

# Understanding the socioeconomic trends in areas impacted by flooding during Hurricane Harvey in Houston, Texas

By: Luisa Florez  
GIS in Water Resources CE 394K  
Final Project  
The University of Texas at Austin  
12/07/2018



David J. Phillip/Associated Press (Photograph). (2018)

# Table of Contents

## Contents

<b>1 Introduction</b> .....	3
1.1 Background .....	3
1.2 Site Information .....	4
1.3 Project Objectives .....	4
<b>2 Methodology</b> .....	5
2.1 Data Collection .....	5
2.1.1 County Level Data .....	5
2.1.2 Flooding Data .....	5
2.1.3 Socioeconomic Data .....	6
2.2 Data Processing .....	7
2.2.1 Flood Indicator Data .....	7
2.2.2 Socioeconomic Data .....	7
2.3 Limitations of data .....	8
<b>3 Results</b> .....	8
3.1 Flooding .....	8
3.2 Socioeconomic Trends .....	10
3.2.1 Income .....	10
3.2.2 Social Vulnerability Index .....	12
3.1.3 Unemployment .....	15
3.3 Future Work .....	18
<b>4 Conclusion</b> .....	19
<b>Work Cited</b> .....	19

# 1 Introduction

## 1.1 Background

On August 25th, 2017 category 3 hurricane Harvey made landfall in Rockport, Texas. 12 hours after its landfall, Hurricane Harvey was downgraded to a tropical storm (Blake & Zelinsky, 2017). A feature that differentiated Hurricane Harvey from previous storms was its extreme slow movement. From August 26 to August 30th, Hurricane Harvey hovered over southeast Texas, dropping 20-40 inches of rain (Figure 1). A city that suffered greatly from this large amount of rain was Houston, Texas. It resulted in 36 deaths and at least 160,000 flooded structures in Harris and Galveston Counties alone (Blake & Zelinsky, 2017).

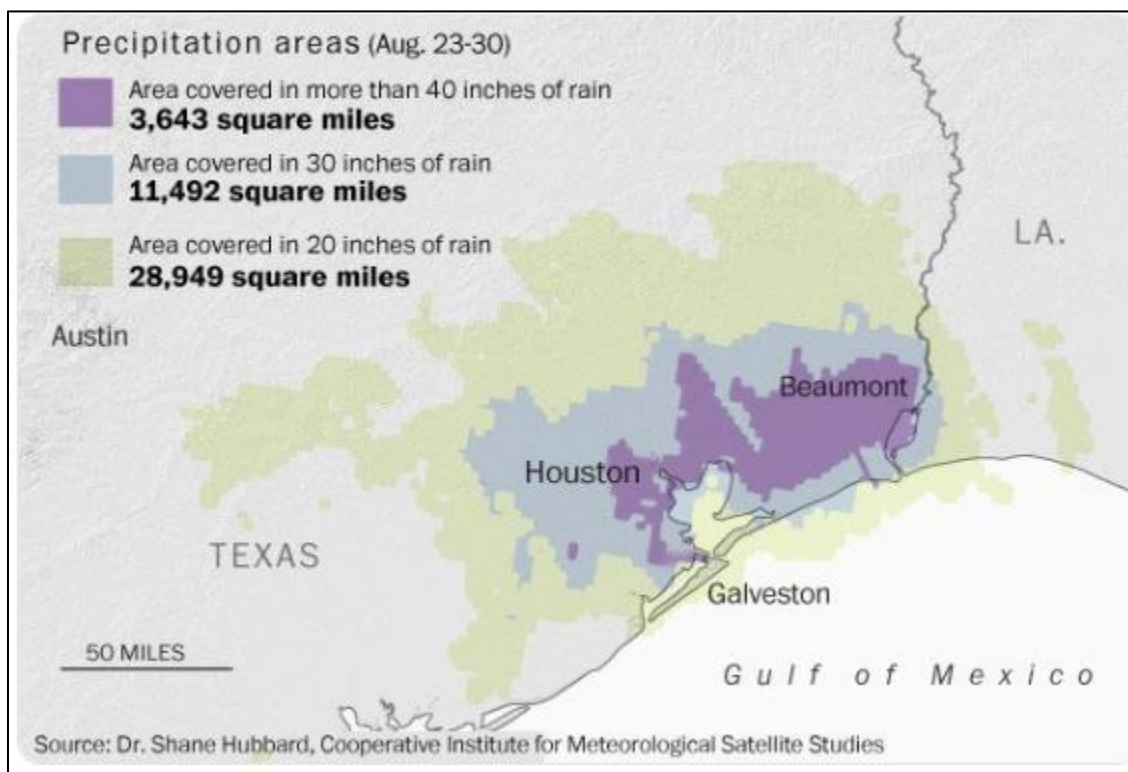


Figure 1. Precipitation areas from Aug.23 – Aug. 30 (@NSWBrownsville 2017).

On August 31st, 2017 a The New York Times article (Turkewitz & Burch, 2018) titled 'Storm With 'No Boundaries' Took Aim at Rich and Poor Alike' categorized Harvey as a storm that impacted both the rich and poor neighborhoods of the Houston metropolitan area. For this project, I aim to probe this statement and analyze the socioeconomic trends of the people who were affected by the floods.

## 1.2 Site Information

The area of interest for this project can be found in Figure 2. The study area consists of 10 counties including Harris, Fort Bend, Galveston, Montgomery, Liberty, Chambers, Brazoria, Waller, Austin and Walker counties. It has an area of approximately 8,000 square miles and a population of approximately 6.2 million (U.S. Census, 2012).

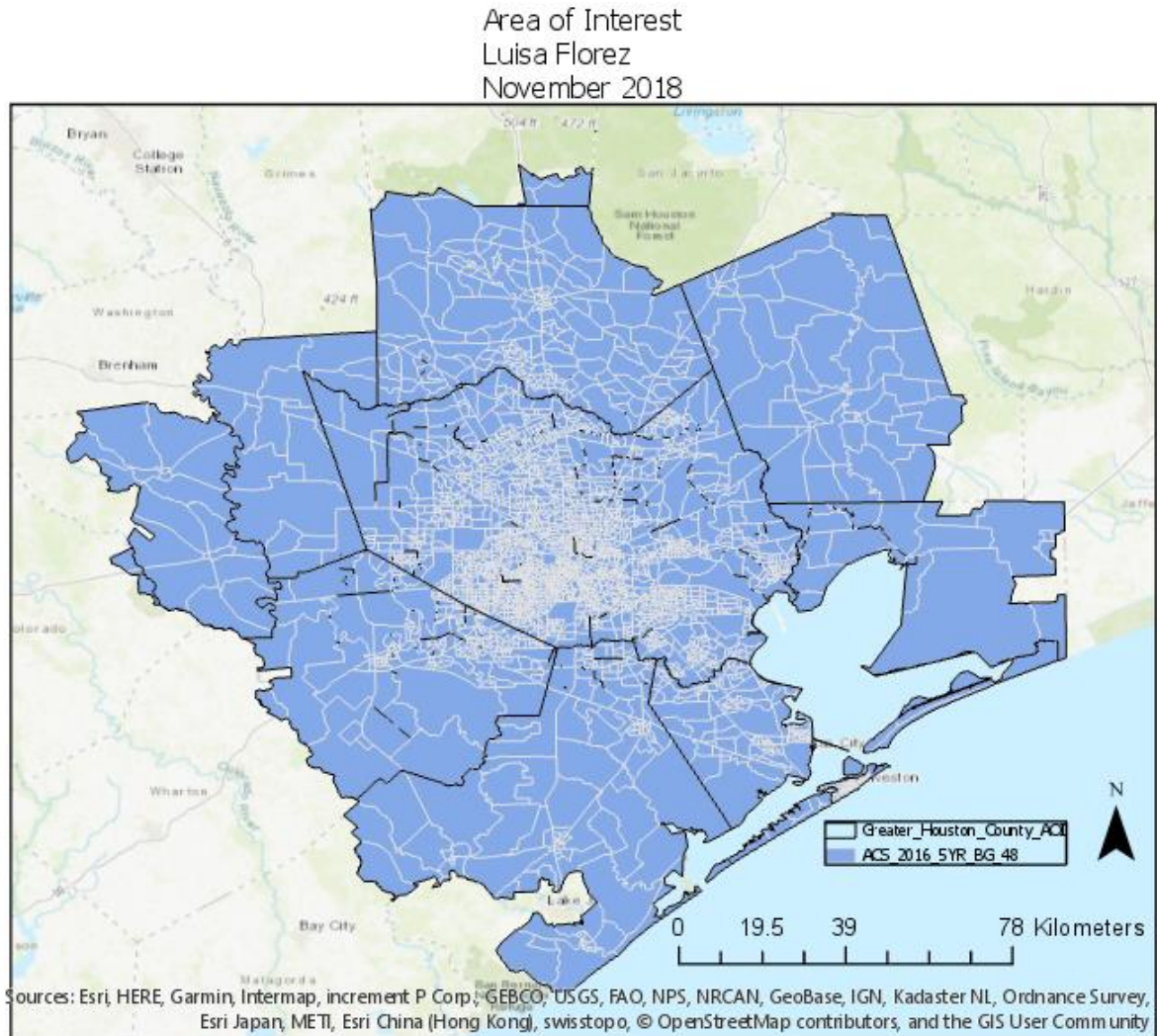


Figure 2. Area of interest.

## 1.3 Project Objectives

The objective for this project is to develop an understanding of the areas that were directly affected by the flooding produced from Hurricane Harvey by considering their income levels, unemployment levels and social vulnerability indices. Having a socioeconomic understanding of the area is important because it can help government agencies predict how a community will react to a disaster in terms of its resilience. This

analysis can help government bodies be proactive and forward-looking when determining the use of their resources.

## 2 Methodology

### 2.1 Data Collection

#### 2.1.1 County Level Data

The county level shapefile used is obtained from the City of Houston COGIS Open Data Portal (City of Houston, 2018a).

#### 2.1.2 Flooding Data

In order to get a spatial understanding of where the flooding occurred, “flood indicator” data points are used as a proxy. In this project, three sources are used as “flood indicators” (Figure 3). The first source is obtained from 2017 Fema Building Damage Assessments values (FEMA, 2017). A total of 187,031 (99.9% of all data) data points are obtained from this source. The second source, which provided 202 data points (0.1% of all data), comes from the number of calls made to the 311 Harris County Service Line to report flooding and drainage issues (City of Houston, 2018b). Lastly, the third source is obtained by floods that were reported via social media platforms. 2 data points are obtained from this category via Twitter (Homeland Infrastructure Foundation-Level Data, 2018) and consisted of 0.001% of all data.

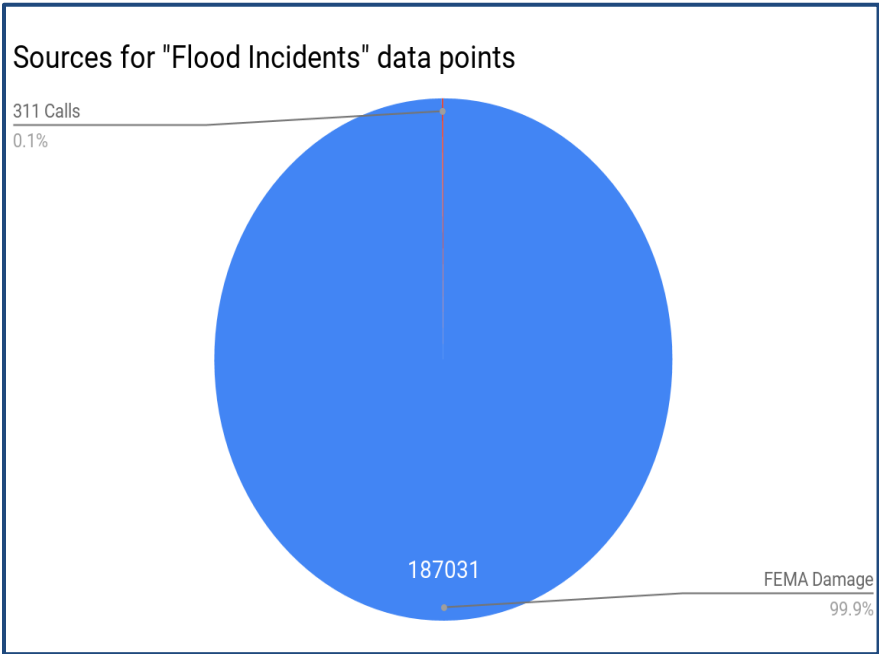


Figure 3. Sources for “Flood Incidents” data points.

In addition to the “flood indicator” data points, shapefiles for FEMA designated floodplains (Houston-Galveston Area Council, 2015) are obtained in order to develop an understanding of where flooding is usually expected throughout the Houston metropolitan area.

### 2.1.3 Socioeconomic Data

Three different socioeconomic parameters are explored in this project. The first parameter is income level and is obtained from the American Community Survey 5-year estimates for 2012 - 2016 for block group Texas (U.S. Census Bureau, 2011). In this project, the definition for income is defined to be the median household income for the past 12 months (in 2016 inflation-adjusted dollars). Income levels have been grouped into three groups that can be found in Table 1. The ranges for each level are determined using the Pew Research Center’s definition for middle-income which defines it as  $\frac{2}{3}$  to 2 times the U.S. median household income (Fry & Kochhar, 2018). For this study, the 2017 U.S. median household income of \$61,372 is used (Konish, 2018).

Table 1. Income level ranges used in this project

Income Level	Range
Low Income	<\$45,200
Middle Income	\$45,200 – \$135,600
High Income	\$135,600 – \$250,001

The second parameter, unemployment rates, is also collected from American Community Survey 5-year estimates for 2012 - 2016 for block group Texas (U.S. Census Bureau, 2011). The unemployment rates used for this analysis are calculated by dividing the number of unemployed individuals by the total number of individuals in the labor force (employed and unemployed). The four ranges used for the unemployment rate are as follows: <4.4%, 4.4% - 15%, 15% - 30%, 30% - 46%. The first range was determined by equaling it to the 2017 national unemployment rate. The next three ranges were determined through increments of approximately 10-15 percent units.

The last parameter used is the Center for Disease Control and Prevention’s (CDC) 2014 Social Vulnerability Index (SVI) and is obtained from Living Atlas (Data\_CDC, 2017). The SVI refers to the “resilience of the communities when confronted by external stresses on human health” (ATSDR n.d.). The flooding experienced due to Harvey qualifies as an external stress. The SVI is divided into four categories that combined represent the overall vulnerability (Figure 4). For the purpose of this project, data associated from only the socioeconomic status category will be used. The socioeconomic status category includes poverty levels, unemployment, income and education. Areas that possess a SVI between

0.00-0.25 have less social vulnerability and the area closest to a SVI value of 1.00 have the highest vulnerability.

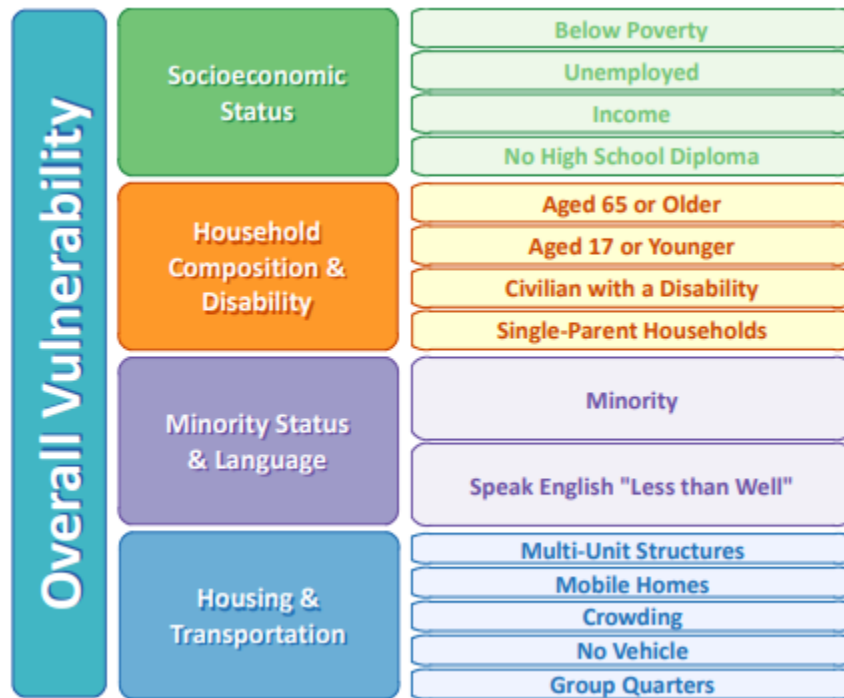


Figure 4. Categories that encompass the overall vulnerability considered in SVI (Hallisey et al. n.d.)

## 2.2 Data Processing

All shapefiles were uploaded into ArcGIS Pro. The scope of the data was limited to areas included within the counties boundary lines, which is depicted in Figure 2. Datasets that expanded to areas larger than the area of interest, such as the original SVI shapefile, were clipped to the county area.

### 2.2.1 Flood Indicator Data

The three data sets were combined using the “merge” analysis tool in ArcGIS Pro. In order to delete any repeated data entries, the “delete identical” feature was used.

### 2.2.2 Socioeconomic Data

Prior to data processing, different ranges for income levels, SVI, and unemployment levels were determined. Once each range was established, a query was done to select the polygons that were in each range, and afterwards, the polygons were exported to create a new shapefile. The purpose of isolating every range into its own layer was to be able to apply the select “completely within” feature in ArcGIS to the flood indicators. By using the select “completely within” query, the program is able to identify the number of flood

indicators that fall within each range. It is important to note that by selecting “completely within” it is possible that ArcGIS will not take flood indicators that are at a boundary of two polygons into account.

### 2.3 Limitations of data

It is important to acknowledge the limitations associated with this data. A major limitation is incomplete data sets, specifically in shapefiles obtained from the different agencies. Since each shapefile is obtained from a different city, state or federal agency, data may be missing for certain areas within the Houston metropolitan area. In addition, another limitation lies in the fact that the flood indicators that are used as a proxy to identify flooded areas are 1D and therefore may not capture the total spatial extent of the flooding. In addition, even though the majority of data points for the flood incidents come from FEMA determined damage assessments where FEMA employees survey the neighborhoods in person (Peterson, 2012), it is possible that a bias may exist in flooding incidents that are reported based on socioeconomic levels. Lastly, it is important to note that this data does not capture the scale of impact, such as the number of people, houses or structures affected.

## 3 Results

### 3.1 Flooding

Figure 5 depicts the FEMA designated floodplains in the Houston metropolitan area. As expected, most of the 100- and 500- year floodplains follow the path of the major rivers that cross throughout the area. When the floodplain locations are compared to the areas where significant flooding incidents occurred (Figure 6), we find that flood incidents do occur along the banks of the rivers. However, an interesting observation to note is that the areas with the densest concentration of flood indicators in Figure 6 occur in the western part of the city, and not necessarily on the banks of a major river.



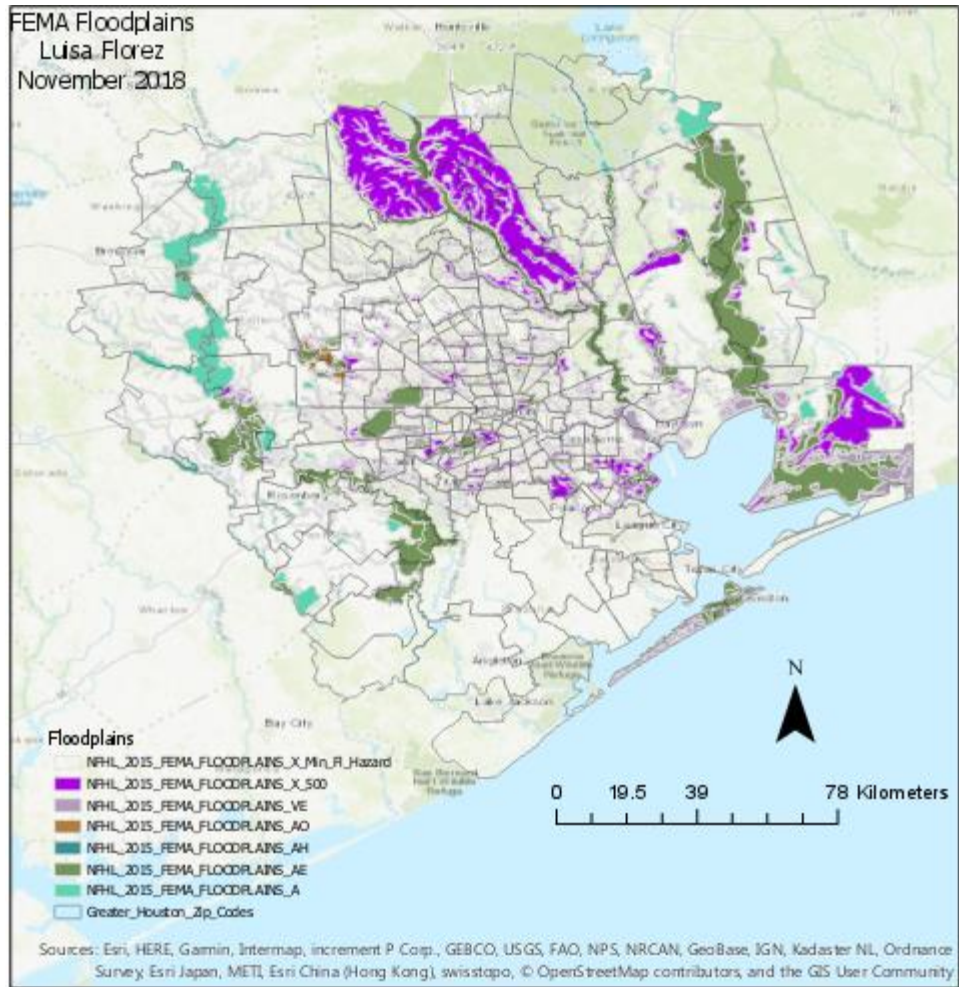


Figure 5. Floodplains for the Houston metropolitan areas

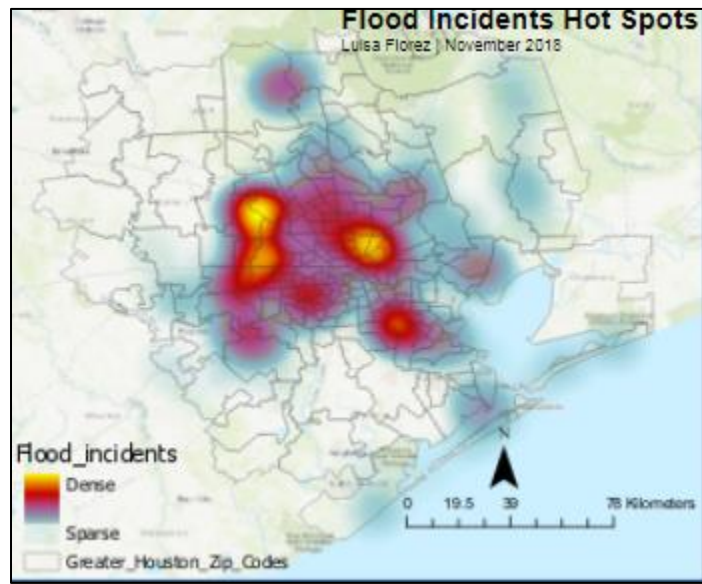


Figure 6. Hot spot map of flood incidents.

## 3.2 Socioeconomic Trends

### 3.2.1 Income

Figure 7 shows the income distribution for the Houston metropolitan area. The largest area belongs to the middle income level and is spread throughout the city. It is followed in area by the lowest income level, which is concentrated mainly in the northeast of the metropolitan area, and is then followed by the area of highest income levels, which is located mainly in the southwest area of the city. A graph comparing all three areas can be seen in Figure 8. Once the location of the flood indicators are overlaid on the income levels (Figure 9), as expected, the income level with the largest number of flood incidents is the middle income level (Figure 10). However, when the density of flood indicators is calculated in a per square mile basis, we see the high income level experiencing almost 3 times as many incidents per square mile than the middle income level (Figure 11).

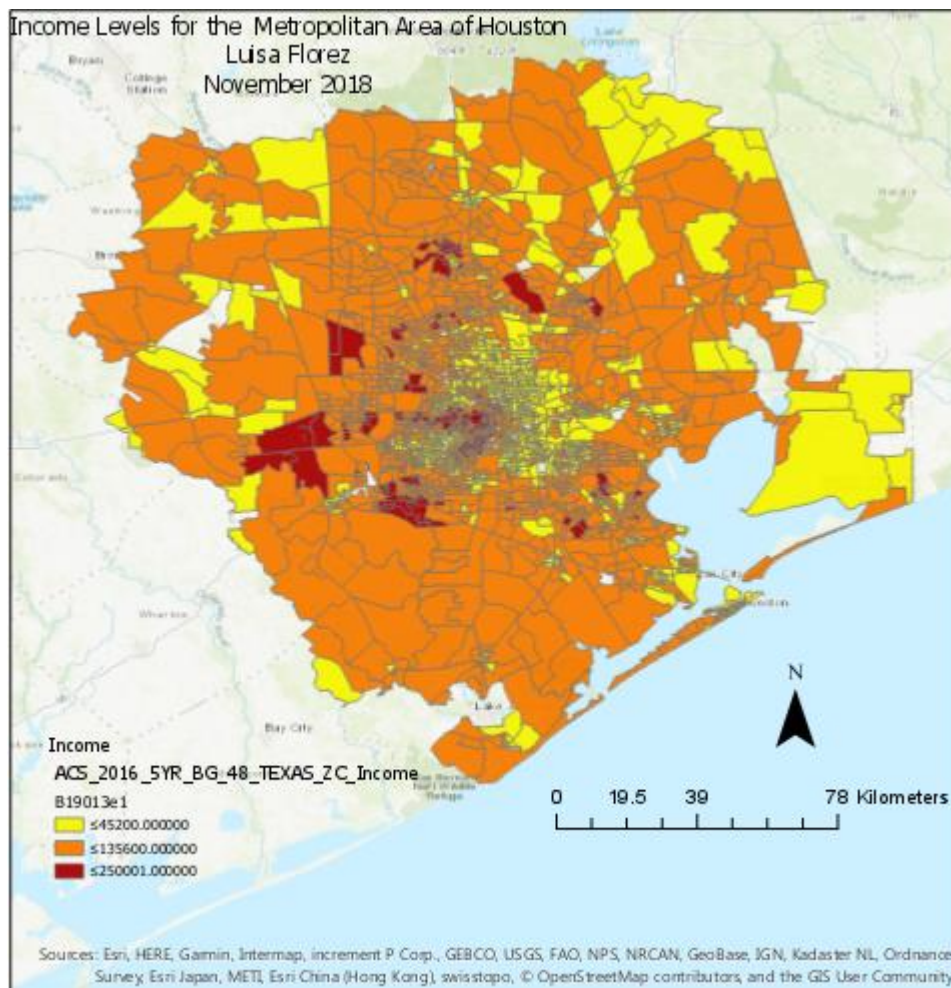


Figure 7. Income levels for the Houston metropolitan area.

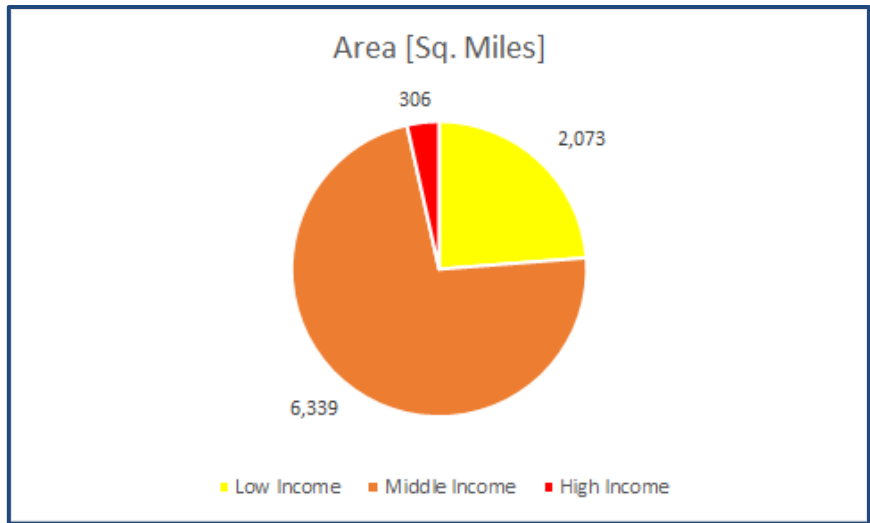


Figure 8. Area based on income level.

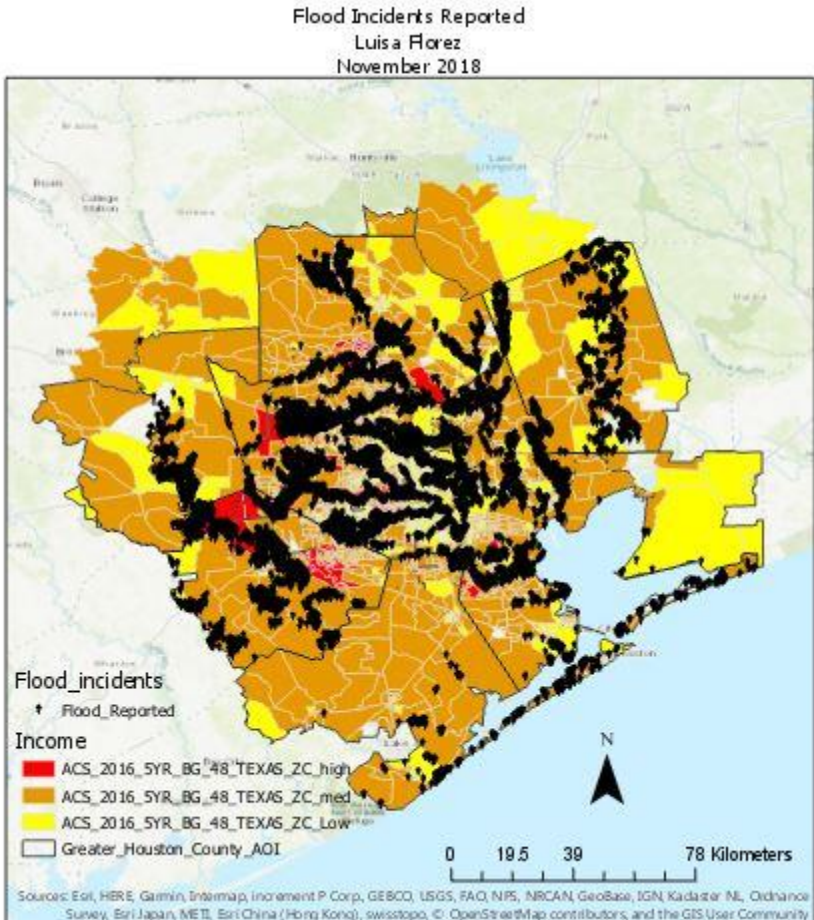


Figure 9. Flood incidents and income levels.

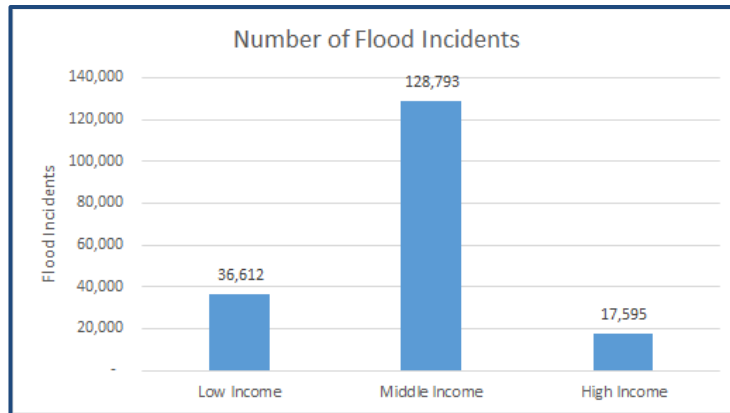


Figure 10. Number of total flood incidents.

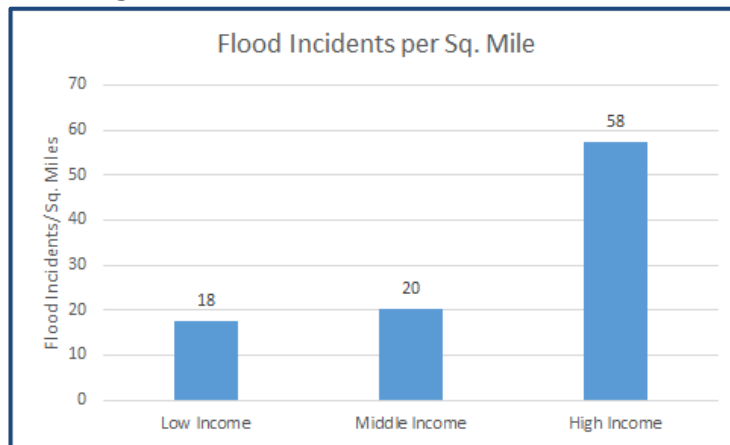


Figure 11. Number of flood incidents per square mile.

### 3.2.2 Social Vulnerability Index

According to Figure 12, the population with the highest vulnerability in the Houston metropolitan area is located predominantly in the northeast and northwest of the city. The population with the lowest social vulnerability seems to circle the downtown area of Houston with a higher concentration on the west side of town. As is seen in Figure 13, the majority of the area is comprised of the second most vulnerable group which has a SVI between 0.5-0.75. Interestingly, the groups with the lowest vulnerability (SVI=0.00-0.25) and the highest vulnerability (SVI=0.75-1.00) have similar total areas. Nonetheless, the total number of flood incidents (Figure 14) and flood incidents per square mile (Figure 15) is still larger in the lowest vulnerability group than in any other group. A map of the location of flood indicators in respect to the SVI distribution can be found in Figure 16.

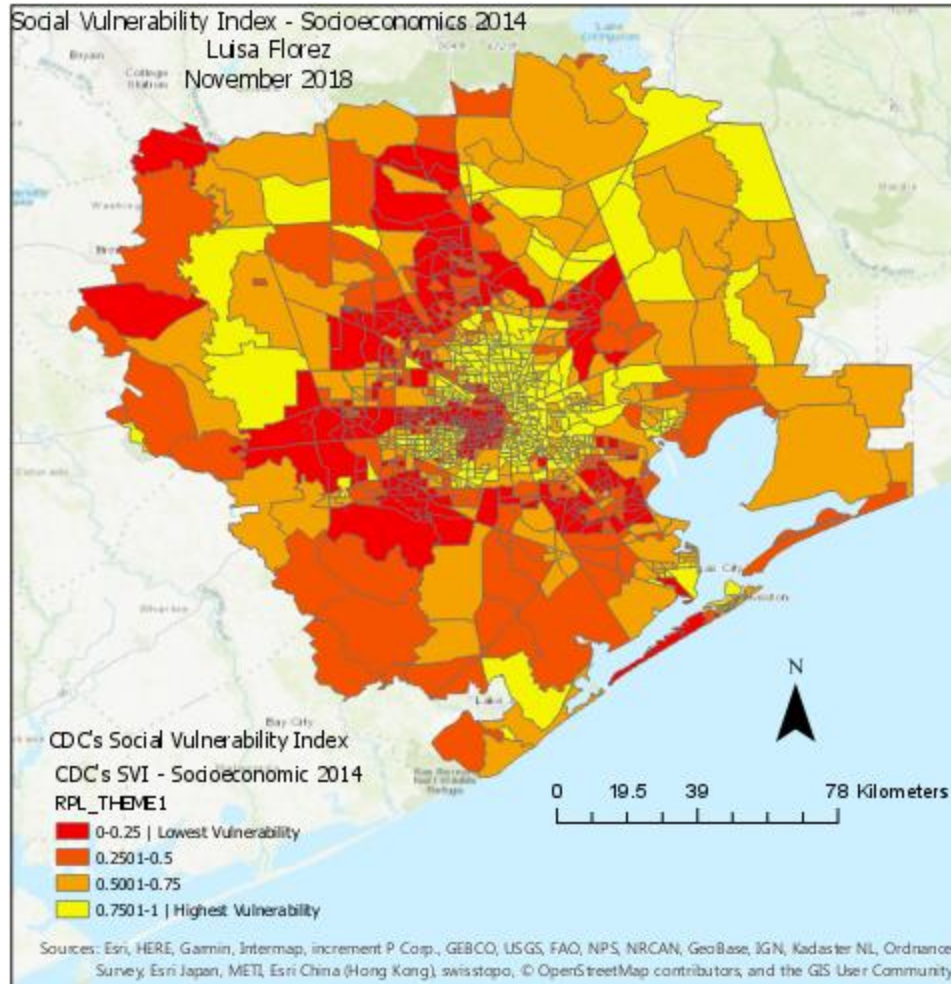


Figure 12. Social vulnerability index for the Houston Metropolitan area.

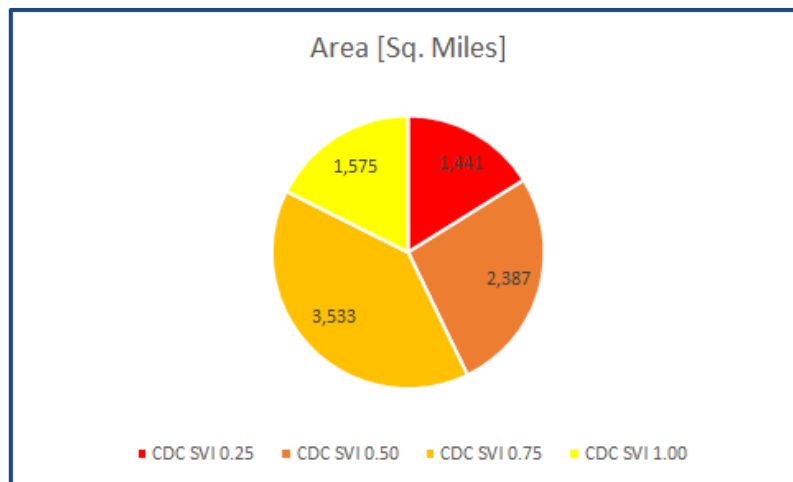


Figure 13. Area based on SVI values.

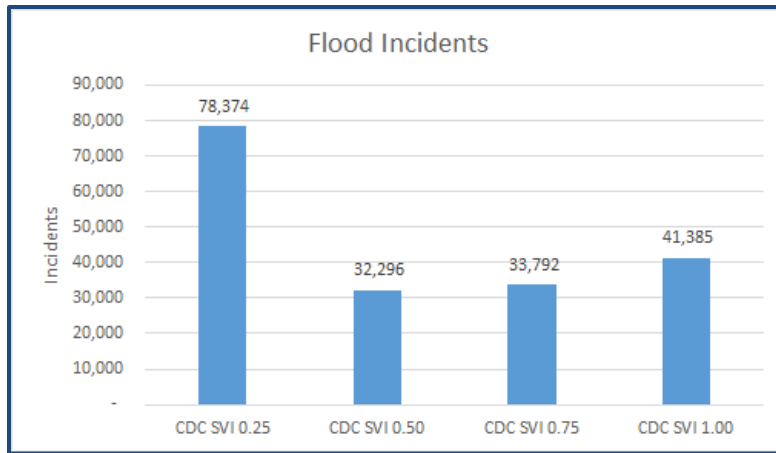


Figure 14. Number of total flood incidents based on SVI value.

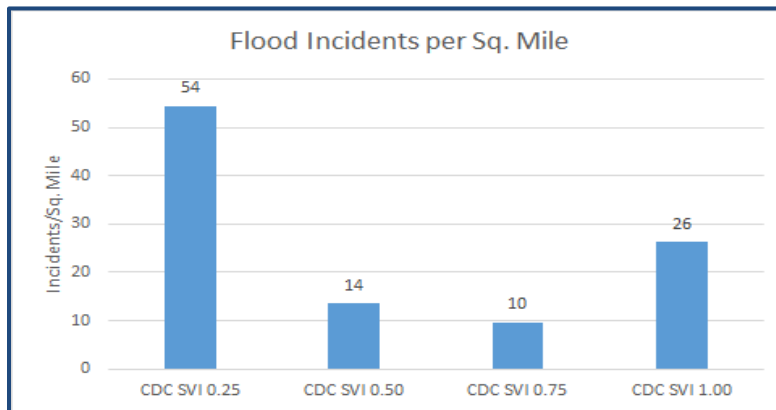


Figure 15. Flood incidents per square mile based on SVI value.

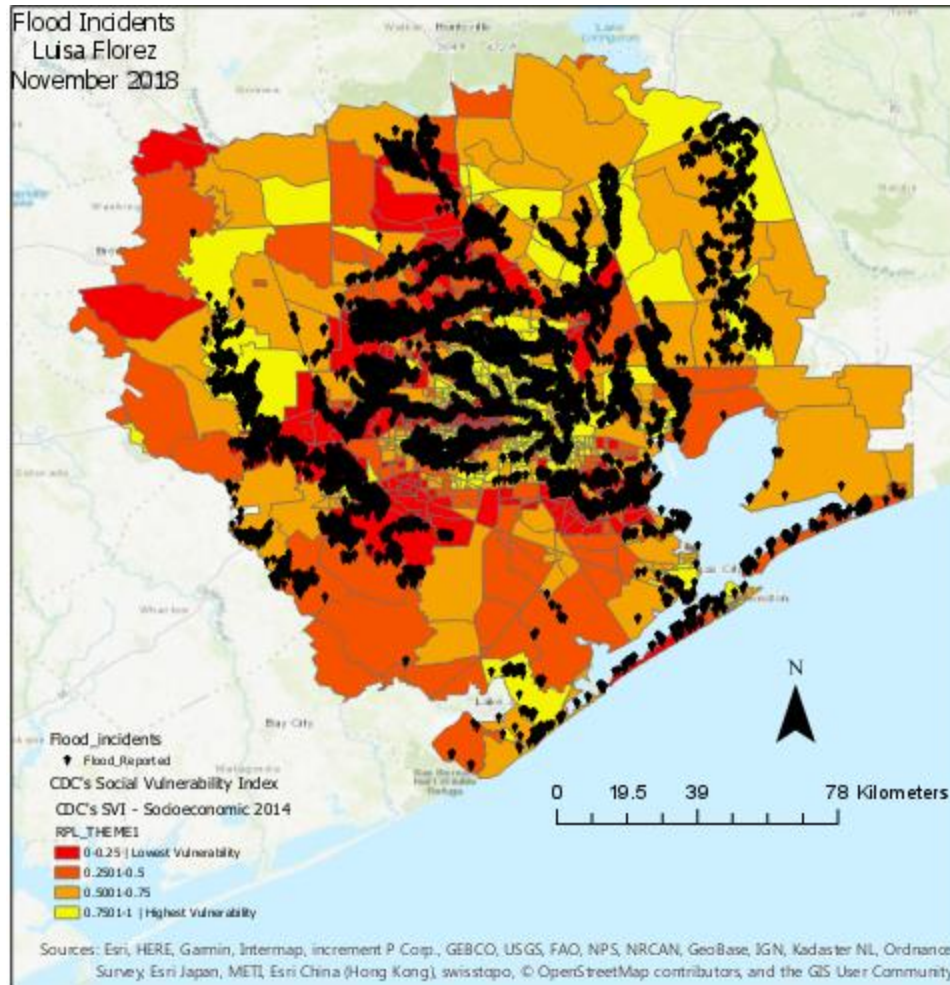


Figure 16. Flood incidents and SVI levels.

### 3.1.3 Unemployment

In terms of unemployment, the majority of the Houston metropolitan area has an unemployment percentage less than 4.4% or between 4.4% -15%. This can be visually appreciated in Figure 17. Figure 18 demonstrates the location of the flood incidents in regards to unemployment percentages. Since the total area for these two ranges are very similar (Figure 19), it is expected that the total number of flood incidents for these two groups will be similar. Figure 20 reinforces this expectation. However, in terms of flood incidents per square mile, the area with the highest rate of unemployment (30%-46%) has approximately two times as many incidents than the other groups (Figure 21).

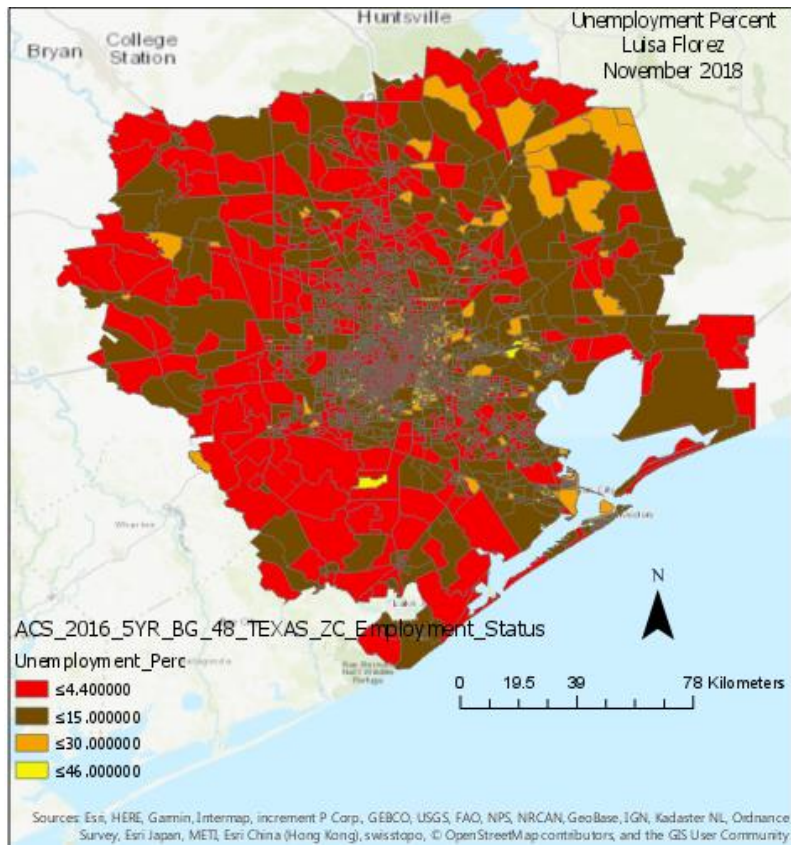


Figure 17. Unemployment rates in the Houston metropolitan area.

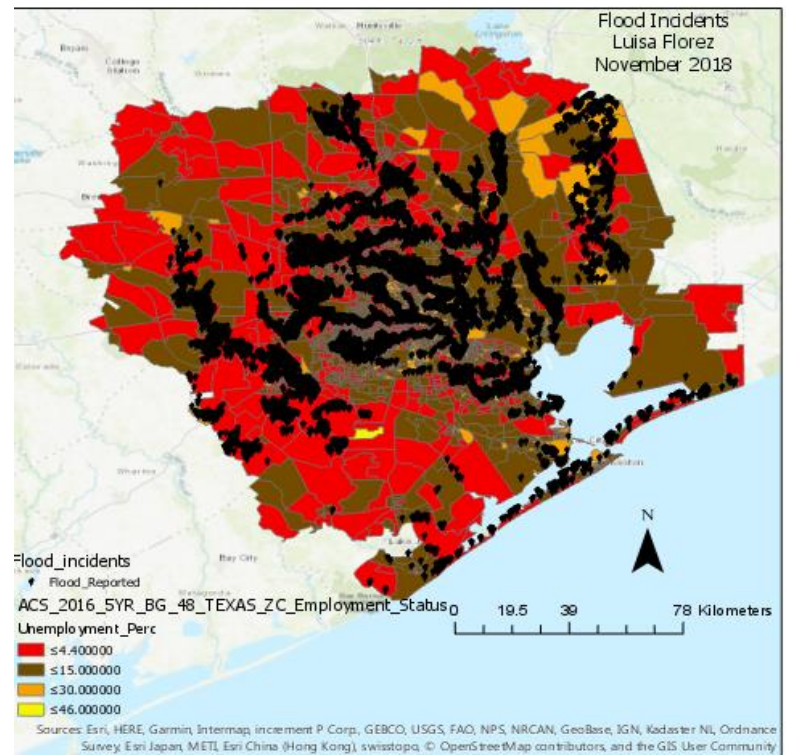


Figure 18. Flood incidents with respect to unemployment percentages.



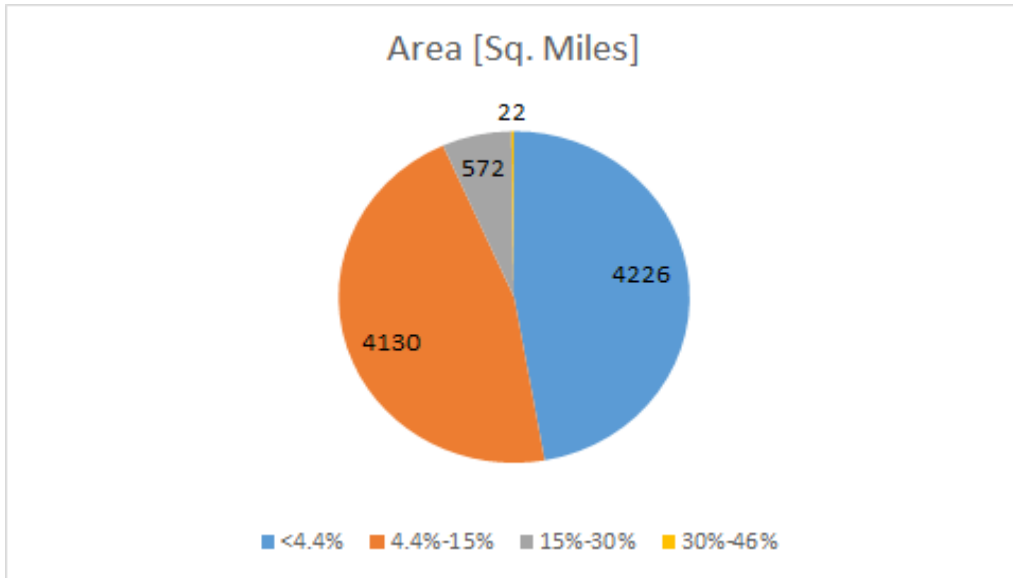


Figure 19. Area in terms of unemployment percentages.

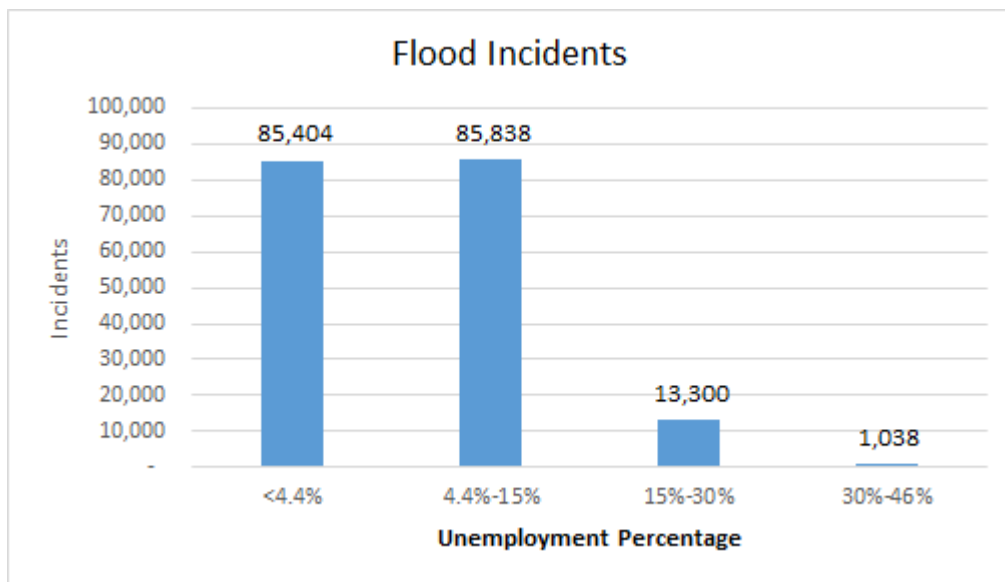


Figure 20. Number of total flood incidents based on unemployment percentage.

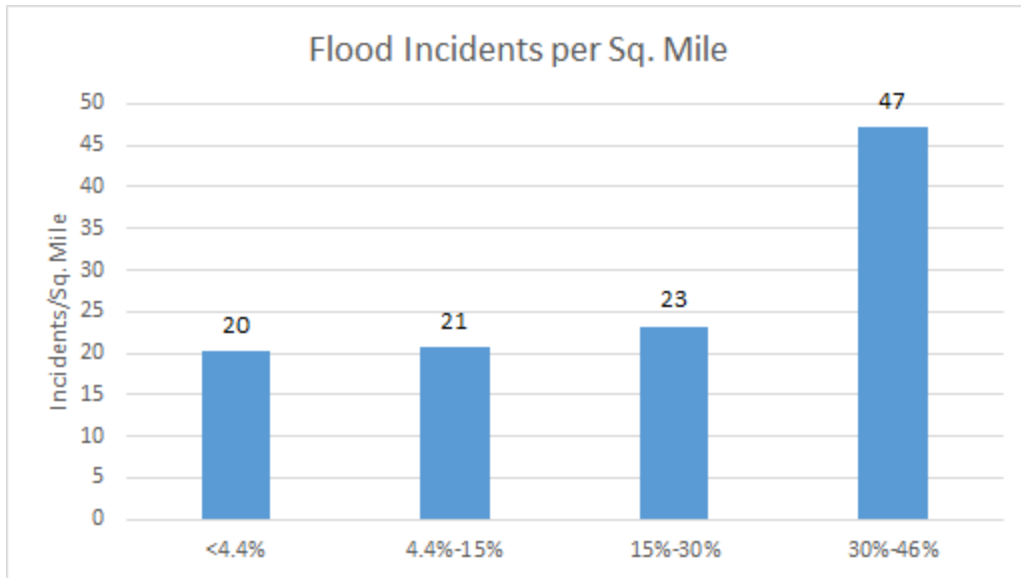


Figure 21. Flood incidents per square mile based on unemployment percentage.

### 3.3 Future Work

As was seen with the income level, SVI and unemployment percentage analysis, the group that would generally fall in the higher socioeconomic level experienced a large number of flooding incidents, a finding that may not have been expected. A potential hypothesis for this may be due to the flooding that occurred as a result of opening the flood gates at the Barker and Addicks Reservoirs. The areas affected by the release of the reservoirs are generally higher socioeconomic status neighborhoods. Had these reservoirs not been on the verge of overflowing, perhaps the number of flood incidents in high socioeconomic status areas would have been less. A way to investigate this would be to eliminate incidents in the proximity of the two reservoirs and perform the analysis again. In addition, another exercise worth exploring would be to develop an understanding of what socioeconomic groups are located in the current 100- and 500-year FEMA floodplains.

To further understand the socioeconomic trends in areas affected by Hurricane Harvey's flooding, it would be interesting to see the total number of flooding incidents and flooding indicator densities in terms of the distribution of education levels and poverty levels across the city. With these two extra parameters, all four of the subcategories used to determine the socioeconomic status in the CDC's SVI would be analyzed. In addition, it would also be interesting to see the spatial distribution of households with and without flood insurance and identify how many flood incidents are in each.

As mentioned in the "limitations of data" section, there may be a bias in regards to which flooding events are reported in terms of the socioeconomic status of the

neighborhood. For example, it may be possible that neighborhoods with higher socioeconomic status have more resources to report flooding than neighborhoods with lower socioeconomic status, and therefore resulting in more flood indicator data points being concentrated in the higher socioeconomic area. This could give the impression that more flooding occurred in specific areas. Therefore, it would be beneficial to investigate if any correlations exist between reporting flood incidents and socioeconomic status in order to address any potential bias in the data.

#### **4 Conclusion**

This study looked at three different parameters, income level, SVI and unemployment percentages, in order to understand the socioeconomic trends in areas that were affected by the flooding caused by Hurricane Harvey. Looking at the placement of the flood indicators used for this analysis in comparison to the distribution of the three before-mentioned parameters, generally, those with a higher income, lower SVI value, and lower unemployment percentage experienced similar or larger amounts of flood incidents than those in the opposite side of the spectrum. Though the analysis may suggest that areas of higher wealth were affected at a higher frequency than areas of lower wealth, it is important to note that the resilience and reaction of each community can be drastically different depending on the socioeconomic status of its residents. This idea was addressed in a September 3, 2018 The New York Times article (Fernandez, 2018) that acknowledged the rapid reconstruction that is going on in Houston is not uniform throughout the entirety of the city, specifically in poorer areas where rebuilding has been very slow. By understanding who exactly is impacted by natural disasters such as this one, city, state, and federal officials can better utilize, allocate and distribute their resources for the benefit of their residents.

#### **Work Cited**

@NWSBrownsville. (2017). "Just astounding the amount and extent of rainfall that #Harvey2017 produced in Southeast Texas. From a NOAA/NWS research partner @UWCIMSS." *Twitter*, 31 Aug. 2017, 6:23 p.m., <https://twitter.com/NWSBrownsville/status/903428053870800897>

ATSDR (n.d.). *CDC's social vulnerability Index (SVI): What is the SVI?*. ATSDR. Retrieved from <https://svi.cdc.gov/>

Blake, E.S. & Zelinsky, D.A. (2018). National hurricane center tropical cyclone report Hurricane Harvey. *NOAA*. Retrieved from [https://www.nhc.noaa.gov/data/tcr/AL092017\\_Harvey.pdf](https://www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf).

CDC (2017). *CDC's Social Vulnerability Index (SVI) – 2014 socioeconomic SVI, county level*. Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry/ Geospatial Research, Analysis, and Services Program. Retrieved from <https://www.arcgis.com/home/item.html?id=ecb87d1b0e8e41ddb55c153e5c2d6db1>

City of Houston (2018a). *Zip codes* [Data file]. Retrieved from <https://cohgis-mycity.opendata.arcgis.com/datasets/zip-codes?geometry=-100.762%2C27.983%2C-90.424%2C31.323>

City of Houston (2018b). *311 service request dashboard for flooding and drainage requests between August 26<sup>th</sup> and August 30<sup>th</sup>* [Data file]. Retrieved from <https://www.houstontx.gov/innovation/311dashboards.html>.

David J. Phillip/Associated Press (Photograph). (2018, August). Retrieved from <https://www.nytimes.com/interactive/2018/08/30/magazine/hurricane-harvey-houston-floods-texas-emergency.html>

FEMA (2017). *National disasters Hurricane Harvey data damage assessments* [Data file]. Retrieved from <https://data.femadata.com/NationalDisasters/HurricaneHarvey/Data/DamageAssessments/>

Fernandez, M. (2018). *A year after Hurricane Harvey, Houston's poorest neighborhoods are slowest to recover*. The New York Times. Retrieved from <https://www.nytimes.com/2018/09/03/us/hurricane-harvey-houston.html>

Fry, R. & Kochhar, R. (2018). *Are you in the American middle class? Find out with our income calculator*. Pew Research Center. Retrieved from <http://www.pewresearch.org/fact-tank/2018/09/06/are-you-in-the-american-middle-class/>

Hallisey, E., Flanagan, B., Kolling, J. & Lewis, B. (n.d.). *A social vulnerability Index (SVI) from the CDC*. Center for Disease Control and Prevention Agency for Toxic Substances and Disease Registry. Retrieved from [https://svi.cdc.gov/Documents/Publications/CDC\\_ATSDR\\_SVI\\_Materials/SVI\\_Poster\\_0703\\_2014\\_FINAL.pdf](https://svi.cdc.gov/Documents/Publications/CDC_ATSDR_SVI_Materials/SVI_Poster_0703_2014_FINAL.pdf)

Homeland Infrastructure Foundation-Level Data (2018). *Emergency Help Request from Twitter* [Data sets]. Retrieved from <https://respond-harvey-geoplatform.opendata.arcgis.com/datasets/DataCapableInc::emergency-help-requests-from-twitter>

Houston-Galveston Area Council (2015). *FEMA Floodplains NFHL 2015* [Data set]. Retrieved from <https://h-gac.sharefile.com/share/view/s3afeac2a3314e4bb>

Konish, L. (2018). *US median household income climbs to new high of \$61,372*. CNBC. Retrieved from <https://www.cnbc.com/2018/09/12/median-household-income-climbs-to-new-high-of-61372.html>

Peterson, M. (2012). *What preliminary damage assessments really mean*. FEMA. Retrieved from <https://www.fema.gov/blog/2012-03-20/what-preliminary-damage-assessments-really-mean>

Turkewitz, J. and Burch, A. D.S. (2017, August 31). *Storm with 'no boundaries' took aim at rich and poor alike*. The New York Times. Retrieved from [www.nytimes.com/2017/08/31/us/hurricane-harvey-houston-homes-flooded.htm](http://www.nytimes.com/2017/08/31/us/hurricane-harvey-houston-homes-flooded.htm)

U.S. Census Bureau (2011). *Block group Texas Geodatabase 2012-2016 American Community Survey 5-year estimates* [Data file]. Retrieved from <https://www.census.gov/geo/maps-data/data/tiger-data.html>

U.S. Census Bureau (2012). *Census block groups* [Data file]. Retrieved from <http://www.cae.utexas.edu/prof/maidment/giswr2014/census/census.gdb.zip>.