

How to Build a Neighborhood House Price Index

While a primary objective of community revitalization is to raise property values, measuring that change can be difficult. One tool that can be used to track property values is a price index. Like a stock market index, a house price index allows investors to compare the change in value of one investment to others in the market. A good house price index allows buyers and sellers to estimate the current value of homes. Although most house price indices are used on a national scale, neighborhood level indices are also used by realtors and economists.

Step 1: Define the neighborhood. Are there at least 30 home sale transactions each year? Is each section of the neighborhood fairly similar? Are there significant differences in value in sub-areas, putting aside the quality of homes? Are the neighborhood's boundaries similar to those used by the city (plat number) or post office (zip code)? (this makes gathering data easier)

Step 2: Look at the housing stock. Were most of the houses built around the same time? Are they of similar size quality? If so the median price method may work well. If not, a constant quality index may be needed (see Types of Neighborhood House Price Indices).

Step 3: Find home sales data. Does realtor multiple listing service (MLS) data include all of the transactions located in the neighborhoods? (MLS data usually misses non-broker deals) How accessible is county or city data? Does this data contain many unusual transactions? (deeds signed years in advance, \$1 sales, etc) Is there data on the quality of the home? (floor space, baths, bedrooms, age)

Step 4: Transfer the data into a spreadsheet. Can you obtain the data electronically? Can pages of data be optically scanned by a computer? Will you have to hand enter data? How accurate is the data entry or optical scan?

Step 5: Choose an index method: median sales, repeat sales or constant quality hedonics. The following framework can help you choose an appropriate methodology.

- (a) If the neighborhood is homogenous, use a median sales price method.
- (b) If the neighborhood is heterogeneous (meaning there is a lot of variation), *and* historic home sales data is available, use a repeat sales method.
- (c) If the neighborhood is heterogeneous, *and* sales data has property characteristics *and* historic home sales data is not available, use a constant quality hedonic method.

Step 6: Adjust for inflation. Typically, use the Consumer Price Index for All Urban Consumers (CPI-UX) by the U.S. Department Of Labor Bureau of Labor Statistics (available on the internet at <http://www.bls.gov>).

For example:

CPI-UX U.S. City Average All items (1982-84=100)

1995 Q1:	150.9
1995 Q2:	152.2
1995 Q3:	153.0
1995 Q4:	153.9
1996 Q1:	155.1
1996 Q2:	156.5
1996 Q3:	157.5
1996 Q4:	158.8
1997 Q1:	159.6
1997 Q2:	160.2
1997 Q3:	160.9
1997 Q4:	161.8
1998 Q1:	162.0
1998 Q2:	162.5

CALCULATING THE CHANGE IN CPI	
<i>Assume it is 1998 and you want to view data from 1995 in real dollars, adjusted for inflation.</i>	
Take the CPI for current period:	162.5
Less CPI for base period:	<u>150.9</u>
Equals index point change:	11.4
Divided by base period CPI:	150.9
Equals	0.076
Result multiplied by 100:	0.076 x 100
Equals percent change:	7.6%

Step 7: Choose a benchmark period. Measure change from a particular year or quarter. Many indices choose one year and set it equal to 100. Hence, a 2% change in year 2 will result in an index value of 102. This can be easier to interpret than an absolute dollar value. Moreover, this allows a way to show prices without the effects of currency inflation.

Step 8: Compare your index to the entire market: Changes in neighborhood prices *must* be viewed in the context of the regional housing market. Compare the percentage change in the neighborhood index to the NAR or Freddie Mac index for the metropolitan area. If the increase in home values in the neighborhood beat the region, the market values the neighborhood more highly than in the past.

Step 9: Historic analysis. Calculate the rate of change for each period relative to the region. Does your index lead or lag the general trend in regional prices? Look at the peaks and troughs of the house price cycle. Is the change in your community more dramatic in downturns or less dramatic in upturns than the region? This tells you the relative risk and returns associate with house prices in the area.

Step 10: Maintenance. Every quarter, or year, depending on sales volume, update sales data. Continue to compare recent periods to historic ones and measure change over time. Every few years, re-benchmark the index to 100 and re-visit the methodology and assumptions used.

Types of Neighborhood House Price Indices

There are at least three ways these price indices are constructed, using: (1) median sales prices, (2) repeat-sales prices, and (3) constant-quality prices.

Median Sales Price: One of the most widely known house price indices is the National Association of Realtors (NAR) median sales price index for single-family, existing (as opposed to new) homes. The NAR tracks the sale price of all homes, including land, that are listed and sold by realtors.

Median sales prices are relatively easy to calculate. The only data needed is the date of the sale and final sale price. Once sorted from smallest to largest, the median value is the one in the middle (the average value could also be used, but an average is likely to be skewed by very high or very low-priced homes). Median values from previous periods can be compared to the current period. The difference between periods, once reduced by the value of inflation (using the government's Consumer Price Index for Urban Consumers, the CPI-UX) is the change in house price for the area covered, in real dollars. One advantage of a median sales price index is that it is easy to interpret - it is the price of homes and/or the change in the inflation adjusted prices of homes.

Most median price indices do not adjust for the quality of houses sold. If in one period, a high percentage of large homes sold, while in another period, more small homes than large ones sold, a median sales price index will change *simply based on the type of homes sold*. Since the NAR index includes so many homes nationally, this is usually not a problem. At the neighborhood level, with smaller numbers of homes selling in a given year, the median sales price calculation can be distorted by the sale several unusually small or large homes.

In neighborhoods where the housing stock is highly homogenous—that is almost all homes are of a similar size and quality, the median sales price method is likely to be reliable. However, if most houses in a neighborhood are very different from each other, the median sales price method can lead to uneven results. In periods when more large, expensive homes sell, the index will rise; when more small homes sell, it will fall. Yet, this change is not related to underlying price trends, but rather a result of the type of house sold.

Repeat Sales: Other house price indices control for, or hold constant, the quality of homes sold. One such index is the Freddie Mac Weighted Repeat Sales Index. This index determines the median change in house prices for homes that are financed through Freddie Mac's secondary market. Many homes pass through its portfolio several times as they are sold. Each time a home passes through its portfolio for both a purchase and a sale (over the space of years, usually) Freddie Mac calculates the median change in price.¹ Since change are measured for each house, the quality and size of the unit usually is constant. Prices are less effected by the types of homes sold, as opposed to a median sale price index. Of course, it is likely some homes will be improved between purchase and sale, perhaps adding a new

¹ Peek, Joe and Wilcox, James, *The Determinants of Single-Family House Prices* Federal Reserve Bank of Boston WP 91-7, December 1991.

kitchen or bath, for example. Unfortunately, these changes will cause a repeat-sales index to inflate moderately over time. However, this method provides a more consistent measure than median sales prices.

Calculating repeat-sales prices can be complicated. The simplest way is to create an un-weighted index of change in home values for the area. The index will go up or down based on the average, or median, annualized difference between the purchase and selling prices of homes.

For example, a home could sell in 1998 for \$60,000. Suppose the home was purchased for \$40,000 in 1988. Thus, this home had a \$20,000 increase in price. But, that \$20,000 includes inflation—the gain is not in ‘real’ dollars.’ That amount must be adjusted by inflation (using the CPI-UX). A \$20,000 gain may be worth \$10,000 after 50% in inflation over 10 years (5% annually). That gain must also be divided by the number of years the home was owned (10), or \$1,000 per year. That amount is then converted to a percentage gain from the initial investment, or \$1,000 divided by \$40,000, equaling 2.5%. The median gain for the neighborhood could be compared to other areas or used to inflate an index. For example, if the base year was 1997 at 100, 1998 would be 102.5.

Freddie Mac weights each sale so it represents a larger population of home sales, thus resulting in a national estimate (for example, the 100 two-bedroom homes sold in its portfolio could be weighted to represent the entire 12 million such homes that exist in the U.S.). This is usually too difficult to pursue at the neighborhood level, however.

Home purchase and sale prices can be obtained from the county recorder or city assessor’s office. Such data needs to extend at least 15 years in time to be used in a repeat-sales index. Even then, the original purchase price of many homes may be unknown. Beyond the address, sale price, and date, no property characteristics need to be tracked.

Constant Quality (Hedonics): A third approach is the constant-quality method, used by the Census’s Constant Quality C-27 Series (now part of the C-25). Using a statistical technique called hedonic regression, the prices of new homes sold are adjusted for 10 physical characteristics, including: floor area (ft²), number of stories, number of bathrooms, air conditioning, type of parking, type of foundation, geographic location, proximity to metropolitan area, land area (ft²), and the presence of a fireplace.² Other constant quality hedonic indices add the age of the unit, presence of gas heat, a pool, and a basement. By taking the average price, holding these factors constant, an index can be produced that controls for variation in the types of homes sold.

Creating a constant quality index requires detailed data. Calculations can be complicated, although most spreadsheet software systems can perform the linear regression required. Most of all, hedonics requires careful thought about what influences house prices.

In general, the price of a house is dependent on the supply and demand for houses. At a particular level of supply and demand, however, a number of factors influence house prices:

² US Bureau of the Census *Price Index of New One Family Homes Sold. Current Construction Reports Series C-27. 1991*

Price = (Land, Structure, Neighborhood, Accessibility, Location)

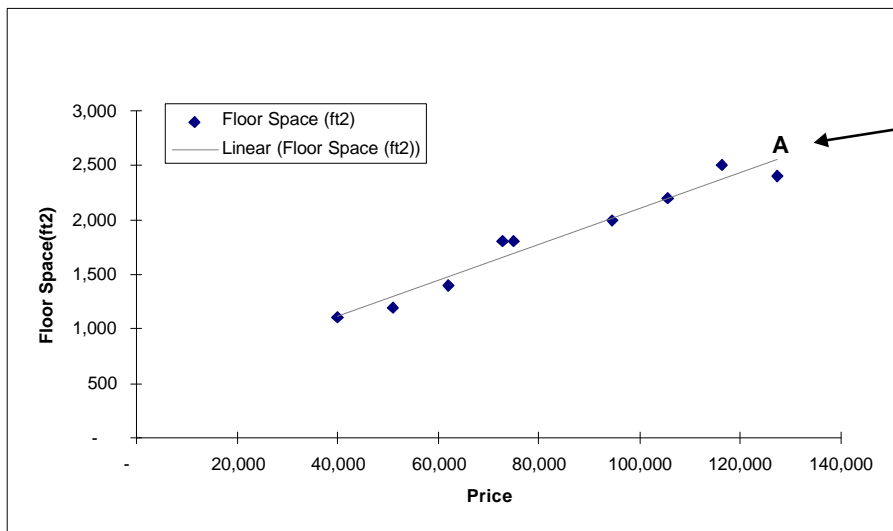
- I. Land is the quality and size of the property
- II. Structure is the quality and size of building
- III. Neighborhood includes people—such as the income, occupation and education of residents, and place—such as the share of units that are single-family detached homes, or the quality of local schools
- IV. Accessibility is the proximity to transit, shopping, recreation and employment
- V. Location is the proximity to enhancing qualities, such as parks, or to detracting qualities, such as crime and air and noise pollution

Neighborhood level price indices assume factors III, IV, and V are consistent from house to house. Only factors related to the house itself (I and II) will vary.

This method assumes that houses with more square feet, more bedrooms, a garage, or more baths, command higher prices. This relationship is assumed to be linear, as shown in diagonal line 'A' in the graph below. Each additional square foot of space, on average, raises home prices \$40.

This relationship can be shown as a linear equation: Price = the number square feet (b), times the cost per square foot (m), plus the value of the house over and above its size, described as a constant (x). The slope of the line (m) is the rate (\$ per foot).

Thus, at \$40 per foot, a 1,000 square foot house costs \$40,000, plus a constant, in this case \$10,000. Comparing cost per square feet, or the constant, across neighborhoods allows a comparison of similar prices, despite variation in the size and quality of house—it allows us to compare “apples to apples.”



Linear Regression:

Slope of line: $y=mb+x$

\$40 per ft² + x

1,000 ft² * \$40

=\$40,000 + x

x = constant value,
regardless of size.

By controlling for age, number of bedrooms, lot size and floor size, much of the variation in house prices can be explained, or predicted, by a linear equation. In fact, hedonic models sometimes explain 70% to 90% of all variation in house prices.

The table below shows a series of home sales data. The table on the right shows the results of a simple regression of price on the type of house. Based on the characteristics of these homes, the constant quality price of a home is \$68,381.

For each bedroom, add \$3,020. For an additional square foot, add \$3.64. For every year of age, *subtract* \$797 dollars. But holding these qualities constant, \$68,381 captures the value of an average home. To this value, the qualities of an average house could be added: For example:

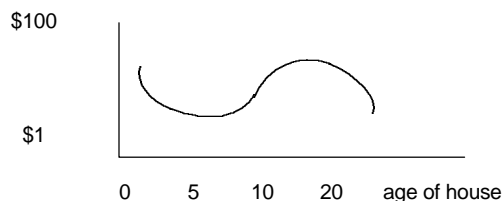
$$\begin{aligned}
 &\text{A 45 year old, 2 bedroom, 1,500 ft}^2 \text{ house on a 3,000 ft}^2 \text{ lot would cost:} \\
 &\$68,381 + 45*(-797) + 2*3020 + 1500*3.64 + 3,000*1.18. \\
 &\$68,381 - 35,865 + 6040 + 5460 + 3540 = \$47,556
 \end{aligned}$$

Sale Date	Address	Age	Bedrooms	Lotsize	Sqfeet	Price
2/3/98	27 Smith	80	2	2,900	1,100	40,087
2/6/98	99 Buck	57	3	3,000	1,800	72,772
2/8/98	98 Concord	45	3	4,000	2,000	94,562
2/4/98	54 Rob	60	2	3,340	1,200	50,982
2/5/98	9 Craig	55	3	3,390	1,400	61,877
2/9/98	8 Bond	30	4	6,940	2,200	105,457
2/7/98	20 Ford	50	3	6,240	1,800	75,000
2/11/98	47 Maple	10	5.5	6,000	2,400	127,247
2/10/98	3 Eagle	20	6	5,590	2,500	116,352

Regression Table	
<i>Regression of price on: age bedrooms lotsize sqfeet</i>	
price	Coefficient
-----	-----
age	-797.31
bedrooms	3020.32
lotsize	1.18
sqfeet	3.64
constant	68381.06
-----	-----

Microsoft Excel runs regressions as part of its Data Analysis Tools, or with the LINEST function. Using the data above, the column of house prices would be regressed on the columns age, bedrooms, lot size and square feet. The coefficient on the constant could be used to measure changes in the price index—or the average age, bedrooms and sizes could be multiplied by the appropriate coefficient to estimate the value of a ‘typical’ house. Like other methods, prices over time must also be adjusted for inflation by the CPI-UX.

One problem with this method is that it assumes linear relationships between houses and their characteristics. For example, consider the influence of the age of a home on its price. Newer homes are usually more valuable, since they are in better condition. Yet, homes tend to decline in value for their first 10 years, as they wear and age. Then, as homes are renovated or improved, they increase in value. Then they decline in value as these renovations depreciate. And so on. The graph (below) shows a cycle of prices to age.



As important as age is to determining house price, its relationship to average house prices is not linear. This violates the assumption shown in the “price and size” graph above.³ Nonlinear relationships can be corrected for by using exponents and other techniques. However, using such mathematical instruments requires a thorough understanding of how each variable impacts home prices, and can be complicated. Yet, this approach has the advantage of producing an index of house prices that truly reflects the market, without requiring the historic data used in repeat sales calculations.⁴

(c) Michael Collins 1998

³ Goodman, Allen and Thomas, Thibodeau, *Dwelling-Age-Related Heteroskedasticity in Hedonic House Price Equations* Journal of Housing Research 8(2) 1997.

⁴ The best way to use this approach is first to scatter graph price and each individual variable. Look at the plot and see what shape the relationship takes. If it is mostly linear, the hedonic method will work. If the plot of a variable has a U or wave shape, plot price against the square of that variable. If a linear form does not appear in this plot, try that variable cubed. A regression for a hedonic index may have the following form: regress price on lotsize floorsize bedrooms bathrooms garage age age² age³ lotsize² floorsize². The constant will produce an average home price, controlling for the quality of the home.