

DATA.STAT.770 Dimensionality Reduction and Visualization, Spring 2021, Exercise set 5

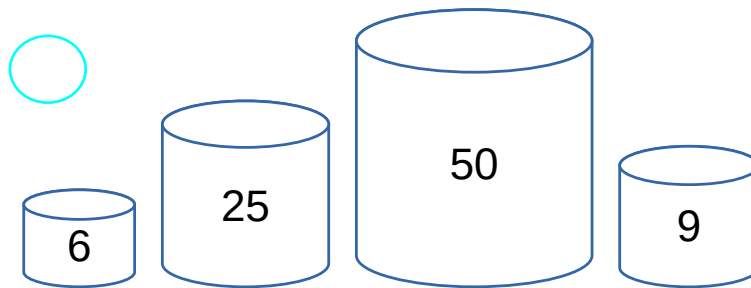
Part C: Graphical Excellence

Problem C1: Appealing Visualization

Look for an example of scientific visualization that you find particularly beautiful in a recent issue of a high profile journal (Nature, Science, etc...). Try to explain what makes it appealing or useful.

Problem C2: Lie Factor

Calculate the lie factor of the following graph. Do you think the concept of a lie-factor is useful or relevant in practice ?



Problem C3: Chartjunk and Data-Ink Ratio

Kenneth and Jeffrey are running a high-tech company. Unfortunately, business has been slow and the next interim report is not going to please the shareholders. Sales are declining and operating costs are going through the roof.

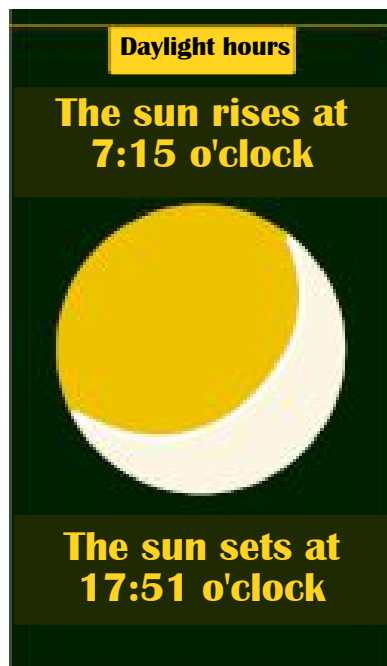
- Create some artificial data about sales and operating costs that corresponds to the description above.
- Help Kenneth and Jeffrey create a visualization of their sales and expenses from the past four quarters that makes the situation look less dramatic. You can use chartjunk, optical illusions, “creative layout”, etc. , to display the data.
- Kenneth and Jeffrey got busted by the SEC for providing incorrect information about the company’s financial situation to the shareholders. The new management asks you to create a truthful visualization of the sales and expenses. Do this by improving a standard graph from MS Excel, OpenOffice, Matlab, etc., according to Tufte’s principles. Present both the standard

graph and the improved one. Discuss what improvements you made and why they are useful. Use the same data as in the previous part.

Problem C4: Improving Real-world Visualizations

The visualizations in the pictures on the next page are translated into English from real infographics that have appeared in the widely distributed HS Metro magazine (<http://www.hs.fi/metro/>).

- Analyze the figures. Do they represent good data visualizations? Do they involve chartjunk? Can the graphics cause misleading inferences about the data? Do they misrepresent the size of some effect - if so, what is the lie factor?
- For each figure, create an improved visualization.
- Consider the figure published by Finland's Ministry of Finance on their Twitter account on February 28, 2017: <https://pbs.twimg.com/media/C5vOmxkWYAEzBRh.jpg>. This figure represents amount of unemployed job seekers (purple curve) and amount of open jobs (yellow curve) from 2007 to 2016. Discuss the figure - can it misrepresent the size of some effect? If so, can you suggest an improved visualization?



From HS Metro, March 2, 2016, translated into English



From HS Metro, January 14, 2016, translated into English



From HS Metro, October 17, 2016, translated into English