



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?

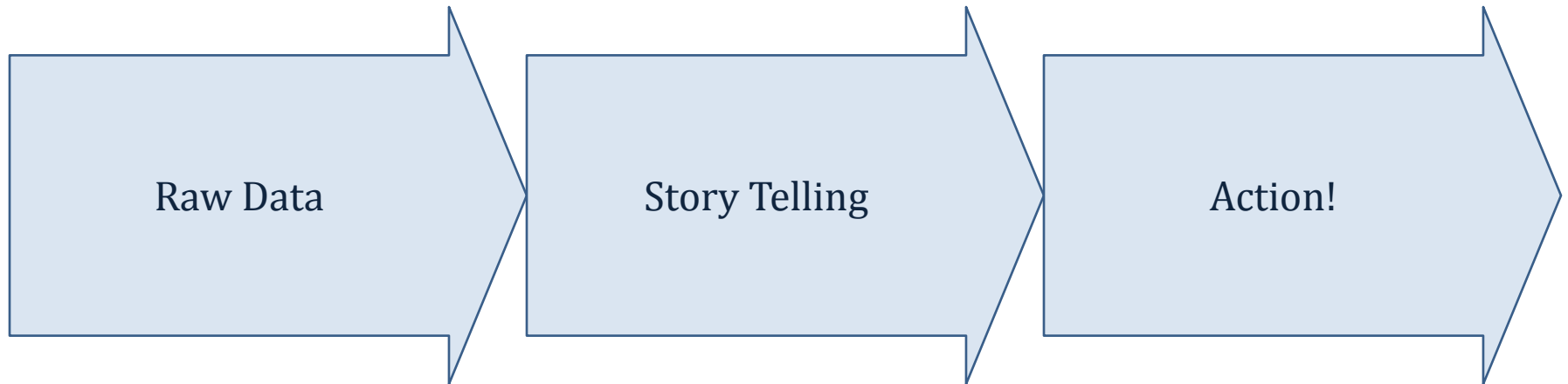
- Much of the research done is wasted
- People who are researched deserve to see the results
- The bigger the audience is, the more information you will draw from the research

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

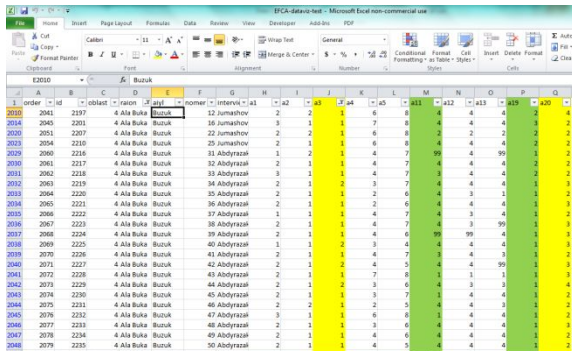


Data Visualization

CONNECTING RESEARCH TO THE WORLD

The art and science of data visualization

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



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	order	id	oblast	reason	pdf	nomer	interv	a2	a3	a4	a5	a11	a12	a13	a19	a20	a
2018	2041	2207	4	Alia Buka	Butuk	12	Jumathov	2	2	1	6	8	4	4	4	4	4
2019	2041	2207	4	Alia Buka	Butuk	16	Jumathov	3	1	1	7	8	4	4	4	4	4
2020	2051	2207	4	Alia Buka	Butuk	22	Jumathov	2	2	1	6	8	4	4	4	4	4
2023	2054	2210	4	Alia Buka	Butuk	29	Jumathov	2	1	1	6	8	4	4	4	4	4
2029	2060	2216	4	Alia Buka	Butuk	31	Abdyrasol	1	2	1	4	7	99	4	99	1	2
2030	2061	2217	4	Alia Buka	Butuk	32	Abdyrasol	2	1	1	4	7	4	4	4	4	2
2031	2062	2218	4	Alia Buka	Butuk	33	Abdyrasol	3	1	1	4	7	4	4	4	4	2
2032	2063	2219	4	Alia Buka	Butuk	34	Abdyrasol	2	1	2	3	7	4	4	4	4	2
2033	2064	2220	4	Alia Buka	Butuk	35	Abdyrasol	2	1	1	2	6	4	3	3	2	2
2034	2065	2221	4	Alia Buka	Butuk	36	Abdyrasol	2	1	1	2	6	4	4	4	4	2
2035	2066	2222	4	Alia Buka	Butuk	37	Abdyrasol	1	1	1	4	7	4	3	4	4	2
2036	2067	2223	4	Alia Buka	Butuk	38	Abdyrasol	2	1	1	4	7	4	3	99	4	2
2037	2068	2224	4	Alia Buka	Butuk	39	Abdyrasol	2	1	1	4	7	4	3	99	4	2
2038	2069	2225	4	Alia Buka	Butuk	40	Abdyrasol	1	1	2	3	6	4	4	4	4	2
2039	2070	2226	4	Alia Buka	Butuk	41	Abdyrasol	2	1	1	4	7	4	4	4	4	2
2040	2071	2227	4	Alia Buka	Butuk	42	Abdyrasol	2	1	2	4	3	4	4	4	4	2
2041	2072	2228	4	Alia Buka	Butuk	43	Abdyrasol	2	1	1	7	8	4	3	3	3	2
2042	2073	2229	4	Alia Buka	Butuk	44	Abdyrasol	2	1	1	3	6	4	3	3	3	2
2043	2074	2230	4	Alia Buka	Butuk	45	Abdyrasol	2	1	1	3	7	4	4	4	4	2
2044	2075	2231	4	Alia Buka	Butuk	46	Abdyrasol	2	1	2	5	4	4	4	4	4	2
2045	2076	2232	4	Alia Buka	Butuk	47	Abdyrasol	1	1	1	6	8	4	4	4	4	2
2046	2077	2233	4	Alia Buka	Butuk	48	Abdyrasol	2	1	1	3	6	4	4	4	4	2
2047	2078	2234	4	Alia Buka	Butuk	49	Abdyrasol	2	1	1	4	6	4	4	4	4	2
2048	2079	2235	4	Alia Buka	Butuk	50	Abdyrasol	2	1	1	4	3	4	4	4	4	2

Raw data



Computer programming



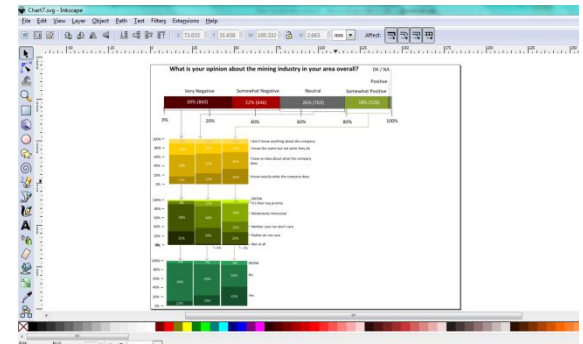
Graphic design

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

def webshow(img):
    savefig(img,dpi=500)
    print 'Content-Type: text/html\n'
    print ''

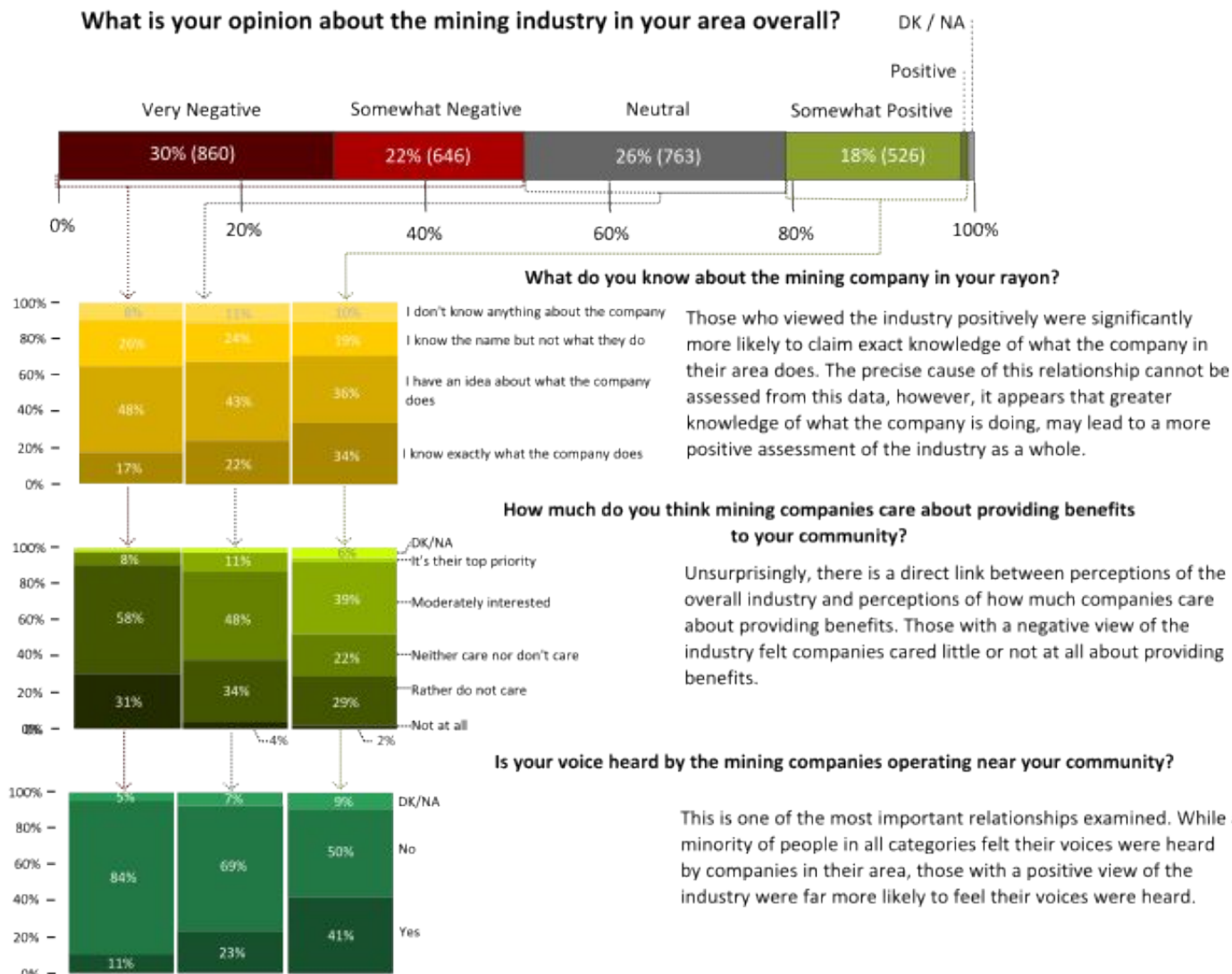
genres = []
n = 0
for c in sorted_list:
    genres.append(sorted_list[n][0])
    n += 1

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



Data Visualization

CONNECTING RESEARCH TO THE WORLD



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

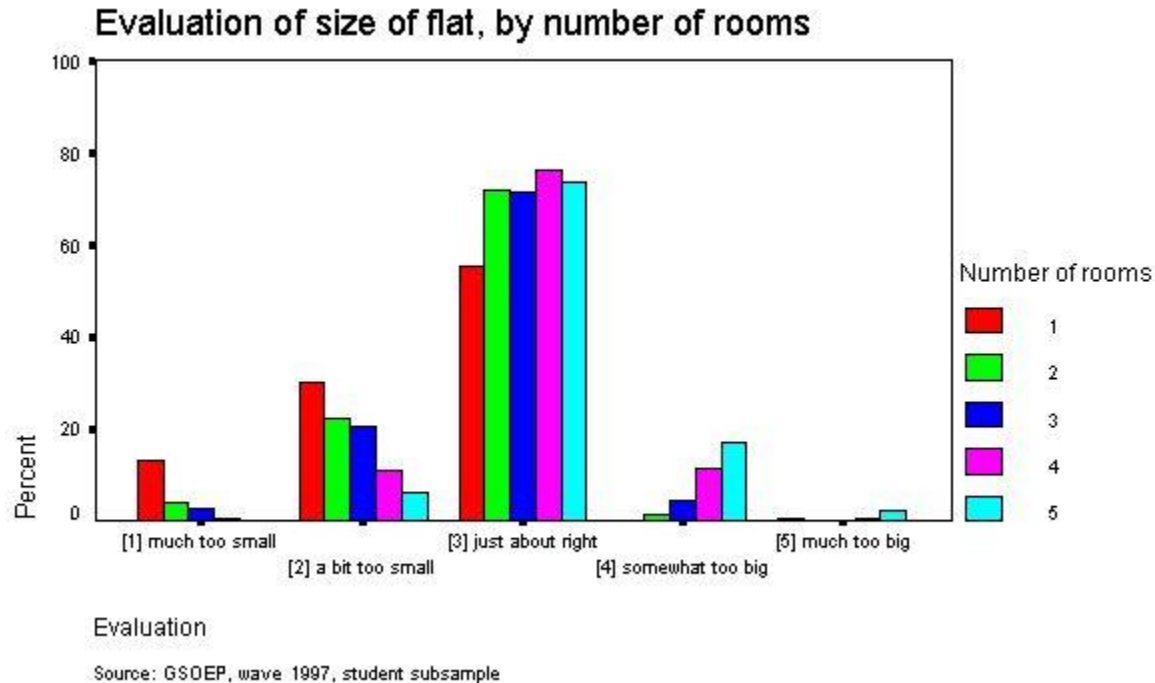
- Sociologist generally use SPSS
- Many other analysts use Excel
- Statisticians use “R”
- Graphic designers use Illustrator or Inkscape

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

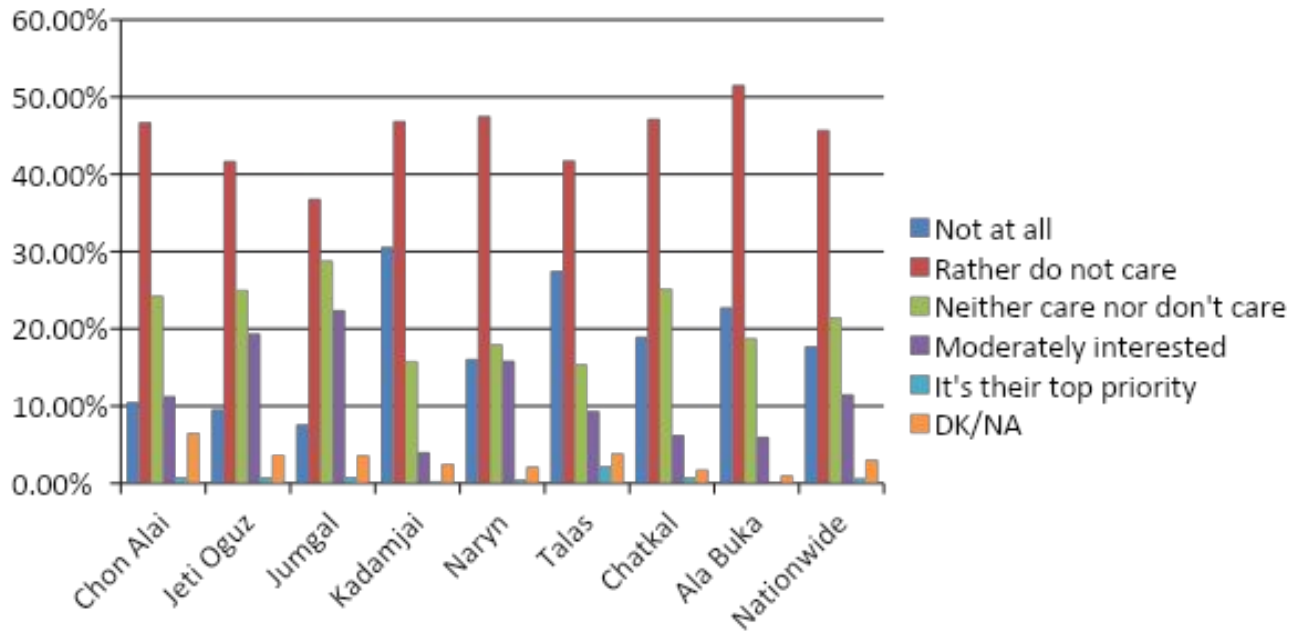
Disadvantages:

- 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

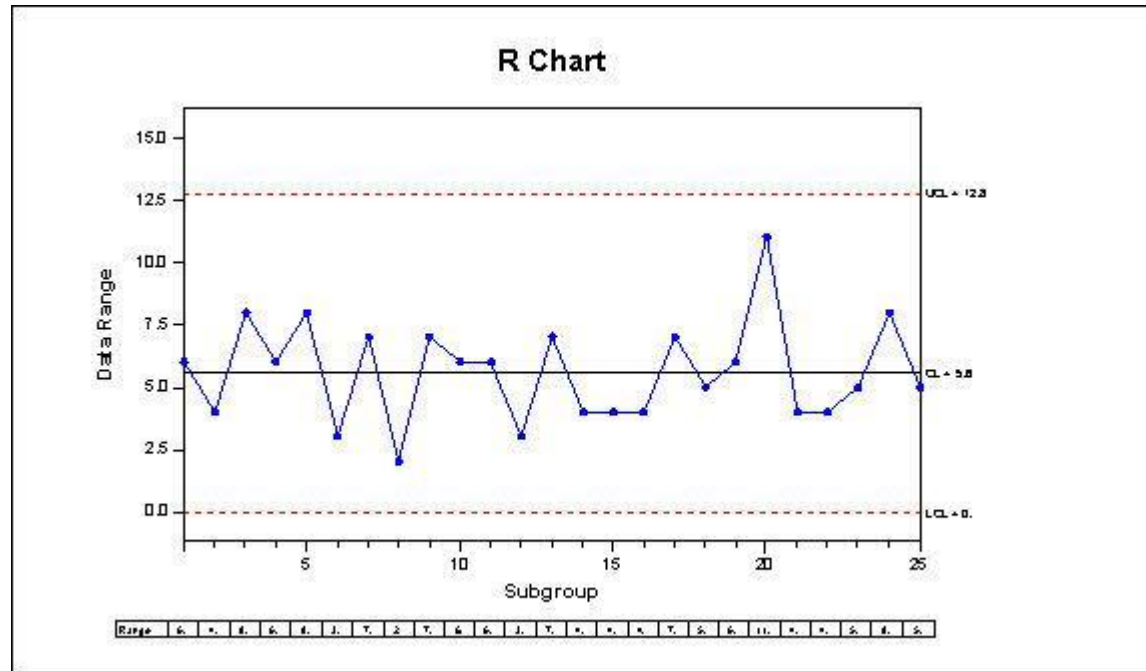
Disadvantages:

- 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

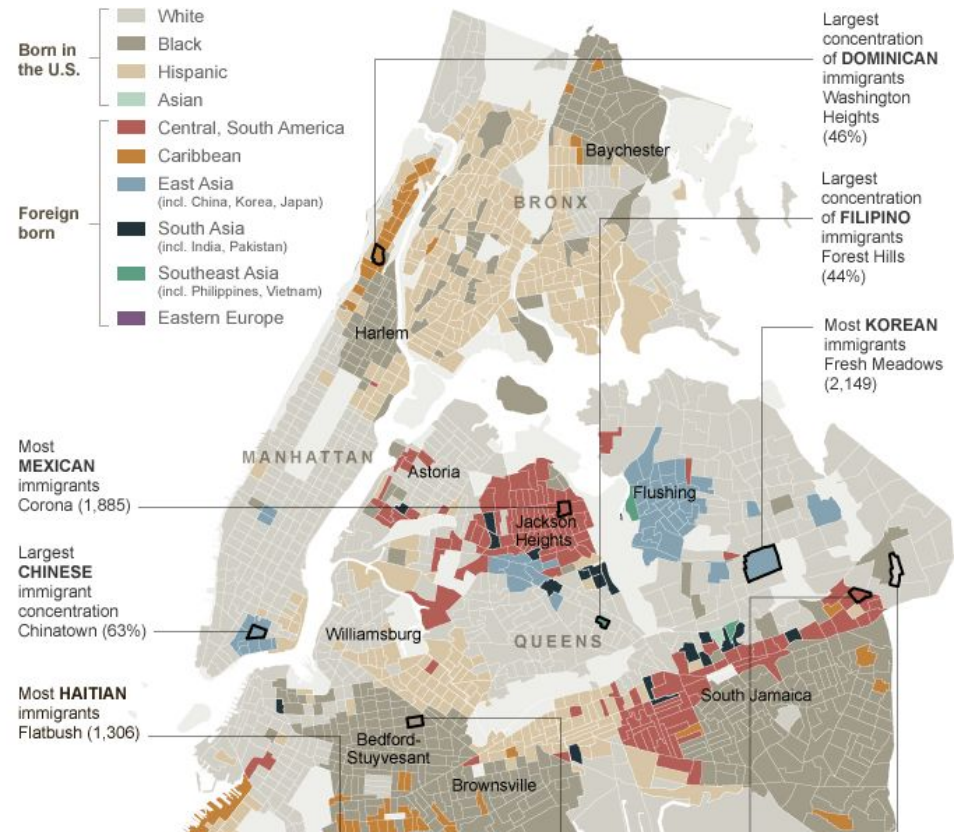
Disadvantages:

- 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:

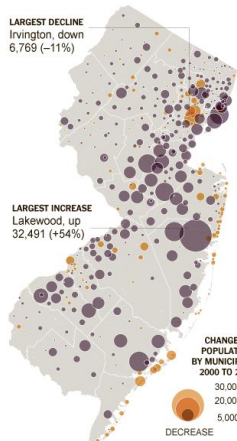
- Java
- Javascript
- HTML
- Python
- PHP
- Flash
- HTML

Data Visualization

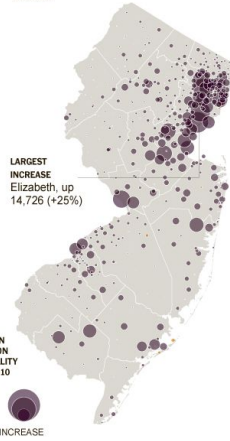
CONNECTING RESEARCH TO THE WORLD

Static examples:

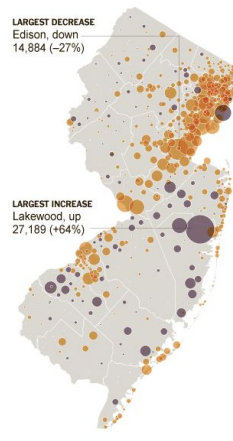
Total population declined in some urban cores, but not in Newark, and grew in places like Lakewood, Franklin and Hoboken.



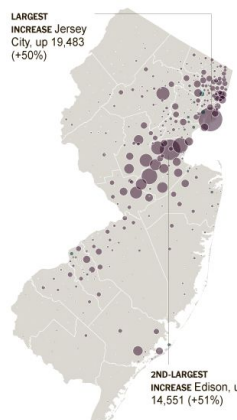
Hispanic population grew widely. In inner cities like Paterson, Hispanics replaced blacks. In some smaller places, their numbers doubled.



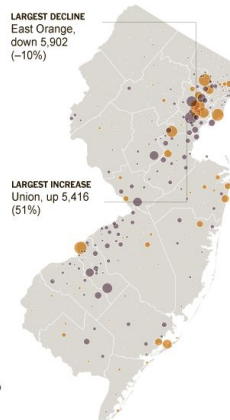
White population fell across the state, giving way to Asians in Woodbridge and Hispanics in many places. It grew in Ocean County.



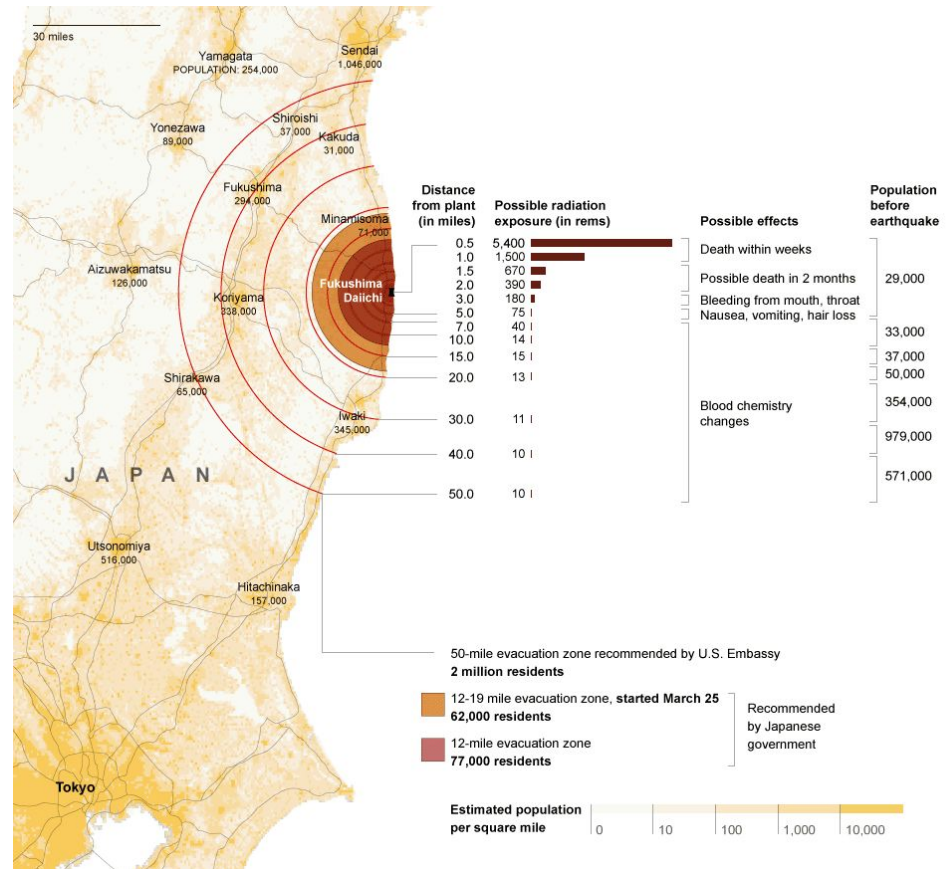
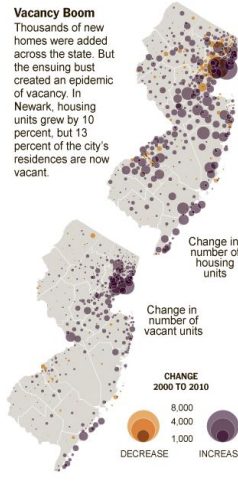
Asian population grew in most places, but the greatest increases were along Interstate 95, which also saw declines in white population.



Black population declined in cities like Paterson, where Hispanic numbers grew, and increased in nearby areas.



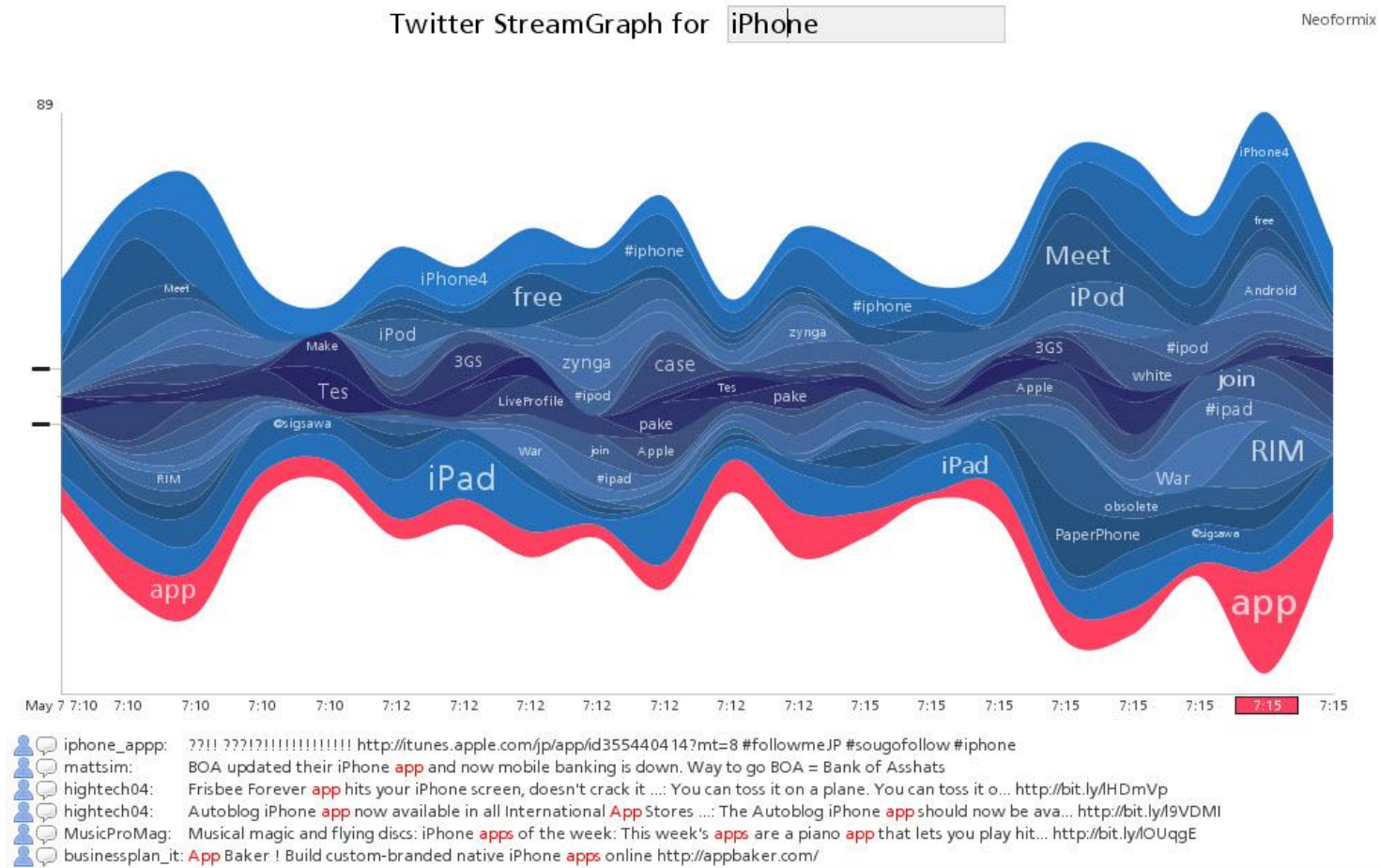
Vacancy Boom
Thousands of new homes were added across the state. But the ensuing bust created an epidemic of vacancy. In Newark, housing units grew by 10 percent, but 13 percent of the city's residences are now vacant.



Data Visualization

CONNECTING RESEARCH TO THE WORLD

Static examples:



Data Visualization

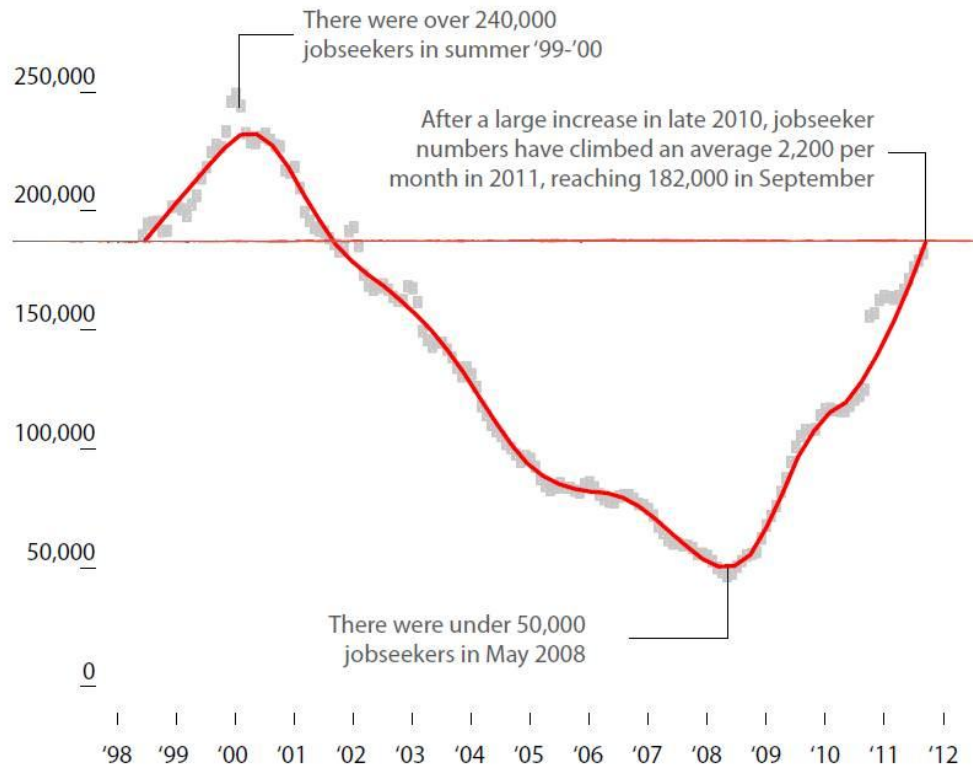
CONNECTING RESEARCH TO THE WORLD

Static examples:

NZ Unemployment Continues to Climb

Work and Income New Zealand greeted 18,000 new jobseekers between January and September this year, bringing the total number of jobseekers to 182,000. The last time unemployment reached this level was 1998-2001, when it continued to climb to almost 250,000.

300,000 registered jobseekers



Source: Stats NZ | Babbling Wren

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?