Info W18 - Project 2 - July 2020

Data Science for Climate Justice in Texas

Enoch Luk, Frank Escalante, Eugene Wong

Datasets:

Climate Justice:
Affected by Hurricane Harvey
http://data.houstontx.gov/dataset/city-of-houston-harvey-damage-assessment-open-data
Contains 2.5k damage assessment surveys describing, by census tract:

- Group population count
- Racial demographics
- Age demographics
- District location data

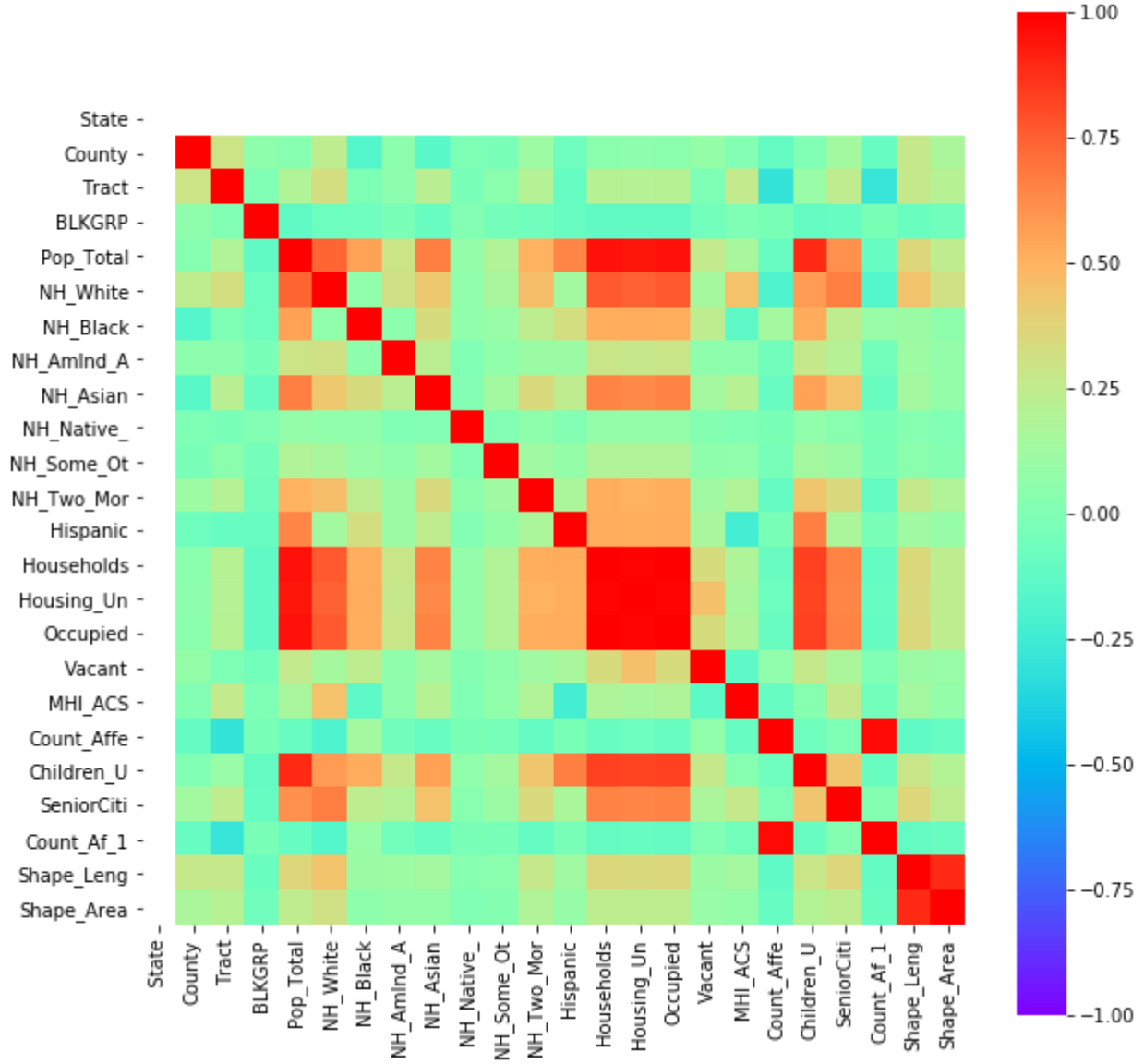
Social vulnerability by county
https://svi.cdc.gov/data-and-tools-download.html
Contains 254 rows of data per county for estimates and margin of errors for:

- Total population and population density
- Unemployment rate
- Age demographics
- General socioeconomic indicators

Other possible datasets:

Global Temperatures
Hurricane/Flooding events

Initial exploration:



```
In [1]: import pandas as pd
        import matplotlib.pyplot as plt
        import seaborn as sns
        import numpy as np
        from sklearn.metrics import mean_absolute_error
        from sklearn.linear_model import LinearRegression
        from sklearn.model_selection import train_test_split
        df = pd.read_csv (r'C:/Users/enoch/Desktop/project2/datasets_32193_41794_Harvey_BG.csv')
```

```
In [6]: corr=df.corr()
        plt.figure(figsize=(10,10))

        thing = sns.heatmap(
            corr,
            vmin=-1, vmax=1, center=0,
            cmap='rainbow',
            square=True,
        )
```

What we plan to cover in the final report:

Identify climate change trends and correlation

1. How does rising average global temperatures affect the frequency of hurricanes and flooding in Texas?
2. How can we create interactive visualization such as graphs and plots with pandas
3. How do disasters disproportionately affect vulnerable communities?
4. How do we define vulnerable communities?
5. What is/are the most prevalent of races in areas affected by the increases in frequencies of hurricanes/flooding.

Additional Possibilities

1. Applying machine learning
2. Create maps
3.