Higher income means higher green?

<Jiaxu ZHOU, Ann16253470, jz2308>

Abstract:
Through the regressions of green space size and median household income in Manhattan's community, we may conclude that high income community means high green community in Manhattan is not the truth.

Introduction:
The inequality always exists in everywhere. Basically, the economy inequality, the social inequality and the environmental inequality are the 3 main factors which we meet in city life. It is obvious that the 3 inequality have strong connections among each other. In this project, the relationship between economy inequality and environmental inequality will be discussed. 2 indexes are selected to represented the 2 factors. The Median household income per community reflect economy inequality and the green space size and ratio reflect environmental inequality.
People with higher income has prior right to choose where to live. On the hand, they prefer places with better environment, which means more green space. On the other hand, they may living to choose the center of cities to easily access to other resources but means higher density at the same time.
This project will discuss the situation in Manhattan NYC.

Data:
1.The Median household Income of the American Community Surveys.
The American Community Survey (ACS) is an ongoing statistical survey by the U.S. Census Bureau. It regularly gathers information previously contained only in the long form of the decennial census, such as ancestry, educational attainment, income, language proficiency, migration, disability, employment, and housing characteristics.
The median household income is chosen as the index to reflect the economy inequality. In the problem, I select 2014 5-year estimate (the most recent 5-year dataset) and select the data at the zip code level.
Data resource: http://factfinder.census.gov/faces/nav/jsf/pages/download_center.xhtml
2.The land use information from the DCP Community Portal Project
The DCP Community Portal offers public access to tools, resources and information on a variety of topics related to New York City's land use, community planning and budget processes and demographic trends. Data and resources within the portal are updated on a periodic basis determined by information availability.
To concentrate the problem in Manhattan, I chose the land use data of communities with the format 1**. The area of different categories of land are listed.
Limits: The open space and recreation index is not equal to green space completely.
Data resource: http://www.nyc.gov/html/dcp/html/neigh_info/nhmap.shtml
3.PLUTO Data
PLUTO data is extensive land use and geographic data at the tax lot level in comma–separated values (CSV) file format. The PLUTO files contain more than seventy fields derived from data maintained by city agencies.

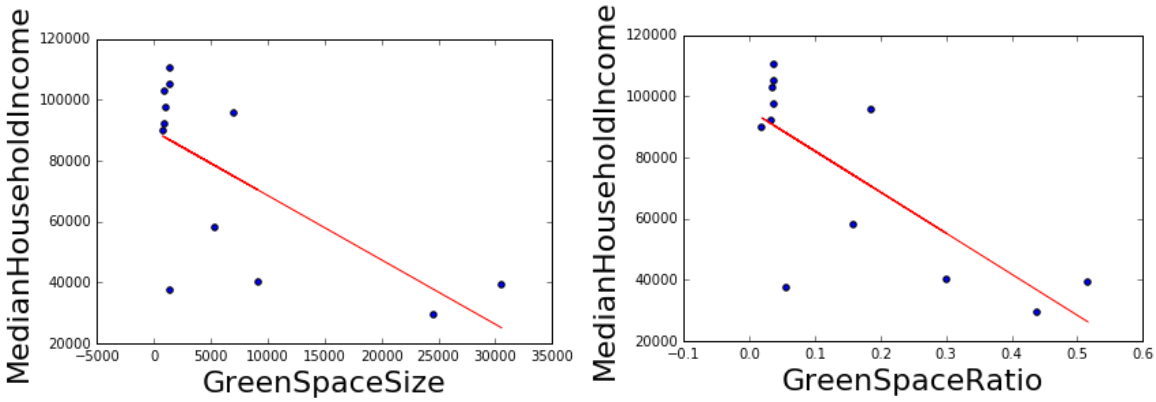
In this problem, I use the zip code and community code columns to transfer the Median income I got. After that, I am able to know the median household income of every community. Thus, the income index and the green space index are on the same level. I could try to find out the relationship between them.
Limits: City-level parks like Central-park is not in any community.
Data resource: http://www.nyc.gov/html/dcp/html/bytes/applbyte.shtml#pluto

Methodology:
1.Data Collection:
To find out the relationship among economy inequality and the environmental inequality, I chose 2 index to reflect them.
The Median Household Income was chosen to reflect the economy inequality since the income would directly influent the economy level of a family. The green space area and ratio was selected to reflect the environment inequality. The 'green space' is a great way to improve air quality, noise pollution and so on. The inequality in green space strongly related to the inequality in environment.
2.Data Cleaning:
The data cleaning process basically contains 2 steps. First is to transfer the median income per zip code to median household income per community, since I didn't get the income data grouped by community. To reach goal, I used the PLUTO data which contained almost all fields deriving ways of the city agency. The next step is to merge the income data per community in Manhattan with the land use data per community.

redundant with previous description. but at the same time you did not mention the assumptions you made to go from zip code to community level in the income: all communities in the same zip code are assumed to have same income?

3.Data Analysis:
To find out the relationship between the economy inequality, which represented by Median household income per community, and the environment inequality, which represented by the green space, I did 2 regressions.
The regressors of one in the median household income per community and the green space size. The regressors of another in the median household income per community with the ratio of green space size and total community land size.

Conclusion:



From the 2 plots above, we can clear find out that with the increase of green space size and green space ratio, the income of people in community decrease. This result is totally different from the prior expectation, which is that the people with higher income intend to live in communities with more green space. Since the people with more income may willing to spend more money on their living, they may be able to choose a community with better environment.

you should see if this gets normalized by the number of units per floor: higher income communities in Manhattan have high rises

However, the conclusion I got have several limits and bias.

1.The mean level of green space ratio in Manhattan are quite high. The environment factor may not by one of the most important influential factors when people decide where to live.

2.Central Park was not counted in the green space of any community. However, the community surround central park was great influenced by it. For example, the Upper East Side community only has 1408.7 sq of green space. But people live in there may use central park as their green space for exercise and recreation.

Future Work:

1.One more regression could be done with the regressors of median household income and green space size per person in each community.

2.The analysis of relationship among open green space and median household income could be applied in all the communities among NYC.

3.The median household income and green space size and ratio of the 5 boroughs in NYC could be listed.

4.The data of income and green space of NYC could be compared with other cities.

Bibliography:

[1]*Thomas Astell-Burt, Xiaoqi Feng, Suzanne Mavoa, Hannah M Badland, Billie Giles-Corti,* Do low-income neighbourhoods have the least green space? A cross-sectional study of Australia's most populous cities

[2]American Community Survey https://www.census.gov/programs-surveys/acs/

[3]Department of City Planning, NYC http://www.nyc.gov/html/dcp/