Accessibility Data Documentation 2018

Prepared by the
Accessibility Observatory at the University of Minnesota

June 26, 2018

**ACCESSIBILITY
OBSERVATORY**

UNIVERSITY OF MINNESOTA
Driven to Discover℠

1 Overview

This document describes the 2018 accessibility data released by the Accessibility Observatory at the University of Minnesota. The data are organized into individual files for each core-based statistical area (CBSA) included in the Access Across America project for 2018, as well as for each state in the US. Data files can be accessed at http://access.umn.edu/research/america/transit/2018/

The following sections describe the format, naming, and content of the data files.

2 Data Formats

The data files are provided in one data format: Geopackage. Geopackage (.gpkg) files are an open-source, geospatial filetype that can contain multiple layers of data in a single file, and can be opened in most GIS software, including both ArcGIS and QGIS.

3 Geography and Coverage

The data files provide accessibility records for individual Census blocks within each CBSA. Each record is uniquely identified by its 15-digit [geoid] field, which corresponds to an individual block's GEOID code based on the U.S. Census Bureau's 2010 geography definitions.

Geopackage data files include a geospatial data layer, which contains a geometry for each record which describes the geographical boundary of the Census block. This is based on 2010-vintage TIGER/Line data from the U.S. Census Bureau, and is stored in the WGS84 coordinate system.

Data are provided for each block in the selected CBSAs, and in all states, that include any amount of land area. Blocks with a land area of zero (e.g. blocks that are entirely water) are omitted.

4 File Naming Conventions

Accessibility data files are named according to the following template:

[CBSA or STATE]_[MODE]_[YEAR]_[STIME]-[ETIME].[EXT]

- [CBSA or STATE] is the 5-digit code identifying a core-based statistical area, or 2-digit code identifying a state.

- [MODE] is a 2-character code identifying the transportation mode. Currently defined codes are tr for public transit and au for automobile.

- [YEAR] identifies the year of the accessibility data.

- [STIME] indicates the first departure time included in averaged accessibility data.

- [ETIME] indicates the last departure time included in averaged accessibility data.

- `[EXT]` is the file extension, which is determined by the data format. Geopackage files have the extension `.gpkg`, and are compressed into `.zip` files.

For example, the file named `33460_tr_2018_0700-0859.zip` contains transit accessibility data for CBSA 33460 (Minneapolis-St. Paul-Bloomington, MN-WI) in 2018, averaged between 7:00 AM and 8:59 AM (inclusive). The data are in a compressed geopackage format.

As another example, the file named `27_au_2018_0800.zip` contains auto accessibility data for the state of Minnesota in 2018, for a departure time of 8:00 AM. The data are in a compressed geopackage format.

5   Data Field Definitions

As of 2018, two years of LEHD data are used in tandem, to allow more year-to-year comparisons to be made. The below table summarizes the data field definitions and codes for this accessibility dataset; the "fieldname_descriptions" layer within the geopackage file describes the data fields in full.

Data Field Definitions

Field Position	Field Name	Type	Definition
1	blockid	string	Origin Census block ID
2	threshold	integer	Travel time threshold in seconds
3	w_c000_16	integer	Total number of jobs that can be reached from the location within threshold seconds, 2016 LEHD
4	w_ca01_16	integer	Number of jobs for workers age 29 or younger that can be reached from the location within threshold seconds, 2016 LEHD
…	…	…	…
53	w_cns01_16	integer	Number of jobs for workers at firms with Firm Size: 500+ Employees that can be reached from the location within threshold seconds, 2016 LEHD
54	r_c000_16	integer	Total number of workers that can be reached from the location within threshold seconds, 2016 LEHD
55	r_ca01_16	integer	Number of workers age 29 or younger that can be reached from the location within threshold seconds, 2016 LEHD
…	…	…	…
94	r_cs02_16	integer	Number of workers with Sex: Female that can be reached from the location within threshold seconds, 2016 LEHD
95	w_c000_15	integer	Total number of jobs that can be reached from the location within threshold seconds, 2015 LEHD
…	…	…	…
186	r_cs02_15	integer	Number of workers with Sex: Female that can be reached from the location within threshold seconds, 2015 LEHD

Note: Fields 3 through 52 correspond to the Workplace Area Characteristics (WAC) LEHD tables, and fields 53-94 correspond to the Residence Area Characteristics (RAC) LEHD tables. Fields 3-94 correspond to LEHD 2016; fields 95-186 correspond to LEHD 2015. Full documentation of the NAICS sector codes and descriptions used in the LODES dataset is available at https://lehd.ces.census.gov/data/lodes/LODES7/LODESTechDoc7.4.pdf.

For example, the Geopackage data layer record

```
id,blockid,threshold,w_c000_16,w_ca01_16,…,w_cs02_16
  1000,490099601001120,1800,661354,1723,…,25181
```

indicates that 661,354 total jobs can be reached from origin Census block 490099601001120 within 30 minutes by transit (see filename for parameter information), using LEHD 2016 data; likewise, 1,723 jobs for workers age 29 or younger can be reached from this origin within 30 minutes, and 25,181 women workers can be reached from this origin within 30 minutes.

Data files for all modes follow the format and field definitions as above; for transit, there is one file describing the average accessibility between 7:00 AM and 8:59 AM; for auto, there is one file for each hourly departure time; for biking, there are four files, corresponding to the four Level of Traffic Stress (LTS) values, each with a 12:00 noon departure time; for walking, there is one file, corresponding to a 12:00 noon departure time.

6   Opening and Using Geopackages in ArcGIS

1. Unzip the zip archive

2. Use the "Add Data" function in Arc to select the .gpkg file

3. Select which layer(s) are needed – always select "main.blocks" as this layer contains the Census block shapes; select any other attribute data layers as well

There are three types of layers in the geopackage file – the "main.blocks" layer is the spatial features layer, and all other layers are either numerical attribute data tables, or the "fieldname_descriptions" metadata layer. The numerical attribute layers are named with the following format:

[mode]_[threshold]_minutes
- mode: two-letter code indicating the transport mode used - threshold: travel time threshold used for this data layer

To use the data spatially, perform a join between the "main.blocks" layer and the desired numerical data layer, using either the numerical "id" fields, or 15-digit "blockid" fields, as join fields.