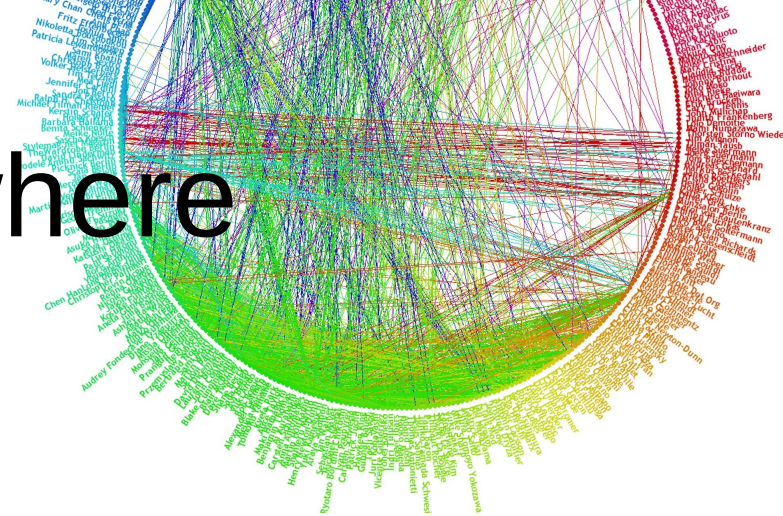# Eric Pitman Summer Workshop in Computational Science
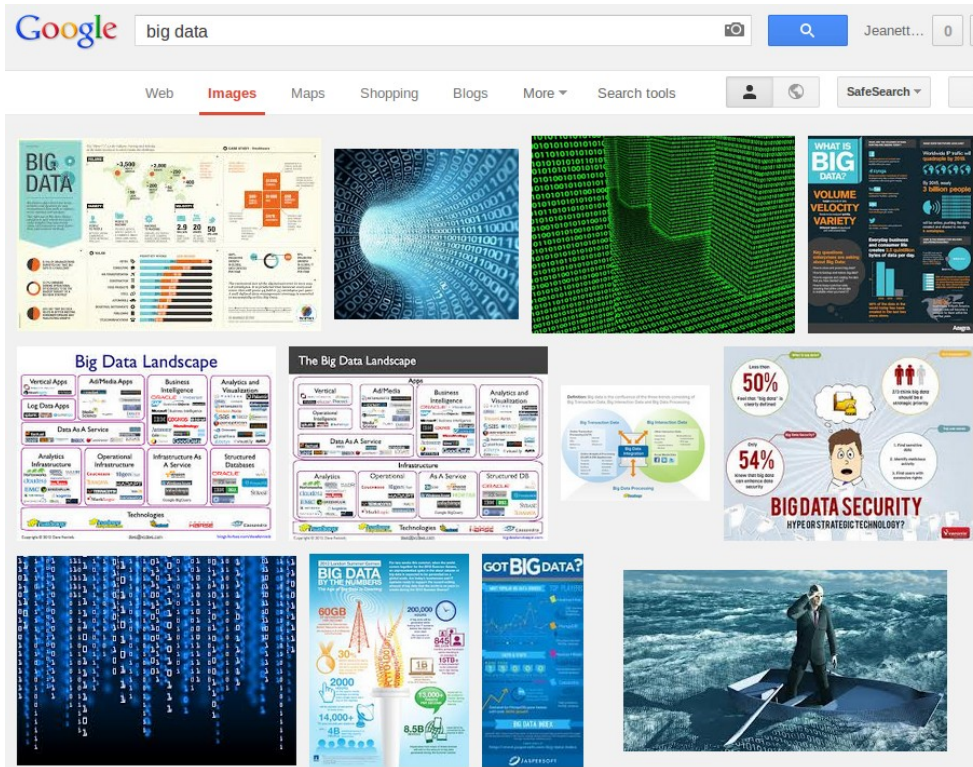
Intro to R

...an introduction to R, statistics, programming, and getting to know datasets

Data is Everywhere

- For example:

  – Science/engineering/medicine

  – Environmental science/Social science/Law enforcement

  – Finance and marketing

  – Social media

- How do we come to an understanding of what a dataset contains?

- Can we draw conclusions from a dataset?

- Let's taste the complexities for ourselves

"Big Data" means Three Things

- *Volume*: lots of data

- *Velocity*: coming at you fast—Twitter, 7TB/day

- *Variety*: text, pictures, video, etc.

Our Plan for the Workshop

- Introduce the R language

- Do some programming

- Tours, demos, guest lectures: sample some science & engineering

- Examine, model, and visualize datasets

- Project: explore and characterize protein crystallization data from HWI