

MIS 0855 Spring 2015 – Data Science

Day 34 – Beyond Numbers






Min-Seok Pang

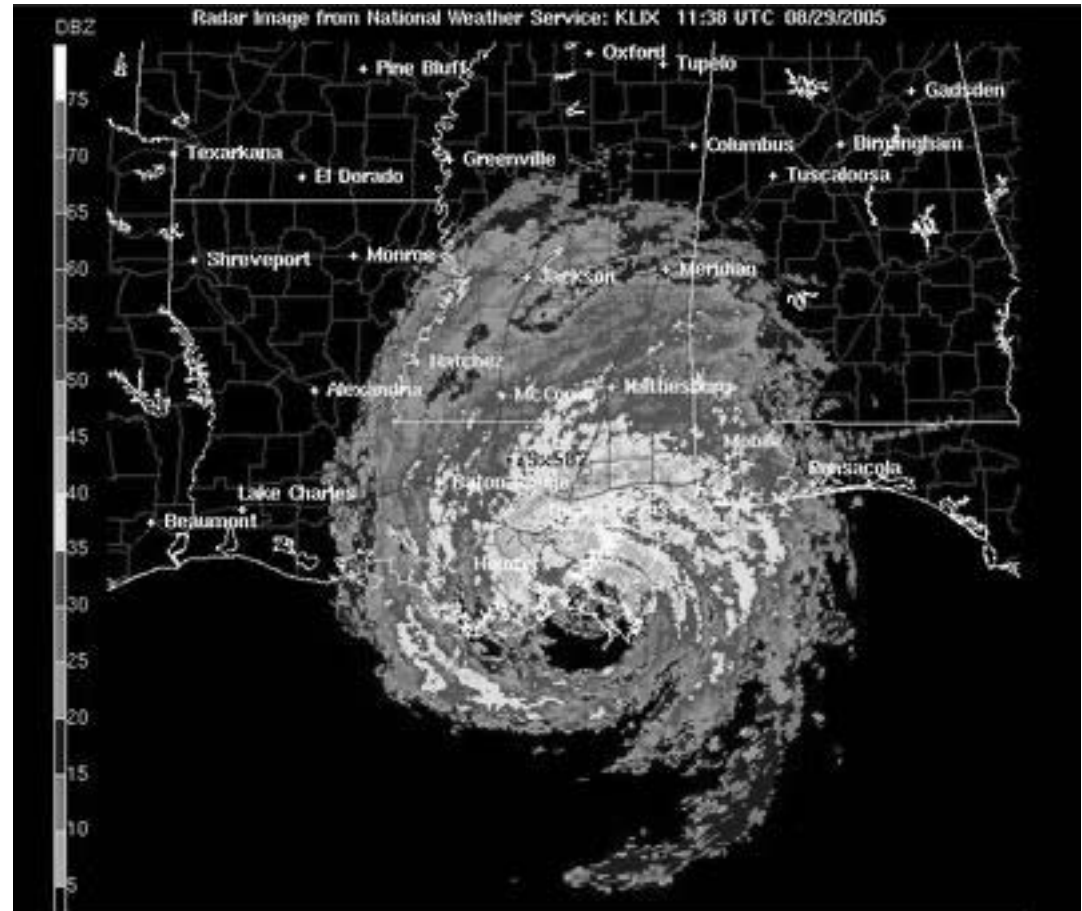
**Management Information Systems
Fox School of Business, Temple University
minspang@temple.edu**

Apr. 8th, 2015

Structured vs. Unstructured Data

My 10 Day Forecast Updated: Oct 11, 2014, 1:34pm EDT Desktop App

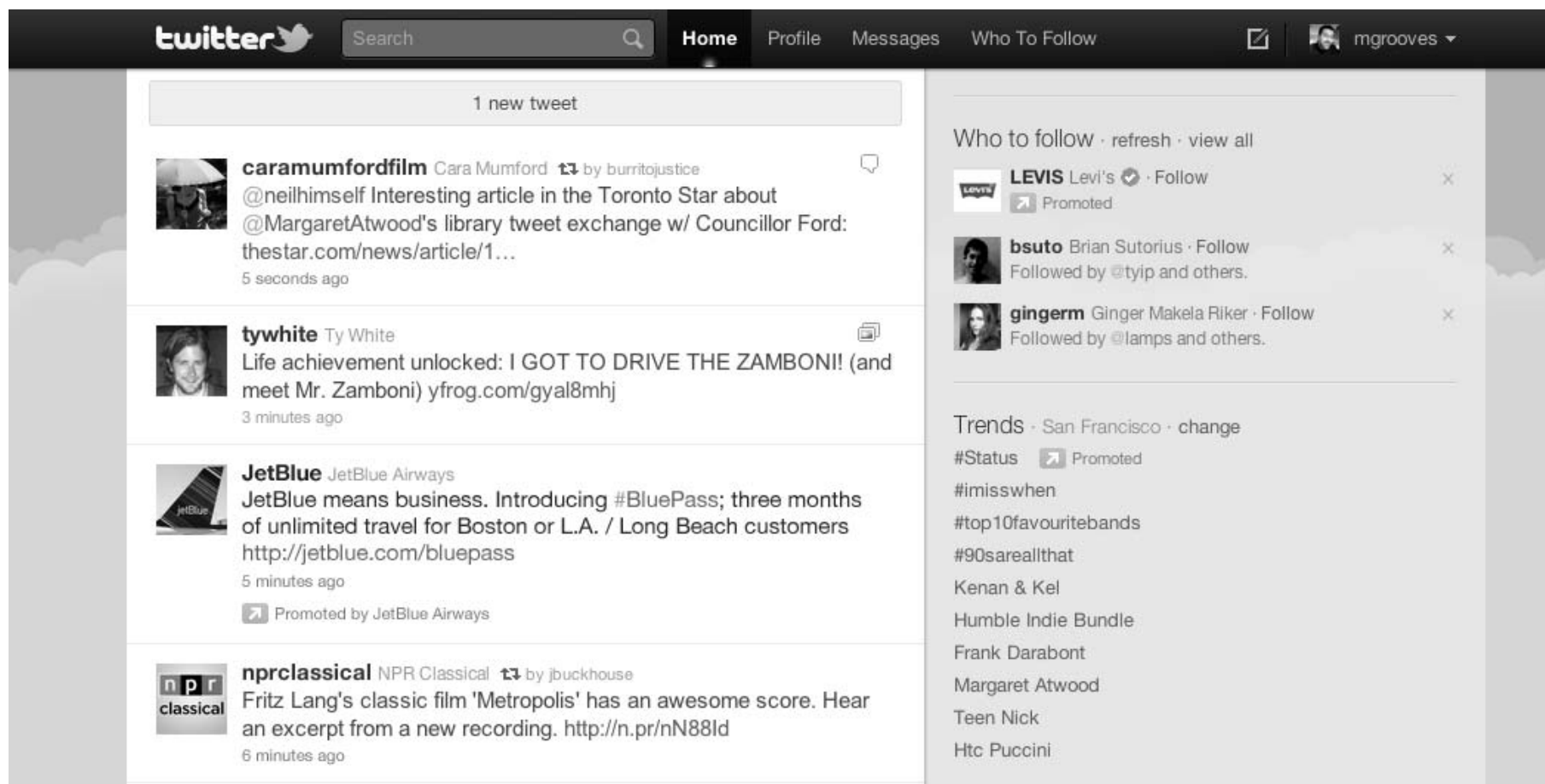
Today Oct 11		57° 47° ^F	CHANCE OF RAIN: 20%	WIND: N at 8 mph	Details
Cloudy					
Sun Oct 12		65° 53°	CHANCE OF RAIN: 0%	WIND: ENE at 7 mph	Details
Mostly Sunny					
Mon Oct 13		72° 64°	CHANCE OF RAIN: 50%	WIND: SSE at 9 mph	Details
Showers					
Tue Oct 14		77° 65°	CHANCE OF RAIN: 10%	WIND: SSE at 13 mph	Details
Partly Cloudy					
Wed Oct 15		73° 56°	CHANCE OF RAIN: 70%	WIND: SSE at 14 mph	Details
Light Rain					
Thu Oct 16		67° 54°	CHANCE OF RAIN: 20%	WIND: WSW at 11 mph	Details
Partly Cloudy					



● Other examples of unstructured data?

Twitter Tweets

- What is the structure data and what is unstructured?



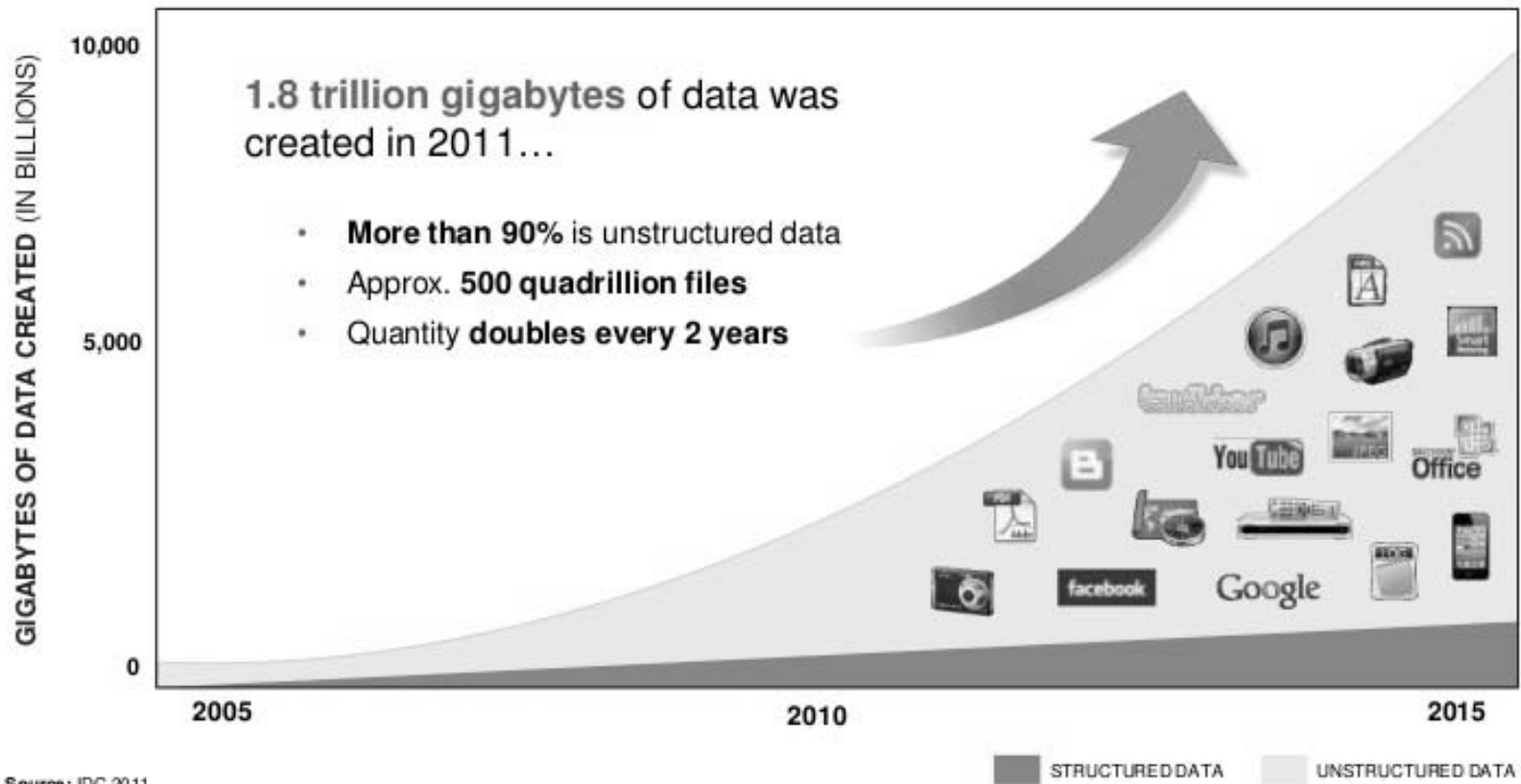
The screenshot shows the Twitter interface with a search bar, navigation tabs (Home, Profile, Messages, Who To Follow), and a user profile (mgrooves). The main feed displays four tweets:

- caramumfordfilm** (Cara Mumford) retweeted by burritojustice: "@neilhimsself Interesting article in the Toronto Star about @MargaretAtwood's library tweet exchange w/ Councillor Ford: thestar.com/news/article/1..." (5 seconds ago)
- tywhite** (Ty White): "Life achievement unlocked: I GOT TO DRIVE THE ZAMBONI! (and meet Mr. Zamboni) yfrog.com/gyal8mhj" (3 minutes ago)
- JetBlue** (JetBlue Airways): "JetBlue means business. Introducing #BluePass; three months of unlimited travel for Boston or L.A. / Long Beach customers http://jetblue.com/bluepass" (5 minutes ago, promoted by JetBlue Airways)
- nprclassical** (NPR Classical) retweeted by j buckhouse: "Fritz Lang's classic film 'Metropolis' has an awesome score. Hear an excerpt from a new recording. http://n.pr/nN88ld" (6 minutes ago)

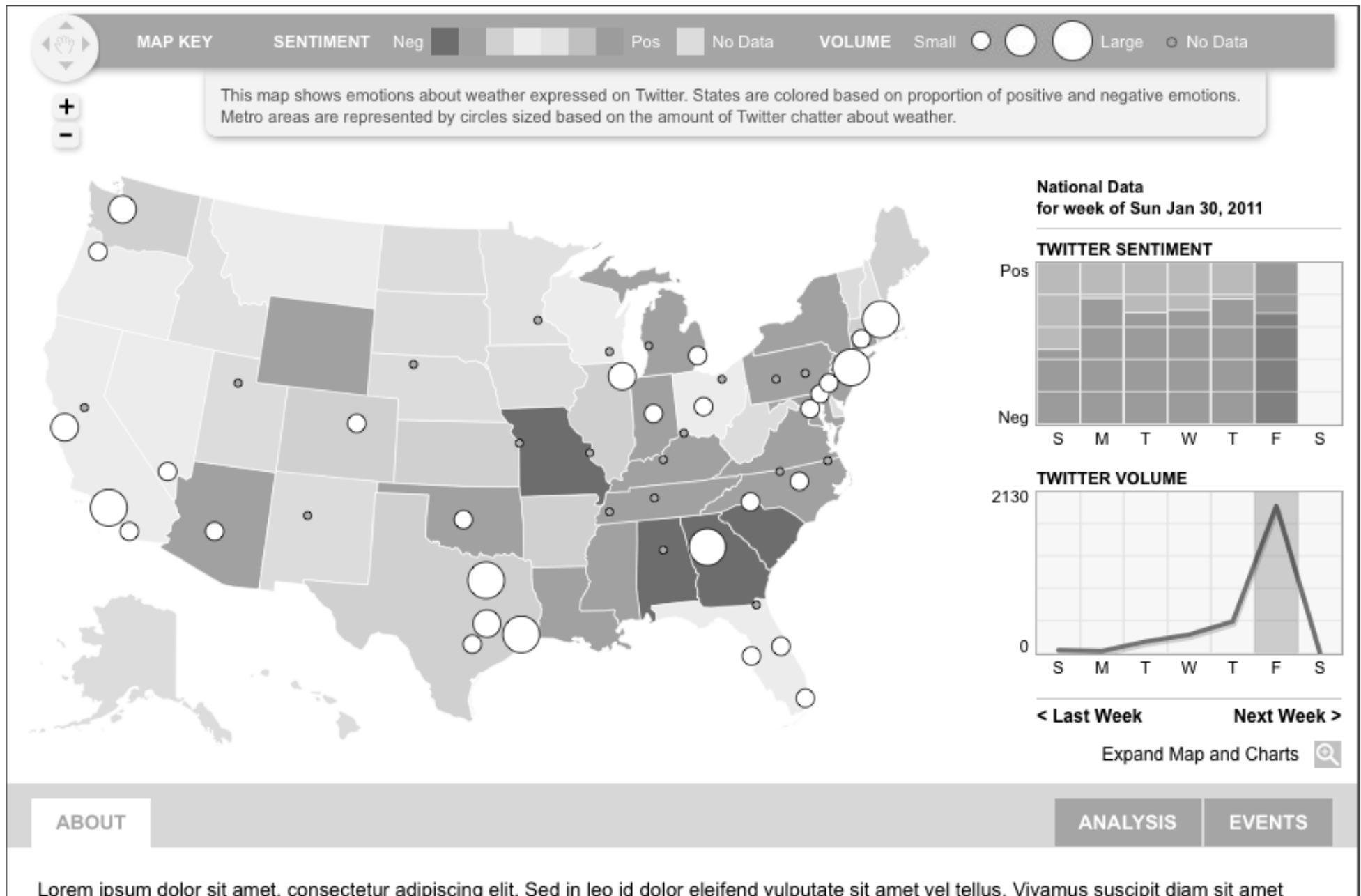
The right sidebar features "Who to follow" (refresh, view all) with users LEVIS (Levi's, promoted), bsuto (Brian Sutorius), and gingerm (Ginger Makela Riker). Below is a "Trends" section for San Francisco, listing #Status (promoted), #imisswhen, #top10favouritebands, #90sareallthat, Kenan & Kel, Humble Indie Bundle, Frank Darabont, Margaret Atwood, Teen Nick, and Htc Puccini.

<http://www.cnet.com/news/twitter-starts-injecting-paid-tweets-into-the-stream/>

Explosive Data Growth



Twitter Sentiment Analysis



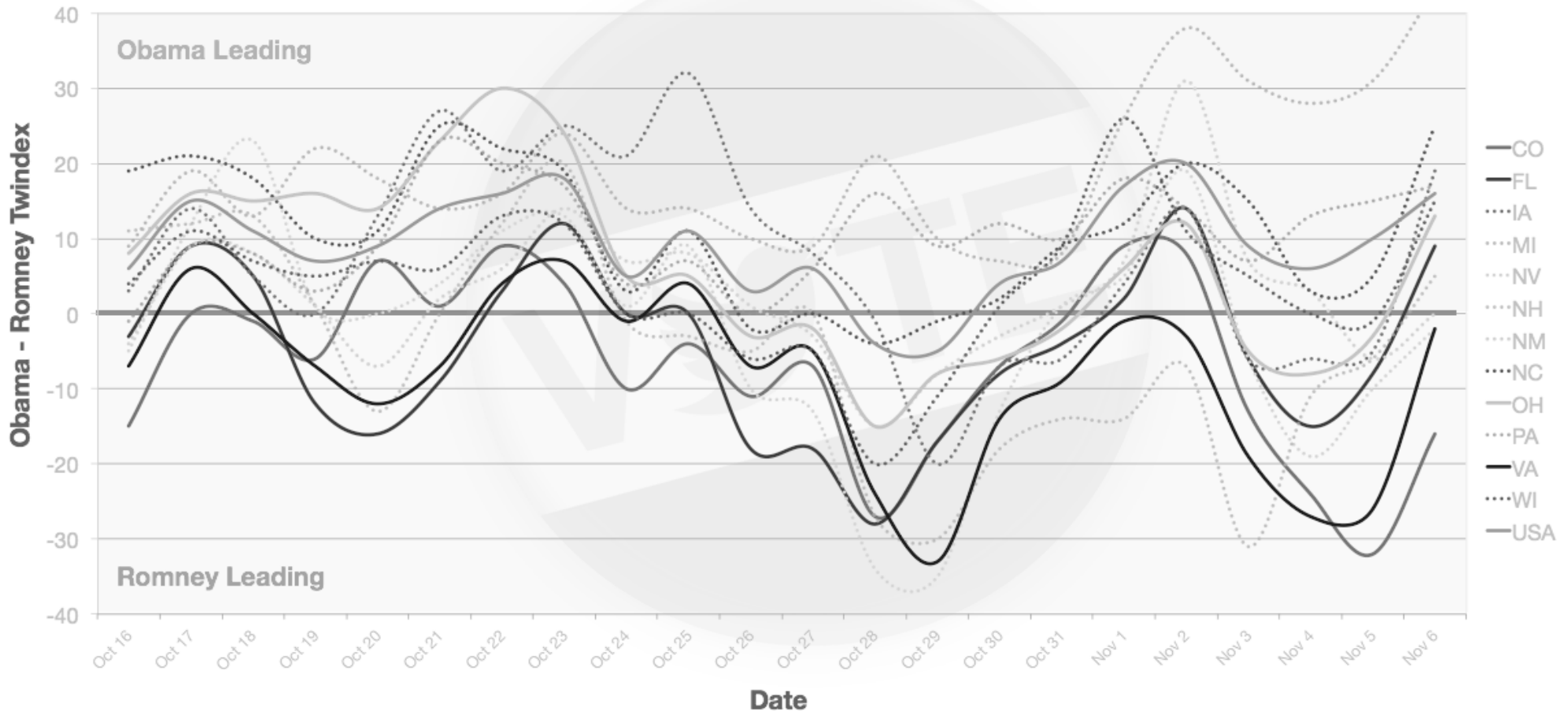
<http://www.dialogueearth.org/2011/05/13/hope-for-human-sentiment-analysis-coding/>

Twitter Political Index: The Swing States

Net difference in @BarackObama vs. @MittRomney #Twindex scores in @USAToday- and @Gallup-defined “swing states.”
October 16, 2012 to November 6, 2012 (partial day)

Follow @gov for more about government & politics on Twitter.

ANALYTICS BY **TOPSY**



<http://mashable.com/2012/11/06/twitter-sentiment-romney-obama/>

Yelp Reviews for Poi Dog Snack Shop

“Positive” Word Library



Gabby M.
Philadelphia, PA
17 friends
88 reviews

★★★★★ 7/11/2014

Sooo good! The pork is amazingly flavorful and tender! The staff is very friendly. I wish they came around University City more often. Cannot wait to have it again!



Sean P.
Philadelphia, PA
25 friends
12 reviews

★★★★★ 10/10/2013

By far my favorite food truck at Temple.

I've eaten here numerous times. They operate an odd schedule, moving to different locations out the city. I haven't pinned down exactly which days they're on campus, but I crack a huge smile every time I see their little cart.

I've never had a bad meal at Poi Dog. The tacos are superb, katsu chicken quite tasty, and even the spam musabi, which I was very hesitant about, was a wonderful surprise. The mochi nori chicken is among the best meals that I've ever had from a food truck. It's crunchy and sweet with nuanced spices that I've yet to pin down.

Great

Fantastic

Best

Wonderful

Outstanding

Amazing

Share review

Compliment

Send message

Follow Sean P.

A sentiment analysis is not an easy task.

Your class was **terrific**

Your class was **terrifically bad**

Your class was **awfully good**

I'd **like** to tell you that your
class was **great**, but I **can't**!

How we'll do a sentiment analysis



Retrieve Tweets
using a spreadsheet
add-in in Google
Drive

Copy the tweets
from Google
Drive to a special
Excel workbook

Excel classifies
the Tweets as
positive or
negative