**APPLIED DATA SCIENCE**

**FOUNDATION PROJECT PROPOSAL**

**TOPIC**

*Impact of Hurricane Sandy: 311 open data analysis of pre and post*

**DATASET**

1. 311 open data

* Oct 1st 2011 - Dec 31st 2011 (before Hurricane Sandy)
* Oct 1st 2012 - Dec 31st 2012 (during Hurricane Sandy)
* Oct 1st 2013 - Dec 31st 2013 (after Hurricane Sandy)

1. Annual Census Data (2011-2013)

**BRIEF OVERVIEW**

Determine impact of Hurricane Sandy by analyzing correlation and patterns between 311 data and census data of before, during, and after the event with statistical models and tools that we have learned from this course.

**CONTENTS**

1. Visualization

* The number of complaints by zip codes (sector or precinct) and agencies
* Separate original data into different datasets according to geological location (e.g. zip code, city, or borough) to observe the impact of Hurricane Sandy by measuring the amount of complaint calls or other variables.
* Time series plot of number of complaints based on different time resolutions (e.g. by year and by seasons)
* Map the complaint onto CartoDB or other GIS viewers using shapefiles (manually using matplotlib)

1. Finding correlation

* Use/compare linear and non-linear regressive models to find correlation among variables from 311 data and census data (e.g. types and agencies from 311 data, population, income, household from census data)
* Try to define variables that we can measure to evaluate the impact of Hurricane Sandy by comparing pre and post event.

1. Test and compare the accuracy of the models we previously found

* If it is it not applicable for 2012, then why?
* Explain how we can say statistically that we should not apply model to 2012
* R square and f statistic comparison
* See if 2011 model can be applied in 2013 (see if Sandy affects the 2013)
* Pearson test
* Multivariate regression

1. Optimization

* Based on correlation that we found via measuring methods, optimize any field that we can adjust accordingly to responsive inputs.

**TEAM MEMBER**

1. Kim, Tae (thk301@nyu.edu)
2. Putro, Dimas (drp354@nyu.edu)
3. Zhuang, Yuzheng (yz2611@nyu.edu)