Guide to Geocoding w/ GB

I. Before you Geocode

The files you work with (in "NYC_geocode"):
- **basemaps**: William Perris Fire Insurance Maps for 1852-54 and 1857. The address locator is based on the 1852 map plates. These are displayed in ArcMap via a server connection. If you ever lose that connection or need to reset it, see "GEO_AddingBaseHistoricalImagery.pdf" in the NYC_geocode folder for step-by-step directions.
- **street directories**: "1854-1855_combined.pdf" is the Directory from which all addresses have been derived and the one you'll use as a reference. There is also the "1845-1846_combined.pdf" directory that furnishes 1845 addresses. You'll rarely have to use this older directory. The 1854 Street Directory begins on pdf page 805 and the 1845 Street Directory begins on pdf page 434.
- **building footprints:** in */NYC_geocode/shapefiles/buildings*. This is a polygon shapefile of all building footprints as marked on the 1852 Perris basemap. Each building has been digitized and assigned an address. This shapefile is the foundation of the address locator. This shapefile was crowdsourced and therefore contains many errors (see section 7 below on crowdsourced building footprints). There are two files, "perris_build_final_proj" and "perris_build_final_proj_Copy." We work from the "Copy" version which has been edited to amend some of the crowdsourced errors and serves as the foundation for the "building_addresses_2" address locator. However, many errors remain so mark them as you see them and plan to amend this shapefile in the future.
- **address locator:** located in the "address_locator" folder in the NYC_geocode Dropbox folder. This includes two address locators, "building_addresses" and "building_addresses_2." The second one "building_addresses_2" is the most recently updated address locator and the one that we have used to geocode most feature classes up to this point (May 2015). Feel free to make a new one based on your edits to the building footprints feature class. Be sure it's in a projected coordinate system (New York State Plane Long Island FIPS).

General tips
- Red buildings are brick, yellow buildings are either wood or stone (ask GB), and green buildings are industrial buildings.
- Home addresses should not be matched to green buildings as they do not constitute residences
- The 1854 Street Directory begins on page 805 of the address directory
- Always write down anything unusual. It's extremely important to keep an archive of your decision-making; it may solve problems that come up in the future.
- ***Always mark where you left off if you do not finish a batch in one session!***
- Geocode matched addresses first ("Matched addresses score 80-100), taking care to unmatch any address that does not fit. Then move to the unmatched addresses and go through. When all possible matches have been matched, check for rear buildings: go to "All Addresses" and sort descending by either the home_position or the bus_position field (depending on which type your geocoding), and check to see if any of the buildings are rear buildings.

Geocoding Settings
- use address locator's spatial reference
- set all score minimums at 80
- uncheck the "Match tied candidates" setting: we don't want to match tied candidates

● display XY coordinates and keep Reference ID checked

Interactive Rematch dialog
"search" field

Scroll over in the interactive rematch box to the "ARC_Single" field and sort that field by ascending addresses. Then scroll over to the "search" field and click directly on the first entry in the field so that you can easily scroll down with the keyboard. It's important to geocode on this field because it's a direct reference to the original address as written in the original directory, rather than the address as written by the employee who creates these spreadsheets. In other words, it's a field that is unaffected by human error, only perhaps the PDF error of digitizing all the text. But the main reason why it's important to have this field handy is because it can answer your questions faster whenever you happen upon an error or discrepancy in the geocoding .

Address position

Address position tells you where a building address is located within a building lot or at a street corner. Sometimes an address will have multiple buildings assigned to it with one main building facing the street, but an individual may reside in one of the rear buildings. Address position may also tell you whether an address is located at the end of block (i.e.corner building). These do not come up frequently but there are at least a few within each feature class. In the "search" field in each excel spreadsheet, address positions are denoted with abbreviations ("r." for rear building; "c." for corner building); address positions are also isolated into their own fields: "home_position" and "bus_position" fields which correspond to the respective positions of home and business addresses.

As this project requires high accuracy and precision in geocoding, it is important to check for the address positions. This is typically done at the end of a geocoding session after all matched and unmatched addresses have been sorted through and matched. Scroll over to either the "home_position" or "bus_position" field (depending on whether you're geocoding home or business addresses) and sort the field by descending.Then, edit each address match according to its position if it is specified.

Do not match an address in a rear building if the rear building is not on the map.

II. Solutions to Common Issues

1. Addresses missing from the map

Some addresses can't be geocoded because they cannot be found on the original 1854 basemap. These include addresses with house numbers that may exceed the highest number on a certain street, vanity addresses (see section below on vanity addresses), Brooklyn and New Jersey addresses, streets that have been renamed or that are not in Manhattan, and addresses that extend beyond the northern extent of the map (addresses located higher than the southern side of 42nd street). Most of the time, these addresses will not be matched at all. Still, you should note these instances as you geocode and record the reason why they cannot be matched to familiarize yourself with the data and for future reference.

Before you write off an address, first compare its address on the 1852 Perris basemap to the 1857 Perris basemap, or consult with the 1854 Street Directory to determine its cross streets. Sometimes, this comparison reveals the location of an address that may not have been initially apparent on the 1852 basemap. An example: an address on the 1852 basemap is inconsistent with the 1854 Street Directory but consistent with 1857 basemap. If an address on the 1857 basemap is corroborated by the 1854 directory, you can safely geocode according to the 1857 basemap. This method may also work when geocoding an address for a building or block that may be lacking house numbers altogether.

Another issue is when a series of consecutive buildings skips a number that should be there. An example is Broadway 180 and 184. There is no building for 182 on the basemap. But we can assume that its location, if it were reflected on the basemap, would be between 180 and 184. In this case, we can match Broadway 182 to Broadway 184 because it is the correct location even though the numbering may say something else.

Another way to address this particular issue, or any instance of a missing address, is to simply search for a nearby or consecutive address within the interactive rematch dialog. To use the example of Broadway 182, rather than scrolling around the map for this specific address, search for Broadway 184 and "Zoom to candidates." This is a much faster method for dealing with unmatched addresses.

As a general rule, when in doubt, consult the 1857 Perris map and the 1854 directory. If you still can't figure it out, consult with GB and/or leave it unmatched.

For future reference, some renamed streets include Factory Street (actually an extension of Waverly Place in the West Village that starts at 151) and Walnut Street (becomes Jackson Street).

2. Missing building footprints

It is fairly common to find addresses that do not have a building footprint on the 1852 basemap. Footprints in a particular location may appear on the 1857 basemap, but there is ultimately nothing we can do for addresses without footprints, so we keep these unmatched. Below is running list of common addresses that are missing footprints:
- Trinity Buildings at 111 Broadway
- Lamartine Place on 29th Street between 8th and 9th Avenue

3. **Missing basemaps**

For whatever reason, the 1852 Perris basemap is missing a few plates. Addresses that correspond to the location of a missing map plates are left unmatched because there are no building footprints to match. Below is a running list of missing map plates denoted by their boundaries:

- **East Village**: 6th, 7th, and 8th Streets between Avenue C and Avenue D
- **City Hall/City Hall Place:** The east side of Broadway from 218 north to "Stewart's Store"/Broadway 290; the south side of Chambers Street from 15 to 76; the northern side of Park Row numbers 1 through 34/"Trustees of the Church"; and the northern side of Nassau Street from 154 and Chatham Street 9 to 67.
- **St. John's Lane** - *see 1854 directory*

4. **Existing institutions**

The original 1852 basemap labels landmarks and important institutions such as churches, hotels, coal and lumber yards and other manufacturing institutions, post offices, schools, and railroad freight depots. In addition, individuals' addresses may also make specific reference to an institution within the address field. There is no general rule for geocoding on these addresses and geocoding should therefore occur on a case-by-case basis.

You should never match a home address to a church, manufacturing institution, post office, or school if the given home address matches that of the institution. Dealing with hotels is a little less straightforward. An common example is when the given home address for individuals in elite and upper middle-class professions (such as lawyers, bankers, and doctors) ends up being an existing hotel. It was fairly common for these types of professionals to take residence in hotels so you can match them. Another common issue is when lower-class professionals such as tailors, clothiers, and shoes and boots retailers have business addresses in existing hotels. We are not sure how common this was in 1854, but it comes up enough in the data to argue that these addresses can be matched to a hotel. I have matched these addresses in the past and noted when they have been hotels. Ultimately, geocoding these cases is mostly up to GB's and your own discretion.

Often, individuals may list a specific institution such as a food market or City Hall as their business address without naming the actual street or building number. Sometimes, these buildings are labeled directly on the basemap. For example, many merchants will name "Merchants' Exchange" as their business address, which is labeled on the 1852 basemap as "The Exchange" (366 Wall Street). When you don't know the building's location on the basemap, run a search in the 1854 directory for that institution and it will most likely have the address. If it's not listed in the directory, record the address and institution in your notes and consult with GB.

In all cases concerning an existing institution, you should always note the individual's address and the institution to which you did or did not match it for future reference.

A final note: GB has excluded individuals who have explicitly given hotels as their home addresses. These are compiled into a series of excel sheets by profession that he will eventually have you geocode.

Here is a running list of existing institutions (and their addresses) that are either in the 1854 directory or labeled on the 1852 basemap:

- Post Building
- The Mercantile Library Society
- Christian
- The Exchange: Merchants Exchange, Wall Street 366
- Trinity Buildings: Broadway 111 (there is no building footprint so these addresses are left blank
- City Hall: this is located on one of the missing map plates, so these addresses are typically left blank
- Centre Street Market, Washington Square Market

5. **Vanity Addresses**

Vanity addresses: a customized address that differs in street name and numbering layout of the original street plan (some contemporary examples: 30 Rock, Lincoln Center, One57). These are arbitrarily assigned to buildings by developers both in the past and in the present. Within our data, vanity addresses typically come in the form of renaming a portion of a street as a "place," such as Abingdon Place, Carroll Place, or St. Mark's Place. Most of these vanity addresses go undetected by the address locator because they are not marked on the 1852 basemap. Therefore, you will routinely consult the 1854 directory to locate them. Below is a running list of vanity addresses and their actual locations:

- Clinton Place: 8th Street, west from 755 Broadway
- Abingdon Place: in Troy Street between
- Abingdon Square
- Amity Place
- Carroll Place
- Lafayette Place
- Lamartine Place: West 29th Street—there was no development there in 1852 so these addresses are mostly left blank
- St. Mark's Place: 8th Street

6. **Duplicates**

Duplicate entries and addresses are very common in the 1854 data but they come in a number of forms that are treated differently. These include normal duplicates, typo duplicates, family name duplicates, institutional duplicates, and duplicates resulting from an individual with a single home address owning multiple business addresses. To detect duplicates, you should keep the "search" field in the Interactive Rematch dialog box in view.

- *normal duplicate:* This is when an individual and his/her address has been entered two or more times into the original excel file, usually in succession. In the "search" column, the entries should be identical down to the letter. These entries should not be coded more than once so unmatch entries wherever necessary. Be attentive while you geocode so that you can detect these issues.
- *typo duplicate:* This occurs when an individual was entered twice in the directory under multiple names, at least one of which is misspelled. For example, you may find a "M'Kaeng William" at 254 Sullivan Street and a "M'Kaeg William" at the same address. The names are close enough to assume that one is a typo and the other is the true address. This may not be true in every instance so again, matching or unmatching addresses where necessary is up to your discretion. Still, this is a fairly common issue so be attentive.

- ○ *duplicates in home feature classes:* Sometimes, an individual will have multiple business addresses but only one home address. Geocoding the business addresses is not a problem since each business entry is unique. The problem arises when you need to geocode the home addresses: you end up with duplicated home addresses that correspond to each unique business address. You will typically detect these the same was you would a normal duplicate, two addresses in succession; the difference is that the original 1854 entry in the "search" field will reveal multiple business addresses registered to the same person. Verify that this is the case and unmatch addresses where necessary, writing down the individual's name, the duplicated address, and the reason why the entry has been excluded.
- ○ *family name duplicate:* Sometimes, family relatives such as a father and a son may live or work in the same place. Often they have the same name, distinguished only by an initial. These addresses that may appear to be duplicated are actually unique. Therefore, always check the original 1854 entry for name initials before you unmatch it. If the entries are unique (as in, there's a slight different in name), match both.
- ○ *institutional duplicate:* Certain individuals may list their business addresses twice: once for their personal entry and once for the business itself if the business is named after them. For example, law partners Bob Smith and John Moore may each have unique business entries as well as a third entry that reads "Smith, Moore & Associates" with the same business address. While these are duplicated entries, GB prefers to geocode both the individual and business entries in order to model the size and volume of a business. This is common to lawyers, tailors, shoes & boot retailers, hat retailers, and clothiers.

7. **Issues in crowdsourced building footprints: mismarked footprints and picking points**
Because the building footprints shapefile was developed through crowdsourcing, it is prone to frequent and varied errors. Generally these errors come as typos, missing addresses for a footprint, or an incorrect street name (for example, in the original footprints shapefile, West 4th Street was input as 4th Avenue which created frequen incorrect address matches during geocoding). To deal with these issues, you will often have to manually locate the address to find these addresses and, in some cases, pick a point to assign to a building address.

You typically discover typo issues when you have to manually search for an address on the map that does not come up in the interactive rematch dialog. When you find the typo, input the typo address into the interactive rematch and match the typo'd address (it should be 100%).

Many typos are actually alternate spellings of a street. For example, Catherine Street in Lower Manhattan is spelled with an "a" (Catharine Street) in the 1854 Street Directory, meaning that every individual living on Catharine Street will go undetected by the interactive geocoder. To address the issue, simply search for your address with the map's spelling of Catherine Street (with an "e") and match the entry to this address.

Sometimes, a building footprint is missing an assigned address altogether, likely because the person who input the addresses accidentally skipped it. In this case, manually locate the address on the map. When you find it, verify that a building footprint for the address has been drawn into the crowdsourced feature class. Then, pick a point for the address: click the "Pick Address from Map" button at the bottom of the interactive rematch box, scroll over to the location where you want to place the point, right-click directly on this location, and select "Pick Address" from the menu. *Important:* You only pick a point if its building footprint has been drawn in—if it exists on the basemap but doesn't have a corresponding footprint, leave the address unmatched.

Below is a list of errors with the crowdsourced feature class:
- ○ Catherine Street is spelled "Catharine Street" with an "a" in the directory but "Catherine Street" on the Perris Fire Insurance Map and the corresponding building footprints shapefile
- ○ Delancy versus Delancey
- ○ 4th Street versus 4th Avenue
- ○ The first block of 1st Avenue and 2nd Avenue are written as Allen and Chrystie Streets. Or maybe clinton. That's wrong, but we have fixed the building footprint feature class, so you just have to know.

8. **Street renumbering**
Some street layouts can be confusing and warrant explanation. Below is a running list of such instances that is not at all exhaustive. Add to this list as you encounter these issues:
- ○ **Bowery:** Between East 6th Street and East 14th Street, Bowery ceases to be Bowery and becomes 4th Avenue. According to the 1854 Street Directory, **East 6th Street** is the boundary of transition between Bowery and 4th Avenue. However, the 1852 basemap marks **East 14th Street** as the boundary of transition between Bowery and 4th Avenue. As a result of this difference, the first hundred or so street addresses on 4th Avenue are missed entirely by the address locator. Lucky for us, it turns out the the Bowery and 4th Avenue Street numbering schemes are consistent: 1 4th Avenue begins at 401 Bowery, 2 4th Avenue is equivalent to 402 Bowery, and so on and so forth. So if you can't find a single-digit or low double-digit 4th Avenue address, search for it as a Bowery instead and add 400 to the address in question.
- ○ **Union Place:** Union Place is actually two streets as noted in the 1854 Street Directory: 4th Avenue and Broadway between 14th and 17th Streets. Even Union Place addresses are on 4th Avenue, odd Union Place addresses are on Broadway. Be careful not to mistake 1 Union Place (in the address locator as 1 Broadway) for the true 1 Broadway located at Battery Place at the bottom of Manhattan.
- ○ **Madison Square:** Similar to Union Place, Madison Square addresses are also on multiple streets. The addresses go around the square, which is located between 5th Avenue and Madison Avenue, and East 23rd and East 26th Streets. For even Madison Square addresses, code on East 23rd Street. For odd addresses, code on East 26th Street.
- ○ **Gramercy Park:** Gramercy Park is a little more straightforward than Madison Square because the numbers follow the perimeter of the park. East 20th Street addresses however are a bit confusing. Use your discretion.

III.   Editing

As mentioned briefly above, the building footprints shapefile has been edited to fix issues such as address typos that can make geocoding more difficult. This is not a top priority and you should only pursue this if you don't have any other pressing material to geocode.

The process is very simple. As you geocode, mark down addresses where you notice typos, alternate names, or missing addresses in the footprints feature class. Make sure that these notes are distinct (I used an abbreviation "AL" for address locator in my geocoding notes and after a while, I maintained a separate running listed for address locator edits). You should be ready to make edits after surveying each plate of the basemap a number of times through geocoding and you have a substantial amount of edits.

To edit, you'll proceed as you would with any other feature class. First, make a copy of the "perris_build_final_proj_Copy" and rename it in order to keep the original in tact. Save this new shapefile in the same "/NYC_geocode/shapefiles/buildings" folder. Begin an editing session in ArcMap and select the copied building footprint feature class to work from. Sort the table by ascending according to the "street" field. Then, pinpoint your addresses one by one within the table and make the necessary edits. You'll only edit 4 fields for each entry you edit: "updated_at," "number," "street," and "address_1." For every edited entry, replace the date (2013-10-17 or 2014-08-21) with the current date in the same format. The "number" and "street" fields are respectively the house number and street of the building in question. This last field, "address_1" is the field that is used to label the building footprints as you geocode. Be sure to input the edited address in the same Street-Number format. Repeat this for every edited entry and **be sure to save your edits along the way and at the very end.**

Once you are done editing, **and you've saved your edits**, you will create an address locator. I believe there is a "Create Address Locator" tool that you can search in the ArcMap search bar. If not, navigate to your edited building footprints shapefile in ArcCatalog, right-click, scroll to "New" and select "Address Locator." See the following inputs for the dialog box:
- "Address Locator Style" set to "General - Single Field."
- "Reference Data": navigate to the "/NYC_geocode/shapefiles/buildings" folder and select your edited building footprints shapefile
- "Role" set to "Primary Table"
- Verify that the fields properly correspond to the address locator fields
- Select "/NYC_geocode/address_locator" folder as the save destination and name the address location to "building_addresses_3"

Click "OK" and you should have a new address locator that reflects your edits. When you geocode via the Interactive Rematch tool, be sure to set the address locator as the new one you just created "building addresses_3." Ideally, the percentage of automatic matches should increase when you geocode via the Interactive Rematch dialog.

IV. Geocoding profiles by profession (in order of completion date)

1. **Bankers**
   a. **bankers_geo**: 100% (52 Total, 52 matched, 0 unmatched)
   b. **bankers_bus_geo**: 96% (94 Total, 91 matched, 3 unmatched) Unmatched addresses include those missing footprints on the actual map or those lying beyond the extent of the map.
   c. **bankers_bus_geo_2**: 100% (56 Total, 56 matched, 0 unmatched)

2. **Brokers**
   a. **brokers_geo**: 97% (433 Total, 418 matched, 15 unmatched)
   b. **brokers_bus_geo**: 97% (767 Total, 743 matched, 24 unmatched) Unmatched addresses include those missing footprints on the actual map or those lying beyond the extent of the map.
   c. **brokers_bus_geo_2:** 98% (380 Total, 376 matched, 4 unmatched) Unmatched addresses include those missing footprints on the actual map or those lying beyond the extent of the map.

3. **Merchants.** As noted above, many merchants will name "Merchants' Exchange" as their business address, which is labeled on the 1852 basemap as "The Exchange" (search "Wall Street 366"). Jauncey Court is another commonly cited location (search "Wall Street 37-41").
   a. **merchants_home_geo**: 95% (314 Total, 299 matched, 15 unmatched)
   b. **merchants_bus_geo:** 98% (580 Total, 568 matched, 12 unmatched). 504 blank entries were carried over from the Excel sheet into ArcMap. These have been excluded from totals.
   c. **merchants_bus_geo_2**: 98% (422 Total, 414 matched, 8 unmatched) Unmatched addresses include those missing footprints on the actual map or those lying beyond the extent of the map.

4. **Fish**
   a. **fish_st_house_geo**: 89% (18 Total, 16 matched, 2 unmatched)
   b. **fish_store_geo**: 96% (49 Total, 47 matched, 2 unmatched)
   c. **fish_market_geo**: 95% (56 Total, 53 matched, 3 unmatched)

5. **Milk**
   a. **milk_house_geo**: 97% (38 Total, 37 matched, 1 unmatched)
   b. **milk_store_geo**: 97% (332 Total, 321 matched, 11 unmatched)

6. **Lawyers.** Duplicates were common to the lawyers feature class, particularly family name duplicates as well as duplicates due to co-ownership of a firm or multiple firms owned by a single person.
   a. **lawyers_home_geo**: 97% (785 Total, 763 matched, 22 unmatched)
   b. **lawyers_bus_geo_1**: 95% (1177 Total, 1122 matched, 55 unmatched) 481 blank entries were accidentally appended to the excel table. These have been excluded from the final percentage. The remaining addresses are unmatched because of missing building footprints and basemaps or because they're duplicates.
   c. **lawyers_bus_geo_2**: 94%. (500 Total, 472 matched, 28 unmatched) Almost all of the remaining addresses are missing building footprints on the original basemap. Excluding, these, I've coded about 99% of the feature class.

7. **Grocers 1845**. grocer_store(1&2)_geo_edit: 99% (1940 Total, 1913 matched, 27 unmatched. This is a feature class of the business addresses for grocers in 1845. The shapefile attribute table notes that there are 2040 entries, but 100 blank entries were accidentally appended to the excel table, so there are actually 1940 entries.The inaccuracies in geocoding this data squarely result from the mismatch in the house numbering over one decade between the 1852 Perris basemap and the 1845 directory data. Also, this data is lower quality because it's from an earlier directory. The main issue in geocoding this feature class is with the change in street numbering between 1845 and 1852,

which in many parts of the city shifts by a block or more. This feature class was geocoded according to the 1852 basemap, with certain addresses relocated to reflect the change in street numbering during the decade. In order to figure out which addresses had to be relocated, the 1845 Street Directory was cross-checked against the 1852 basemap and the differences in numbering were noted for every street intersection on the basemap. Below is a list of streets in the 1845 feature classes will have to be recoded or reassessed. For each of the following streets, the error in numbering persisted for 3 or more blocks. Those in parentheses have negligible errors. Finally, "**" indicates that numbering is off by >5 homes.

Ann Street
(Avenue 1)
Avenue 3rd
Avenue 6th
Avenue 8th**
Avenue 9th**
Canal Street**
Catharine Street
(Charlton Street)
Clinton Street**
Division Street**
Eldridge Street
Grand Street**
Greenwich Street**
Hester Street
Houston Street**
Madison Street**
Market Street
Monroe Street**
Mott Street**
Orchard Street
Prince Street
Spring Street**
Suffolk Street**
West Broadway** (off by two blocks)
4th Street**
5th Street**
10th Street**
12th Street** (between Greenwich Ave and 7th Ave)
West 13th Street
West 14th Street
West 19th Street
West 20th Street**
West 21st Street**

8. **Grocers 1854**. grocers_home_1854_geo: 96% (1424 Total, 1361 matched, 63 unmatched) 10 of the unmatched addresses exist beyond the map's northern extent. The remaining missing addresses were missing basemaps.

9. **Clerks**
    a. **clerks_home_geo**: 95% (504 Total, 480 matched, 24 unmatched) Main issues include missing basemaps, streets that don't exist, and street numbering that extends beyond the extent of the map.
    b. **clerks_bus_geo**: 93% (544 Total, 507 matched, 37 unmatched) The missing City Hall basemap accounts for the relatively low matched percentage. 21 of the 37 remaining unmatched addresses are City Hall addresses. The remaining unmatched addresses exist because they don't have a footprint, they lie beyond the extent of the street, or they're located at various irrelevant institutions i.e. coal yards, American Bible Society, and a saloon.
10. **Laborers.** 6000+ laborers addresses were originally collected but less than 200 entries actually explicitly distinguish between home and business addresses and the geocoding result yielded little content for analysis. Then two random samples of 1,000 laborers each were taken of the 6,000+ cases in which the given address is assumed to be the home address. It is listed under "bus_address" within the excel spreadsheet, but in reality is home address. Many laborers also live in areas where the building footprint has not been drawn on the map or in informal/ungridded structures.
    a. **laborers_home_geo_1**: 95% (924 Total, 879 matched, 45 unmatched). At least 18 out of the 45 unmatched addresses exist beyond the northern extent of the 1854 basemap.The remaining addresses for industrial buildings or have missing footprints/basemaps.
    b. **laborers_home_geo_2**: 97% (946 Total, 888 matched, 58 unmatched) This percentage excludes all addresses that lie beyond the northern extent of the basemap. If we include these addresses, only 94% of addresses have been coded.
11. **Tailors.** Duplicates were common to the lawyers feature class. Similar to lawyers, they were particularly family name duplicates as well as duplicates due to co-ownership of a firm or multiple firms owned by a single person.
    a. **tailors_home_geo**: 96% (415 Total, 373 matched, 42 unmatched). The tailors feature class contained about 25 addresses that went beyond the extent of the basemap (23 of them to the north, 1 to the west, and 1 at Avenue A). Before factoring these out, 90% of tailors homes were coded (42 addresses unmatched out of 415 total). Excluding the addresses that extend beyond the map, 96% have been coded (373 out of 390). The remaining unmatched addresses almost all have missing basemaps or footprints.
    b. **tailors_bus_geo**: 96% (417 Total, 402 matched, 5 unmatched) 8 addresses are located beyond the northern extent of the map. Many of the tailors' business addresses are located in hotels. A few are located at "Barnum's Museum" on Park Row, at churches, and at the Mechanics Institute at the start of Bowery and have therefore been left unmatched.
12. **Physicians**
    a. **physicians_home_geo**: 96% (168 Total, 161 matched, 7 unmatched)
    b. **physicians_bus_geo**: 98% (185 Total, 181 matched, 4 unmatched)
13. **Mechanics**
    a. **mechanics_home_geo**: 97% (37 Total, 36 matched, 1 unmatched)
14. **Hats**
    a. **hats_home_geo**: 96% (57 Total, 55 matched, 2 unmatched)
    b. **hats_bus_geo**: 97% (150 Total, 146 matched, 4 unmatched)
15. **Shoes and Boots**
    a. shoes&boots_home_geo: 90% (178 Total, 161 matched, 17 unmatched). Most of the unmatched entries are duplicated home addresses where there were multiple business addresses.

b. **shoes&boots_bus_geo**: 96% (771 Total, 740 matched, 31 unmatched) I finished shoes and boots business addresses. All of the unmatched addresses exist beyond the extent of the map, miss a footprint, or are home addresses. There were also a number of addresses located in hotels and one located at the New York Herald (99 Nassau Street), which was marked down but matched.

16. **Clothiers**
   a. **clothiers_home_geo**: 96% (440 Total, 389 matched, 51 unmatched). 34 home addresses had locations but were left unmatched because they were duplicated for individuals with multiple businesses. The remaining entries are missing from or exist beyond the extent of the basemap.
   b. **clothiers_bus_geo:** 97% (1048 Total, 1012 matched, 36 unmatched)Unmatched addresses include those with missing basemaps and footprints and those located in industrial areas (Iron Yard, for example), or duplicates (very few).

17. **Dry Goods**
   a. **drygoods_home_geo**: 96% (522 Total, 472 matched, 50 unmatched). 31 home addresses had locations but were left unmatched because they were duplicated for individuals with multiple businesses.
   b. **drygoods_bus_geo**: 96% (1234 Total, 1182 matched, 52 unmatched)

18. **Jewelers**
   a. **jewelers_home_geo**: 97% (229 Total, 222 matched, 7 unmatched)
   b. **jewelers_bus_geo**: 97% (714 Total, 694 matched, 20 unmatched). A fair number of business addresses are listed under bus_position with "rear" addresses. Also, there are many duplicated business addresses that have been matched in order to model size of a business. Uncoded business addresses for furniture and jewelers are typically missing a basemap, footprint, or address or the location is an unrelated institution, such as a church or coal yard.

19. **Furniture**
   a. **furniture_home_geo**: 95% (133 Total, 127 matched, 6 unmatched). This feature class excludes duplicate entries for individuals with more than one brick-and-mortar store. Duplicates have been in recorded in notes.
   b. **furniture_bus_geo**: 96% (248 Total, 238 matched, 10 unmatched). are typically missing a basemap, footprint, or address or the location is an unrelated institution, such as a church or coal yard.

20. **Bookkeepers**
   a. **bookkeepers_home_geo**: 94%, (110 Total, 103 matched, 7 unmatched)
   b. **bookkeepers_bus_geo**: 97% (437 Total, 424 matched, 13 unmatched)