Milken Institute School of Public Health
THE GEORGE WASHINGTON UNIVERSITY

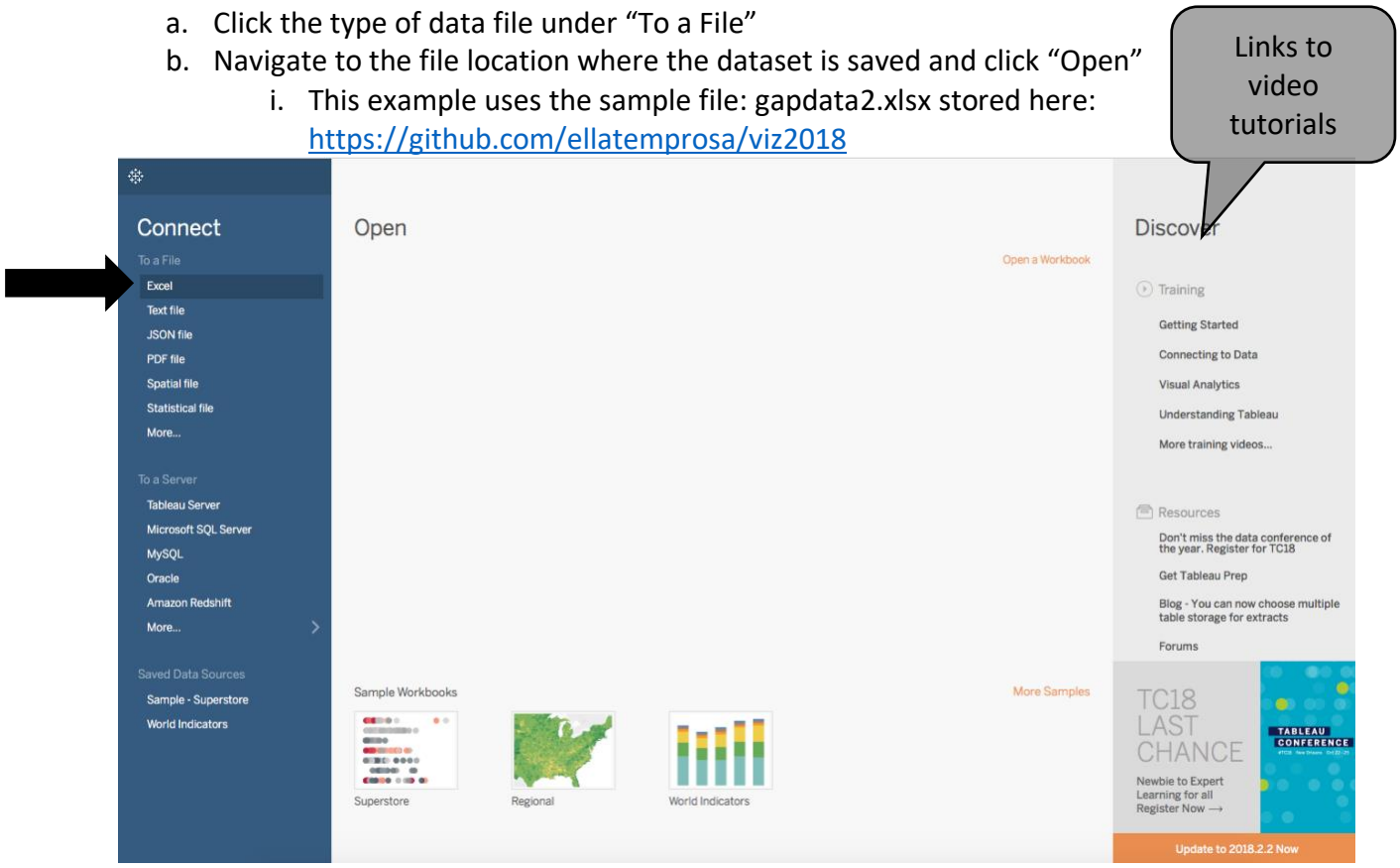
PUBH6299: Visual Exploration of Public Health Data
Handout #1

Fall 2018

Table of Contents

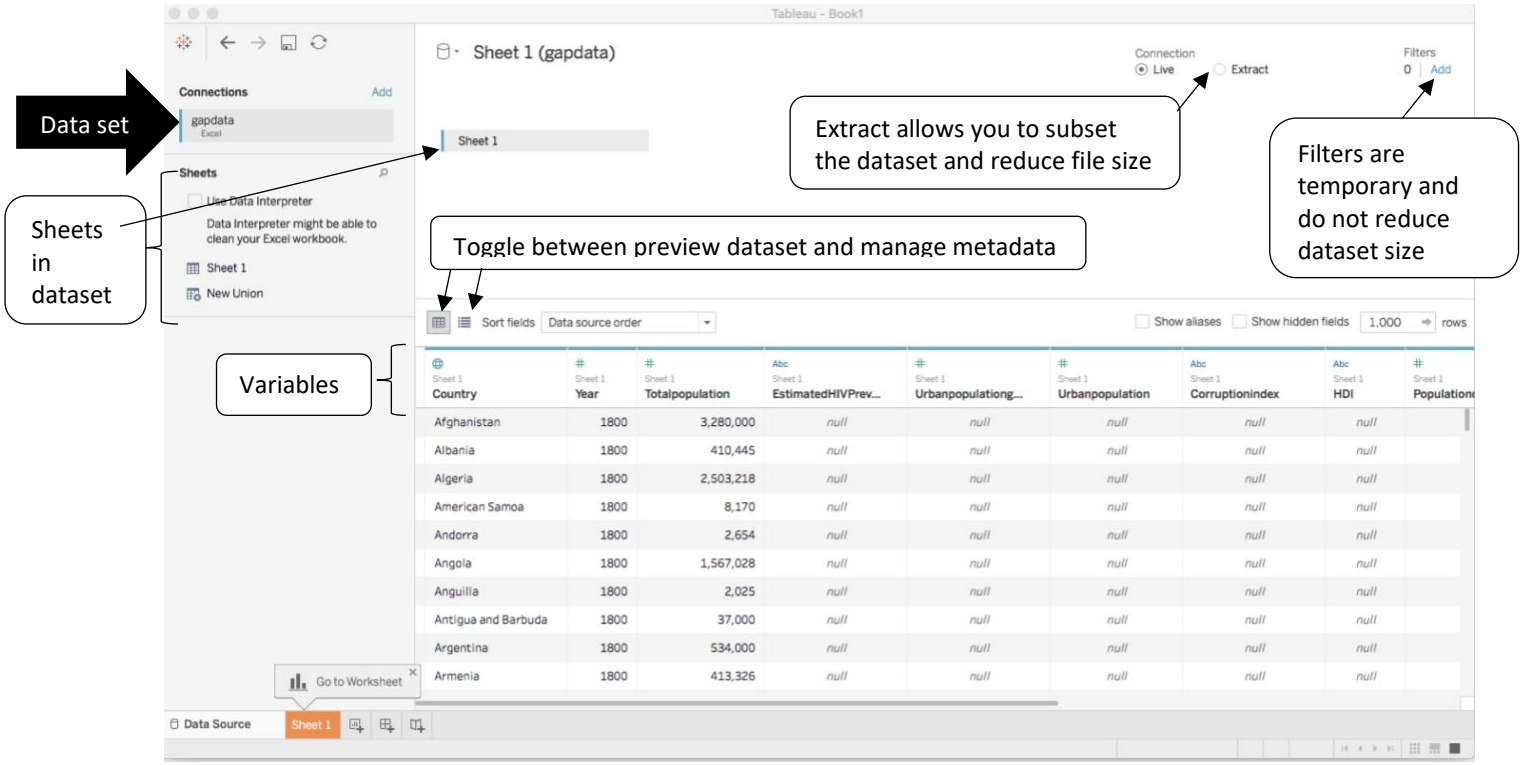
1. Connect to your data

a. Click the type of data file under "To a File"
b. Navigate to the file location where the dataset is saved and click "Open"
    i. This example uses the sample file: gapdata2.xlsx stored here: https://github.com/ellatemprosa/viz2018



2. Orientation to Data Source and Metadata

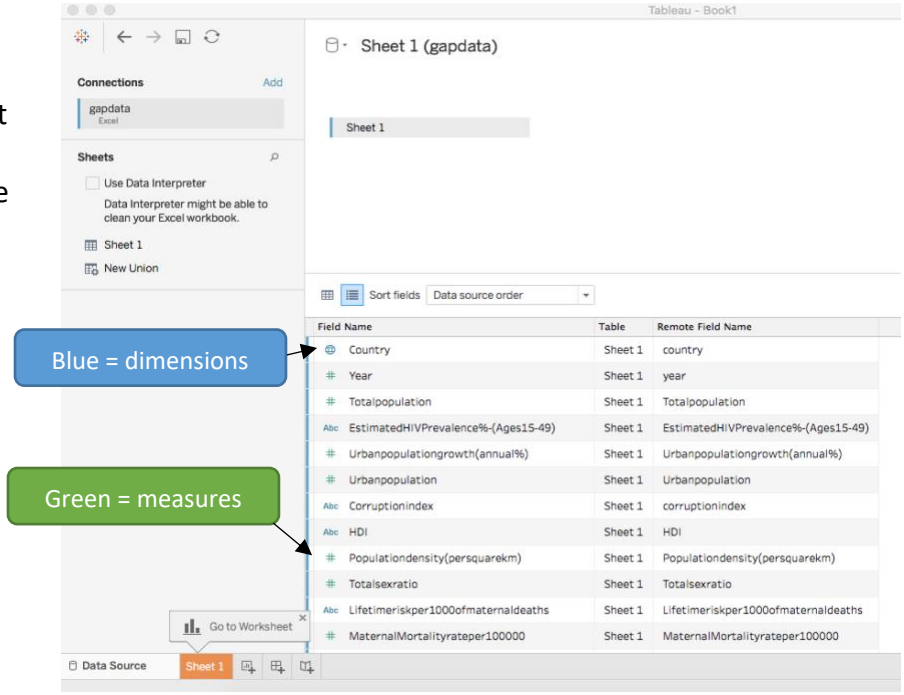
Once you've connected to a data source, you will be redirected to the Data Source tab with an overview of the dataset. The Data Source tab allows you the filter, extract and rename variables and data types, and view/edit metadata.



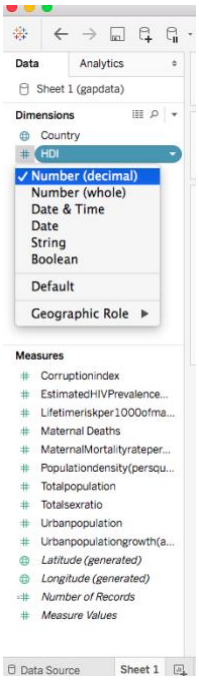
Manage metadata

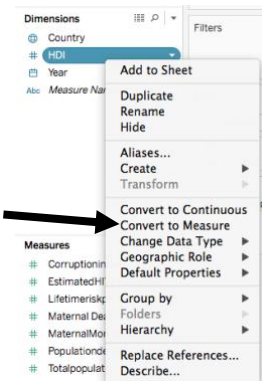
Data is stored as two main types in Tableau. In this view, you can correct data types that were automatically assigned by Tableau or you can make the corrections in Sheet 1.

- **Dimensions**: categorical, qualitative. **Abc,**
  - Examples: Country, State, Indicator, Patient ID
- **Measures**: quantitative, numeric. **#**
  - Examples: Results, Targets, Cost



Exercise 1: Update the metadata to correct data types





a) Change data type: Right click on a variable's data type icon to change the type of data.

   We're changing HDI (Human Development Index) from a String **(Abc)** to a Number (decimal). This converts the icon to a blue **#**.

b) Convert from Dimension to Measure: Right click on the variable name and select "**Convert to Measure".**

   This will move the variable from **Dimensions** to **Measures** and change the data type icon to a green **#**

c) Complete this step for other variables as needed.

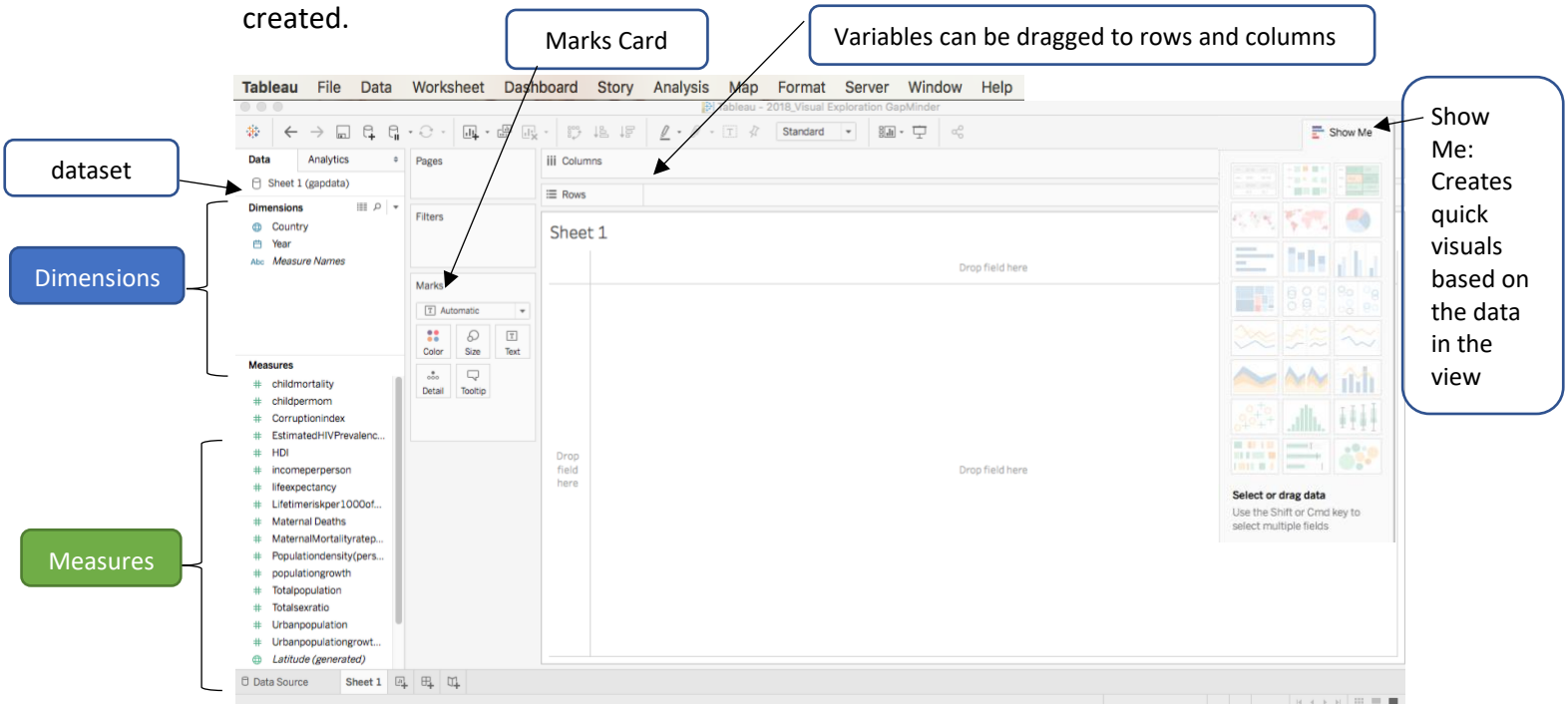
How to save:

a) click File → Save as Packaged Workbook (.twbx)

   By saving as a .twbx, this saves the dataset within Tableau and will allow easy transfer and sharing of files with your colleagues. If you save only as a .twb, then anyone you share the workbook with must have access to the same dataset and structure that you have to view what you've created.

3. Orientation to the interface

Click on Sheet 1. This is the main interface where graphs, tables, and visualizations are created.

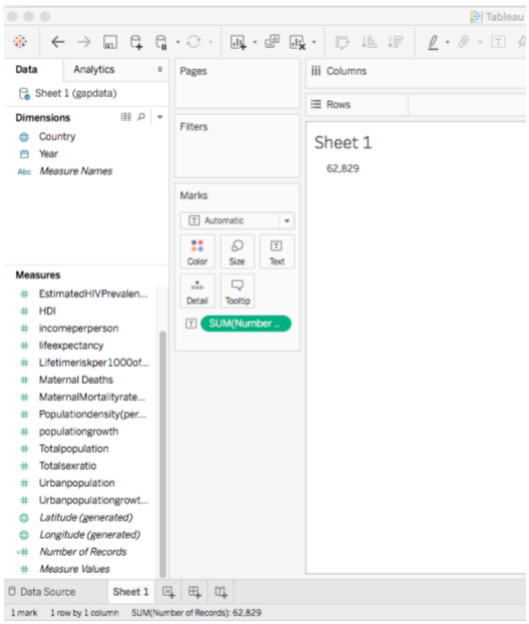


Data Import Check

It's always important to double check the data that was imported, including data types and number of records. Click on Sheet 1 at the bottom of the Tableau window.

1. In the Measures pane, drag *Number of Records* to Text in the Marks Card. By doing this, you'll there are 23,803 records in the dataset.

   This should correspond to the number of rows of data in your dataset and is always a good check to ensure the full dataset imported.

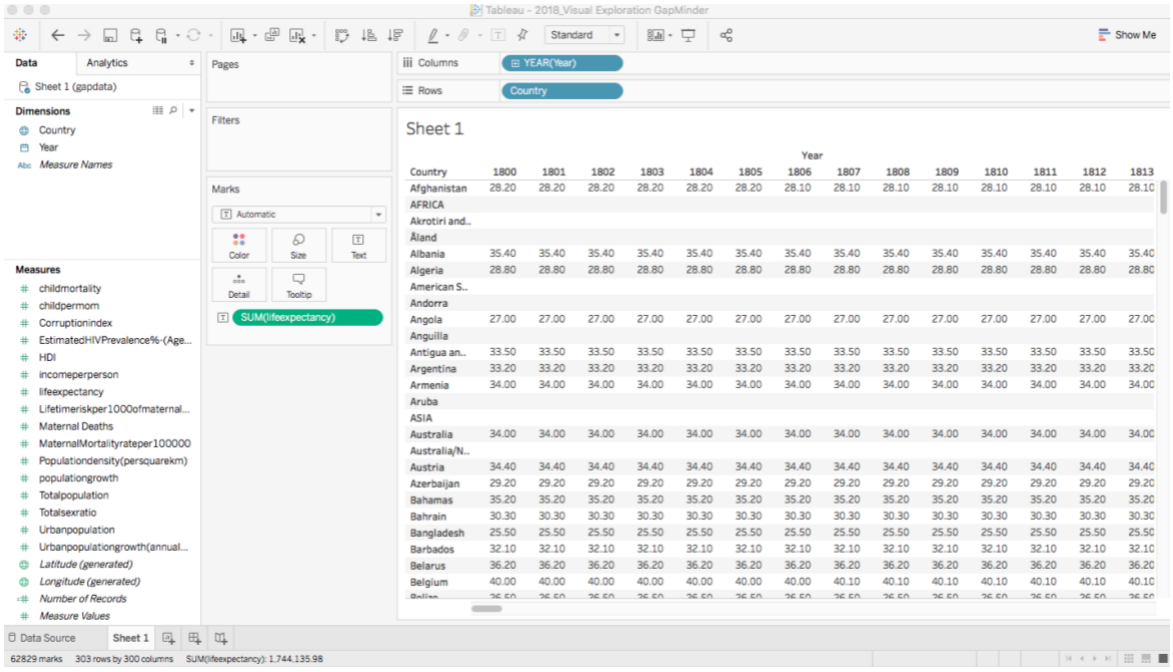


4. Create a table

Display the life expectancy per country.

a) From the Dimensions pane, drag Country to **Rows**
b) From the Dimensions pane, Drag Year to **Columns**
c) From the Measures pane, drag lifeexpectancy to **Text** under the Marks Card

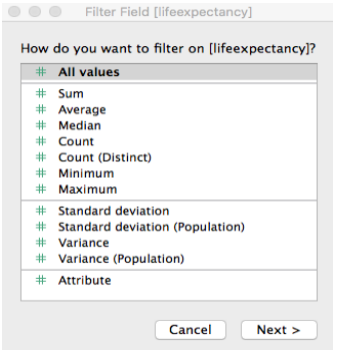
Your table should look similar to the one below. This is great, but this is a lot of data to digest.

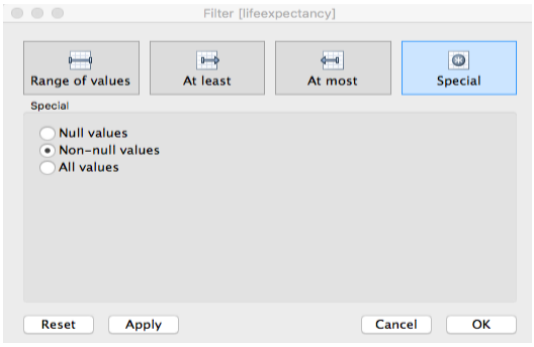


Add Filters to the table

Filter 1: Life Expectancy.

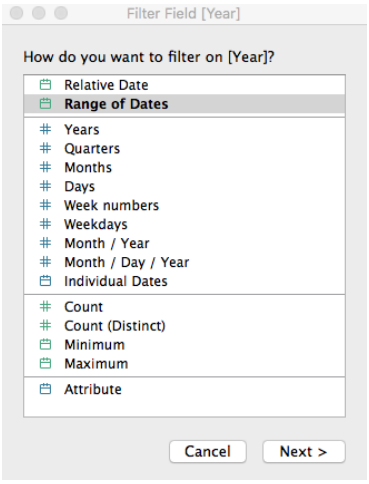
d) From the Measures pane, drag lifeexpectancy to the **Filters Card**
e) Select All Values, then click Next in the pop-up window
f) Click on Special at the top of the next window, and select **Non-null values**. This will show only data points for which there are values. Blanks will be excluded. Click **OK.**

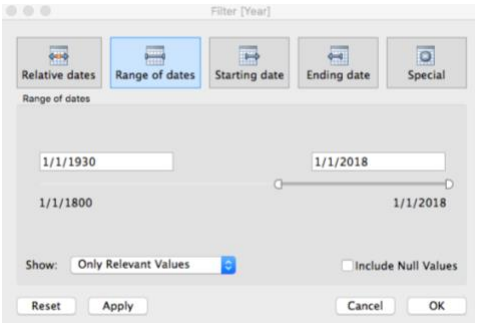




Filter 2: Year

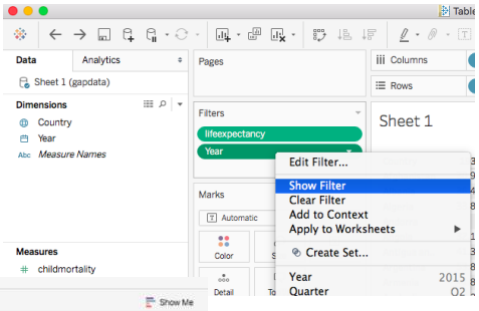
g) From the Dimension pane, drag Year to the **Filters Card**
h) Select **Range of Dates,** click Next
i) Change the dates to start from 1930 to 2018.
j) Click **OK**

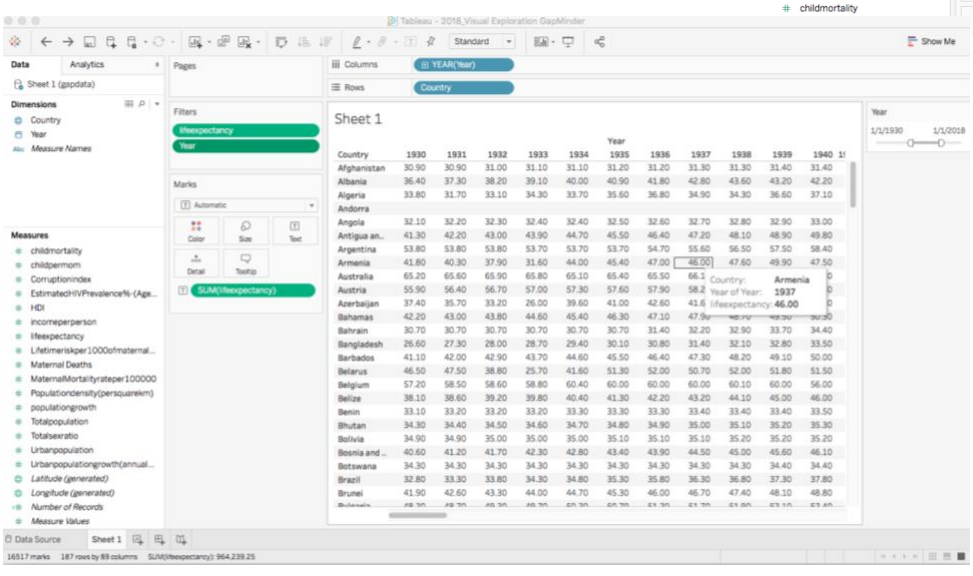




k) Right click on Year in the Filters Card or under Dimensions, select **Show Filter.**

This will display the Filter for Years in the top right of your view. You can now change the filters to show specific years or a range of years for the table.

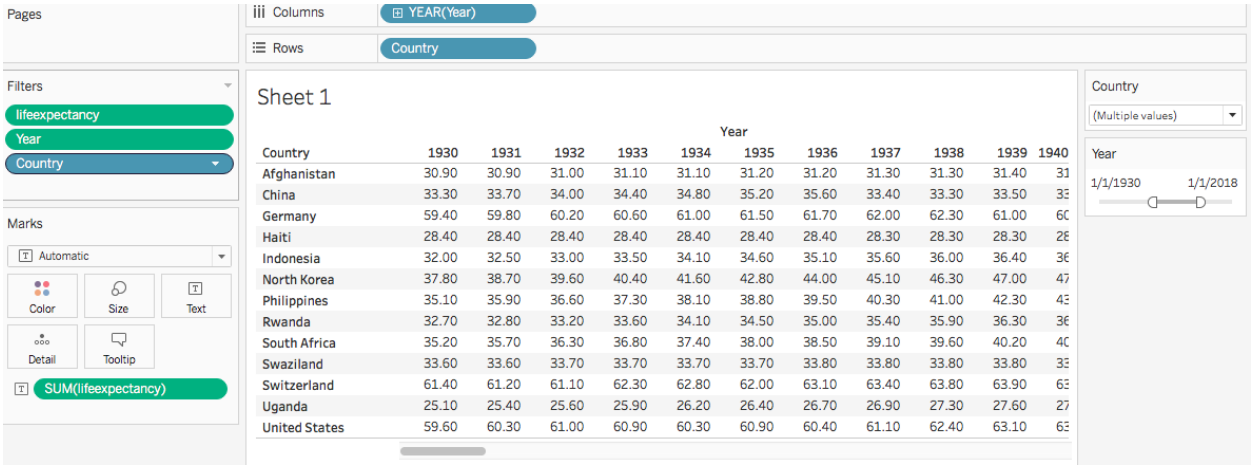




Exercise 2: Customize with Filters and Colors

Using the same process as we did for Years, create a filter that will limit the number of countries displayed.

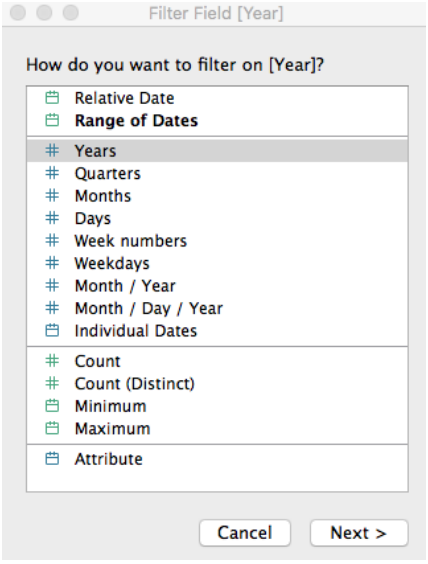
Let's only show the following countries: Afghanistan, China, Germany, Haiti, Indonesia, North Korea, Philippines, Rwanda, South Africa, Swaziland, Switzerland, Uganda, United States. Your Sheet should now look like the below:

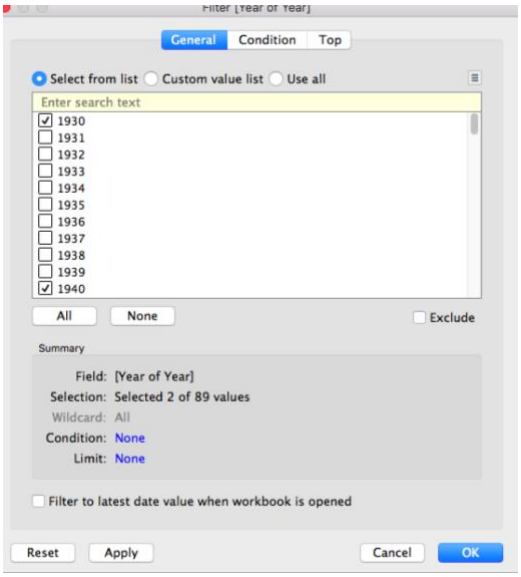


This is a lot of information and hard to decipher.

Let's filter the year to the following: 1920, 1930, 1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010, 2018.

a) From the Dimensions pane, drag Year to the **Filters card**
b) Select **Years**, click **Next**
c) Select only the relevant dates, and click **OK**

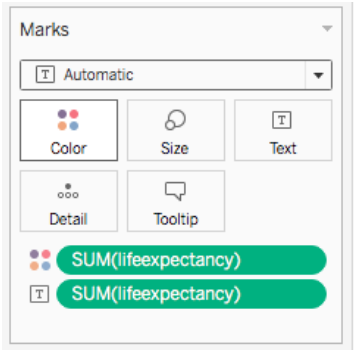




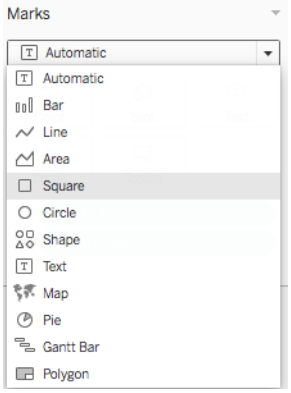
d) Your Sheet should now look as follows:

Transform to a highlighted table

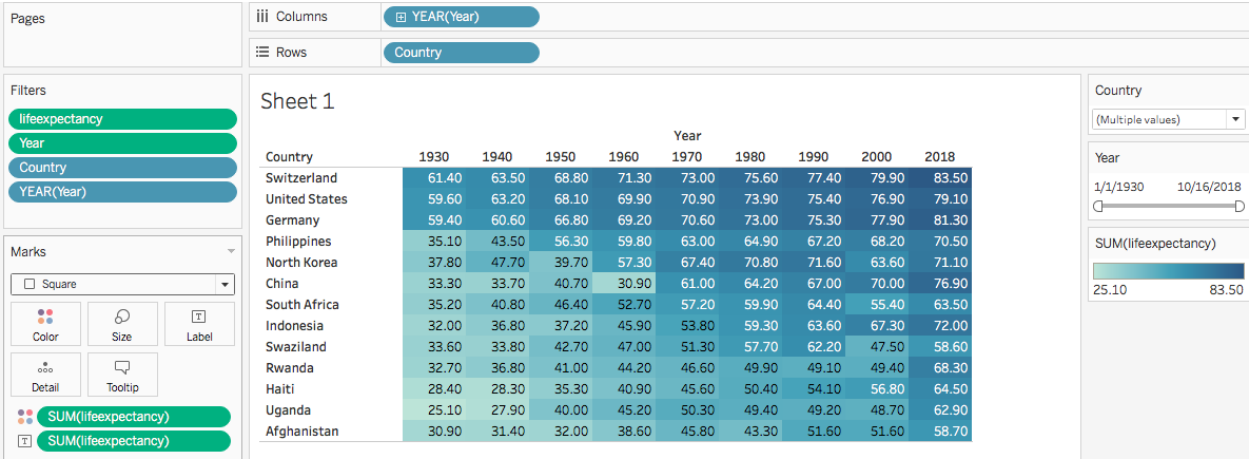
a) From the Measures pane, drag lifeexpectancy to Color under the Marks Card. This color codes the life expectancy values based on a color range.



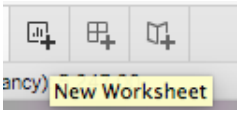
b) Click on the drop-down under the Marks Card and select "Square"



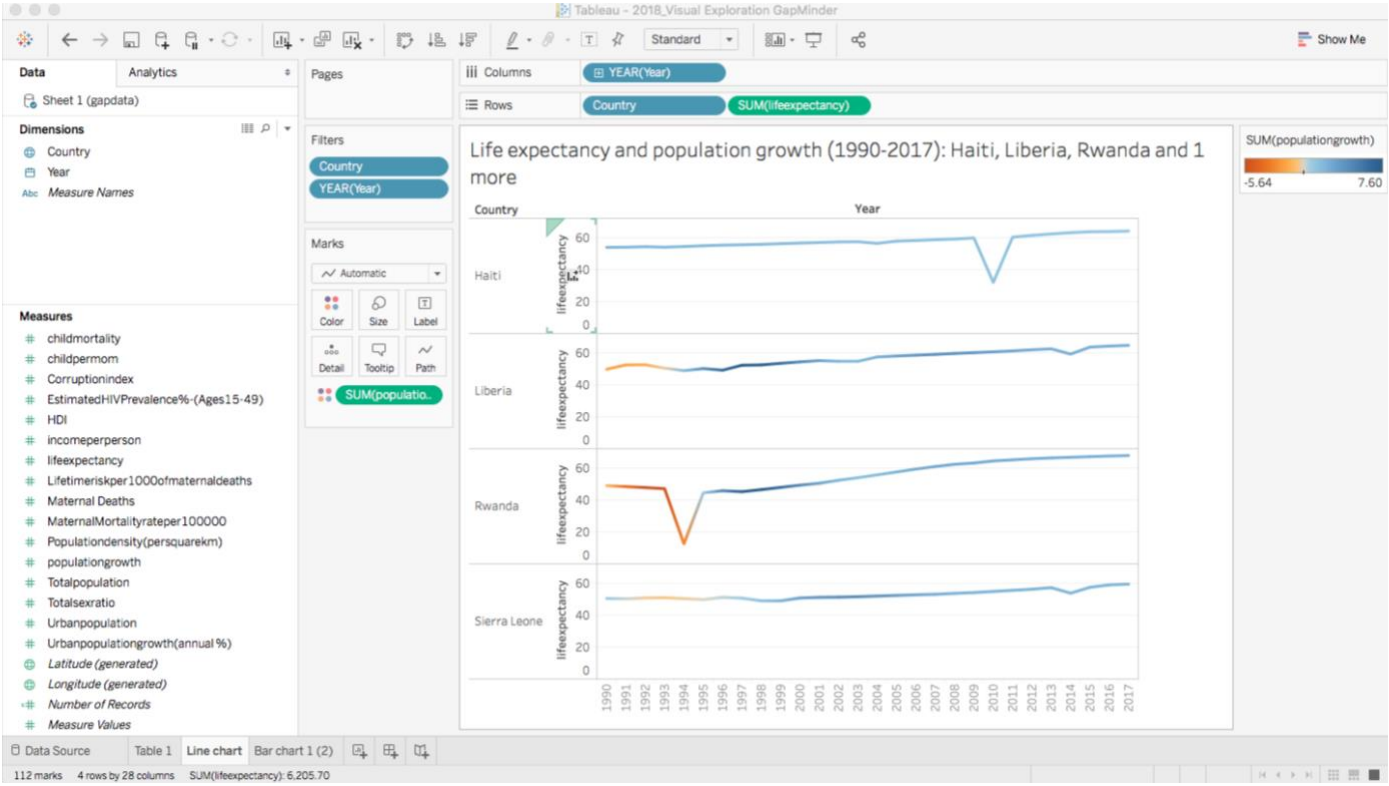
Now you have created a highlighted table, based on conditional formatting for the life expectancy values across all countries for the selected years.



5. Create a line graph: life expectancy and population growth by country

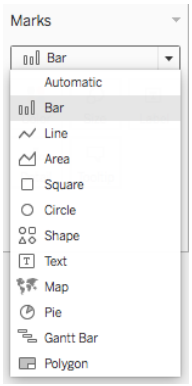


a) Create a new sheet and give it a title
b) Drag **Country to Rows**
c) Drag **Year to Columns**
d) Drag **Country to Filters** and filter for: Haiti, Liberia, Rwanda, Sierra Leone
e) Drag **Year to Filters** and filter for: all years between 1990-2017
f) Drag **#lifeexpectancy** to **Rows**
g) Drag **#populationgrowth** to **Color** in the Marks Card
h) Give your sheet a descriptive title

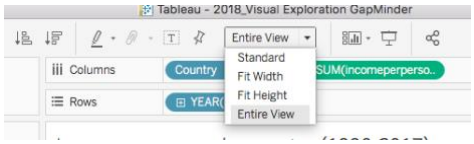


6. Create a bar chart: income per person by country

a) Drag **Country** to **Columns**
b) Drag **Year** to **Rows**
c) Drag **incomeperperson** to **Columns**
d) Drag **Country to Filters** and filter for: Haiti, Liberia, Rwanda, Sierra Leone
e) Drag **Year to Filters** and filter for: all years between 1990-2017
f) Change type of chart under Marks Card to: Bar

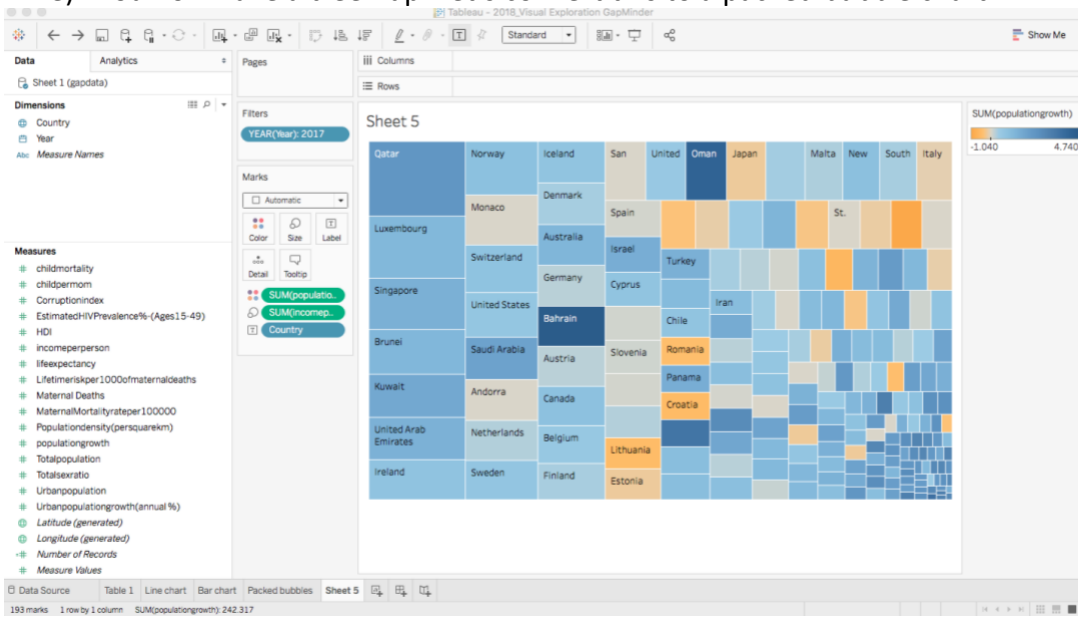


g) Change the view type to **Entire View** so the full bar chart is visible
h) Give your sheet a descriptive title

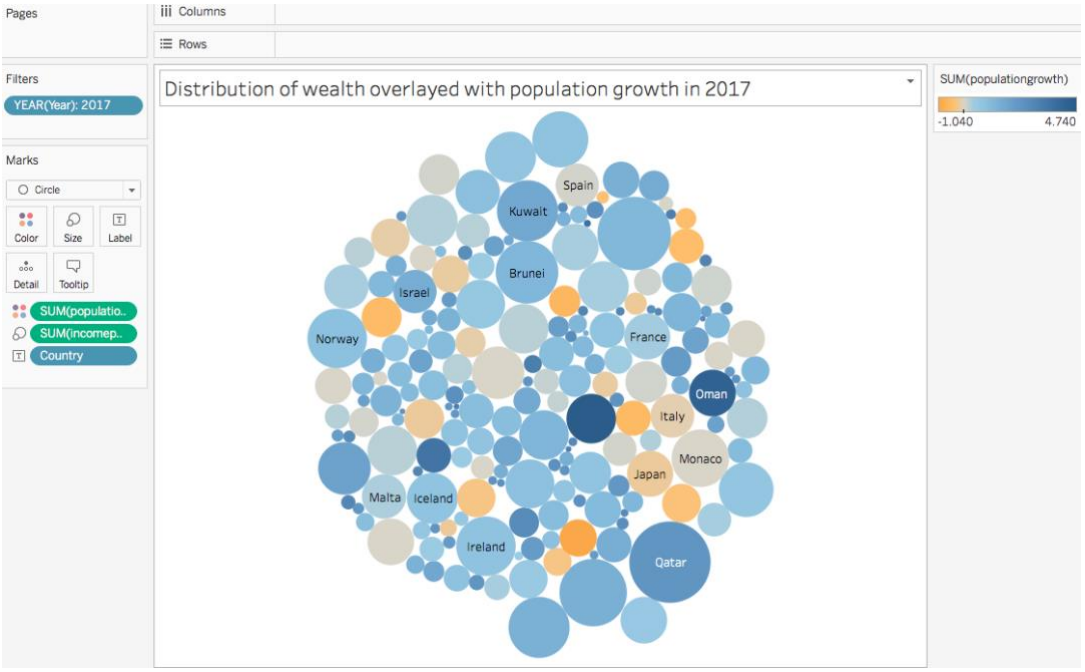


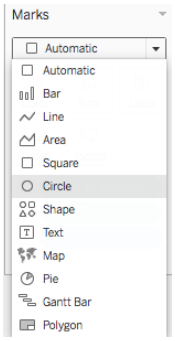
7. Create a packed bubble chart to display distribution of wealth across countries

a) Drag **Year** to the **Filters card** and filter for 2017
b) Drag **populationgrowth** to **Colors** in the Marks Card
c) Drag **incomeperperson** to **Size** in the Marks Card
d) Drag **Country** to **Label** in the Marks Card
e) You now have a treemap. Let's convert this to a packed bubble chart.



f) Change the type of Mark by selecting the drop-down button in the Marks Card and selected **Circle**





Exercise 3: Speed visualization!

Create a visualization using at least 2 variables in the dataset. Start with a question – what do you want to know about population growth? Are you interested in a particular country and trends over time? Perhaps a certain variable across a select few countries? Practice your Tableau skills by picking 2 variables to visualize.

Be ready to explain why you chose your specific type of visualization for the variables of interest and any new things you learned while pulling together the visualization in Tableau.