




ARTICLE



<https://doi.org/10.1057/s41599-023-02312-7>

OPEN

# Lifestyle pattern analysis unveils recovery trajectories of communities impacted by disasters

Natalie Coleman <sup>1</sup>✉, Chenyue Liu<sup>1</sup>, Yiqing Zhao<sup>1</sup> & Ali Mostafavi<sup>1</sup>

Lifestyle recovery captures the collective effects of population activities as well as the restoration of infrastructure and business services. This study uses a novel approach to leverage privacy-enhanced location intelligence data, which is anonymized and aggregated, to characterize distinctive lifestyle patterns and to unveil recovery trajectories after 2017 Hurricane Harvey in Harris County, Texas (USA). The analysis integrates multiple data sources to record the number of visits from home census block groups (CBGs) to different points of interest (POIs) in the county during the baseline and disaster periods. For the methodology, the research utilizes unsupervised machine learning and ANOVA statistical testing to characterize the recovery of lifestyles using privacy-enhanced location intelligence data. First, primary clustering using k-means characterized four distinct essential and non-essential lifestyle patterns. For each primary lifestyle cluster, the secondary clustering characterized the impact of the hurricane into four possible recovery trajectories based on the severity of maximum disruption and duration of recovery. The findings further reveal multiple recovery trajectories and durations within each lifestyle cluster, which imply differential recovery rates among similar lifestyles and different demographic groups. The impact of flooding on lifestyle recovery extends beyond the flooded regions, as 59% of CBGs with extreme recovery durations did not have at least 1% of direct flooding impacts. The findings offer a twofold theoretical significance: (1) lifestyle recovery is a critical milestone that needs to be examined, quantified, and monitored in the aftermath of disasters; (2) spatial structures of cities formed by human mobility and distribution of facilities extend the spatial reach of flood impacts on population lifestyles. These provide novel data-driven insights for public officials and emergency managers to examine, measure, and monitor a critical milestone in community recovery trajectory based on the return of lifestyles to normalcy.

<sup>1</sup>Texas A&M University, College Station, USA. ✉email: [ncoleman@tamu.edu](mailto:ncoleman@tamu.edu)

## Introduction

Disasters not only cause physical damage to infrastructure systems and building facilities but also affect the societal dynamics of communities by disrupting the lifestyles of populations. Lifestyles can be characterized by essential and non-essential activities, such as visiting grocery stores to meet food needs or visiting local shopping centers to buy clothing (FEMA, 2020; Li et al., 2021; Podesta et al., 2021). The disruption of essential lifestyles negatively impacts populations, as they are unable to access critical resources to maintain their health and well-being. The restoration of non-essential services can also signal a return to normalcy, as people are more able and comfortable to resume their earlier activities (Horney et al., 2017; Links et al., 2018). During the rebuilding and recovery process following a disaster, people may incrementally resume their standard lifestyles. The return to pre-disruption lifestyle standards could represent an important facet of the societal impacts of disasters, and thus, has the potential to unveil a critical milestone in community recovery (Ma et al., 2022; Yuan et al., 2021).

Recovery is a vital component of the disaster management cycle. According to FEMA (2016), disaster recovery involves the return of all formal and informal systems to a normal state. Eid and El-adaway (2018) show that the disaster recovery process is multidimensional and nonlinear which involves a multitude of strategies and capacities to rebuild community livelihoods. Indeed, it is well-established that recovery is a complex process of multiple dimensions (Horney et al., 2017; Platt et al., 2020). Finucane et al. (2020) discuss how recovery planning should be adaptive to community needs and establish formal mechanisms for assessment. Recent literature has also stressed the importance of considering social connections (Delilah Roque et al., 2020; Masud-All-Kamal & Monirul Hassan, 2018; Sadri et al., 2018) and social inequities (Emrich et al., 2022; Peacock et al., 2014; Sovacool et al., 2018) in varying levels of recovery. According to Mayer (2019), the recovery process continues to be enhanced by interventions and capacity building in individuals and communities, which is being shown in programs across the globe.

Although there is no concrete consensus on recovery, several recovery frameworks have been developed. For example, Cutter et al. (2008) presents a system perspective on recovery using factors of ecological, social, economic, infrastructure, institutional, and community competence. Based on the ratings of a Delphi-panel to expert opinion, Jordan and Javernick-Will (2013) viewed recovery through infrastructure, social, economic, and environmental components. Horney et al. (2017) recognized the importance of developing potential indicators for recovery including but not limited to housing, infrastructure, and community planning and capacity building.

Recovery frameworks must also include elements of social recovery through institutions and services. These have singled out the significance of measuring people's life activities as part of the rebuilding phase after a disaster (Links et al., 2018; Tatsuki, 2007). The framework by Links (2018) focuses on the variety of human needs by presenting the concepts of community functioning, prevention and mitigation, and population. In addition, Tatsuki (2007) mentioned several factors of long-term recovery process such as community rebuilding and government relations. Other research studies have examined measured recovery through the amount of population returns (Finch et al., 2010), the functionality of infrastructure services (Chang et al., 2014), the extent of housing damage (Peacock et al., 2014), and the rates of neighborhood poverty (Lee, 2020); however, to the authors' knowledge, few studies have examined the lifestyle activities of populations at the granular scale of census block group (CBG) level. Therefore, the analysis of lifestyle patterns before and after disasters could bring unique

insights about a population's return to normalcy and supplement established recovery frameworks.

However, there is a gap of knowledge related to the specification, characterization, and quantification of recovery milestones (FEMA, 2016; Mayer, 2019; Tierney & Oliver-Smith, 2012). First, most recovery indicators are qualitative which can lead to a vague understanding of the disaster impact and recovery efforts. Quantitative measures of recovery are needed to aid in the allocation of resources and facilitate the restoration of critical systems. In fact, there is a need to produce quantifiable measures of recovery to not only help us better understand the recovery process but also evaluate the effectiveness of recovery efforts (Cutter et al., 2013). Second, recovery indicators are limited from a spatiotemporal perspective. They fail to capture the changing nature of individuals, populations, and communities in a post-disaster setting. Thus, the research introduces the concept of lifestyles to bring a much-needed quantitative and dynamic aspects of recovery. It also proposes the use of location intelligence data to quantify lifestyle patterns.

Location intelligence data can reveal important insights about complex disaster phenomena pertaining to population activities and interactions with critical facilities (Huang et al., 2022; Yuan, Esmalian, et al., 2022). For example, location intelligence data has examined disaster impacts (Rosenblum et al., 2021; Wang & Taylor, 2018; Yu et al., 2018), damage reporting (Yuan & Liu, 2020; Yuan, Yang, et al., 2022), evacuation patterns (Deng et al., 2021; Han et al., 2019; Hong et al., 2021), and rapid impact assessments (Kontokosta & Malik, 2018; Salley et al., 2019). Hong et al. (2021) categorized the evacuation patterns caused by Hurricane Harvey into shelter-in-place, stable, distressed, and abandoned households. Certain research studies found that human mobility has distinctive patterns during stable and disruptive periods (Wang & Taylor, 2014; Yabe et al., 2020). Podesta et al. (2021) measured periods of disruption to points of interest (POIs) in Harris County after the landfall of Hurricane Harvey; the study concluded that POI systems have differential recovery rates to their baselines. To expand on the understanding of accessibility to essential services, Esmalian et al. (2022), utilized location intelligence data to determine equitable access to critical facilities during disruptive events. Despite these recent advances, little of the existing work has leveraged location intelligence data for specifying and characterizing community recovery. Unlike traditional datasets (e.g., surveys) used for recovery, location intelligence data has greater spatial granularity, larger samples, and shorter lags in data collection. It could also relieve the burden of data collection from affected populations. Thus, examining lifestyle recovery as a critical recovery milestone can yield understanding about recovery trajectories at a granular level.

In this study, we leveraged location intelligence data, specifically POI data and human mobility data, to examine lifestyle activities at the CBG level. In doing so, this study presents a novel methodology to specify lifestyle patterns and understand lifestyle recovery as a critical milestone following a disaster. Due to their holistic and dynamic nature, lifestyles encompass routine changes across various spatial units after the disaster impact, and lifestyle recovery is a critical milestone based on capturing people's interactions with infrastructure, businesses, and social interactions in the community (Cutter et al., 2008; Podesta et al., 2021). We processed and analyzed datasets from Spectus and SafeGraph, two prominent commercial companies that collect and process location-based data at a granular scale. We combined that data with publicly available building footprint data to specify population visitations to POI. The study context was Hurricane Harvey in Harris County, Texas, from August 2017 through November 2017. The objectives of the research are to (1) assess

the extent of impact of disasters on lifestyle activities, (2) identify patterns of recovery trajectories, and (3) examine the variations of recovery given flooding and demographic attributes. The analysis focused on answering the following research questions: (1) What are distinctive patterns in typical lifestyles of populations during normal periods?; (2) To what extent, do disasters impact lifestyle activities and how quickly do lifestyle patterns recover?; and (3) To what extent, do the duration and trajectory of lifestyle recovery vary across different areas with different flooding impacts and demographic attributes? The aim of the research is to guide decision-makers such as emergency managers, public leaders, and service companies with more quantitative understanding of lifestyle recovery and bring attention to the potential disparities after a hazard event.

**Data and methods**

**Study context.** We analyzed and quantified the recovery of lifestyles using location intelligence data captured before, during, and after the 2017 Hurricane Harvey in Harris County, Texas. Hurricane Harvey was a Category 4 storm that made landfall in August 2017 and greatly impacted the Texas coastal region. Extensive flooding led to the inaccessibility of essential and non-essential facilities (CBS/AP. (2017)). The Texas Department of Transportation reported nearly 350 Houston road locations being affected (Dong et al., 2022; Zipline’s Logistics consultants, 2017), and the Electric Reliability Council of Texas (ERCOT, 2017) reported thousands of power disruptions. These disruptions had significant impacts on people’s lifestyle activities in the aftermath of Harvey. The impacts of Harvey and the metropolitan setting of Harris County provide a suitable context for examining lifestyle recovery patterns (Blake & Zelinsky, 2018).

**Data sources.** To specify daily lifestyle patterns, we processed multiple datasets to determine daily visitations from each home CBG to different POIs. Figure 1 depicts an overview of the methods for processing and analyzing the datasets. Human mobility data was primarily obtained from Spectus, a location intelligence data company that provided location data (e.g., Global Positioning System data) of high spatial accuracy and a very high frequency of observations (Spectus, n.d.). Every day, more than 100 data points on average are collected from each anonymous device. Currently, there are roughly 15 million daily active users in the United States. The Spectus panel consists of adults who provided informed consent to anonymized data collection for research purposes, through a General Data Protection Regulation (GDPR) and California Consumer Privacy Act

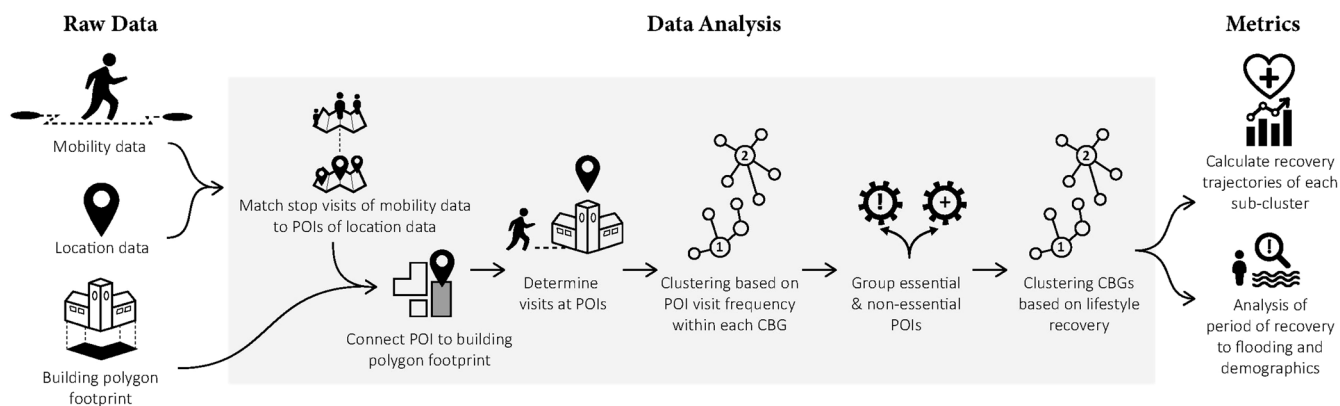
(CCPA) compliant framework. Compared to traditional Call Detail Record (CDR) data, which provides general information about telephone exchanges and interactions, location intelligence data with precise GPS information provides exact data about destinations, which is extremely useful for detecting human lifestyle patterns in detail. Spectus’s main database is built using third-party apps that capture opted-in users’ anonymized location points if they consent to share their location with these apps. Data is further anonymized and erased from system, then aggregated at the CBG level to ensure consumer privacy and confidentiality.

For 2017, Spectus data lacked North American Industry Classification System (United States Office of Management and Budget, 2022). NAICS codes to distinguish the type of each POI (stop points of unique device IDs). To address this, we utilized the location and type of different POIs consistent with NAICS codes from the SafeGraph dataset (SafeGraph). Safegraph’s core dataset contains spatial coordinates and addresses of POIs, as well as basic information about each POI, such as brand, operating time, latitude, and longitude, and NAICS code. SafeGraph gathers POIs by scraping open store locators on the web using publicly available APIs and scraping open web domains that provide updated locations. To fill in the gaps, SafeGraph independently performs processing and modeling as well as licenses third-party data to infer additional attributes. After collecting the POIs from Safegraph’s core dataset, we used Microsoft’s open-building footprint dataset to determine the polygon of each POI (Microsoft Open Source). The 1,542,887 building footprint polygon geometries in the Houston metropolitan area were created by Microsoft using their computer vision algorithms on satellite imagery.

To evaluate the spatial variation of lifestyle impacts and recovery based on flooding status, we used flooding data from the estimated flood depths on August 29, 2017 (FEMA, 2018). The data had a gridded horizontal resolution of three meters, which was processed appropriately for the CBG scale analysis. The percentage of CBG area flooded during Hurricane Harvey was an indicator of flooding extent. Demographic information, including total population, median income, percentage of the white population, and percentage of elderly residents above 65 years old, were collected at a CBG level from the US Census data (US Census Data (2017)).

**Data processing of human visitation to POIs**

*Home\_CBG to stop table.* Data processing focused on the period of August to November 2017. The cut-off date was the end of



**Fig. 1 Integrating multiple data sources to specify population visitation to POIs for characterizing lifestyle patterns.** Notes: Raw data includes multiple data sources. Data analysis includes the merging of data sources for visits at POIs and the overall methods for clustering the data. Metrics are the final outputs of the methods including calculating the recovery trajectories.

November to minimize the effect of the holiday season on population lifestyle patterns. To associate each anonymous device's home CBG in Harris County, we first retrieved data from the device matrix table in the Spectus core database. We recognized a CBG as a user's home\_CBG if the device consistently resided there for more than one day. The device matrix includes the user\_id and home\_CBG\_id. Second, we used the Spectus core dataset's stop table to capture anonymous human mobility. The stop points were captured in the dataset if individuals stopped at a location for a suitable time scale. User\_id, latitude and longitude, stop-by date, and stop-by time were all stored in the stop database. By merging the device matrix table and stop table through the user\_id, home\_CBG was acquired.

*POI polygon table.* To specify the polygon for each POI, the place\_id in Spectus's dataset was extracted using the POI table in Spectus's core datasets. This table has basic information about each POI, such as place id, place name, place address, and POI's latitude and longitude. We obtained each place id with its associated building polygon by merging the POI table with the building footprint dataset.

*Home\_CBG to POI table.* By merging the latitude and longitude of the POI table to the building footprint dataset, we matched each stop point of POI with a unique place\_id. Then we classified each POI using the NAICS code, which provides information on the associated sector, subsector, industry group, and industries. Since the 2017 Spectus data lacks the NAICS code, we can match the polygon information (latitude and longitude) to Safegraph's data, which does include the NAICS code. Informed by prior studies (Li et al., 2021; Podesta et al., 2021), we identified eleven POIs which contained five essential (gasoline stations, grocery and merchandise, health and personal care stores, medical facilities, and education) and six non-essential (banks, stores and dealers, restaurants, entertainment, recreation and gym centers and beauty care). The classification of essential was further influenced by FEMA definition of community lifelines, where FEMA discusses how community lifelines promote the food, water, health, and shelter needs of the community (FEMA, 2020). Table SI.1 in the Supplementary Information has the detailed POI-NAICS table with essential and non-essential lifestyles. By doing so, the daily visit data was generated from the home\_CBG to POI, and weekly data was generated by aggregating seven days of daily data. The rationale for weekly aggregation is that lifestyle patterns might vary between weekdays but are usually consistent on weekly basis (Sevtsuk & Ratti, 2010).

To analyze and quantify lifestyle recovery for different CBGs, we first established a baseline to compare against the lifestyle impacts. The baseline was calculated by averaging the number of visits during the first two weeks of August 2017, the period before Hurricane Harvey made landfall. Second, we grouped POIs into essential and non-essential POIs to establish the baseline number of visits to each group of POIs. Then, we compared the number of visits in the subsequent weeks against the baseline to measure lifestyle recovery.

We will briefly acknowledge the limitations of this data processing. Due to the data availability and scope, we were not able to collect data from before 2017 and did not feel it was appropriate to use data after 2018 to establish the baseline. Published work has used similar processing methods (Esmalian et al., 2022; Hong et al., 2021). As the field and analysis of mobility data grows, future studies could control for seasonal and temporal patterns by observing previous periods of annual activity. To ensure the validity of the data and results, we also need to ensure that the population activity and mobility datasets used are representative. Regarding the representativeness of Spectus data, researchers from the University of Washington

evaluated the utility of this data by mapping trips within an origin-destination transportation analysis (Wang et al., 2019). By analyzing the aggregated mobility patterns of more than 500,000 anonymous Spectus users (representing 12.5% of the population of the Puget Sound region under analysis), the researchers determined that Spectus data, as compared to the cellular network and in-vehicle GPS data, benefited from a superior combination of large scale, high accuracy, precision, and observational frequency. Beyond validating scale and accuracy, the research found that Spectus data is highly demographically representative (Wang et al., 2019). Additional existing studies on Spectus data have also demonstrated the representativeness of the data (Aleta et al., 2020; Vugrin et al., 2010). SafeGraph data have also been used to study consumer preferences (Athey et al., 2018).

A CBG is considered as recovered at the specific week if the number of visits to POIs is 90% of the baseline value. The concept of recovery of essential and non-essential lifestyles was modified from studies observing the resilience of critical systems (Nan & Sansavini, 2017; Vugrin et al., 2010). CBGs that did not return to the 90% threshold are labeled at 15 weeks of recovery which is the maximum amount of time for the data analysis. A sensitivity analysis of recovery to the baseline (thresholds of 90%, 80%, 70%) was performed and is found in Figure SI.4 of the Supplementary Information.

**Primary and secondary k-means clustering.** K-means clustering is an unsupervised learning method for creating clusters and cluster centers in a set of unlabeled data (Lenormand et al., 2015; Pappalardo et al., 2015; Thuillier et al., 2018; Toole et al., 2015). Instead of analyzing the lifestyle patterns of each CBG separately, we used k-means clustering to group CBGs with similar POI visitation patterns (i.e., the lifestyle patterns in the primary clusters) and recovery rates (i.e., differential recovery trajectories in the secondary clusters). In other words, k-means clusters similar CBGs based on the available data and does not depend on manual input to establish the boundaries of each dataset.

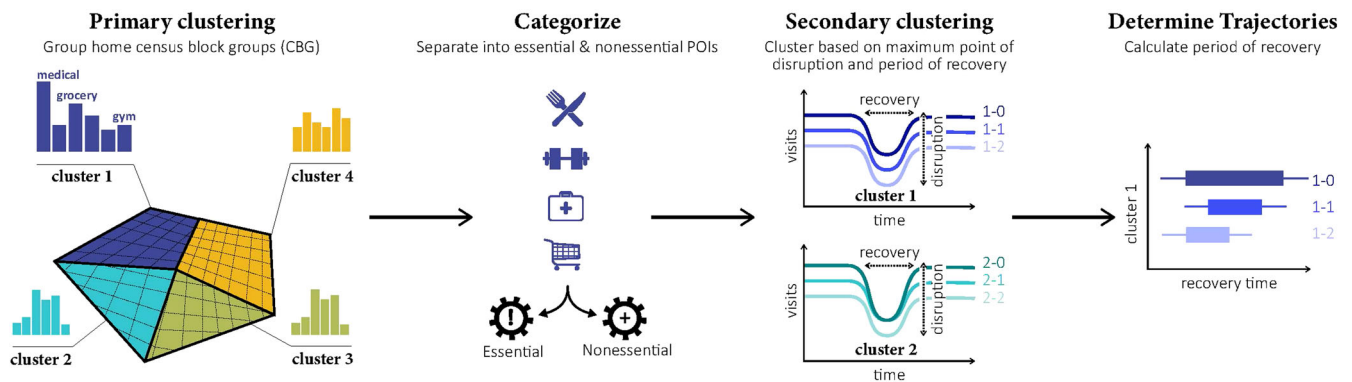
As shown in Fig. 2, k-means clustering is used in two steps. The first step created the clusters of CBGs based on their non-disrupted lifestyle patterns before the disruption of Hurricane Harvey; the second step clustered the CBGs within each primary cluster based on the lifestyle recovery after disruption. A minimum-maximum formula standardized the visitation values. The primary clustering was obtained from the POI visitations during the baseline period of the first two weeks of August 2017 (before the landfall of Hurricane Harvey). Vector data consisted of the aggregate frequency of visits to POIs over the total number of visits to all POIs. For example, the first input in the vector would consider the number of visits to grocery stores divided by the total number of visits to all POIs for a particular CBG within the same period. The secondary clustering of the lifestyle recovery is based on four features: the maximum point of disruption in essential POI visits, the maximum point of disruption in non-essential POI visits, the duration of recovery in essential POI visits, and the duration of recovery in non-essential POI visits. The maximum point of disruption is defined as the greatest percent change in visits to essential and non-essential POIs following the baseline period. Duration of recovery is the number of weeks for CBGs to recover to 90% of the baseline. In both clustering methods, the number of clusters is found through the elbow method as shown in Figure SI.1 in the Supplementary Information.

## Results

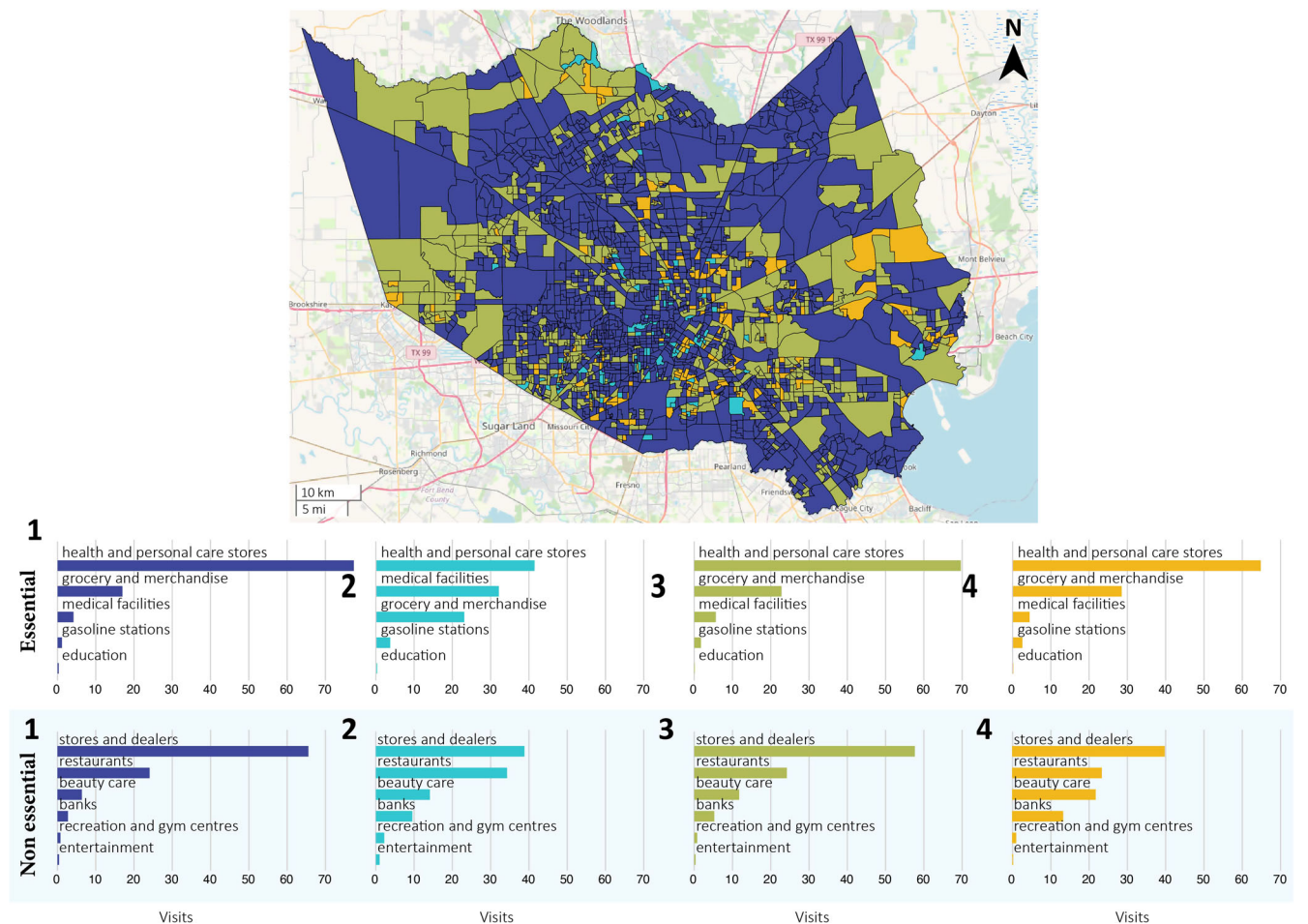
Primary clustering of k-means characterizes the lifestyles of CBGs for the baseline period before Hurricane Harvey. It shows the level of dependence that the cluster normally has on different

essential and non-essential lifestyles based on the frequency of visits. The elbow method (Figure SI.1 in Supplementary Information) found that four clusters were the optimal number of k-clusters. The median values and distribution of the demographic characteristics of the four clusters are found in Figure SI.2 and Table SI.2 in the Supplementary Information. In essence, 60.26% of CBGs were in Cluster 1, 3.68% in Cluster 2, 27.94% in

Cluster 3, and 8.12% in Cluster 4. According to the merged census data, Cluster 1 contained about 3.09 million people within an area of  $2.86 \times 10^9$  square miles; Cluster 2 contained  $1.42 \times 10^5$  people within an area of  $6.46 \times 10^7$  square miles; Cluster 3 contained about 1.48 million people within an area of 2.87 with  $1.48 \times 10^9$  square miles; and Cluster 4 contained about  $3.21 \times 10^5$  people within an area of  $2.05 \times 10^8$  square miles. Figure 3 shows



**Fig. 2 Primary and secondary clustering to determine the extent of lifestyle characterization and disruption.** Notes: Primary clustering groups home CBGs from prominent visits to POIs. POIs are then grouped into essential and non-essential lifestyles. Secondary clusters group within each of the primary clusters based on the maximum point of disruption and duration of recovery for essential and non-essential POIs. Finally, trajectories of recovery are examined based on mean and distribution.



**Fig. 3 Results show the frequency distribution of the essential and non-essential lifestyles across the four primary clusters.** Notes: POIs with greater visits means that populations are more dependent on those POIs. The bar plots present the prominent weekly POI visitations (essential and non-essential) within each lifestyle cluster. The x-axis shows the number of weekly visitations, and the y-axis shows the type of POI. In the map, Clusters 1 and 3 represent the majority of primary lifestyle clusters with 60.26% of CBGs and 27.94% of CBGs respectively.

**Table 1 ANOVA testing on the frequency of lifestyles for primary clusters.**

Essential POIs	Essential F-value	Non-essential POIs	Non-Essential F-value
Education	32.61*	Banks	1480.89*
Gasoline	89.18*	Entertainment	155.24*
Grocery and Merchandise	509.71*	Stores and Dealers	2335.61*
Health and Personal	674.54*	Beauty	733.07*
Medical Facilities	178.43*	Gyms	136.75*
		Restaurant	226.42*

\*Values are statistically significant at  $p < 0.05$

the spatial distribution and relative frequency of visits to the essential and non-essential lifestyles. The exact values of the visit frequency to the essential and non-essential POI categories are tabulated in the Supplementary Information for Tables SI.3a and SI.3b. Non-essential categories had consistent rankings across the four clusters with the top three rankings as stores and dealers, restaurants, and beauty care. Essential categories had consistent rankings for Cluster 1, 3, and 4 with the top three being health and personal care stores, grocery and merchandise, and medical facilities. However, Cluster 2 had a slight adjustment with the top three ordered rankings as health and personal care stores, medical facilities, and grocery and merchandise.

Table 1 displays how the relative frequency of each lifestyle varied across the clusters, which was confirmed by analysis of variance (ANOVA) testing with a statistical  $p$ -value of  $< 0.05$ . This indicates that each primary cluster has distinct dependence on its lifestyles when compared to the other clusters. It is important to recognize the subtle differences which make up the essential and non-essential lifestyles of the primary clusters. Cluster 1 had the highest relative frequency for essential lifestyles with health and personal care stores (77.40%) and education (0.32%) as well as for non-essential lifestyles with stores and dealers (65.57%). When compared to other clusters, Cluster 2 had the highest relative frequency for essential lifestyles with medical facilities (31.97%) and gasoline stations (3.60%) as well as for non-essential lifestyles with restaurants (34.27%), recreation and gyms (2.21%), and entertainment (0.99%). Cluster 3 had no distinct high frequency to the essential and non-essential lifestyles. Lastly, Cluster 4 had the highest relative frequency for essential lifestyles with grocery and merchandise (28.40%) and non-essential lifestyles of beauty care (21.86%) and banks (13.37%).

### Secondary clustering of essential and non-essential lifestyles.

Secondary clustering of  $k$ -means within each primary cluster will determine the state of lifestyles during disrupted periods (e.g., Hurricane Harvey) and detect disproportionate rates of recovery within the same lifestyle. Using the elbow method, the analysis found three sub-clusters for each primary cluster. Among the total twelve sub-clusters, we identified four patterns of recovery trajectories: immediate recovery duration (IRD) which had 0 weeks of lifestyle disruption, short recovery duration (SRD) which had 1–6 weeks of disruption, moderate recovery duration (MRD) which had 7–14 weeks of disruption, and extreme recovery duration (ERD) which had 15 weeks or more of disruption. Figure 4 shows the spatial distribution of these secondary clusters. Figure 5 shows the data distribution and median values of the recovery trajectories of essential and non-essential lifestyles for each of the sub-clusters, where each CBG in the sub-cluster is

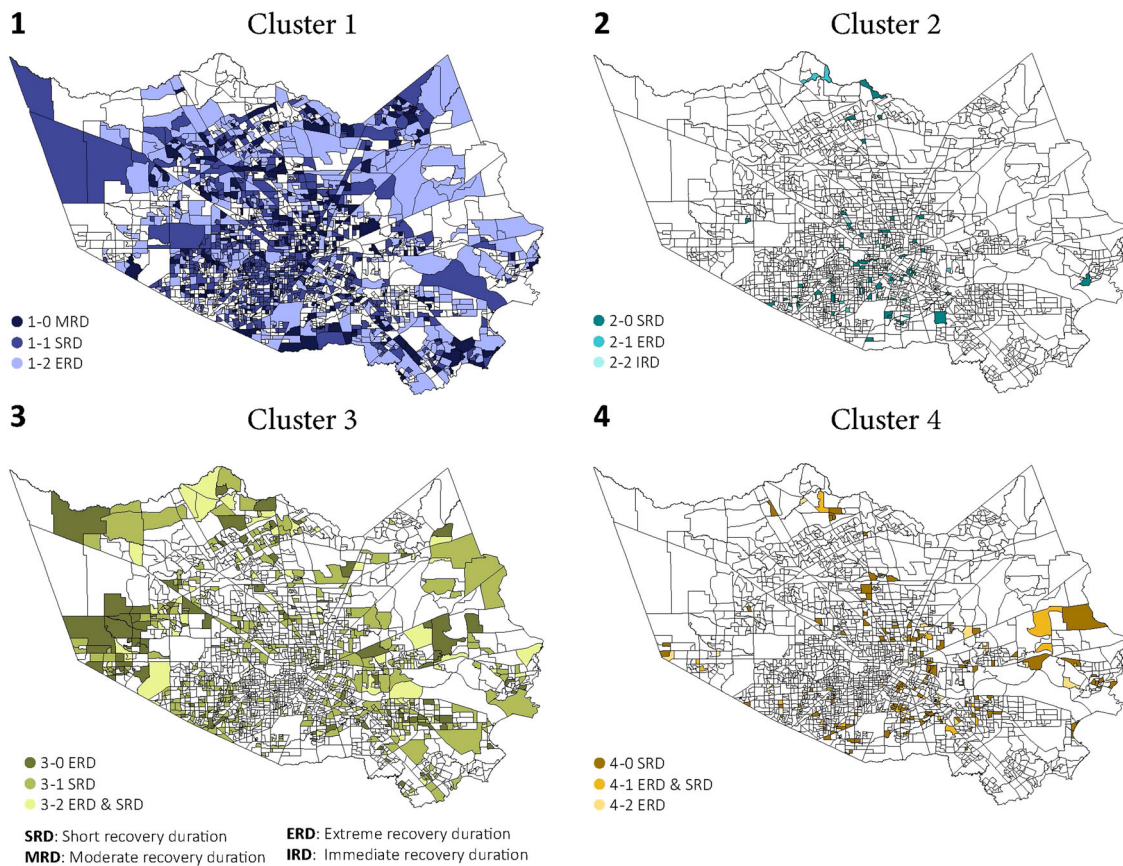
represented by a point. Examples of the individual recovery rates of CBG is on Figure SI.5 in the Supplementary Information.

ANOVA testing found that the differences between the recovery trajectories of the secondary clusters were statistically significant at  $p < 0.05$  (Table 2). This result shows that within the same primary cluster, or the same lifestyle, there could be disproportionate rates of recovery. For example, Cluster 2 has a higher frequency of visits to medical facilities when compared to other clusters which suggest that medical facilities could be a distinguishing feature of the lifestyle of Cluster 2. However, Cluster 2 had three rates of recovery meaning that certain CBGs had an immediate recovery duration, a short recovery duration, or an extreme recovery duration. For certain clusters, essential lifestyle recovery occurred before non-essential lifestyles and vice versa, which shows that lifestyle recovery may not be sequential.

### Exploration of demographics and flooding to lifestyle recovery.

The research also examined lifestyle recovery patterns in connection to demographics and flooding. For the secondary clusters, Table 3 summarizes the count of CBGs, the median values and percentages of demographics across the CBGs, and the percentage of CBGs that have a 1% flooding threshold. The 1% flooding threshold means that at least 1% of the CBG area experienced flooding impacts. To clarify, the research examined the disruption of lifestyles from people's home CBG to their ability to accommodate their essential and non-essential lifestyles. We consider the impact of flooding at the home CBG level, and we do not consider the flooding disruptions in accessibility to essential and non-essential lifestyles (e.g., flooding at critical roads to reach lifestyles).

The results showed some variation between the demographic attributes of populations and their lifestyle recovery; however, the results did not have a general pattern of lifestyle recovery for different race and income groups to yield statistically conclusive findings. ANOVA testing found no statistical differences between the mean or median values of the demographic characteristics for the three lifestyle recovery trajectories within each lifestyle cluster. It is worth noting the demographic differences from an exploratory perspective, which can be further investigated in future research. For instance, Cluster 2-1 experienced extreme recovery duration (15 or more weeks) along with had a \$32,674 median income and a 60% minority population. Meanwhile, Cluster 2-2 had an immediate recovery duration (0 weeks) but a median income of \$67,857 and a 19% minority population. Thus, Cluster 2-1 had an extreme recovery duration coupled with lower income and higher percentage of minority population while Cluster 2-2 had an immediate recovery duration with higher income and lower percentage of minority population. This demonstrates the possible social disparities of recovery trajectories associated to vulnerable groups within similar lifestyles. However, the opposite trend is observed for Cluster 3. Cluster 3-0 had an extreme recovery duration (15 or more weeks) with a median income of \$61,319 while Cluster 3-1 had a short recovery duration (2 weeks) with a median income of \$47,721. Furthermore, the results indicate that the flooding of the home CBGs are not always connected with impacts of lifestyle recovery. Approximately 89% of non-flooded home CBGs still experienced at least 1 week of disruption in lifestyles, which emphasizes that the impacts on lifestyle can extend beyond direct flooding of the household. Particularly, Cluster 2-2 had 33% of home CBGs with at least 1% flooding and had immediate recovery duration (0 weeks) while Cluster 4-2 which similarly experienced 33% of home CBGs with at least 1% flooding had extreme recovery duration (15 or more weeks). However, Cluster 2-2 did hold a higher median income of \$67,857 while Cluster 4-2 held a lower



**Fig. 4 Spatial distribution of the twelve secondary clusters and the four types of recovery trajectories.** Notes: The maps show the spatial distribution of the secondary clusters including the immediate recovery duration (IRD) which had 0 weeks of lifestyle disruption, short recovery duration (SRD) which had 1–6 weeks of disruption, moderate recovery duration (MRD) which had 7–14 weeks of disruption, and extreme recovery duration (ERD) which had 15 weeks or more of disruption.

median income of \$37,283. In this particular instance, lower income and not flooding percentage could be associated with the longer recovery duration. Though the research study did not find statistically significant relationships, these exploratory connections bring attention to the role of demographics and flooding in the household. Thus, future research should investigate the interrelationships of demographics and flooding for the lifestyle analysis of other disasters.

**Closer examination of extreme recovery duration of lifestyles.**

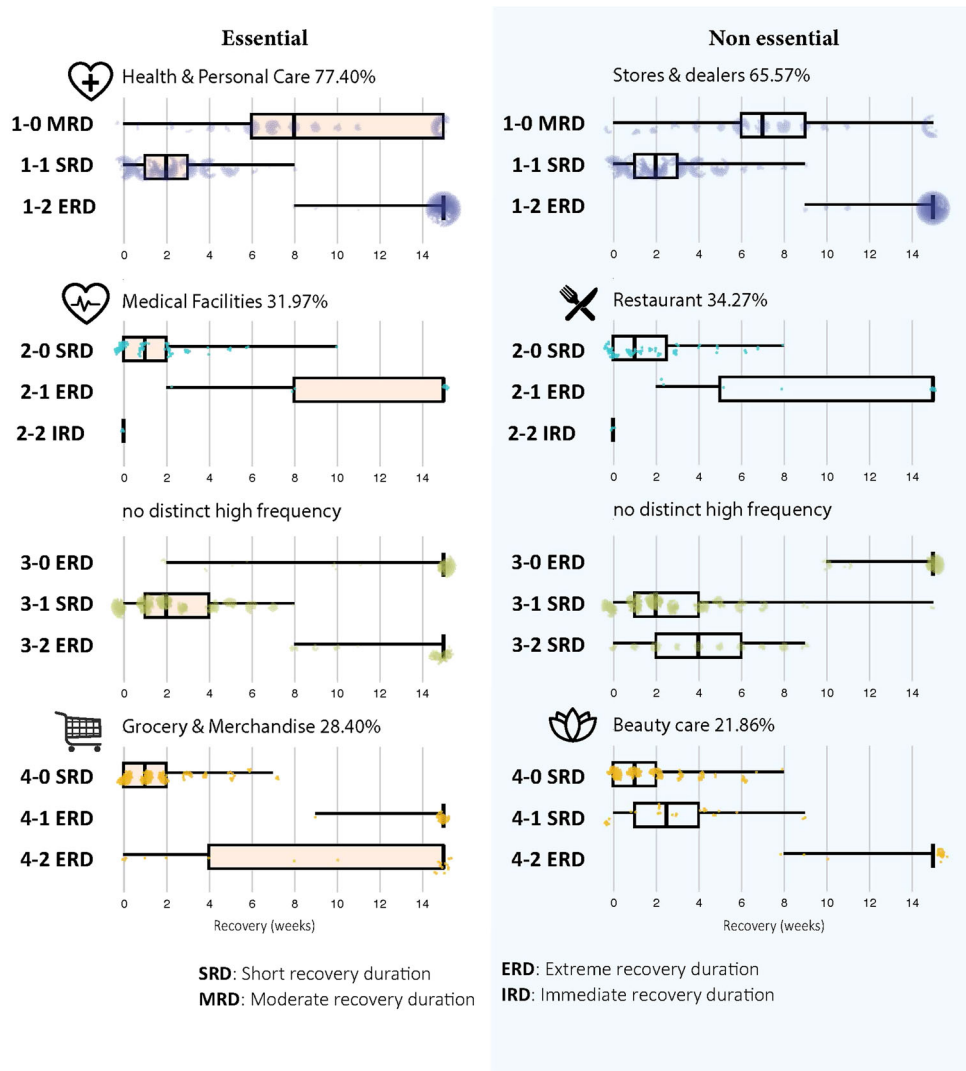
We further examined disparities in lifestyle recovery by focusing on the CBGs in secondary clusters which had extreme recovery duration (15 or more weeks). Approximately 500 CBGs from Cluster 1-1, 4 CBGs from Cluster 2-1, 92 CBGs from Cluster 3-0, and 12 CBGs from Cluster 4-0 did not recover to 90% of their essential and non-essential lifestyles after 15 weeks. Table 4 shows the median values of the demographics from these CBGs as well as the count of CBGs, the percentage from the secondary and primary clusters, and the percentage of CBGs with 1% flooding.

Figure 6 displays the spatial distribution of the CBGs with extreme recovery duration (15 or more weeks) including their associated secondary clustering. For example, Cluster 2-1 and Cluster 4-2 had extreme recovery durations (15 or more weeks) but these clusters had median incomes below the median incomes of all CBGs. However, Cluster 1-2 and Cluster 3-0 also had extreme recovery durations (15 or more weeks), and these clusters had median incomes above the median income of all CBGs. Though these secondary clusters had different ranges of income, they still exhibited the same lifestyle recovery. This result

suggests that people with different demographic attributes can have similar lifestyle patterns as well as similar lifestyle recovery. Furthermore, home CBGs with an extreme recovery duration may not always be subjected to any direct flooding. In total, 297 CBGs in Cluster 1-2, 2 CBGs in Cluster 2-1, 50 CBGs in Cluster 3-0, and 8 CBGs in Cluster 4-2, which equates to almost 59% of the CBGs with extreme recovery duration, did not experience at least 1% flooding of the area. The findings further indicate that the mechanisms driving the recovery of lifestyles may be beyond the extent of flooding in the household even in the most extreme circumstances. It demonstrates that the spatial structure of communities, which is primarily formed by human mobility and distribution of POI facilities, extends the spatial reach of flood impacts on people’s lifestyles regardless of the flood status of their home CBG. Due to the scope of the dataset, we do acknowledge that the research did not investigate flooding in relation to people’s ability to access essential and non-essential lifestyles. As location-intelligence data grows for the disaster research field, future studies can examine how to trip distance to lifestyles from home CBGs can play a role in the recovery process.

**Discussion**

Current strategies for measuring community recovery rely heavily on qualitative data which can be prone to subjectivity and biases, put the burden of data collection on impacted people, and provide lagging indicators of recovery (Kontokosta & Malik, 2018; Rosenblum et al., 2021). Recognizing these knowledge gaps, the research presents a transformative way to efficiently quantify and proactively monitor community recovery through lifestyle



**Fig. 5** Box-whisker plots of the duration of recovery for secondary clusters which show prominent essential and non-essential POIs for the primary clusters. Notes: The box-whisker plots display the distribution of the duration of recovery for essential and non-essential lifestyles of the CBGs in the secondary clusters. The icons and percentages show the relative highest frequency of visits to the essential and non-essential POIs for the primary clusters. Each CBG is represented by a point on the plot.

**Table 2** ANOVA testing on essential and non-essential lifestyle recovery for secondary clusters within primary clusters.

Clusters	Essential F-value	Non-essential F-value
Cluster 1	4129.18*	4474.89*
Cluster 2	80.10*	41.59*
Cluster 3	1740.17*	1350.98*
Cluster 4	313.39*	365.59*

\*Values are statistically significant at  $p \leq 0.05$

analysis. It argues that lifestyles could be a critical milestone of the overall recovery process. The recovery of lifestyles captures the combined effects of people’s activities, restoration of infrastructure, and access to businesses and services in the community. In the research methodology, we leveraged location intelligence data at the CBG level to specify distinct patterns of lifestyles and to quantify the extent of disaster impacts and speed of recovery. While location intelligence data has been utilized for examining other disaster-related phenomena such as evacuation (Han et al.,

2019; Wang & Taylor, 2014, 2016), flooding impacts (Farahmand et al., 2022; Farahmand et al., 2021; Fotovatikhah et al., 2018), community resilience (Yabe et al., 2020; Yuan et al., 2021), and access to facilities (Lenormand et al., 2015), Fan et. al (2022), no prior work had attempted to leverage location intelligence data in the context of recovery of lifestyles. In particular, the integration of population lifestyle patterns with other relevant datasets is novel. It overlays human mobility, POI facility location, building footprint, sociodemographic, and flooding datasets to uncover lifestyle patterns over space and time. The following will discuss the implications of the significant findings of the research as well as connections to how decision-makers can interpret and adjust the methodology.

**Characterizing communities based on dependence on essential and non-essential lifestyles.** Restoration of these essential lifestyles, such as those which provide food, water, shelter, and health services, can promote overall community recovery. However, individuals may have distinct relationships with these essential and even non-essential lifestyles. In the case of Harris County, the clustering analysis identified four distinctive lifestyles with subtle



**Table 3 Median values of the demographics for the secondary clusters.**

	Count of CBGs	Total Pop	Elderly (%)	White (%)	Median Income (\$)	Flood >=1% (% of CBGs)	Essential: Duration of Recovery	Essential: Maximum Point of Disruption	Non-essential: Duration of Recovery	Non-essential: Maximum Point of Disruption
All*	2144	1,843	10.01	68.21	54,539	38.66	5	71.69	4	85.78
1-0	252	1,723	10.47	69.00	56,295	44.05	8	70.97	7	89.08
1-1	521	1,804	10.77	70.29	60,444	35.12	2	65.82	2	82.41
1-2	519	2,040	10.03	68.26	61,655	40.85	15	75.12	15	89.15
2-0	67	1,126	10.80	57.35	43,000	38.80	1	45.76	1	51.17
2-1	9	1,097	10.90	40.07	32,674	55.56	15	68.31	15	68.62
2-2	3	1,453	9.93	81.76	67,857	33.33	0	-1.18	0	-4.43
3-0	115	2,233	9.20	68.96	61,319	42.61	15	65.08	15	82.84
3-1	389	1,853	9.55	65.81	47,721	40.36	2	54.05	2	77.62
3-2	95	2,093	8.77	70.55	49,821	34.74	15	62.42	4	82.25
4-0	127	1,441	9.00	62.23	38,005	29.13	1	55.08	1	53.13
4-1	26	1,793	8.94	56.51	35,256	30.77	15	63.05	2.5	70.64
4-2	21	1,349	11.76	61.71	37,283	33.33	15	67.03	15	72.43

- The maximum point of disruption is defined as the greatest percentage change in visits to essential and non-essential POIs, while the duration of recovery is the number of weeks for a CBG to recover to 90% of its baseline value.
- Yellow shading indicates that the secondary cluster had extreme recovery duration (15 or more weeks) in only their essential lifestyles to 90% of the baseline
- \*All referring to all CBGs in Harris County

• The maximum point of disruption is defined as the greatest percentage change in visits to essential and non-essential POIs, while the duration of recovery is the number of weeks for a CBG to recover to 90% of its baseline value.  
 • Yellow shading indicates that the secondary cluster had extreme recovery duration (15 or more weeks) in only their essential lifestyles to 90% of the baseline  
 • \*All referring to all CBGs in Harris County

**Table 4 Median values of the demographics for CBGs with extreme recovery duration (15 or more weeks for essential and non-essential lifestyles).**

Count of CBGs	CBGs from the Secondary Cluster (%)	CBGs from Primary Cluster (%)	Total Pop	Elderly (%)	White (%)	Median Income (\$)	Flood >=1% (% of CBGs)	
All*	2144	—	—	1843	10.01	68.21	54,539	38.66
1-2	500	96.34	38.70	2054	9.97	68.08	61,674	40.60
2-1	4	44.44	5.06	901	11.80	31.90	29,854	50.00
3-0	92	80.00	15.35	2272	9.00	69.57	61,637	45.65
4-2	12	57.14	6.90	1435	11.71	69.03	45,774	33.33

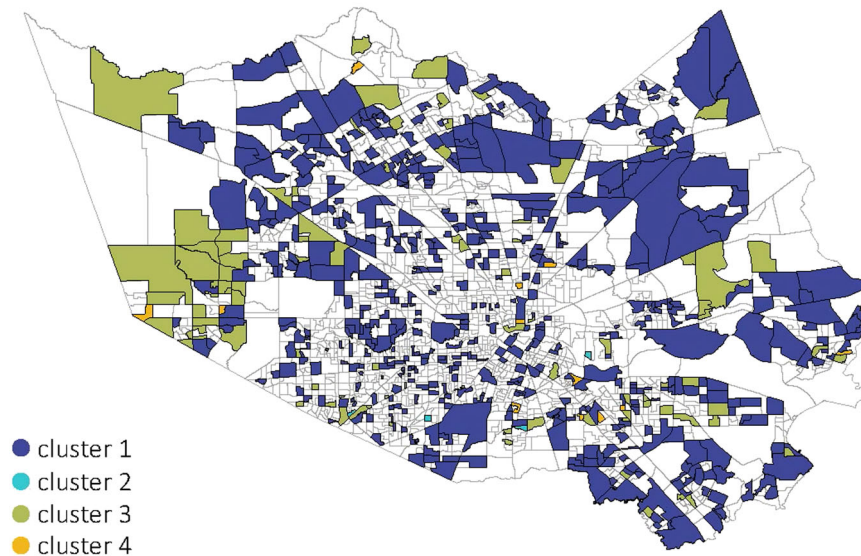
\*All referring to all CBGs in Harris County

differences in the dependencies to POIs. When compared to other clusters, Cluster 2 is more dependent on medical facilities (32%) which may also suggest that the recovery of Cluster 2 is also more linked to the recovery of medical facilities. In addition, Cluster 1 has the highest relative frequency of visits to health and personal stores (77%) and stores and dealers (66%) while Cluster 4 has the highest relative frequency of visits to grocery stores (28%) and beauty care (22%). These findings can help emergency managers understand which lifestyles are critical components of the standard lifestyles within a community. Such information can better prioritize the restoration of essential and non-essential lifestyles to efficiently maintain the well-being of the disrupted areas. For example, resources could be directed to restoring medical facilities for Cluster 2 and grocery stores for Cluster 1. Though the classification of essential and non-essential lifestyles is supported

by previous studies (FEMA, 2020; Podesta et al., 2021), these could be adjusted by decision-makers based on the community's needs. For example, banks could be considered critical to the community's economic activity based on the decision-makers.

**Detecting disproportionate rates of recovery trajectories.** Furthermore, the results advance our understanding of differential rates of recovery among locations and populations by analyzing the recovery of lifestyles at the CBG level. In examining Harris County, there were four diverse recovery trajectories which included immediate recovery duration, short recovery duration, moderate recovery duration, and extreme recovery duration. These recovery trajectories also indicate that the recovery of essential and non-essential lifestyles does not have a fixed

## Extreme Clustering



**Fig. 6 Spatial distribution of CBGs with extreme recovery duration (15 or more weeks of essential and non-essential) at the secondary cluster level.**

Notes: The spatial distribution of the CBGs with extreme recovery duration (15 or more weeks) labeled with their associated primary cluster.

Approximately 500 CBGs from Cluster 1-1, 4 CBGs from Cluster 2-1, 92 CBGs from Cluster 3-0, and 12 CBGs from Cluster 4-0 did not recover to 90% of their essential and non-essential lifestyles.

sequence meaning that essential lifestyle recovery can occur before or after non-essential lifestyle recovery. Within the same essential and non-essential lifestyle clusters, there are disproportionate rates of recovery. Thus, decision-makers must not only consider the unique lifestyles in a community but also the different abilities and capacities to recover from these lifestyles. For example, medical facilities may be prioritized for Cluster 2; however, further prioritization may be directed to areas that have extreme recovery duration. The current approach to recovery is rather reactive; however, by building a foundational knowledge of lifestyle, decision-makers could be more proactive in detecting potentially disproportionate rates of recovery and efficiently allocating resources.

**Considering the impacts on vulnerable populations.** The distinct demographics of these home CBGs could inform how emergency managers and public officials prioritize and manage the recovery of these areas as they can thoroughly understand which populations are being impacted. Disaster research argues that socially vulnerable populations, including those of low income and a high percentage of minority residents, have unique challenges to recovering from disasters (Mitsova et al., 2019; Mitsova et al., 2018). Literature has shown that such populations experience greater hardships and have lower capacities to withstand disrupted services (Coleman et al., 2020a, 2020b). Indeed, demographic characteristics are rooted in equitable recovery literature and lifestyles signal the direct needs and dependencies of populations. Though the data did not show an overall statistically significant pattern, it did uncover instances of disparity such as the extreme duration recovery and low median incomes in Clusters 2 and 4. Decision-makers should be on alert that delays in lifestyle recovery for vulnerable populations could lead to significant well-being impacts. Combining location-intelligence data with publicly available data on demographic information could assist in detecting communities that may face exacerbated impacts from disrupted lifestyles.

**Understanding the extent of flooding on lifestyle patterns.** The findings imply that the flooding impact on lifestyles extends beyond the direct flood exposure to home CBGs. The spatial structures of communities may be more formed by human mobility and facility distribution. Hence, populations living in non-flooded home CBGs could still experience significant disruption to their essential and non-essential lifestyles. This is demonstrated as approximately 89% of non-flooded home CBGs still experienced 1 week of disruption and 58% of home CBGs with extreme recovery duration were also non-flooded home CBGs. The finding reveals that greater investment should be made to study the network dynamics and spatial spillover effects of flooding and lifestyle populations. Decision-makers should take note of the systemic effects of flooding on the broader community. This means that even home CBGs that do not experience direct flooding can be negatively impacted and must be considered in the recovery process. Flooding could instead impact the ability of residents to meet their lifestyle needs by disrupting the normal route of travel. The connection between trip distance and the inability to access essential and non-essential lifestyles is outside the scope of this research; however, future studies can examine this as a potential avenue of the impact on lifestyles.

**Steps for future research.** The research findings will advance our understanding of lifestyle recovery as a critical milestone of overall community recovery. Future research could evaluate the underlying mechanisms of lifestyle recovery and other influencing factors such as demographics and flooding in several methods. First, future research could investigate how the trip distance and travel time to essential and non-essential facilities could play a role in lifestyle recovery. Second, supervised learning methods (such as regression, decision trees, and neural networks) could predict how sociodemographic features of the population and the attributes of a hazard event influence the recovery of lifestyles. Third, the analysis of lifestyle patterns will increase in reliability as more data is collected for population activity and human mobility on a broader temporal and finer spatial

granularity. With a combined effort, future research will further the understanding of the social, physical, and environmental disparities associated with lifestyle recovery.

### Concluding recommendations

This study provides a novel approach to quantifying and examining community recovery based on the fluctuations of population lifestyles. The characterization of community recovery would enable proactive monitoring of when communities bounce back at granular spatial scales. The results also could inform decision-makers and emergency managers in the following aspects:

- First, the findings show that decision-makers should understand the relationship of communities to essential and non-essential services during normal periods. Such knowledge can set an initial state of the community to compare future disruptive events. By characterizing lifestyles early-on, decision-makers can prioritize the restoration of important services based on how the population depends on certain essential and non-essential services in their lifestyle.
- Second, even within similar lifestyles, there are differential rates of recovery in the aftermath of disruptive events. This insight can bring attention to vulnerable populations and areas in a community that may not have the capabilities to recover from lifestyle disruptions.
- Third, the research shows that demographics and flooding extent could partially explain the differential rates of lifestyle recovery. Using publicly available datasets, decision-makers could preemptively know which areas are prone to extreme recovery durations based on vulnerable demographic characteristics and flooding extent.

### Data availability

All data sources were collected through a CCPA- and GDPR-compliant framework and used for research purposes. The data that support the findings of this study are available from Spectus Inc., but restrictions apply to the availability of these data, which were used under license for the current study. The data can be accessed upon request submitted on Spectus.ai. Other data (shapefiles, demographics, flooding) used in this study are all publicly available and are cited accordingly.

Received: 14 July 2022; Accepted: 24 October 2023;

Published online: 10 November 2023

### References

- Aleta A, Martín-Corral D, Pastore y Piontti A, Ajelli M, Litvinova M, Chinazzi M, Dean NE, Halloran ME, Longini IM, Merler S, Pentland A, Vespignani A, Moro E, Moreno Y (2020) Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* 4(9):964–971. <https://doi.org/10.1038/s41562-020-0931-9>
- Athey S, Blei D, Donnelly R, Ruiz F, Schmidt T (2018) Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. *AEA Papers and Proceedings* 108:64–67. <https://doi.org/10.1257/pandp.20181031>
- Blake ES, Zelinsky DA (2018) *National Hurricane Center Tropical Cyclone Report: Hurricane Harvey*. U. S. N. O. a. Administration & U. S. N. W. Service. <https://www.hssl.org/?abstract&did=807581>
- CBS/AP (2017) *Hurricane Harvey: Texas power outages affect more than quarter-million*. <https://www.cbsnews.com/news/hurricane-harvey-texas-power-outages-affect-more-than-255000/>
- Chang SE, McDaniels T, Fox J, Dhariwal R, Longstaff H (2014) Toward disaster-resilient cities: Characterizing resilience of infrastructure systems with expert judgments. *Risk Analysis* 34(3):416–434. <https://doi.org/10.1111/risa.12133>
- Coleman N, Esmalian A, Mostafavi A (2020a) Anatomy of susceptibility for shelter-in-place households facing infrastructure service disruptions caused by natural hazards. *International Journal of Disaster Risk Reduction* 50:101875. <https://doi.org/10.1016/j.ijdrr.2020.101875>
- Coleman N, Esmalian A, Mostafavi A (2020b) Equitable Resilience in Infrastructure Systems: Empirical Assessment of Disparities in Hardship Experiences of Vulnerable Populations during Service Disruptions. *Natural Hazards Review*, 21(4). [https://doi.org/10.1061/\(asce\)nh.1527-6996.0000401](https://doi.org/10.1061/(asce)nh.1527-6996.0000401)
- Cutter SL, Ahearn JA, Amadei B, Crawford P, Eide EA, Galloway GE, Goodchild MF, Kunreuther HC, Li-Vollmer M, Schoch-Spana M, Scrimshaw SC, Stanley EM, Whitney G, Zoback ML (2013) Disaster resilience: a national imperative. *Environment* 55(2):25–29. <https://doi.org/10.1080/00139157.2013.768076>
- Cutter SL, Barnes L, Berry M, Burton C, Evans E, Tate E, Webb J (2008) A place-based model for understanding community resilience to natural disasters. *Global Environmental Change* 18(4):598–606. <https://doi.org/10.1016/j.gloenvcha.2008.07.013>
- Delilah Roque A, Pijawka D, Wutich A (2020) The role of social capital in resiliency: Disaster recovery in Puerto Rico. *Risk, Hazards & Crisis in. Public Policy* 11(2):204–235
- Deng H, Aldrich DP, Danziger MM, Gao J, Phillips NE, Cornelius SP, Wang QR (2021) High-resolution human mobility data reveal race and wealth disparities in disaster evacuation patterns. *Humanities and Social Sciences Communications* 8(1):144. <https://doi.org/10.1057/s41599-021-00824-8>
- Dong S, Gao X, Mostafavi A, Gao J (2022) Modest flooding can trigger catastrophic road network collapse due to compound failure. *Communications Earth & Environment*, 3(1). <https://doi.org/10.1038/s43247-022-00366-0>
- Eid MS, El-adaway IH (2018) Decision-Making Framework for Holistic Sustainable Disaster Recovery: Agent-Based Approach for Decreasing Vulnerabilities of the Associated Communities. *Journal of Infrastructure Systems*, 24(3). [https://doi.org/10.1061/\(asce\)is.1943-555x.0000427](https://doi.org/10.1061/(asce)is.1943-555x.0000427)
- Emrich CT, Aksha SK, Zhou Y (2022) Assessing distributive inequities in FEMA’s Disaster recovery assistance fund allocation. *International Journal of Disaster Risk Reduction* 74:102855. <https://doi.org/10.1016/j.ijdrr.2022.102855>
- ERCOT (2017) *ERCOT Responds to Hurricane Harvey*. <https://www.ercot.com/help/harvey>
- Esmalian A, Coleman N, Yuan F, Xiao X, Mostafavi A (2022) 12, 20203. Characterizing equitable access to grocery stores during disasters using location-based data. *Scientific Reports*. <https://doi.org/10.1038/s41598-022-23532-y>
- Fan C, Jiang X, Lee R, Mostafavi A (2022) Equality of access and resilience in urban population-facility networks Abstract. *npj Urban Sustainability* 2:103805. <https://doi.org/10.1038/s42949-022-00051-3>
- Farahmand H, Wang W, Mostafavi A, Maron M (2022) 49, 1893-1911. Anomalous human activity fluctuations from digital trace data signal flood inundation status. *Environment and Planning B: Urban Analytics and City Science*, 23998083211069990. <https://doi.org/10.1177/23998083211069990>
- Farahmand H, Xu Y, Mostafavi A (2021) A spatial-temporal graph deep learning model for urban flood nowcasting leveraging heterogeneous community features, A Spatial-temporal Graph Deep Learning Model for Urban Flood Nowcasting Leveraging Heterogeneous Community Features. *arXiv preprint arXiv:2111.08450*
- FEMA (2016) *National Disaster Recovery Framework, Second Edition* (National Response and Disaster Recovery Frameworks, Issue
- FEMA. (2018). *FEMA- Harvey Flood Depths*. <http://www.hydroshare.org/resource/e8768f4cb4d5478a96d2b1cbd00d9e85>
- FEMA (2020) *Community Lifelines*. <https://www.fema.gov/emergency-managers/practitioners/lifelines>
- Finch C, Emrich CT, Cutter SL (2010) Disaster disparities and differential recovery in New Orleans. *Population and Environment* 31(4):179–202. <https://doi.org/10.1007/s11111-009-0099-8>
- Finucane, ML, Acosta, J, Wicker, A, & Whipkey, K (2020). Short-term solutions to a long-term challenge: Rethinking disaster recovery planning to reduce vulnerabilities and inequities. *International Journal of Environmental Research and Public Health*, 17(2). <https://doi.org/10.3390/ijerph17020482>
- Fotovatikhah F, Herrera M, Shamshirband S, Chau KW, Ardabili SF, Piran MJ (2018) Survey of computational intelligence as basis to big flood management: Challenges, research directions and future work. *Engineering Applications of Computational Fluid Mechanics* 12(1):411–437. <https://doi.org/10.1080/19942060.2018.1448896>
- Han SY, Tsou M-H, Knaap E, Rey S, Cao G (2019) How Do Cities Flow in an Emergency? Tracing Human Mobility Patterns during a Natural Disaster with Big Data and Geospatial Data Science. *Urban Science* 3(2):51. <https://doi.org/10.3390/urbansci3020051>
- Hong B, Bonczak BJ, Gupta A, Kontokosta CE (2021) Measuring inequality in community resilience to natural disasters using large-scale mobility data. *Nature Communications* 12(1):1870. <https://doi.org/10.1038/s41467-021-22160-w>
- Horney J, Dwyer C, Aminto M, Berke P, Smith G (2017) Developing indicators to measure post-disaster community recovery in the United States. *Disasters* 41(1):124–149. <https://doi.org/10.1111/disa.12190>

- Huang X, Lu J, Gao S, Wang S, Liu Z, Wei H (2022) Staying at Home Is a Privilege: Evidence from Fine-Grained Mobile Phone Location Data in the United States during the COVID-19 Pandemic. *Annals of the American Association of Geographers* 112(1):286–305. <https://doi.org/10.1080/24694452.2021.1904819>
- Jordan E, Javernick-Will A (2013) Indicators of Community Recovery: Content Analysis and Delphi Approach. *Natural Hazards Review* 14(1):21–28. [https://doi.org/10.1061/\(asce\)nh.1527-6996.0000087](https://doi.org/10.1061/(asce)nh.1527-6996.0000087)
- Kontokosta CE, Malik A (2018) The Resilience to Emergencies and Disasters Index: Applying big data to benchmark and validate neighborhood resilience capacity. *Sustainable Cities and Society* 36:272–285. <https://doi.org/10.1016/j.scs.2017.10.025>
- Lee D (2020) The Impact of Natural Disasters on Neighborhood Poverty Rate: A Neighborhood Change Perspective. *Journal of Planning Education and Research* 40(4):447–459. <https://doi.org/10.1177/0739456X18769144>
- Lenormand M, Louail T, Cantú-Ros OG, Picornell M, Herranz R, Arias JM, Barthelemy M, Miguel MS, Ramasco JJ (2015) Influence of sociodemographic characteristics on human mobility [corrected]. *Scientific reports* 5:5
- Li Q, Tang Z, Coleman N, Mostafavi A (2021) Detecting Early-Warning Signals in Time Series of Visits to Points of Interest to Examine Population Response to COVID-19 Pandemic. *IEEE Access* 9:27189–27200. <https://doi.org/10.1109/ACCESS.2021.3058568>
- Links JM, Schwartz BS, Lin S, Kanarek N, Mitrani-Reiser J, Sell TK, Watson CR, Ward D, Slemple C, Burhans R, Gill K, Igusa T, Zhao X, Aguirre B, Trainor J, Nigg J, Inglesby T, Carbone E, Kendra JM (2018) COPEWELL: A Conceptual Framework and System Dynamics Model for Predicting Community Functioning and Resilience after Disasters. *Disaster Medicine and Public Health Preparedness* 12(1):127–137. <https://doi.org/10.1017/dmp.2017.139>
- Ma, J, Li, B, & Mostafavi, A (2022). Characterizing Urban Lifestyle Signatures Using Motif Properties in Network of Places. *ArXiv, abs/2204.01103*
- Masud-All-Kamal M, Monirul Hassan SM (2018) The link between social capital and disaster recovery: evidence from coastal communities in Bangladesh. *Natural Hazards* 93(3):1547–1564. <https://doi.org/10.1007/s11069-018-3367-z>
- Mayer B (2019) A Review of the Literature on Community Resilience and Disaster Recovery. In *Current environmental health reports* 6:167–173
- Microsoft Open Source. US Building Footprints. *GitHub*. <https://github.com/microsoft/USBuildingFootprints>
- Mitsova D, Escaleras M, Sapat A, Esnard AM, Lamadrid AJ (2019) The effects of infrastructure service disruptions and socio-economic vulnerability on Hurricane recovery. *Sustainability (Switzerland)* 11(2):516. <https://doi.org/10.3390/su11020516>
- Mitsova D, Esnard AM, Sapat A, Lai BS (2018) Socioeconomic vulnerability and electric power restoration timelines in Florida: the case of Hurricane Irma. *Natural Hazards* 94(2):689–709. <https://doi.org/10.1007/s11069-018-3413-x>
- Nan C, Sansavini G (2017) A quantitative method for assessing resilience of interdependent infrastructures. *Reliability Engineering and System Safety* 157:35–53. <https://doi.org/10.1016/j.res.2016.08.013>
- Pappalardo L, Simini F, Rinzivillo S, Pedreschi D, Giannotti F, Barabási AL (2015) Returners and explorers dichotomy in human mobility. *Nature Communications* 6:8166. <https://doi.org/10.1038/ncomms9166>
- Peacock WG, Van Zandt S, Zhang Y, Highfield WE (2014) Inequities in long-term housing recovery after disasters. *Journal of the American Planning Association* 80(4):356–371. <https://doi.org/10.1080/01944363.2014.980440>
- Platt S, Gautam D, Rupakhty R (2020) Speed and quality of recovery after the Gorkha Earthquake 2015 Nepal. *International Journal of Disaster Risk Reduction* 50:101689. <https://doi.org/10.1016/j.ijdrr.2020.101689>
- Podesta C, Coleman N, Esmalian A, Yuan F, Mostafavi A (2021) Quantifying community resilience based on fluctuations in visits to points-of-interest derived from digital trace data. *Journal of the Royal Society Interface* 18(177):20210158. <https://doi.org/10.1098/rsif.2021.0158>
- Rosenblum AJ, Wend CM, Akhtar Z, Rosman L, Freeman JD, Barnett DJ (2021) Use of Big Data in Disaster Recovery: An Integrative Literature Review. In *Disaster Medicine and Public Health Preparedness* 17:68
- Sadri AM, Ukkusuri SV, Lee S, Clawson R, Aldrich D, Nelson MS, Seipel J, Kelly D (2018) The role of social capital, personal networks, and emergency responders in post-disaster recovery and resilience: a study of rural communities in Indiana. *Natural Hazards* 90(3):1377–1406. <https://doi.org/10.1007/s11069-017-3103-0>
- SafeGraph. *POI Places Data for a Changing World*. <https://www.safegraph.com/>
- Salley C, Mohammadi N, Taylor JE (2019) Semi-Supervised Machine Learning Framework for Fusing Georeferenced Data from Social Media and Community-Driven Applications. *Computing in Civil Engineering* 2021:114–122
- Sevtsuk A, Ratti C (2010) Does Urban mobility have a daily routine? Learning from the aggregate data of mobile networks. *Journal of Urban Technology* 17(1):41–60. <https://doi.org/10.1080/10630731003597322>
- Sovacool BK, Tan-Mullins M, Abrahamse W (2018) Bloated bodies and broken bricks: Power, ecology, and inequality in the political economy of natural disaster recovery. *World Development* 110:243–255. <https://doi.org/10.1016/j.worlddev.2018.05.028>
- Spectus (n.d.) *Spectus - Data Clean Room for Human Mobility Analysis*. <https://spectus.ai/>
- Tatsuki S (2007) Long-term Life Recovery Processes Among Survivors of the 1995 Kobe Earthquake: 1999, 2001, 2003, and 2005 Life Recovery Social Survey Results. *Journal of disaster research* 2:484–501
- Thuillier E, Moalic L, Lamrous S, Caminada A (2018) Clustering Weekly Patterns of Human Mobility Through Mobile Phone Data. *IEEE Transactions on Mobile Computing* 17(4):817–830. <https://doi.org/10.1109/TMC.2017.2742953>
- Tierney K, Oliver-Smith A (2012) Social Dimensions of Disaster Recovery. *International Journal of Mass Emergencies and Disasters* 30(2):123–146
- Toole, JL, Herrera-Yaque, C, Schneider, CM, & González, MC (2015). Coupling human mobility and social ties. *Journal of the Royal Society Interface*, 12(105). <https://doi.org/10.1098/rsif.2014.1128>
- United States Office of Management and Budget. (2022). *North American Industry Classification System*. [https://www.census.gov/naics/reference\\_files\\_tools/2022\\_NAICS\\_Manual.pdf](https://www.census.gov/naics/reference_files_tools/2022_NAICS_Manual.pdf)
- US Census Data. 2017 Census Data. <https://www.census.gov/data.html>
- Vugrin, ED, Warren, DE, Ehlen, MA, & Camphouse, RC (2010), 77-116. A framework for assessing the resilience of infrastructure and economic systems. In *Sustainable and Resilient Critical Infrastructure Systems: Simulation, Modeling, and Intelligent Engineering*. [https://doi.org/10.1007/978-3-642-11405-2\\_3](https://doi.org/10.1007/978-3-642-11405-2_3)
- Wang F, Wang J, Cao J, Chen C, Ban X (2019) Extracting trips from multi-sourced data for mobility pattern analysis: An app-based data example. *Transportation Research Part C: Emerging Technologies* 105:183–202. <https://doi.org/10.1016/j.trc.2019.05.028>
- Wang Q, Taylor JE (2014) Quantifying human mobility perturbation and resilience in hurricane sandy. *PLoS ONE* 9(11):112608. <https://doi.org/10.1371/journal.pone.0112608>
- Wang Q, Taylor JE (2016) Patterns and limitations of urban human mobility resilience under the influence of multiple types of natural disaster. *PLoS ONE* 11(1):0147299. <https://doi.org/10.1371/journal.pone.0147299>
- Wang Y, Taylor JE (2018) Coupling sentiment and human mobility in natural disasters: a Twitter-based study of the 2014 South Napa Earthquake. *Natural Hazards* 92(2):907–925. <https://doi.org/10.1007/s11069-018-3231-1>
- Yabe T, Tsubouchi K, Fujiwara N, Sekimoto Y, Ukkusuri SV (2020) Understanding post-disaster population recovery patterns. *Journal of the Royal Society Interface* 17(163):20190532. <https://doi.org/10.1098/rsif.2019.0532>
- Yu M, Yang C, Li Y (2018) Big data in natural disaster management: a review. In *Geosciences (Switzerland)* 8:165
- Yuan F, Esmalian A, Oztekin B, Mostafavi A (2022) Unveiling spatial patterns of disaster impacts and recovery using credit card transaction fluctuations. *Environment and Planning B: Urban Analytics and City Science* 49(9):2378–2391. <https://doi.org/10.1177/23998083221090246>
- Yuan F, Fan C, Farahmand H, Coleman N, Esmalian A, Lee C-CB, Patrascu FI, Zhang C, Dong S, & Mostafavi A (2021) Smart Flood Resilience: Harnessing Community-Scale Big Data for Predictive Flood Risk Monitoring, Rapid Impact Assessment, and Situational Awareness. *Environmental Research Infrastructure and Sustainability*
- Yuan, F, & Liu, R (2020). Mining Social Media Data for Rapid Damage Assessment during Hurricane Matthew: Feasibility Study. *Journal of Computing in Civil Engineering*, 34(3). [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000877](https://doi.org/10.1061/(asce)cp.1943-5487.0000877)
- Yuan F, Yang Y, Li Q, Mostafavi A (2022) Unraveling the Temporal Importance of Community-Scale Human Activity Features for Rapid Assessment of Flood Impacts. *IEEE Access* 10:1138–1150. <https://doi.org/10.1109/ACCESS.2021.3137651>
- Zipline's Logistics consultants. (2017). *Transportation Impacts from Hurricane Harvey*. <https://ziplineanalytics.com/blog/hurricane-harvey/>

## Acknowledgements

We also would like to acknowledge the data support from Spectus Inc and SafeGraph. This material is based in part upon work supported by the National Science Foundation under Grant CMMI-1846069 (CAREER) and the support of the National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, Spectus Inc, or SafeGraph.

## Author contributions

All authors critically revised the manuscript, gave final approval for publication, and agree to be held accountable for the work performed therein. NC was the lead Ph.D. student researcher and first author, who was responsible for supervising data collection, performing the main part of the analysis, interpreting the significant results, and writing most of the manuscript. CL was responsible for guiding data collection and assisting in the methodology sections of the project and manuscript. YZ assisted in analyzing and

processing the location intelligence data. AM was the faculty advisor for the project and provided critical feedback on the project development and manuscript.

### Competing interests

The authors declare no competing interests.

### Ethical Approval

This article does not contain any studies with human participants performed by any of the authors.

### Informed Consent

This article does not contain any studies with human participants performed by any of the authors.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1057/s41599-023-02312-7>.

**Correspondence** and requests for materials should be addressed to Natalie Coleman.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2023