PUI2015 Extra Credit Project
Spatial Homogeneity of NYC Neighborhoods
<Yuan Lai, https://github.com/ylurban/PUI2015_ylai, yl3250>

Problem Description: socio-spatial inequality in NYC is a long time discussion regarding city planning, economic development and social justice. This project aims to cross check spatial and social homogeneity to better understand socio-spatial characters of different neighborhoods in NYC.

Data: The major datasets include PLUTO data on building information and U.S. Census data on demographic data.

Analysis: By data cleaning, filtering, and aggregation, each neighborhood will be assign value based on it social-spatial homogeneity. The study focuses on the features' correlation and spatial autocorrelation of neighborhoods.

References: There are previous studies focusing on social-spatial pattern correlation in Paris and Israel by using data analysis and GIS visualization. However, few study focused on New York City, especially through PLUTO data.

Deliverable: The final output will be a report with statistical conclusions and data visualization showing the degree of social-spatial homogeneity in New York City. This information would be valuable for both public agencies and private sectors to better evaluate neighborhood characters and real estate development, and better informed risk management and capital planning. Especially, this study will provide a metric to balance property investment and social good for the City.

Progress: So far I have calculated the building, lot, and land use variety score by zip code and normalized them for future calculation. I also calculated social diversity by zip code by using the percentage of foreign-born residents. Later I will visualize the data on NYC map.

	Num Floors	Bldg Front	Bldg Depth	Irr Lot Code	Total Score
ZipCode					
10307	0.000000	0.000072	0.000029	0.544602	0.032117
11207	0.045713	0.000086	0.000158	0.397846	0.026106
11208	0.024037	0.000139	0.000147	0.524757	0.032370
11416	0.004880	0.000004	0.000031	0.000000	0.000000
	0.005233	0.000132	0.000103	1.000000	0.059801

Table 1: Data frame with building diversity score.

	zip	pop
0	10001	0.210245
1	10002	0.424290
2	10003	0.196511
3	10004	0.266158
4	10005	0.273472

Table 2: Data frame with population diversity score.