MAKING PEOPLE CARE

PAIRING NARRATIVE WITH VISUALIZATION

ILIA BLINDERMAN

STRUCTURE

My background

Some theory

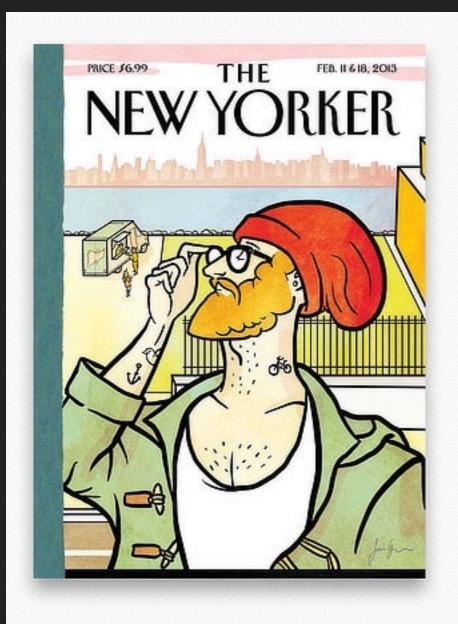
Outline of missed connections project + code examples

Discussion of pairing narrative

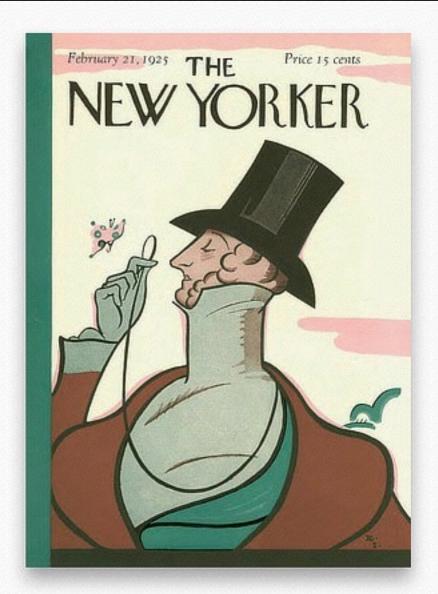
MY BACKGROUND

- Psychological research
- Journalism
- Data + viz

THE NEW YORKER MAGAZINE







EXAMPLES OF NEW YORKER NARRATIVES

Structure of long form writing - 7,597 words (38 min.)

Commuting - 6,705 words (34 min.)

Elevators - 8,100 words (41 min.)

WHY READ ABOUT ELEVATORS FOR AN HOUR?

Inherent interest: Because you're all about elevators

 Emotional resonance: Because the story garners some emotional response

Which of the two is the superior strategy?

ELEVATORS?



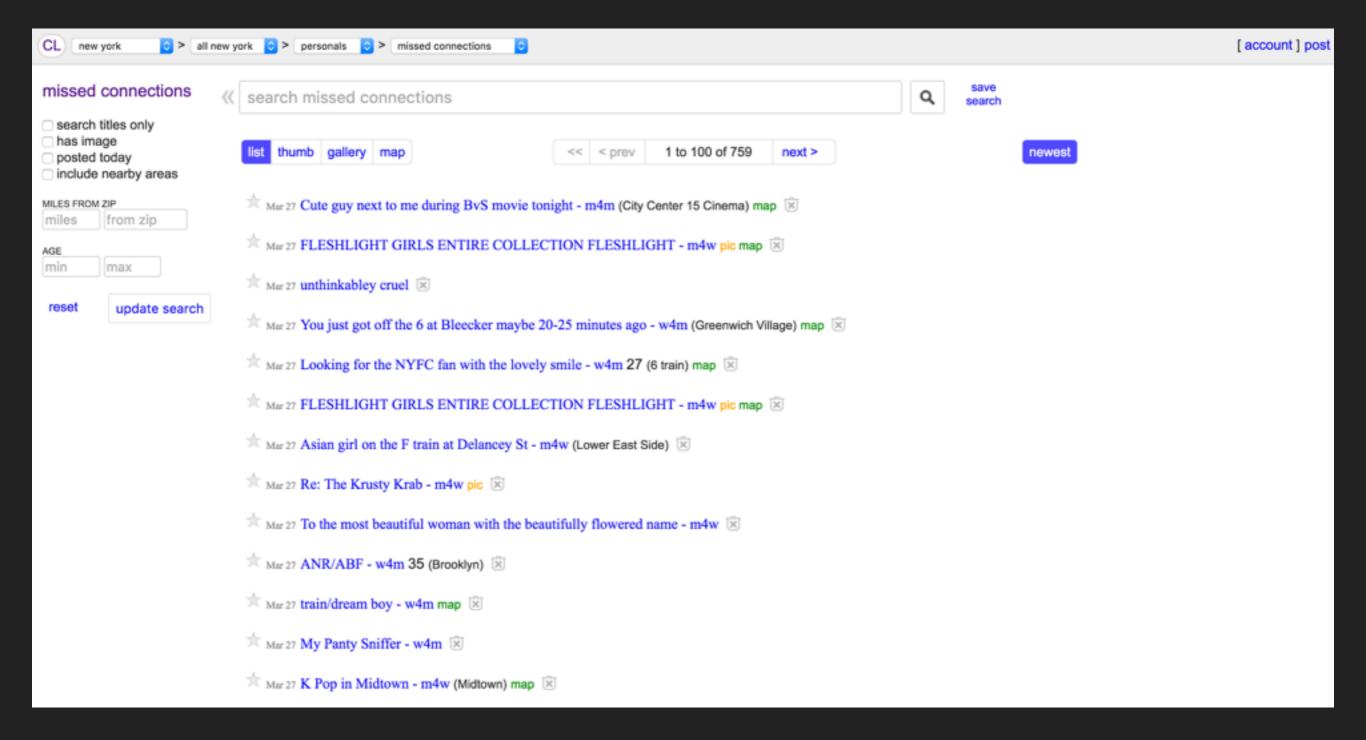
THE WAY YOU COMMUNICATE FACTS MATTERS

Communicating > presenting facts

▶ 1) Capture audience attention

> 2) Retain it.

WAIT, WHAT ABOUT THESE MISSED CONNECTIONS?



WHY GO DEEP ON MISSED CONNECTIONS?

Interesting

No one had done it

Rich dataset

Free (ish) from observer effects

HOW?

- Getting data: Python, BeautifulSoup
- Cleaning + analyzing data: R; Python's Natural Language
 Toolkit (NLTK) Package
- Visualizing data: <u>D3.js</u>

COLLECTING DATA

Python script (c/o Brian Abelson)

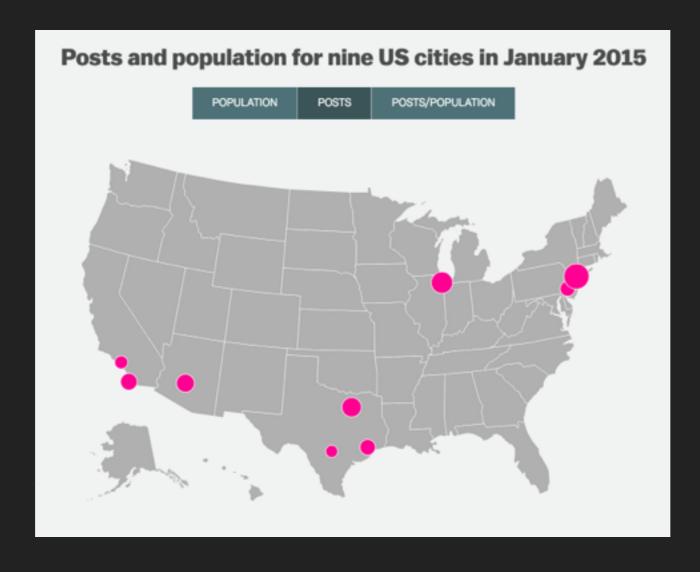
Daily manual scraping during January 2015

~12,000 missed connections records

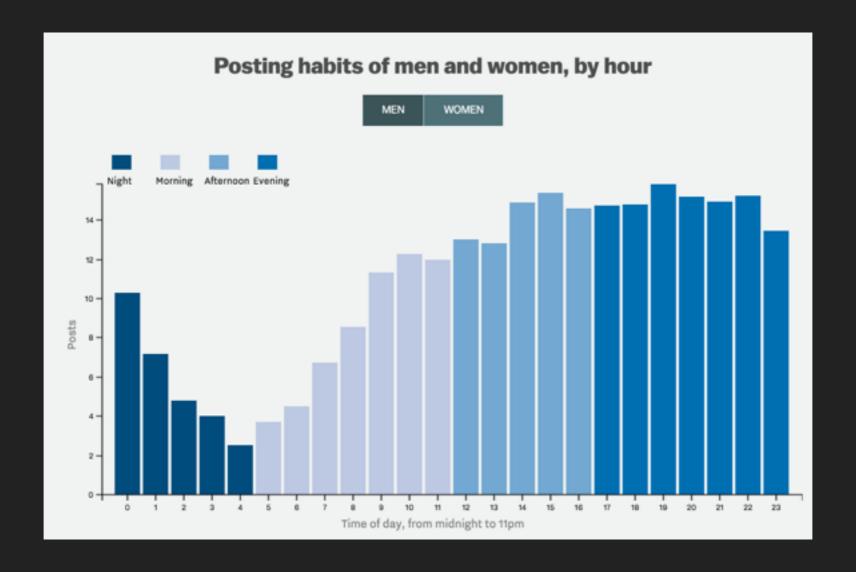
DATA POINTS

- Who was looking for whom
- Age of poster
- Time of posting
- City
- Date
- Length + contents of message

 NYC has the most posts, but Dallas has the highest posts per capita



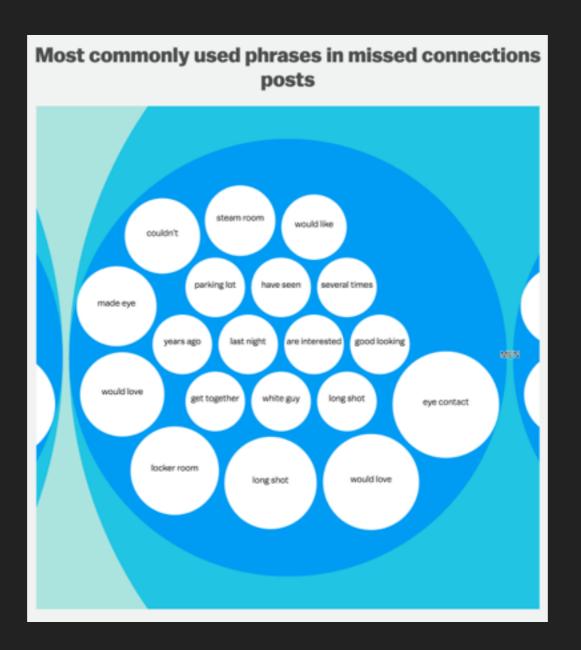
Men post all day, women post after work



Men post much, much, much more than women.



Everyone uses the term "eye contact"



VIZ - INDIVIDUAL CHARTS

- How to visualize each part of the story?
 - ▶ This is hard.
 - ▶ Even Mike Bostock, the creator of D3, wrestles with this.
- http://www.datavizcatalogue.com/
- http://bl.ocksplorer.org/
- http://christopheviau.com/d3list/

EXAMPLE

- How to visualize pervasive differences in
 - 1. Age
 - 2. Individual post length
 - 3. Number of posts
 - 4. Missed connection poster sex
 - 5. Missed connection subject sex
 - 6. City

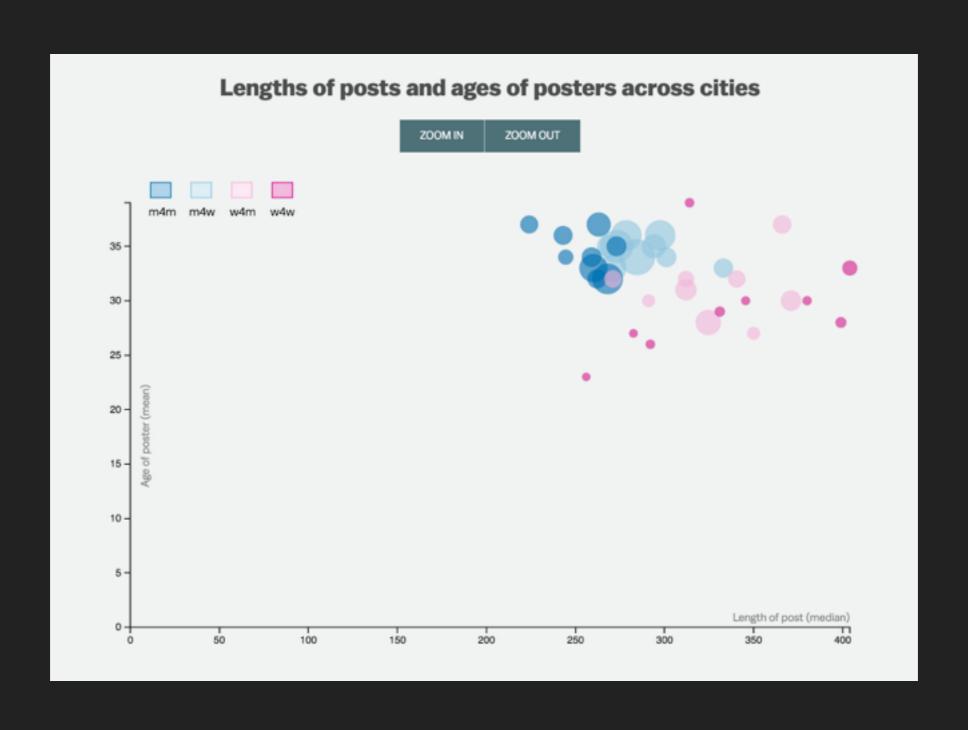
UHHHH



SCATTERPLOTS ARE AMAZING - LOTS OF VISUAL CHANNELS

- X axis
- Y axis
- Circle radius
- Circle color
- Interactive aspects

SCATTERPLOT



MAKING CIRCLES

SCATTERPLOT DATA - JOINED TO VISUAL ELEMENTS

	city	seeker	Median_body_length	mean_age	median_topic_lengt	count
1	Chicago	m4m	260	33	25	579
2	Chicago	m4w	273	35	22	800
3	Chicago	w4m	312	31	22	311
4	Chicago	w4w	331	29	23	36
5	Dallas	m4m	263	37	22	415
6	Dallas	m4w	278.5	36	22	693
7	Dallas	w4m	371	30	19	279
8	Dallas	w4w	345.5	30	23	15
9	Houston	m4m	273	35	22	265
10	Houston	m4w	294	35	23	410
11	Houston	w4m	312	32	22	158
12	Houston	w4w	292	26	19	22
13	Los Angeles	m4m	262	32	26	236
14	Los Angeles	m4w	333	33	29	244
15	Los Angeles	w4m	350	27	30	86
16	Los Angeles	w4w	256	23	23	7
17	New York	m4m	268	32	27	701
18	New York	m4w	284.5	34	27	984
19	New York	w4m	324.5	28	27	463
20	New York	w4w	404	33	23	131

ZOOMING IN/ZOOMING OUT: SCALES ARE MAGIC

REMOVING AND ADDING CIRCLES

```
//Add 4 g groups for each of the 4 seekers types
var groupM4W=svg.append("g")
                                .attr("id", "groupM4W");
var groupM4M=svg.append("g")
                                .attr("id", "groupM4M");
var groupW4M=svg.append("g")
                                .attr("id", "groupW4M");
var groupW4W=svg.append("g")
                                .attr("id", "groupW4W");
//Creating selections of circles to append into each of the 4 seeker groups
var rmCircM4W= svg.selectAll("#m4w").remove();
var rmCircM4M= svg.selectAll("#m4m").remove();
var rmCircW4M= svg.selectAll("#w4m").remove();
var rmCircW4W= svg.selectAll("#w4w").remove();
//Removing the 4 groups from DOM and adding it back within the g group.
rmCircM4W.each(function(){
        var that = this;
        groupM4W.append(function(){ return that;});
});
rmCircM4M.each(function(){
        var that = this;
        groupM4M.append(function(){ return that;});
});
rmCircW4M.each(function(){
        var that = this;
        groupW4M.append(function(){ return that;});
});
rmCircW4W.each(function(){
        var that = this;
        groupW4W.append(function(){ return that;});
});
//Adding capacity to toggle categories of seekers on and off
d3.select("rect#m4m").on("click", function(){
        var active = groupM4M.active ? false:true,
        newOpacity = active ? 0:1;
        d3.select("#groupM4M").transition().duration(200).style("opacity", newOpacity);
        groupM4M.active = active;
});
```

DISSECTING THE NARRATIVE: INTRO

- Fact: I ran a study
- Narrative: Missed Connections as a cultural phenomenon

Vox





I analyzed 10,000 Craigslist missed connections. Here's what I learned.

By Ilia Blinderman on November 12, 2015

As much as I've always wanted to be the subject of someone's missed connection, I doubt I'm suited to it. Aboard public transport, my go-to facial expression is a stern, furrowed brow. "Solving the crisis in the Middle East, no-doubt," my fellow passengers probably sympathize, contemplating the task of fixing the world's thorniest geopolitical dilemma from aboard the subway and letting me be.

Not that any of this has stopped me from posting missed connections myself. I can't recall what I wrote, but after some virtual digging I found the subject lines of my posts buried years deep within my inbox:

"GIRL IN BLUE/ WHITE FLORAL DRESS ON N TRAIN HEADED UPTOWN AT 9PM MONDAY" -M4W (MISSED CONNECTION)

"CUTE BRUNETTE GIRL READING WHO SHARED MY TABLE AT THE BEAN ON MONDAY" - M4W (MISSED CONNECTION)

That there exists a digital town square where lonely hearts can declare their feelings without fear of public rejection is both lucky and improbable, but the hit rate, by all accounts, is low. I've yet to hear any firsthand stories of missed connections that have resulted in anything more than a date or two before the romance petered out. Still, if it seems strange that a quirky section of a website that prides itself on an aggressively dial-up-era design has gained such traction in popular culture - all in spite of the scarce likelihood of finding love - look no further than the motivations of gold miners or oil prospectors. Each successive romantic relationship is a failure until it isn't, and the lousy odds of forging a real connection don't have much impact on our inborn optimism.

It may have been my own failures to connect that spurred me to take a closer look at the habits and behaviors of other posters. It also could have been the odd voyeuristic appeal of the whole missed connections section. putting those private and vulnerable declarations of affection into a ruthlessly public, yet anonymous, context. Finally, I suspect it was also rooted in my old psych grad school mentality, making the promise of a large data set, untainted by the specter of observer effects, too tempting to ignore. Who were these people who posted hundreds of messages each day? To see, I gathered the missed connection postings from the nine largest US cities and got to work.

connections from New York, LA, Chicago, Houston, Philadelphia, Phoenix, San Antonio, San Diego, and Dallas. I analyzed the language, the people who looked and whom they looked for, the days they posted, the words they used, their ages, and a dozen of other points of comparison. And so, without

DISSECTING THE NARRATIVE: RATES OF POSTING

- Fact: Missed connections rates, by city
- Narrative: Discussion of solitude and loneliness common to large cities

New York seems to be a city manufactured for missed opportunities to meet strangers. There is the pervasive reach of the subway, the relatively scarce number of drivers, the never-ending throngs of people. There is, furthermore, New York's take on big-city ioneliness that so many newcomers discover: long days at work, bookended by lengthy commutes that take New Yorkers to the far reaches of the vast metropolis. You won't be heading to drinks in Brooklyn if you finish work at 6 in Midtown and catch

savvy enough to change from the express to the local at 96th Street. Indeed of all the cities New York had the highest number of missed connections. Initially, this seemed to be all about numbers: Of course the largest city in the country would have the most posts.

Los Angeles, however, dispels that notion (mouse over the circles in the map below to see each city's population and missed connections count). The second largest city in the US, LA sits at the bottom of the tally, with only a few hundred missed connections posted during January. I should note here that if you're anything like me, your internal sleuth rears up, and you begin to postulate that this is because Angelenos drive everywhere and are never

Posts and population for nine US cities in January 2015



To get a sense of the true likelihood of landing a missed connection in the cities I've mentioned, I worked out the number of postings for every 10,000 inhabitants. The news for lonely New Yorkers isn't especially heartening: New York City sits on the tail end of the rankings, with just over three missed connections per 10,000 inhabitants; Los Angeles comes in dead last. The cities whose residents are most likely to seek out that alluring stranger are Dallas, Phoenix, and San Diego, but even with Dallas's first-place ranking of

once pronounced, "She was one in a million, so there's five more just in New South Wales." Against that backdrop, your chances of a missed connection are orders of magnitude more promising.

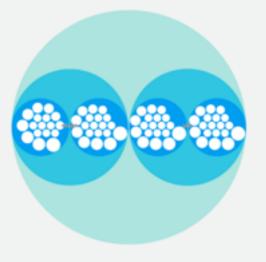
DISSECTING THE NARRATIVE: LANGUAGE USED

- Fact: Most common words
- Narrative: Discussion of the language people tend to use in online personal ads.

In an exceptionally insightful New Yorker piece, Nick Paumgarten, discussing online dating profiles, remarks that "demonstrating the ability, and the inclination, to write well is a rough equivalent to showing up in a black Mercedes" — partly, he says, because "males know that the best way to get laid is to send messages to as many fernales as possible. To be efficient, they put very little work into each message and therefore pay scant attention to each woman's profile." By this logic, I expected people wanting to impress a potential crush to put some hard yards into their messages; no Great American Novels, to be sure, but something witty, beasing, playful; in other words, the verbal equivalent of a black Mercedes. If

connections section, you are, by now, smirking at my naiveté. Much like in real life, the board is populated by a mixture of occasional gems filled with earnest feeling and self-reflection, a mass of posts whose allure ranges from to to on the color scale, and a small but impassioned band of people who seek to reconnect with former neighbors from years back in hopes of finding a foot mistress or offering themselves into indentured sexual labor. Star-crossed lovers, these are not. Below, you can find the most commonly phrases for each group (keep in mind that in a category like "women seeking women," which had few posts to begin with, even an infrequently occurring set of words may get flagged if used, by chance, on a couple of occasions). The larger, light blue circles represent the men and women who post the connections; each contains two darker circles that represent whom they're looking for. The white innermost circles indicate the most commonly used phrases.

Most commonly used phrases in missed connections posts



RECAP

Narrative = ** **

ADDITIONAL RESOURCES

- Dashing D3 Newsletter
- Scott Murray's Interactive Data Visualization for the Web
- Elijah Meeks' <u>D3.js In Action</u>