

Predictive Analytics using Facebook Social Graphs

Mehdi Allahyari and Supa' Mike

These slides available at: <http://cs8380.michaelcotterell.com/>

*Originally presented on February 14th, 2012 in
Dr. Arpinar's Spring 2012 class on Advanced Topics in Information Systems
at the University of Georgia, Athens, GA, USA*

Predictive Analytics

Encompasses a variety of statistical techniques that analyze current and historical relationships between explanatory variables and the predicted variables from past occurrences, and exploiting it to predict future outcomes.

Techniques:

- modeling,
- machine learning and
- data mining
 - classification
 - clustering
 - frequent pattern mining

See: http://en.wikipedia.org/wiki/Predictive_analytics

Facebook Social Graph

At Facebook's core is the social graph; people and the connections they have to everything they care about.

The Graph API presents a simple, consistent view of the Facebook social graph, uniformly representing:

- the objects in the graph (e.g., people, photos, events, and pages) and
- the connections between them (e.g., friend relationships, shared content, and photo tags).



Friendships in the Facebook Social Graph

What's the *Cool*est Data?

EVERYTHING! (<https://developers.facebook.com/tools/explorer?method=GET&path=4941934>)

Example of checkin data:

```
{ "id": "1",  
  "from": { "name": "Michael Cotterell", "id": "5" },  
  "tags": { "data": [ { "name": "Haseeb Yousaf", "id": "7" },  
                     { "name": "Mehdi Allahyari", "id": "4" } ] },  
  "message": "Dubstep music is cool here.",  
  "place": { "id": "23",  
             "name": "Cool Place",  
             "location": { "street": "1531 Somewhere Dr.", "city": "Athens",  
                           "state": "GA", "country": "United States",  
                           "zip": "30605",  
                           "latitude": 33, "longitude": -84 } },  
  "application": { "name": "Facebook for Android", "id": "350" },  
  "created_time": "2012-02-13T18:52:30+0000" },
```

How do we retrieve the data?

Develop a small application that queries the Facebook Graph API for all of its checkin and event data.

Considerations:

- Is the API *rate limited*? (Does it restrict how many queries you can perform during any given time interval? If so, how do we handle this?)
- Are query results limited?

After we Query the Graph API, how do we Store the JSON Data?

Relational Database

- Small Schema
- Easy to Query

RDF

- Makes Linked Data connections possible / easier (e.g., dbpedia).
- Perhaps use YAGO2 format to support Temporal / Geographic queries?

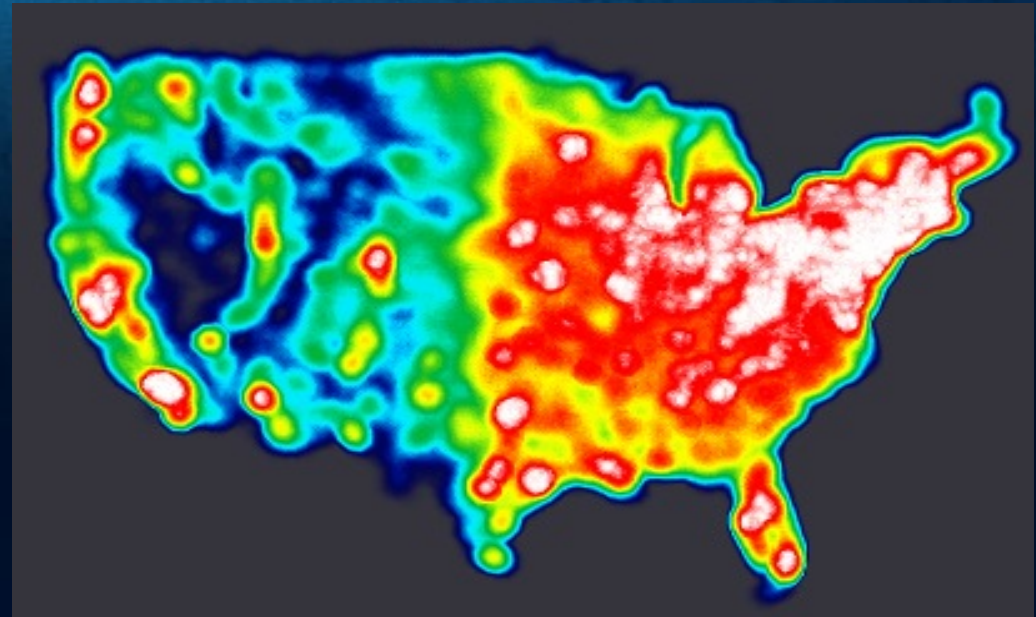
We will probably use RDF...

In addition to *geographic information*, each result from a query to the Graph API contains string data in *tags*, *messages*, *location names*, etc.

Using Linked Data and ontological techniques, we can find **semantically related terms**.

Find locations that mention these related terms.

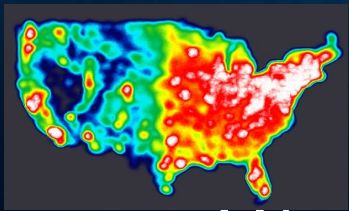
Where geographically, do these terms overlap? We can capture this information in a heat map.



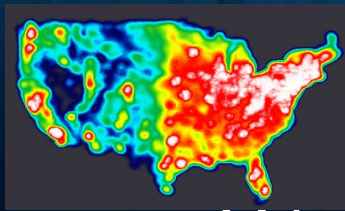
What are we trying to Predict?

More of a question:

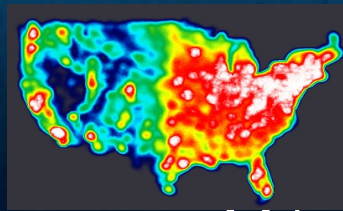
Can trends and correlations between *Facebook Social Graph Data* and *Semantically Related Meta Data* be used to predict the socio-geographic flow of social interests?



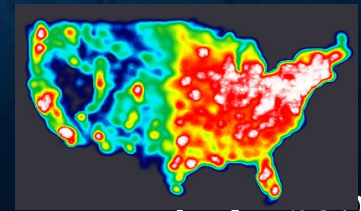
Q1



Q2



Q3



uture?

Potentially Useful

Social Psychologists perform similar studies where they follow the social interactions and interests of a group of people in order to determine future behavior.

- What if they could do this with all Facebook users?

Businesses, Corporations, Lobbyists and Politicians would like to know where to target their advertising.

- Targeted advertising is not only more useful for advertisers but significantly less intrusive to customers.

Questions?

We would like to thank the audience for their patience!
Ya'll are great!