Te Du
PLUTO Data Challenge
2/1/2020

1) The live interactive visualization can be found here. You can adjust the slider at the bottom of the map to select an assessed total value, then the map will be updated to include only BBL has valuation above the selected value. You can also choose the criteria (number of residential units, number of buildings, and total lot area) being displayed. I also have a set of static maps generated for the exercise which can be found here. I also want to point out that for the static choropleth maps I used the script written by Dr.Federica Bianco, who was the lecturer for a graduate class I took. The repository for the script can be found on her github.
2) BBL
3) Record number 5888. There is no building on this lot.
4) The frequent land use type is One & Two Family Buildings
5) There are 143 unique building classes in the dataset
6) The first thing I googled for the challenge is to look for the shapefiles for the . I was also trying to find out which feature will be unique identifiers which will allow the geometries to be joined with.
7) For the data exploration, I mainly used Python with Jupyter Notebook. For the visualization part, I used the Bokeh server to work with Python to create the interactive map and later used Heroku service to deploy map to the open web.
8) I learned that the dataset is very rich with extensive information. Even with the little time I spent diving into the dataset, I realize the dataset comes from very heterogeneous sources and it can take a lot of efforts to ensure the dataset is consistent internally. The first exercise really let me to rethink about visualization strategy. When the data points are sparse and small compared to the entire geography, interactive mapping features become extremely important. Features such as zooming and panning around the map can make the map much more legible and useful. It gives me more ideas to implement for such visualizations.
9) I did not get stuck at any point except spent some extra time on figuring out a solution to the visualization piece. I also probably spent more time to annotated the notebook than writing the codes to find the solution.