

## Data Shop

*Data Shop, a department of Cityscape, presents short articles or notes on the uses of data in housing and urban research. Through this department, the Office of Policy Development and Research introduces readers to new and overlooked data sources and to improved techniques in using well-known data. The emphasis is on sources and methods that analysts can use in their own work. Researchers often run into knotty data problems involving data interpretation or manipulation that must be solved before a project can proceed, but they seldom get to focus in detail on the solutions to such problems. If you have an idea for an applied, data-centric note of no more than 3,000 words, please send a one-paragraph abstract to [chalita.d.brandly@hud.gov](mailto:chalita.d.brandly@hud.gov) for consideration.*

---

# New Data Fields for HUD Aggregated USPS Administrative Data on Address Vacancies

Alexander Din

U.S. Department of Housing and Urban Development

*The views expressed in this article are those of the author and do not represent the official positions or policies of the Office of Policy Development and Research, the U.S. Department of Housing and Urban Development, or the U.S. Government.*

---

## Abstract

Since 2005, the United States Department of Housing and Urban Development (HUD) has worked in partnership with the United States Postal Service (USPS) to receive administrative data on address vacancies. HUD has made that data available to government entities and nonprofit researchers. Since 2012, HUD has received more than 3,100 requests for access to the data. In the most recent agreement between HUD and USPS, new fields have become available regarding (1) the USPS preferred name and preferred state for a ZIP Code, (2) the count of addresses added to the USPS Address Management System (AMS) during the quarter, and (3) drop counts for entities such as mobile home communities and gated communities where mail is delivered to a single recipient but no data are collected for the addresses using that node. The purpose of acquiring those extra data was to better understand address vacancy and neighborhood change. It is expected that these new data fields will continue to be available for future datasets.

## Background

The Department of Housing and Urban Development (HUD) has been acquiring administrative data on address vacancies from the United States Postal Service (USPS) since 2005. HUD makes the administrative data on address vacancies available to government entities and nonprofit researchers.<sup>1</sup> Since 2012, more than 3,100 requests for the data have been reviewed. Data have been provided to local and state governments, federal agencies, universities, research institutions, and nonprofit organizations.<sup>2</sup> Requests for the data peaked in 2019, with more than 400 requests.

The interagency agreement between HUD and USPS provides HUD with ZIP+4 data regarding a large variety of variables describing the addresses at that particular point. The 2021 Quarter 1 (2021 Q1) dataset had more than 36,000,000 records. A ZIP+4 is a very granular level of geography and can be thought of as a cul-de-sac, a block of rowhomes, or a single floor of an apartment building.<sup>3</sup> As letter carriers deliver mail, they also collect information regarding whether an address is a residence, business, or other and whether the address is occupied and collecting the mail. To protect potentially identifying information and in accordance with the interagency agreement, data provided to researchers are aggregated to census tracts; HUD does not provide researchers with the raw ZIP+4 data. The data are timely because data for the previous quarter are typically released within a month of that quarter ending.

The data are in demand. In addition to the several thousand requests for access to the data, a search of Google Scholar for “usps hud address vacancies” produces nearly 800 results. That many results may be modest to some but, overall, indicates that the dataset is being referenced. Determining the exact number of citations presents some difficulty because of the variance in citations. Typically, the dataset has been used for analysis of blight and vacancy but also as a component for various spatial modeling analyses.

Beginning in 2021 Q1, the following data points have become available. These data points represent several new fields each:

1. The USPS Preferred City Name and USPS Preferred State Name.
2. The number of addresses served by a drop stop, a place where the USPS letter carrier drops off mail, which is then delivered to its final destination by someone else, such as at a group quarters facility or gated community.
3. The number of addresses new to the Address Management System (AMS) over the previous quarter.

---

<sup>1</sup> The website can be accessed here: <https://www.huduser.gov/portal/datasets/usps.html>.

<sup>2</sup> Requests for data were only available dating back to 2012.

<sup>3</sup> For example, the author once lived in an apartment building that had not only a ZIP+4 for each floor but a ZIP+4 for each third of each floor. The building had 53 ZIP+4 designations in total, 51 for each section of each floor, one for the ground floor retail area, and one for the leasing office.

The new fields to the HUD-USPS administrative data on address vacancy were added to enable researchers to understand place names, neighborhood characteristics, neighborhood change, and vacancy.

## **New Fields**

### **Preferred City and State**

The Zone Improvement Plan (ZIP) Code was launched in 1963 to sort mail more efficiently (USPS OIG 2013). ZIP Codes are five-digit codes that describe where a piece of mail is to be sent. The first three characters describe the group of states and sectional center facility (SCF)—a USPS mail sorting center—where the letter is to go. From the SCF, the fourth and fifth digits of the ZIP Code further describe the area to which to send mail. ZIP Codes were designed for mail delivery and do not adhere to political boundaries. As a result, ZIP Codes frequently do not align with political or administrative boundaries. Furthermore, not all places have their own ZIP Codes; some communities do not have their own post office and are instead served by a facility in a neighboring jurisdiction. The result is that the community may not use its name on addresses but instead uses the community where the post office is located.

The use of ZIP Codes has grown to surpass delivering mail. ZIP Codes are frequently used to describe areas, to engage with populations, for marketing, and to perform spatial analyses. Considerable research has examined the errors of using ZIP Codes as units of analysis in geospatial health research (Beyer, Schultz, and Rushton, 2007; Cudnick et al., 2012; Grubestic and Matisziw, 2006; Krieger et al., 2002; Sadler, 2019). In addition to those issues, the Congressional Research Service (CRS) cited the following six common problems caused by the misalignment between ZIP Code boundaries and municipal boundaries (Ginsberg, 2011):

- Higher automobile insurance rates for drivers who live in the suburbs but are charged city rates on the basis of their ZIP Codes.
- Residents who are confused about where to vote in municipal elections because they do not distinguish between their voting and mailing addresses.
- Sales tax revenues rebated by states to the cities where they are collected often being misdirected because they are collected by merchants with ZIP Codes in different jurisdictions or by merchants who mail their products to customers knowing only their ZIP Codes.
- Individuals being sent jury duty notices when they are not eligible to serve on the basis of their residences.
- Emergency service vehicles being misdirected by confusion over what town a call has come from on the basis of mailing address information.
- Homeowners in expensive neighborhoods complaining that their housing values are diminished because their mailing addresses place them in less prestigious communities.

Due to those problems, USPS has received inquiries about whether new ZIP Codes can be created, if addresses can be added to or removed from certain ZIP Codes, and if ZIP Code names can be changed. CRS indicated that USPS believes many of those problems were due to community identity issues. Rather than issuing new ZIP Codes, which would require considerable investment and financial resources, USPS recognizes multiple names for a ZIP Code but also recommends that some names should not be used.

Beginning in 2021 Q1, data acquired by HUD from USPS will have the USPS Preferred City Name, which is not the only recognized name but the name USPS primarily recognizes. USPS also keeps records of names to avoid; however, only USPS preferred name is provided, not the other recognized names or the names to avoid.

Exhibit 1 shows Mount Airy, Maryland, a small community approximately 30 miles west of Baltimore City, Maryland, and 35 miles northwest of Washington, DC, using both the Census Designated Place boundaries (Mount Airy CDP) and the ZIP+4 centroids for the 21771 ZIP Code, which has the preferred name Mount Airy<sup>4</sup> (Mount Airy ZIP). The Mount Airy CDP has a far smaller area, about 4.1 square miles, whereas the minimum bounding geometry<sup>5</sup> for the Mount Airy ZIP is nearly 20 times larger, about 96.6 square miles. By area, the Mount Airy CDP is split between Frederick County and Carroll County; approximately 34.3 percent and 65.7 percent of the land area are in each county, respectively. The Mount Airy ZIP Code expands far outside the Mount Airy CDP; however, only about 35.6 percent of its addresses are within the Mount Airy CDP. The map in exhibit 1 shows that the Mount Airy ZIP extends beyond Frederick and Carroll Counties into western Howard and far northern Montgomery counties. Exhibit 2 lists the ratios for the Mount Airy ZIP.

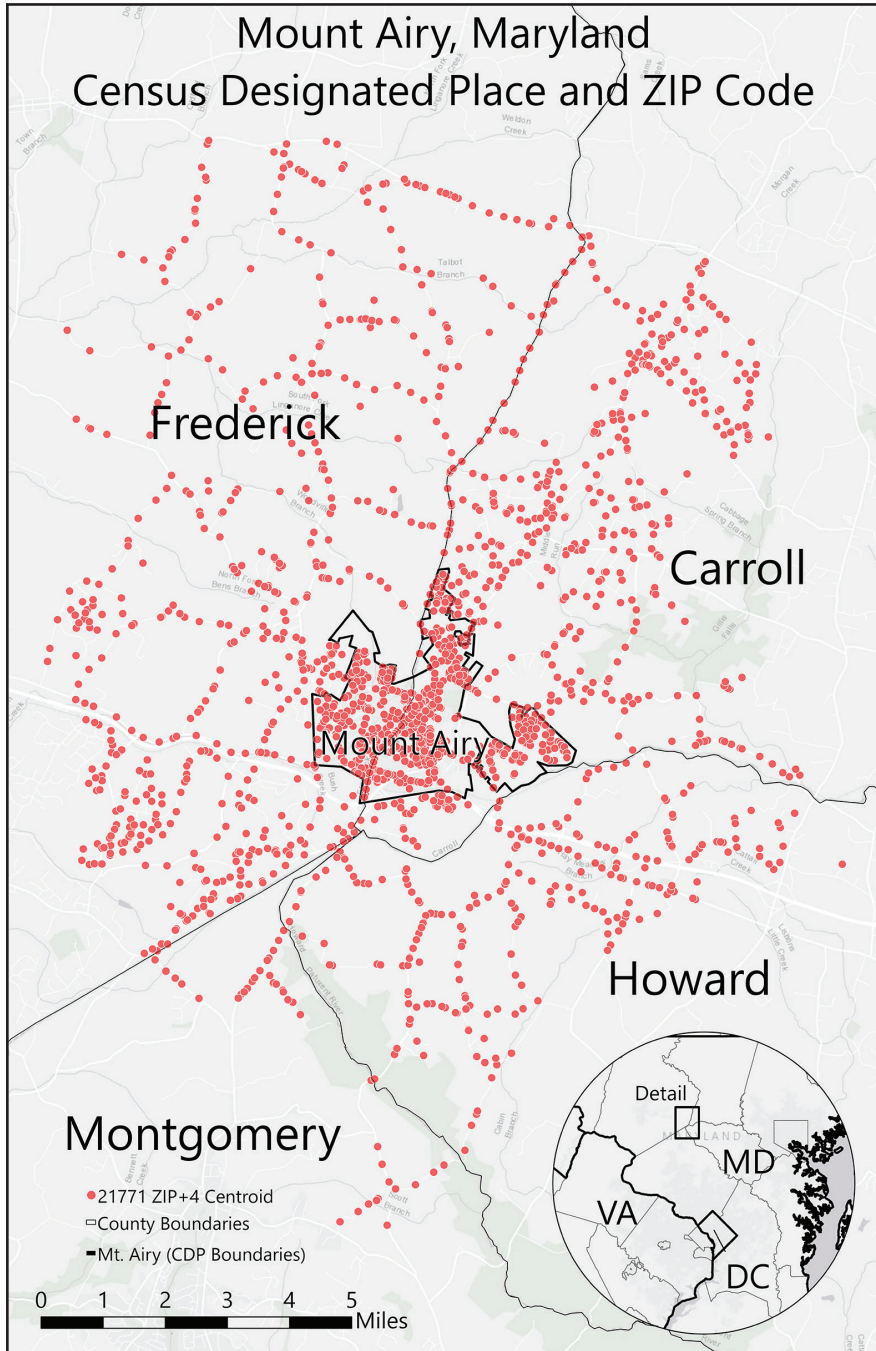
---

<sup>4</sup> The 21771 ZIP Code does not have any other recognized names or names to avoid.

<sup>5</sup> The convex hull method for the minimum bounding geometry tool was used to calculate the area of the Mount Airy ZIP Code: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/minimum-bounding-geometry.htm>.

**Exhibit 1**

**Mount Airy, Maryland**



CDP = Census Designated Place.

Source: Compiled by author based on 2021 Quarter 1 HUD Aggregated USPS Administrative Data on Address Vacancies and Census Bureau data

**Exhibit 2**

Ratios for the Mount Airy, Maryland, ZIP Code

ZIP	COUNTY	USPS_PREFERRED_CITY	USPS_PREFERRED_STATE	RES_RATIO	BUS_RATIO	OTH_RATIO	TOT_RATIO	COUNTY_NAME
21711	24021	MOUNT AIRY	MD	0.420422418	0.149490374	0.221428571	0.399954019	Frederick
21711	24031	MOUNT AIRY	MD	0.0114751	0.00339751	0.0	0.0108054	Montgomery
21711	24027	MOUNT AIRY	MD	0.107018127	0.0351076	0.0214286	0.101233811	Howard
21711	24013	MOUNT AIRY	MD	0.461084317	0.81200453	0.757142857	0.488006744	Carroll

*Note: The county name field was added for clarity.*

*Source: 2021 Quarter 1 HUD-USPS ZIP Code ZIP-to-County Crosswalk File*

Understanding geographic differences between USPS-defined places and CDPs is important for researchers, spatial analysts, policymakers, and others so that when conducting analysis or creating policy, they understand the area for which the analysis is being performed or for which the policy is written. The new fields describing USPS-defined places will better inform researchers and practitioners about geography.

The new data are outlined by new field descriptions in the documentation tab<sup>6</sup> on the web page. The documentation is as follows:

**Exhibit 3**

**USPS\_ZIP\_PREF\_CITY –USPS preferred city name**

**USPS\_ZIP\_PREF\_STATE –USPS preferred address state**

*Note: ZIP Code preferred city names frequently do not align with administrative names; for more information, please see USPS City Versus Census Geography.<sup>7</sup>*

*Source: HUD-USPS ZIP Code Crosswalk Files Codebook*

**Drop Counts**

The number of addresses served by a single drop stop were also added to the data (hereafter, addresses served by a drop stop are referred to as “drop counts” and the points are referred to as “drop stops”). They are not addresses dropped from AMS. Counts of addresses served by a drop stop are available by residential, business, or other addresses but are not broken out by vacant or no-stat and their child categories. Addresses associated with a drop stop are addresses where multiple residences are served by a single drop stop. These drop counts may include, but are not limited to, apartment buildings, care facilities, group quarters, and other facilities where multiple people live and mail is delivered to a single node. Not all nor most apartment buildings or other group facilities would necessarily be served by a drop stop.

<sup>6</sup> Access the web page here: [https://www.huduser.gov/portal/datasets/usps\\_crosswalk.html#codebook](https://www.huduser.gov/portal/datasets/usps_crosswalk.html#codebook).

<sup>7</sup> That page can be found here: [https://www.census.gov/programs-surveys/geography/guidance/geo-areas/usps\\_census\\_city.html](https://www.census.gov/programs-surveys/geography/guidance/geo-areas/usps_census_city.html).

The USPS introduced three new fields for drop counts, which are broken down by address type, such as residential, business, or other. Data for vacant or no-stat drop count addresses are not available. Exhibit 4 shows the breakdown of the number of total addresses in each category, the number of addresses in that category served by a drop stop, and the share served by a drop stop. Nationally, slightly more than one percent and less than seven percent of residential and business addresses, respectively, are served by a drop stop; no addresses in the Other category are served by a drop stop.

**Exhibit 4**

Address Type	Number of Addresses	Addresses Served by Drop Stop	Addresses Served by Drop Stop (Share)
Residential	154,038,372	2,072,937	1.34%
Business	14,097,183	967,247	6.86%
Other	7,189,491	0	0.00%

Source: 2021 Quarter 1 HUD Aggregated USPS Administrative Data on Address Vacancies

**Geographic Variation of Drop Counts**

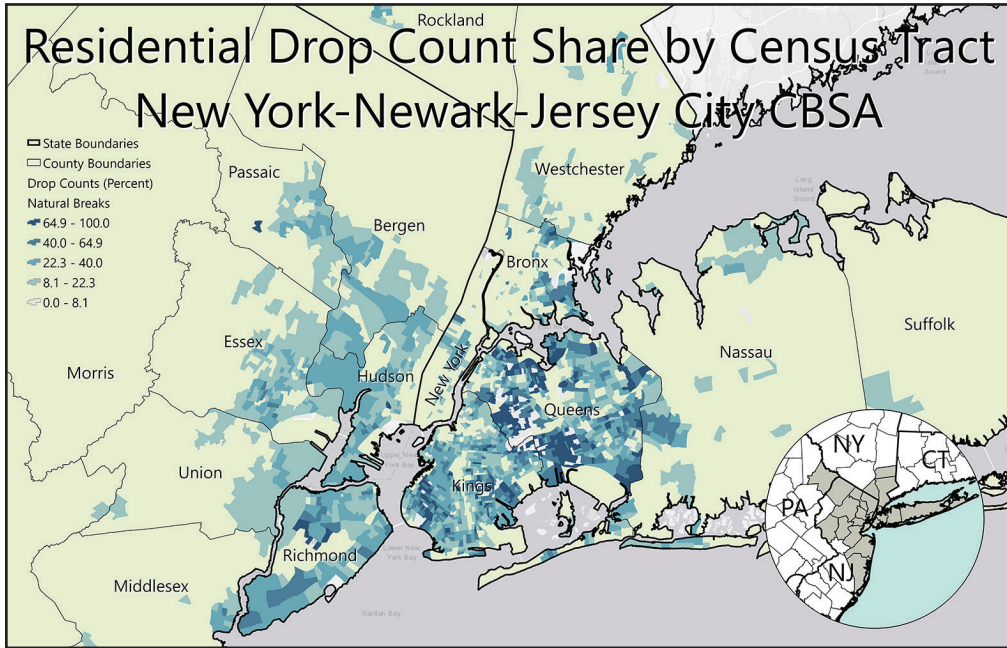
Drop counts display geographic disparities. Although 1.34 percent of residential addresses nationally are served by a drop stop, variation exists depending on the location. For example, of the 939 Core-Based Statistical Areas (CBSAs), which are made up of metropolitan and micropolitan counties, the overwhelming majority of CBSAs (834, or 88.8 percent) have less than the 1.34-percent benchmark of residential addresses served by drop stops. Of the remaining 105 CBSAs, 97 have between 1.34 percent and 10.0 percent of residential addresses served by a drop stop. Seven of the remaining eight CBSAs have 10.1 percent through 16.1 percent of addresses served by drop stops. The Athens, Ohio CBSA, a micropolitan area of roughly 65,000 people composed solely of Athens County, Ohio, had nearly one-half (47.2 percent) of addresses served by drop stops. This geographic variation implies that although drop stops are described to serve particular address types, either they are not recorded as such in AMS or this practice is not common throughout the United States.

Exhibit 5 visualizes the share of residential addresses served by a drop stop in the New York-Newark-Jersey City CBSA (NYC CBSA), the 23-county area around and including the five counties of New York City. The map focuses on the core counties of the NYC CBSA because many of the outlying counties have few to no residential addresses served by a drop stop. The majority of the census tracts served by drop stops appear to be in New York City but not in Manhattan.<sup>8</sup> Manhattan, or New York County—the densest county in the United States (U.S. Census Bureau, 2018)—has relatively few neighbors served by drop stops. Numerous neighborhoods throughout Staten Island, Brooklyn, and Queens use drop stops, and some, but fewer, drop stops are in the Bronx. Some neighborhoods in the suburban counties have drop stops, particularly in New Jersey in Bergen, Essex, Passaic, and Union Counties. A small number of neighborhoods in suburban New York counties, such as Nassau and Westchester Counties, are served by drop counts.

<sup>8</sup> Manhattan is shown as New York County on the map. Kings County and Richmond County are more commonly referred to as Brooklyn and Staten Island, respectively.

**Exhibit 5**

Residential Drop Count Share by Census Tract in New York-Newark-Jersey City CBSA



CBSA = core-based statistical area.

Source: 2021 Quarter 1 HUD Aggregated USPS Administrative Data on Address Vacancies

As shown in exhibit 6, the data dictionary for HUD Aggregated USPS Administrative Data on Address Vacancies has been updated for the three new fields.

**Exhibit 6**

New Drop Count Field Descriptions

Field	Description
DROP_ADD_R	Count of residences served at a drop site. A drop site is defined as single delivery serving multiple residences/households.
DROP_ADD_B	Count of the businesses served at a drop site. A drop site is defined as single delivery serving multiple businesses.
DROP_ADD_O	Count of the other addresses served at a drop site. A drop site is defined as single delivery serving multiple other addresses.

Source: HUD Aggregated USPS Administrative Data on Address Vacancies Data Dictionary<sup>9</sup>

**New Addresses**

The HUD Aggregated USPS Administrative Data on Address Vacancies will now include the number of addresses by type that were identified as new to AMS in that quarter. Those fields are a

<sup>9</sup> The Data Dictionary for HUD Aggregated USPS Administrative Data Since Quarter 4, 2020, can be found at [https://www.huduser.gov/portal/datasets/usps/USPS\\_HUD\\_Address\\_Vacancy\\_Data\\_Dictionaries.xlsx](https://www.huduser.gov/portal/datasets/usps/USPS_HUD_Address_Vacancy_Data_Dictionaries.xlsx).



count of addresses that are new to AMS, not a comparison of change. Of the 73,470 census tracts in the vacancy data, 48,294 (65.7 percent) recorded zero new residential addresses, and 9,899 census tracts (13.5 percent) recorded only one new residential address. The remaining 15,277 tracts (20.8 percent) contain the vast majority of the new residential addresses (79,468 of 89,367 new residential addresses). Exhibit 7 shows the total number of addresses and the number of addresses new to AMS by address type. Addresses new to AMS constitute a very small component of total residential or business addresses—0.06 and 0.07 percent, respectively. Exhibit 8 shows the new address count fields in the data dictionary. Anecdotally, USPS expressed that when updating the AMS, addresses are typically added after they have become active.

### Exhibit 7

Addresses Identified as New to the Address Management System (AMS) 2021 Quarter 1

Address Type	Addresses	Addresses New to AMS	Addresses New to AMS (Share)
Residential	154,038,372	89,367	0.06%
Business	14,097,183	10,520	0.07%
Other	7,189,491	0	0.00%

Source: 2021 Quarter 1 HUD Aggregated USPS Administrative Data on Address Vacancies

### Exhibit 8

New Count Field Descriptions

Field	Description
NEW_ADDR_R	Number of new residential addresses added to the AMS database in the previous quarter (occupied, no-stat, or vacant).
NEW_ADDR_B	Number of new business addresses added to the AMS database in the previous quarter (occupied, no-stat, or vacant).
NEW_ADDR_O	Number of new other addresses added to the AMS database in the previous quarter (occupied, no-stat, or vacant).

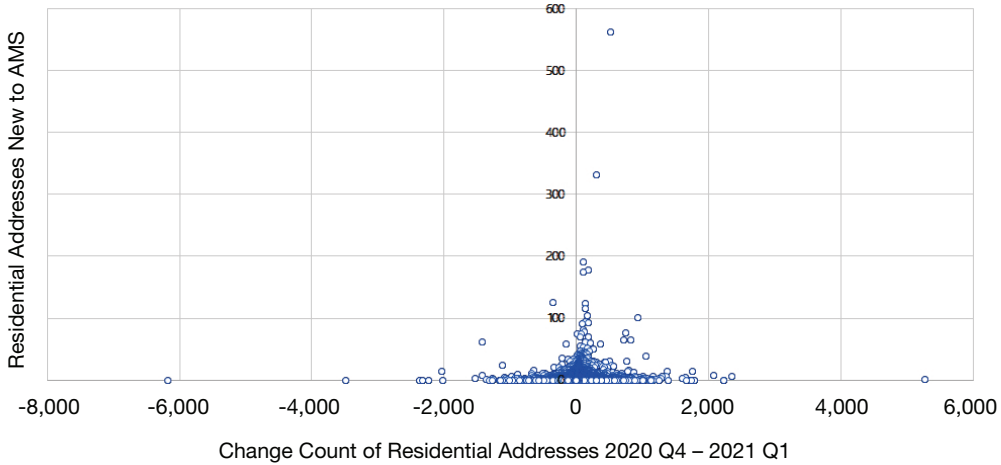
Source: HUD Aggregated USPS Administrative Data on Address Vacancies Data Dictionary

Comparing the count of total residential addresses in 2020 Quarter 4 (2020 Q4) with 2021 Q1 reveals 445,607 new residential addresses, with total residential addresses each quarter of 153,592,765 and 154,038,372, respectively. This discrepancy is just under five times the reported number of addresses new to AMS and includes tracts that have fewer residential addresses. Of the 73,470 census tracts in the dataset, 17,161 tracts (23.3 percent) had a decrease in the number of residential addresses, 23,587 tracts (32.1 percent) had no change in residential addresses, and 32,716 tracts (44.5 percent) had an increase in residential addresses.

Performing a linear regression between the calculated difference in residential addresses from 2020 Q4 to 2021 Q1 and the reported residential addresses new to AMS in 2021 Q1 reveals little information. Although the p-value is very small at 1.2533e-104, the slope is 0.0039 and the r-value is 0.08, suggesting that the relationship is quite insignificant and not meaningful. Exhibit 9 visualizes this weak relationship.

**Exhibit 9**

Comparison Between Count of Residential Addresses and Residential Addresses New to AMS, 2021 Q1



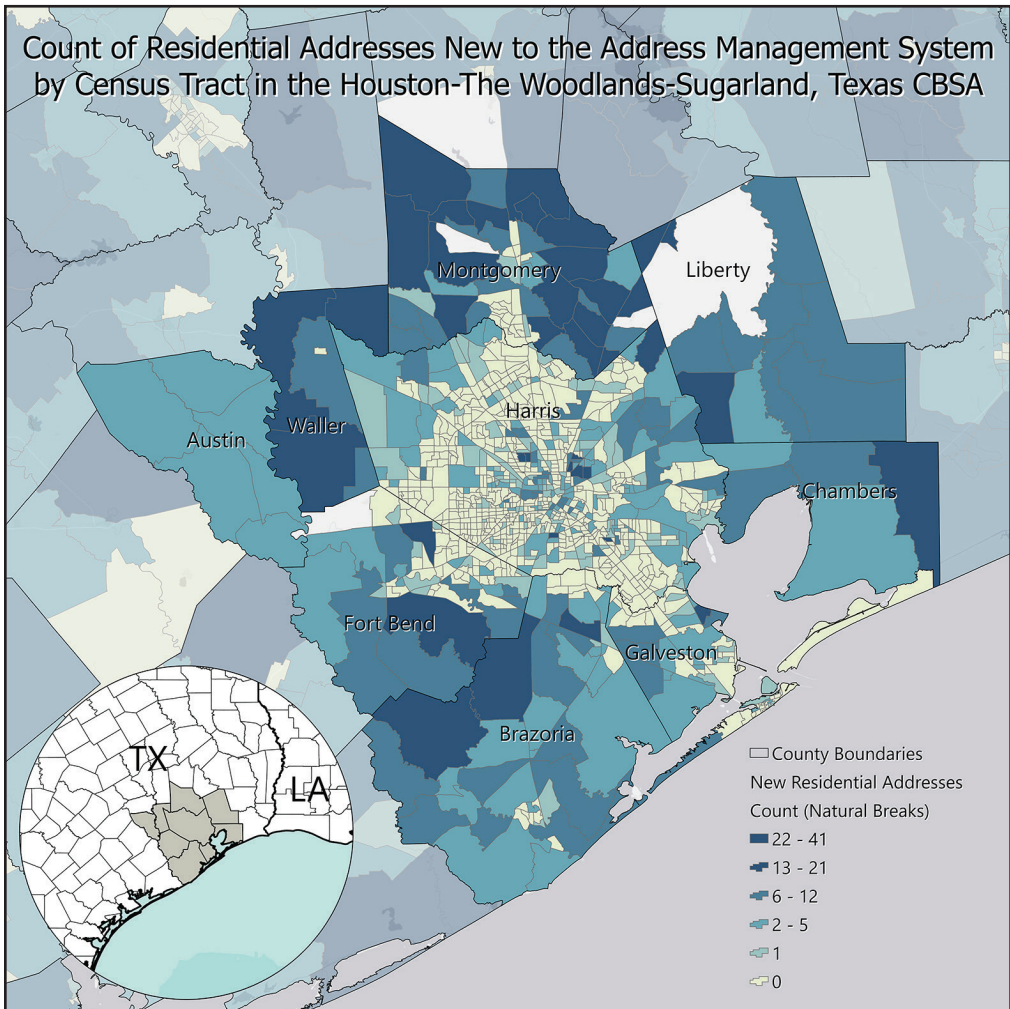
AMS = Address Management System. Q1 = quarter 1. Q4 = quarter 4.  
 Source: 2021 Quarter 1 and 2020 Quarter 4 HUD Aggregated USPS Administrative Data on Address Vacancies

Cartographic inspection of counts of residential addresses new to the AMS reveals that the tracts tend to be on the outskirts of metropolitan/micropolitan areas or in nonmetropolitan areas. The vast majority of residential addresses new to the AMS in 2021 Q1 (60,492 or 67.7 percent), however, are in counties described as Central by the Office of Management and Budget.<sup>10</sup> Of the 25,176 census tracts that had one or more new addresses recorded, the majority are concentrated in a few Central census tracts. Nonmetropolitan and outlying census tracts tend to be much larger because the polygons must be larger to cover similar levels of population due to lower population densities in rural areas. The result is that the cartographic analysis shows city centers with few to no new addresses to AMS, whereas outlying and nonmetropolitan census tracts, which are small shares of the overall number of census tracts, appear to have many more new residential addresses. For example, exhibit 10 shows the Houston-The Woodlands-Sugar Land CBSA. The CBSA has nine counties and is centered around Harris County, where Houston is. Overall, USPS reported 2,383 new residential addresses, the largest share of which (720, or 30.2 percent) were in Harris County; however, only 27.7 percent of census tracts had any new addresses. This percentage is the lowest of any county in the Houston-The Woodlands-Sugar Land CBSA. The other eight counties had between 36.8 percent and 100.0 percent of tracts having new residential addresses, and three counties had 100.0 percent (Austin, Chambers, and Liberty counties).

<sup>10</sup> See the March 2020 file at <https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/delineation-files.html>.

**Exhibit 10**

Count of Residential Addresses New to the Address Management System by Census Tract in the Houston-The Woodlands-Sugar Land, Texas CBSA



CBSA = core-based statistical area.

Source: 2021 Quarter 1 HUD Aggregated USPS Administrative Data on Address Vacancies

**Conclusion**

It is important to revisit datasets and update as necessary while maintaining the consistency that analysts, researchers, policymakers, and others rely on. The goal of these new data fields is to provide new points of information. The ZIP Code Crosswalk Files containing the USPS Preferred Name will help researchers understand (1) where data are being aggregated to or aggregated from and (2) that places with differing names may overlap with each other depending on who is creating the names. Drop counts are intended for researchers to potentially understand the characteristics

of a neighborhood. Including addresses new to the AMS is envisioned to be a convenient method to understand recent growth over the previous quarter. The new fields will likely not be the last ones added to the HUD Aggregated USPS Administrative Data on Address Vacancies as more information becomes available or future data linkages become possible.

## Acknowledgments

I thank the United States Postal Service for providing these new data fields.

## Author

Alexander Din is a social science analyst at the U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

## References

- Beyer, Kirsten, Alan Schultz, and Gerard Rushton. 2007. "Using ZIP Codes as Geocodes in Cancer Research." In *Geocoding Health Data: The Use of Geographic Codes in Cancer Prevention and Control, Research and Practice*, edited by G. Rushton, M.P. Armstrong, J. Gittler, B.R. Greene, C.E. Pavlik, M.M. West, and D.L. Zimmerman. Boca Raton, FL: CRC Press: 37–64.
- Cudnik, Michael T., Jing Yao, Dana Zive, Craig Newgard, and Alan T. Murray. 2012. "Surrogate Markers of Transport Distance for Out-of-Hospital Cardiac Arrest Patients," *Prehospital Emergency Care* 16 (2): 266–272.
- Ginsberg, Wendy R. 2011. *Changing Postal ZIP Code Boundaries*. Washington, DC: Congressional Research Service.
- Grubestic, Tony H., and Timothy C. Matisziw. 2006. "On the Use of ZIP Codes and ZIP Code Tabulation Areas (ZCTAs) for the Spatial Analysis of Epidemiological Data," *International Journal of Health Geographics* 5 (1): 58.
- Krieger, Nancy, Pamela Waterman, Jarvis T. Chen, Mah-Jabeen Soobader, S.V. Subramanian, and Rosa Carson. 2002. "Zip Code Caveat: Bias Due to Spatiotemporal Mismatches Between Zip Codes and U.S. Census-Defined Geographic Areas—The Public Health Disparities Geocoding Project," *American Journal of Public Health* 92 (7): 1100–1102.
- Sadler, Richard. 2019. "Misalignment Between ZIP Codes and Municipal Boundaries: A Problem for Public Health," *Cityscape* 21 (3): 335–340.
- U.S. Census Bureau. 2018. "Population Density by County: 2010." <https://www.census.gov/library/visualizations/2010/geo/population-density-county-2010.html>
- United States Postal Service Office of Inspector General. 2013. "The Untold Story of the ZIP Code." Arlington, VA: USPS OIG. <https://permanent.access.gpo.gov/gpo47009/rarc-wp-13-006.pdf>.