Assignment 9: Data Visualization

Team: Venkatesh Duvvuri ([VED14@pitt.edu](mailto:VED14@pitt.edu)); Haifa Alnasser ([HIA11@pitt.edu](mailto:HIA11@pitt.edu)); Gopi Tata ([GKT3@pitt.edu](mailto:GKT3@pitt.edu))

Find at least three examples of graphical visualization of complex data.

Write a brief report that shows these examples (screen shots) and criticizes them from the point of view of some of the principles that we have talked about in class.

Are there any opportunities for improvement? Please make sure to include proper attribution of the screen shots (i.e., addresses where they can be found on the web).

1. Extreme Water Shortages are expected to Hit These Countries by 2040

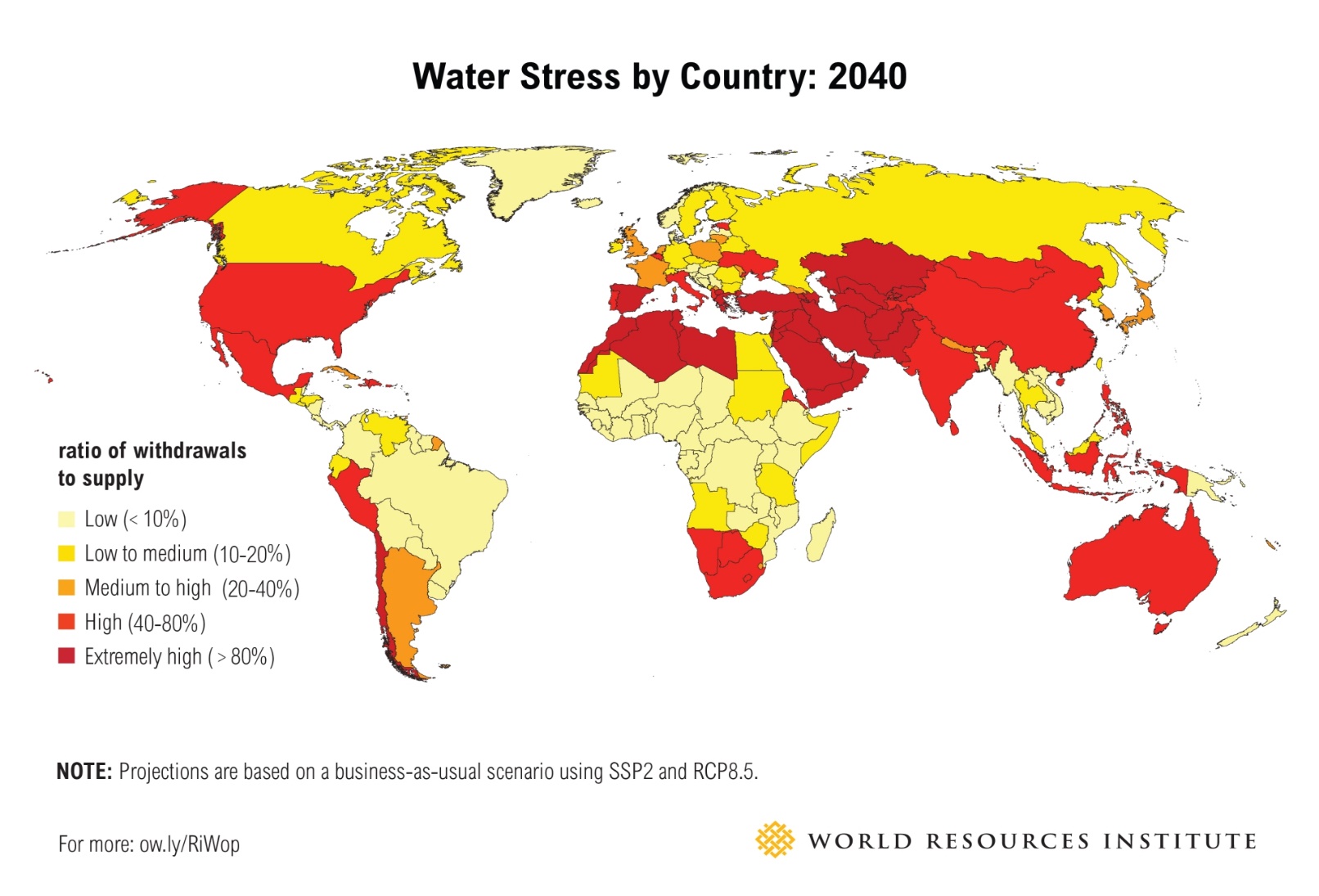
We strongly believe that the following water shortage forecast was depicted as below by World Resources Institute is very informative. The data visualization is about the water shortage or stress that would likely to be experienced by various countries of the globe by 2040. The graphical representation of the spatial data of ratio of water withdrawals to the supply by country.

Various countries have been categorized into five different categories namely Low, Low to Medium, Medium to High, High and Extremely High based on the percentage of water withdrawals to the supply. The categorization Labels have been provide showing the percentage and the color coding; like if the percentage is below 10% categorized as Low and light Yellow colored ,if the percentage is more than 80% is categorized as Extremely High with Dark Red and so on. The countries have been uniquely color coded for each category.

That being said one can:

1. Easily count how many countries fall under a given category.
2. Name the countries whichever country one is interested to know
3. Identify the regions at broader level than countries like continents, eastern vs western hemi-spheres or north vs southern hemi-spheres.
4. Identify or compare the situation among developed countries, or developing countries or by countries of large economies etc.

http://www.visualcapitalist.com/extreme-water-shortages-are-expected-to-hit-these-countries-by-2040/



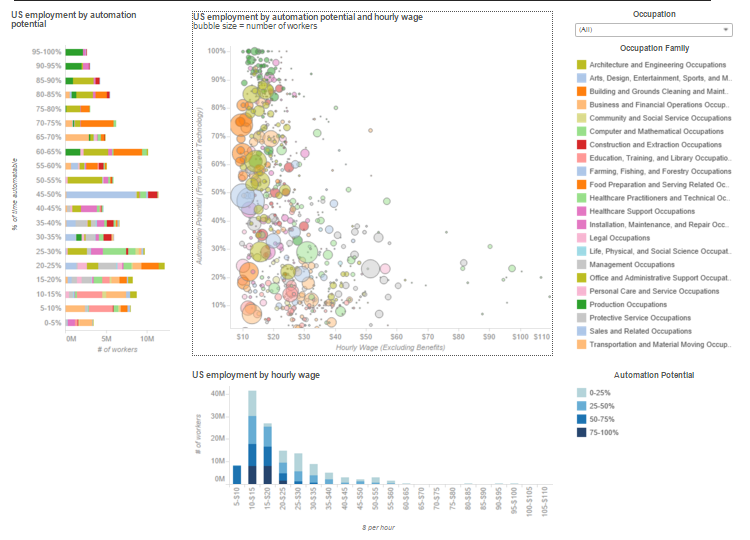
1. Charting the Automation Potential of U.S. Jobs

The following is the graphical represenation of the Automation of job in US for immediate future for a span of four years. The data visualization is bout US employment by Automation and Hourly wagge (three variables). There have been two graphical representations one by US employment and hourly wage and the other one is number of workers and hourly wage. We believe that the data charting of automation potential of US jobs is not useful. The reason is because:

1. It is rather difficult to follow various categorizations.
2. There are too many categorizations annd difficult to follow or summarize though each category has been color coded and labelled.
3. One cannot practically distinguish the color difference as there are many categories.
4. There are too many bubbles representing the number of workers and made the graph rather cluttered. Many bubbles have been over-lapped.

Improvements (Automation Potential vs Hourly Wage):

1. Use a different graphical representation like bar chart or Histogram for showing the Automation Potential vs Hourly wage.
2. May use multiple graphs for different Occupation Families or use broader categorization of Occupation Families.
3. Few more colors could be used for better visualization.

<http://www.visualcapitalist.com/charting-automation-potential-of-u-s-jobs/>  
  


1. The Periodic Table of Commodity Returns

The following is the graphical represenation of the periodic returns of commodity returns. The data representing the returns for various commodities from 2005 till 2015. The performance of each commodity has been depicted for each year. This is an example of representing Temporal data.

We believe this is a good presentation of temporal data; performance by each year. Each commodity has been distinguished by a different coloer and labelled. The Peroformance percentage has been mentioned right in the box that represents each commodity along with it symbol or Name.

From this data visualization one can:

1. Easily identify performance for a given commodity.
2. Compare performance for a given commodity across the years mentioned
3. Compare performance of different commodities for a given year or across the years
4. Easily follow legends , the color coding ; very high readability and useful.
5. Easily identify if the performance of a commodity is improved or reduced noted by the sign coding (poositive or negative).
6. Understand the type of commodity as the Name or Symbol of the commodity was mentioned.

Criticism:

1. Since this is representing the Performance over a period of time a Line chart (Trend chart) with Legends of the color coding and performance values mapped on the Line chart would be easier to understand the information quicker.

<http://www.visualcapitalist.com/the-periodic-table-of-commodity-returns/>  
  
