LODES: Updating Historical Data to 2020 Tabulation Geography

April 4, 2023

With the release of Version 8, the LEHD Origin-Destination Employment Statistics (LODES) public-use data product[1] is being updated to 2020 tabulation geography as defined for the 2020 Decennial Census. In Versions 6 and 7, LODES used the 2010 tabulation geography from the 2010 Decennial Census and prior to that, it used 2000 tabulation geography from the 2000 Decennial Census. This document discusses the method by which the data have been updated.

Comparing 2010 to 2020

The source of geography for LODES is the U.S. Census Bureau's TIGER/Line data product. For more detailed technical information, see the references section at the end of this document.

The 2010 tabulation census blocks were created by the Census Bureau as the smallest geography for which demographic/economic data are tabulated for the 2010 Decennial Census. Because the tabulation blocks change only once every decade, they serve as a good base for consistent tabulation of a dataset such as LODES over many years. In 2010 about 11.1 million census blocks were defined for the states and territories.

In 2020 – to track the growth and movement of the population – the tabulation blocks were redefined based upon data from the 2020 Decennial Census. While the population increased between 2010 and 2020, the block set contracted by over 26% between 2010 and 2020 to 8.2 million blocks as many unpopulated blocks (e.g. highway median strips) were consolidated.

In general, some transitions between 2010 and 2020 were very simple. The boundaries of a block could remain the same and only the code by which it was identified would change. Or even more simply, nothing at all might have changed. About 5.0 million blocks did not change their boundaries in any way from 2010 to 2020.

Some changes could be more complex – a new housing development might cause a large, formerly unpopulated block to split into many blocks now filled with people. Formerly populous blocks could also be combined in areas losing population. And in the most complex examples, sets of blocks are transformed into other sets of blocks with no easy splits or combinations. Some basic representations of these changes are diagrammed as part of the methodology description below.

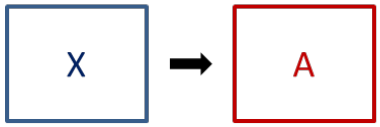
[1] Note: The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release. CBDRB-FY21-249. Project Number: P-6000266.

Transforming LODES Data from 2010 to 2020 Tabulation Blocks

LODES data from 2019 and prior years (Version 7) have been published in 2010 census blocks. To present these older data in 2020 census blocks, a transformation process was applied to these historical data. The general algorithm[2] is to take each job in a 2010 block and assign it to a 2020 block with a probability based upon the areal overlap between the 2010 block and the intersecting 2020 block(s). The areal intersections are taken from the block relationship files published by the Census Bureau.

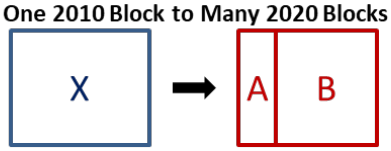
If a 2010 census block is exactly the same as a 2020 census block, then they share 100% of each other's area and any jobs in the 2010 census block will have a 100% chance of being assigned to the corresponding 2020 census block. But if a 2010 census block is split into two parts, one with 25% of the original area and another with 75%, then any jobs in the 2010 census block will be allocated into the first part with 25% chance and into the second part with a 75% chance. Fractional job counts are not allowed. A few additional examples of this transformation are illustrated below.

One 2010 Block to One 2020 Block



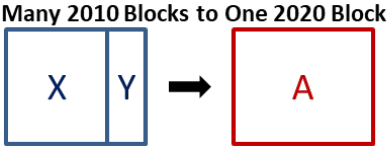
Census block X in 2010 becomes census block A in 2020, with A having 100% of X's area. Each job in X has a 100% chance of being placed in A.

One 2010 Block to Many 2020 Blocks



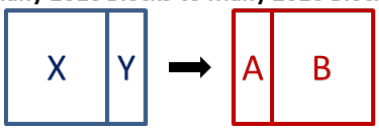
Census block X in 2010 becomes census blocks A and B in 2020, with A having 25% of X's area and B having 75% of X's area. Each job in X has a 25% chance of being placed in A and a 75% chance of being placed in B.

Many 2010 Blocks to One 2020 Block



Census blocks X and Y in 2010 becomes census block A in 2020, with A having 100% of X's area and 100% of Y's area. Each job in X has a 100% chance of being placed in A and each job in Y has a 100% chance of being placed in A.

Many 2010 Blocks to Many 2020 Blocks



Census blocks X and Y in 2010 become census blocks A and B in 2020, with A having 33% of X's area and 0% of Y's area and B having 67% of X's and 100% of Y's area. Each job in X has a 33% chance of being placed in A and a 67% chance of being placed in B. Each job in Y has no chance of being placed in A and a 100% chance of being placed in B.

A few additional notes about this method:

First, the allocation is a statistical process and may not result in a distribution of jobs that exactly matches the areal distribution. If, for example, in the "One-to-Many" case above, block X has 100 jobs, we might expect to see 25 jobs allocated to block A and 75 jobs allocated. This is possible, but so are

[2] The method described is the same one that was used to transform LODES data from 2000 census blocks to 2010 census blocks. That transformation was previously documented in lehd.ces.census.gov/doc/help/onthemap/OnTheMap2010Geography.pdf

allocations of 24:76, 27:73, and 0:100, even though the last is extremely unlikely (about 1 chance in 3 trillion).

Second, this transformation does not allow for jobs to be moved between blocks that do not share some areal intersection. For example, if 2020 block C were some distance from A and B, it could inherit *none* of block X's jobs, because it does not overlap block X in any way.

Finally, some geographic areas may exhibit a shift in job count when they are compared between LODES Version 7 and LODES Version 8. This could happen for two main reasons:

1. A geographic area has been redefined. Examples of this are census tracts that have been completely redrawn from 2010 to 2020 as well as cities or other political geographies that redefined their boundaries intercensally.
2. In certain cases the redefinition of blocks in 2020 could cause the allocations of some jobs from a 2010 block to be inside an political/administrative area (city, Congressional District, etc.) and some to be outside of the area. This would cause a different total if that same block in 2010 were allocated wholly into or out of the area. This can only happen if the political/administrative area split a 2010 block into two separate parts.

References

TIGER/Line Shapefiles: www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html

2010-2020 Relationship Files: www.census.gov/geographies/reference-files/time-series/geo/relationship-files.2020.html

Additional information on LODES: lehd.ces.census.gov/data/#lodes

LODES Feedback: CES.OnTheMap.Feedback@census.gov