Web Mapping Design:

Spatial and Temporal Analysis of Community Garden in New York City

Ziqin Wang

April 28, 2017

Web-Mapping Design:
Spatial and Temporal Analysis of Community Garden

Abstract

Community gardens in the city have multiple roles, including health-promoting, economic, environmental, and cultural. These roles are particularly important in terms of urban issues and social concerns, such as differential access to recreation, parks, fresh product in poorer communities. The garden has the ability to improve neighborhood quality. This project will examine the spatial relationship between the presence of the community garden and neighborhood characteristics in Manhattan area, New York City. Finally, based on the spatial variables performance, a web-based mapping is built as a reference for garden selection in other parts of the New York City.

Introduction

During the financial crisis of the 1970s, many parts of the New York City suffered and there were plenty of vacant and abandoned lots, especially in Manhattan neighborhoods such as the Lower East Side, Hell's Kitchen, and East Harlem. A nonprofit environmental group started to plant seeds, fertilize, water in the vacant lots. This movement not only changed people's attitude towards vacant lots but soon became a program that fostered neighborhood participation.

When realizing the benefit of outsourcing the maintenance of the vacant lot to willingness community groups. New York City initiated the GreenThumb program to provide assistance and coordination. Any community groups that are more than 10 people can identify a vacant lot and secure a licensing agreement with GreenThumb. Whether through fresh food plots or flower gardens, residents transformed the unattractive or sometimes unsafe spaces into activated open spaces, providing recreation space in especially underserved areas.

Today there are more than 600 gardens in the city while 8000 remained lots may still need change, especially to those neighborhood that lack of recreation areas. The community gardens may have a positive role on impacting the physical and social makeup of a neighborhood. The project aims to use GIS analysis tool and regression model to identify the spatial characteristic between gardens' presence and their

neighborhoods, and act as a reference for planner to locate the public area in the neighborhood.

Hypotheses

This paper aims to analyze the relationship between community gardens and neighborhood characteristics. Specially, the main research questions are: what are the neighborhood characteristic of community gardens in New York City? The hypothesis here is that the presence of community gardens in urban neighborhoods in New York City would more likely to fail in areas with worse socioeconomic characteristics, like low income levels, low housing values. However, it is possible that these community gardens were found in areas with limited subway access. This project would link the neighborhood characteristics to the community garden and also describe and evaluate the relationship.

Data

Data Gathering:

The data used in this project was taken from a various of sources: US Census, New York Open Data Portal, General OASIS Data, Tableau Online, and many other databases. The full list of data type and descriptions are provided in table1. Background literature review on New York City community gardens was provided through the website of GreenThumb.

Feature Name	Alternate Name
PLUTO Dataset	All Land Use Type
Subway Entrance	Each Entrance of Subway
Boundary	Zip Code Boundary
Income	Income by Block In 2016
Rent	Rent by Block In 2016
Green Policy Zone	New York Vege Plan Zone
General OASIS Maps	The Richest Source of Community Map Of NYC

Table 1

Data Preparation:

The main data format is in shapefiles and CSVs. The analysis of the data is in the scope of census block level so all the data is classified to this level. Depending on the content

of the data and the transformations between different datasets, I described how and why the data was cleaned and classified for the purpose of the project.

US Census Data

The US Census data came from census block data in 2015. The census block data all had similar attribute tables where only some of the census tract numbers had changed due to the reorganization of the block tract. Based on the block number and block name, spatial join was performed to geocode the csv data to shapefile.

New York Open Data Portal and OASIS Open Data

Most of the shapefiles and CSVs in this project were downloaded through these two platforms. All the files have already geocoded and the attribute table is clear with name reference. To be useful in the project, some of the datasets were cleaned and assembled and the changes are listed below:

The New York land use dataset came in the form of a very large CSV file which contains 24 variables where each variable was clearly divided into different categories. I filtered the dataset to only keep the Manhattan area as training dataset and Brooklyn as test dataset. Additionally, I took out vacant lot that were authority by the government as these lots can transfer to community gardens which later was spatial joined to census block level. To get the vacant lot rate in every census block, I calculated all the lots in every census block using spatial join and calculated the vacant lot rate by dividing lot number to vacant lot number. The remaining variables like building type, dwelling type were cleaned as the dataset was too large.

The existing community gardens in Manhattan area was generated in a point shapefile from OASIS dataset. In this project, I am interested in using the number of community garden in one census block as the dependent variable. I achieved this by using spatial join in ArcGIS.

Tableau Dataset

The online tableau dataset has massive information where rent and income by census block in New York City is well visualized, cleaned and stored in CSV format. The dataset is manually added tract number in Excel to perform geocoding in ArcGIS. As rent burden is one primary variable in the project, I also created an empty column in the attribute table and calculated the value by dividing income to rent.

Web Presentation

The final step of this project required an interactive website where CartoDB hosts all the data. The data format had converted from shapefiles/CSVs to Geojson which would be helpful as it appeals visually and let user to navigate the places where they interest..

Methodology

Scope

The scope of the project is in both spatial and temporal features. The project will examine the differences between census blocks with focus on the economic, environmental, and physical conditions. In addition, the study area would be mainly in the Manhattan area.

Exploratory Analysis

We have hypothesized that community garden would likely to have spatial pattern with neighborhood characteristic and especially related to social and economic factors. Before analyzing the relationship between and neighborhood, we began by mapping the distribution of the gardens and vacant lots to see if there is a spatial pattern in general without moving into the statistic model. We first dig into the visualization of the gardens distribution in 2016.

Community Garden Distribution

figure 1

The points in figure 1 above highlight the distribution of the community gardens. It is clearly that the gardens are more likely to occur in the lower southeast side and the upper west area in Harlem neighborhood. In the upper west sides, there are a few gardens in different neighborhood that more likely a random but not cluster pattern.

In order to get a more clearly explanation of gardens' existence in the Manhattan area in terms of whether vacant lot being a crucial part in the existing of garden. I overlaid the garden layer with the vacant lot layer, as depicted below in figure 2.

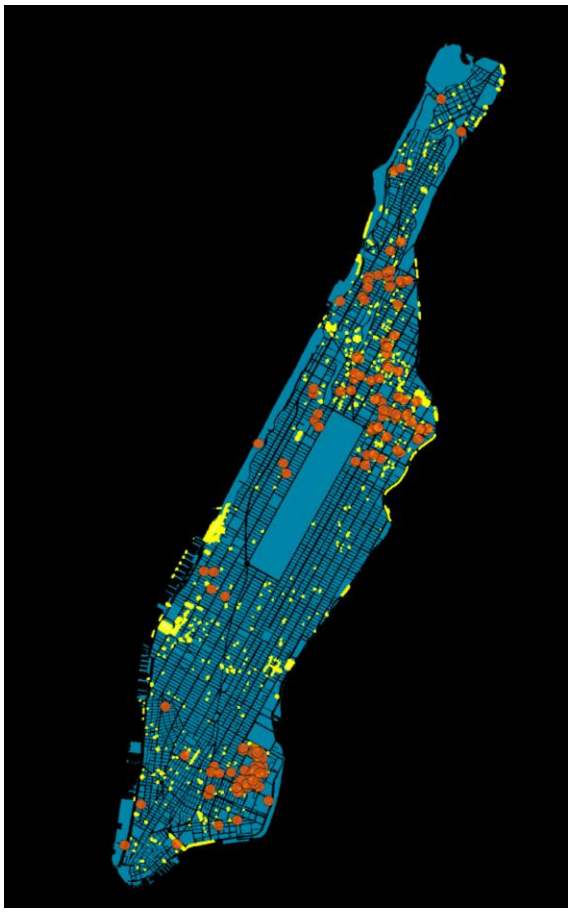

figure 2

The map of vacant lot with community garden in the Manhattan area shows that the high concentration of gardens in the lower east side does not accompany with massive vacant lots. While the upper side has a higher concentration of gardens and vacant lots.

Network Analysis

One of the hypothesis of the spatial pattern of community garden is that the inadequate subway access might impact the transit statue within neighborhood. Community gardens' arising in the city would be a part of a revival of nearby open space, food access, and have the potential to alleviate the issues of inadequate transit access. Instead of generating geodesic buffer to create service areas, the network approach allows us to produce road analysis with real travel pattern. The 5 and 10-minutes subway access map is depicted below in figure 3.

figure 3

Network analysis are displayed here and weighted by walking distance. We see an inadequate metro access in the lower east area where community gardens are dense. While on the upper sides, most of the community gardens are in the 10 minutes walking distance and some of them are within the 5-minute zone which is quite different from the lower side.

Rent and Income

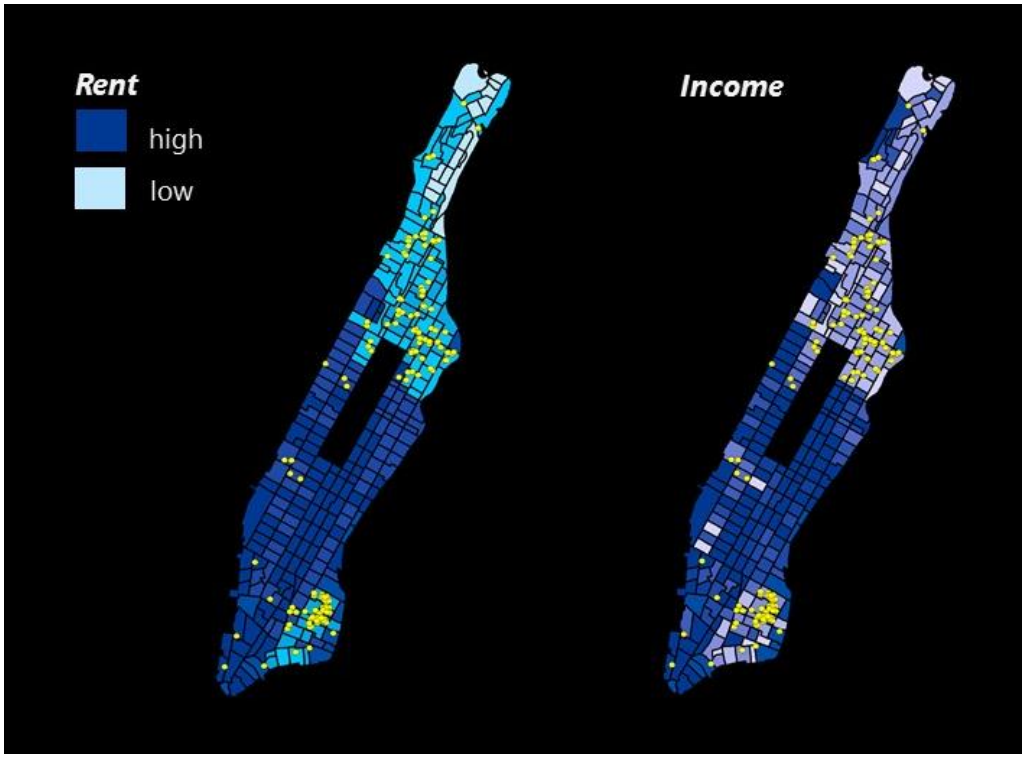

figure 4

Based on the census block, the map illustrates the output of income and rent distribution. Here, we are most interested in the areas where community garden are dense, which is characterized by the yellow dot. It is clearly that the blocks where rent and income are relatively low have more community garden than other blocks.

Vacant lot rate and rent burden

Once we had the resulting rent, income, vacant lot, lot data from every census block, we were able to relate the information to build vacant lot rate and rent burden map by creating resulting column with calculation, as depicted in figure 5.

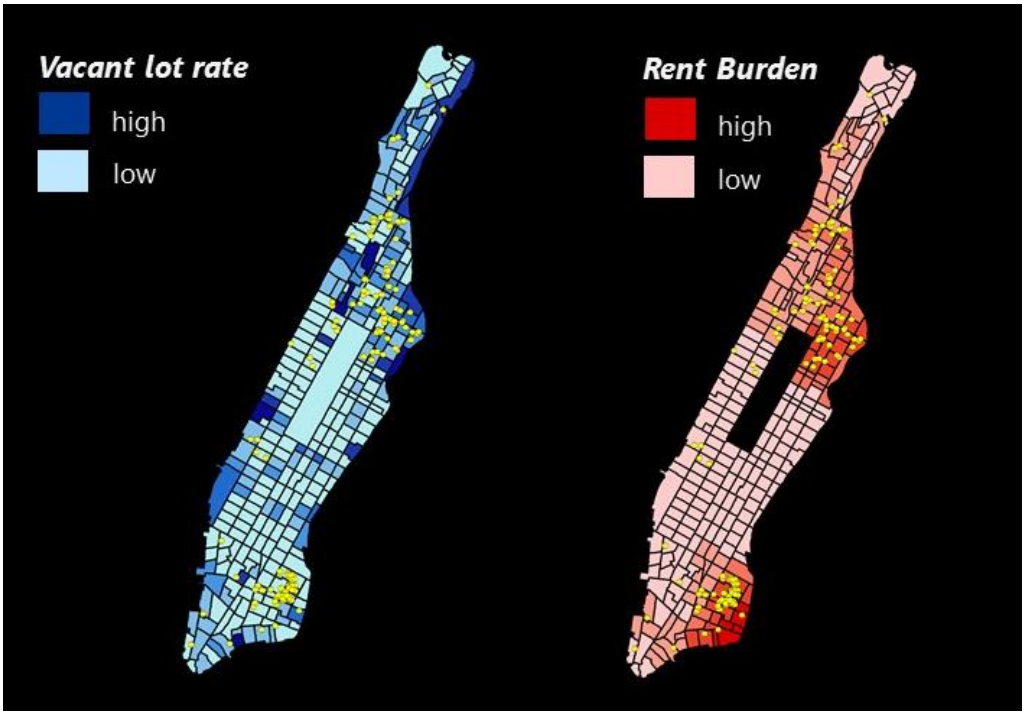

figure 5

The vacant lot rate in the map did not show great correlation with community gardens, as the lower east side has a low vacant lot rate with dense garden while the upper areas appeared a random vacant lot rate result. The rent burden refers to the proportion of rent spend in income, the map clearly showed that community gardens showed up more often in block that has a high rent burden, in other words, these neighborhoods spent much higher proportion of their income in paying the rent. Considering these area as show in figure 4 also had relatively less income and lower rent, the living standard in these neighborhoods can be quite tough, and community gardens might be their way to gain access to recreation, food access and increase social capital.

Land use and Green zone

The last part of the map show the land use type and green zone area. To be more specific, I classified the land use type to residential area and none-residential area to show a clear spatial pattern of the community gardens. Green zone policy refers to the

policy that residential areas in certain district can use their lots to grow vegetable or other fresh.

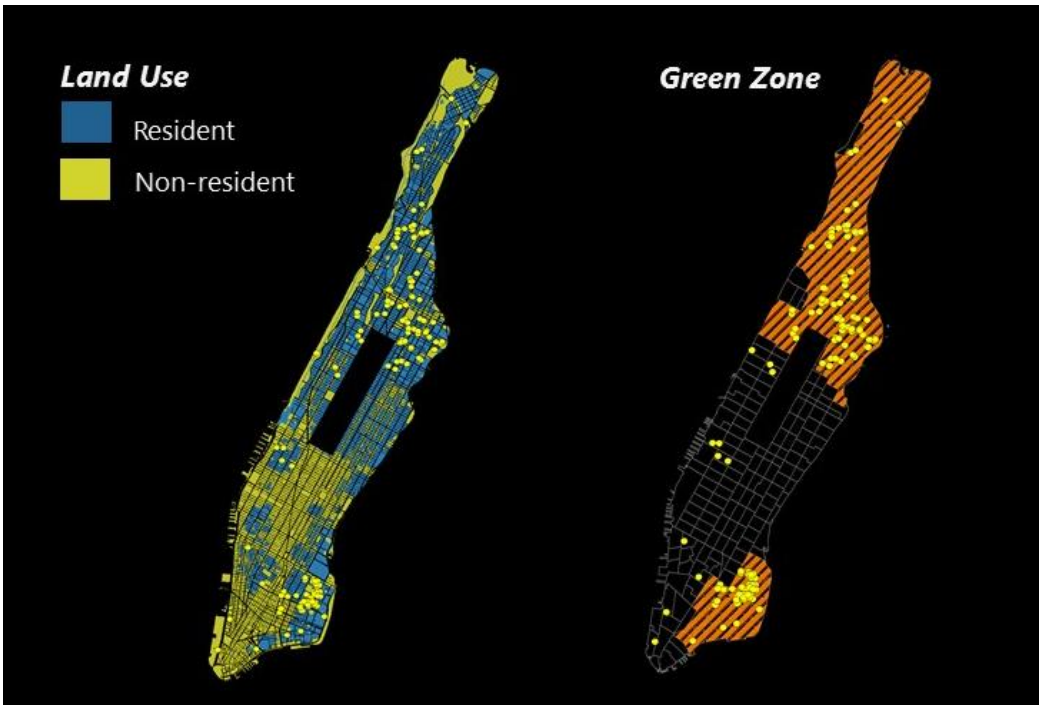

figure 6

Based on the map, it is clearly that the community gardens show up in areas that are most residential blocks and the policy plays a great role in the existing of the gardens. The majority of the gardens are located in areas that allow to plan vegetables.

It is useful to take a look at these social, economic maps to gain overall awareness about the distribution of the community gardens and the factors that might relate to those occurrences in terms of GIS data. Our hope is that the variable would be the same significant in the regression model.

Formulating A Preliminary Regression Model

After performing the preliminary analysis on the community garden data, I used the variable present in the map to construct a regression model. To start with, I performed a series of tedious data management steps that first merge all variables into census block group. Predictors such as the average rent in 2016, average income in 2016, rent burden in 2016, whether in or out of green policy zone, population change number in 2016, whether in the 5 minutes walking distance zone, whether in the 10 minutes walking distance zone, whether the average above land use is residential. This step

tried to include as much variables as possible and test whether they are statistically significant when running the model.

As which regression model to choose, it depends on the distribution of the dependent variable, and more so the residuals. The dependent variable is the number of garden in a block. If the distribution of the number of gardens look more or less continuous (i.e., there's a lot of variability and the numbers range say, from 0 to 100) then it is comfortable using OLS model. If the possible values are 0,1,2,3,4,5,6 then Poisson regression would be more appropriate.

When checking with the data, we discovered that that out of 281 tracts, 63 have gardens in their tracts and 27 of them have one garden, 16 has three, 12 has four and some random numbers from four to eight which suggest the dependent variable is not continuous. We then used passion regression in the project.

After running cross-tabulation to test whether there is serious multicollinearity between predictors, we narrow down to 6 predictors which called rent burden, fresh policy zone, vacant lot rate, population change number, income and rent. Below is the summary output in R. All the predictors are statistically significant after running stepwise regression.

```
Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -23.74673 1611.63441  -0.015   0.9882
RENT_INCOM    0.02399    0.01020   2.351   < 2e-16 ***
FRESH         0.59400    0.58973   1.007   0.0023 **
VACANT_RAT    0.02470    0.01672   1.477   0.0022 **
POPchange    -1.32101    1.48077  -0.892  1.80e-10 ***
sub_5         0.43162    0.36729   1.175   0.2399
sub_10       -0.43579    0.37086  -1.175   0.2400
Income2016    1.25387    0.41140   3.048   0.0023 **
Rent2016     -1.19318    0.39146  -3.048  1.26e-09 ***
land_use     22.49289 1611.63413   0.014   0.9889
```

figure 7

Conditional on the variables in our model, for every rent burden increase one percent, the expected number of community garden will increase by 0.02399. For every population drop one percent, the expected number of community garden will increase by 1.32101. and for every rent drop one dollar, the expected number of community garden will increase by 1.19318. It is extremely hard to evaluate a Poisson regression

model's accuracy as the R^2 is hard to tell. And that is the part we are not getting into right now.

Web Mapping Interface

There are two basic visualization functions here which are finding the existing garden and assigning the future garden. All the visualization and interaction are built with JavaScript and SQL. The method of finding garden relies on the geographic algorithm within certain areas where people draw the polygon. The map would show the existing gardens in the Manhattan area. For the second part, with the preliminary model we built above and all the same data format we got in other part of New York City's boroughs, we can estimate the potential garden numbers in the block level.

For locating the future community garden, right now the interaction map assigns the gardens in block level. As different blocks have different number of vacant lots and different potential gardens, the transfer rate (predict/lots-number) of every block is different. The model we build is at a block level, which means that every vacant lot in the same block has the same possibility to become a garden though different blocks have different transfer rate. For example, the transfer rate of Williamsburg is different than Greenpoint in Brooklyn, but every vacant lot in Williamsburg shares the same transfer rate.

Under this hypothesis, we introduce random selection in choosing the vacant lot in the lot level, as listed below. That is to say if we want one hundred gardens in Brooklyn where they have one thousand vacant lots right now, we calculate every lot's transfer rate which would be same in the same block level while different in different block. With the assigned 10% assigned rate, we compare the value with the random math result which would be a random number between 0 to 1. If the calculated number is larger than random number, it transfers successful to the garden. Since we introduce random selection in here, the map would follow the standard distribution which means that the expected garden number is 10, it could be less or more.

*Math.random ~ (number assigned/predict number sum) * block transfer rate*

Figure 8

Figure 9

Conclusion

Areas with community gardens show different spatial pattern than other parts of the city. There are more residents living in poverty. There are more population decline rate and less rent value in blocks with community gardens. In light of these findings, the project advocates for a restructuring of community garden to alter existing community garden policy to better support the low income and inadequate access to metro, fresh food neighborhood. The project used ArcGIS to process spatial data, R to build a Poisson regression model and JavaScript to build a web application, which cover almost all the skill I learned in the past one year. For a full view of the project and the interactive map, the project is hosted on Github (https://github.com/Ziqinwang/assigngarden).

Discussion

While the web mapping look good and runs smoothly, the model we built may need to investigate more and gain a better understanding of spatial temporal analyses like how to check the accurate of the Poisson regression. In working with the data, nine predictors are used in the model and there might be much more useful data that need to explore. Also, the data is quite mess when we were analyzing in ArcGIS. By creating a database might contribute to order the data tidy and clean and update much easier when new data are imported in the future.

Data source

Building footprint: http://bit.ly/2gj5p1m
Hydrography: http://bit.ly/1GXHvSp
Open Space: http://bit.ly/1xp4KSu
Parking Lots: http://bit.ly/1rZDCof
Landmarks: http://bit.ly/1ClycWj
Subway entrance: https://data.cityofnewyork.us/Transportation/Subway-Entrances/drex-xx56
Boundary: https://data.cityofnewyork.us/Business/Zip-Code-Boundaries/i8iw-xf4u
General OASIS maps**:** http://www.oasisnyc.net/pages/data.aspx

References

GreenThumb: http://www.greenthumbnyc.org/get_involved.html

How You Can Turn New York City's Vacant Lots into Community Gardens: https://www.dnainfo.com/new-york/20140924/astoria/how-you-can-turn-new-york-citys-vacant-lots-into-community-gardens

Chris Whong: http://chriswhong.com/local/this-ugly-map-got-a-lot-of-attention-today/

Armstrong, Donna. "A survey of community gardens in upstate New York: Implications for health promotion and community development." Health & place 6.4 (2000): 319-327.

Saldivar-Tanaka, Laura, and Marianne E. Krasny. "Culturing community development, neighborhood open space, and civic agriculture: The case of Latino community gardens in New York City." Agriculture and human values 21.4 (2004): 399-412.

Shinew, Kimberly J., Troy D. Glover, and Diana C. Parry. "Leisure spaces as potential sites for interracial interaction: Community gardens in urban areas." Journal of leisure research 36.3 (2004): 336.

Wang, Haoluan, Feng Qiu, and Brent Swallow. "Can community gardens and farmers' markets relieve food desert problems? A study of Edmonton, Canada." Applied Geography 55 (2014): 127-137.

Eizenberg, Efrat. From the ground up: Community gardens in New York City and the politics of spatial transformation. Routledge, 2016.

Angotti, Tom. "Urban agriculture: long-term strategy or impossible dream?: Lessons from prospect farm in Brooklyn, New York." public health 129.4 (2015): 336-341.

Plunz, Richard. A history of housing in New York City. Columbia University Press, 2016.

Owen, David. "Green Manhattan: Everywhere Should be More like New York." The City Reader (2015): 414.

McPhearson, Timon, Peleg Kremer, and Zoé A. Hamstead. "Mapping ecosystem services in New York City: Applying a social–ecological approach in urban vacant land." Ecosystem Services 5 (2013): 11-26.