DS 100: Principles and Techniques of Data Science	**Date: February 9, 2018**

Discussion #3

Name:

Reading and Evaluating Figures

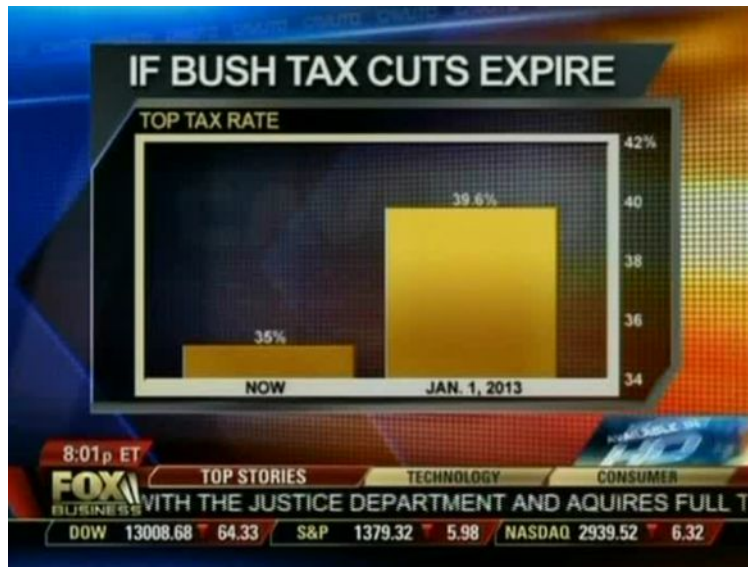1. First answer the questions regarding Figure 1 below and discuss your answers.

Figure 1: Fox Business chart from 2012, comparing the top marginal tax rate for 2012 and 2013.

 (a) What message does this plot convey?

 (b) How do they physically convey this message?

 (c) Would you change this chart? If yes, how so? If not, why not?

2. Watch Vox's video given at this link watch `http://bit.ly/data100-yaxis`. Discuss your opinions on this video.

3. Discuss the problems with keeping the visualizations below as they are. Color versions are given in the document found on the course website.
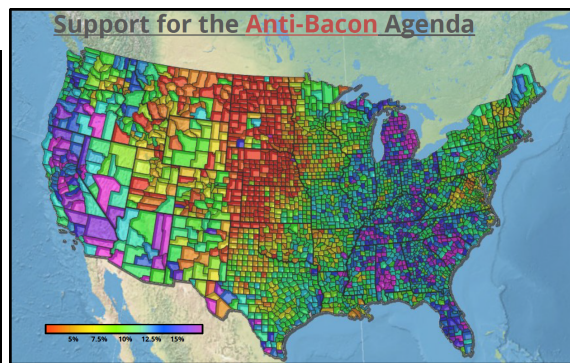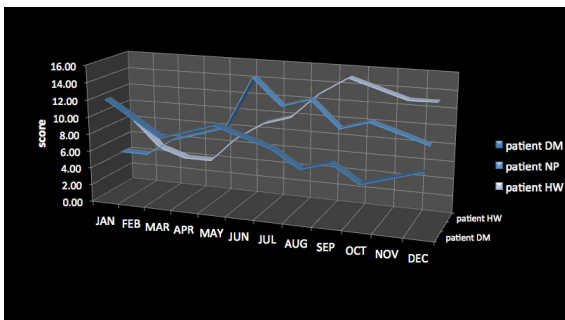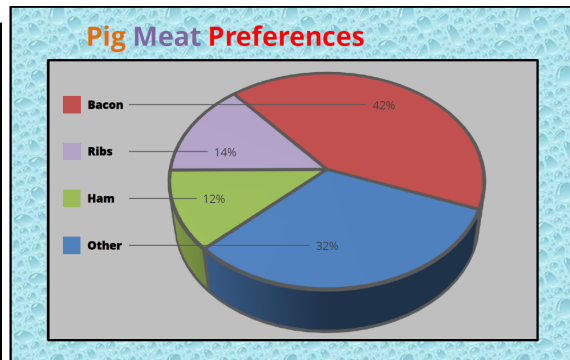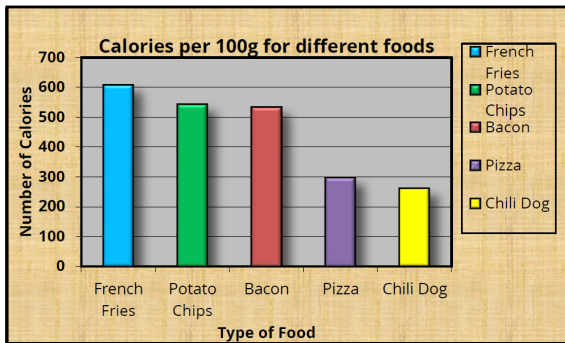
4. The chips that are present in your computer contain electrical components called transistors. Intel is one of the leading manufacturers of these chips; they released the first chip for home computers in 1979.

   We'd like to visualize the improvements in chips that Intel has made since 1979.

   Download a copy of the discussion from the course website to see the image in color.
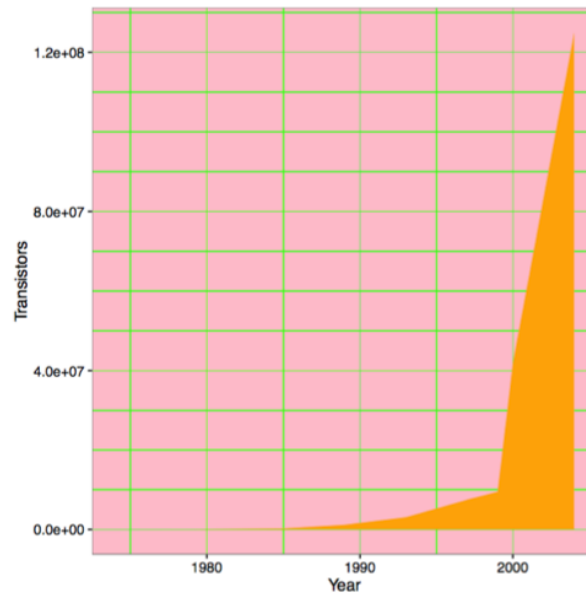

   (a) What is this plot trying to communicate?

   (b) How does it communicate it? Discuss what data are being mapped to what physical properties of the plot.

   (c) Are there visual elements that seem unnecessary? List them.

5. Now we look at an attempt to improve the plot.
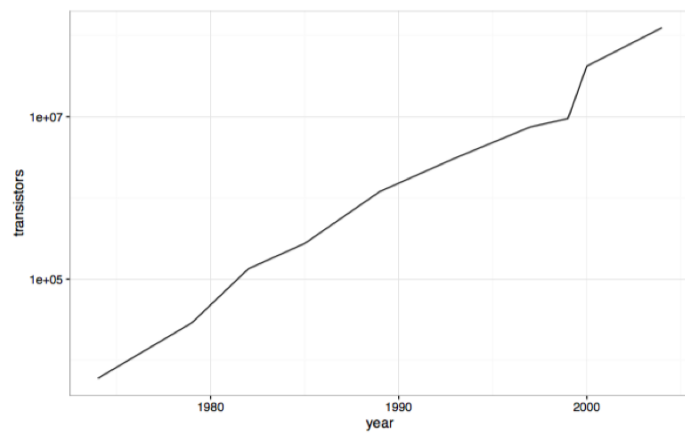
(a) What is clear in this plot that wasn't clear in the plot before?

(b) What design decisions made understanding the data easier?

(c) Are there still properties of the plot that seem like they need more context?

6. Here's a third attempt to improve the plot. MIPS stands for Millions of Instructions Per Second. Download a copy of the discussion from the course website to see the image in color.
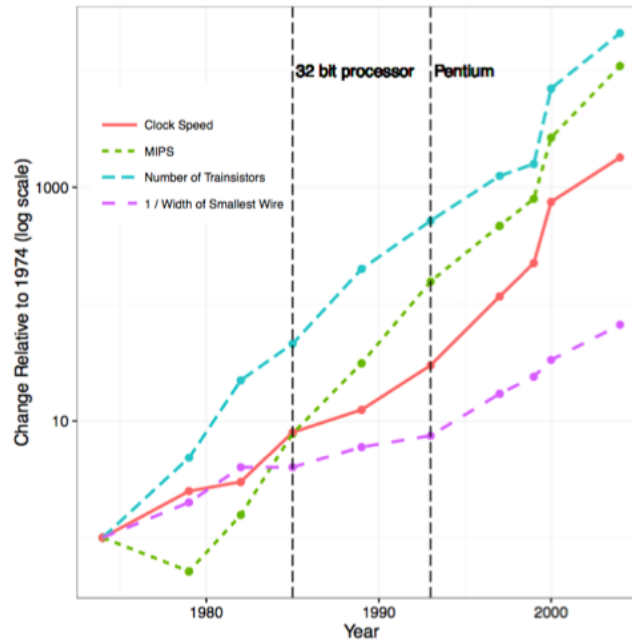
(a) What is clear in this plot that wasn't clear in the plot before?

(b) What design decisions made understanding the data easier?

(c) Are there still elements that seem like they deserve context?

(d) Why do we get to draw the four curves on the same plot? The four curves seem to represent data with different units.

Logarithmic Transformations

7. In our discussion, $\log$ will refer to the natural logarithm (also written as ln in other texts). A logarithm in base b will be written as $\log_b$. Sometimes, it will be convenient to write exponential as $\exp$ i.e. $\exp(x) = e^x$. Let $a, b \in \mathbb{R}$ be real numbers and $x, y \in \mathbb{R}_+$ be positive real numbers. Which of these statements about logarithms are true?

   A. $\log xy = \log x \log y$

   B. $\log(x + y) = \log x + \log y$

   C. $\log 1^a = 0$

   D. $\log \frac{x^a}{y^b} = a \log x - b \log y$

   E. $\log_x y^a = a \log y / \log x$

   F. $a^x \exp(b^y) = \exp(x \log a + by)$

8. When visualizing the data, what are some reasons for performing log transformations on the data?