

Using Linear Regression in Real Estate Market Analysis and Projection

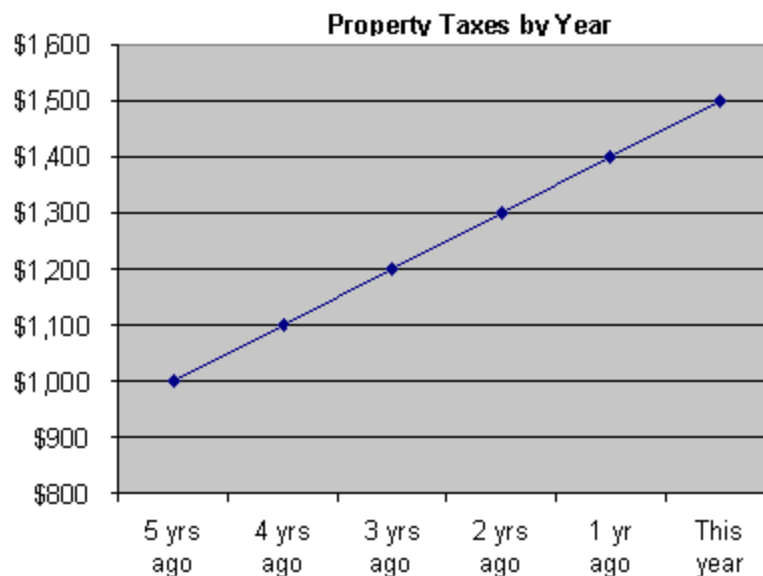
By Frank Gallinelli

Regression - no, it's not what your spouse accuses you of when you want to trade in the mini-van for a two-seater stick-shift convertible (well, maybe it is, but that's a topic for a different article).

Linear regression is a statistical technique of which we can make good use in our real estate analysis and projections. To put it into non-technical terms, it lets us look at a situation where we can take some facts that we know (dare we call them real data?) and use them to identify a trend. If a trend really does exist, it, in turn, allows us to predict the value of something otherwise unknown.

All right, that was vague enough to be part of the tax code. Now let's get concrete. Five years ago my property taxes were \$1,000. Four years ago they were \$1,100. Three years ago, \$1,200. Two years ago, \$1,300 and last year \$1,400. Given this trend, what we can we reasonably predict we'll pay this year? Right. \$1,500.

How did we guess? We probably had a flashback to our junior high school algebra class (talk about regression!). In the graph paper of our mind, we plotted a perfectly straight line. The line was formed by a series of data points and it clearly suggested a trend.



Each data point on this graph represents two pieces of information, or "variables:" an independent variable - time - plotted along the horizontal x-axis and a dependent variable - the tax amount - plotted along the vertical or y-axis. The first data point, therefore, is a dot that appears where "5 yrs ago" and "\$1,000" intersect. The second point lands where "4 yrs ago" and "\$1,100" intersect and so on. The tax amount is the dependent variable because it changes as a function of time. In other words the tax bill depends on the year, not the other way around.

When we play connect-the-dots as in the graphic above (hence the name linear regression), we see that those dots form a perfectly straight line. If we extend that line beyond our known data points a bit, we can see that in the current year, assuming that the trend line holds up, we could reasonably expect the taxes to be \$1,500.

Of course, in real life our ducks don't always line up so nicely in a row. When they look like this, we'll probably need computer software to fit the best possible line to the series of points. Then we can use the resulting straight line to make our predictions.



Using Linear Regression in Real Estate Market Analysis and Projection

By Frank Gallinelli

(Page 2 of 4)



There are numerous ways that we can use linear regression in real property analysis. We invite you to download a RealData® model, "Value" to give the concept a spin. "Value" is a Microsoft Excel® worksheet designed to help us estimate a property's worth using the market data, or comparable sales, approach to valuation.

This approach assumes that recent sales of properties that are nearby and are comparable to the subject provide the best indicators as to the value of the subject.

While we might use this model with almost any type of real estate, let's assume for the sake of example that we want to estimate the value of a single-family residence. Although previously sold homes may be comparable they are unlikely to be identical, either to each other or to the subject being appraised. One may have more land; another may offer more interior space; a third may boast a better layout and so on.

As a rule such differences are generally reflected in the selling prices of the homes. Properties that are otherwise similar sell for more or less as a function of their distinguishing features. If we can identify some measure (index) of the appeal or amenities of the properties in a given neighborhood, then we may also be able to discern a pattern between that measure and the value of the properties -- our trend line again. We can then use the pattern to predict the values of other properties in the same locale. Our model will permit us to determine by regression analysis whether or not a linear relationship exists between selling price and some independent variable that we define.

One possible technique is to use the property tax assessment as an index of value. Although assessments seldom reflect true market price, they often provide a good indication of relative value so they're worth a try. If the assessments and prices from a number of recent home sales in a neighborhood define a linear relationship, "Value" can measure the strength of that relationship and use it to estimate the worth of a home not yet sold.

After we open this model we can enter the address, an index and an adjusted selling price for as many as fifteen comparable sold properties. (Regarding the term "adjusted:" We may want to correct for price inflation whenever a sale is more than a few months old.) At the bottom (after #15, we'll enter the address and the index amount of the subject property. The program will fill in the field for the number of comparables used and compute the subject property's estimated selling price.

Property	Index	Adjusted Selling Price	# of Comparables Used
1 125 Wisteria	47,000	152,500	9
2 145 Wisteria	36,000	135,000	
3 325 Katalpa	29,000	114,900	
4 125 Elm	44,000	148,500	
5 680 Wisteria	41,000	146,000	
6 200 Elm	39,000	140,000	
7 240 Elm	37,500	133,900	
8 590 Wisteria	33,500	127,500	
9 455 Katalpa	31,000	125,000	
10			
11			
12			
13			
14			
15			
Subject 150 Wisteria	42,500	145,916	

Using Linear Regression in Real Estate Market Analysis and Projection

By Frank Gallinelli

(Page 3 of 4)

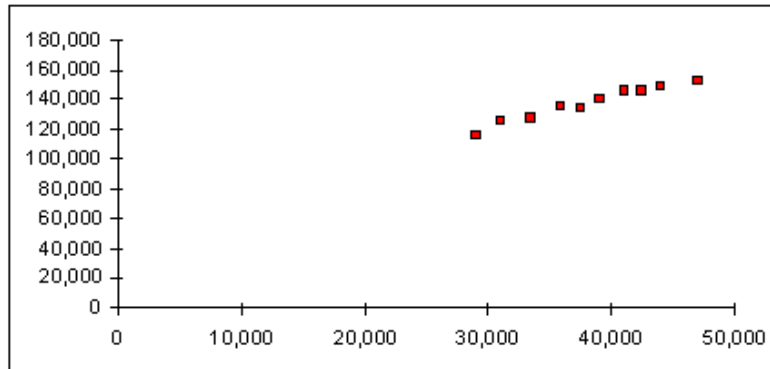
The results appear in a report and graph, in the section below.

Market Analysis by Linear Regression

Property	Index	Actual Selling Price (adj.)	Projected Selling Price	Diff.	% Diff.
125 Wisteria	47,000	152,500	155,012	-2,512	-1.65%
145 Wisteria	36,000	135,000	132,778	2,222	1.65%
325 Katalpa	29,000	114,900	118,629	-3,729	-3.25%
125 Elm	44,000	148,500	148,948	-448	-0.30%
680 Wisteria	41,000	146,000	142,884	3,116	2.13%
200 Elm	39,000	140,000	138,842	1,158	0.83%
240 Elm	37,500	133,900	135,810	-1,910	-1.43%
590 Wisteria	33,500	127,500	127,725	-225	-0.18%
455 Katalpa	31,000	125,000	122,672	2,328	1.86%

Subject Property **Index** **Projected Selling Price**
150 Wisteria **42,500** **\$145,916**

Correlation Coefficient: **0.981**



Notice that the program will specify a correlation coefficient. This is a new bit of terminology we didn't see in our simplified explanation above. This number is a statistical measurement of the reliability of the relationship between the index and the adjusted selling price. To put it another way, it's a numerical way of expressing how straight our dots line up.

A correlation of 1.00 is a perfect relationship, while zero indicates that we have completely random data. In most cases, a correlation coefficient of less than 0.90 means that the relationship between index and selling price is not sufficiently reliable for use as the basis of a prediction.

As an interesting sidebar, we can see how accurately this regression analysis would have predicted the values of the homes whose actual selling prices we know. That is because the program computes and displays the selling prices that the analysis



Using Linear Regression in Real Estate Market Analysis and Projection

By Frank Gallinelli

(Page 4 of 4)

would have predicted for each of the comparables. We also see the dollar and percentage differences between the projected and actual prices. This section provides a very graphic demonstration of the accuracy - or inaccuracy - of "Value's" prediction.

We need to keep in mind that, as with most projections, the quality of our output is entirely dependent on the quality of our input. We certainly have to make appropriate choices for our comparables. Otherwise we can't reasonably expect to achieve meaningful results. In addition, the kind of index we select must relate consistently to value. If we find tax assessments to be unreliable, we may want to try gross living area or experiment with a scoring system (X points for each bedroom, Y points for each bath, etc.).

A regression analysis like the one provided in this model can be very useful because of its ability to provide statistical support to what might otherwise be a subjective estimate of value. Property sellers and buyers can use it to support price negotiations; and agents can use it to enhance the effectiveness of their listing presentations. With a bit of imagination, linear regression can be used in many ways to poke and prod our analyses and projections. It's name notwithstanding, it can take us a big step forward.