Problem Set #2 – due 11/14 – submit assignment to NYU Classes
Civic Analytics and Urban Intelligence
NYU - Fall 2017
Prof. Kontokosta

Energy use in buildings accounts for approximately 79% of all of New York City's energy consumption and carbon emissions. In an effort to improve its sustainability and efficiency, the City adopted an ambitious goal of reducing carbon emissions by 80% by the year 2050. The New York City Mayor's Office of Sustainability (MOS) is responsible for implementing this plan, which calls for major reductions in building energy use (The City of New York, 2016). One significant analytical challenge is to develop a robust method for peer comparison and benchmarking of buildings. This is necessary to understand relative building energy performance, given the specific characteristics of a building, and to design regulations that require buildings to meet a certain energy performance standard. Another concern revolves around environmental justice and the impact of building energy use and emissions on local communities. The question is whether low-income communities disproportionately bear the burden of inefficient buildings and possible negative effects on health.

Using the datasets posted to NYUClasses, plus Census data and other data sources as you determine, please address one of the following two questions:

Q1. How can we develop a metric for understanding energy performance and peer building comparison that accounts for the differences in building and occupancy characteristics that impact consumption?

Q2. What are the social/community impacts of building energy use and emissions? Specifically,
-What is the relationship between building energy use and public health, measured by the local asthma hospitalization rate?
-What is the relationship between neighborhood income and residential building energy use?

To do so, you will need to complete these steps:

1. Clean (including outliers) and merge the Local Law 84 data, PLUTO data, and Census data (and other datasets as needed)
2. Provide descriptive statistics for the integrated dataset of key variables in your analysis
3. Create an appropriate model to test your hypothesis
4. Develop a visualization (map) for your results, as well as plots, graphs, and model output
5. Discuss your findings in terms of how they address the respective question you set out to answer. Discuss the limitations of your approach and justify any assumptions used in your analysis.

Please use the IEEE conference proceedings template for your paper (a link has been posted on NYUClasses). The paper should be no more than 6 pages in this format.

As a reminder, please use the following structure:

- Introduction
- Literature Review (brief background of relevant studies)
- Data and Methods
- Results
- Discussion and conclusion