Collin DeVore

Nicholas Kaukis

STAT 3013

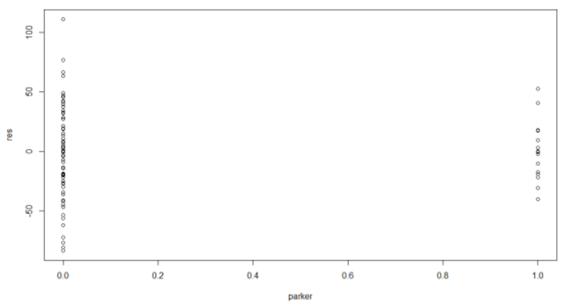
11/16/15

Linear Regression Analysis

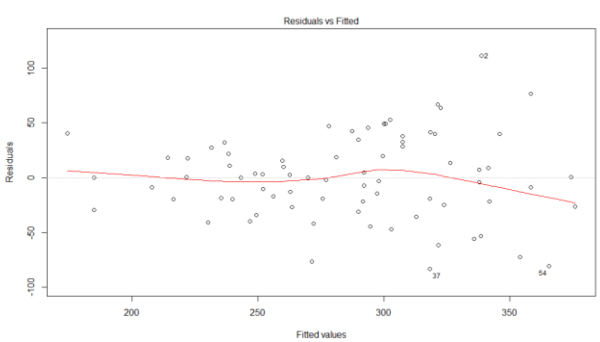
When analyzing the Homes model, I first used regression on the model so that I could get an idea of what it looked like with multicollinearity and have something to measure the progress of subsequent models off of. This gave an overall p-value of 8.906(e^-8), showing that there was some multicollinearity issues. I then created and analyzed the interactions that could be taking place between the factors using the "pairs" function on RStudio. Many of the plots did not seem to have much interaction between them, even though there were some that had an obvious interaction between them, such as Age and Agesq. These two factors seem to have a hook or parabolic shape between them. Also, Bathbed had an anticipated interaction with both Bath and Bed. None of the factors, however, had an obvious and unanticipated interaction with the other factors, so they were all left in at this point in the analysis. In order to find a model that more accurately represented the price, the p-values of the factors were calculated next to see what variables are fairly unrelated to price.

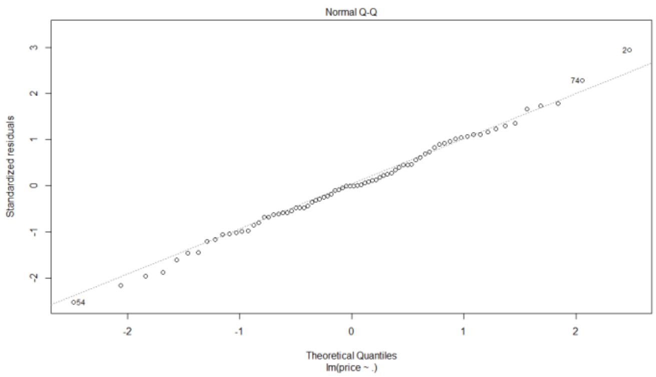
Looking at the p-values, there were some factors that had too high of p-values to be used. These variables were Age (0.24784), Garage (0.22460), Adams (0.30998), Crest (0.68266), and Parker (0.38942). Out of the five dummy variable sets regarding the different elementary schools that were nearby, only Harris (0.01685) and Edison (0.00187) had a significant impact on the price of the home, with Edison being more significant than Harris. This could mean that these two are prestigious schools, or that they are located near more expensive houses, though it does not imply that these schools are necessarily the cause of the price raise.

After looking at the p-values, I plotted the residuals in order to see if there is any problem regarding the data with the lower p-values. The residuals for the age and the garage appear to be spread well enough that there may just not be a correlation, but the residuals associated with the dummy variables seemed to be a bit heavy on one side and not well spread on the other, suggesting that they may not contain an accurate depiction of the dummy variables. An example of the problems with the residuals of the dummy variables can be seen with the Parker residuals graph shown below.



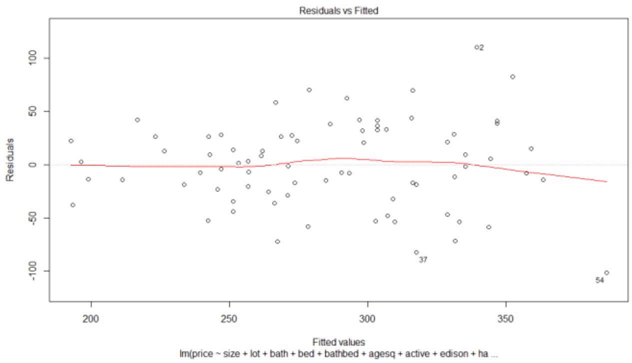
When I was done looking at all of the p- values separately, I looked at the graph of the residuals of the whole equation. The mean of the residuals had a fairly constant mean of zero, but the variance appears to have a funnel or cone shape that gets much larger to the right of the graph. The qq line kept showed that the residuals were normally distributed. The graphs of the residuals and the qq line are shown below.

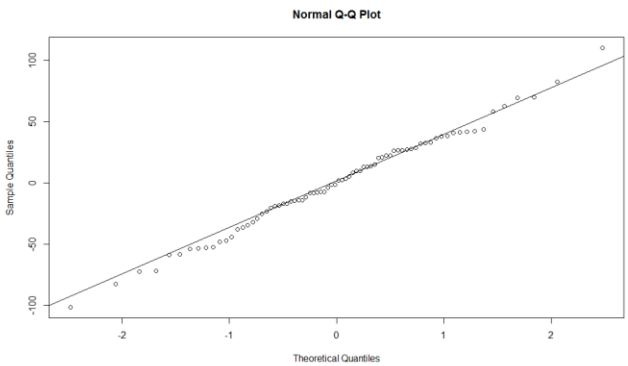




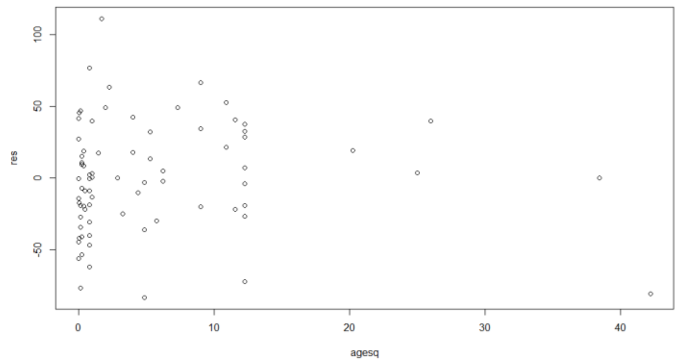
After looking at both the p-values and the residuals of the variables, a new model was used based on the factors with the lower p-values. The factors in this new model and their original p-values are Size (0.049), Lot (0.00518), Bath (0.01961), Bed (0.01961), Bathbed (0.0338), Agesq (0.01905), Active (0.01614), Edison (0.00187), and Harris (0.01685). These factors together give a p-value of 8.821(e^9). Although there is not much of a change in the overall p-value, it seems that the individual p-values are low enough to support the idea that these values fit the line the best. It may also signify that there are many values outside of this dataset that have not yet been explored. The new p-values for each of the factors are Size (0.004664), Lot (0.002776), Bath (0.061327), Bed (0.005313), Bathbed (0.026411), Agesq (0.091283), Active (0.005476), Edison (0.000315), and Harris (0.001001).

After the p-values of the new model were calculated, the new interaction plot was analyzed. The change in the factors gave more residuals that had a mean that appeared to equal zero. The variance was not well spread, though there appeared to be almost no interactions between factors excluding the Bathbed, Bed, and Bath factors' interaction. With many of the factors, there is an issue with constant variance, though this issue can be argued for each factor, and therefore these variables can be left in the formula. When all of the residuals are taken together, the variance appears to have improved due to the fact that the cone shape of the residuals are not as drastic. A model of the residuals when taken together is shown below, along with a model of the qq line, which showed that the residuals stayed normally distributed.





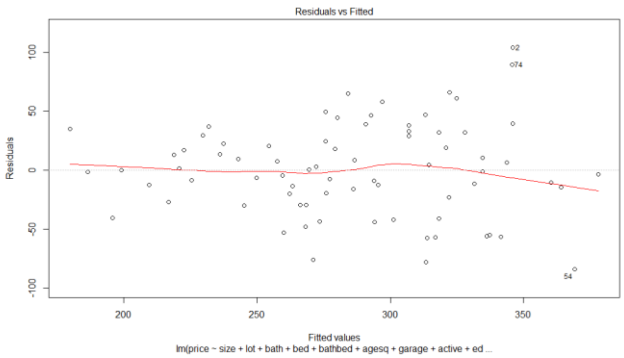
The rise in the p-value of Agesq suggests that a closer inspection of the factor is appropriate. Looking at the residuals of Agesq, five points can be seen that appear to be off to the side. These can skew the mean of the residuals and influence the values of the factor, which shows that it may not be a good depiction of the actual value of age squared. After that, there is the fact that there is a cluster of data points near the y axis, which does not appear to be significant since they are centered somewhere around 0. Even though these problems exist, the data is fairly distributed and there appears to be a constant variance, leading me to believe that the rest of the data points account and make up for this deviation of the later residuals. The graph of the residuals of Agesq is shown below.

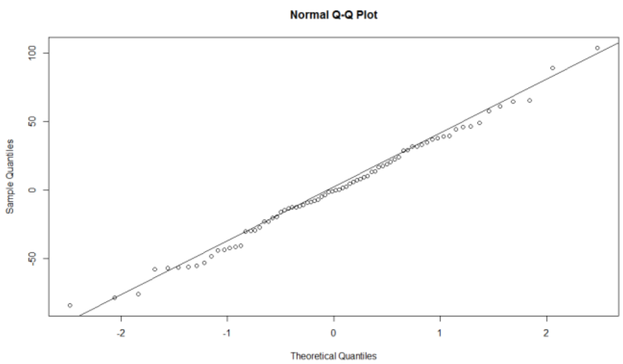


After analyzing the residuals associated with the Agesq factor, I analyzed the residuals associated with the other factors. For the dummy variables, there appears to be a larger spread and more variables on the right side, suggesting that they are better indicators for the model than the other dummy variables mentioned earlier. Some of the residuals for the other factors are questionable, such as the size, in which there is a slight chance that there may not be a constant variance. Most of the variance problems are caused by a single data point, however, so it still seems that this factor can be used as a fair predictor of price.

After the first two equations had been analyzed, I made one last equation using stepwise regression. Using the Aikaike Information Criterion, I obtained a new equation where the factors Size, Lot, Bath, Bed, Bathbed, Agesq, Garage, Active, Edison, and Harris were used. The interesting part of this equation is that all of the same factors are used in the second equation, except for garage, which must have had a higher p-value when only the other factors were used. This could be due to some multicollinearity issues that can be resolved when the other factors are taken out. Taken together, the Aikaike Information Criterion (AIC) of the first equation used is 800.4267, the AIC of the second equation used is 798.6595, and the AIC of the third equation is 795.6837. This shows a slight improvement from the original equation, though not by much.

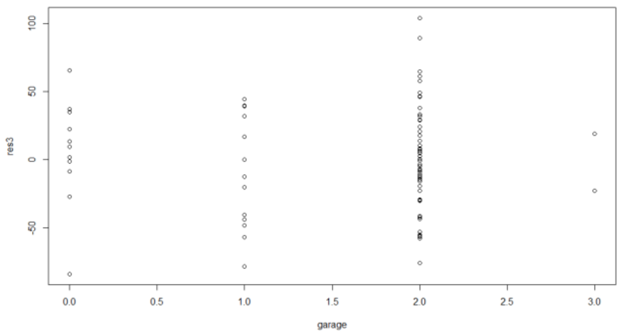
Looking at the residuals, the last equation fixes many of the issues associated with the first two models. The mean stays at about zero, the residuals follow the normality assumption as shown by the qq plot, and there is almost a constant variance. There is still a slight cone shape, but, for the most part, that has been fixed, making it more usable. The residual plot and qq line are shown below.





After analyzing all of the residuals together, I analyzed them separately.

While there were still some issues associated with the spread of the residuals and some clustering problems, especially with Agesq which seemed to be having the same issues that it had earlier, the overall spread of these factors had improved. This may have occurred because stepwise regression took out the factors that were having the most issues, thus insuring that the factors that predicted the price best were used. In many of the residuals, there was not a constant variance, yet all other assumptions appeared to have been met. For instance, many had a mean that was equal to zero, though, to some extent, it could be argued that knowing one residual may help a person to better predict the location of another residual due to the shape of them. An example of one of the residual plots is shown below, in which the residuals of the factor Garage are shown.



The last procedure that was done was to make one last regression line model in which the factors were squared in order to see if there was any interaction between or within the factors during the analysis. Unfortunately, this yielded no results, showing that the optimal equation that can be used is equation three. None of the equations could account for most of the prices, showing that there may be other factors at work when deciding the price of the home. Many of these factors do still appear to have an effect, however. More studies can be done in order to see what other equation can influence the price of a home, though this is the best that can be done with the information provided.

Appendix

Code

```
attach(homesr)

pairs(~ ., data=homesr)

homes.model1 <- lm(price~., data = homesr)
summary(homes.model1)

plot (homesr$age, homesr$price)
plot (homesr$agesq, homesr$price)

res <- residuals(homes.model1)
plot (size, res)
plot (lot, res)
plot (bath,res)
plot (bed,res)
plot (bathbed,res)
plot (age,res)
plot (agesq,res)
plot (garage,res)
plot (active,res)
plot (edison,res)
plot (harris,res)
plot (adams,res)
plot (crest,res)
plot (parker,res)

qqnorm(res)
qqline(res)

homes.model2 <- lm(price~size+lot+bath+bed+bathbed+agesq+active+edison+harris, data=homesr)
summary(homes.model2)

pairs(~size+lot+bath+bed+bathbed+agesq+active+edison+harris)

plot(homes.model1)
plot(homes.model2)

step (homes.model1)

homes.model3 <- step (homes.model1)
summary(homes.model3)

plot(homes.model3)

pairs(~size+lot+bath+bed+bathbed+agesq+garage+active+edison+harris)

AIC(homes.model1)
AIC(homes.model2)
AIC(homes.model3)

res2 <- residuals(homes.model2)
plot (size, res2)
plot (lot, res2)
plot (bath,res2)
```

plot (bed,res2)
plot (bathbed,res2)
plot (age,res2)
plot (agesq,res2)
plot (garage,res2)
plot (active,res2)
plot (edison,res2)
plot (harris,res2)
plot (adams,res2)
plot (crest,res2)
plot (parker,res2)

res3 <- residuals(homes.model3)
plot (size, res3)
plot (lot, res3)
plot (bath,res3)
plot (bed,res3)
plot (bathbed,res3)
plot (age,res3)
plot (agesq,res3)
plot (garage,res3)
plot (active,res3)
plot (edison,res3)
plot (harris,res3)
plot (adams,res3)
plot (crest,res3)
plot (parker,res3)

qqnorm(res2)
qqline(res2)

qqnorm(res3)
qqline(res3)

homes.model4 <-lm(price~.^2,homesr)
summary(homes.model4)

Model Summaries

```
> homes.model1 <- lm(price~., data = homesr)
> summary(homes.model1)

Call:
lm(formula = price ~ ., data = homesr)

Residuals:
    Min      1Q  Median      3Q     Max
-83.284 -22.628  -0.066  27.790 111.323

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 337.6628   124.9621   2.702  0.00891 **
size         58.7689    29.2569   2.009  0.04900 *
lot          10.3619     3.5731   2.900  0.00518 **
bath        -98.7362    47.9507  -2.059  0.04376 *
bed         -77.4817    32.3252  -2.397  0.01961 *
bathbed      29.6573    13.6582   2.171  0.03380 *
age           3.7771     3.2371   1.167  0.24784
agesq         1.8236     0.7571   2.409  0.01905 *
garage       10.6773     8.7030   1.227  0.22460
active       30.3572    12.2685   2.474  0.01614 *
edison       59.2149    18.2076   3.252  0.00187 **
harris       40.2345    16.3717   2.458  0.01685 *
adams       -28.8890    28.2176  -1.024  0.30998
```

```
crest         -8.8819    21.6213  -0.411  0.68266
parker       -13.9336    16.0736  -0.867  0.38942
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 42.37 on 61 degrees of freedom
Multiple R-squared:  0.5989,	Adjusted R-squared:  0.5068
F-statistic: 6.505 on 14 and 61 DF,  p-value: 8.906e-08
```

```
> homes.model2 <- lm(price~size+lot+bath+bed+bathbed+agesq+active+edison+harris, data=
homesr)
> summary(homes.model2)

Call:
lm(formula = price ~ size + lot + bath + bed + bathbed + agesq +
    active + edison + harris, data = homesr)

Residuals:
     Min       1Q   Median       3Q      Max
-101.578  -23.721    0.133   27.577  110.508

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  319.093    108.943   2.929 0.004664 **
size          71.358     27.978   2.550 0.013086 *
lot           10.617      3.416   3.108 0.002776 **
bath         -82.136     43.148  -1.904 0.061327 .
bed          -82.000     28.443  -2.883 0.005313 **
bathbed       27.523     12.119   2.271 0.026411 *
agesq          1.237      0.722   1.714 0.091283 .
active        31.853     11.090   2.872 0.005476 **
edison        62.787     16.509   3.803 0.000315 ***
harris        51.584     14.978   3.444 0.001001 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43 on 66 degrees of freedom
Multiple R-squared:  0.553,	Adjusted R-squared:  0.492
F-statistic: 9.072 on 9 and 66 DF,  p-value: 8.821e-09
```

```
summary(homes.model3)

Call:
lm(formula = price ~ size + lot + bath + bed + bathbed + agesq +
    garage + active + edison + harris, data = homesr)

Residuals:
    Min      1Q  Median      3Q     Max
-84.234 -24.091  -0.581  29.075 104.106

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 322.7285   106.2569   3.037 0.003435 **
size         59.6082    27.8541   2.140 0.036112 *
lot           9.3038     3.3893   2.745 0.007815 **
bath        -92.8312    42.3867  -2.190 0.032111 *
bed         -79.6453    27.7610  -2.869 0.005549 **
bathbed      29.5347    11.8574   2.491 0.015308 *
agesq         1.4557     0.7118   2.045 0.044878 *
garage       16.2220     7.7353   2.097 0.039875 *
active       27.7465    10.9913   2.524 0.014040 *
edison       61.9646    16.1050   3.848 0.000275 ***
harris       49.9654    14.6272   3.416 0.001100 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 41.93 on 65 degrees of freedom
Multiple R-squared:  0.5813,	Adjusted R-squared:  0.5169
F-statistic: 9.025 on 10 and 65 DF,  p-value: 4e-09
```

```
> homes.model4 <-lm(price~.^2,homesr)
> summary(homes.model4)

Call:
lm(formula = price ~ .^2, data = homesr)

Residuals:
         1          2          3          4          5          6 
-9.820e-15  7.819e-15  8.766e-15  5.991e-16 -3.992e-16  5.703e-16 
         7          8          9         10         11         12 
-2.241e-15 -3.763e-14  8.227e-15  7.178e-15 -3.850e-14 -7.929e-15 
        13         14         15         16         17         18 
 5.925e-14 -1.218e-14 -5.978e-16  1.083e-14 -1.884e-15  1.347e-13 
        19         20         21         22         23         24 
-4.215e-13 -6.712e-15  2.457e-16 -1.004e-14  3.711e-14  4.334e-14 
        25         26         27         28         29         30 
 5.942e-14  1.569e-14 -2.419e-14  1.587e-14 -4.389e-14 -1.045e-14 
        31         32         33         34         35         36 
 1.733e-16 -1.649e-16  2.330e-14 -6.043e-15 -2.574e-15  4.755e-15 
        37         38         39         40         41         42 
 6.612e-15 -6.411e-14 -1.354e-14  3.064e-14  1.798e-16  1.460e-14 
        43         44         45         46         47         48 
-4.235e-14 -3.183e-15  4.972e-14 -1.129e-14  2.139e-14 -9.345e-14 
        49         50         51         52         53         54 
-1.160e-16  2.469e-16 -1.074e-14 -1.610e-14  9.085e-15  1.085e-15 
        55         56         57         58         59         60 
-3.346e-15 -1.597e-14  3.316e-14  1.044e-14  5.625e+00 -4.333e+00 
        61         62         63         64         65         66 
-5.625e+00 -3.333e-01  4.667e+00  7.965e-15 -1.786e-14 -1.774e-14 
        67         68         69         70         71         72 
 3.690e-13 -5.787e-14 -1.958e-16 -3.632e-14  3.233e-16  7.426e-16 
        73         74         75         76 
 4.438e-15  2.160e-14  2.012e-14 -1.086e-16 

Coefficients: (33 not defined because of singularities)
                Estimate Std. Error t value Pr(>|t|)   
(Intercept)     30674.55   69920.30   0.439  0.69055   
size           -42013.34   29343.43  -1.432  0.24761   
lot             -3865.14    4652.54  -0.831  0.46705   
bath            16682.49   36127.34   0.462  0.67567   
bed             -9210.35   20147.24  -0.457  0.67863   
bathbed          -507.76    9807.71  -0.052  0.96197   
age             -9862.19    4350.87  -2.267  0.10825   
agesq            1955.93     830.69   2.355  0.09989 . 
garage          16761.83    6086.66   2.754  0.07051 . 
active          -6524.18    1680.91  -3.881  0.03030 * 
edison          20869.93    7045.49   2.962  0.05944 . 
harris         154178.77   51793.94   2.977  0.05875 . 
adams           11430.87    4831.29   2.366  0.09886 . 
crest            3630.37    6095.82   0.596  0.59341   
parker          -1985.29    5832.97  -0.340  0.75603   
size:lot         1226.20     293.55   4.177  0.02499 * 
size:bath        6197.07   14029.61   0.442  0.68860   
size:bed        10011.33    8515.16   1.176  0.32452   
size:bathbed    -2068.51    4061.57  -0.509  0.64565   
size:age         -594.70     146.81  -4.051  0.02709 * 
size:agesq        184.12      27.62   6.666  0.00688 **
size:garage      2840.82     692.95   4.100  0.02626 * 
size:active      3127.10     785.70   3.980  0.02838 * 
size:edison     -1987.39     410.57  -4.841  0.01682 * 
size:harris     -6804.34    1895.59  -3.590  0.03703 * 
size:adams      -5863.16    2515.39  -2.331  0.10206   
size:crest       1426.01    2165.55   0.658  0.55724   
size:parker     -5981.39    1683.58  -3.553  0.03802 * 
lot:bath         -384.98    2310.86  -0.167  0.87828   
lot:bed           378.31    1433.43   0.264  0.80893   
lot:bathbed        18.01     707.46   0.025  0.98129   
lot:age            36.92      33.22   1.111  0.34746   
lot:agesq          17.87      24.12   0.741  0.51243   
```

lot:garage	374.14	112.38	3.329	0.04474	*
lot:active	-157.72	57.10	-2.762	0.07002	.
lot:edison	910.82	265.18	3.435	0.04139	*
lot:harris	508.00	137.70	3.689	0.03454	*
lot:adams	281.50	126.03	2.234	0.11162	
lot:crest	2013.32	686.57	2.932	0.06088	.
lot:parker	-143.24	131.72	-1.087	0.35640	
bath:bed	NA	NA	NA	NA	
bath:bathbed	-457.60	201.29	-2.273	0.10759	
bath:age	5107.28	2127.53	2.401	0.09583	.
bath:agesq	-1654.59	570.97	-2.898	0.06261	.
bath:garage	-11795.23	3622.55	-3.256	0.04727	*
bath:active	-104.81	612.53	-0.171	0.87502	
bath:edison	-5932.15	2849.01	-2.082	0.12873	
bath:harris	-70086.72	23770.60	-2.948	0.06010	.
bath:adams	NA	NA	NA	NA	
bath:crest	-6642.94	1467.22	-4.528	0.02016	*
bath:parker	7948.23	4124.82	1.927	0.14962	
bed:bathbed	-97.59	60.03	-1.626	0.20248	
bed:age	2299.10	1091.01	2.107	0.12569	
bed:agesq	-680.43	257.98	-2.638	0.07782	.
bed:garage	-5979.90	1878.84	-3.183	0.04999	*
bed:active	273.35	253.91	1.077	0.36055	
bed:edison	-6674.89	2498.34	-2.672	0.07558	.
bed:harris	-46978.71	16484.12	-2.850	0.06511	.
bed:adams	NA	NA	NA	NA	
bed:crest	-887.93	594.57	-1.493	0.23217	
bed:parker	4753.38	2466.37	1.927	0.14956	
bathbed:age	-1263.95	562.50	-2.247	0.11024	
bathbed:agesq	465.56	166.24	2.800	0.06783	.
bathbed:garage	3078.07	938.80	3.279	0.04647	*
bathbed:active	48.62	143.68	0.338	0.75739	
bathbed:edison	1788.62	939.69	1.903	0.15312	
bathbed:harris	23010.28	7943.94	2.897	0.06268	.
bathbed:adams	NA	NA	NA	NA	
bathbed:crest	187.81	370.05	0.508	0.64674	
bathbed:parker	-2651.38	1312.46	-2.020	0.13664	
age:agesq	73.89	22.30	3.313	0.04530	*
age:garage	430.62	114.42	3.764	0.03281	*
age:active	-261.29	68.35	-3.823	0.03151	*
age:edison	-332.89	95.12	-3.500	0.03949	*
age:harris	40.40	46.99	0.860	0.45319	
age:adams	NA	NA	NA	NA	
age:crest	NA	NA	NA	NA	
age:parker	51.72	54.63	0.947	0.41360	
agesq:garage	89.36	35.23	2.537	0.08493	.
agesq:active	NA	NA	NA	NA	
agesq:edison	NA	NA	NA	NA	
agesq:harris	NA	NA	NA	NA	
agesq:adams	NA	NA	NA	NA	
agesq:crest	NA	NA	NA	NA	
agesq:parker	NA	NA	NA	NA	
garage:active	NA	NA	NA	NA	
garage:edison	NA	NA	NA	NA	
garage:harris	NA	NA	NA	NA	
garage:adams	NA	NA	NA	NA	
garage:crest	NA	NA	NA	NA	
garage:parker	NA	NA	NA	NA	
active:edison	NA	NA	NA	NA	
active:harris	NA	NA	NA	NA	
active:adams	NA	NA	NA	NA	
active:crest	NA	NA	NA	NA	
active:parker	NA	NA	NA	NA	
edison:harris	NA	NA	NA	NA	
edison:adams	NA	NA	NA	NA	
edison:crest	NA	NA	NA	NA	
edison:parker	NA	NA	NA	NA	
harris:adams	NA	NA	NA	NA	
harris:crest	NA	NA	NA	NA	
harris:parker	NA	NA	NA	NA	
adams:crest	NA	NA	NA	NA	
adams:parker	NA	NA	NA	NA	

```
crest:parker           NA          NA       NA        NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.886 on 3 degrees of freedom
Multiple R-squared:  0.9996,      Adjusted R-squared:  0.9905
F-statistic: 109.4 on 72 and 3 DF,  p-value: 0.00121
```

Stepwise Regression

```
> homes.model3 <- step (homes.model1)
Start:  AIC=582.75
price ~ size + lot + bath + bed + bathbed + age + agesq + garage +
    active + edison + harris + adams + crest + parker

          Df Sum of Sq    RSS    AIC
- crest    1     302.9 109809 580.96
- parker   1    1349.0 110855 581.68
- adams    1    1881.6 111387 582.04
- age      1    2444.0 111950 582.43
- garage   1    2702.0 112208 582.60
<none>                 109506 582.75
- size     1    7243.5 116749 585.62
- bath     1    7611.5 117117 585.86
- bathbed  1    8464.2 117970 586.41
- bed      1   10313.9 119820 587.59
- agesq    1   10415.1 119921 587.65
- harris   1   10842.2 120348 587.92
- active   1   10991.2 120497 588.02
- lot      1   15097.1 124603 590.56
- edison   1   18987.2 128493 592.90

Step:  AIC=580.96
price ~ size + lot + bath + bed + bathbed + age + agesq + garage +
    active + edison + harris + adams + parker

          Df Sum of Sq    RSS    AIC
- parker   1    1088.5 110897 579.71
- adams    1    1679.8 111489 580.11
- age      1    2387.2 112196 580.59
<none>                 109809 580.96
- garage   1    3273.2 113082 581.19
- size     1    6948.1 116757 583.62
- bath     1    7406.3 117215 583.92
- bathbed  1    8289.3 118098 584.49
- bed      1   10151.8 119961 585.68
- agesq    1   10237.6 120046 585.73
- active   1   11394.5 121203 586.46
- harris   1   12196.6 122005 586.96
- lot      1   14794.5 124603 588.56
- edison   1   20834.3 130643 592.16

Step:  AIC=579.71
price ~ size + lot + bath + bed + bathbed + age + agesq + garage +
    active + edison + harris + adams

          Df Sum of Sq    RSS    AIC
- adams    1    1498.2 112395 578.73
- age      1    2011.5 112909 579.07
<none>                 110897 579.71
- garage   1    3666.5 114564 580.18
- size     1    6304.2 117202 581.91
- agesq    1    9400.7 120298 583.89
- active   1   10589.0 121486 584.64
- bath     1   10689.6 121587 584.70
- bathbed  1   12934.4 123832 586.09
- lot      1   13850.1 124747 586.65
```

```
- bed      1   15696.2 126593 587.77
- harris   1   16508.4 127406 588.25
- edison   1   24361.9 135259 592.80

Step:  AIC=578.73
price ~ size + lot + bath + bed + bathbed + age + agesq + garage + 
    active + edison + harris

          Df Sum of Sq    RSS    AIC
- age      1    1905.4 114301 578.01
<none>                 112395 578.73
- garage   1    4403.5 116799 579.65
- size     1    7219.8 119615 581.46
- agesq    1    8799.0 121194 582.46
- bath     1    9294.7 121690 582.77
- active   1   10938.1 123334 583.79
- bathbed  1   11495.9 123891 584.13
- bed      1   14198.6 126594 585.77
- lot      1   14612.8 127008 586.02
- harris   1   17810.5 130206 587.91
- edison   1   27910.3 140306 593.58

Step:  AIC=578.01
price ~ size + lot + bath + bed + bathbed + agesq + garage + 
    active + edison + harris

          Df Sum of Sq    RSS    AIC
<none>                 114301 578.01
- agesq    1    7356.1 121657 580.75
- garage   1    7733.8 122035 580.98
- size     1    8053.2 122354 581.18
- bath     1    8434.6 122735 581.42
- bathbed  1   10910.0 125211 582.93
- active   1   11206.1 125507 583.11
- lot      1   13250.4 127551 584.34
- bed      1   14473.9 128775 585.07
- harris   1   20518.9 134820 588.55
- edison   1   26031.6 140332 591.60
```