

What is Data?

Author: Nicholas G Reich

*This material is part of the **statsTeachR** project*

Made available under the Creative Commons Attribution-ShareAlike 3.0 Unported License: http://creativecommons.org/licenses/by-sa/3.0/deed.en_US

Newton showed that the book of nature is written in the language of mathematics. Some chapters ... boil down to a clear-cut equation; but scholars who attempted to reduce biology, economics, and psychology to neat Newtonian equations have discovered that these fields have a level of complexity that makes such an aspiration futile.

This did not mean, however, that they gave up on mathematics.

A new branch of mathematics was developed over the last 200 years to deal with the more complex aspects of reality: statistics.

- Yuval Noah Harari

Sapiens: A Brief History of Humankind

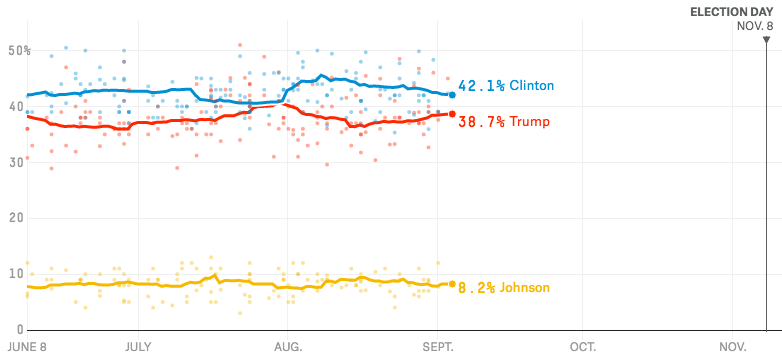
Data are measurements from our
imperfect, noisy world.

In practice, what are some sources of the noise?

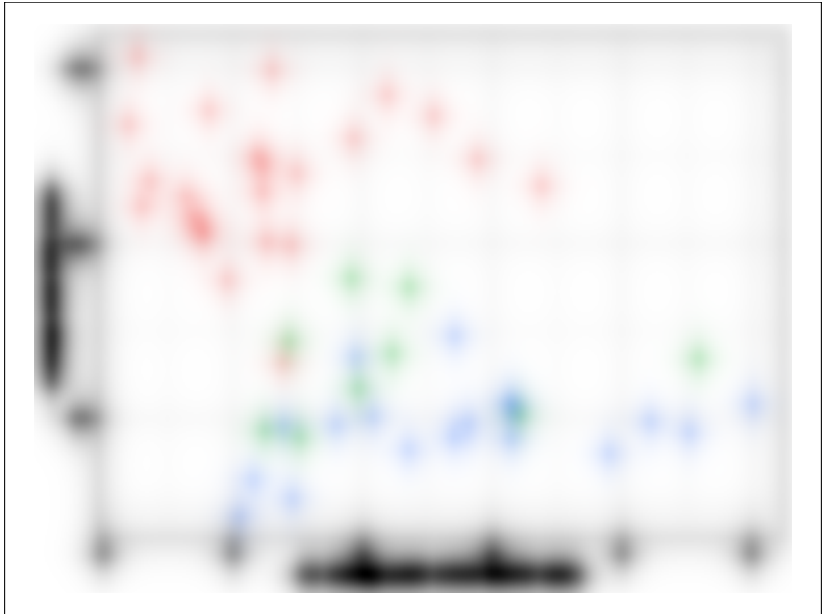
Whence the noise?

Plot of M/D heights data?

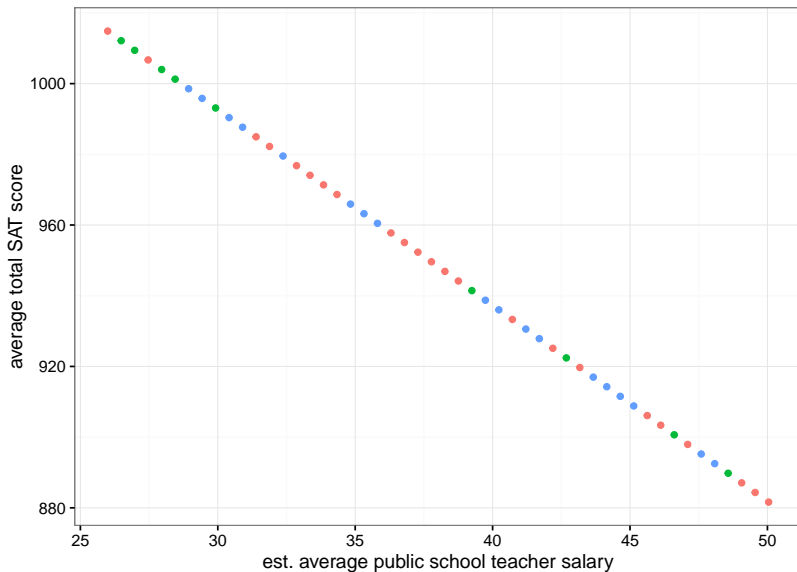
Whence the noise?



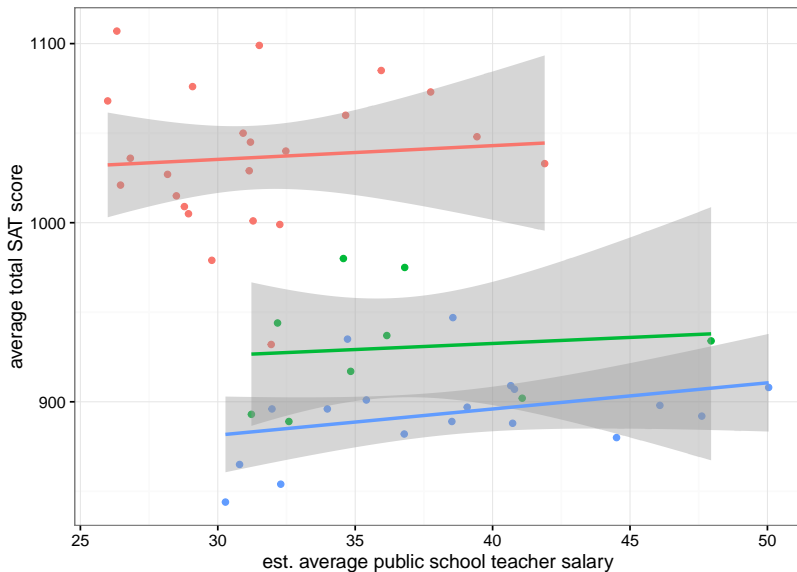
Statistics brings data into focus



Statistics does not eliminate noise



Statistics speaks a language of uncertainty



Your project

What types of variables are you collecting?

Who are you collecting data on?

What population are you trying to draw conclusions about?

Sample vs. Population

Do you expect your sample to be representative of the population?
Why or why not?