Viral Social Media Movements

Claire Huang and Heejung Chung

Introduction

Introduction: Importance of studying social media

- Social media has **broad reach** across demographics, locations
  - **Facebook** = 1.65 billion users
  - **Twitter** = 310 million users
  - **Youtube** = 4.95 billion videos viewed per day

- Quickly becoming **medium of choice** for information spread

- **Nonprofit** organizations don’t have many marketing resources
  - maximize effectiveness by studying previous movements

Introduction: Movements analyzed

- #BlackLivesMatter (**BLM**)
  - Started July 2013, after George Zimmerman's acquittal
  - Multiple successive, domestic **"spark"** events
  - Relatively **successful** (controversial)
- #BringBackOurGirls (**BBOG**)
  - Started April 2014, after Chibok schoolgirl kidnapping
  - One foreign spark event
  - Very **unsuccessful** (girls are still missing, hashtag out of use)

Images
http://www.cnn.com/2015/04/14/opinions/sesay-bring-back-our-girls-one-year-on/
https://solidarity-us.org/node/4316

Introduction: Areas of focus & Defining virality

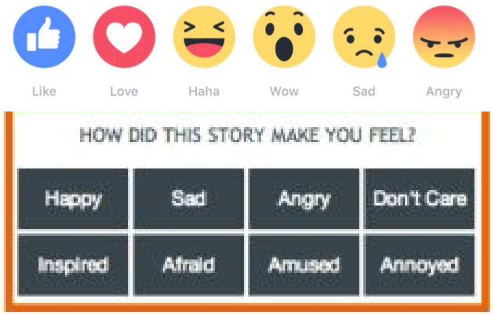
- Individual tweet engagement
  - *Virality = # favorites and # retweets*
  - *What factors make an individual tweet go viral?*

- Popularity over time
  - *Virality = volume, ie. # times #BlackLivesMatter or #BringBackOurGirls is tweeted per day*
  - *What trends are in a movement's virality over time?*

Background

Background

- Emotional response
  - **Terms**: valence-arousal-dominance, narrow/broadcasting
  - high valence (positive) ⇒ ↑ sharing
  - high arousal (action) ⇒ ↑ narrowcasting
  - high dominance (in control) ⇒ ↑ broadcasting



Images: Guerini and Saiano, http://www.hlntv.com/shows/the-daily-share/articles/2016/02/24/facebook-rolls-out-reaction-buttons

- Author and audience background
  - <u>audience</u>: medium vs. high education level
  - <u>author</u>: “high popularity [# followers] does not imply high influence” (Romero, et al)

Background

- Multiple networks over time
  - predictable content migration patterns
    - ex. Reddit ⇒ Twitter ⇒ Facebook

- Concurrent news coverage
  - many movements are either <u>sparked</u> or <u>spread</u> by news coverage
    - “It Gets Better” or “ALS Ice Bucket Challenge” videos featured
    - #BringBackOurGirls: delayed news coverage, negative stories
    - #BlackLivesMatter: multiple news events

Methodology

Methodology: Original Plan

- 8 movements with different:
  - media types (video, image, text)
  - durations and times
  - success levels

- For each movement, we planned to analyze…
  - view frequency
  - response frequency
  - sharing frequency
  - duration of popularity

	Movement	Social Media Platform		
1	Dove Beauty Campaign			Youtube
2	Ice Bucket/ALS	Facebook		
3	Red Equals Sign			
4	Movember		Twitter	
5	Kony 2012			
6	It Gets Better			
7	Black Lives Matter			
8	Bring Back Our Girls			

Our eight movements during the planning phase

Methodology: Obstacles

Facebook

- Privacy controls would introduce bias
- Instagram had similar controls

Youtube

- Data API limits (eg. max # of retrievable video objects)

Twitter

- Search API limits (eg. couldn’t retrieve tweets more than a week old)
- Jefferson Henrique script was inconsistent => time limitation

Google News Search API

- Ceased operations in 2014
- no alternative news search with the same breadth

Methodology: Actual Plan

- Found tweet-collecting script (by Jefferson Henrique)
  - bypasses API limitations by querying Twitter website directly

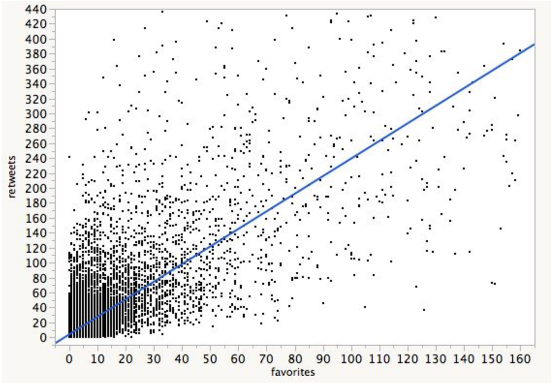
- Narrowed focus...
  - Twitter only
  - Two movements only
  - Opposite success-levels

```python
for tweetHTML in tweets:
    tweetPQ = PyQuery(tweetHTML)
    tweet = models.Tweet()

    usernameTweet = tweetPQ("span.username.js-action-profile-name b").text();
    txt = re.sub(r"\s+", " ", re.sub(r"[^\x00-\x7F]", "", tweetPQ("p.js-tweet-text").text())
    # print txt.decode('utf-8')
    retweets = int(tweetPQ("span.ProfileTweet-action--retweet span.ProfileTweet-actionCount"
    favorites = int(tweetPQ("span.ProfileTweet-action--favorite span.ProfileTweet-actionCoun
    if retweets <= 0 and favorites <= 0:
        continue
    dateSec = int(tweetPQ("small.time span.js-short-timestamp").attr("data-time"));
    id = tweetPQ.attr("data-tweet-id");
    permalink = tweetPQ.attr("data-permalink-path");

    geo = ''
    geoSpan = tweetPQ('span.Tweet-geo')
    if len(geoSpan) > 0:
        geo = geoSpan.attr('title')

    tweet.id = id
    tweet.permalink = 'https://twitter.com' + permalink
    tweet.username = usernameTweet
    tweet.text = txt
```

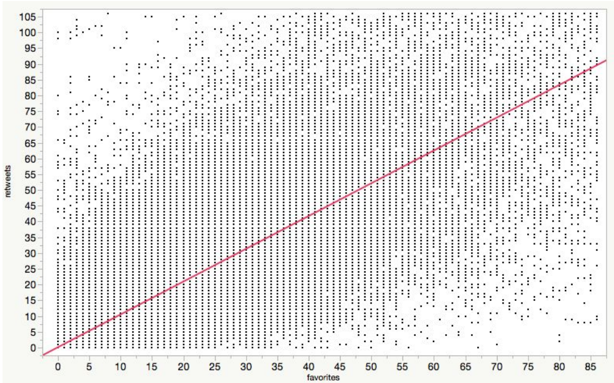
	Movement	**Media Platform**
1	**Black Lives Matter**	**Twitter**
2	**Bring Back Our Girls**	

Results

Results: Engagement with Individual Tweets

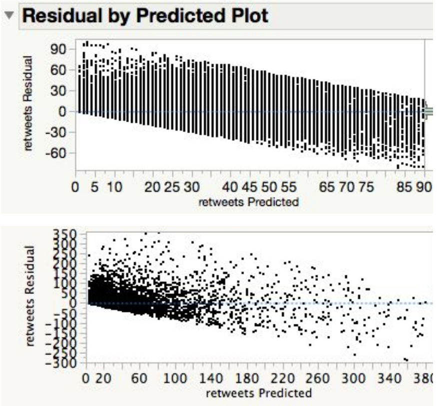


slope = 2.36



slope = 1.04

both slopes are significantly **greater than 1**

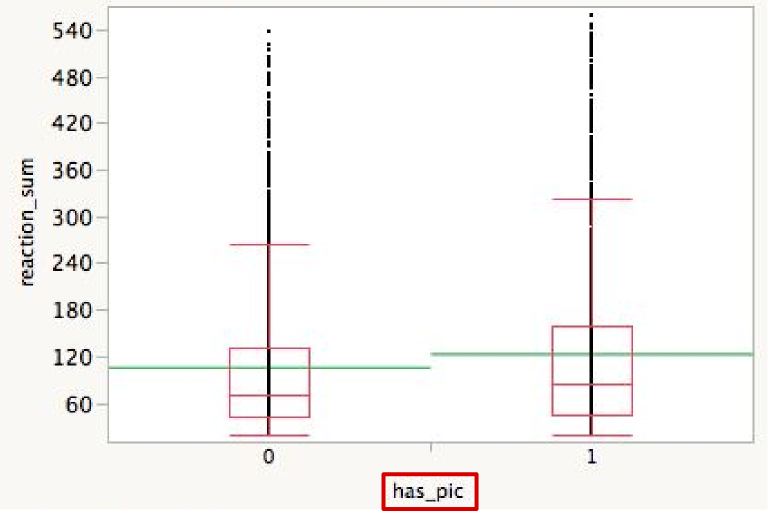


unexplained downward **pattern** in residuals

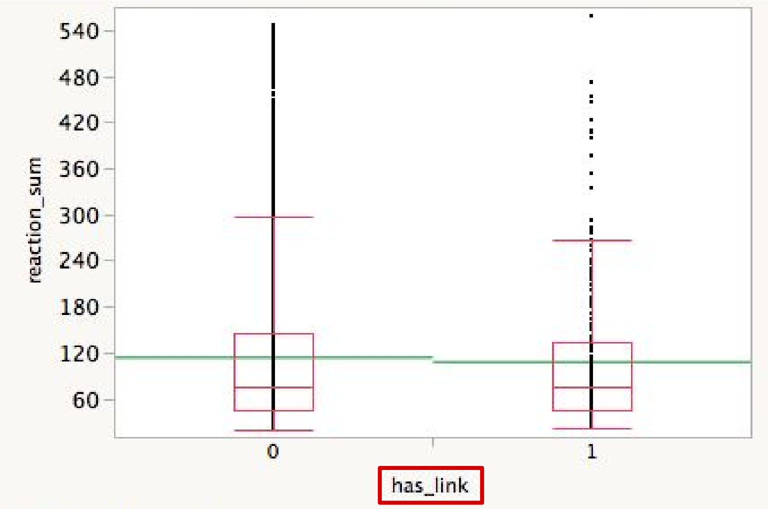
Bivariate analysis of # favorites versus # retweets

Results: Engagement with Individual Tweets

LEGEND:
quantiles
mean



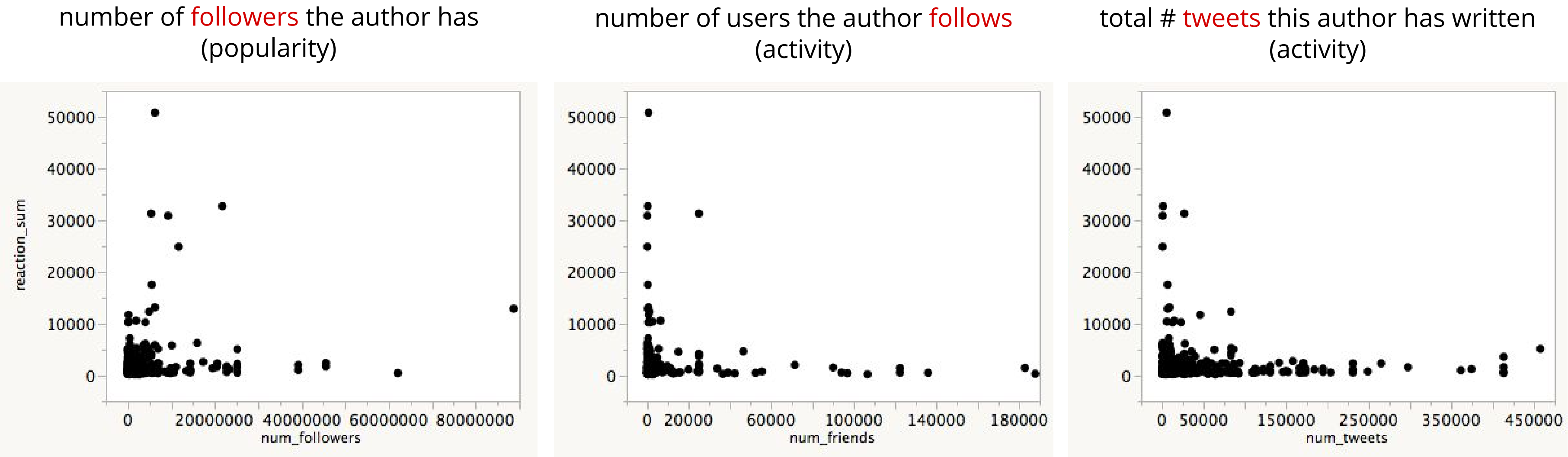
significantly **greater engagement** with picture



not significantly different engagement with or without external link

Bivariate analysis of picture or link presence in #BBOG tweets versus engagement

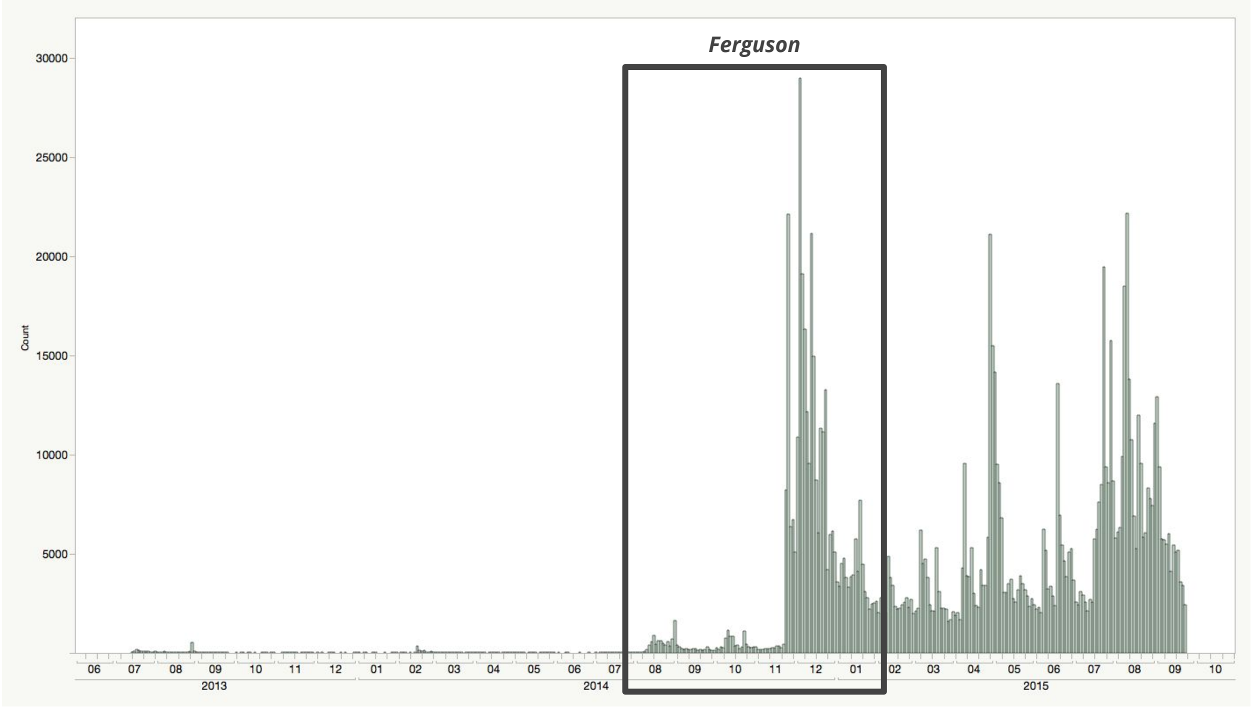
Results: Engagement with Individual Tweets



no clear patterns for any author traits

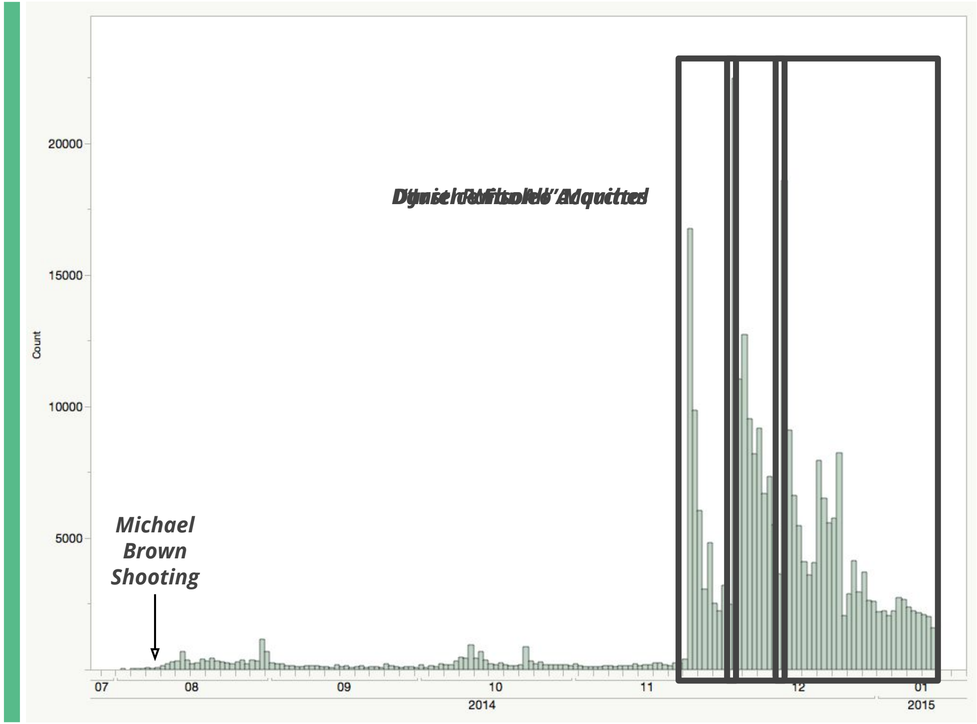
Bivariate analyses of author traits and engagement for the top 0.5% #BBOG tweets

Results: Popularity Over Time



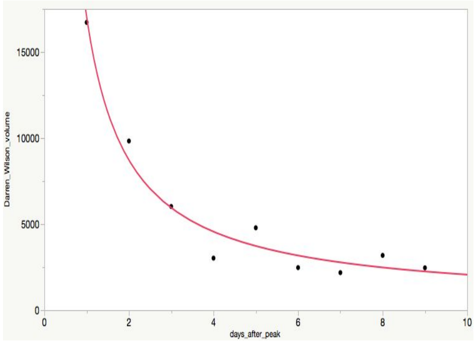
*Time Series of **#BlackLivesMatter** Usage, July 2013 to September 2015*

Results: Popularity Over Time



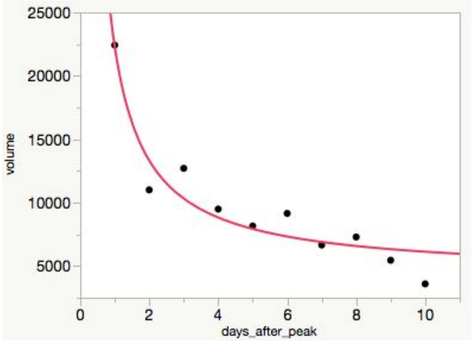
*Time Series of **#BlackLivesMatter** Usage During Ferguson, July 2014 to January 2015*

Results: Popularity Over Time (mostly **reciprocal**)



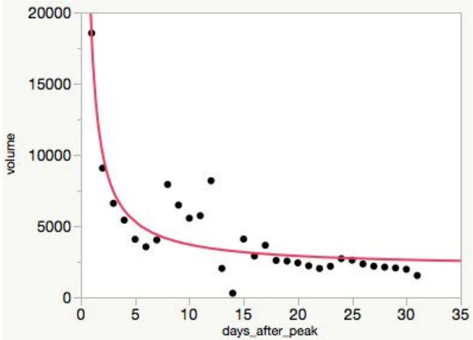
Wilson Acquittal

volume = 399.6 + 16634.7 * ***Recip***(days_after_peak)



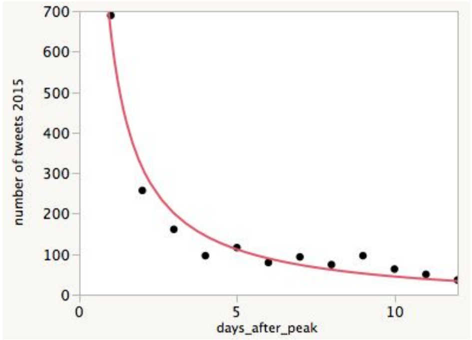
Pantaleo Acquittal

volume = 4332 + 17952 * ***Recip***(days_after_peak)



"Justice For All"

volume = 2098.9 + 15994.2 * ***Recip***(days_after_peak)

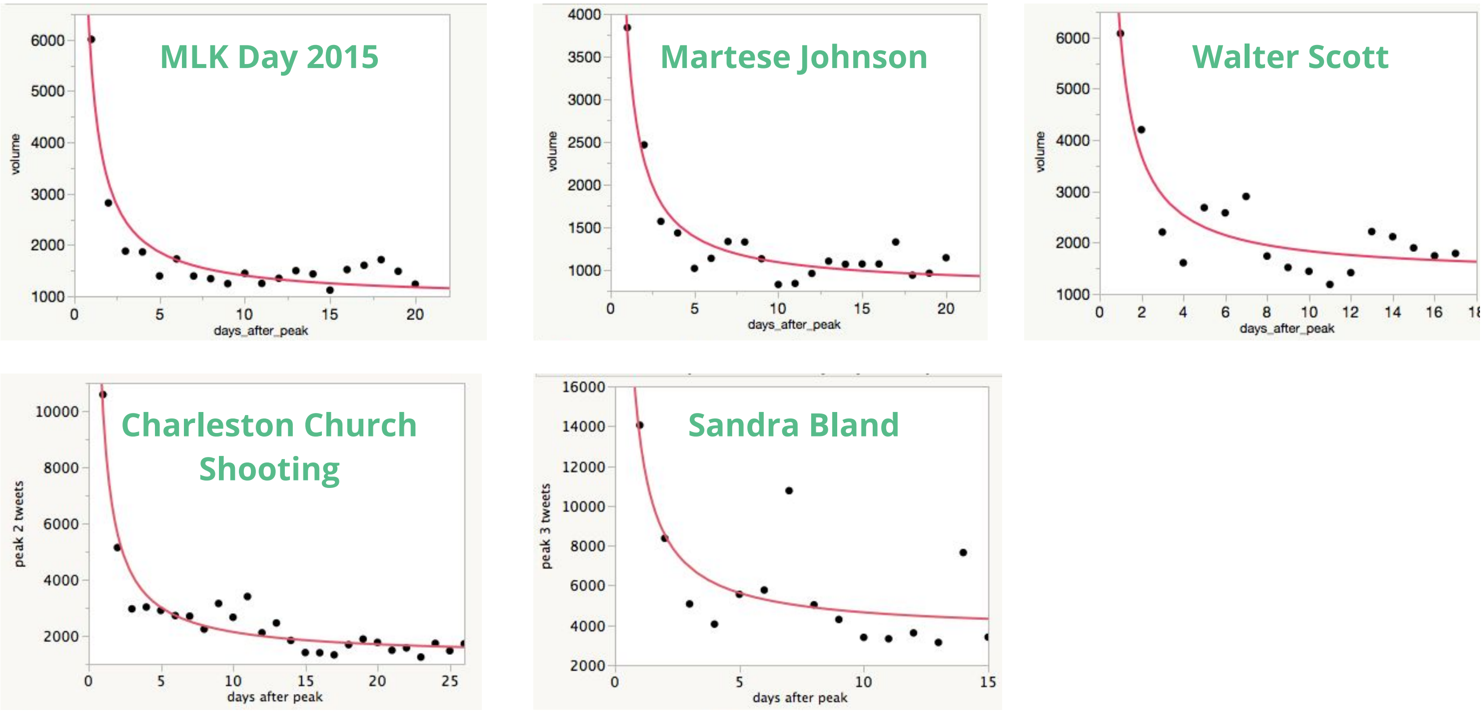


BBOG 1st Anniversary

volume = -19.9 + 674* ***Recip***(days_after_peak)

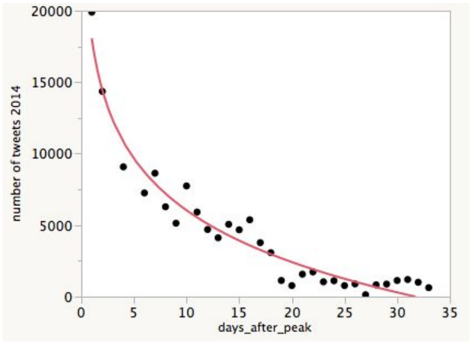
***Reciprocal** Regressions on Time Series Data After Spark Events*

Results: Popularity Over Time (mostly **reciprocal**)



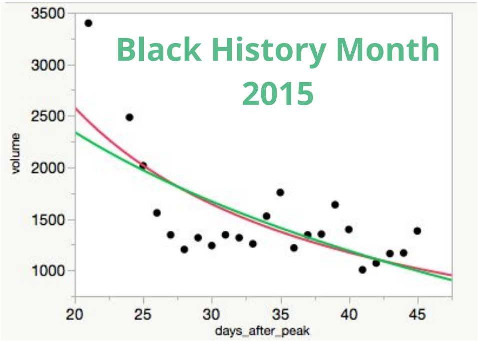
Reciprocal *Regressions on Time Series Data After Spark Events*

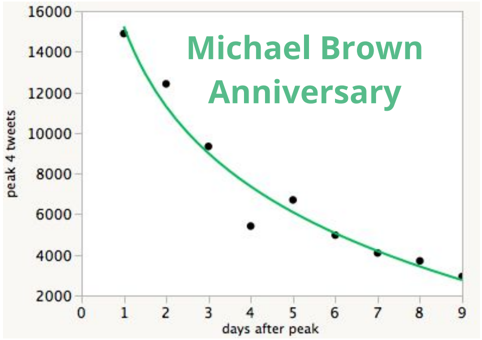
Results: Popularity Over Time (some **logarithmic**)

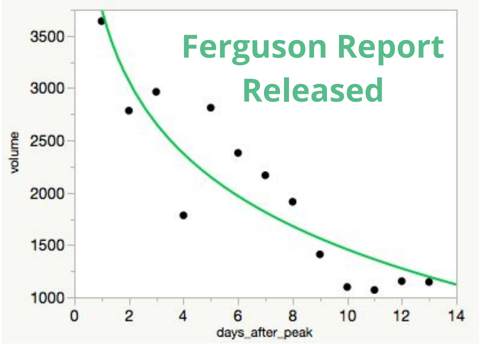


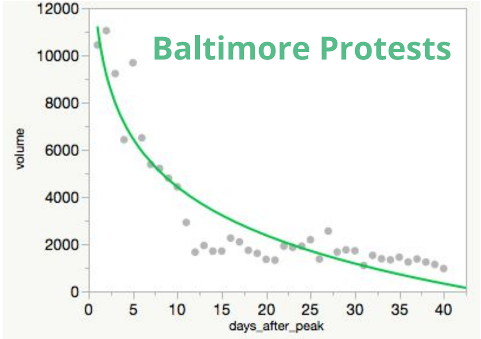
BBOG Kidnapping

volume = 18067.8 + 5525.8831 * **Log**(days_after_peak)









Logarithmic *Regression on Time Series Data After Spark Events*

Discussion & Conclusions

Discussion: Interpreting Findings

- Individual tweet engagement
  - The more viral the tweet, the more **retweets (broadcasting)** per favorite (narrowcasting)
  - Tweets with **pictures** correlate with more engagement
  - For the most viral tweets, **author traits** do not correlate with engagement

- Trends in popularity over time
  - **9** sparks: **reciprocal** fit
  - **4** sparks: **logarithmic** fit
  - **1** spark was **ambiguous** (both fits were mediocre)
  - All cases: Huge coefficient (ie **sharp decline**) ⇒ **unsustainable**

Discussion: Recommendations

- For organizations looking to go viral
  - Short-term engagement: photos
  - Long-term engagement: engineer ***repeated "sparks"*** (aka. spice things up)

- Request to social media companies (esp. hosting public content)
  - Run/store data on visual and textual ***sentiment analysis as posts are being created***
  - Easier data collection and/or analysis

Discussion: Further Questions

- Engagement with individual content
  - Author
    - What **traits** about an author correlate with higher engagement?
    - **Retention**: how long and often does each author post on the same topic?
  - Emotional content
    - **Sentiment analysis** on text and photos

- Trends in popularity over time
  - Concurrent news coverage
    - Did news coverage **peak** before or after social media? Are these peaks longer or shorter?
  - Can we explain reciprocal and logarithmic models of decline?
    - Potential factors: familiarity of spark? domestic vs foreign spark?

Thank You!