A Beginner’s Guide to Accessing and Using Home Mortgage Disclosure Act Data
1. Purpose

Mortgage lending data reported in accordance with the Home Mortgage Disclosure Act (HMDA), commonly called “HMDA data,” is the largest source of publicly available data on mortgage lending in the United States.

The purpose of this guide is to introduce this data to potential users who have little to no experience with handling and analyzing raw data. This guide covers the basics of what HMDA data are and how to access the data, as well as a step-by-step guide for using HMDA data. The guide will instruct the beginner user how to find and download HMDA data, select subsets and filters for the data, and begin to analyze the HMDA data using pivot tables, grouping data together, and matching data across datasets. For these exercises, you will need a computer with internet access and Microsoft Excel.

The methods described are examples only of how to understand the HMDA data and do not describe all the ways the data can be used and understood. In addition, we caution against drawing legal conclusions from a particular analysis suggested in this guide. Legal compliance may depend on a variety of factors that may not be accounted for in an analysis provided using the methods described here.
2. What are HMDA data?

The Home Mortgage Disclosure Act requires financial institutions to maintain, report, and publicly disclose loan-level information about mortgages. These data help show whether lenders are serving the housing needs of their communities; give a range of stakeholders information that helps them make recommendations, decisions and policies; and shed light on lending patterns that could be discriminatory. The CFPB modifies publicly released data to protect applicant and borrower privacy.

HMDA was originally enacted by Congress in 1975 and is implemented by Regulation C.¹ HMDA has been amended by Congress several times since it was enacted, and federal agencies have issued multiple regulatory changes to Regulation C as well. Collectively, the Board of Governors of the Federal Reserve System (FRB), the Federal Deposit Insurance Corporation (FDIC), the National Credit Union Administration (NCUA), the Office of the Comptroller of the Currency (OCC), and the CFPB comprise the Federal Financial Institutions Examination Council (FFIEC), a governmental interagency body that facilitates public access to the data.

Over the years, amendments to HMDA and agency rulemakings have changed which financial institutions are required to report, and the data they are required to report under HMDA.

2.1 Where to find HMDA data

HMDA data is available at https://ffiec.cfpb.gov/. They can be accessed by using:

- The **HMDA Data Browser**; the entire static HMDA dataset or custom datasets and summary tables are available at this site.

- The **Dynamic National Loan Level Dataset**.

These resources help provide additional background on the HMDA data:

- **HMDA Public Data Fields with Values and Definitions** provides an overview of the fields and the content.

- **Public LAR Schema** indicates the maximum length of each field.

¹ 12 C.F.R. Part 1003.
2.2 Who reports HMDA data

HMDA requires financial institutions², including depository (e.g. banks) and non-depository (e.g. non-bank mortgage companies) institutions, to report HMDA data. However, not every institution that issues or originates a home mortgage is required by HMDA to report its mortgage data.

For example, an institution must have had a home or branch office location in a metropolitan statistical area (MSA) on the preceding December 31st in order to be covered by HMDA’s requirements.³ In this example, a non-depository financial institution is deemed to have a branch office in an MSA if, in the preceding calendar year, it received applications for, originated, or purchased five or more HMDA-reportable loans related to property located in that MSA, even if it does not have an office in that MSA.⁴

Similarly, the depository or non-depository institution must have originated at least 100 closed-end mortgage loans in each of the two preceding calendar years or have originated at least 200 open-end lines of credit in each of the two preceding calendar years in order to be covered by HMDA’s reporting requirements.

Other institutional coverage rules apply. A chart summarizing HMDA institutional coverage is available here. In addition, for those institutions that are covered, only loans that meet Regulation C’s transactional coverage requirements must be reported.⁵ A chart summarizing HMDA transactional coverage is found here.

² Regulation C defines “financial institutions” at 12 CFR § 1003.2(g).
³ 12 CFR 1003.2(g)(1)(ii).
⁴ 12 CFR 1003.2(c)(2).
⁵ Regulation C’s transactional coverage criteria is generally found within the definition of “Covered Loan,” located at 12 CFR 1003.2(e) and the associated commentary.


2.3 HMDA data points

Financial institutions subject to HMDA’s requirements report up to 110 different data points for each mortgage application they receive. Collectively, the data points for each application are referred to as a “loan/application register” or “LAR.” Among the data points are loan type, loan purpose, demographic information of the borrower or co-borrowers including race, ethnicity, sex and age, the location of the dwelling, the action the financial institution took on the application, the loan amount, the interest rate, any points and fees charged in connection with the loan, and the property value.

For some transactions, a certain data point may not be applicable. For example, if a borrower withdrew a mortgage application before the institution made a credit decision, the institution would report the code for “not applicable” for the interest rate data point. HMDA also provides that certain financial institutions may be eligible for partial exemptions for certain transactions. If a financial institution is eligible for a partial exemption for a specific transaction, then the financial institution is not required to collect, record, or report certain data points for the transaction.

2.4 Public access to HMDA data

Financial institutions are required to submit annual HMDA LAR in electronic format by March 1 of the year following the calendar year for which the data are collected. The CFPB releases the data to the public on or before March 31 following data collection.

In order to protect borrower and applicant privacy, the publicly-released data exclude or modify several data points reported by financial institutions, such as the universal loan identifier, the date the application was received or the date shown on the application form, the address of the property, the credit score or scores relied on in making the credit decision, and applicant borrower or ethnicity free-form text field.

---

6 12 CFR 1003.5(a)(1)(i).
7 CFPB, Disclosure of Loan-Level HMDA Data, 84 FR 649 (Jan. 31, 2019). Note that the above is not a comprehensive list of excluded data points.
3. Step-by-Step Guide

Below is a sample exercise to explain how to find, download, and begin to analyze available HMDA data. You’ll need access to the internet and Microsoft Excel to follow this example.

In this exercise, we will use HMDA data to find out how many applications HMDA reporters in Birmingham, Alabama received from minority neighborhoods. For the purpose of this exercise, a minority neighborhood is a census tract where at least 50% of the residents are identified by the U.S. Census Bureau as a racial or ethnic minority (i.e., Black or African American, Hispanic, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, or some other race)

3.1 Using the Data Browser to filter and download a dataset

STEP 1: Go online and find the HMDA Data Browser at [https://ffiec.cfpb.gov/data-browser/](https://ffiec.cfpb.gov/data-browser/).

STEP 2: Click “Select, Summarize, Download”
STEP 3: Select the year for which you want HMDA data. For our example, click 2020.

STEP 4: Select a Geography, by State, County, Metropolitan Statistical Area/Metropolitan Division (MSA/MD), or Nationwide. For our example, select “MSA/MD” and start typing “Birmingham,” and you will see the Birmingham-Hoover-Alabama MSA appear for selection.

STEP 5: Select one or more financial institutions, or leave it blank to see all financial institutions that reported HMDA Data in that geography. For our example, leave this blank.

STEP 6: Select any filters. The filters you apply will directly affect the results you receive. When selecting which filters to apply, you may consider:

- Am I interested only in originated loans? Or all applications? Or all HMDA records including preapprovals and post-origination secondary market loan purchases? The Action Taken filter will be helpful if these are concerns.
- Am I focused only on conventional loans? Or Federal Housing Administration (FHA), Department of Veterans Affairs (VA), or Rural Housing Service (RHS) loans? Find the Loan Type filter if these questions are important to your analysis.
- Am I interested in HMDA records only for purchasing a home? Or refinancing, home improvement loans, or loans for other purposes? The Loan Purpose filter is helpful here.
- Do I want to analyze only manufactured housing in my analysis? Find the Construction Method filter for this.

For our example, both “Action Taken” and “Loan Purpose” filters are selected.

Selections for each filter will appear. For our example, under the Action Taken column, select Loan Originated, Application approved by not accepted, Application denied, Application withdrawn by applicant, and File closed for incompleteness. Under the Loan Purpose column, select Home Purchase.
STEP 7: Click “View Summary Table” to see a summary of the data you are about to download. Make sure that the total record count does not exceed Excel’s capacity (which is normally a few hundred thousand records). Beyond this number, you will need other software to download and analyze the data. Here is the summary that will appear, based on the choices you made thus far:

![Data Summary Table]

The filtered data contains **28,000** rows, each with all 99 public data fields.

STEP 8: Next, you can click the “Download Dataset” button to download the dataset to conduct further analysis. The default download is into a comma separated values (CSV) file that can be opened and used in Excel and similar types of software. Note that it may take a few minutes to download the file.

![Download Dataset Button]

The filtered data contains **28,000** rows, each with all 99 public data fields.
3.2 Using Excel to understand your data: Introduction to pivot tables

One way to work with large amounts of HMDA data in a manageable way is with “pivot tables” in Microsoft Excel. There are many online resources to learn how to use Excel, including free tutorial videos.

STEP 1: Open the CSV file you just downloaded by clicking on Excel file that appears in the lower left hand of your screen after you click the “Download Dataset” button.
An Excel file should open that looks like this:

![Excel file screenshot]

8 Note that different versions of Microsoft Excel may look slightly different from the images in this guide.
STEP 2: Click on the triangle in the upper left-hand corner of the table. This will select all of the data in the file.

STEP 3: Click “Insert”
STEP 4: Click “Pivot Table.”

You will see a dialog box pop up:
If you are new to using pivot tables, use all of the pre-selected default settings and click “OK.” With experience, you may want to consider changing some of these settings.

The pivot table will open in a new tab and will look like this:
STEP 5: Using the PivotTable Fields section appearing along the right side of your screen, select and drag the fields you want into the rows, columns, and values section to see summary tables of your dataset.

For this example, we want to determine the number of applications that each lender in the Birmingham MSA received from minority neighborhoods. To do so, you will need the unique identifier for each financial institution—called the Legal Entity Identifier (LEI). To add it, click on the “lei” field and drag it into the “Rows” box below.

To find whether an application came from a minority census tract, select the “tract_minority_pop_percent” field and drag that into the “Columns” box.

Finally, you will need to select the “Values” you want displayed in the pivot table. Here, we are going to select LEI because each application will have that value. Click and drag “lei” to the Values box.

The Custom Name should pre-populate with “Count of lei” (as in the photo immediately above) but if it displays something different, click on the small down arrow to the right and select “Value Field Settings.” A box will appear. Select “Count” from the list of options and click “OK.”
3.3 Grouping in Excel

At this point, you should see a table something like the one below. Each column represents a census tract identified with its percentage of minority residents. Each row represents a mortgage lender who reported loans for the Birmingham MSA. The values in the table represent how many applications were received and loans were originated in each census tract by each lender.

To better visualize the data, you can group data together in Excel. In this example, we will group the “tract minority population” columns to analyze whether loan applications are coming from low minority neighborhoods (0-49% minority), majority minority neighborhoods (50-79% minority), or high minority neighborhoods (80-100% minority).
STEP 1: Select all of the values that you want to group together. In this case, for the first group that includes census tracts for what we are calling low minority neighborhoods (meaning 0-49% minority), click and drag your cursor on the values in the “Row Labels” rows from 0 all the way up to (but not including) 50. Right click on the selected group, and select “Group.”

A new row will appear with a place to name that group.

Click on the cell that says “Group1” and insert your new group name. You may want to refer to it as “0-49% Minority.” Click on the minus sign to collapse the table and consolidate all of the values that you just grouped together into the new “0-49% Minority” column.
STEP 2: Repeat this step with the other minority groupings for 50-79% and 80-100%. This will result in a list of each financial institution (by LEI) and the number of applications they received from minority neighborhoods.

However, the HMDA data only have the legal entity identifier (LEI), not the institution name.

3.4 Using the VLOOKUP function in Excel to match different datasets

In order to see the names of all the institutions, you may use the VLOOKUP function, which matches data in your dataset with another.

STEP 1: First, you need to find the dataset with institution names that you want to use to match to the LEI in your spreadsheet. This can be found in the HMDA transmittal sheet. Go back to https://ffiec.cfpb.gov/ and select “Data Publication.”
Then select “Snapshot National Loan-Level Dataset.”
Select the year (and make sure that you are using HMDA data from the same year). For our example, click “2020.”

Select the Transmittal Sheet (TS) CSV file to download.

This file will download into a zip file. Open and unzip the file.
STEP 2: Move the new file into the original file you used to build your pivot table as a new tab. Do that by right clicking on the tab at the bottom of the spreadsheet labeled “2020_ts_csv.” Then click “Move or Copy.”

For the top box, you want to move the selected sheet to book “msa_13820 ...” For the “Before sheet” box, it does not matter where you move it to as long as you can keep track of the sheets, but for ease, choose “move to end.” Click “OK.”
STEP 3: In the original data sheet, called “msamd_13820 …”, insert a blank column where you want to add the name of the financial institution. Right click on the column, select “Insert Column.”

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>activity_y:lei</td>
<td>Institution Name</td>
<td>derived_r_state_cod</td>
<td>county</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2020 549300PL8ER6H23P0Z91</td>
<td>13820</td>
<td>AL</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

STEP 4: Name the column (e.g., “Institution Name”).

STEP 5: Begin typing the VLOOKUP statement (=VLOOKUP) in the first empty cell below the column header.

When you enter the formula into Excel, you will need to include four parts:

1. Lookup value (what field are you using to match between the datasets)
2. Table array (where should Excel look in order to find the new data)
3. Column index number (which column should the new data come from)
4. Range lookup (you can just write “false” here until you get more advanced with Excel)

STEP 6: Select as the lookup value the “lei” (Legal Entity Identifier) by simply clicking on the B2 cell (or by typing “B2”). Type a comma and a space.
STEP 7: Select the Table Array – Navigate to the “2020_ts_csv” sheet and select the columns where Excel should be matching the LEI and the institution’s name. For our example, select four columns, from “lei” to “respondent_name,” by clicking on the “C” and drag your cursor to “F.” Type a comma and a space.
STEP 8: Now add the column of those selected where the name can be found, in this case column 4 by typing “4”. Type a comma and a space.

For the last part of the VLOOKUP, just type “FALSE.” Your complete command should read:  

Hit enter to submit the formula and you will be able to see the matched entity name.

STEP 9: Paste the formula into the entire column to match all names. Right click on the cell with the formula (cell C2) and select “Copy” (note: you can also copy by typing Ctrl + C)
Then right click on the column header “C” and select the Paste icon.

You should now have the name of the institution in column C.
STEP 10: To update your pivot table so that the names of the institutions appear in place of the LEIs, go to the pivot table sheet (here called “Sheet1”). Right click anywhere on the table and select “Refresh.”

STEP 11: Click on your new “name” field (called “respondent_name”) and drag it to the “Rows” box to replace the LEI with the institution names.
You may need to click on the small black arrow next to “lei” in the Rows box and select “Remove field” in order to remove the “lei.”

Your final pivot table should look something like the below.

This table may be more user-friendly than the one downloaded originally. The rows of the table represent each HMDA-reporting financial institution with mortgage lending activity in the Birmingham MSA during 2020. The grouped columns are low-, majority-, and high-minority census tracts, and the table numbers represent the number of mortgage applications and originations in each group of tracts.
4. Conclusion

We hope users have found this guide useful in understanding how HMDA data may be accessed and analyzed using widely available technology tools. In addition to the techniques described here, we encourage users to explore the many resources provided for HMDA reporters and users alike. FFIEC’s [HMDA Maps tool](https://www.consumerfinance.gov/hmda-maps/) allows you to explore and visualize online subsets of HMDA data, filtered by popular variables. The CFPB also has an array of webinars for HMDA data users available on YouTube at [https://www.youtube.com/user/cfpbvideo/videos](https://www.youtube.com/user/cfpbvideo/videos).

Given the importance of accurately reported HMDA data to the CFPB’s fair lending mission, the CFPB maintains a comprehensive suite of resources on its public website to help filers fulfill their reporting requirements under HMDA and Regulation C and allowing others to evaluate and study mortgage lending. These resources include: an Executive Summary of HMDA rule changes;[^9] Small Entity Compliance Guide;[^10] Key Dates Timeline;[^11] Institutional and Transactional Coverage Charts;[^12] Reportable HMDA Data Chart;[^13] sample data collection form;[^14] and FAQs,[^15] in addition to downloadable webinars,[^16] which provide an overview of the HMDA rule. The CFPB also provides on its website an interactive version of Regulation C that is easier to access and navigate than the printed version of Regulation C.[^17]


Together with the FFIEC, the CFPB also routinely updates its HMDA resources throughout the year to ensure HMDA reporters have the most up-to-date information. The agency also works with the FFIEC to publish data submission resources for HMDA filers and vendors on its Resources for HMDA Filers website.

In addition, HMDA users and reporters can ask questions about HMDA and Regulation C, including how to submit HMDA data, by emailing the CFPB’s HMDA Help at HMDAHelp@cfpb.gov. The agency also offers financial institutions, service providers, and others, informal staff guidance on specific questions about the statutes and rules the CFPB implements, including ECOA and Regulation B and HMDA and Regulation C, through its Regulation Inquiries platform.