Using Social Media Monitoring to Guide Inspections

Lauren DiPrete, REHS
Senior Environmental Health Specialist
Southern Nevada Health District

Outline

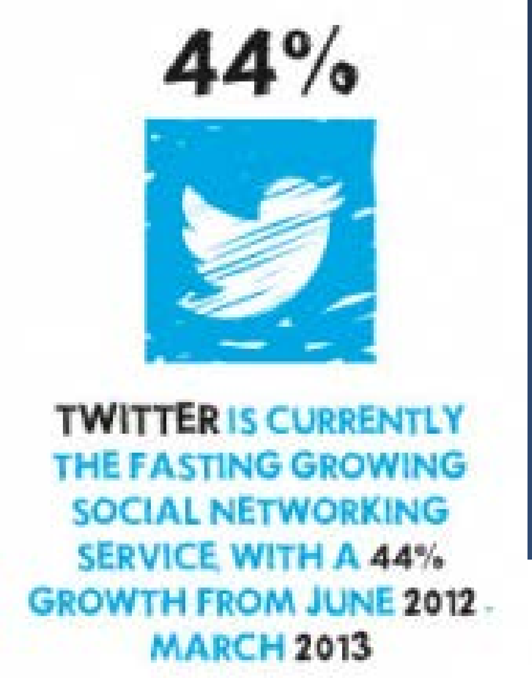
- Growth of social media
- Overview of social media monitoring for public health
- Formation of partnership
- Study design and obstacles
- Results
- Future work

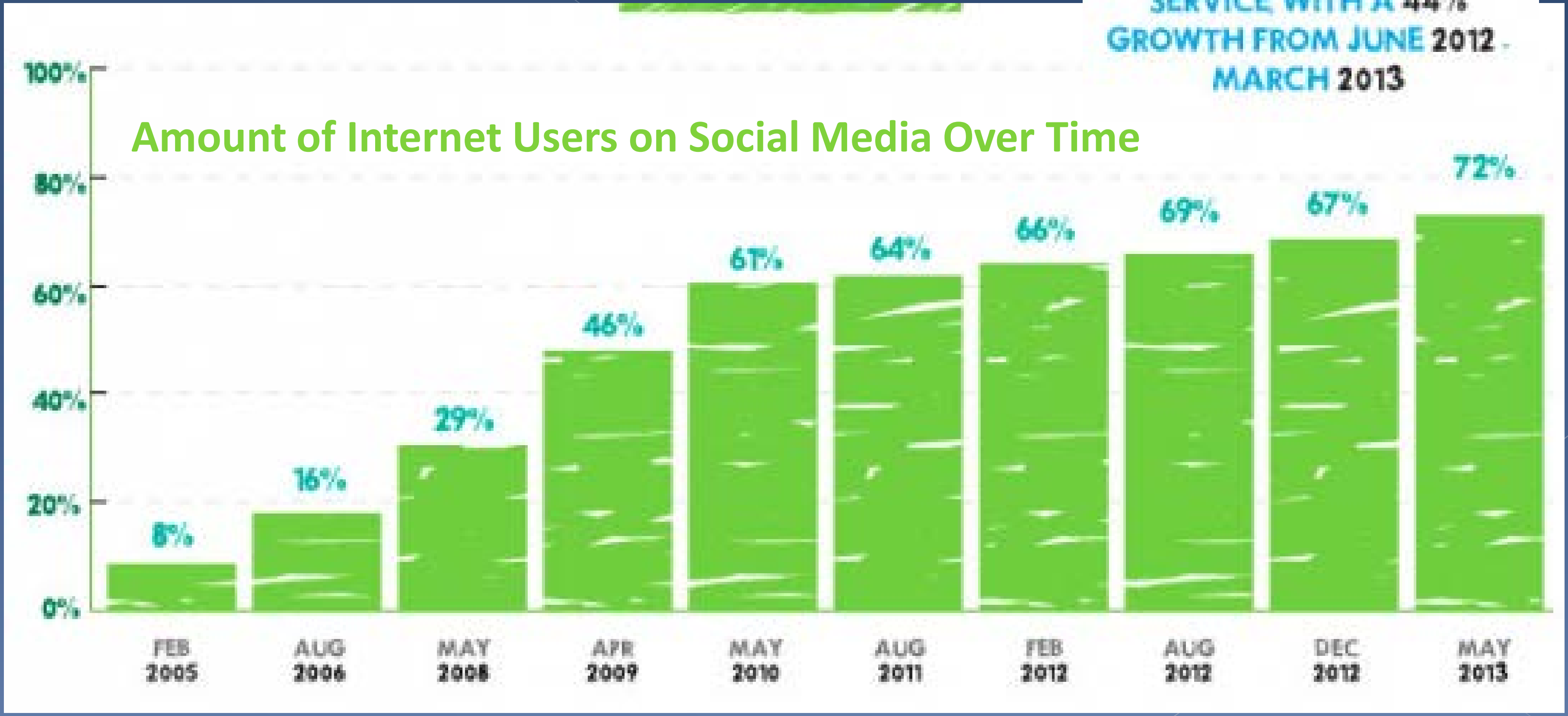
Foodborne Illness

- Common
  - 1 in 6 Americans sickened
  - 3,000 deaths annually
- Costly
  - $15.6 billion annually per USDA 2014
- Preventable

Growth of Social Media

- Image credit: Search Engine Journal





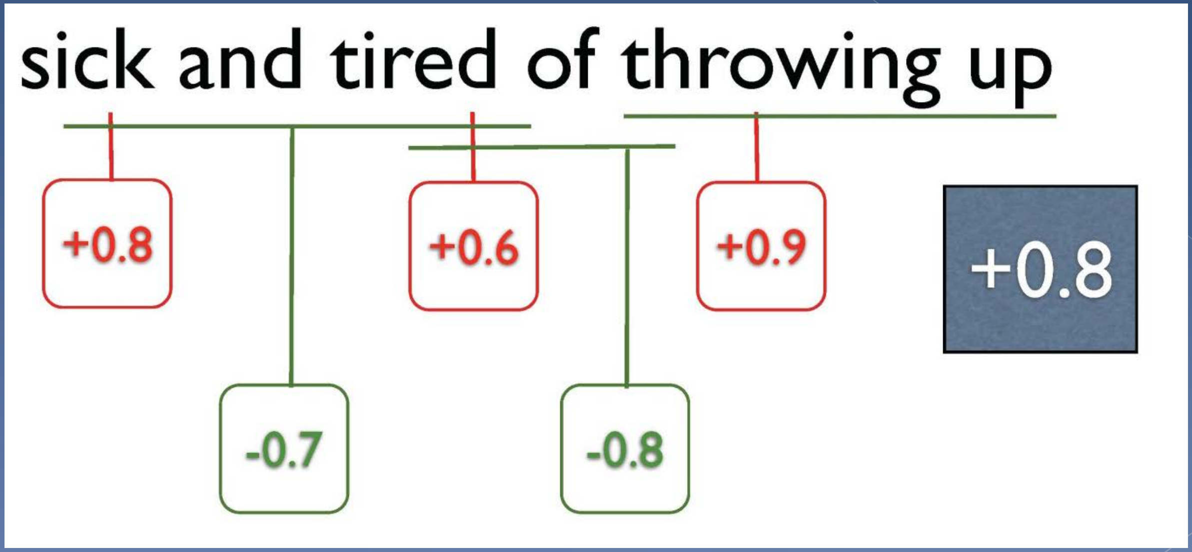
How it works

- Listens to Twitter and flags restaurants

- Examples of “sick” tweets picked up:
  - Still throwing up I can't go to work stupid food poisoning I hate it
  - I have THE WORST stomach ache right now
  - I'm gonna throw up
  - WOW NOW I HAVE DIARRHEA WHAT THE HECK MAN
  - My stomach is ... up fam! What in the ... did I eat?

How it works

- Advanced Language Model
- Statistical Word Trigram Model
  - “So sick” vs “So sick of homework”
  - “Under the weather”



How it works

- 1. Tweet from a restaurant
  - (If you didn’t turn geolocation services off, then….)
- 2. Match tweet to restaurant using Google Place
- 3. Follow user for 2 weeks
- 4. Link any subsequent “sick” tweets to restaurant
- 5. Score restaurants based on number of sick tweets
- 6. List of “sick” restaurants flagged by software presented to health department

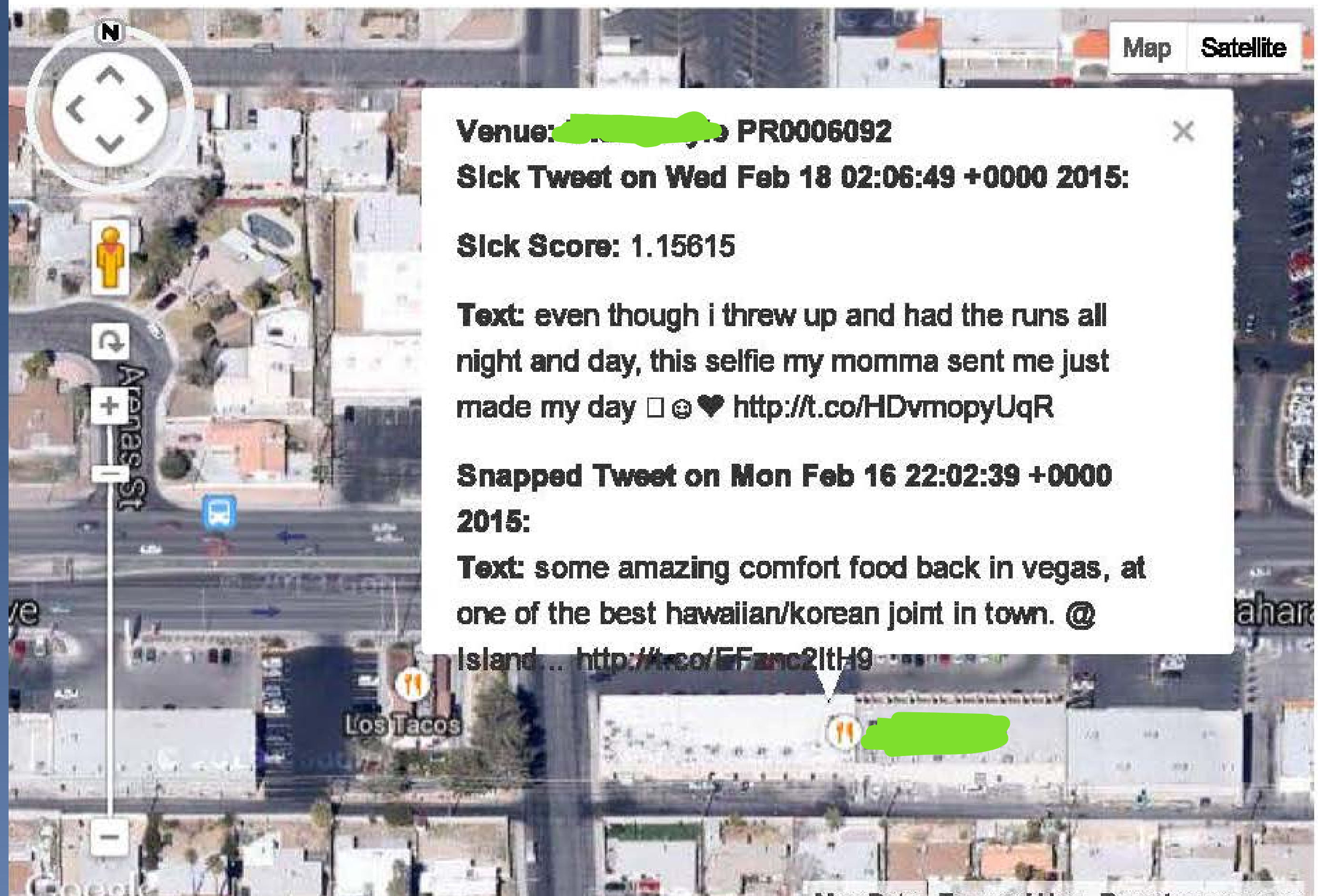
Multiple Restaurants

- What if a user posts from multiple restaurants before their sick tweet?
  - Sick tweet snaps back to all of them
  - Other users will help identify which one is the culprit

Turning This…

Into This

Google Name	Sick Score Avg	#Recent Tweets > 1.0 score	#Midterm Tweets > 1.0 score	#Longterm Tweets > 1.0 score	LaV Health Name	Permit Nbr	Followup
[redacted]	1.46	0	0	2	[redacted]	PR0006092	☑



Forming a Research Question

- Does the software guide inspectors toward facilities with uncontrolled risk factors?
- Goal: Intervene more quickly to prevent sickness

Study Launch

- Checked nEmesis daily
- Selected facilities to follow up on
- Identified matched controls based on district and permit type
- Dispatched blind inspectors
- Conducted routine inspections
- Collected results

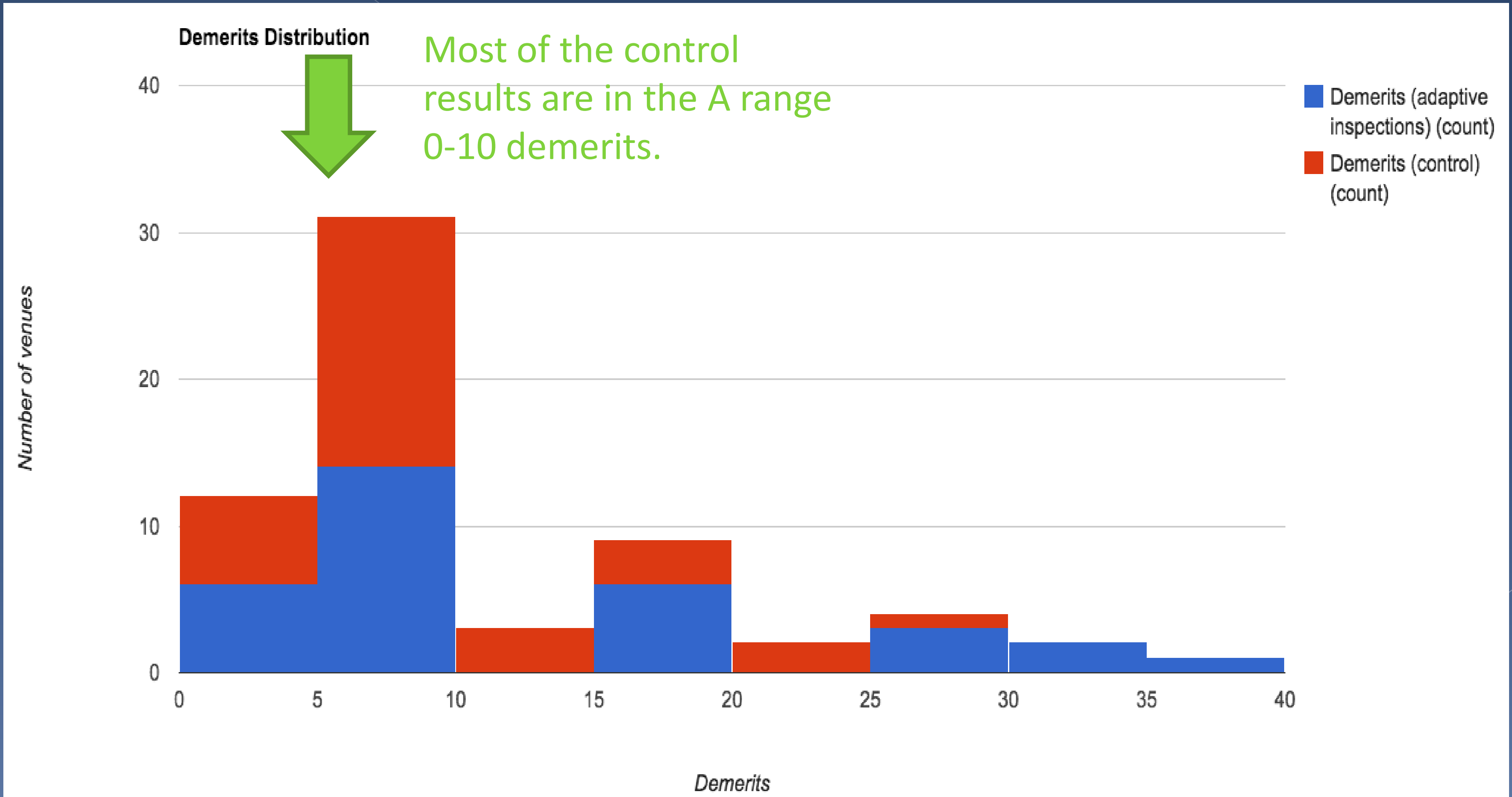
Results

- On a typical day:
  - 16,000 geo-tagged tweets
  - 3,600 users
  - 1,000 tweets from 600 users snap to a restaurant

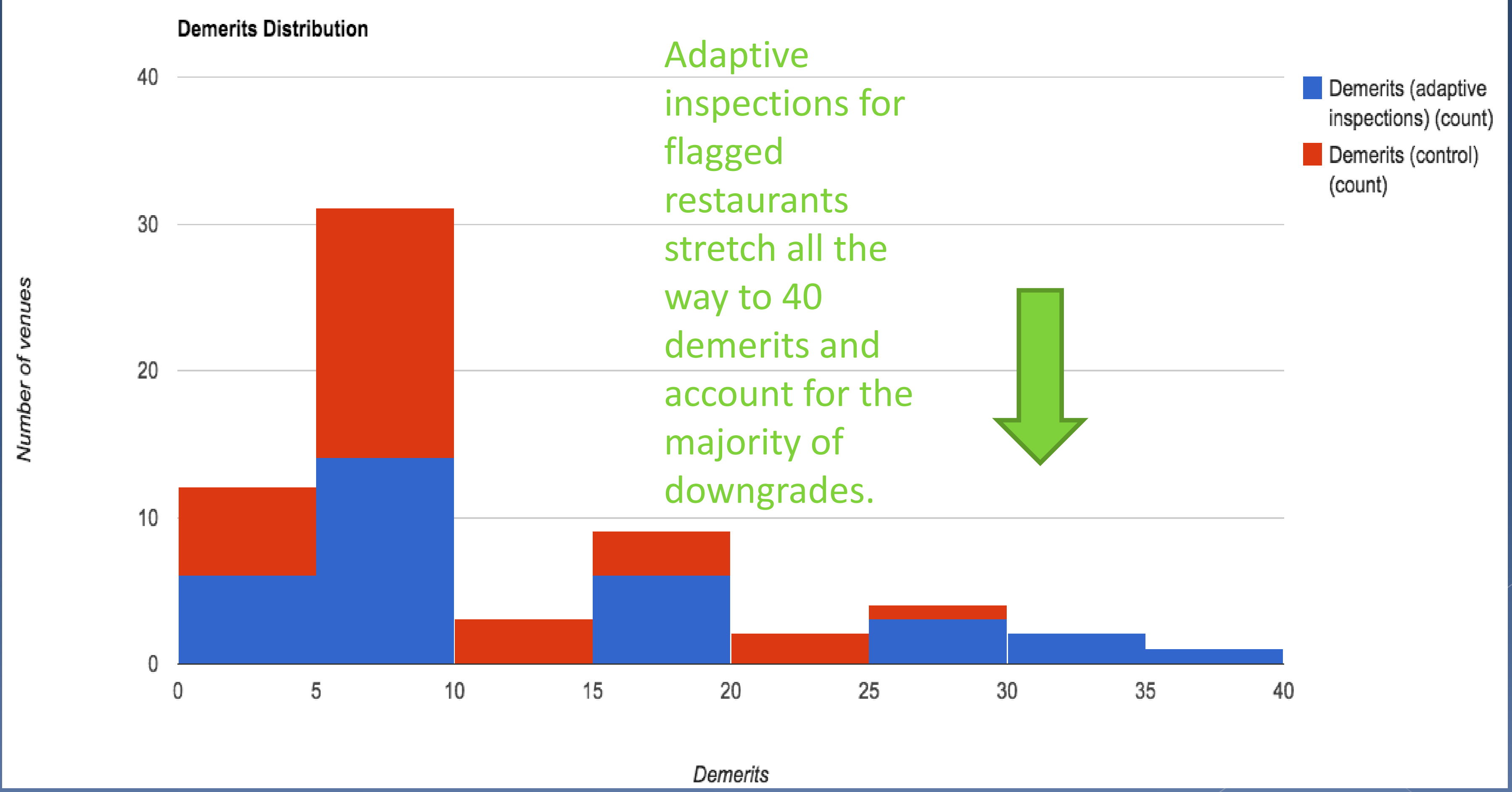
Results

- Study conducted from

  Jan 2, 2015 - Mar 31, 2015
- Inspected 72 flagged facilities and 72 control
- Flagged facilities earned more demerits than control facilities: 9 vs. 6
  - P-value of 0.019
- Flagged facilities 64% more likely to earn a C-downgrade than control



Demerits Distribution
Most of the control results are in the A range 0-10 demerits.
40
30
20
10
0
Number of venues
0
5
10
15
20
25
30
35
40
Demerits
Demerits (adaptive inspections) (count)
Demerits (control) (count)



Adaptive inspections for flagged restaurants stretch all the way to 40 demerits and account for the majority of downgrades.

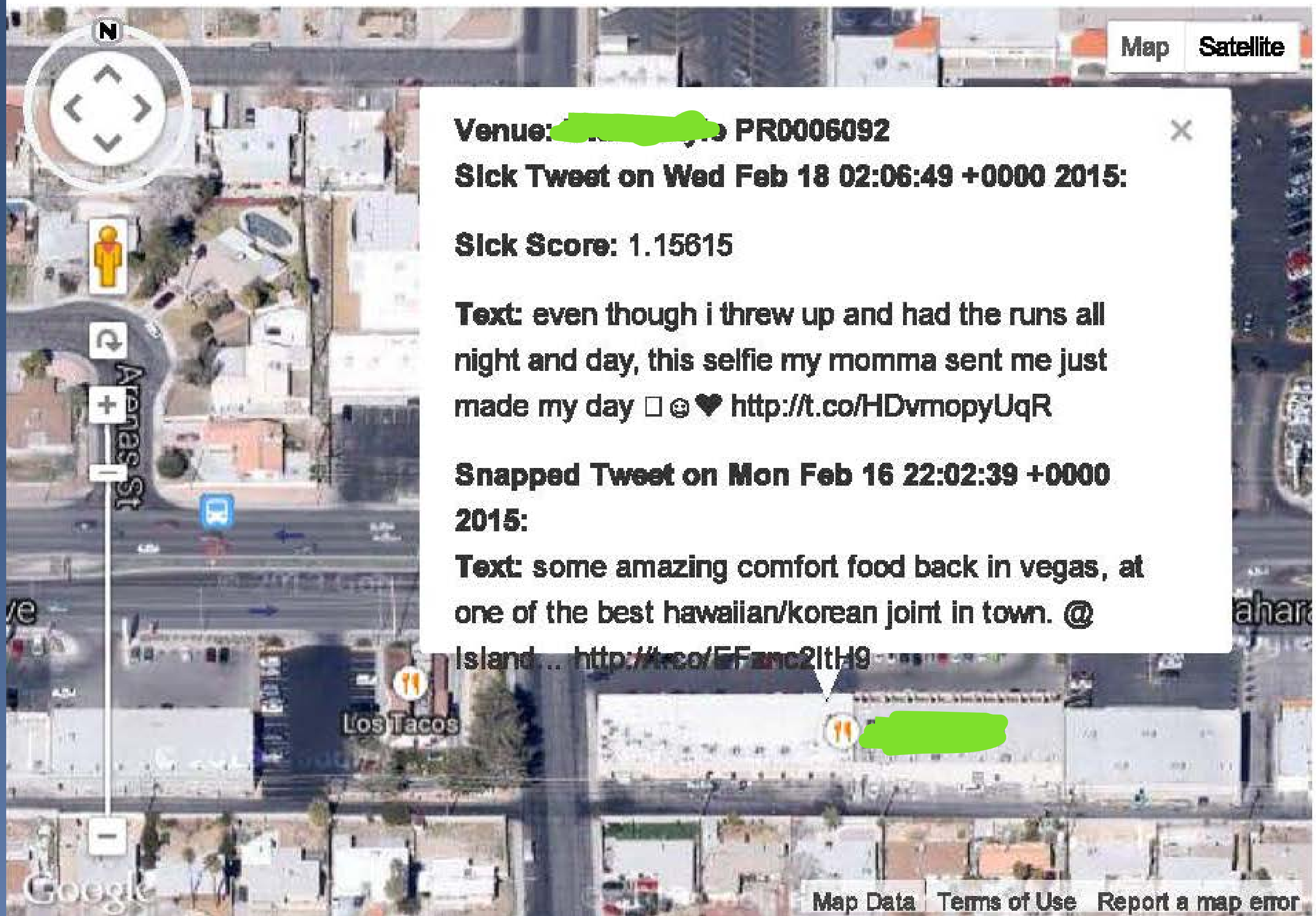
Unexpected Results

- Fewer FBI investigations performed during the study period

Time Period	Illness Investigations
Study period	5
Same period the year before	11
Average over 7 years	7.3

  - *Unable to determine statistical significance
- Inspection at flagged facility identified foodhandler working while ill
- Unpermitted special processes
- nEmesis identified possible unpermitted food establishments unknown to SNHD

Google Name	Sick Score Avg	#Recent Tweets > 1.0 score	#Midterm Tweets > 1.0 score	#Longterm Tweets > 1.0 score	LaV Health Name	Permit Nbr	Followup
[redacted]	1.46	0	0	2	[redacted]	PR0006092	✔



39 DEMERIT C DOWNGRADE

Limitations

- Unable to detect tweets in other languages
- Unable to detect evidence of illness posted to platforms other than Twitter

Next Steps

- CDC EHS-Net Cooperative Agreement
- Upgraded software
- Full launch

Conclusions

- Effective in flagging appropriate facilities
- Social media monitoring could be a useful tool for inspectors to conduct adaptive inspections

Deploying nEmesis: Preventing Foodborne Illness by Data Mining Social Media

Adam Sadilek*[1], Henry Kautz[1], Lauren DiPrete[2], Brian Labus[2],
Eric Portman[1], Jack Teitel[1], and Vincent Silenzio[3]

[1]Department of Computer Science, University of Rochester, Rochester, NY
[2]Southern Nevada Health District, Las Vegas, NV
[3]School of Medicine & Dentistry, University of Rochester, Rochester, NY

Abstract

Foodborne illness afflicts 48 million people annually in the U.S. alone. Over 128,000 are hospitalized and 3,000 die from the infection. While preventable with proper food safety practices, the traditional restaurant inspection process has limited impact given the predictability and low frequency of inspections, and the dynamic nature of the kitchen environment. Despite this reality, the inspection process has remained largely unchanged for decades. We apply machine learning to Twitter data and develop a system that automatically detects venues likely to pose a public health hazard. Health professionals subsequently inspect individual flagged venues in a double blind experiment spanning the entire Las Vegas metropolitan area over three months. By contrast, previous research in this domain has been limited to indirect correlative validation using only aggregate statistics. We show that adaptive inspection process is 63% more effective at identifying problematic venues than the current state of the art. The live deployment shows that if every inspection in Las Vegas became adaptive, we can prevent over 9,000 cases of foodborne illness and 557 hospitalizations annually. Additionally, adaptive inspections result in unexpected benefits, including the identification of venues lacking permits, contagious kitchen staff, and fewer customer complaints filed with the Las Vegas health department.

Introduction

The fight against foodborne illness is complicated by the fact that many cases are not diagnosed or traced back to specific sources of contaminated food. In a typical U.S. city, if a food establishment passes their routine inspection, they may not see the health department again for up to a year. Food establishments can roughly predict the timing of their next inspection and prepare for it. Furthermore, the kitchen environment is dynamic, and ordinary inspections merely provide a snapshot view. For example, the day after an inspection, a contagious cook or server could come to work or a refrigerator could break, either of which can lead to a food poisoning. Unless the outbreak is massive, the illness is unlikely to be traced back to the venue.

Figure 1: nEmesis web interface. The top window shows a portion of the list of food venues ranked by the number of tweeted illness self-reports by patrons. The bottom window provides a map of the selected venue, and allows the user to view the specific tweets that were classified as illness self-reports.

*Adam is now at Google.

- Presented at the 2016 AAAI Conference
- Presented at the 2016 NEHA Conference

References

- Harris, J., Mansour, R., Choucair, B., Olson, J., Nissen, C., & Bhatt, J. (2014) Health Department Use of Social Media to Identify Foodborne Illness. *Centers for Disease Control and Prevention Morbidity and Mortality Weekly Report,* 63(32) 681-685.
- Harrison, C., Jorder, M., Stern, H., Stavinsky, F., Reddy, V., Hanson, H., Balter, S. (2014). Using online reviews by restaurant patrons to identify unreported cases of foodborne illness. *Centers for Disease Control and Prevention Morbidity and Mortality Weekly Report*, 63(20), 441-445.
- Jones, Kelsey (November 15, 2013). The Rise of Social Media v2.0. *Search Engine Journal*. Retrieved from https://www.searchenginejournal.com/growth-social-media-2-0-infographic/77055/.
- Sadilek, A., Kautz, H., DiPrete, L., Labus, B., Portman, E., Teitel, J., Silenzio, V.(2016). Deploying nEmesis: Preventing foodborne illness by data mining social media. *Twenty Eighth Annual Conference on Innovative Applications of Artificial Intelligence*.
- Sadilek, Brennan, Kautz, and Silenzio.
nEmesis: Which Restaurants Should You Avoid Today?
First AAAI Conference on Human Computation and Crowdsourcing, 2013

Contact

Lauren DiPrete

diprete@snhdmail.org

(702) 759-1504

Southern Nevada Health District

Senior Environmental Health Specialist