­­­­**Tufte, Chapter 4: Data-Ink**

By Z. Trozenski

**Introduction and Key Terms**

Graphics should represent the substance and sense of the data and not draw attention away from the data and towards the design. You don't want your aesthetics to compete with the information you're trying to present. **Data-ink** refers to the ink dedicated towards representing the actual data or, according to Tufte, “the **non-erasable** core of a graphic, the **non-redundant** ink arranged in response to variation in the numbers represented”.

**Concepts**

#1 – “Above all else, show the data”

A large share of ink on a graphic should present data-information, changing as the data changes.

#2 Maximizing share of Data Ink

Data-ink ratio = data ink / total ink used to print the graphic

= proportion of a graphics ink devoted to the non-redundant display of data information

= 1.0 - proportion of a graphic that can be erased without loss of data information

(In essence you want to devote the most ink to displaying the information found within your data, with as little as possible devoted to non-data (i.e. legends, ticks, axes, etc.))

The more ink dedicated to the data itself the better: " Maximize the data-ink ratio within reason"

Every blot of ink on a graphic requires a reason to be there, the most compelling reason is that the ink shows new information. Better yet, is to think about representing the most amount of data-information with the least amount of ink.

#3 Two Erasing Principles

"Erase non-data ink, within reason"

"Erase redundant data-ink, within reason."

When you have ink devoted to items in your graphic which is not explicitly devoted to depicting statistical information it clutters the graphic and muddies the clarity of the graphic.

Redundant data ink: depicts the same number over and over (i.e. labels over shaded bars, shaded bars, labels in bar charts, etc.)

Gratuitous decoration and reinforcement of the data generate too much redundant ink

Bilateral symmetry is also redundant (though I disagree with this point) in box plots, open bars, faces

Redundancy does have a usefulness in that in can order complexity and establish context, allowing for comparisons over different parts of the data, and creating an aesthetic balance. Sometimes cycles in repeated data allow for a reader to navigate the graphic better without having to start over at the beginning of the graphic.

#4 Application of Principles in Editing and Redesign

You should feel empowered to revisit your graphics and prune out unnecessary data. Some best practices to consider: over multiple iterations of the graphic, remove different elements of reference ink (gridlines, precisions marks, axes, etc.) one at a time and examine the graph. If the removed element creates confusion for the reader either by revoking context or establishing order, than that element should remain. Any element, which when removed, is not hindering the comprehension of the data can be left out. This may also lead to more efficient uses of reference ink.