**MEMORANDUM**

**To:** Madera & Associates

**From:** Theo Carr

**Date:** 20 March 2018

**Re:** Effect of UFFI on Florida property values

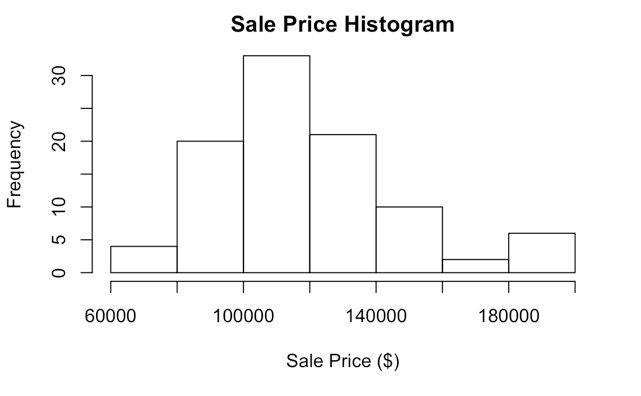
**Summary:** Based on real estate sales data collected from 2009 to 2016, UFFI has a statistically significant negative effect on the sales price of a property. Comparing the average sales price of properties with and without UFFI revealed that historically the chemical has decreased the value of a property by about $7000, on average. Although UFFI is not statistically significant when predicting the value of a house with other factors, on its own it can be shown to have a meaningful effect.

**Analysis:** Analysis was performed on 99 property sales from 2009 to 2014. The following independent variables were investigated as possible predictors for sale price of a given property:

* Year sold
* Basement area
* Lot area
* Living area
* Presence of UFFI
* Presence of brick exterior
* Presence of central air conditioning
* Presence of pool
* Number of enclosed parking spaces
* Whether older than 45 years

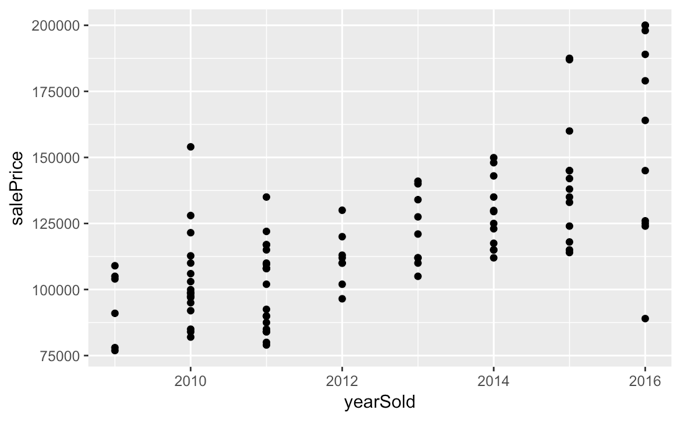
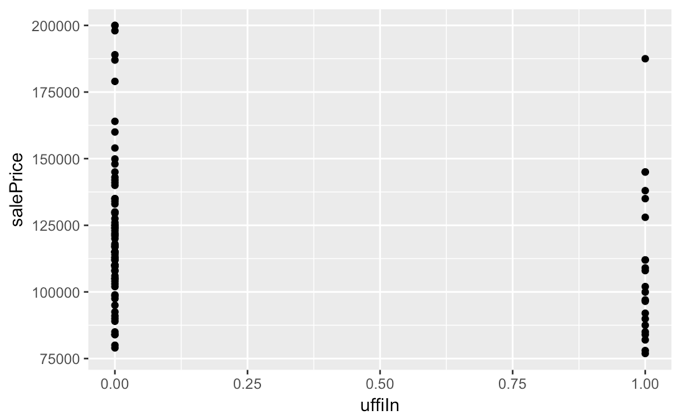
*Outlier detection:* For the four continuous variables, there were three cases that fell farther than three standard deviations from the mean. These cases were removed for the following analysis.

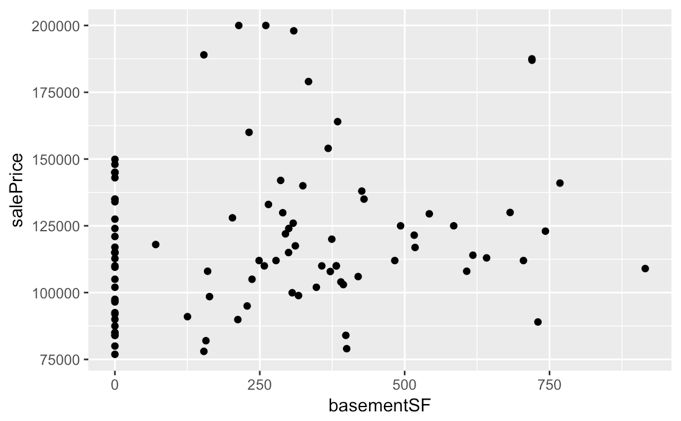
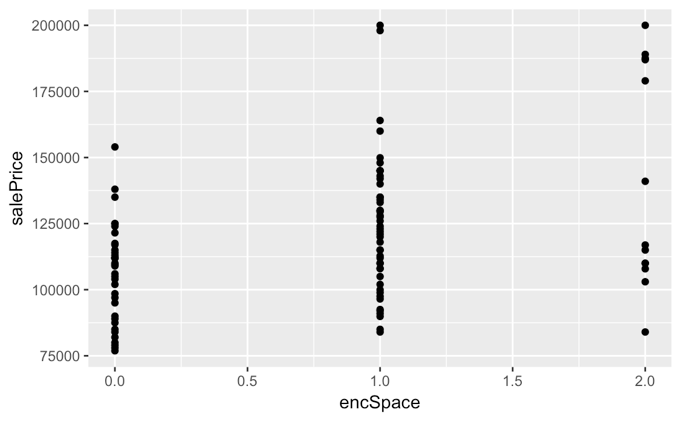
*Distribution assumption:* Creating a histogram of Sale Prices confirmed that the data is normally distributed, allowing us to perform parametric statistical analysis:

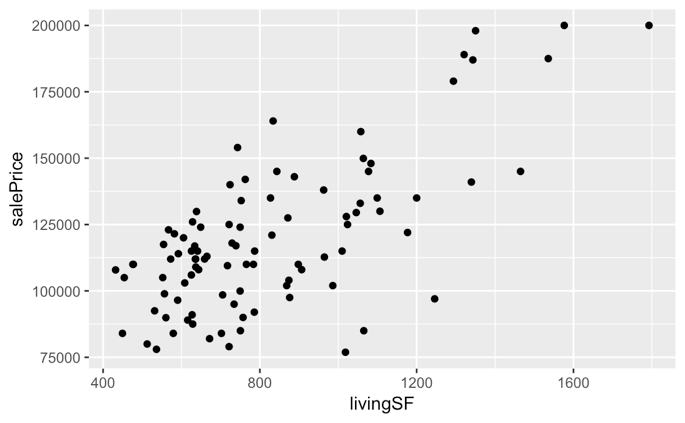


*Significance of UFFI:* Two-tailed z and t-tests were used to test the null hypothesis that the mean sale price of homes with UFFI (~$109,000) was equal to the mean sale price of homes without UFFI (~$123,000). In both cases the null hypothesis was rejected at a significance level of 0.05. The t-test presents a 95% confidence interval with lower bound $420 and upper bound $27,000. This result implies that the difference in mean sale price of houses with and without UFFI is not coincidental, and should be investigated further.

*Correlation:* A multiple regression model was constructed to determine whether UFFI was a significant predictor variable of selling price when taken with the full set of variables. Prior to building the model, the independent variables’ relationship with sale price was visually inspected.Several examples are shown below:



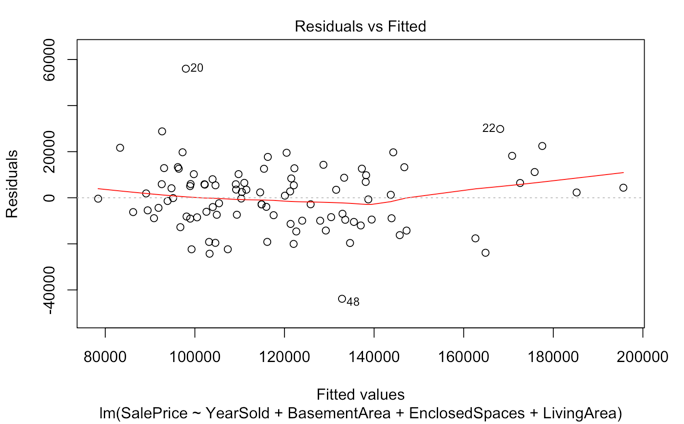
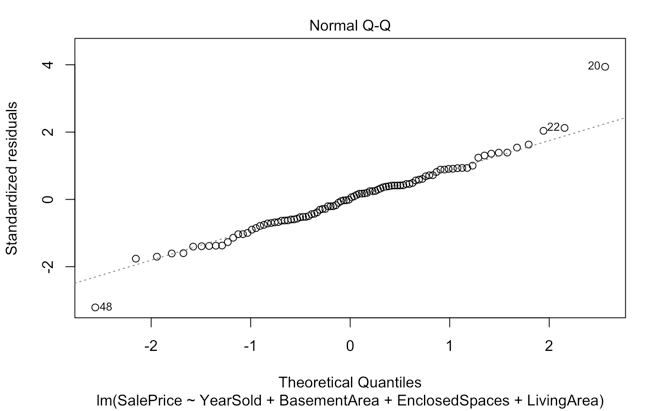
The graphs indicate that the strongest predictors of sale price appear to be year sold and living area, although each of the other variables display correlations as well.

*Model Construction:* The multiple regression model was formed using backward elimination, whereby the variable with the highest p-value greater than 0.05 was eliminated, until all remaining variables had p-values less than 0.05. Presence of UFFI was the last variable to be eliminated, with a p-value of 0.063 in the penultimate model. This result indicates that UFFI is on the cusp of being considered a significant predictor in our model. Notably, a larger sample size would likely reduce the p-value and allow for the inclusion of UFFI in our “ideal” model. The four significant predictor variables were identified as year sold, basement area, living area, and enclosed parking spaces.

While we cannot confidently conclude that the coefficient of our model’s UFFI variable is not 0, we may use the model to estimate the effect of UFFI’s value on the sale price of a home. According to the model, UFFI has negative effect of about $7000 (with standard error of $3500) when taking into account the four significant predictor variables identified above.

*Model Validation:* Note that due to the small sample size, the model was validating using the same data as for training. This type of analysis may lead to overfitting, and could be remedied by performing k-fold cross-validation as part of a deeper analysis.

The errors (residuals) from comparing our model’s estimates to actual values do not appear to follow any particular pattern, indicating that the model does not have any obvious flaws. A quantile-quantile plot offers supporting evidence that the underlying distribution of our model’s predictions matches that of the actual sale price data.

*Applying the model:* To understand the potential flaws and usefulness of the model, consider a house with the following characteristics:

* older than 45 years
* doesn’t have a finished basement
* has a lot area of 5000 square feet
* has a brick exterior
* 2 enclosed parking spaces
* 1700 square feet of living space
* central air
* no pool

A key missing data point in this case is the year in which the property was sold. The price of a house depends heavily on the time when it is sold - for example, a house sold in 2006 shortly before the housing market crash of 2008 would have a much different valuation two years later. For this reason, we should specify a year. To illustrate this point, we examine the expected sale price for the same home in 2009, 2012, and 2016, according to our model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Without UFFI | | | | |
| Year Sold | Estimate | Lower bound | Upper bound |
| 2009 | 149,129.00 | 117,419.10 | 180,838.80 |
| 2012 | 165,702.90 | 135,015.50 | 196,390.20 |
| 2016 | 187,801.40 | 157,514.70 | 218,088.10 |

|  |  |  |  |
| --- | --- | --- | --- |
| With UFFI | | | |
| Year Sold | Estimate | Lower bound | Upper bound |
| 2009 | 142,361.30 | 110,589.10 | 174,133.40 |
| 2012 | 158,935.20 | 127,978.40 | 189,891.90 |
| 2016 | 181,033.70 | 150,199.30 | 211,868.10 |

Clearly, the expected sales price of a house depends heavily on when it was purchased. If $215,000 was paid for this particular home, the amount of overpayment depends largely on how we define overpayment, and is complicated by when a house is sold. For example, we must take into account the real value of money: $1000 in 1980 was worth more than $1000 today. If two houses sold for the same price, but 20 years apart, the "real" sales price (adjusted for inflation) would be different.

We must also take into consideration market cycles. As discussed previously, the sale price of a house would be dramatically different if sold in 2006, before the market collapsed, versus 2009, in the midst of the downturn. This makes determining the actual value of a home tricky, and means it is difficult to separate the sale price of a home from when it was (or will be) sold.

Additionally, we must take into consideration other external market factors - if a bidding competition drives up the Sale Price of a house, does this mean that the winning bidder should be compensated for overpayment? In a capitalist market, the answer is likely no, because another buyer was willing to purchase the house at a slightly lower price.

*Further analysis:* To examine the average sales price over time, three forecasting models were constructed: weighted moving average, exponential smoothing, and time-series linear regression. The linear regression model had the lowest Mean Squares Error (MSE), indicating it was the best model for predicting future average sale prices. We expect this result because the trend of sale price consistently increased from 2009 to 2016. Therefore, a model which places maximum weight on the recent (and therefore greatest) value is closest to the next value. In practice, this means we should only look at the most recent value, ignoring all older values. Over a longer time period, where we don't see a relatively constant upward trend in average sales price, a different weighting might make more sense.

**Conclusions:** It is highly likely that the presence of UFFI has a negative effect on the sale price of a home, following from the z and t-tests described in the analysis section. While the presence of UFFI was not found to be a significant predictor variable in a multiple regression model, the p-value close to 0.06 indicated that a larger sample size would likely make UFFI a statistically significant factor in the model. According to estimates gleaned from this model, UFFI could be expected to decrease the value of a home by about $7000, with a standard error of $3500.

**Recommendations:** Our firm recommends presenting the evidence that historically, there is a statistically significant difference in sale price between houses with and without UFFI. Combined with the knowledge that the previous homeowner knowingly withheld this information (while also knowing that this information would likely lower the sale price of the home), it seems likely that your firm’s client is entitled to a compensation of several thousand dollars. At the very least, there is sufficient statistical evidence to indicate that further investigation should be done for this case. A larger sample size would likely confirm our preliminary findings that UFFI has a negative effect on the sale price of a home, and would help to quantify the effect. To this end, we also recommend that additional data is collected.