VAST Challenge 2019
Reviewer Guide: Mini-Challenge 3

This document provides information to support peer review of submissions to VAST Challenge 2019, Mini-Challenge 3. It covers background about the submission structure, the challenge problem, tasks and questions presented to participants, potential answers, and evidence found in the Challenge data supporting these answers. For a full description of the challenge problems and to access the data provided to the participants, please visit https://vast-challenge.github.io/2019/MC2.html.

Submissions

Participants are required to submit their entries on a standard answer form, along with a video explaining how visual analytics were used to help solve the challenges. Please consider both parts of the submission in your review. If you have difficulty reading the answer form or playing the video, please contact us at vast_challenge@ieee.org for assistance.

Scenario

The VAST Challenge 2019 presents three mini-challenges and a grand challenge for participants to apply visual analytics research and technologies to help a city grapple with the aftermath of an earthquake that damages their nuclear power plant.

Overview

St. Himark is a vibrant community located in the Oceanus Sea. Home to the world-renowned St. Himark Museum, beautiful beaches, and the Wilson Forest Nature Preserve, St. Himark is one of the region's best cities for raising a family and provides employment across a number of industries including the Always Safe Nuclear Power Plant. Well, all that was true before the disastrous earthquake that hits the area during the course of this year's challenge. Mayor Jordan, city officials, and emergency services are overwhelmed and are desperate for assistance in understanding the true situation on the ground and how best to deploy the limited resources available to this relatively small community.

Mini-Challenge 3: Voice from the People

Seismic and survey data are useful for capturing the objective damage that the earthquake has caused St. Himark. However, this data has limitations. First, official surveys are time consuming and do not stay current in a rapidly changing situation. Second, they don't establish how citizens are reacting to the current crisis. Third, they are often insufficiently granular, providing little insight into differences between neighborhoods. In other words, the seismic and survey data do not provide an up-to-date view of the structural and humanitarian impact caused by the earthquake on a neighborhood-by-neighborhood basis. The City has concluded that this knowledge is necessary to determine where to allocate emergency resources.

City Officials have identified a subset of Y*INT, a community-based social media platform, as a potential source for revealing the current state of St. Himark's neighborhoods and people. Knowing that

participants are skilled in visual analytics, the City has asked participants to analyze Y*INT messages in order to determine the appropriate actions it must take in order to assist the community in this disaster.

Ground Truth

The dataset spans from April 6th to April 11th. There is an embedded day/night cycle that persists across the dataset. The majority of the data is auto-generated based on a hidden Markov model, and artifacts of that model should not be the focus of the results. Throughout the entire timeline there are hand-generated stories with much lower volume but much greater novelty.

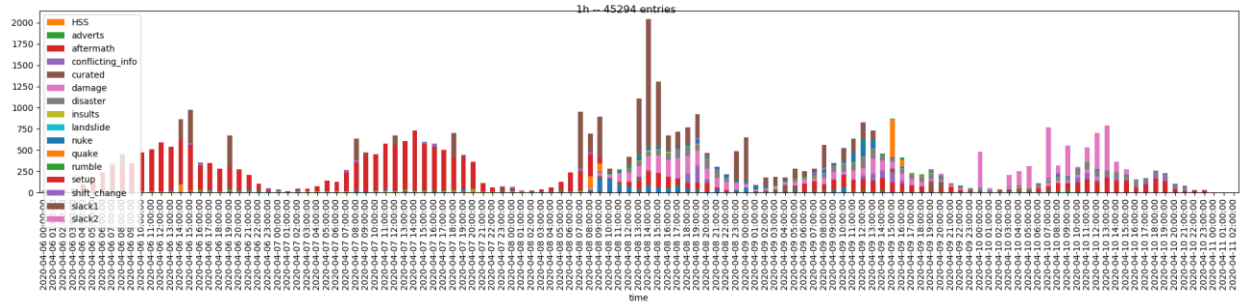


Figure 1. Pattern of messages over time

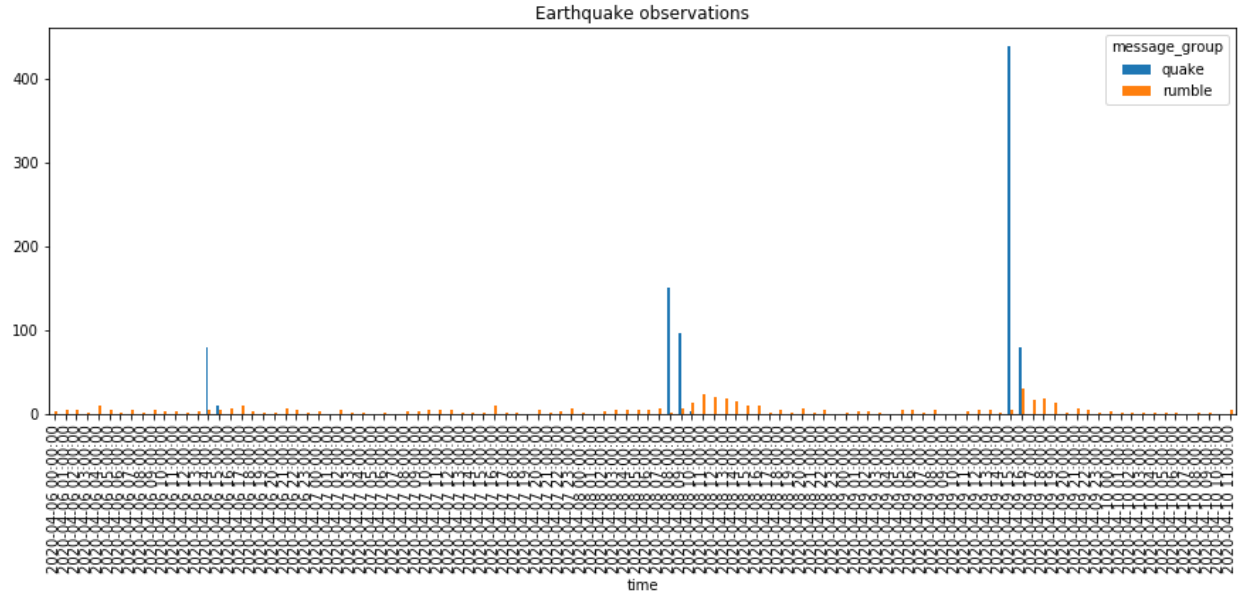


Figure 2. Pattern of earthquakes followed by Rumble data.

There were three earthquakes. A foreshock occurs on April 6, a main quake on April 8, and an aftershock on April 9. The foreshock gets little press, but leads messages about general earthquake safety. The main quake is the strongest, but the aftershock gets the most direct references in the dataset.

Earthquakes are closely followed by messages encouraging the use of the “Rumble” app for reporting. There is also a general background of reminders to get Rumble from an official account.

Questions to Participants

The following are the questions they were asked to respond to. Following each question is information to support your evaluation of their answers.

The Y*INT social media data includes many patterns of life, reports on breaking events, and information about the aftermath of events. There are also some red herrings, jokes and other off-topic content. This guide covers the primary patterns pertinent to this mini-challenge, but the data contains many more interesting patterns than are described here. Consider the evidence that the submitters present. Is it consistent with the types of evidence presented here? Was it discovered in a meaningful/interesting way? Does the submission focus on patterns relevant to the earthquake and its response?

Note: A mistaken version of the readme said the data goes to April 13th. If somebody points that out, they are correct (but it isn't terribly interesting).

As you review the submissions, keep in mind that we are looking for submissions that go above and beyond rudimentary visualization techniques and shallow analyses. We are looking for visual analytics approaches that provide context for the data, support interactive exploration to identify and understand patterns. We encourage creativity. We are also particularly interested in representations of uncertainty in the visualization. Analyses should provide both answers and evidence or rationales, and should demonstrate critical thinking about the data. It is even better if assumptions, multiple competing hypotheses are documented as well.
As you develop your review of the submission, please also consider the following questions.

- If there were hypotheses that remained unresolved, did the submission specify actions to be taken to resolve them?
- Did the team develop an innovative visual analytic tool? Alternatively, did they use an existing tool in an innovative way?
- Did visualizations enable the analysis process? Or did they merely illustrate conclusions? Did the submission rely more heavily on non-visual analytic approaches?
- Did their tool allow useful interactions?
- Did they use all the available data?
- Was the submission clear?

Sample Answers and Notes for Question 1

1. **Using visual analytics, characterize conditions across the city and recommend how resources should be allocated at 5 hours and 30 hours after the earthquake. Include evidence from the data to support these recommendations. Consider how to allocate resources such as road crews, sewer repair crews, power, and rescue teams.**

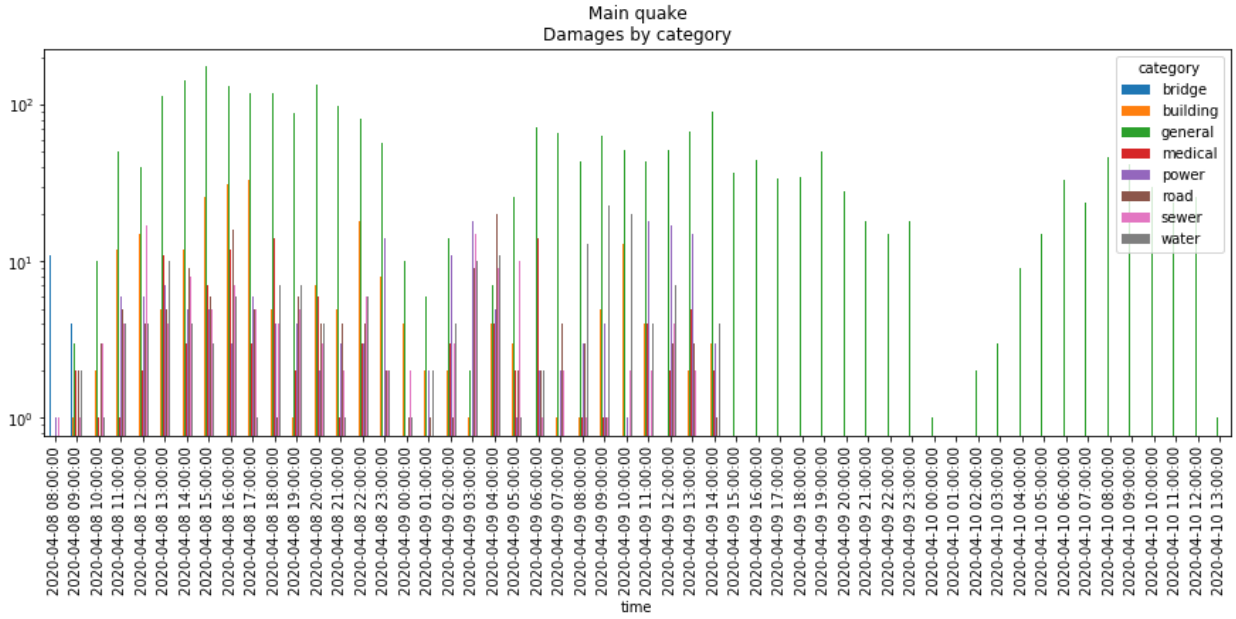
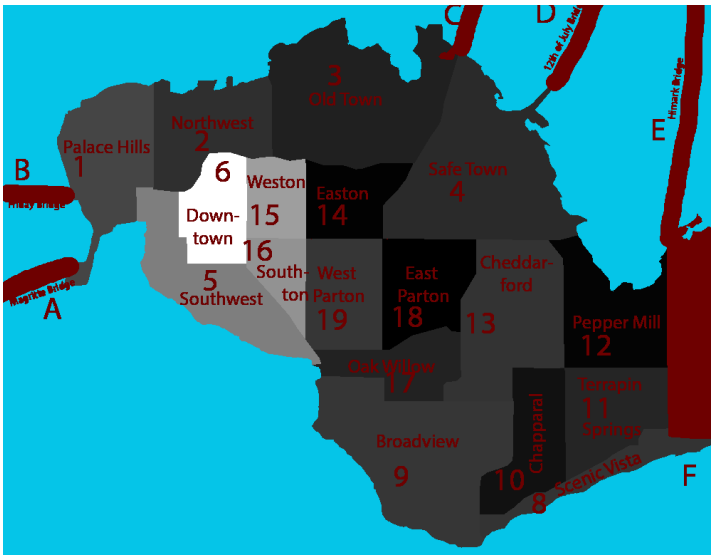


Figure 3. Damage reports by category over time.

Damage reports follow the earthquake events. They are broken into a number of different categories, with the more specific reports closer to the earthquake and more general reports persisting. The quantity follows a general day/night cycle. There is a small echo of repetition following the aftershock, but much smaller than after the initial quake. There are also messages about liquefaction/landslides after the aftershock.

The different damage types (building, sewer, water, etc.) each follow slightly different temporal patterns (building damages and road damages being reported over a wider span than water/sewer issues). They follow geospatial patterns driven by the population distribution. Comparisons between places, between times or between damage types are all valid.



There are some location-specific damage reports as well. For example, a significant story is the closure/rumored-collapse of the bridges (large and small) around St. Himark. The map to the left shows the density of those messages (darker means more messages). Regions with more bridges are more active than regions without them. For example, down-town has very few bridges, while Pepper Mill has many and is adjacent to the Himark Bridge (Bridge E). The Wilson Forest (region 7) is experiencing a blackout at the time this map was created.

Figure 4: Relative density of bridge closure messages.

Other events to be aware of include:

- Concerns about architecture in Old Town (region 3)
- Liquefaction and landslides concentrated in Wilson Forest, Scenic Vistas, Chapparal, and Terrapin Springs (regions 7, 9, 10, 11)
- Nuclear concerns start near the plant in Safe Town (region 4) and then generally rise all over town.
- Fires in the more densely populated regions.

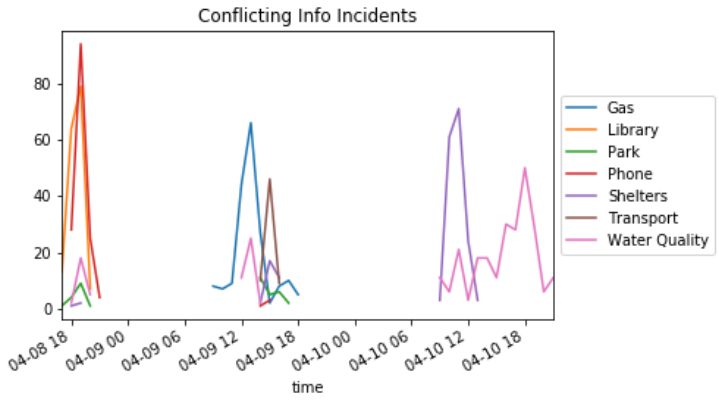


Figure 5. Conflicting information over time.

There are a few times when information is in conflict in the dataset about a specific topic. Those are illustrated in the figure to the left. For example, in the "Parks" incidents, people report that space in parks is both available and that they are full. Similarly, the Library times include reports that libraries are open as shelters while others concurrently report that the libraries are closed.

Sample Answers and Notes for Question 2

2. Identify at least 3 times when conditions change in a way that warrants a re-allocation of city resources. What were the conditions before and after the inflection point? What locations were affected? Which resources are involved?

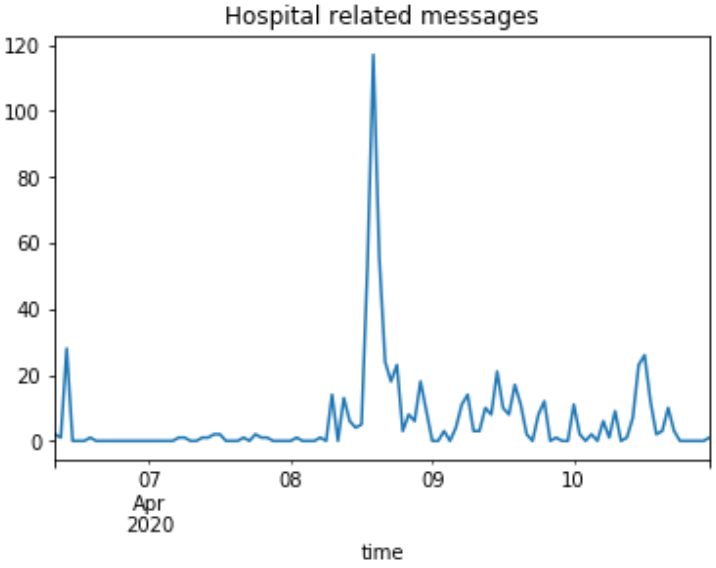


Figure 6. Hospital messages peak following the April 8 request for more information.

The evidence for this this question is closely related to the evidence for question 1. While question 1 looks at absolute conditions, this question is looking for comparisons between two points in time. Significant events can be pulled from question 1, but the responses should focus on the contrast between time points.

An example of a before/after behavior change can be seen in the figure to the left. This is an illustration of discussion of hospitals. There are two interesting events. First, before the quake, there is very little mention of hospitals and health care, while after it is a consistent part of the chatter. Second, the large

spike late on the 8th is driven by a public safety request for more information about the hospitals and a flood of citizen and employee reports.

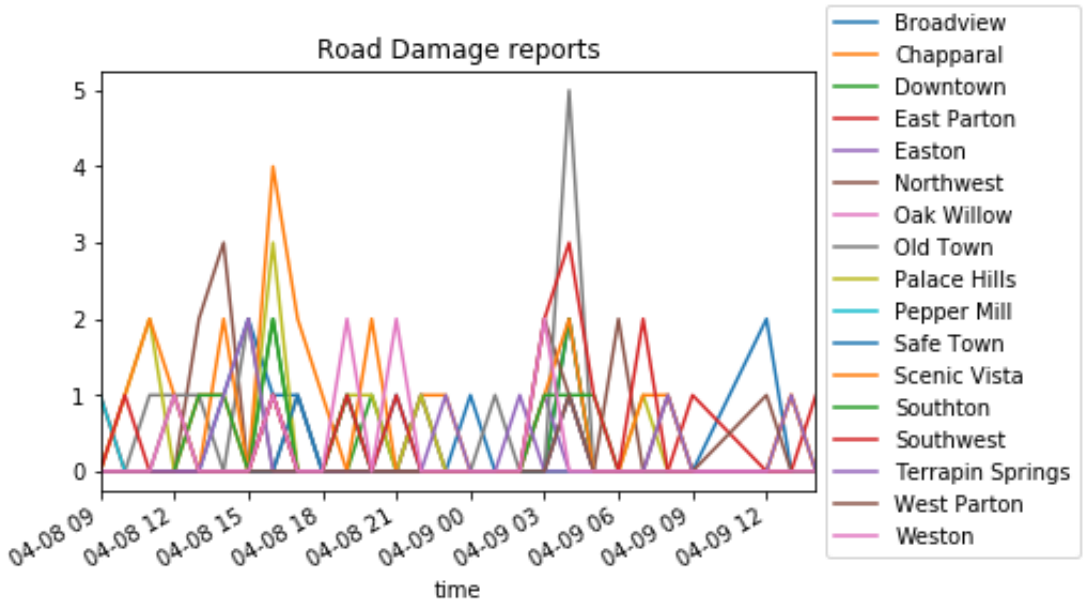


Figure 7: Road damage reports by region over time

The most natural inflection points are shortly before and after the various earthquakes. There are public safety announcements that are responses to other key events in the dataset. These may be cited as evidence of resource redistribution events. Example events have evidence in the dataset include:

- Road damage related to liquefaction after the aftershock (see Figure 7)
- There are discussions of animals (dogs, cats, etc.) in the streets at several points having escaped from various shelters.
- Citizens are concerned about nuclear contamination grows after each quake (see details in question 3).
- Several areas have increased reporting of building damage and the need for people to be rescued.
- Group/elder care facilities express shortages of supplies

Sample Answers and Notes for Question 3

3. Take the pulse of the community. How has the earthquake affect life in St. Himark? What is the community experiencing outside the realm of the first two questions? Show decision makers summary information and relevant/characteristic examples.

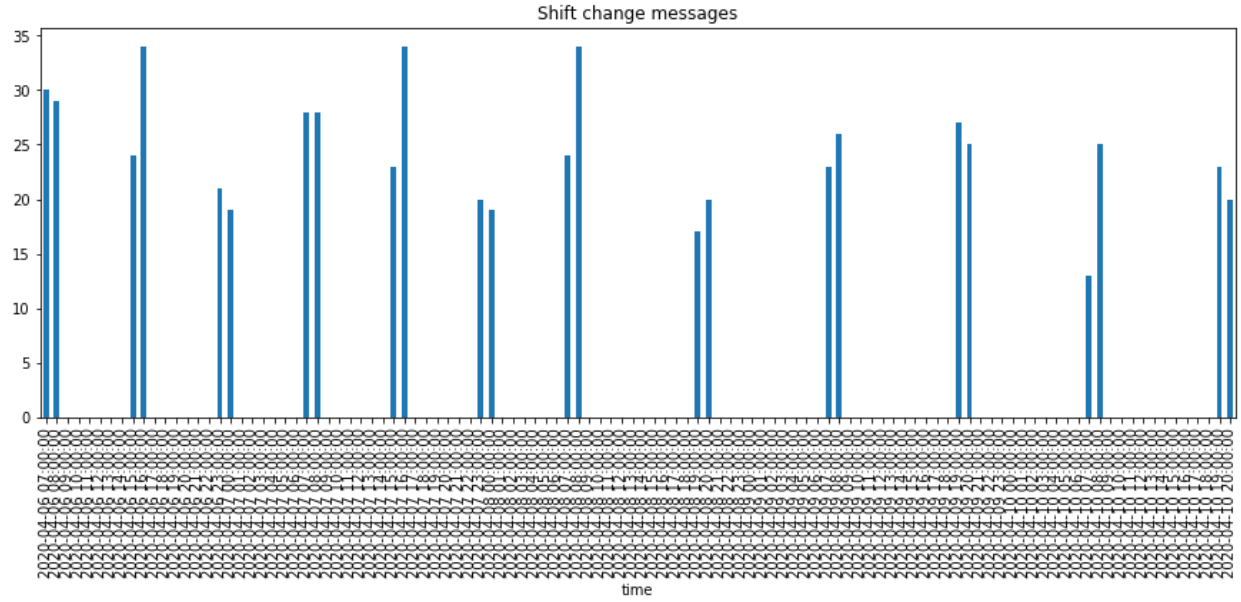


Figure 8. Shift change messages over time.

Pre- quake the nuclear plant is on an 8 hour shift. After the quake it is on a 12 hour shift. Workers use a variety of names for the power plant ("Always Safe" is the cannonical one, but the workers have others). This is the most clear example of changing work behaviors. Theses messages are mainly in Safe Town (region 4)

Additionally, the population distribution during the day changes post-quake (as evidenced by the geographic distribution of messages). The pre-quake daytime message distributions are very similar to each other and fairly distinct from the evening distribution. Post quake, the daytime distribution looks more like the night-time distribution.

There is a shift in the vocabulary used pre-quake vs. post-quake. What are people talking about and where they are should be rich targets for comparisons and should form the bulk of evidence for where to allocate resources. There are two shifts bulk shifts in vocabulary:

- Sentiment shifts from generally positive pre-quake to generally negative post-quake
- Attitudes and focus shifts from long-term (neutral/happy) to more immediate concerns to more long term (questioning) as the timeline progresses

Other shifts that may make an appearance:

- The pre-quake Y*INT stream includes many references to civic events (there is a "Lacki Dasicle" concert, a circus coming to town and the "annual pipe and drum competition"). After the quake, these are referred to, but the events are canceled.
- The Department of Transportation provides traffic reports pre-prequake and looks at traffic condition, post-quake it is about road conditions and bridges.

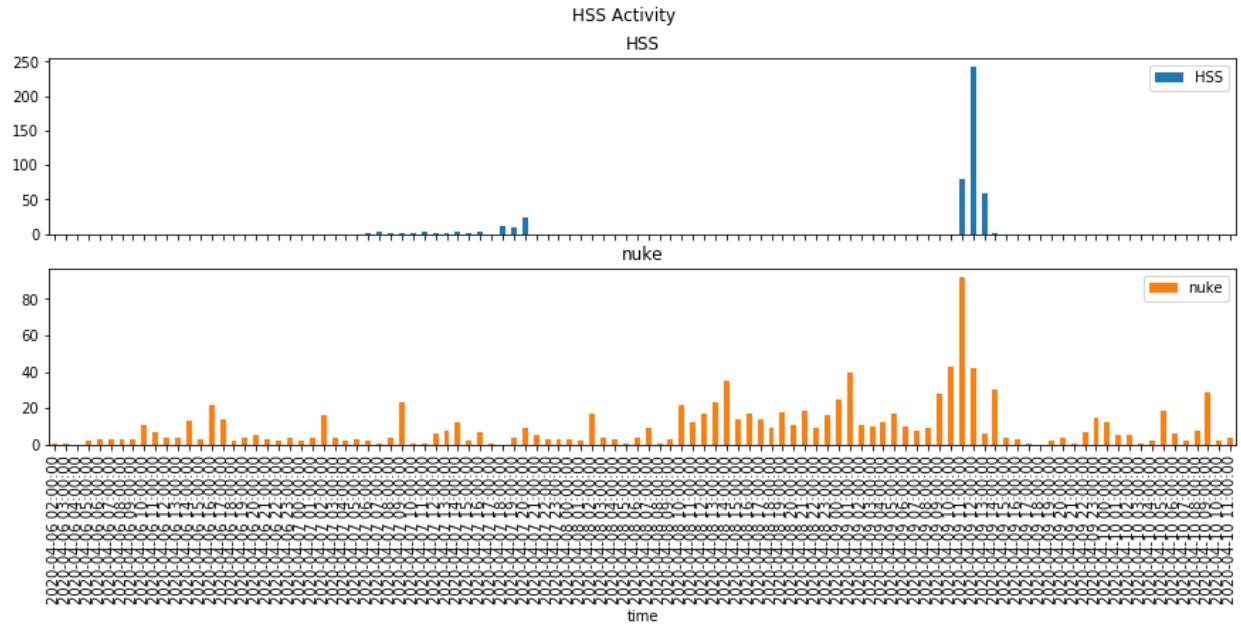


Figure 9. Activity by the retired scientists group HSS.

HSS is a retired scientists group in town who are active in promoting citizen science. They have a regular meet-up that happens shortly before the quake (with Y*INT activity increasing beforehand). After the quake, people are hearing rumors of nuclear issues and talking more and more about them. HSS teams up with the city and to put out a press release. This team-up gets a great deal of social media attention, and generally calms the growing fears. Key HSS accounts are: WatchAmbroseLEarn, CanAdoreThedore23, WieldsMAttoc, JoliotKnows4, PeterLovesCrystal54, RadCurieus, GermanWrinkler43 and CookWhatWhen81. The accounts may be cited as influential around HSS and nuclear concerns.

Other examples of shifting concerns include:

- Changes in how parks are being used (music festival vs. tent city)
- Scarcity issues around fuel
- Poorly functioning phone systems
- Where to get health care

Sample Answers and Notes for Question 4

4. The data for this challenge can be analyzed either as a static collection or as a dynamic stream of data, as it would occur in a real emergency. Describe how you analyzed the data - as a static collection or a stream. How do you think this choice affected your analysis?

Responses may vary. We are looking for thoughtful answers regarding how the form of the data may affect their insights. Do they assert that patterns may be easier to identify in streaming or static data? Do they provide rationale for their assertion?

This question also invites reflection on how data processing decisions impact the answers produced. Things you might see addressed include:

1. How was sequence of events used in the answers?
2. How do they think sequence impacted the answers?

3. Does answer confidence/accuracy change over time?
4. How many passes over the data were required to get their answer?
5. If they used machine learning techniques, did they use time in setting up training data?

Our concern here is that they are introspective about their own process and recognize the impact of the choices that they made.

Other Notes on the Dataset

There is little significant information in the network structure of this dataset.

Not all stories are relevant the challenge problems. There are a number of stories in the data that are present but generally not useful. For example:

- The world never stops advertising. Since dataset represents a subset that was deemed relevant, these advertisements tend to refer to quake-related words ("Shake up your decorations with our amazing sale!").
    - Advertising comes from accounts named: Syndicated4, Syndicated5, Syndicated348, _____3333_____, J0rdanWantsBacon, JordanWantsBac0n, CantonCoordon2, Opportunities2, Opportunities1
- Social messaging platforms aren't always polite. There is a set of randomly generated insults sprinkled through the dataset.
- A number of Y*INT accounts have running jokes between them.

Stories like these may show up in submissions. If they enrich or are used to contrast a larger point, that is a generally positive point. However, they should not be the focus of an answer. Use your judgement to decide if the way the story is found and how it is used to support a question's answer is useful to the spirit of the question.

Account content comes from several different sources. There are three main categories of accounts. Table 1 lists relevant includes city, media and utility accounts. These accounts are key to understanding the operation of the city. Table 2 lists hand-generated accounts for general citizens. These may be important in the analysis as they are often the origin of various stories. The third category of accounts is the majority of accounts and they are built an automated process. These accounts amplify hand-generated messages, but may have artifacts of the generation process.

Table 1. Important city, media, and utility accounts

StHimarkEMA	Emergency Management System (EMS) and Emergency Operations Center (EOC)
DOT-StHimark	Department of transportation
StHimark	General City account
Survey-SHM	Dept of land use
SHM-WaterSewer	Water & Sewer
CityHealth-SMH	City health department
KRAKTV	Local TV station
AlwaysSafePowerCompany	The company that runs the nuclear power plant
SHM-PowerC	The power utility company
ThesaurusGasCompany	Gas utility company
SHM-SaveAnimals	St. Himark Animal rescue
StHimark-Learns	School system

Table 2. Informative Individual Accounts

Sal	Nicky77	WhatNowHow	theaterlover34
kyle	ElwinJR	Hiker4Life	AlwaysUpToEat34
You2	Runner23	TakeABreak	FrancineHiggins
ICYou	Rescue14	Mommy2Four	FireCaptainJack
Doood	Rescue12	CityEvents	LivesToHike79
firefighter87	SunnyDay	RunningTree	Mrs.Tillenbottom
Bopper	LoopyLou	HurtsToPurr	IfIOnlyHadATrain
Booper	DoNotWant	OldTownGirl	MyNameIsYourName2
BBGirl	RoseWyatt	BoogieBunny	HimarkGrocersAssc
Olivia	EllaChang	Bonanza4Evr	RhyminSimonJBread
BarryB	CiloFlame	GreenTMouse	StHimarkMusicFest
Milo99	MindyLevy	ChapparelFD	TommyNotThatTommy
Oppsie	wftrubble	Bonanze4Evr	DareYouToPetTheCat
Westie	RayOfRain	localburger	HopeSpringsEternal
ghosty	OtherJoey	BroadviewBoi	ArtistLalisaManoban
kristen	BoringJoe	ComeBackJack	retiredfiremarshall
ImNotUR	Phone4Life	SlimRodents2	AlwaysSafePowerCompany
Chappie	DerekNolan	biscuitJtoot	StHimarkMutualInsuranceCo
ElwinJr	NoFearCher	VanessaCorwin	UnitedChurchOfJohnsonville
CityEOC	WalkerTree	LivesToHike79	EmielRegis
BB Girl	RockyFlame	2NDammendment	FireCpn