

# Survey of Construction Microdata Files

---



- Include single-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 2011
- More information about the Survey of Construction is available online at [www.census.gov/starts](http://www.census.gov/starts).
- Collection of these data is partially funded by the Department of Housing and Urban Development (HUD).

Files are available in SAS or Microsoft Excel format.

Files can be sent via e-mail or on a cd-rom.

Download at [www.census.gov/construction/chars/microdata.html](http://www.census.gov/construction/chars/microdata.html) or call (301) 763-5160.

# General Information

---

Annual microdata files contain all Survey of Construction (SOC) sampled single-family 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 characteristics are collected throughout the construction process, thus the highest response rate is associated with completed homes. The Department of Housing and Urban Development (HUD) funds the collection of these characteristics and allows the Census Bureau to provide the data to the public free of charge. 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 include 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 2011) can be done by using one of the available dates (start, sale, or 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) that indicates how many houses a sampled 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 that would identify the specific addresses or builders of the sampled 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 that 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/starts](http://www.census.gov/starts) 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.

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
1	Central Air-conditioning? (ACS)	1 = Yes 2 = No 0 = Not reported	None.
2	Age-Restricted? (AGER)	Indicates whether the house is in a development that is age-restricted. 1 = Yes 2 = No 0 = Not reported	Age-restricted developments are intended for occupancy by persons of a specific age group, usually 55 and older or 62 and older.  <i>Not available in files for years prior to 2009.</i>
3	Community Association? (ASSOC)	Indicates whether the house is part of a community or homeowner's association. 1 = Yes 2 = No 0 = Not reported	Community or homeowner's associations are formal legal entities created to maintain common areas of a development and to enforce private deed restrictions; these organizations are usually created when the development is built, and membership is mandatory.  <i>Not available in files for years prior to 2009.</i>
4	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.
5	Category Code (CAT)	1 = Built for sale/sold 2 = Contractor-built 3 = Owner-built 4 = Built for rent	None.
6	Closing Costs in Sales Price? (CLOS)	1 = Yes 2 = No 0 = Not applicable or not reported	Applies only to houses sold.  <i>Not available in files for years after 2007.</i>
7	Condominium? (CON)	1 = Yes 2 = No 0 = Not applicable or not reported	Applies only to houses built for sale/sold.

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
8	Deck? (DECK)	1 = Yes 2 = No 0 = Not reported	None.
9	Design of House (DET)	1 = Detached 2 = Attached 0 = Not reported	None.
10	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.
11	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.
12	Finished Basement? (FNBS)	1 = Yes 2 = No 0 = Not applicable or not reported	Applies only to houses reporting a full or partial basement.
13	Two-Story Foyer (FOYER)	Indicates whether the house has a two-story foyer. 1 = Yes 2 = No 0 = Not reported	A two-story or "open" foyer means that the entranceway inside the front door of the house has a ceiling that is at the level of the second-floor ceiling.  <i>Not available in files for years prior to 2009.</i>

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
14	Framing Material (FRAME)	Primary material used for the structural framing of the house. 1 = Wood 2 = Steel 3 = Concrete/masonry other than insulated concrete forms 4 = Insulated concrete forms (or SIP) 0 = Not reported	Structural framing refers to the studs used to frame exterior walls.  <i>Not available in files for years prior to 2009.</i>
15	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.
16	Primary Space Heating System (HEAT)	01 = Air-source heat pump or 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, and no heat.
17	Second Space Heating System (HEAT2)	01 = Air-source heat pump or ground-source heat pump 02 = Forced-air furnace without heat pump 03 = Hot water or steam system 04 = Other 00 = Not reported or no second heating system	"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, and other heat sources.  <i>Not available in files for years prior to 2009.</i>

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
18	Laundry Location (LNDR)	Location of the washer and dryer connections in the house. 1 = In the basement 2 = On the first floor 3 = On the second or higher floor 4 = In the garage or carport 5 = No connection planned 6 = Multiple locations 0 = Not reported	<i>Not available in files for years prior to 2009.</i>
19	Metropolitan Area (METRO)	1 = In a Metropolitan Statistical Area (MSA) 2 = Not in an MSA	None.
20	Construction Method (MFGS)	1 = Modular 2 = Panelized / Precut 3 = Site-built 0 = Not reported	None.
21	Patio? (PATI)	1 = Yes 2 = No 0 = Not reported	None.
22	Porch? (PRCH)	1 = Yes 2 = No 0 = Not reported	None.
23	Sewage Disposal System (SEWER)	1 = Public sewers (including community or shared sewage/septic system) 2 = Individual septic system 3 = Other	<i>Not available in files for years prior to 2009.</i>
24	Stories (STOR)	1 = 1 story 2 = 2 or more stories (including 1½ stories) 3 = Split-level 0 = Not reported	None.

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
25	Primary Exterior Wall Material (WAL1)	01 = Wood or wood products 02 = Brick or brick veneer 03 = Aluminum siding (not covered with vinyl) 04 = Stucco 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 or Hardiboard) 09 = Other 00 = Not reported	<p>“Other” includes poured concrete, steel siding, and other materials.</p> <p><i>Beginning with the file for 2008, Aluminum siding is also included with “09-Other” rather than being shown separately.</i></p>
26	Secondary Exterior Wall Material (WAL2)	01 = Wood or wood products 02 = Brick or brick veneer 03 = Aluminum siding (not covered with vinyl) 04 = Stucco 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 or Hardiboard) 09 = Other 00 = Not reported or no secondary wall material	<p>Applies only to houses reporting “yes” to having a secondary exterior wall material.</p> <p>“Other” includes poured concrete, steel siding, and other materials.</p> <p><i>Beginning with the file for 2008, Aluminum siding is also included with “09-Other” rather than being shown separately.</i></p> <p><i>In files for years prior to 2009, WAL2 is “10” for houses with no secondary wall material (those with WAL5=“2”). Beginning with the file for 2009, houses with WAL5=“2” have WAL2=“00”.</i></p>
27	Secondary Exterior Wall Material? (WALS)	1 = Yes 2 = No 0 = Not reported	<p><i>In files for years prior to 2009, WAL2 is “10” for houses with no secondary wall material (those with WAL5=“2”). Beginning with the file for 2009, houses with WAL5=“2” have WAL2=“00”.</i></p>

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
28	Water Supply Source (WATER)	1 = Public water (including community or shared water supply/well) 2 = Individual well 3 = Other 0 = Not reported	<i>Not available in files for years prior to 2009.</i>
29	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.
30	Bedrooms (BEDR)	2 = 2 bedrooms or less 3 = 3 bedrooms 4 = 4 bedrooms 5 = 5 bedrooms or more 0 = Not reported	None.
31	Completion Date (COMP)	YYYYMM = Year/month, such as 200901 0 = Not completed	None.
32	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.
33	Final Square Foot Area of Finished Basement at Completion (FFNSQ)	Final square foot area of finished basement reported at the time the house was completed; in square feet.	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.  <i>Not available in files for years prior to 2009.</i>



<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
34	Fireplace (FPLS)	0 = None 1 = 1 fireplace 2 = 2 or more fireplaces 9 = Not reported	None.
35	Full Bathrooms (FULB)	1 = 1 bathroom or less 2 = 2 bathrooms 3 = 3 bathrooms 4 = 4 bathrooms or more 9 = Not reported	None.
36	Half Bathrooms (HAFB)	0 = 0 half bathrooms 1 = 1 half bathrooms 2 = 2 half bathrooms or more 9 = Not reported	None.
37	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.
38	Permit value (PVALU)	Estimated value of the house as shown on the building permit; whole dollars to nearest \$100. 0 = Not applicable or not reported	Within each division, top 3% and bottom 1% of all reported values were changed to those limits. Unchanged values less than \$15,000 were changed to \$15,000.  <i>Not available in files for years prior to 2009.</i>
39	Sale Date (SALE)	YYYYMM = Year/month, such as 200901 0 = Not sold	Applies only to houses built for sale/ sold.
40	Final Square Foot Area of House at Completion (FSQFS)	Final square foot area reported at the time the house was completed; in square feet.	Top and bottom 1% of all reported values have been changed to values at those limits.  <i>Not available in files for years prior to 2009.</i>

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
41	Start Date (STRT)	YYYYMM = Year/month, such as 200901 0 = Not started	None.
42	Contract Price (CONPR)	Price on the original contract awarded to the general contractor; 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 \$15,000 were changed to \$15,000.
43	Sales Price (SLPR)	Sales price at the time the first sales contract was signed or deposit made; includes the price of the improved lot; 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 \$15,000 were changed to \$15,000.
44	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.
45	Weight (WEIGHT)	Number of housing units represented by this sample case.	None.
46	Primary Space Heating Fuel (FUEL)	01 = Electricity 02 = Natural gas 03 = Bottled or liquefied petroleum gas (including propane) 04 = Oil (including heating oil or kerosene) 05 = Other or no heat 00 = Not reported	“Other” includes wood, pellets, solar, coal, and other fuels.  <i>In the files for 2008 and 2009, Bottled gas is included with “02-Natural gas” rather than being shown separately.</i>
47	Second Space Heating Fuel (FUEL2)	01 = Electricity 02 = Natural gas 03 = Bottled or liquefied petroleum gas (including propane) 04 = Oil (including heating oil or kerosene) 05 = Other 00 = Not reported or no second heating system	“Other” includes wood, pellets, solar, coal, and other fuels.  <i>Not available in files for years prior to 2009.</i>  <i>In the file for 2009, Bottled gas is included with “02-Natural gas” rather than being shown separately.</i>

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
48	Final Contract Price at Completion (FCONPR)	Final contract price reported at the time the house was completed; 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 \$15,000 were changed to \$15,000.  <i>Not available in files for years prior to 2009.</i>
49	Final Sales Price at Completion (FSLPR)	Final sales price reported at the time the house was completed; 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 \$15,000 were changed to \$15,000.  <i>Not available in files for years prior to 2009.</i>
50	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.
51	Square Foot Area of Finished Basement Flag (FNSQ_F)	1 = Final square foot area of finished basement at time of completion has been changed 0 = No changing done	Relationships of changed values to other values will be different.
52	Final Square Foot Area of Finished Basement at Completion Flag (FFNSQ_F)	1 = Final square foot area of finished basement at time of completion has been changed 0 = No changing done	Relationships of changed values to other values will be different.  <i>Not available in files for years prior to 2009.</i>
53	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.
54	Final Sales Price Flag (FSLPR_F)	1 = Final sales price reported at the time the house was completed has been changed 0 = No changing done	Relationships of changed values to other values will be different.  <i>Not available in files for years prior to 2009.</i>
55	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.

<i>Item</i>	<i>Variable Name</i>	<i>Description</i>	<i>Remarks</i>
56	Final Contract Price Flag (FCONPR_F)	1 = Final contract price reported at the time the house was completed has been changed 0 = No changing done	Relationships of changed values to other values will be different.  <i>Not available in files for years prior to 2009.</i>
57	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.
58	Square Foot Area of House Flag (SQFS_F)	1 = Square foot area has been changed 0 = No changing done	Relationships of changed values to other values will be different.
59	Final Square Foot Area of House at Completion Flag (FSQFS_F)	1 = Final square foot area of house at time of completion has been changed 0 = No changing done	Relationships of changed values to other values will be different.  <i>Not available in files for years prior to 2009.</i>
60	Permit Value Flag (PVALU_F)	1 = Permit value has been changed 0 = No changing done	Relationships of changed values to other values will be different.  <i>Not available in files for years prior to 2009.</i>
61	Permit Authorization Date (AUTH)	YYYYMM = Year/month, such as 200901	Usually the month the building permit for the house 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.