Uniform Appraisal Dataset Appraisal-Level Public Use File

Version 1.0

Data Documentation

October 16, 2023

Table of Contents

Federal Housing Finance Agency

1. Overview

The Uniform Appraisal Dataset (UAD) Appraisal-Level Public Use File (PUF)—hereinafter referred to as the UAD Appraisal-Level PUF—is a data file created for public users on a nationally representative five percent random sample of single-family appraisal records associated with mortgages acquired by Fannie Mae and Freddie Mac (the Enterprises). It contains selected information from the single-family UAD merged with race and ethnicity of borrowers obtained from mortgage documents associated with the appraisals in the UAD. The Federal Housing Finance Agency (FHFA) created the UAD Appraisal-Level PUF from the data obtained from the Enterprises to allow public users the flexibility to produce numerous estimates of interest while reducing the likelihood that they could re-identify a borrower, property owner, or appraiser.

This document's purpose is to provide users with background information on how FHFA created the UAD Appraisal-Level PUF. It includes information about the universe of UAD appraisal records, the subset of UAD appraisal records eligible to be included in the UAD Appraisal-Level PUF, disclosure avoidance, sample design, geographic areas, file structure, and information quality.

This document is designed to work in tandem with the UAD Appraisal-Level PUF Data Dictionary. The Data Dictionary contains the full list of data fields, their allowable values, additional information about disclosure avoidance, and helpful notes for users.

2. About the UAD

The UAD is a standardized industry dataset for appraisal information that the Enterprises receive electronically through the Uniform Collateral Data Portal® (UCDP®). The Enterprises, at the direction of FHFA, developed the first UAD in 2010 using the Mortgage Industry Standards Maintenance Organization (MISMO®) Reference Model 2.6. MISMO models provide consistent data standards for mortgage loans acquired by the Enterprises.[1]

A UAD appraisal record includes information collected by appraisers using the Uniform Residential Appraisal Report (URAR). The current version of the URAR for one-unit single-family homes is Fannie Mae Form 1004 and Freddie Mac Form 70.[2,3] Other federal agencies also use the URAR.

Not all mortgage loans acquired by the Enterprises will have a UAD appraisal record. For instance, both Enterprises waive traditional appraisals in certain instances where they deem automated appraisal methods as adequate. Thus, UAD appraisal records include only mortgage loans requiring traditional appraisals. The UAD appraisal records that the Enterprises collect may include some appraisals related to other lending sources, including Federal Housing Administration (FHA) and portfolio loans, as well as

---

[1] This information is sourced from the Enterprises websites describing the UAD: https://singlefamily.fanniemae.com/delivering/uniform-mortgage-data-program/uniform-appraisal-dataset and https://sf.freddiemac.com/tools-learning/uniform-mortgage-data-program/uad.
[2] Single family includes detached, attached, or semi-detached structures and associated lots.
[3] There are similar forms for single-unit condominiums, manufactured homes, and small multifamily residential income (rental) properties. However, only single family and condo currently utilize UAD standardization and mapping.

appraisal not connected to any mortgage loan, such as those associated with a denied loan application. As noted in Section 3, these UAD appraisal records are not included in the UAD Appraisal-Level PUF.

FHFA receives UAD appraisal records from the Enterprises. The Agency currently possesses records from 2013 through the most recent quarter. As of June 2023, this collection includes more than 80 million records about the "subject" property and more than 360 million records about "comparable" properties. However, it is important to note that not all these records reflect a property's final appraisal(s). The UCDP is a transactional business system. As such, it contains UAD appraisal records that may be incomplete or out of date.

3. Universe of UAD Appraisal Records Eligible for UAD Appraisal-Level PUF

3a. Overview of Universe for the UAD Appraisal-Level PUF

In creating the UAD Appraisal-Level PUF, FHFA considered the universe of *unique* appraisals for mortgage loans acquired by the Enterprises. In most instances, a loan has only one appraisal. In other instances, a loan has a second appraisal. However, the UCDP is a transactional business system containing far more appraisal records than "final" appraisals. To build the UAD Appraisal-Level PUF, it was important to keep a unique appraisal for each loan by carefully removing incomplete or out-of-date appraisals.

FHFA staff, in consultation with UAD experts at the Enterprises, developed a protocol to identify the final appraisal record(s) for each loan. This protocol is described in detail in Appendix 1.

Once FHFA determined the final appraisal record(s) for each loan, they applied additional criteria for inclusion and exclusion. The inclusion and exclusion criteria, also described in Appendix 1, are as follows:

1. Property appraisals conducted between January 1, 2013, and December 31, 2021 are included. Property appraisals conducted after December 31, 2021 are excluded, but FHFA intends to add them to the UAD Appraisal-Level PUF in the future.[4]
2. Appraisals for single-family properties appraised using Fannie Mae Form 1004 or Freddie Mac Form 70 are included. Appraisals for condominiums, manufactured homes, and small multifamily rental properties as well as other appraisals are excluded.
3. Appraisals for mortgage loans acquired by the Enterprises are included. Appraisals for loans not acquired by the Enterprises—such as denied loans and loans funded by other institutions—are excluded.
4. Appraisal records with the purchase or refinance transaction type that is inconsistent with the mortgage purpose code of FHFA's Enterprise mortgage acquisition data are excluded.

---

[4] Typically, there is a gap of 1-3 months from an appraisal to loan origination, another gap of 1-6 months between loan origination and loan acquisition by the Enterprises, and a gap of about 45 days from loan acquisition to data delivery. Due to this reason, the mortgage data for 2022 originations, including the information on whether mortgages associated with the appraisal were originated and acquired by the Enterprises, are not yet fully available.

5. For appraisals with purchase transaction type, only appraisals for arm's length transactions are included. Appraisals for real estate owned (REO), short sale, and foreclosure purchase transactions are excluded.
6. Appraisal records for the properties in Guam and the U.S. Virgin Islands are excluded because of insufficient record counts.

3b. Comparison with UAD Aggregate Statistics Data File

FHFA created and continues to maintain a UAD Aggregate Statistics Data File. The Aggregate Statistics Data File includes millions of statistics aggregated from UAD appraisal records. Users of both the Aggregate Statistics Data File and the Appraisal-Level PUF might wonder whether aggregate statistical estimates created using the Appraisal-Level PUF will match similar estimates in the Aggregate Statistics Data File.

The answer is "usually not". The primary reason estimates will differ is the universe of UAD records eligible for inclusion in each product is different. In fact, the universe of UAD records eligible for inclusion in the UAD Appraisal-Level PUF is smaller than the universe of UAD records eligible for inclusion in UAD Aggregate Statistics. The following table describes the similarities and differences in the universes of each product.

Inclusion/Exclusion Criteria	UAD Aggregate Statistics Data File Eligible Universe	UAD Appraisal-Level PUF Eligible Universe
Property type	Includes only single-family properties	Includes only single-family properties
Purchase type	Includes appraisals for all purchase or refinance mortgage loans.	Includes appraisals for arms-length purchase mortgage loans, all refinance mortgage loans, and all other (non-purchase and non-refinance) mortgage loans.
Address quality	Property address of appraised property must be of sufficient quality such that geocoding software can determine the county.	Property address of appraised property must be of sufficient quality such that the geocoding software or loan originator can determine the state.
Geographic scope	Properties in the 50 U.S. states, the District of Columbia, and Puerto Rico.	Properties in the 50 U.S. states, the District of Columbia, and Puerto Rico.
Enterprises acquisition status	Appraisal for any submission to the Enterprises, including mortgage loans not acquired by the Enterprises.	Appraisal for mortgage loans acquired by the Enterprises.

The second reason estimates may differ slightly is because the UAD Appraisal-Level PUF is a five-percent sample of eligible UAD appraisal records while the UAD Aggregate Statistics Data File includes aggregate statistics based on all (*i.e.*, 100 percent) of eligible UAD appraisal records. The differences due to sampling will generally be larger with more granular estimates (*i.e.*, estimates based on few UAD Appraisal-Level PUF records).

4. Disclosure Avoidance and Privacy Harm Reduction Techniques

A UAD appraisal record contains information about the property owner, borrower, and appraiser(s), including Personally Identifiable Information (PII). A UAD appraisal record also includes a significant amount of information about the features, neighborhood, quality, and condition of the appraised property and comparable properties. While much of the information about properties contained in a UAD appraisal record is also available in public records or real estate listings, or is observable from a publicly accessible location, some information is only available through a site visit during a traditional appraisal.

FHFA designed the UAD Appraisal-Level PUF to balance the need to include property, borrower, and neighborhood characteristics most useful to analyze with the need to meet two privacy goals. The first privacy goal was to reduce the possibility of identifying (sometimes referred to as re-identifying) a property address based on the available information about the appraised property. FHFA is aware that, due to the significant number of property characteristics common to UAD and public sources, a person could perform a re-identification attack that might reveal the property address of a UAD record.

The second privacy goal was to reduce the privacy harm to borrowers, property owners, and appraisers, *should a user be able to re-identify a property address successfully*. To achieve these two privacy goals, FHFA adopted several disclosure avoidance techniques described below.

4a. Sampling

FHFA used sampling as one of the disclosure avoidance methods to reduce potential re-identifying of borrowers, property addresses, and appraisers. FHFA's sample design objectives included the following:

- A sampling rate low enough to achieve re-identification reduction benefits meaningfully but high enough to produce statistically reliable estimates for key estimates.
- A sample design prioritizing production of estimates for the nation and for the nation's largest housing markets, as defined by metropolitan areas.
- A sample design prioritizing statistical precision in estimates for purchase appraisals, and specifically, estimates commonly used as indicators of potential appraisal bias.

Section 5 describes the sample design.

4b. Remove All Personally Identifiable Information

For the UAD Appraisal-Level PUF, FHFA did not include any data fields that are commonly considered PII. These included, but were not limited to, property owner(s) name and address, borrower(s) name and contact information, and appraiser name(s) and contact information.

4c. Remove All Precise Location Information

For the UAD Appraisal-Level PUF, FHFA did not include any data fields that provided precise or near-precise location information. In doing so, the Agency considered individual data fields as well as combination of data fields. These included, but were not limited to, mailing addresses, assessor parcel

number, lot number, legal description, zoning code, and Federal Emergency Management Agency flood zone.

4d. Remove Certain Information That Contributed to Re-Identification

FHFA evaluated each property characteristic in a UAD appraisal record to determine whether it was typically available in another public source such as public records or real estate listings. For property characteristics common to the UAD appraisal record and another public source, FHFA balanced the utility of including the property characteristic in the UAD Appraisal-Level PUF with its contribution to increasing the probability of re-identification. For the UAD Appraisal-Level PUF, FHFA removed numerous property characteristics that did not meet its balancing test.

4e. Remove Certain Information That Contributed to Privacy Harm

FHFA evaluated each property characteristic in a UAD appraisal record to determine whether it was typically available in another public source such as public records or real estate listings. For property characteristics NOT common to the UAD and another public source, FHFA balanced the utility of including the property characteristic in the UAD Appraisal-Level PUF with its contribution to privacy harm if the property address was to be re-identified. For the UAD Appraisal-Level PUF, FHFA removed numerous property characteristics that it identified, at this time, as having privacy harm that outweighs public utility or as having insufficient public utility to warrant disclosure in the initial dataset. FHFA plans to continue to evaluate appraisal data under this approach.

4f. Minimum Record Count Threshold

For the UAD Appraisal-Level PUF, FHFA adopted a minimum record count threshold of 11 records based on the Department of Housing and Urban Development's (HUD's) "Privacy Rule of Eleven." [5] In short, for census tracts with fewer than 11 appraisal records in a given year, the tract number was suppressed (primary suppression). In instances where only one tract in a county required suppression, FHFA suppressed the tract number for an additional tract (complementary suppression). Complementary suppression helps avoid the "disclosure by subtraction" issue that could compromise privacy.

FHFA applied the same protocol for counties with fewer than 11 appraisal records in a given year.

4g. Rounding

Some property characteristics in the UAD appraisal records were originally continuous data fields representing dollar values. For the UAD Appraisal-Level PUF, where appropriate, FHFA rounded the data field to the midpoint of each $10,000 buckets. Examples included appraised value (appraised_value) and contract value (contract_price).

4h. Binning

Some property characteristics in the UAD appraisal records were originally continuous data fields with precise values. For the UAD Appraisal-Level PUF, where appropriate, FHFA binned continuous data

---

[5] https://www.huduser.gov/portal/portal/sites/default/files/pdf/PDR-Privacy-Rule-of-Eleven.pdf

fields into categorical data fields with value range groups. For example, FHFA binned the lot size continuous data field into a categorical data field (lot_size) with five categories representing value ranges.

4i. Recoding

Some property characteristics in the UAD appraisal records were originally categorical data fields with numerous categories. For the UAD Appraisal-Level PUF, where appropriate, FHFA combined some categories to reduce the overall number of categories. For example, FHFA reduced the property condition data field (condition) from six to five categories.

4j. Bottom Coding and Top Coding

Some property characteristics in the UAD appraisal records were originally continuous data fields with especially low or high values. For the UAD Appraisal-Level PUF, where appropriate, FHFA capped these values at or near the 1st, 95th, or 99th percentile, often in combination with rounding, binning, or recoding. For example, the contract value data field (contract_price) was top coded at $1.7 million for high-cost areas and at $720,00 for other areas.[6]

5. Sample Design, Weighting, and Variance Estimation

5a. Sample Design

As mentioned in Section 4a, FHFA used sampling as one of the disclosure avoidance methods to reduce potential re-identification of borrowers, property addresses, and appraisers. To achieve the objectives described in Section 4a, FHFA elected to use a stratified simple random sample approach. FHFA drew a nationally representative simple random sample of five percent of all eligible appraisal records by year. FHFA determined that an annual sample of five percent of appraisal records balanced the need for re-identification reduction while maintaining adequate statistical power for numerous types of uses.

The following table shows the total number of appraisal records eligible for the sample, and the total sample size.

Year	Total Number of Records Eligible for the Sample	Total Sample Size
2013	2,736,580	136,829
2014	2,106,940	105,347
2015	2,692,520	134,626
2016	3,282,980	164,149
2017	2,648,740	132,437
2018	2,374,120	118,706
2019	2,948,520	147,426
2020	4,545,200	227,260
2021	4,135,320	206,766

---

[6] High-cost areas are defined as part of the FHFA Conforming Loan Limits regulations.
https://www.fhfa.gov/DataTools/Downloads/Pages/Conforming-Loan-Limit.aspx

5b. Weighting

In the UAD Appraisal-Level PUF, the data field "weight" represents the inverse probability of selecting the appraisal record for a given year. Because FHFA used the same sampling rate (5 percent) across all years, all UAD Appraisal-Level PUF appraisal records will have the same value (20) for the weight data field.

Users must utilize the weight data field to produce accurate appraisal counts and variance estimates. However, users do not need to utilize the weight data field for other types of estimates such as means, medians, or percentages. Users may want to include the weight in regression analyses but may also choose not to.

5c. Variance Estimation

Due to the simplicity of the stratified simple random sample design, users can estimate variances using standard formulas without the need to consider the sample design. In certain circumstances, users may wish to incorporate strata into their variance estimation. For those circumstances, the table in Section 4a provides population totals for each stratum.

6. Geography in the UAD Appraisal-Level PUF

The UAD Appraisal-Level PUF includes state, county, and 2010 census tract, subject to the disclosure avoidance techniques described in Section 4f. This combination of geographic indicators allows users to create estimates for other types of geographic areas, including Census Divisions and Core Based Statistical Areas (CBSAs).

As noted in Section 2, the UAD Appraisal-Level PUF will include a sample of appraisal records from 2013 through 2021. FHFA recognizes that this time period spans two Decennial Census periods (i.e., 2010 and 2020), and that some users may wish to merge year-specific information derived from other data sources, such as the Home Mortgage Disclosure Act (HMDA) data or the American Community Survey (ACS) to the Appraisal-Level PUF. FHFA considered two options to facilitate data merges. Option 1 was to include the vintage of Census Tract that reflected the year of the appraisal. For instance, a 2016 Appraisal-Level PUF record would have the 2010 vintage of Census Tract while a 2021 Appraisal-Level PUF record would have the 2020 vintage of Census Tract. Option 2 was to include the 2010 vintage of Census Tract for all Appraisal-Level PUF records.

FHFA elected option 2 to maximize the ability of users to merge HMDA data to the UAD Appraisal-Level PUF. The HMDA data file for 2020 and 2021 includes the 2010 vintage of Census Tracts. FHFA anticipates switching to the 2020 vintage of Census Tracts for future additions of records to the Appraisal-Level PUF (e.g., 2022 or 2023 records).

7. UAD Appraisal-Level PUF Format

The UAD Appraisal-Level PUF format is described below. Users should consult the UAD Appraisal-Level PUF Dictionary for a full list of data fields.

7a. File Format

The UAD Appraisal-Level PUF includes one record per appraisal of a subject property associated with a mortgage loan acquired by the Enterprises. In other words, the rows represent an appraisal while the columns include information about the property, comparable properties, borrower(s), neighborhood, and appraisal process. The UAD Appraisal-Level PUF is available as year-specific files as well as a combined file.

The UAD UAD Appraisal-Level PUFs are available in CSV, R, SAS, and Stata formats.

7b. Key Administrative and Geographic Data Fields

Users should note the following important data fields in the UAD Appraisal-Level PUF

Field	Definition
record_id	A unique ID for each appraisal record. FHFA created this data field that has no relationship to any unique id data fields in the original UAD appraisal records.
year	The year in which the appraisal was conducted. This data field is used to stratify the sample.
tract_2010	The 2010 census tract for the appraised property for appraisals conducted from 2013 through 2021. Due to potential disclosure issues, FHFA cannot provide 2020 census tract values.
weight	The sampling weight for the appraisal record, based on stratification by year and sample random sampling within a year. Currently, the weight is the same value for all records in a given year.

7c. Borrower Demographic Data Fields

A UAD appraisal record does not include demographic information about borrowers because the current version of the URAR does not have a field to collect it. However, to allow users to create key appraisal estimates for different borrower types, FHFA merged borrower demographic information from mortgage loan data to the UAD appraisal records in the UAD Appraisal-Level PUF. The demographic data field included is:

Field	Definition
race_ethnicity	Data field describing the race and ethnicity of the borrower(s)

7d. Information About Comparable Properties

A UAD appraisal record includes information about the subject property and information about one or more comparable properties (referred to as "comps") used to estimate value. The UAD Appraisal-Level PUF does not have a separate table for comps. Rather, it includes seven data fields that represent aggregate statistics about the comps. These data fieldss include:

Field	Description
same_tract_percent	Percent of comps located within same census tract as subject property
adjusted_price	Average adjusted sale price of all settled comps
calculated_proximity	Average proximity of comps to subject property (calculated distance) (in miles)
reported_proximity	Average proximity of comps to subject property (reported distance) (in miles)
median_adjustment	Median net (total) adjustments of all comps
average_adjustment	Average net (total) adjustments of all comps
number_comparables	Number of comps used in appraisal

8. Information Quality

The table below describes aspects of information quality based on FHFA's criteria.

Information Quality Criteria	Response
1. Describe the underlying source of any data used to create the product, including whether FHFA or a different agency collected the data.	Described in Sections 2 and 3.
2. Describe the statistical methods or models used to create the product.	The UAD Appraisal-Level PUF represents appraisal records. There was no modeling applied to the data fields. As described in Section 7, some data fields represent aggregate statistics for a subject property's comps records. The UAD Appraisal-Level PUF Data Dictionary describes how FHFA performed the statistical aggregation.
3. Describe the intended uses of the product, and, if applicable, uses not recommended.	FHFA intends for the UAD Appraisal-Level PUF to be used to understand patterns and geographic variation in appraisals. FHFA does not recommend it be used as a basis for appraising an individual property or to infer characteristics of an appraiser or the quality of an appraisal. FHFA does not recommend using it to draw conclusions about the prevalence of certain property characteristics among all properties in an area (such as accessory dwelling units) because the appraisal records reflect appraisals performed over a specific period and do not reflect data for all properties in an area.

4. Describe the time period presented in the event or phenomenon reflected in the product.	The UAD Appraisal-Level PUF includes a sample of appraisal records for appraisals conducted between January 1, 2013, and December 31, 2021 associated with mortgage loans acquired by the Enterprises.
5. Describe the granularity (*i.e.*, amount of disaggregation) of any key estimates. Granularity can be expressed in units of time, level of geographic detail available, or the amount of detail available on any number of characteristics.	The UAD Appraisal-Level PUF includes a sample of appraisal records. Users can utilize the data fields in the file to produce estimates by year, various geographic levels, and numerous characteristics.
6. Where applicable, describe any known major or significant errors in the underlying source data used to create the product.	FHFA is not aware of any major or significant errors in the UAD appraisal records. Where applicable, FHFA made fixes to correct minor data quality errors.
7. Where applicable, describe how users can estimate errors, such as errors from sampling.	Described in Section 5.
8. Where applicable, describe the consistency or comparability with estimates contained in other products published by FHFA.	The overall count of appraisal records does not directly compare to other Enterprise loan counts FHFA publishes. Some loans acquired by the Enterprises do not receive an appraisal. As noted in Section 3, the universe of appraisal records used to produce the UAD Appraisal-Level PUF includes only loans acquired by the Enterprises while the universe of appraisal records used to produce the UAD Aggregate Statistics Data File does not have this restriction.
9. Where applicable, describe the steps taken to ensure the product protects the privacy and confidentiality of underlying entity (*e.g.*, borrower, business) reflected in the source data, where applicable.	Described in Section 4.
10. Where applicable, describe the verification and validation steps taken to ensure errors are not introduced in the production process.	FHFA internally verified the appraisal records in the UAD Appraisal-Level PUF. The primary method was "code replication," whereby two different persons or teams produced the UAD Appraisal-Level PUF, then compared them to ensure consistency. Where inconsistency was discovered, differences were resolved.
11. Where applicable, describe the "chain of custody" of the product from its verification and validation to when it is posted on the website.	FHFA's Office of Congressional Affairs and Communications posts final products on FHFA's website. Once the final products post, FHFA staff download the product and perform a check to ensure the correct version is on the website.

Appendix 1. Protocol for Identifying Unique Final Appraisals in the UAD

This process below describes how FHFA created the universe of appraisal records eligible to include in the UAD Appraisal-Level PUF. These steps were carried out sequentially to identify unique final appraisals in the UAD.

1. Keep only active appraisals using the active document identification data fields.
2. Keep only appraisals that were successfully submitted through the UCDP, as denoted by the document status indicator.
3. For each "folder" within each "drawer," keep only the most recent appraisal. The most recent appraisal is the one with the largest "drawer" value.
4. For each appraisal "drawer," if there are two or more appraisals by the same appraiser and for the same amount, keep the most recent appraisal. The most recent appraisal is the one with the largest "folder" value.

In addition, the following steps were carried out to make the final inclusion and exclusion for the UAD Appraisal-Level PUF as mentioned in Section 3a:

1. Remove any appraisal conducted before January 1, 2013 and after December 31, 2021.
2. Only keep appraisals for single-family properties appraised using Fannie Mae Form 1004 or Freddie Mac Form 70.
3. Remove any appraisal where there is no Fannie Mae or Freddie Mac loan number. These are appraisals for loans that were not acquired by the Enterprises.
4. Remove any purchase or refinance appraisal where the UAD transaction type is inconsistent with the mortgage purpose code in FHFA's Enterprise mortgage acquisition data.
5. Remove any purchase appraisal that is a non-arm's length transaction.
6. Remove any appraisal in Guam or the U.S. Virgin Islands.