

Chapter 3

Data

As an example we will map census tracts in Maryland and Baltimore City with limited access to COVID-19 vaccination sites using open access datasets.

3.1 Data

The following datasets were retrieved from different open access sources on April 4, 2021:

1. Food Access Research Atlas Data (Data Download 2015) (Excel Workbook format [.xlsx])
2. Census tract boundaries (ESRI shapefile collection [.zip])
3. List of COVID-19 vaccination sites in Maryland, with coordinates (comma separated values file, [.csv])
4. Maryland Physical Boundaries - County Boundaries (ESRI shapefile collection [.zip])

Access to Data

The datasets were archived in a zip compressed file (**SpatialAnalysisData.zip**) that can be downloaded at **SpatialAnalysisData** (URL: https://mega.nz/file/hYk02AyK#4knQ1zcaIxKTN_GX3JN8ZyOPD41XjynIbA3PvpvSaG4). The link will connect you to a cloud storage service (MEGA, <https://mega.nz>) and ask you to download the file. By accessing the cloud storage service and downloading the file/data you agree to the [terms of service](#) of MEGA and to the [terms of use](#) of the code and data.

Data Retrieval

1. Create a directory called **data** in your R project folder.
2. Click on the **SpatialAnalysisData** link (URL: https://mega.nz/file/hYk02AyK#4knQ1zcaIxKTN_GX3JN8ZyOPD41XjynIbA3PvpvSaG4).
3. Click on **Download** and save the file in the **data** directory. A zip compressed archive file (**SpatialAnalysisData.zip**; about 81.5 MB large) will be downloaded.
4. Unzip the archive file with the following R command:

```
destfile <- "data/SpatialAnalysisData.zip"  
unzip(destfile, exdir = "data/")
```

If you get a warning message, such as

```
Warning:  
In unzip(destfile, exdir = "data/") : error 1 in extracting from zip file"))
```

then the file may have not been downloaded into your directory.

Check:

- if you have a typo in the **destfile**,
- if you are in the correct working directory (with the command **getwd()**), or
- if the zip archive file is in the **data** directory of your working directory.

The latter can be checked with the R command **file.exists()**:

```
file.exists("data/SpatialAnalysisData.zip")
```

If the outcome is `[1] "FALSE"` then the file is missing. Find the file and move it into the **data** directory.

Once successful, you should have the following files in your **data** directory:

- FoodAccess2015.xlsx
- gz_2010_24_140_00_500k.zip
- gz_2010_24_140_00_500k.dbf
- gz_2010_24_140_00_500k.prj
- gz_2010_24_140_00_500k.shp
- gz_2010_24_140_00_500k.shx
- gz_2010_24_140_00_500k.xml
- Maryland_Physical_Boundaries_-_County_Boundaries_(Detailed).cpg
- Maryland_Physical_Boundaries_-_County_Boundaries_(Detailed).csv
- Maryland_Physical_Boundaries_-_County_Boundaries_(Detailed).dbf
- Maryland_Physical_Boundaries_-_County_Boundaries_(Detailed).prj
- Maryland_Physical_Boundaries_-_County_Boundaries_(Detailed).shp
- Maryland_Physical_Boundaries_-_County_Boundaries_(Detailed).shx
- Maryland_Physical_Boundaries_-_County_Boundaries_(Detailed).xml
- MD_counties_CT.dbf

- MD_counties_CT.prj
- MD_counties_CT.shp
- MD_counties_CT.shx
- MD_Covid19_VacSites_2021-04-04.csv
- SpatialAnalysisData.zip

3.2 Data Sources and Description

Food Access Research Atlas Data (2015 Dataset)

In 2008 the U.S. Congress directed the USDA to assess the extend of areas with limited access to fresh food (U.S. Department of Agriculture, 2009). Subsequently, the USDA Economic Research Service (USDA-ERS) developed a mapping tool called “the Food Desert Locator” (Ver Ploeg *et al.*, 2011). The Food Desert Locator allows to identify census tracts that qualify as food deserts [as defined by USDA, Treasury, and HHS (Ver Ploeg *et al.*, 2011)].

We are using this dataset as it lists (among more than 140 parameters) low-income and low-vehicle access census tracts.

The data is compiled into an Excel workbook (download size approx. 75 Mb). The workbook has three sheets, **Read me**, **Variable Lookup**, and **Food Access Research Atlas**. The latter contains the data aggregated based on data from the 2010 census.

Definitions and data sources can be accessed at <https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/>.

Source:	USDA Economic Research Service (USDA-ERS), available at https://www.ers.usda.gov/data-products/food-access-research-atlas/download-the-data/
Source URL	https://www.ers.usda.gov/webdocs/DataFiles/80591/DataDownload2015.xlsx?v=5276.5 , retrieved on Apr 4, 2021

Census Tracts

Food Atlas data is based on census tracts of the 2010 census. Census tracts are (relatively stable) statistical subdivisions of counties that ideally have 4,000 people (range: 1,200–8,000 people) (U.S. Census Bureau, 2019). To map, census tracts polygons are required.

Source:	Census Tracts for Maryland, United States Census Bureau, Cartographic Boundary Files – Shapefile, available at https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.2010.html
---------	--

URL (Maryland): https://www2.census.gov/geo/tiger/GENZ2010/gz_2010_24_140_00_500k.zip, retrieved on Apr 4, 2021

Census tract polygons are stored in an ESRI shapefile collection as a **zip** archive file. The **zip** file contains 5 files, a shapefile (**.shp**), 3 necessary auxiliary files (**.dbf**, **.prj**, **.shx**), and an additional **.xml** file. The **.xml** file is not necessary for spatial analysis.

List of COVID-19 Vaccination Sites

A list of COVID-19 vaccination sites in Maryland was retrieved from the Maryland Department of Health Coronavirus website (<https://coronavirus.maryland.gov>)

Source: All Maryland Vaccination Sites,
<https://coronavirus.maryland.gov/datasets/all-maryland-vaccination-sites-1>

URL (.csv) https://opendata.arcgis.com/datasets/d677f143334648a1a40b84d94df8e134_4.csv, retrieved on April 4, 2021

Credit: Maryland Department of Health COVID-19 Testing Task Force

Metadata: <https://www.arcgis.com/sharing/rest/content/items/d677f143334648a1a40b84d94df8e134/info/metadata/metadata.xml?format=default&output=html>

The dataset was slightly modified for educational purposes. An additional variable (column) — **Location** — was added and variables rearranged. The **Location** variable contains GPS coordinates of the vaccination sites in a (**latitude**, **longitude**) format. This format is often encountered in datasets that provide GPS information in tabular form.

Maryland Physical Boundaries — County Boundaries (Detailed)

Shapefile collection containing boundaries of Maryland counties.

Source: Maryland Physical Boundaries - County Boundaries (Detailed)
<https://data.imap.maryland.gov/datasets/maryland-physical-boundaries-county-boundaries-detailed?geometry=-80.968%2C38.061%2C-73.569%2C39.559><https://data.imap.maryland.gov/datasets/maryland-physical-boundaries-county->

URL (.zip) https://opendata.arcgis.com/datasets/2315ef0b071a4ec59420e3d342dbcfe2_0.zip?outSR=%7B%22latestWkid%22%3A3857%2C%22wkid%22%3A102100%7D, retrieved on Apr 4, 2021

Credit: MD iMAP, SHA

Metadata: <https://www.arcgis.com/sharing/rest/content/items/2315ef0b071a4ec59420e3d342dbcfe2/info/metadata/metadata.xml?format=default&output=html>

Maryland County Boundaries

Shapefile collection containing boundaries of Maryland counties. Based on unified census tract boundaries (source of the original data: Census Tracts).