Data, Perception and Ignorance

USHA KRISHNAMOORTHY

"Let the dataset change your mindset" – Hans Rosling

We live in an era of data and information. Right from deciding what to read, what to wear, which restaurant to go to, which city to visit, whom to vote for, we consider ourselves rational human beings who rely on data to make all our decisions. How much of this data is based on facts rather than opinions and/or perceptions?

This review looks at two websites, **Gapminder** and **Our World in Data**, which attempt to provide reliable global statistics and promote a fact-based worldview. While Our World in Data is targeted towards users who have a basic understanding of economics and statistics, Gapminder requires little or no background of either. Examples of my usage of these websites in the classroom are described below.

Gapminder

While looking for information about lifestyles of children around the world as part of a school project, I came across the website Gapminder (https://www.gapminder.org/). The website was co-founded by Hans Rosling with his son and daughter-in-law. Rosling, a physician, statistician and public speaker, was a professor of International Health at Karolinska Institute, Sweden. He held several presentations around the world to promote the use of data in exploring issues and trends in global development.

The mission of Gapminder Foundation, as stated on the website, is:

"Fighting devastating ignorance with fact-based worldviews everyone can understand."

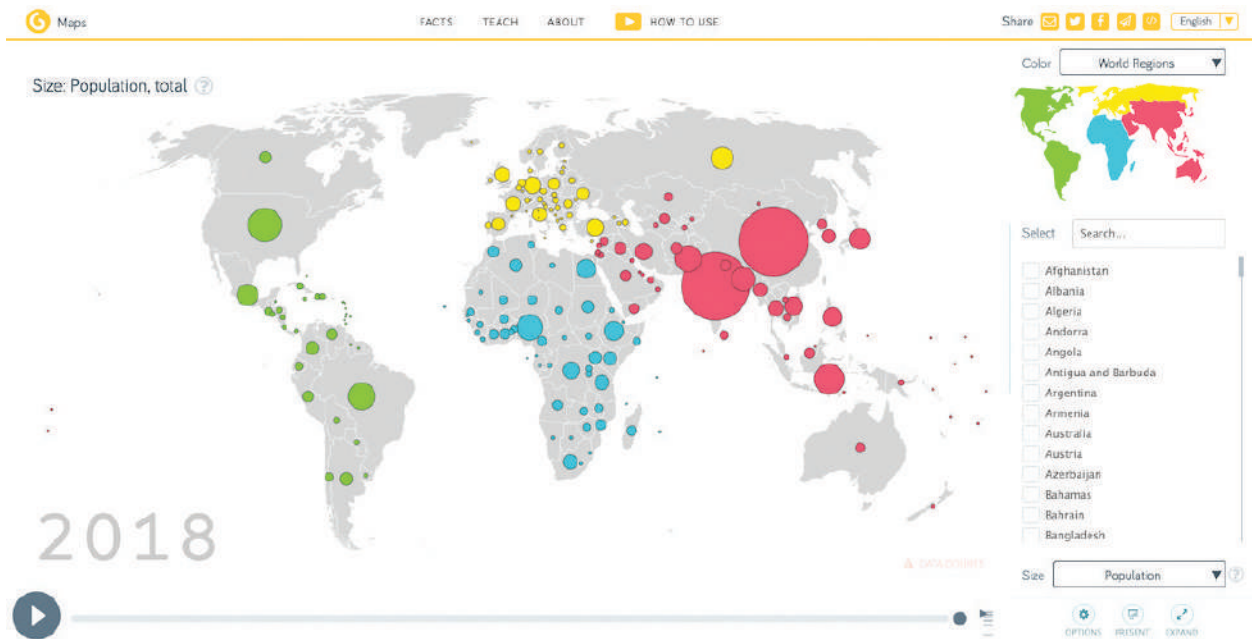*Keywords: Data, statistics, economics, perception, context, facts*
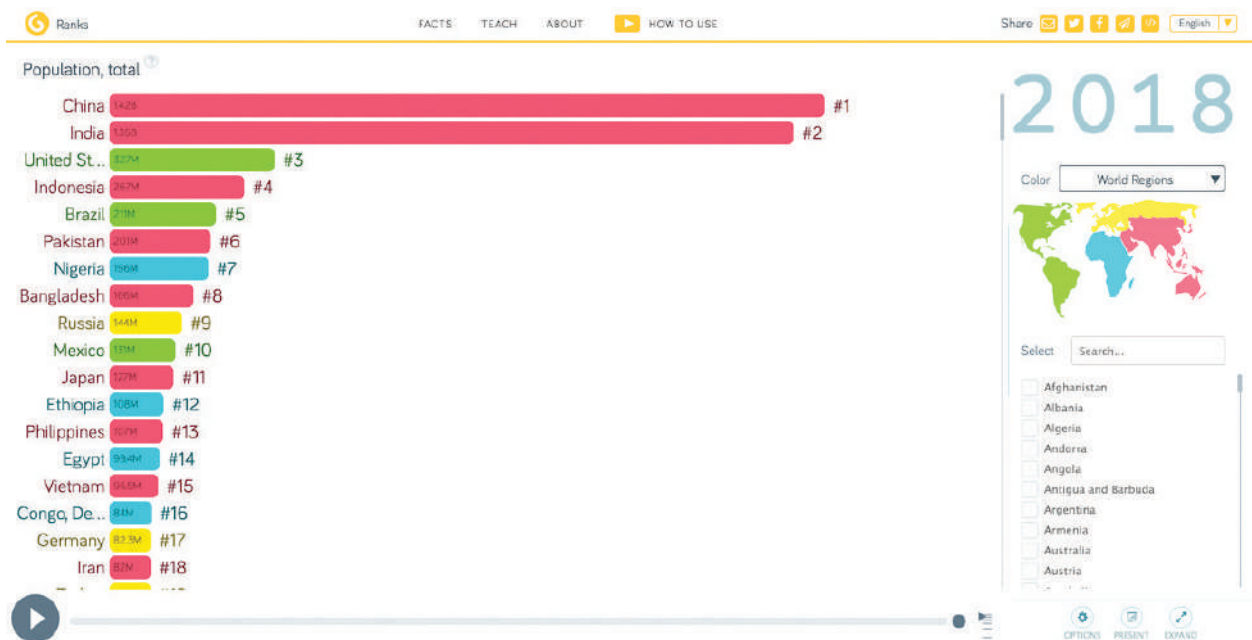
Image source: Gapminder Tools


Image source: Gapminder Tools

There is a belief amongst the creators, who are statisticians and/or computer programmers, that people tend to have an overdramatic image and prejudiced notions about the world that can actually lead to bad choices and decisions. In his 2006 TED talk "The Best Stats You've Ever Seen", Hans Rosling stated that university students from Sweden know statistically significantly less about the world than chimpanzees (who are equally likely to pick the right or wrong answer in response to certain questions). He conducted numerous surveys involving the world's top decision makers in both the public and private sector and here again he found the "global ignorance quotient" to be quite high. The website is a result of this mission to eradicate widespread ignorance about global statistics.

The website provides a variety of data visualization graphics (some examples below) that allow the user to get a worldview backed by data as well as tools to analyze data for oneself. This greater use and understanding of global statistics, the creators believe, is a way to promote sustainable global development and promote a way of thinking about society and the world that is based on concrete facts, which is what led them to create the Gapminder Foundation in 2005 and subsequently the website.

Gapminder Data

The website combines publicly available data from multiple sources and makes it available in time series format making it possible to analyze trends through the years. Some of the sources of data include the International Labor Organization, the World Bank, the World Health Organization and the OECD QWIDS (https://stats.oecd.org/qwids/). There are some gaps in terms of coverage of the data, which could also be due to a lack of formal records being available for certain countries. While the data is available for download in structured tabular formats, the more interesting use of it is with the tool provided by the website itself. While at first glance they seem to look like simple line graphs or bar graphs plotting a few parameters, the stories that one can get from the data become more apparent as you start to take a closer look. Here's one example.

The parameters shown in this example:

- GDP per capita adjusted to inflation on the x axis
- Life expectancy on the y axis
- Population depicted as the size of the bubbles
- World regions (Americas, Africa, Europe, Asia + Australia) depicted using the color of the bubble

Image source: Gapminder
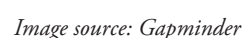
A slider provided at the bottom of the screen can be used to look at a particular year or you can click on the "play" icon to view change over a period of time.

This section of the website also comes with a set of videos demonstrating how the tool can be used. They do so by taking up questions such as "How does income relate to life expectancy?", "Will saving poor children lead to overpopulation?" etc. The videos feature Hans Rosling who in a very animated manner demonstrates how data and the tools available can be used to answer the question.

There is an offline version of the tool which includes all data available on the Gapminder website and which also allows you to plug in your own data to use with the visualizations. The offline tool can be periodically updated as new data becomes available on the website. All the material is freely available under the Creative Commons Attribution 4.0 International license, which (amongst other things) permits usage in schools for educational purposes.

Classroom Usage: I have used data (life expectancy and income per person) from Gapminder with senior school students (Class XI) as part of their statistics curriculum. Looking at the data, they were able to not only look at the correlation of these parameters, but also look at the impact of specific events in history (World War II, Indian Independence, economic liberalisation) on these parameters. We also watched the video on the website, "How does income relate to life expectancy". The Gapminder offline tool opened up a lot of interesting ways in which data can be represented for students, and the whole time-series animation brought in a new dimension to the statistical analysis of data! The biggest caveat however, was the gaps in data which we found while exploring some African countries. While we were able to look at worldwide trends, patterns and impact, zooming in on specific countries made data gaps quite glaring.

Impromptu conversations on how a mental process is quickly triggered to fill in the gaps to agree with the trends brought in an interesting twist to the topic.

Dollar Street

Another key section of the website is "Dollar Street". While the data visualization tools of Gapminder help visualize trends and correlations between global statistics, they still do not help understand how people in various corners of the world live. To address this, Anna Rosling Rönnlund has come up with Dollar Street.

The idea is to put every family in the world on a street with houses lined up by income: the poorest living on one end, the richest on the other and the rest else in between. A team of photographers have documented over 264 homes in 50 countries so far. In each home, the photographer spends a day taking photos of up to 135 objects, such as the family's toothbrushes or favourite pair of shoes and also body parts such as teeth, hands. All photos are then tagged by household function, family name and income.

The images allow one to create a mental representation/model of living conditions of a specific income group in a particular country as well as to contrast the conditions across different countries. The website includes details of how the researchers went about the rather complex process of calculating income and gathering data for this. There are definitely gaps, as one would expect in an undertaking of such a scale, but if the project continues, it has the potential to bring out a comprehensive description of the living standards in many countries.

Classroom Usage: This section brings the (however clichéd) saying "a picture is worth a thousand words" to life! Students (Class IX) enjoyed looking at this page and explored everything from toilets to toothbrushes to homes and cars. Some of the pictures confirmed their
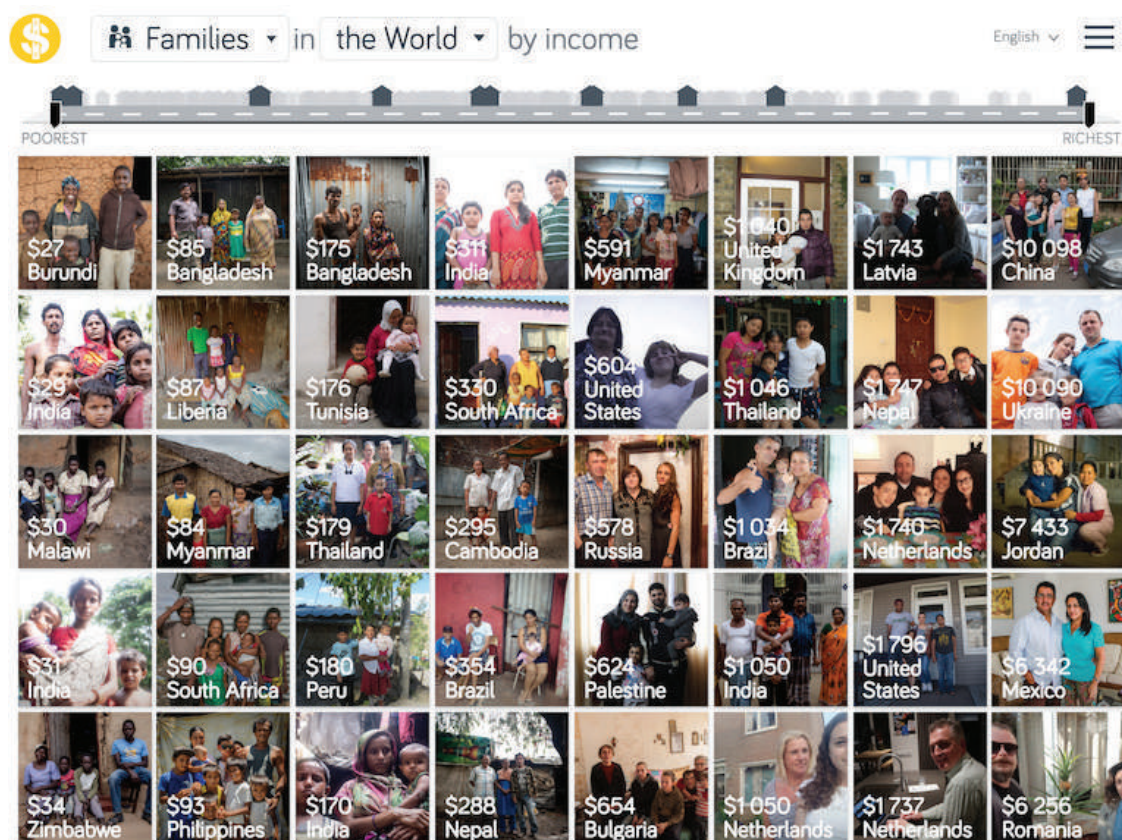
Image source: Screen capture

stereotypes while some left them baffled. While looking at countries that they were familiar with, such as India, they got a sense of the gaps in data. Having said that, it is definitely a very unique way of getting a picture of the world and its diversity. The commonly held idea of data as numbers and visuals as graphs and charts was questioned with the use of images as an interesting way of understanding living conditions.

Teaching Resources

In addition to the tools and data, there is a section called "TEACH" for educators who would like to use Gapminder in the classroom. This section includes presentations, videos and quizzes that can be used to explain global development over the years. In addition, there are also lesson plans and handouts availabl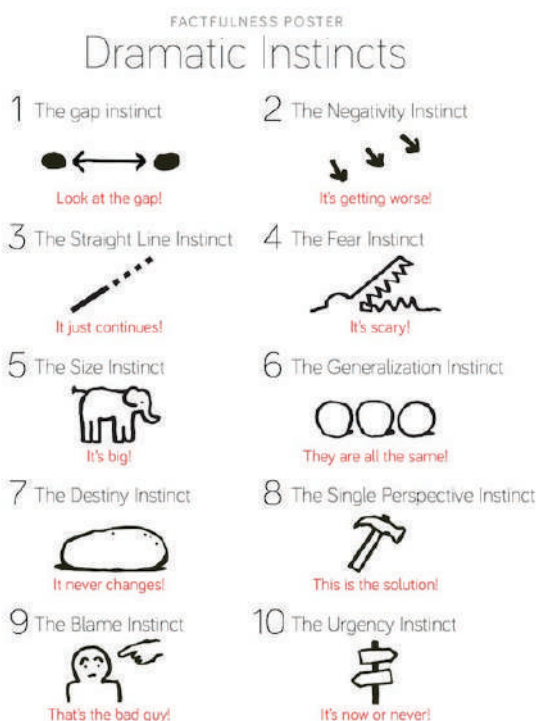e for download. I have used life expectancy and income data from this website with high school students and we analysed trends for specific countries. We then looked at the same data using Gapminder tools to understand how these parameters correlate with each other.

Factfulness

This section is related to the book with the same name written by the creators of the website. It is described as:

"The stress-reducing habit of only carrying opinions for which you have strong supporting facts."

The interesting parts about this section are the posters that depict what stories in the media get our attention, trigger our dramatic instincts and how we can look at them. There is also a presentation available that can be used to go over each of these "instincts" in detail.

FACTFULNESS POSTER
Dramatic Instincts

1 The gap instinct — Look at the gap!

2 The Negativity Instinct — It's getting worse!

3 The Straight Line Instinct — It just continues!

4 The Fear Instinct — It's scary!

5 The Size Instinct — It's big!

6 The Generalization Instinct — They are all the same!

7 The Destiny Instinct — It never changes!

8 The Single Perspective Instinct — This is the solution!

9 The Blame Instinct — That's the bad guy!

10 The Urgency Instinct — It's now or never!

It probably summarises the raison d'etre of the foundation and all the tools and material they have developed.

Website Features

- **Visual appeal:** The layout makes it easy to browse and review various sections. While the homepage doesn't summarise what the webpage is about, it has details about various sections and a video introducing them. In keeping with the current trend of visualisation-driven websites, there is minimal use of text and images and videos are used to explain various tools and concepts. The menus however are not consistent across pages, so the browser "Back" button has to be used a lot to navigate across sections. Dollar Street and the Gapminder tool are easy to access and are probably what one would visit most often. Both these pages render well on mobile devices as well, with no apparent loss of functionality.

- **Loading speed:** Even though the website is graphics intensive, the media seems to be optimised for a variety of bandwidths and it worked seamlessly in all instances. The embedded videos are hosted on YouTube or TED.

- **Advertisements:** The website is ad free. On some pages there is a section through which people can donate to the Gapminder foundation.

- **Plug-ins/Additional Software:** I have used the website on Google Chrome and no additional software is required except if you want to download the offline version of the tool.

- **Search:** There is a search function that allows you to do a keyword search on the entire website. Additionally you can search for specific datasets on the DATA section of the website.

- **Help/FAQ:** The Help section is mostly FAQ and covers a wide range of questions. In addition to this, the LAB section explains usage of various APIs (Application Program Interface) for advanced users familiar with web programming that can be used to integrate tools within websites.

To summarise, the website provides numerous innovative ways of narrating stories with data. Hans Rosling's enthusiasm in the videos makes data look exciting and beautiful and is a joy to watch! In the classroom, I have found it useful to introduce statistics as a subject and to get students excited about its applications!

What to watch out for?

The data itself. While the sources of data are mostly the experts in the subject matter and official data owners, one still needs to keep an eye out for all the "C"s of data quality: completeness, consistency, credibility, cleanliness! Do not mistake non-availability of a picture of a high-income group person in a particular country to be an indicator of poverty levels in the whole country!

Our world in data

Another website with very similar objectives is Our World in Data (https://ourworldindata.org/). Their purpose, as stated in the website:

"Through interactive data visualizations we can see how the world has changed; by summarizing the scientific literature we explain why. Understanding how and why the world has changed up to now allows us to see that a better future is possible."

The project was started in 2011 by economist Max Roser from the University of Cambridge and it focuses on poverty, health, and the distribution of incomes.

Data Sources

The website used 3 sources for all the data:

- Specialized institutes – such as the Peace Research Institute Oslo (PRIO)

- Research articles – such as Bourguignon & Morrison – 'Inequality Among World Citizens: 1820-1992' in the American Economic Review [2002]

- International institutions or statistical agencies – such as the OECD, the World Bank, and UN institutions.

Every chart/visualization has the source clearly indicated and includes details about when it was collated. Some charts, for example "Population by country" (https://ourworldindata.org/world-population-growth) indicate Gapminder as the source for certain years. The data is available for download in CSV (comma separated value) format.

Analysis

With every data type, the website provides detailed analysis by a team of researchers. Unlike Gapminder, which has short videos to explain data trends and to answer FAQs, Our World in Data has detailed reports with each data type. The report is accompanied by charts and other visualizations such as maps to illustrate various statistics that have been computed and analysed.

While Gapminder provides tools to review and analyse data by oneself, this website provides that analysis compiled from many sources including news agencies and the United Nations. Some sections like the one on Child Labor have a detailed section on data and research gaps and also on why those gaps exist.
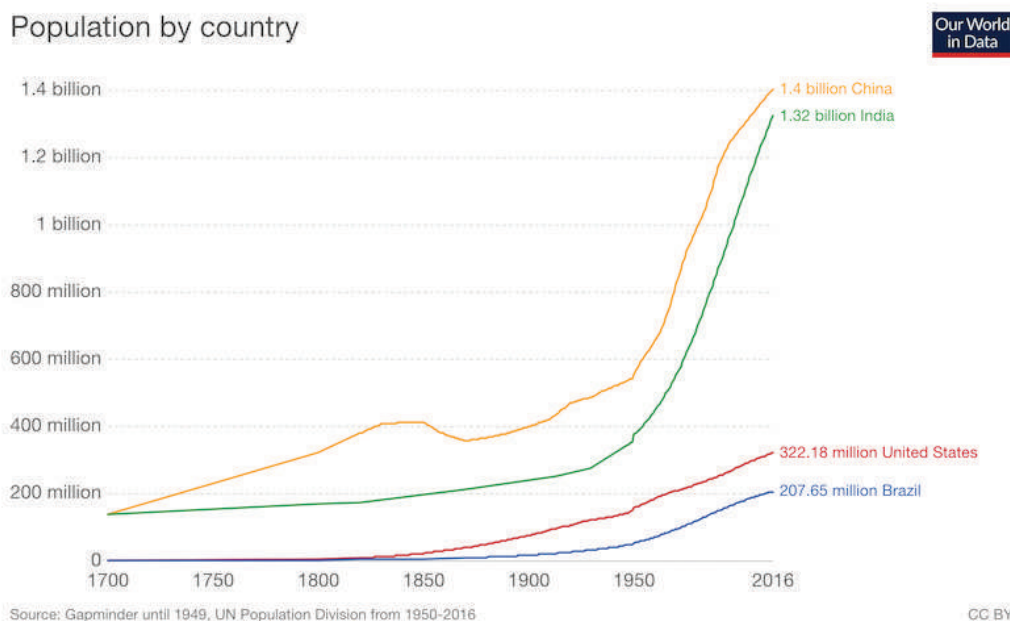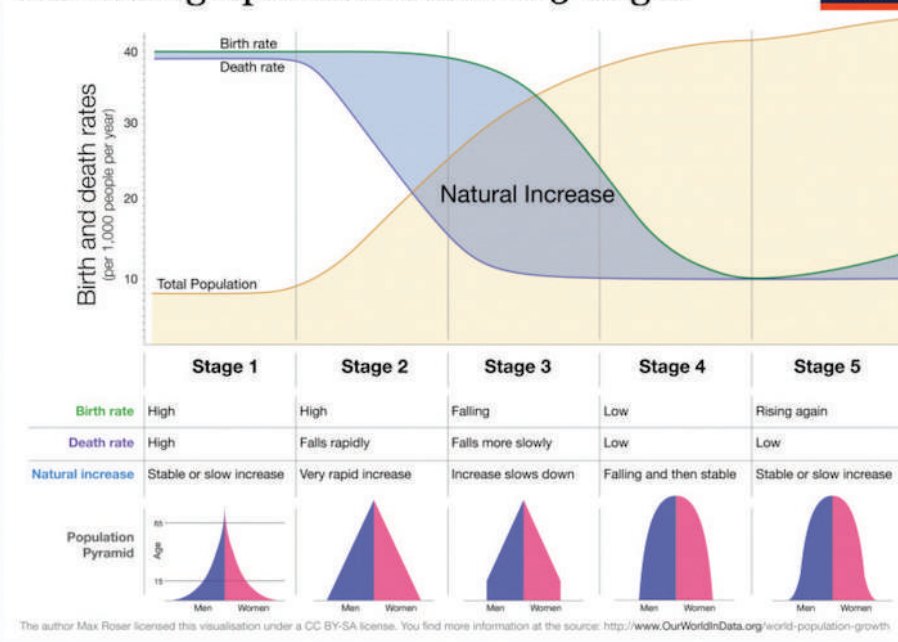

Image source: Screen capture

Image source: Screen capture

Some of the interesting statistics and statements mentioned include:

- *Recent estimates suggest that today's population size is roughly equivalent to 6.9% of the total number of people ever born*
- *A country's level of education attainment is a key determinant of the emergence and sustainability of democratic political institutions*
- *It is estimated that North Korea spends roughly one-third of their national income on defense.*

Just like Gapminder, all the material in this website is also freely available under the Creative Commons Attribution 4.0 International license.

Teaching Resources

This section provides notes, presentations, links to related blog posts and a list of books/articles related to various topics. The presentation includes relevant charts to illustrate various ideas.

Classroom Usage: We used this website while researching child labour for a school project with class X students. It not only provides data but also provides pointers on what correlations one should focus on and how data should be

read. However, some parts are for a much older audience and require a background in economics and/or statistics to interpret.

Website Features

- **Visual appeal:** The website is very clean with just a few menus which are clear and easy to navigate through. Unlike Gapminder, this website uses a lot of text interspersed with charts and other graphics. For this reason, it's probably less appealing to a younger audience. This website, just like Gapminder, renders well on a mobile device, but it is a lot easier to interact with the charts on a computer.

- **Loading speed:** The website loads without any obvious lags, and navigation is quite seamless.

- **Advertisements:** The website is ad free. There is a separate section on how one can donate.

- **Plug-ins/Additional Software:** I have used the website on Google Chrome and no additional software is required.

- **Search:** There is a search function that allows you to do a keyword search on the entire website.

- **Help/FAQ:** The "How to use" section answers some FAQs on how the website, charts and data can be used.

Both these websites demonstrate unique stories data can narrate. While accuracy and completeness of data of an event or a situation is crucial, it is perhaps appropriate visualization and representation that provide insight and understanding of the worlds that this data is meant to portray.

Many independent initiatives are also being undertaken to produce eye-catching visuals to recalibrate popular assumptions and beliefs about global statistics. Commemorating this is the Kantar *Information is Beautiful* Awards, which recognises excellence and beauty in data visualizations, infographics, interactives and information art. In February 2019, they announced the winners of the World Data Visualization Prize

(https://informationisbeautiful.net/2019/winners-of-the-world-data-visualization-prize/). This year's topic (much like the two websites reviewed above) focused on how governments around the world are improving citizens' lives, and the innovations that drive and measure success in this realm. (https://wdvp.worldgovernmentsummit.org/). Award categories included interactive, static and hand-drawn "napkin" and a final grand prize, the winner of which is An Alternative, Data-Driven, Country Map by Nikita Rokotyan. It uses an artificial intelligence technology called t-SNE (t-Distributed Stochastic Neighbor Embedding) to discover clusters of nations that may not be physically next to each other but are related by happiness score, health expenditure, investment in education and many other variables. This technology provides a methodology for visualising high-dimensional data in a two or three dimensional map. The aim of dimensionality reduction is to preserve as much of the significant structure of the high-dimensional data as possible in the low-dimensional map. This is done by using all dimensions to allocate a location to each datapoint. You then end up with a different image of the world, in which Japan, UK, USA, France are neighbours and so are Russia, Brazil and Argentina!

Just like Gapminder and Our World in Data, all the entries to this contest have truly innovative ways of representing data, making information, however dismal, truly beautiful!

USHA KRISHNAMOORTHY currently works at Centre For Learning, Bangalore, where she teaches mathematics and statistics. She has an undergraduate degree in Statistics from Lady Shri Ram college, Delhi University and a Master's in Business Administration in finance and marketing from IMI, New Delhi. She has worked for 18 years in the hydrocarbon exploration sector in the fields of data management, software engineering & marketing. She is interested in the area of mathematics and exploring the use of technology in developing logical reasoning and analytical skills. She may be contacted at ushakmoorthy@gmail.com