

PUBH6299: Visual Exploration of Public Health Data

Handout #1

Fall 2018

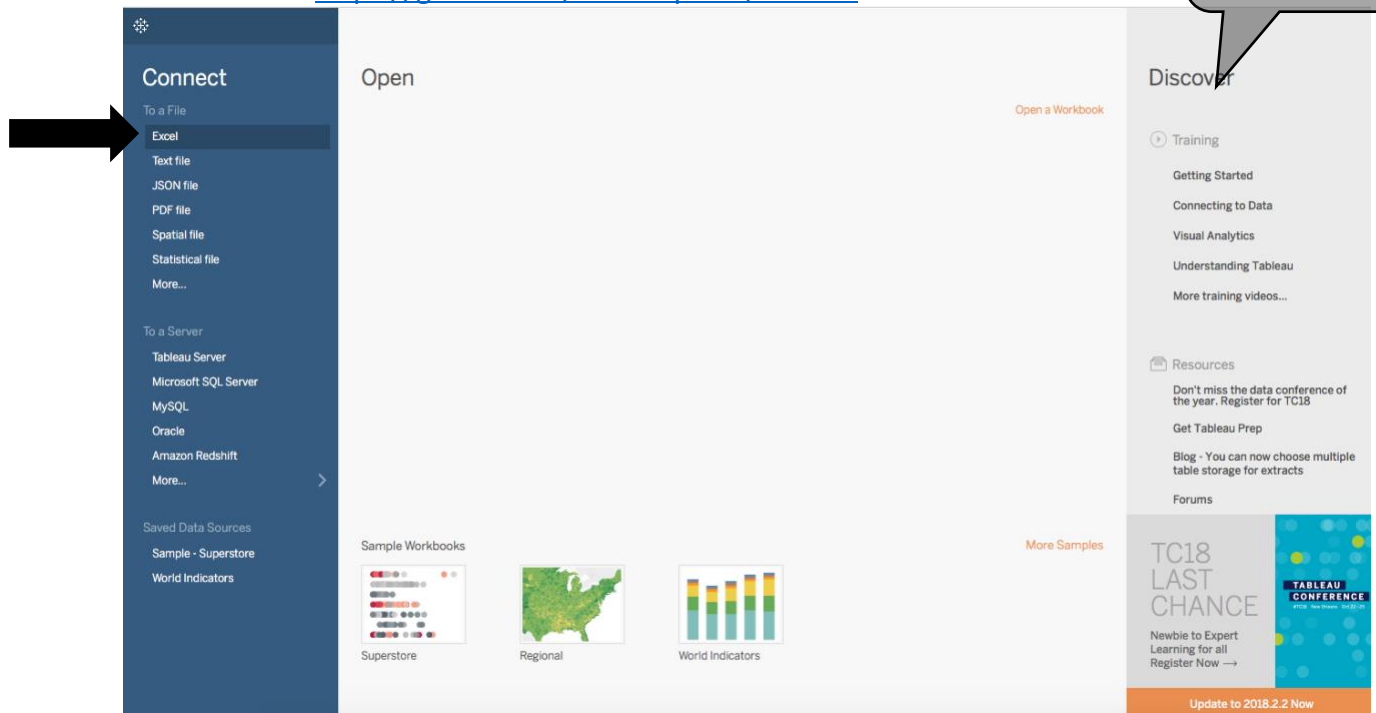
Table of Contents

1. CONNECT TO YOUR DATA	2
2. ORIENTATION TO DATA SOURCE AND METADATA.....	3
MANAGE METADATA.....	3
EXERCISE 1: UPDATE THE METADATA TO CORRECT DATA TYPES.....	4
3. ORIENTATION TO THE INTERFACE	5
DATA IMPORT CHECK.....	5
4. CREATE A TABLE.....	6
DISPLAY THE LIFE EXPECTANCY PER COUNTRY.	6
ADD FILTERS TO THE TABLE.....	6
EXERCISE 2: CUSTOMIZE WITH FILTERS AND COLORS.....	8
<i>Let's filter the year.....</i>	<i>8</i>
<i>Transform to a highlighted table.....</i>	<i>9</i>
5. CREATE A LINE GRAPH: LIFE EXPECTANCY AND POPULATION GROWTH BY COUNTRY.....	10
6. CREATE A BAR CHART: INCOME PER PERSON BY COUNTRY	11
7. CREATE A PACKED BUBBLE CHART TO DISPLAY DISTRIBUTION OF WEALTH ACROSS COUNTRIES.....	12
EXERCISE 3: SPEED VISUALIZATION!.....	13

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>

Links to
video
tutorials



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.

Data set

Sheets in dataset

Variables

Extract allows you to subset the dataset and reduce file size

Filters are temporary and do not reduce dataset size

Toggle between preview dataset and manage metadata

Sheet 1	#	#	Abc	#	#	Abc	Abc	#
Country	Year	Totalpopulation	EstimatedHIVPrev...	Urbanpopulation...	Urbanpopulation	Corruptionindex	HDI	Population
Afghanistan	1800	3,280,000	null	null	null	null	null	
Albania	1800	410,445	null	null	null	null	null	
Algeria	1800	2,503,218	null	null	null	null	null	
American Samoa	1800	8,170	null	null	null	null	null	
Andorra	1800	2,654	null	null	null	null	null	
Angola	1800	1,567,028	null	null	null	null	null	
Anguilla	1800	2,025	null	null	null	null	null	
Antigua and Barbuda	1800	37,000	null	null	null	null	null	
Argentina	1800	534,000	null	null	null	null	null	
Armenia	1800	413,326	null	null	null	null	null	

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

Blue = dimensions

Green = measures

Field Name	Table	Remote Field Name
Country	Sheet 1	country
Year	Sheet 1	year
Totalpopulation	Sheet 1	Totalpopulation
Abc EstimatedHIVPrevalence%-(Ages15-49)	Sheet 1	EstimatedHIVPrevalence%-(Ages15-49)
Urbanpopulationgrowth(annual%)	Sheet 1	Urbanpopulationgrowth(annual%)
Urbanpopulation	Sheet 1	Urbanpopulation
Abc Corruptionindex	Sheet 1	corruptionindex
Abc HDI	Sheet 1	HDI
Populationdensity(persquarekm)	Sheet 1	Populationdensity(persquarekm)
Totalsexratio	Sheet 1	Totalsexratio
Abc Lifetimeriskper1000ofmaternaldeaths	Sheet 1	Lifetimeriskper1000ofmaternaldeaths
MaternalMortalityrateper100000	Sheet 1	MaternalMortalityrateper100000

Exercise 1: Update the metadata to correct data types

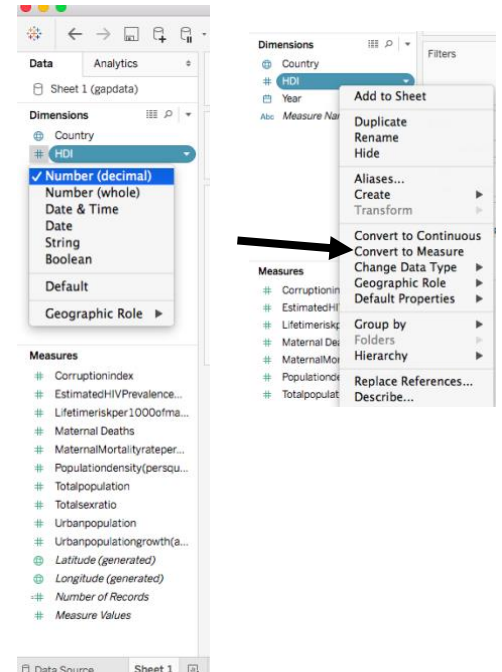
- Change data type: Right click on a variable's data type icon to change the type of data.

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

- 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 #

- Complete this step for other variables as needed.



How to save:

- 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.

The screenshot shows the Tableau desktop application interface. Annotations include:

- dataset**: Points to the 'Data' pane on the left.
- Dimensions**: Points to the 'Dimensions' section in the left pane, which includes 'Country' and 'Year'.
- Measures**: Points to the 'Measures' section in the left pane, which lists various health indicators like 'childmortality' and 'lifeexpectancy'.
- Marks Card**: Points to the 'Marks' shelf in the center pane.
- Variables can be dragged to rows and columns**: Points to the 'Columns' and 'Rows' shelves at the top of the center pane.
- Show Me**: Points to the 'Show Me' button in the top right corner.

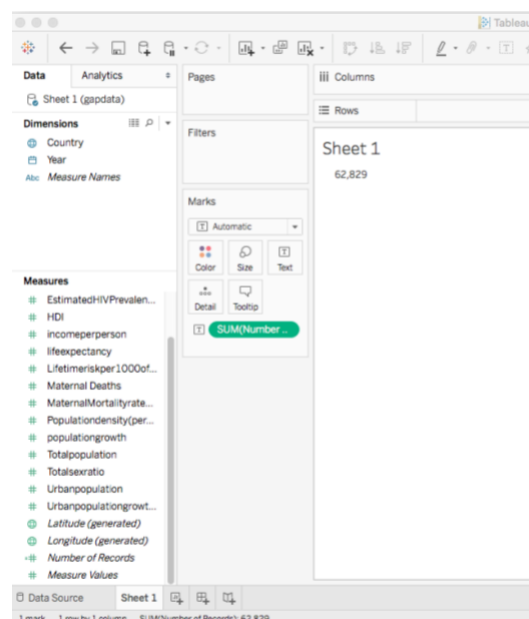
The 'Show Me' button is described as: 'Creates quick visuals based on the data in the view'.

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.

- From the Dimensions pane, drag **Country** to **Rows**
- From the Dimensions pane, Drag **Year** to **Columns**
- 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.

Tableau - 2018_Visual Exploration GapMinder

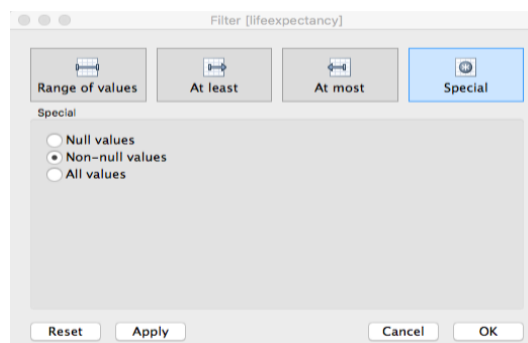
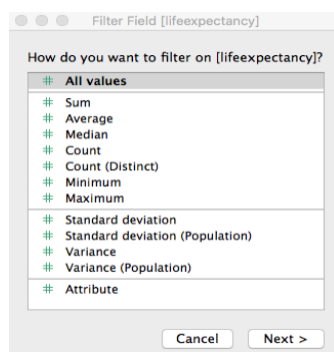
Sheet 1

Country	1800	1801	1802	1803	1804	1805	1806	1807	1808	1809	1810	1811	1812	1813
Afghanistan	28.20	28.20	28.20	28.20	28.20	28.20	28.10	28.10	28.10	28.10	28.10	28.10	28.10	28.10
AFRICA														
Akrotiri and...														
Åland														
Albania	35.40	35.40	35.40	35.40	35.40	35.40	35.40	35.40	35.40	35.40	35.40	35.40	35.40	35.40
Algeria	28.80	28.80	28.80	28.80	28.80	28.80	28.80	28.80	28.80	28.80	28.80	28.80	28.80	28.80
American S...														
Andorra														
Angola	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00
Anguilla														
Antigua an...	33.50	33.50	33.50	33.50	33.50	33.50	33.50	33.50	33.50	33.50	33.50	33.50	33.50	33.50
Argentina	33.20	33.20	33.20	33.20	33.20	33.20	33.20	33.20	33.20	33.20	33.20	33.20	33.20	33.20
Armenia	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00
Aruba														
ASIA														
Australia	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00
Australia/N...														
Austria	34.40	34.40	34.40	34.40	34.40	34.40	34.40	34.40	34.40	34.40	34.40	34.40	34.40	34.40
Azerbaijan	29.20	29.20	29.20	29.20	29.20	29.20	29.20	29.20	29.20	29.20	29.20	29.20	29.20	29.20
Bahamas	35.20	35.20	35.20	35.20	35.20	35.20	35.20	35.20	35.20	35.20	35.20	35.20	35.20	35.20
Bahrain	30.30	30.30	30.30	30.30	30.30	30.30	30.30	30.30	30.30	30.30	30.30	30.30	30.30	30.30
Bangladesh	25.50	25.50	25.50	25.50	25.50	25.50	25.50	25.50	25.50	25.50	25.50	25.50	25.50	25.50
Barbados	32.10	32.10	32.10	32.10	32.10	32.10	32.10	32.10	32.10	32.10	32.10	32.10	32.10	32.10
Belarus	36.20	36.20	36.20	36.20	36.20	36.20	36.20	36.20	36.20	36.20	36.20	36.20	36.20	36.20
Belgium	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.10	40.10	40.10	40.10	40.10	40.10	40.10
Belize	36.60	36.60	36.60	36.60	36.60	36.60	36.60	36.60	36.60	36.60	36.60	36.60	36.60	36.60

Add Filters to the table

Filter 1: Life Expectancy.

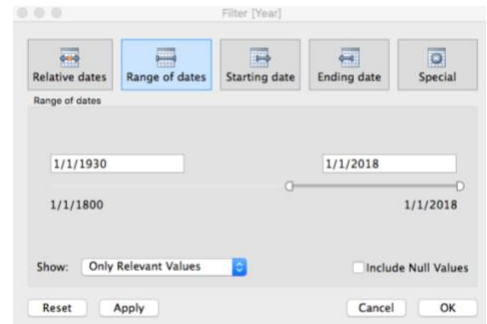
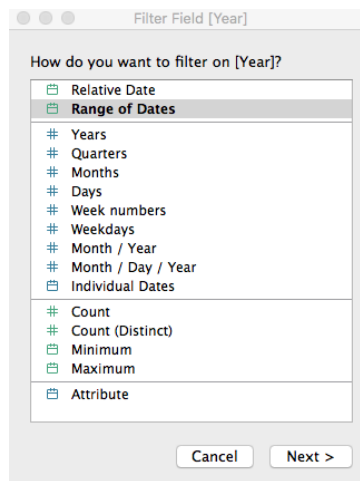
- From the Measures pane, drag lifeexpectancy to the **Filters Card**
- Select All Values, then click Next in the pop-up window
- Click on Special at the top of the next window, and select **Non-null values**. This will show only data points for which there are values. Blanks will be excluded. Click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



GWU PUBH6299: Visual Exploration of Public Health Data Handout #1

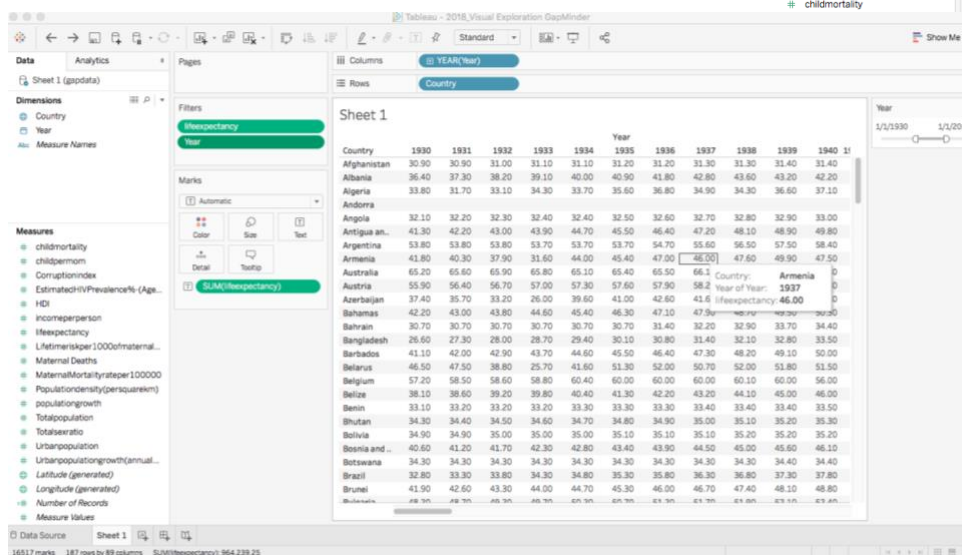
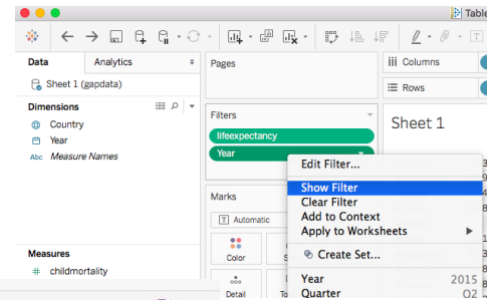
Filter 2: Year

- 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:

Sheet 1

Country	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940
Afghanistan	30.90	30.90	31.00	31.10	31.10	31.20	31.20	31.30	31.30	31.40	31.50
China	33.30	33.70	34.00	34.40	34.80	35.20	35.60	33.40	33.30	33.50	33.60
Germany	59.40	59.80	60.20	60.60	61.00	61.50	61.70	62.00	62.30	61.00	60.90
Haiti	28.40	28.40	28.40	28.40	28.40	28.40	28.40	28.30	28.30	28.30	28.30
Indonesia	32.00	32.50	33.00	33.50	34.10	34.60	35.10	35.60	36.00	36.40	36.80
North Korea	37.80	38.70	39.60	40.40	41.60	42.80	44.00	45.10	46.30	47.00	47.50
Philippines	35.10	35.90	36.60	37.30	38.10	38.80	39.50	40.30	41.00	42.30	43.00
Rwanda	32.70	32.80	33.20	33.60	34.10	34.50	35.00	35.40	35.90	36.30	36.80
South Africa	35.20	35.70	36.30	36.80	37.40	38.00	38.50	39.10	39.60	40.20	40.80
Swaziland	33.60	33.60	33.70	33.70	33.70	33.70	33.80	33.80	33.80	33.80	33.80
Switzerland	61.40	61.20	61.10	62.30	62.80	62.00	63.10	63.40	63.80	63.90	64.00
Uganda	25.10	25.40	25.60	25.90	26.20	26.40	26.70	26.90	27.30	27.60	27.90
United States	59.60	60.30	61.00	60.90	60.30	60.90	60.40	61.10	62.40	63.10	63.50

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.

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

Filter Field [Year]

How do you want to filter on [Year]?

- Relative Date
- Range of Dates
- # Years**
- # Quarters
- # Months
- # Days
- # Week numbers
- # Weekdays
- # Month / Year
- # Month / Day / Year
- Individual Dates
- # Count
- # Count (Distinct)
- Minimum
- Maximum
- Attribute

Cancel Next >

Filter [Year of Year]

General Condition Top

Select from list Custom value list Use all

Enter search text

- ☒ 1930
- ☐ 1931
- ☐ 1932
- ☐ 1933
- ☐ 1934
- ☐ 1935
- ☐ 1936
- ☐ 1937
- ☐ 1938
- ☐ 1939
- ☒ 1940

All None Exclude

Summary

Field: [Year of Year]

Selection: Selected 2 of 89 values

Wildcard: All

Condition: None

Limit: None

☐ Filter to latest date value when workbook is opened

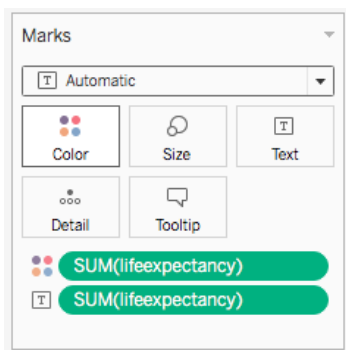
Reset Apply Cancel OK

d) Your Sheet should now look as follows:

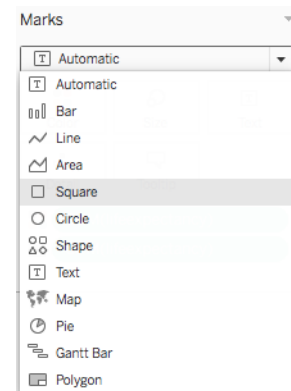


Transform to a highlighted table

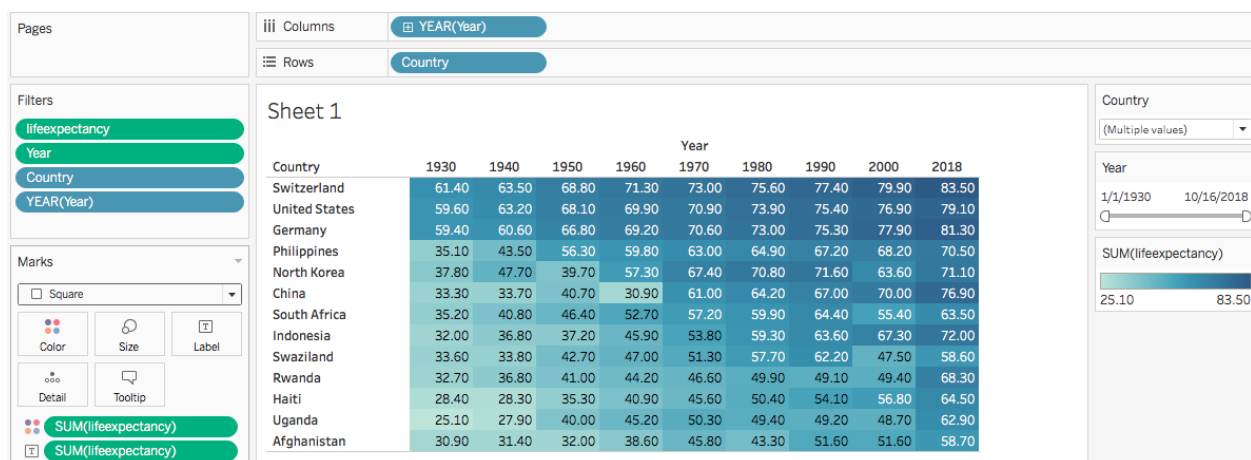
- a) From the Measures pane, drag lifeexpectancy to Color under the Marks Card. This color codes the life expectancy values based on a color range.



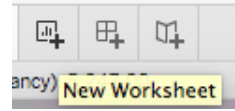
- b) Click on the drop-down under the Marks Card and select "Square"



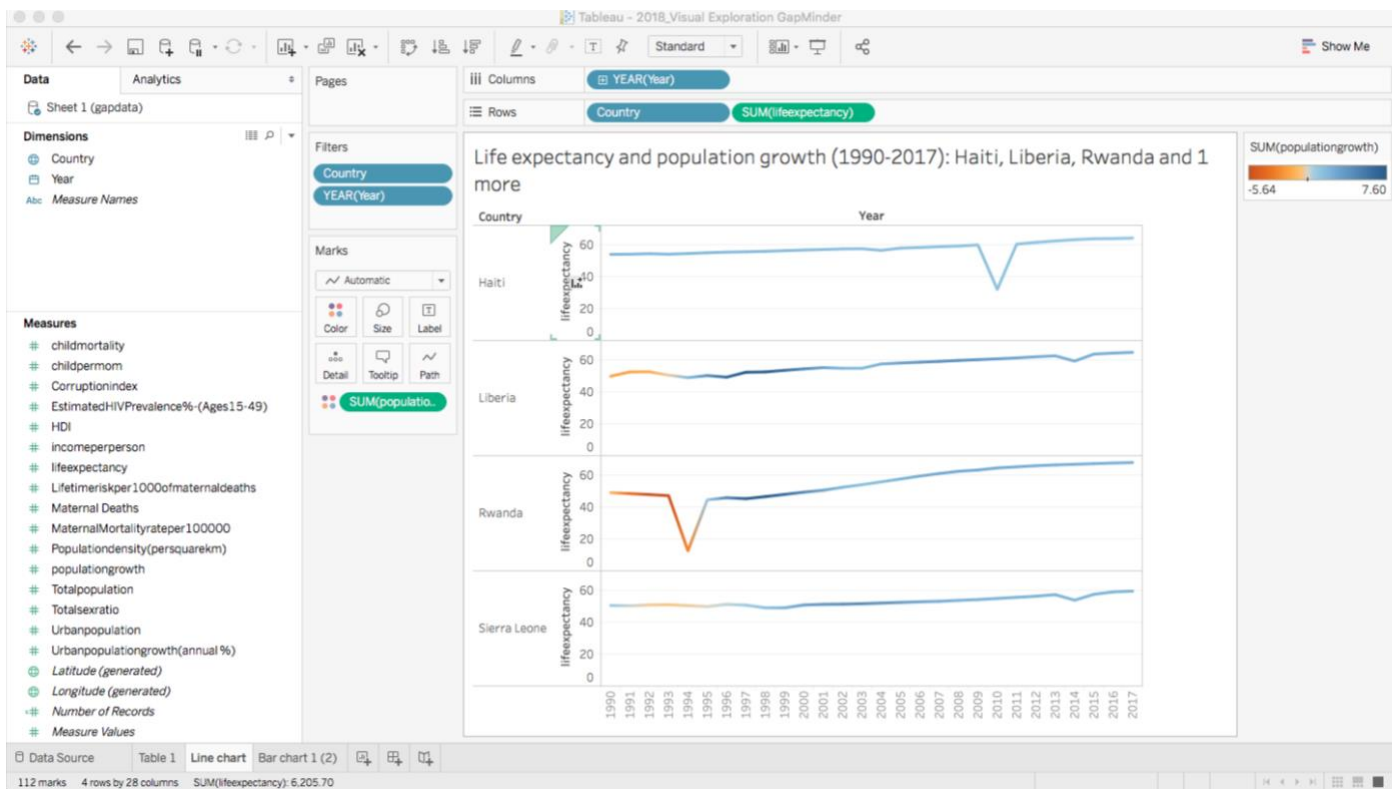
Now you have created a highlighted table, based on conditional formatting for the life expectancy values across all countries for the selected years.



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

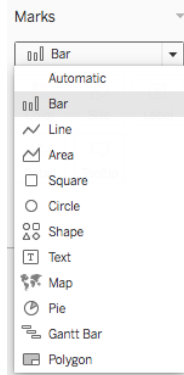


- Create a new sheet and give it a title
- Drag **Country** to Rows
- Drag **Year** to Columns
- Drag **Country** to **Filters** and filter for: Haiti, Liberia, Rwanda, Sierra Leone
- Drag **Year** to **Filters** and filter for: all years between 1990-2017
- Drag **#lifeexpectancy** to Rows
- Drag **#populationgrowth**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

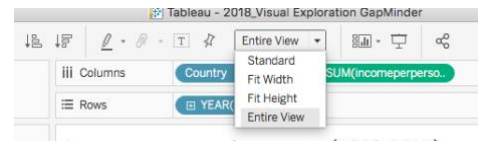


6. Create a bar chart: income per person by country

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

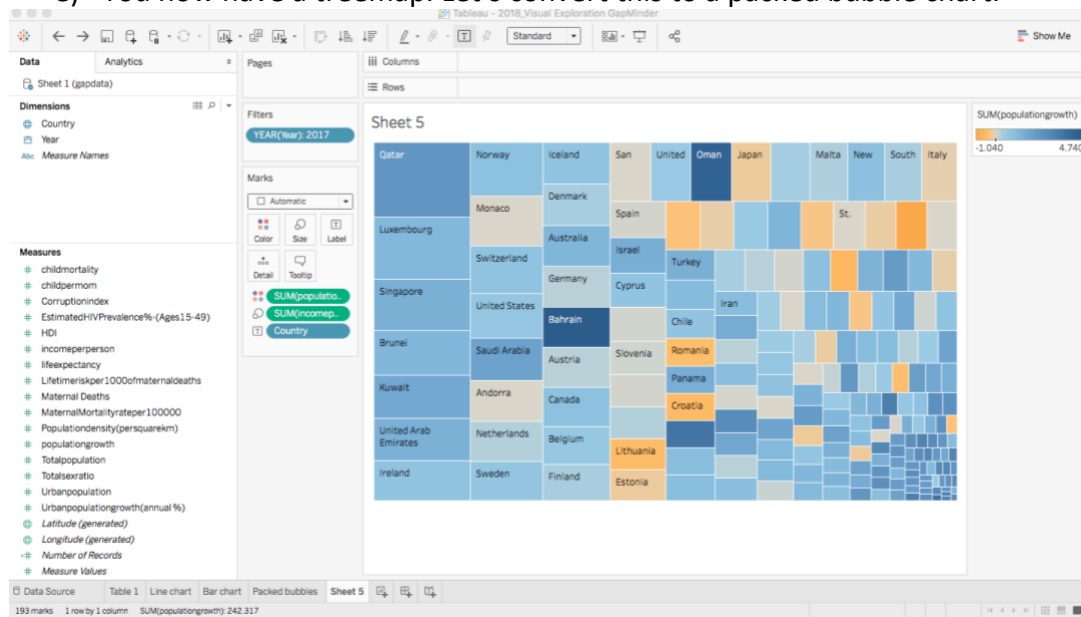


- Change the view type to **Entire View** so the full bar chart is visible
- Give your sheet a descriptive title

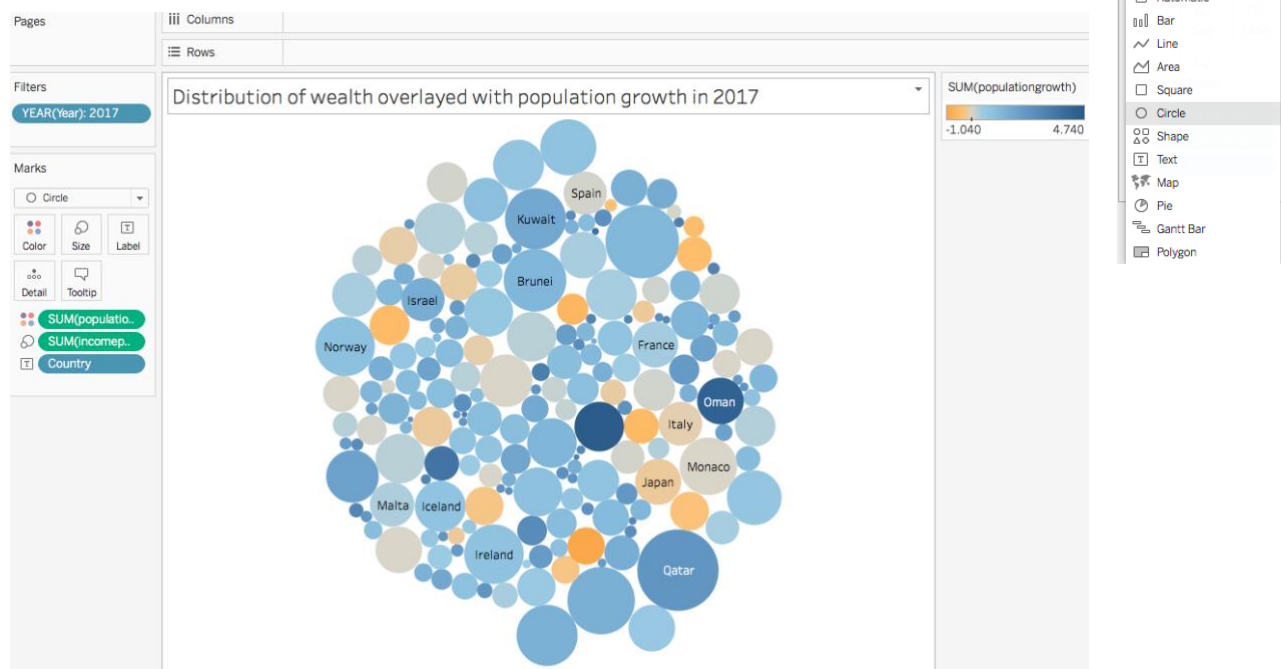


7. Create a packed bubble chart to display distribution of wealth across countries

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



- 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.