4. 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.

4.1 Address Ranges

Address range information is available in the following relationship file:

Address Ranges County-based Relationship File (Current)

The term “address range” refers to the collection of all possible structure numbers from the first structure number to the last structure number, including all numbers of a specified parity in between. Address ranges fall along an edge side relative to the coded direction of the edge. The 2020 Census Prototype 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 5).

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 permanent edge identifier (TLID) attribute. Multiple address ranges can apply to the same edge because addresses with different number sequences (e.g., 101, 103, 1622, 1624…) or non-numeric characters (e.g., N101, N103, S099, S97) can appear along that edge. Note that the most inclusive address range associated with each side of a street edge appears in the All Lines shapefile.

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

ZIP Codes and Address Ranges

The address numbers used to create address ranges are commonly known as 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; for example, 213 Main Street 90210. In the 2020 Census Prototype 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 Census Prototype 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.

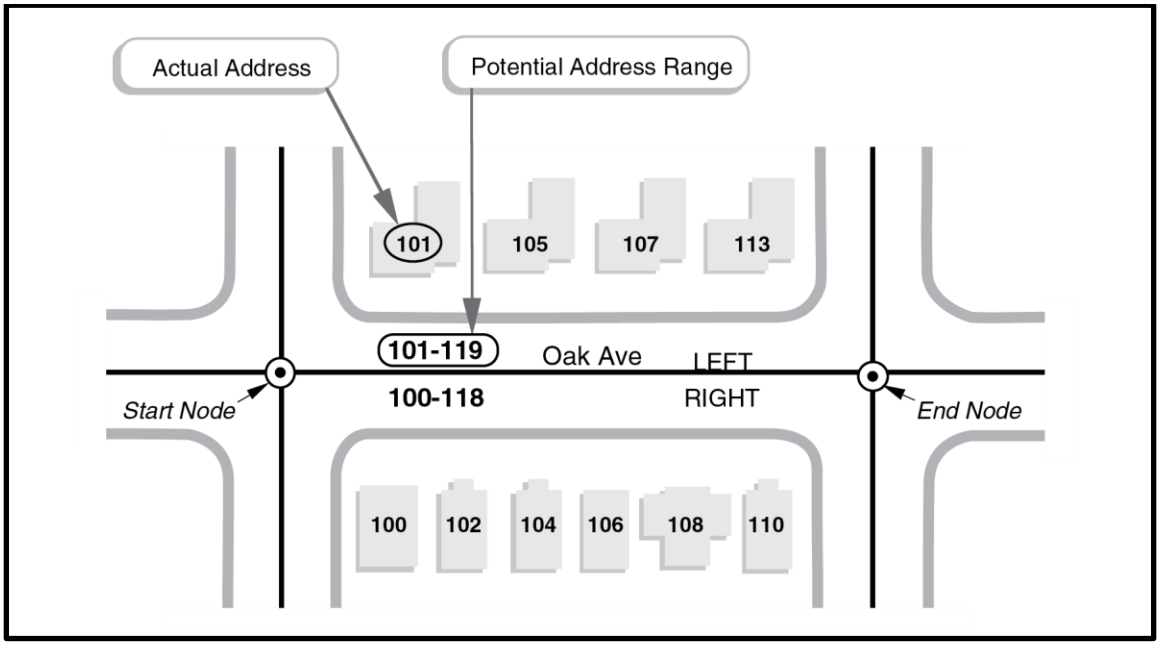

Figure 5: TIGER/Line Shapefiles Address Range Basics

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 most inclusive address range has the largest range of potential house number values of all address ranges associated with the side on an edge. It is not a composite of the available address ranges.

Table 4: Address range product comparison table

Layer Name	Filename	Spatial Data	Address Ranges	Geocoding Usability
All Lines Shapefile	edge.shp	Yes	Most inclusive address ranges	Limited 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

Some basic characteristics of address ranges are as follows:

- The 2020 Census Prototype 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 Census Prototype Shapefiles do include address ranges and ZIP Codes in some small places where the USPS provides only post office box service. These address ranges represent the structure numbers collected during the 2000 and 2010 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 post office 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 such as out-of-parity or out-of-sequence addresses. The Census Bureau does not provide any single address-address ranges in the 2020 Census Prototype Shapefiles, including out-of-parity and out-of-sequence address ranges 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 is 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 Census Prototype Shapefiles: 700-798 Main Street, 701-707 Main Street, and 711- 799 Main Street. Based on the information provided, data users cannot tell where 709 Main Street is located. Suppression of single address-address ranges is to protect the confidentiality of individual addresses as specified by Title 13 of the U.S. Code.
- 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 have been identified and resolved.
- Address ranges in the 2020 Census Prototype Shapefiles may be associated with one or more of the street names that belong to an edge. Caution: 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.

Imputed Address Ranges

Imputed address ranges occur during the process of updating the MAF/TIGER database 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 6 and 7). 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 are evenly distributed 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.

Geocoding

To get the best geocoding match results in ArcGIS, the Census Bureau advises data users to use the Address Range Feature Shapefile (ADDRFEAT.shp) to geo-reference/geocode addresses. Address ranges in the MAF/TIGER database may be separated into multiple address ranges on the same edge because of 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 of 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.

Limitations

Users of the address ranges in the 2020 Census Prototype Shapefiles should be aware that address range overlaps, gaps, odd/even reversals, and low-high orientation reversals may exist in the data. With the exception of overlaps, these may be valid. While the Census Bureau continues to edit and correct data errors, it is possible that some still exist. The U.S. 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.

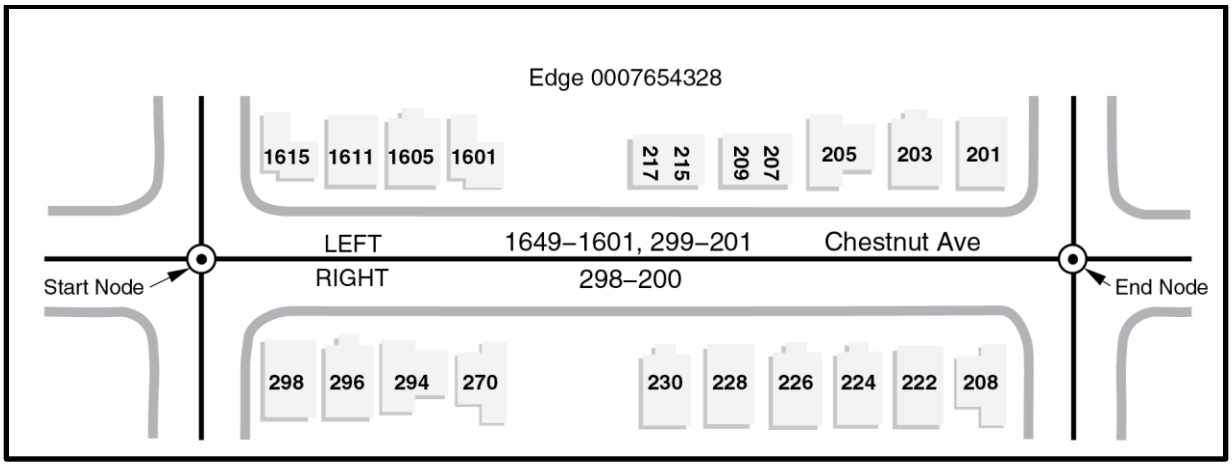

Figure 6: TIGER/Line Shapefile Address Range Imputes - Before Split

The MAF/TIGER database 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.

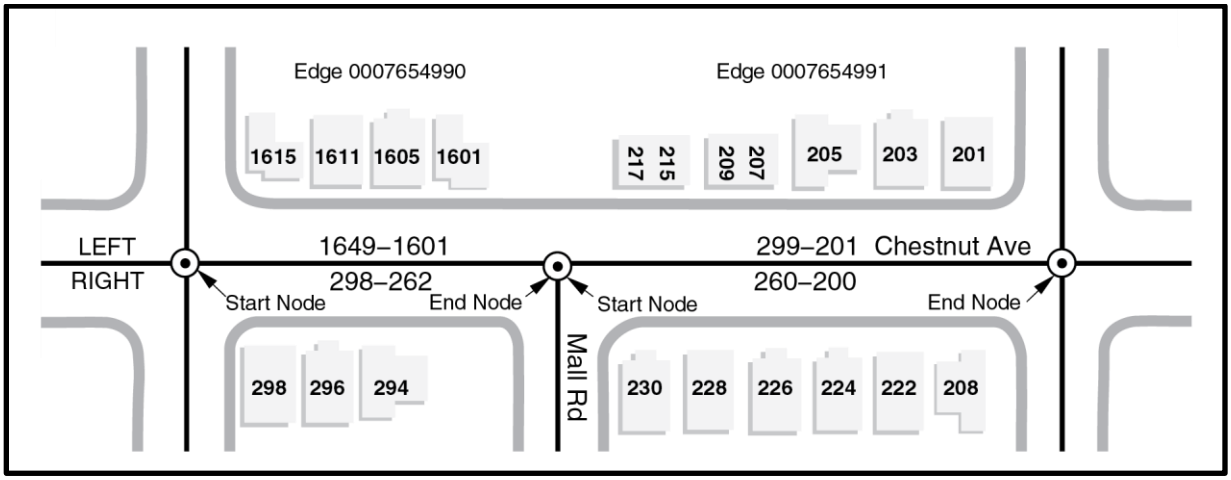

Figure 7: TIGER/Line Shapefile Address Range Imputes - After Split

In the diagram above, Mail Rd. has split the edge into two parts. The MAF/TIGER database will assign a new TIGER/Line identification number (TLID) to each part and delete the old number. It 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 MAF/TIGER database divides those address ranges that contain the split point and assigns a part to each of the edges.

Note: The 2020 Census Prototype Shapefiles cover Providence County, RI only. The shapefiles contain 2018 geography in the anticipated 2020 Census format.

4.1.1 Address Ranges Relationship File Record Layout (Current)

File Name: tl_2018_<state + county FIPS>_addr.dbf

Field	Length	Type	Description
TLID	10	Integer	permanent edge ID
FROMHN	12	String	From house number
TOHN	12	String	To house number
SIDE	1	String	side indicator flag
ZIP	5	String	5-digit ZIP code
PLUS4	4	String	ZIP+4 code
FROMTYP	1	String	From address range end type
TOTYP	1	String	To address range end type
ARID	22	String	Address range identifier
MTFCC	5	String	MAF/TIGER feature class code

4.2 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 (Current)

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 address range identifier (ARID) attribute. The linear feature identifier (LINEARID) attribute identifies the linear feature name and relates the address range back to the Feature Names Relationship File.

Note: The 2020 Census Prototype Shapefiles cover Providence County, RI only. The shapefiles contain 2018 geography in the anticipated 2020 Census format.

4.2.1 Address Range-Feature Name Relationship File Record Layout (Current)

File Name: tl_2018_<state + county FIPS>_addrfn.dbf

Field	Length	Type	Description
ARID	22	String	Address range identifier
LINEARID	22	String	Linear feature identifier

4.3 Feature Names

Feature name information is available in the following relationship file:

Feature Names County-based Relationship File (Current)

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 the permanent edge identifier (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 linear feature identifier (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.

Note that the MTFCC in this relationship file refers to the specific MAF/TIGER feature class code associated with this linear feature and feature name. 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.

Appendices B, C, and D of this document include additional information about feature name components.

Note: The 2020 Census Prototype Shapefiles cover Providence County, RI only. The shapefiles contain 2018 geography in the anticipated 2020 Census format.

4.3.1 Feature Names Relationship File Record Layout (Current)

File Name: tl_2018_<state + county FIPS>_featnames.dbf

Field	Length	Type	Description
TLID	10	Integer	permanent edge ID
FULLNAME	100	String	Concatenation of expanded text for prefix qualifier, prefix direction, prefix type, base name, suffix type, suffix direction, and suffix qualifier (as available) with a space between each expanded text field
NAME	100	String	Base name portion of the standardized name
PREDIRABRV	15	String	Prefix direction description component of the feature name
PRETYPABRV	50	String	Prefix type description component of the feature name

Field	Length	Type	Description
PREQUALABR	15	String	Prefix qualifier description component of the feature name
SUFDIRABRV	15	String	Suffix direction description component of the feature name
SUFTYPABRV	50	String	Suffix type description component of the feature name
SUFQUALABR	15	String	Suffix qualifier description component of the feature name
PREDIR	2	String	prefix direction code component of the feature name
PRETYP	3	String	prefix type code description component of the feature name
PREQUAL	2	String	prefix qualifier code component of the feature name
SUFDIR	2	String	suffix direction code component of the feature name
SUFTYP	3	String	suffix type code description component of the feature name
SUFQUAL	2	String	suffix qualifier code component of the feature name
LINEARID	22	String	Linear feature identifier
MTFCC	5	String	MAF/TIGER feature class code
PAFLAG	1	String	primary/alternate flag

4.4 Topological Faces-Area Landmark Relationships

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

Topological Faces-Area Landmark County-based Relationship File (Current)

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 permanent face identifier (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.

Note: The 2020 Census Prototype Shapefiles cover Providence County, RI only. The shapefiles contain 2018 geography in the anticipated 2020 Census format.

4.4.1 Topological Faces-Area Landmark Relationship File Record Layout (Current)

File Name: tl_2018_<state + county FIPS>_facesal.dbf

Field	Length	Type	Description
TFID	10	Integer	permanent face ID
AREAID	22	String	Area landmark identifier

4.5 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 (Current)

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 permanent face identifier (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.

Note: The 2020 Census Prototype Shapefiles cover Providence County, RI only. The shapefiles contain 2018 geography in the anticipated 2020 Census format.

4.5.1 Topological Faces-Area Hydrography County-based Relationship File Record Layout (Current)

File Name: tl_2018_<state + county FIPS>_facesah.dbf

Field	Length	Type	Description
TFID	10	Integer	permanent face ID
HYDROID	22	String	Area hydrography identifier