Data Visualization

Connecting research to the world

Data Visualization

CONNECTING RESEARCH TO THE WORLD

Why is visualizing data important?

A big problem with social research (and many other types of research) today is that it doesn't reach many of the most important audiences (the public, research participants etc.)

What does research generally produce?

HUGE BORING REPORTS NOBODY WANTS TO READ!

Data Visualization
CONNECTING RESEARCH TO THE WORLD

Why the disconnect?

- Most researchers communicate in language only researchers use
- Most organizations paying for research don't know to ask for something different
- Like all other industries, researchers have their own dogma – change is difficult

Why the does it matter?

- Much of the research done is wasted
- People who are researched deserve to see the results
- The bigger the audience is, the more information you will draw from the research

Data Visualization
CONNECTING RESEARCH TO THE WORLD

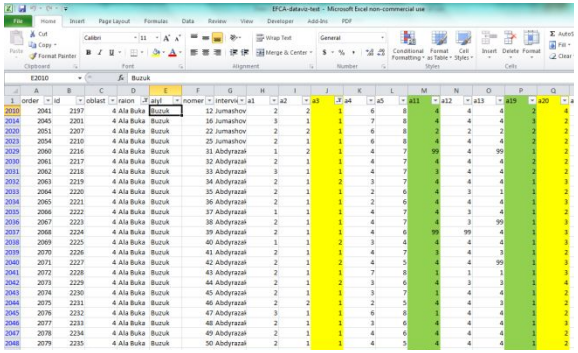
Visualizing data is about telling a story

- Every data point has a story – just like a character in a book
- There are relationships, and interactions with data – just like characters in a book
- How you present data can help people remember
- Good visualization can make people take action

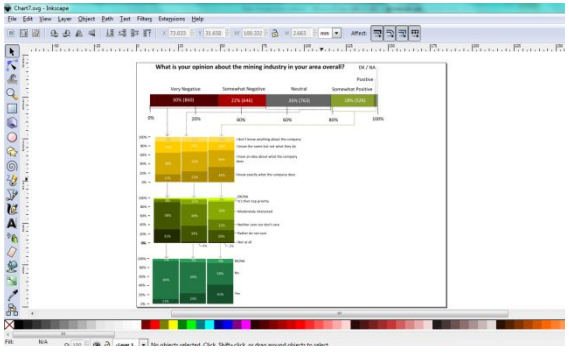
Data Visualization
CONNECTING RESEARCH TO THE WORLD

The art and science of data visualization

- Part scientist / researcher, part computer programmer, part graphic designer, part story teller



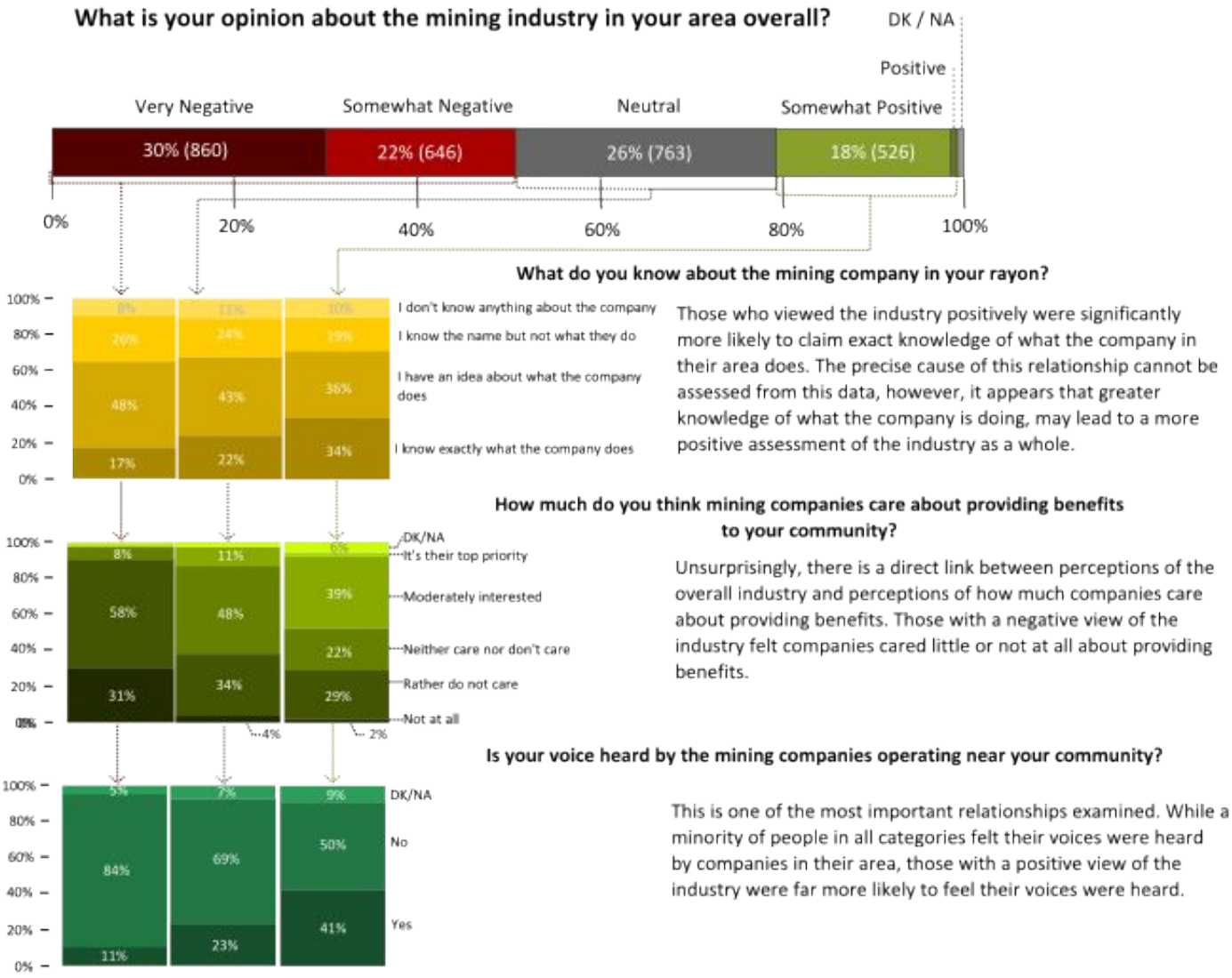
```
import matplotlib
matplotlib.use('Agg')
from pylab import *
import calendar

def webshow(img):
    savefig(img,dpi=500)
    print 'Content-Type: text/html\n'
    print '<img width="800" height="400" src="'+img+'" />'

genres = []
n = 0
for c in sorted_list:
    genres.append(sorted_list[n][0])
    n += 1

grosses = []
a = 0
for c in sorted_list:
    grosses.append(sorted_list[a][1])
```



Raw data → Computer programming → Graphic design

Data Visualization

CONNECTING RESEARCH TO THE WORLD



Those who viewed the industry positively were significantly more likely to claim exact knowledge of what the company in their area does. The precise cause of this relationship cannot be assessed from this data, however, it appears that greater knowledge of what the company is doing, may lead to a more positive assessment of the industry as a whole.

Unsurprisingly, there is a direct link between perceptions of the overall industry and perceptions of how much companies care about providing benefits. Those with a negative view of the industry felt companies cared little or not at all about providing benefits.

This is one of the most important relationships examined. While a minority of people in all categories felt their voices were heard by companies in their area, those with a positive view of the industry were far more likely to feel their voices were heard.

Data Visualization
CONNECTING RESEARCH TO THE WORLD

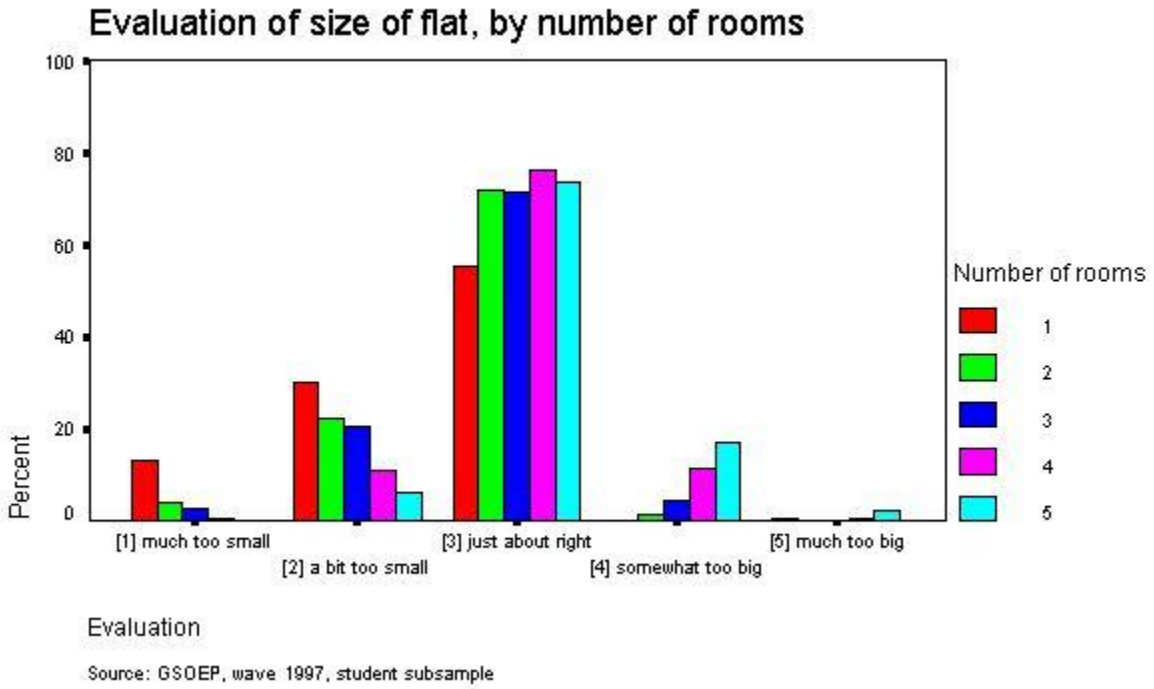
Common types of “out-of-the-box” visualizations

- Sociologist generally use SPSS
- Many other analysts use Excel
- Statisticians use “R”
- Graphic designers use Illustrator or Inkskape

All have their advantages and disadvantages as seen next…

Data Visualization
CONNECTING RESEARCH TO THE WORLD

SPSS:

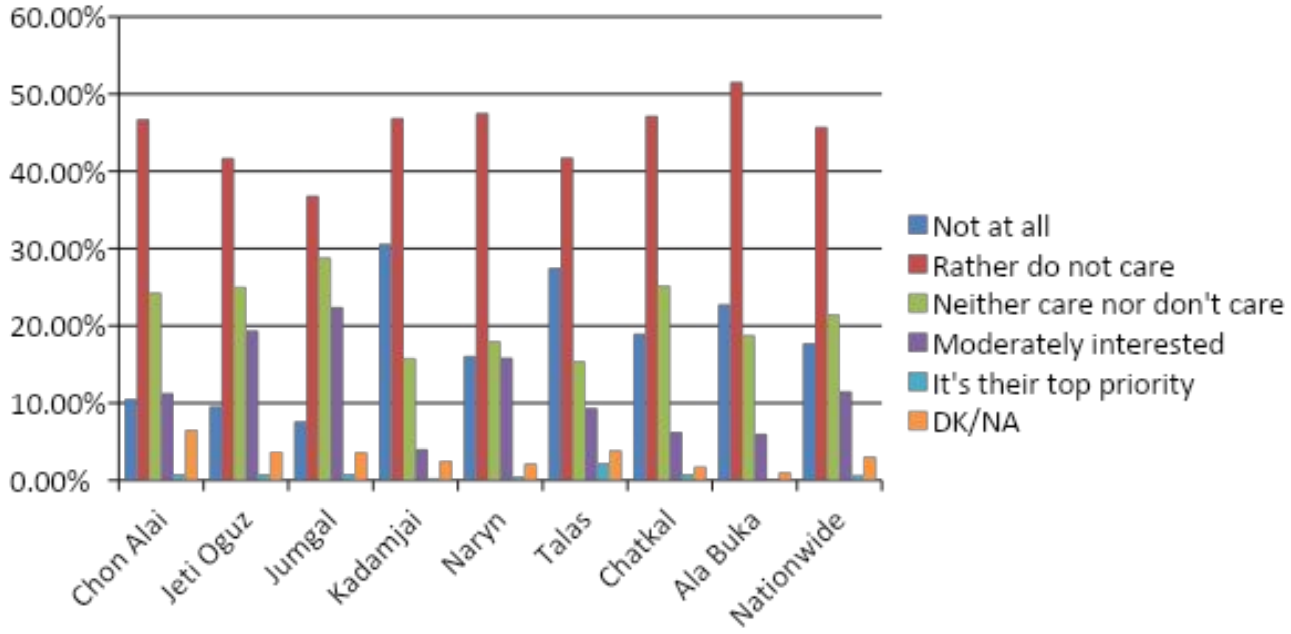


Advantages:
- Widely used in social sciences
- Extensive number of features
- Very rigorous analytical capabilities
- Able to handle large amounts of data

Disadvantages:
- UGLY charts
- Limited number of visualizations
- Incredibly expensive
- Not used by anyone but sociologists

Data Visualization
CONNECTING RESEARCH TO THE WORLD

Excel:



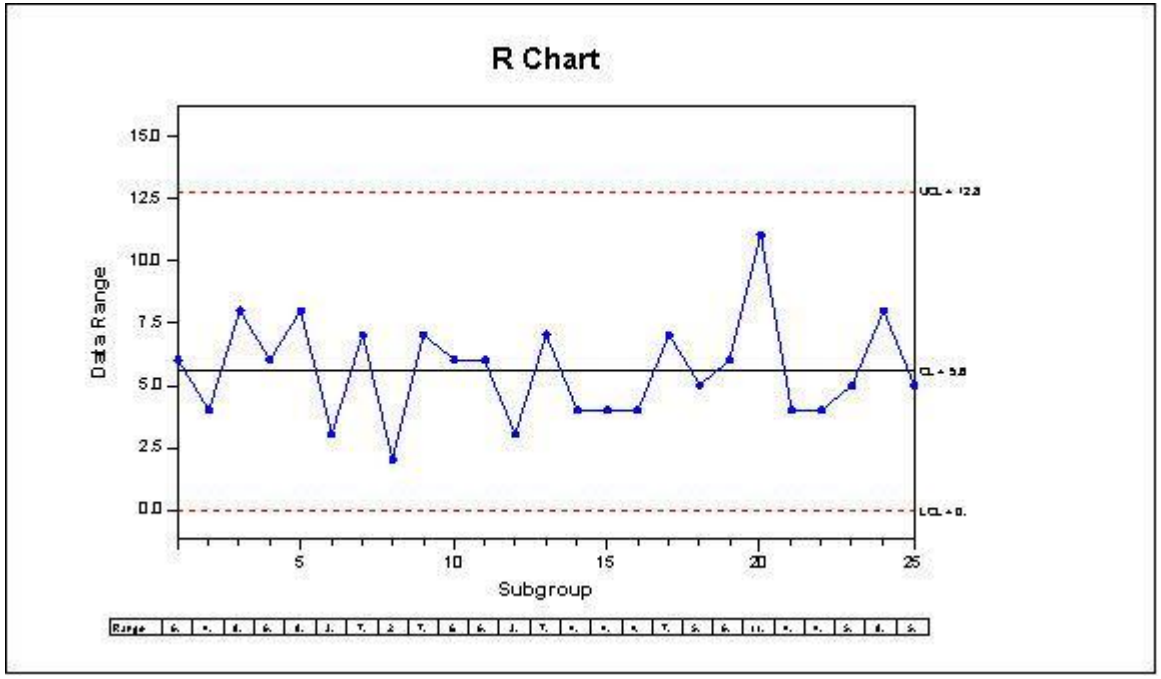
Advantages:
- Used in nearly every office
- Very easy
- Able to handle somewhat large datasets
- Decent graphic quality

Disadvantages:
- Limited chart types
- Not the best graphics

Data Visualization

CONNECTING RESEARCH TO THE WORLD

R:



Advantages:

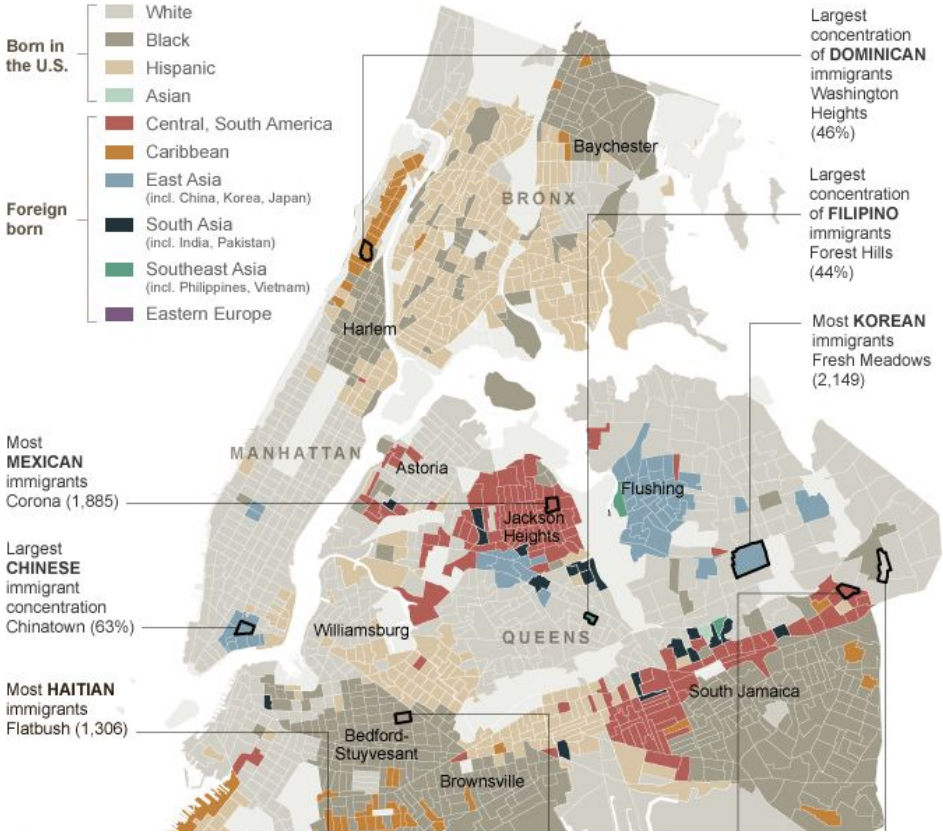
- Free! (legally)
- Able to handle large datasets
- Countless analytical functions

Disadvantages:

- Requires coding knowledge
- Ugly graphics

Data Visualization
CONNECTING RESEARCH TO THE WORLD

Illustrator / Inkscape:



Advantages:
- Amazing graphic quality
- Limitless flexibility for types of graphics

Disadvantages:
- No analysis functions
- Easier to make mistakes

Data Visualization
CONNECTING RESEARCH TO THE WORLD

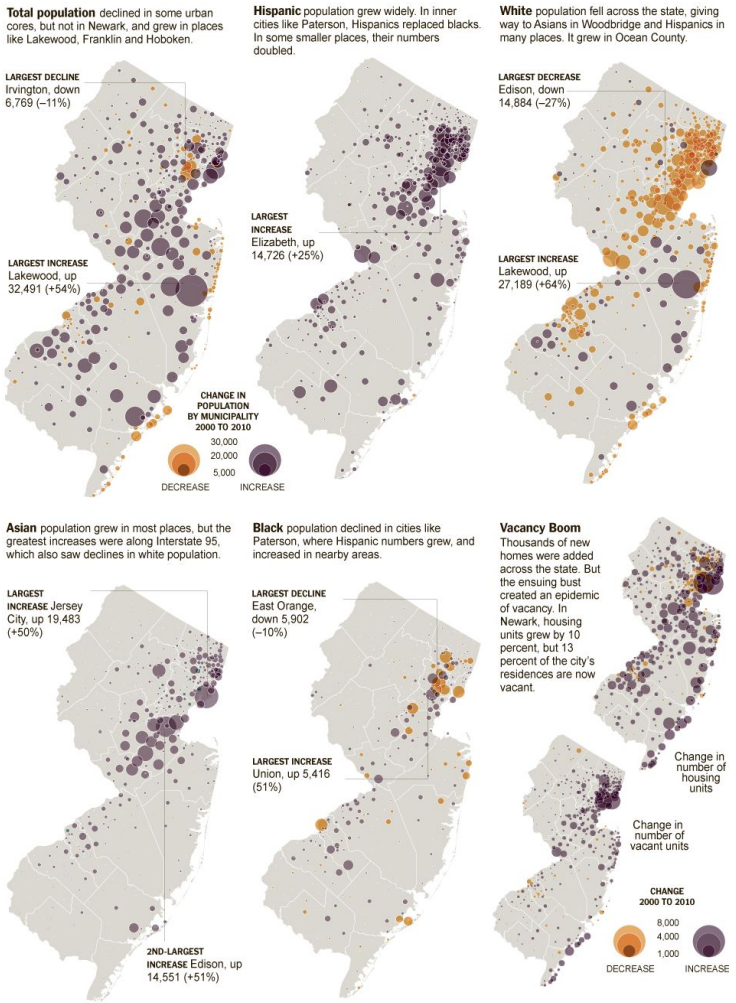
The best outcomes are a mix of programming and graphic design

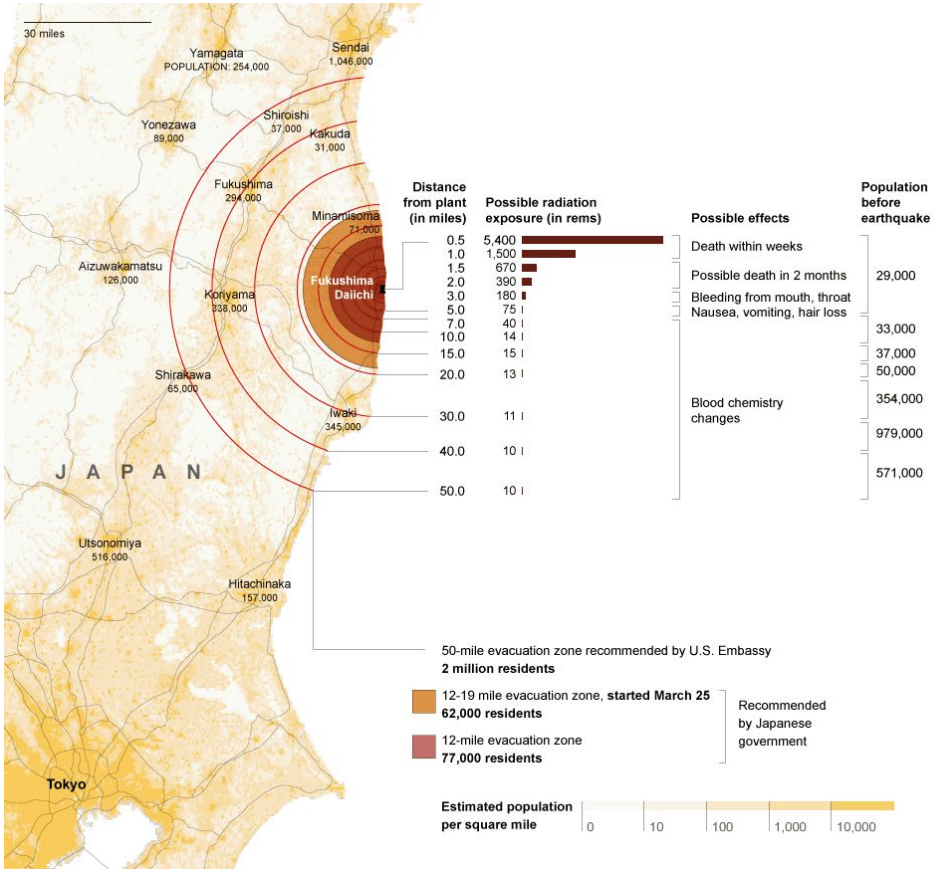
Programming Options:

- Java
- Javascript
- HTML
- Python
- PHP
- Flash
- HTML

Data Visualization
CONNECTING RESEARCH TO THE WORLD

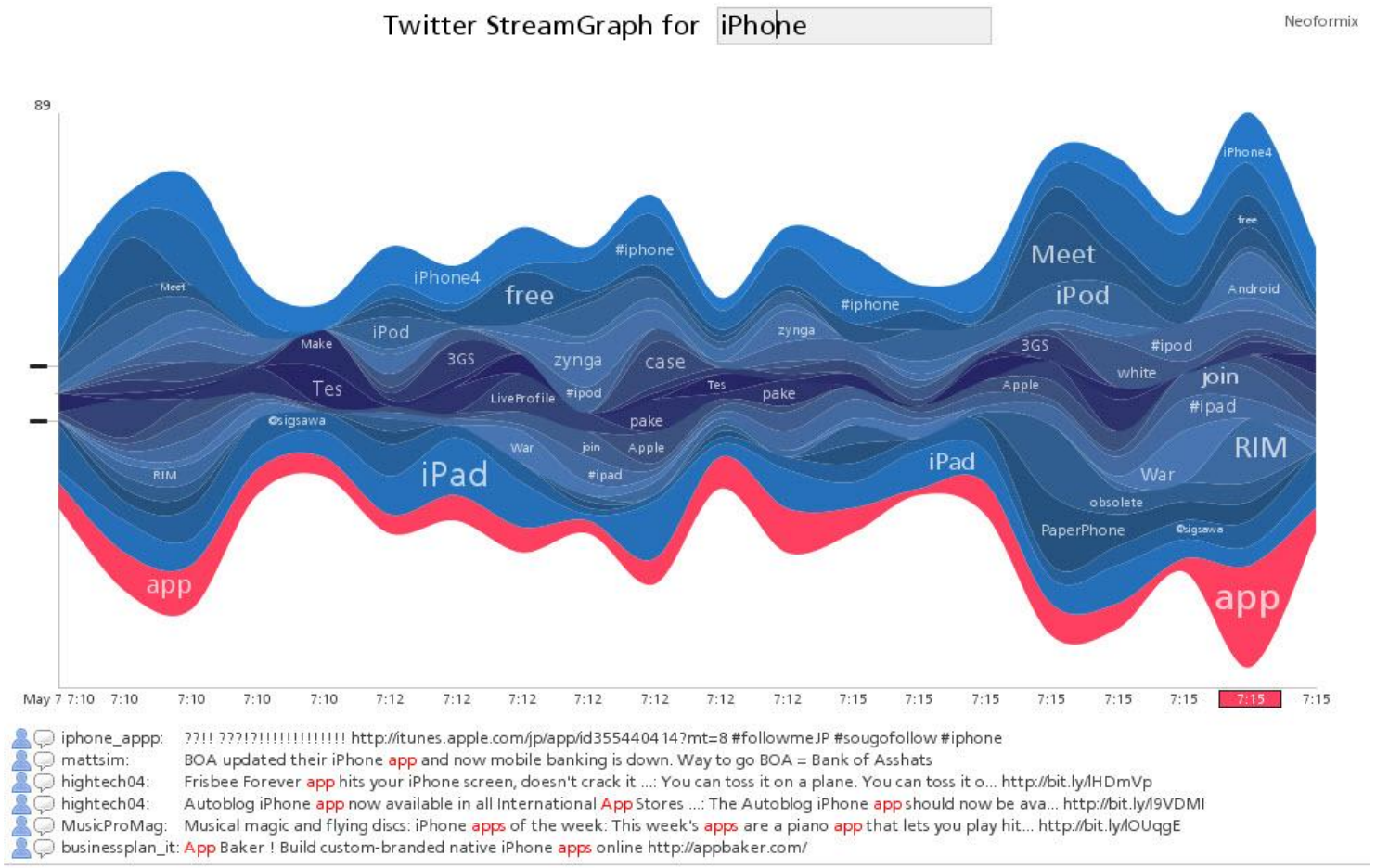
Static examples:





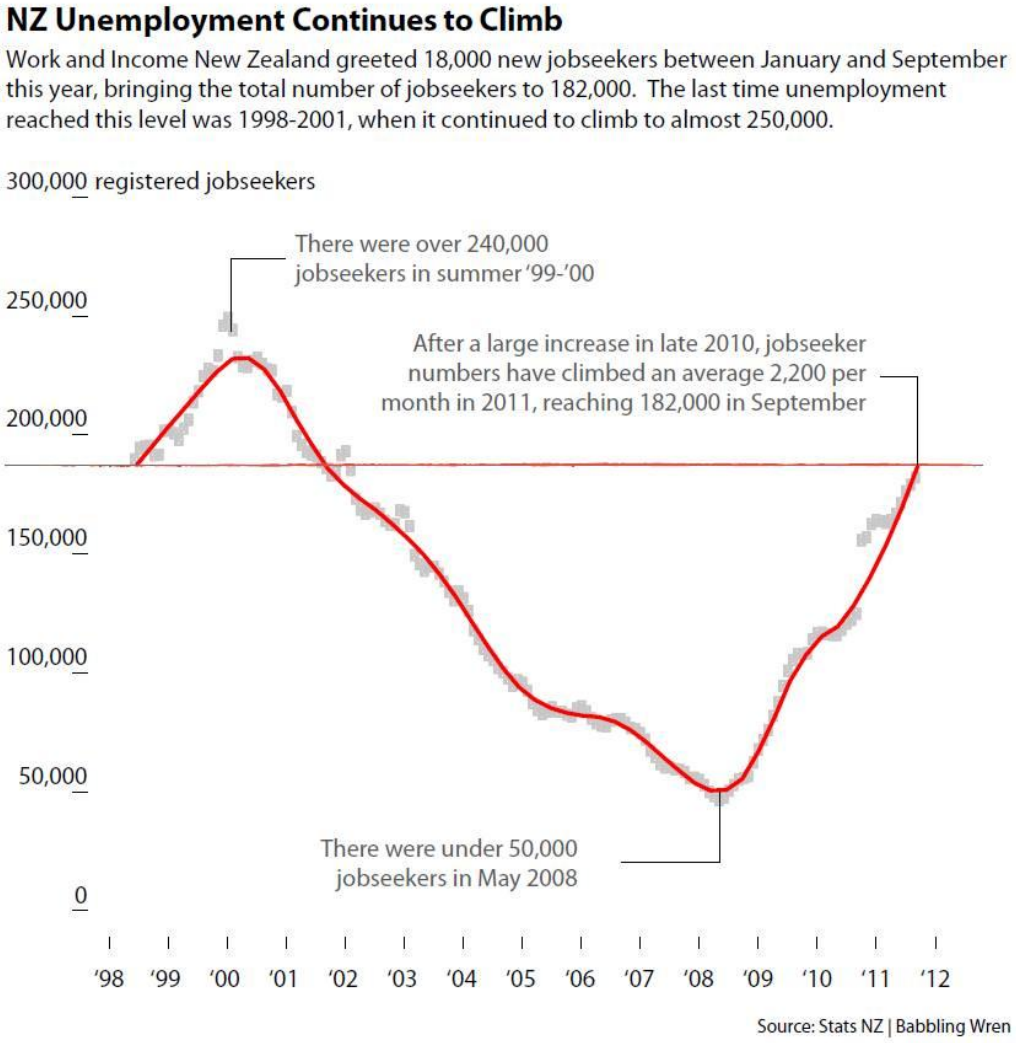
Data Visualization
CONNECTING RESEARCH TO THE WORLD

Static examples:



Data Visualization
CONNECTING RESEARCH TO THE WORLD

Static examples:



Data Visualization
CONNECTING RESEARCH TO THE WORLD

Dynamic examples: Tableau and Gapminder

http://www.tableausoftware.com/learn/gallery/high-school-reading-and-math-scores

http://www.gapminder.org/world/

http://www.gapminder.org/videos/200-years-that-changed-the-world-bbc/

Data Visualization

What story do you want to tell?