

# Tableau Academic for Research Week 2020

## Table of Contents

<b>The Data .....</b>	<b>1</b>
<b>Opening the data in Tableau.....</b>	<b>1</b>
<b>Visualisation 1 – World cases by day bar graph .....</b>	<b>3</b>
<b>Visualisation 2 – Top 10 Countries by Cases.....</b>	<b>8</b>
<b>Visualisation 3 – Australia Cases and Deaths Double line graph.....</b>	<b>11</b>
<b>Visualisation 4– Deaths in the top 7 worst affected countries since first 100 deaths .....</b>	<b>17</b>
<b>Visualisation 5 – Map .....</b>	<b>25</b>
<b>Stretch Goal 1 – Testing Animated bar graph .....</b>	<b>32</b>
<b>Stretch Goal 2 - Dashboard .....</b>	<b>34</b>

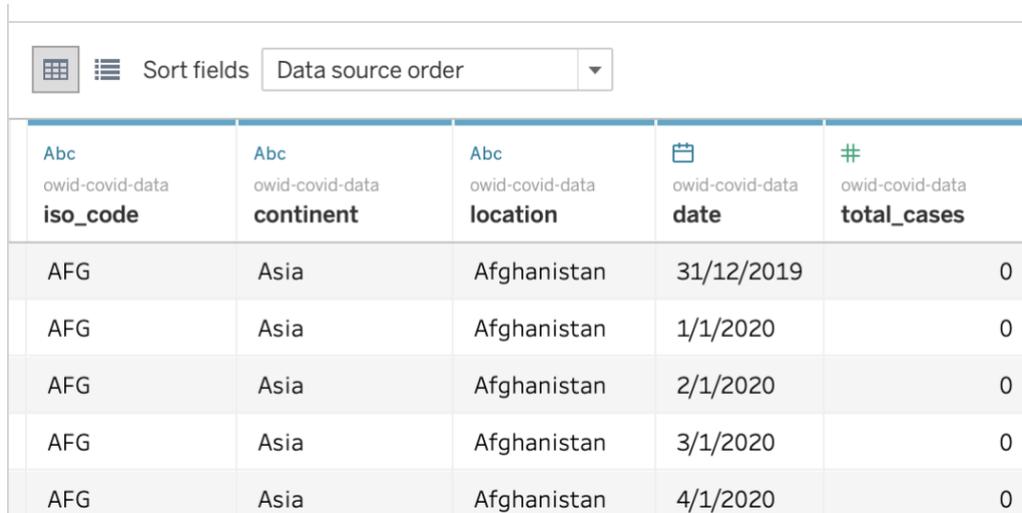
## The Data

Our data is in .csv form and comes from ouroworldindata.org

They have a ‘complete our world in data COVID 19 dataset’. It’s day by day data from each country so a huge dataset.

## Opening the data in Tableau

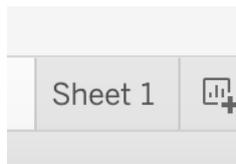
In Tableau - Connect > To a File > and then open the csv file we supplied you during the class. You’ll see information in a table:



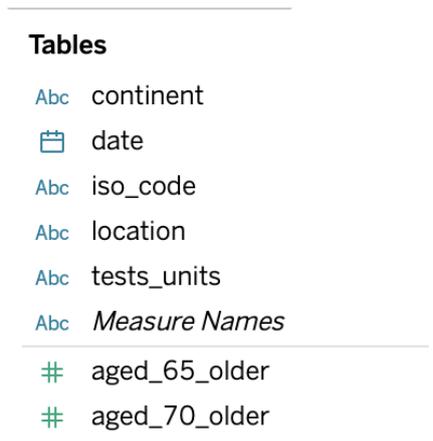
The screenshot shows the Tableau interface with a table of data. At the top, there is a 'Sort fields' dropdown menu set to 'Data source order'. The table has five columns: 'iso\_code', 'continent', 'location', 'date', and 'total\_cases'. The data rows show Afghanistan (AFG) in Asia for the years 2019 and 2020, with zero total cases recorded for all dates shown.

iso_code	continent	location	date	total_cases
AFG	Asia	Afghanistan	31/12/2019	0
AFG	Asia	Afghanistan	1/1/2020	0
AFG	Asia	Afghanistan	2/1/2020	0
AFG	Asia	Afghanistan	3/1/2020	0
AFG	Asia	Afghanistan	4/1/2020	0

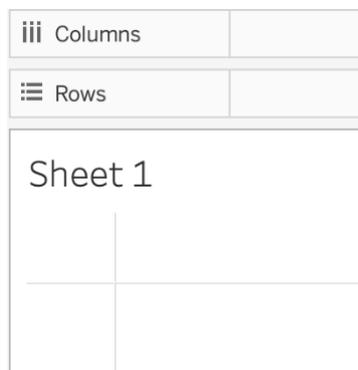
To graph the data, open a worksheet.



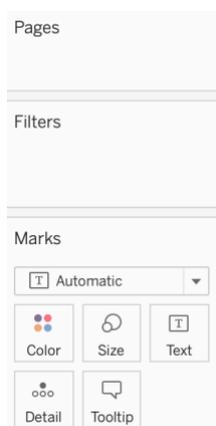
Within the worksheet you'll see the fields from the uploaded spreadsheet on the left of screen



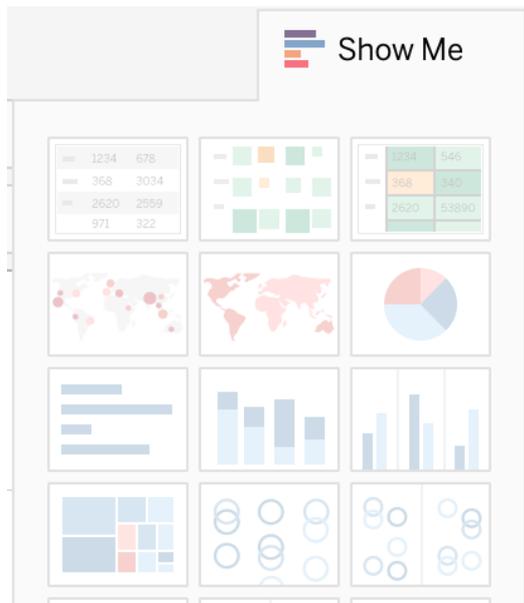
The worksheet itself is a form of pivot table. You drag the data from the table of values on the left over to columns, rows, or onto the worksheet itself to display it.



The pages filters and marks boxes let you control how the data is displayed:



The Show Me area lets you toggle between visualisation types



Mostly, these four elements are all you'll need to make visualisations.

## Visualisation 1 – World cases by day bar graph

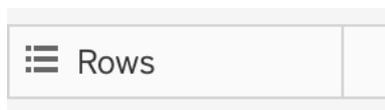
Begin by finding the location field on the left of screen

Abc iso\_code

Abc location

Abc tests\_units

Then drag the location value to the Rows area of the sheet



It should now look like this:

The screenshot shows the software interface with the 'Rows' area populated. The 'location' field is now visible in the 'Rows' header. Below the header, a table displays data for 'Sheet 1'.

location	
Afghanistan	Abc
Albania	Abc
Algeria	Abc

Now, say we want to make a graph of World cases by day.

The first thing we could do is look to see if there is a World tally within location.

It turns out there is! That makes our job pretty easy. Click on World and then tick the Keep Only option.



Now we just have this

## Sheet 1

location
World

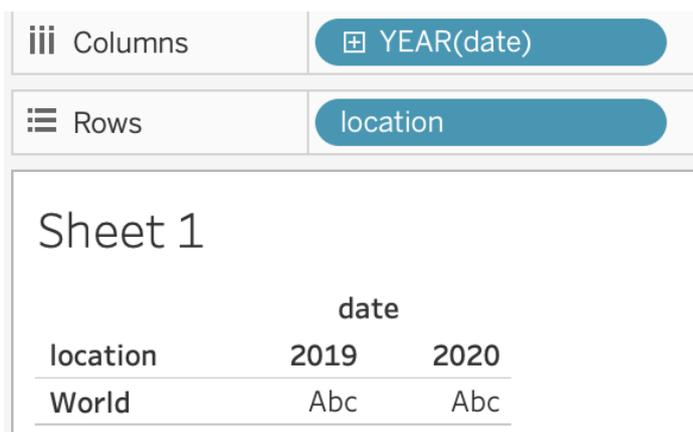
Then find the date value from the left hand table

Abc continent

📅 date

Abc iso\_code

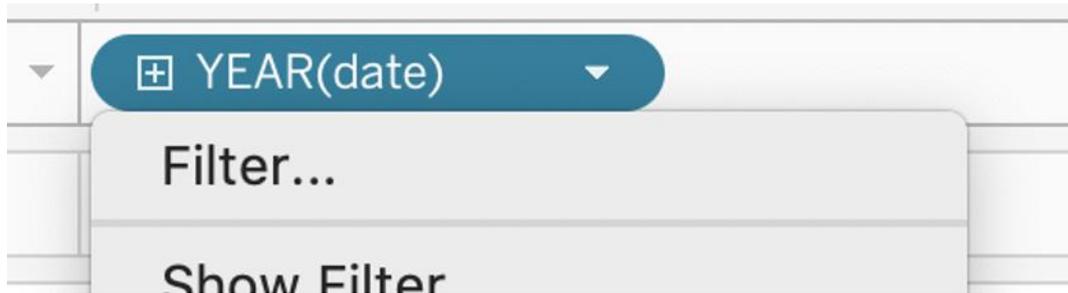
Drag the date value to the columns part of the worksheet. It should now look like this:

A screenshot of a worksheet interface. The 'Columns' section contains 'YEAR(date)' and the 'Rows' section contains 'location'. Below this, the worksheet content is shown as a table with the following structure:

	date	
location	2019	2020
World	Abc	Abc

What we want to do now is break down the time increment from years into days so we can better see the data.

To do this click on the down arrow by the year value



And then select day

Year	2015
Quarter	Q2 2015
Month	May 2015
Week Number	Week 5, 2015
Day	May 8, 2015

Now we have this very fine range of data points waiting for some data

## Sheet 1



Now find the new\_cases measure from the left of screen

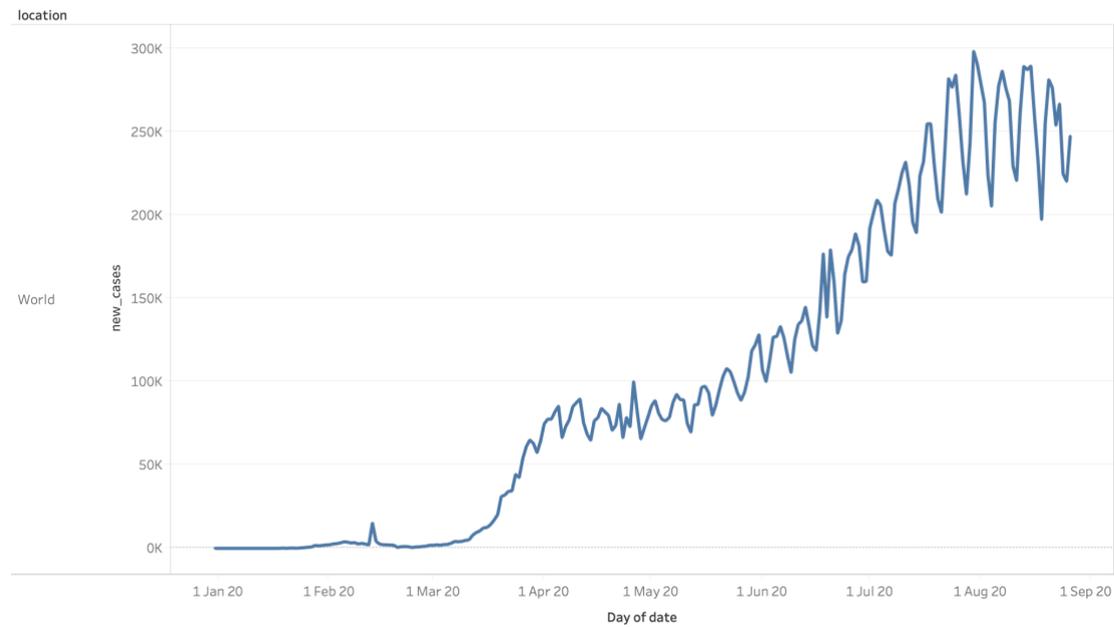
- # median\_age
- # new\_cases
- # new\_cases\_per\_million

Drag new cases next to location on the Rows area

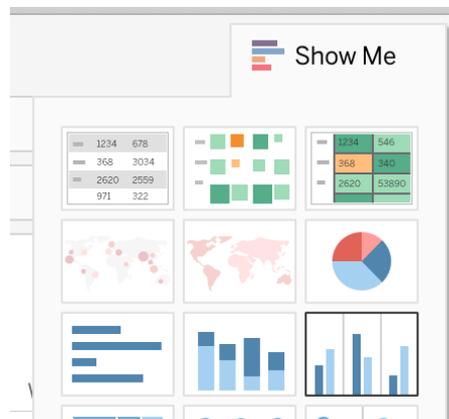


You should now see a line graph showing that world wide cases are trending upwards with a characteristic oscillating flow

Sheet 7



If we want to ,we can change the visualisation to a bar graph. To do this go to the show me area on the right of screen and choose bar graph:



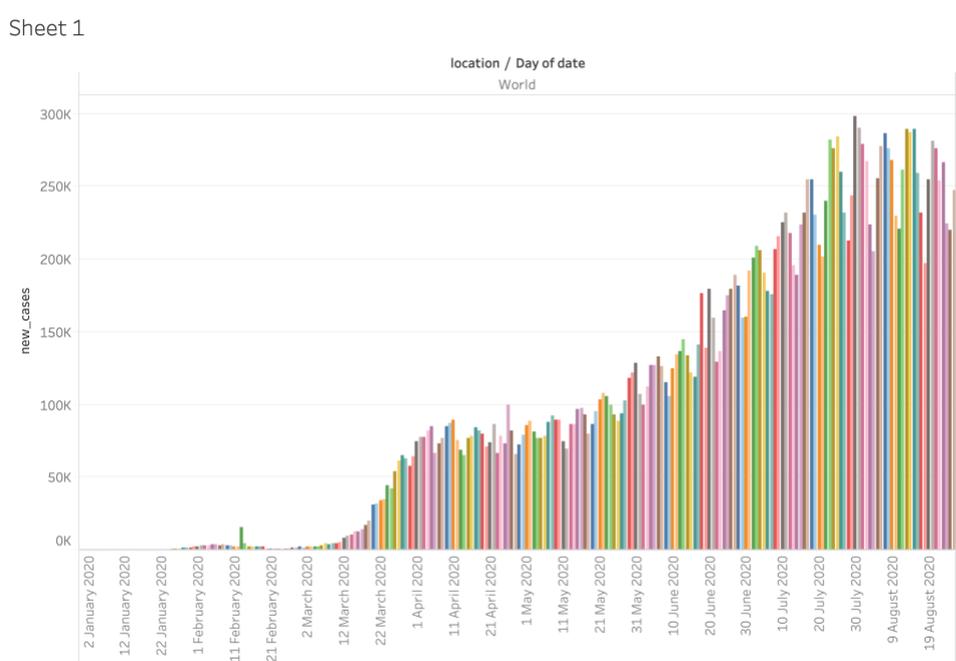
The base of the visualisation will now look like this:



All of these data points won't fit on the page! So, let's go up to the area above the worksheet and find the drop down labelled Standard. Use the drop down to select Fit Width.



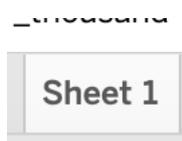
Hey presto! We now have a bar graph:



If you mouse over the graph it will give you information



Now let's label this graph. Double click the sheet at the bottom of screen



Rename Sheet 1 as World Cases/Day

Then click the plus sign next to World Cases/Day to open a new sheet (Sheet 2)

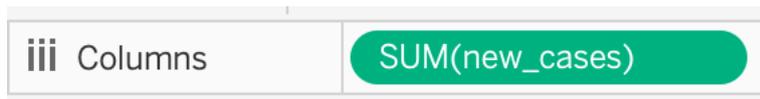


## Visualisation 2 – Top 10 Countries by Cases

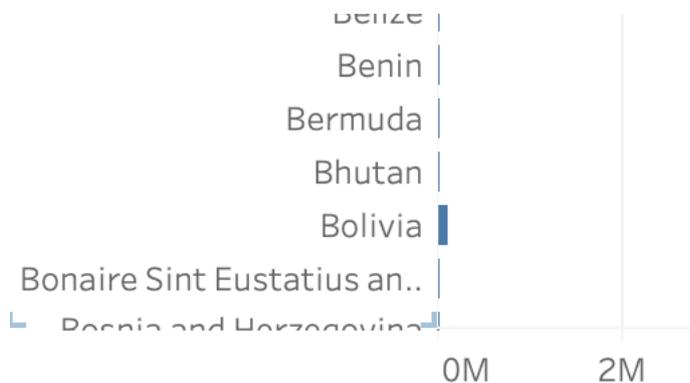
In the new worksheet take the location value from the left of screen and drag it to rows



Now drag new\_cases to columns

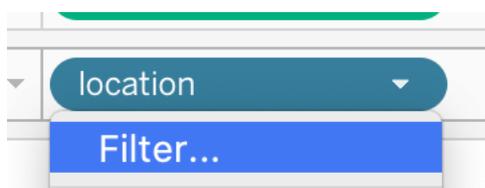


You now see all the countries of the world. Notice the scale is very small.

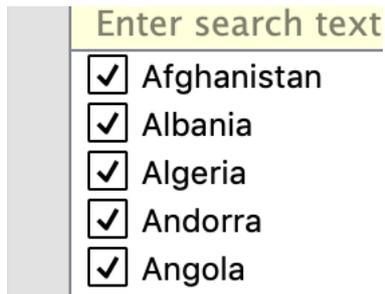


That's because the 'World' value is still in there and is throwing the numbers off. Let's filter the World out.

To do this go to the location pill and use the drop down to select filter



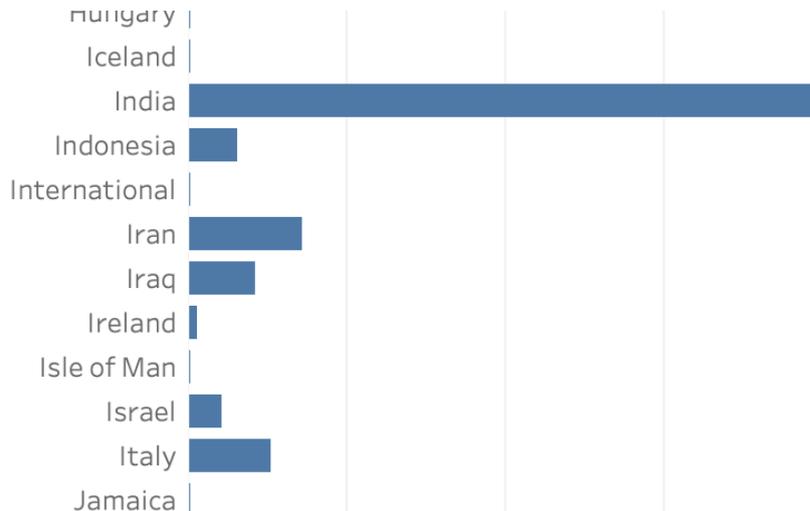
You'll now see check boxes for all the countries



Scroll down to find World and untick

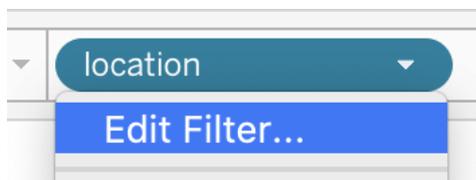


Then press ok. Now the scale is better.

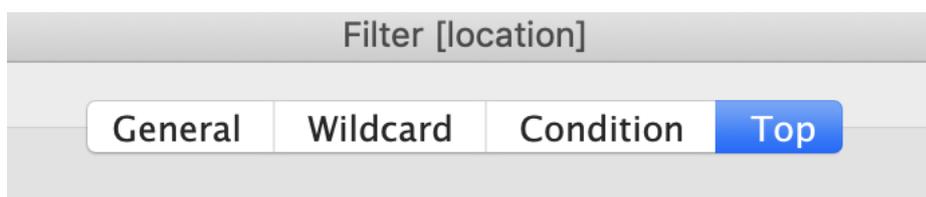


Now all we need to do is restrict the graph to the countries with the 10 highest cases.

To do this go back to the location pill and use the drop down to choose Edit Filter



In the filter pop up window, select Top

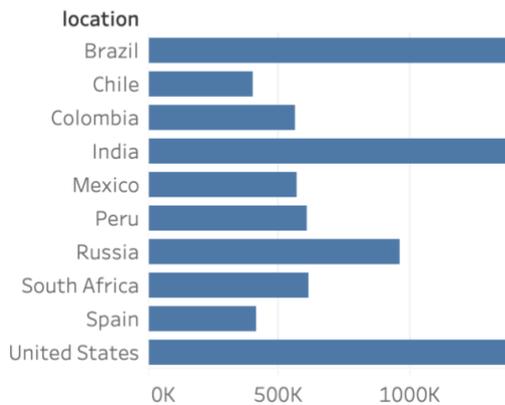


Then tick the radio button next to By field. The default setting should be Top 10 by new\_cases.

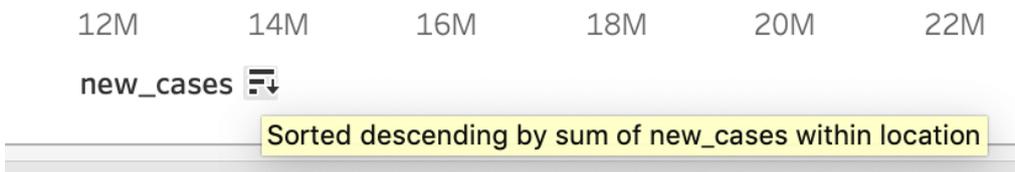
The screenshot shows a filter window with the following settings:  
- Radio button: **By field:** (selected)  
- Dropdown 1: **Top**  
- Input field: **10**  
- Dropdown 2: **by**  
- Dropdown 3: **new\_cases**  
- Dropdown 4: **Sum**

Press OK to close the filter window. You should now see this

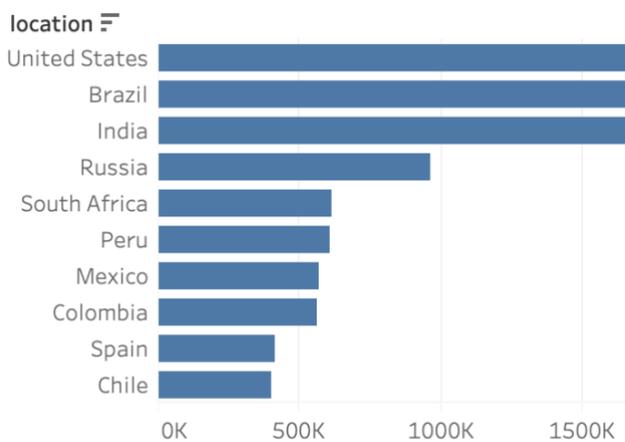
## Sheet 2



You can sort these top to bottom by going to the bottom of the page and choosing the sort option



After you sort descending the top of the graph looks like this:

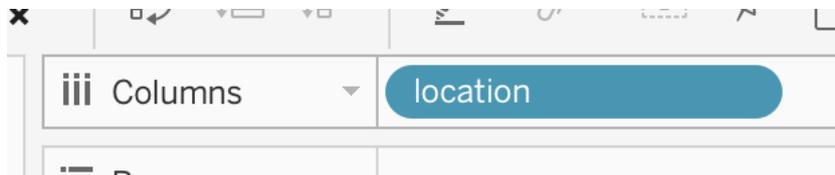


Label the sheet Top 10 Countries / Cases and open a new worksheet.

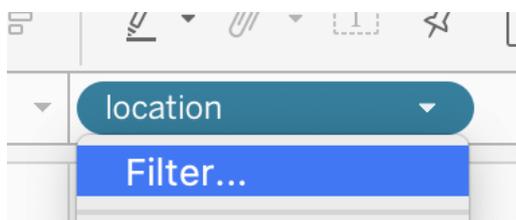
## Visualisation 3 – Australia Cases and Deaths Double line graph.

Start a new worksheet

Find location on the left of screen and drag it to the Columns area



Now let's filter the locations so we only see Australia. To do this use the drop down button on the location pill and select Filter



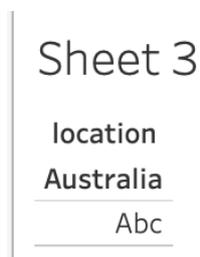
In the top up window first deselect all countries by pressing None



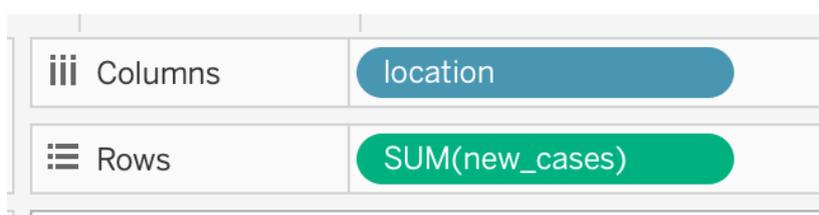
Then tick Australia and press OK.



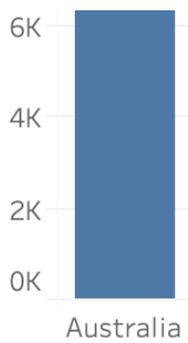
You should now see this



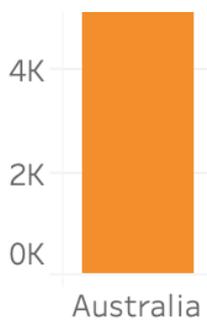
Now find new\_cases on the left of screen and drop it in Rows



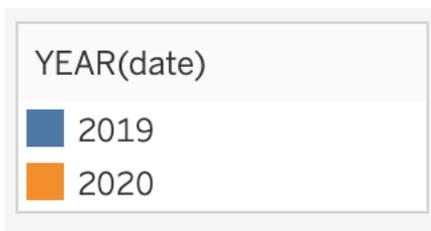
You should now see one column of data.



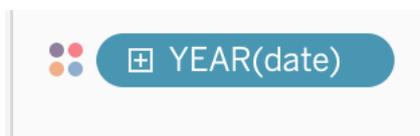
We want to convert this data into a line graph charting cases over time, so take the Date field from the left of screen and drag it over the blue graph and let it go. The column should turn orange



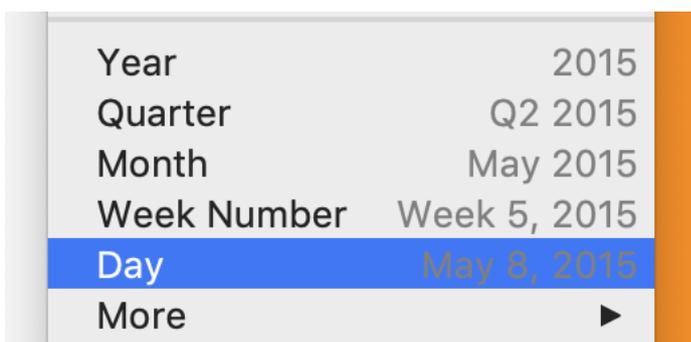
You'll notice on the top right there is now a key indicating 2019 and 2020.



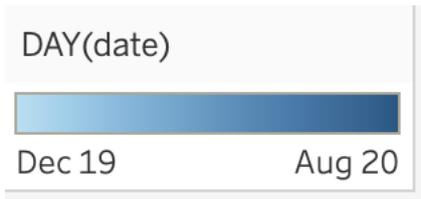
Now, go to the Year pill in the marks area



And use the drop down to select Day



There will still be a single bar column, but it will now be shaded different shades of blue.

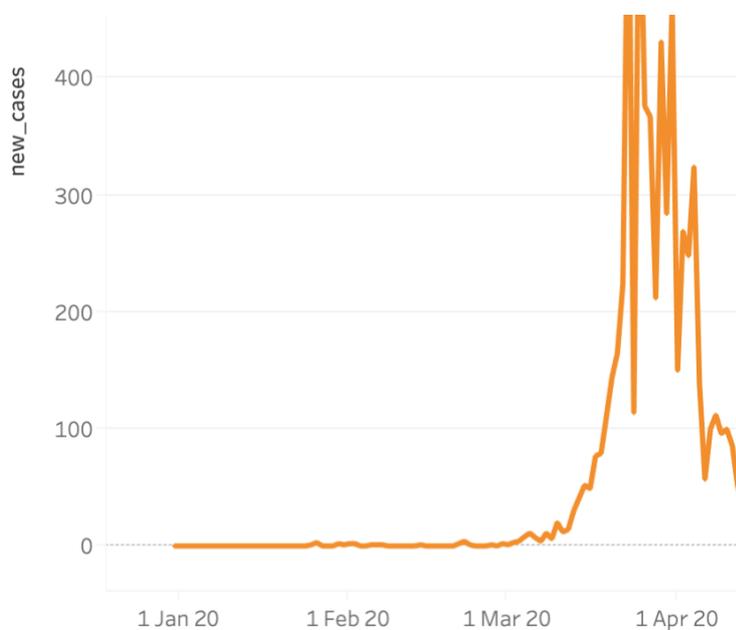


Now, go to Show Me and select line graph:

And switch to line graph



We now have a line graph of new cases by Date

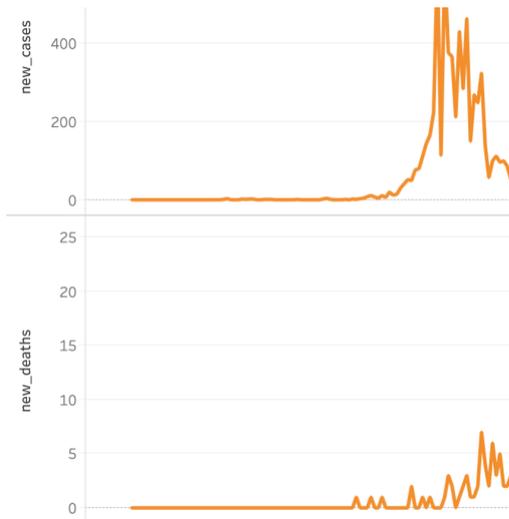


Now we want to add the cases of new deaths to this graph. To represent this information best it might be nice to have a double axis graph, because otherwise the deaths will be hard to read on this graph (because thankfully the deaths are much less than the cases)

So, find the new\_deaths measure of the left of screen and drag it next to new\_cases in rows



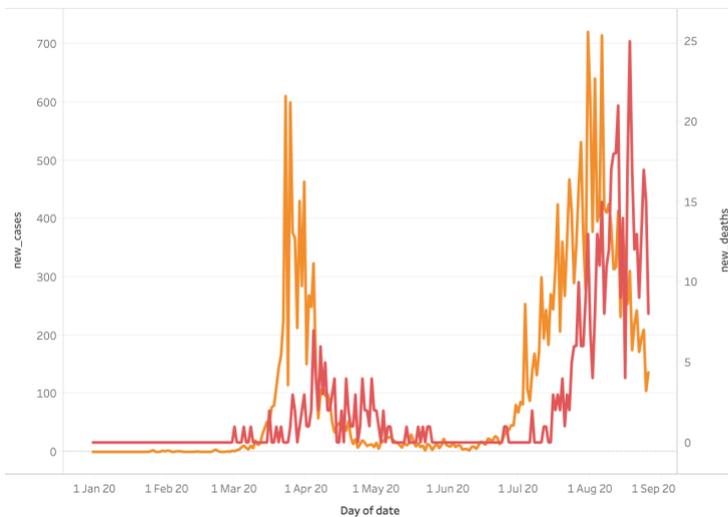
You'll now see two graphs side by side:



Now, go to Show Me and choose a double axis graph



The graphs are now overlaid, with cases in orange and deaths in red. Notice the scale on the right for deaths.



This scale is a bit deceptive because it makes it look like deaths and cases are equivalent. However we don't want to have cases so small they can't be read at all. So, we'll make the cases scale go from 0-1000 and deaths go from 0-100.

To do this double click on the label saying new\_cases on the left of the graph. An edit axis pop up window should open. In this window, select the 'Fixed' range radio button and then enter 0 as the Fixed Start point and 1000 as the fixed end point.

Edit Axis [new\_cases] ×

**General** Tick Marks

---

**Range**

- Automatic  Include zero
- Uniform axis range for all rows or columns
- Independent axis ranges for each row or column
- Fixed

Fixed start ▼	Fixed end ▼
0	1000

Now do the same thing for new\_deaths, with the end points being 0 and 100.

Fixed start ▼	Fixed end ▼
0	100

If you want to change the colour to something with a better contrast than red and orange choose the Color option in the marks area

Automatic ▼

<p style="margin: 0;">Color</p>	<p style="margin: 0;">Size</p>	<p style="margin: 0;">Label</p>
---------------------------------	--------------------------------	---------------------------------

Then select Edit Colors

Color - Drag dimensions or measures

**Color**

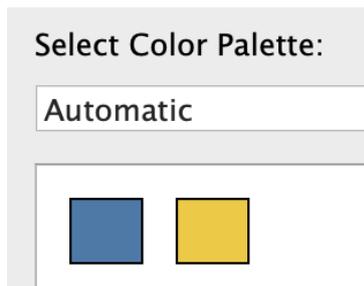
Edit Colors...

**Opacity**

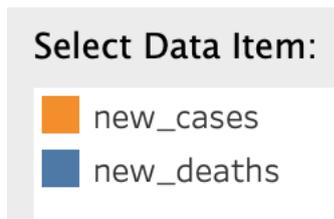
Click on the measure you want to change, eg: new\_deaths

	new_cases
	new_deaths

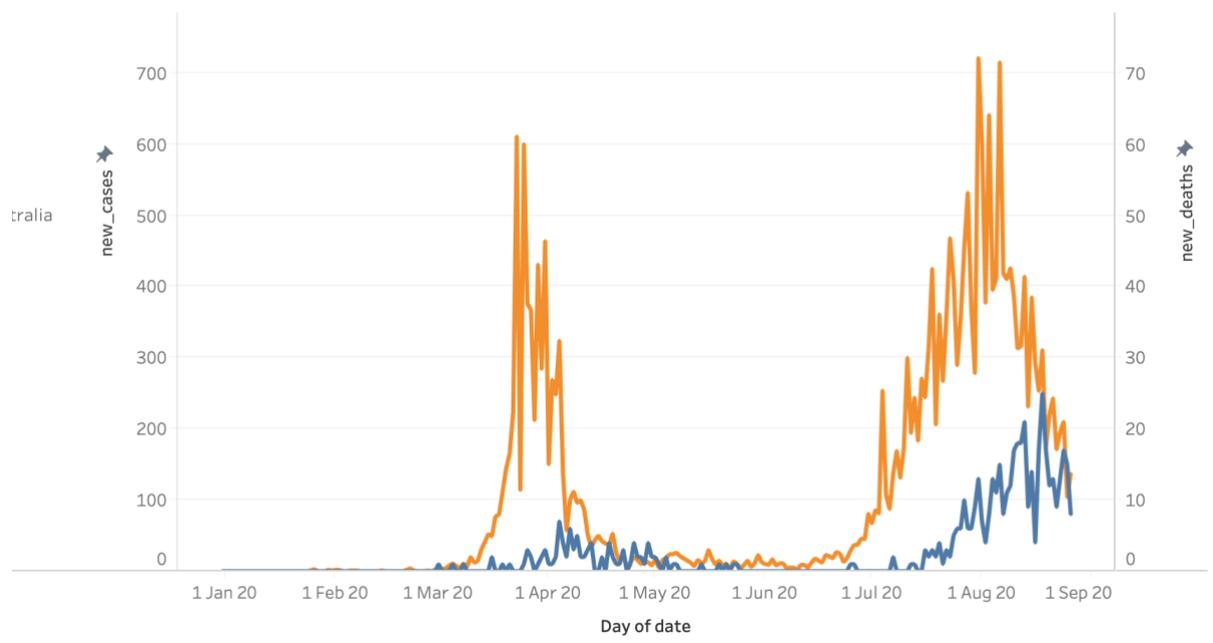
Then select a colour from the right of the Edit Color window, eg: blue



The fields should now look like this



And the graph should look like this:



## Visualisation 4– Deaths in the top 7 worst affected countries since first 100 deaths

In this graph we'll chart the way different countries have had their death counts change over time.

Start a new worksheet

Then begin by dragging date to columns and locations to rows.

	date	
location	2019	2020
Afghanistan	Abc	Abc
Albania		Abc

Now find the total\_deaths measure on the right of screen and drop it over the table.

The table should now look like this

	date	
location	2019	2020
Afghanistan	0	87,967
Albania		11,627
Algeria	0	116,211
Andorra		6,943
Angola		3,190

Again we'll have to exclude the World figure from the stats.

To do this scroll down in the list until you see World then click on it and press exclude

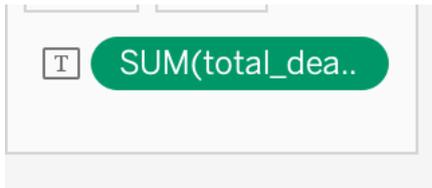
western Sahara		
World	0	64,46
Yemen		
Zambia		

✓ Keep Only   ✗ Exclude

Looking at these stats we can see they look unusually high.

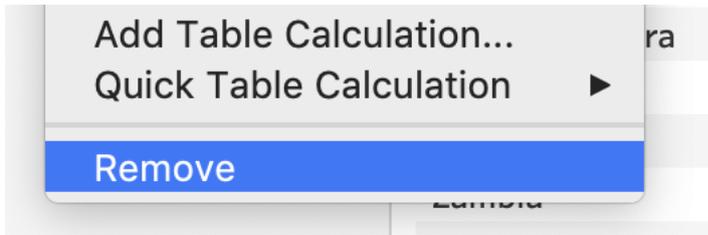
Ukraine		141,265
United Arab Emirates	0	35,810
United Kingdom	0	4,978,495
United States	0	15,654,925

That's because the deaths measure is set to sum, meaning it keeps adding up the running total every day, creating an erroneous count.

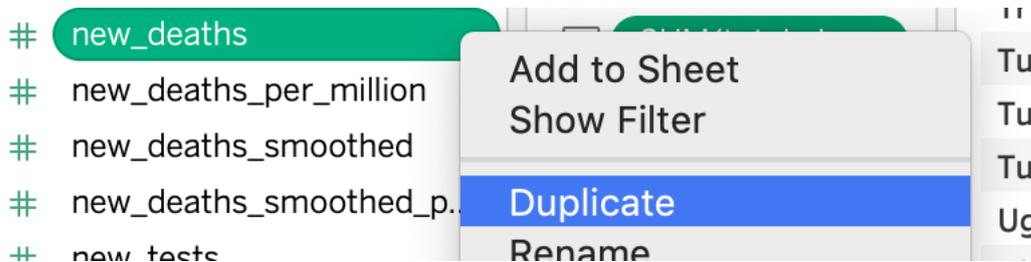


So what we might have to do is configure the way this data point is being tallied.

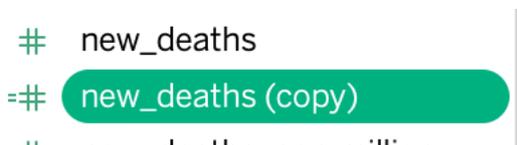
The first thing we'll do is remove the sum pill by using the drop down and choosing remove



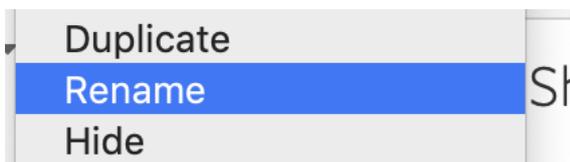
Then what I'll do is make a duplicate of the new\_deaths measure on the left so I don't foul up that data. To do this go to new\_deaths and use the drop down to select Duplicate



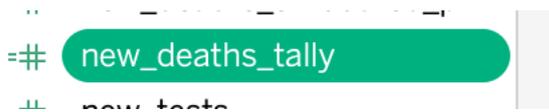
You should now have two measures – new\_deaths and new\_deaths\_copy



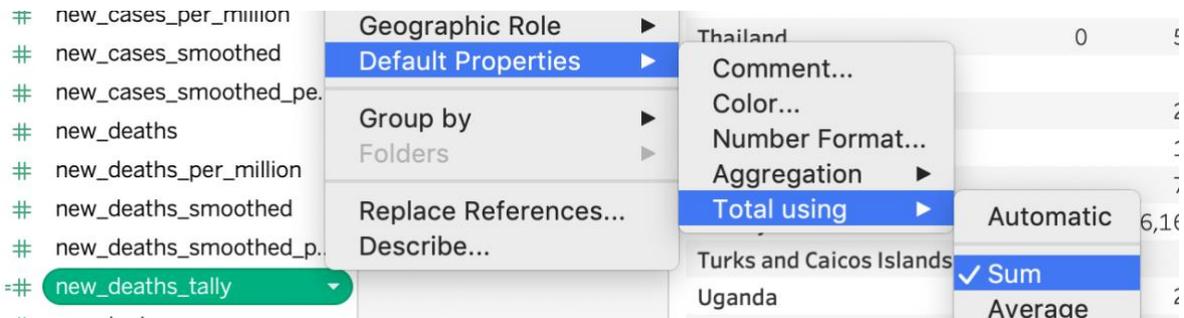
Now use the drop down on the new\_deaths\_copy measure to choose Rename



And rename the new measure new\_deaths\_tally



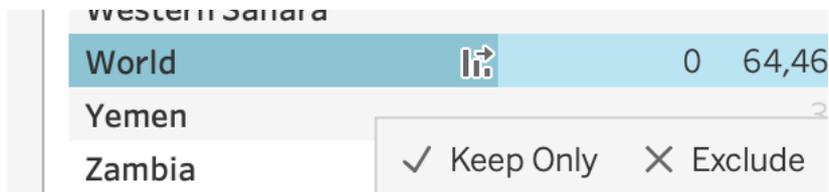
Now use the drop down by new\_deaths\_tally and select Default Properties>Total Using>Sum



Then drop the new\_deaths\_tally measure back over the graph. The numbers should look more accurate (if a bit out of date now)

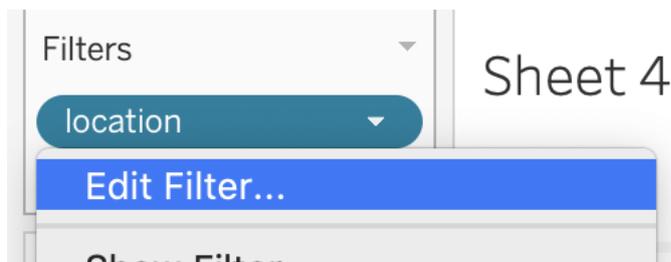
United Arab Emirates	0	377
United Kingdom	0	41,449
United States	0	178,486
United States Virgin Islan..		14
..		12

We'll need to exclude World data from the list again:



Now we want to restrict the results to the countries with the 7 highest total deaths. The number 7 is arbitrary, and was just chosen so it doesn't crowd the graph too much.

To limit to the top 7 go to the Location pill in the Filters section, and use the drop down to select Edit Filter.



In the filter pop up box choose Top, then select the By Field radio button. Select 7 to be the Top amount. Ensure the top 7 is for the new\_deaths\_tally measure. Press OK

General Wildcard Condition **Top**

None  
 By field:

Top 7 by

new\_deaths\_tally Sum

We should now see only a few countries list

location	date	
	2019	2020
Brazil	0	116,580
France	0	30,544
India	0	59,449
Italy	0	35,445
Mexico	0	61,450
United Kingdom	0	41,449
United States	0	178,486

We need to do one more thing before turning this into a graph. Notice at the moment the only time value we have is the year 2020. What this means is that if we make a graph we'll only have a single data point for time, and so we won't get a line graph.

location	Year of d..
	2020
Brazil	116,580
France	30,544
India	59,449

So, let's change the way the date is displayed to show month rather than year. To do this go to the Year pill in Columns and use the drop down to select month.

✓ Year	2015
Quarter	Q2 2015
Month	May 2015

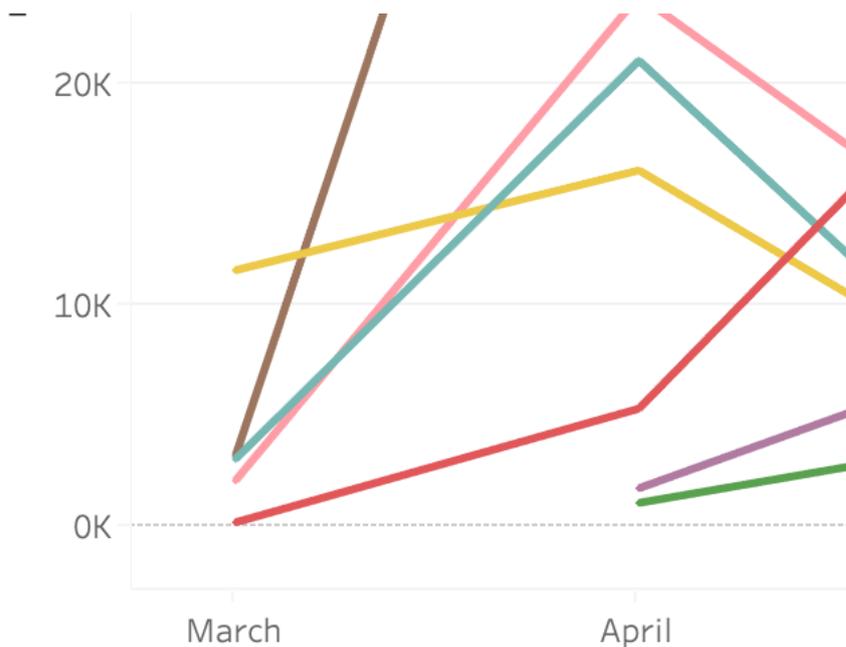
Our data now looks like this

location	Month of date					
	March 2020	April 2020	May 2020	June 2020	July 2020	August 2020
Brazil	159	5,307	23,368	29,480	32,949	25,317
France	3,022	21,063	4,684	1,042	441	290
India		1,042	4,090	11,729	18,854	23,702
Italy	11,570	16,091	5,658	1,404	388	313
Mexico		1,704	8,047	17,342	18,879	15,450
United Kingdom	2,050	23,999	11,336	2,956	828	280
United States	3,170	57,796	42,815	22,359	25,930	26,416

Now go to Show Me and select Line graph



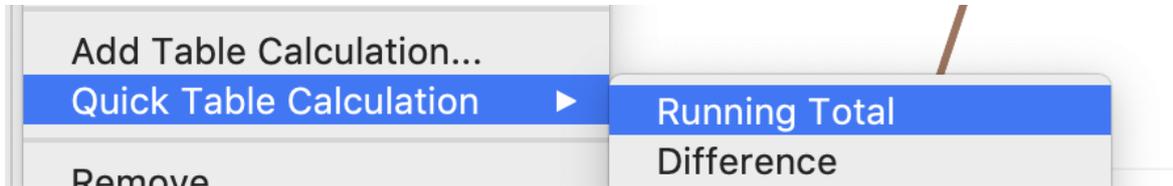
We're almost there now. The only problem is the numbers are unique for each month, whereas we wanted them to tally (ie: only go up, and not go up and down)



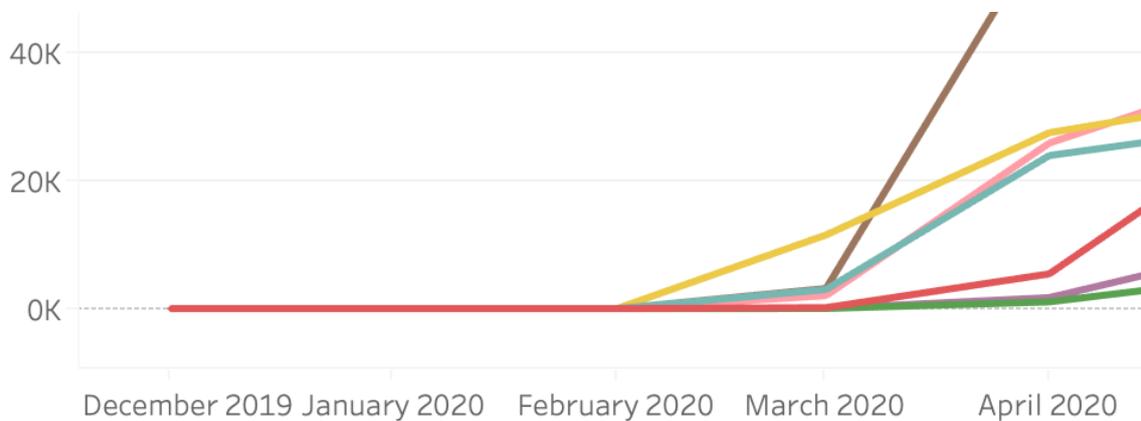
To fix this, go to the Sum Pill in Rows



and use the drop-down to select Quick Table Calculation > Running Total



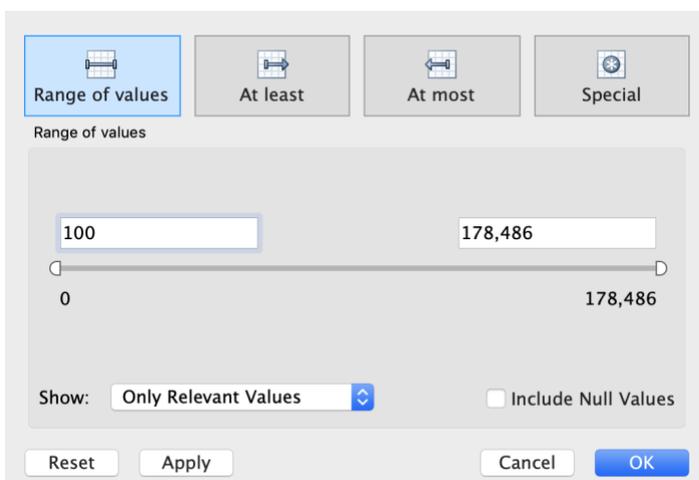
You can see that there's a lot of blank space so we can make the baseline 100 deaths, as with other graphs you might have seen online.



To do this go to the Sum pill in Marks and use the drop down to select Filter

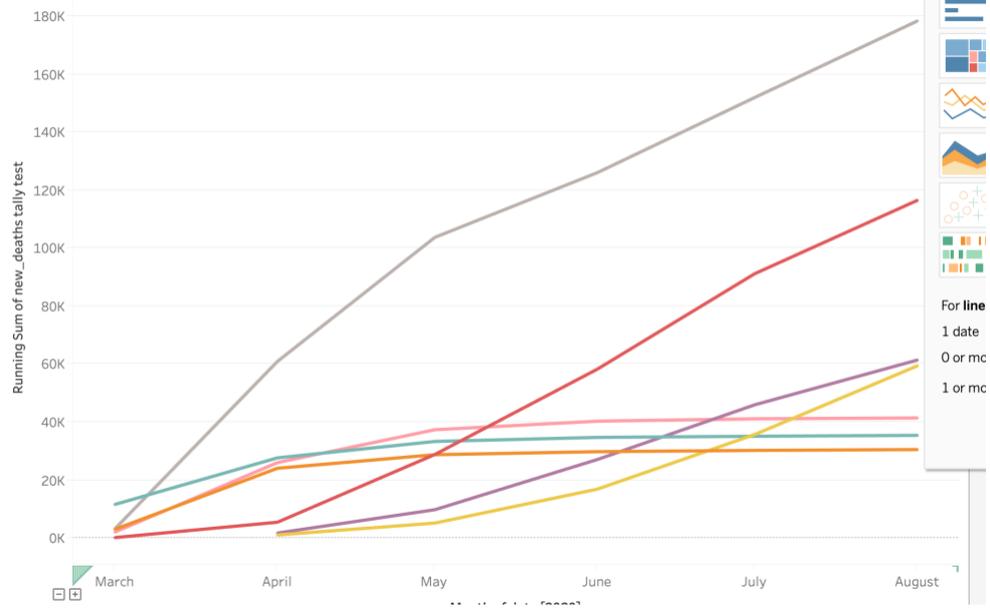


In the pop up window change the zero in the left hand value to 100 and press OK.



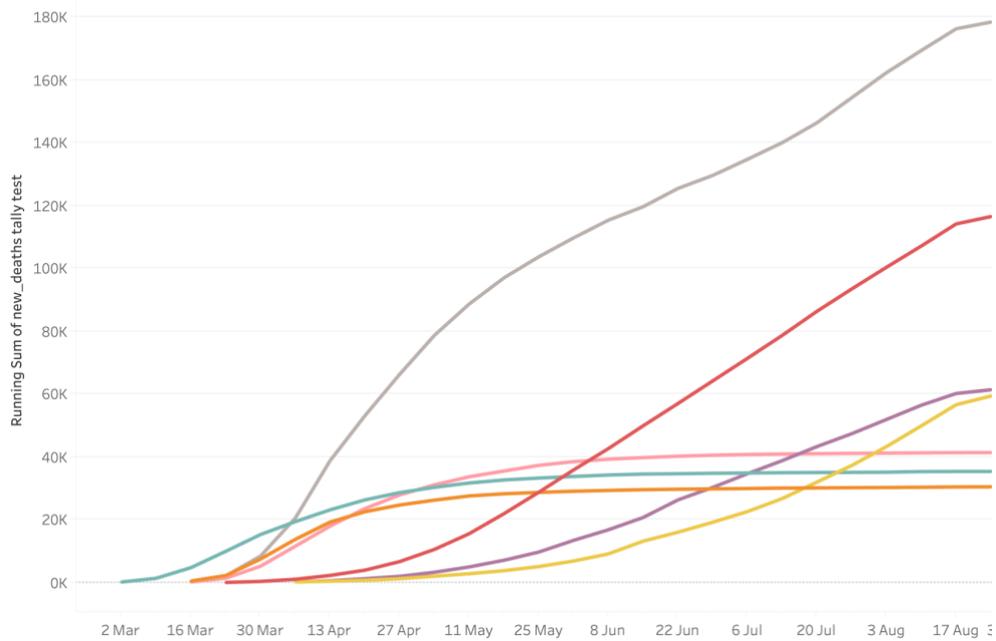
## After filtering

death count test



If you want the data to be more granular you can change the date from months to weeks using the drop down on the new\_deaths pill:

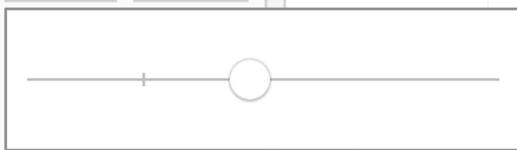
death count test



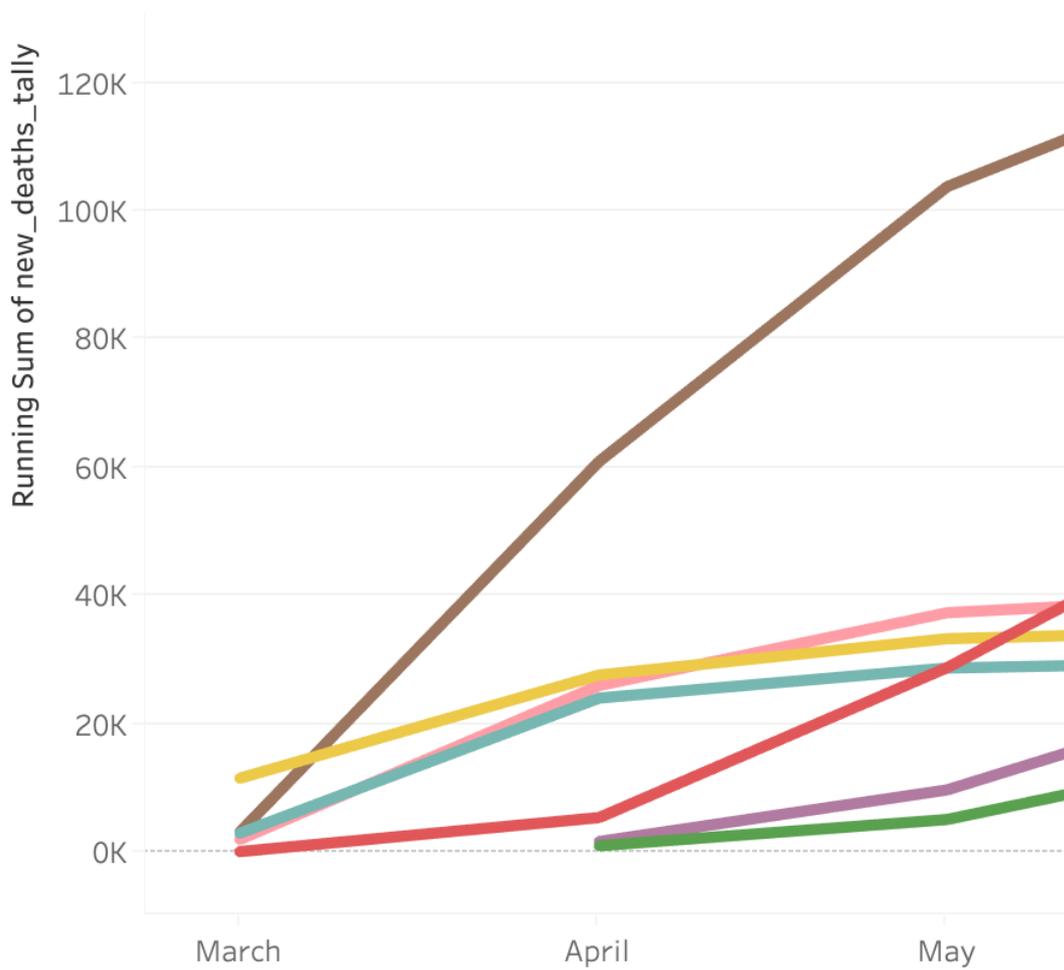
You can also change the thickness of the lines if you want. To do this go to Size in the marks area



Then drag the slider to the right



Graph now looks like this:



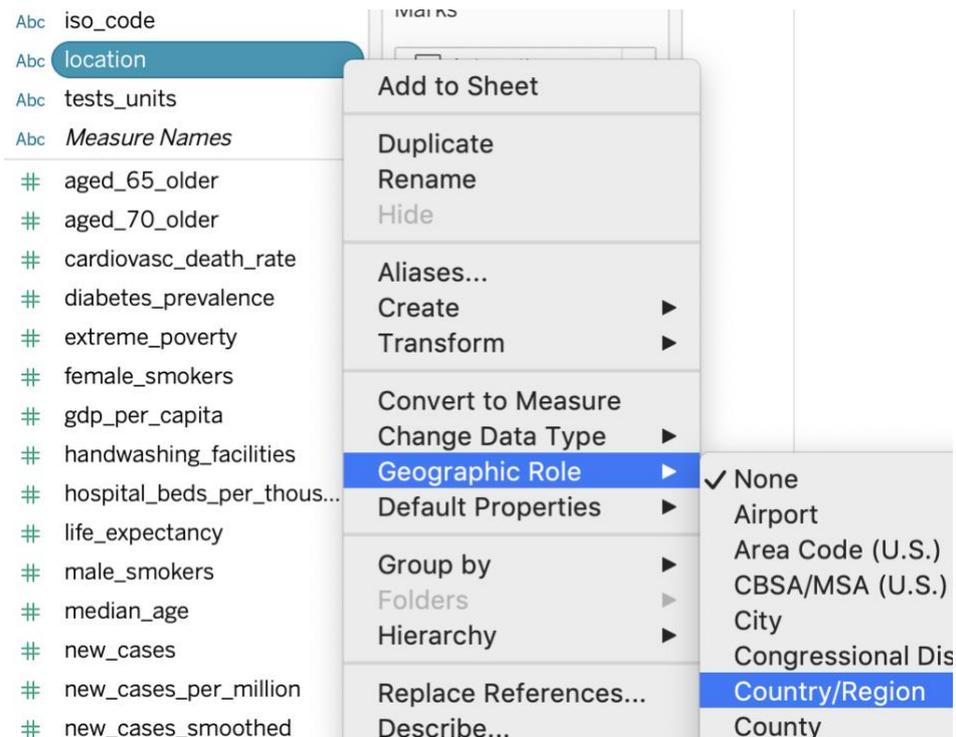
## Visualisation 5 – Map

Start a new worksheet

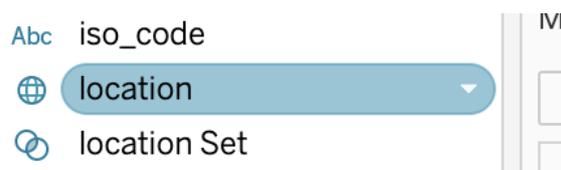
To make a map the first thing we need to do is make the value of location change from text into Geocoding.

To do this use the drop down inside the location pill on the left of screen and select

Geographic Role>Country



Now it's a geo-dimension (indicated by the globe icon).



Now drag the location pill over the large empty space beneath the rows area

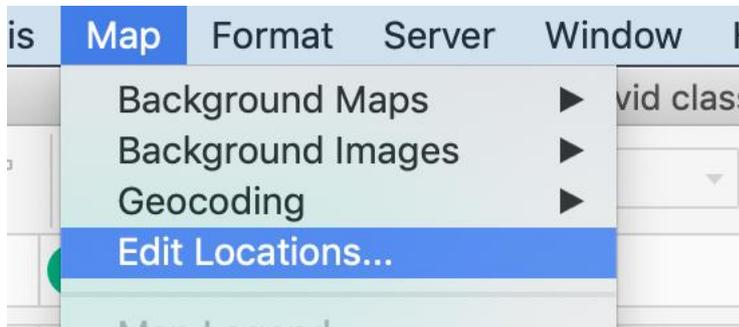
Drop field here

You should now see a world map. The blue dots represent countries



It's worth checking to see if there are any bad matches. To do this go to

Map>Edit Locations



There are 4 bad matches

Your Data	Matching Location ▲
International	<b>Unrecognized</b>
Timor	<b>Unrecognized</b>
Vatican	<b>Unrecognized</b>
World	<b>Unrecognized</b>
Afghanistan	Afghanistan

Tidy these up if you can

Filter location and exclude World

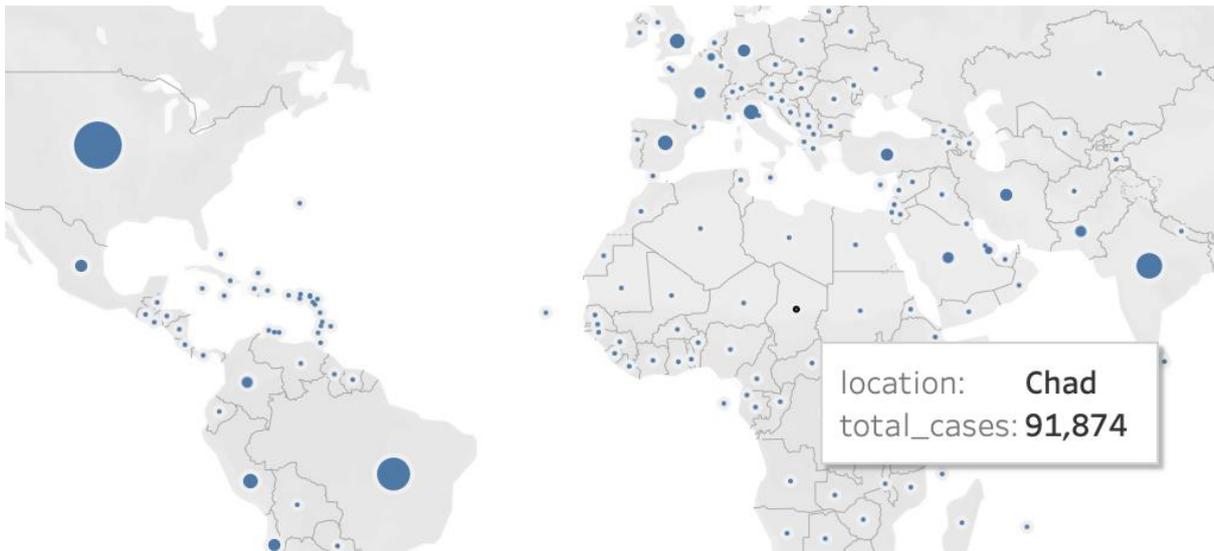


Then go back to map>edit locations and fix

Your Data	Matching Location
Timor	<b>Timor Leste</b>
Vatican	<b>Vatican City</b>
Afghanistan	Afghanistan

Now that we've done that, let's drop the total\_cases measure onto the map.

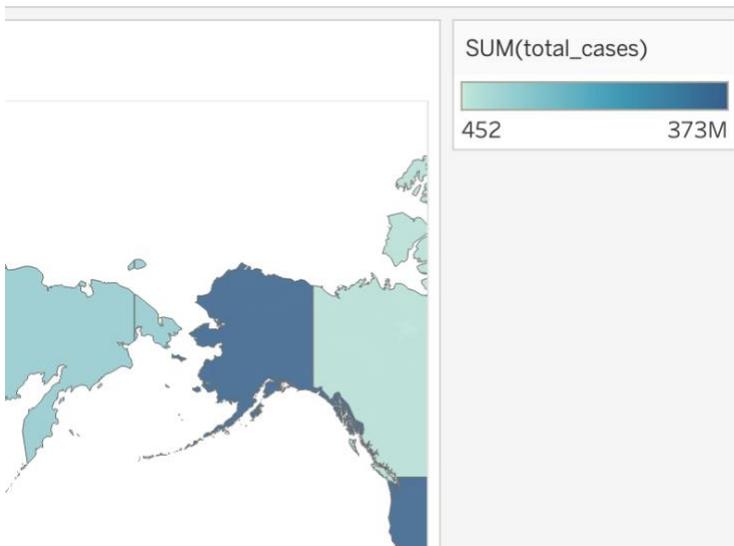
Total\_cases-map now looks like this, with the dot size representing large amounts of cases.



If you prefer you can go to Show Me to view a heat map (the one on the right below)

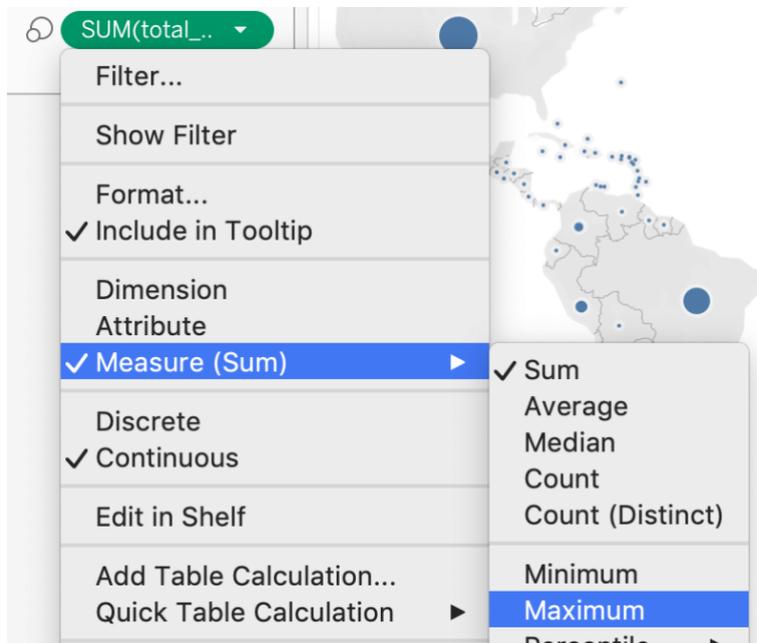


The map will now be shaded in. If you minimise the Show Me button you'll see the scale – which looks off (too high)

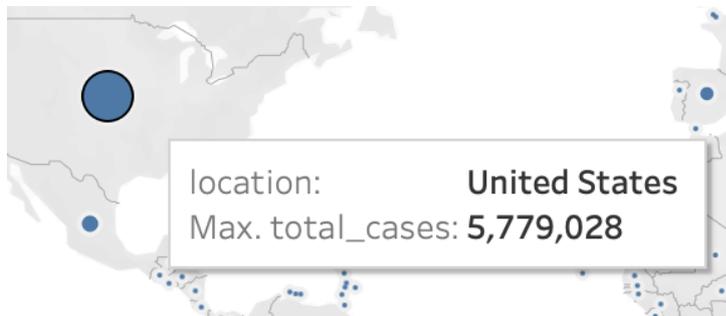


So, let's try and change this to a running tally so it doesn't add every day together.

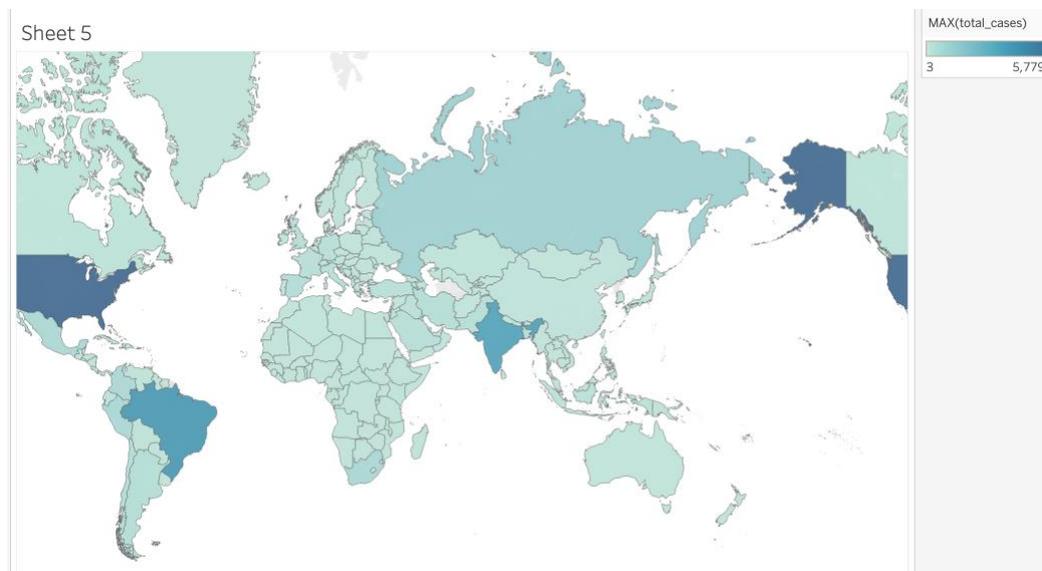
To do this let's try using the Drop down by the SUM pill to change the measure from sum to maximum.



Data looks better now (remember this data is from earlier in the year)

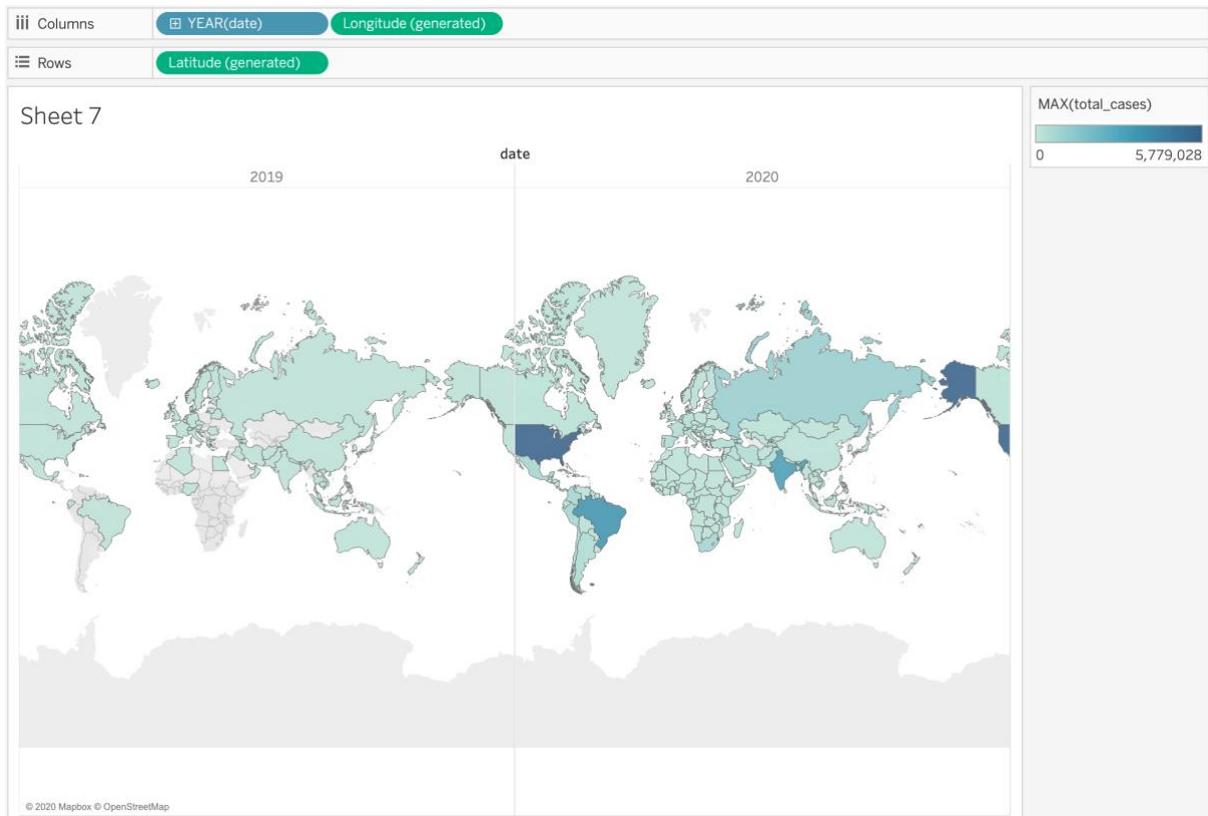


Use show me to Change back to a heat map

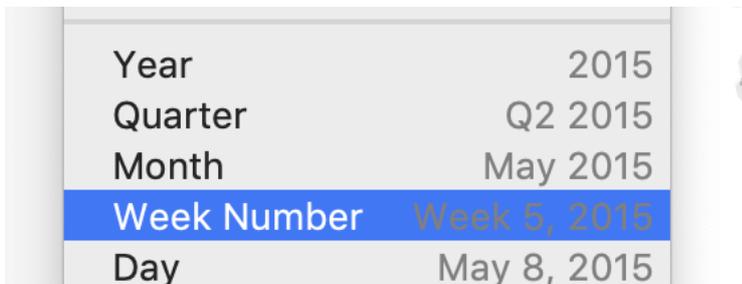


Now let's see if we have time to do something cool.

Drag the date pill onto columns to the left of longitude. You'll see screen split into 2 – one for 2019 and one for 2020



If you use the drop down by the Year pill in Marks and choose weeks



Now we see many pages of the heat graph



This isn't useful to us in this form but it might be as an animation!

First thing we need to do is have this show as one map per page.

To do this drag Date from the Columns area...



...into the pages area

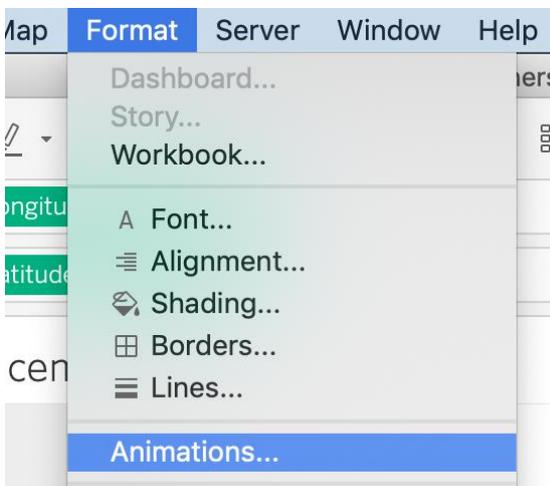


Now we have 1 sheet per week

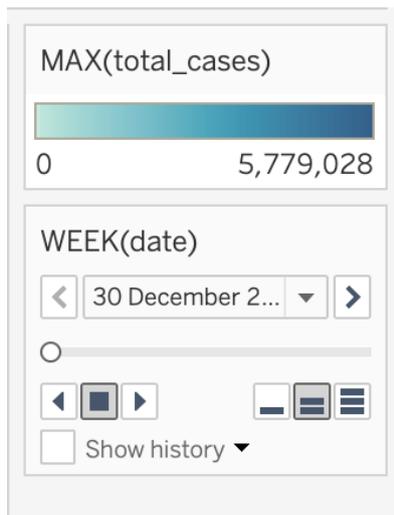
Sheet 5 - 30 December 2019



Then do Format>Animations

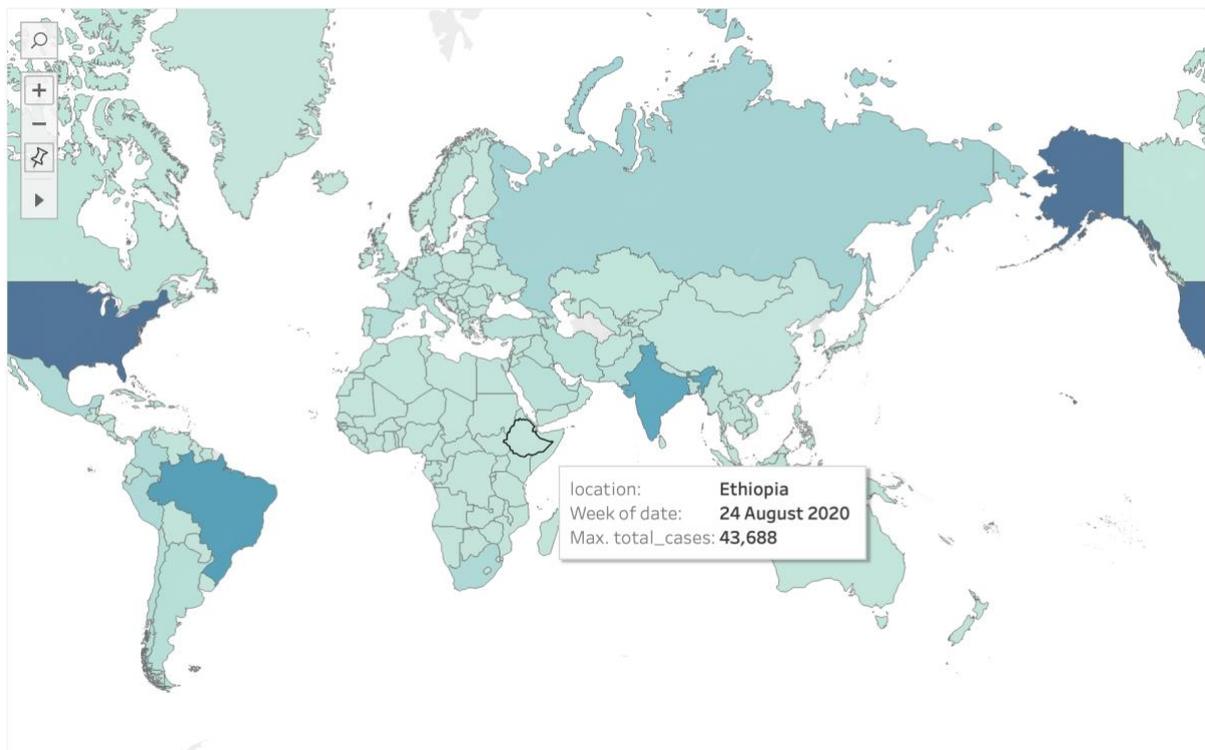


You'll now see this control panel on the right.



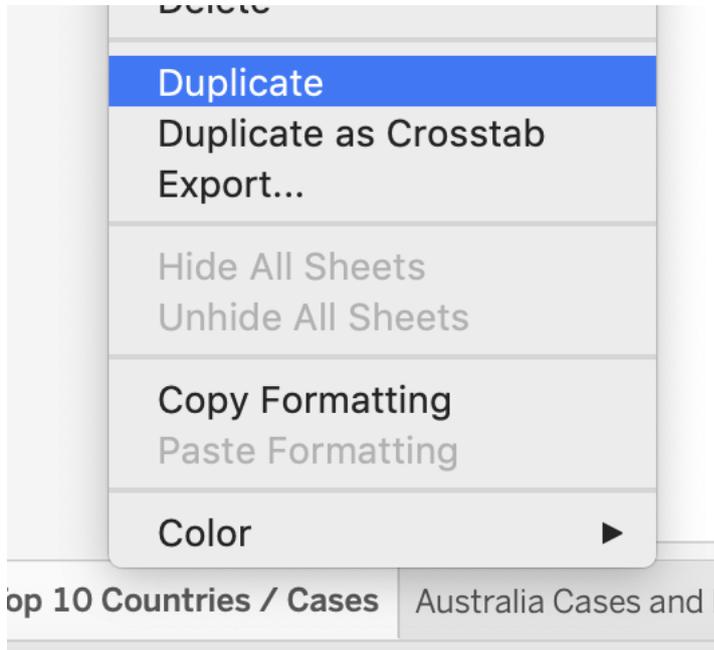
Press play and watch it animate week by week

Sheet 5 - 24 August 2020

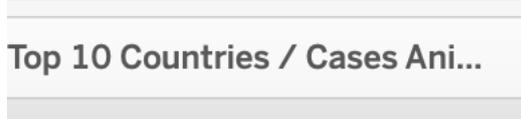


## Stretch Goal 1 – Testing Animated bar graph

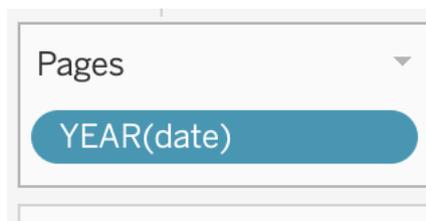
Let's begin by duplicating the Top 10 Countries by Cases worksheet



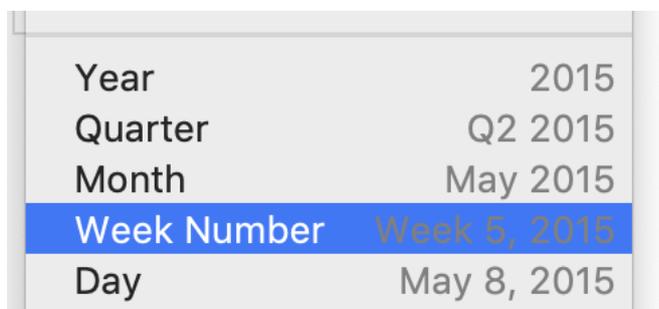
And rename it Top Countries/ Cases Animation



Then drop date to pages



Use the drop down by Year to change the Measure to weeks



You can press play on the button on the right now but it will show week by week stats only.

So, we need to do a cool calculation:

Use the drop down next to sum and choose Add Table Calculation, then fill it out like this:

Table Calculation ✕

Running Sum of new\_cases

**Calculation Type**

Running Total ▼

---

Sum ▼

---

**Compute Using**

Table (down)

Cell

**Specific Dimensions**

Week of date

location

Restarting every ▼

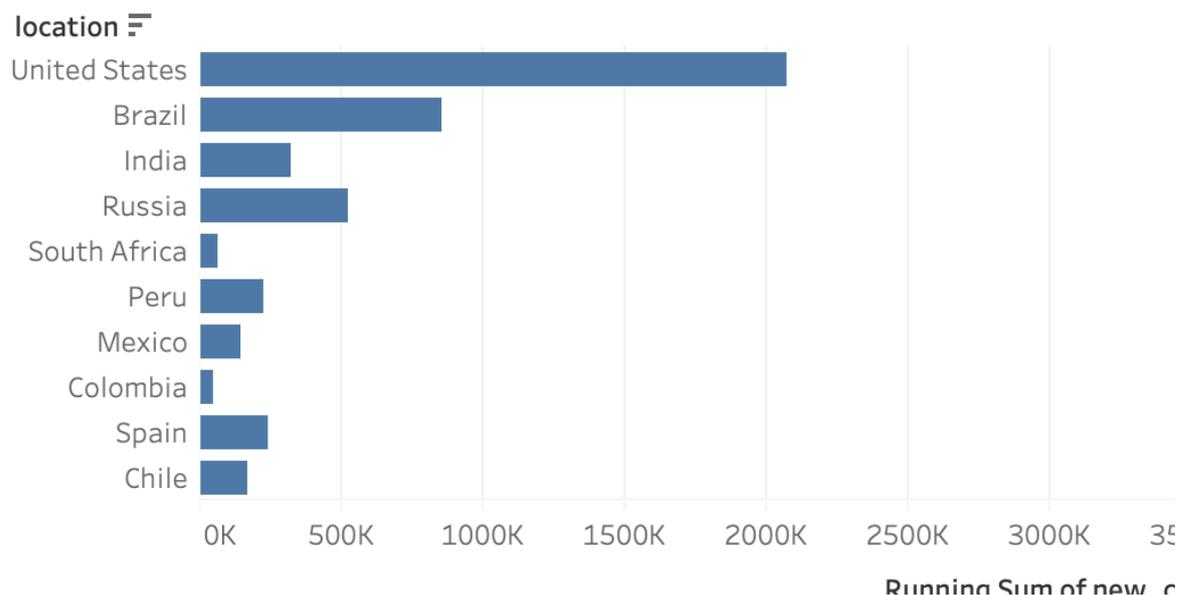
Sort order Specific Dimensions ▼

Add secondary calculation

Show calculation assistance

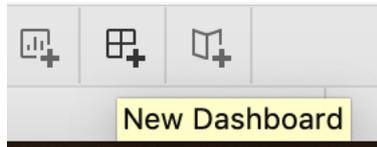
Now go and press play on the right and watch the graph grow

## Top 10 Countries / Cases Animation - 8 June 2020



## Stretch Goal 2 - Dashboard

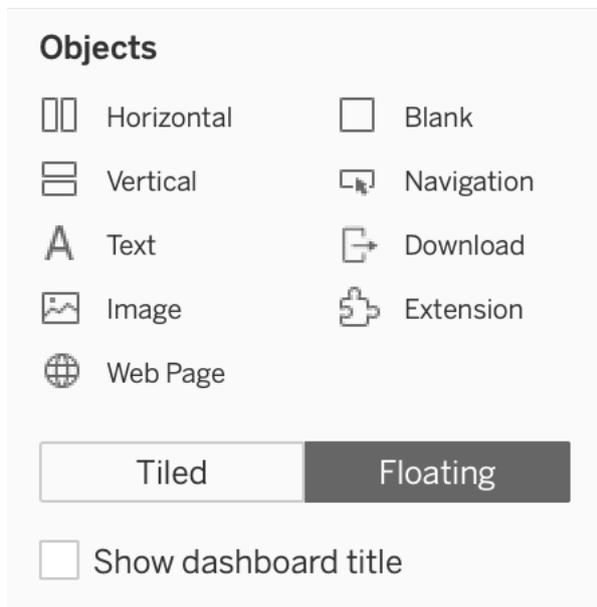
Instead of using New sheet, this time use New Dashboard



On the left are our visualisations

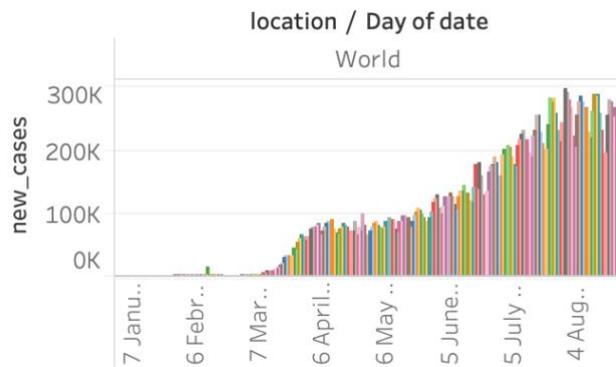


Before beginning, change the presentation mode in the Objects menu on the left of screen so that it reads Floating



Drag one of the sheets onto the big white space – a visualisation should appear in a floating window

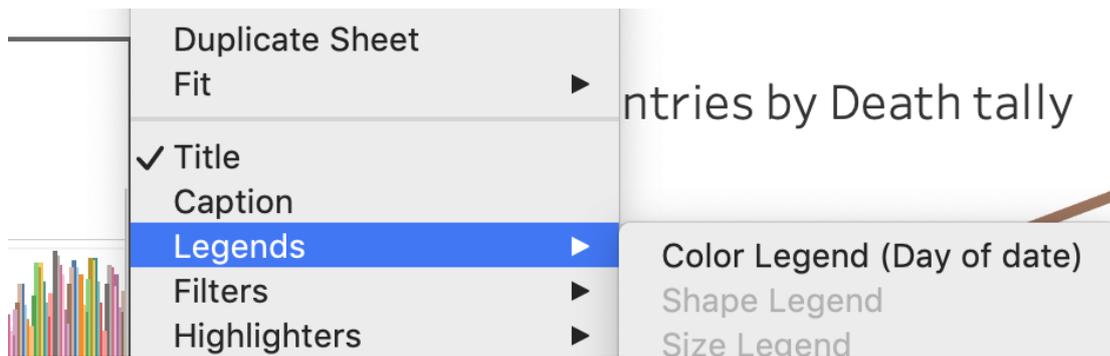
## World Cases/Day



You can remove an unneeded graph legend by clicking on the down arrow on the left of the box



And then selecting Legends> then untick Color Legend



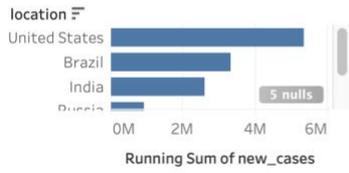
If you drag an animated visualisation onto the dashboard you'll see an attached player

## Week of date



You can use this to drive both animations at once!

Top 10 Countries / Cases Animation - 10 August 2020

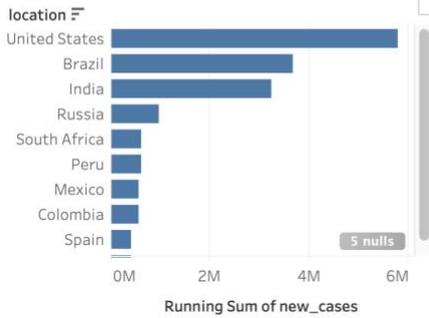


World Map Cases - 10 August 2020



The finished product!

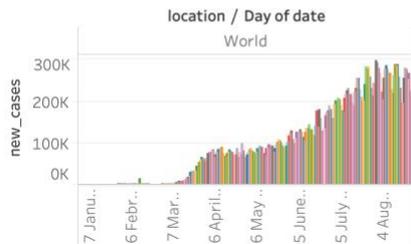
Top 10 Countries / Cases Animation - 24 August 2020



World Map Cases - 24 August 2020



World Cases/Day



Top 7 Countries by Death tally

