Prerequisites

- You have installed Tableau Desktop on your computer.
  Available here: http://www.tableau.com/academic/students
- You have downloaded the data.
  Available here: https://data.nasa.gov/view/angv-aquq (export as .csv) or from our server: http://vda.univie.ac.at/Teaching/Vis/16w/Tutorials/data/Global_Landslide_Catalog_Export.csv

Motivation

With Tableau you can quickly create visualization of your data. Its ease of use makes it valuable tool for initial data exploration and it also allows you to create complex interactive visualizations (e.g. for prototyping).
You can create dashboards consisting of multiple views on the data and options for interaction. Examples are:

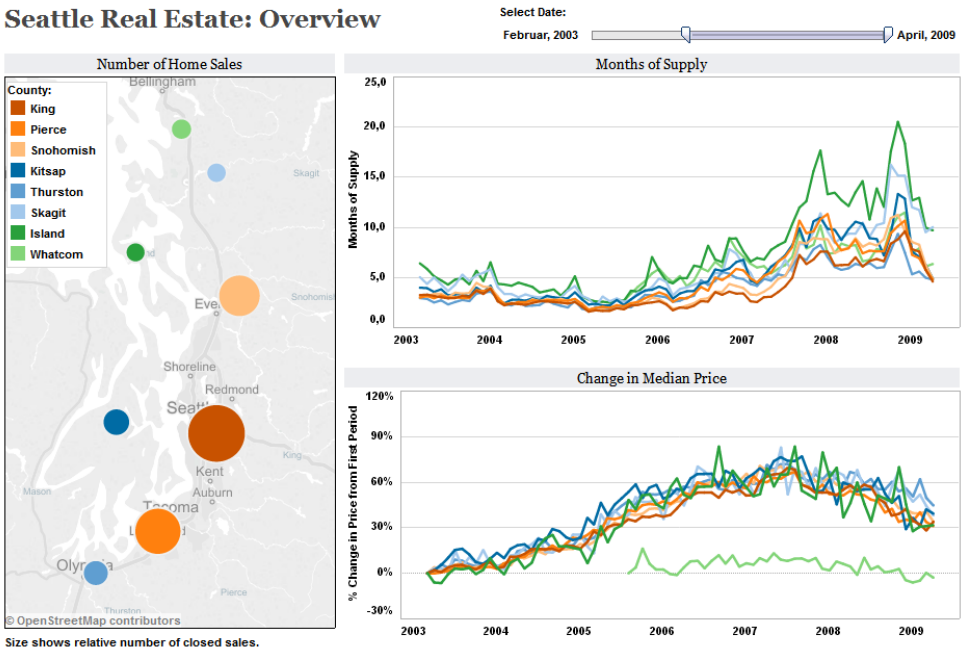


Figure 1: http://www.tableau.com/stories/gallery/real-estate-prices

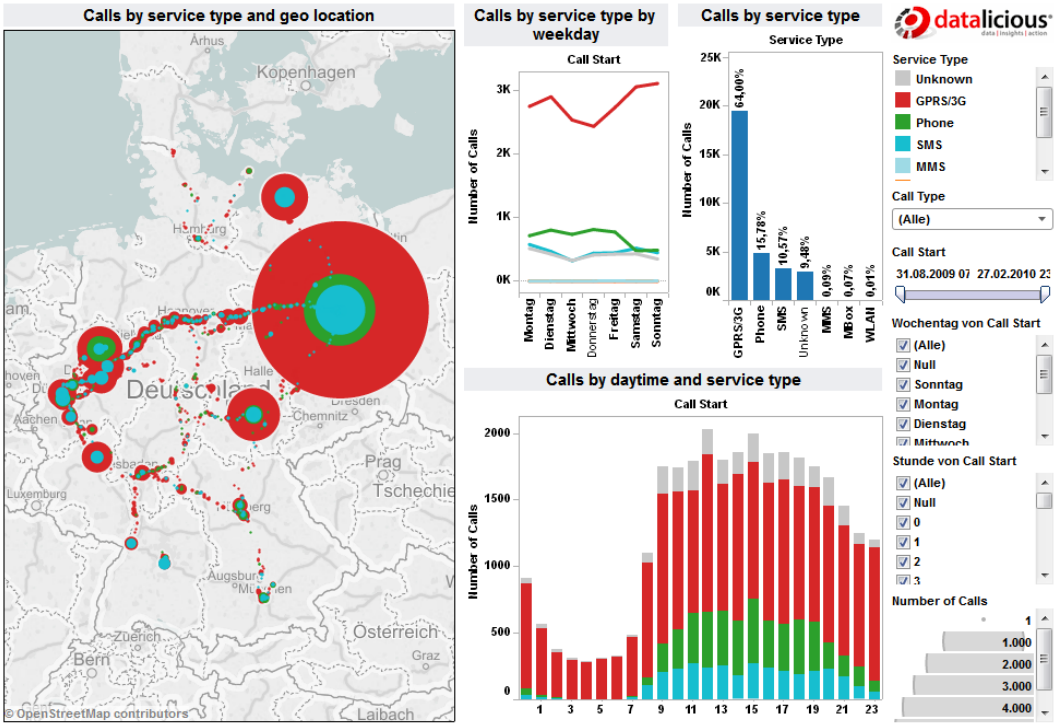


Figure 2: http://public.tableau.com/profile/datalicious.pty.ltd#!/vizhome/MalteSpitzCallData/MalteSpitzcalldatadashboard

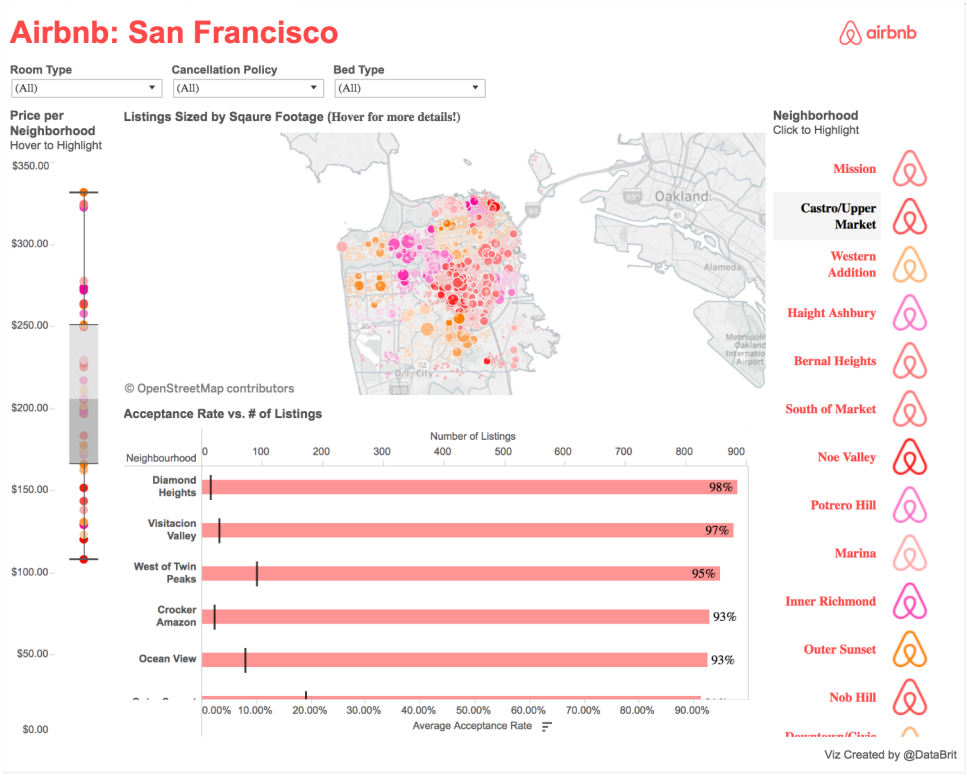
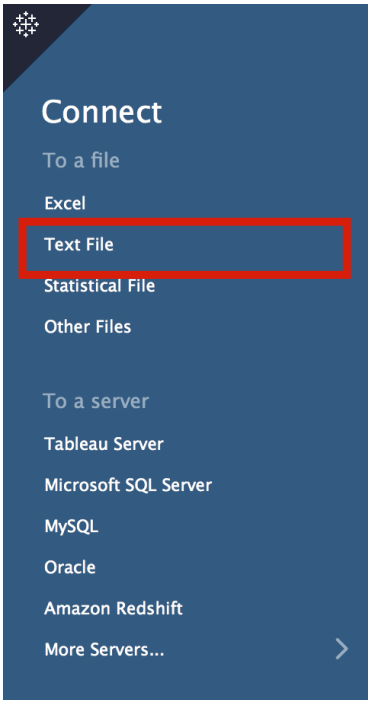


Figure 3: https://public.tableau.com/views/AirbnbSanFranciscoAnalysis/Airbnb?:embed=y&:loadOrderID=0&:display_count=yes&:showTabs=y

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**.

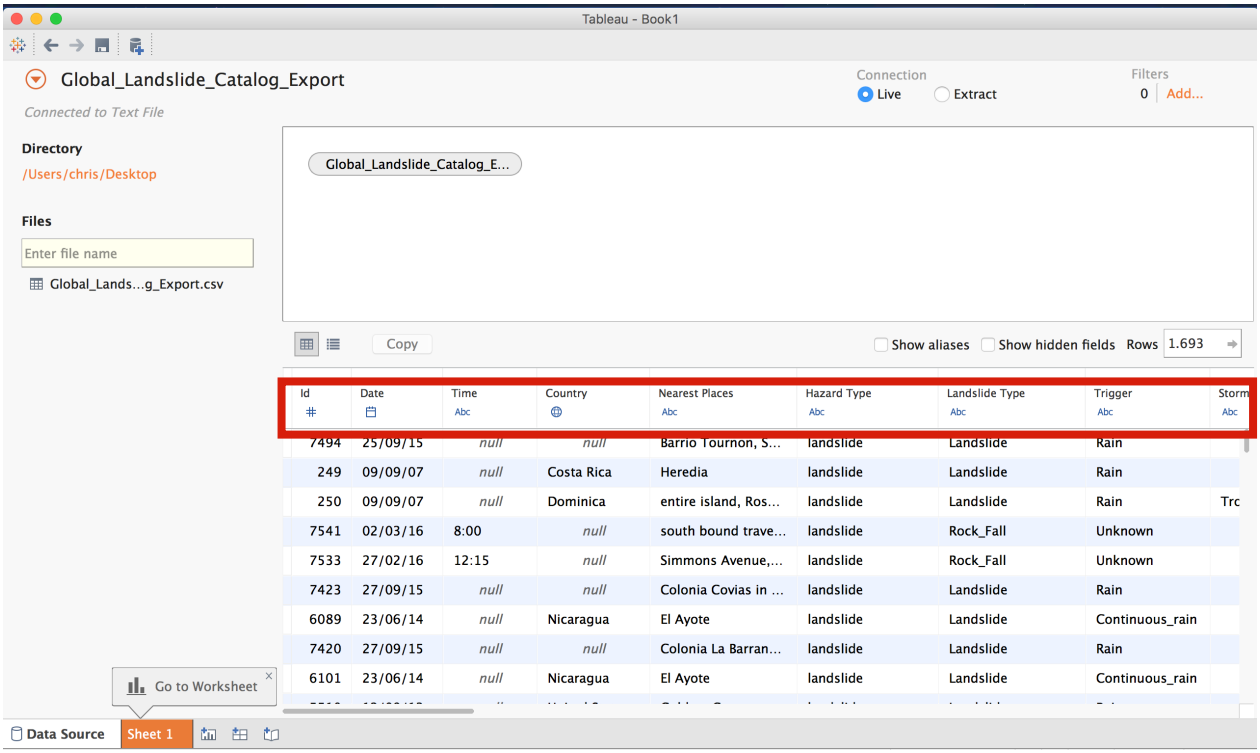
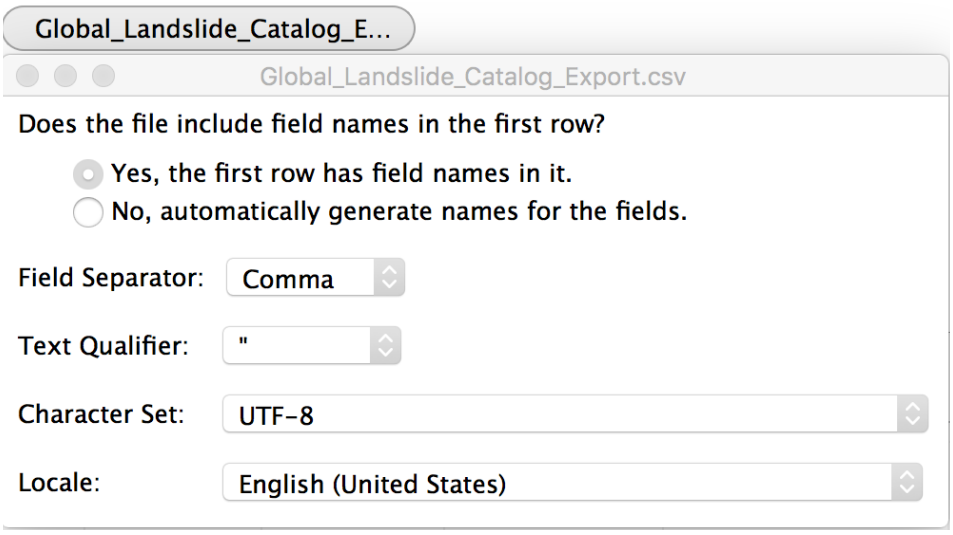


Tableau evaluates automatically what data type is in each column. Always check if the automatic data types are correct by controlling the column headers marked in the figure above.



Mouseover the dataset name and click the little gear to open the settings for the dataset. Choose the correct settings:

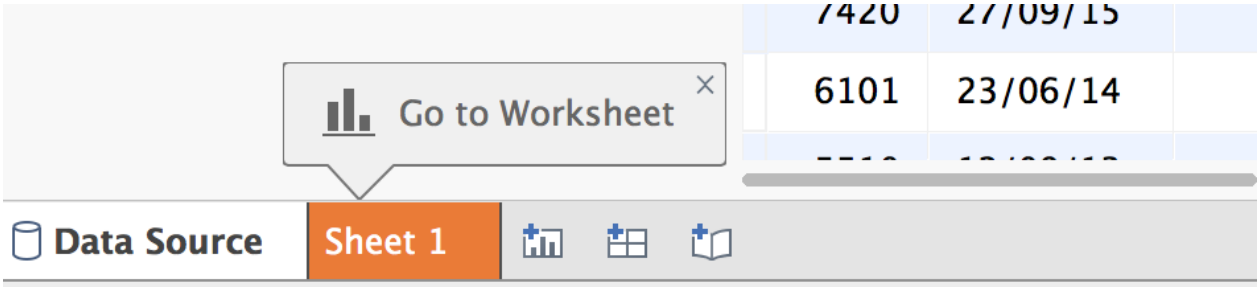
Field Separator - Comma, tab or other

Text Qualifier - How can strings be detected

Character Set - Encoding

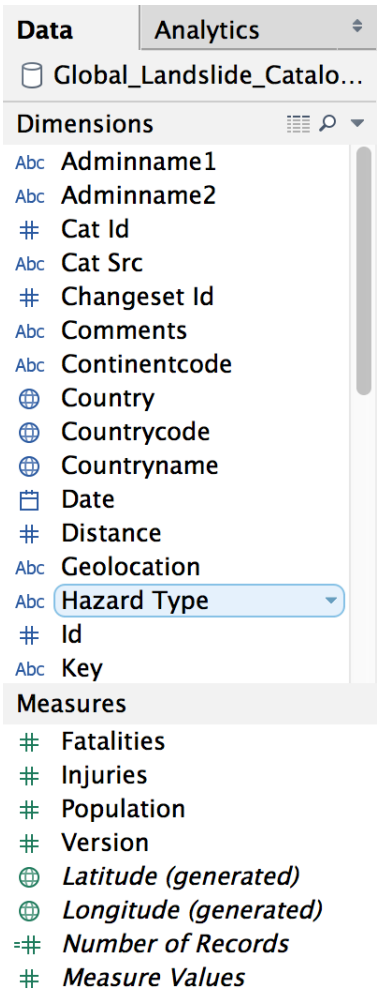
Locale -Where is the data from

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.



Basics

Dimensions & Measures



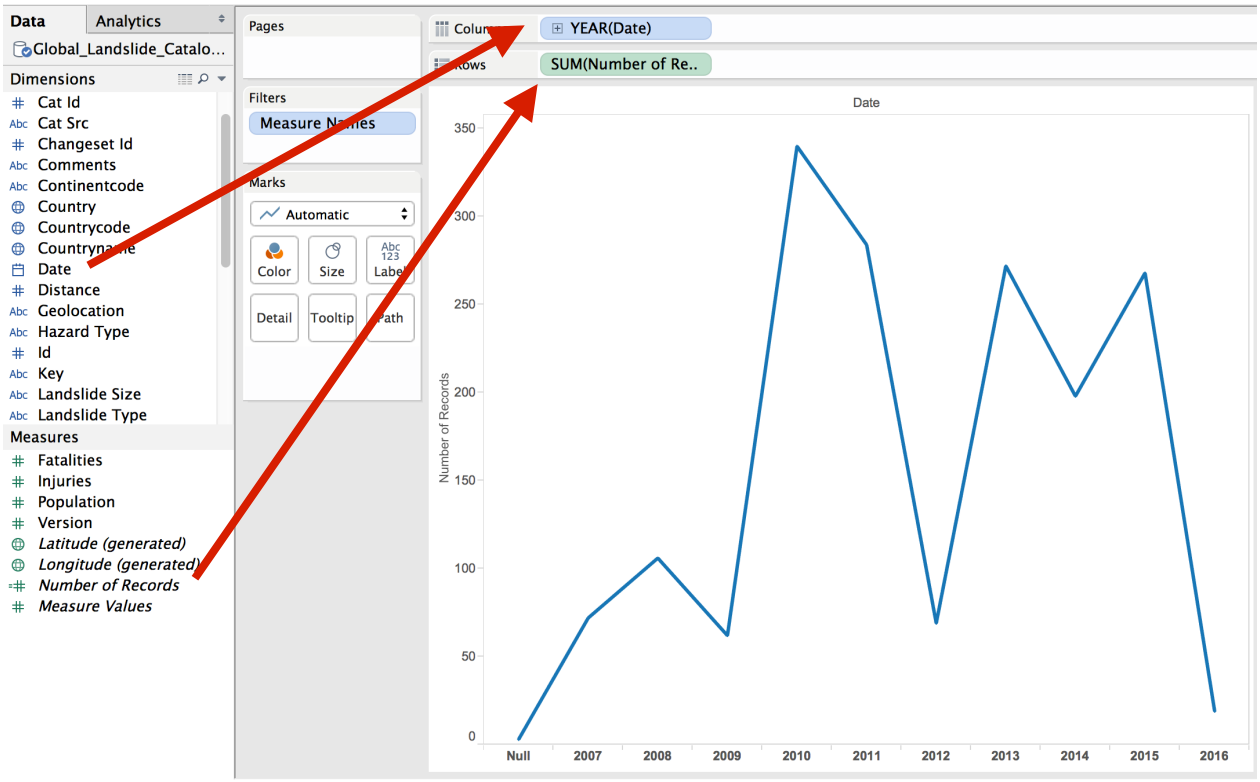
In the leftmost panel you will see the columns of your table as either dimensions or measures.

Dimensions are usually categorical datatypes. They can be used to separate your data by discrete tags.

Measures are the quantitative data that you will encode in your marks.

You can drag and drop both measures and dimensions to the central panel to create plots.

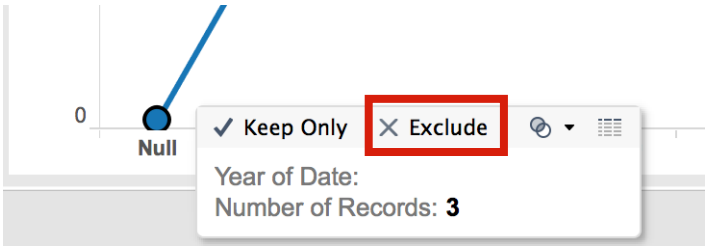
Building Charts



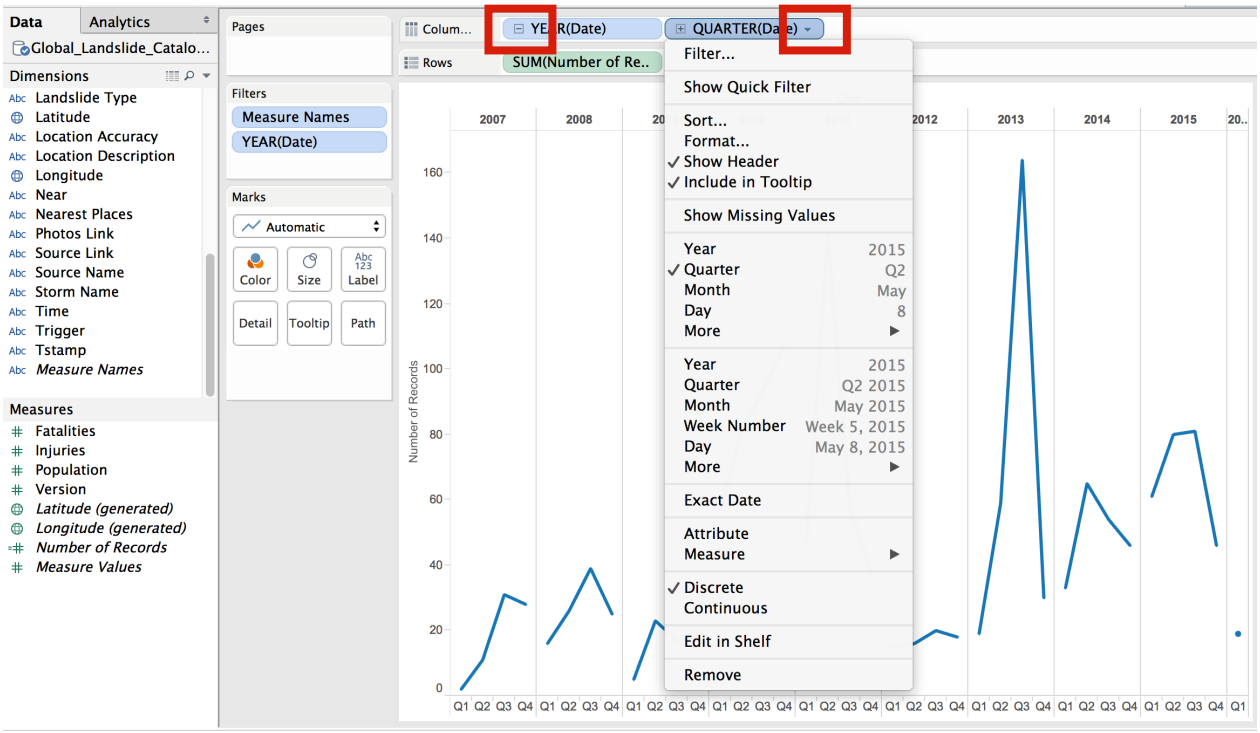
Let's suppose we want to view the development of the number of landslides over time.

Simply **drag** the dimension **Date** to the columns and the measure **Number of Records** to rows.

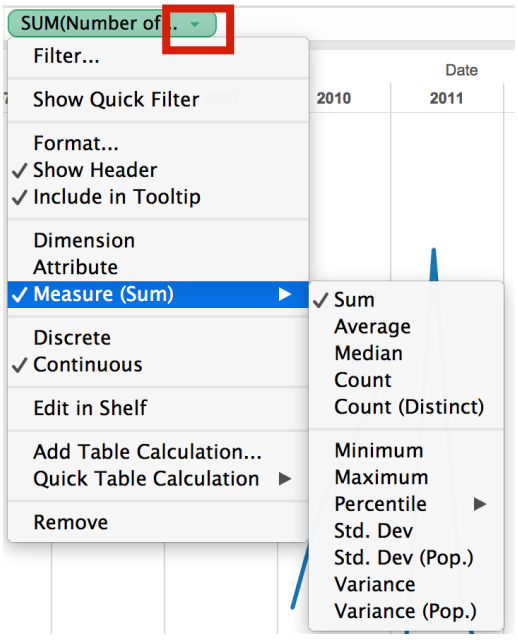
You will see that the records are automatically summed up and the date is set to YEAR. Hence the plot shows the number of landslides per year.



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



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.

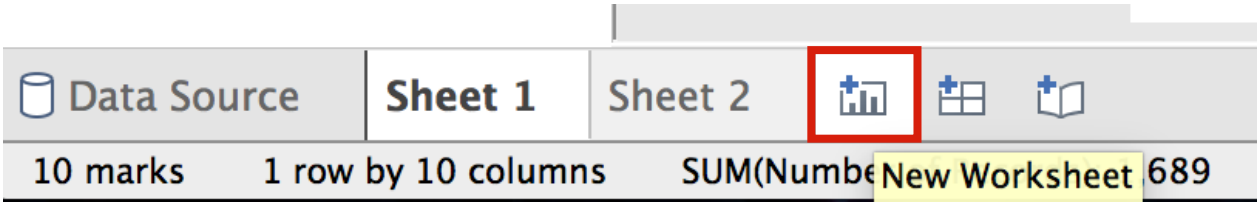


The small triangle is available for all datatypes. The measure uses per default **Sum** and you can change it to **Average** or **Variance**. This depends on your dataset and the questions you want to answer/things you want to show.

Hands-On

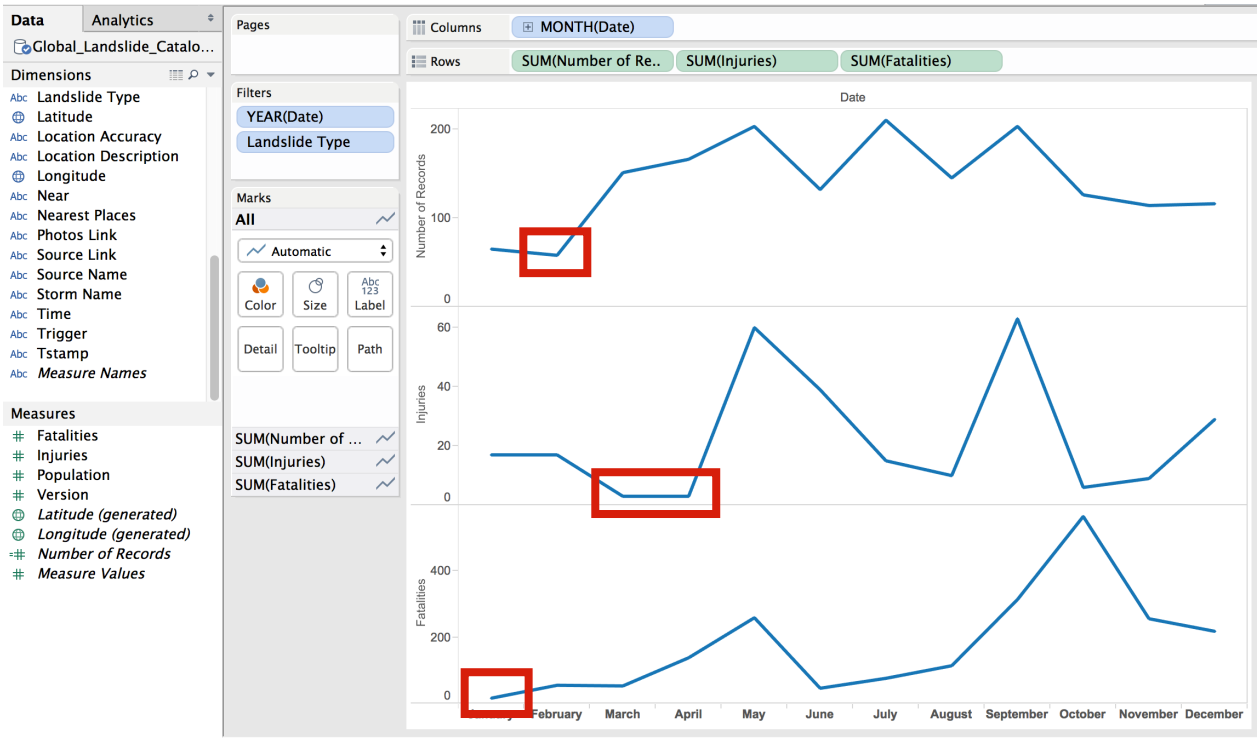
Question: Which month is the **safest month**? Show the total number of landslides, the number of injured persons and the number fatalities.

Create a new sheet by clicking on the **New Worksheet** button and answer the question with help of Tableau:



If you need help ask the TA's!

One possible solution:

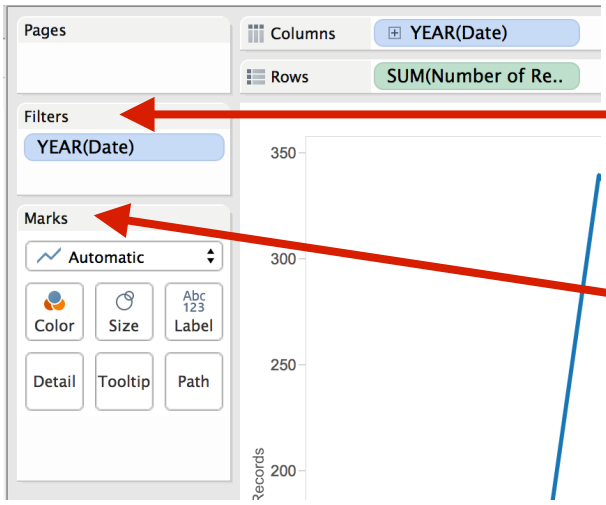


From a **total number of events** points of view: February is the safest month.

From a **number of injured persons** points of view: March and April are the safest months.

From a **number of fatalities** point of view: January is the safest month.

Marks & Appearance

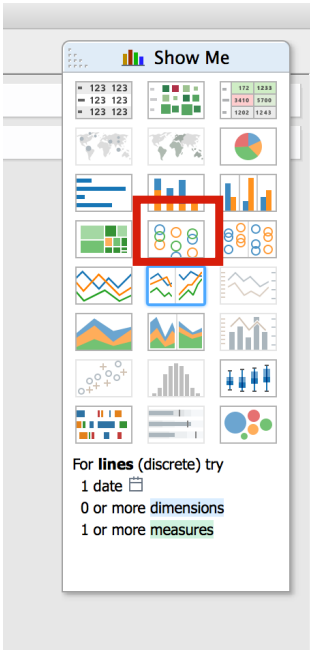


Back to **Sheet 1**.The excluded data is not lost, it is just filtered away. You can always see what filters have been set in the **Filters** panel.

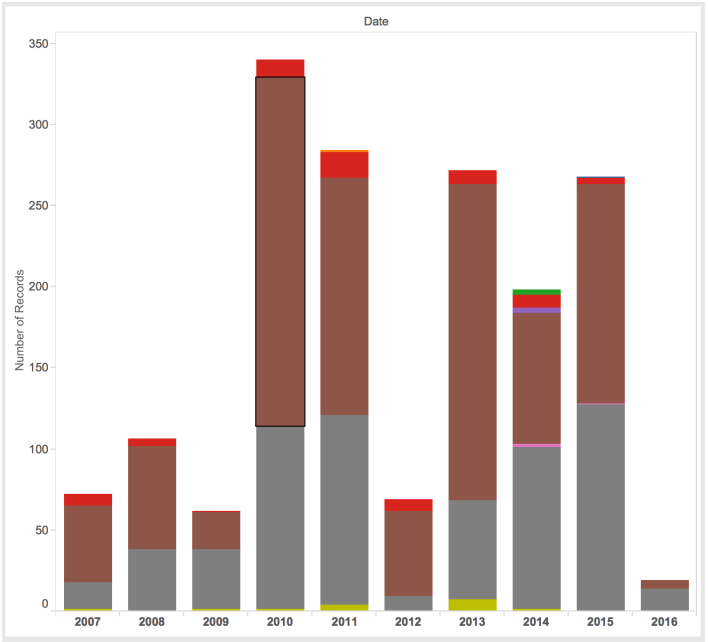
Next let's take a look at the **Marks** panel below it. Here you can modify the appearance of your plot. We want to use marks to encode more information in view.

Lets 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 that:



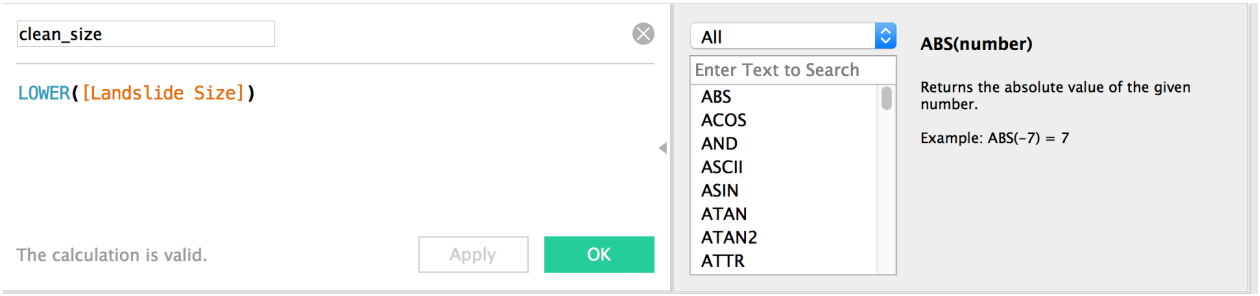
The default color encoding is not really good. Lets change it!

Exclude **Null** and the **2011** one by right clicking them and select **Exclude**.

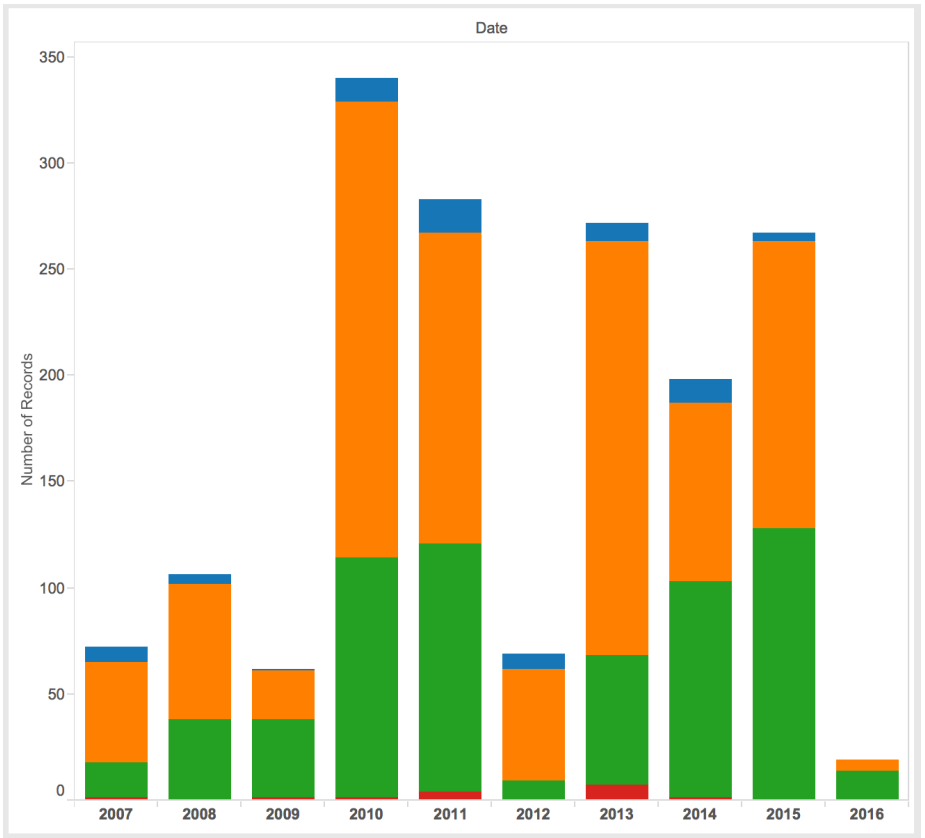
Reorder the entries by **drag and drop**.

We can see here another data quality issue -> Small and small should be in one group.

We can fix this by creating a calculated field. To do so right-click anywhere on the white area in the **Data panel.** Click **Create Calculated Field…** and use the following calculation to create a new measure we will call **clean_size**:



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



Reorder the entries again by **drag and drop**.

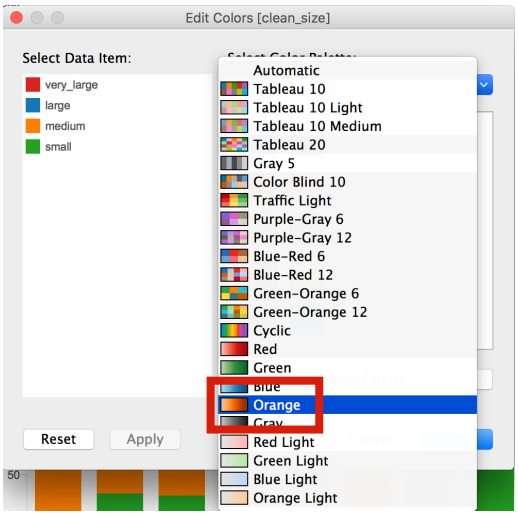
Click the small triangle which is visibly when you mouse over and select **Edit Colors…**

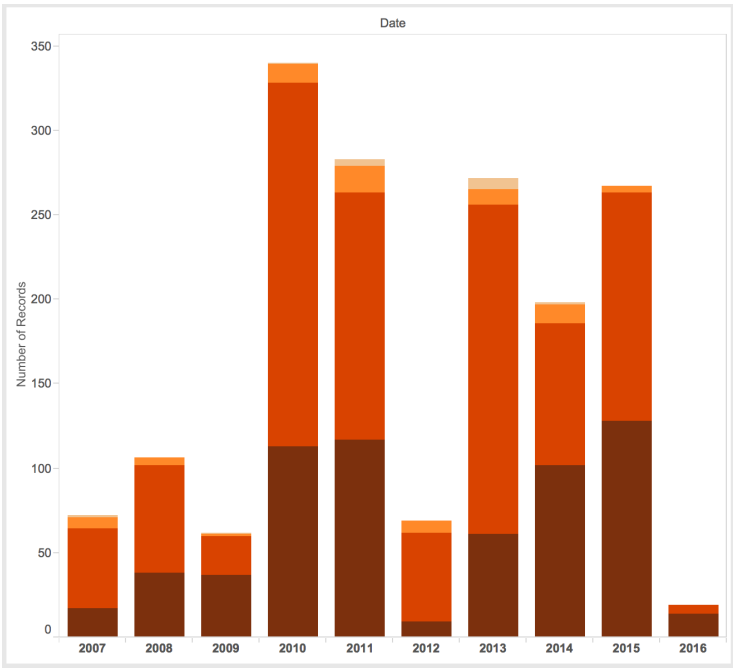
Select **Orange**

-> **Assign Palette**

-> **Apply**

-> **OK**

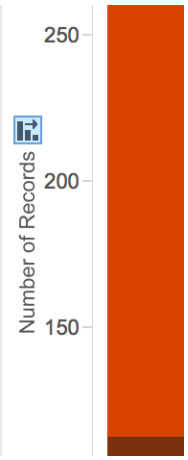




It should look like this now!

We can see that most of the Landslides are small or medium.

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



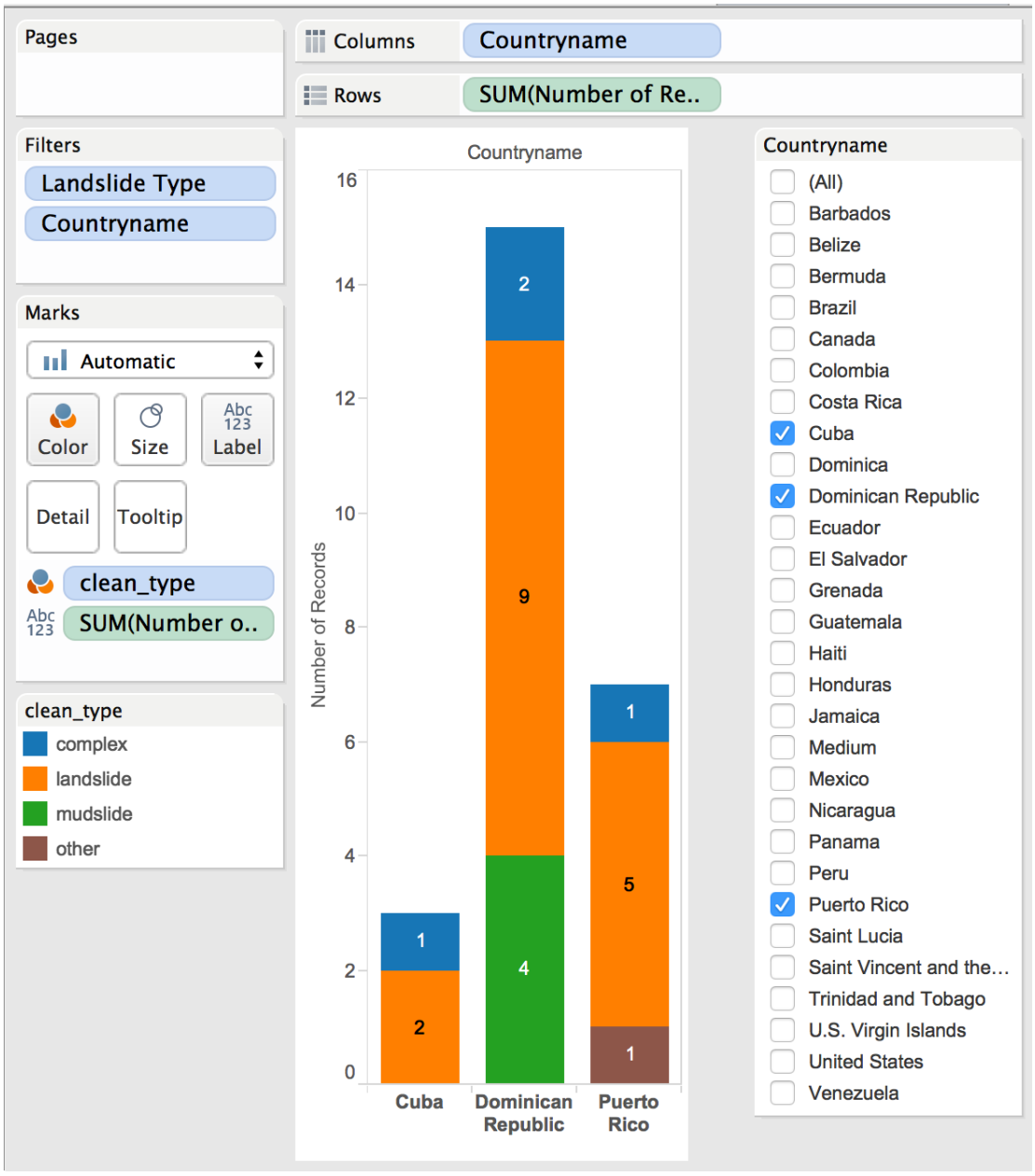
You can add more information by dragging the **Number of Records** to **Label** in the **Marks Area**.

Hands-On

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

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

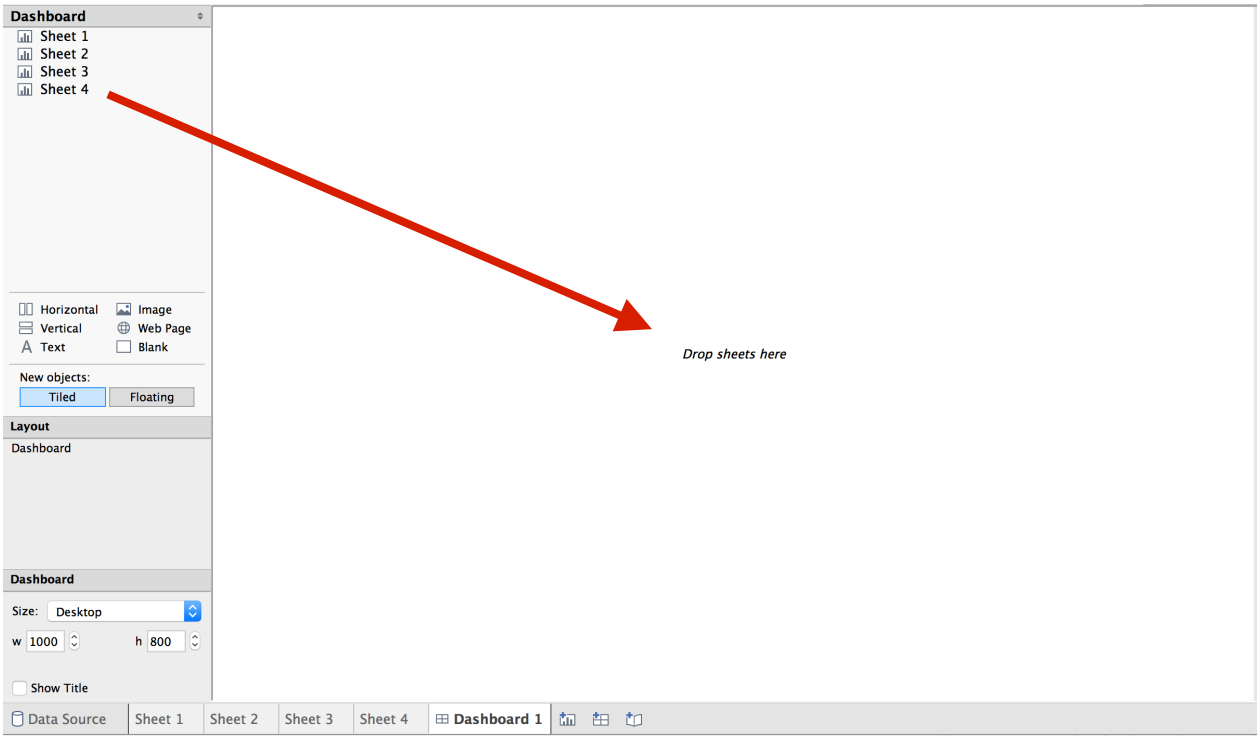
One possible solution:



Dashboard

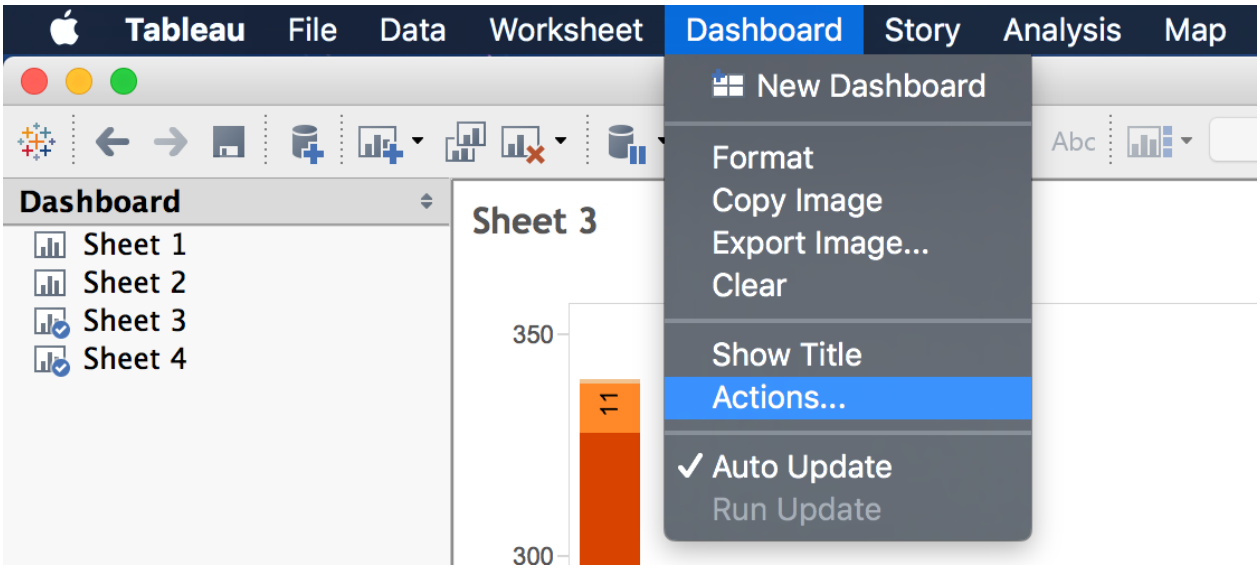
A Dashboard connects multiple single views with each other. This allows more complex questions to be analysed. Do create a dashboard click the **new Dashboard** button at the bottom.

Drag Sheets from the left side into the middle area:

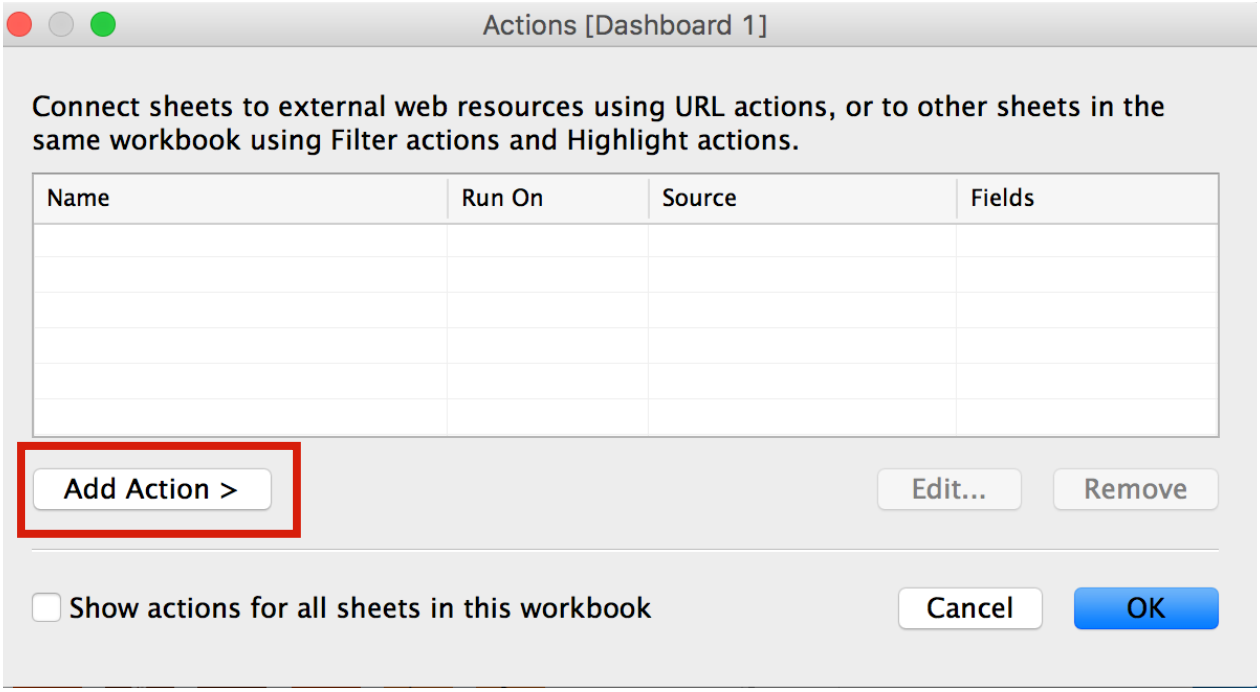


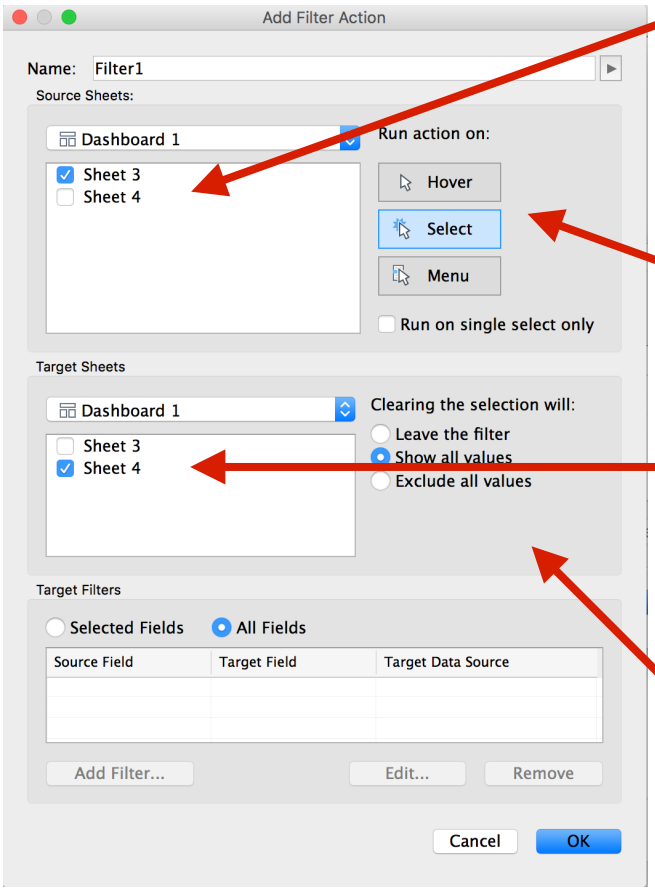
You can now define Dashboard wide actions like Filters and Highlights. We will add a filter so that can select in one view data and the second view will be updated accordingly.

In the Dashboard, add a new Action by selecting Dashboard -> **Actions…**



Add a new action with **Add Action >** and select **Filter…**





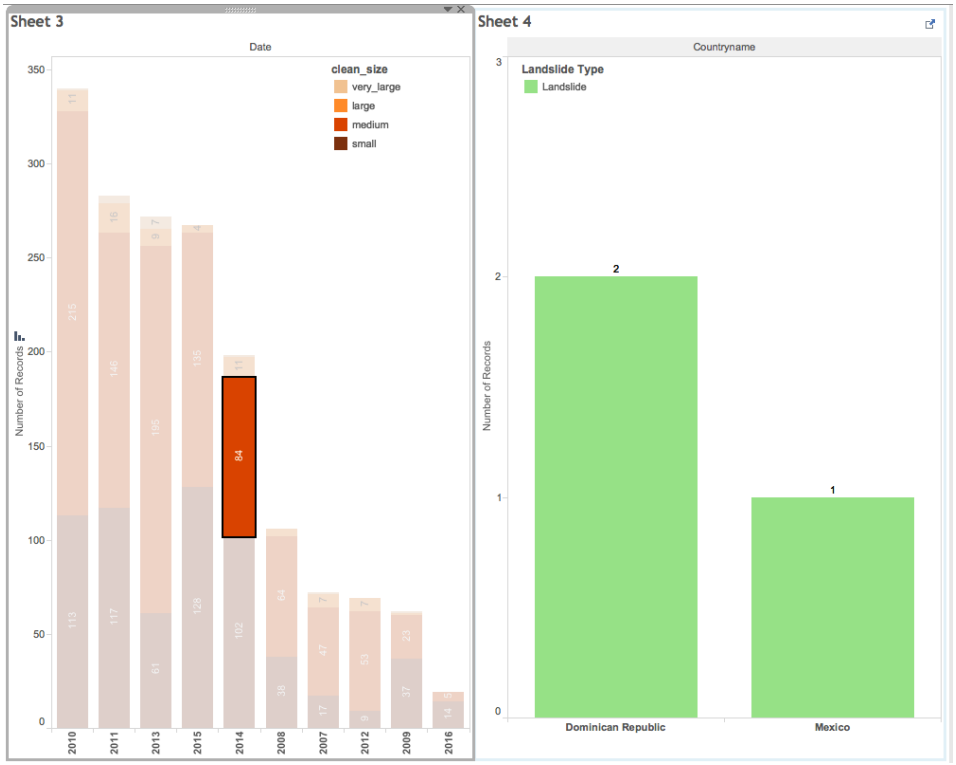
Source Sheets define the sheet where the filter action is recognised. In our case is Sheet 3 the source.

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

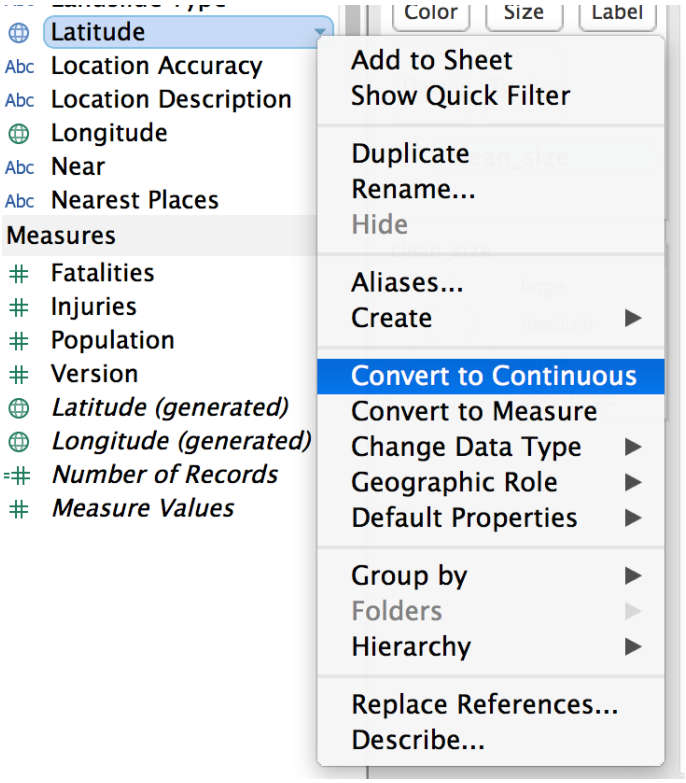
Clearing the selection will provides option what happens when you clear the selection.

The picture on the right shows the action in action. I have selected medium size landslides in 2014 and we can see that from the three countries we selected in the second hands-on section, only two have valid entries.



Explore

Exercise: Use the mapping functionality of Tableau and create a interactive Dashboard with at least 2 views, to explore the dataset further. Show us what you can find out about triggers, size, type, location, impact and other things!



To use the location data you can use the country information, but there are **null** values in this column. To get **all** data points, you have to **Convert Latitude** and **Longitude** to **Continuous**. After that double click both and the map is ready to go.

If you have problems, ask the Ta's!

A possible solution: Storm Watcher