Initial documentation for Historical Multifamily Supply Analysis

Notes: need to entirely re-do the AG Data script

multifam munge and plotting (main script):

* Read in condo rental metric (relies on outside work, likely HVS)
* Create condo/coop conversion dataframe {relies on outside work: AG Data}
  + Generate aggregate conversion statistics
* Create pluto augmented (pluto with extra field showing condo or coop)
  + *Note: it appears that building class A and B are lumped in with condos*
  + Read in condocoop dataframe {relies on outside work: new key create script (rename)}
  + Read in pluto
  + Join pluto with the condo coop dataframe and munge
  + Bring in handcoded data (?) to fill out pluto
  + TCO lag
  + Final munge
* Create conversions dataframes
  + Pluto.aug had been joined with the AG data, so going through by year and area and quantifying the number of conversions
* Step backward through time with pluto augmented to get the number of units by year
  + Listwise for area
  + Join into dataframe and create the coop rentals etc variables
  + Do for Neighborhood, Borough, NYC
  + Bind into single dataframe

npv\_key\_create script:

* load and munge pluto
* load and munge expanded pad
* join pluto and pad
* create initial notice of property value key
  + includes bbl, document year, file size, filename, year (year for which the lot is classified: document year -1)
* join npv key with pad/pluto
* join pad/pluto with npv key (this results in the full list but with missing npv values)
* scrape pdf building classes using initial npv key
  + turn into dataframe the main components of which are the condo level bbl and building class
* join with pad/pluto
* for each billing bbl, get the internal building classes

NEW ORDER OF SCRIPTS

Npv key create

AG Data Address Cleaning

Pluto\_augment\_create

Back\_in\_time