**hmda\_harmonizer supplementary sources:**

To execute the hmda\_harmonizer script, the user must download the following files and place them in subfolders with names specified by the section headers.

The hmda\_harmonizer.do file automatically processes zip files when needed – no further action required beyond downloading these files and putting them in properly named folders.

**hmda\_lender\_panels:**

1. Downloading 2010-2017 data:
   1. Go to the link here: <https://www.consumerfinance.gov/data-research/hmda/historic-data/>
   2. Find the grey box titled “Select the HMDA data you want,”
   3. Set “geographic area” to “Nationwide”
   4. Set “records included” to “All records”
   5. Set “variable descriptions” to “Plain language labels and HMDA codes”
   6. Click “search” to ensure these filters are applied
   7. Scroll to “Related nationwide institution data,” find the “Panel” column, and click the download links to each year’s CSV files.
2. Downloading 2018-2020 data:
   1. Repeat these instructions for each year, 2018-2020
      1. Go to the link here: <https://ffiec.cfpb.gov/data-publication/snapshot-national-loan-level-dataset>
      2. Under “Select a year,” click the given year
      3. Under “Reporter Panel,” select the CSV version of the file

Every file from 2010-2017 is downloaded as a zipped folder. 2018-2020 files are downloaded as CSVs. All are saved to the directory hmda\_loan\_application\_register/input/original\_hmda.

In total, we download files called:

1. hmda\_yyyy\_panel (yyyy = years 2007-2017) – these are the 2007-2017 lender panels
2. yyyy\_public\_panel\_csv (yyyy = years from 2018-2021) – these are the 2018-2021 lender panels

**avery\_file:**

1. Go to the link here: <https://sites.google.com/site/neilbhutta/data>
2. Scroll down to the header “HMDA lender file (a.k.a. ‘The Avery file’)” and click both links, next to “Zipped SAS and Stata datasets, and documentation 1990-2017” and “Zipped Stata datasets and documentation, 2018-19”
3. On each Dropbox tab that opens, click the “Download” button at top left to download the complete folder at each link

*Continued on next page*

**nic:**

1. Go to the link here: <https://www.ffiec.gov/npw/FinancialReport/DataDownload>
2. Download all 5 of the zip files listed under “CSV Download”:
   1. “Attributes – Active”
   2. “Attributes – Closed”
   3. “Attributes – Branches”
   4. “Relationships”
   5. “Transformations.”
3. Optional: click the expandable header “About the Files,” and click on the link “Data Dictionary” to download the documentation.

**hmda\_to­\_lei\_xwalk:**

1. Go to the link here: <https://ffiec.cfpb.gov/documentation/2017/identifiers-faq/>
2. Scroll down to the section called “How can I map 2017 institution identifiers to 2018 institution identifiers?”
3. Click the link marked “ARID2017 to LEI Reference Table (CSV)”

**See next page for optional instructions for finding and downloading the HMDA loan-level data.**

**Optional: HMDA loan-level data.**

The following are instructions for finding and downloading the HMDA loan-level data. This is not required for the construction of the HMDAHarmonizer panel itself. But, since the primary function of the panel is to use it with the loan-level dataset, below are instructions for finding the loan-level datasets.

1. Downloading 2007-2017 data:
   1. Go to the link here - <https://www.consumerfinance.gov/data-research/hmda/historic-data/>
   2. LAR data
      1. Find the grey box titled “Select the HMDA data you want,”
      2. Set “geographic area” to “Nationwide”
      3. Set “records included” to “All records”
      4. Set “variable descriptions” to “Plain language labels and HMDA codes”
      5. Click “search” to ensure these filters are applied
      6. Scroll to “Showing nationwide records,” and click the download links to each year’s CSV files.
2. Downloading 2018-2021 data:
   1. LAR data - repeat these instructions for each year, 2018-2021
      1. Go to the link here - <https://ffiec.cfpb.gov/data-publication/dynamic-national-loan-level-dataset/2018>
      2. Under “Select a year,” click the given year
      3. Under “[year] Dynamic Datasets,” click the link marked “Loan/Application Records (LAR).

Note that the link above sends the user to a page to download the Dynamic National Loan Level Dataset. There is also the Snapshot National Loan Level Dataset. I personally have not found official documentation explicitly distinguishing between the two, but my understanding is that the Dynamic version is updated weekly with filer resubmissions, whereas the Snapshot version is updated less frequently/is subject to a “data freeze” after a certain date.