

Documentation on HOUSE.XLS

HOUSE.XLS is an Excel spreadsheet dataset consisting of a sample of 980 single family homes sold in Santa Clara County, California in June 1987. Below is a brief description the variables included in HOUSE.DAT.

ID	Unique identifier for each observation
CAMPBELL	Dummy variable for the city of Campbell
LOSGATOS	Dummy variable for the city of Los Gatos
PALOALTO	Dummy variable for the city of Palo Alto
MILPITAS	Dummy variable for the city of Milpitas
SUNNYVALE	Dummy variable for the city of Sunnyvale
LOSALTOS	Dummy variable for the city of Los Altos
SARATOGA	Dummy variable for the city of Saratoga
GILROY	Dummy variable for the city of Gilroy
MOUNVIEW	Dummy variable for the city of Mountain View
SNTACLRA	Dummy variable for the city of Santa Clara
CITYPOP	Population of city
OWNOCC80	Percent of housing in relevant census tract that is owner occupied, 1980
MFPCT87	Percent of housing in relevant census tract that is multi-family, 1987
PRICE	Sale price of property (in dollars)
LIVAREA	Total floorspace in dwelling (in square feet)
LOTAREA	Area of lot (in square feet)

BATHRMS	Number of full bathrooms
FIREPLCE	Number of fireplaces
POOLBIN	Dummy variable indicating existence of one or more swimming pools
SPABIN	Dummy variable indicating existence of one or more spas
PARKSPCE	Number of parking spaces available in garage or carport
AGE	Age of structure (in years)
STORIES	Number of stories
BEDRMS	Number of bedrooms
ROOMS	Total number of rooms
PSALEYR	Year in which last previous sale occurred Note: missing data/no previous sale coded as -9.
PPRICE	Previous sale price (in dollars). Note: if no previous sale, coded as 0.
MINLAREA	Minimum allowable lot area in 1987 (in square feet)
MINLWIDE	Mimimum allowable lot width in 1987 (in feet) Note: missing data coded as -9.
MINFYARD	Minimum allowable front yard set back in 1987 (in feet)
MINSYARD	Minimum allowable side yard set back in 1987 (in feet)
MINRYARD	Mimimum allowable rear yard set back in 1987 (in feet)
MAXHIGH	Maximum allowable height of structure in 1987 (in feet)
NOCONFRM	Dummy variable indicating actual lot size is more than 100 feet the minimum allowed in 1987
MILLRATE percent).	Legislated tax rate for relevant tax rate area (in
TOTACRE	Total acreage in relevant tax rate area (includes acreage with unknown land use).

TOTAWLU	Total acreage in relevant tax rate area for which land use is known.
PCTSFR	Percent of TOTAWLU that is single-family residential (land use code 01).
PCTMFR	Percent of TOTAWLU that is multi-family residential (land use codes 02,03,04,06).
PCTOTHR use	Percent of TOTAWLU that is other residential uses (land codes 05,07,08,09).
PCTIND	Percent of TOTAWLU that is industrial use (land use codes 10-40).
PCTCOMM 50-59).	Percent of TOTAWLU that is commercial (land use codes 50-59).
PCTREGSC	Percent of TOTAWLU that is regional shopping centers (land use code 50).
PCTSMKT 52).	Percent of TOTAWLU that is supermarkets (land use code 52).
PCTBUS	Percent of TOTAWLU that is business districts (land use codes 53-57).
PCTLPRK	Percent of TOTAWLU that is local parks (land use codes 80-84).
PCTRPRK 85).	Percent of TOTAWLU that is regional parks (land use code 85).
PCTOS 86-88).	Percent of TOTAWLU that is open space (land use codes 86-88).
PCTAG codes	Percent of TOTAWLU that is agricultural use (land use codes 90-99).
PCTVACNT code	Percent of TOTAWLU that is vacant urban land (land use code 69).

PCTMAJBD	Percent of TOTAWLU that is major business districts (land use code 55).
PCTLOCBD use	Percent of TOTAWLU that is local business districts (land code 56).
PCTDRVIN	Percent of TOTAWLU that is "auto-oriented and drive-in businesses and strip commercial" (land use code 58).
PCTOFC (land	Percent of TOTAWLU that is offices, banks, and clinics use code 59).
PCT2FAM use	Percent of TOTAWLU that is two-family residential (land code 02).
PCT34FAM residential	Percent of TOTAWLU that is three and four family residential (land use code 03).
PCT5FAM (land	Percent of TOTAWLU that is 5 or more family residential use code 04).
PCTCONDO (land	Percent of TOTAWLU that is condominiums and townhouses use code 06).
PCTR use	Percent of TOTAWLU that is railroad transportation (land code 43).
PCTHGHWY use	Percent of TOTAWLU that is limited access streets (land code 41).
PHWKID**	Percentage of households in relevant census tract with own children under 18 for year indicated.
PWHITE**	Percentage of persons in the relevant census tract who are white for year indicated.
PANGL080	Percentage of persons in the relevant census tract who are white, but not of Spanish origin in 1980

P4COLL**	Percentage of persons 25 and over in the relevant census tract who completed 4 or more years of college in the year indicated
PHSGRD**	Percentage of persons 25 and over in the relevant census tract who are high school graduates.
HERF** census	Herfindahl index of income homogeneity in the relevant tract for the year indicated.
MEDY**	Median income in the relevant census tract for the year indicated.
MEDAGE80	Median age of persons in the relevant census tract, 1980.
TIMECNTR	Peak morning commute time to geographic center of "high tech." manufacturing area (area bounded by Hwys. 101, 880, 280, and 85), 1980 (in minutes)
TIMESJ	Peak morning commute time to downtown San Jose, 1980 (in minutes)
TIMEAPPL Computer,	Peak morning commute time to headquarters of Apple 1980 (in minutes)
TIMELOCK Corp.,	Peak morning commute time to headquarters of Lockheed 1980 (in minutes).
EXPPPUP by	Total operating expenditures for all grades K-12 divided by average daily attendance for all grades, K-12 (regular and adult students)
NFEXPPC	Total city expenditures, less expenditures on fire protection, divided by city population
POLEXPPC	Total city expenditures on police divided by city population
PRKEXPPC	Total city expenditures on parks and recreation divided by city population

TOTEXPPC	Total city expenditures divided by city population
TOTSCRE3	Average district total score (sum of reading, writing, and math scores) on the California Assessment Program battery of tests administered to third graders, 1986/87 school year
DR	Dummy variable indicating presence of an area in census tract designated as having a high risk of fault rupture. (1 if high risk, 0 otherwise)
DS	Dummy variable indicating presence of an area in census tract designated as having a high risk of slides (1 if high risk zero otherwise)
DC	Dummy variable indicating presence of an area in census tract designated as having a high risk of compressible soils (1 if high risk, zero otherwise)
ALT1	Altitude of municipality in feet above sea level.
GRADIENT	Change in altitude from municipality to nearest municipality (in feet per mile).
RAIN87	Rainfall in 1987 (in inches).
AGGY87	Aggregate income in relevant census tract, 1987 (in millions of dollars).

MIT OpenCourseWare
<http://ocw.mit.edu>

14.42 / 14.420 Environmental Policy and Economics

Spring 2011

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.