

Historical Multifamily Supply High Level Overview

Concept

The core idea of our historical multifamily supply analysis is to determine the supply of rentals, condos and coops per year from 1980 to the last year. We do this by starting with the most current version of pluto and stepping back through time by looking at what buildings existed according to year built. We further attach condo/coop conversion data from the NY Attorney General's office. Because many buildings that are currently condo or coop were originally rental this allows us to determine at what point the building was converted and allocate those units to the appropriate category.

Caveat

It should be noted that we are likely undercounting supply in past years because our methodology does not account for buildings which are demolished. However, we ignore this for the following reasons. 1) Large residential buildings being torn down to build fewer units is unlikely. 2) The novel finding of our analysis is that there are fewer rental units today than there were in 1980. If we could find a way to account for buildings that were demolished it would actually bolster our findings. As a result, though our estimate may be biased, the argument is honest.

Output

The final output of this project is an excel file of summary statistics in a pivot table. Rows are years and columns denote the classification of units (rental, condo, coop, etc). The pivot table has a filter of AREA. By filtering on AREA (NYC, Queens, Astoria) the pivot table shows the number of units by year from 1980 to 2017 for that geographic area.

Day to day use of our multifamily analysis should be through this pivot table. The collection of scripts in this repository should be used for yearly updates, troubleshooting or bespoke requests. For instance, if historical supply for areas within two blocks of Central Park was requested it would require using the R code instead of the excel file.

Steps

In the final stage of the process a dataframe called pluto augmented (pluto.aug) is utilized. This dataframe is pluto with the addition of the following fields:

- Type (condo, coop, rental, small housing, other)
- ConversionYear (year in which condo or coop buildings were converted)
- DECLARE_NEWBUILD (binary flag for if building was initially constructed as a coop/condo)

To create pluto augmented, two other dataframes must first be generated.

The first of these is a dataframe that identifies all buildings as condo, coop, rental, small (single or two family housing) or other. This is done by using the Department of Finance Notice of Property Value documents. While PLUTO shows building classification at the billing BBL level, in some circumstances (particularly RM class buildings) the condo lots inside can be either rental, condo or coop. NOPV documents allow us to see the classification of individual condo lots. To get this information we join a dataframe showing all NOPV classifications with PLUTO via PAD. If the billing BBL class indicates condo, we assign Type based on the most prevalent class at the condo level BBL within it. For instance, if a billing BBL classified as RM are a mixture of residential and commercial. If we look at the NOPV documents attached to it and see a few RK class lots (stores) along with an RR type lot (condo rental) which has 100 units, we can safely call this a rental building. However, if it has 100 R3 class lots inside of it, we can call it a condo building.

To create the conversion year dataframe we start with an excel file of all condo/coop conversions registered with the NY Attorney General's office. This file was procured from the NY Attorney General's office via FOIL request and will need to be updated yearly. Unfortunately, neither BBL nor any other NYC identifier is attached to the conversion documents. In the AG data the primary key is PLAN_ID and the foreign key is address. Many addresses were originally written on a paper form and are not standardized. In a number of instances multiple addresses are combined into a single plan. So, to attach these to pluto we must normalize the addresses. This is done through a combination of the NYC Geoclient API, Google Maps API, manual address cleansing and manually finding the BBLs of given addresses with Zola. The resulting dataframe is then put into a format which is friendly both to the final stage of the analysis as well as coming back later to troubleshoot or update.

In the last stage we join both of these dataframes with pluto to create pluto augmented. For each year we use pluto augmented to calculate:

The number of rentals -

Sum of residential units in buildings built that year or earlier + sum of residential units in condo/coop buildings that were built that year or earlier and converted at a later date

The number of Coops (or condos):

Sum of residential units in coop (or condo) buildings built that year or earlier and were either constructed as coop (or condo) or were converted after that year.

To calculate the number of individually owned condos being rented we apply a proportion derived from the New York City Housing and Vacancy Survey raw data. This is a triennial survey conducted jointly by the city and Census Bureau.

Finally, the data is joined with a separate file that contains time series population for neighborhood, borough and city. This is saved as a csv and then an excel file with a pivot table

in it is created. The easiest way to create the pivot table will be simply to mimic the previous file.

An additional aspect of the analysis is the proportion of units which are rented at fair market rates and those which are 421-a. The number of fair market units is also a proportion derived from NYC HVS. The number of 421-a units are derived from NYC Rent Guidelines Board.

Notes on NYC HVS:

The New York City Housing and Vacancy Survey is a triennial survey which asks a weighted representative sample of New York households a variety of questions regarding their housing circumstances. It should be noted that the unit of observation is a housing unit, NOT the people living in the housing unit.

The first wave was conducted in 1968. We have data dating back to 1980 but if so desired the Census Bureau would provide us earlier data as well. Our contacts at Census were incredibly friendly and happy to both provide data and answer questions.

Updating our results from NYC HVS will need to be almost entirely re-done by the data team member who is most familiar with survey methodology. The reason being that the survey is altered to lesser and greater extents every fourth wave. This is largely manifested in the way in which categorical responses are coded. However, particularly when comparing surveys which are temporally quite distant, the wording of questions changes and some questions are dropped entirely. Census will most often say flat out that they are not comparable. However, because we are using the data in a way that was not intended and do not need the level of accuracy other users require this is not necessarily the case for us. Judgement needs to be used as to whether questions are compatible and how to match the response coding.

The question of rent stabilization program should not be trusted. Many buildings can be stabilized under more than one program and the response only shows one of these. Precedence of the program changes. Most importantly our friends Melissa Klein and Bob (last name not known) at the Census Bureau indicated that this question was for internal use and the values should not be trusted.

Currently we calculate our proportions city-wide. The survey would allow us to do so based on more granular geographies if we so chose to do so. This should be approached carefully. We already need to combine multiple questions and might quickly run out of statistical power. Because the survey is representative and not random degrees of freedom calculations are somewhat complex. Instructions are indicated in the NYC HVS documentation but do not directly apply to our usage of the data.

Currently the most recent wave of NYC HVS we use is 2014. The 2017 wave will be available shortly

<https://www.census.gov/programs-surveys/nychvs/data/datasets.html>

If more detailed information is required contact Melissa Kresin at the Census Bureau. Let her know you work for Hodges Ward Elliott and have taken over for Bill Bachrach. Be sure to offer profuse thanks for their assistance and let them know that the data has been extremely valuable to us.

Melissa.Kresin@census.gov

301.763.9825

In the event Melissa no longer works in there or is unable to help, her colleague Bob (didn't ever catch his last name) was also of great assistance. The department they work in is:

Financial & Market Characteristics Branch
Social, Economic & Housing Statistics Division
U.S. Census Bureau

Notes on RGB 421-a calculation:

To calculate the number of 421-a units we initially used estimates from the rent guidelines board. We found this in a series of (annual) research reports issued by RGB. The URL below provides an example of such a research report. These research reports only extend back to the 1990s so the number of units was then back-casted.

<https://www1.nyc.gov/assets/rentguidelinesboard/pdf/changes18.pdf>

However, this is problematic because 421-a stabilized units cease to be stabilized after 10-15 years (depending on the specifics of the building). Moreover, we received pushback on the low estimate of 421-a units. Tim spent a good deal of time contacting city officials and found a better source. Unfortunately I have been unable to find documentation on this source.

General Notes

- The increase in Fair Market stock comes primarily from de-regulation, not from new construction.
- Will counts only 15% of 421-a units as stabilized.