

# **Analysis of the possible contribution of socioeconomic status to flood risk in Harris County, Texas**

**By: Brooke Fry**

**GEO327G**

**GIS & GPS Applications in Earth Science**

**Fall 2020**

**Dr. Mark Helper**

## **1 Background**

Harris County, Texas is prone to flooding that has the potential to cause severe damage. In April 2016, the historic ‘Tax Day’ flood occurred, dumping 12 to 16 inches of rain across the northwest portion of Harris County over a span of 12 hours. The Tax Day flood caused over \$1 billion in property damage and trapped people in their homes upwards of a week. In August 2017, flooding due to Hurricane Harvey displaced more than 30,000 people and caused \$125 billion in damage. With increased urbanization in Harris County and the Houston area and increased impervious cover associated with it, it is likely that flooding in Houston will continue to be a threat and will become even more severe over time. Also, in the face of a changing climate resulting in changes to the global hydrologic cycle, it is likely that rainfall intensity and patterns will change in the future. With Houston being the fourth largest city and one of the most diverse cities in the United States, it is important to understand how communities are and will be affected by flooding based on socioeconomic status. It is important to know whether the risk of flood is affected by a household’s income level. Households that are considered low income would likely not have the means to immediately repair damage caused by flooding or have the opportunity to miss work for days at a time.

## **2 Problem**

The goal of this study is to determine if there is an apparent link between socioeconomic status and risk of flood in Harris County, Texas.

Understanding the relationship between socioeconomic status and flood risk is essential for preparing communities for flood events and for the allocation of resources post flood events. Low income households and communities are at a disadvantage when it comes to flood event preparation and recovery and repair post flood events.

The relationship between socioeconomic status and risk of flood is investigated using the Federal Emergency Management Agency (FEMA) Flood Insurance Rate Maps (FIRM) and income data from the United States Census Bureau (USCB). FEMA flood maps are the standard maps to use in the United States to determine flood insurance rates. The data were analyzed using operations in ArcMap and Microsoft excel.

## **3 Methods**

To investigate the relationship between socioeconomic status and risk of flood, FEMA flood maps and income data from the USCB were analyzed. After data preprocessing (discussed in section 5.1), the resulting data was analyzed in ArcMap. To visualize the flood map data, a flood map was created (Figure 1). To visualize census tracts and areas depending on the median

household income, an income by census tract map was made (Figure 2). The shapefiles containing flood zone data and census income data were intersected to create a new shapefile with flood zone and income data. This shapefile was used in the creation and symbolization of a map with flood zone and census data combinations (Figure 3). Each of these steps is discussed in greater depth below.

#### **4 Data Collection**

Harris Co. shapefile:

[https://hub.arcgis.com/datasets/b0223bbaff4041d29eb79bb5a17d9f30\\_0](https://hub.arcgis.com/datasets/b0223bbaff4041d29eb79bb5a17d9f30_0)

Source: ArcGIS Hub

Type: Polyline Shapefile

Projection: WGS84

Census Income Data (2018):

<https://data.census.gov/cedsci/table?q=harris%20county%20income&g=0500000US48201.140000&tid=ACSST5Y2018.S1903&hidePreview=true>

Source: US Census Bureau

Type: .csv (converted to .xls to make processing in ArcMap easier)

Downloaded table for Harris County income data

Used Column: S1903\_C03\_001E containing median household income data (USD)

Census Tract Data:

<https://h-gac.sharefile.com/share/view/sc1857dcfa4248eea>

Source: Houston-Galveston Area Council

Type: Polygon Shapefile

Projection: NAD 1983 State Plane Texas South Central FIPS 4204 Feet

FEMA FIRM data:

<https://h-gac.sharefile.com/share/view/s3afeac2a3314e4bb>

Source: Houston-Galveston Area Council

Type: Polygon Shapefile

Projection: NAD 1983 State Plane Texas South Central FIPS 4204 Feet

## 5 Data Processing

### 5.1 Data Preprocessing

The Harris County shapefile was originally projected as WGS84. The county line shape file was reprojected to NAD 1983 State Plane Texas South Central FIPS 4204 Feet to match the other shapefiles being used.

Upon downloading USCB income data, there were a lot of unnecessary columns. Using Microsoft Excel, the extra columns were deleted. The only columns retained were the ones containing census tract GEOID data and the column containing median household income data. The GEOID values originally were preceded by text that needed to be removed. This preceding text was removed using a Python pandas dataframe and the `.replace()` function.

The FEMA flood zone shapefile and the USCB tract shapefile were each clipped using the reprojected Harris County shapefile.

The cleaned census income data was joined to the census tract shapefile based on GEOID to result in a table with data containing information on census tract and the income within that census tract.

### 5.2 Data Processing

Using the clipped FEMA flood zone data, the data was symbolized based on the `FLD_ZONE` field. The resulting map displays the various flood zones within Harris County (Figure 1). The description of the flood zone IDs is found in Table 1. For the purposes of the study, field zones A, AE, AH, and AO were combined into one zone that will be referred to as A\*. These zones were combined because they are all within the floodplain and only differ based on measurement methods.

Table 1: Description of the flood zones based on FEMA's "How to Read a Flood Insurance Rate Map Tutorial".

FLD_ZONE	Description
X	Not within the 100-year floodplain
A, AE, AH, AO	Within the 100-year floodplain
VE	Within the 100-year floodplain with additional coastal hazards

Using clipped and joined USCB tract and income data, a map displaying the income level of each tract was created (Figure 2). The map was symbolized by quantities that I specified and the ranges were changed to fit the defined income level ranges (Table 2). These income level ranges were simply chosen as easy cut-off points.

Table 2: Income level distinctions based on median household income ranges.

<b>Income Level</b>	<b>Median Household Income Range (USD)</b>
Low	< \$40,000
Medium	\$40,000 to \$125,000
High	>= \$125,000

To visualize the relationship between median household income and flood risk, a map containing both was created (Figure 3). To create this map, the floodplain shape file and census data were intersected using the Intersect tool within the Overlay section in the Analysis toolbox. The resulting shapefile contains new polygons with the same flood zone and census tract. Select by attributes was used on the intersecting shapefile to extract the areas of flood zone and income level combinations (Table 3). The areas were found using the statistics function of the Area\_sqmi column. The percentage of total area was calculated in Microsoft Excel. The ranks attached to the varying combinations of flood zone and income level should be understood as 1-3 being the lowest risk (outside of the floodplain), ranks 4-6 as being at risk (within the 100 year floodplain) and ranks 7-9 as having the highest risk. There were no low income and high income levels located within the highest risk flood zone, VE.

Table 3: Information regarding flood zone and income level combinations, the ranks attached to the combinations (used in Figure 3), and the areas and percentage of total area for each combination.

<b>FIRM Flood Zone</b>	<b>Income Level</b>	<b>Rank</b>	<b>Area (sq. mi.)</b>	<b>% of Total Area</b>
X	Low	1	3876.45996	3.70
X	Medium	2	54750.4832	52.22
X	High	3	5252.26898	5.01
A, AE, AH, AO	Low	4	2311.74286	2.20
A, AE, AH, AO	Medium	5	33438.5888	31.89
A, AE, AH, AO	High	6	4334.21696	4.13
VE	Low	7	0	0
VE	Medium	8	883.148561	0.84
VE	High	9	0	0

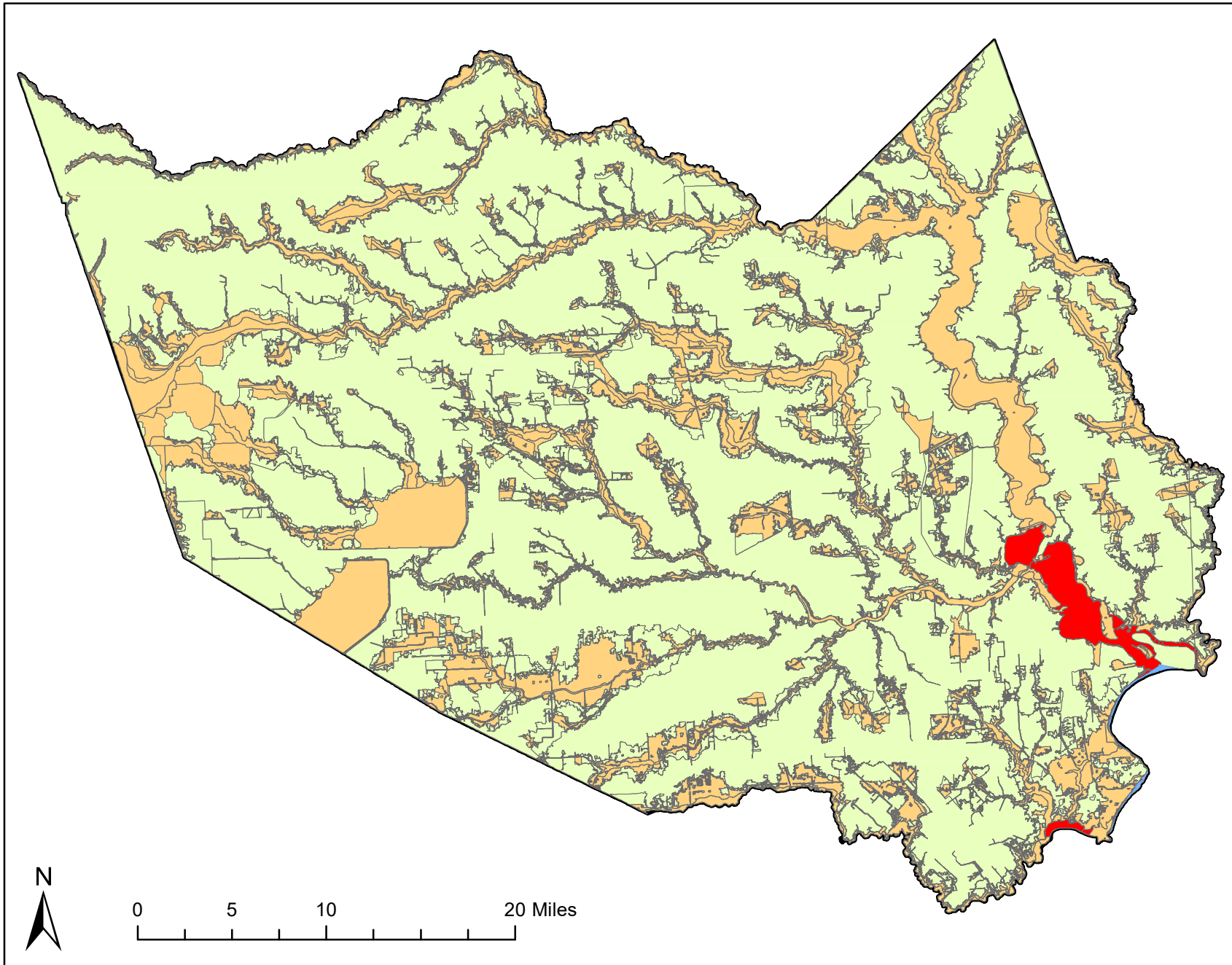
## 6 Results and Conclusion

The flood zone and income level with the highest percentage of total area is Medium income and Flood zone X, with 52.2% of the total area (Table 3). For flood zones located within the 100 year floodplain, zone A\*, the medium level income group has the highest percentage within this risk. Table 3 shows that the largest percentage of area in Harris County is occupied by middle income households. There were no low income households located in the highest risk zone, VE. According to the analysis done here and the areas in Table 3, there does not appear to be a noticeable trend between income level and flood risk.

Future studies may find relationships between socioeconomic status and flood risk by looking at other socioeconomic factors such as race. Limitations of this investigation include using tract level data, perhaps other results would be seen looking at smaller neighborhoods with smaller scale income and flood data.

# Figure 1: Flood insurance rate zones for Harris County, Texas

Brooke Fry  
12/7/2020



## Legend

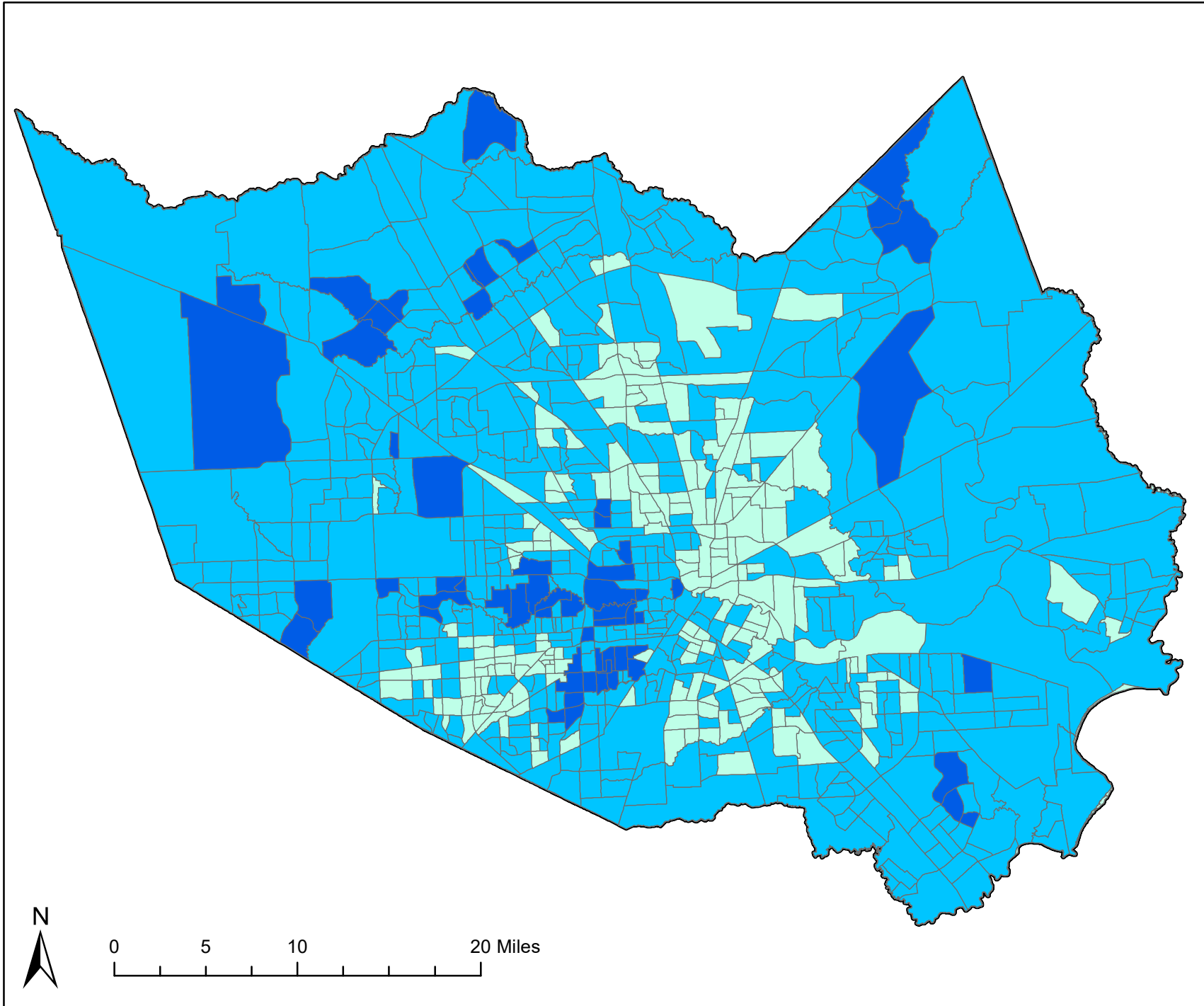
### FIRM Flood Zones

- A
- AE
- AH
- AO
- OPEN WATER
- VE
- X

Data Sources:  
FEMA's NFHL  
ArcGIS Hub  
Harris Co. Boundary

Datum: NAD 1983  
Projection: Texas  
State Plane  
Coordinate System

# Figure 2: Median Household Income by Census Tract in Harris County, TX



### Legend

#### Median Income (\$)

- < 40,000
- 40,000 - 125,000
- >= 125,000
- Census Tract

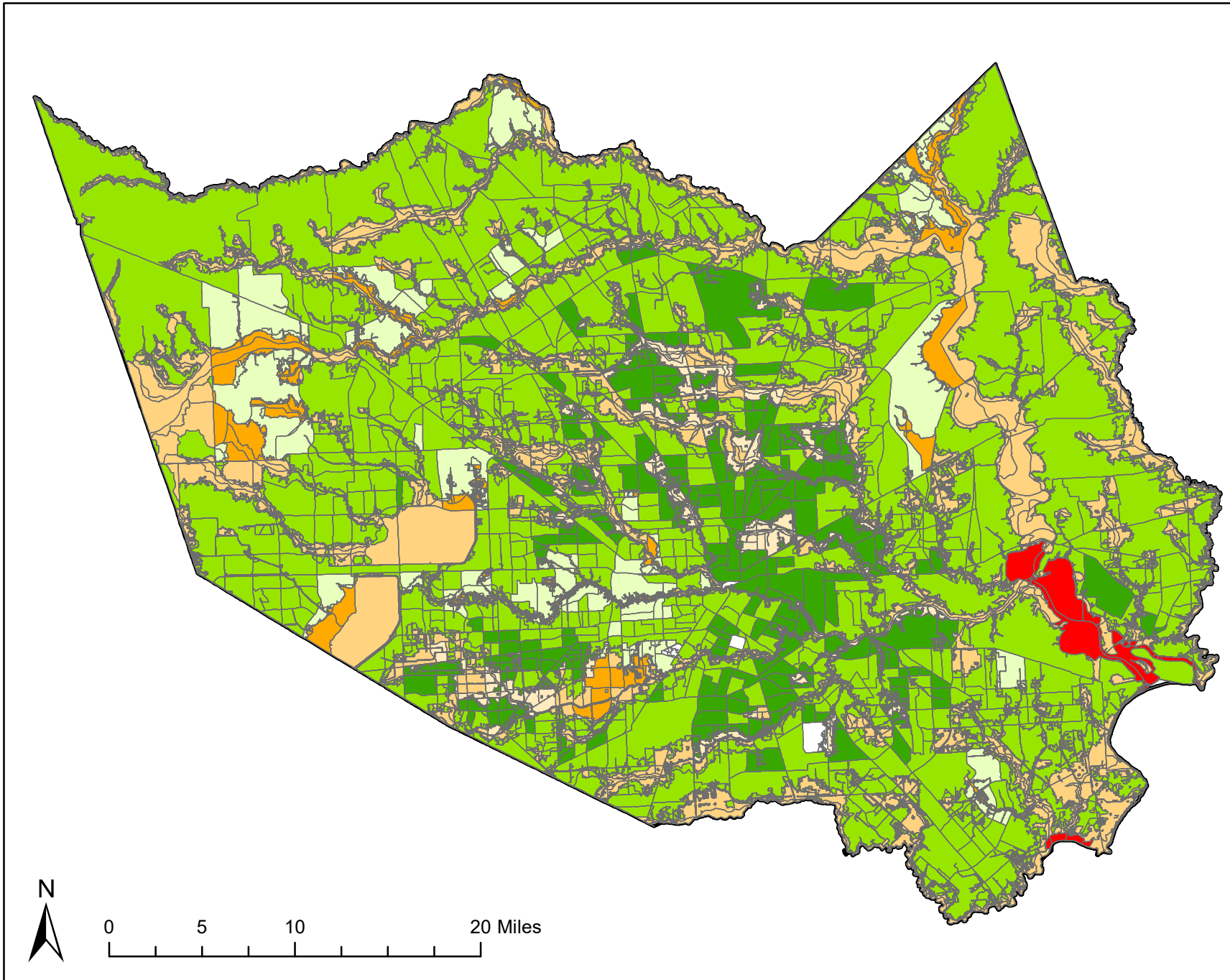
Data Sources:  
US Census  
Bureau

Datum: NAD 1983  
Projection: Texas  
State Plane  
Coordinate System



# Figure 3: Flood Risk and Household Income Level

Brooke Fry  
12/7/2020



## Legend

- No income data
- Low income, Zone X
- Medium income, Zone X
- High income, Zone X
- Low income, Zone A\*
- Medium income, Zone A\*
- High income, Zone A\*
- Medium income, Zone VE

Data Source:  
US Census Bureau  
FEMA's NFHL

Datum: NAD 1983  
Projection: Texas  
State Plane  
Coordinate System