





# Health and Economic Factor Visualization

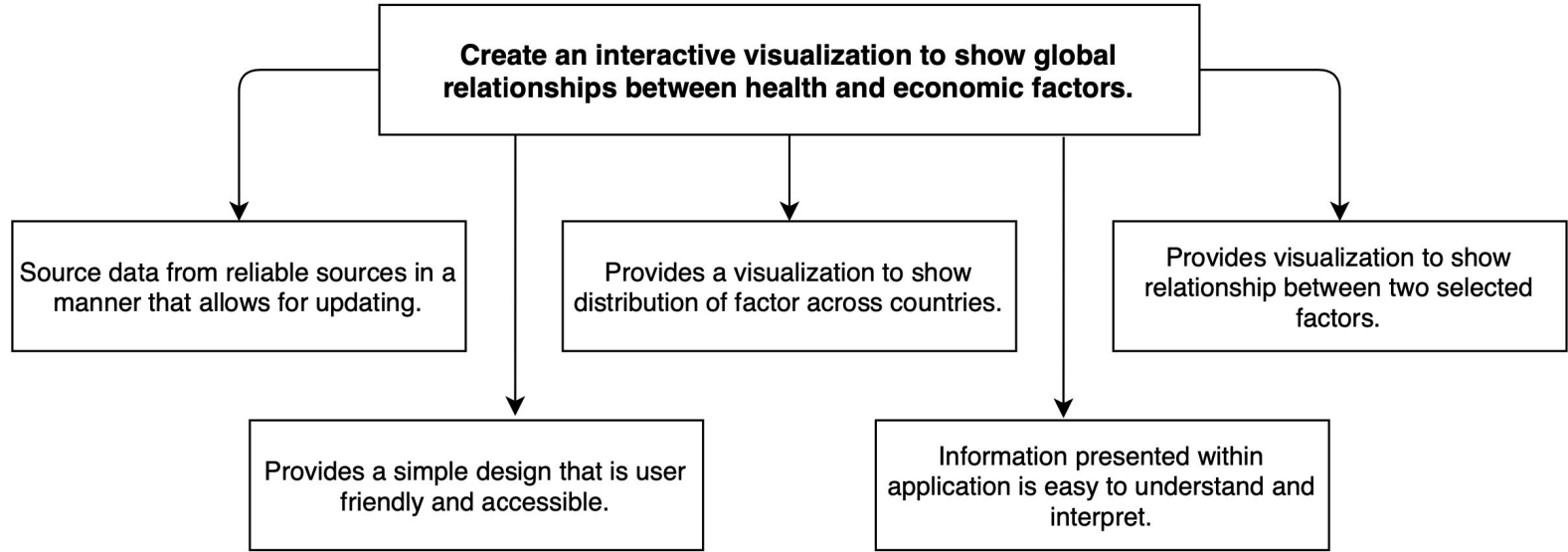
Nathaniel Barrington, Gabriela Guzman, Bryce  
Huffman, Zachery Key & Thomas Twomey

University of Virginia, Spring 2020  
Professor Afsaneh Doryab  
SYS 2202 Final Project: Group 11





# Objectives



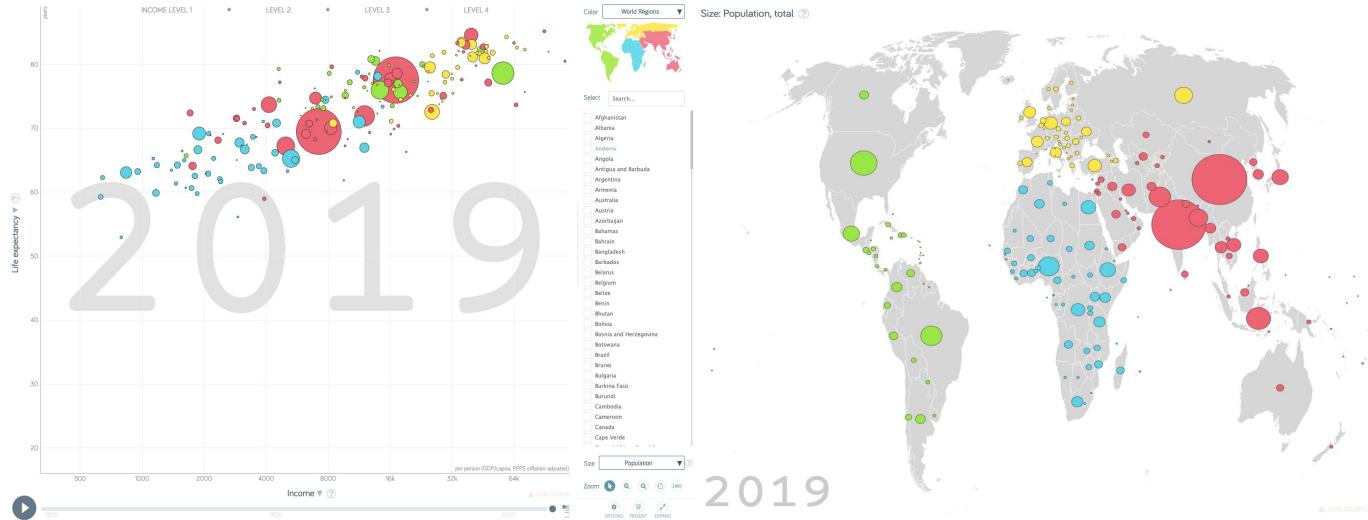


# Prior Art

## Prior Art: Example 1

# Gapminder

Figure 1: Gapminder Tool Screenshots



### Pros:

- Extensive set of data from numerous sources
- Ability to filter countries and color code by region

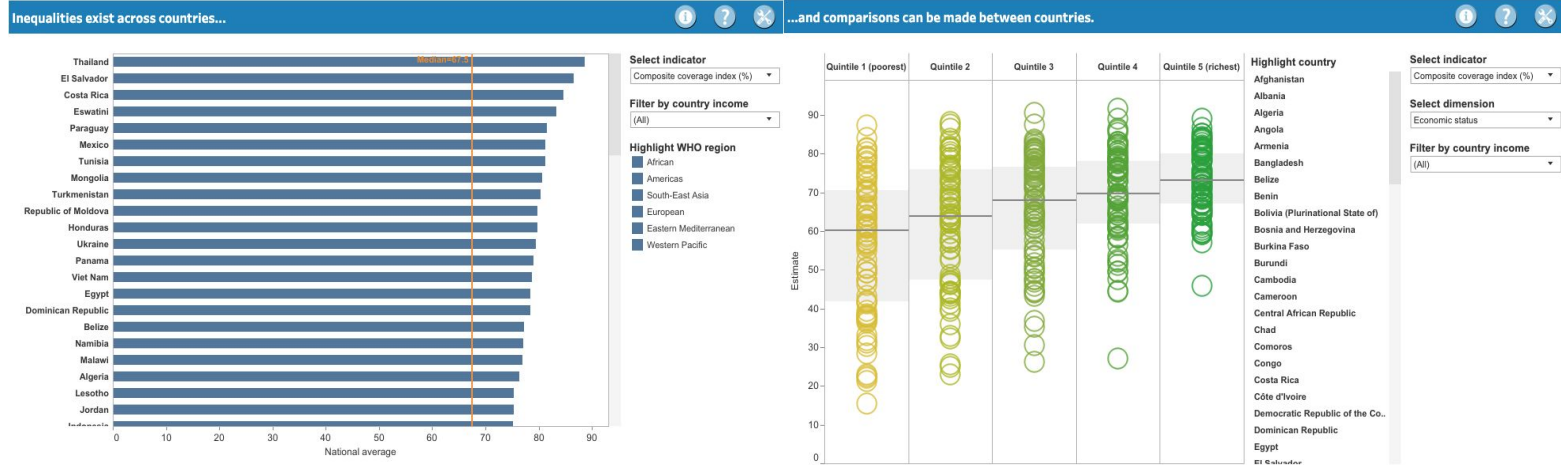
### Cons:

- One factor world map lacks meaningful visualization
- Plots can not be viewed simultaneously

## Prior Art: Example 2

# World Health Organization Interactive Visualizations

Figure 2: WHO Interactive Visualization Screenshots



### Pros:

- Numerous plot types
- Heavy use of statistical information

### Cons:

- Visualizations are not user-friendly or easily understandable
- Lack of ability to examine multiple factors



# Data Sourcing

## Online Database and APIs

Data was imported from the **World Bank** and the **World Health Organization** using R's "WDI" and "url" function respectively. These data sources were selected on the basis of their **reliability**, **accessibility** and **preprocessing** that led to clean and understandable data formatting.

- World Bank
  - *Selected Attributes:* Literacy Rates (Male/Female over 15 years old), GDP per capita (2010 \$ US), Life expectancy (Male/Female), GDP % of Trade & GDP % of Services
- World Health Organization
  - *Selected Attributes:* Infant Mortality Rates (Male/Female/Overall, under 5 years old)





# Data Processing

# General Process Overview

1. Import data from World Bank & WHO using APIs
2. Remove NA values and eliminate redundant columns
3. Join datasets with country codes (ISO3 standard)
4. Restructure data frames using “spread” function on years

Country	ISO3	1980	1981	...	2017	2018
Aruba	ARW	89	92	...	94	95
Afghanistan	AFG	92	87	...	90	92
...	...	...	...	...	...	...

**Figure 3:** Sample of Data Frame Post-Processing

1. Merge cleaned data frames with Latitude and Longitude coordinates for each country by ISO3 country code.
2. Create custom “fill” columns in the data frame for choropleth coloring.
3. Write algorithms for fill column to get ratio values for arguments to coloring the graph.

Long	Lat	Country	ISO3	1980	1981	fill
-67.879	12.452	Aruba	ABW	89	92	.67
-70.832	12.528	Aruba	ABW	89	92	.67
...	...	...	...	...	...	
74.892	37.361	Afghanistan	AFG	54	56	.5

**Figure 4:** World Map Data Frame for Literacy Rate Head

$\text{Fill} = \text{value} / \max(\text{all values in attribute column})$

$\text{InFill} = \min(\log(1.5 + \text{factors}), .6)$

**Figure 5:** Algorithm Used to Determine World Map Fill

## Two-Factor Scatterplot

1. Removed columns from each factor's data table that correspond to the selected year
2. Perform inner join to create data frame for plotting based on selected inputs
3. Remove any records with NA values

<b>i</b>	<b>x</b>	<b>y</b>
Afghanistan	98.2	15600.14
Albania	86.3	10512.98
...	...	...
Zimbabwe	96.7	12415.19

**Figure 6:** Scatterplot Data Frame Head



# Interface and Visualization

# R Shiny UI and Reactivity

- App is developed in R using Shiny and a reactive environment
  - Allows for visualizations to only update if input is changed
  - Enabled ability to have multiple plots interact with one another
- General data processing is conducted once at app run, including API queries
- World Map and Scatterplot are linked together through factor and year selection
- Dynamic text updating for data sources

# Interactive Plots

- Choropleth World Map
  - Uses “plotly” package with “ggplot” to create world map
  - Allows for scaling, zooming, and country selection
  - Ability to download the Choropleth map as a png
- Two-factor Scatterplot
  - Uses “ggvis” to create interactive scatterplot
  - Conducts data merge for each plot selection
  - Allows for scaling and plot download
  - X-Axis corresponds to factor displayed in World Map
  - Average trendline for visual relationship analysis



# R Shiny Demonstration

[GitHub Repository](#)





# Final Summary



# Reflection

- The final R Shiny App encompasses the desired goal from the objectives.
  - No sacrifices were made from original idea to end product
- Further improvements could be made to enhance experience
  - Dynamic year selection based on available data
  - Addition of country specific visualizations



# References

# References

Gapminder. (2020). *Gapminder tools: maps*. Retrieved from  
[https://www.gapminder.org/tools/#\\$chart-type=bubbles](https://www.gapminder.org/tools/#$chart-type=bubbles)

GitHub. (2020). *Sys\_2202finalproject*. Retrieved from  
[https://github.com/brycehuffman/sys\\_2202\\_finalproject](https://github.com/brycehuffman/sys_2202_finalproject)

World Health Organization. (2020). *Global Health Observatory (GHO) data: Interactive data visualizations*.  
Retrieved from [https://www.who.int/gho/health\\_equity/interactive\\_data\\_visualizations/en/](https://www.who.int/gho/health_equity/interactive_data_visualizations/en/)