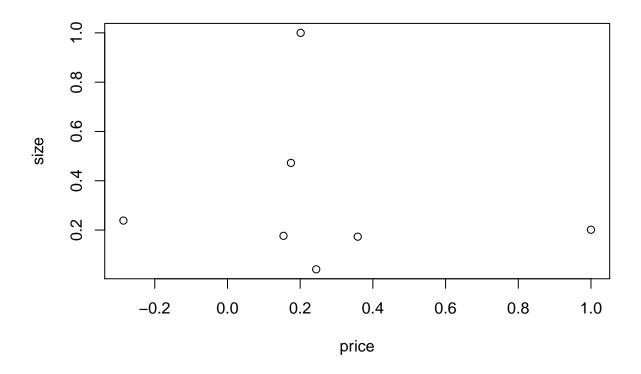
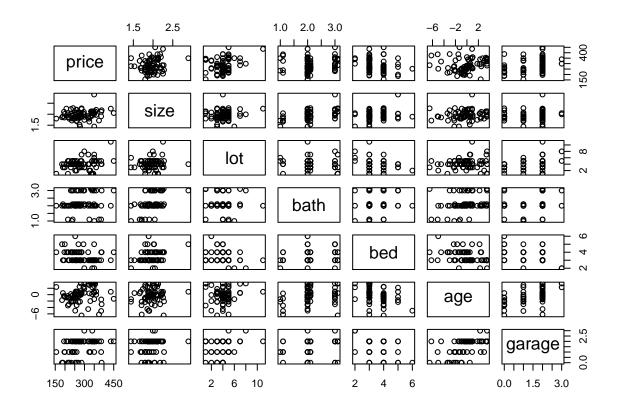
## Home Price Model

```
#This section of the code focused on looking at correlations between predictors, doing different regres
rm(list = ls())
library(readxl)
library(olsrr)
## Warning: package 'olsrr' was built under R version 4.0.5
## Registered S3 methods overwritten by 'tibble':
##
     method
                from
##
     format.tbl pillar
##
     print.tbl pillar
##
## Attaching package: 'olsrr'
## The following object is masked from 'package:datasets':
##
##
homeprice <- read excel("C:/Users/Daniel Huang/Desktop/Junior/IE Project/homeprice (1).xls")
homeprice2 <- subset(homeprice, select = -c(id, elem, status, year, agesq, bathbed))
View(homeprice)
#Correlation between predictors
numeric_data <- subset(homeprice, select = c(price, size, lot, bath, bed, age, garage))
cor(numeric_data)
##
                           size
                                         lot
                                                   bath
                                                               bed
## price
           1.0000000 0.20143783 0.24423228 0.1746578 -0.2861975
                                                                    0.15412476
           0.2014378 1.00000000 0.04079199 0.4725406 0.2384530
## size
                                                                    0.17656934
           0.2442323\ 0.04079199\ 1.00000000\ -0.1709961\ -0.2138298\ -0.03933975
## lot
## bath
           0.1746578 0.47254061 -0.17099609 1.0000000 0.1468258 0.33452385
## bed
          -0.2861975 0.23845303 -0.21382977 0.1468258 1.0000000 -0.36315062
           0.1541248 \ 0.17656934 \ -0.03933975 \quad 0.3345239 \ -0.3631506 \quad 1.00000000
## age
## garage 0.3583861 0.17315776 0.23581664 0.2017635 -0.4038039 0.53718464
              garage
## price
           0.3583861
## size
           0.1731578
## lot
           0.2358166
## bath
           0.2017635
## bed
          -0.4038039
           0.5371846
## age
## garage 1.0000000
```

plot(cor(numeric\_data))



pairs(numeric\_data)



#Forward Stepwise Regression with R^2
model <- lm(price ~ ., data = homeprice2)
summary(ols\_step\_both\_p(model))</pre>

```
##
              Length Class
                                 Mode
## orders
                      -none-
                                 character
## method
               6
                                 character
                      -none-
## steps
               1
                      -none-
                                 numeric
## predictors
               6
                      -none-
                                 character
## rsquare
               6
                      -none-
                                 numeric
## aic
               6
                      -none-
                                 numeric
## sbc
               6
                      -none-
                                 numeric
## sbic
               6
                      -none-
                                 numeric
## adjr
               6
                      -none-
                                 numeric
## rmse
               6
                      -none-
                                 numeric
## mallows_cp
               6
                      -none-
                                 numeric
## indvar
              12
                      -none-
                                 character
## betas
              27
                      -none-
                                 numeric
## lbetas
               6
                      -none-
                                 numeric
## pvalues
              27
                      -none-
                                 numeric
## beta pval
               4
                      data.frame list
## model
              12
                      lm
                                 list
```

```
#Backward Regression
ols_step_backward_p(model, details = TRUE)
```

## Backward Elimination Method

```
##
## Candidate Terms:
##
## 1 . size
## 2 . lot
## 3 . bath
## 4 . bed
## 5 . age
## 6 . garage
## 7 . active
## 8 . edison
## 9 . harris
## 10 . adams
## 11 . crest
## 12 . parker
##
## We are eliminating variables based on p value...
##
## - crest
##
## Backward Elimination: Step 1
## Variable crest Removed
##
                     Model Summary
## -----
                    0.728 RMSE
0.530 Coef. Var
0.450 MSE
## R
                                               44.757
## R-Squared
                                               15.661
## Adj. R-Squared
                                              2003.210
               0.273
## Pred R-Squared
                               MAE
                                               33.325
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                           ANOVA
## -----
##
                Sum of
                          DF Mean Square F Sig.
##
              Squares
## -----
## Regression 144797.033 11
## Residual 128205.438 64
                                 13163.367 6.571 0.0000
                                  2003.210
## Total
             273002.471
                          75
##
                              Parameter Estimates
## -----
                                             t Sig
      model Beta Std. Error Std. Beta
                                                            lower
                                                                    upper
## -----
## (Intercept) 96.323 57.615 1.672 0.099 -18.776 211.422

## size 69.119 30.164 0.243 2.291 0.025 8.858 129.379

## lot 10.687 3.598 0.293 2.970 0.004 3.499 17.874

## bath 4.739 11.608 0.045 0.408 0.684 -18.450 27.929
```

##	bed	-12.488	9.046	-0.153	-1.380	0.172	-30.561	5.584		
##	age	1.619								
##	garage	10.292								
##		30.264					6.909			
##		80.121								
##		46.740		0.302						
##		-6.938			-0.244		-63.810			
##	parker			-0.124						
##										
##										
##										
##	- adams									
##										
##	Backward Elim	ination: Ste	p 2							
##										
##	Variable adam	ns Removed								
##										
##			Model Summar	У						
##										
##	R	0	.728 F	MSE	44.4	132				
	R-Squared			Coef. Var		547				
	Adj. R-Squared				1974.2	222				
	Pred R-Squared				33.4	106				
	RMSE: Root Me MSE: Mean Squ MAE: Mean Abs	are Error								
##		Sum of								
		Sum of Squares	 DF	Mean Square	F	Sig.				
## ##		Squares	DF	Mean Square	F	Sig.				
## ## ##		Squares		Mean Square 14467.807						
## ## ## ##		Squares  144678.066	10							
## ## ## ## ##	Regression Residual Total	Squares 144678.066 128324.405 273002.471	10 65 75	14467.807 1974.222						
## ## ## ## ##	Regression Residual	Squares 144678.066 128324.405 273002.471	10 65 75	14467.807 1974.222						
## ## ## ## ## ##	Regression Residual Total	Squares 144678.066 128324.405 273002.471	10 65 75	14467.807 1974.222	7.328					
## ## ## ## ## ## ##	Regression Residual Total	Squares 144678.066 128324.405 273002.471	10 65 75	14467.807 1974.222	7.328					
## ## ## ## ## ## ##	Regression Residual Total	Squares 144678.066 128324.405 273002.471	10 65 75	14467.807 1974.222	7.328 nates	0.0000				
## ## ## ## ## ## ##	Regression Residual Total	Squares 144678.066 128324.405 273002.471	10 65 75	14467.807 1974.222	7.328		lower	 upper		
######################################	Regression Residual Total model	Squares  144678.066 128324.405 273002.471  Beta	10 65 75 F Std. Error	14467.807 1974.222	7.328 nates	0.0000 Sig				
## ## ## ## ## ## ##	Regression Residual Total model(Intercept)	Squares 144678.066 128324.405 273002.471	10 65 75	14467.807 1974.222	7.328 nates	0.0000	lower -17.040 10.179	upper  201.504 129.336		
######################################	Regression Residual Total model(Intercept)	Squares  144678.066 128324.405 273002.471  Beta  92.232	10 65 75 F Std. Error	14467.807 1974.222 Parameter Estin	7.328  nates  t  1.686	0.0000 Sig	-17.040	201.504		
## ## ## ## ## ## ## ##	Regression Residual Total model(Intercept) size	Squares  144678.066 128324.405 273002.471  Beta  92.232 69.758	10 65 75 F Std. Error 54.714 29.832	14467.807 1974.222 Parameter Estin Std. Beta	7.328  nates  t  1.686 2.338	0.0000 Sig 0.097	-17.040 10.179	201.504 129.336		
## ## ## ## ## ## ## ## ##	Regression Residual Total model (Intercept) size lot	Squares  144678.066 128324.405 273002.471  Beta  92.232 69.758 10.709	10 65 75 F Std. Error 54.714 29.832 3.571	14467.807 1974.222 Parameter Estin Std. Beta 0.246 0.293	7.328  nates  t  1.686 2.338 2.999	0.0000 Sig 0.097 0.022 0.004	-17.040 10.179 3.578	201.504 129.336 17.840		
## ## ## ## ## ## ## ## ##	Regression Residual Total model (Intercept) size lot bath	Squares  144678.066 128324.405 273002.471  Beta  92.232 69.758 10.709 4.993	10 65 75 Std. Error 54.714 29.832 3.571 11.477	14467.807 1974.222 Parameter Estin Std. Beta 0.246 0.293 0.047	7.328  nates  t  1.686 2.338 2.999 0.435	0.0000 Sig 0.097 0.022 0.004 0.665	-17.040 10.179 3.578 -17.928	201.504 129.336 17.840 27.915		
## ## ## ## ## ## ## ## ##	Regression Residual Total model (Intercept) size lot bath bed	Squares  144678.066 128324.405 273002.471  Beta  92.232 69.758 10.709 4.993 -12.189	10 65 75 Std. Error 54.714 29.832 3.571 11.477 8.898	14467.807 1974.222 Parameter Estin Std. Beta 0.246 0.293 0.047 -0.149	7.328  nates  t  1.686 2.338 2.999 0.435 -1.370	0.0000 Sig 0.097 0.022 0.004 0.665 0.175	-17.040 10.179 3.578 -17.928 -29.959	201.504 129.336 17.840 27.915 5.581		
######################################	Regression Residual Total model (Intercept) size lot bath bed age	Squares  144678.066 128324.405 273002.471  Beta  92.232 69.758 10.709 4.993 -12.189 1.609	10 65 75 Std. Error 54.714 29.832 3.571 11.477 8.898 3.274	14467.807 1974.222 Parameter Estin Std. Beta 0.246 0.293 0.047 -0.149 0.063	7.328  nates  t  1.686 2.338 2.999 0.435 -1.370 0.492	0.0000 Sig 0.097 0.022 0.004 0.665 0.175 0.625	-17.040 10.179 3.578 -17.928 -29.959 -4.929	201.504 129.336 17.840 27.915 5.581 8.147		
######################################	Regression Residual Total  model  (Intercept) size lot bath bed age garage	Squares  144678.066 128324.405 273002.471  Beta  92.232 69.758 10.709 4.993 -12.189 1.609 10.636	10 65 75 F Std. Error 54.714 29.832 3.571 11.477 8.898 3.274 8.781	14467.807 1974.222 Parameter Estin Std. Beta 0.246 0.293 0.047 -0.149 0.063 0.136	7.328  t  1.686 2.338 2.999 0.435 -1.370 0.492 1.211	0.0000 Sig 0.097 0.022 0.004 0.665 0.175 0.625 0.230	-17.040 10.179 3.578 -17.928 -29.959 -4.929 -6.902	201.504 129.336 17.840 27.915 5.581 8.147 28.173		
######################################	Regression Residual Total  model  (Intercept) size lot bath bed age garage active	Squares  144678.066 128324.405 273002.471  Beta  92.232 69.758 10.709 4.993 -12.189 1.609 10.636 30.267	10 65 75 Std. Error 54.714 29.832 3.571 11.477 8.898 3.274 8.781 11.606	14467.807 1974.222 Parameter Estin Std. Beta 0.246 0.293 0.047 -0.149 0.063 0.136 0.237	7.328  t  1.686 2.338 2.999 0.435 -1.370 0.492 1.211 2.608	0.0000 Sig 0.097 0.022 0.004 0.665 0.175 0.625 0.230 0.011	-17.040 10.179 3.578 -17.928 -29.959 -4.929 -6.902 7.088	201.504 129.336 17.840 27.915 5.581 8.147 28.173 53.445		
######################################	Regression Residual Total model (Intercept) size lot bath bed age garage active edison	Squares  144678.066 128324.405 273002.471  Beta  92.232 69.758 10.709 4.993 -12.189 1.609 10.636 30.267 80.856	10 65 75 F Std. Error 54.714 29.832 3.571 11.477 8.898 3.274 8.781 11.606 16.774	14467.807 1974.222 Parameter Estin Std. Beta 0.246 0.293 0.047 -0.149 0.063 0.136 0.237 0.492	7.328  t  1.686 2.338 2.999 0.435 -1.370 0.492 1.211 2.608 4.820	0.0000 Sig 0.097 0.022 0.004 0.665 0.175 0.625 0.230 0.011 0.000	-17.040 10.179 3.578 -17.928 -29.959 -4.929 -6.902 7.088 47.357	201.504 129.336 17.840 27.915 5.581 8.147 28.173 53.445 114.356		

```
##
##
## - bath
##
## Backward Elimination: Step 3
## Variable bath Removed
##
                       Model Summary
##
## R
                      0.727
                                RMSE
                                                 44.158
## R-Squared
                      0.529
                                Coef. Var
                                                 15.451
## Adj. R-Squared
                      0.464
                                MSE
                                               1949.971
## Pred R-Squared
                                                 33.234
                      0.354
                                MAE
   RMSE: Root Mean Square Error
## MSE: Mean Square Error
  MAE: Mean Absolute Error
##
##
                             ANOVA
##
  ______
                 Sum of
##
                         DF Mean Square F Sig.
                Squares
                                 -----
                          9
                                  16033.819 8.223 0.0000
## Regression 144304.369
## Residual 128698.102
                           66
                                   1949.971
## Total
             273002.471
                            75
##
                                Parameter Estimates
       model
                 Beta
                        Std. Error
                                   Std. Beta
                                               t
                                                       Sig
                                                                lower
                                                                         upper
## (Intercept)
               90.505
                                               1.669
                                                       0.100
                                                              -17.776
                                                                       198.786
                           54.234
                                     0.264 2.773 0.007
0.289 2.988 0.004
                                     0.264
                                                              21.005
                                                                      129.098
##
      size
              75.051
                           27.070
                                                                       17.584
##
              10.540
                           3.528
        lot
                                                               3.497
##
        bed -11.421
                           8.667
                                     -0.140 -1.318 0.192 -28.725
                                                                        5.883
##
               1.938
                           3.166
                                      0.075 0.612
                                                     0.542
                                                              -4.382
                                                                        8.259
         age
                                     0.139
              10.875
                                              1.249
                                                              -6.516
##
                           8.710
                                                      0.216
                                                                       28.265
      garage
##
      active 30.638
                                     0.240 2.663
                                                      0.010
                                                               7.671
                          11.503
                                                                       53.605
##
      edison 80.965
                          16.669
                                     0.493 4.857
                                                      0.000
                                                              47.685 114.245
##
              47.939
                          15.305
                                      0.310
                                              3.132
                                                       0.003
                                                              17.382
                                                                       78.495
      harris
                                                                        8.118
      parker
              -19.489
                           13.828
                                     -0.129
                                              -1.409
                                                       0.163
                                                              -47.097
##
##
##
## - age
##
## Backward Elimination: Step 4
## Variable age Removed
##
##
                       Model Summary
```

```
0.725 RMSE
0.526 Coef. Var
0.469 MSE
## R
                                                    43.952
                       0.526
## R-Squared
                                                    15.379
## Adj. R-Squared
                      0.469
                                                   1931.781
## Pred R-Squared
                                MAE
                      0.377
                                                     33.358
## -----
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                               ANOVA
##
                  Sum of
                 Squares
                             DF Mean Square
                                                  F
                                                            Sig.
## -----
## Regression 143573.123
                            8 17946.640
                                                   9.29 0.0000
## Residual 129429.348
                              67
                                      1931.781
## Total
              273002.471
                              75
##
##
                                  Parameter Estimates
                Beta Std. Error
                                     Std. Beta
                                                  1.634 0.107 -19.471
## (Intercept)
                87.956
                             53.821
                                                                             195.384
                                        0.274 2.931 0.005 24.843 130.857
##
   size 77.850
                            26.557
                                                                   3.244 16.988
##
        lot
               10.116
                            3.443
                                        0.277 2.938 0.005
                                       -0.156 -1.534 0.130 -29.425
##
        bed -12.788
                             8.335
                                                                             3.849
                         7.882
11.437
15.139
15.003
13.737

    0.168
    1.661
    0.101
    -2.638
    28.829

    0.243
    2.707
    0.009
    8.131
    53.789

    0.467
    5.072
    0.000
    46.572
    107.006

##
      garage 13.096
      garage 13.096
active 30.960
##
##
      edison 76.789
                                        0.321 3.303 0.002 19.613 79.507
               49.560
##
      harris
                                        -0.126 -1.381 0.172 -46.393
      parker -18.973
                             13.737
                                                                             8.447
##
##
##
##
## No more variables satisfy the condition of p value = 0.3
##
##
## Variables Removed:
## - crest
## - adams
## - bath
## - age
##
##
## Final Model Output
##
                        Model Summary
                        0.725 RMSE
0.526 Coef. Var
## R
                                                    43.952
                                              15.379
## R-Squared
```

```
## Adj. R-Squared
                      0.469 MSE
                                                1931.781
## Pred R-Squared
                      0.377
                                MAE
                                                 33.358
## -----
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                             ANOVA
##
                 Sum of
##
                Squares
                           DF Mean Square
                                               F
                                                      Sig.
## -----
                          8
67
                                17946.640
## Regression 143573.123
                                               9.29 0.0000
## Residual 129429.348
## Total 273002.471
                                   1931.781
                         75
## Total
              273002.471
##
##
                               Parameter Estimates
       model
                Beta
                      Std. Error
                                   Std. Beta
                                                t
                                                        Sig
                                                                lower
                                                                        upper
              87.956
## (Intercept)
                           53.821
                                               1.634
                                                       0.107
                                                              -19.471
                                                                        195.384
                                    0.274 2.931 0.005
0.277 2.938 0.005
              77.850
      size
                                                              24.843
##
                           26.557
                                                                       130.857
                          3.443
                                                                      16.988
        lot
              10.116
                                                               3.244
##
##
        bed -12.788
                           8.335
                                     -0.156 -1.534 0.130 -29.425
                                                                        3.849
##
      garage 13.096
                           7.882
                                      0.168 1.661 0.101 -2.638
                                                                       28.829

      0.243
      2.707
      0.009
      8.131
      53.789

      0.467
      5.072
      0.000
      46.572
      107.006

##
      active
              30.960
                          11.437
                         15.139
      edison 76.789
##
##
              49.560
                          15.003
                                      0.321 3.303 0.002 19.613 79.507
      harris
                                      -0.126 -1.381 0.172 -46.393
                                                                        8.447
      parker
             -18.973
                           13.737
##
##
##
##
                         Elimination Summary
##
         Variable
                              Adj.
## Step
       Removed R-Square R-Square C(p)
                                                AIC
                                                           RMSE
## -----
                 0.5304
                           0.4497 11.0134
0.4576 9.0719
##
    1
       crest
                                                806.4085
                                                          44.7572
   2 adams
##
                    0.530
                                                804.4790 44.4322
##
    3 bath
                                       7.2556
                   0.5286
                             0.4643
                                                802.7000 44.1585
                            0.4693
                                     5.6150
                                                        43.9520
     4 age
##
                    0.5259
                                                801.1306
#Forward Regression
model2 <- lm(price ~ ., data = homeprice2)</pre>
ols_step_forward_p(model2, details = TRUE)
## Forward Selection Method
##
## Candidate Terms:
```

##

## 1. size

```
## 2. lot
## 3. bath
## 4. bed
## 5. age
## 6. garage
## 7. active
## 8. edison
## 9. harris
## 10. adams
## 11. crest
## 12. parker
## We are selecting variables based on p value...
##
##
## Forward Selection: Step 1
##
## - garage
##
##
                    Model Summary
## -----
                   0.358 RMSE
0.128 Coef. Var
## R-Squared
                                            19.841
## Adj. R-Squared
                   0.117
                           MSE
                                           3215.377
## Pred R-Squared
                  0.082
                           MAE
                                            45.622
## -----
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                          ANOVA
##
               Sum of
##
             Squares
                        DF Mean Square
                                          F
                                                 Sig.
                     74
                         1 35064.597
## Regression 35064.597
                                        10.905 0.0015
## Residual 237937.874
                               3215.377
                      75
## Total
            273002.471
##
                           Parameter Estimates
## -----
                                          t
                                                  Sig
      model Beta Std. Error Std. Beta
                                                         lower
## (Intercept) 241.927
                       14.791
                                        16.356 0.000
                                                        212.455
                                                                271.399
                        8.484 0.358 3.302 0.001 11.112
  garage 28.017
##
                                                                44.922
##
##
##
## Forward Selection: Step 2
## - edison
##
```

:		Model Summa	•				
R R-Squared Adj. R-Squar Pred R-Squar	0 0 red 0 red 0	.514 .264 .244 .195	RMSE Coef. Var MSE MAE	52.4 18.3 2752.3 42.1	357 304 .80		
RMSE: Root MSE: Mean S	Mean Square E	rror	/A				
:	Sum of Squares	DF	Mean Square	F			
Regression Residual Total		2 73 75	36042.141 2752.304				
: : :			Parameter Estir				
model	Beta 	Std. Error	Std. Beta	t 	Sig	lower	uppe
garage edison	224.044 33.235 61.512	7.977 16.772	0.425 0.374	4.166	0.000	195.091 17.336 28.085	49.13
: :	ection: Step 3						
:							
R R-Squared Adj. R-Squar Pred R-Squar	0 red 0	.349 .322	RMSE Coef. Var MSE MAE	49.6 17.3 2469.6 37.5	889 887		
RMSE: Root MSE: Mean S MAE: Mean	Mean Square E	rror					
:		ANOV					
: :	Sum of						
:	Squares		Mean Square		Sig.		
Regression		3	31728.331 2469.687		0.0000		

	Parameter Estimates							
model	Beta	Std. Error		t	Sig	lower	uppe	
	217.842			15.661	0.000	 190.113	245.5	
garage	30.831	7.597	0.394	4.058	0.000	15.686	45.9	
	70.718			4.373	0.000	38.483	102.9	
harris	46.205 	15.108	0.299	3.058	0.003	16.088 	76.3	
Forward Col	ostion: Stop /	1						
	ection: Step 4	ŧ						
- lot								
		Model Summar						
	C		MSE	47.2				
				16 h	.')6'			
R-Squared		).420 (		16.5				
R-Squared Adj. R-Squa Pred R-Squa  RMSE: Root	red C red C  Mean Square E	).387 N	ISE IAE	2230.8 37.6	851			
Pred R-Square  RMSE: Root MSE: Mean	red ( red (	).387 M ).309 M 	ISE IAE	2230.8	851			
R-Squared Adj. R-Squa: Pred R-Squa: RMSE: Root MSE: Mean	red ( red (  Mean Square E Square Error	).387 M ).309 M 	ISE IAE 	2230.8	851			
R-Squared Adj. R-Squa: Pred R-Squa: RMSE: Root MSE: Mean	red ( red (  Mean Square E Square Error	0.387 N 0.309 N Error	ISE IAE 	2230.8	851			
R-Squared Adj. R-Squa: Pred R-Squa: RMSE: Root MSE: Mean	red ( red ( Mean Square F Square Error Absolute Error	0.387 M 0.309 M Error ANOVA	ISE IAE 	2230.8 37.6	851			
R-Squared Adj. R-Squared Pred R-Squared RMSE: Root MSE: Mean	red ( red ( Mean Square E Square Error Absolute Error Sum of	0.387 M 0.309 M Error ANOVA	Mean Square	2230.8 37.6	851 851 			
R-Squared Adj. R-Squared Pred R-Squared RMSE: Root MSE: Mean MAE: Mean	red Cored Co	0.387 M 0.309 M Error ANOVA	Mean Square	2230.8 37.6	Sig.			
R-Squared Adj. R-Squa: Pred R-Squa: RMSE: Root MSE: Mean MAE: Mean Regression Residual	red Cored Co	DF	Mean Square 28653.008	2230.8 37.6	Sig.			
R-Squared Adj. R-Squared Pred R-Squared RMSE: Root MSE: Mean MAE: Mean Regression Residual Total	red (0 red (0 mean Square F Square Error Absolute Error Sum of Squares	DF 4 71 75	Mean Square 28653.008 2230.851	2230.8 37.6  F 	Sig.			
R-Squared Adj. R-Squa: Pred R-Squa: RMSE: Root MSE: Mean MAE: Mean  Regression Residual Total model	red (ored (o	DF  4 71 75  Std. Error	Mean Square 28653.008 2230.851 Parameter Estim	2230.8 37.6 	Sig.		 upp	
R-Squared Adj. R-Squared Pred R-Squared RMSE: Root MSE: Mean MAE: Mean Regression Residual Total model	red (content of the content of the c	DF  4 71 75  Std. Error	Mean Square 28653.008 2230.851  Parameter Estim	F 12.844	Sig.	lower		
R-Squared Adj. R-Squared Pred R-Squared RMSE: Root MSE: Mean MAE: Mean  Regression Residual Total  model  (Intercept) garage	red (0 red (0 red (0 Mean Square E Square Error Absolute Error Sum of Squares 114612.033 158390.438 273002.471  Beta  182.243 24.594	DF  ANOVA  71  75  Std. Error  17.897  7.524	Mean Square 28653.008 2230.851  Parameter Estim Std. Beta	2230.8 37.6 F 12.844 10.183 3.269	Sig Sig 0.0000 Sig	lower 146.557 9.592	217.9 39.5	
R-Squared Adj. R-Squared Pred R-Squared RMSE: Root MSE: Mean MAE: Mean  Regression Residual Total  model  (Intercept) garage	red (0 red (0 red (0 Mean Square E Square Error Absolute Error Sum of Squares 114612.033 158390.438 273002.471 Beta	DF  ANOVA  71  75  Std. Error  17.897  7.524	Mean Square 28653.008 2230.851  Parameter Estim	2230.8 37.6 F 12.844 10.183 3.269	Sig Sig 0.0000 Sig	lower 146.557 9.592	217.9 39.5	
R-Squared Adj. R-Squared Pred R-Squared RMSE: Root MSE: Mean MAE: Mean  Regression Residual Total  model  (Intercept) garage edison	red (0 red (0 Mean Square F Square Error Absolute Error Sum of Squares 114612.033 158390.438 273002.471 Beta 182.243 24.594 72.519	DF  ANOVA  71  75  Std. Error  17.897  7.524  15.381	Mean Square 28653.008 2230.851  Parameter Estim Std. Beta	2230.8 37.6 F 12.844 10.183 3.269 4.715 4.014	Sig Sig 0.0000 Sig 0.000 0.002 0.000 0.000	lower 146.557 9.592 41.850	217.9 39.5 103.1 91.4	

11

## Forward Selection: Step 5

```
##
## - active
##
##
                          Model Summary
                        0.673 RMSE
0.453 Coef. Var
0.414 MSE
0.320 MAE
## R
                                                         46.178
## R-Squared
                                                          16.158
## Adj. R-Squared
                                                         2132.372
## Pred R-Squared
                                                         36.059
  RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                                  ANOVA
##
                    Sum of
##
                  Squares
                                DF Mean Square
                                                       F
                                     24747.280
                               5
70
                                                     11.606 0.0000
## Regression 123736.401
## Residual 149266.070
              273002.471
##
                                    Parameter Estimates
        model
                 Beta Std. Error
                                          Std. Beta
                                                                 Sig
                                                                           lower
                                                       9.994 0.000 141.531 212.106
## (Intercept) 176.818
                              17.693

      0.279
      2.921
      0.005
      6.929
      36.747

      0.477
      5.124
      0.000
      47.909
      108.968

       garage 21.838
##
                                7.475
       edison 78.438
                             15.307
14.883
##
                                            0.390 4.050 0.000 30.594
##
       harris 60.278
                                                                                  89.962
##
       lot 10.855
                                3.520
                                            0.297 3.083 0.003
                                                                         3.833
                                                                                   17.876
      active 24.376
                                             0.191 2.069 0.042
                                                                          0.874
                                                                                  47.879
                               11.784
##
##
##
## Forward Selection: Step 6
##
## - size
##
                          Model Summary
##
                        0.703 RMSE
0.494 Coef.
0.450 MSE
0.360 MAE
                                                          44.735
                                     Coef. Var
## R-Squared
                                                          15.653
## Adj. R-Squared
                                                         2001.184
## Pred R-Squared
                                                          34.954
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
```

ANOVA

##

##								=	
##		Sum of							
##		Squares	D	F	Mean Square	F	Sig.		
								-	
##	Regression Residual	134920.786	6	6	22486.798	11.237	0.0000		
##	Total	273002.471	7	9 5	2001.104				
								=	
##									
##				Par	ameter Estim	ates			
##									
##					Std. Beta		_	lower	upper
##	(T-++)					1 012	0.000	40 440	164 016
##	(Intercept)		51.3	69 61	0 220	1.213	0.229	2 002	104.810
##	garage edison	82 059	14 9	0 <del>4</del> 08	0.239	5 504	0.013	5.903	111 799
##	harris	59.962	14.4	19	0.388	4.159	0.000	31.197	88.727
##	lot	10.877	3.4	10	0.298	3.189	0.002	4.073	17.680
##	harris lot active	29.305	11.6	05	0.230	2.525	0.014	6.154	52.455
##	size	59.487	25.1	63	0.239 0.499 0.388 0.298 0.230 0.209	2.364	0.021	9.288	109.685
##									
##									
##									
##	Forward Select	ion: Stop 7	7						
##	rolward belect	cion. Step i							
	- bed								
##									
##			Model Su	mmary	,				
##									
##			0.716			44.			
##	R-Squared	(	0.512	Co	ef. Var	15.			
##	Adj. R-Squared Pred R-Squared	i (	).462	MS	SE E	1957.			
##		1 (	.368	IM.	1E 	33.	751		
	RMSE: Root Me								
	MSE: Mean Squ	_							
##	MAE: Mean Abs		:						
##									
##				NOVA					
								-	
##		Sum of	Di	-	M	г	Q.:		
##					Mean Square			_	
	Regression								
	Residual					_0.200	3.000		
		273002.471							
##								-	
##									
##					rameter Esti	mates			
##		Po+5						1	,
## ##	model						_	lower	upper
##	(Intercept)							-16.006	

##	garage	13.712	7.922	0.175	1.731	0.088	-2.096	29.520
##	edison	82.074	14.745	0.499	5.566	0.000	52.651	111.496
##	harris	54.513	14.665	0.353	3.717	0.000	25.248	83.777
##	lot	9.703	3.453	0.266	2.810	0.006	2.813	16.592
##	active	29.783	11.481	0.233	2.594	0.012	6.872	52.694
##	size	74.562	26.626	0.263	2.800	0.007	21.432	127.693
##	bed	-13.350	8.381	-0.163	-1.593	0.116	-30.074	3.373
##								
##								
##								
##								
##	Forward Selec	ction: Step 8	3					
##								
##	- parker							
##								
##			Model Summar					
##				RMSE	43.9			
	R-Squared			Coef. Var	15.3			
	Adj. R-Square		0.469		1931.7			
##	# Pred R-Squared 0.377 MAE 33.358 #				358			
	RMSE: Root N	_	Error					
## ##	MSE: Mean So	-	<b></b>					
##	MAE: Mean At	osolute Ello.	Ľ					
##			ANOV	٨				
##			ANOV					
##		Sum of						
##		Squares		Mean Square	F	Sig		
		-						
	Regression			17946.640		0.0000		
##	Residual	129429.348		1931.781	0.20			
	Total							
##								
##			]	Parameter Estim	ates			
##								
##	model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
##								
##	(Intercept)	87.956	53.821		1.634	0.107	-19.471	195.384
##	garage	13.096	7.882	0.168	1.661	0.101	-2.638	28.829
##	edison	76.789	15.139	0.467	5.072	0.000	46.572	107.006
##	harris	49.560	15.003	0.321	3.303	0.002	19.613	79.507
##	lot	10.116	3.443	0.277	2.938	0.005	3.244	16.988

## ## ##

##

##

##

##

##

## No more variables to be added.

30.960

77.850

-12.788

-18.973

size

bed

active

parker

0.243

0.274

-0.156

-0.126

11.437

26.557

8.335

13.737

2.707

2.931

-1.534

-1.381

0.009

0.005

0.130

0.172

8.131

24.843

-29.425

-46.393

53.789

130.857

3.849

8.447

```
##
## Variables Entered:
##
## + garage
## + edison
## + harris
## + lot
## + active
## + size
## + bed
## + parker
##
##
## Final Model Output
##
##
                        Model Summary
## R
                       0.725
                                RMSE
                                                    43.952
                       0.526 Coef. Var
## R-Squared
                                                    15.379
## Adj. R-Squared
                       0.469
                                MSE
                                                  1931.781
                   0.377 MAE
## Pred R-Squared
## -----
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
                              ANOVA
## ---
                  Sum of
                          DF Mean Square F
                 Squares
                                                           Sig.
  ______
## Regression 143573.123 8 17946.640
                                                  9.29
                                                         0.0000
## Residual 129429.348
## Total 273002.471
                             67
                                     1931.781
                              75
##
##
                                 Parameter Estimates
##
                         Std. Error
                                      Std. Beta
                  Beta
                                                  t
                                                  1.634 0.107 -19.471
## (Intercept)
                87.956
                             53.821
                                                                           195.384
                           7.882 0.168 1.661 0.101 -2.638
15.139 0.467 5.072 0.000 46.572
15.003 0.321 3.303 0.002 19.613
3.443 0.277 2.938 0.005 3.244
11.437 0.243 2.707 0.009 8.131
       garage 13.096
##
                                                                            28.829
##
       edison 76.789
                                                                           107.006
                          15.003
3.443
11.437
##
       harris 49.560
                                                                           79.507
                                                                            16.988
      lot 10.116
##
##
       active
               30.960
                                                                            53.789
      size 77.850
bed -12.788
##
               77.850
                           26.557
                                        0.274 2.931 0.005
                                                                  24.843 130.857
##
                             8.335
                                        -0.156
                                                 -1.534 0.130 -29.425
                                                                            3.849
       parker -18.973
                                                 -1.381 0.172 -46.393
##
                            13.737
                                        -0.126
                                                                            8.447
```

Selection Summary

## ##

```
##
           Variable
                                       Adj.
## Step
           Entered R-Square R-Square
                                                 C(p)
                                                             AIC
                                                                           RMSE
  ______
        garage 0.1284 0.1167 44.9473 833.4051 56.7043 edison 0.2640 0.2439 28.7520 822.5526 52.4624
##
      1
##
      2

    0.2640
    0.2439
    28.7320
    822.3320
    32.4624

    0.3487
    0.3215
    19.3979
    815.2699
    49.6959

    0.4198
    0.3871
    11.8495
    808.4771
    47.2319

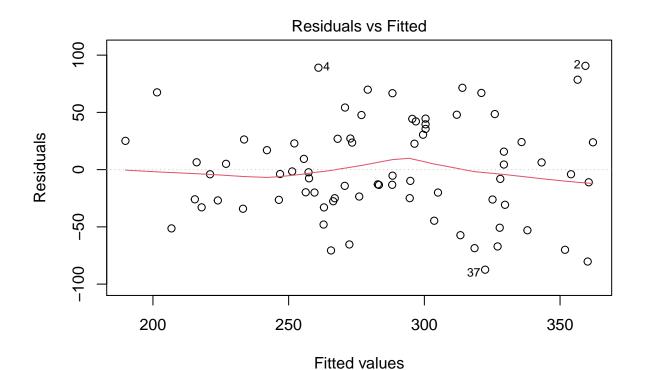
    0.4532
    0.4142
    9.3648
    805.9678
    46.1776

    0.4942
    0.4502
    5.8676
    802.0486
    44.7346

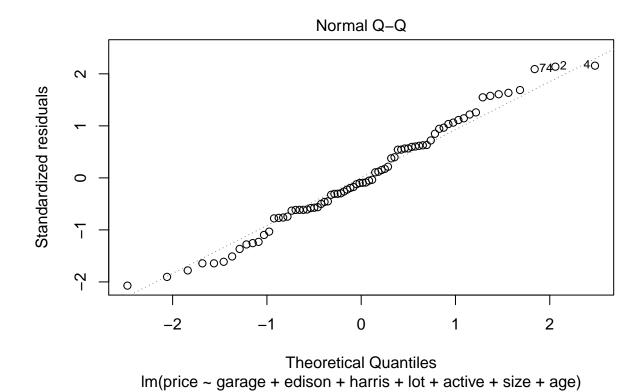
##
      3 harris
         lot
##
      4
##
      5
         active
##
      6 size
##
      7 bed
                         0.5124
                                     0.4622
                                                 5.4261
                                                             801.2641 44.2443
                                                                        43.9520
                          0.5259
                                      0.4693
##
      8
           parker
                                                   5.6150
                                                             801.1306
## ----
#Preliminary Model
model2 <- lm(price~garage+edison+harris+lot+active+size+age,data=homeprice2)</pre>
additive_mod <- lm(price~garage+edison+harris+lot+active+size+age+agesq,data=homeprice)
anova(additive_mod)
## Analysis of Variance Table
##
## Response: price
##
             Df Sum Sq Mean Sq F value
                                          Pr(>F)
           1 35065 35065 18.4348 5.801e-05 ***
## garage
## edison
             1 37020 37020 19.4627 3.813e-05 ***
## harris
             1 23101 23101 12.1449 0.0008728 ***
            1 19427
                         19427 10.2135 0.0021274 **
## lot
## active 1 9124 9124 4.7970 0.0319937 *
## size
            1 11184 11184 5.8801 0.0180160 *
## age
             1 1638 1638 0.8613 0.3567183
           1 9004
                           9004 4.7336 0.0331109 *
## agesq
## Residuals 67 127440 1902
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
#F-test with age and age 2
age_only <- lm(price~age,data=homeprice)</pre>
withagesq <- lm(price~age + agesq,data=homeprice)</pre>
anova(withagesq)
## Analysis of Variance Table
## Response: price
             Df Sum Sq Mean Sq F value
                         6485 2.0325 0.158231
              1
                 6485
                          33598 10.5299 0.001774 **
## agesq
             1 33598
## Residuals 73 232920
                           3191
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
Fcrit \leftarrow qf(0.95, 1, 73)
summary(model2)
##
## Call:
```

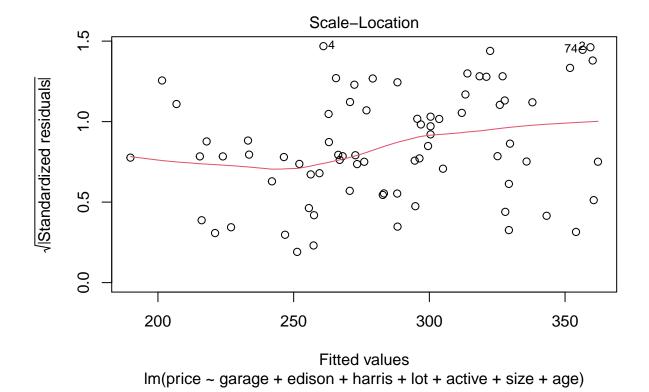
## lm(formula = price ~ garage + edison + harris + lot + active +

```
##
       size + age, data = homeprice2)
##
  Residuals:
##
##
       Min
                1Q
                                 3Q
                    Median
                                        Max
##
   -87.396 -26.532
                    -4.041
                             26.999
                                     90.647
##
## Coefficients:
               Estimate Std. Error t value Pr(>|t|)
##
## (Intercept)
                 70.507
                             52.227
                                      1.350 0.181485
                 14.765
                              8.550
                                      1.727 0.088740 .
  garage
## edison
                 88.302
                             16.449
                                      5.368 1.04e-06 ***
                             14.801
## harris
                 57.017
                                      3.852 0.000262 ***
                 11.302
                              3.447
                                      3.279 0.001646 **
## lot
                 28.863
                             11.630
                                      2.482 0.015552 *
## active
## size
                 57.520
                             25.290
                                      2.274 0.026102 *
## age
                  2.801
                              3.099
                                      0.904 0.369416
##
                   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Signif. codes:
##
## Residual standard error: 44.79 on 68 degrees of freedom
## Multiple R-squared: 0.5002, Adjusted R-squared: 0.4488
## F-statistic: 9.723 on 7 and 68 DF, p-value: 2.507e-08
plot(model2)
```

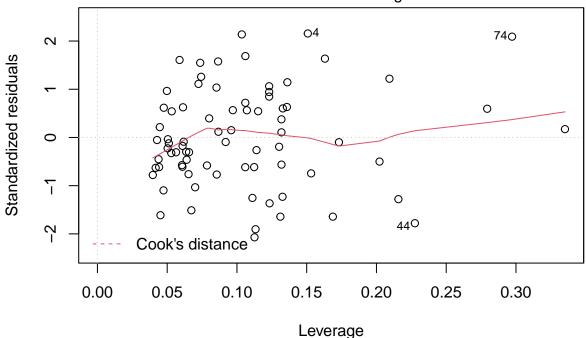


Im(price ~ garage + edison + harris + lot + active + size + age)





## Residuals vs Leverage



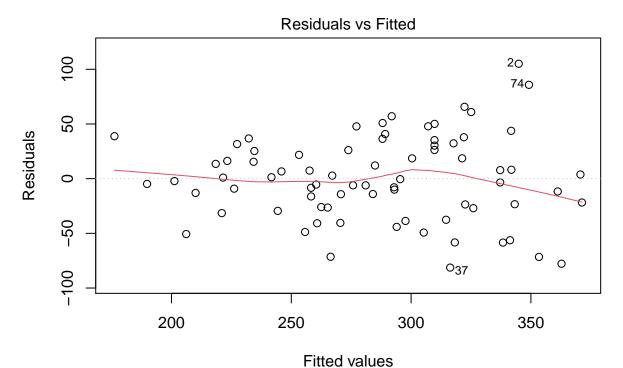
Im(price ~ garage + edison + harris + lot + active + size + age)

```
#Final Model
```

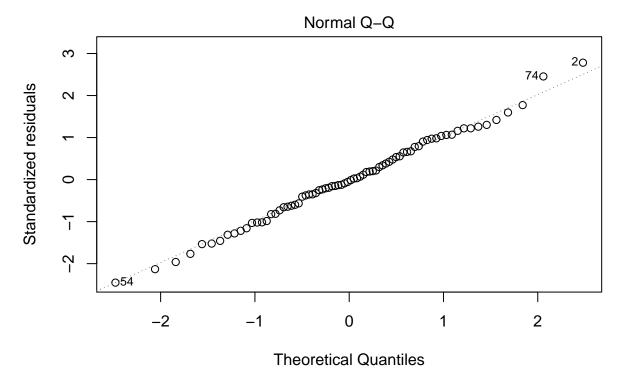
final\_mod <- lm(price~garage+edison+harris+lot+active+size+age+agesq+bath+bed+bathbed,data=homeprice)
summary(final\_mod)</pre>

```
##
## Call:
  lm(formula = price ~ garage + edison + harris + lot + active +
##
       size + age + agesq + bath + bed + bathbed, data = homeprice)
##
##
   Residuals:
##
       Min
                1Q
                    Median
                                 3Q
                                         Max
##
   -81.310 -26.101
                    -1.337
                             27.236 105.083
##
##
  Coefficients:
               Estimate Std. Error t value Pr(>|t|)
##
## (Intercept) 332.4781
                           106.5992
                                       3.119 0.002720 **
                13.1185
                             8.2846
                                       1.583 0.118239
  garage
                67.0619
                            16.8220
## edison
                                       3.987 0.000175 ***
## harris
                47.2732
                            14.8443
                                       3.185 0.002240 **
                             3.4379
                                       2.885 0.005335 **
## lot
                 9.9168
                27.4236
                            10.9885
                                       2.496 0.015160 *
## active
## size
                56.7190
                            27.9738
                                       2.028 0.046772 *
                                       1.042 0.301507
## age
                 3.3014
                             3.1695
## agesq
                  1.6413
                             0.7333
                                       2.238 0.028683 *
## bath
                -98.1564
                            42.6665
                                      -2.301 0.024686 *
               -78.9099
                            27.7519
                                     -2.843 0.005985 **
## bed
```

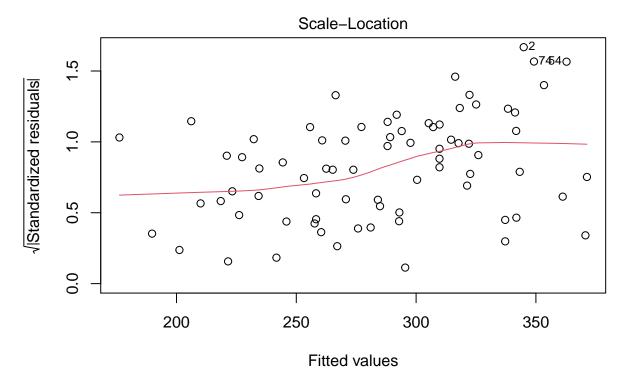
```
## bathbed 30.3901 11.8781 2.559 0.012889 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 41.91 on 64 degrees of freedom
## Multiple R-squared: 0.5883, Adjusted R-squared: 0.5175
## F-statistic: 8.314 on 11 and 64 DF, p-value: 7.698e-09
plot(final_mod)
```



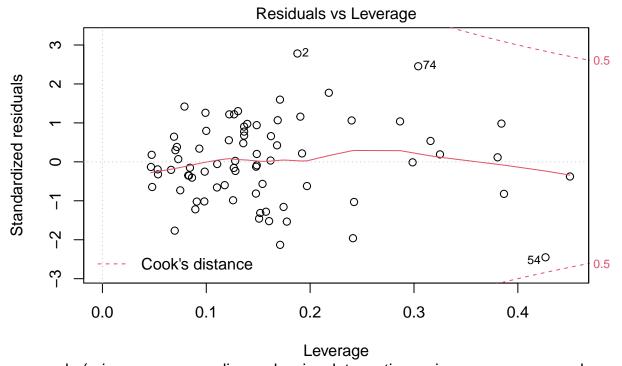
Im(price ~ garage + edison + harris + lot + active + size + age + agesq + b ...



Im(price ~ garage + edison + harris + lot + active + size + age + agesq + b ...



Im(price ~ garage + edison + harris + lot + active + size + age + agesq + b ...



 $Im(price \sim garage + edison + harris + lot + active + size + age + agesq + b ...$  The code in this box is the code used to plot the residuals vs fitted values of our final model, the code run for

our outlier analysis, and the code used to run our AIC and BIC model selection.