

Final Project Appendix

```

lp.initialmodel<-
lm(acrePrice~region+improvements+acres+tillable+financing+crpPct+productivity,dat
a=LandPrices)

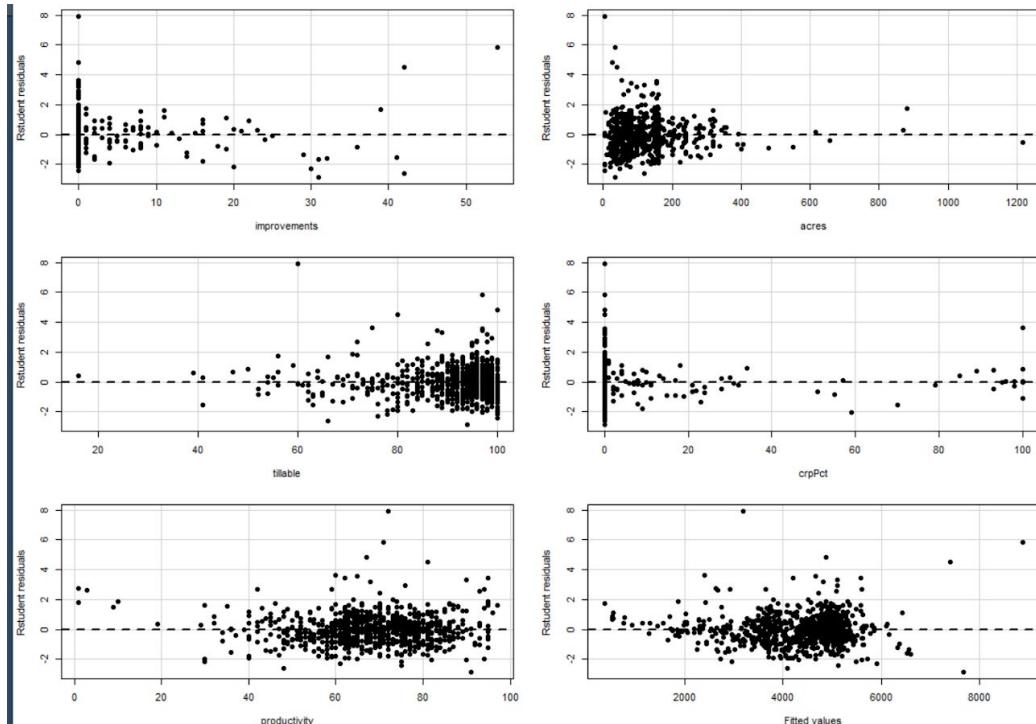
Call:
lm(formula = acrePrice ~ region + improvements + acres + tillable +
    financing + crpPct + productivity, data = LandPrices)

Residuals:
    Min      1Q  Median      3Q     Max 
-2672.2 -623.3  -50.7   471.1  7219.5 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -1429.3062  384.9366 -3.713 0.000219 ***
regionNorthwest -2080.9727 201.8937 -10.307 < 2e-16 ***
regionSouth Central -205.7645 135.1569 -1.522 0.128301  
regionSouth East -391.6871 157.3974 -2.489 0.013030 *  
regionSouth West -205.3499 133.1067 -1.543 0.123288  
regionWest Central -989.7581 143.9293 -6.877 1.23e-11 *** 
improvements       74.2373   6.4835 11.450 < 2e-16 *** 
acres             -0.2598   0.3745 -0.694 0.487936  
tillable           41.8634   3.4870 12.006 < 2e-16 *** 
financingtitle_transfer 259.2991 117.0105  2.216 0.026970 *  
crpPct            -12.0146   2.6013 -4.619 4.50e-06 *** 
productivity        30.9953   2.9367 10.554 < 2e-16 *** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 961.1 on 800 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.5275, Adjusted R-squared:  0.521 
F-statistic: 81.19 on 11 and 800 DF, p-value: < 2.2e-16

```

Figure 0A: Output from running the code of the initial model**Figure 0B:** Residual plots for initial model, checking regression assumptions

```

Test stat Pr(>|Test stat|)
improvements 4.0816    4.922e-05 ***
acres        -0.0405   0.96769
tillable      0.8166    0.41440
crpPct       2.3618    0.01843 *
productivity 4.3313    1.671e-05 ***
Tukey test   5.5135    3.517e-08 ***

---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*'
```

Figure 0C: Test statistic output for residual plots of initial model

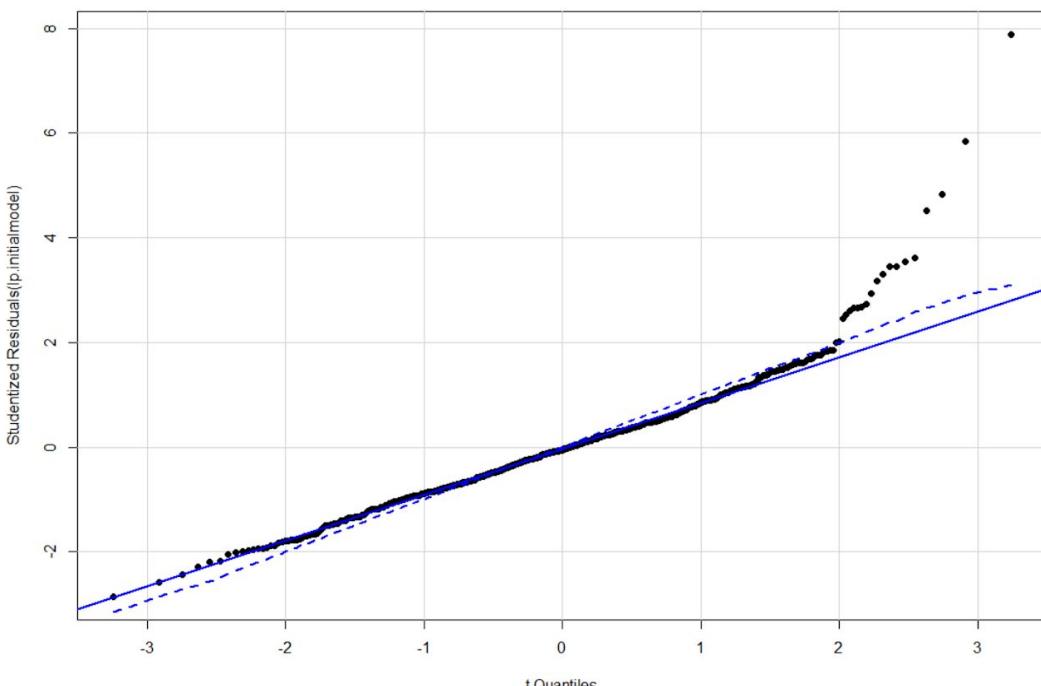


Figure 0D: QQ plot for initial model

```

shapiro-wilk normality test
data: rstudent(lp.initialmodel)
W = 0.93177, p-value < 2.2e-16
```

Figure 0E: Shapiro-Wilk test for Normality output for initial model

```

Variables Tolerance      VIF
1      regionNorthwest 0.6103623 1.638371
2      regionSouth Central 0.3179977 3.144677
3      regionSouth East 0.4659106 2.146335
4      regionSouth West 0.3019739 3.311544
5      regionWest Central 0.3665043 2.728481
6      improvements 0.9519440 1.050482
7      acres 0.9255593 1.080428
8      tillable 0.9182630 1.089013
9      financingtitle_transfer 0.9910576 1.009023
10     crpPct 0.9153410 1.092489
11     productivity 0.7285432 1.372602
```

Figure 0F: Multicollinearity values for initial model

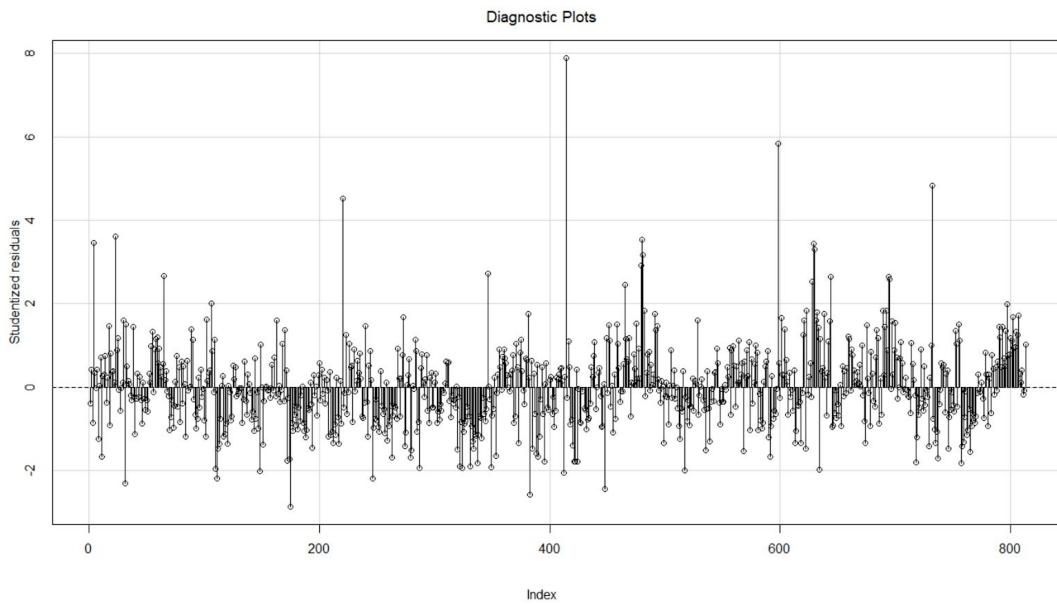


Figure 0G: Plot of studentized residuals for diagnosing outliers

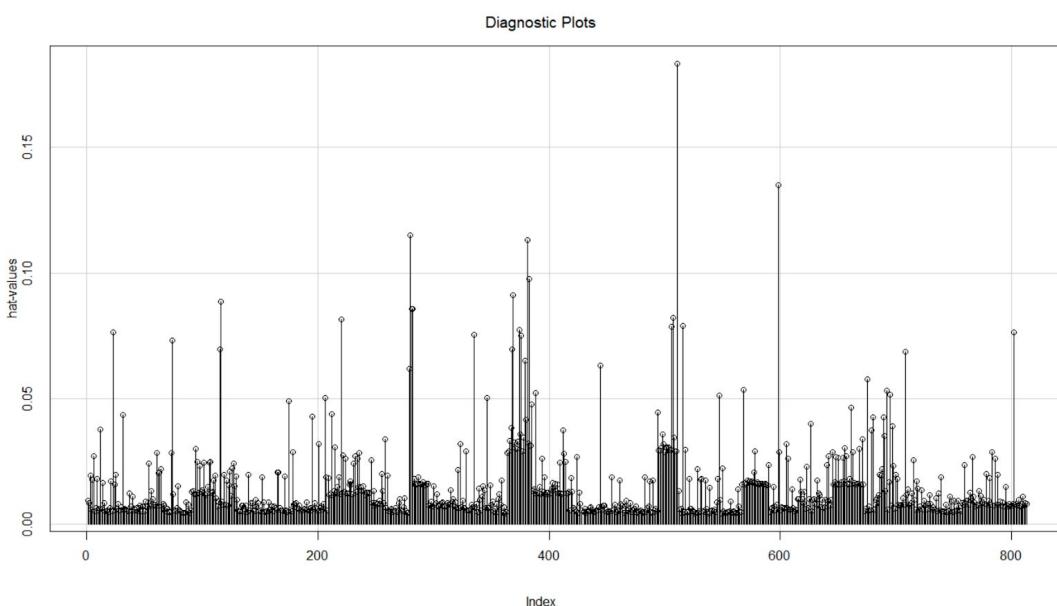


Figure 0H: Plot of hat-values for determining high leverage

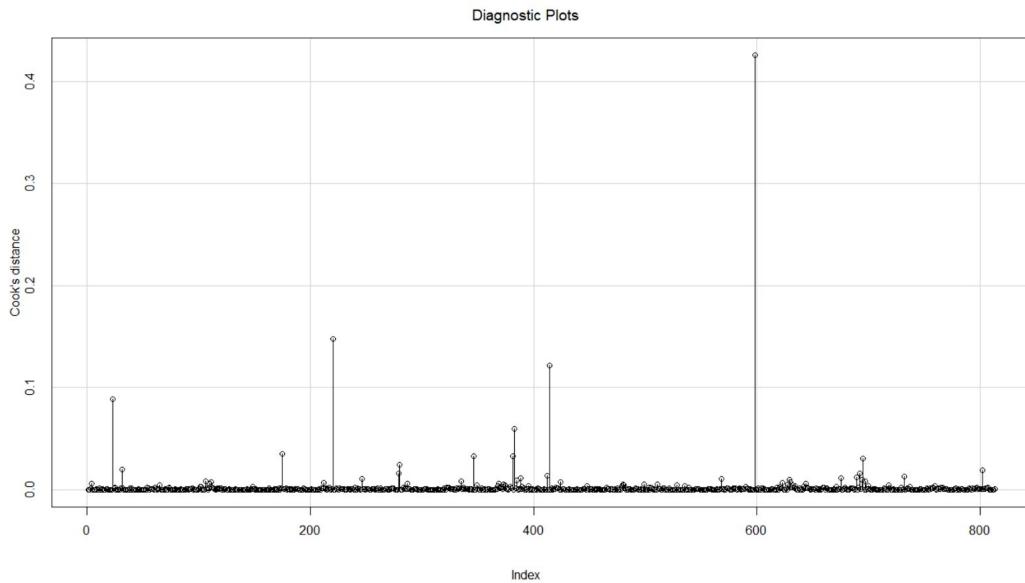


Figure 0I: Plot of Cook's distances for finding influential observations

```
lp.model1 <-
lm(basicPower(acrePrice, 0.5)~improvements+acres+tillable+crpPct+productivity+region+financing,data=LandPrices)
```

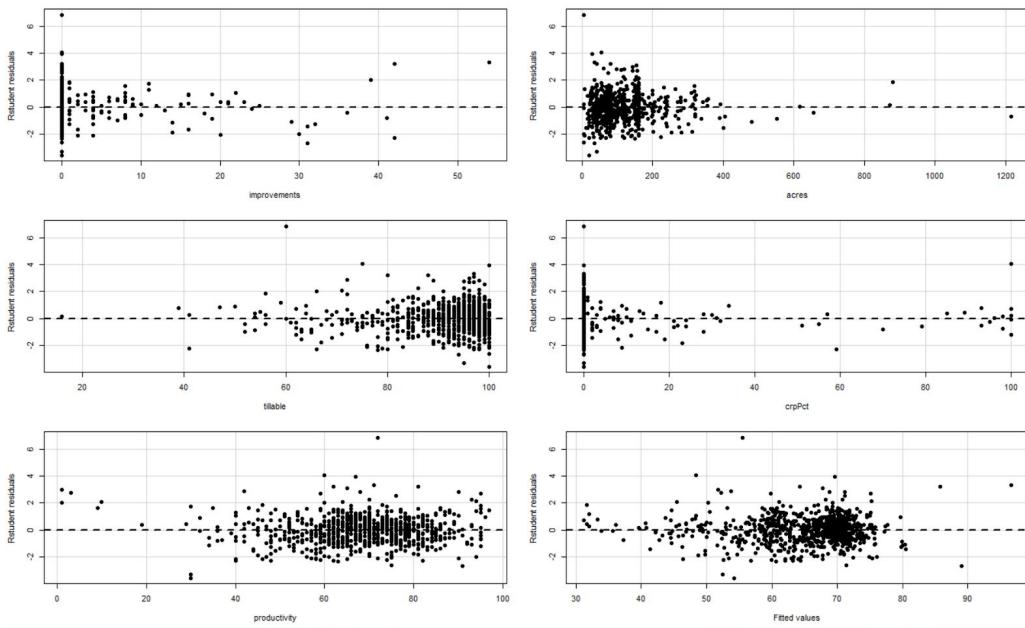


Figure 1A: Residual plots for model 1, checking regression assumptions

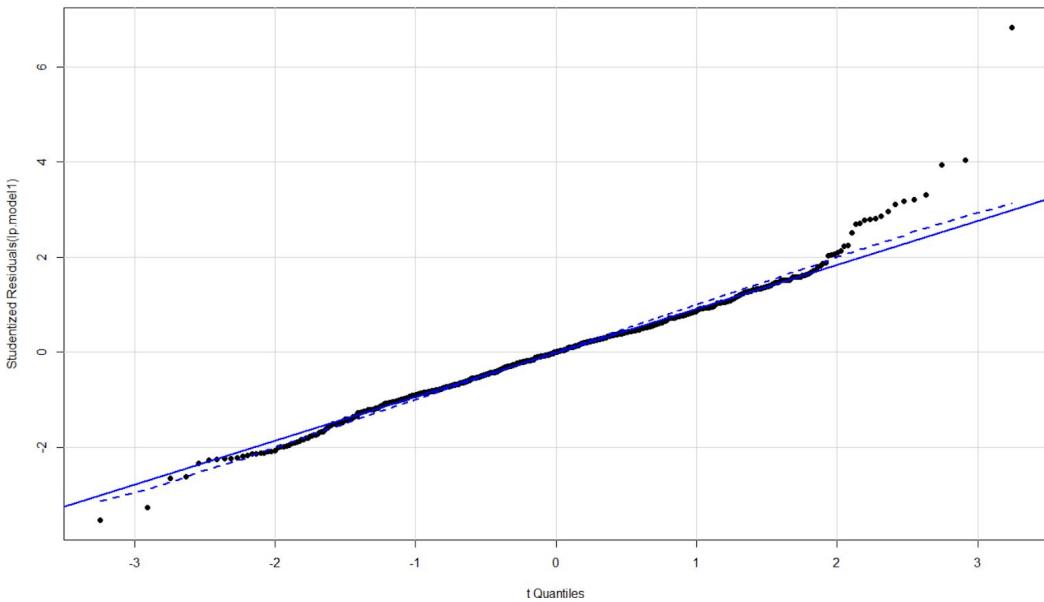


Figure 1B: QQ plot for model 1

```
Shapiro-Wilk normality test
data: rstudent(lp.model1)
W = 0.97078, p-value = 1.126e-11
```

Figure 1C: Shapiro-Wilk test for Normality output for model 1

```
lp.model2 <-
lm(basicPower(acrePrice,0.5)~improvements+basicPower(acres,0)+tillable+crpPct+productivity+region+financing,data=LandPrices)
```

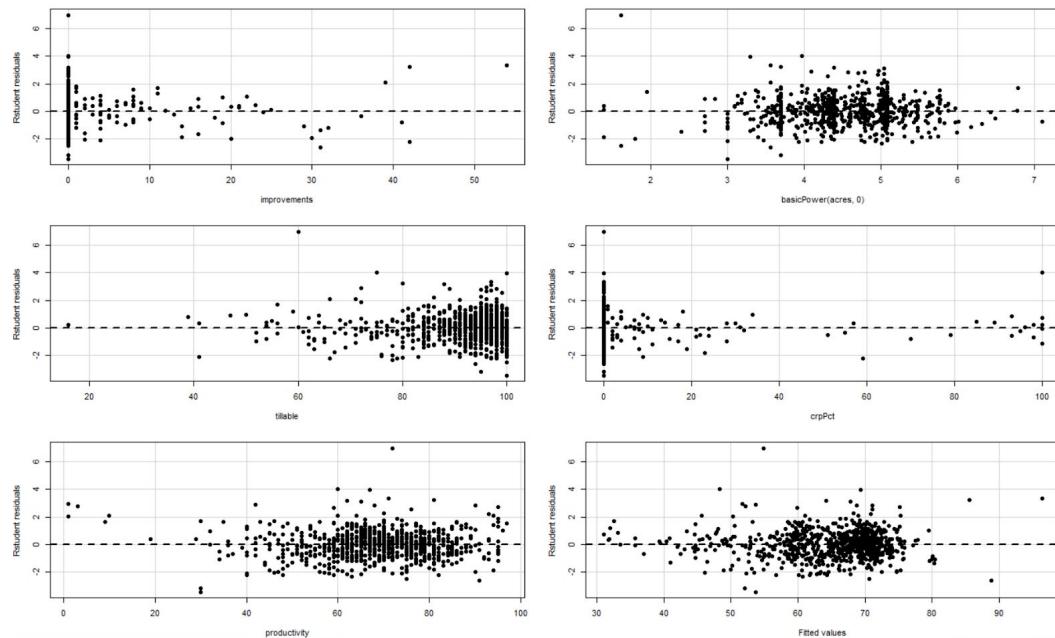


Figure 2A: Residual plots for model 2, checking regression assumptions

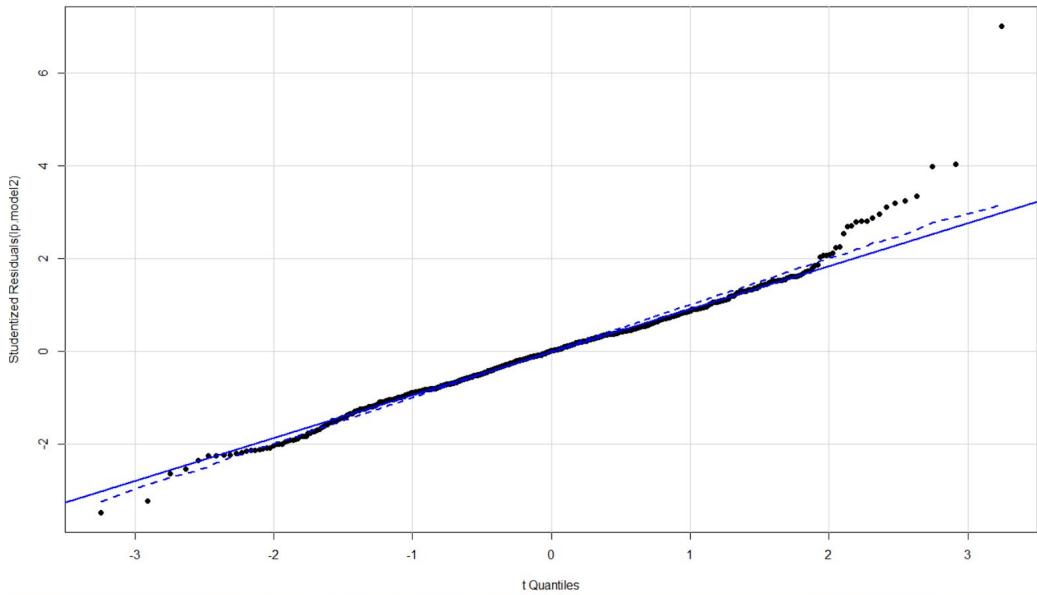


Figure 2B: QQ plot for model 2

```
Shapiro-Wilk normality test  
data: rstudent(lp.model2)  
W = 0.96927, p-value = 4.796e-12
```

Figure 2C: Shapiro-Wilk test for Normality output for model 2

```

lp.model3 <-
lm(sqrt(acrePrice)~(improvements+log(acres)+tillable+crpPct+productivity+region+financing)^2,data=LandPrices)

Call:
lm(formula = sqrt(acrePrice) ~ (improvements + log(acres) + tillable +
    crpPct + productivity + region + financing)^2, data = LandPrices)

Residuals:
    Min      1Q  Median      3Q     Max 
-21.671 -3.987  0.201  3.804 31.512 

Coefficients:
                                         Estimate Std. Error t value Pr(>|t|)    
(Intercept)                         5.458e+01  2.815e+01  1.939 0.052906 .  
improvements                      2.852e-01  6.594e-01  0.433 0.665472  
log(acres)                          -7.687e+00 3.722e+00 -2.065 0.039234 *  
tillable                            -1.778e-01  2.426e-01 -0.733 0.463973  
crpPct                             5.063e-01  3.578e-01  1.415 0.157486  
productivity                       2.822e-01  2.839e-01  0.994 0.320440  
regionNorthwest                     1.683e+00  2.046e+01  0.082 0.934471  
regionSouthCentral                  3.016e+01  1.673e+01  1.803 0.071755 .  
regionSouthEast                     2.179e+01  1.874e+01  1.162 0.245430  
regionSouthWest                     4.738e+00  1.693e+01  -0.280 0.779677  
regionWestCentral                   -7.746e+00  1.775e+01 -0.436 0.666239  
financingtitle_transfer            -1.389e+01  1.347e+01 -1.032 0.302484  
improvements:log(acres)           -1.913e-01  1.065e-01 -1.797 0.072716 .  
improvements:tillable              1.118e-02  7.041e-03  1.589 0.112566  
improvements:crpPct                8.975e-04  4.431e-03  0.203 0.839529  
improvements:productivity          -1.351e-03  5.099e-03 -0.265 0.791156  
improvements:regionNorthwest       -5.248e-02  2.460e-01 -0.213 0.831124  
improvements:regionSouthCentral    2.097e-03  2.110e-01  0.010 0.992071  
improvements:regionSouthEast       2.025e-01  2.068e-01  0.979 0.327691  
improvements:regionSouthWest       1.416e-01  2.341e-01  0.605 0.545556  
improvements:regionWestCentral     5.683e-01  2.377e-01  2.391 0.017051 *  
improvements:financingtitle_transfer 1.819e-02  1.674e-01  0.109 0.913489  
log(acres):tillable               1.425e-01  3.285e-02  4.339 1.63e-05 ***  
log(acres):crpPct                 -5.230e-02  3.762e-02 -1.390 0.164887  
log(acres):productivity            -3.993e-02  3.145e-02 -1.270 0.204623  
log(acres):regionNorthwest         -5.159e-01  2.249e+00 -0.229 0.818607  
log(acres):regionSouthCentral      2.666e+00  1.523e+00 -1.750 0.080516 .  
log(acres):regionSouthEast         3.684e+00  1.752e+00 -2.103 0.035819 *  
log(acres):regionSouthWest         1.280e+00  1.486e+00 -0.862 0.389054  
log(acres):regionWestCentral       1.943e+00  1.554e+00 -1.250 0.211570  
log(acres):financingtitle_transfer 1.595e-01  1.390e+00 -0.115 0.908712  

tillable:crpPct                     -3.304e-03  2.819e-03 -1.172 0.241538  
tillable:productivity                -1.144e-05  1.845e-03 -0.006 0.995053  
tillable:regionNorthwest             -4.943e-01  1.450e-01 -3.408 0.000688 ***  
tillable:regionSouthCentral          -1.976e-01  1.083e-01 -1.825 0.068362 .  
tillable:regionSouthEast             -3.018e-01  1.230e-01 -2.453 0.014394 *  
tillable:regionSouthWest             -1.908e-01  1.083e-01 -1.763 0.078349 .  
tillable:regionWestCentral           -1.717e-01  1.247e-01 -1.377 0.169059  
tillable:financingtitle_transfer    1.475e-01  1.063e-01  1.387 0.165712  
crpPct:productivity                 -3.358e-03  2.583e-03 -1.300 0.194079  
crpPct:regionNorthwest              1.089e-01  8.982e-02  1.213 0.225617  
crpPct:regionSouthCentral           2.176e-01  1.344e-01  1.619 0.105899  
crpPct:regionSouthEast              2.000e-02  1.473e-01  0.136 0.892050  
crpPct:regionSouthWest              5.526e-02  8.773e-02  0.630 0.529009  
crpPct:regionWestCentral            9.154e-02  8.948e-02  1.023 0.306627  
crpPct:financingtitle_transfer     2.300e-02  7.157e-02  0.321 0.748056  
productivity:regionNorthwest        2.901e-01  2.161e-01  1.342 0.179900  
productivity:regionSouthCentral     -4.390e-02  1.684e-01 -0.261 0.794418  
productivity:regionSouthEast        2.539e-01  1.929e-01  1.317 0.188391  
productivity:regionSouthWest        3.338e-01  1.699e-01  1.965 0.049785 *  
productivity:regionWestCentral      3.227e-01  1.806e-01  1.787 0.074348 .  
productivity:financingtitle_transfer 7.533e-03  6.489e-02  0.116 0.907611  
regionNorthwest:financingtitle_transfer 7.824e+00  5.265e+00  1.486 0.137724  
regionSouthCentral:financingtitle_transfer 1.607e+00  3.008e+00  0.534 0.593327  
regionSouthEast:financingtitle_transfer 4.171e-01  3.456e+00  0.121 0.903984  
regionSouthWest:financingtitle_transfer 2.761e+00  2.942e+00  0.938 0.348341  
regionWestCentral:financingtitle_transfer 3.033e+00  3.161e+00  0.960 0.337552  

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.59 on 755 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.6592,    Adjusted R-squared:  0.6339
F-statistic: 26.08 on 56 and 755 DF,   p-value: < 2.2e-16

```

Figure 3: Summary output for model 3 - includes all possible interactions

```

lp.model4 <-
lm(sqrt(acrePrice)~(improvements+log(acres)+tillable+crpPct+productivity+region+financing+improvements*region+log(acres)*region+tillable*region+productivity*region+log(acres)*tillable), data=LandPrices)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 41.237521 16.226210 2.541 0.011233 *
improvements 0.398114 0.120899 3.293 0.001036 **
log(acres) -9.618267 3.094903 -3.108 0.001953 **
tillable 0.009151 0.149930 0.061 0.951345
crpPct -0.079312 0.019595 -4.047 5.69e-05 ***
productivity 0.224519 0.105745 2.123 0.034051 *
regionNorthwest 15.738410 16.409187 0.959 0.337795
regionSouth Central 40.305195 12.967871 3.108 0.001951 **
regionSouth East 30.976686 15.261744 2.030 0.042727 *
regionSouth West 4.910132 13.283861 0.370 0.711758
regionWest Central -0.107540 14.039907 -0.008 0.993891
financingtitle_transfer 1.497239 0.814449 1.838 0.066392 .
improvements:regionNorthwest -0.330681 0.179297 -1.844 0.065517 .
improvements:regionSouth Central 0.119353 0.146279 0.816 0.414791
improvements:regionSouth East 0.162276 0.151769 1.069 0.285297
improvements:regionSouth West 0.046813 0.196779 0.238 0.812024
improvements:regionWest Central 0.452420 0.185246 2.442 0.014817 *
log(acres):regionNorthwest -0.012958 2.051675 -0.006 0.994962
log(acres):regionSouth Central -2.691236 1.463654 -1.839 0.066338 .
log(acres):regionSouth East -3.749015 1.698083 -2.208 0.027549 *
log(acres):regionSouth West -1.194994 1.433078 -0.834 0.404612
log(acres):regionWest Central -1.312839 1.501191 -0.875 0.382098
tillable:regionNorthwest -0.485067 0.139320 -3.482 0.000526 ***
tillable:regionSouth Central -0.204560 0.106279 -1.925 0.054625 .
tillable:regionSouth East -0.293985 0.120886 -2.432 0.015244 *
tillable:regionSouth West -0.183755 0.106816 -1.720 0.085774 .
tillable:regionWest Central -0.166851 0.121652 -1.372 0.170601
productivity:regionNorthwest 0.150823 0.168097 0.897 0.369871
productivity:regionSouth Central -0.155668 0.109820 -1.417 0.156741
productivity:regionSouth East 0.118609 0.143978 0.824 0.410309
productivity:regionSouth West 0.217599 0.110133 1.976 0.048532 *
productivity:regionWest Central 0.205419 0.124916 1.644 0.100485
log(acres):tillable 0.128847 0.030408 4.237 2.53e-05 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.601 on 779 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared: 0.6471, Adjusted R-squared: 0.6326
F-statistic: 44.64 on 32 and 779 DF, p-value: < 2.2e-16

```

Figure 4: Summary output for model 4 - includes only statistically significant interactions

```

lp.model5 <-
lm(sqrt(acrePrice)~improvements+log2(acres)+tillable+productivity+region+improvem
ents*region+log2(acres)*region+tillable*region+productivity*region+log2(acres)*tilla
ble, data=LandPrices)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 37.07052 16.19865 2.288 0.022375 *
improvements 0.36402 0.12160 2.994 0.002843 **
log2(acres) -7.33132 2.16444 -3.387 0.000742 ***
tillable -0.02346 0.15144 -0.155 0.876926
productivity 0.35117 0.10098 3.477 0.000534 ***
regionNorthwest 16.24353 16.57403 0.980 0.327360
regionSouth Central 48.37845 12.90404 3.749 0.000191 ***
regionSouth East 40.77137 15.22639 2.678 0.007569 **
regionSouth West 12.40233 13.22265 0.938 0.348554
regionWest Central 8.03717 13.98506 0.575 0.565662
improvements:regionNorthwest -0.30916 0.18107 -1.707 0.088141 .
improvements:regionSouth Central 0.14669 0.14761 0.994 0.320643
improvements:regionSouth East 0.19930 0.15301 1.303 0.193123
improvements:regionSouth West 0.08108 0.19881 0.408 0.683517
improvements:regionWest Central 0.49951 0.18651 2.678 0.007558 **
log2(acres):regionNorthwest 1.27521 1.40722 0.906 0.365115
log2(acres):regionSouth Central -1.83833 1.02609 -1.792 0.073585 .
log2(acres):regionSouth East -2.48225 1.19023 -2.086 0.037347 *
log2(acres):regionSouth West -0.73784 1.00410 -0.735 0.462665
log2(acres):regionWest Central -0.91065 1.05248 -0.865 0.387172
tillable:regionNorthwest -0.49799 0.14088 -3.535 0.000432 ***
tillable:regionSouth Central -0.19850 0.10748 -1.847 0.065156 .
tillable:regionSouth East -0.31178 0.12221 -2.551 0.010925 *
tillable:regionSouth West -0.19045 0.10801 -1.763 0.078250 .
tillable:regionWest Central -0.17579 0.12295 -1.430 0.153161
productivity:regionNorthwest 0.01123 0.16595 0.068 0.946088
productivity:regionSouth Central -0.28214 0.10519 -2.682 0.007467 **
productivity:regionSouth East -0.01043 0.14135 -0.074 0.941201
productivity:regionSouth West 0.10981 0.10697 1.026 0.304982
productivity:regionWest Central 0.10180 0.12303 0.827 0.408249
log2(acres):tillable 0.09589 0.02127 4.509 7.52e-06 ***

---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.677 on 781 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared: 0.638, Adjusted R-squared: 0.6241
F-statistic: 45.89 on 30 and 781 DF, p-value: < 2.2e-16

```

Figure 5: Summary output for model 5 - includes only statistically significant variables and interactions from model 4 above

	Variables	Tolerance	VIF
1	improvements	0.130638190	7.654729
2	log2(acres)	0.011164030	89.573387
3	tillable	0.023498335	42.556206
4	productivity	0.029740365	33.624335
5	regionNorthwest	0.004371692	228.744385
6	regionSouth Central	0.001683920	593.852457
7	regionSouth East	0.002403124	416.125097
8	regionSouth West	0.001477084	677.009613
9	regionWest Central	0.001873792	533.677051
10	improvements:regionNorthwest	0.404558948	2.471828
11	improvements:regionSouth Central	0.308220659	3.244429
12	improvements:regionSouth East	0.331312534	3.018298
13	improvements:regionSouth West	0.588573576	1.699023
14	improvements:regionWest Central	0.515208002	1.940964
15	log2(acres):regionNorthwest	0.011800180	84.744467
16	log2(acres):regionSouth Central	0.006817725	146.676495
17	log2(acres):regionSouth East	0.009440807	105.923149
18	log2(acres):regionSouth West	0.005527563	180.911540
19	log2(acres):regionWest Central	0.007033631	142.174081
20	tillable:regionNorthwest	0.008108429	123.328457
21	tillable:regionSouth Central	0.002847467	351.189333
22	tillable:regionSouth East	0.004586829	218.015519
23	tillable:regionSouth West	0.002619000	381.825196
24	tillable:regionWest Central	0.002824230	354.078775
25	productivity:regionNorthwest	0.014924969	67.001814
26	productivity:regionSouth Central	0.004472806	223.573281
27	productivity:regionSouth East	0.004725147	211.633620
28	productivity:regionSouth West	0.004424354	226.021687
29	productivity:regionWest Central	0.006748046	148.191057
30	log2(acres):tillable	0.009376550	106.649029

Figure 5A: Multicollinearity VIF values for model 5

```

Call:
lm(formula = sqrt(acrePrice) ~ improvements + log(acres) + tillable +
    crpPct + productivity + region + financing + improvements:log(acres) +
    improvements:tillable + improvements:productivity + improvements:region +
    log(acres):tillable + log(acres):region + tillable:crpPct +
    tillable:region + productivity:region + log(acres):tillable:region +
    improvements:productivity:region, data = LandPrices)

Residuals:
    Min      1Q  Median      3Q     Max 
-20.163 -4.023  0.235  3.741 31.995 

Coefficients:
                                         Estimate Std. Error t value Pr(>|t|)    
(Intercept)                         -77.966703  60.622593 -1.286 0.198797  
improvements                      0.525133  0.597596  0.879 0.379817  
log(acres)                          20.528217 15.285830  1.343 0.179684  
tillable                            1.255930  0.629533  1.995 0.046395 *  
crpPct                             0.438730  0.170259  2.577 0.010157 *  
productivity                        0.265158  0.117722  2.252 0.024579 *  
regionNorthwest                     28.077734 75.224062  0.373 0.709062  
regionSouth Central                 115.968835 64.757573  1.791 0.073718 .  
regionSouth East                    241.610762 66.702532  3.622 0.000311 ***  
regionSouth West                    151.074359 64.367620  2.347 0.019177 *  
regionWest Central                  115.880182 68.400032  1.694 0.090643 .  
financingtitle_transfer            1.643018  0.801202  2.051 0.040637 *  
improvements:log(acres)           -0.318905  0.112894 -2.825 0.004854 **  
improvements:tillable              0.019015  0.006803  2.795 0.005319 **  
improvements:productivity          -0.006131  0.007681 -0.798 0.424961  
improvements:regionNorthwest       -1.134012 2.321681 -0.488 0.625375  
improvements:regionSouth Central   0.409202  0.674307  0.607 0.544131  
improvements:regionSouth East      -3.897202 1.223481 -3.185 0.001504 **  
improvements:regionSouth West      -0.613200 1.071230 -0.572 0.567201  
improvements:regionWest Central    0.904042  2.031324  0.445 0.656409  
log(acres):tillable               -0.193312 0.164642 -1.174 0.240706  
log(acres):regionNorthwest        -9.709491 17.465069 -0.556 0.578416  
log(acres):regionSouth Central    -22.487431 16.249753 -1.384 0.166804  
log(acres):regionSouth East       -54.974487 16.579518 -3.316 0.000957 ***  
log(acres):regionSouth West       -37.615775 15.985955 -2.353 0.018872 *  
log(acres):regionWest Central     -30.778505 16.929699 -1.818 0.069452 .  
tillable:crpPct                   -0.005840  0.001890 -3.089 0.002078 **  
tillable:regionNorthwest           -0.432654 0.818602 -0.529 0.597287  
tillable:regionSouth Central       -0.985790 0.676152 -1.458 0.145267  
tillable:regionSouth East          -2.538320 0.701852 -3.617 0.000318 ***  
tillable:regionSouth West          -0.985790 0.676152 -1.458 0.145267  
tillable:regionSouth East          -2.538320 0.701852 -3.617 0.000318 ***  
tillable:regionSouth West          -2.538320 0.701852 -3.617 0.000318 ***  
tillable:regionWest Central        -1.716872 0.669993 -2.563 0.010582 *  
tillable:regionWest Central        -1.376047 0.719256 -1.913 0.056101 .  
productivity:regionNorthwest      0.102875 0.182874  0.563 0.573909  
productivity:regionSouth Central   0.190051 0.121529 -1.564 0.118272  
productivity:regionSouth East      0.011989 0.154198 -0.078 0.938048  
productivity:regionSouth West      0.167036 0.121644  1.373 0.170107  
productivity:regionWest Central    0.155683 0.135195  1.152 0.249867  
log(acres):tillable:regionNorthwest 0.071884 0.193383  0.372 0.710207  
log(acres):tillable:regionSouth Central 0.210163 0.175103  1.200 0.230422  
log(acres):tillable:regionSouth East 0.568429 0.180329  3.152 0.001684 **  
log(acres):tillable:regionSouth West 0.390408 0.172194  2.267 0.023653 *  
log(acres):tillable:regionWest Central 0.315333 0.183532  1.718 0.086176 .  
improvements:productivity:regionNorthwest 0.025850 0.047312  0.546 0.584969  
improvements:productivity:regionSouth Central -0.006926 0.009749 -0.710 0.477631  
improvements:productivity:regionSouth East 0.054479 0.016514  3.299 0.001015 **  
improvements:productivity:regionSouth West 0.010811 0.014879  0.727 0.467729  
improvements:productivity:regionWest Central -0.005300 0.032852 -0.161 0.871879  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.436 on 765 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.6706,    Adjusted R-squared:  0.6508 
F-statistic: 33.86 on 46 and 765 DF,  p-value: < 2.2e-16

```

Figure 3A: Summary output for stepwise regression testing for model 3

```

1 improvements 5.025109e-03 199.000644
2 log(acres) 4.328495e-04 2310.271862
3 tillable 1.263422e-03 791.500887
4 crpPct 9.582038e-03 104.361933
5 productivity 2.033239e-02 49.182616
6 regionNorthwest 1.971721e-04 5071.710864
7 regionSouth Central 6.212199e-05 16097.359875
8 regionSouth East 1.163426e-04 8595.302438
9 regionSouth West 5.791093e-05 17267.897578
10 regionWest Central 7.277644e-05 13740.711520
11 financingtitle_transfer 9.479618e-01 1.054895
12 improvements:log(acres) 7.911016e-03 126.406017
13 improvements:tillable 5.716135e-03 174.943392
14 improvements:productivity 6.593207e-03 151.671265
15 improvements:regionNorthwest 2.286221e-03 437.403068
16 improvements:regionSouth Central 1.372232e-02 72.873952
17 improvements:regionSouth East 4.814344e-03 207.712616
18 improvements:regionSouth West 1.883559e-02 53.090984
19 improvements:regionWest Central 4.035486e-03 247.801657
20 log(acres):tillable 3.025314e-04 3305.441998
21 log(acres):regionNorthwest 1.481409e-04 6750.329908
22 log(acres):regionSouth Central 5.256769e-05 19023.090349
23 log(acres):regionSouth East 9.408711e-05 10628.448870
24 log(acres):regionSouth West 4.217059e-05 23713.208449
25 log(acres):regionWest Central 5.256690e-05 19023.379455
26 tillable:crpPct 9.703682e-03 103.053669
27 tillable:regionNorthwest 2.231083e-04 4482.127432
28 tillable:regionSouth Central 6.685154e-05 14958.517783
29 tillable:regionSouth East 1.292075e-04 7739.487957
30 tillable:regionSouth West 6.324106e-05 15812.512759
31 tillable:regionWest Central 7.666737e-05 13043.359110
32 productivity:regionNorthwest 1.141933e-02 87.570824
33 productivity:regionSouth Central 3.113104e-03 321.222776
34 productivity:regionSouth East 3.688938e-03 271.080748
35 productivity:regionSouth West 3.178788e-03 314.585325
36 productivity:regionWest Central 5.192041e-03 192.602469
37 log(acres):tillable:regionNorthwest 1.630318e-04 6133.771186
38 log(acres):tillable:regionSouth Central 5.286940e-05 18914.532113
39 log(acres):tillable:regionSouth East 9.652683e-05 10359.813940
40 log(acres):tillable:regionSouth West 4.331951e-05 23084.287795
41 log(acres):tillable:regionWest Central 5.173796e-05 19328.168570
42 improvements:productivity:regionNorthwest 2.319249e-03 431.174049
43 improvements:productivity:regionSouth Central 1.216422e-02 82.208287
44 improvements:productivity:regionSouth East 4.549223e-03 219.817780
45 improvements:productivity:regionSouth West 1.791908e-02 55.806429
46 improvements:productivity:regionWest Central 4.078842e-03 245.167651

```

Figure S1: Multicollinearity VIF values for step 1, the stepwise regression function from above

```

lp.model6 <-
lm(sqrt(acrePrice)~cImp+cAcre+cTill+cPct+cProd+region+financing,data=LandPrices)

Call:
lm(formula = sqrt(acrePrice) ~ cImp + cAcre + cTill + cPct +
    cProd + region + financing, data = LandPrices)

Residuals:
    Min      1Q  Median      3Q     Max 
-24.105 -4.570   0.088   3.835  47.129 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 67.10046   1.13680  59.026 < 2e-16 ***
cImp         0.50152   0.04780  10.493 < 2e-16 ***
cAcre        0.21020   0.36635   0.574  0.56628  
cTill         0.34740   0.02577  13.482 < 2e-16 ***
cPct          -0.10746   0.01917  -5.606  2.85e-08 ***
cProd         0.24668   0.02167  11.382 < 2e-16 *** 
regionNorthwest -19.25499  1.48641 -12.954 < 2e-16 *** 
regionSouth Central -1.66634  0.99670  -1.672  0.09494 .  
regionSouth East -3.20375  1.16105  -2.759  0.00592 ** 
regionSouth West -2.08067  0.98350  -2.116  0.03469 *  
regionWest Central -8.19600  1.06181  -7.719  3.50e-14 *** 
financingtitle_transfer 1.85959  0.86314   2.154  0.03150 * 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 7.087 on 800 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.5824,    Adjusted R-squared:  0.5766 
F-statistic: 101.4 on 11 and 800 DF,  p-value: < 2.2e-16

```

Figure 6: Summary output of model 6 - all quantitative predictor variables centered

```

lp.model7 <-
lm(sqrt(acrePrice)~(cImp+cTill+cPct+cProd+region+financing)^2,data=LandPrices)

Call:
lm(formula = sqrt(acrePrice) ~ (cImp + cTill + cPct + cProd +
    region + financing)^2, data = LandPrices)

Residuals:
    Min      1Q  Median      3Q     Max 
-22.450 -3.972  0.175  3.891 43.052 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept) 6.871e+01 2.357e+00 29.147 < 2e-16 ***
cImp         3.710e-01 1.685e-01  2.201 0.028019 *  
cTill        4.772e-01 1.389e-01  3.434 0.000626 *** 
cPct         -2.538e-01 1.264e-01 -2.007 0.045069 *  
cProd        1.032e-01 1.769e-01  0.583 0.559920    
regionNorthwest -2.214e+01 5.466e+00 -4.051 5.62e-05 *** 
regionSouth Central -1.705e+00 2.870e+00 -0.594 0.552762  
regionSouth East   -4.086e+00 3.396e+00 -1.203 0.229326  
regionSouth West    -4.665e+00 2.796e+00 -1.668 0.095649 .  
regionWest Central  -8.517e+00 3.074e+00 -2.771 0.005726 ** 
financingtitle_transfer -6.568e-01 2.519e+00 -0.261 0.794367  
cImp:cTill       5.546e-03 6.123e-03  0.906 0.365299    
cImp:cPct        2.810e-03 4.288e-03  0.655 0.512440    
cImp:cProd       2.156e-03 4.926e-03  0.438 0.661760  
cImp:regionNorthwest -1.527e-01 2.332e-01 -0.655 0.512732  
cImp:regionSouth Central 1.212e-01 1.944e-01  0.623 0.533362  
cImp:regionSouth East   1.704e-01 2.069e-01  0.824 0.410392  
cImp:regionSouth West    9.563e-02 2.337e-01  0.409 0.682445  
cImp:regionWest Central  5.726e-01 2.389e-01  2.397 0.016775 *  
cImp:financingtitle_transfer 2.208e-02 1.681e-01  0.131 0.895556  
cTill:cPct        -1.849e-03 2.657e-03 -0.696 0.486682  
cTill:cProd       -4.985e-04 1.864e-03 -0.267 0.789213  
cTill:regionNorthwest -3.475e-01 1.395e-01 -2.491 0.012945 *  
cTill:regionSouth Central 1.848e-01 1.080e-01 -1.711 0.087498 .  
cTill:regionSouth East   -3.061e-01 1.219e-01 -2.511 0.012248 *  
cTill:regionSouth West    -1.280e-01 1.069e-01 -1.198 0.231416  
cTill:regionWest Central  -1.434e-01 1.245e-01 -1.152 0.249843  
cTill:financingtitle_transfer 5.050e-02 1.049e-01  0.481 0.630397  
cPct:cProd        -3.812e-03 2.515e-03 -1.516 0.129897    
cPct:regionNorthwest 1.089e-01 9.022e-02  1.207 0.227921  
cPct:regionSouth Central 2.670e-01 1.329e-01  2.009 0.044908 *  
cPct:regionSouth East   6.283e-02 1.478e-01  0.425 0.670938  
cPct:regionSouth West    3.778e-02 8.873e-02  0.426 0.670411  
cPct:regionWest Central  7.885e-02 9.003e-02  0.876 0.381405  
cPct:financingtitle_transfer 8.383e-03 7.127e-02  0.118 0.906394  

cProd:regionNorthwest 3.003e-01 2.139e-01  1.404 0.160664  
cProd:regionSouth Central -7.329e-02 1.701e-01 -0.431 0.666747  
cProd:regionSouth East   1.696e-01 1.936e-01  0.876 0.381309  
cProd:regionSouth West    2.929e-01 1.716e-01  1.707 0.088209 .  
cProd:regionWest Central  3.205e-01 1.829e-01  1.753 0.080048 .  
cProd:financingtitle_transfer 3.452e-02 6.259e-02  0.551 0.581454  
regionNorthwest:financingtitle_transfer 6.267e+00 5.303e+00  1.182 0.237674  
regionSouth Central:financingtitle_transfer 1.819e+00 3.044e+00  0.598 0.550193  
regionSouth East:financingtitle_transfer 8.424e-01 3.488e+00  0.242 0.809206  
regionSouth West:financingtitle_transfer 3.080e+00 2.980e+00  1.033 0.301782  
regionWest Central:financingtitle_transfer 3.431e+00 3.198e+00  1.073 0.283713  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.698 on 766 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.6428, Adjusted R-squared:  0.6218 
F-statistic: 30.63 on 45 and 766 DF,  p-value: < 2.2e-16

```

Figure 7: Summary output for model 7- all interactions included, after insignificant term from model 6 was eliminated

```

lp.model8 <-
lm(sqrt(acrePrice)~(cImp+cTill+cPct+region+region*cImp+region*cTill+region*cPct)
,data=LandPrices)

Call:
lm(formula = sqrt(acrePrice) ~ (cImp + cTill + cPct + region +
region * cImp + region * cTill + region * cPct), data = LandPrices)

Residuals:
    Min      1Q  Median      3Q     Max 
-32.552 -4.512   0.224   4.168  42.209 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 68.365154  0.942915 72.504 < 2e-16 ***
cImp         0.388304  0.135787  2.860  0.00435 **  
cTill        0.519261  0.107654  4.823 1.69e-06 *** 
cPct        -0.175801  0.050300 -3.495  0.00050 *** 
regionNorthwest -22.658535  1.683784 -13.457 < 2e-16 *** 
regionSouth Central 0.037305  1.074208  0.035  0.97231    
regionSouth East   -1.203011  1.243636 -0.967  0.33367    
regionSouth West   -1.420195  1.060963 -1.339  0.18109    
regionWest Central -10.002561  1.129402 -8.857 < 2e-16 *** 
cImp:regionNorthwest -0.274889  0.201634 -1.363  0.17318    
cImp:regionSouth Central 0.116330  0.164851  0.706  0.48060    
cImp:regionSouth East   0.150665  0.171199  0.880  0.37910    
cImp:regionSouth West   0.104759  0.222399  0.471  0.63774    
cImp:regionWest Central 0.572202  0.208069  2.750  0.00610 ** 
cTill:regionNorthwest -0.263790  0.146208 -1.804  0.07158 .  
cTill:regionSouth Central 0.152667  0.118391 -1.290  0.19760    
cTill:regionSouth East   -0.284195  0.133851 -2.123  0.03405 *  
cTill:regionSouth West   -0.052535  0.117378 -0.448  0.65459    
cTill:regionWest Central -0.003236  0.133514 -0.024  0.98067    
cPct:regionNorthwest  0.090860  0.061257  1.483  0.13841    
cPct:regionSouth Central 0.224199  0.077829  2.881  0.00408 **  
cPct:regionSouth East   -0.049524  0.102771 -0.482  0.63002    
cPct:regionSouth West   -0.086914  0.067767 -1.283  0.20003    
cPct:regionWest Central 0.018607  0.072436  0.257  0.79734  
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.485 on 788 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.5411,    Adjusted R-squared:  0.5277 
F-statistic: 40.39 on 23 and 788 DF,  p-value: < 2.2e-16

```

Figure 8: Regression summary output for model 8 - only significant variables and interactions included - used as final model

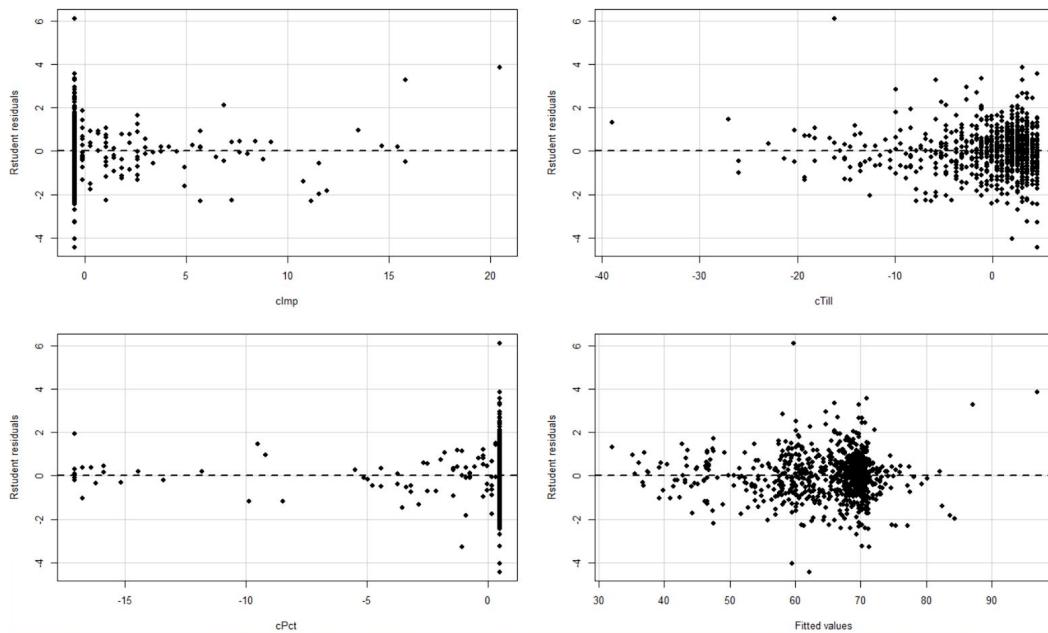


Figure 8A: Residual plots for final model

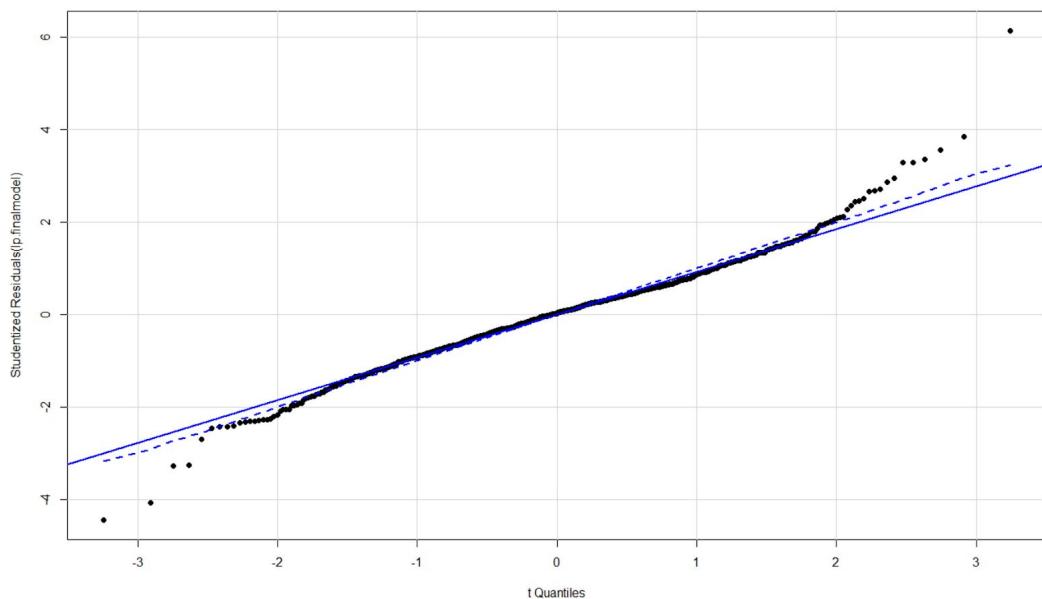


Figure 8B: QQ plot for final model

```
Shapiro-Wilk normality test
data: rstudent(lp.finalmodel)
W = 0.97394, p-value = 7.377e-11
```

Figure 8C: Shapiro-Wilk test for Normality output for final model

	variables	Tolerance	VIF
1	cImp	0.13164269	7.596320
2	cTill	0.05843469	17.113122
3	cPct	0.14849043	6.734441
4	regionNorthwest	0.53227462	1.878729
5	regionSouth Central	0.30535067	3.274923
6	regionSouth East	0.45267472	2.209092
7	regionSouth West	0.28830006	3.468608
8	regionWest Central	0.36104004	2.769776
9	cImp:regionNorthwest	0.43940083	2.275826
10	cImp:regionSouth Central	0.32054005	3.119735
11	cImp:regionSouth East	0.35698498	2.801238
12	cImp:regionSouth West	0.60555468	1.651379
13	cImp:regionWest Central	0.53146017	1.881609
14	cTill:regionNorthwest	0.32239859	3.101751
15	cTill:regionSouth Central	0.16819141	5.945607
16	cTill:regionSouth East	0.33949755	2.945529
17	cTill:regionSouth West	0.15731378	6.356722
18	cTill:regionWest Central	0.32319262	3.094130
19	cPct:regionNorthwest	0.28291744	3.534600
20	cPct:regionSouth Central	0.55710034	1.795009
21	cPct:regionSouth East	0.74452119	1.343145
22	cPct:regionSouth West	0.44425068	2.250981
23	cPct:regionWest Central	0.51673566	1.935225

Figure 8D: Multicollinearity values for final model

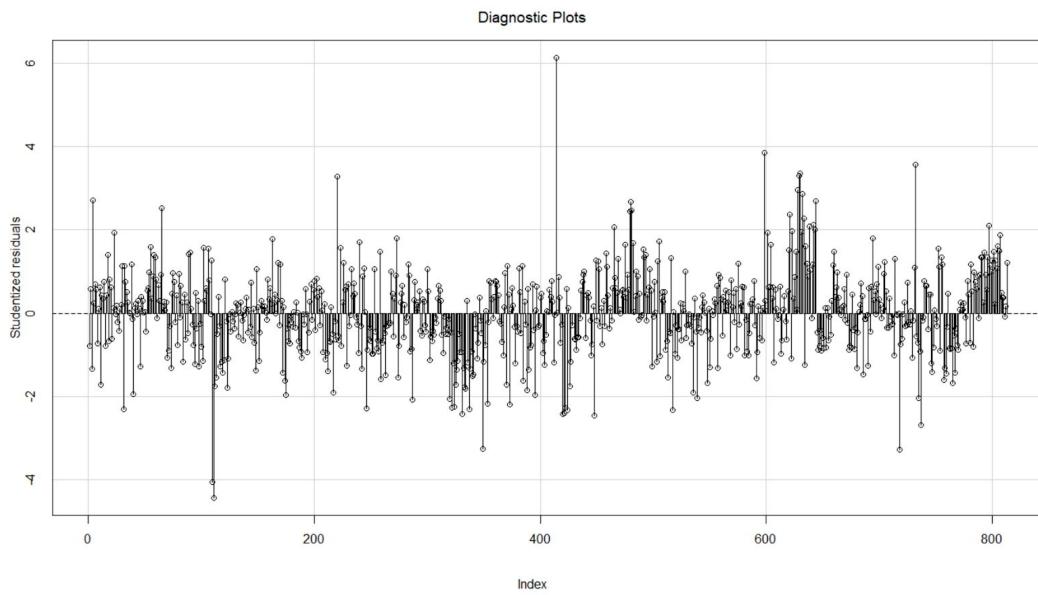


Figure 8E: Plot of studentized residuals for diagnosing outliers

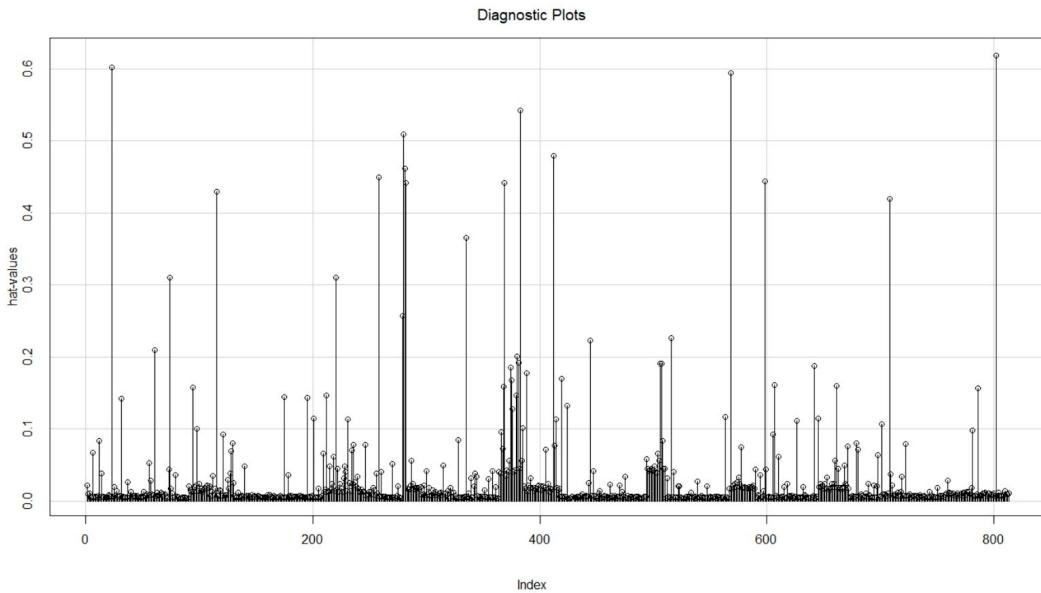


Figure 8F: Plot of hat-values for determining high leverage

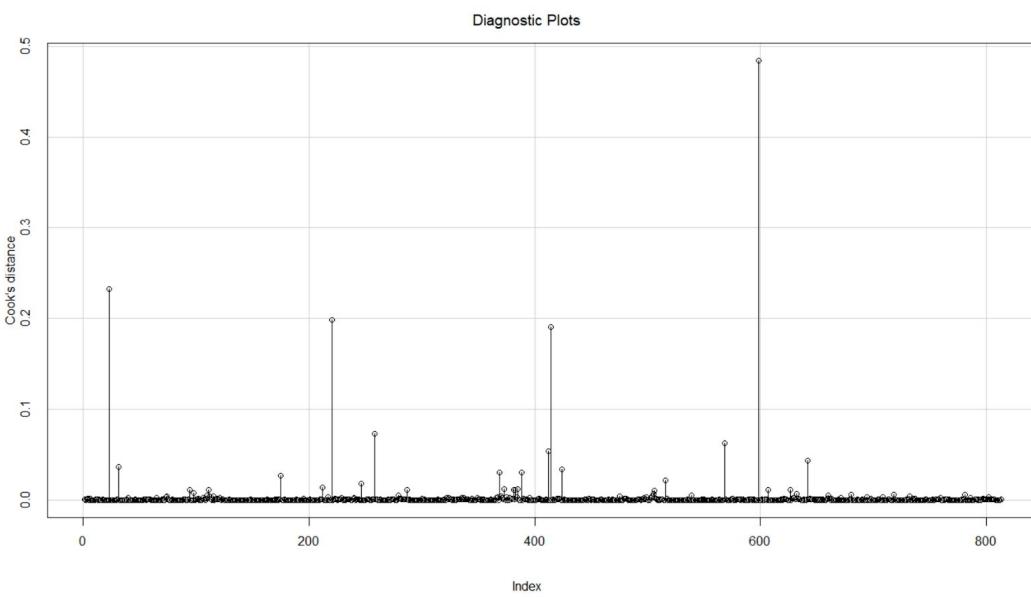


Figure 8G: Plot of Cook's distances for finding influential observations

#	n	predictors	rsquare	adjr	predrsq	cp	aic	sbic	sbc
1	1	region	0.3126265	0.3083677	0.3030480	368.290850	5897.264	3581.516	5930.169
2	2	region cTill:region	0.4369279	0.4291953	0.4144745	157.395221	5747.095	3422.265	5808.205
3	3	region clmp:region cTill:region	0.5016573	0.4909875	0.4611767	49.496566	5653.864	3323.017	5743.154
4	4	region clmp:region cTill:region cPct:region	0.5410661	0.5276708	0.4922080	-15.997699	5598.970	3259.446	5716.457
5	5	clmp region clmp:region cTill:region cPct:region	0.5410661	0.5276708	0.4922080	-13.997699	5598.970	3262.049	5716.457
6	6	clmp cTill region clmp:region cTill:region cPct:region	0.5410661	0.5276708	0.4922080	-11.997699	5598.970	3264.652	5716.457
7	7	cTill cPct region clmp:region cTill:cPct cTill:region cPct:region	0.5414795	0.5274966	0.4931362	-10.705683	5600.238	3266.540	5722.425
8	8	clmp cTill cPct region clmp:region cTill:cPct cTill:region cPct:region	0.5414795	0.5274966	0.4931362	-8.705683	5600.238	3269.145	5722.425
9	9	clmp cTill cPct region clmp:cTill clmp:region cTill:cPct cTill:region cPct:region	0.5415368	0.5269546	0.4867534	-6.803841	5602.137	3271.652	5729.023
10	10	clmp cTill cPct region clmp:cTill clmp:cPct clmp:region cTill:cPct cTill:region cPct:region	0.5416513	0.5264703	0.4841084	-5.000000	5603.934	3274.061	5735.520

Figure 8H: Highest adjusted R² and lowest Cp, AIC, and SBC (#4 - same as our final)