

## HDS 2010 Data Contents and Structure

The data deliverable contains files and documentation for data collected as part of the Housing Discrimination Study 2010. Data files are provided as SAS datasets. These datasets are provided for raw data (collected directly from testing), intermediate data (cleaned data used for analysis), and technical data (universe flags and weights).

Raw data includes the following datasets:

- Advancecall (advanced contact form)
- Apptcall (appointment contact form)
- Assignment (test assignment form)
- Availunit (available rental unit form)
- Followup (agent follow-up form)
- Neighborhood (sales neighborhood form)
- Rechomes (sales recommended homes form)
- Rental (rental site visit form)
- Rentnarr (rental site visit narrative)
- Rhgeo (sales recommended homes geocoded to Census tracts)
- Sales (sales sit visit form)
- Salenarr (sales site visit narrative)
- Taf (test assessment form)
- Tester (tester application form)

Intermediate data includes the following datasets:

- Appointment\_intermediate (data used for appointment analysis)
- Identifiability (survey data for testers' perceived race and ethnicity)
- Pre\_rental\_multi (data used for rental appointment multivariate analysis)
- Pre\_sales\_multi (data used for sales appointment multivariate analysis)
- Rental\_intermediate (data used for rental analysis)
- Rental\_multi (data used for rental multivariate analysis)
- Sales\_intermediate (data used for sales analysis)
- Sales\_multi (data used for sales multivariate analysis)
- Steering\_intermediate (data used for steering analysis)
- Skintone (data identifying the darkness of testers' skin)

Technical data includes the following datasets:

- Rentalflags (universe flags for rental tests)
- Salesflags (universe flags for sales tests)
- Rentalweights (weights for rental tests)
- Salesweights (weights for sales tests)

We have included a codebook for each set of data that indicates variable names, variable lengths, labels, and non-negative allowable values, if applicable.

Additional detail about the structure of the data tables, the level of analysis for each table, and how they fit together is provided below.

## Raw Data Tables

### Table: Advancecall

- Description: Documents each call made by the advance caller/emailer (ACE) to inquire about the availability of an advertised housing unit.
- Tenure: Rental and Sales
- Record-level: Each record represents an advanced contact made by the ACE to inquire about housing (multiple calls per test are possible).

### Table: Apptcall

- Description: Documents each contact made by the HDS tester attempting to make an appointment with the agent.
- Tenure: Rental and Sales
- Record-level: Each record represents an appointment contact made by the tester (multiple contacts per test are possible).

### Table: Assignment

- Description: Takes housing advertisements and assigns them to HDS testers, along with characteristics for testers to use during the test.
- Tenure: Rental and Sales
- Record-level: Each record represents an assignment made to a tester (each test has up to two records, one per tester).

### Table: Availunit

- Description: Documents each available rental unit that agents told testers about.
- Tenure: Rental
- Record-level: Each record represents an available rental unit (multiple available units per tests are possible).

### Table: Followup

- Description: Documents each follow-up contact made by a housing agent.
- Tenure: Rental and Sales
- Record-level: Each record represents a follow-up contact by a housing agent (multiple follow-up contacts per test are possible).

### Table: Neighborhood

- Description: Documents comments made by housing agents about a particular neighborhood.
- Tenure: Sales
- Record-level: Each record represents a single neighborhood that the agent made comments about (multiple neighborhoods per test are possible).

### Table: Rechomes

- Description: Documents each home that agents recommended to testers.
- Tenure: Sales
- Record-level: Each record represents a recommended home (multiple recommended homes per test are possible).

Table: Rental

- Description: Documents each rental site visit.
- Tenure: Rental
- Record-level: Each record represents a rental site visit (each test has up to two records, one per tester).

Table: Rentnarr

- Description: Documents narratives written by testers.
- Tenure: Rental
- Record-level: Each record represents a rental narrative (each test has up to two records, one per tester – not all tests required narratives).

Table: Rhgeo

- Description: Geocoded version of the Rechomes table.
- Tenure: Sales
- Record-level: Each record represents a geocoded recommended home (multiple recommended homes per test are possible). Addresses are geocoded to Census tracts using ESRI software. Demographic data from the 2010 Census and 2006-2010 American Community Survey are merged onto each record.

Table: Salenarr

- Description: Documents narratives written by testers.
- Tenure: Sales
- Record-level: Each record represents a sales narrative (each test has up to two records, one per tester – not all tests required narratives).

Table: Sales

- Description: Documents each sales site visit.
- Tenure: Sales
- Record-level: Each record represents a sales site visit (each test has up to four records, one or two per tester).

Table: Taf

- Description: Contains unique information about each test, including advertisement information and whether the test passed quality review.
- Tenure: Rental and Sales
- Record-level: Each record represents one test.

Table: Tester

- Description: Contains information that each tester is required to submit prior to participating as an HDS tester.
- Tenure: N/A
- Record-level: Each record represents one HDS tester.

## Intermediate Data Tables

### Table: Appointment\_intermediate

- Description: Data table used for appointment (pre-test) analysis.
- Tenure: Rental and Sales
- Record-level: Each record represents one test in which both testers attempted to make an appointment with a housing agent (or dropped-in to a housing provider).
- Additional notes: Merge rentalflags and rentalweights, or salesflags and salesweights, by the CONTROL variable, in order to complete the analysis.

### Table: Identifiability

- Description: Survey data used to determine the perceived race of each HDS tester.
- Tenure: N/A
- Record-level: Each record represents one survey response identifying the perceived race of a tester (up to 3 responses are available for each tester for name and voice analysis; up to 3 responses are available for each tester for name, voice and photo analysis).

### Table: Pre\_rental\_multi

- Description: Data table used for rental appointment (pre-test) multivariate analysis.
- Tenure: Rental
- Record-level: Each record represents one test in which both testers attempted to make an appointment with a housing agent (or dropped-in to a housing provider).

### Table: Pre\_sales\_multi

- Description: Data table used for sales appointment (pre-test) multivariate analysis.
- Tenure: Sales
- Record-level: Each record represents one test in which both testers attempted to make an appointment with a housing agent.

### Table: Rental\_intermediate

- Description: Data table used for rental analysis.
- Tenure: Rental
- Record-level: Each record represents one test completed test that passed quality review.
- Additional notes: Merge rentalflags and rentalweights by the CONTROL variable, in order to complete the analysis.

### Table: Rental\_multi

- Description: Data table used for rental multivariate analysis.
- Tenure: Rental
- Record-level: Each record represents one completed test that passed quality review.

### Table: Sales\_intermediate

- Description: Data table used for sales analysis.
- Tenure: Sales
- Record-level: Each record represents one test completed test that passed quality review.
- Additional notes: Merge salesflags and salesweights by the CONTROL variable, in order to complete the analysis.

Table: Sales\_multi

- Description: Data table used for sales multivariate analysis.
- Tenure: Sales
- Record-level: Each record represents one completed test that passed quality review.

Table: Skintone

- Description: Rating of the darkness of testers' skin.
- Tenure: N/A
- Record-level: Each record represents one tester.

Table: Steering\_intermediate

- Description: Data table used for steering analysis.
- Tenure: Sales
- Record-level: Each record represents one completed test that passed quality review and was successfully geocoded to a Census tract.
- Additional notes: Census variables are averaged over recommended or inspected tracts. Merge salesflags and salesweights by the CONTROL variable, in order to complete the analysis.

## Technical Data Tables

### Table: Rentalflags

- Description: Flag data used to identify subsamples of tests used for rental analysis.
- Tenure: Rental
- Record-level: Each record represents one test.

### Table: Rentalweights

- Description: Weight data used for rental analysis
- Tenure: Rental
- Record-level: Each record represents one test.

### Table: Salesflags

- Description: Flag data used to identify subsamples of tests used for sales analysis.
- Tenure: Sales
- Record-level: Each record represents one test.

### Table: Salesweights

- Description: Weight data used for sales analysis
- Tenure: Sales
- Record-level: Each record represents one test.