



Getting Started with Data Visualization

Geoff McGhee
 Tooling Up for Digital Humanities Seminar
 May 6, 2011

Dealing with Data



Explosion of Electronic Information

- 2003 estimate: 5 exabytes/day* new info
- Open government/transparency movements
- E-commerce, electronic record-keeping
- Digitization of media (photos, music, books...)
- Remote sensors, RFID tags, POS systems
- Plummeting Cost of Storage
- Data formats (XML, JSON, RDF...), APIs
- Social media

* *Exabyte = 1 million terabytes*

Dealing with Data

The Promise of Data Visualization



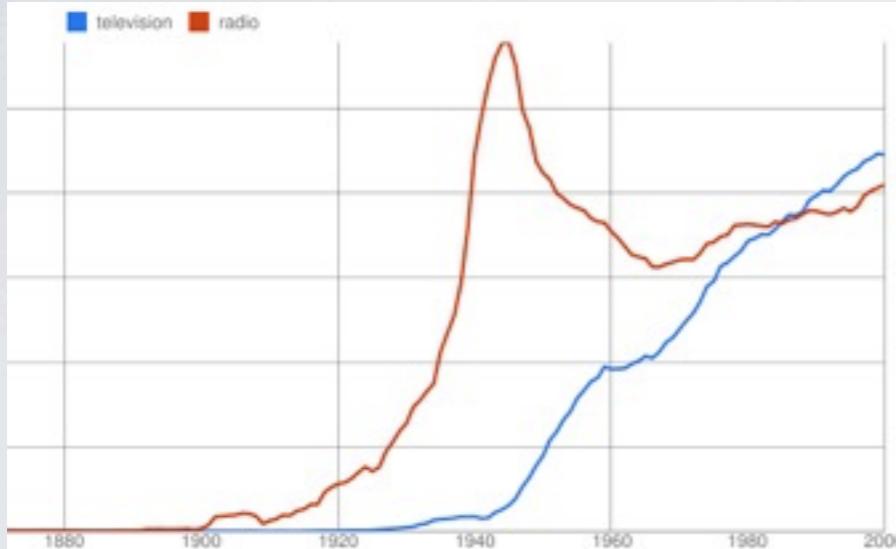
Map of New Brainland by Unit Seven via Flickr

Using the Eye-Brain Connection

- Bypass language centers, go direct to the visual cortex
 - Leverage ability to recognize patterns, visual sense-making
 - Powerful graphics chips enable animation, live data processing possible

Roles of Data Visualization

Roles of Data Visualization

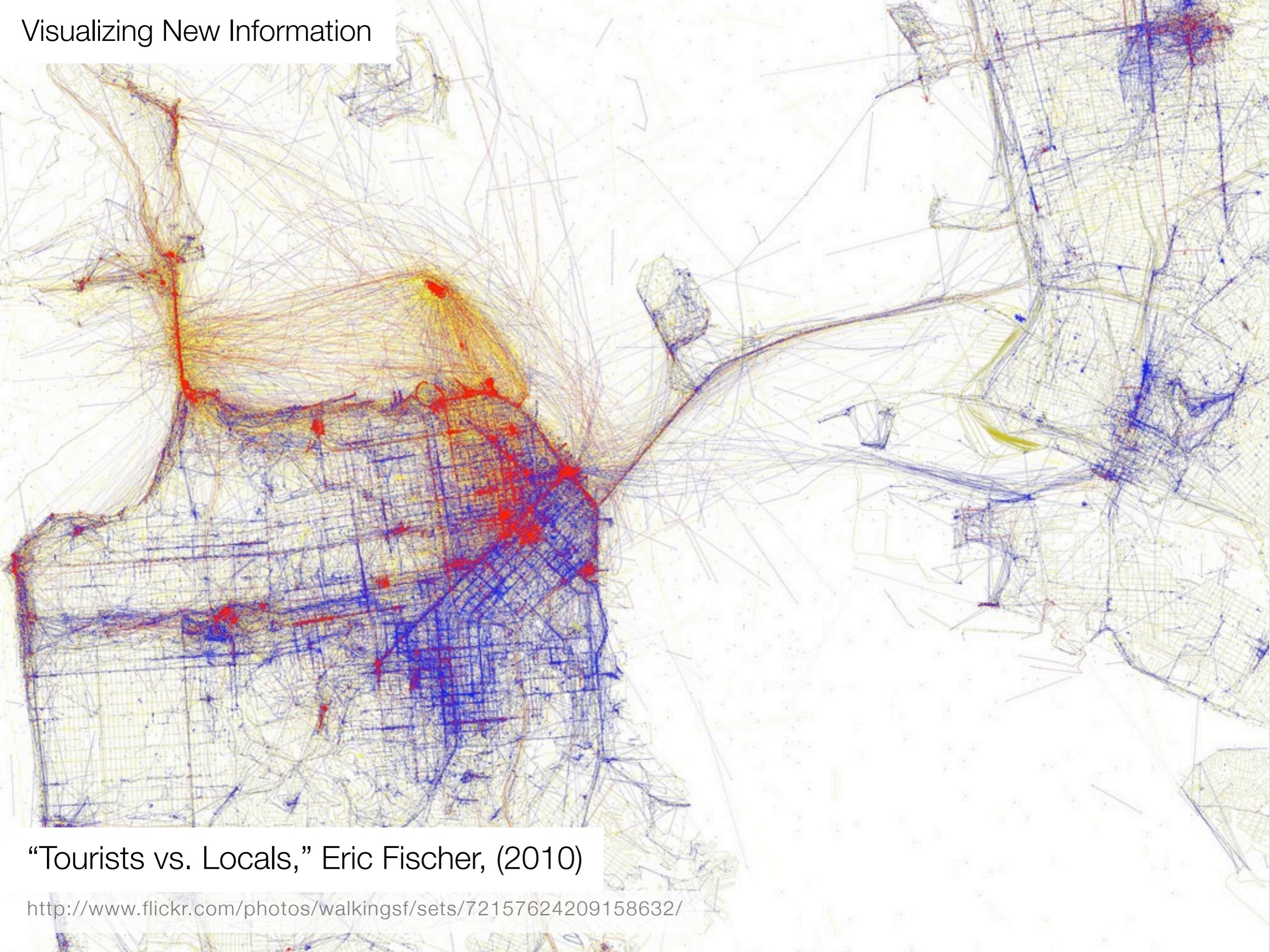


Google N-Gram Viewer

Making Sense of New Information

- Data that reveals previously unknown insights into patterns of life
- Visualization as a way to “throw things on the wall” and examine
- **Things that used to be unknown, unknowable, or impractical to know**
- Less about visualization than the **data**

Visualizing New Information



“Tourists vs. Locals,” Eric Fischer, (2010)

<http://www.flickr.com/photos/walkingsf/sets/72157624209158632/>

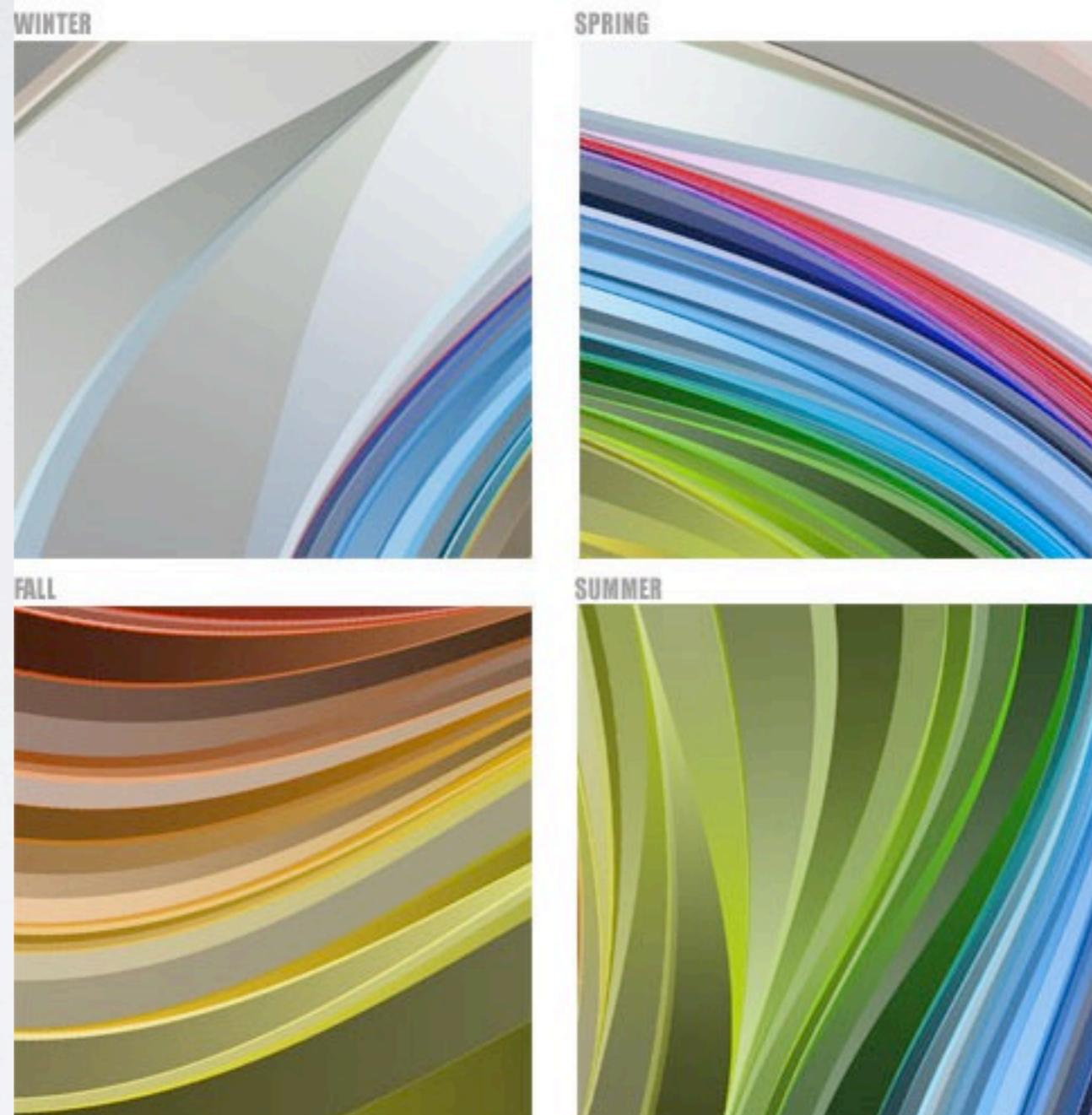
Visualizing New Information

 THE METRIC ...Taking a Measured View of the Hub

FLICKR OF HOPE It may not feel like springtime quite yet, but just beneath the lingering snowbanks the city is coming to life—and Flickr has the photos to prove it. To chart the ebb and flow of Boston's seasonal hues, Examiner asked Fernanda Viégas and Martin Wattenberg of the IBM Visual Communication Lab in Cambridge to analyze the more than 1,200 images of Boston Common found on the popular photo-sharing website (Flickr's informal tagging system means a few Public Garden shots may have sneaked in, too). Their software scanned the photos for the 216 different colors in the so-called Web-safe palette, and tallied how frequently each one appeared. Each stripe on the wheel below represents one color; the thickness of the stripe shows how often that color occurred in a given month. Just look at that green band starting to assert itself in March. You can practically feel the grass between your toes.



BOSTON MARCH 2009 55



“Flickr Flow,” Fernanda Viégas and Martin Wattenberg (2009)

<http://hint.fm/projects/flickr/>

Visualizing New Information

Gameday Audio
MLB.TV
Classic
Gameday

DET	5
CLE	5
Diamond - 8	

NYM	0
PHI	10
Diamond - 8	

TOR	VIEWING
LAA	7
TB	4
Diamond - 8	

FLA	7
CIN	4
Diamond - 9	

SEA	5
BOS	4
Diamond - 8	

STL	2
ATL	3
Diamond - 8	

MIL	4
HOU	0
Diamond - 6	

BAL	3
CWS	3
Diamond - 6	

MIN	3
KC	1
Diamond - 6	

PIT	1
COL	0
Diamond - 5	

CHC	0
ARI	0
Diamond - 1	

TEX	0
OAK	0
Diamond - 1	

SD	0
LAD	0
Diamond - 1	

SF	0
WSH	3
FINAL	

Friday April 29, 2011 | Previous | Today | Next |

1	2	3	4	5	6	7	8	9	R	H	E
Y	0	0	2	1	0	2	0		5	8	1
Y	0	1	0	1	0	0			2	5	1

Previous batter: Brett Gardner walks.
None out.

Pitching

Casey Janssen - #44 RHP
0-0, 0.96 ERA
11 Pitches - 4 Strikes, 7 Balls

Batting

Derek Jeter - #2 SS
.264 AVG, 0 HR, 5 RBI
1-for-3: Flyout (2), Single

Result	MPH	Pitch	NF
1 Called Strike	92	4-Seam Fastball	47
2 Ball	78	Curveball	15
3 Swinging Strike	90	4-Seam Fastball	45
4 Ball In Dirt	76	Curveball	30
5 Ball	90	4-Seam Fastball	60

Preferences

Loves to face: Sinker
Season: .000 0 0

Hates to face: Cutter
Last 5 Years: .000 0 0

Hot / Cold Zones

ZONE	AB	Avg	HR	RBI	BB	SO	OPS
YTD	87	.264	0	5	9	7	.617

STATS FILTERS

None | Pitch Type | Count | Outs
Runners | L/R

vs. Janssen

Hot / Cold Zones

3	7	4	5
7	4	5	3
3	11	8	9
1	4	15	1

Swisher On Deck

Teixeira In Hole

6

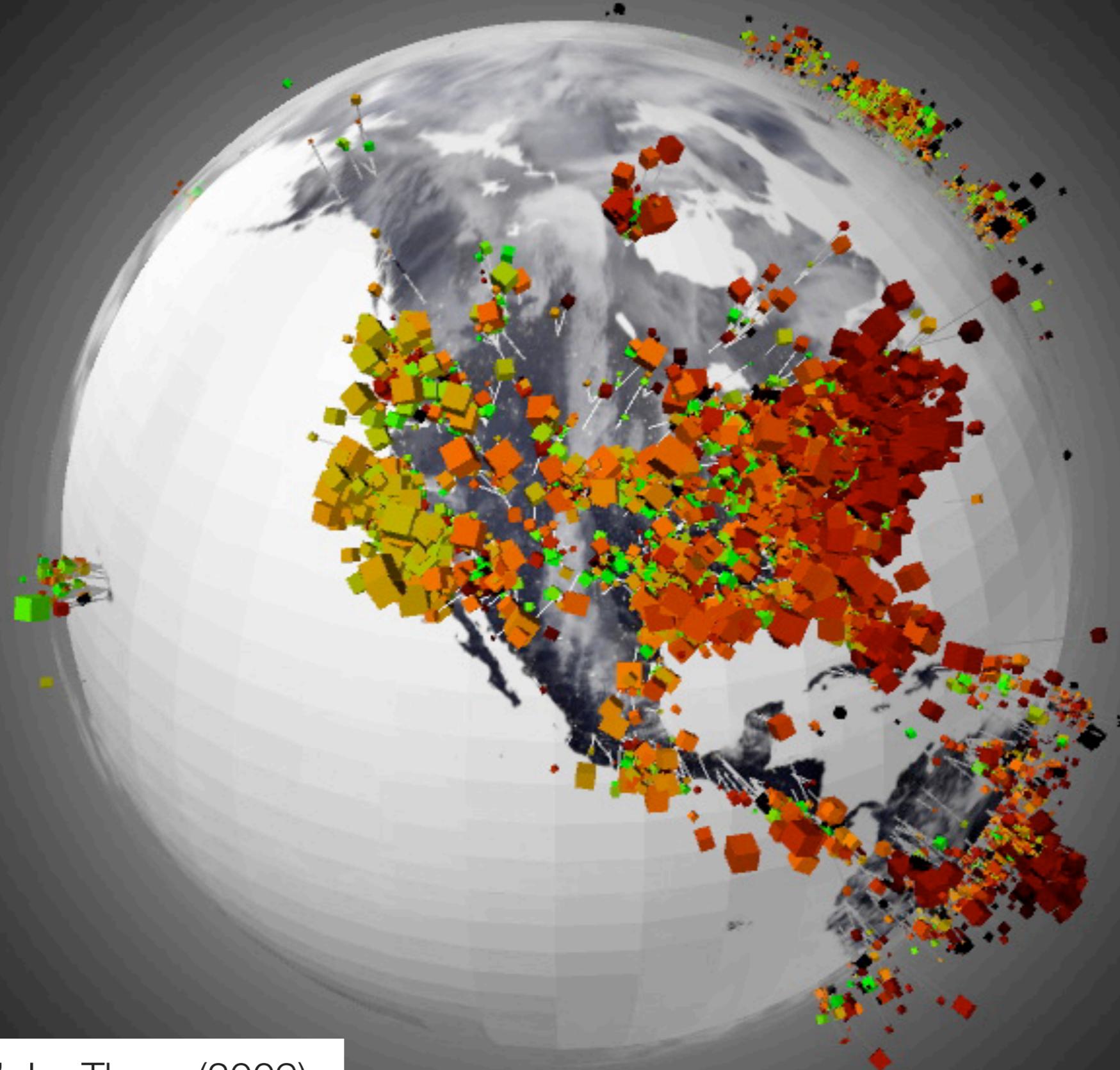
MIN @ KC Watch live

See Also: [Game Preview](#) | [Gameday View](#) | [Gameday Mini](#)

<http://mlb.mlb.com>

“GameDay,” Major League Baseball (2011)

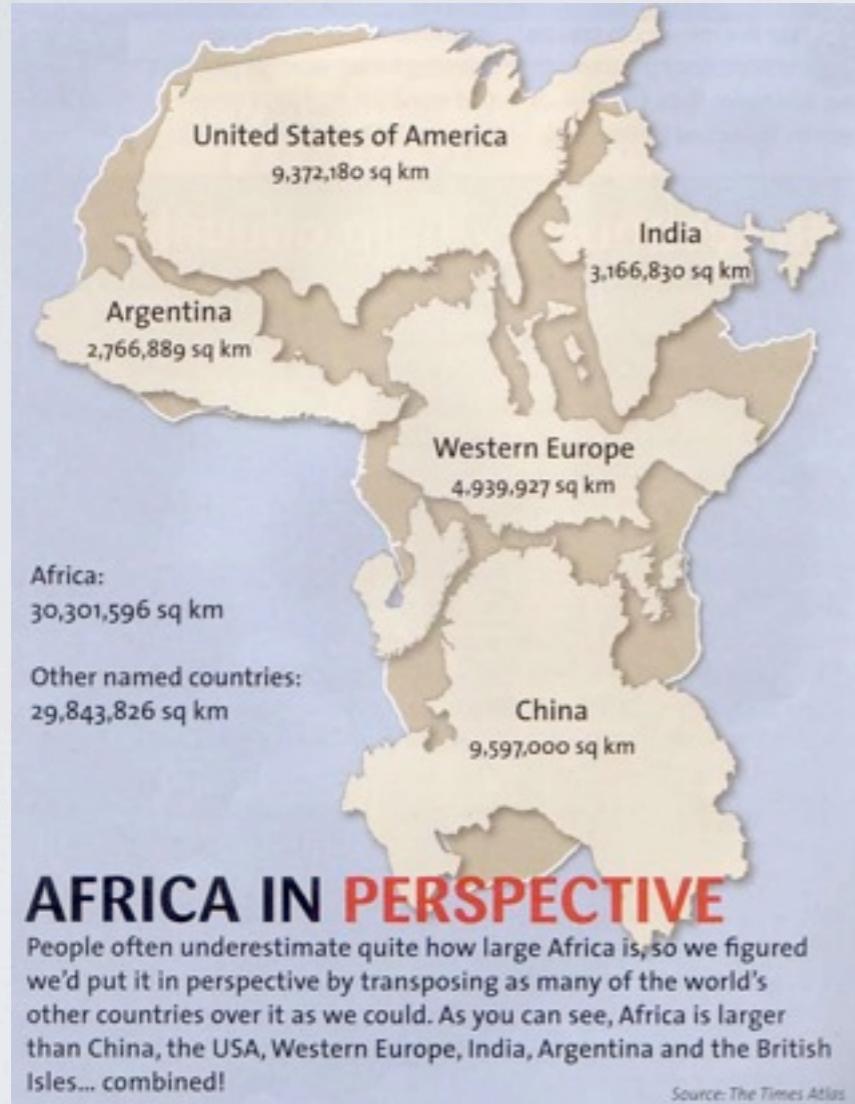
Visualizing New Information



“Good Morning,” Jer Thorp (2009)

<http://blog.blprnt.com/blog/blprnt/goodmorning>

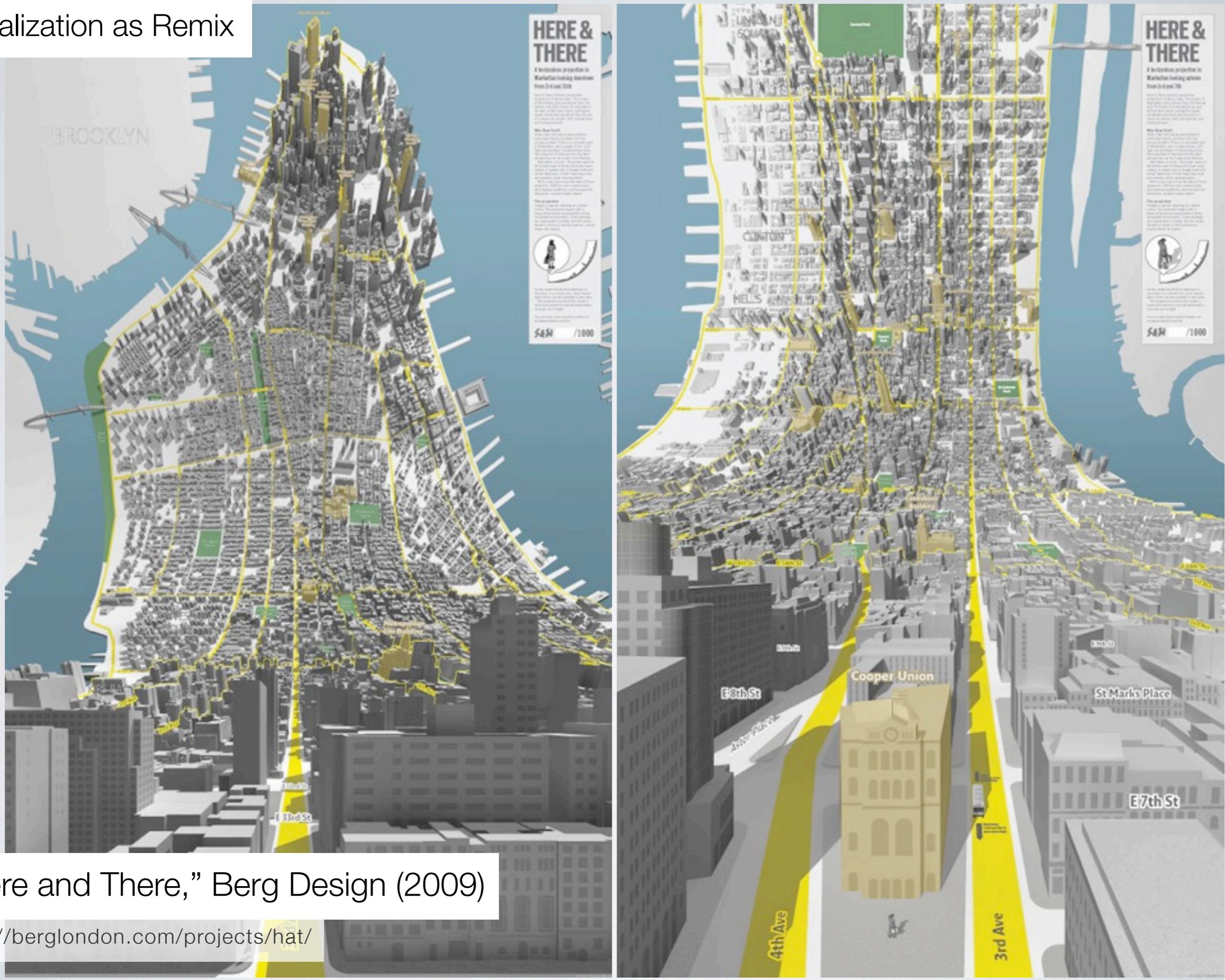
Roles of Data Visualization



Remix: The Familiar Through a New Lens

- Innovations in graphic display can change how we experience an idea
- Less about data than the **visualization**
- “Now I see it”

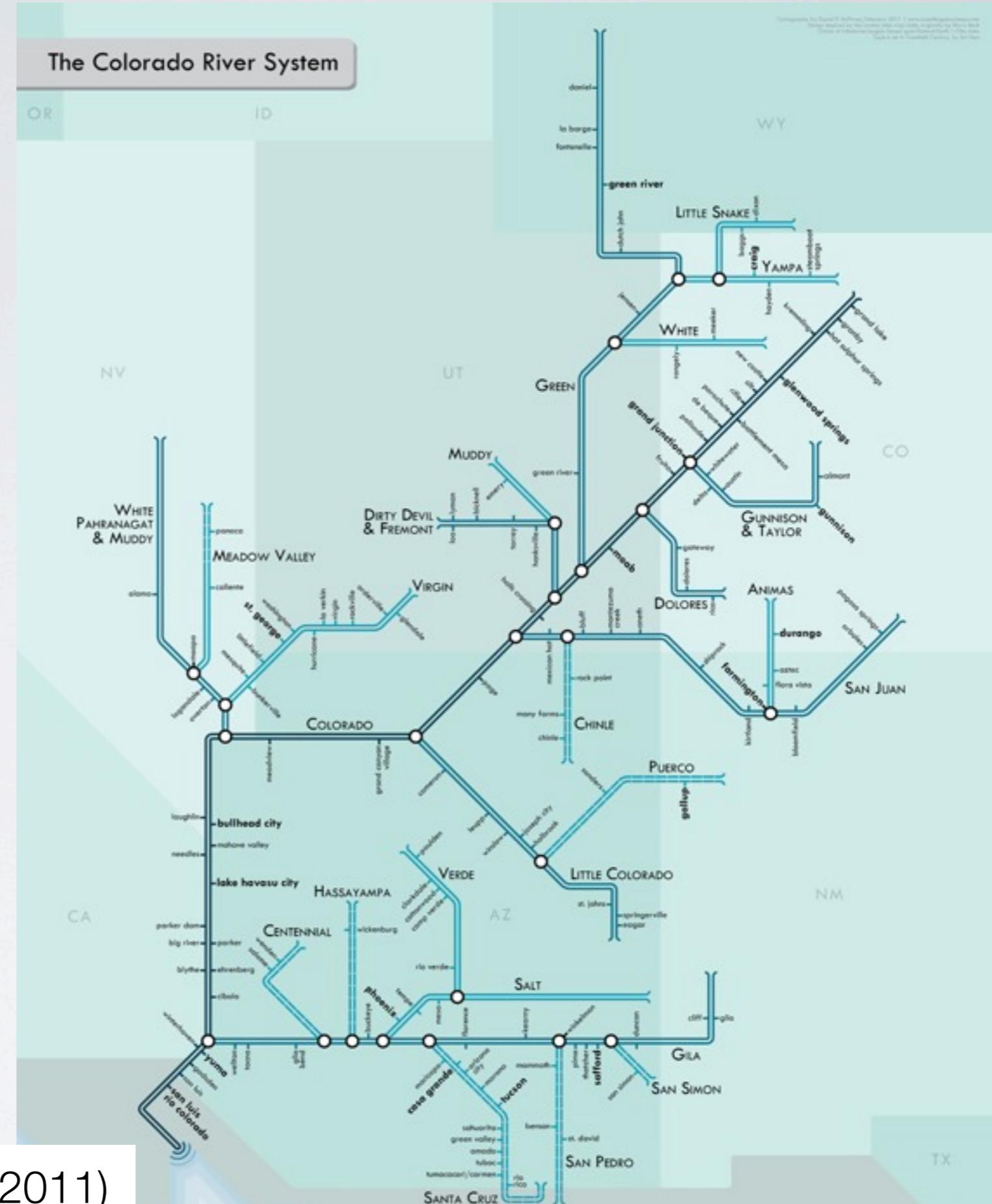
Visualization as Remix



"Here and There," Berg Design (2009)

<http://berglondon.com/projects/hat/>

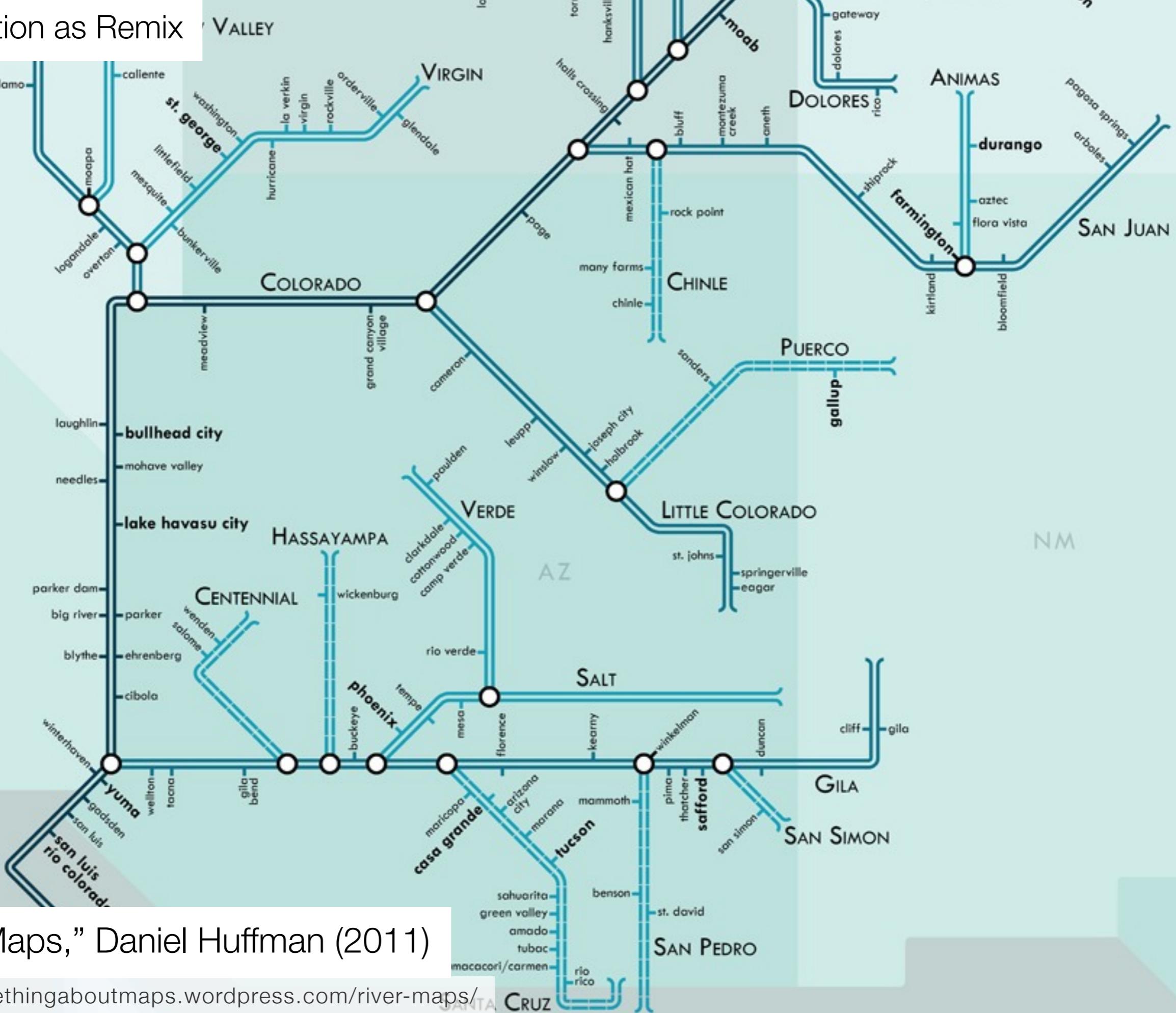
Visualization as Remix



“River Maps,” Daniel Huffman (2011)

<http://somethingaboutmaps.wordpress.com/river-maps/>

Visualization as Remix



“River Maps,” Daniel Huffman (2011)

<http://somethingaboutmaps.wordpress.com/river-maps/>

Published: November 6, 2009

The Jobless Rate for People Like You

Not all groups have felt the recession equally.



UNEMPLOYMENT RATE,
12 MONTH AVG. ENDING SEPT. '09

3.6%

For white women ages 25 to 44 with
a college degree

All men and women

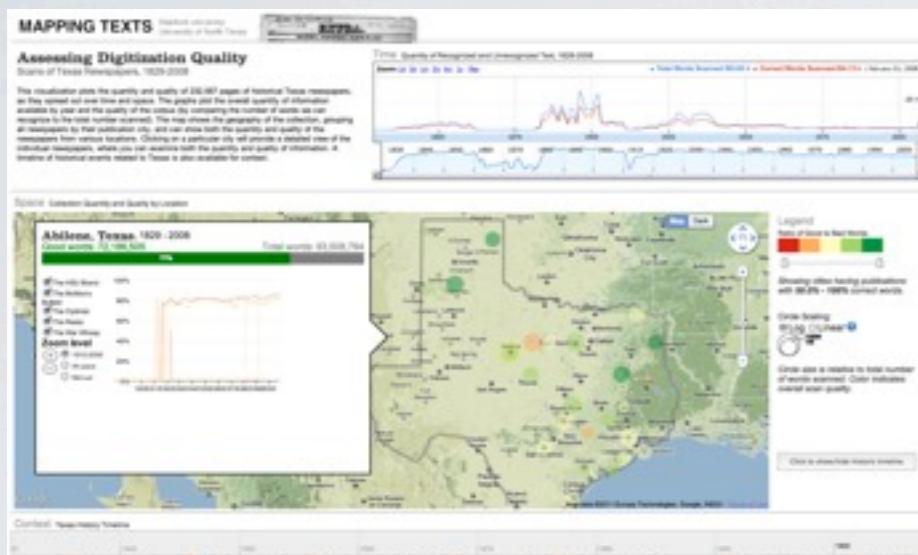
The New York Times (2009)

Jan. 2008

White women ages 25 to 44 with a college
degree
Click to select

Sept. 2009

Roles of Data Visualization



Analyzing OCR Quality of Newspapers

Environment for Exploration

- Tool for individual or collective exploration
- Can show same data in multiple dimensions, like time/space
- Search, filter, drill down to details
- Ideally, mark and share discoveries **within the tool**

Environment for Exploration

Community Survey, based on samples from 2005 to 2009. Because these figures are based on
largely in places with a low population, and are best regarded as estimates.

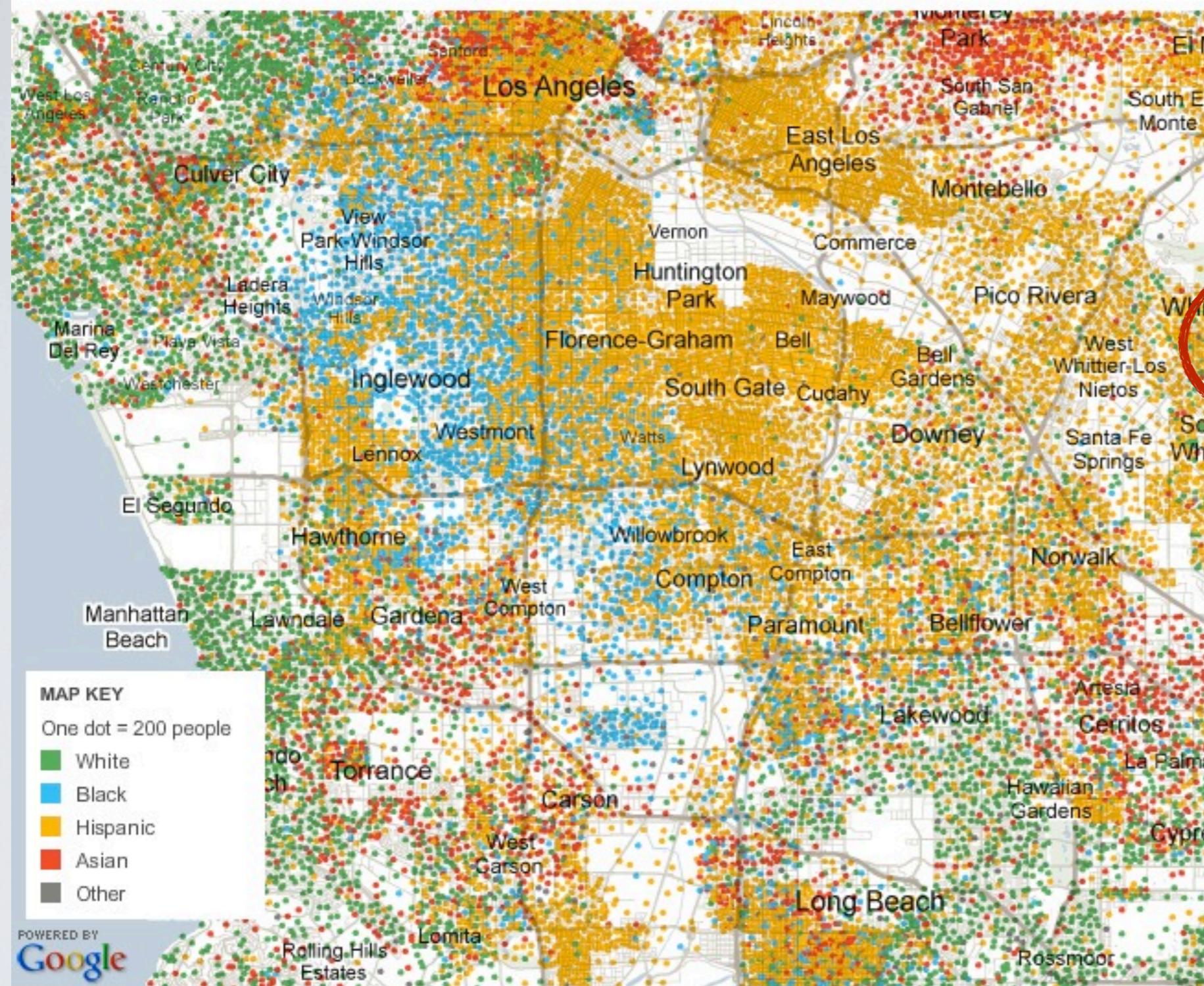
[View Readers Maps \(49\)](#)

Distribution of racial and ethnic groups

[View More Maps](#)

Milwaukee, WI

Go



By MATTHEW BLOCH, SHAN CARTER and ALAN McLEAN | Source: 2005-9 American Community Survey, Census Bureau; [socialexplorer.com](#)

Note: Dots are evenly distributed across each Census tract
and scaled for inflation.

Readers' Maps

A selection of maps shared by
Times readers on Twitter

[My Boise Census map shows rich vs.
poor neighborhoods...](#)
<http://t.co/2nD444a> #nytcensus

© Popkin, Dec. 17 1:10 PM ET

The almost complete Balkanisation of
places like LA and particularly NY is
astonishing. #nytcensus
<http://t.co/bcTeJ9K>

@hughsheehy, Dec. 21 6:16 AM ET

USC's big asian and white student
population makes the campus look
like christmas <http://t.co/7379YjE>
#nytcensus <http://twitpic.com/3hw5he>

@peard33, Dec. 20 6:47 PM ET

Census maps emerging #nytcensus
<http://t.co/9UZpCbn> Durango betwn
30-40% homeowners spend more
than 30% income on mortgage-
unsustainable

@heathererb2, Dec. 18 5:52 PM ET

Amazing infographic to explore! My
census map shows income of
Columbus neighborhoods. Explains
mall locations #nytcensus
<http://t.co/2SvmzQj>

© Popkin, Dec. 19 7:04 PM ET

"Mapping America: Every City, Every Block," The New York Times (2010)

<http://projects.nytimes.com/census/2010/explorer>

Environment for Exploration

MAPPING TEXTS

Stanford University
University of North Texas



Assessing Digitization Quality

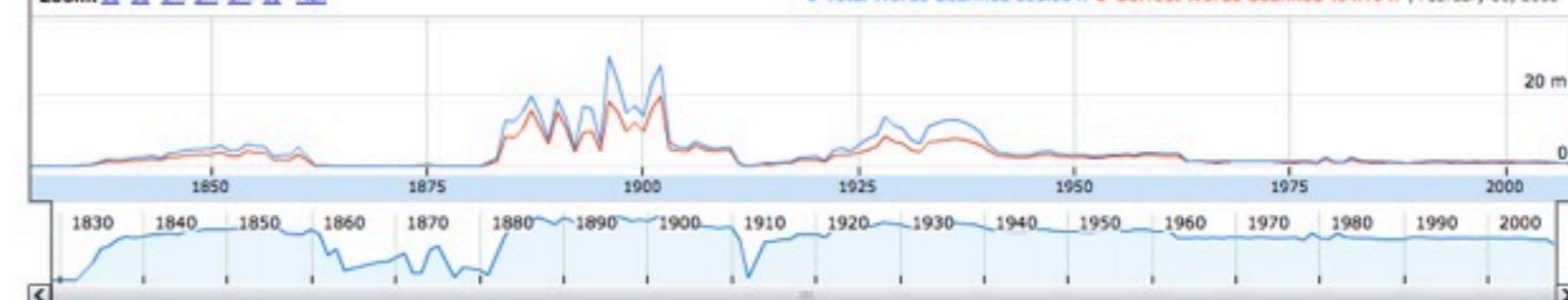
Scans of Texas Newspapers, 1829-2008

This visualization plots the quantity and quality of 232,567 pages of historical Texas newspapers, as they spread out over time and space. The graphs plot the overall quantity of information available by year and the quality of the corpus (by comparing the number of words we can recognize to the total number scanned). The map shows the geography of the collection, grouping all newspapers by their publication city, and can show both the quantity and quality of the newspapers from various locations. Clicking on a particular city will provide a detailed view of the individual newspapers, where you can examine both the quantity and quality of information. A timeline of historical events related to Texas is also available for context.

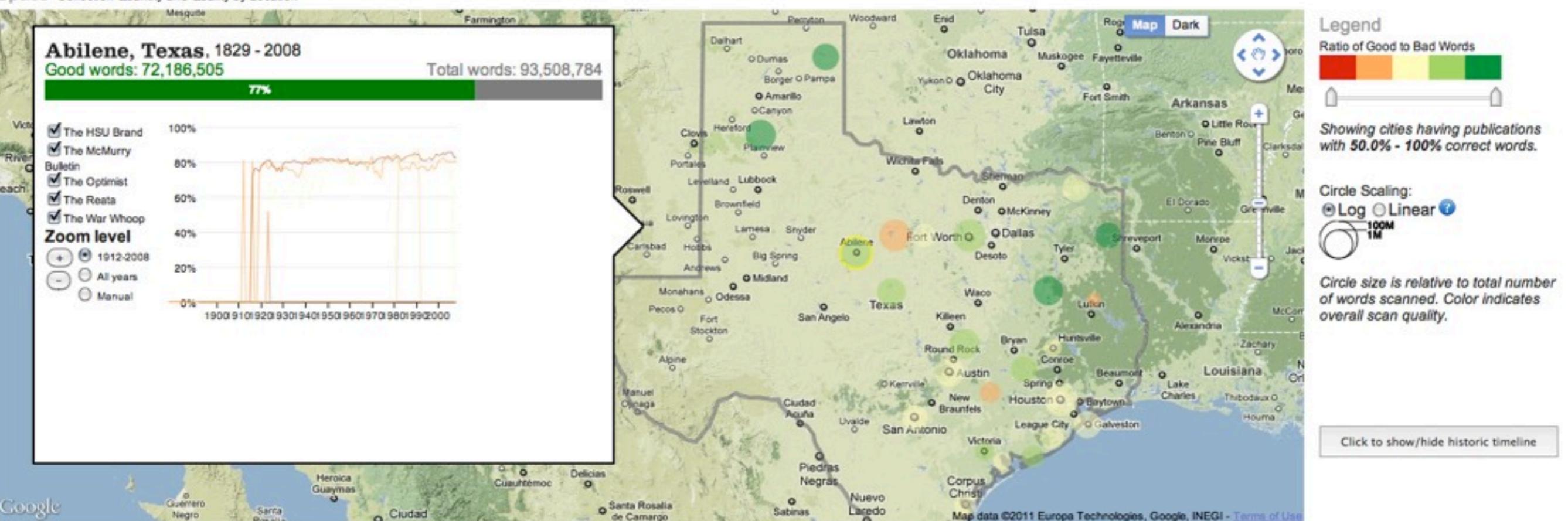
Time Quantity of Recognized and Unrecognized Text, 1828-2008

Zoom: 1d 5d 1m 3m 6m 1y Max

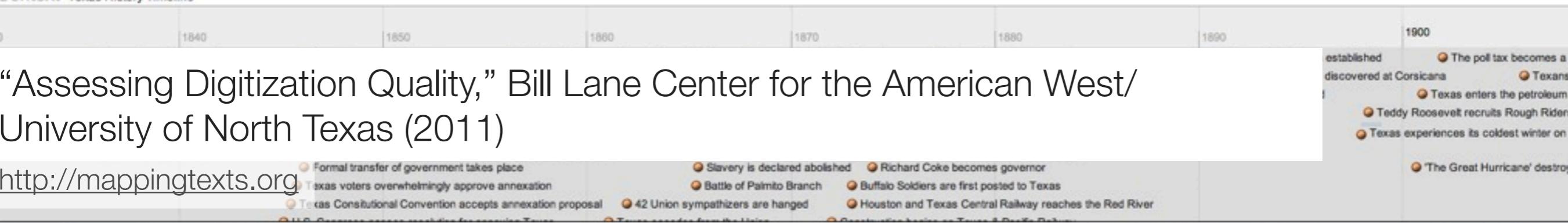
Total Words Scanned 563.85 k • Correct Words Scanned 454.13 k | February 01, 2008



Space Collection Quantity and Quality by Location



Context Texas History Timeline



"Assessing Digitization Quality," Bill Lane Center for the American West/
University of North Texas (2011)

<http://mappingtexts.org>

Life Cycle of Visualizations

Gathering Data

Discovery/
Acquisition

Analyzing It

Cleaning/
“Munging”

Analysis/Exploratory
Visualization

Sharing Findings

Publication



Life Cycle of Visualizations

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research

Spreadsheets
Databases
Digitized Media

Other Downloads

Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

[Junar](#)
[Outwit Hub](#)
[ScraperWiki](#)

Life Cycle of Visualizations

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research

Spreadsheets
Databases
Digitized Media

Other Downloads

Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

[Junar](#)

[Outwit Hub](#)

[ScraperWiki](#)



Search for data streams. Try 'GDP', 'soccer', 'Apple'

Search

Examples: earthquakes, apple stock, GDP+US, "NBA Standings".

Welcome to Junar!

[Sign out](#)



Hi mcgeoff!

[Edit your account details here.](#)

What now?

- › [Go to your Dashboard](#)
- › [Create a Data Stream](#)
- › [Lost? Help & Tips here!](#)

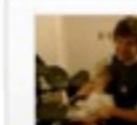
What can I do with Junar?

Junar will save you time searching for data. Gain easy access and get ready to track latest values of data around the web!

Top Data Streams

We have over 11248 Data Streams in our catalog. [Create your own](#) and contribute!

[Explore our catalog](#)



Peru Elections 2011 | Candidates

Peru elections 2011 candidates

WORLD | by birdman 3 weeks, 6 days ago

Nominee	Ollanta Humala	Keiko Fujimori	Pedro Pablo Kuczynski
Party	Gana Perú	Fuerza 2011	Alianza por el Gran Cambio
Popular vote	4,466,275[1]	3,307,369[1]	2,605,514[2]
Percentage	31.73%	23.49%	18.51%



Most Influential

This entry shows a random sample.

SCIENCE | by melissa

1	Isaac Newton	the Newtonian Revolution
2	Albert Einstein	Twentieth Century Physics
3	Kurt Gödel	The Incompleteness Theorem

Top Dashboards

[See our Dashboard demo](#)

1 Peru Elections 2011

Peru elections 2011 facts and figures
by jonmurmur

Discover Data on

Exchange Rates total area met

USA People Country

LiteracyRate costa rica 2009

Weekday Service (Monday - Friday)

- Does not operate on [university holidays](#)
- Serves additional stops along route (check sign at stop/ask driver)

Vi (Classic Hyatt)	Stanford West Apts.	Pasteur @ Sand Hill	Hospital Fountain	Stock Farm / PS5	LKSC	Main Quad	Schwab / Knight	Bowdoin @ Pine Hill	Olmsted @ Wellesley	Olmsted @ Hulme	Knight / Schwab	Main Quad	LKSC	Stock Farm / PS5	Holiday
6:00 a	6:01 a	6:04 a	6:06 a	6:13 a	6:16 a	6:19 a	6:22 a			6:27 a	6:30 a	6:33 a	6:36 a	6:39 a	6:42 a
6:30 a	6:31 a	6:34 a	6:36 a	6:43 a	6:46 a	6:49 a	6:52 a			6:57 a	7:00 a	7:03 a	7:06 a	7:09 a	7:12 a
7:00 a	7:01 a	7:04 a	7:06 a	7:13 a	7:16 a	7:19 a	7:22 a	7:26 a	7:29 a	7:33 a	7:36 a	7:39 a	7:42 a	7:46 a	7:49 a
					7:30 a	7:33 a	7:36 a	7:40 a	7:43 a	7:47 a	7:50 a	7:53 a	7:56 a		
7:30 a	7:31 a	7:34 a	7:36 a	7:43 a	7:46 a	7:49 a	7:52 a	7:56 a	7:59 a	8:03 a	8:06 a	8:09 a	8:12 a	8:16 a	8:20 a
					8:00 a	8:03 a	8:06 a	8:10 a	8:13 a	8:17 a	8:20 a	8:23 a	8:29 a		
8:00 a	8:01 a	8:04 a	8:06 a	8:13 a	8:16 a	8:19 a	8:22 a	8:26 a	8:29 a	8:33 a	8:36 a	8:39 a	8:42 a	8:46 a	8:50 a
					8:30 a	8:33 a	8:36 a	8:40 a	8:43 a	8:47 a	8:50 a	8:53 a	8:59 a		
8:30 a	8:31 a	8:34 a	8:36 a	8:43 a	8:46 a	8:49 a	8:52 a	8:56 a	8:59 a	9:03 a	9:06 a	9:09 a	9:12 a	9:16 a	9:20 a
					9:00 a	9:03 a	9:06 a	9:10 a	9:13 a	9:17 a	9:20 a	9:23 a	9:29 a		
Vi (Classic Hyatt)	Stanford West Apts.	Pasteur @ Sand Hill	Hospital Fountain	Stock Farm / PS5	LKSC	Main Quad	Schwab / Knight	Bowdoin @ Pine Hill	Olmsted @ Wellesley	Olmsted @ Hulme	Knight / Schwab	Main Quad	LKSC	Stock Farm / PS5	Holiday
9:00 a	9:01 a	9:04 a	9:06 a	9:13 a	9:16 a	9:19 a	9:22 a	9:26 a	9:29 a	9:33 a	9:36 a	9:39 a	9:42 a	9:46 a	9:49 a
					9:30 a	9:33 a	9:36 a	9:40 a	9:43 a	9:47 a	9:50 a	9:53 a	9:59 a		
9:30 a	9:31 a	9:34 a	9:36 a	9:43 a	9:46 a	9:49 a	9:52 a	9:56 a	9:59 a	10:03 a	10:06 a	10:09 a	10:12 a	10:16 a	10:20 a
					10:00 a	10:03 a	10:06 a	10:10 a	10:13 a	10:17 a	10:20 a	10:23 a	10:29 a		
10:00 a	10:01 a	10:04 a	10:06 a	10:13 a	10:16 a	10:19 a	10:22 a	10:26 a	10:29 a	10:33 a	10:36 a	10:39 a	10:42 a	10:46 a	10:50 a
					10:30 a	10:33 a	10:36 a	10:40 a	10:43 a	10:47 a	10:50 a	10:53 a	10:59 a		
10:30 a	10:31 a	10:34 a	10:36 a	10:43 a	10:46 a	10:49 a	10:52 a	10:56 a	10:59 a	11:03 a	11:06 a	11:09 a	11:12 a	11:16 a	11:20 a
11:00 a	11:01 a	11:04 a	11:06 a	11:13 a	11:16 a	11:19 a	11:22 a	11:26 a	11:29 a	11:33 a	11:36 a	11:39 a	11:42 a	11:46 a	11:50 a
11:30 a	11:31 a	11:34 a	11:36 a	11:43 a	11:46 a	11:49 a	11:52 a	11:56 a	11:59 a	12:03 p	12:06 p	12:09 p	12:12 p	12:16 p	12:20 a
Vi (Classic Hyatt)	Stanford West Apts.	Pasteur @ Sand Hill	Hospital Fountain	Stock Farm / PS5	LKSC	Main Quad	Schwab / Knight	Bowdoin @ Pine Hill	Olmsted @ Wellesley	Olmsted @ Hulme	Knight / Schwab	Main Quad	LKSC	Stock Farm / PS5	Holiday
12:00 p	12:01 p	12:04 p	12:06 p	12:13 p	12:16 p	12:19 p	12:22 p	12:26 p	12:29 p	12:33 p	12:36 p	12:39 p	12:42 p	12:46 p	12:50 a
12:30 p	12:31 p	12:34 p	12:36 p	12:43 p	12:46 p	12:49 p	12:52 p	12:56 p	12:59 p	1:03 p	1:06 p	1:09 p	1:12 p	1:16 p	1:20 a
1:00 p	1:01 p	1:04 p	1:06 p	1:13 p	1:16 p	1:19 p	1:22 p	1:26 p	1:29 p	1:33 p	1:36 p	1:39 p	1:42 p	1:46 p	1:50 a
1:30 p	1:31 p	1:34 p	1:36 p	1:43 p	1:46 p	1:49 p	1:52 p	1:56 p	1:59 p	2:03 p	2:06 p	2:09 p	2:12 p	2:16 p	2:20 a
2:00 p	2:01 p	2:04 p	2:06 p	2:13 p	2:16 p	2:19 p	2:22 p	2:26 p	2:29 p	2:33 p	2:36 p	2:39 p	2:42 p	2:46 p	2:50 a
2:30 p	2:31 p	2:34 p	2:36 p	2:43 p	2:46 p	2:49 p	2:52 p	2:56 p	2:59 p	3:03 p	3:06 p	3:09 p	3:12 p	3:16 p	3:20 a
3:00 p	3:01 p	3:04 p	3:06 p	3:13 p	3:16 p	3:19 p	3:22 p	3:26 p	3:29 p	3:33 p	3:36 p	3:39 p	3:42 p	3:46 p	3:50 a
3:30 p	3:31 p	3:34 p	3:36 p	3:43 p	3:46 p	3:49 p	3:52 p	3:56 p	3:59 p	4:03 p	4:06 p	4:09 p	4:12 p	4:16 p	4:20 a
4:00 p	4:01 p	4:04 p	4:06 p	4:13 p	4:16 p	4:19 p	4:22 p	4:26 p	4:29 p	4:33 p	4:36 p	4:39 p	4:42 p	4:46 p	4:50 a
Vi (Classic Hyatt)	Stanford West Apts.	Pasteur @ Sand Hill	Hospital Fountain	Stock Farm / PS5	LKSC	Main Quad	Schwab / Knight	Bowdoin @ Pine Hill	Olmsted @ Wellesley	Olmsted @ Hulme	Knight / Schwab	Main Quad	LKSC	Stock Farm / PS5	Holiday
4:30 p	4:31 p	4:34 p	4:36 p	4:43 p	4:46 p	4:49 p	4:52 p	4:56 p	4:59 p	5:03 p	5:06 p	5:09 p	5:12 p	5:16 p	5:20 a
5:00 p	5:01 p	5:04 p	5:06 p	5:13 p	5:16 p	5:19 p	5:22 p	5:26 p	5:29 p	5:33 p	5:36 p	5:39 p	5:42 p	5:46 p	5:50 a
5:30 p	5:31 p	5:34 p	5:36 p	5:43 p	5:46 p	5:49 p	5:52 p	5:56 p	5:59 p	6:03 p	6:06 p	6:09 p	6:12 p	6:16 p	6:20 a

Create Data Streams

Create your own!

You can use Junar to extract data from webpages *

If you find a table on a website you want to extract data, just use your bookmarklet (see below on how to install it) or copy / paste the URL (web address) into this text box and click on start!:

[Start ➤](#)

(*) Expect improvements in our Data Stream creation tool in coming weeks.

[Start](#)

Help & Tips

[Don't you know what to do here?](#)

The first time you create a Data Stream, it's normal to feel like you're lost in the jungle!

So, to make it easier for you, we've made a video that explains you how to create Data Streams!. Watch it now!

Or... You can request to us!

Don't find what you are looking for? Don't know what a URL is, and you still need data?

Don't worry. Send us a request describing what data you need, and we will create a Data Stream for you!

I need:

[Send](#)

And please, check out the [Help & Tips](#) section to clear out all your doubts.

Bookmarklet

[Skip the copy paste of the URL!](#)

Just drag and drop the "Send to Junar" link into your browser's bookmarks or favorites and click it to create a data stream.

Advertising



C Line C

Schedule effective 3/1/2011

Weekday

- Does not stop at stations
- Serves a limited number of stops (stop/ask driver)

Tip: Tables highlighted in orange contain the data that can be extracted. Click on the cells you want to extract or Right click on the table for more extracting options. Gray Cells are not selectable.

Stop/ask driver)

VI (Classic Hyatt)	S	Finish	Clean	Cancel	Stock Farm / PS5	LKSC	Main Quad	Schwab / Knight	Bowdoin @ Pine Hill	Olmsted @ Wellesley	Olmsted @ Hulme	Knight / Schwab	Main Quad	LKSC	Stock Farm / PS5	Hospital Fountain
6:00 a		6:01 a	6:04 a	6:06 a	6:13 a	6:16 a	6:19 a	6:22 a			6:27 a	6:30 a	6:33 a	6:36 a	6:39 a	6:45 a
6:30 a		6:31 a	6:34 a	6:36 a	6:43 a	6:46 a	6:49 a	6:52 a			6:57 a	7:00 a	7:03 a	7:06 a	7:09 a	7:15 a
7:00 a		7:01 a	7:04 a	7:06 a	7:13 a	7:16 a	7:19 a	7:22 a	7:26 a	7:29 a	7:33 a	7:36 a	7:39 a	7:42 a	7:46 a	7:52 a
						7:30 a	7:33 a	7:36 a	7:40 a	7:43 a	7:47 a	7:50 a	7:53 a	7:59 a		
7:30 a		7:31 a	7:34 a	7:36 a	7:43 a	7:46 a	7:49 a	7:52 a	7:56 a	7:59 a	8:03 a	8:06 a	8:09 a	8:12 a	8:16 a	8:22 a
						8:00 a	8:03 a	8:06 a	8:10 a	8:13 a	8:17 a	8:20 a	8:23 a	8:29 a		
8:00 a		8:01 a	8:04 a	8:06 a	8:13 a	8:16 a	8:19 a	8:22 a	8:26 a	8:29 a	8:33 a	8:36 a	8:39 a	8:42 a	8:46 a	8:52 a
						8:30 a	8:33 a	8:36 a	8:40 a	8:43 a	8:47 a	8:50 a	8:53 a	8:59 a		
8:30 a		8:31 a	8:34 a	8:36 a	8:43 a	8:46 a	8:49 a	8:52 a	8:56 a	8:59 a	9:03 a	9:06 a	9:09 a	9:12 a	9:16 a	9:22 a
						9:00 a	9:03 a	9:06 a	9:10 a	9:13 a	9:17 a	9:20 a	9:23 a	9:29 a		
VI (Classic Hyatt)	Stanford West Apts.	Pasteur @ Sand Hill	Hospital Fountain	Stock Farm / PS5	LKSC	Main Quad	Schwab / Knight	Bowdoin @ Pine Hill	Olmsted @ Wellesley	Olmsted @ Hulme	Knight / Schwab	Main Quad	LKSC	Stock Farm / PS5	Hospital Fountain	
9:00 a		9:01 a	9:04 a	9:06 a	9:13 a	9:16 a	9:19 a	9:22 a	9:26 a	9:29 a	9:33 a	9:36 a	9:39 a	9:42 a	9:46 a	9:52 a
						9:30 a	9:33 a	9:36 a	9:40 a	9:43 a	9:47 a	9:50 a	9:53 a	9:59 a		
9:30 a		9:31 a	9:34 a	9:36 a	9:43 a	9:46 a	9:49 a	9:52 a	9:56 a	9:59 a	10:03 a	10:06 a	10:09 a	10:12 a	10:16 a	10:22 a
						10:00 a	10:03 a	10:06 a	10:10 a	10:13 a	10:17 a	10:20 a	10:23 a	10:29 a		
10:00 a		10:01 a	10:04 a	10:06 a	10:13 a	10:16 a	10:19 a	10:22 a	10:26 a	10:29 a	10:33 a	10:36 a	10:39 a	10:42 a	10:46 a	10:52 a
						10:30 a	10:33 a	10:36 a	10:40 a	10:43 a	10:47 a	10:50 a	10:53 a	10:59 a		
10:30 a		10:31 a	10:34 a	10:36 a	10:43 a	10:46 a	10:49 a	10:52 a	10:56 a	10:59 a	11:03 a	11:06 a	11:09 a	11:12 a	11:16 a	11:22 a
11:00 a		11:01 a	11:04 a	11:06 a	11:13 a	11:16 a	11:19 a	11:22 a	11:26 a	11:29 a	11:33 a	11:36 a	11:39 a	11:42 a	11:46 a	11:52 a
11:30 a		11:31 a	11:34 a	11:36 a	11:43 a	11:46 a	11:49 a	11:52 a	11:56 a	11:59 a	12:03 p	12:06 p	12:09 p	12:12 p	12:16 p	12:22 a
VI (Classic Hyatt)	Stanford West Apts.	Pasteur @ Sand Hill	Hospital Fountain	Stock Farm / PS5	LKSC	Main Quad	Schwab / Knight	Bowdoin @ Pine Hill	Olmsted @ Wellesley	Olmsted @ Hulme	Knight / Schwab	Main Quad	LKSC	Stock Farm / PS5	Hospital Fountain	
12:00 p		12:01 p	12:04 p	12:06 p	12:13 p	12:16 p	12:19 p	12:22 p	12:26 p	12:29 p	12:33 p	12:36 p	12:39 p	12:42 p	12:46 p	12:52 a
12:30 p		12:31 p	12:34 p	12:36 p	12:43 p	12:46 p	12:49 p	12:52 p	12:56 p	12:59 p	1:03 p	1:06 p	1:09 p	1:12 p	1:16 p	1:22 a
1:00 p		1:01 p	1:04 p	1:06 p	1:13 p	1:16 p	1:19 p	1:22 p	1:26 p	1:29 p	1:33 p	1:36 p	1:39 p	1:42 p	1:46 p	1:52 a
1:30 p		1:31 p	1:34 p	1:36 p	1:43 p	1:46 p	1:49 p	1:52 p	1:56 p	1:59 p	2:03 p	2:06 p	2:09 p	2:12 p	2:16 p	2:22 a
2:00 p		2:01 p	2:04 p	2:06 p	2:13 p	2:16 p	2:19 p	2:22 p	2:26 p	2:29 p	2:33 p	2:36 p	2:39 p	2:42 p	2:46 p	2:52 a
2:30 p		2:31 p	2:34 p	2:36 p	2:43 p	2:46 p	2:49 p	2:52 p	2:56 p	2:59 p	3:03 p	3:06 p	3:09 p	3:12 p	3:16 p	3:22 a
3:00 p		3:01 p	3:04 p	3:06 p	3:13 p	3:16 p	3:19 p	3:22 p	3:26 p	3:29 p	3:33 p	3:36 p	3:39 p	3:42 p	3:46 p	3:52 a
3:30 p		3:31 p	3:34 p	3:36 p	3:43 p	3:46 p	3:49 p	3:52 p	3:56 p	3:59 p	4:03 p	4:06 p	4:09 p	4:12 p	4:16 p	4:22 a
4:00 p		4:01 p	4:04 p	4:06 p	4:13 p	4:16 p	4:19 p	4:22 p	4:26 p	4:29 p	4:33 p	4:36 p	4:39 p	4:42 p	4:46 p	4:52 a
						4:30 p	4:33 p	4:36 p	4:40 p	4:43 p	4:47 p	4:50 p	4:53 p	4:59 p		
VI (Classic Hyatt)	Stanford West Apts.	Pasteur @ Sand Hill	Hospital Fountain	Stock Farm / PS5	LKSC	Main Quad	Schwab / Knight	Bowdoin @ Pine Hill	Olmsted @ Wellesley	Olmsted @ Hulme	Knight / Schwab	Main Quad	LKSC	Stock Farm / PS5	Hospital Fountain	
4:30 p		4:31 p	4:34 p	4:36 p	4:43 p	4:46 p	4:49 p	4:52 p	4:56 p	4:59 p	5:03 p	5:06 p	5:09 p	5:12 p	5:16 p	5:22 a
5:00 p		5:01 p	5:04 p	5:06 p	5:13 p	5:16 p	5:19 p	5:22 p	5:26 p	5:29 p	5:33 p	5:36 p	5:39 p	5:42 p	5:46 p	5:52 a
						5:30 p	5:33 p	5:36 p	5:40 p	5:43 p	5:47 p	5:50 p	5:53 p	5:59 p		

Dashboard Manager

[Add or create Dashboards](#)[Geoff's Tasks](#) [Help & Tips](#)

Geoff's Tasks by me | [Edit](#)

List of tasks and info from other sites

Actions: Like (0) | Export Dashboard | View C

Share Dashboard: Share | Tweet | 0 URL: <http://bit.ly/...>[+ Add Data Streams](#)

Cabot Oil and Gas Corp

Analysts Info

ECONOMY | by marir1411 4 months, 3 weeks ago

Annual EPS Est (Dec-11): 1.12

Quarterly EPS Est (Jun-11): 0.24

Mean Recommendation*: 2.2

PEG Ratio (5 yr expected): 2.13

Cabot Oil

Last Trade

ECONOMY | by marir1411 4 months, 3 weeks ago

52.44

Marguerite Line C Schedule

Bowdoin Pine Hill Departures

EDUCATION | by mcgeoff 1 hour, 55 minutes ago

Bowdoin @ Pine Hill 6:33 a 7:03 a 7:26 a 7:39 a

[Add a Data Stream](#)

Life Cycle of Visualizations

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research

Spreadsheets
Databases
Digitized Media

Other Downloads

Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

[Junar](#)
[Outwit Hub](#)
[ScraperWiki](#)

Life Cycle of Visualizations

Gathering Data

Discovery/
Acquisition

Original Research

Spreadsheets

Databases

Digitized Media

Other Downloads

Public Data

Archives/Libraries

Academic Partners

Purchase

Scraping

Junar

Outwit Hub

ScraperWiki

Analyzing It

Cleaning/
“Munging”

Normalization,
Format
Conversion

Google Refine

Data Wrangler

Mr. Data Converter

Sharing Findings

Publication



Life Cycle of Visualizations

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping
Junar
Outwit Hub
ScraperWiki

Normalization,
Format
Conversion
Google Refine
Data Wrangler
Mr. Data Converter



[Project Home](#) [Downloads](#) [Wiki](#) [Issues](#) [Source](#)

«

Downloads

Updated Nov 10, 2010 by dfhu...@goo...

[Getting Started - FAQ](#)[Documentation for Users](#)[Documentation for Developers](#)[How to Participate](#)[+ Other resources](#)For all downloads, see [this list](#).**Here are [instructions for upgrading from 1.1 to 2.0](#).**

Windows



Mac

google-refine-2.0-r1836.dmg

Download, open, drag icon into the Applications folder and double click on it.



Linux

The list of all previous releases is [here](#).

Google Refine demo

Contractor Name						Type of Contract	Date of Award	Start Date	End Date	Total value of Con
ASAP SOFTWARE EXPRESS INC DELL MARKETING L.P.		Facet	Text facet			06/30/2011				
BMC SOFTWARE DISTRIBUTION INCORPORATED GOVCONNECTION INCORPORATED		Text filter	Numeric facet			03/31/2010				
ITS CORPORATION		Edit cells	Timeline facet			04/30/2011				
SENET INTERNATIONAL CORPORATION		Edit column	Scatterplot facet			12/30/2011				
		Transpose	Custom text facet...							
		Sort...	Custom numeric facet...							
		View	Customized facets							
		Reconcile	05/04/2009	05/05/2009	07/03/2009					0.04
IT FEDERAL SALES LIMITED LIABILITY COMPANY		firm fixed price	01/26/2009	01/26/2010	09/30/2010					
		firm fixed price	10/01/2009	10/01/2009	09/25/2010					
		firm fixed price	09/30/2009	10/01/2009	09/30/2010					
		firm fixed price	11/05/2009	11/05/2009	09/30/2010					
REDHAWK IT SOLUTIONS LLC		firm fixed price	01/22/2009	01/01/2010	12/31/2010					

Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

[Junar](#)
[Outwit Hub](#)
[ScraperWiki](#)

Analyzing It

Cleaning/
“Munging”

Normalization,
Format
Conversion
[Google Refine](#)
[Data Wrangler](#)
[Mr. Data Converter](#)

Sharing Findings

Publication



Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

Junar
Outwit Hub
ScraperWiki

**Normalization,
Format
Conversion**

Google Refine
Data Wrangler

Mr. Data Converter



Mr. Data Converter

I will convert your Excel data into one of several web-friendly formats, including HTML, JSON and XML.

Email me at mr.dataconverter@gmail.com.

News and updates available via [Twitter](#).

Source available on [github](#).

SETTINGS

First row is the header

Transform: downcase upcase none

Include white space in output

Indent with: tabs spaces

Input CSV or tab-delimited data. Using Excel? Simply copy and paste. No data on hand? [Use sample](#)

NAME	VALUE	COLOR	DATE
Alan	12	blue	Sep. 25, 2009
Shan	13	"green	blue"
John	45	orange	Sep. 29, 2009
Minna	27	teal	Sep. 30, 2009

Output as XML - Nodes

```
<?xml version="1.0" encoding="UTF-8"?>
<rows>
  <row>
    <NAME>Alan</NAME>
    <VALUE>12</VALUE>
    <COLOR>blue</COLOR>
    <DATE>Sep. 25, 2009</DATE>
  </row>
  <row>
    <NAME>Shan</NAME>
    <VALUE>13</VALUE>
    <COLOR>green\tblue</COLOR>
    <DATE>Sep. 27, 2009</DATE>
  </row>
  <row>
    <NAME>John</NAME>
    <VALUE>45</VALUE>
    <COLOR>orange</COLOR>
    <DATE>Sep. 29, 2009</DATE>
  </row>
  <row>
    <NAME>Minna</NAME>
    <VALUE>27</VALUE>
    <COLOR>teal</COLOR>
    <DATE>Sep. 30, 2009</DATE>
  </row>
</rows>
```

Mr. Data Converter

I will convert your Excel data into one of several web-friendly formats, including HTML, JSON and XML.

Email me at mr.dataconverter@gmail.com.

News and updates available via [Twitter](#).

Source available on [github](#).

SETTINGS

First row is the header

Transform: downcase upcase none

Include white space in output

Indent with: tabs spaces

Input CSV or tab-delimited data. Using Excel? Simply copy and paste. No data on hand? [Use sample](#)

NAME	VALUE	COLOR	DATE
Alan	12	blue	Sep. 25, 2009
Shan	13	"green blue"	Sep. 27, 2009
John	45	orange	Sep. 29, 2009
Minna	27	teal	Sep. 30, 2009

Output as JSON - Properties ▼

```
[{"NAME": "Alan", "VALUE": 12, "COLOR": "blue", "DATE": "Sep. 25, 2009"},  
 {"NAME": "Shan", "VALUE": 13, "COLOR": "green\\tblue", "DATE": "Sep. 27, 2009"},  
 {"NAME": "John", "VALUE": 45, "COLOR": "orange", "DATE": "Sep. 29, 2009"},  
 {"NAME": "Minna", "VALUE": 27, "COLOR": "teal", "DATE": "Sep. 30, 2009"}];
```

Mr. Data Converter

I will convert your Excel data into one of several web-friendly formats, including HTML, JSON and XML.

Email me at mr.dataconverter@gmail.com.

News and updates available via [Twitter](#).

Source available on [github](#).

SETTINGS

First row is the header

Transform: downcase upcase none

Include white space in output

Indent with: tabs spaces

Input CSV or tab-delimited data. Using Excel? Simply copy and paste. No data on hand? [Use sample](#)

NAME	VALUE	COLOR	DATE
Alan	12	blue	Sep. 25, 2009
Shan	13	"green	blue"
John	45	orange	Sep. 29, 2009
Minna	27	teal	Sep. 30, 2009

Actionsript
ASP/VBScript
HTML
JSON - Properties
JSON - Column Arrays
JSON - Row Arrays

Output as: MySQL

PHP
Python - Dict
Ruby
XML - Properties
XML - Nodes

```
CREATE TABLE MrDataConverter (
    id INT NOT NULL AUTO_INCREMENT,
    NAME VARCHAR(255) NOT NULL,
    VALUE VARCHAR(255) NOT NULL,
    COLOR VARCHAR(255),
    DATE VARCHAR(255)
);
INSERT INTO MrDataConverter
    (NAME,VALUE,COLOR,DATE)
VALUES
    ('Alan',12,'blue','Sep. 25, 2009'),
    ('Shan',13,'green\tblue','Sep. 27, 2009'),
    ('John',45,'orange','Sep. 29, 2009'),
    ('Minna',27,'teal','Sep. 30, 2009');
```

Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

[Junar](#)
[Outwit Hub](#)
[ScraperWiki](#)

Analyzing It

Cleaning/
“Munging”

Normalization,
Format
Conversion
[Google Refine](#)
[Data Wrangler](#)
[Mr. Data Converter](#)

Sharing Findings

Publication



Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping
Junar
Outwit Hub
ScraperWiki

Normalization,
Format
Conversion
Google Refine
Data Wrangler
Mr. Data Converter

Web Services
Google Spreadsheets
Google Fusion Tables
IBM ManyEyes

Applications
Tableau/Tableau Public
MS Office, OpenOffice
Gephi
Node XL (plug-in for Excel)
Spotfire
R
Processing



Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping
[Junar](#)
[Outwit Hub](#)
[ScraperWiki](#)

Normalization,
Format
Conversion
[Google Refine](#)
[Data Wrangler](#)
[Mr. Data Converter](#)

Web Services
[Google Spreadsheets](#)
[Google Fusion Tables](#)
IBM ManyEyes

Applications
Tableau/Tableau Public
MS Office, OpenOffice
[Gephi](#)
[Node XL](#) (plug-in for Excel)
[Spotfire](#)
[R](#)
[Processing](#)



Getting Started with Visualization

Free and Web-Based Applications

The screenshot shows the IBM ManyEyes homepage. At the top, there's a navigation bar with 'Many Eyes' on the left, 'Log in' and 'IBM' on the right. Below the navigation is a banner stating 'An experiment brought to you by IBM Research and the IBM Cognos software group'. The main content area features a large title 'Create a visualization in three easy steps' with three numbered steps: 1. CHOOSE A DATA SET, 2. CHOOSE A VISUALIZATION, and 3. CUSTOMIZE & PUBLISH. Each step has a brief description. To the left of the steps is a sidebar with sections for 'Explore' (Visualizations, Data sets, Comments, Topic centers), 'Participate' (Create a visualization, Upload a data set, Create a topic center, Register), 'Learn more' (Quick start, Visualization types, Data format and style, About Many Eyes, FAQ, Blog), 'Contact us' (Contact, Report a bug), 'Legal' (Terms of use, Privacy, Provide Feedback), and 'Popular tags' (Visualizations, Data sets, 2008, 2009, 2010, OECDFactbook, Obama, Patent, US, USPTO, budget, canada, census, crime, education). The central area also includes a 'Start with a data set' section with instructions and a link to upload a dataset.

IBM ManyEyes

<http://manyeyes.alphaworks.ibm.com>

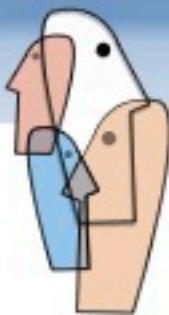
Pros: Easy as pie

Many different chart forms

Interactivity

Bring your own data, or
use existing data set

Cons: Java applets are slow, clunky
Little design control



Visualization Options Available in Many Eyes

Explore

[Visualizations](#)
[Data sets](#)
[Comments](#)
[Topic centers](#)

Participate

[Create a visualization](#)
[Upload a data set](#)
[Create a topic center](#)
[Register](#)

Learn more

[Quick start](#)
[Visualization types](#)
[Data format and style](#)
[About Many Eyes](#)
[FAQ](#)
[Blog](#)

Contact us

[Contact](#)
[Report a bug](#)

Legal

[Terms of use](#)
[Privacy](#)
[Provide Feedback](#)

Popular tags:

[Visualizations](#) [Data sets](#)

2008 2009 2010
OECD Factbook

Finding the right way view your data is as much an art as a science. The visualizations provided on Many Eyes range from the ordinary to the experimental.

See relationships among data points

- Network Diagram
- Scatterplot
- Matrix Chart

Compare a set of values

- Bar Chart
- Block Histogram
- Bubble Chart

Track rises and falls over time

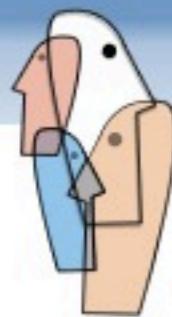
- Line Graph
- Stack Graph
- Stack Graph for Categories

See the parts of a whole

- Pie Chart
- Treemap
- Treemap for Comparisons

Analyze a text

- Word Tree
- Tag Cloud



Visualizations : Degrees earned, by Gender over 5 decades

Explore

[Visualizations](#)
[Data sets](#)
[Comments](#)
[Topic centers](#)

Participate

[Create a visualization](#)
[Upload a data set](#)
[Create a topic center](#)
[Register](#)

Learn more

[Quick start](#)
[Visualization types](#)
[Data format and style](#)
[About Many Eyes](#)
[FAQ](#)
[Blog](#)

Contact us

[Contact](#)
[Report a bug](#)

Legal

[Terms of use](#)
[Privacy](#)
[Provide Feedback](#)

Popular tags:

[Visualizations](#) [Data sets](#)

2008 2009 2010

Sex

Male
 Female

Expert Options

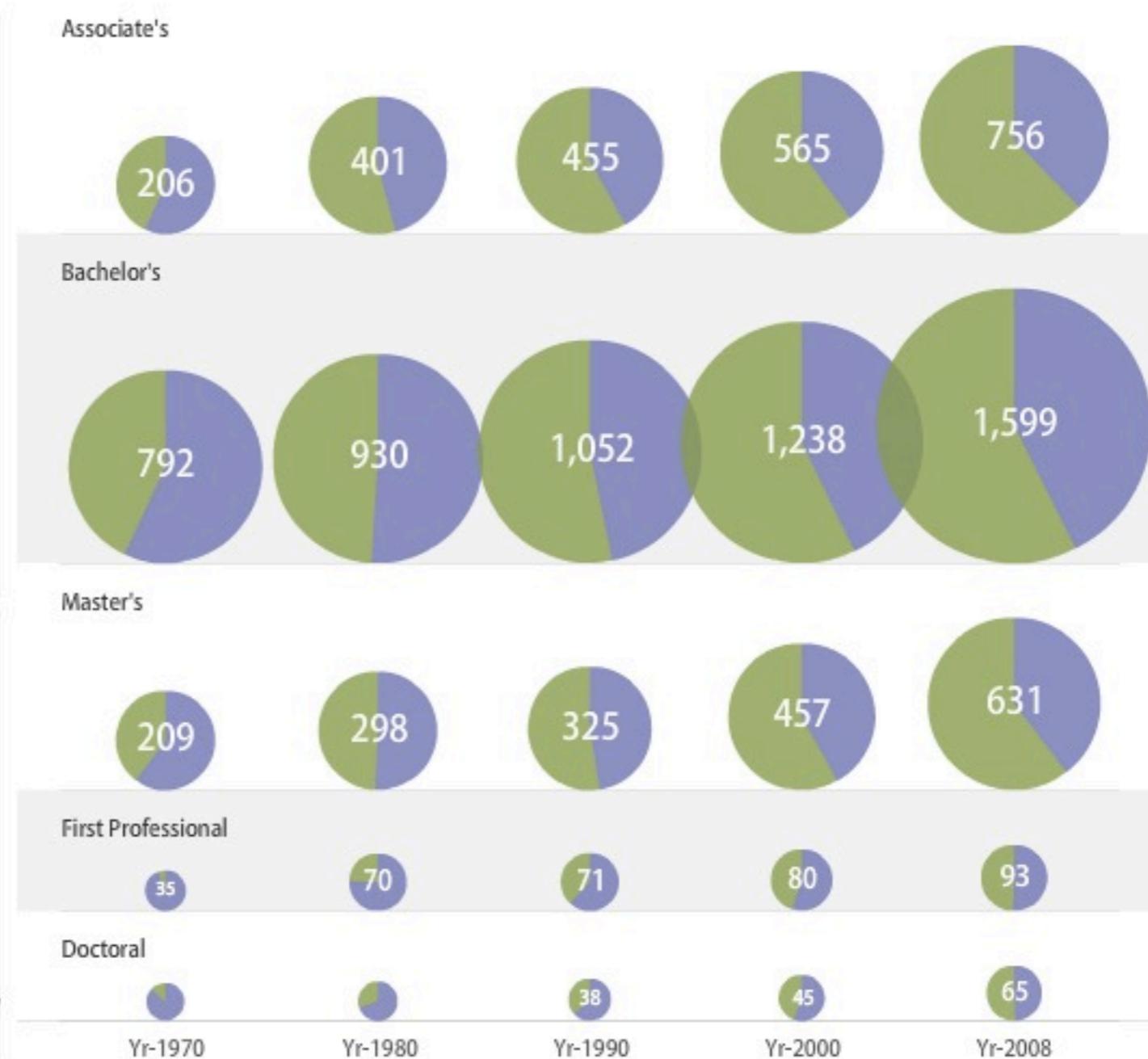
Bubbles
 Bars

Aggregate By

Total
 Average

Size By

Value
 Percent of Row
 Same Size





Visualizations : Degrees earned, by Gender over 5 decades

Explore

Visualizations
Data sets
Comments
Topic centers

Participate

Create a visualization
Upload a data set
Create a topic center
Register

Learn more

Quick start
Visualization types
Data format and style
About Many Eyes
FAQ
Blog

Contact us

Contact
Report a bug

Legal

Terms of use
Privacy
Provide Feedback

Popular tags:

Visualizations Data sets

2008 2009 2010

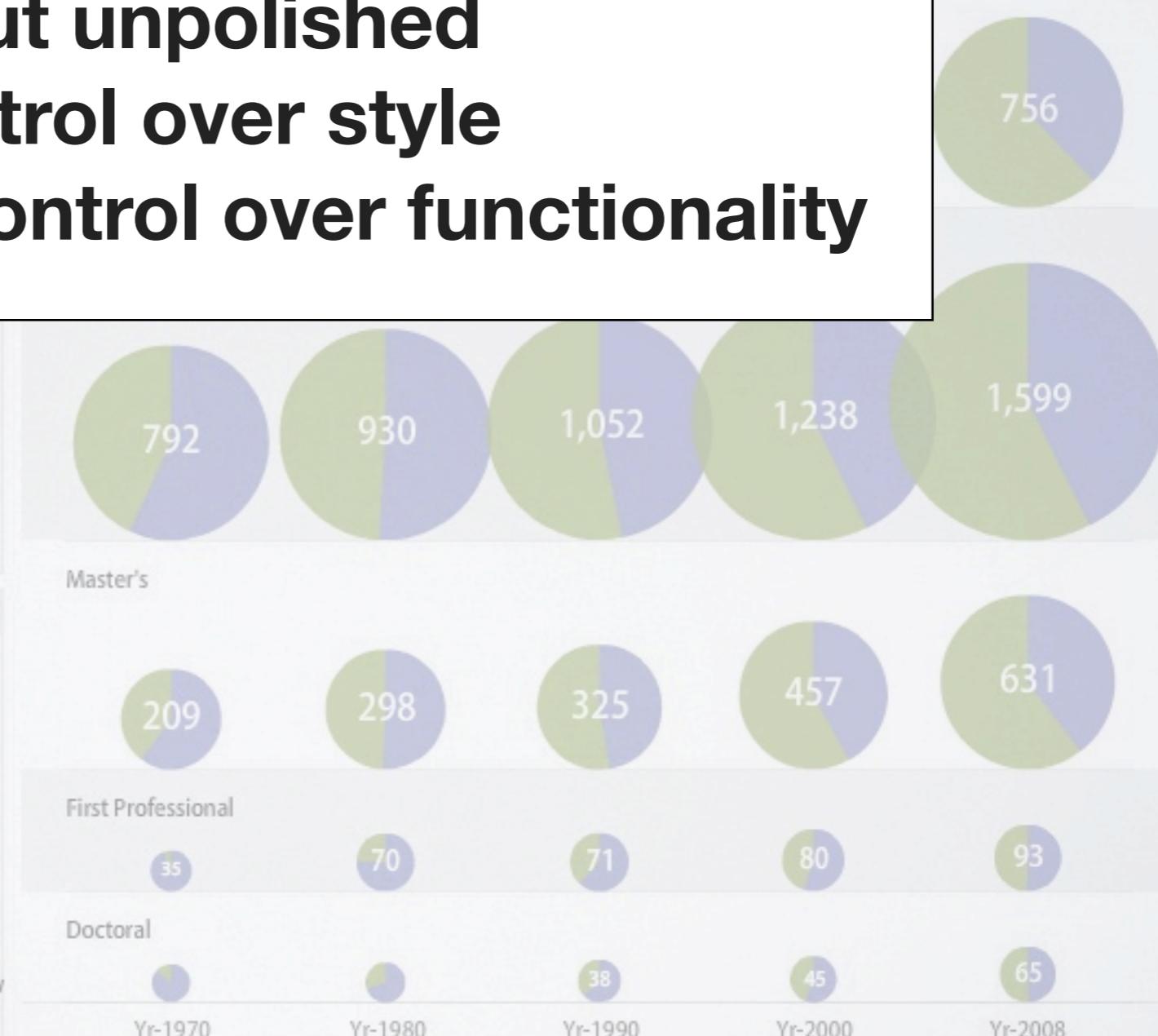
Uploaded by: vkn13

Created at: Feb 7 2011

Description: Females have comfortably overtaken Males in total number of degrees earned

Tags: education degree sex

**Easy but unpolished
No control over style
Little control over functionality**



Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping
Junar
Outwit Hub
ScraperWiki

Normalization,
Format
Conversion
Google Refine
Data Wrangler
Mr. Data Converter

Web Services
Google Spreadsheets
Google Fusion Tables
IBM ManyEyes

Applications
Tableau/Tableau Public
MS Office, OpenOffice
Gephi
Node XL (plug-in for Excel)
Spotfire
R
Processing



Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping
[Junar](#)
[Outwit Hub](#)
[ScraperWiki](#)

**Normalization,
Format
Conversion**
[Google Refine](#)
[Data Wrangler](#)
[Mr. Data Converter](#)

Web Services
[Google Spreadsheets](#)
[Google Fusion Tables](#)
[IBM ManyEyes](#)

Applications
[Tableau/Tableau Public](#)
MS Office, OpenOffice
[Gephi](#)
[Node XL](#) (plug-in for Excel)
[Spotfire](#)
[R](#)
[Processing](#)

Static Visualizations
Previous tools +
Adobe Illustrator
Adobe Photoshop

Animated Visualizations
[Processing](#)
Adobe Flash
Adobe After Effects

Interactive/Web Visualizations
HTML5
[Protopis](#)
[D3](#)
<http://processingjs.org/>
Adobe Flash or Flex
[Processing](#)

Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research
Spreadsheets
Databases
Digitized Media

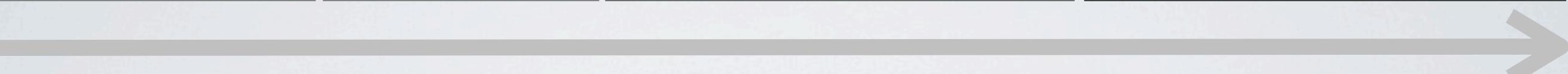
Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping
Junar
Outwit Hub
ScraperWiki

**Normalization,
Format
Conversion**
Google Refine
Data Wrangler
Mr. Data Converter

Web Services
Google Spreadsheets
Google Fusion Tables
IBM ManyEyes

Applications
Tableau/Tableau Public
MS Office, OpenOffice
Gephi
Node XL (plug-in for Excel)
Spotfire
R
Processing

Static Visualizations
Previous tools +
Adobe Illustrator
Adobe Photoshop



Animated Visualizations
Processing
Adobe Flash
Adobe After Effects

Interactive/Web Visualizations
HTML5
Protopis
D3
http://processingjs.org/
Adobe Flash or Flex
Processing

Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research
Spreadsheets
Databases
Digitized Media

Other Downloads
Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping
Junar
Outwit Hub
ScraperWiki

Normalization,
Format
Conversion
Google Refine
Data Wrangler
Mr. Data Converter

Web Services
Google Spreadsheets
Google Fusion Tables
IBM ManyEyes



Static Visualizations
Previous tools +
Adobe Illustrator } refinement
Adobe Photoshop

Animated Visualizations
Processing
Adobe Flash COMMONLY USED FOR NEWS
Adobe After Effects

Interactive/Web Visualizations
HTML5
Protopis
D3
http://processingjs.org/
Adobe Flash or Flex
Processing

Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research

Spreadsheets
Databases
Digitized Media

Other Downloads

Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

Junar
Outwit Hub
ScraperWiki

Normalization, Format Conversion

Google Refine
Data Wrangler
Mr. Data Converter

Web Services

Google Spreadsheets
Google Fusion Tables
IBM ManyEyes

Applications

Tableau/Tableau Public
MS Office, OpenOffice
Gephi
Node XL (plug-in for Excel)
Spotfire
R
Processing

Static Visualizations

Previous tools +
Adobe Illustrator }
Adobe Photoshop } **refinement**

Animated Visualizations

Processing
Adobe Flash
Adobe After Effects

COMMON FOR ART



Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research

Spreadsheets
Databases
Digitized Media

Other Downloads

Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

Junar
Outwit Hub
ScraperWiki

Normalization, Format Conversion

Google Refine
Data Wrangler
Mr. Data Converter

Web Services

Google Spreadsheets
Google Fusion Tables
IBM ManyEyes

Applications

Tableau/Tableau Public
MS Office, OpenOffice
Gephi
Node XL (plug-in for Excel)
Spotfire
R
Processing

Static Visualizations

Previous tools +
Adobe Illustrator }
Adobe Photoshop } ***refinement***

Animated Visualizations

~~Processing~~ ***WON'T WORK***
~~Adobe Flash~~ ***ON IPHONE/IPAD***
Adobe After Effects

Interactive/Web Visualizations

HTML5
Protopis
D3
<http://processingjs.org/>
~~Adobe Flash or Flex~~
~~Processing~~

Data Visualization Workflow

Gathering Data

Discovery/
Acquisition

Cleaning/
“Munging”

Analyzing It

Analysis/Exploratory
Visualization

Sharing Findings

Publication

Original Research

Spreadsheets
Databases
Digitized Media

Other Downloads

Public Data
Archives/Libraries
Academic Partners
Purchase

Scraping

Junar
Outwit Hub
ScraperWiki

Normalization, Format Conversion

Google Refine
Data Wrangler
Mr. Data Converter

Web Services

Google Spreadsheets
Google Fusion Tables
IBM ManyEyes

Applications

Tableau/Tableau Public
MS Office, OpenOffice
Gephi
Node XL (plug-in for Excel)
Spotfire
R
Processing

Static Visualizations

Previous tools +
Adobe Illustrator }
Adobe Photoshop } *refinement*

Animated Visualizations

Processing
Adobe Flash
Adobe After Effects

Interactive/Web Visualizations

HTML5
Protopis
D3
http://processingjs.org/
Adobe Flash or Flex
Processing

Video Documentary

Journalism in the Age of Data

E-mail this to a friend | Printable version

In graphics: Eurozone in crisis

Debt | Economy | Recession | Jobs

Debt □ Deficit

National debt as a percentage of GDP, 2009

Country	Total debt (% GDP)	Total debt (€ m)	2009 deficit (% GDP)
Ireland	64.0	1,790,755	5.5
Greece	115.1	273,437	13.6
Belgium	96.7	306,806	6.6
France	77.8	1,483,025	7.6
Portugal	75.8	125,810	9.4
Germany	73.2	1,782,211	3.3
Malta	69.1	3,945	3.8
UK	58.1	1,067,819	11.5
Austria	66.5	184,126	3.4
Ireland	64.0	104,807	14.3
Netherlands	60.3	347,021	5.3
Cyprus	56.2	19,507	6.1
Spain	53.2	558,850	11.2
Portugal	48	79,217	2.2
Slovenia	35.9	12,519	5.5
Slovakia	35.7	22,885	6.8
Luxembourg	14.5	5,494	0.7

SOURCE: Eurostat, All figures for 2009

One of the main causes of the currency crisis in the eurozone is that

5:08

RELATED INFORMATION FOR THIS CHAPTER

Jeffrey Heer Steve Duenes Amanda Cox J. Paige West Scott Byrne-Fraser



Scott Byrne-Fraser, BBC News

Scott Byrne-Fraser
Creative Director/Team Leader
BBCNews.com

Byrne-Fraser is the Creative Director of the BBC News website. He manages a team of about eight designers, eight journalists and four developers who create maps, infographics, charts and data visualizations. He comes from a background producing 3D and motion graphics for broadcast news at Sky Television. He wants to foster collaboration between the traditionally separate on-air and online graphics teams at the BBC. "We've got a huge TV design team, a huge online design team, and they're currently totally separate silos, and I want to bring the two together so we can do more videos online." He is also working to re-orient the group's focus towards more data analysis and visualization.

CHAPTERS

I. Introduction

II. Data Vis in Journalism

III. Telling "Data Stories"

IV. A New Era in Infographics

V. Life as a Data Stream

VI. Exploring Data

VII. Technologies and Tools

VIII. First Steps

RELATED LINKS

- BBC News Special Reports
- Mapping UK's Teen Murder Toll
- Bank of England Interest Rates, 1694-2010

SHARE |

OPTIONS FOR VIEWING THIS VIDEO

ANNOTATED VERSION

NON-FLASH

DESCRIPTION

KEY POINTS

Journalists are coping with the rising information flood by borrowing data visualization techniques from computer scientists, researchers and artists. Some newsrooms are already beginning to retool their staffs and systems to prepare for a future in which data becomes a medium. But how do we communicate with data, how can traditional narratives be fused with sophisticated, interactive information displays?

ABOUT ME

ACKNOWLEDGEMENTS

GEOFF McGHEE is an online journalist specializing in multimedia and information graphics. Over the past decade he has worked at The New York Times and ABCNews.com, and in France at Le Monde Interactif. In 2009-2010, he spent a John S. Knight Journalism Fellowship at Stanford University studying data visualization. In July 2010, McGhee began a new job developing visualizations and interactive content for the Bill Lane Center for the American West at Stanford.



Journalism in the Age of Data by Geoff McGhee is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.

Video Documentary

Journalism in the Age of Data

A video report on data visualization as a storytelling medium
Produced during a 2009-2010 Knight Journalism Fellowship
Total Running Time: 54 Minutes; with related information and links

E-mail this to a friend | Printable version

In graphics: Eurozone in crisis

Debt Economy Recession Jobs

Debt Deficit

National debt as a percentage of GDP, 2009

Click on a heading to sort by that measure

	Total debt (% GDP)	Total debt (K m)	2009 deficit (% GDP)
Italy	115.8	1,750,755	5.5
Greece	116.1	273,437	13.6
Ireland	66.7	305,806	6.6
France	77.8	1,483,025	7.6
Portugal	76.3	125,810	9.4
Germany	73.2	1,762,211	3.3
Austria	81.1	379,414	3.8

One of the main causes of the currency crisis in the eurozone is that

5:08

RELATED INFORMATION FOR THIS CHAPTER

Jeffrey Heer  Steve Duenes  Amanda Cox  J. Paige West  Scott Byrne-Fraser 



Scott Byrne-Fraser, BBC News

Byrne-Fraser is the Creative Director of the BBC News website. He manages a team of about eight designers, eight journalists and four developers who create maps, infographics, charts and data visualizations. He comes from a background producing 3D and motion graphics for broadcast news at Sky Television. He wants to foster collaboration between the traditionally separate on-air and online graphics teams at the BBC. "We've got a huge TV design team, a huge online design team, and they're currently totally separate silos, and I want to bring the two together so we can do more videos online." He is also working to re-orient the group's focus towards more data analysis and visualization.

Scott Byrne-Fraser
Creative Director/Team Leader
BBCNews.com

CHAPTERS

I. Introduction

II. Data Vis in Journalism

III. Telling "Data Stories"

IV. A New Era in Infographics

VII. Technologies and Tools

VIII. First Steps

RELATED LINKS

- [BBC News Special Reports](#)
- [Mapping UK's Teen Murder Toll](#)
- [Bank of England Interest Rates, 1694-2010](#)

SHARE |      

OPTIONS FOR VIEWING THIS VIDEO

ANNOTATED VERSION

NON-FLASH

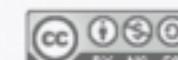
DESCRIPTION

KEY POINTS

Journalists are coping with the rising information flood by borrowing data visualization techniques from computer scientists, researchers and artists. Some newsrooms are already beginning to retool their staffs and systems to prepare for a future in which data becomes a medium. But how do we communicate with data, how can traditional with sophisticated, interactive ?

ACKNOWLEDGEMENTS

GEOFF McGHEE is an online journalist specializing in multimedia and information graphics. Over the past decade he has worked at The New York Times and ABCNews.com, and in France at Le Monde Interactif. In 2009-2010, he spent a John S. Knight Journalism Fellowship at Stanford University studying data visualization. In July 2010, McGhee began a new job developing visualizations and interactive content for the Bill Lane Center for the American West at Stanford.



Journalism in the Age of Data by Geoff McGhee is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.

Thanks!

gmcghee@stanford.edu

@mcgeoff

