

Comparing the Interests of People in Different Cities Using Topic Analysis and Social Media

Jake Lasley, Alan Motta

November 28, 2018

Social Media

What is social media?

Twitter & Instagram



Figure: Instagram Logo

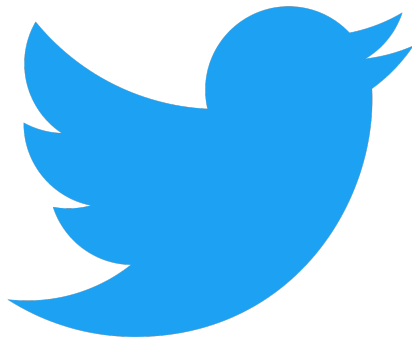


Figure: Twitter Logo

The Problem With Instagram

- ▶ Restrictive API
- ▶ Low API request rates

Solution: Twitter

- ▶ Twitter offers developer accounts (good API access)
- ▶ Multiple search API end points
- ▶ High request rates

Data Gathering Process

Steps:

1. Find users from different cities
2. Collect tweets from each user
3. Analyze tweets
4. Compare cities

Data Gathering Process

Two main API endpoints:

- ▶ **users/search** : Returns pages of users (max 1k results, across max of 50 pages)
- ▶ **statuses/user_timeline** : Returns most recent 3,200 tweets of specified user (if account is not private)

Data Gathering Process

Querying For Data

twurl: cURL like tool which wraps OAuth

```
\$ twurl /1.1/users/search.json?q=City
```

```
\$ twurl /1.1/statuses/timeline.json?id=user\_id
```