

data•visualization

at Twitter

data
at Twitter

“Tweets”

#events

TV Shows New Year
Earthquake Protest
Oscars Super Bowl
World Cup Election
Breaking news ...

#curiosity

Sleep pattern
Human behavior
Language ...

What could we learn from the Tweets?

data•
at Twitter

“Tweets”

Get data

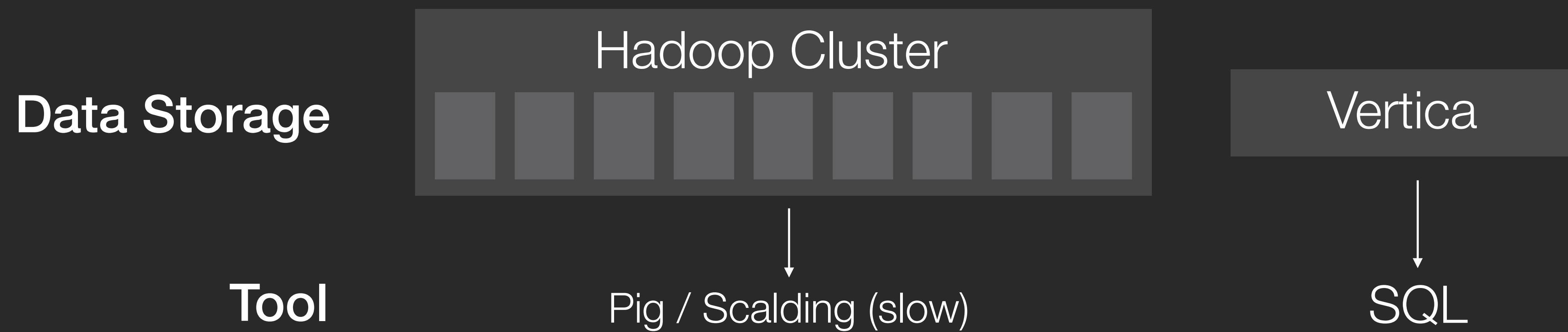
1

•vis

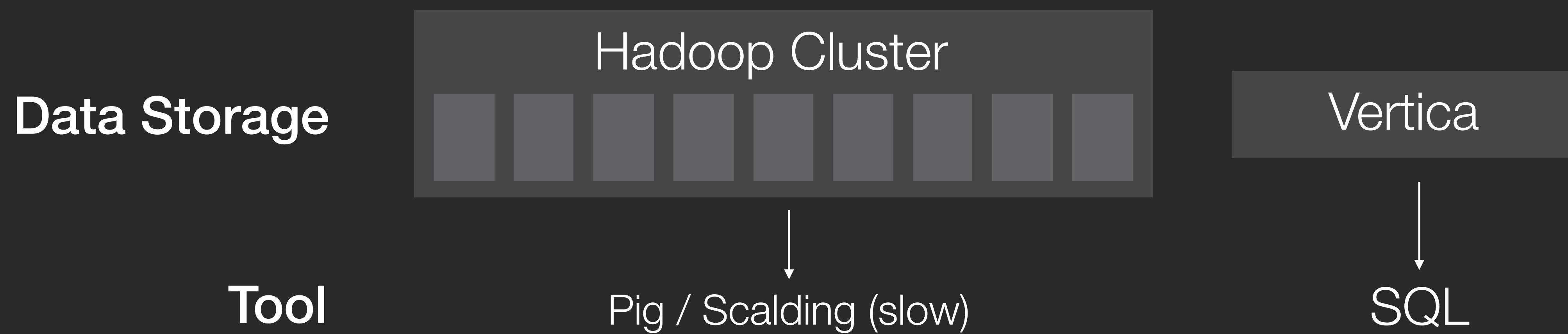
Challenges

- Too much data
 - Want only relevant Tweets
 - hashtag: #BRA
 - keywords: “goal”
 - Need to aggregate & reduce size
 - Long processing time (hours)

Workflow



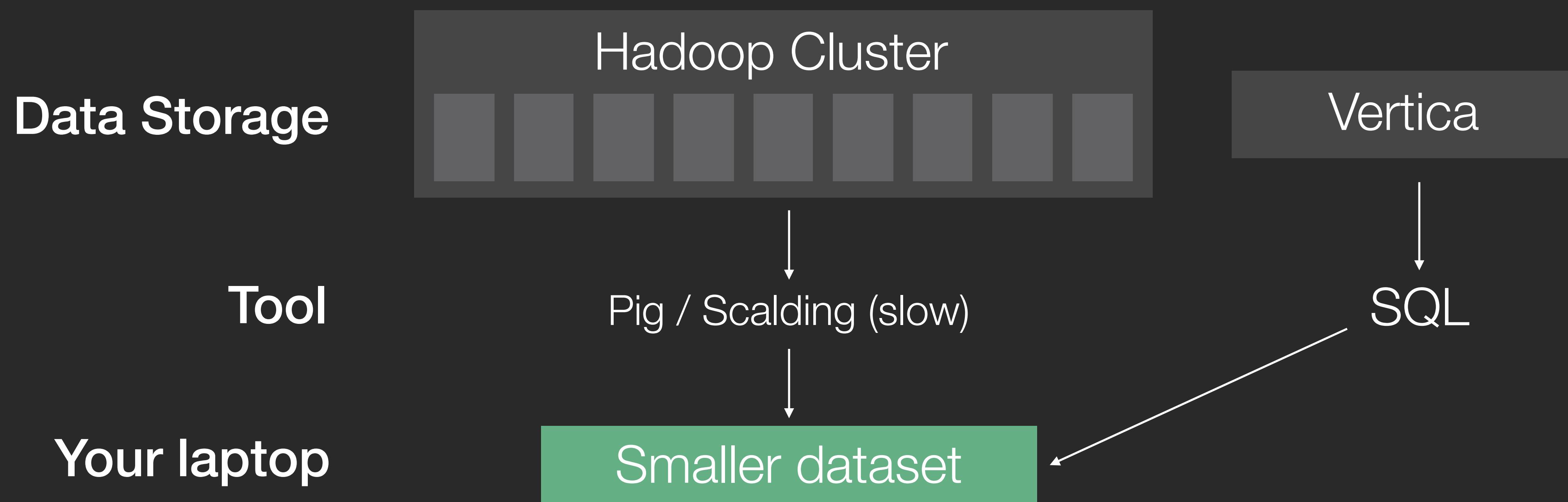
Workflow



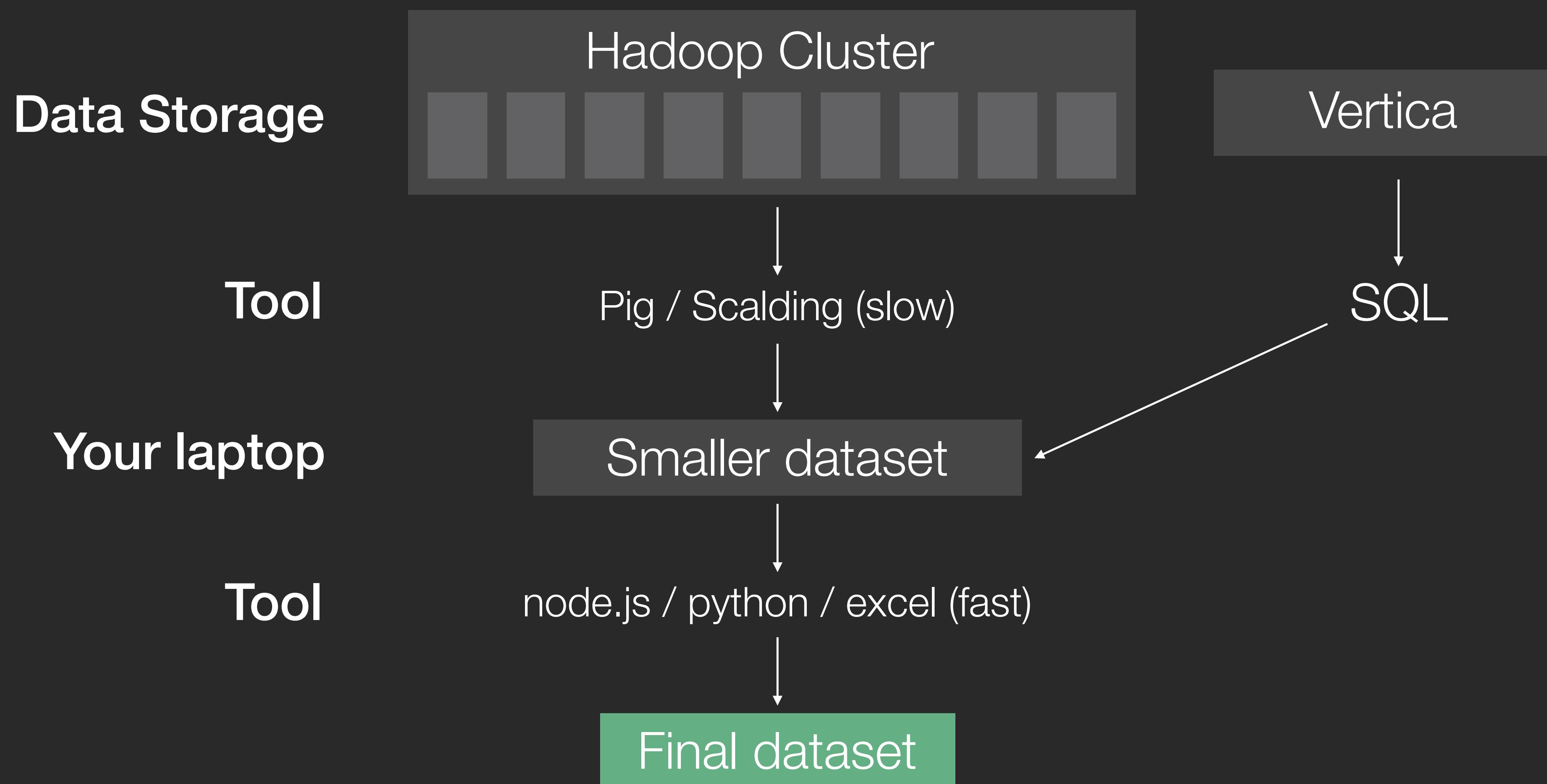
--- PIG

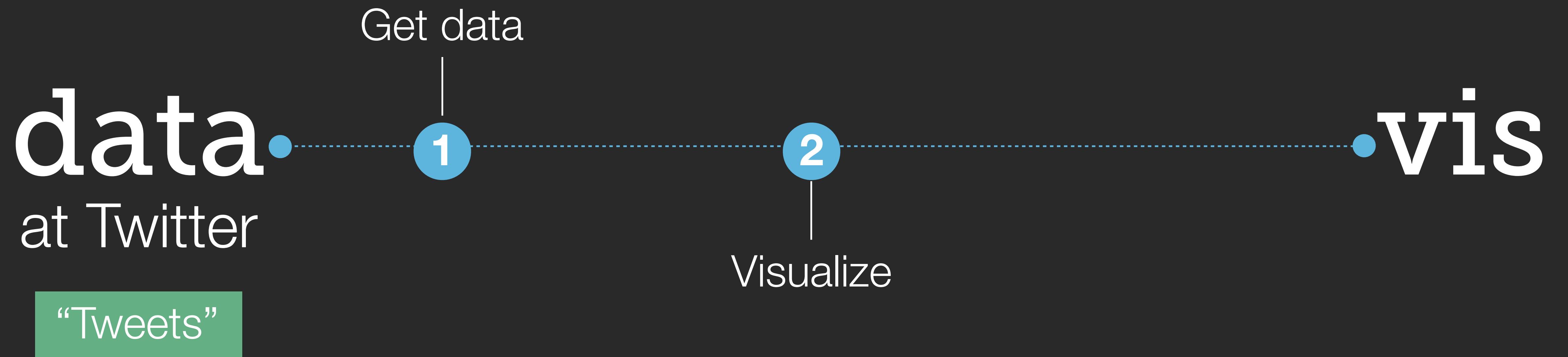
```
tweets = load '/tables/statuses/$DATE' using TwadoopLoader();  
  
filtered_tweets = filter tweets by twtext matches '.*#BRA.*';  
filtered_tweets = foreach filtered_tweets  
    generate '#BRA' as name:chararray, extract_minute(created_at) as minute;  
counted_hashflag = foreach (group filtered_tweets by minute)  
    generate FLATTEN(group) as (name), COUNT(filtered_tweets) as count;
```

Workflow



Workflow

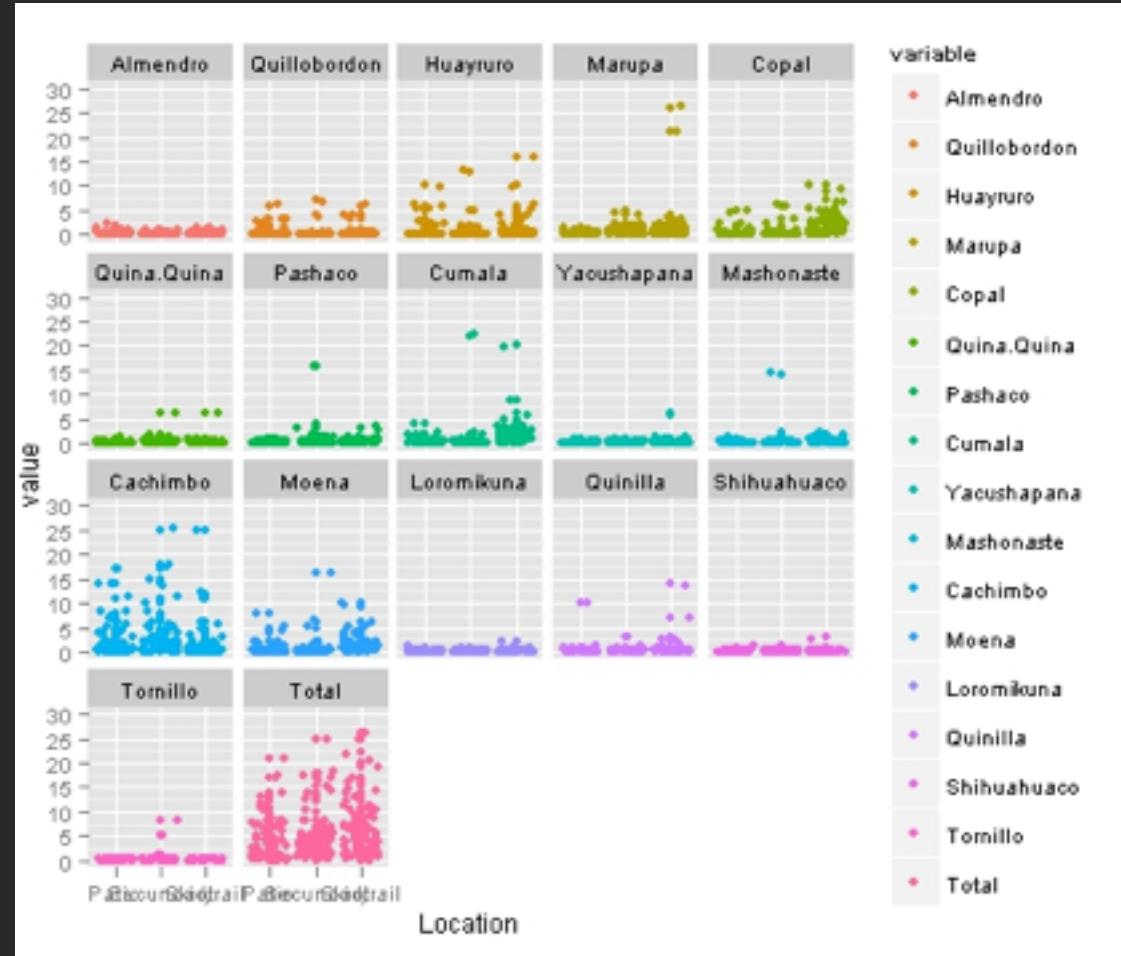




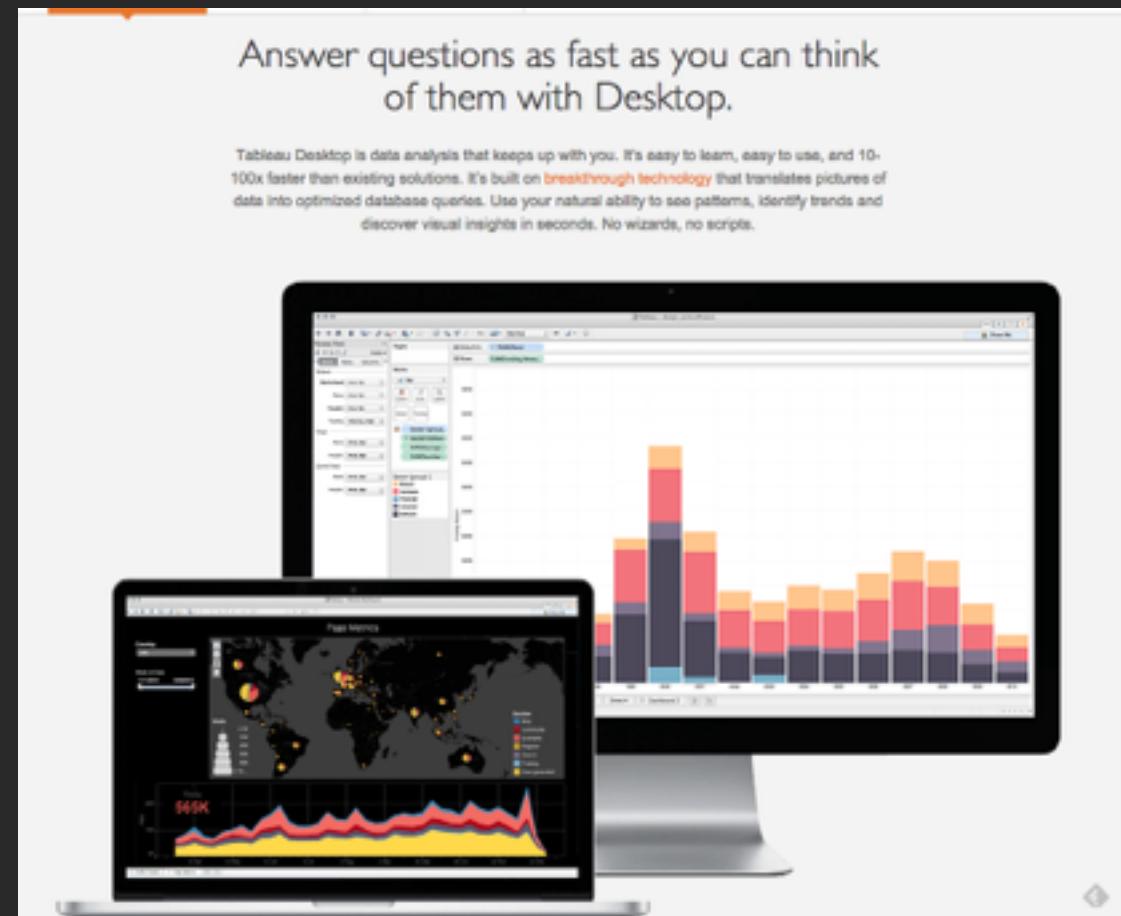
Visualize

- Peek into data
 - Check data & test ideas
- Decide how to visualize
 - Guided by data type
- Choose tools
- Start building

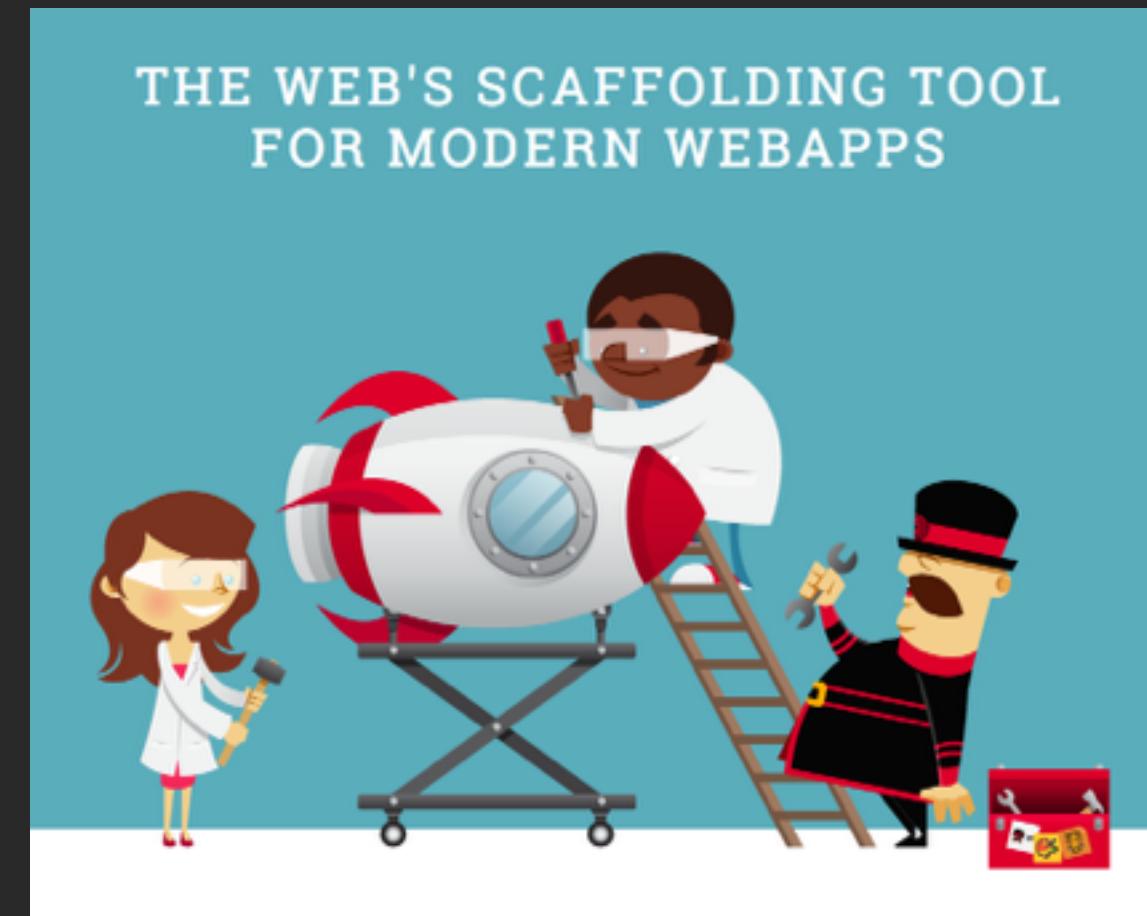
R



d3

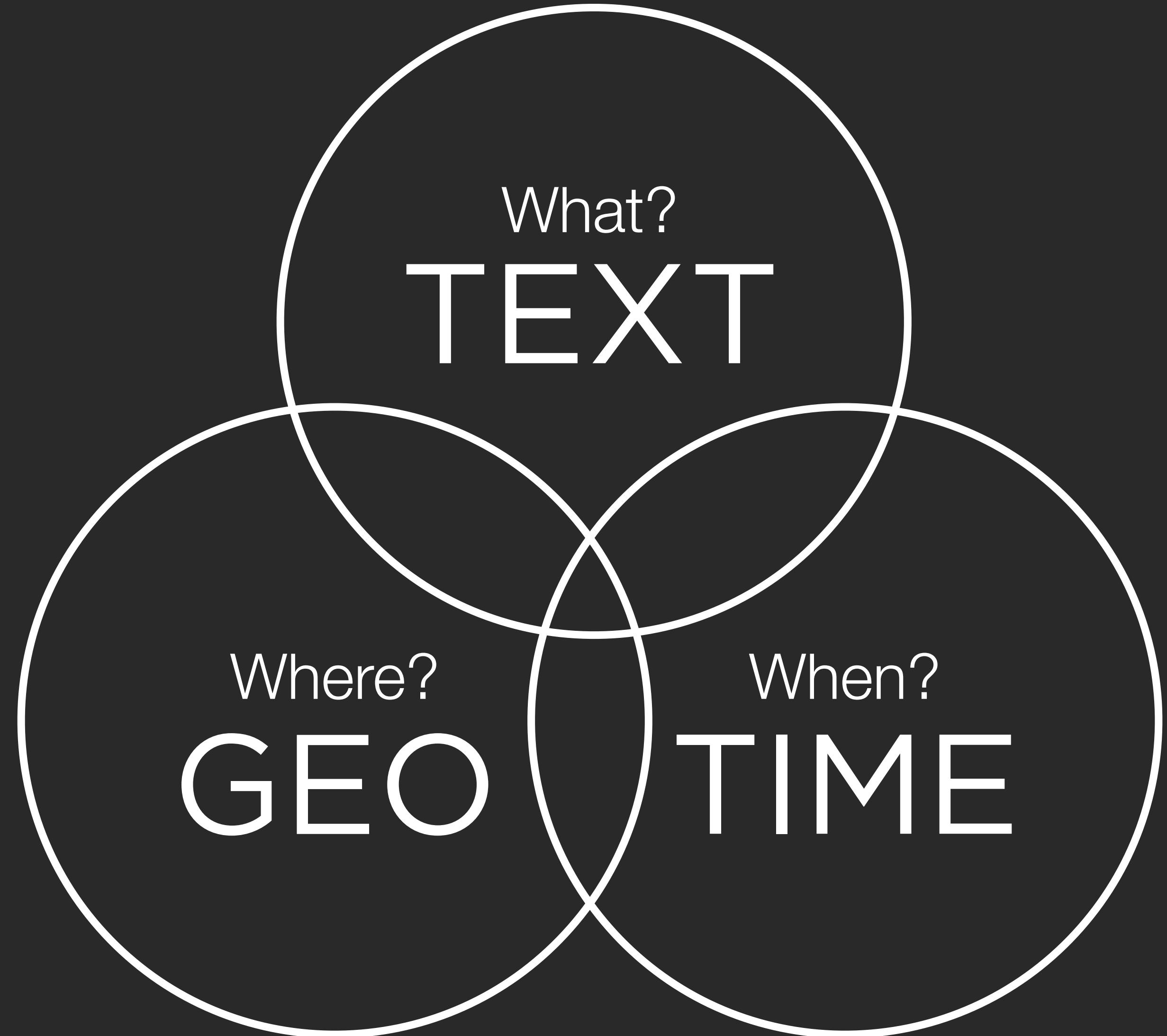


Tableau



Yeoman

Visualize Data



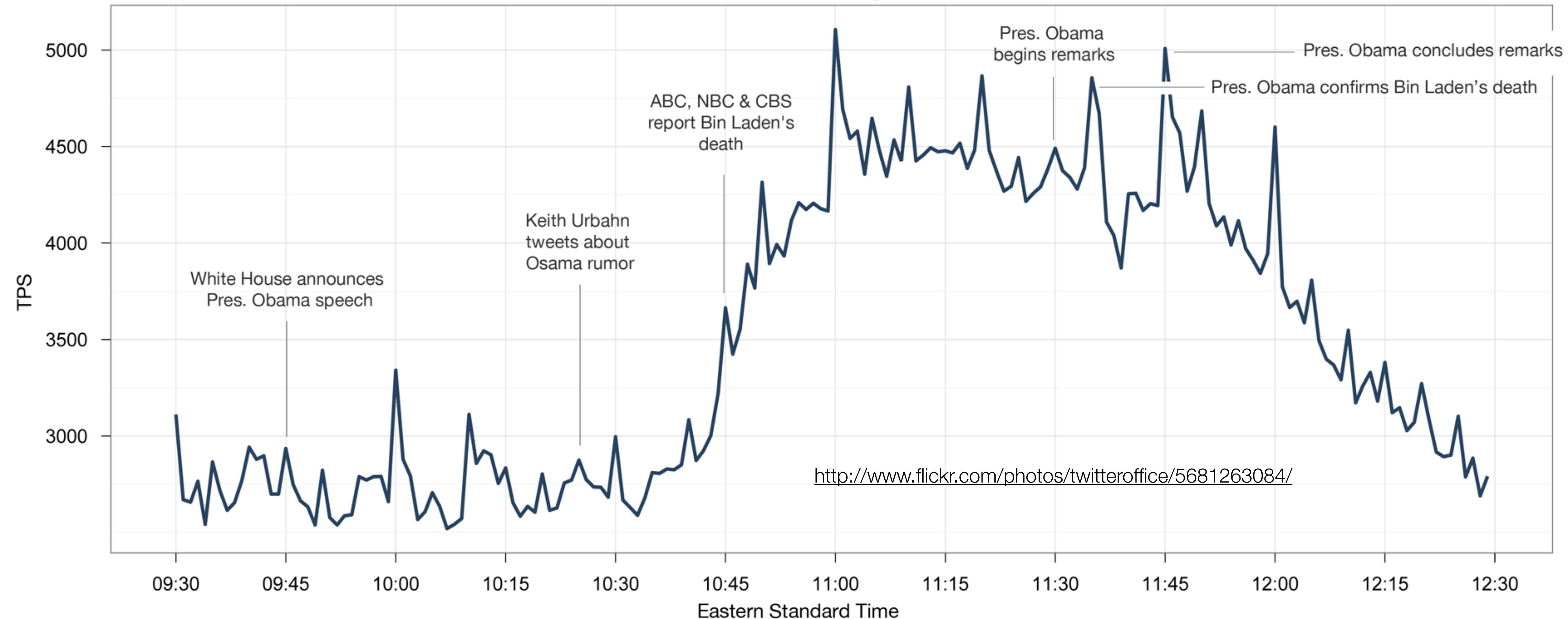
Visualize Data



Time

Tweets/second + Annotation

Tweets Per Second: Evening of May 1, 2011



Visualize Data



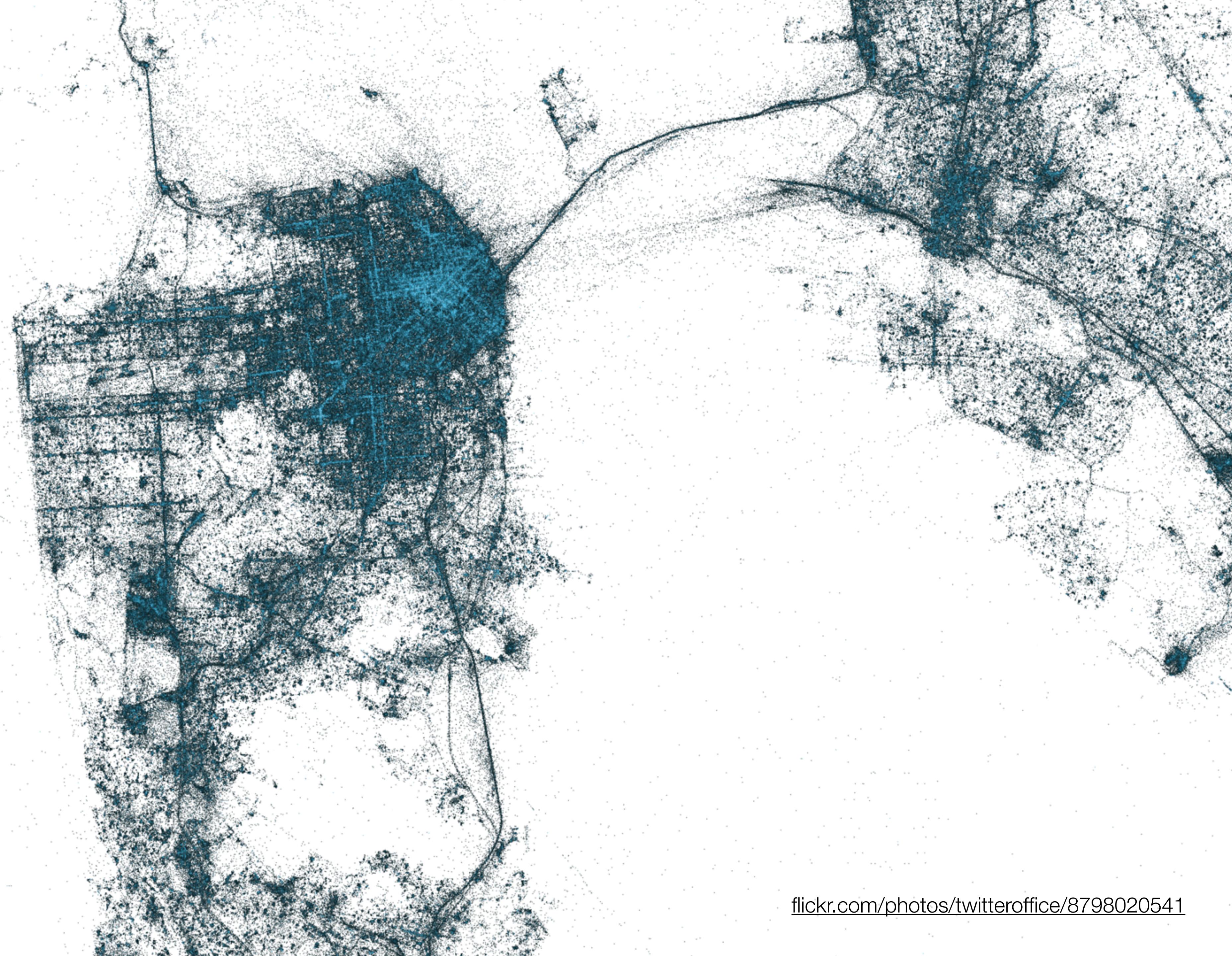
Geo

San Francisco

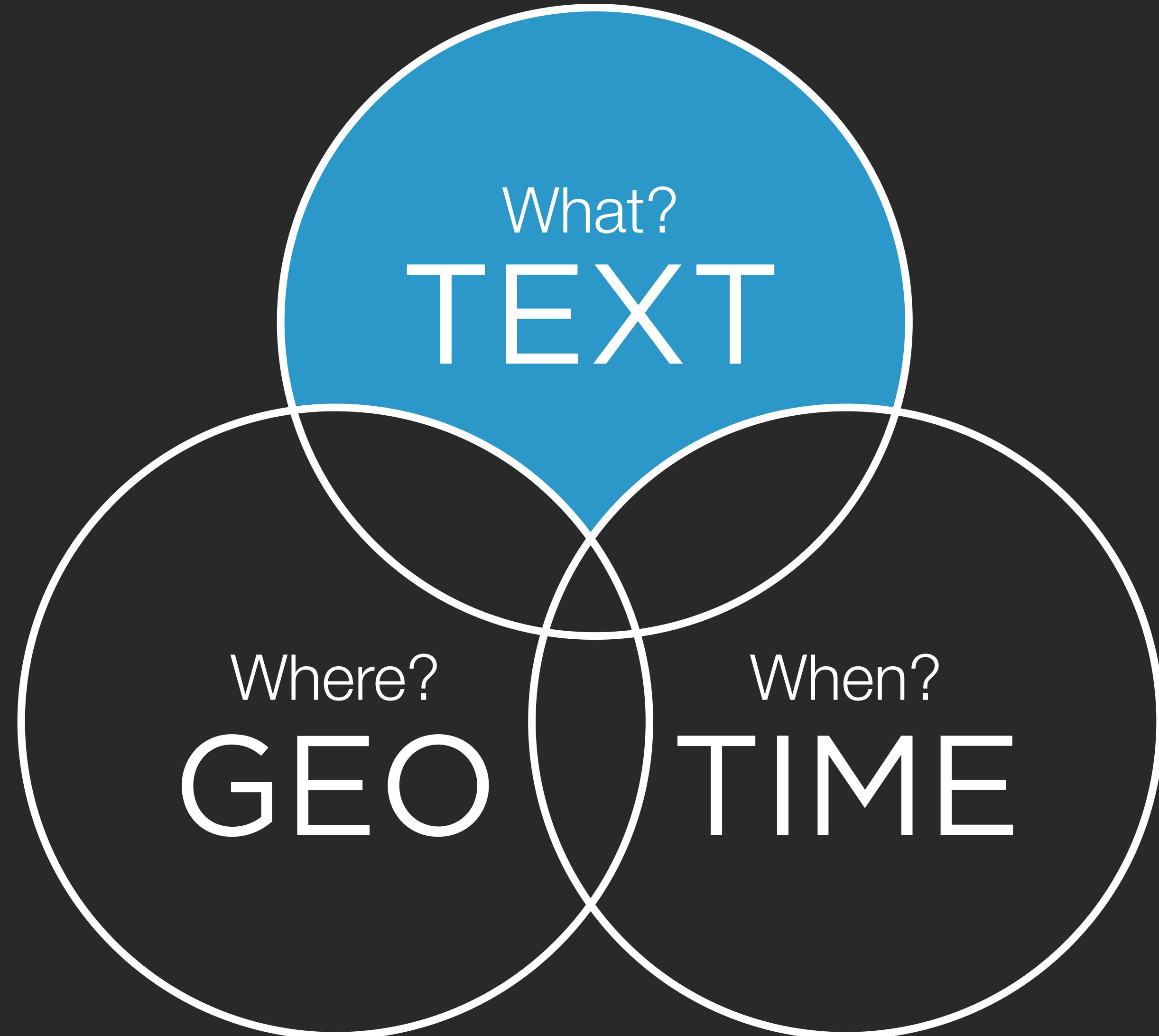
Low density



High density

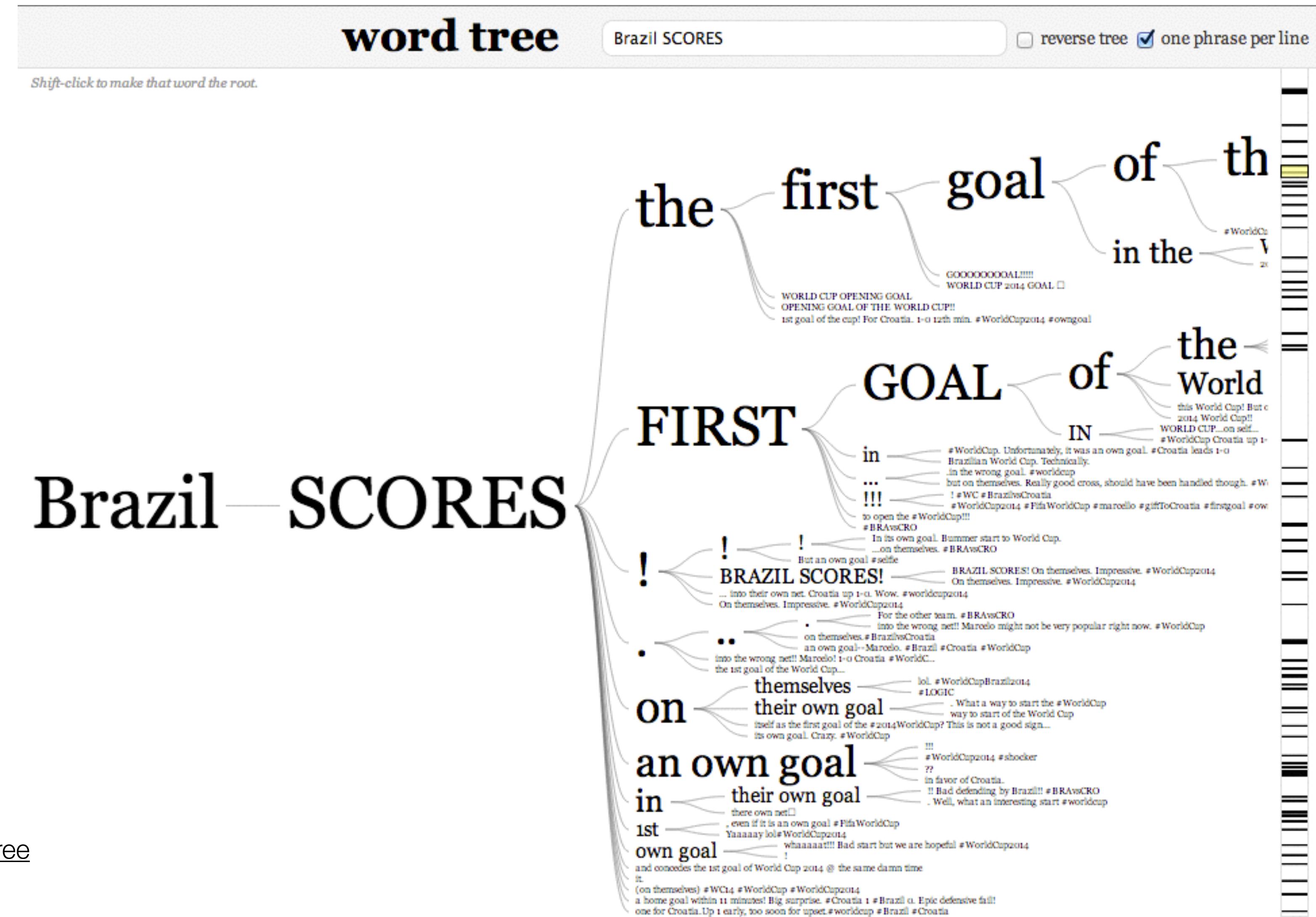


Visualize Data

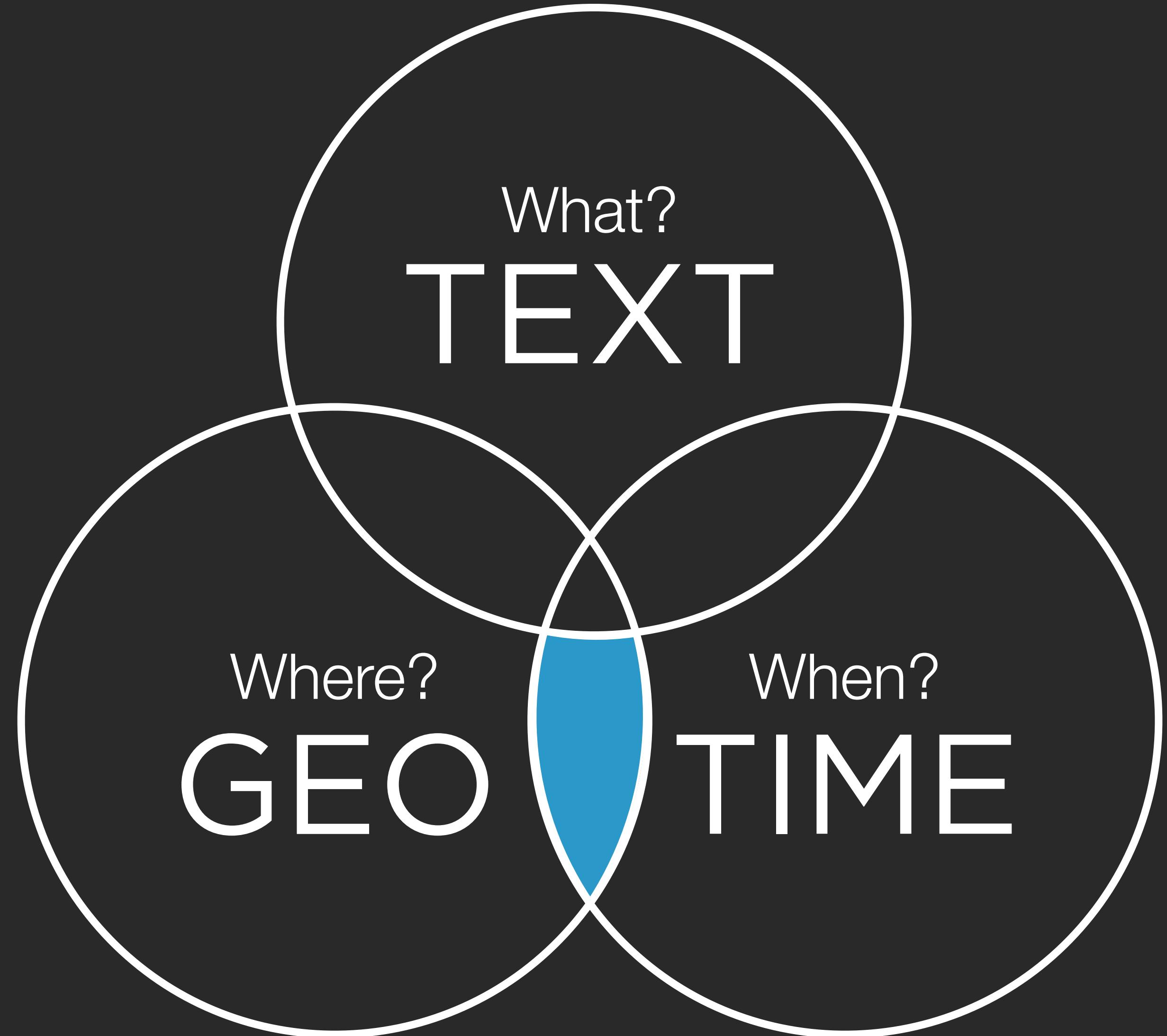


Text

WordTree [Wattenberg & Viégas 2008]

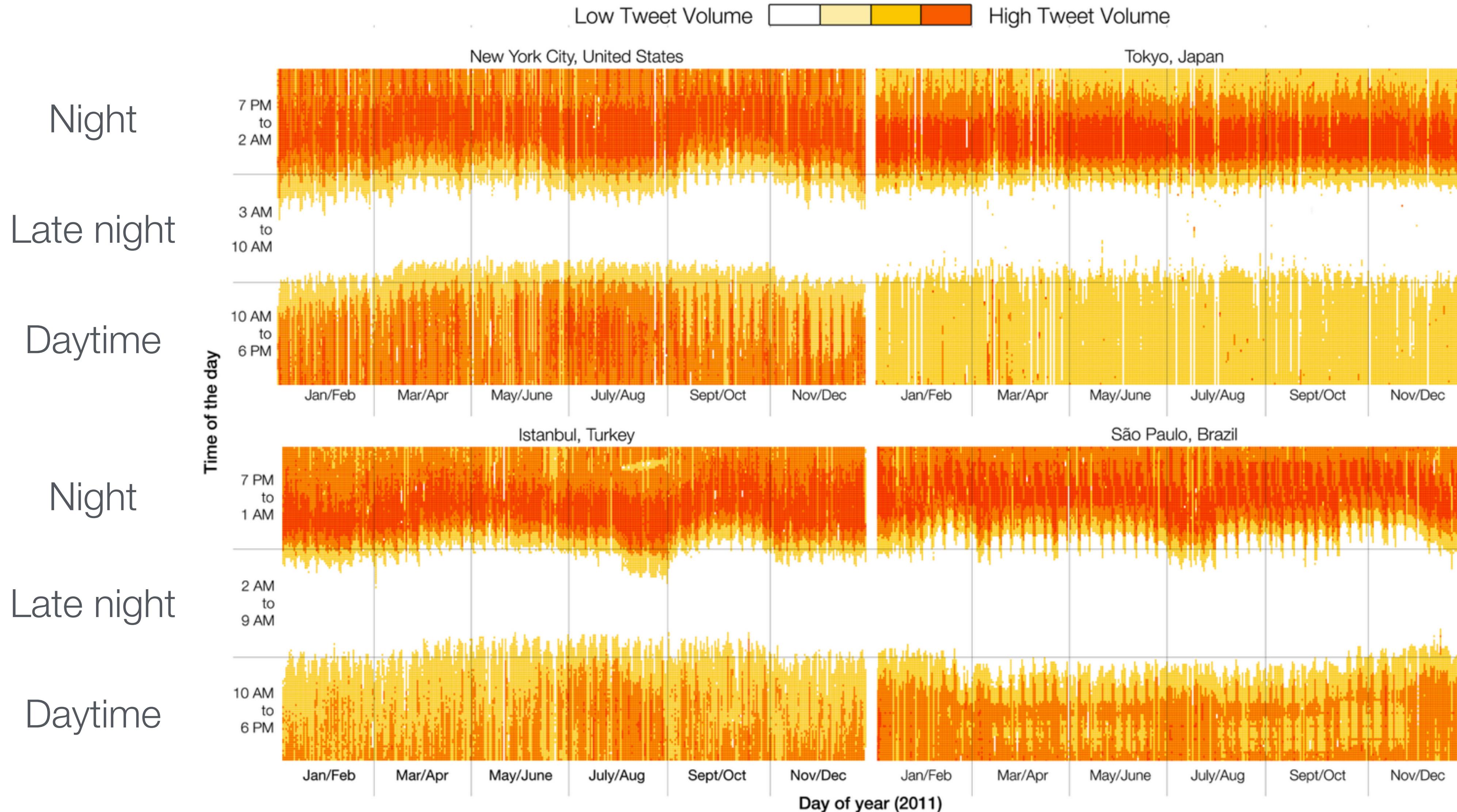


Visualize Data

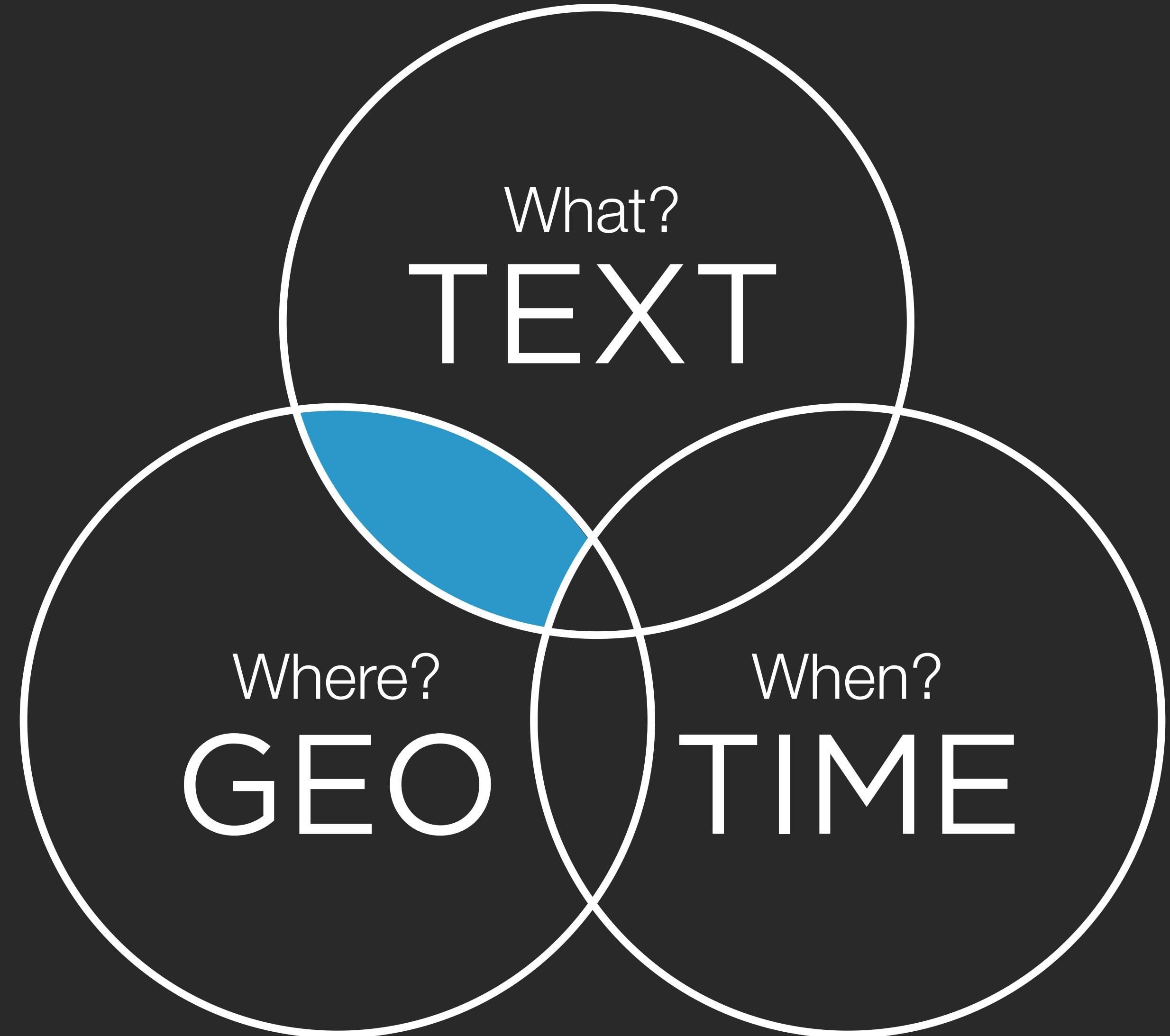


Time + Geo

Tweet pattern [Rios & Lin 2012]

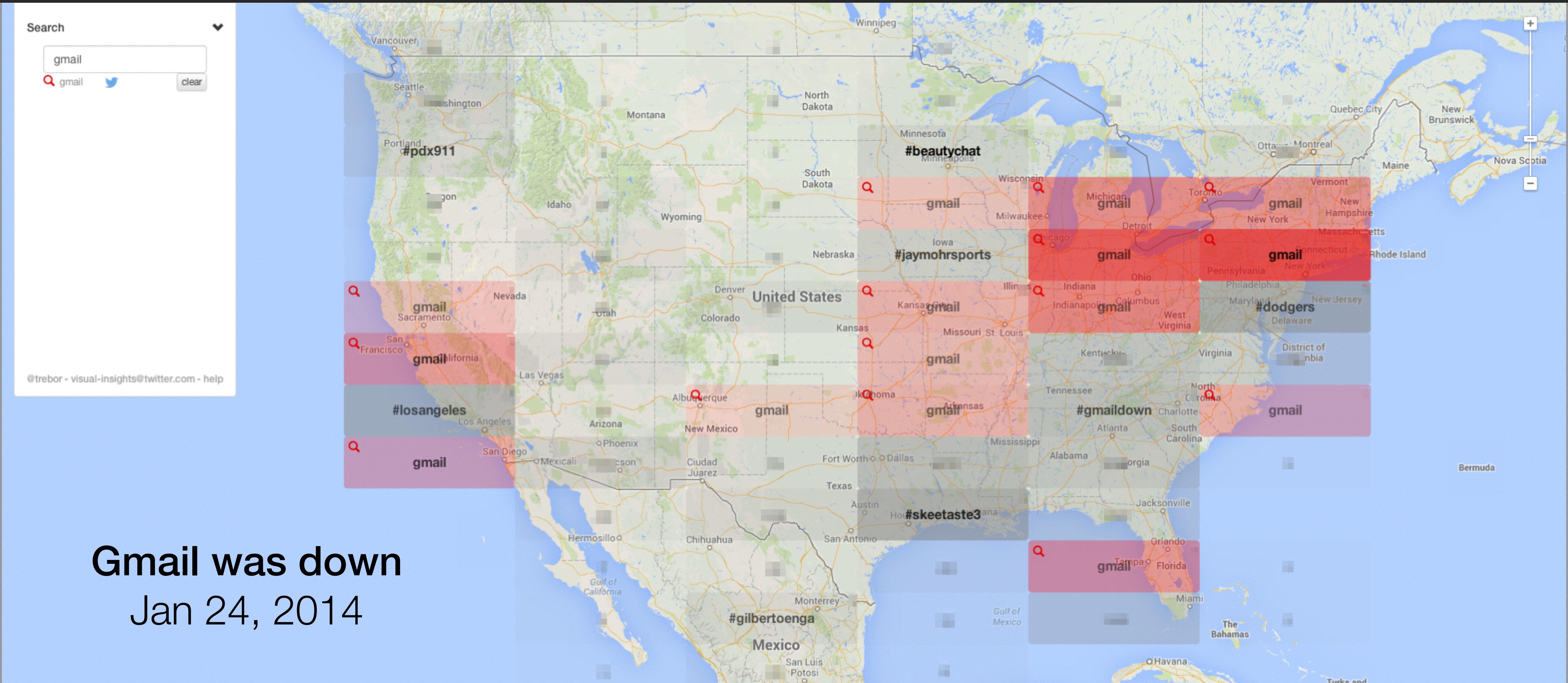


Visualize Data

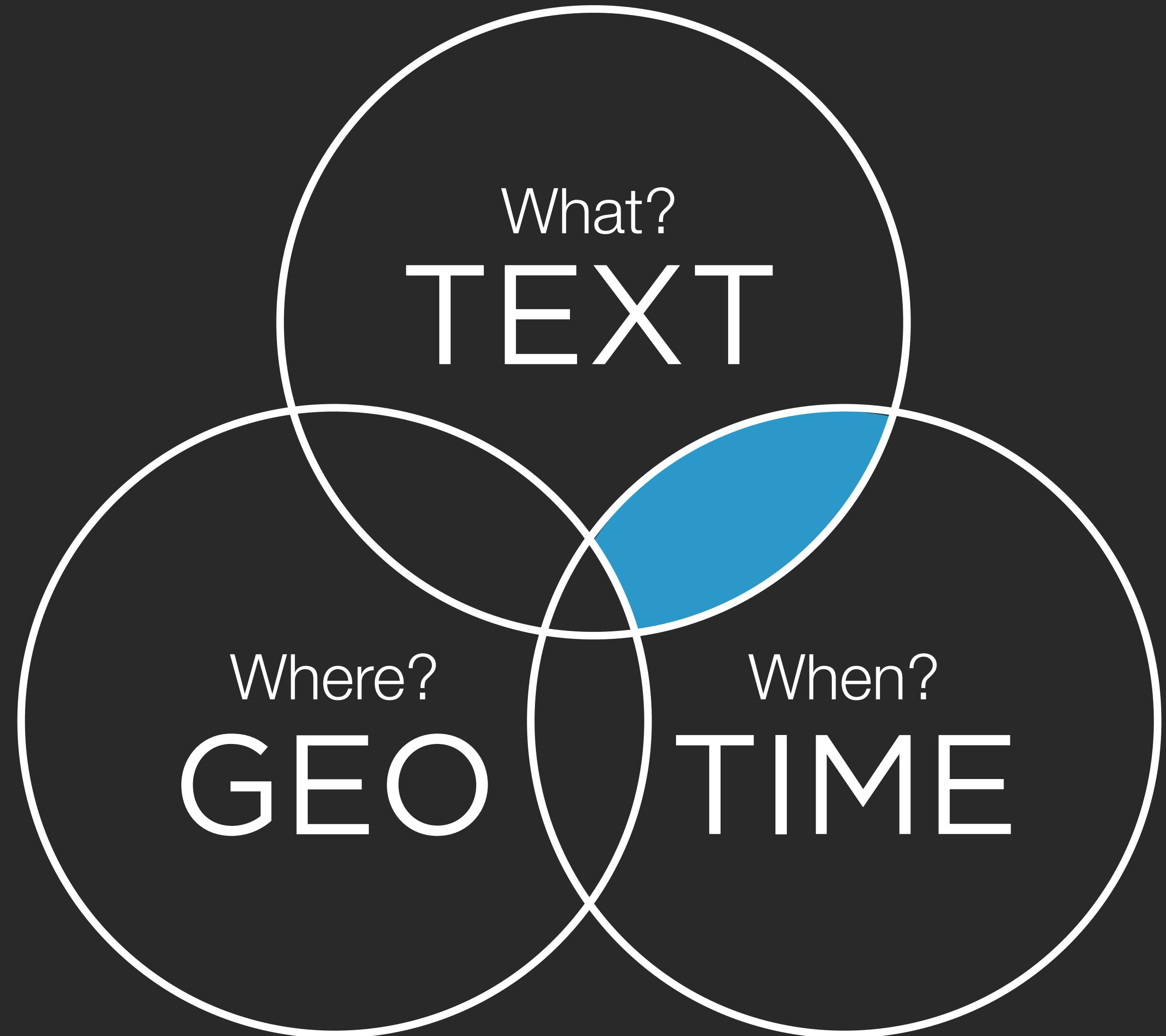


Geo + Text

Real-time Tweet map

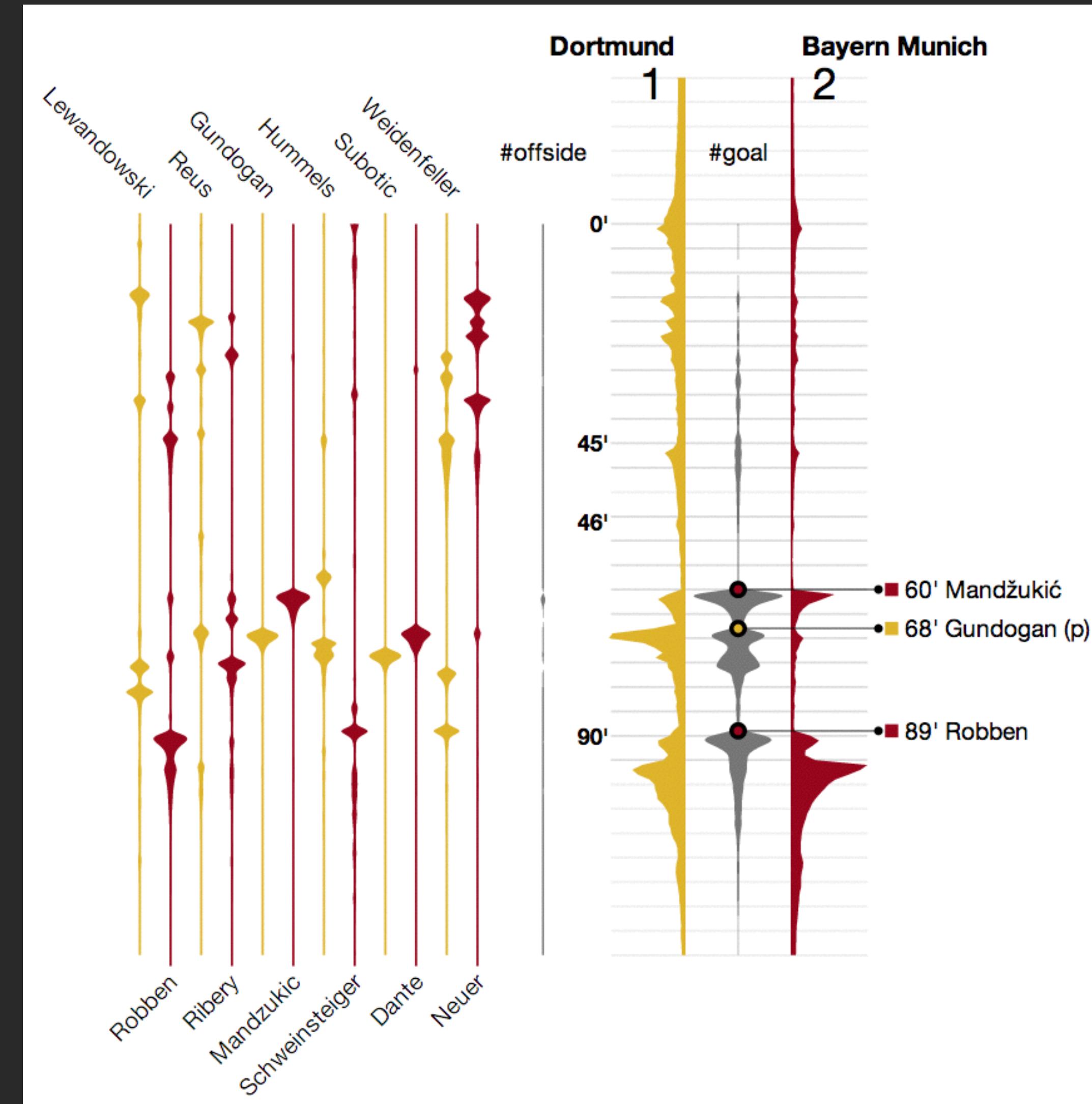


Visualize Data



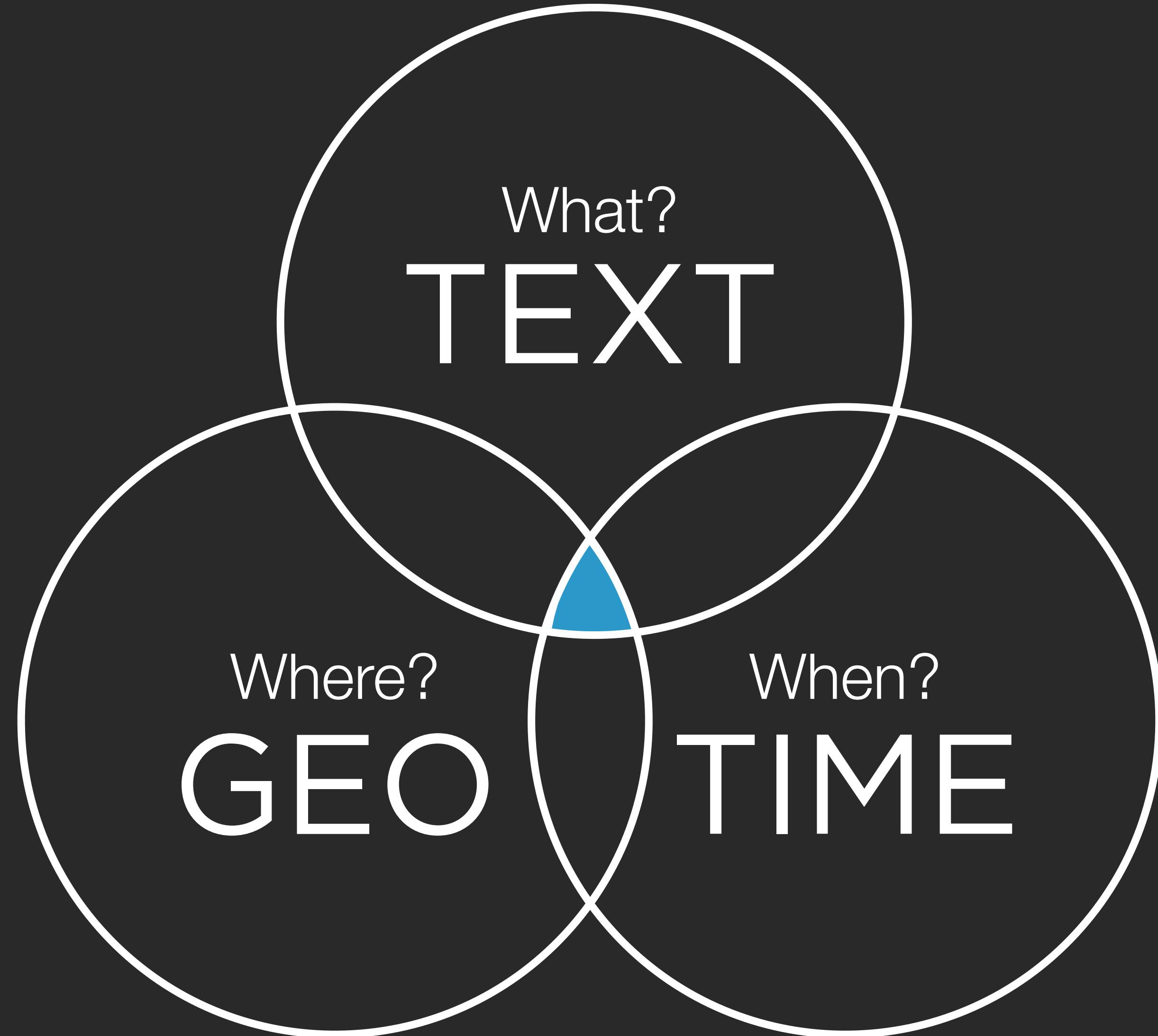
Time + Text

UEFA Champions League



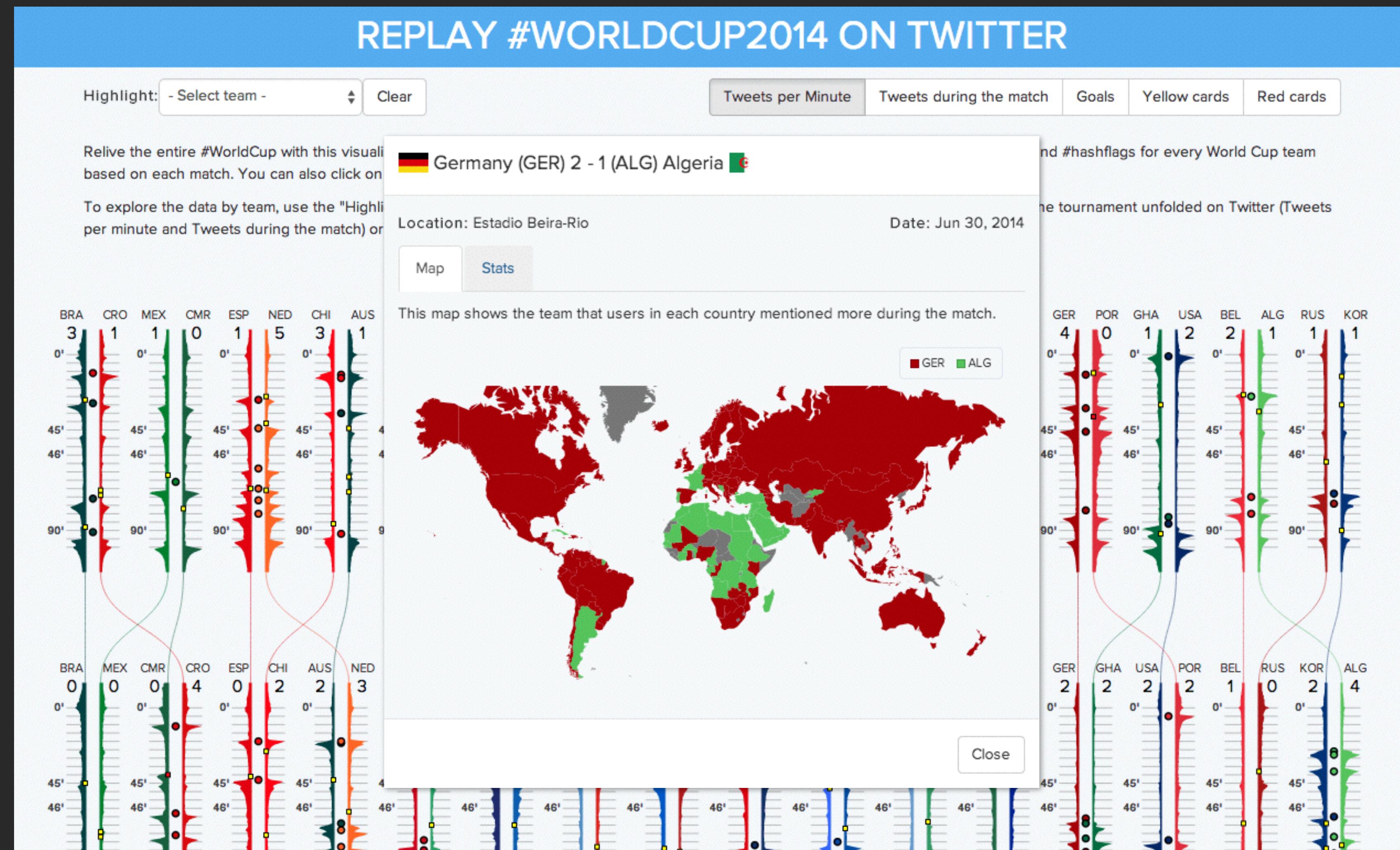
+ players

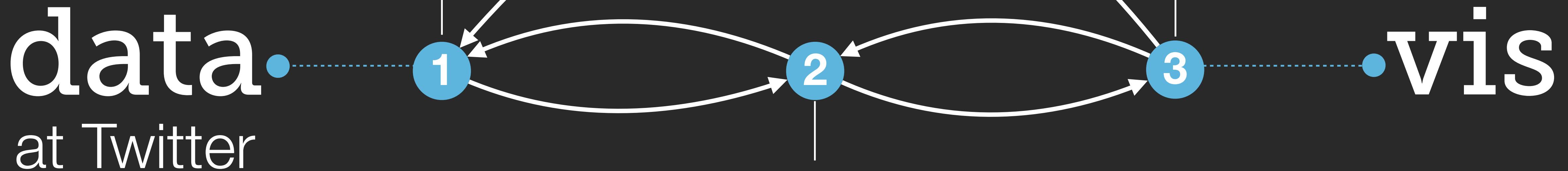
Visualize Data



Time + Text + Geo

World Cup 2014





big data => small data

Get data

data•
at Twitter

“Tweets”

- users
- followers graph
- logs
- etc.

self, peer, external
Evaluate

• vis
(with deadline)

@kristw / <https://interactive.twitter.com>

+ visualizations by @philogb, @miguelrios & @trebor