

Santa Barbara Housing Data Description and Dictionary

SPACE Workshop

July 2004

Stuart Sweeney

The data sets and an accompanying GeoDa lab were created for the Santa Barbara SPACE workshop. The lab focuses on a housing sales data set because of the ease of model interpretation of multiple regression hedonic models. If housing price is regressed against a set of housing characteristics (square footage, presence of a view, number of bathrooms, etc.) the coefficients have a very natural interpretation for an undergraduate. For example, given an additional bathroom, how much would we expect the sales price to increase?

The data and the lab are provided purely as a teaching tool. The observations are an ad hoc sample of convenience. Several observations were excluded due to missing values. The analysis in the lab is also provided more as a means of highlighting the functionality of the software than as a correct spatial analysis. The spatial econometric model specifications are trivial and have not been subjected to a full set of diagnostics.

The housing sales data is provided in three sets of files. The first set (sb_blkatt_p.*) contains the point locations of houses with some census data attributes attached to each point at the block group level. The second (sb_blkgrp_p.*) and third (sb_tract_p.*) sets report average values of the housing characteristics and census data at the block group level and tract level, respectively. The attached table and figures provide an overview of the data. Table 1 contains all of the variable definitions and their labels in the data sets. Note that the binary indicator for proximity to Highway 101 is only included on the first data set. Figures 1 and 2 provide some context for the observations. Figure 3 has three panels showing the point level, block group level, and tract level data. The three aggregation levels are included so that you can experiment with stability of exploratory and regression results of a single set of observations under two alternative aggregation and frame constraints.

Table 1. Housing sales variables and descriptions**I. Household level data**

<i>Variable Name</i>	<i>Description</i>
ID	Numeric ID used for spatial weights matrix
HH_TYPE	HH_type=1 for single family, =0 otherwise
DAYS	Number of days on the market
TIME	Time of sale (number of days from t=0)
PRICE	Sales price in dollars
H_SQFT	Square footage of the house
L_SQFT	Square footage of the lot
BEDS	Number of bedrooms
BATHS_F	Number of full bathrooms
BATHS_H	Number of half bathrooms
BATHS	Total bathrooms (BATHS_F+0.5*BATHS_H)
BUILT	Year the house was built
PPSFT_H	Housing price per square foot (PRICE/H_SQFT)
PPSFT_L	Lot price per square foot (PRICE/L_SQFT)

II. Region indicators

<i>Variable Name</i>	<i>Description</i>
<u>Elementary School (district)</u>	
S_ADAM	Adams (SBESD)
S_BRAN	Brandon (GUSD)
S_CLEV	Cleveland (SBESD)
S_EL_C	El Camino (GUSD)
S_EL_R	El Rancho (GUSD)
S_ELLW	Ellwood (GUSD)
S_FOOT	Foothill (GUSD)
S_FRAN	Franklin (SBESD)
S_HARD	Harding (SBESD)
S_HOLL	Hollister (GUSD)
S_HOPE	Hope (HSD)
S_ISLA	Isla Vista (GUSD)
S_KELL	Kellogg (GUSD)
S_LA	La Patera (GUSD)
S_MCKI	McKinley (SBESD)
S_MONR	Monroe (SBESD)
S_MONT	Monte Vista
S_MTN	Mountain View (GUSD)
S_OTHE	Not listed
S_PEAB	Peabody (SBESD)
S_ROOS	Roosevelt (SBESD)
S_VIEJ	Veija Valley
S_WASH	Washington (SBESD)

<u>Region</u>	
R_SBEAST	East of State Street
R_SBWEST	West of State Street
R_HOPE	Hope Ranch
R_GOLETA_S	South Goleta
R_GOLETA_N	North Goleta
HWY101	Within 0.2 miles from Hwy 101

Table 1. Housing sales variables and descriptions (cont.)

III. Block-group level or tract level data

<i>Variable Name</i>	<i>Description</i>
AREA	Block-group (or tract) land area
POP2000	Population in 2000 Census
POP00_SQMI	Population per square mile
HISPANIC	Population listing Hispanic as ethnicity
HOUSEHOLDS	Number of households
AVE_HH_SZ	Average household size
FHH_CHILD	Female headed households with children
AVE_FAM_SZ	Average family size
HSE_UNITS	Total housing units
VACANT	Vacant housing units
OWNER_OCC	Owner-occupied housing units
RENTER_OCC	Renter-occupied housing units
PCT_RENT	Percent of housing units that are rented
PCT_HISP	Percent of population that is Hispanic
PCT_FHHC	Percent of households that are FHHC

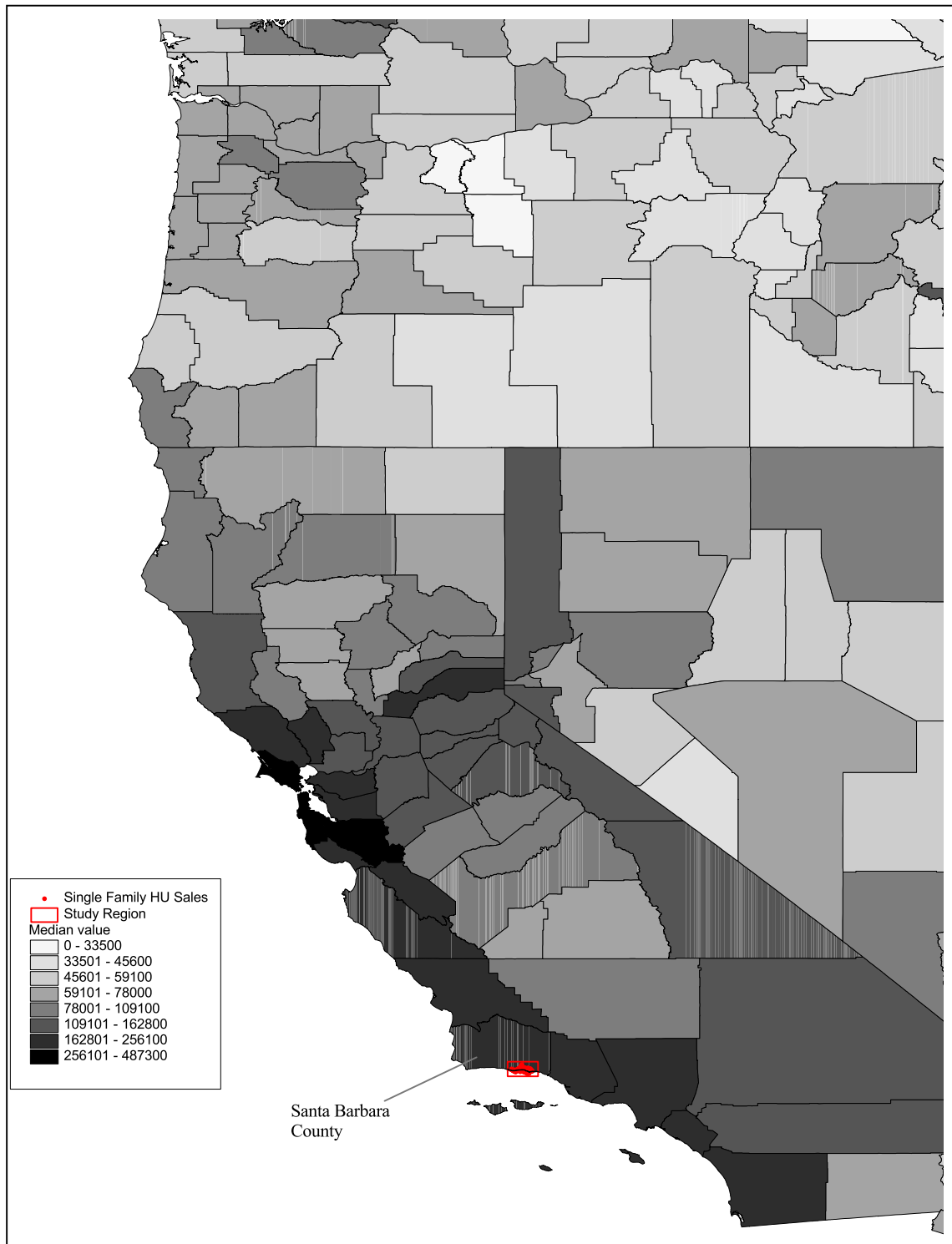


Figure 1: Median Housing Values

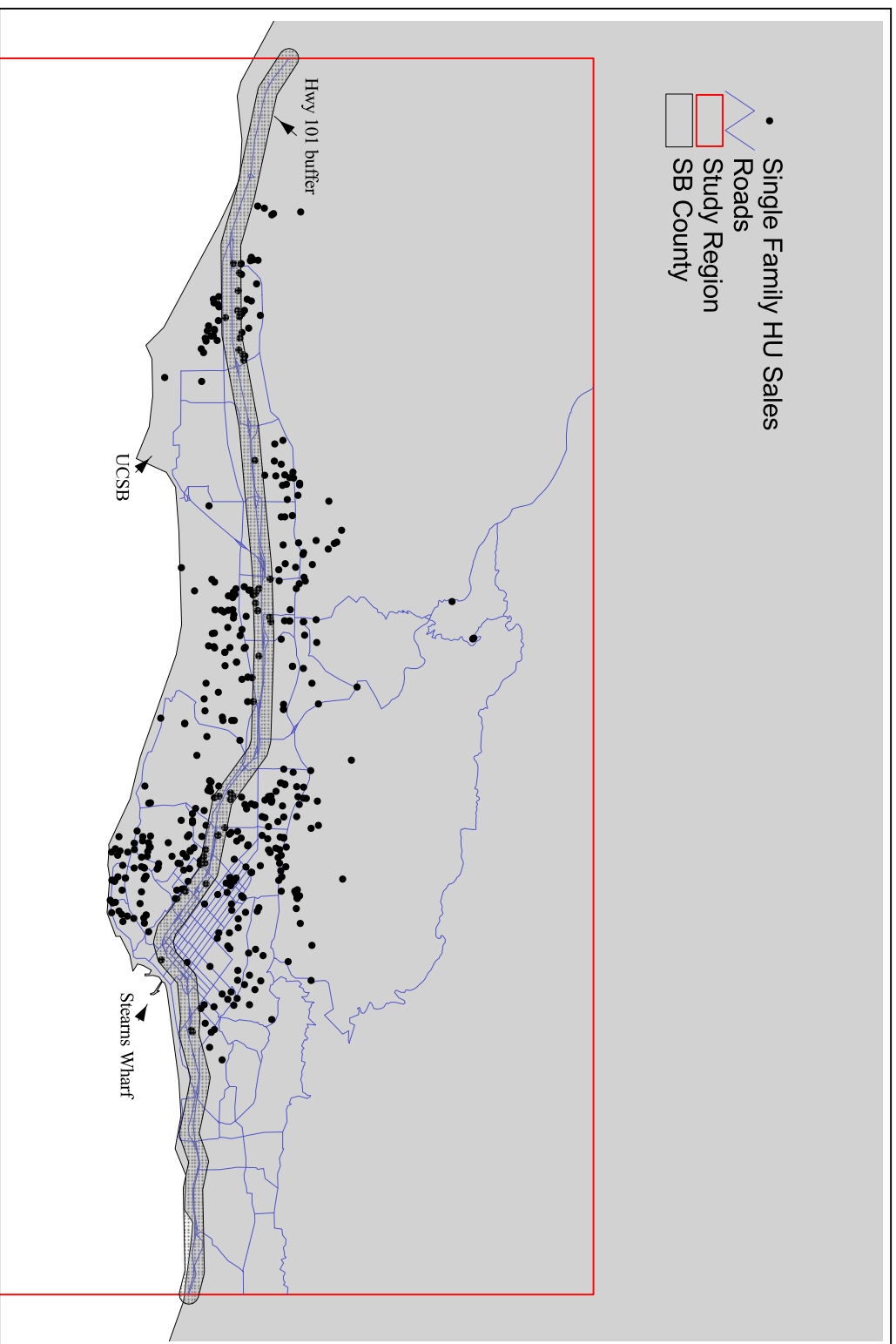


Figure 2: South Coast single family home sales

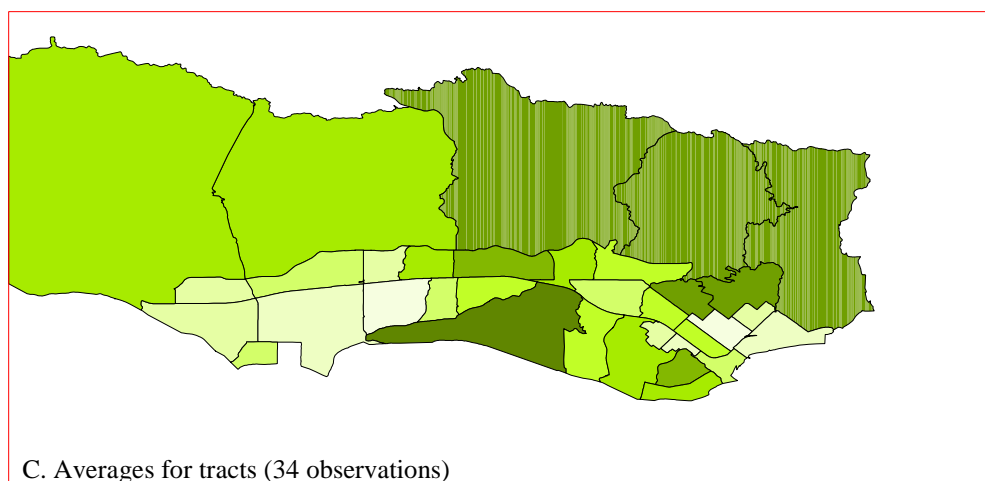
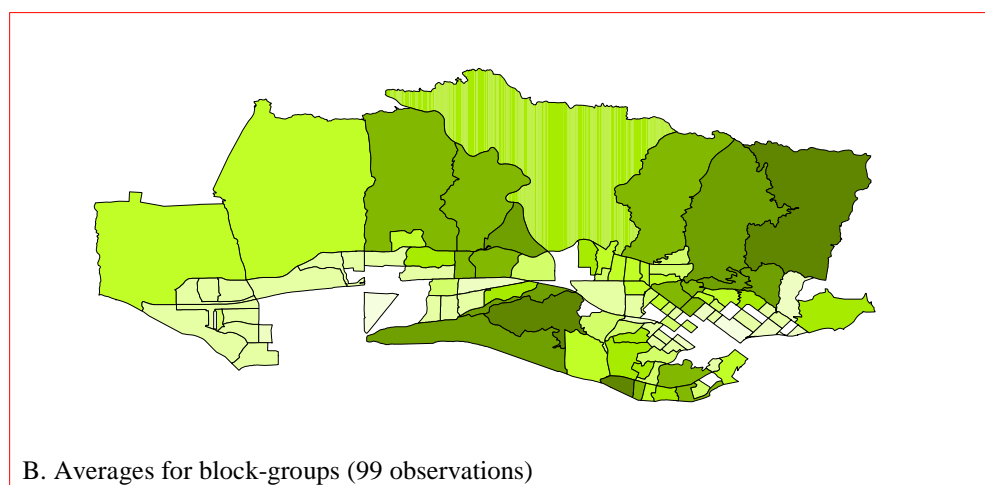
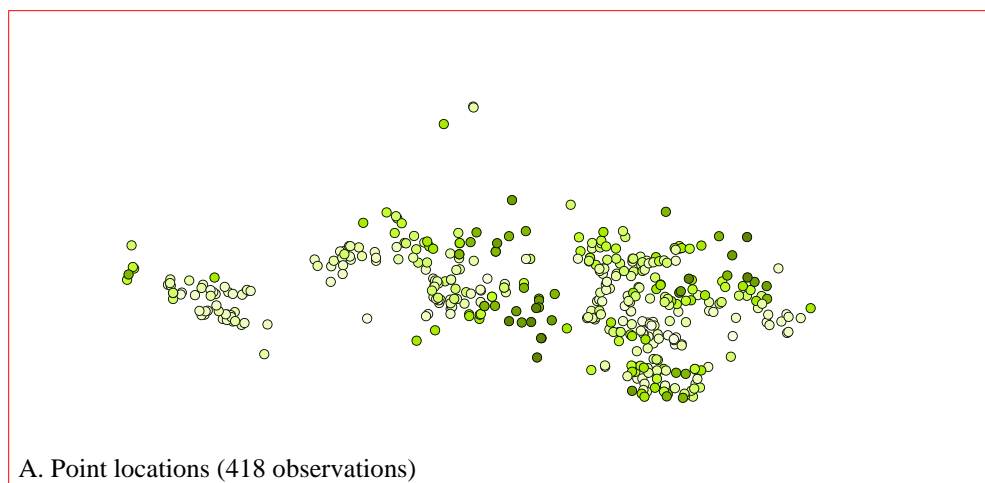


Figure 3: Housing price sales gradients for individual observations (A) and two areal aggregates (B and C)