



USPS Vacancy Data - Frequently Asked Questions

Why have the number of residential addresses (and total addresses) risen so dramatically since 2011?

There is a phenomenon that HUD has been studying that involves a sudden, sharp increase in the number of addresses in the USPS No-Stat category, which also manifests itself in an increase in the total number of addresses. The increase in addresses is over 7 million since 2011, with the bulk of those increases happening in two quarters – one in 2011 and the other in 2014. There is no evidence to be found in any of the available national administrative or survey data sources that would make that kind of increase in the number of residential housing units plausible.

HUD has raised the issue with USPS and they confirmed that the reason for these increases can be traced to a relatively new program called [Move to Competitive \(MTC\) Street Addressing for PO Boxes](#) which enables USPS customers to register their PO Boxes as a street address in order to receive packages and deliveries from private carriers who require a street address for delivery, such as UPS and FedEx. The implementation of this program has had the undesired effect of essentially bloating the number of addresses in the USPS's Address Management System (AMS). Each time a customer registers their PO Box with the MTC program a "new" address record is created, and that address is coded as a No-Stat. This has potential ramifications for our users depending on how they are applying the data. For example, if a user calculates a vacancy rate for residential addresses by dividing the number of vacant residential addresses by the total number of residential addresses, the rate might be artificially depressed because the denominator might include a rather large number of MTC addresses (i.e. PO Boxes that are not physical addresses but now show up in the inventory of total addresses). We know that some users are already subtracting the No-Stat address count from the total address count before calculating a vacancy rate, which would help to address this issue.

[Large Fluctuations in Addresses: White Paper on the Increases in the USPS No-Stat Category](#)

What happened in 2010 that caused an increase in the vacant address counts?

In March 2010, USPS began implementing new procedures to improve the accuracy of the vacant indicator. First, the USPS deployed a new application that allowed for delivery unit/carrier to answer yes or no (an address is either vacant or not) which then updates database automatically (nightly). Along with the application are data provided to the carrier from other data sources including the USPS Change of Address and USPS Address Change Service to assist the carrier in determining occupancy status. Second, Census 2000 mailings were also used as a "one time" source for the application. Because the Census mailing involved persistent mailing to non-respondent addresses, postal

carriers were better able to determine if an address was vacant or not. Carriers counted an address as vacant if they saw repeated Census mailings to that address, even if the address was not flagged as vacant for the preceding 90 days—which is the normal USPS protocol for determining vacant status. As a result, the vacancy data show an increase in the number of vacant addresses in 2010. The data improvements also resulted in a noticeable increase in the number of vacant addresses in the 0-3-month category in 2010.

What types of addresses are in the “Other” category?

The “Other” category was introduced into the USPS AMS to categorize what the USPS calls the “default” address of a structure that contains multiple units with the same street address but different unit numbers, such as most multifamily residential buildings. The default address may have many deliverable unit addresses associated with it, but the USPS created the “other” category to be able to isolate the default front door address and mark it as a No-Stat. The default address of a building will have its own ZIP+4 record. Theoretically, that ZIP+4 record should always have a total AMS count of ‘1’ and that 1 address should be categorized as a No-Stat. However, there are anomalies in the system where this is clearly not the case and USPS are unable to explain what they are. HUD has no other guidance regarding the Other category to provide to users of the Vacancy Data.

How closely does the number of residential addresses in the USPS data track with the number of housing units reported by the Census Bureau?

HUD compared the 2006 ACS housing unit count (a special tabulation from Census) and residential USPS address counts at the tract level (including No-Stats) and found that, in general, the USPS count was consistently lower in non-metro counties than the Census’ housing unit count. Meanwhile, in metro counties, the USPS count was consistently higher, but closer to the Census count than in the non-metro counties. In one county, for which HUD had excellent local data that included zoning designations, HUD compared the county’s residential address count with both USPS and Census and found that the Census count was consistently lower than the county’s count at the census tract level, while the USPS count was consistently higher, but closer to what the county was reporting.

I’m observing an implausibly large fluctuation in the number of addresses in a census tract over time. What’s going on here?

HUD recently discovered that the geocoding process that assigns a raw ZIP+4 record to the correct Census tract has been producing some anomalies in the data over time. First, HUD is not able to determine a tract code for every ZIP+4. The clear majority (over 99%) of records are successfully geocoded, but sometimes a particular ZIP+4 is not locatable by HUD’s geocoding engine on first try. However, the geocoder’s base data are updated regularly and some ZIP+4 records that did not geocode in one quarter will geocode in the next and get added to the mix. As a result, you may see spikes in the total address count of a tract that cannot necessarily be attributed to growth since the previous quarter.

Somewhat related to this problem, is the issue of changing ZIP Code boundaries. Sometimes the USPS will split a ZIP code, which results in new ZIP+4 records being created or reassigned. This can also manifest itself in weird spikes or drops in total addresses from quarter to quarter that cannot necessarily be attributed to an actual growth or decline in the number of addresses. When a ZIP code is split, our geocoding takes a while to catch up. The USPS publishes every year a product that identifies the major ZIP Code boundary adjustments that were made. You can download it at <http://ribbs.usps.gov/>. Refer to this document if you suspect that this problem may be affecting your area of analysis.

Should I include No-Stats when calculating a vacancy rate?

It can be very difficult to calculate a vacancy rate because technically some No-Stats are vacant too. No-Stats are addresses that have been, more or less, abandoned or addresses that are under construction and are not yet ready to be occupied. There is no real way to tell the two apart in the dataset. That's why it helps to have local intimate knowledge of the area you are studying. This can help you decide on whether to include No-Stats as part of the vacancy calculation. In lieu of local knowledge, you might consider using other neighborhood characteristics to get an idea of what kind of neighborhood it is. For example, if you are dealing with a high poverty neighborhood, there is a good chance that many of the No-Stats are vacant or abandoned structures. Another indication may come from some other fields in the dataset. For example, you might look at the No-Stat 12 to 24-month count. If you see increasing numbers of addresses that are staying in the No-Stat category over time, it might signify a neighborhood in decline. In another example, you might look at AMS (total addresses). If AMS is growing and you see a similar growth in the No-Stat 0 to 3-month category, it is possible that there is new construction occurring. But if AMS is holding steady or declining and you see increase in No-Stat, then that could be a sign of a neighborhood in trouble. A straight up decline in AMS may suggest demolition.

Will the business/residential/other distinction be made available for all extracts made prior to June 30, 2008?

No.

Are P.O. Boxes counted as addresses?

In urban areas, addresses whose residents collect mail at P.O. Boxes instead of through regular delivery, are counted as No-Stats, even if there may be someone living there. Addresses on rural routes that use P.O. Boxes instead of delivery are not counted as addresses.

Can I obtain the non-aggregated version of the data or a level of aggregation finer than Census Tract?

No. The ability to provide vacancy and No-Stat information to HUD was thoroughly reviewed by the USPS Privacy and Legal departments in coordination

with Address Management. This review determined that the information could be provided only if use was aggregate to a Census Tract level. Aggregates at a Census Tract level virtually eliminate the chance of informational disclosure of individual addresses and their occupancy status.

Why am I getting a Microsoft Jet database engine error when I try to import the DBF file into Access?

For some reason, Access only recognizes a DBF file if the name of the file is 8 characters or less. If you want to import the DBF files into Access, then rename them using 8 characters or less.

Why are there two separate lists of data files – one for 2010 Census Tracts and one for 2000? And why aren't the same quarters of data available for both sets of data files?

When HUD originally received the USPS Vacancy Data in 2005, the 2010 Census Tract universe had not yet been created. It wasn't until the first quarter of 2013 that HUD developed the capability to geocode the raw USPS ZIP+4 records to the correct 2010 Census Tract. Due to the costs associated with managing multiple versions of a fully geocoded ZIP+4 universe, HUD decided not to continue tabulating the USPS Vacancy Data at the 2000 Census Tract level after the last quarter of 2012. That's why the two lists of quarterly files, one for 2010 and another for 2000 tracts, are mutually exclusive. HUD warns users to take care in comparing 2000 and 2010 tract summaries of the USPS Vacancy Data because tract boundaries have changed significantly between the two vintages. The Census provides a [relationship file between 2000 and 2010 Census tracts](#) for advanced users.

Why is the Quarter 4 Ending December 31, 2005 file missing data in most of the vacancy and No-Stat columns?

The starting point for counting the number of days that an address is in No-Stat or vacancy status is November 18, 2005. Therefore, all of the No-Stat and vacancy counts in the December 2005 tract-level summary file fit into the 0 to 3 month No-Stat and vacant categories. Because the USPS started counting in November there is no data for the remaining vacancy and No-Stat categories in the December 2005 file.

What is a GEOID and how do I decode it to understand what state, county or Census Tract the data are relevant to?

The GEOID field in each tract-level file contains the concatenation of the state FIPS code, the county FIPS code and the Census Tract code. By querying this field you can pick out states, counties, or specific tracts within a certain county. If you are unsure what your county or state FIPS code is, go to FIPS County Lookup page on the U.S. Census Bureau's website at <https://www.census.gov/geo/reference/codes/cou.html>. So, for example, if you would like to select the Census tracts in Bergen County, NJ you would construct a query that selects all the records where GEOID begins with '34003'

because the state FIPS code for New Jersey is '34' and the county FIPS code for Bergen is '003'.

How do I read the Census Tract-level summary files?

The Census Tract-level summary files are downloaded as zipped DBF files. Once you've unzipped them, you can use a variety of programs to read the files. Excel is the most popular reader for DBF files. There are many other software packages that read DBF files, such as SPSS, SAS, or Access. Most GIS software will also read in DBF files.