Assignment 6

The dataset I select is New York City air quality surveillance data provided by Department of Health and Mental Hygiene (DOHMH) on the research of different air pollutants’ effect on various diseases in different districts of NYC (UHF42) between 2005-2007 and 2009-2011. For my visualization, I chose to study the effect of PM2.5 on Asthma, which is reflected in the data with PM2.5-Attributable Asthma ED Visits, indicating the number of hospital ED visits in the given district that is caused by PM2.5.

**Vis1:** Question: What is the difference of PM2.5 effects on Asthma ED Visits between adults and children in the two time periods?

Data: For this question, I need the data of PM2.5-Attributable Asthma ED Visits by child and by adult in the two research periods: 2005-2007 and 2009-2011.

Visual: I chose to use bar charts to display because it can show each category clearly, and made three different views so the user could switch between them. The first view is clustered bar chart which shows the numbers of ED Visits in each category by each district. The second view is stacked bar chart which shows the total number of ED Visits of each districts as well as the breakdown (what period of time and what population). The third view is percentage stacked bar chart which shows which population in which period of time are most susceptible to the effect of PM2.5 on Asthma.

**Vis2:** Question: How do the PM2.5 effects on Adult Asthma ED Visits change between two periods?

Data: For this question, I need the data of PM2.5-Attributable Asthma ED Visits by adult in the two research periods: 2005-2007 and 2009-2011.

Visual: I chose to use gradient line charts to display to show

1) the absolute scale of each district’s Adult PM2.5-Attributable Asthma ED Visits, so the user can easily see which district has the most significant impact by PM2.5

2) the change of each district’s visits number, so the user can see the change of effects in different districts in between the two time period

**Vis3:** Question: What’s the geographic distribution of the effects of PM2.5 on Adult Asthma ED Visits?

Data: For this question, I need to transform the original data which contains PM2.5-Attributable Asthma ED Visits of two populations in two periods into one data value. I chose to take the average of the four values as an indicator of the overall effect of PM2.5 in the district. I also obtained the geojson file for UHF42 districts that contains the shape file for the 42 districts of NYC.

Visual: I chose to use geographic map to display the attribute and set the color scale from light green to red in the same way how Air Quality data is usually scaled in color. From the visualization we can see that the green area which are less affected by PM2.5 are the Williamsburg Area in central Brooklyn as well as the Upper East Area in Manhattan, which has better greening, while the red area which are more affected by PM2.5 are in Queens and Staten Island where residential populations are dense and lack of greening. (Some district are not shown because of the mismatch of district names between the data file and the json file.)