

## Survey of Construction Microdata Files



- Include one-family houses started, completed, and/or sold in the U.S. each year
- Also include houses authorized but not started, under construction, or for sale at the end of the year
- Files contain more than 40 physical and financial characteristics
- Available geography: Census Division and inside/outside metropolitan area
- Annual files are available for 1999 through 2008
- More information about the Survey of Construction is available online at [www.census.gov/mcd](http://www.census.gov/mcd).



Files are available in SAS or Microsoft Excel format.  
Files can be sent via e-mail or on a cd-rom.  
The cost is \$50.00 for each annual file.

## General Information

Annual microdata files contain all Survey of Construction (SOC) sample houses started, sold, and/or completed during the year. Houses authorized by building permits but not started at the end of the year, under construction at the end of the year, or for sale at the end of the year are also included. Houses for which construction was abandoned after permit issuance or after start are not included.

SOC is a national sample survey of new houses selected from building permits and a canvassing of areas not requiring permits. Builders or owners of the houses selected are interviewed for information including start, sale, and completion dates, and more than 40 physical and financial characteristics of the houses. The overall national sampling rate is about 1 in 50 new houses, although this varies considerably by individual survey location based on activity.

Available geography on the files includes the nine Census Divisions (DIV) and whether a house is inside or outside a metropolitan area (METRO).

All new houses are classified into one of four categories (CAT): (1) **houses built for sale** (whether sold or not) where the house and land are sold as a package; (2) **contractor-built houses** where the owner of the land hires a general contractor to build the house; (3) **owner-built houses** where the owner of the land builds the house him/herself or acts as the general contractor; and (4) **houses built for rent**. Only houses “built for sale” can have a sale date.

Tabulating records on the files by characteristics for a given year (for example, all houses completed in 2007) can be done by utilizing one of the dates available (start, sale, completion) which appear in year/month (YYYYMM) format.

The files also allow tabulations of the number of units authorized but not started, under construction (started but not completed), or for sale (built for sale but not sold) at the end of the year.

Each record contains a sampling weight (WEIGHT) which indicates how many houses a sample house represents. Tabulated totals from the files will not exactly match published estimates, primarily because the files were produced at different times. The files do not provide accurate tabulations of total units authorized by permits for a number of reasons.

The files contain no information which would identify the specific addresses or builders of the sample houses. In addition, certain variables such as sales price (SLPR) and square footage (SQFS) have been “top and bottom-coded” to further remove the possibility of disclosure. Cases which have been modified as such are identified with “flags.” The relationship of modified variables with other variables will be altered.

Please refer to [www.census.gov/mcd](http://www.census.gov/mcd) for a full description of SOC including a discussion of the reliability of the data and measures of sampling error.

Information on each variable contained in the files is shown in the following table.

Variable Number	Variable Name	Description	Remarks
1	Division (DIV)	1 = New England 2 = Middle Atlantic 3 = East North Central 4 = West North Central 5 = South Atlantic 6 = East South Central 7 = West South Central 8 = Mountain 9 = Pacific	Order of houses within each division was randomly scrambled.
2	Metropolitan Area (METRO)	1 = In a Metropolitan Statistical Area (MSA) 2 = Not in an MSA	None.
3	Weight (WEIGHT)	Number of housing units represented by this sample case	None.
4	Category Code (CAT)	1 = Built for sale/sold 2 = Contractor-built 3 = Owner-built 4 = Built for rent	None.
5	Permit Authorization Date (AUTH)	Month the building permit for the house was issued YYYYMM = Year/month, such as 200601	Usually the month the permit was issued, but can vary based on the availability of permits for sampling for SOC. Set to the start date if the house is in an area where permits are not required or if the start date was before the issuance of the final building permit.

Variable Number	Variable Name	Description	Remarks
6	Start Date (STRT)	YYYYMM = Year/month, such as 200601 0 = Not started	None.
7	Completion Date (COMP)	YYYYMM = Year/month, such as 200601 0 = Not completed	None.
8	Sale Date (SALE)	YYYYMM = Year/month, such as 200601 0 = Not sold	Applies only to houses built for sale/ sold.
9	Design of House (DET)	1 = Detached 2 = Attached 0 = Not reported	None.
10	Construction Method (MFGS)	1 = Modular 2 = Panelized / Precut 3 = Site-built 0 = Not reported	None.
11	Square Foot Area of House (SQFS)	In square feet. 0 = Not reported	Top and bottom 1% of all reported values have been changed to values at those limits.
12	Square Foot Area of House Flag (SQFS_F)	1 = Square foot value has been changed 0 = No changing done	Relationships of changed values to other values will be different.
13	Foundation (BASE)	01 = Full or partial basement 02 = Crawl space 03 = Slab 04 = Other 00 = Not reported	“Other” includes raised supports, earthen, and other foundation types.

Variable Number	Variable Name	Description	Remarks
14	Finished Basement? (FNBS)	1 = Yes 2 = No 0 = Not applicable or not reported	Applies only to houses reporting a full or partial basement.
15	Square Foot Area of Finished Basement (FNSQ)	In square feet 0 = Not applicable or not reported	Applies only to houses with a full or partial finished basement and reporting square foot areas of house (unchanged) and finished basement. Values not allowed to exceed 50% or be less than 5% of area of house. Values outside those limits changed to those limits.
16	Square Foot Area of Finished Basement Flag (FNSQ_F)	1 = Square foot value has been changed 0 = No changing done	Relationships of changed values to other values will be different.
17	Bedrooms (BEDR)	2 = 2 bedrooms or less 3 = 3 bedrooms 4 = 4 bedrooms 5 = 5 bedrooms or more 0 = Not reported	None.
18	Full Bathrooms (FULB)	1 = 1 bathroom or less 2 = 2 bathrooms 3 = 3 bathrooms 4 = 4 bathrooms or more 9 = Not reported	None.

Variable Number	Variable Name	Description	Remarks
19	Half Bathrooms (HAFB)	0 = 0 half bathrooms 1 = 1 half bathrooms 2 = 2 half bathrooms or more 9 = Not reported	None.
20	Stories (STOR)	1 = 1 story 2 = 2 or more stories (including 1½ stories) 3 = Split-level 0 = Not reported	None.
21	Primary Exterior Wall Material (WAL1)	01 = Wood or wood products (including masonite or T111) 02 = Brick or brick veneer 03 = Aluminum siding (not covered with vinyl) 04 = Concrete stucco (such as Shotcrete) 05 = Vinyl siding (including vinyl-covered aluminum) 06 = Concrete block (not including stucco) 07 = Stone, rock, or other stone materials 08 = Fiber cement siding (such as Hardiplank, Hardiboard) 09 = Other 00 = Not reported	“Other” includes poured concrete, synthetic stucco, steel siding, and other materials.  <b>Beginning with the file for 2008, Aluminum siding is also included with “09-Other” rather than being shown separately.</b>
22	Secondary Exterior Wall Material? (WALS)	1 = Yes 2 = No 0 = Not reported	None.

Variable Number	Variable Name	Description	Remarks
23	Secondary Exterior Wall Material (WAL2)	01 = Wood or wood products (including masonite or T111) 02 = Brick or brick veneer 03 = Aluminum siding (not covered with vinyl) 04 = Concrete stucco (such as Shotcrete) 05 = Vinyl siding (including vinyl-covered aluminum) 06 = Concrete block (not including stucco) 07 = Stone, rock, or other stone materials 08 = Fiber cement siding (such as Hardiplank, Hardiboard) 09 = Other 10 = No secondary wall material 00 = Not reported	Applies only to houses reporting “yes” to having a secondary exterior wall material.  “Other” includes poured concrete, synthetic stucco, steel siding, and other materials.  <b>Beginning with the file for 2008, Aluminum siding is also included with “09-Other” rather than being shown separately.</b>
24	Parking Facility (GAR)	1 = 1 car garage 2 = 2 car garage 3 = 3 or more car garage 4 = Other 0 = Not reported	“Other” includes carport, other offstreet parking (including a driveway with no garage or carport), and other parking facilities.
25	Fireplace (FPLS)	0 = None 1 = 1 fireplace 2 = 2 or more fireplaces 9 = Not reported	None.

Variable Number	Variable Name	Description	Remarks
26	Deck (DECK)	1 = Yes 2 = No 0 = Not reported	None.
27	Patio? (PATI)	1 = Yes 2 = No 0 = Not reported	None.
28	Porch? (PRCH)	1 = Yes 2 = No 0 = Not reported	None.
29	Primary Space Heating Fuel (FUEL)	01 = Electricity 02 = Natural gas (from underground pipes) 03 = Bottled gas (including propane or tank gas) 04 = Oil (including heating oil or kerosene) 05 = Other or no heat 00 = Not reported	“Other” includes wood, pellets, solar, coal, and other fuels.  <b>Beginning with the file for 2008, Bottled gas is included with “02-Natural gas” rather than being shown separately.</b>
30	Primary Space Heating System (HEAT)	01 = Air-source heat pump/ Ground-source heat pump 02 = Forced-air furnace without heat pump 03 = Hot water or steam system 04 = Other or no heat 00 = Not reported	“Other” includes electric baseboard or wall panels, fireplace with insert, stove that burns coal or wood, non-portable room heater that burns liquid fuel and is connected to a flue, vent, or chimney, passive solar system, other heat sources, or no heat.



Variable Number	Variable Name	Description	Remarks
31	Central Air-conditioning? (ACS)	1 = Yes 2 = No 0 = Not reported	None.
32	Sales Price (SLPR)	Whole dollars to nearest \$100 0 = Not applicable or not reported	Applies only to houses sold. Within each division, top 3% and bottom 1% of all reported values were changed to those limits. Unchanged values less than \$15000 were changed to \$15000.
33	Sales Price Flag (SLPR_F)	1 = Sales price has been changed 0 = No changing done	Relationships of changed values to other values will be different.
34	Contract Price (CONPR)	Whole dollars to nearest \$100. 0 = Not applicable or not reported	Applies only to contractor-built houses. Within each division, top 3% and bottom 1% of all reported values were changed to those limits. Unchanged values less than \$15000 were changed to \$15000.
35	Contract Price Flag (CONPR_F)	1 = Contract price has been changed 0 = No changing done	Relationships of changed values to other values will be different.
36	Condominium? (CON)	1 = Yes 2 = No 0 = Not applicable or not reported	Applies only to houses built for sale/sold.
37	Closing Costs in Sales Price? (CLOS)	1 = Yes 2 = No 0 = Not applicable or not reported	Applies only to houses sold.  <b>Beginning with the file for 2008, Closing Costs is no longer available.</b>

Variable Number	Variable Name	Description	Remarks
38	Lot Value (LOTV)	Whole dollars to nearest \$100. 0 = Not applicable or not reported	Applies only to houses sold with reported sales prices (unchanged) and lot values. Top and bottom 3% of all ratios of lot value to sales price by division were determined. Lot values with ratios outside those limits were changed such that recalculated ratios equal the limits.
39	Lot Value Flag (LOTV_F)	1 = Lot value has been changed. 0 = No changing done	Relationships of changed values to other values will be different.
40	Square Foot Area of the Lot (AREA)	In square feet 0 = Not reported	For all houses, values less than 2,000 were changed to 2,000. For houses built for sale/sold and houses built for rent, all values greater than 87,120 were changed to 87,120. For houses that are contractor-built and owner-built, all values greater than 435,600 were changed to 435,600.
41	Square Foot Area of the Lot Flag (AREA_F)	1 = Lot size has been changed 0 = No changing done	Relationships of changed values to other values will be different.
42	Type of Financing (FINC)	01 = Conventional 02 = FHA 03 = VA 04 = Cash 05 = Other financing 00 = Not reported	“Other” includes Rural Housing Service, cash, Habitat for Humanity, loan from an individual, State or local government mortgage-backed bonds and other types of financing.