**Urban Institute National Data Repository**

**Home Mortgage Disclosure Act Data Series Documentation**

Years of data available from UI: 1997-2011

Source data release month: September

Source data release frequency: Annual

Source data URL: <http://www.ffiec.gov/hmda/>

**Overview**: The Home Mortgage Disclosure Act (HMDA) requires most lending institutions to report mortgage loan applications, including the outcome of the application, information about the loan and applicant, and location of the property. FFIEC collects the data in order to determine whether financial institutions are meeting a community’s housing credit needs; to target community development funds to attract private investment; and to identify possible discriminatory lending patterns. UI has summarized the loan-level data for various geographic levels into indicators on the racial and income distribution of borrowers, denial rates by race and income, and loans from subprime lenders by race. In 2004, FFIEC expanded the data to include structure type, lien status, and the interest rate spread.

These data files were initially designed by the Urban Institute (UI) to be displayed on the DataPlace web site. The site was first national community indicators portal with dynamic maps, charts, and profiles. After the closing of the Fannie Mae Foundation, DataPlace was spun off to an independent nonprofit, which then closed in 2009. The Urban Institute is no longer associated with the website.

**Documentation**: For a detailed list of fields and license information, see [Excel metadata file](http://www.neighborhoodindicators.org/library/catalog/ui-national-data-repository-home-mortgage-disclosure-act-metadata). [*A Guide to Home Mortgage Disclosure Act Data*](http://www.urban.org/publications/1001247.html)provides more information about the source data, UI’s methodology, interpretation of indicators, and examples of local use of the indicators. Below we include descriptions related to geography and selected variables to help the user interpret and use the data.

**Licensing:** The data are licensed under the Open Database License. See [license text](http://www.neighborhoodindicators.org/open-data-base-license) for the exact terms of use. You are free to share and adapt these datasets with the following conditions:

* For any public use or redistribution of the data files, or works produced from them, you must make clear to others the license (ODbL) of the data files and keep intact any notices on the original data files (see citation requirements).
* Public use of any adapted version of these data files, or works produced from an adapted database must also be offered under the ODbL.

*Required Attribution*: The [UI HMDA Summary data files](http://www.neighborhoodindicators.org/urban-institute-data-repository) and the procedures for constructing them were initially developed by the Urban Institute to support [DataPlace](http://www.dataplace.org/) (Hyperlinks or plain text of links required).

**Notes on data structure**: The data includes records for all of the geographic summary levels (tract, place, etc.). Each level identified by the formatted field GEOSCALEID (see geography section below). The data file uses semi-colons as delimiters, which was necessary because the text fields contained commas.

**Notes on geographies:** All tract data are for 2000 census tract definitions. For 1997-2002 data, the published data used 1990 census tract definitions, which were converted to 2000 definitions using the Neighborhood Change Data Base weighting system. Place codes are assigned through a tract-to-place crosswalk, so smaller cities/towns are not included in the place-level data. The data files for 2004 and after include geographies (census tract through state-level) for Puerto Rico.

|  |  |
| --- | --- |
| **Geographic Levels** | **GEOSCALEID** |
| Census tracts | 1 |
| Places | 2 |
| Counties | 3 |
| States | 4 |
| County | 6 |
| Census Divisions | 8 |
| Census Regions | 9 |
| Post-2003 Metros | 11 |
| Post-2003 Metro Divisions | 12 |

**Notes on the relative income variables:** From 1997 to 2003, the Federal Financial Institutions Examinations Council (FFIEC) did not provide the HUD area median family income (HAMFIs). To calculate the income category for a given borrower, UI merged the HAMFI file obtained from the [HUDUser website](http://www.huduser.org/portal/datasets/il.html). We continued this practice through the 2011 file, even though the HAMFI was included on the source HMDA data since 2004.

A recent diagnostic analysis showed differences between our merged HAMFI figure and the FFIEC-provided HAMFI figure for 33 percent of owner-occupied home purchase and refinancing loans. In 11 percent of the cases, the difference exceeded ten percent. The median percent difference was about zero, so half of the time the UI figure was larger than the FFIEC-provided number and half of the time it was smaller. Further research showed that the FFIEC indicators did not match the HUDUser public file. For example, on the 2011 file, the HAMFI assigned by FFIEC to the District of Columbia (based on the metropolitan median) was $104,300. The median income on the HUDUser site for FY2011 is $106,100. Montgomery County, MD was assigned a median income higher than its metro - $111,900. We have submitted an inquiry into FFIEC about their methodology for assigning HAMFIs and will update this document when we have more information.

We have posted the data files with the counties and the income limits used for each year for your reference on the [UI Dataverse web site](http://www.neighborhoodindicators.org/urban-institute-data-repository).

**Notes on Decennial Census/ACS variables:** Beginning with the 2010 file, we included a series of variables based on the [2000-to-2010 census tract correspondence data](http://www.census.gov/geo/maps-data/data/relationship.html) and the American Community Survey 2005-2009 estimates, and no longer included the 2000 figures. Please see the full [metadata](http://www.neighborhoodindicators.org/library/catalog/ui-national-data-repository-home-mortgage-disclosure-act-metadata) for more detail.

|  |  |  |
| --- | --- | --- |
| **Field name** | **Field label** | **Data Source** |
| MedianHshldIncomeadj | Median household income 2000, adjusted for inflation | Decennial Census 2000, long-form |
| MedianOwnerHshldIncomeadj | Median owner household income 2000, adjusted for inflation | Decennial Census 2000, long-form |
| |  | | --- | | MedianHshldIncomeadj2009 | | Median household income 2005-2009, adjusted for inflation | American Community Survey, 2005-2009 Five Year data file |
| MedianOwnerHshldIncomeadj2009 | Median owner household income 2005-2009, adjusted for inflation | American Community Survey, 2005-2009 Five Year data file |
| NumHsngUnitsTotal | Number of housing units, 2000 | Decennial Census 2000, long-form |
| NumHsngUnits1\_4Fam | Number of housing units, 1 to 4 families (HMDA def.), 2000 | Decennial Census, long form, 2000 |
| NumHsngUnits5plusFam | Number of housing units, 5 or more families (HMDA def.), 2000 | Decennial Census, long form, 2000 |
| NumHsngUnitsTotal2010 | Number of housing units, 2010 | Census 2000-to-2010 tract correspondence file |
| NumHsngUnits1\_4Fam2010 | Number of housing units, 1 to 4 families (HMDA def.), 2010 | Census 2000-to-2010 tract correspondence file, American Community Survey 2005-2010 |
| NumHsngUnits5plusFam2010 | Number of housing units, 5 or more families (HMDA def.), 2010 | Census 2000-to-2010 tract correspondence file, American Community Survey 2005-2010 |

**Notes on variable series over time:** Because of the original DataPlace technical requirements, the subprime lender fields (derived from a list of subprime lenders published on [HUDUser](http://www.huduser.org/portal/datasets/manu.html) through 2005) are present in all of the years, but were only calculated up to 2006. The list was discontinued by HUD because the new interest rate spread variable that was introduced with the 2004 HMDA data better identified subprime loans. We used the 2005 list for the 2006 data, but discontinued this series in later years because of the massive changes in the mortgage industry.

Since these files were developed over time, we have made additions to the variable set over the years. In the case of new variables, we did not have the resources to go back and calculate the fields for earlier years. Please see the full [metadata](http://www.neighborhoodindicators.org/library/catalog/ui-national-data-repository-home-mortgage-disclosure-act-metadata) for more detail.

**Changes to the UI HMDA Summary Data File**

|  |  |  |
| --- | --- | --- |
| **Series reference** | **Series Description** | **Years** |
| Various | Many sections related to lien position and structure type enabled by major revisions to the source file.  Indicators marked as “Post\_2004” in the metadata | 2004-present |
| h39 | Owner-occupied home purchase originations by income and race (1-4 + manufactured housing) | 2007 – present |
| h40 | First-lien owner-occupied home purchase originations by income and race (1-4 + manufactured housing) | 2007 - present |
| H41 | First-lien applications by 2004 purpose\*structure type | 2009 - present |
| H15A (expanded) | First-lien investor home purchase loans  Added:  MrtgFirstOrigPurchNotOwn1\_4m  MrtgFirstOrigPurchNA1\_4m | 2009 - present |

**Other Resources:** For more examples of local use of HMDA data, see the [HMDA section of the National Neighborhood Indicators Partnership](http://www.neighborhoodindicators.org/data-source/4) website.

Documentation Date: 7-16-2013

Documentation Staff: Kathy Pettit