5-Digit ZIP Code Tabulation Area (ZCTA) National Shapefile (2010 Census)

ZCTAs are approximate area representations of U.S. Postal Service (USPS) 5-digit ZIP Code service areas that the Census Bureau creates using census blocks to present statistical data from censuses and surveys. The Census Bureau defines ZCTAs by allocating each block that contains addresses to a single ZIP Code tabulation area, usually to the ZCTA that reflects the most frequently occurring ZIP Code for the addresses within that block. Blocks that do not contain addresses but that are completely surrounded by a single ZIP Code tabulation area (enclaves) are assigned to the surrounding ZCTA; those surrounded by multiple ZCTAs will be added to a single ZCTA based on the longest shared border. The Census Bureau identifies ZCTAs using a 5-character code that represents the most frequently occurring USPS ZIP Code within that ZCTA. This code may contain leading zeros.

Users should not use ZCTAs to identify the official USPS ZIP Code for mail delivery. The USPS makes periodic changes to ZIP Codes to support more efficient mail delivery. ZIP Codes that cover primarily nonresidential or post office box addresses may not have a corresponding ZCTA because the delineation process uses primarily residential addresses, resulting in a bias towards ZIP Codes used for city-style mail delivery.

For more information on ZCTAs, visit:

<https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html>

See Appendix O-3 5-digit ZIP Code Tabulation Area (ZCTA) National Record Layout.

5. Relationship File Concept Overview

Relationships files are database files that provide additional attribute information that users can join to the TIGER/Line Shapefiles. The following sections describe, in alphabetical order, the geographic entity type displayed in each relationship file as well as the record layout for each file.

5.1 Address Ranges

Address range information is available in the following relationship file:

Address Ranges County-based Relationship File

Address ranges fall along an edge side relative to the coded direction of the edge. The 2020 TIGER/Line Shapefiles contain potential address ranges, not individual addresses. Potential ranges include the full range of possible structure numbers even though the actual structures might not exist (see Figure 26).

The address ranges relationship file contains the attributes of each address range. Each address range applies to a single edge side and has a unique address range identifier (ARID) value. A user can determine the edge to which an address range applies by linking the address range to the all lines shapefile using the TLID attribute. Multiple address ranges can apply to the same edge because addresses with different number sequences (e.g., 101, 103, 1622, 1624, etc.) or alphanumeric characters (e.g., N101, N103, S099, S97) can appear along that edge. The most inclusive address range associated with each side of a street edge appears in the all lines shapefile.

The most inclusive address range (not a composite of the available address ranges) has the largest range of potential house number values of all address ranges associated with the side of an edge. The Census Bureau provides the most inclusive address ranges for users looking for data comparable to the address ranges supplied in the Record Type 1 (RT1) of the TIGER/Line data files.

5.1.1 ZIP Codes and Address Ranges

The address numbers used to create address ranges are house number-street name style addresses (or city-style addresses). A house number-street name style address minimally consists of a structure number, street name, and a 5-digit ZIP Code (e.g., 213 Main Street 90210). In the 2020 TIGER/Line Shapefiles, ZIP Codes are only associated to address ranges.

The ZIP Code is an attribute of the address ranges. The address ranges relationship file has a 5-digit ZIP Code field containing a numeric code that may have leading zeroes. Both sides of a street typically have the same ZIP Code, but this is not always true. Different ZIP Codes may serve different sides of a street or cover addresses at each end of a street. Nearly all address ranges will have a ZIP Code, but there are a few instances where unknown ZIP Codes result in null/blank values in the ZIP Code field.

The 2020 TIGER/Line Shapefiles may not contain all street delivery ZIP Codes and may contain some non-delivery ZIP Codes. In some cases, P.O. Box delivery ZIP Codes may be associated with house number-street name style addresses that are not used for mail delivery (see below). The distribution of ZIP Codes in the TIGER/Line Shapefiles may not reflect the exact USPS ZIP Code service area. Likewise, the address range ZIP Codes may not match the ZIP Code Tabulation Area (ZCTA) for the area.

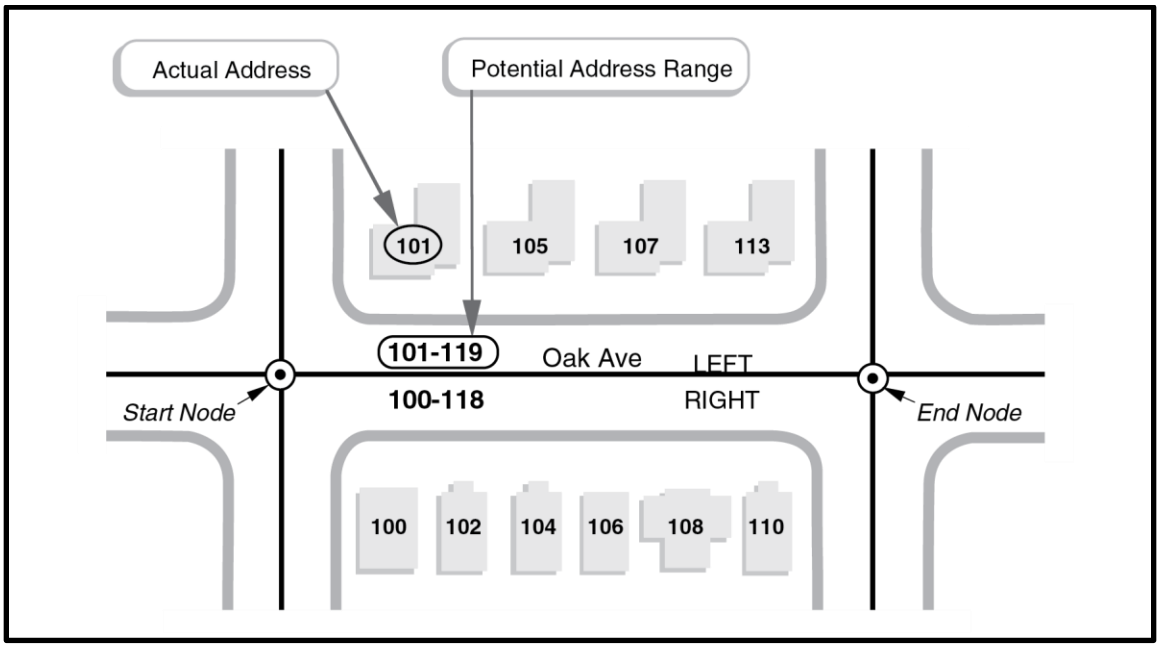

Figure 28: TIGER/Line Shapefiles Address Range Basics

The TIGER/Line Shapefiles contain potential address ranges for city-style addresses. The edge (between the start node and the end node) in the diagram above has two address ranges; the left side has odd-

numbered addresses and the right side has the complementary even-numbered addresses. Potential address ranges along an edge have values that encompass the addresses of existing structures, as well as those not yet built.

Note: The all lines shapefile has the largest range of potential house number values of all address ranges associated with the side on an edge.

Layer Name	Filename	Spatial Data	Address Ranges	Geocoding Usability
All Lines Shapefile	edge.shp	Yes	Most inclusive address ranges	Limited geocoding
Address Range Feature Shapefile	addrfeat.shp	Yes	All address ranges	Best source for geocoding
Address Range Table	addr.dbf	No	All address ranges	No geocoding
Address Range to Feature Name Relationship Table	addrfn.dbf	No	No address ranges	No geocoding

Figure 29: Address Range Product Comparison

Some basic characteristics of address ranges are as follows:

- The 2020 TIGER/Line Shapefiles generally contain address ranges with only house number-street name style addresses. They do not show rural route and post office box addresses. They may contain structure numbers assigned in select areas for use by local emergency services, but not for mail delivery. The 2020 TIGER/Line Shapefiles do include address ranges and ZIP Codes in some small places where the USPS provides only P.O. Box service. These address ranges represent the structure numbers collected during census field operations, supplemented with addresses provided through local participant programs and intercensal Census Bureau activities and updates. These structure-number addresses may have ZIP Codes associated only with P.O. Box addresses. The USPS does not recognize these street addresses as valid mailing addresses and does not assign a ZIP+4 Code to them or include them in the ZIP+4 file. The address ranges may be used to geocode a structure to the census block, but users should be aware of potential conflicts with similar or duplicate mailing street addresses.
- Gaps may exist between multiple ranges for a single edge. A gap may be significant because any numbers missing from one edge may actually appear on another edge. This situation occurs in cases where there are address anomalies (e.g., out-of-parity or out-of-sequence addresses) that cover a single house number. For example, address 709 Main Street is in the middle of the even side of the 700 block of Main Street and suppressed because it is a single address-address range. The following address ranges for the 700 block of Main Street will appear in the 2020 TIGER/Line Shapefiles: 700-798 Main Street, 701-707 Main Street, and 711-799 Main Street. Users cannot tell where 709 Main Street is located based on the information provided. Suppression of single address-address ranges is to protect the confidentiality of individual addresses as specified by Title 13 of the U.S.C.

- Address ranges can include numbers with alphabetic characters. These characters help uniquely identify addresses within a county. For instance, certain unincorporated areas of Genesee County, Michigan, add a letter G prefix to the address number. The characters maintain a consistent column placement within the address range field; for example, the letter G maintains a consistent column placement in the range G1 to G99.
- Some address systems use a hyphen to separate avenue numbers, private road designators, and grid cell numbers from the structure numbers; for example, 10-01 Reynolds St. uses a hyphen to separate the avenue number (i.e., Tenth Avenue) from the structure number. Depending on the locality, the hyphen may be unnecessary for address matching.
- Address ranges exist only for street features, and in some cases, geographic corridor and geographic offset boundary features adjacent to street features. When these boundaries exist, the address ranges move from the street centerline to the boundary to ensure that addresses will geocode to the correct entity.
- Address ranges (consisting of a unique combination of structure number, ZIP Code, feature name, feature type, and directional) should not overlap; addresses should belong to only one address range. The Census Bureau edits address ranges to locate possible overlaps but cannot guarantee that all possible overlap situations are found and have been resolved.
- Address ranges in the 2020 TIGER/Line Shapefiles may be associated with one or more of the street names that belong to an edge. Address range overlap conflicts may occur if the address ranges are associated with some street names or route numbers not intended for use in locating addresses. A route number may traverse several streets with similar house numbers, but different common names used for mail delivery.

5.1.2 Imputed Address Ranges

Imputed address ranges occur during the process of updating the MTDB when a new edge intersects an existing edge with address ranges. The intersection splits the existing edge and produces two new edges connected by a new node located at the intersection point. The update program divides the old address ranges between the two new edges and imputes the address range ends at the new node.

The impute process allocates either all or part of each original address range to each of the new edges in proportion to their lengths (see Figures 28 and 29). For each side of the original edge, the process considers all address ranges appearing on the side and determines the overall low and high addresses. The process assumes addresses have an even distribution along the length of the edge and applies the proportion of edge lengths to the overall address range to calculate a split-point address for each side. Address ranges that fall entirely above or below the split-point address move intact to one of the new edges. The process divides any address ranges that contain the split-point address and allocates each part to one of the new edges. The new address range ends created from the split are imputed values and have the from address range type (FROMTYP) or to address range type (TOTYP) set to imputed value. Some intermediate address range ends also may carry the impute flag. These address range ends fall between the overall high and low address for edge sides that have more than one address range. In current practice, the imputation process will assign the entire address range to one of the edges if the other is very small and would receive only a single address using the proportional division of address ranges.

5.1.3 Geocoding

To get the best match results, the Census Bureau advises users to use the address range feature shapefile (ADDRFEAT.shp) to geo-reference/geocode addresses. Address ranges in the MTDB may separate into multiple address ranges on the same edge (e.g., ZIP Code differences or to establish gaps created by address anomalies located elsewhere). Some address ranges may also include embedded

alphanumeric characters or hyphens that make them distinct from the other address ranges on the same edge side. The ADDRFEAT.shp contains all the address range to edge and street name relationships for a county to increase the number of potential geocoding matches. In comparison, the most inclusive address range in the all line shapefile (EDGES.shp) can also be used for geocoding but a single pair of left and right side address ranges and the primary street name on the edge may not always provide complete address range coverage.

5.1.4 Limitations

Users of the address ranges in the 2020 TIGER/Line Shapefiles should be aware that address range overlaps, gaps, odd/even reversals, and low-high orientation reversals may exist in the data. Apart from overlaps, these may be valid. While the Census Bureau continues to edit and correct data errors, it is possible that some still exist. The Census Bureau defines address ranges on a county-by-county basis. Streets that cross county boundaries may have overlapping address ranges. The Census Bureau is implementing checks to identify and correct these issues. Geocoders often return the address range located in the first county, alphabetically.

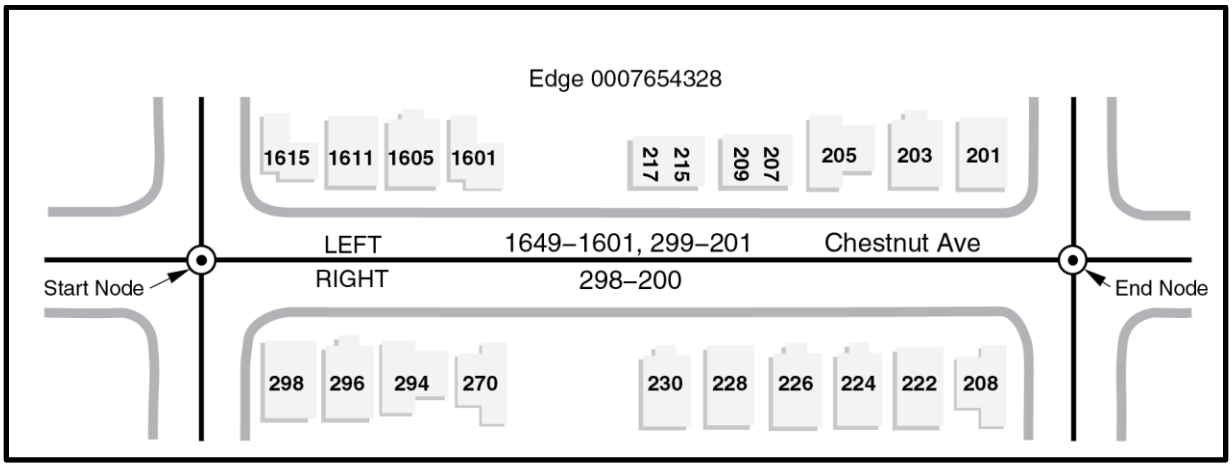

Figure 30: TIGER/Line Shapefile Address Range Imputes - Before Split

The MTDB uses impute flags to indicate that the one or both ends of an address range reflect calculations instead of known values. Imputed address situations generally occur when a new edge splits an edge with existing address ranges. The illustration above shows the address ranges on Chestnut Ave. before a split (see Figure 30).

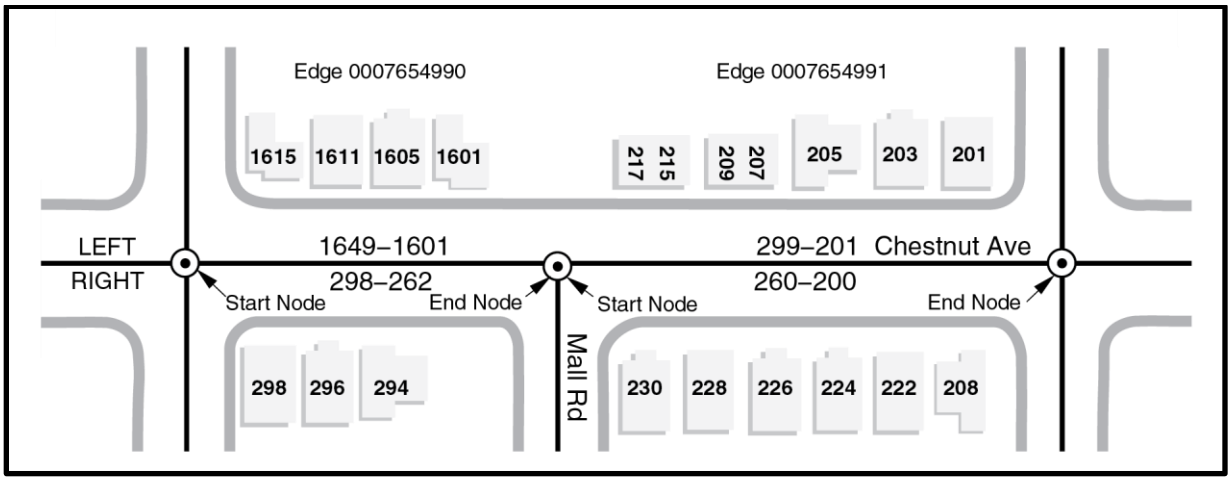

Figure 31: TIGER/Line Shapefile Address Range Imputes - After Split

In the diagram above (see Figure 31), Mail Rd. has split the edge into two parts. The MTDB will assign a new TIGER/Line Identification Number (TLID) to each part and delete the old number. The MTDB next determines the overall address range for each edge side (1649 to 201 on the left side and 298 to 200 on the right side) and the split points for each affected address range (approximately 1088 on the left side and 261 of the right side). Address ranges that fall entirely above or below the split point belong to one of the two new edges and do not get an impute flag. The MTDB divides those address ranges that contain the split point and assigns a part to each of the edges.

See Appendix R-1 Address Ranges County-based Relationship Record Layout.

5.2 Address Range-Feature Name Relationships

Address range-to-feature name relationship information is available in the following relationship file:

Address Range-Feature Name County-based Relationship File

The address range-feature name relationship file contains a record for each address range-linear feature name relationship. The purpose of this relationship file is to identify all street names associated with each address range. An edge can have several feature names and an address range located on an edge can be associated with multiple feature names. A user can link to the address ranges relationship file by using the ARID attribute. The LINEARID attribute identifies the linear feature name and relates the address range back to the Feature Names Relationship File.

See Appendix R-2 Address Range-Feature Name County-based Relationship Record Layout.

5.3 Feature Names

Feature name information is available in the following relationship file:

Feature Names County-based Relationship File

The feature names relationship file contains a record for each feature name-edge combination and includes the feature name attributes. The edge to which a feature names relationship file record applies can be determined by linking to the all lines shapefile using TLID attribute. Multiple feature names relationship table records can link to the same edge.

For example:

A road edge could link to U.S. Hwy 22 and Rathburn Road.

The LINEARID attribute identifies the linear feature to which the feature name applies. Multiple feature names may exist for the same edge. Linear features are not included in the data set, but users can construct them using the all lines shapefile and the relationship tables.

If the edge is both a road and a rail feature, the name associated with the rail feature will carry a rail feature MTFCC. If there are any address ranges on the edge, they apply only to the designated street features.

See Appendix B Feature Name Directions

See Appendix C Feature Name Qualifiers

See Appendix D Feature Name Types

See Appendix R-3 Feature Names County-based Relationship Record Layout

5.4 Topological Faces-Area Landmark Relationships

Topological faces-to-area landmark relationship information is available in the following relationship file:

Topological Faces-Area Landmark State-based Relationship File

The topological faces-area landmark relationship file contains a record for each face-area landmark relationship. The face to which a topological faces-area landmark relationship file record applies can be determined by linking to the topological faces (polygons with all geocodes) Shapefile using the TFID attribute. The area landmark to which a topological faces-area landmark relationship table record applies can be determined by linking to the area landmark shapefile using the area landmark identifier (AREAID) attribute. A face may be part of multiple area landmarks. An area landmark may consist of multiple faces.

See Appendix R-4.1 Topological Faces-Area Landmark County-based Relationship File Record Layout.

5.5 Topological Faces-Area Hydrography Relationships

Topological faces-to-area hydrography relationship information is available in the following relationship file:

Topological Faces-Area Hydrography County-based Relationship File

The topological faces-area hydrography relationship file contains a record for each face-area hydrography feature relationship. The face to which a topological faces-area hydrography relationship file record applies can be determined by linking to the topological faces (polygons with all geocodes) using the TFID

attribute. The area hydrography feature to which a topological faces-area hydrography relationship file record applies can be determined by linking to the area hydrography shapefile using the area hydrography identifier (HYDROID) attribute. A face may be part of multiple area water features. An area water feature may consist of multiple faces.

See Appendix R-4.2 Topological Faces-Area Hydrography County-based Relationship Record Layout.

5.6 Topological Faces-Military Installation Relationships

Topological faces-to-military installation relationship information is available in the following relationship file:

Topological Faces-Military Installation National Relationship File

The topological faces-military installation relationship file contains a record for each face-military installation relationship. To find out more information about the face the military installation relates to, use TFID in the topological faces (polygons with all geocodes) Shapefile. To determine the military installation the record applies to use the area id (AREAID) attribute found in the military installation shapefile. A military installation feature may consist of multiple faces.

See Appendix R-4.3 Topological Faces-Military Installation National Relationship Record Layout.