Using NYBBLES Data in GUPS

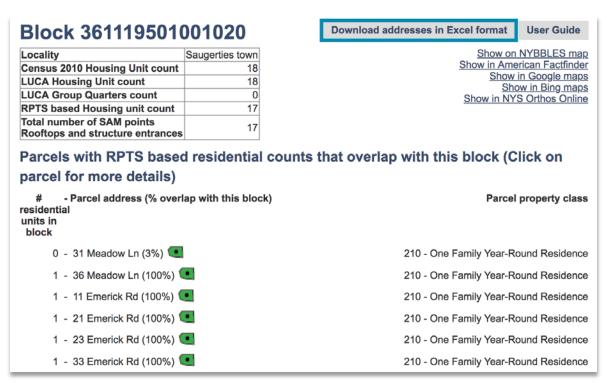
By: Elizabeth Womack, Cornell Program on Applied Demographics

The purpose of this guide is to show how to use NYBBLES generated data in GUPS. The guide is outlined in the following way: 1. Preparing NYBBLES Data for GUPS, 2. Importing Data into GUPS, and 3. Downloading Data from GUPS. Section 1 uses real examples from NYBBLES, sections 2 and 3 use fake data and a test version of GUPS.

1. Preparing NYBBLES Data for GUPS

1.1 Downloading Data

When looking at a particular block in the NYBBLES block browser, individuals can download individual excel files for each blocks by clicking the "Download addresses in Excel format" button. We are interested in this particular block because it appears the Census Bureau has one more record of housing units located in the Census block than estimated by the RPTS data.



After downloading the file, an excel file will appear with a variety of headers as seen in the table below, which are described in more detail in the NYBBLES User Guide. In order to prepare this data for GUPS, I will focus on information that is populating the headers that are bolded:

- **parcellD:** The ID number or order number of the parcel address for which an entry is located in. ID numbers are not unique between blocks, only within blocks.
- MAPSPOT: The order number ID of the SAM point located within
- Source: The source for the address entry, RPTS or SAM

- **Suggested action**: A suggested action in order to help make the downloaded address file useable in GUPS, such as deleting duplicates or verifying the existence of an address within a block
- STATE
- COUNTY
- TRACT
- BLOCK
- GEOID
- GQ
- HOUSE NUMBER
- STREET NAME
- APARTMENT/UNIT #
- ZIP CODE
- GQ NAME
- FACILITY NAME
- LOCATION DESCRIPTION
- NONCITY-STYLE ADDRESS
- NONCITY-STYLE ZIP CODE
- USE
- LATITUDE
- LONGITUDE
- CITY_STYLE ADDRESS

Looking at these particular fields in Excel, my data has the following information:

parcelID	MAPSPOT	source	Suggested action	HOUSE NUMBER	STREET NAME	ZIP CODE
1		RPTS (3% in block)	Check presence of HU in parcel in block (0 estimated)		Meadow Ln	
2		RPTS (100% in block)	Compare this and next SAM record - keep 1	36	Meadow Ln	
2	1	SAM	Compare this and previous parcel record - keep 1	36	Meadow Lane	12477
3		RPTS (100% in block)	Reduce to 1 address records	11	Emerick Rd	
3	1	SAM	Check address and remove some	15	Emerick Road	12477
3	2	SAM	Check address and remove some	11	Emerick Road	12477

4		RPTS (100% in block)	Compare this and next SAM record - keep 1	21	Emerick Rd	
4	1	SAM	Compare this and previous parcel record - keep 1	21	Emerick Road	12477
5		RPTS (100% in block)	Compare this and next SAM record - keep 1		Emerick Rd	
5	1	SAM	Compare this and previous parcel record - keep 1	23	Emerick Road	12477
6	RPTS (100% in block) Reduce to 1 address records		33	Emerick Rd		
6	1	1 SAM Check address and remove some			Emerick Road	12477
6	2	SAM	Check address and remove some		Emerick Road	12477
7		RPTS (100% in block)	Fill estimated 1 address records	25	Old Rt 9W	
8		RPTS (100% in block)	Fill estimated 1 address records	37	Emerick Rd	
9	9 RPTS (100% in block)		Fill estimated 1 address records	31	Emerick Rd	
10	10 RPTS (100% in block)		Reduce to 1 address records	29	Emerick Rd	
10	1	SAM	Check address and remove some		Emerick Road	12477
10	2	SAM	Check address and remove some		Emerick Road	12477

1.2 Cleaning the NYBBLES Data: Example A

In order for the information from NYBBLES to be useable in GUPS, there is a bit of data cleaning required. Let us explore the table more closely. Below, I have grouped properties located within a particular parcel address using color

Note, this in only a partial listing of the Block Excel file which has 39 rows.

Ultimately, because there are 17 housing units estimated using the Real Property Tax data, and 18 housing units estimated by LUCA, therefore there should be at least 17 unique housing units at a minimum. We need to fix the excel file so it reflects this.

parcelID	MAPSPOT	source	Suggested action		STREET NAME	ZIP CODE
:	1	RPTS (3% in block)	Check presence of HU in parcel in block (0 estimated)	31	Meadow Ln	
:	2	RPTS (100% in block)	Compare this and next SAM record - keep 1	36	Meadow Ln	

			-	,		
2	1	SAM	Compare this and previous parcel record - keep 1	36	Meadow Lane	12477
3		RPTS (100% in block)	Reduce to 1 address records	11	Emerick Rd	
3	1	SAM	Check address and remove some	15	Emerick Road	12477
3	2	SAM	Check address and remove some	11	Emerick Road	12477
4		RPTS (100% in block)	Compare this and next SAM record - keep 1		Emerick Rd	
4	1	SAM	Compare this and previous parcel record - keep 1		Emerick Road	12477
5		RPTS (100% in block)	Compare this and next SAM record - keep 1	23	Emerick Rd	
5	1	SAM	Compare this and previous parcel record - keep 1		Emerick Road	12477
6		RPTS (100% in block)	Reduce to 1 address records		Emerick Rd	
6	1	SAM	Check address and remove some		Emerick Road	12477
6	2	SAM	Check address and remove some		Emerick Road	12477
7		RPTS (100% in block)	Fill estimated 1 address records	25	Old Rt 9W	
8		RPTS (100% in block)	Fill estimated 1 address records	37	Emerick Rd	
9		RPTS (100% in block)	Fill estimated 1 address records	31	Emerick Rd	
10		RPTS (100% in block)	Reduce to 1 address records	29	Emerick Rd	
10	1	SAM	Check address and remove some	29	Emerick Road	12477
10	2	SAM	Check address and remove some		Emerick Road	12477

Let us go through each parcel property address to understand the appropriate action suggested in order to clean the data.

Example 1

parcellD	MAPSPOT	source	Suggested action	HOUSE NUMBER	STREET NAME	ZIP CODE
1		RPTS (3% in block)	Check presence of HU in parcel in block (0 estimated)	31	Meadow Ln	

For the parcel address above, there are 0 housing units estimated by the Real Property Data, as stated in the "Suggested Action" field. Since there is only 3% of the parcel

address located in the block, local knowledge would be needed to see if this address of 31 Meadow Lane exists and if its located in the block.

Example 2

parcelID	MAPSPOT	source	Suggested action	HOUSE NUMBER	STREET NAME	ZIP CODE
2		RPTS (100% in block)	Compare this and next SAM record - keep 1	36	Meadow Ln	
2	1	SAM	Compare this and previous parcel record - keep 1	36	Meadow Lane	12477

For the parcel address above at 36 Meadow Lane, there is 1 SAM point located in the parcel address—36 Meadow Lane. As a result, including both addresses would be a duplicate. Therefore, the suggested action for the first and second entry is to compare them to one another and to only keep 1. Since 100% of the parcel is located in the block, we can assume that this address is also within the block. Users should be more inclined to keep the SAM points when dealing with duplicates, because their mapspot is typically that of a rooftop while the RPTS latitude and longitude typically map to a parcel centroid.

Example 3

parcellD	MAPSPOT	source	Suggested action	HOUSE NUMBER	STREET NAME	ZIP CODE
3		RPTS (100% in block)	Reduce to 1 address records	11	Emerick Rd	
3	1	SAM	Check address and remove some	15	Emerick Road	12477
3	2	SAM	Check address and remove some	11	Emerick Road	12477

For the parcel address above at 11 Emerick Rd, there are 2 SAM points located within the parcel address: 15 Emerick Road and 11 Emerick Road. Since the RPTS data only estimates 1 housing address, the suggested action code with the parcel address is to "Reduce to 1 address records." Local knowledge is required to know if there are not actually two housing units located in the parcel address and the block. Here there could be a miscount between SAM and RPTS because, the RPTS data could be incorrectly predicting the number of housing units, there could be new development not reflected in the RPTS data, or perhaps a set of the SAM coordinates were digitized incorrectly.

Example 4

parcelID		MAPSPOT	source	Suggested action	HOUSE NUMBER	STREET NAME	ZIP CODE
	7		RPTS (100% in block)	Fill estimated 1 address records	25	Old Rt 9W	

For the parcel address above, there is an estimated 1 housing unit located inside. However, there are no other SAM points which share parcelID 7. As a result, the suggested action to complete the entry would be to fill out the addressing of the housing unit located within 25 Old Rt 9W.

Example 5

parcelID	MAPSPOT	source	Suggested action	HOUSE NUMBER	STREET NAME	ZIP CODE
9		RPTS (100% in block)	Fill estimated 1 address records		Emerick Rd	
10		RPTS (100% in block)	Reduce to 1 address records		Emerick Rd	
10	1	SAM	Check address and remove some		Emerick Road	12477
10	2	SAM	Check address and remove some	31	Emerick Road	12477

For the parcel addresses above, we can note for parcelID 9 with address 31 Emerick Road, the suggested action is to fill estimated 1 address record. However, for the subsequent parcelID 10 with address 29 Emerick Road, the suggested action is to reduce to 1 address record. Note, within the second parcel address, there is a SAM point of 31 Emerick Road, which may in fact coordinate with the parcelID 9 parcel address of 31 Emerick Road. Local knowledge may be required remediate whether the SAM point was incorrectly located. However, ultimately both property addresses of 29 Emerick Road and 31 Emerick Road should both be included in the final Excel List to be put into GUPS.

1.2 Cleaning the NYBBLES Data: Example B

Let us now turn our attention to the following Census Block which has a LUCA count of 71 Housing Units and 2 Group Quarters, but from the RPTS data, an estimated 76 housing

units. This Census Block is also of interest, because It includes apartment complexes which often will have several housing units located in one particular building or parcel address.



Locality	Ithaca city
Census 2010 Housing Unit count	25
LUCA Housing Unit count	71
LUCA Group Quarters count	2
RPTS based Housing unit count	76
Total number of SAM points Rooftops and structure entrances	86

Download addresses in Excel format

User Guide

Show on NYBBLES map Show in American Factfinder Show in Google maps Show in Bing maps Show in NYS Orthos Online

Parcels with RPTS based residential counts that overlap with this block (Click on parcel for more details)

- Parcel address (% overlap with this block) residential units in block

Parcel property class

4 - 115 Buffalo St W (100%) 6 - 113-13.5 Buffalo St W (100%) 2 - 204 Cayuga St N (100%)

50 - 100 Seneca St W (100%)

5 - 114 Seneca St W (100%)

4 - 116 Seneca St W (100%)

1 - 207 Geneva St N (100%)

2 - 211 Geneva St N (100%) 2 - 117 Buffalo St W (100%)

411 - Apartments

411 - Apartments

483 - Converted Residence

411 - Apartments

411 - Apartments 411 - Apartments

483 - Converted Residence

483 - Converted Residence

220 - Two Family Year-Round Residence

The Excel file for this particular block has 96 rows. Ultimately, it should have somewhere between 71 and 76 addresses as these correspond to the LUCA Housing Unit count and the RPTS estimated housing units numbers.

Let us look at three parcelIDs from the Census Block:

parcellD	MAPSPOT	source	Suggested	HOUSE	STREET NAME	APARTMENT/	ZIP CODE	LOCATION DESCRIPTION
1		RPTS (100% in block)	Reduce to 4 address records	115	Buffalo St W			Parcel centroid
1	1	SAM	Check address and remove some		West Buffalo Street	Apartment	14850	Rooftop

1	2	SAM	Check address and		West Buffalo	Apartment	14850	Rooftop
			remove some		Street	5		
1	3	SAM	Check address and		West Buffalo	Apartment	14850	Rooftop
			remove some		Street	2		
1	4	SAM	Check address and		West Buffalo	Apartment	14850	Rooftop
			remove some		Street	3		
1	5	SAM	Check address and		West Buffalo		14850	Rooftop
			remove some		Street			
2		RPTS	Reduce to 6 address	113-	Buffalo St W			Parcel
		(100% in	records	13.5				centroid
		block)						
2	1	SAM	Check address and		West Buffalo		14850	Rooftop
			remove some		Street			·
2	2	SAM	Check address and		West Buffalo		14850	Rooftop
			remove some		Street			
2	3	SAM	Check address and		West Buffalo	Apartment	14850	Rooftop
			remove some		Street	3		
2	4	SAM	Check address and		West Buffalo	Apartment	14850	Rooftop
			remove some		Street	1		
2	5	SAM	Check address and		West Buffalo	Apartment	14850	Rooftop
			remove some		Street	2		
2	6	SAM	Check address and		West Buffalo	Apartment	14850	Rooftop
			remove some		Street	4		
2	7	SAM	Check address and		West Buffalo	Apartment	14850	Rooftop
			remove some		Street	5		
3		RPTS	Fill estimated 2	204	Cayuga St N			Parcel
		(100% in	address records					centroid
		block)						
3	1	SAM	Check address (fill		North Cayuga		14850	Rooftop
			apartment/unit		Street			
			field) and keep					
			una Recp					

Example 1

parcelID	MAPSPOT	source	Suggested	HOUSE	STREET NAME	APARTMENT/	ZIP CODE	LOCATION DESCRIPTION
1		RPTS (100% in block)	Reduce to 4 address records	115	Buffalo St W			Parcel centroid
1	1	SAM	Check address and remove some		West Buffalo Street	Apartment	14850	Rooftop
1	2	SAM	Check address and remove some		West Buffalo Street	Apartment 5	14850	Rooftop
1	3	SAM	Check address and remove some		West Buffalo Street	Apartment 2	14850	Rooftop
1	4	SAM	Check address and remove some		West Buffalo Street	Apartment 3	14850	Rooftop
1	5	SAM	Check address and remove some		West Buffalo Street		14850	Rooftop

For the parcel address above at 115 Buffalo St SW, the parcel address has an estimated 4 housing units estimated from the RPTS as seen in the "Suggested Action" field of the Excel file, as well as the block browser. However, there are 5 SAM point addresses included in this Excel file. Surprisingly, there is an Apartment 5, which may argue that the RPTS data estimated incorrectly the number of housing units, or that the SAM data misread the apartment numbers. In any case, local knowledge is required to eliminate any duplicates and to make sure every entry is unique. For instance, MAPSPOT 5 and the first RPTS entry (which is a parcel centroid location) may be unnecessary and could be replaced with a 115 Buffalo St W, Apartment 1.

Example 2

_			1	ı		ı	1		
3	parceIID	MAPSPOT	source	Suggested action	HOUSE NUMBER	STREET NAME	APARTMENT/ UNIT #	ZIP CODE	LOCATION DESCRIPTION
	2		RPTS (100% in	Reduce to 6 address records	113- 13.5	Buffalo St W			Parcel centroid
			block)						
	2	1	SAM	Check address and		West Buffalo		14850	Rooftop
				remove some		Street			

2	2	SAM	Check address and	West Buffalo		14850	Rooftop
			remove some	Street			
2	3	SAM	Check address and	West Buffalo	Apartment	14850	Rooftop
			remove some	Street	3		
2	4	SAM	Check address and	West Buffalo	Apartment	14850	Rooftop
			remove some	Street	1		
2	5	SAM	Check address and	West Buffalo	Apartment	14850	Rooftop
			remove some	Street	2		
2	6	SAM	Check address and	West Buffalo	Apartment	14850	Rooftop
			remove some	Street	4		
2	7	SAM	Check address and	West Buffalo	Apartment	14850	Rooftop
			remove some	Street	5		

For the parcel address above, the RPTS data estimates 6 housing units. There are 7 SAM points associated with this parcel address, with 5 of them denoting a specific apartment. The other two points, MAPSPOT 1 and 2 should be investigated to see if they are in fact the 6th apartment or should be replaced with a better geo-location.

Example 3

parcellD	MAPSPOT	source	Suggested	HOUSE	STREET NAME	APARTMENT/	ZIP CODE	LOCATION
3		RPTS (100% in block)	Fill estimated 2 address records	204	Cayuga St N			Parcel centroid
3	1	SAM	Check address (fill apartment/unit field) and keep		North Cayuga Street		14850	Rooftop

For the parcel address, above, the RPTS data estimates 2 addresses, bt there is only 1 SAM point. However, because there are 2 address records located in a parcel with only 1 property, this is probably a two-family home. Therefore, the SAM point's suggested action of "check address (Fill apartment/unit field) and keep" should be followed. A resulting entry of could be desired (of course with corresponding latitude and longitude):

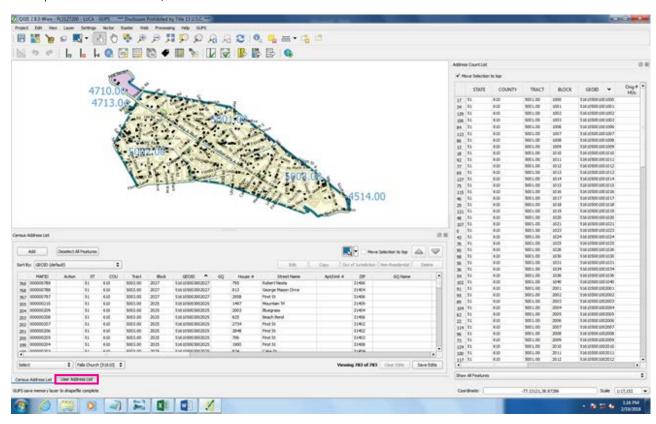
parcelID	MAPSPOT	source	Suggested	HOUSE	STREET NAME	APARTMENT/	ZIP CODE	LOCATION DESCRIPTION
----------	---------	--------	-----------	-------	-------------	------------	----------	----------------------

3	1		204	Cayuga St N	Apartment A		Rooftop
3	2		204	Cayuga St N	Apartment B	14850	Rooftop

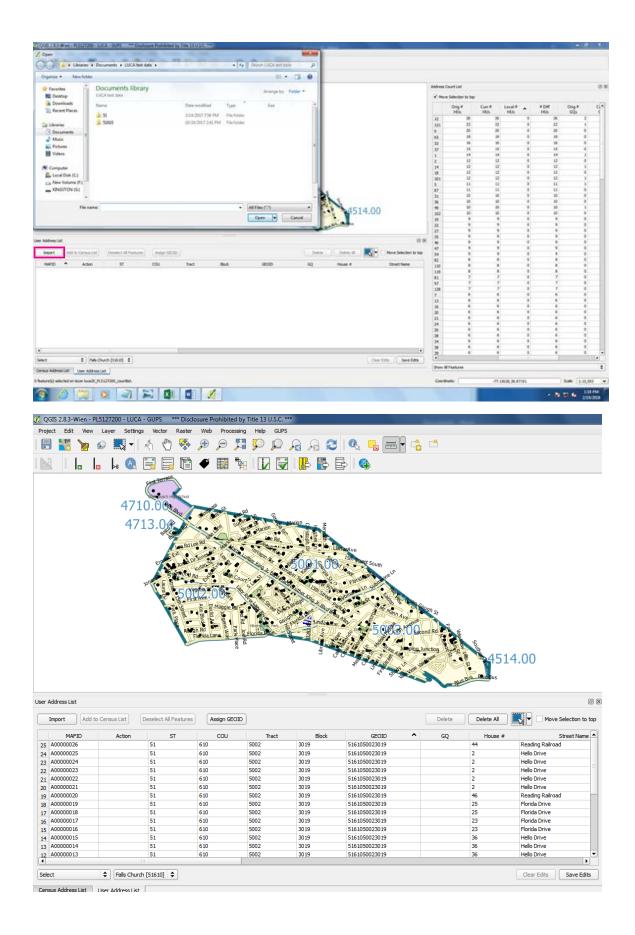
2. Importing Data into GUPS

2.1 Example 1

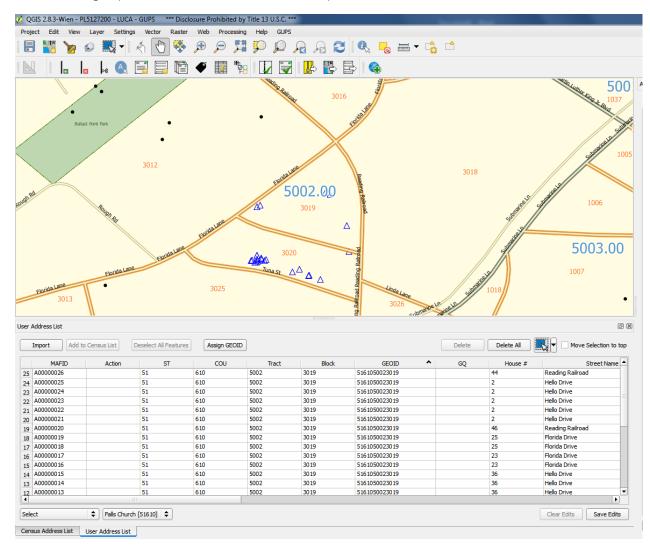
To import data into GUPS, click the "User Address List" tab found on the bottom of the screen.



Notice that compared to the Census Address List, the User Address List is empty until we find our csv file. I select the "Import" button and choose the correct CSV file with my cleaned NYBBLES data. The address list then populates.

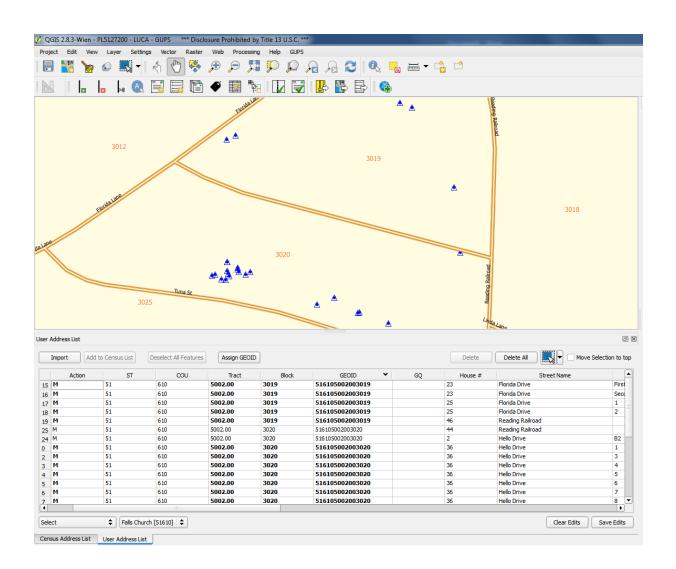


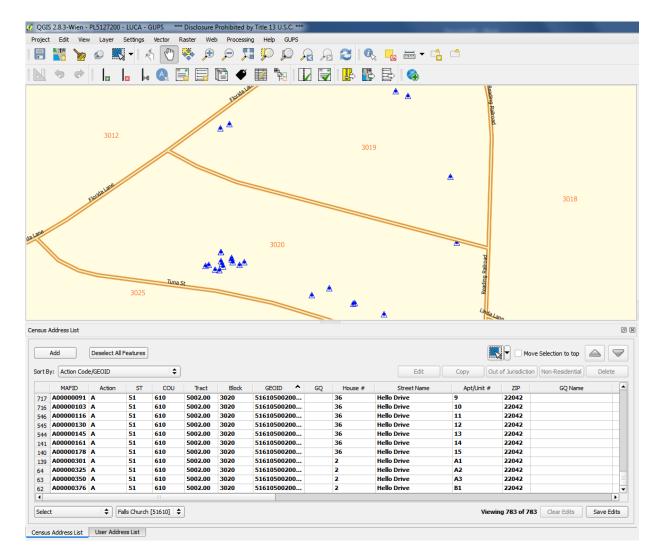
If I zoom into my block of interest, I notice that my fake data strides two different blocks. The User Address List addresses haven't been added to the Census list, because as we can see—there are hollow triangles. If the Census had these addresses, there would be a black dot intersecting the said triangle. (Notice the black dots in Block 3012).



There may have been an error in the NYBBLES data. GUPS can fix any geo-coding errors. By selecting an address line, I will then have an option to "Add to Census List." If for an entry the Block number I have listed is incorrect compared to what GUPS believes is correct, GUPS will fix this automatically and prompt me to select the "Save Edits" button. Or, if I am missing data or want to fix it before said prompt, I can select the "Assign GEOID" button. By selecting the "Assign GEOID" button, I am essentially giving GUPS permission to rewrite the State, County, Tract, Block, and GEOID information for my entry. There is a "C" on the action list for these items standing for "Corrected."

Once I add the addresses to the Census List, I can see that there is a blue dot inside the triangles on the map. Additionally on the User Address List tab there is an "M" on the action list for "Moved," while there is an "A" on the action list for "Added" on the Census List.

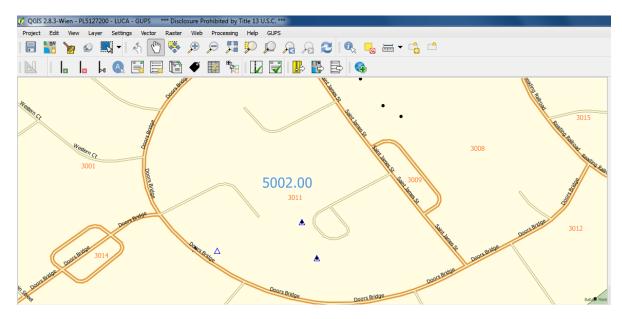




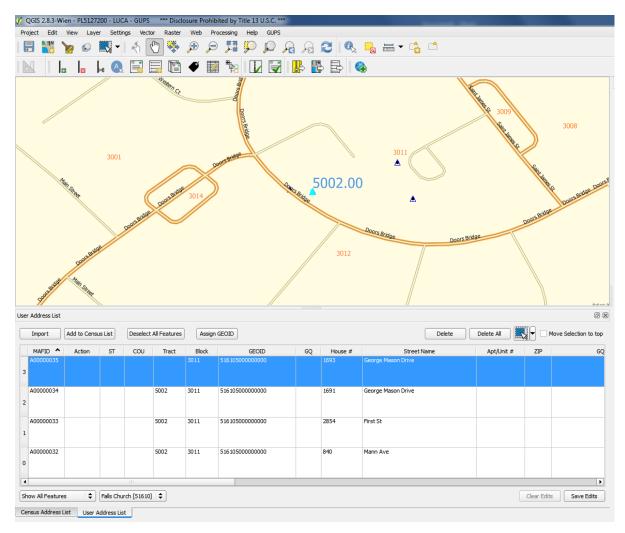
2.2 Example 2

After finishing with the block above, I wish to import data from another block from NYBBLES into GUPS. I have already cleaned the file and saved it as a csv file.

For this block, there are 4 housing units estimated in the block using RPTS data, but LUCA only has 3 housing units recorded. Once I import the csv file following the same procedure as above, I notice in the map above that for three of the address there is an intersected dot and triangle, while one triangle is hollow. This one hollow triangle is the housing unit that has not been recorded in LUCA.

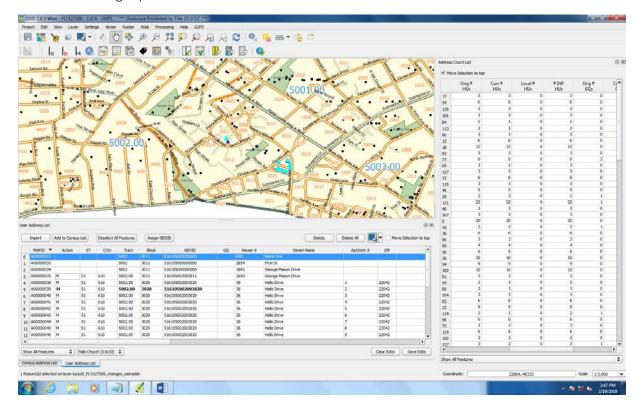


After selecting different entries on my User Address List, I find the entry for the missing housing address.



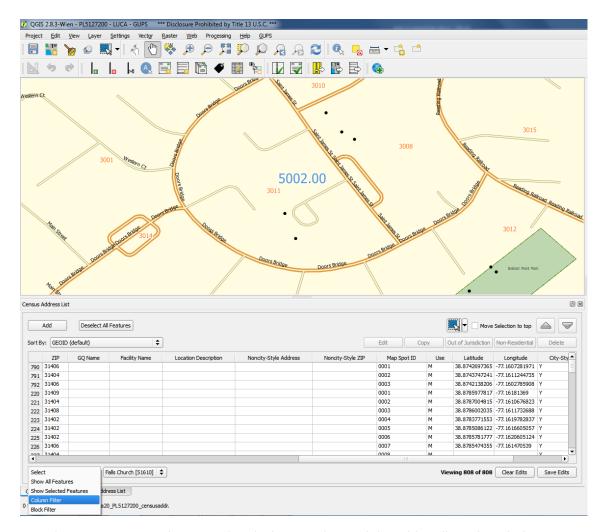
I then select "Add to Census List," and as before there is an "M" for the entry on the "User Address List" tab and "A" on the Census Address List.

We can see there is no action for the entries that that the Census Bureau already has, because I didn't choose those addresses to be added to the Census Bureau (i.e. the 3 housing units with the intersected triangles).

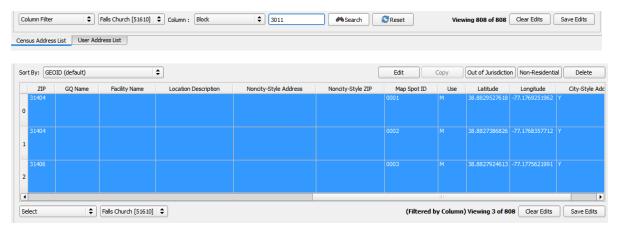


3. Downloading Data from GUPS

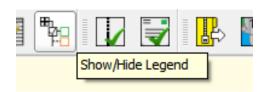
Let's say I was interested in downloading the Census data from Block 3011 from Section 2.2 before I imported my data. In order to download data from GUPS, I can select the "Census Address List" tab and under the "Select" drop-down menu choose "Column filter."

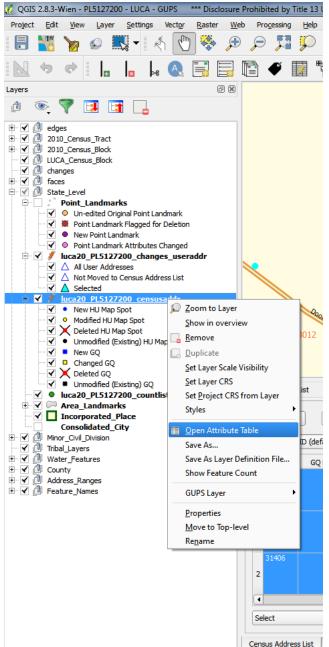


From here, I can type in the particular Block ID number and the table will result with the 3 housing unit entries. These should automatically be highlighted.



Pressing the "Show/Hide Legend" button, located in the top of the window reveals the legend.

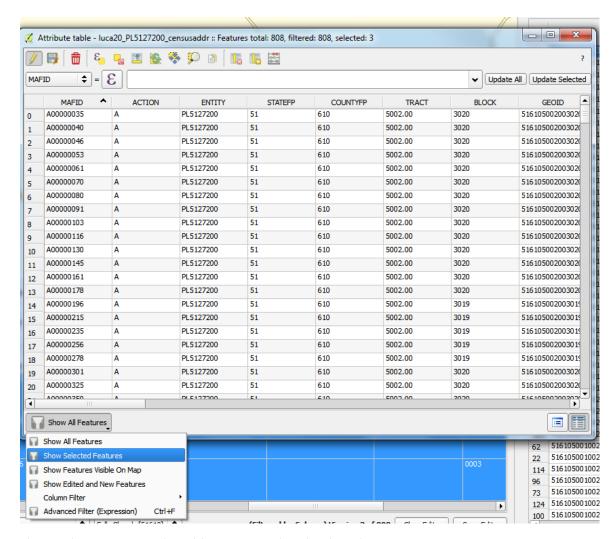




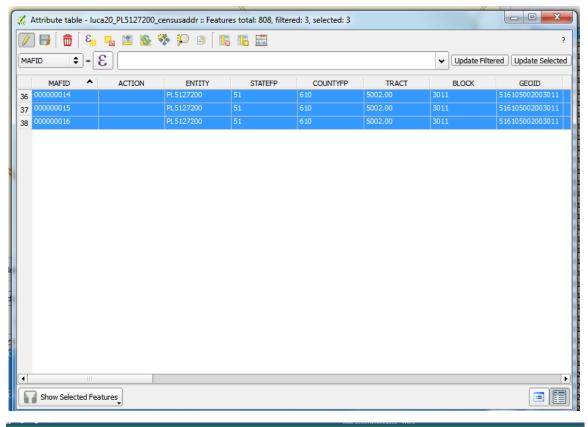
From here, a window will appear on the left-hand side of the GUPS browser. Make sure your list of addresses is still highlighted. In the layers window, right click on the corresponding census address list.

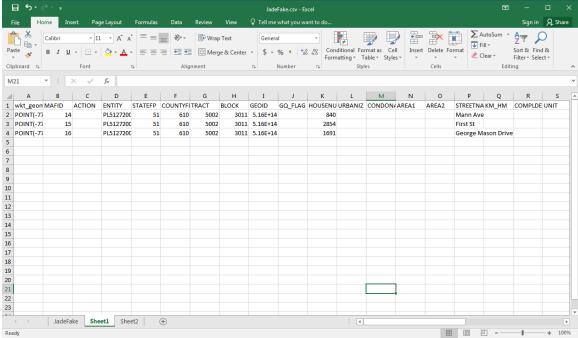
Note that there are two lists of data with identical icons, however one has "censusadd" in the name and other is "changes_useraddr." Right click on the former and select "Open Attribute Table."

Once the table pops up, in the bottom left corner, select the "Show All Features" drop down menu and choose "Show Selected Features."



The resulting entries in the table correspond to the three housing unit LUCA entries. To copy them into an Excel file, hold down the "CTRL" and "C" keyboard button simultaneously. Then paste into Excel.





The downloaded addresses can be combined with the NYBBLES data in Excel to check what NYBBLES addresses are duplicates of Census Addresses and what NYBBLES addresses need to be added.

4. Resources

NYBBLES Links

- NYBBLES Homepage: https://pad.human.cornell.edu/census2020/nybbles.cfm
- NYBBLES User Guide: http://pad.human.cornell.edu/census2020/docs/NYBBLES UserGuide.pdf
- Single address geocoders: https://pad.human.cornell.edu/tools/geocoder.cfm

Census & LUCA Related Links

- Census Glossary: https://www.census.gov/glossary
- Census LUCA Homepage: https://www.census.gov/programs-surveys/decennial-census/about/luca.html
- Census GUPS presentation: https://www2.census.gov/geo/pdfs/partnerships/luca/2020LUCA Training GUPS.pdf
- GUPS webinar by Geography Department of the New York Regional Census Center https://census.webex.com/census/lsr.php?RCID=9c74b5bab78b2b80212980e075c4ee76
- Census LUCA Frequently Asked Questions:
 https://www2.census.gov/geo/pdfs/partnerships/luca/2020LUCA_FAQ.pdf
- LUCA Frequently asked questions at NY State Data Center: https://labor.ny.gov/nys-data-center/luca-frequently-asked-questions.shtm
- LUCA training webinars organized by NY Council on Children and Families: http://ccf.ny.gov/council-initiatives/census-2020-get-prepared/luca-training-webinars/
- GUPS webinar by Geography Department of the New York Regional Census Center https://census.webex.com/census/lsr.php?RCID=9c74b5bab78b2b80212980e075c4ee76