WHAT SHOULD WE TEACH IN AN INTRO STAT COURSE?

Webster West
North Carolina Stat University
Integrated Analytics LLC

WHY SHOULD WE CHANGE?

- There are now some very large data sets consisting of millions of observations.

- There are millions of data sets consisting of a smaller number of observations.

- Most of these data sets are readily accessible online.

- Almost none of these data sets are appropriate for classical inference!!

DROP INFERENCE?

- Perhaps inference should be minimized.

- Do we really need to go though all the permutations of one-sample and two-sample tests and confidence intervals for means and proportions?

- Do students lose a true appreciation for the concepts?

SIMULATION/RANDOMIZATION TO THE RESCUE?

- Some instructors have opted to rely on technology to introduce computationally intensive methods like simulation and resampling.

- What are the potential benefits?
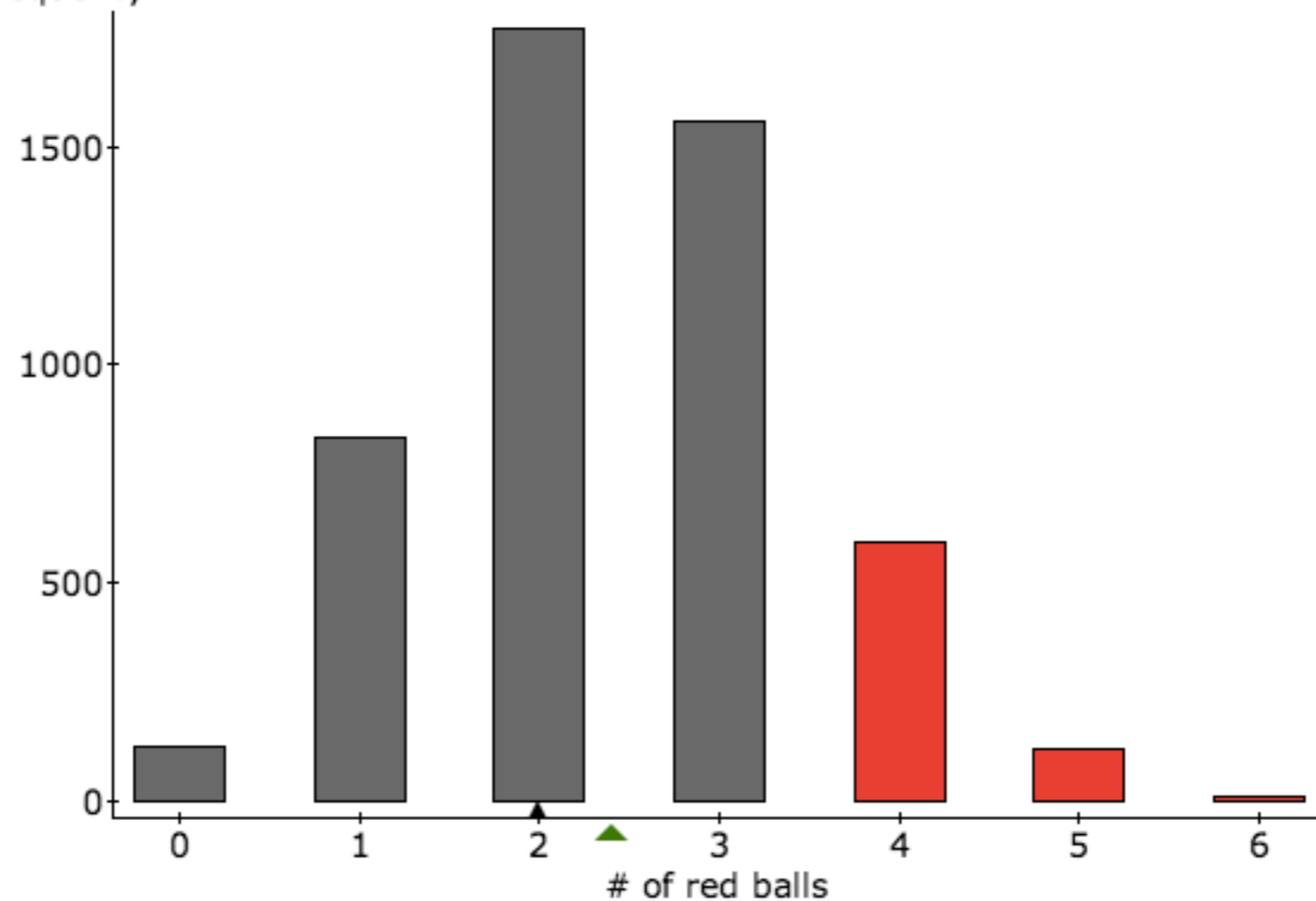
- What are the potential pitfalls?

# SIMULATION EXAMPLE

RANDOMIZATION EXAMPLE
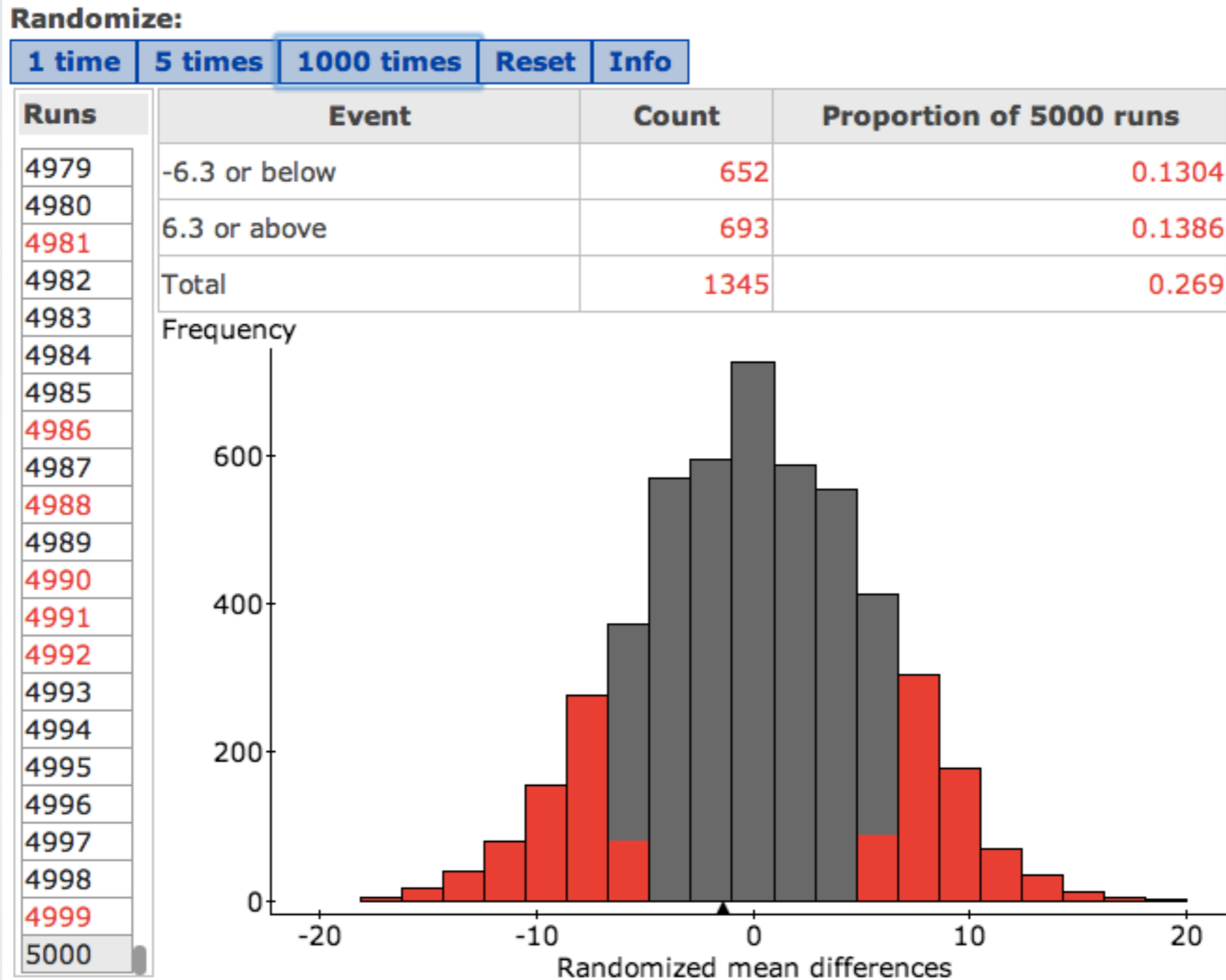
HOW ABOUT A BLACK BOX?

- Why not skip theoretical development and formulas altogether?

- Simply let software do the heavy lifting!

- Students focus on assumptions and interpretations.

- What are the potential benefits?

- What are the potential pitfalls?

WHAT SHOULD WE ADD TO THE CURRICULUM?

- More emphasis on technical skills needed to really analyze data in the real world: Data manipulation, formatting, subsetting and grouping

- Data visualization and data-based story telling

- More advanced statistical methods: logistic regression, multiple regression, applied time series, analyzing textual data, etc…

PRACTICAL EXAMPLES
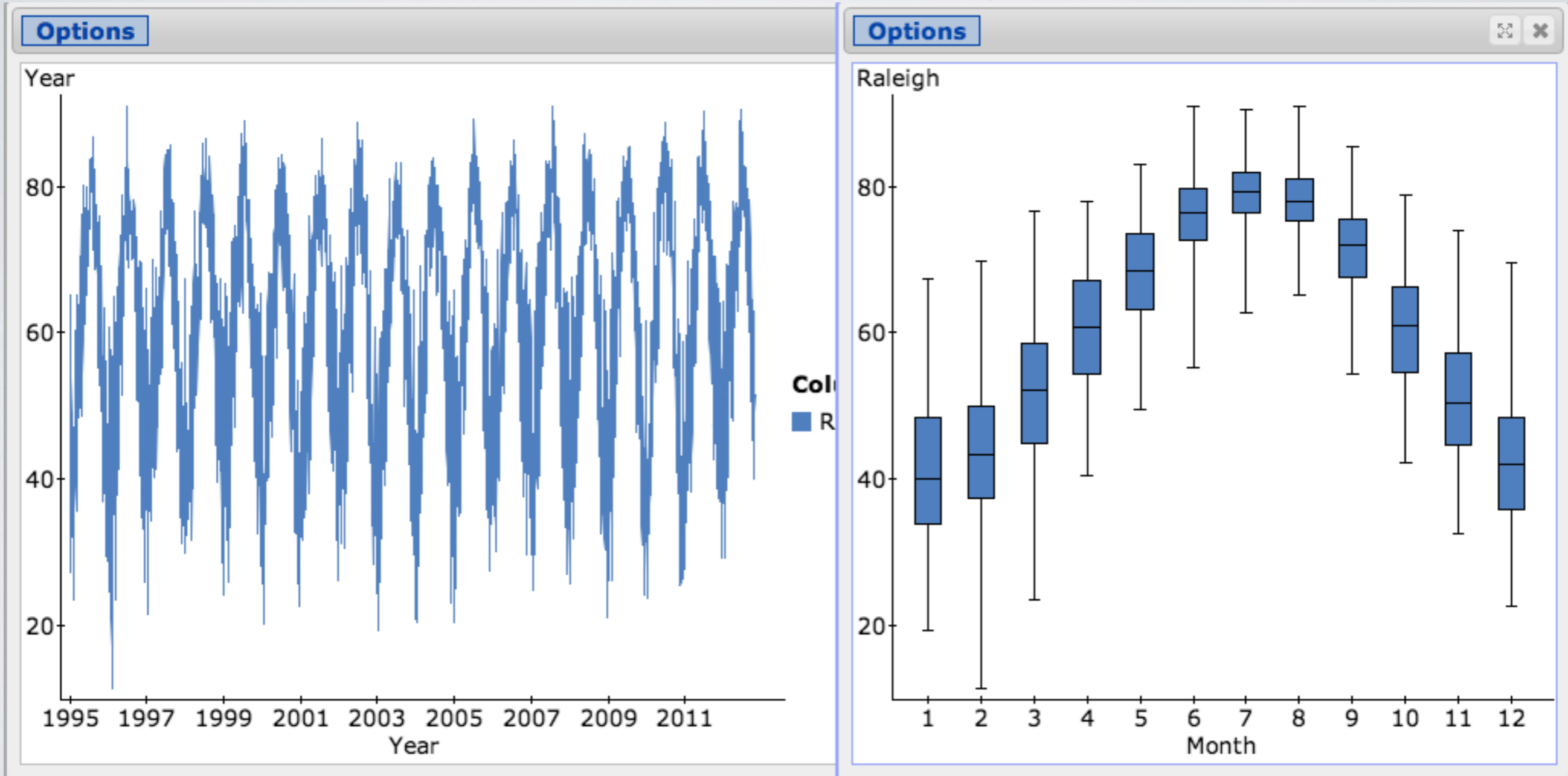
TEXT EXAMPLE
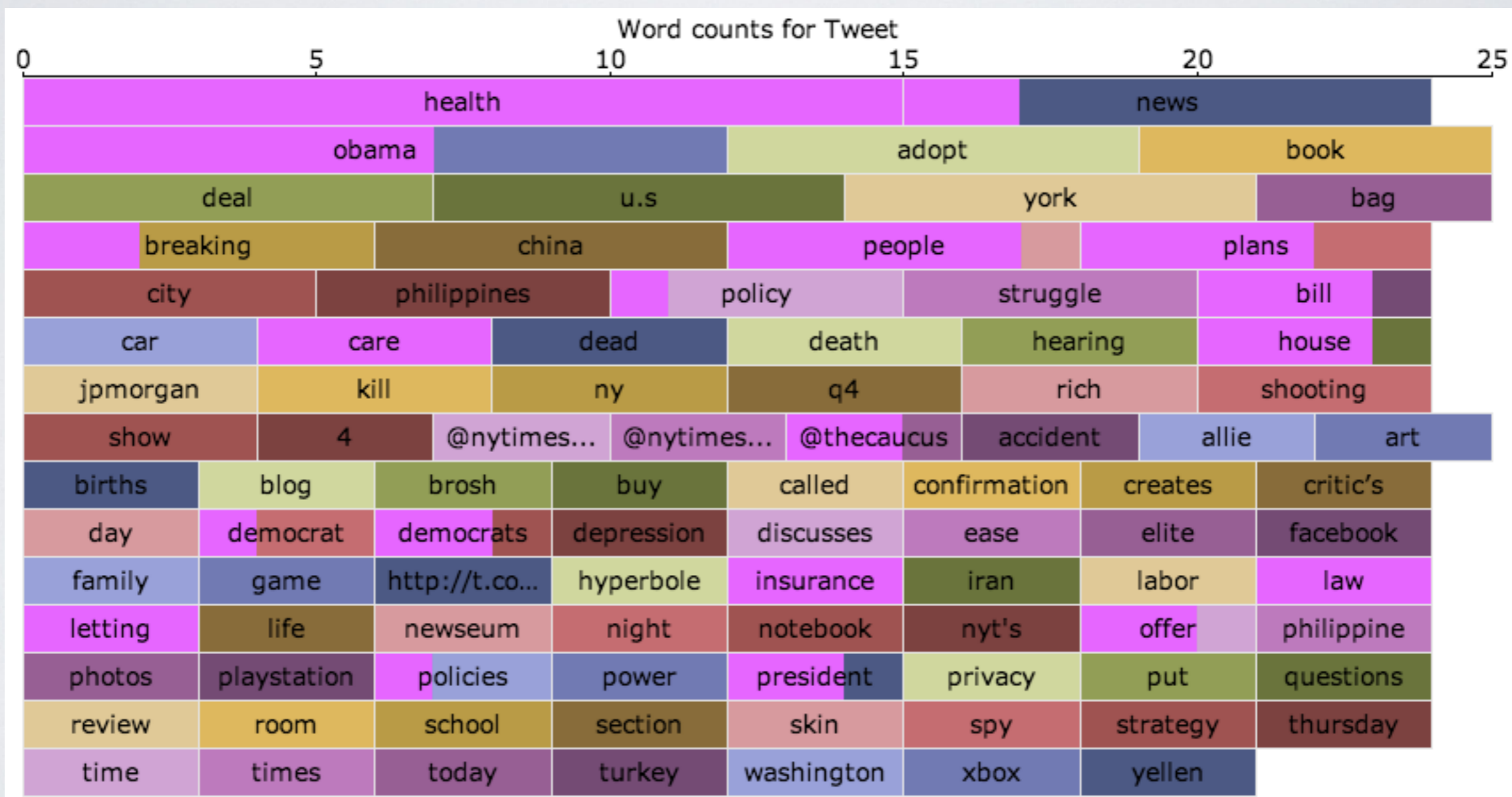
A FEW QUESTIONS

- What else can we trim down? Probability perhaps?

- What are the overarching principles that we should teach related to practical data analysis?

- What are the overarching principles that we should teach related to data visualization/story telling?