

# User Guide & Codebook

yyyy.mm.dd\_abortionaccess\_countyxmonth

Caitlin Myers, Middlebury College

Last updated January 1, 2024

## User Guide

---

The data `yyyy.mm.dd_abortionaccess_countyxmonth`, uploaded to Open Science Framework (OSF) in Stata and CSV formats, is a panel of United States county-by-month distances to the nearest abortion provider. These data were prepared and published by Caitlin Myers at Middlebury College for academic research purposes, and have been regularly updated through the most recent indicated vintage published at OSF.

These public-use data are based on the Myers Abortion Facility Database,<sup>1</sup> a restricted-use dataset I document at OSF but only provide through an application process and data use agreement. This database identifies the names and addresses of all facilities—including private physician offices, hospitals, and freestanding clinics—that publicly advertised the provision of abortion services or are otherwise likely to be identifiable to a large fraction of women seeking an abortion. This database is not intended to capture private physicians and hospitals that provide a small number of abortions each year and do not advertise their services. The database covers the period January 1, 2009 through the most recent indicated vintage published at OSF.

All data files associated with the Myers Abortion Facility Database have a prefix in the form `yyyy.mm.dd` which indicates the vintage of the data. For instance, the latest vintage of the data as of the writing of this guidebook is January 1, 2024, and files from this most recent vintage have the prefix `2024.01.01`. Please cite the county-by-month distance data using the indicated citation at OSF:

Caitlin Myers. [yyyy]. County-by-month travel distances to nearest abortion provider, Vintage [Month, Day, Year]. Retrieved from [osf.io/pfxq3](https://osf.io/pfxq3).

These are county-by-month panel data, with the unit of observation defined by the county fips code (`origin_fips_code`) and monthly date (`monthlydate`). The variable `distances_origintodest` measures the travel (driving) distance from the origin county to the nearest abortion facility.

For each origin county, the origin point for the trip is defined as the population center of the county published by the United States Census Bureau.<sup>2</sup> The destination point for a trip is the geocoordinates of the abortion facility, which are calculated using their street addresses and the Stata `georoute` module, which relies on the HERE API.<sup>3</sup> Again using the `georoute` module, I identify the nearest facility by travel distance, which takes into account road networks and traffic conditions. A lack of road networks in some locations, most notably Alaska and Hawaii, prevents calculation of travel distances to some facilities. In those cases I instead calculate and

---

<sup>1</sup> Caitlin Myers. 2024. Myers Abortion Facility Database (Restricted), Vintage January 1, 2024. Retrieved from [osf.io/tj4ud](https://osf.io/tj4ud).

<sup>2</sup> United States Census Bureau. Centers of Population by County for the 2010 Census. Online dataset available at <https://www.census.gov/geographies/reference-files/2010/geo/2010-centers-population.html>.

<sup>3</sup> Sylvain Weber & Martin Péclat, 2016. "[GEOROUTE: Stata module to calculate travel distance and travel time between two addresses or two geographical points](#)," Statistical Software Components S458264, Boston College Department of Economics, revised 28 Apr 2020.

## Introduction

record geodesic distance using the Stata `geonear` module.<sup>4</sup> The data record the population of women aged 15-19 in the origin county based on data from SEER and the location (city, county, and state) of the nearest facility.

In addition to travel distance, the data also report the average service population (`asp`) at the destination city, a proxy for facility congestion developed and introduced by Lindo et al. (2019).<sup>5</sup> The average population is the total population of women aged 15-44 sharing a common destination city divided by the number of facilities in that destination city.

For the purpose of calculation ASP, destination cities are defined by the Core-Based Statistical Area (CBSA). If a destination is not part of a CBSA, it is alternatively defined by the county.

If you believe you have found an error in the data, please contact Caitlin Myers at [cmyers@middlebury.edu](mailto:cmyers@middlebury.edu).

## Codebook

---

-----  
`origin_fips_code`

Origin county fips code  
-----

Type: String (str5)

Unique values: 3,142

Examples:    "17067"  
              "26051"  
              "36115"  
              "47171"

-----  
`origin_county_name`

Origin county name  
-----

Type: String (str48)

Unique values: 3,142

Examples:    "Cook County (IL)"  
              "Hernando County (FL)"

---

<sup>4</sup> Robert Picard, 2010. "[GEONEAR: Stata module to find nearest neighbors using geodetic distances](#)," Statistical Software Components S457146, Boston College Department of Economics, revised 14 Sep 2019.

<sup>5</sup> Jason Lindo, Caitlin Myers, Andrea Schlosser, and Scott Cunningham. 2019. "[How far is too far? New Evidence on abortion clinic closures, access, and abortions](#)." *Journal of Human Resources* 55: 1137-1160.

## Introduction

"Meigs County (TN)"

"San Patricio County (TX)"

## Introduction

---

origin_state	Origin state
--------------	--------------

---

Type: String (str2)

Unique values: 51

Examples:     "AK"  
              "DC"

---

origin_population	Population of women aged 15-44 in origin county
-------------------	---

---

Type: Numeric (double)

---

monthlydate	Monthly date
-------------	--------------

---

Type: Numeric (float)

---

year	Year
------	------

---

Type: Numeric (float)

---

month	Month
-------	-------

---

Type: Numeric (float)

---

dest_fips_code	Destination provider county FIPS
----------------	----------------------------------

---

Type: String (str5)

Additional detail: This is the county fips code corresponding to the location of the nearest abortion facility.

## Introduction

Examples: "15009"  
"24510"  
"37063"

---

destination_county_name	Destination county name
-------------------------	-------------------------

---

Type: String (str48)

Unique values: 3,142

Additional detail: This is the county name corresponding to the location of the nearest abortion facility.

Examples: "Cook County (IL)"  
"Hernando County (FL)"  
"Meigs County (TN)"

---

dest_loc_id	Unique identifier for the locality of the nearest facility
-------------	--

---

Type: String (str5)

Additional detail: This uniquely identifies the “city” where the nearest abortion facility is located. It is a CBSA code if the city is part of a CBSA, and otherwise a county fips code.

---

dest_loc_name	Name of the locality of the nearest provider
---------------	--

---

Type: String (str46)

Additional detail: This uniquely identifies the “city” where the nearest abortion facility is located. It is a CBSA name if the city is part of a CBSA, and otherwise a county name.

Example: “Montgomery, AL”

---

dest_loc_type	Records whether nearest location is a CBSA or County
---------------	--

---

Type: String (str6)

Values: “CBSA”  
“County”

## Introduction

---

distance_origintodest	Distance from origin county to destination provider (miles)
-----------------------	---

---

Type: Numeric (float)

Additional detail: This is the driving distance in miles from the population centroid of the county to the geocoordinates of the nearest abortion facility calculated with the Stata georoute program. For a handful of origin-destination pairs, nearly all involving coordinates in Alaska or Hawaii, driving distances cannot be calculated due to limited or non-existent road networks. In these cases, identified by the distance\_type flag below, I substitute geodesic (“as the crow flies”) distance.

---

distance_type	Distance calculation: travel or geodesic
---------------	--

---

Type: String (str17)

Values: "Geodesic"  
"Travel (HERE API)"

---

dest_asp	Average service population in destination
----------	---

---

Type: Numeric (float)

Additional detail: This is defined by the destination CBSA and is equal to the sum of populations of women aged 15-44 in all counties for which the destination CBSA is the nearest location to obtain an abortion divided by the number of facilities in that CBSA.

.