ProxCode Analysis

Goal: To identify illogical values in a lot's proximity code, based on a building's proximity to other buildings, detected by spatial queries performed against building footprints.

Result: About 1.26% of lots marked as attached or semi-attached have buildings with no attached building. About 8.37% of lots marked as having freestanding buildings have buildings that appear to be attached.

Valid proximity code values include 0, 1, 2, and 3. 375 PLUTO records have values of NULL.

Attached and Semi-Attached Lots (ProxCode '2' or '3')

The idea here is to get buildings on lots designated as attached, where there is no attached building. First, I joined PLUTO to building footprints on MPLUTO_BBL, selecting only those lots where the proximity code was equal to '2' or '3', excluding lots where NumBldgs is greater than 1. This returned 331,926 buildings.

I then joined this result to building footprints again, this time looking for buildings intersecting with the building on the primary lot, where the BINs were not equal. The idea behind the query was to get a count of abutting buildings; if the query did not locate any abutting buildings, it returns a count of zero, and therefore should be among the result set I was interested in.

I loaded the results to a Postgres table that I would be able to review in QGIS. It seemed that the query was still picking buildings with abutting buildings. The distance (ST_Distance) between these buildings was not zero, however, so I ran an additional query to filter out buildings that were within 1 foot of the primary buildings. This returned 4169 buildings, or 1.26% of the 331,926 buildings designated as attached.

Detached Lots (ProxCode '1')

Here I did the opposite: I wanted buildings on lots designated as detached, which nevertheless have an attached building.

As before, I joined PLUTO to building footprints, this time to identify buildings on lots designated as standalone (162,869 buildings). From there, I looked for abutting buildings, buildings zero feet away from the primary building, and yet not sharing the same geometry. This returned 13627 buildings. I created two shapefiles: one of the primary buildings on detached lots having abutting buildings, and one with the abutting buildings.

Conclusion

I'm comfortable that the query works as I intended. I think it would be possible to correct the proximity code for lots marked as attached where the building actually appears to be freestanding using the query I've written. It is more complicated to correct lots marked as

freestanding where the buildings appear to be attached, because I think it would be difficult to differentiate between attached and semi-attached buildings.

Supporting Documents

Jupyter Notebook

Shapefiles available on Sharepoint