

Introduction

Visualizations aid us in understanding the data, find patterns and inspire action among many other use cases. To ensure that the specific use case is met, we require good visualizations, that are able to represent the data coherently and correctly. In this assignment we will see 2 examples - 1 good visualization and 1 bad, and try to understand the reason for this categorization.

Example of a Good Visualization

This visualization shows the trends in Data Breach incidents over the years, and hence is relevant to a large set of people who access the internet everyday or access services that are connected to the internet.

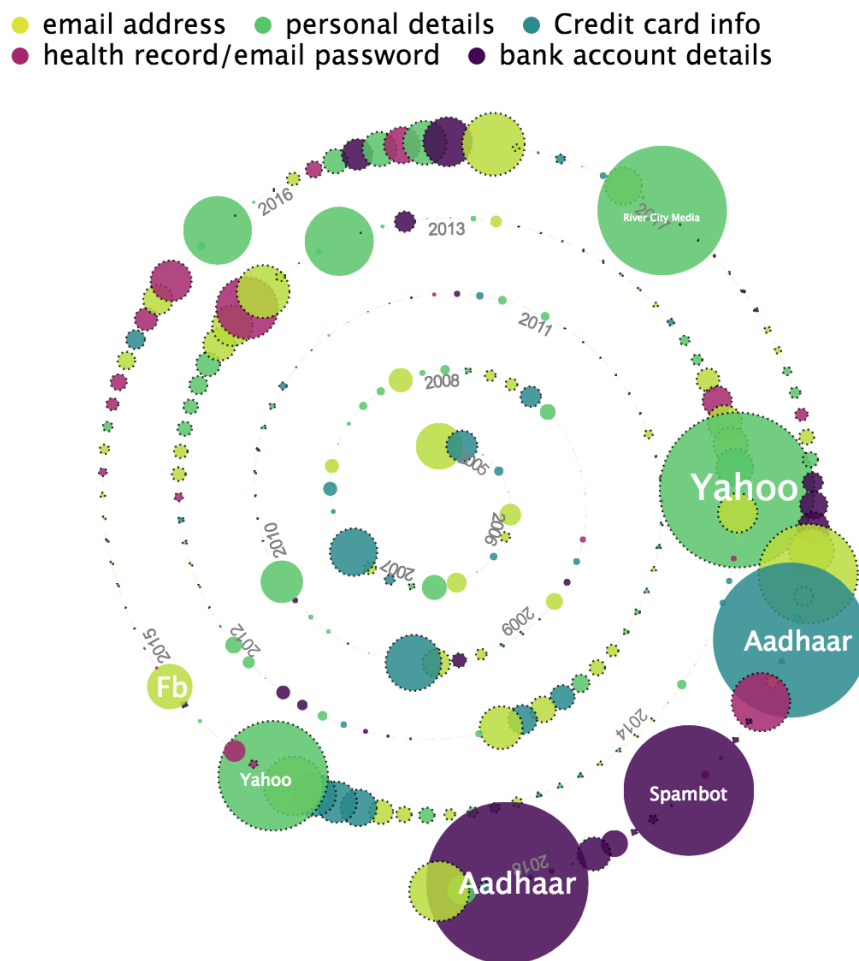


Figure 1: Chain of Data Breach Incidents

According to me, this is a good example as one requires only basic human cognitive capacity to understand it, and hence it can be understood by a wider audience and not a small set of technical experts. For data encoding, it aptly uses perceptual properties such as Position (for Time of occurrence), Area (size of Data Breach) and Hue (for Different Categorical Variables). Lastly, it is a unique way to present the information and definitely captures the viewer's attention.

Example of a Bad Visualization

The visualization given below depicts the Proportion of the Indian population having indoor toilets, along with its 12 Largest States. It is of importance as the National government had launched a campaign to build indoor toilets for 100% of the population in 2014.

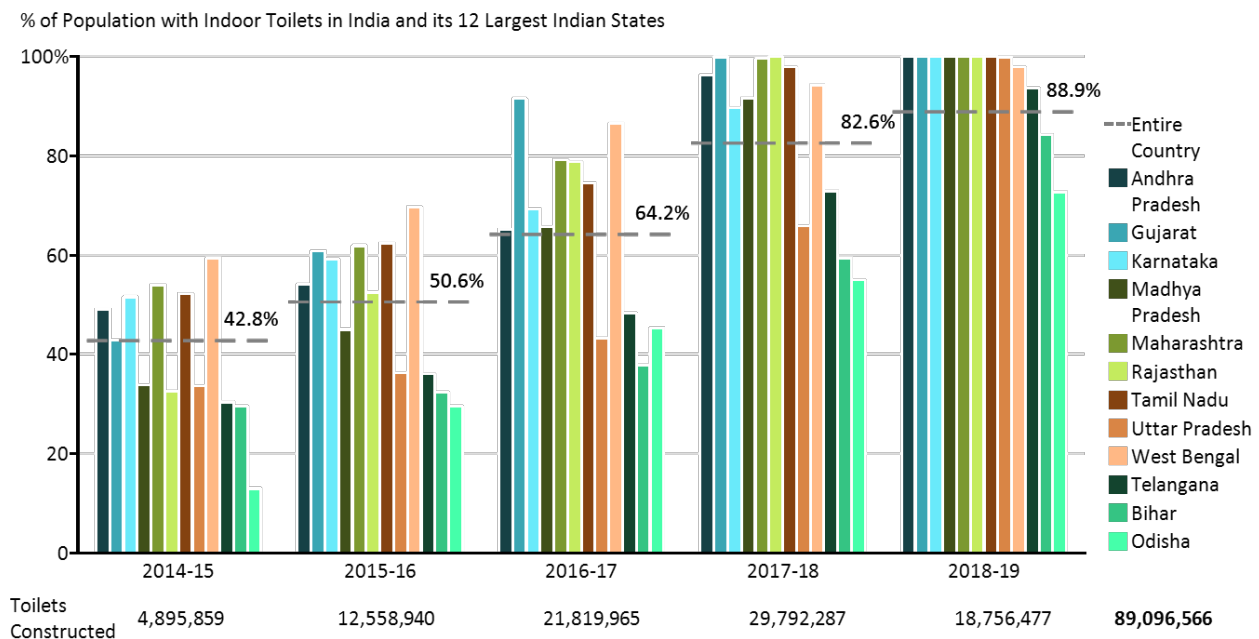


Figure 2: Percent of Population with Indoor Toilets in India and its 12 Largest States

I believe this is a bad visualization as it tells a story, but does not cover the whole picture. It shows that 88.9% (in 2018-19) of the Indian population has indoor toilets and only includes the largest states which also happen to be the most prosperous states. It does not include the states that are lagging behind in development and hence paints a more positive picture than reality. While Hues are generally used for Categorical variables, this visualization also makes the use of Color Intensity which makes the data harder to differentiate.

Sources

1. <https://medium.com/@yanhann10/the-chain-of-data-breach-incidents-2a98658c9db3>
2. <https://www.mekkographics.com/indias-toilet-building-initiative/>