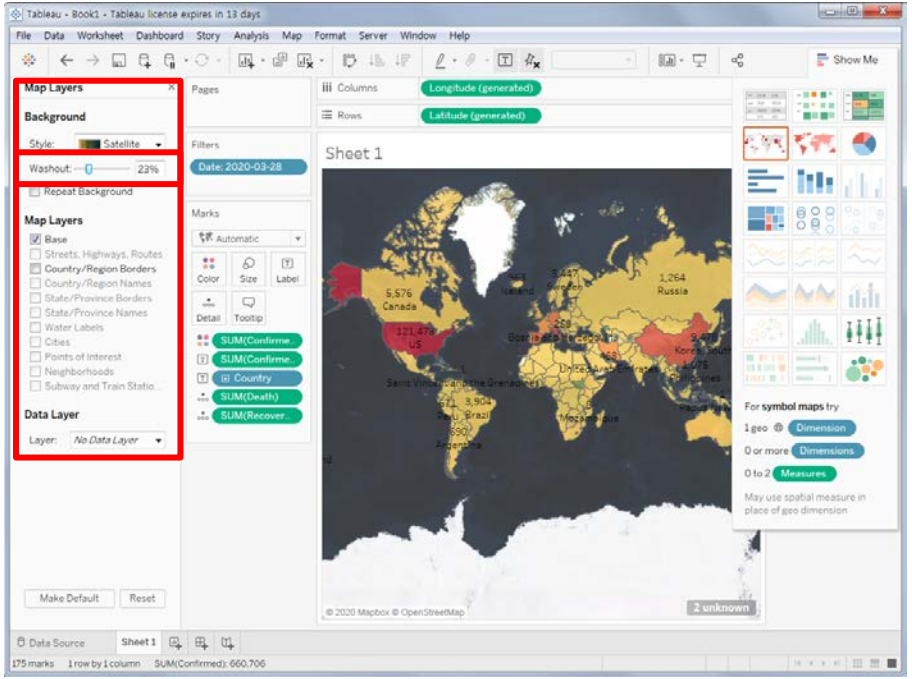


Exercise the Tableau with COVID-19 Dataset

We can change the background of the “Map” by pressing the “Map” button and selecting “Background Maps”. Let's change the map background to Satellite.

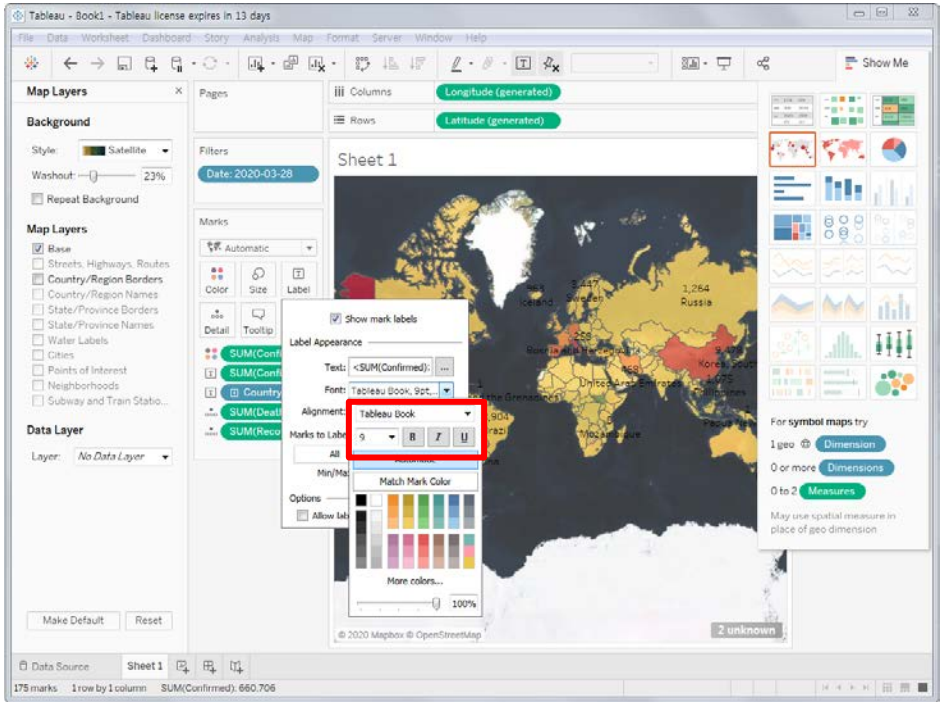


When we click the Map layers button, we can adjust the Background, Map layers, and Data layers on the left. Adjust the Background Washout to 23%.

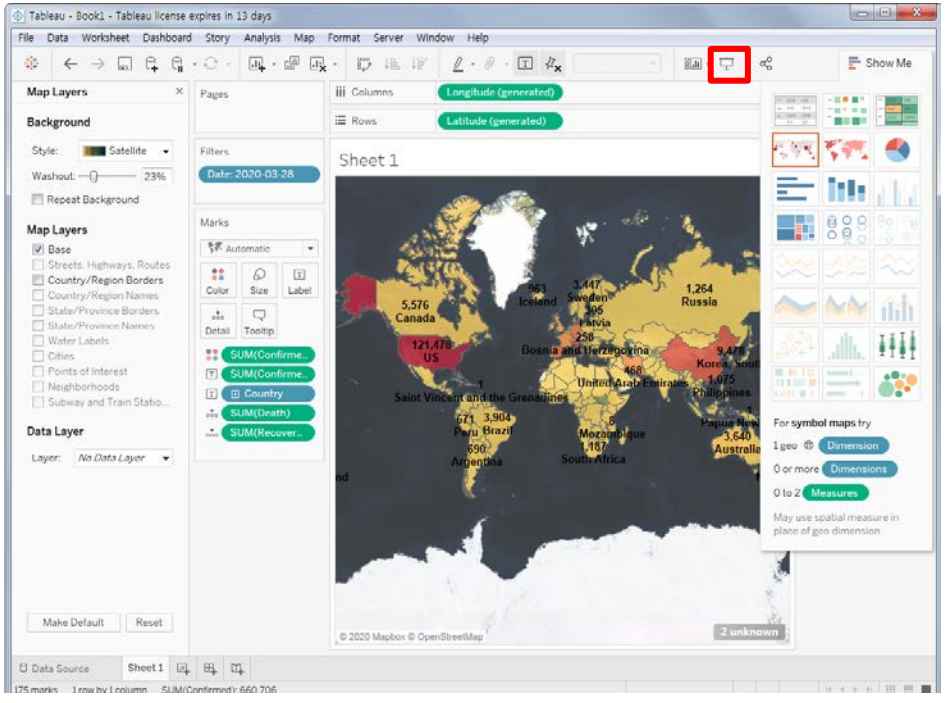


Exercise the Tableau with COVID-19 Dataset

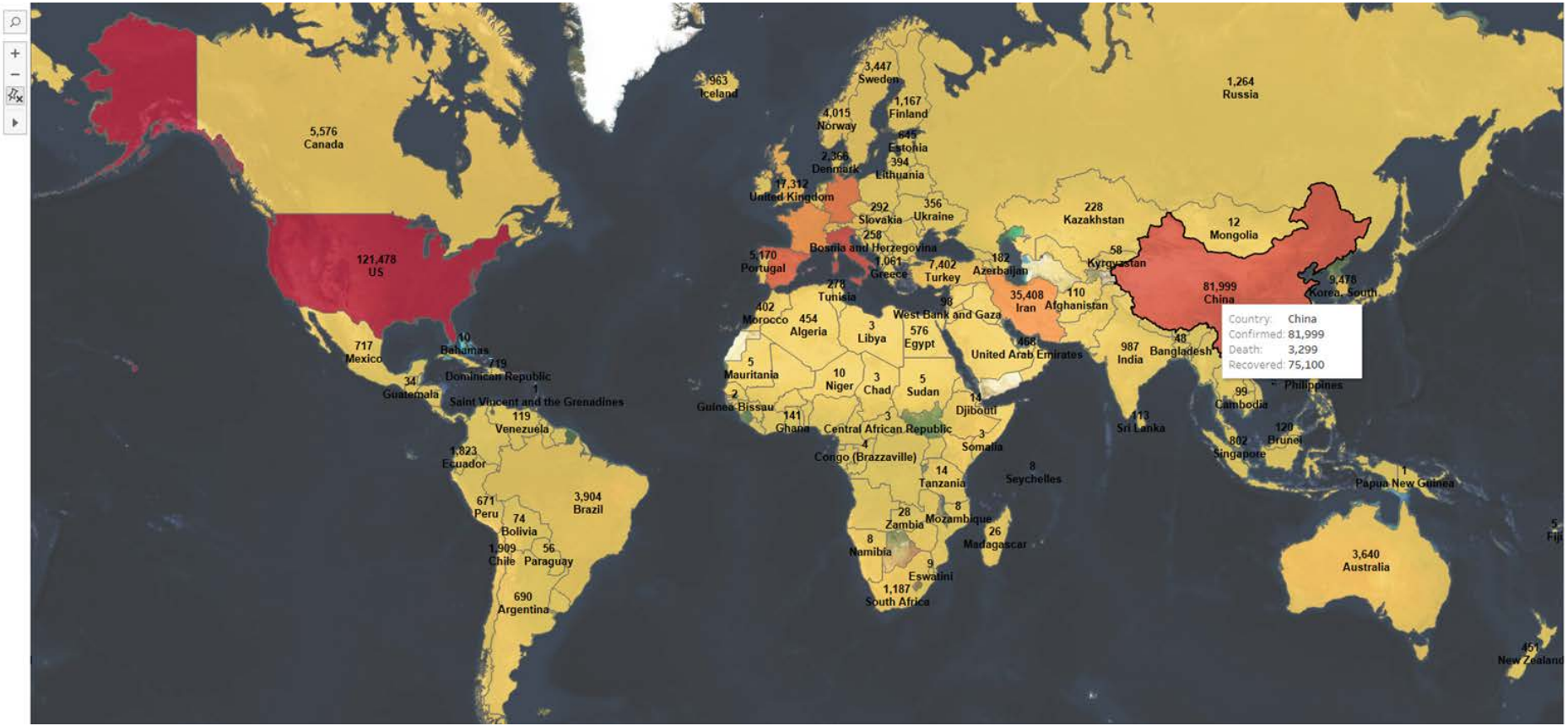
You can change “Text”, “Font”, and “Alignment” in “Label in Marks”.
Let’s Change the Label Font from "Tableau Book" to "Arial" and the text size from 9 to 10.



You can also watch the result with Presentation



Exercise the Tableau with COVID-19 Dataset



Reference Materials

You can also make COVID - 19 dashboard with Tableau. Many institutions already are using Tableau for making Dashboard (link include)

Johns Hopkins Center for Systems Science and Engineering (CSSE)

HealthMap (Harvard/Boston/Northeastern/Oxford Martin School) animated dashboard

World Health Organization (WHO) COVID-19 dashboard

Reference Text Books