Principles of Spatial Data Science

What drove the success of the pro-democracy camp in the Hong Kong 2019 District Council Elections?

Due: 26 April 2021

Student number: 1931393

1977 words

(All figures are in the appendix)

1 Context

In November 2019, the pro-democracy camp won a landslide with unprecedented turnout in Hong Kong's District Council elections. Despite being a low level of government dedicated to local community service, it was the only city-wide election with universal suffrage, so it acted as a referendum on the pro-democracy protests, the democratic movement, and the greater relationship between Hong Kong and China. Despite increasing violence, the pro-democracy camp won a massive landslide, controlling 17 out of 18 Districts with 57% of the vote and over 80% of the seats.

This report aims to discern driving forces in spatial and non-spatial patterns, keeping it in context with the protests. It is still too early for any literature to be published; the only academic source written after the elections is Vukovich (2020), which was just a commentary without peer review. There has been research on previous elections, including the 2016 Legco election (Tang et al., 2018), and the 2015 District Council elections (Lam, 2017). This report aims to continue the academic research, to further understand the pro-democracy movement and Hong Kong's quickly changing political scene.

The 2015 District Council Elections were also just after the Occupy Central pro-democracy protests. Lam wondered whether the relative success of the democrats in 2015 could be maintained; 2019 strongly affirmed the people's support for democracy. The analysis indicates that a wide coalition and popular consensus lead to the landslide, only resisted by cultural support for the pro-Beijing camp. It gives hope that there will be a long-lasting resistance to authoritarianism in Hong Kong, despite the clampdown using the National Security Law.

2 Methodology

For non-spatial analysis, an OLS regression will be used to test for correlation between pro-democracy vote share to age, education, and income. The purpose is to look for immediately discernible trends that might be relevant to spatial autocorrelation. The global and local Moran's I will be used for spatial autocorrelation of the pro-democracy vote share. Because of Hong Kong's high population density, there will be many interaction between small constituencies, so the weights matrix will use 7-nearest neighbours.

The 2019 District Council Constituency boundaries are from the Electoral Affairs Commission (2019a). The polygons are however expanded so every constituency boundary touches at least one other, so a shapefile of the coastline is also needed to clip the constituency boundaries. The coastline data are sourced from the Center for Spatial Sciences (2019) at the University of California, Davis.

The 2019 District Council Election results are sourced from the Electoral Affairs Commission (2019b). However, the results does not record party or political affiliation of the candidates. Even if party membership is known, it is not a good indicator because of the fragmented landscape. Some parties are exclusively local, for example Sai Kung Commons only compete in the Sai Kung District Council. The largest party by city-wide seats is the Democratic Party, who received only 12% of the vote. Therefore, the analysis will only use affiliation between the two camps rather than party membership.

Various news sources, candidates' social media pages, and public party membership records are used to corroborate evidence. Most "independents" are actually affiliated with a camp. The District Councils are local councils, not a city-wide legislature, so some independents do not explicitly or overtly associate themselves with city-wide issues. Additionally, small local parties often just has a Facebook page as their only official source. The lack of reliable media sources that is not affiliated with either camp is a major limitation in amplifying uncertainties. In practice it also risks fake candidates claiming to be affiliated with one camp but is actually running as a spoiler effect (Tong, 2015). Even international sources remain stubbornly vague on the exact count. To minimize the effect of such errors, the regression should involve city-wide percentages rather than by District or constituency, because errors in each constituency are relatively small compared to the entire city. For example, the highest vote count any candidate received was just 6824 in Heng On.

The 2016 population by-census data for DCCAs are from the Census and Statistics Department (2016). It is the latest and most up-to-date official source of Hong Kong's demographics. However it uses the 2016 DCCAs, which are different from the 2019 DCCAs. Constituency names often do not change, so the census and the election results are joined by their English name. Constituencies that were changed were excluded from regression analysis; only 31 being excluded is deemed acceptable. This problem can be avoided by only using spatial autocorrelation analysis on the 2019 elections, so the conclusion should

rely on the latter more.

3 Findings and discussion

Independent variable	Effect size	p-value	R-squared
log(65 and older)	-2.2774	0.039	0.010
log(With a university degree)	-0.2223	0.712	0.00
log(With a secondary education diploma)	2.7738	0.022	0.013
log(Total monthly median income)	3.6441	0.014	0.014
Households with 3 rooms	0.0029	0.00	0.119

Table 1: Regression results against pro-democracy vote share

As Table 1 and Figure 3 shows, the effect size and r-squared value between pro-democracy vote share and age, gender, and education is very small. For people with a university degree, it did not even meet the 0.05 p-value threshold. It indicates that none of them are particularly important and there is no leading cause for pro-democracy support. According to figure 4, for all of the independent variables, their standardized residuals are reasonably normal and homoscedastic. Households with 3 rooms had the largest R-squared value in the census but the lowest effect size in the table. However it is the most common number of rooms (about 38%) in Hong Kong, and the ordinal mean is 3.23 rooms. A small correlation with a frequent independent variable could indicate that the coalition was general and diverse, capturing broad popular support.

The null hypothesis for spatial autocorrelation is that the democratic vote share is randomly distributed spatially. The global Moran's I is 0.133, with a simulated significance of 0.001. It means there is a very small positive spatial autocorrelation in the democratic vote share. Using Figure 5, the low global spatial autocorrelation can be explained by the significant LL and HL values. For constituencies whose neighbours have low democratic support, their actual results are either even higher or even lower, but not in a consistent manner. Moreover, democrats has under-performed in a few constituencies, although some of their neighbours suggests that the surrounding area is already unfavourable for democrats, making it statistically insignificant.

While the null hypothesis can be rejected in favour of positive autocorrelation, figures 6 and 7 shows that while democratic support in most constituencies are randomly distributed, there is an unusual amount of negative autocorrelation. It indicates that most pro-democracy votes are not particularly related to their neighbours, and some pro-democracy areas are close to some pro-Beijing areas.

Hong Kong's high population density could result in support for both camps to be concentrated spatially, considering that many households live in apartments (see figure 8). People might be distributed in vertical rather than horizontal space, and spatial autocorrelation might manifest vertically rather than horizontally. However, Moran's I can only measure spatial autocorrelation by latitude and longitude, not altitude. This can also explain the census showing no strong correlation with any other variable.

Figure 9 shows the local LISA clusters. Sha Tin, Tai Wai, Ma On Shan, and Tai Po are examples of urban, mixed-residential HH clusters that confidently ousted the pro-Beijing camp. Sha Tin had a general strike rally (Regan et al., 2019) and numerous protests (Pang and Zhou, 2019), some of which turned violent. During the Siege of the Chinese University of Hong Kong near Tai Po, metro services were suspended due to protestors throwing objects onto the tracks (Pang and Smith, 2019). Despite the material obstruction to daily life, Tai Po was a strong HH area.

Tuen Mun is an example of pan-democratic unity in 2019 preventing vote splitting. The historical dominance of the pro-Beijing camp was not just due to low turnout, but also lack of coordination (Wong, 2017) (Wei, 2017). Lok Tsui was a relatively safe seat held by the Democratic Party, but the pan-democrats split their votes and narrowly lost in 2015 to Junius Ho, an independent pro-Beijing candidate. Ho then became a controversial figure during the protests as he was allegedly involved in the Yuen Long attacks (Choy, 2020). The controversy could have added fuel to the fire and energized people to vote him out. This pattern is consistent with the protests in Sha Tin and CUHK.

Tsuen Wan and Kwai Tsing (Figure 10) indicated an incumbency advantage helped over-perform. K19, S02, and S03 are safe pro-democratic seats and one of the longest continuously held seats in the District Councils. However, Tseung Kwan O (Figure 11) had a more inconsistent past. Q13, Q21, and Q18 were strong democratic seats historically, but Q15 and Q22 are swing constituencies, and the rest are pro-Beijing. It represents how much pro-Beijing incumbents in seemingly safe seats lost their seats. Remarkably,

Q24 was so strongly pro-Beijing that 4 out of 6 past elections were uncontested, but 2019 saw almost an 8-point swing to the democrat, who won by 13 points. This indicated that the pro-democracy environment was already enough to flip seats without needing incumbents.

Democrats under-performed in areas with a cultural pro-Beijing lean. San Tin and Pat Heung are rural and culturally conservative villages. The winners are backed by the powerful Heung Yee Kuk, a committee specifically for rural interests (Ng, 2015), known for its pro-Beijing lean. The villages were also allegedly involved in the nearby Yuen Long station attack targeting pro-democracy protests in July (RTHK, 2021). The Fujianese immigrants in North Point also lean pro-Beijing (Kuah et al., 2000), and during the protests there were several clashes with pro-democracy protestors (Ives et al., 2019). The Southern District is also a mountainous but a wealthy district, being the residence of the 19 richest persons in the city (Robert, 2015). While the regression analysis did not show a strong correlation with income, income has the highest effect size, which is consistent with the Southern District being a cold-spot. However, it does not have the highest mean income of all households, and because the correlation was still very small, wealthy people in the Southern District cannot be generalized to the city as a whole.

4 Summary and conclusion

An OLS regression with the census found some significant but weak correlation with pro-democracy vote share. The independent variable with the highest effect size was income, while households with 3 rooms has the highest r-squared value. The spatial autocorrelation of pro-democracy vote share showed that there were multiple factors leading to the landslide, indicating that a wide consensus of different voters were the most essential in building the coalition for the blowout. Previous literature was more cautious because of mixed performance, but the broad coalition in 2019 gives hope that there will be continued support for democracy in Hong Kong.

The analysis took a relatively aggressive approach in coercing independents into a camp because a binary affiliation is easier to analyse. However, voters may specifically value the relative independence or focus on local issues, making it inaccurate if they were treated equally to a more radical and partisan candidate. Furthermore, data analysis is not the

only approach to understanding the drivers of the results. A more qualitative approach can be used instead, by analysing social discourse in context of the protests. It is also important to note that intensity of protests do not directly correlate or cause pro-democracy support. For example, on August 31 in Prince Edward station, something similar to the Yuen Long attack happened, with comparable publicity, yet there were no significant clusters. There were many marches in Central and Admiralty near the government offices and CBD, but without significant clusters again.

As some reasoning referred to previous election results informally, a spatial autocorrelation analysis should be repeated on past District Council elections to compare the changes over time in a formal way. OLS regression with the census should also be repeated to compare differences. Further research in English is also needed to explain the cold-spots in Kowloon City and Sau Mau Ping.

A Appendix

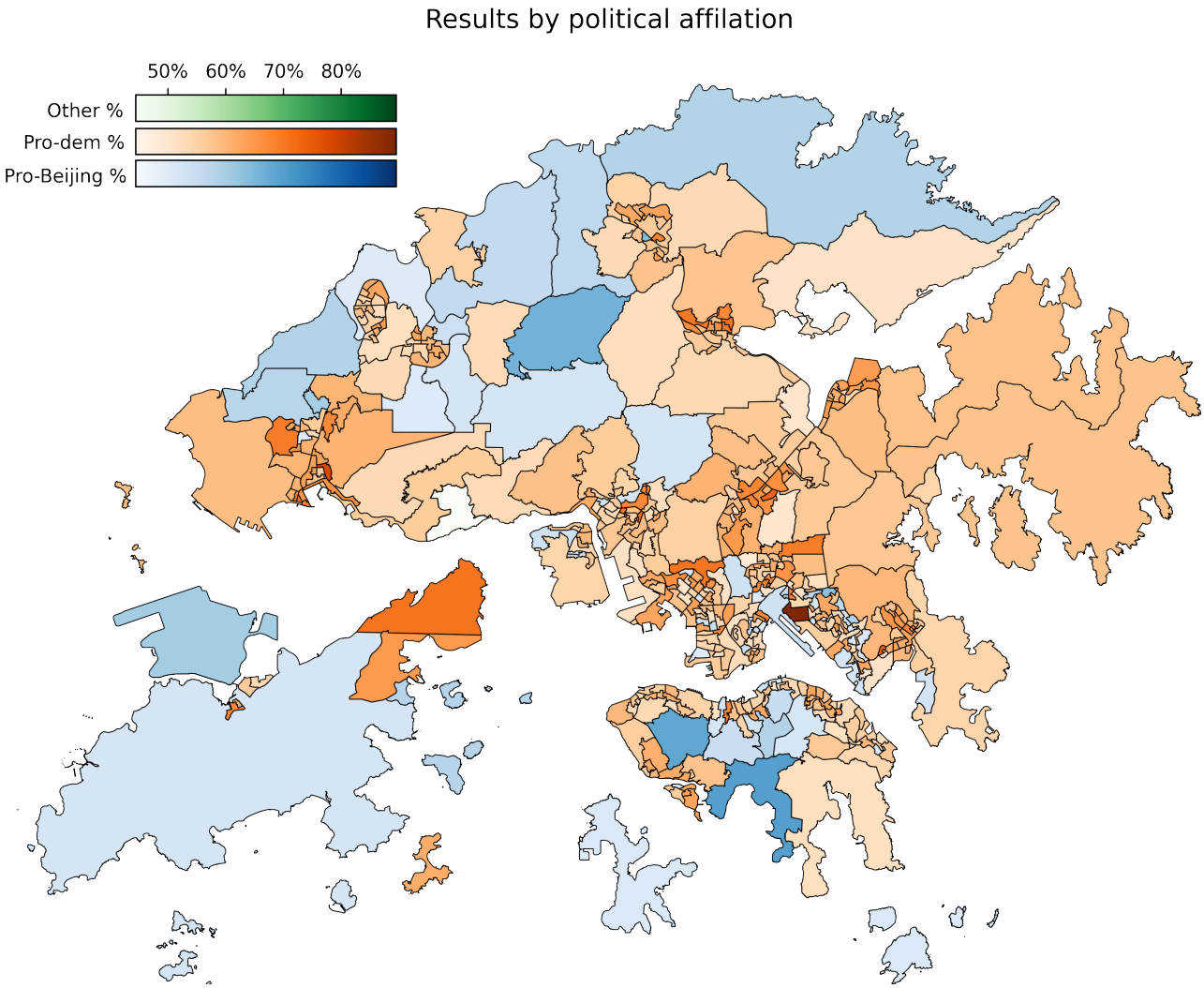

Figure 1: Results of the 2019 District Council Elections: vote share of winning political affiliation

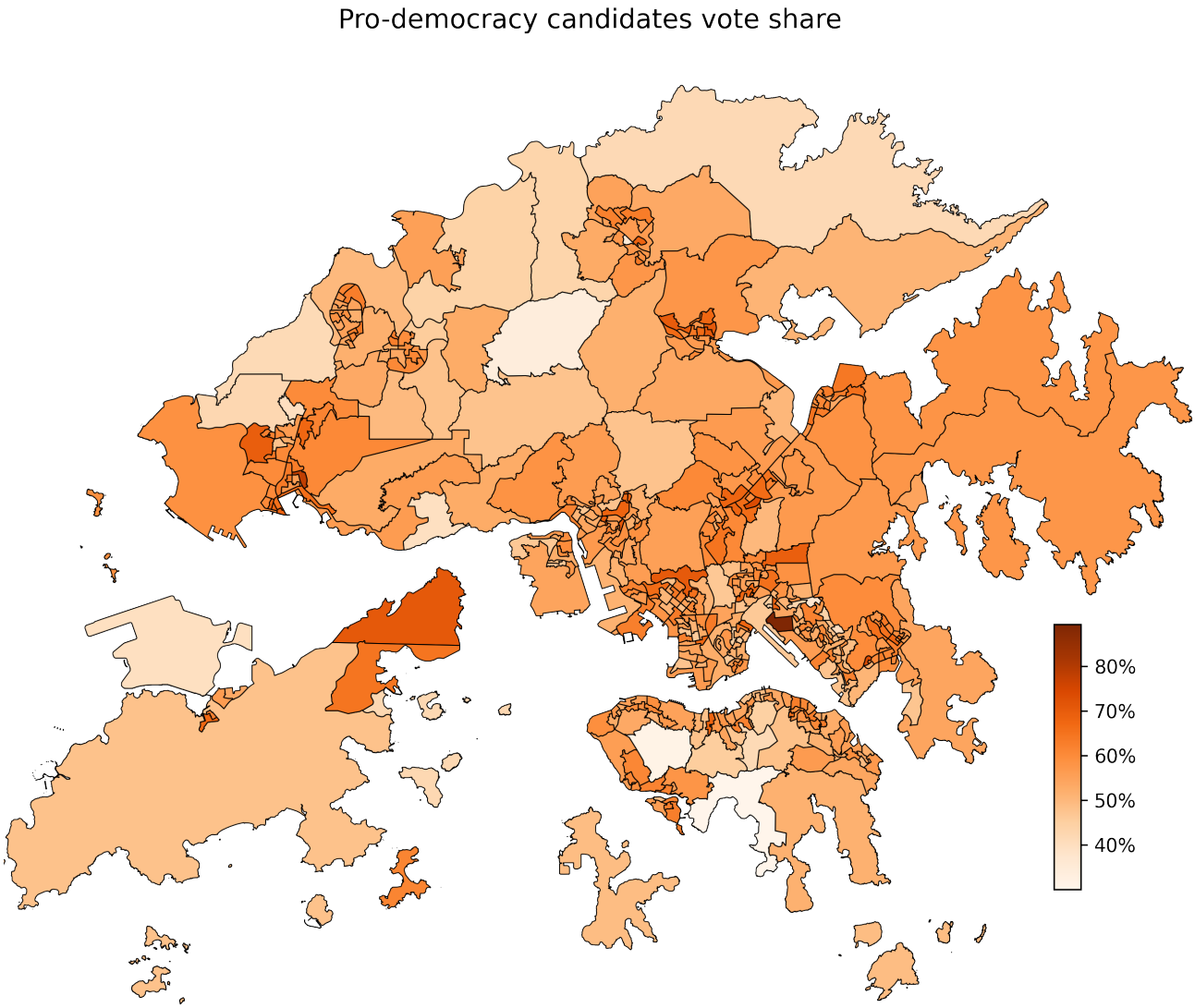

Figure 2: Percentage vote share of pro-democracy candidates

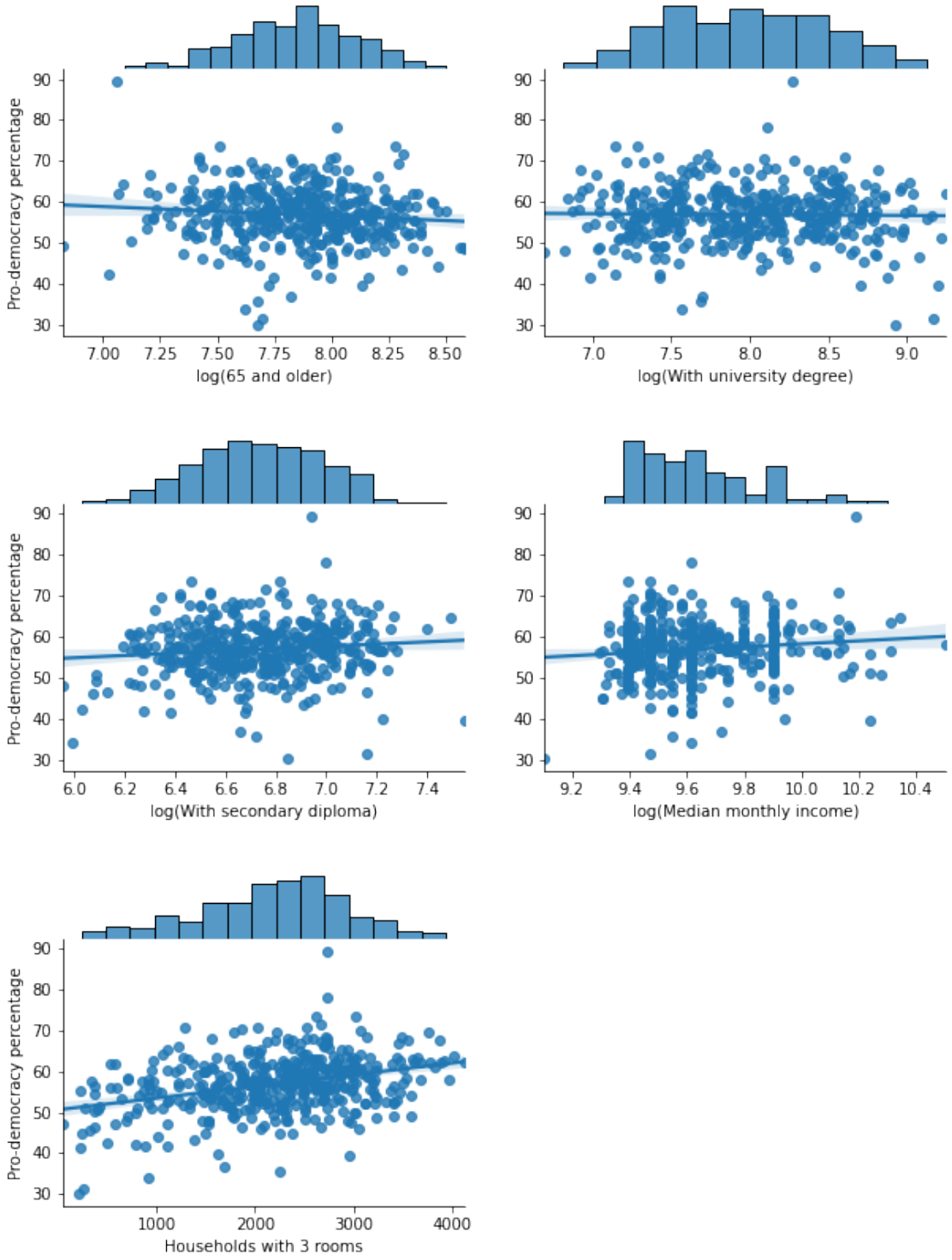

Figure 3: OLS regression of pro-democracy vote share against age, education, and income

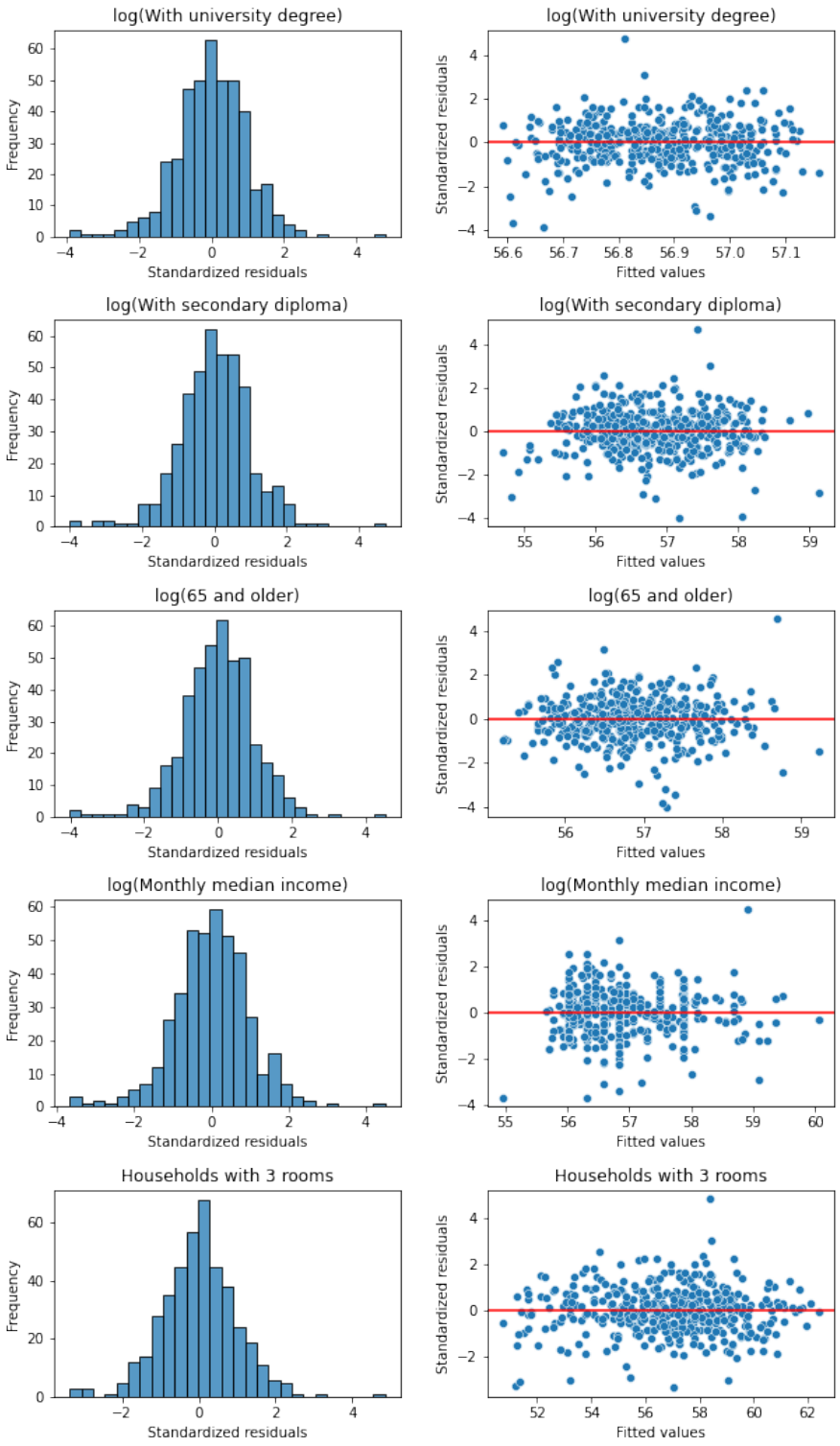

Figure 4: Histograms of standardized residuals, and scatterplot of standardized residuals against fitted values

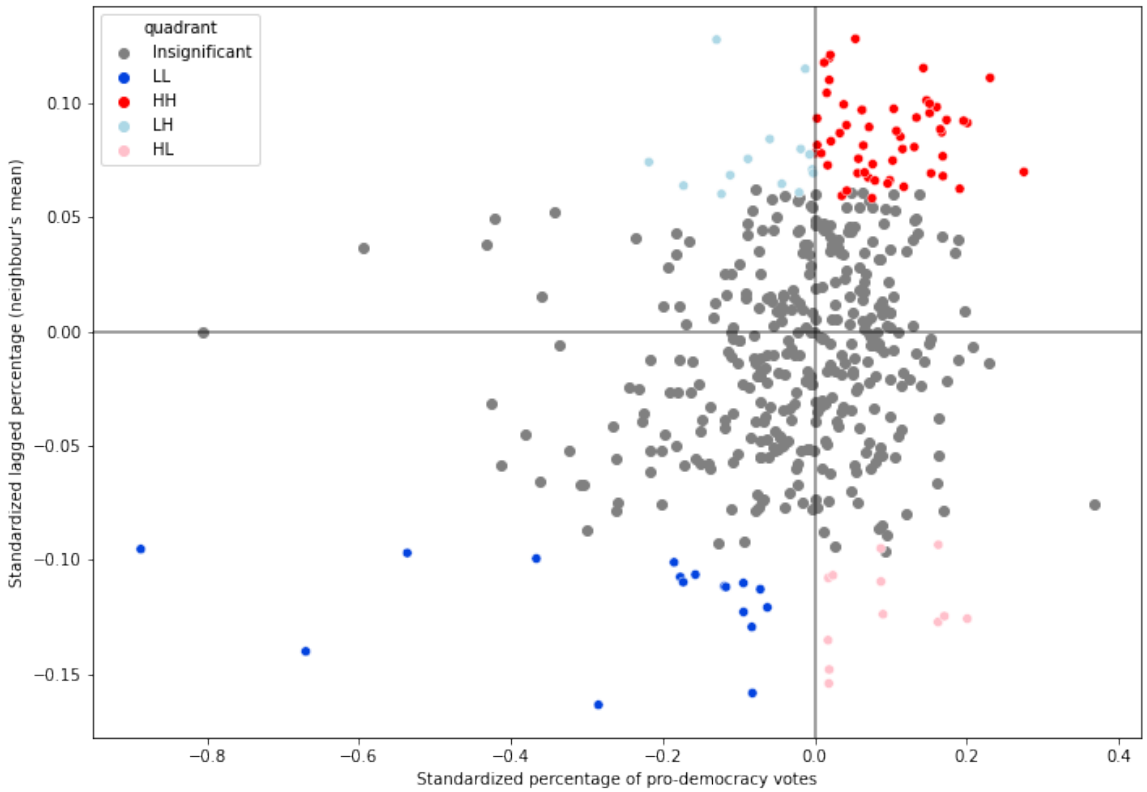

Figure 5: Moran's plot

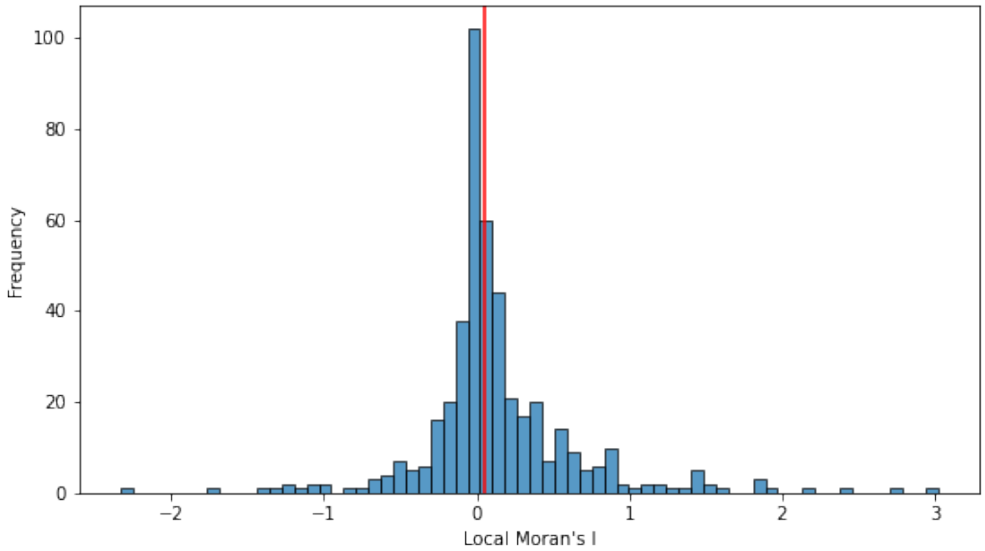

Figure 6: Histogram showing the distribution of Local Moran's I values.
The red line is the global Moran's I

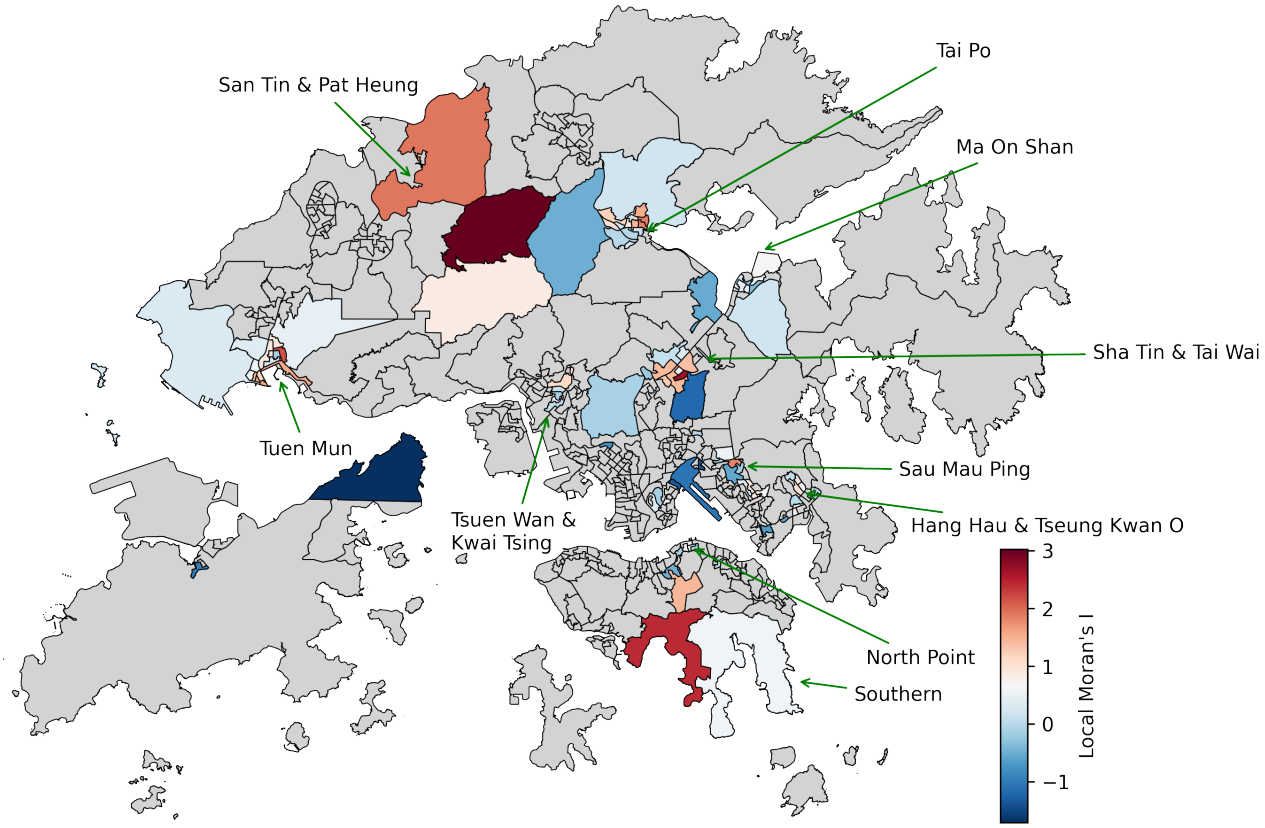

Figure 7: Map of the significant local Moran's I values

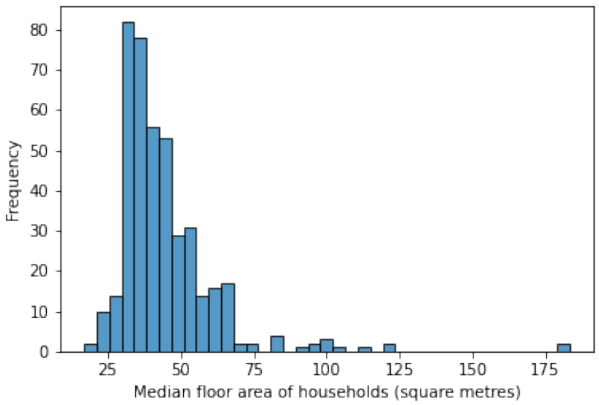

Figure 8: Histogram of median domestic household floor area

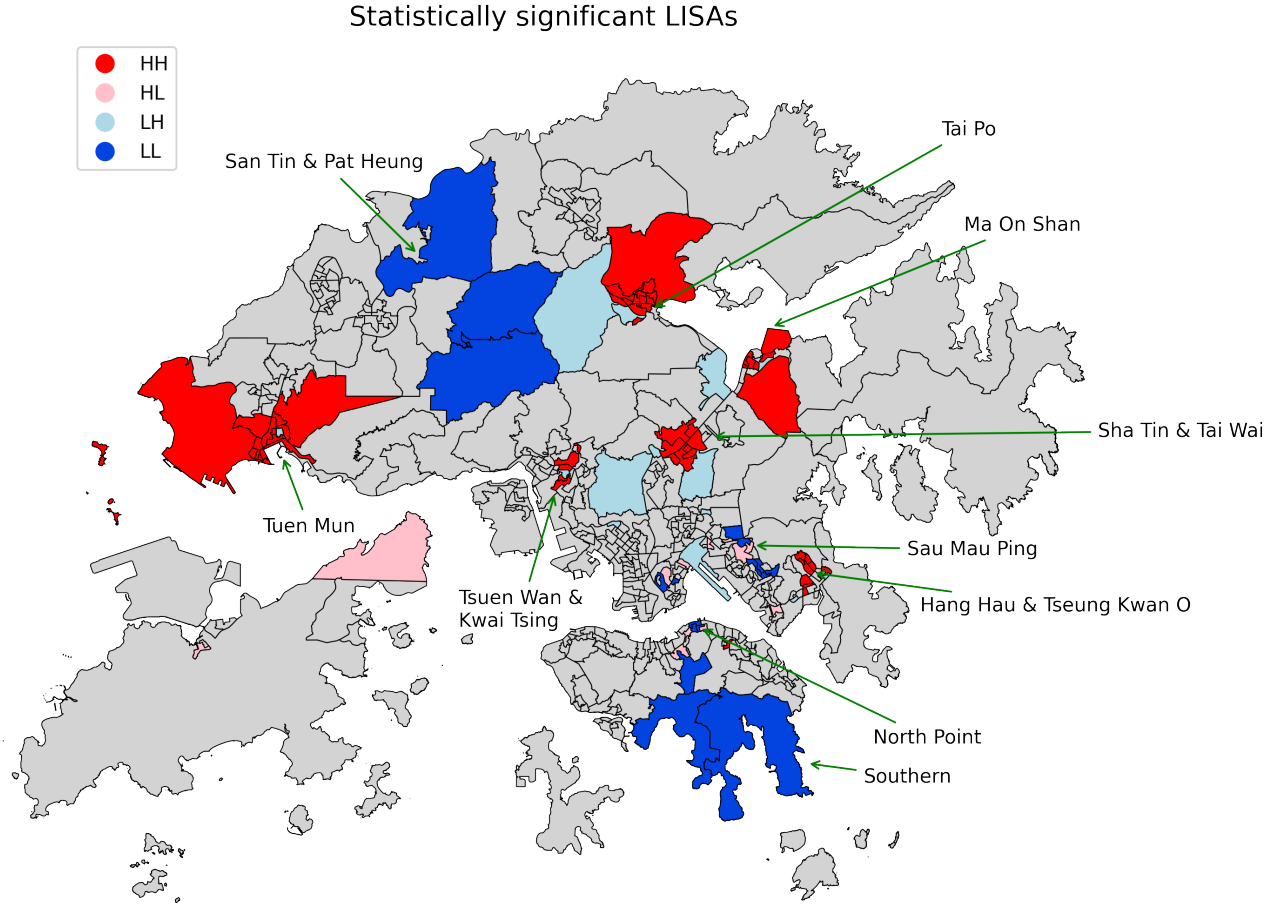

Figure 9: Map of constituencies with a significant LISA value by quadrant

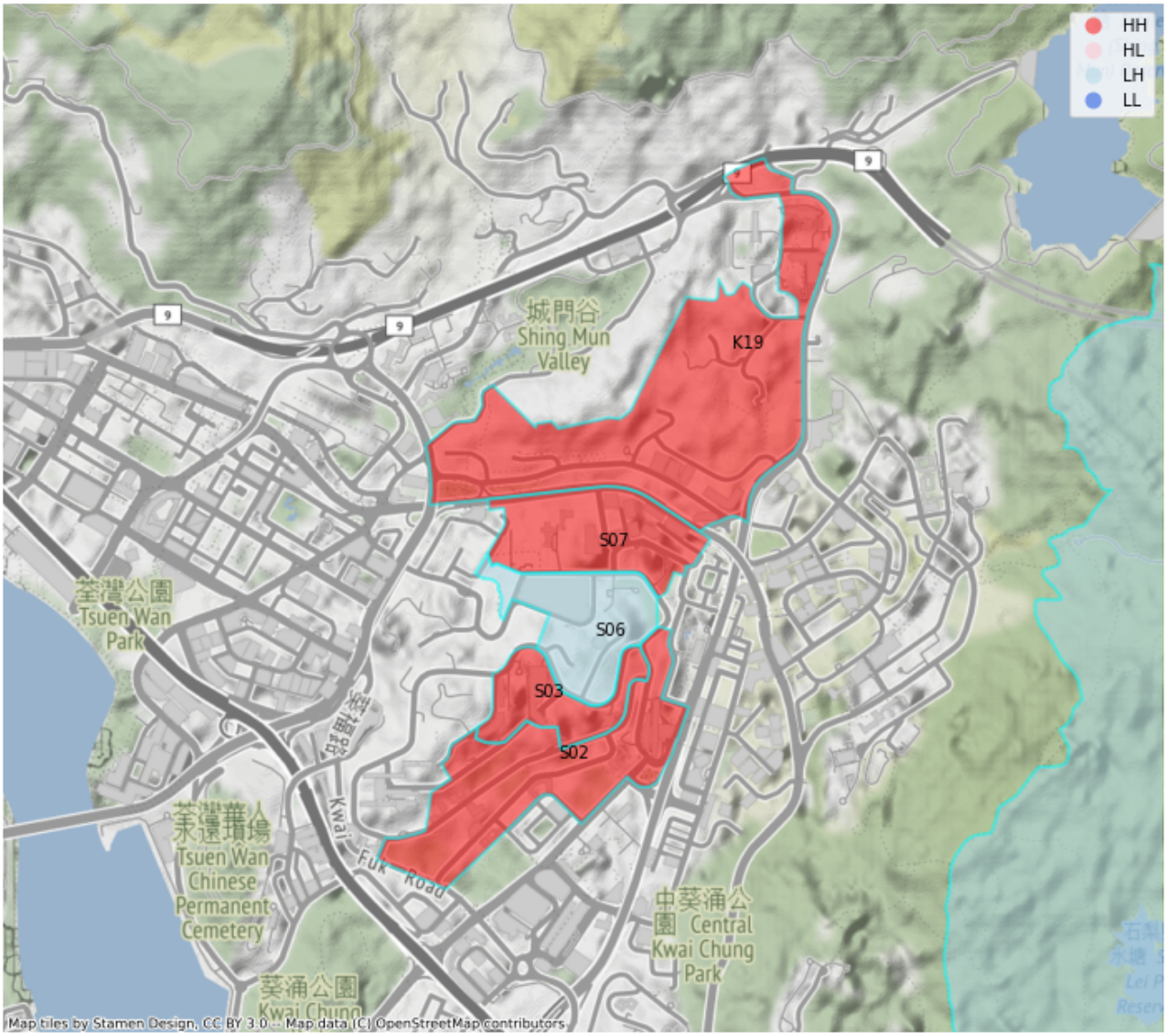

Figure 10: Map of significant LISA clusters in Tsuen Wan and Kwai Tsing

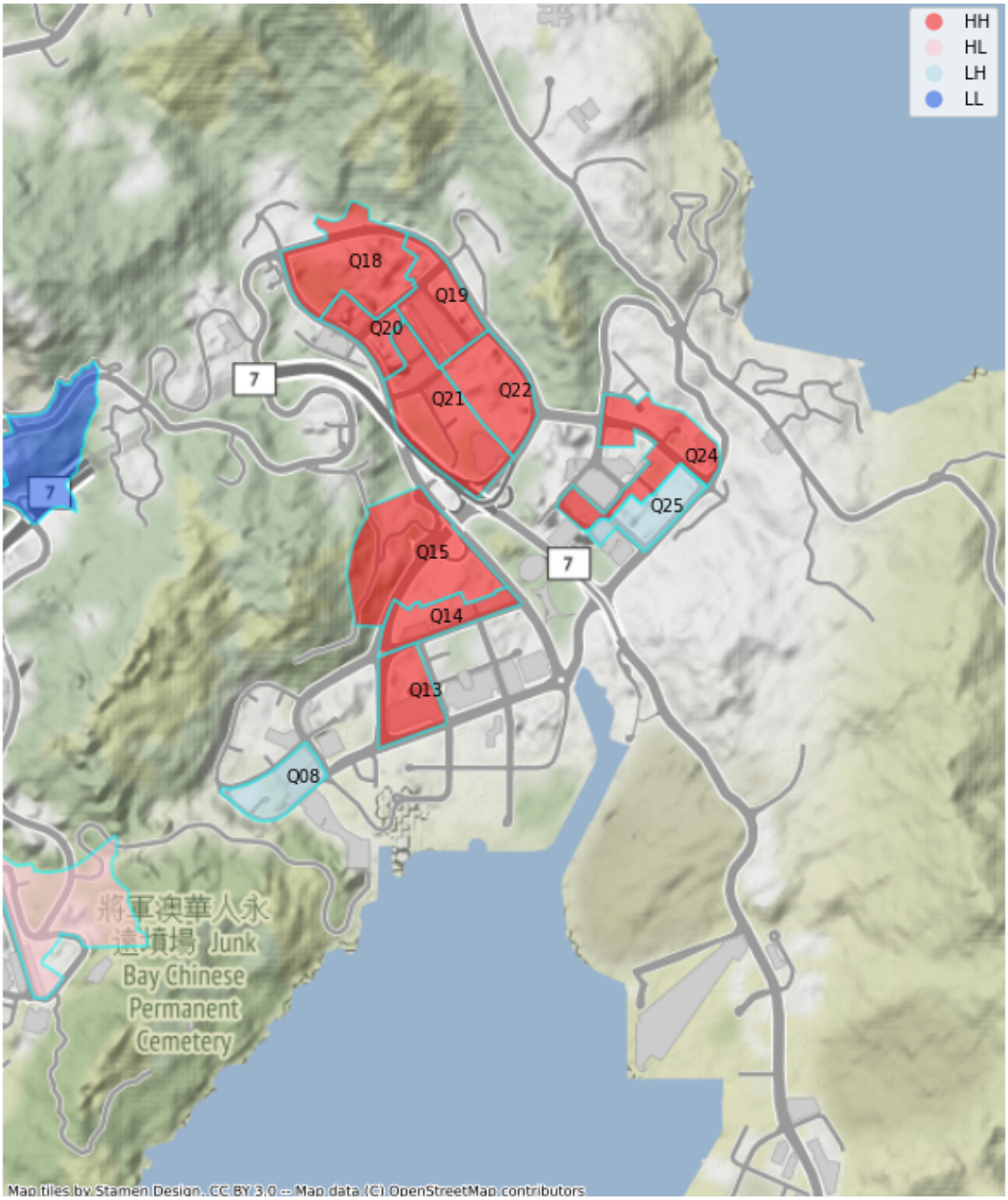

Figure 11: Map of significant LISA clusters in Tseung Kwan O and Hang Hau

References

Census and Statistics Department (2016). *2016 Population By-census Statistics (By District Council Constituency Area)*. URL: https://data.gov.hk/en-data/dataset/hk-censtatd-census_geo-2016-population-bycensus-by-dcca.

Center for Spatial Sciences (2019). *GADM data*. University of California, Davis. URL: https://gadm.org/download_country_v3.html.

Choy, B. (2020). *Hong Kong Connection: 7.21 Who Owns the Truth*. RTHK documentary. URL: https://www.youtube.com/watch?v=mrHywuxPMV0.

Electoral Affairs Commission (2019a). *2019 District Council Ordinary Election Constituency Boundaries*. URL: https://data.gov.hk/en-data/dataset/eac-eacpsi01-dcca-boundaries-2019.

— (2019b). *Election Results*. URL: https://www.elections.gov.hk/dc2019/eng/results_hk.html.

Ives, M., E. Cheng, and K. Li (2019). *Hong Kong Convulsed by Protest as Police Fire Tear Gas Into Subway*. New York Times.

Kuah, K. and S. Wong (2000). "As local and global cultural brokers: A Fujianese territorial-based association (Fujian Tongxiang Hui) in Hong Kong". *Intercultural Relations, Cultural Transformation, and Identity: The Ethnic Chinese*. Ed. by See, TA. International Conference on the Ethnic Chinese, Manila, Philippines, Nov 26-28, 1998. Int Soc Study Chinese Overseas, 226–235.

Lam, J. T. M. (2017). "Hong Kong District Council elections 2015 A political litmus test for the Occupy Central Movement". *Asian Education and Development Studies* 6.4, SI, 354–371.

Ng, K.-c. (2015). *Heung Yee Kuk chairman steps into father's shoes saying he'll seek his advice*. South China Morning Post. URL: https://www.scmp.com/news/hong-kong/politics/article/1814463/heung-yee-kuk-chairman-steps-fathers-shoes-saying-hell-seek.

Pang, J. and J. Smith (2019). *Hong Kong universities become 'battlefields' as citywide violence spreads*. Reuters. URL: https://www.reuters.com/article/us-hongkong-protests-idUSKBN1XM00X.

Pang, J. and J. Zhou (2019). *Violence spreads across Hong Kong's New Territories in 24th weekend of unrest*. Reuters. URL: https://www.reuters.com/article/us-hongkong-protests-idUSKBN1XK04Q.

Regan, H. et al. (2019). *City-wide strikes bring Hong Kong to a standstill*. CNN. URL: https://edition.cnn.com/2019/08/05/asia/hong-kong-strike-august-5-intl-hnk/index.html.

Robert, O. (2015). *The Wealthiest Neighborhood On Earth: Deep Water Bay*. Forbes. URL: https://www.forbes.com/sites/robertolsen/2015/07/29/the-wealthiest-neighborhood-on-earth-deep-water-bay/.

RTHK (2021). *RTHK's Bao Choy found guilty over Yuen Long report*. URL: https://news.rthk.hk/rthk/en/component/k2/1587096-20210422.htm.

Tang, G. and F. L. F. Lee (2018). "Social media campaigns, electoral momentum, and vote shares: evidence from the 2016 Hong Kong Legislative Council election". *Asian Journal of Communication* 28.6, 579–597.

Tong, E. (2015). *Fake candidates and 'vote-snatching': a new era of electoral fraud for Hong Kong?* URL: https://hongkongfp.com/2015/11/27/fake-candidates-and-vote-snatching-a-new-era-of-electoral-fraud-for-hong-kong/.

Vukovich, D. (Jan 2020). "A city and a SAR on fire: as if everything and nothing changes". *Critical Asian Studies* 52.1, 1–17.

Wei, J. C. Y. (2017). "The Strategic Coordination under Quasi-SNTV: A Case Study of Hong Kong". *Japanese Journal of Political Science* 18.1, pp. 155–175.

Wong, S. H.-W. (2017). "Gerrymandering in Electoral Autocracies: Evidence from Hong Kong". *British Journal of Political Science* 49.2, pp. 579–610.