Introduction to Spatial Data Science

Assignment 2

Chen Liang

Data: NYC Education + Socio-Demographics (reprojected to EPSG 4269)

https://geodacenter.github.io/data-and-lab/NYC_Tract_ACS2008_12/

Variables:

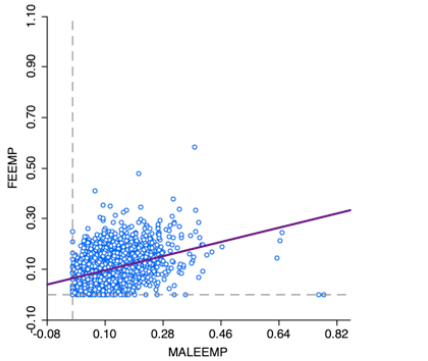
PUBASSPERC: Households with Public Assistance Income (PAI) (***withpubass***)/ Total Households (***households***)

FEEMP: Unemployed female population (***femaleuem***)/ Female Population in Labor Force (***femaleinla***)

MALEEMP: Unemployed male population (***maleunempl***)/ Male Population in Labor Force (***maleinlabo***)

Interpretations

1. Hypothetically, if unemployment rates simply represent the macroeconomic condition, then female and male unemployment rates in the same region should be similar. But in our observation, while the two rates do have a significantly positive and linear relationship, the correlation has fairly low R^2 (0.142) and a coefficient (0.311) far away from 1 as expected. That is, regions with similar female unemployment rates have a wide range of male unemployment rates, and a unit decrease in the male unemployment rate corresponds to only about 0.3 unit decrease in the female unemployment rate on average. One possible explanation is that, women are more likely to work in low-wage, hour-based jobs such as childcare workers, maids and cashiers, which tend to be less sensitive to regional economic differences.
2. As shown in the conditional maps, the percentage of households with PAI is the highest in Bronx and Brooklyn regions and is positively related to male and female unemployment rates. It is interesting that, while the Sunset Park - Bay Ridge area (red circle) and the Brownsville - East New York area (blue circle) in Brooklyn both have high concentration of households with PAI, the male unemployment rate is relatively low in the former area. I suspect that this spatial difference is related to racial demographics, given that Sunset Park has more Hispanic and Asian residents and Brownsville has mostly African American residents. It is possible that the Hispanic and Asian communities tend to have a stronger cultural expectation that males should work to raise the family. It is also possible that, since Brownsville neighborhood has held one of the highest poverty and crime rates in New York City, it is harder for either male or female labors to find job opportunities.
3. As shown below, for both genders, a unit increase in unemployment rates can lead to significantly higher increase in the percentage of households with PAI in northwestern regions than in the rest of the New York City. A possible interpretation is that households with unemployed labors--no matter male or female-- are more likely to live on public assistance income in northwestern regions than those who live elsewhere. If this interpretation is true, my hypotheses are that first, northwestern regions might provide more public assistance programs. Second, with probably higher cost of living in northwestern regions, households with unemployed labors might find it harder to survive without public assistance.
4. Based on the Averages Chart, the percentage of households with PAI and male unemployment rate in East Brooklyn are significantly higher than elsewhere. However, female unemployment rate does not exhibit significant differences. This observation further illustrates the argument that male and female unemployment rates might represent different socio-economic and regional conditions.
5. The comparison between the Bronx and the rest of New York City shows that the percentage of households with PAI is not simply determined by unemployment rates. With same male or female unemployment rates, households in the Bronx are significantly more likely to receive PAI. This indicates that other socio-economic factors such as ethnic groups, crime rates, local industries may also explain some of the variations in the regional distribution of households with PAI.



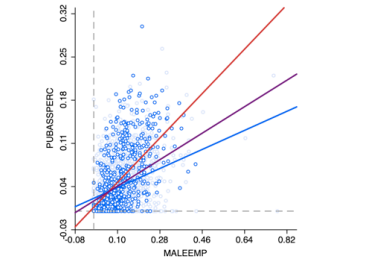
#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
2114	0.142	0.065	0.002	29.176	0	0.311	0.017	18.670	0
0	0	0	0	0	0	0	0	0	0
2114	0.142	0.065	0.002	29.176	0	0.311	0.017	18.670	0

Chow test for sel/unsel regression subsets: can't compute

For interpretation 1

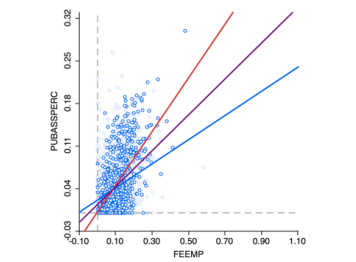
For Interpretation 2

For interpretation 3 (Map theme: Quantile Map for PUBASSPERC)



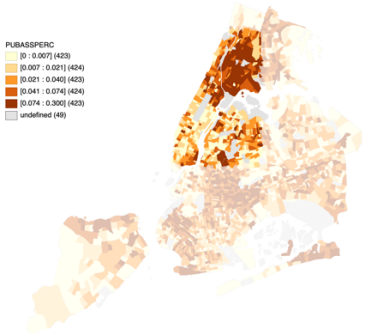
#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
2114	0.176	0.017	0.002	11.014	0	0.239	0.011	21.227	0
656	0.336	0.006	0.003	2.165	0.031	0.401	0.022	18.201	0
1458	0.116	0.021	0.002	12.391	0	0.172	0.012	13.794	0

Chow test for sel/unsel regression subsets: distrib=F(2,2162), ratio=65.7859, p-val=0

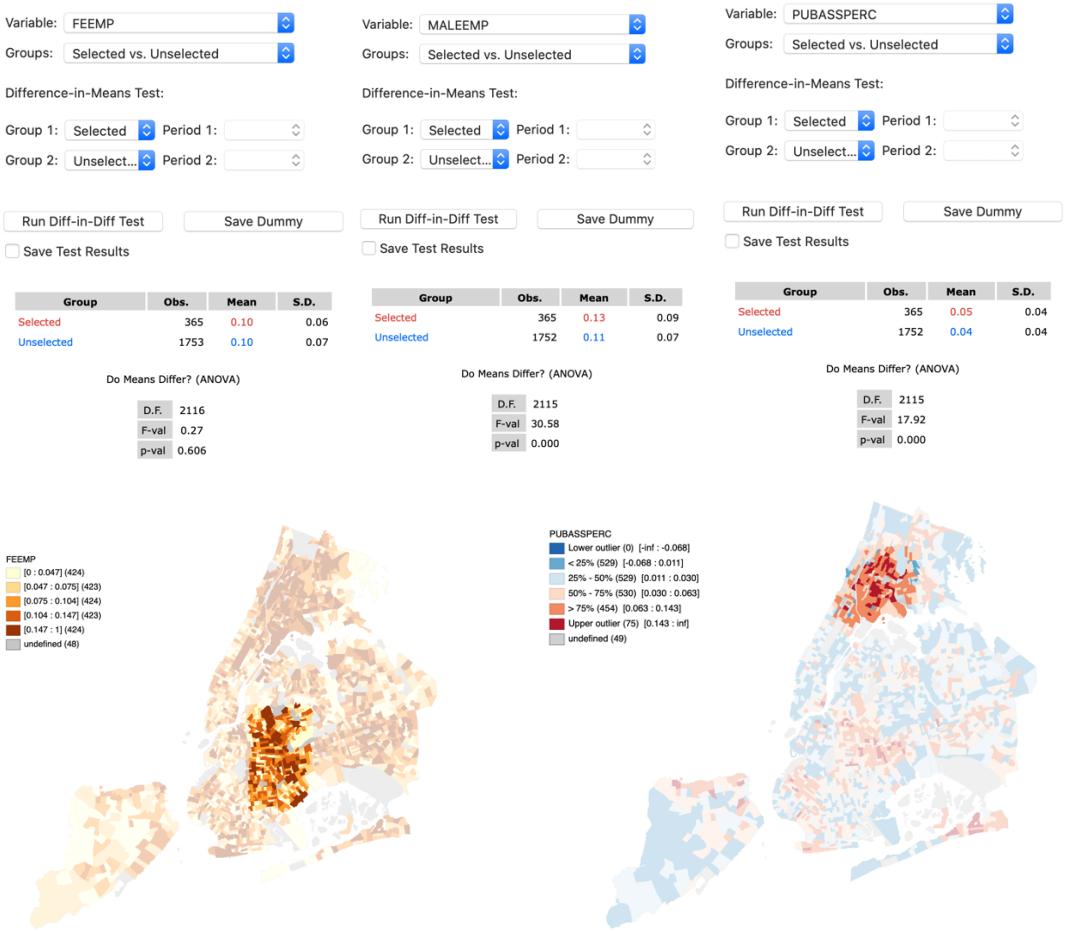


#obs	R^2	const a	std-err a	t-stat a	p-value a	slope b	std-err b	t-stat b	p-value b
2114	0.191	0.013	0.002	8.600	0	0.297	0.013	22.315	0
655	0.368	0.002	0.003	0.709	0.479	0.443	0.023	19.509	0
1459	0.095	0.021	0.002	11.509	0	0.199	0.016	12.395	0

Chow test for sel/unsel regression subsets: distrib=F(2,2162), ratio=49.1204, p-val=0



For interpretation 4 (Map Theme: Quantile Map for FEEMP, lower left)



For interpretation 5 (Map Theme: Box Map Hinge = 1.5 for PUBASSPERC, upper right)