DC4: Exploratory Data Analysis
Group: Adriana Beltran Andrade, Catherine Park, Casey Perez, Rachel Yan, Elaine Ye

Disaster Management in St. Himark

Introduction:

St. Himark, with a population of 246,830 people, is a city that has a place for everyone. The city is divided into nineteen neighborhoods, each having a unique community and lifestyle. St. Himark is guided by Mayor Jordan and the city council, and each council member holds the responsibility of meeting the citizens' needs. One of the most significant locations in St. Himark is the Always Safe nuclear power plant, the primary provider for electricity in the city. The power plant also employs a large number of citizens and provides revenue for the city, making it an important part of St. Himark. Additionally, the Always Safe power plant is carefully monitored for dangerous radiation. However, as a result of the recent earthquake in St. Himark, disaster has hit almost every neighborhood, has caused a potential leakage of radiation from the power plant, and has caused the scarcity of resources. The entire community has been shaken, with city council and emergency services scrambling to assist those who need it most.

Methodology:

While considering the extent of the damage done to the city from the earthquake, it was essential to narrow down the communities in the most potential need. Direct damage from the earthquake, contamination of food/water resources, and lack of shelter were all points of discussion that had to be narrowed down from the social media data set. Key topics were identified based on official messages on social media in order to create a timeline as to what may be happening in the city. Damage reports were used along with this timeline to specify which communities may have been hit hardest by the earthquake. It was also important to consider the possible radiation contamination from the Always Safe power plant in order to evacuate to possibly affected communities. The sensor data helped to narrow down which neighborhoods may have greater risk of radiation exposure through the location data provided by both the mobile and static sensors as well as the extent of the radiation contamination that has occurred.

Data Visualizations & Analysis:

Radiation Sensors (Mobile & Static):

The earthquake that hit all of St. Himark clearly appears to have damaged the Always Safe power plant, leading to dangerous leakage of radioactive material. Due to the construction of the power plant originally not meeting international standards, the plant was carefully monitored with static sensors throughout the city and supplemented by homemade mobile sensors that are attached to citizens' cars. These mobile sensors were originally designed to educate the public and reassure citizens that there was no risk associated with nuclear power. These mobile and static sensors are now being used to monitor the city for potential radioactive leakage.

There are nine static sensors that are located throughout St. Himark. As shown on Figure 1**,** Palace Hills, Southwest, Downtown, Broadview, Old Town, and Cheddarford each have one sensor located within the respective neighborhood while Safe Town has three sensors. Most of the sensors are located in neighborhoods towards the northern part of St. Himark, likely due to their proximity to the Always Safe power plant.

Based on the static sensors, on Figure 2, we were able to take the total and average of the readings for each unique sensor over the 5-day period. The average radiation readings increased as more time passed across all sensors, indicating that the risk of radiation contamination was becoming steadily worse. It was also noted that Sensor 15 went completely offline on April 9th and returned on April 10th, but began to collect readings at a much lower pace, indicating potential damage to the sensor or corruption of data. Sensor 15 is also located directly where the Always Safe power plant is, potentially indicating a possible concern of greater damage in the power plant.

Furthermore, there are 50 mobile sensors that move throughout every neighborhood of St. Himark, each attached to unique cars. Each of these cars appear to have designated paths which are followed in a nearly-identical manner every day, barring any issues. As each car moves around, the sensors collect counts of radiation particles per minute (otherwise known as CPM), which are recorded every 5 seconds in the data. From a brief investigation into the average CPM levels across all mobile sensors for each day, we came across nine sensors that recorded abnormally high CPM levels: 10, 21, 22, 24, 25, 27, 28, 29, and 45 [Figure 3]. It was also noted that three sensors (34, 48, 49) go offline on April 9th and 10th.

Regarding Sensor 10, abnormally high CPM levels (100+) were noted to begin at around 16:00 on April 8th, where the car appears to remain at the entrance for Jade Bridge (on the border between the Old Town and Safe Town neighborhoods) [Figure 4]. These readings began to peak at around 18:00 (consistently 1000+), ceasing altogether at around 22:00 and returning to more regular levels as the car drove back into Old Town. This pattern of readings by Sensor 10 indicates the estimated initial time of radioactive leakage arrival around the Jade Bridge entrance area, as the readings around the area prior to 16:00 were relatively average. From this, it is likely that the Always Safe power plant began to leak radioactive material earlier on April 8th, spreading throughout the neighborhood of Safe Town as the day progressed. A reading from Sensor 22 potentially supports this idea; when the car for Sensor 22 was at the bottom border of Safe Town at around 12:00, there was a reading of 900 CPM. The rest of Sensor 22's readings were relatively standard, so it is likely that this reading is reliable and not due to the sensor malfunctioning. In the aftermath of these readings, it was noted from the social media data that parts of Old Town had been told to not enter their homes for fear of destabilized homes/buildings. However, it was not made directly clear if the decision to prevent citizens from entering their homes may have also been motivated by potential radioactive contamination.

As for the remaining eight abnormal sensors, each exhibited unusual patterns of movement and CPM readings beginning on April 9th. At around 6:00, Sensor 29 began displaying unusual movement patterns unlike those seen from it on the previous two days, potentially as a result of reading higher CPM levels in the hours prior. The car for Sensor 29 drove toward the Wilson Forest Hwy (on the bottom border of the Wilson Forest neighborhood), going all the way to the edge of the city's limits. It remained there until around 19:00, when the car drove to the inner-city end of the highway and stayed at that spot for the rest of the day. Sensors 21, 22, 24, 25, 27, 28, and 45 all also displayed similar patterns of movement, with each sensors' car driving to the edge of the city's limits in the hours following Sensor 29's departure. By around 10:30, all of the aforementioned sensors were gathering readings at the edge of the city's limits, staying there for the vast majority of the day. Some of these sensors also followed suit with Sensor 29's movement later in the night, moving to the inner-city end of the highway from 19:00 onward and gathering readings there for the rest of the day. From this, we can assume that there was some large incident that began around or before 6:00 AM on April 9th, potentially causing the residents of the city to be evacuated throughout the following hours.

All of the sensors that moved to the inner-city end of the Wilson Forest Hwy recorded high CPM levels (1000+) throughout the night [Figure 5], potentially indicating that radioactive particles were present throughout most of the city by the end of April 9th. Those living in the westernmost neighborhoods may not have been exposed to radioactive materials by this point, but it is likely that the rest of the neighborhoods have been at risk since then, if not as early as April 8th. The data provided by Sensor 10 on this day supports this idea, as it also appears to have evacuated the city through Jade Bridge at around the same time as the other sensors mentioned above. However, just like Sensor 29 and a few of the other mobile sensors, it drove back into the city during the nighttime. Sensor 10's readings during this time were relatively standard, remaining below 100 CPM; however, these readings were still higher, on average, than the readings that it picked up in the early morning hours of April 8th. This may indicate that the radioactive materials quickly spread throughout the city after the assumed evacuation in the morning of April 9th, persisting through the night.

On April 10th, all of these sensors continued to display irregular movement. This was likely due to roadblocks being placed at the entrances to the city as a result of either damage or concern for radioactivity. Each sensor continued to gather readings at similar levels from the day prior. Sensor 10 continued to have sub-100 CPM readings throughout the day as it moved through the city, indicating that radioactivity was still likely present throughout the city. It appears to have had to evacuate once more at around 9:00, remaining on Jade Bridge until around 20:30, when it drove back into the city and gathered readings from then on. The other eight sensors also appeared to have evacuated by 9:00, ceasing readings for the rest of the day after exiting the city. That being said, most of the readings by these sensors prior to the time of evacuation recorded very high CPM levels (1000+). From this, it is likely that there was substantial damage and/or radioactivity within the city; it can be assumed that most of the city's inhabitants were outside of the city's limits or seeking refuge within emergency shelters for the rest of the available time frame for safety.

Despite this analysis, there are also limitations on the information we had. It was noted that due to the disasters, monitors may have been potentially disrupted; thus, some of the readings are likely unreliable/incorrect. However, it was difficult to rule out any outliers due to the sheer amount of high readings. We were only able to rule out two readings in which there were negative values for CPM, which we believed to be unlikely to have happened. All other outliers were kept in the data due to concerns about potentially taking out readings that are correct and should have been included. Additionally, interpretation of the unit CPM is difficult due to it being a measure of counting the detection events of radioactive particles. This means that we have no indication of the actual strength of radiation present, only gauging whether radioactive particles are present or not. That being said, higher levels of CPM (100+) indicates a higher risk of radiation exposure; values reading 1000+ are urgent warnings for evacuating the affected area in order to reduce risk of harm to the city's inhabitants.

Social Media:

Assuming that messages announced by "KRAK TV" and city departments were reliable, we created a rough timeline that identifies major events since the magnitude 6.7 earthquake that happened at 8:36 am on April 8th (Figures 6 & 7). The timeline is referenced and filled in with more details when cross-validated in other datasets. The earthquake has had devastating effects on the infrastructures in the city and citizens' daily life. At 9:00 am on April 8th, multiple bridges were closed due to damage, followed by computer system damage at 10:15 and power outage at 15:00. What's worse, floods and water contamination due to broken water and sewer pipes were reported in multiple neighborhoods since the morning. As evidenced in the YInt messages, it was a difficult situation for both rescuers and citizens because 1) the roads were blocked by closed bridges and rubbles, which slowed down the rescuing/repairing process and prevented people from getting food; 2) people couldn't report local damages on the Rumble app when the computer system was down, which could potentially hinder timely emergency response; 3) the floods resulted in many people in need of shelter; 4) the broken water and sewer infrastructure could require months to repair, but the current water contamination has led to water shortage and bacteria infection/poisoning due to contaminated food as reported at 12:30 on April 9th, which could develop into a public health crisis if not addressed appropriately; 5) houses in danger were red-tagged and residents could not enter them, creating discontent among residents who had not found a shelter yet. In terms of government response, all libraries were open as shelters at 15:00, April 8th; The Federal rescue team arrived on 2:00, April 9th, and 80% of power and internet was restored at 7:00 on the same date.

Here's a more in depth look at the issues including water contamination, floods, damaged infrastructure, and residents seeking shelter according to the YInt messages. In the Downtown and Old Town area, there was a large amount of reposts for messages expressing the lack of room at disaster shelters and that some libraries were not offering shelter. Though YInt we can see that in both areas, survivors are sheltering in parks even after being told they were not allowed to by city sources. It wasn't until April 9th that the city decided to allow people to camp at parks. However, due to a rainstorm and the chance of flooding, the city decided to vacate the parks to the shelters by the morning of April 10th. On April 10th, we also saw that the city established Emergency Distribution Centers in each neighborhood local community center.

Due to the earthquake and existing problems with the infrastructure, there was a broken water main that flooded areas such as Old Town and created more difficulty for people to evacuate buildings and find shelter. On April 8th, a sinkhole was also created on a main street of the Oak Willow neighborhood with suspicion of a water main burst. The sinkhole collapsed, taking several cars immediately and more as the hole kept filling with water. The retweets and messages suggest that districts around OT and OW such as Broadview, Weston and Easton have been impacted by the floods.

On April 8th, a recommendation of boiling water due to risk of contamination from the broken pipes was sent out for the neighborhoods of Old Town, Safe Town, Scenic Vista, Broadview, Chaparral and Easton. Members from other neighborhoods were voicing their concerns about their drinking water supply and people in their neighborhoods becoming sick. There was also confusion revolving around which of the recommendations should be followed, as they were recommended to turn off their gas, something necessary to boil their water. On April 9th, two more health announcements came out: one warning that the extent of damage to the water and sewer system is more severe than initially thought and that all St. Himark residents should boil their water, and the other announcing an outbreak of food poisoning from volunteer kitchens.

A sentiment analysis was created from the entries from YInt we had from this incident to further analyze the general feelings within certain areas during this crisis [Figure 8]. Sentiment analysis analyzes the emotion of text by identifying words with emotions and produces positive/negative scores. We could see negative emotions peaked after a few hours of the earthquake in multiple neighborhoods, especially the neighborhoods affected by the floods. Interestingly, we see distinct magnitudes of positive emotions around 12:00 on April 12 when the famous singer Lacki Dasical, who was thought to be missing, appeared in the media.

That being said, several uncertainties in the social media messages should be noted. First of all, we established the timeline based on the assumption that messages containing "KRAK TV" and city departments were reliable, but we are not 100% certain since these messages were sent by different accounts and the locations of the senders might not be at the place where the events took place. Secondly, we were not sure if people retweet because they sympathized with the original sender or because they experienced the same troubles within their local communities. Third, while we acknowledge that individuals' reports on social media are very likely to represent their true personal experience, it is hard to measure the magnitude and range of damage across communities solely based on social media, given that only 4901 accounts use social media in a city with 246,830 people.

Damage Control:

Looking at the damage rating across all five days, we made the following observations. Although the earthquake took place on April 8th, issues related to the water and sewer system were spotted on April 6th between 13:00 and 17:00 in several neighborhoods [Figure 9]. On April 7th, damage ratings exhibited a more random pattern, fluctuating across a wider range (from 0-10) across most of the neighborhood.

Starting on April 8th at around 8:00, the damage reports became more active, which was reflected by more consistent damage ratings and clusters of data [Figure 10]. Interestingly enough, there was evidence indicating a server malfunction, so certain neighborhoods (such as Old Town, Chapparal, and Terrapin Springs) had a halt on receiving reports from the citizens around 12:00. The missing report may be a result of a power outage, which corroborated with what people have reported on social media. Looking more closely at the reported damage in each neighborhood, we found that: Easton, Northwest, and Oak Willow had reported most of the damage on their roads; Broadway, Safe Town, and Scenic Vista seemed to be primarily experiencing issues with their water and sewer system; and East Parton, Oak Willow, Pepper Mill, Terrapin Spring and Southwest reported most damage on their power issues. All neighborhoods reported damage to their infrastructures to some extent, with those near the epicenter of the earthquake being more impacted.

The aftermath of the earthquake continued to affect the neighborhoods on April 9th, with consistent reporting coming in after 15:00. Certain neighborhoods, such as Broadway, took a heavier toll with their water and sewer system, roads, and power being damaged/unavailable until the next morning.

Things started to quiet down a little bit on April 10th, with less damage reports received and damage rating having more random variations. The decrease in reports could be due to the evacuation that took place in the morning. Although neighborhoods were slowly recovering from the earthquake, places such as Broadway, Downtown, Scenic Vista, and Weston were sending out more reports than other areas.

Notably, it seemed that Wilson Forest had the least amount of damage reports received across all five days. In fact, the server did not pick up any activities for Wilson Forest on April 7th. There are three possible explanations. First, the neighborhood itself could be less populated and thus had less individual reporting on damage. Second, Wilson Forest could have been less impacted in general, considering it was in a relatively secluded area. Third, given that citizens must rely on apps to send in updates, inhabitants of Wilson Forest may be less comfortable with technology, meaning they would not be able to report the damages they saw. It was also shown in the social media data that Wilson Forest in general sends out less messages, supporting the first or third idea.

Emergency Response:

In response to all the issues that unfolded in St. Himark and the limited amount of resources/emergency personnel available, it is essential to prioritize for those who need it most. People in immediate danger, such as areas that have suffered great destruction or floods and may have physically trapped/harmed citizens, need to be addressed first. Emergency responders need to focus on areas that have been noted to take spikes of damage in terms of buildings, roads, etc. Areas with severe reported damage on the water and sewer system should also be prioritized, given that water access is essential to a citizen’s survival. Afterwards, those without shelter/who are unable to return to their homes need to be given priority aid due to many of them having to leave their homes without any supplies and facing the risk of more potential rainstorms and flooding. Neighborhoods that are facing great amounts of destruction and with citizens who may struggle to communicate their need for aid (such as neighborhoods with older populations or with a greater proportion of children) should also be addressed as soon as possible.

On the other hand, this earthquake has exposed issues that could guide the city to prepare for potential natural disasters on a regular basis. First off: the basic infrastructures should have regular maintenance to ensure proper functions. For example, the water and sewer system had shown damages before the earthquake and were not addressed in time, leading to an exacerbated situation as the earthquake hit. In addition, since bridges are crucial to connect many parts of the city, their qualities need to be ensured. Alternatively, the city could consider having resources delivered by air to avoid constraints on ground transportation. Second, the city should have regular earthquake drills at homes, schools, and workplaces so that people know what to do and where to go should one hit. Meanwhile, potential shelters should be designated and communicated with the citizens. Third, while the Rumble app and the YInt app to some extent provide useful information to identify needs in the communities, we should be aware that people who have digital literacy and are comfortable with technology might not be representative of the whole population, and moreover internet damage could make these apps dysfunctional. To make the YInt app more effectively communicate and reflect ongoing situations, the city departments and major media outlets could have official accounts to spread reliable updates and plans. In terms of the Rumble app, we propose that the app could go beyond reporting and enable citizens to view overall damages in different places. Each block or small sector in the neighborhood should have community workers to keep people updated, especially vulnerable populations. Finally, the city should consider having more objective reports on disasters, such as using drones with cameras to scan the disaster zones, to better educate and inform the public.

In regards to the Always Safe power plant, the city should continue to keep a close eye on the CPM levels detected by both the static sensors and mobile sensors throughout the city in order to gauge the risk of radiation sickness. If reports of 100+ are recorded, the city should consider evacuating the residents in the immediate area until further notice. If reports of 1000+ are recorded, the city should immediately evacuate the residents of the immediate neighborhood and those surrounding it, preparing them for long-term housing in case the radiation levels are found to be strong. Constant upkeep of the sensors are important to maintain accuracy since it is vital to take immediate action once radioactive particles are detected to ensure the safety of the people. That being said, unnecessary evacuations should be avoided to keep the public sentiment high. Providing these radiation reports to the public is essential to maintain transparency and allow citizens to prepare for evacuation ahead of time.

Conclusion:

The earthquake that has struck St. Himark and caused massive damage and disruption has led to City Council and emergency services scrambling to meet everyone’s needs. All three datasets were used to help map out what is going on in St Himark and pinpoint what may be the priority issue to mitigate and where resources need to be allocated. It was also essential to consider future steps and possible methods or changes to prevent another scenario where understanding the extent of the situation is difficult. Small changes can be made to the Rumble app, more sources of information can be added for the wider population, and better regulation of infrastructure can greatly help for potential future disasters. Overall, both the analysis and considerations for the future are essential for ensuring the safety of the citizens in St. Himark.

Data References:

1. Damage Report Files
2. Sensor Data and Map Files
3. Y*Int Social Media Data Files
4. St Himark Description File

Appendix:

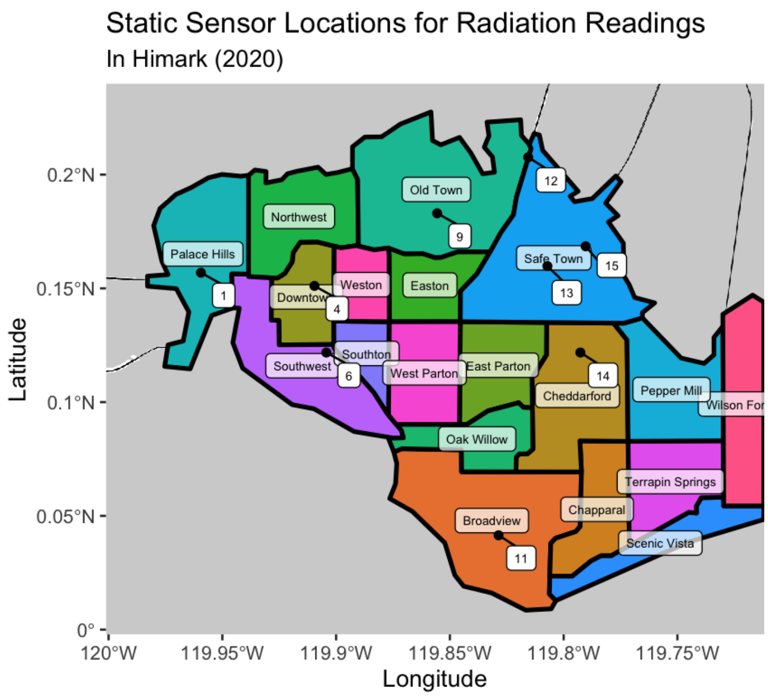

Figure 1. Map of St. Himark and locations of Static Sensors.

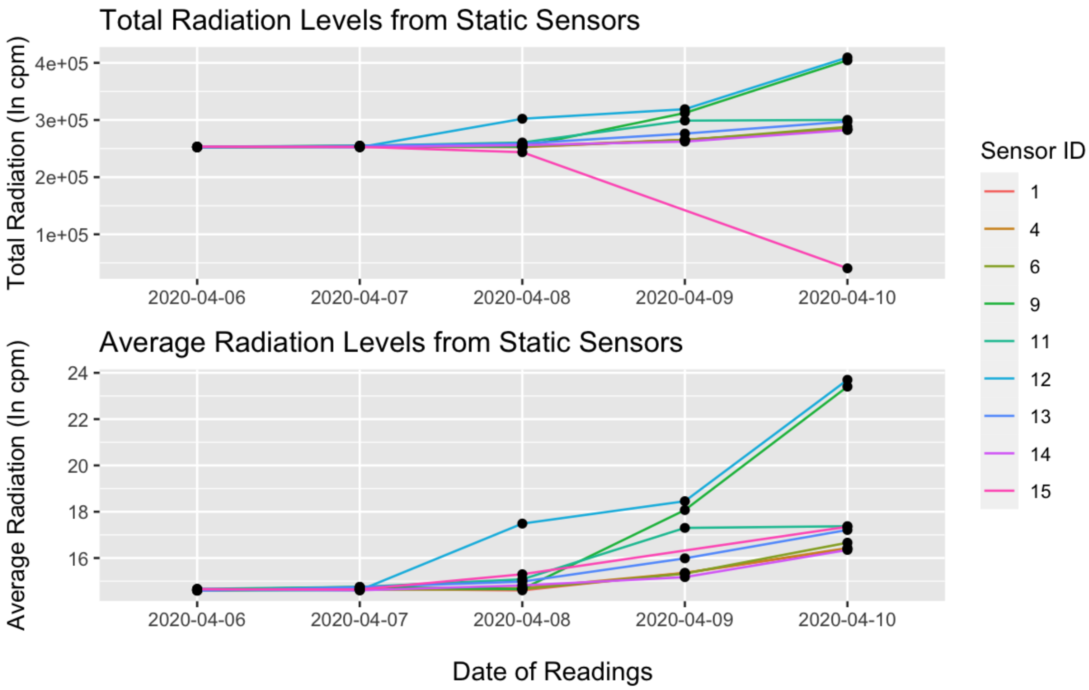

Figure 2. Total and Average Radiation Levels based on Static Sensors.

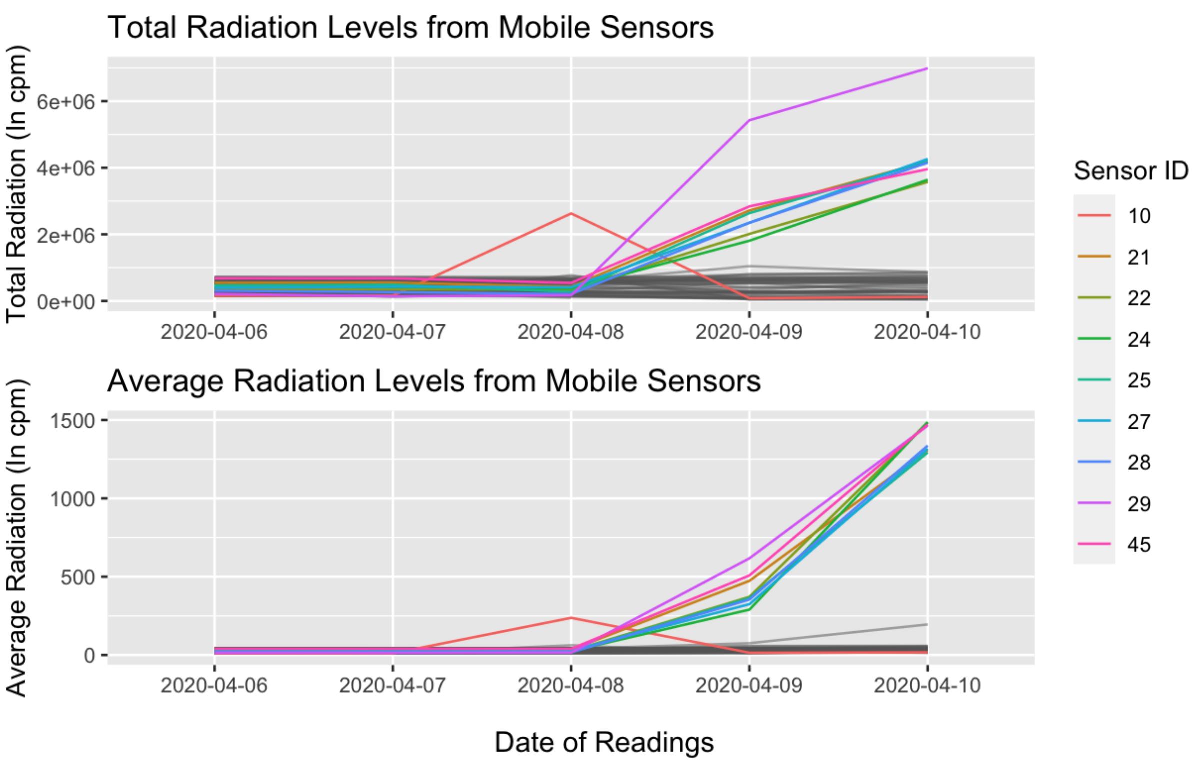

Figure 3. Total and Average Radiation Levels based on Mobile Sensors.

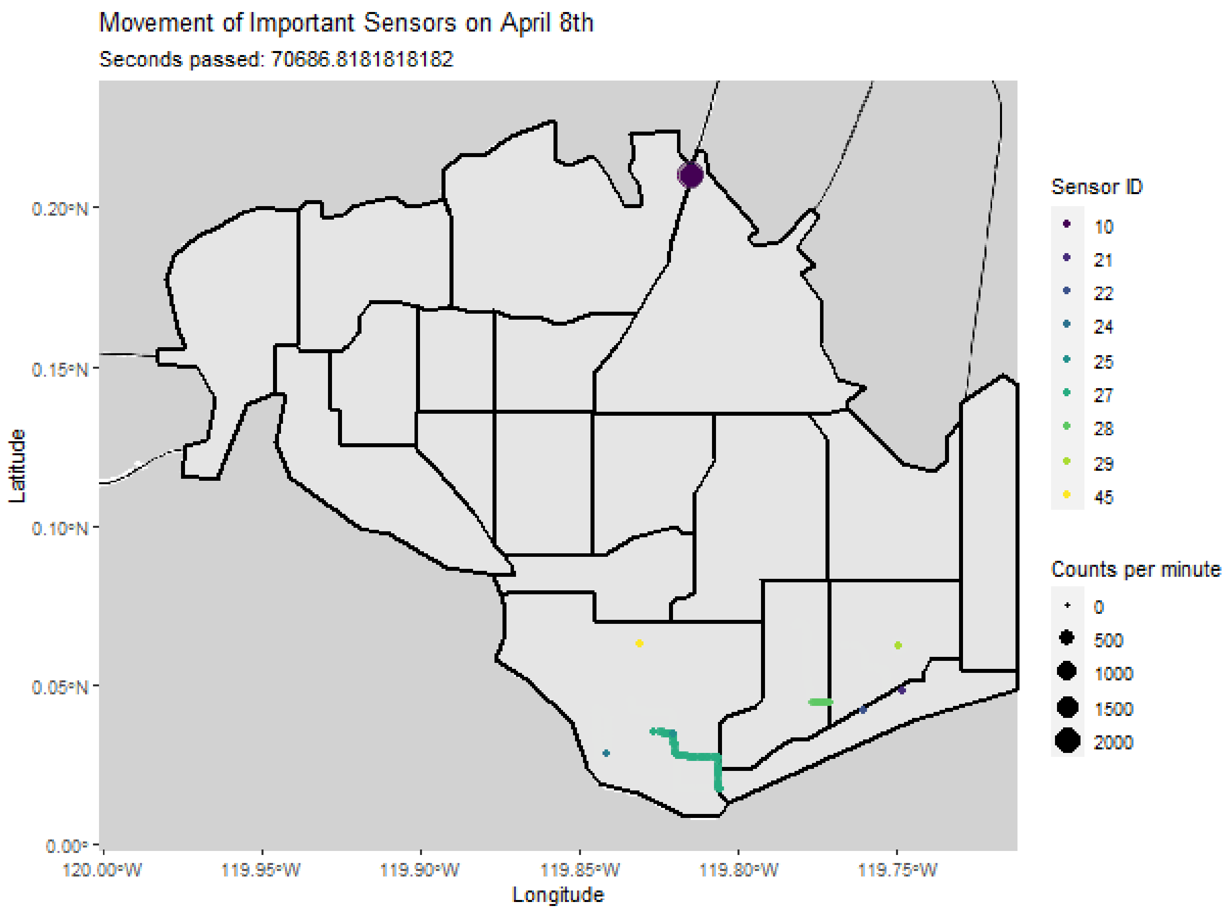

Figure 4. Snapshot of Animated Map of Important Mobile Sensor Movement on April 8th.
(see animatedMap_apr8ImpSensors.gif for movement data for the whole day)

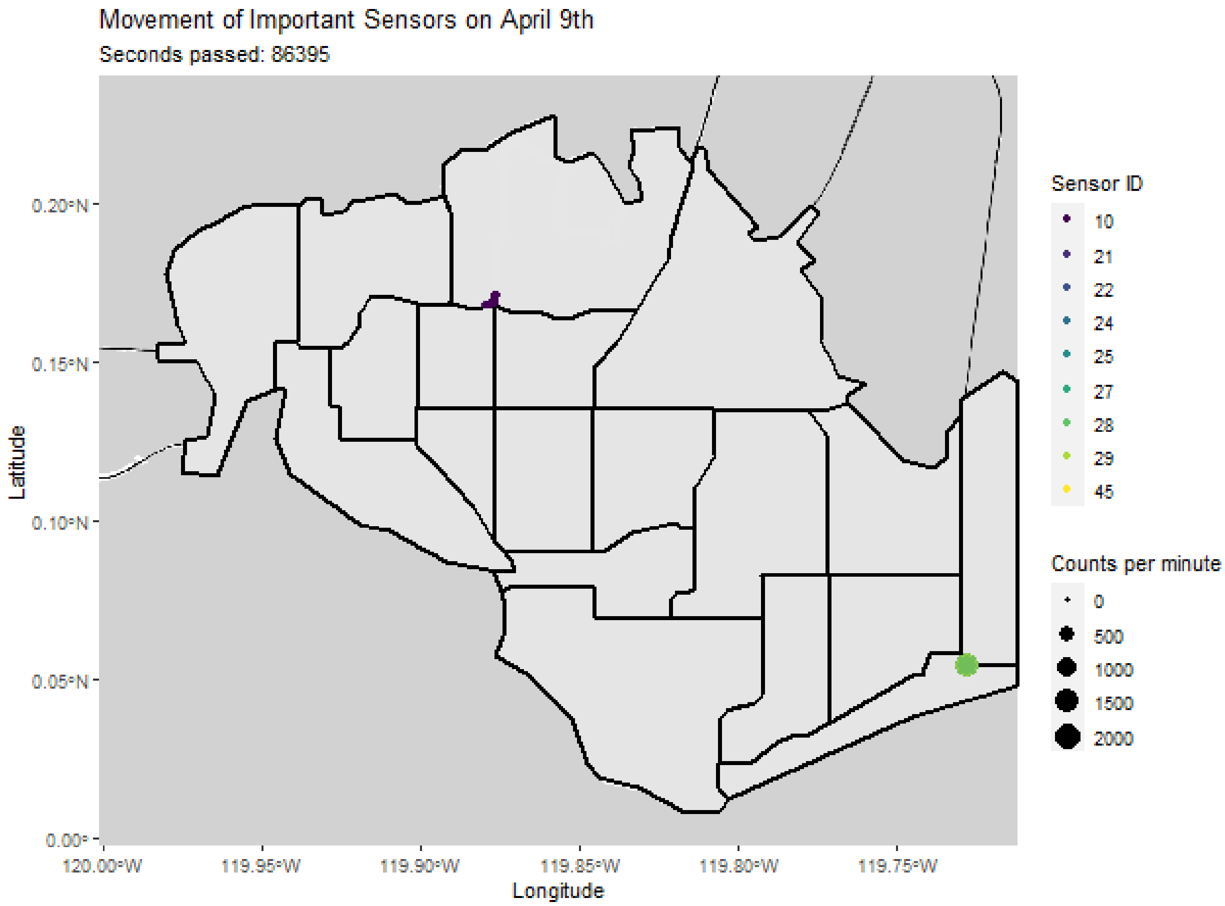

Figure 5. Snapshot of Animated Map of Important Mobile Sensor Movement on April 9th.
(see animatedMap_apr9ImpSensors.gif for movement data for the whole day)

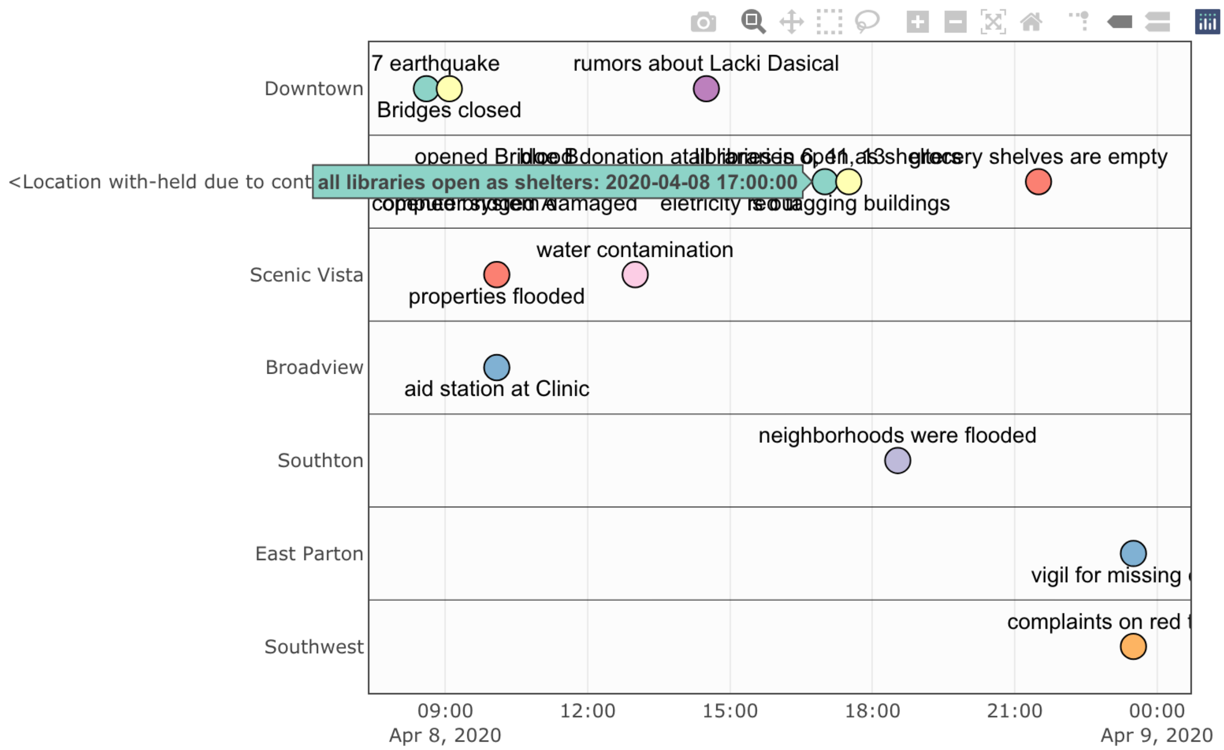

Figure 6. Timeline on April 8, 2020 based on YInt messages.
(see YInt_tl.html for interactive plot; the reliable messages we identified were more comprehensive than what we could plot)

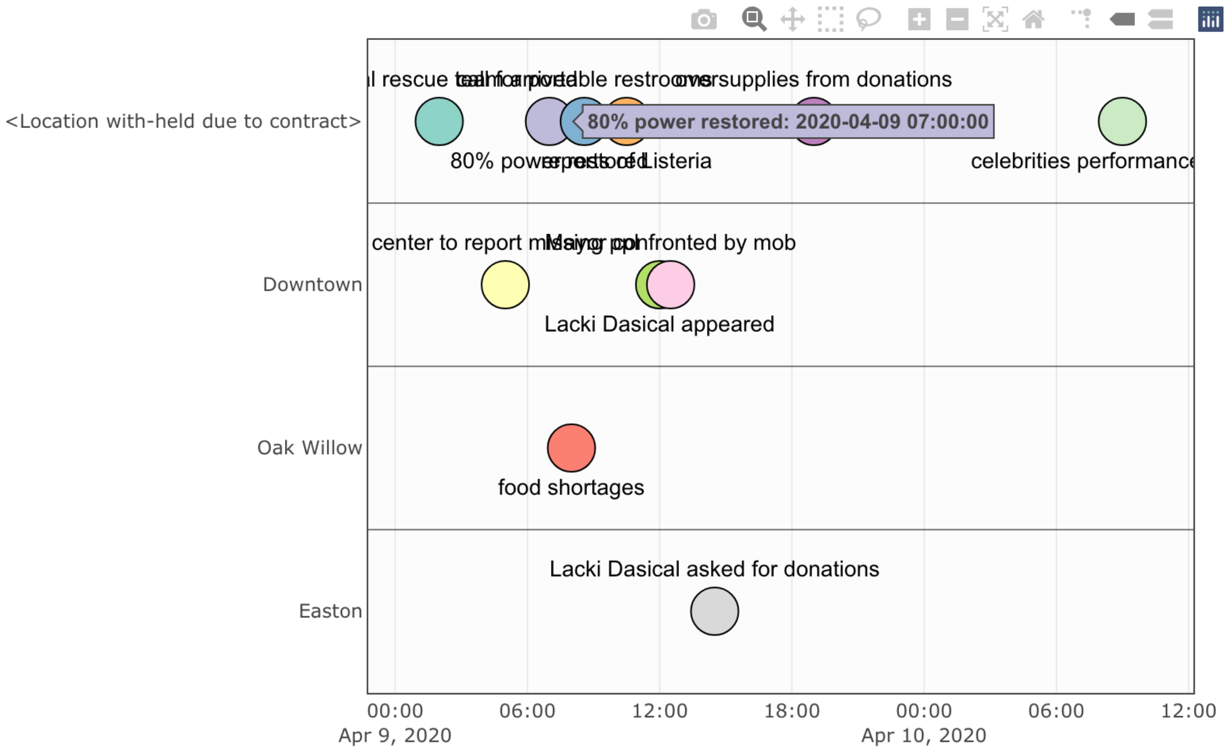

Figure 7. Timeline on April 9 & 10, 2020 based on YInt messages.
(see YInt_tl.html for interactive plot; the reliable messages we identified were more comprehensive than what we could plot)

Figure 8. Sentiment visualization of positive/negative emotions. (see Yint_tl.html for reference)

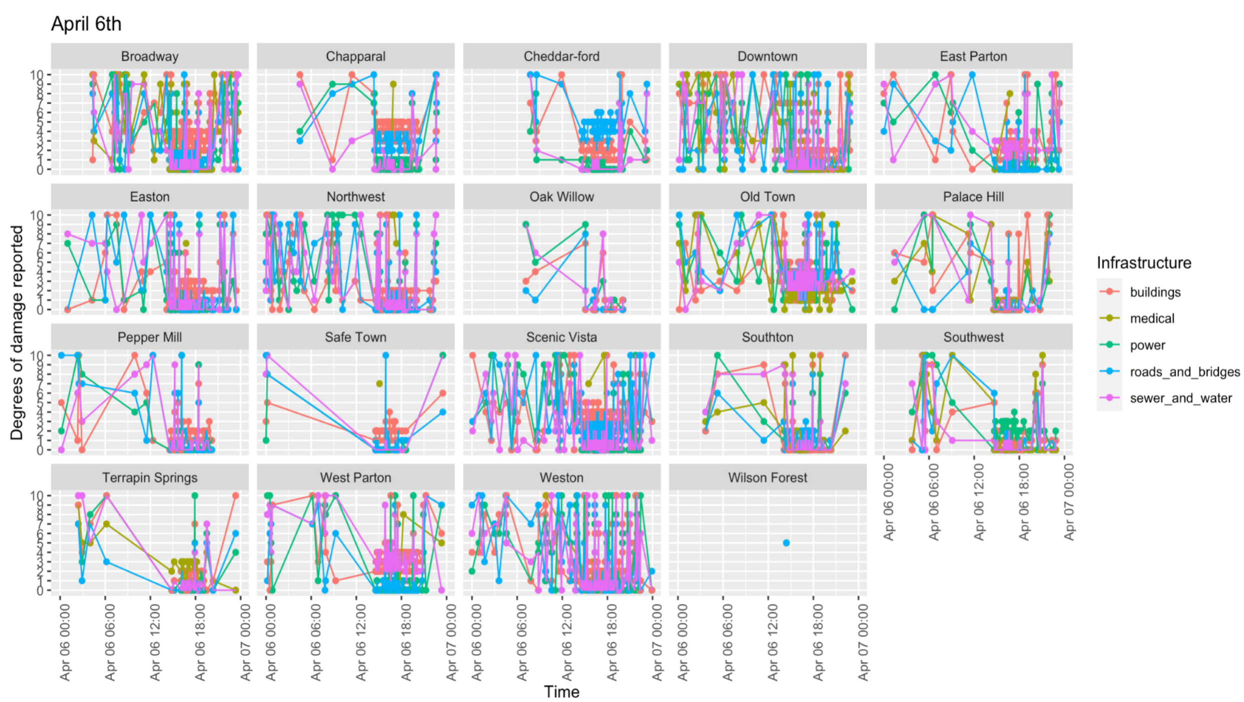

Figure 9. Damage report on April 6th.

Figure 10. Damage report on April 8th from 8:30 - 13:30.