

India's Economic History: An Exercise in Visual Storytelling

Geet Kalra*

Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139

ABSTRACT

This paper explores the use of visual storytelling technique of “scrolly-telling” to convey the conundrum of India being one of the richest countries in the world while many of its citizens are amongst the poorest people in the world. The paper relies on depicting this by first highlighting the stark difference in GDP (Gross Domestic Product) and GDP per capita and then allowing the exploration of various factors that affect these two across countries. There are many visualizations layered on top of the scrolly-telling narrative to aid in knowledge assimilation.

Keywords: GDP, GDP per capita, scrolly-telling

1 INTRODUCTION

The growing wealth gap in the world is a big problem, however, a bigger problem might be the associated apathy. One clear example of this problem can be seen in the author's home country, India, where many of the world's poorest people live. Despite that, in terms of economic standing, the nation has been in top ten for a while.

It would also not be fair to say that people who are not victims of the wealth gap suffer from no other problem and should feel guilty all the time for not doing anything about it (if they aren't). It is this conflicting requirement, of sensitizing the people towards the problem, while not over-burdening them with guilt (which might lead to feelings of hopelessness, and probably more apathy), that the author has tried to address.

A narrative visualization has been designed which takes the user on a journey of discovery of the problem and immediately shows the user options of understanding the problem and its probable solutions better. The user is first sensitized towards the growing wage gap by means of highlighting the GDP and GDP per capita of India. The user is then shown visualizations where highly correlated parameters with GDP are presented. Finally, the user is presented a state-wise breakdown of economy of the nation and how the states harness their wealth, in hopes of allowing the user to better see why some places fare differently than others. The intention was to present the user with a problem, not overwhelm them and give them hope to deal with it.

One of the key tasks of visualization designers is packing large amounts of information for consumption by an audience without overwhelming them. The researchers also have another task, one of ensuring the audience takes away a message (if not the intended message). These are the two key foundation principles for designing the current visualization.

2 RELATED WORK

The design decisions in this work are influenced by the works of visualization researchers who have explored ways to improve information retention and interpretation by users.

Bateman et al [1] concluded in their papers that the long term recall of charts by users was much better when they were shown something that was embellished as opposed to plain information. Although they found that the recall of information in those charts was similar for a bland and “embellished” visualization, the user would be much more likely to go back to the source of information that they remember.

As a visualization designer, ensuring memorability can be just one part of the story. The bigger part can sometimes also be ensuring the intended message goes through (and the unintended doesn't!). Jessica Hullman and Nicholas Diakopoulos explored how the way in which a visualization is designed can affect end-user interpretation [2]. They explore certain dimensions of interpretations, for instance, the psychological affect that transparency clues or proper acknowledgements can have on user perception/interpretation of the visualization.

There are many articles and visualizations online that deal with the economic growth of India, or the low per capita GDP. There are visualizations that compare India with China and try to ascertain why things are the way they are. There are many visualizations which present a dashboard for the user to see where many different parameters have been plotted. The common ways of sharing information are lengthy news articles and static pie chart/bar chart visualizations.

Scrolly-telling (a combination of “scrolling” and “story-telling”) has also picked up as a means for communication a message and there are many websites based on this method of data sharing.

3 METHODS

The methodology can be broadly divided into the following steps

3.1 Data Preparation

To begin with, curated world bank datasets [3] on the change in parameters related to the state of economy, science and technology, and education for the countries of the world were studied. These datasets were joined together using a python script, based on the powerful pandas library [4]. The nature of the join can also be referred to as the “inner join”, where only those rows were selected from the tables which had values for a given country in a given year.

As a next step in data processing, it was realized that even with just three datasets, there were over 80 parameters to be assessed. The easiest way to manage such a large amount of information was determined to be a correlation analysis of all the parameters with the value of interest, the GDP in this case. Figure 1 shows a

* geet@mit.edu

correlation heatmap where the effect of different parameters was studied on the GDP of India. A similar heatmap was also generated for China. The parameters that had a high correlation with GDP for both countries were selected for representation. Note that due to the high density of parameters, they had to be replaced with placeholder alpha-numeric labels.

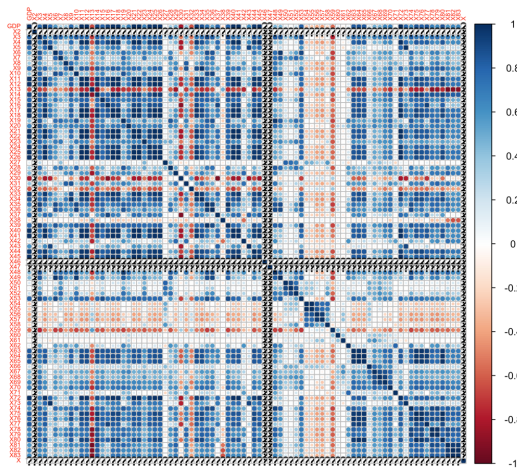


Figure 1: Correlation heatmap for India

3.2 Data Representation

After having decided the parameters of interest, the next question was the mode of communicating the information to the end user. It was ascertained that such information is usually of interest to researchers in the field and laypeople may find it un-interesting. Scrolly-telling was thus selected as the primary mode of sharing these details with the user. The powerful D3.js [5] JavaScript library was used to make an interactive and animated webpage for the users to first experience the motivating factors and then the associated information itself.

Memorability and visual appeal were decided to be a key parameter and thus a captivating image was first designed for the landing page of the visualization. As the data available was from 1960 to 2017, it made sense to add the images of India's first and current leaders, as overseers of this growth. Figure 2 shows the landing page for the viz.



Figure 2: Landing page of the visualization

4 RESULTS

On top of the scrolly-telling platform, four key visualizations were designed:

The user first sees an animated bar chart race where they can see how India and China came at the top of the GDP table starting in the late 80s and early 90s. Figure 3 shows a snapshot of the bar chart race.

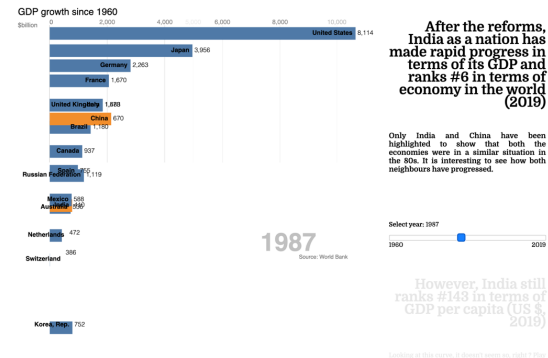


Figure 4: Bar chart race between GDP of countries from 1960 to 2019

Next the user is shown a simple line chart which shows the increasing GDP per capita of India. The user is encouraged to engage with the visualization and given some cues on how to do so. There are buttons that clearly indicate the next logical steps i.e. compare India to the poorest nations in the world and then to the richest. As the user compares India with the richest, they realize the stark difference between the lives of Indian people and those who are the richest. The intended affect is of curiosity and desire to go further.

Next the users are presented with the highly correlated factors in a "Hans Rosling" styled visualization where GDP is plotted on Y axes and the correlated factors on the X-axis. The intention is to allow the user to see changes in GDP with changes in parameters like Foreign investment, trade in goods and services, investment in education, promotion of research, etc.

Finally, the user is allowed to see the state of affairs in India by means of an interactive map. A Choropleth map of India based on State-wise GDP is shown. The user can see clusters of "rich" and "poor" states here and can also click on a state to produce a pie chart depicting how that state earned its revenues for the same given year. It is important to ensure that the color gradient in such maps is not misleading for the users. To ensure this, a custom color scheme was generated by using online websites that allow conversion of colors to hex codes that can be directly used in the code. Figure 5 shows a screenshot from cssgradient.io/ and Figure 6 shows the map generated using these gradients. Figure 6 also shows the interactivity feature where the user can click on any state and see the distribution of the state's GDP into different sectors of the economy.



Figure 5: Generating custom color scheme for the Choropleth map

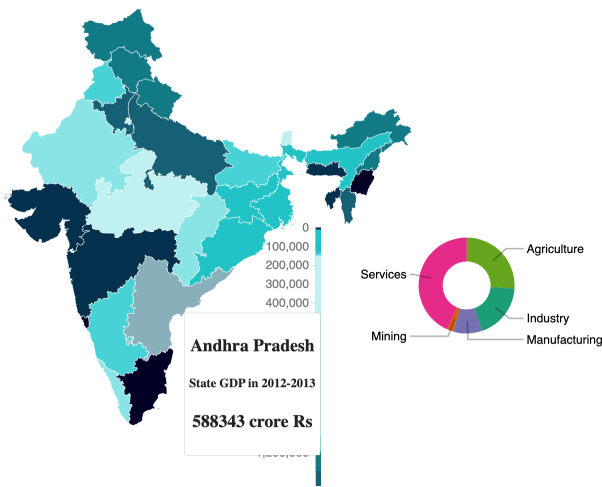


Figure 6. Visualizing of state-wise GDP and its segmentation in the state

5 DISCUSSION

It has been attempted to ensure that the user walks away with something that they can later recall as well as sensitize them towards the growing moral malaise in society. Avenues are also provided for the user to explore ways to understand why something like that happens.

The visualization was presented to some colleagues and friends of the author. The first comment from all the users was on the memorability of the visualization. Before people talked about the content, they commented on the fact that they liked the scrolly telling feature. The author would also like to share that a few weeks after the author presented a first draft of this work in his design class, a classmate was able to recall the work by means of the landing page. These instances allow the author to believe that the decision to implement scrolly-telling achieved the first intended aim, engaging the users.

It must also be noted that the information presented to the user is by no means exhaustive and there is still a lot left to be covered. However, sensitizing people towards the problem is the first step.

5.1 Future Work

Not all the correlated factors were included from the correlation study for the purpose of this system. This decision was mainly driven by time constraints and partly by the desire to first get feedback. Tamara Munzer has discussed extensively why timely feedback is of utmost importance [6].

The natural first step is including all the remaining parameters in the visualization as well. However, that needs to be done along with a summary of the effect of those parameters. Based on user feedback, a short-targeted statement upfront is valued by people.

The next steps would be to increase the interactivity and level of information provided by the last visualization. The user can be allowed to further explore what different sectors of the economy mean and should also be allowed to see the GDP per capita distribution across Indian states. Comparing that with GDP will allow us to highlight the states with possibly large income disparities and understand if the decisions that they have been making historically have had any bearing on those results.

The ultimate aim of the visualization is to be used by people to make policy decisions that can allow for targeted interventions in targeted locations. An ideal future case scenario could be that a districts past could be assessed based on multiple parameters and targeted interventions could be suggested to help people in those areas.

REFERENCES

- [1] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks, "Useful junk? the effects of visual embellishment on comprehension and memorability of charts," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, Apr. 2010, pp. 2573–2582. doi: 10.1145/1753326.1753716.
- [2] J. Hullman and N. Diakopoulos, "Visualization Rhetoric: Framing Effects in Narrative Visualization," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 12, pp. 2231–2240, Dec. 2011, doi: 10.1109/TVCG.2011.255.
- [3] Zening, *ZeningQu/World-Bank-Data-by-Indicators*. 2021. Accessed: May 19, 2021. [Online]. Available: <https://github.com/ZeningQu/World-Bank-Data-by-Indicators>
- [4] "pandas - Python Data Analysis Library." <https://pandas.pydata.org/> (accessed May 19, 2021).
- [5] M. Bostock, "D3.js - Data-Driven Documents." <https://d3js.org/> (accessed May 19, 2021).
- [6] T. Munzner, "A Nested Model for Visualization Design and Validation," *IEEE Trans. Vis. Comput. Graph.*, vol. 15, no. 6, pp. 921–928, Nov. 2009, doi: 10.1109/TVCG.2009.111.