ASSIGNMENT 8.2

Sharath Chandra Tummanapally

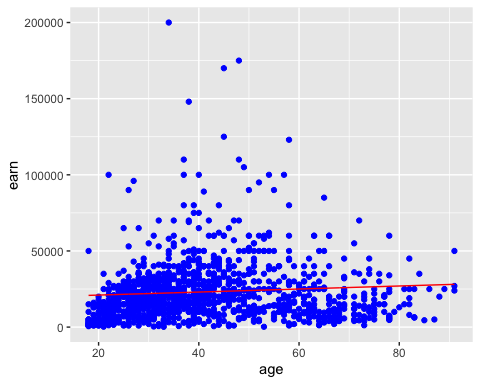
2021-10-20

# 1.

## View the summary of your model using summary()

##   
## Call:  
## lm(formula = earn ~ age, data = heights\_df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -25098 -12622 -3667 6883 177579   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 19041.53 1571.26 12.119 < 2e-16 \*\*\*  
## age 99.41 35.46 2.804 0.00514 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 19420 on 1190 degrees of freedom  
## Multiple R-squared: 0.006561, Adjusted R-squared: 0.005727   
## F-statistic: 7.86 on 1 and 1190 DF, p-value: 0.005137

## Plot the predictions against the original data



## R Squared R^2 = SSM

## [1] 0.006561482

## Corrected Degrees of Freedom for Model (p-1)

## [1] 1

## Degrees of Freedom for Error (n-p)

## [1] 1190

## F Statistic F = MSM/MSE

## [1] 7.859735

## Adjusted R Squared R2 = 1 - (1 - R2)(n - 1) / (n - p)

## [1] 0.005726659

## Calculate the p-value from the F distribution

## [1] 0.005136826

# 2.

## View the summary of your model

##   
## Call:  
## lm(formula = earn ~ ed + race + height + age + sex, data = heights\_df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -39423 -9827 -2208 6157 158723   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -41478.4 12409.4 -3.342 0.000856 \*\*\*  
## ed 2768.4 209.9 13.190 < 2e-16 \*\*\*  
## racehispanic -1414.3 2685.2 -0.527 0.598507   
## raceother 371.0 3837.0 0.097 0.922983   
## racewhite 2432.5 1723.9 1.411 0.158489   
## height 202.5 185.6 1.091 0.275420   
## age 178.3 32.2 5.537 3.78e-08 \*\*\*  
## sexmale 10325.6 1424.5 7.249 7.57e-13 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 17250 on 1184 degrees of freedom  
## Multiple R-squared: 0.2199, Adjusted R-squared: 0.2153   
## F-statistic: 47.68 on 7 and 1184 DF, p-value: < 2.2e-16

## R Squared

## [1] 0.2198953

## Corrected Degrees of Freedom for Model

## [1] 7

## Degrees of Freedom for Error

## [1] 1184

## F Statistic

## [1] 47.67785

## Adjusted R Squared R2 = 1 - (1 - R2)(n - 1) / (n - p)

## [1] 0.2152832

# 3.

## Housing Data

## a. Work individually on this assignment. You are encouraged to collaborate on ideas and strategies pertinent to this assignment. Data for this assignment is focused on real estate transactions recorded from 1964 to 2016 and can be found in Housing.xlsx. Using your skills in statistical correlation, multiple regression, and R programming, you are interested in the following variables: Sale Price and several other possible predictors.

## If you worked with the Housing dataset in previous week – you are in luck, you likely have already found any issues in the dataset and made the necessary transformations. If not, you will want to take some time looking at the data with all your new skills and identifying if you have any clean up that needs to happen.

## b. Complete the following:

## Explain any transformations or modifications you made to the dataset

### Transformations/Modifications that I made to the dataset are:

* Using the ‘ddply’ function, calculated the Average sale price of a house by year is built. This function splits the data, performs modification to the data, and then brings it back together.
* Performed six different operations to analyze/transform the data i.e; GroupBy, Summarize, Mutate, Filter, Select, and Arrange.
  1. Using GroupBy operation, getting the max sale price by site type.
  2. Using Summarize operation, getting the average sale price , median sale price and maximum sq ft lot.
  3. Using Mutate operation, adding the new ‘site\_area\_occupied’ column to housing date frame.
  4. Using Filter operation, getting the housing data from data frame for the site type’s A1,C1.
  5. Using Select function, getting sale date, sale price, address and square feet of lot from housing data frame.
  6. Using Arrange operation, sorting the columns by average price.
* Using purrr package, performed 2 functions on housing dataset.
  1. Using keep function, getting values sale price greater than 600000.
  2. Using discard function, removing all NA values from sale warning.
* Used the cbind and rbind function on the dataset.

## Create two variables; one that will contain the variables Sale Price and Square Foot of Lot (same variables used from previous assignment on simple regression) and one that will contain Sale Price and several additional predictors of your choice. Explain the basis for your additional predictor selections.

Here, We are trying to predict sale price(dependent variable) based on sale date, sq\_ft\_lot, bedrooms, bath count and year built(independent variables). These variables may have affect on the sale price, that are used to estimate the relationship between variables and effect of additional variables on sale price variations.

## Execute a summary() function on two variables defined in the previous step to compare the model results. What are the R2 and Adjusted R2 statistics? Explain what these results tell you about the overall model. Did the inclusion of the additional predictors help explain any large variations found in Sale Price?

##   
## Call:  
## lm(formula = `Sale Price` ~ sq\_ft\_lot, data = housing\_df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2016064 -194842 -63293 91565 3735109   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.418e+05 3.800e+03 168.90 <2e-16 \*\*\*  
## sq\_ft\_lot 8.510e-01 6.217e-02 13.69 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 401500 on 12863 degrees of freedom  
## Multiple R-squared: 0.01435, Adjusted R-squared: 0.01428   
## F-statistic: 187.3 on 1 and 12863 DF, p-value: < 2.2e-16

##   
## Call:  
## lm(formula = `Sale Price` ~ `Sale Date` + sq\_ft\_lot + bedrooms +   
## bath\_full\_count + year\_built, data = housing\_df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2380845 -148498 -45720 63249 3672187   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -9.271e+06 4.387e+05 -21.133 < 2e-16 \*\*\*  
## `Sale Date` 2.068e-04 3.190e-05 6.483 9.34e-11 \*\*\*  
## sq\_ft\_lot 9.311e-01 5.891e-02 15.806 < 2e-16 \*\*\*  
## bedrooms 7.946e+04 4.002e+03 19.854 < 2e-16 \*\*\*  
## bath\_full\_count 8.558e+04 6.072e+03 14.093 < 2e-16 \*\*\*  
## year\_built 4.621e+03 2.205e+02 20.959 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 374100 on 12859 degrees of freedom  
## Multiple R-squared: 0.1446, Adjusted R-squared: 0.1443   
## F-statistic: 434.8 on 5 and 12859 DF, p-value: < 2.2e-16

R2 assumes that every single variable explains the variation in the dependent variable. The adjusted R2 tells the percentage of variation explained by only the independent variables that actually affect the dependent variable. In our example: R2, adjusted R2 for simple regression model are 0.01435, 0.01428. R2, adjusted R2 for multiple regression model are 0.1446, 0.1443. Based on the results we have seen, both the models are similar but second model has both R2 and adjusted values very close to each other. And the inclusion of additional predictors have slight variation in Sale Price. The more full baths a house has the higher price compared to other additional rooms.

## Considering the parameters of the multiple regression model you have created. What are the standardized betas for each parameter and what do the values indicate?

The standardized betas for each parameter are:

## Loading required package: boot

## Loading required package: MASS

##   
## Attaching package: 'QuantPsyc'

## The following object is masked from 'package:base':  
##   
## norm

## `Sale Date` sq\_ft\_lot bedrooms bath\_full\_count year\_built   
## 0.05296503 0.13109685 0.17216649 0.13772539 0.19676964

A standardized beta coefficient compares the strength of the effect of each individual independent variable to the dependent variable. The values indicate that a change in 1 standard deviation of year\_built has more impact on sale price than a 1 standard deviation change in the other variables i.e; sale date, sq\_ft\_lot, bedrooms, bath\_full\_count.

## Calculate the confidence intervals for the parameters in your model and explain what the results indicate.

## 2.5 % 97.5 %  
## (Intercept) 6.343730e+05 6.492698e+05  
## sq\_ft\_lot 7.291208e-01 9.728641e-01

## 2.5 % 97.5 %  
## (Intercept) -1.013065e+07 -8.410843e+06  
## `Sale Date` 1.442755e-04 2.693401e-04  
## sq\_ft\_lot 8.156679e-01 1.046620e+00  
## bedrooms 7.161903e+04 8.730965e+04  
## bath\_full\_count 7.367459e+04 9.748040e+04  
## year\_built 4.188602e+03 5.052878e+03

* These confidence interval results indicates that our data is representative of our population.

## Assess the improvement of the new model compared to your original model (simple regression model) by testing whether this change is significant by performing an analysis of variance.

## Df Sum Sq Mean Sq F value Pr(>F)   
## sq\_ft\_lot 1 3.020e+13 3.020e+13 187.3 <2e-16 \*\*\*  
## Residuals 12863 2.073e+15 1.612e+11   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Df Sum Sq Mean Sq F value Pr(>F)   
## `Sale Date` 1 7.263e+12 7.263e+12 51.9 6.16e-13 \*\*\*  
## sq\_ft\_lot 1 2.983e+13 2.983e+13 213.2 < 2e-16 \*\*\*  
## bedrooms 1 9.945e+13 9.945e+13 710.7 < 2e-16 \*\*\*  
## bath\_full\_count 1 1.062e+14 1.062e+14 759.0 < 2e-16 \*\*\*  
## year\_built 1 6.147e+13 6.147e+13 439.3 < 2e-16 \*\*\*  
## Residuals 12859 1.799e+15 1.399e+11   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Assessment: Adding sale date, sq\_ft\_lot, bedrooms, bath\_full\_count, year\_built to the model seems to have made the model better: it reduced the residual variance (the residual sum of squares went from 2.073e+15 to 1.799e+15), and both original and new model are statistically significant (p-values < 0.001).

## Perform casewise diagnostics to identify outliers and/or influential cases, storing each function’s output in a dataframe assigned to a unique variable name.

## Calculate the standardized residuals using the appropriate command, specifying those that are +-2, storing the results of large residuals in a variable you create.

## Use the appropriate function to show the sum of large residuals.

## [1] 352

## Which specific variables have large residuals (only cases that evaluate as TRUE)?

## Investigate further by calculating the leverage, cooks distance, and covariance rations. Comment on all cases that are problematics.

## leverage cooks.dist covariance.ratios  
## 14 0.0021694436 0.0015021483 1.0007041  
## 72 0.0006590042 0.0010052135 0.9968617  
## 108 0.0011117497 0.0011952502 0.9985727  
## 115 0.0028391176 0.0040834541 0.9992935  
## 239 0.0005482338 0.0006800420 0.9975462  
## 245 0.0005422971 0.0005831582 0.9980014  
## 246 0.0021755270 0.0048435836 0.9964282  
## 295 0.1123421565 0.9626139789 1.1032981  
## 300 0.0031279164 0.0023731081 1.0014828  
## 342 0.0073735000 0.0059727176 1.0056318  
## 344 0.0085734086 0.0101947799 1.0057923  
## 345 0.0085734086 0.0101947799 1.0057923  
## 385 0.0003636711 0.0013221340 0.9906910  
## 396 0.0011225619 0.0019148200 0.9968229  
## 475 0.0006144506 0.0008610932 0.9971631  
## 508 0.0024896273 0.0032733147 0.9992868  
## 519 0.0003173592 0.0002250535 0.9987997  
## 601 0.0005496294 0.0004550372 0.9987003  
## 612 0.0010115527 0.0013017595 0.9978807  
## 656 0.0004494226 0.0005047123 0.9977752  
## 661 0.0010807800 0.0032766924 0.9930874  
## 662 0.0005589766 0.0006799006 0.9976243  
## 670 0.0003060645 0.0013280817 0.9886806  
## 671 0.0003484141 0.0002572891 0.9987489  
## 679 0.0018343373 0.0014708628 1.0000615  
## 723 0.0003105367 0.0002253108 0.9987470  
## 785 0.0005452663 0.0007048806 0.9973972  
## 811 0.0003109864 0.0019272775 0.9835449  
## 853 0.0009288104 0.0008494144 0.9988381  
## 877 0.0014575568 0.0013156235 0.9994016  
## 916 0.0025473596 0.0017241315 1.0011275  
## 929 0.0005198733 0.0006019313 0.9977485  
## 966 0.0005521132 0.0004235455 0.9988727  
## 1033 0.0004949102 0.0005913736 0.9976199  
## 1155 0.0084839887 0.0270529064 1.0001289  
## 1179 0.0003754332 0.0003963947 0.9978888  
## 1294 0.0005444202 0.0003949767 0.9989814  
## 1305 0.0017883201 0.0018665235 0.9993393  
## 1310 0.0005434940 0.0003942137 0.9989809  
## 1442 0.0018425801 0.0024642573 0.9985735  
## 1633 0.0006475547 0.0005539270 0.9987215  
## 1716 0.0005775363 0.0005097055 0.9985755  
## 1745 0.0004852565 0.0010946373 0.9946508  
## 1870 0.0019008310 0.0075823199 0.9912512  
## 1962 0.0010360038 0.0028159500 0.9939153  
## 1963 0.0003352276 0.0005249981 0.9964237  
## 1964 0.0008477285 0.0017108712 0.9956761  
## 1976 0.0006087393 0.0021116995 0.9913991  
## 1977 0.0003101341 0.0012960414 0.9891291  
## 1978 0.0004703386 0.0017943875 0.9902992  
## 1979 0.0003007764 0.0012630505 0.9890634  
## 1980 0.0004743793 0.0018183456 0.9902533  
## 1981 0.0009723980 0.0030898233 0.9925735  
## 1982 0.0006111774 0.0021265360 0.9913725  
## 2020 0.0012380351 0.0018774266 0.9974669  
## 2021 0.0003333443 0.0003441310 0.9979123  
## 2022 0.0016493534 0.0031550304 0.9967737  
## 2084 0.0005059296 0.0004424249 0.9985263  
## 2099 0.0004891247 0.0007199057 0.9968409  
## 2137 0.0003673763 0.0008101651 0.9946735  
## 2157 0.0005416561 0.0008137904 0.9968082  
## 2173 0.0005607840 0.0006203078 0.9979339  
## 2174 0.0005607840 0.0006203078 0.9979339  
## 2243 0.0012253070 0.0016161895 0.9980054  
## 2264 0.0010338361 0.0023202871 0.9952323  
## 2284 0.0027486550 0.0029223245 1.0002501  
## 2302 0.0052474851 0.0039383448 1.0036441  
## 2317 0.0027475549 0.0029203730 1.0002497  
## 2360 0.0003972628 0.0007156312 0.9958294  
## 2361 0.0016702786 0.0049312562 0.9938997  
## 2413 0.0002409743 0.0001814295 0.9986009  
## 2469 0.0003097635 0.0005710211 0.9956246  
## 2604 0.0005407977 0.0003612538 0.9991385  
## 2684 0.0005797514 0.0019401085 0.9917105  
## 2685 0.0004464571 0.0016214280 0.9907843  
## 2686 0.0005216928 0.0019297934 0.9906727  
## 2687 0.0004472265 0.0016256079 0.9907765  
## 2688 0.0002716936 0.0010966674 0.9894861  
## 2689 0.0005217339 0.0019301150 0.9906718  
## 2690 0.0005819832 0.0019375432 0.9917607  
## 2699 0.0246048248 0.0622269084 1.0186407  
## 2708 0.0007851159 0.0014801318 0.9959840  
## 2709 0.0007611262 0.0015599238 0.9955021  
## 2710 0.0008422823 0.0017008348 0.9956675  
## 2717 0.0011264174 0.0034084321 0.9931495  
## 2742 0.0005211503 0.0005215287 0.9981887  
## 2852 0.0021166217 0.0191035454 0.9775734  
## 2867 0.0013601799 0.0013270873 0.9990997  
## 2934 0.0005183548 0.0004640763 0.9984806  
## 2937 0.0005182183 0.0004443740 0.9985861  
## 3065 0.0011097506 0.0016295729 0.9974724  
## 3097 0.0009794290 0.0028250304 0.9933959  
## 3102 0.0004201136 0.0007696545 0.9957669  
## 3110 0.0008534344 0.0010847909 0.9977664  
## 3132 0.0024658185 0.0029989868 0.9995381  
## 3168 0.0005555610 0.0047193660 0.9774642  
## 3169 0.0004366179 0.0034741205 0.9788230  
## 3170 0.0002324743 0.0017256913 0.9800870  
## 3171 0.0002261272 0.0016880616 0.9799652  
## 3172 0.0002256373 0.0016831338 0.9799802  
## 3173 0.0002256300 0.0016830591 0.9799804  
## 3174 0.0002257001 0.0016837730 0.9799782  
## 3175 0.0002258708 0.0016854984 0.9799728  
## 3176 0.0002331917 0.0017333821 0.9800598  
## 3177 0.0002257062 0.0016838344 0.9799780  
## 3178 0.0002255113 0.0016818444 0.9799843  
## 3179 0.0002253562 0.0016802415 0.9799896  
## 3180 0.0002252722 0.0016793663 0.9799926  
## 3181 0.0002254443 0.0016811542 0.9799866  
## 3182 0.0005555681 0.0047188369 0.9774671  
## 3183 0.0005555564 0.0047198076 0.9774618  
## 3184 0.0005555564 0.0047198076 0.9774618  
## 3185 0.0005555690 0.0047187757 0.9774675  
## 3186 0.0005555610 0.0047193660 0.9774642  
## 3187 0.0005555610 0.0047193660 0.9774642  
## 3188 0.0004359514 0.0034655328 0.9788431  
## 3189 0.0002259966 0.0016867600 0.9799690  
## 3190 0.0002260224 0.0016870183 0.9799683  
## 3191 0.0002260484 0.0016872776 0.9799675  
## 3192 0.0002260745 0.0016875379 0.9799667  
## 3193 0.0002261008 0.0016877992 0.9799660  
## 3194 0.0002261537 0.0016883250 0.9799644  
## 3195 0.0002256081 0.0016828355 0.9799812  
## 3196 0.0002254169 0.0016808707 0.9799875  
## 3197 0.0002323068 0.0017238024 0.9800945  
## 3198 0.0005555874 0.0047178055 0.9774731  
## 3199 0.0005555855 0.0047178886 0.9774726  
## 3200 0.0005555564 0.0047198076 0.9774618  
## 3201 0.0005555564 0.0047198076 0.9774618  
## 3202 0.0005555564 0.0047198076 0.9774618  
## 3260 0.0005878837 0.0004590371 0.9988702  
## 3424 0.0004058919 0.0003292596 0.9986030  
## 3464 0.0002158960 0.0015774214 0.9803919  
## 3465 0.0002232710 0.0016251032 0.9804757  
## 3466 0.0002156157 0.0015746986 0.9804002  
## 3467 0.0002151766 0.0015703489 0.9804144  
## 3468 0.0002157575 0.0015760810 0.9803959  
## 3469 0.0002157654 0.0015761577 0.9803957  
## 3470 0.0002157408 0.0015759184 0.9803964  
## 3471 0.0002231585 0.0016239867 0.9804793  
## 3472 0.0002158975 0.0015774361 0.9803918  
## 3473 0.0002156459 0.0015749942 0.9803993  
## 3474 0.0002154559 0.0015731285 0.9804052  
## 3475 0.0002154004 0.0015725800 0.9804070  
## 3476 0.0004507642 0.0035248287 0.9792145  
## 3477 0.0002155586 0.0015741396 0.9804020  
## 3478 0.0005456847 0.0045437626 0.9779159  
## 3479 0.0005456847 0.0045437626 0.9779159  
## 3480 0.0005456847 0.0045437626 0.9779159  
## 3481 0.0005456847 0.0045437626 0.9779159  
## 3482 0.0005456847 0.0045437626 0.9779159  
## 3483 0.0005456847 0.0045437596 0.9779159  
## 3484 0.0004273208 0.0033306748 0.9792607  
## 3485 0.0002312410 0.0016763822 0.9805639  
## 3486 0.0002158265 0.0015767501 0.9803939  
## 3487 0.0002230928 0.0016233309 0.9804814  
## 3488 0.0002159612 0.0015780504 0.9803900  
## 3489 0.0002161248 0.0015796188 0.9803854  
## 3490 0.0002156353 0.0015748908 0.9803996  
## 3491 0.0005846004 0.0042948899 0.9806500  
## 3492 0.0002232523 0.0016249179 0.9804763  
## 3493 0.0002155158 0.0015737192 0.9804033  
## 3494 0.0002157241 0.0015757563 0.9803969  
## 3495 0.0005456847 0.0045437596 0.9779159  
## 3496 0.0005456847 0.0045437596 0.9779159  
## 3497 0.0005456847 0.0045437596 0.9779159  
## 3523 0.0013578184 0.0031268443 0.9953943  
## 3571 0.0038926404 0.0027858431 1.0023735  
## 3837 0.0004959751 0.0003312162 0.9990942  
## 4055 0.0003442884 0.0007995897 0.9943240  
## 4056 0.0016789969 0.0071206963 0.9903291  
## 4248 0.0011756093 0.0008883198 0.9995297  
## 4285 0.0005017686 0.0003374944 0.9990865  
## 4391 0.0020146328 0.0013476786 1.0006142  
## 4648 0.0102174793 0.1232363526 0.9774789  
## 4649 0.0414806418 0.4546876421 1.0134346  
## 4665 0.0005021979 0.0003674980 0.9989215  
## 4695 0.0004457745 0.0017094500 0.9902200  
## 4696 0.0003609534 0.0009098129 0.9937888  
## 4713 0.0001698491 0.0001490736 0.9981809  
## 4740 0.0003192908 0.0003158073 0.9980191  
## 4748 0.0016333786 0.0011428229 1.0001454  
## 4750 0.0023827878 0.0032904524 0.9989947  
## 4754 0.0003725971 0.0003385841 0.9982971  
## 4821 0.0007477789 0.0009438242 0.9976854  
## 4833 0.0002207539 0.0001798792 0.9984075  
## 4834 0.0015540252 0.0025046644 0.9975181  
## 4840 0.0083045398 0.0281056016 0.9994025  
## 4846 0.0005193829 0.0003599260 0.9990472  
## 4934 0.0005613334 0.0010155815 0.9959718  
## 4935 0.0004129153 0.0004262980 0.9979918  
## 4942 0.0019547474 0.0013837563 1.0004451  
## 4971 0.0001658414 0.0001326734 0.9983940  
## 4972 0.0002662028 0.0001911296 0.9987238  
## 5083 0.0299513709 0.1229589557 1.0199121  
## 5084 0.0301707569 0.0501679806 1.0269420  
## 5195 0.0018651946 0.0013020471 1.0003826  
## 5491 0.0002056055 0.0001433412 0.9987214  
## 5494 0.0002058396 0.0001437880 0.9987177  
## 5495 0.0002060122 0.0001440963 0.9987154  
## 5496 0.0002056477 0.0001434248 0.9987207  
## 5498 0.0002178888 0.0001510659 0.9987444  
## 5559 0.0010088933 0.0008798565 0.9990369  
## 5924 0.0011260005 0.0010583894 0.9989647  
## 6055 0.0013303597 0.0011987266 0.9992784  
## 6230 0.0002146358 0.0011203995 0.9861495  
## 6231 0.0002394908 0.0012391478 0.9863014  
## 6232 0.0003614390 0.0017261215 0.9875265  
## 6233 0.0002266659 0.0011778349 0.9862270  
## 6234 0.0004596329 0.0022216207 0.9874648  
## 6235 0.0004595805 0.0022211000 0.9874663  
## 6236 0.0004595977 0.0022212718 0.9874658  
## 6237 0.0004734968 0.0022780189 0.9875407  
## 6238 0.0002266863 0.0011780042 0.9862263  
## 6239 0.0002395141 0.0012393505 0.9863005  
## 6429 0.0002144504 0.0032856752 0.9585272  
## 6430 0.0001930508 0.0029759671 0.9582516  
## 6431 0.0002147353 0.0032915735 0.9585082  
## 6432 0.0002139205 0.0032738162 0.9585739  
## 6433 0.0004339510 0.0064008819 0.9602911  
## 6434 0.0004595946 0.0067435556 0.9605263  
## 6435 0.0004849827 0.0067542550 0.9625725  
## 6436 0.0002135780 0.0032611361 0.9586677  
## 6437 0.0001932286 0.0029792957 0.9582435  
## 6438 0.0001902800 0.0030573588 0.9564881  
## 6439 0.0002033302 0.0031247763 0.9583901  
## 6440 0.0003588024 0.0052395174 0.9606150  
## 6441 0.0003609240 0.0052769418 0.9605688  
## 6442 0.0002149138 0.0032951764 0.9584975  
## 6443 0.0004895617 0.0059110679 0.9676106  
## 6444 0.0004901377 0.0059199934 0.9676002  
## 6445 0.0003632573 0.0045926027 0.9659369  
## 6446 0.0003638667 0.0046020199 0.9659247  
## 6447 0.0004888928 0.0059006394 0.9676231  
## 6448 0.0002032801 0.0027196605 0.9637771  
## 6449 0.0004589883 0.0058382783 0.9658232  
## 6450 0.0004592863 0.0058447807 0.9658075  
## 6451 0.0003609328 0.0045776926 0.9658255  
## 6452 0.0003640096 0.0046042202 0.9659219  
## 6453 0.0004928379 0.0059281098 0.9677383  
## 6454 0.0004895326 0.0059106167 0.9676111  
## 6455 0.0004876139 0.0058804737 0.9676482  
## 6456 0.0002142928 0.0028569984 0.9639148  
## 6457 0.0004596308 0.0058519863 0.9657911  
## 6487 0.0071748274 0.0062575832 1.0052564  
## 6512 0.0004559806 0.0015914020 0.9911875  
## 6587 0.0001907060 0.0002367824 0.9971850  
## 6664 0.0001944379 0.0001413293 0.9986272  
## 6739 0.0012423390 0.0011893132 0.9990328  
## 6766 0.0071774365 0.0082880193 1.0044695  
## 6781 0.0002640093 0.0002048821 0.9985593  
## 6852 0.0011317060 0.0011923513 0.9986529  
## 6853 0.0011317060 0.0011923513 0.9986529  
## 6938 0.0002303781 0.0006158175 0.9932336  
## 6939 0.0003668203 0.0008719249 0.9941946  
## 6940 0.0002303957 0.0006159103 0.9932330  
## 6941 0.0002299256 0.0006131959 0.9932502  
## 6942 0.0004952576 0.0010478515 0.9950510  
## 6943 0.0010433430 0.0019361945 0.9963251  
## 6944 0.0004748981 0.0011385984 0.9942455  
## 6945 0.0002303781 0.0006158175 0.9932336  
## 6946 0.0002303328 0.0006155754 0.9932350  
## 6947 0.0002302838 0.0006153096 0.9932365  
## 6948 0.0002300293 0.0006138476 0.9932458  
## 7039 0.0009453543 0.0012860652 0.9976091  
## 7167 0.0002434550 0.0019003255 0.9790455  
## 7210 0.0012442375 0.0037559951 0.9932877  
## 7211 0.0009420382 0.0026526317 0.9935490  
## 7446 0.0002358834 0.0019871387 0.9773379  
## 7447 0.0002357940 0.0019860468 0.9773418  
## 7448 0.0002357257 0.0019852060 0.9773448  
## 7449 0.0002214814 0.0019792642 0.9759176  
## 7450 0.0002214846 0.0019793051 0.9759174  
## 7451 0.0002214817 0.0019792687 0.9759176  
## 7452 0.0002357040 0.0019849364 0.9773458  
## 7453 0.0002236478 0.0018900728 0.9772519  
## 7454 0.0002236178 0.0018897155 0.9772531  
## 7455 0.0002236071 0.0018895879 0.9772536  
## 7456 0.0002358074 0.0019862114 0.9773412  
## 7457 0.0002359382 0.0019878022 0.9773356  
## 7458 0.0002212037 0.0019756752 0.9759310  
## 7459 0.0002359762 0.0019882603 0.9773340  
## 7460 0.0002214846 0.0019793051 0.9759174  
## 7461 0.0002359762 0.0019882603 0.9773340  
## 7462 0.0002359750 0.0019882463 0.9773341  
## 7463 0.0002359700 0.0019881859 0.9773343  
## 7629 0.0042573542 0.0064799965 1.0004886  
## 7683 0.0001873752 0.0010762093 0.9846743  
## 7686 0.0011424530 0.0008222719 0.9995968  
## 7722 0.0001879320 0.0006369014 0.9912000  
## 7871 0.0025648326 0.0021987691 1.0006406  
## 8154 0.0004052247 0.0002724761 0.9989904  
## 8320 0.0047877723 0.0040774132 1.0028968  
## 8377 0.0644314545 0.4675577955 1.0492027  
## 8457 0.0002360096 0.0003357145 0.9967253  
## 8458 0.0007875662 0.0014606841 0.9960717  
## 8541 0.0005082410 0.0004811519 0.9983268  
## 8698 0.0002675410 0.0002445494 0.9981770  
## 8710 0.0005223815 0.0021212863 0.9896694  
## 8717 0.0020381987 0.0022567895 0.9994127  
## 8763 0.0004868252 0.0005184649 0.9979748  
## 8887 0.0023137448 0.0144124894 0.9854659  
## 8911 0.0011578987 0.0015897966 0.9977870  
## 8946 0.0011586560 0.0015903616 0.9977889  
## 9369 0.0020771918 0.0017874236 1.0001414  
## 9420 0.0072614289 0.0050586051 1.0058350  
## 9488 0.0013090257 0.0010515853 0.9995301  
## 9507 0.0042781301 0.0061749754 1.0007292  
## 9527 0.0017562433 0.0011998565 1.0003148  
## 9528 0.0012093951 0.0015893955 0.9980028  
## 9546 0.0005060288 0.0003497812 0.9990387  
## 9722 0.0003753690 0.0005768604 0.9965457  
## 9730 0.0016543867 0.0021995448 0.9984065  
## 10072 0.0005165455 0.0005464067 0.9980246  
## 10125 0.0008888267 0.0015685551 0.9964240  
## 10131 0.0011825852 0.0010543129 0.9991568  
## 10624 0.0012064862 0.0013581442 0.9985264  
## 10723 0.0004586197 0.0005101516 0.9978142  
## 10787 0.0009549105 0.0028043141 0.9932256  
## 10844 0.0021862012 0.0016077895 1.0006006  
## 10958 0.0006378172 0.0006345619 0.9983219  
## 10993 0.0007332822 0.0006284844 0.9988027  
## 11165 0.0013549599 0.0015279284 0.9986699  
## 11289 0.0002836585 0.0008824799 0.9920684  
## 11413 0.0013090256 0.0034854493 0.9943434  
## 11558 0.0009069575 0.0036834854 0.9900532  
## 11586 0.0010525063 0.0021250965 0.9958789  
## 11668 0.0002476214 0.0002408568 0.9979932  
## 11728 0.0048728629 0.0058887027 1.0019857  
## 11747 0.0016601395 0.0015220016 0.9995653  
## 11772 0.0002757694 0.0028530885 0.9721122  
## 11822 0.0003039847 0.0016453935 0.9857057  
## 11982 0.0025363021 0.0110736527 0.9908437  
## 11992 0.0043205400 0.0630249924 0.9646365  
## 12098 0.0012951683 0.0018266833 0.9978203  
## 12212 0.0018625522 0.0015689149 0.9999766  
## 12255 0.0016006310 0.0022807405 0.9980863  
## 12256 0.0012694921 0.0010888228 0.9993385  
## 12293 0.0004667657 0.0004441860 0.9982715  
## 12370 0.0004426869 0.0003117059 0.9989393  
## 12438 0.0005563753 0.0003896398 0.9990637  
## 12472 0.0016942000 0.0062641342 0.9918530  
## 12487 0.0005821916 0.0019781907 0.9915701  
## 12528 0.0004298663 0.0003452426 0.9986496  
## 12543 0.0007297494 0.0007410401 0.9983565  
## 12577 0.0014422708 0.0054561121 0.9913625  
## 12582 0.0009639757 0.0111479902 0.9694742  
## 12643 0.0004727347 0.0052458425 0.9702582  
## 12759 0.0003167519 0.0003313726 0.9978571  
## 12764 0.0016126691 0.0063746018 0.9910621  
## 12808 0.0005652698 0.0004268031 0.9989196  
## 12854 0.0004519781 0.0004502024 0.9981325

* No problematics observed.

## Perform the necessary calculations to assess the assumption of independence and state if the condition is met or not.

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:purrr':  
##   
## some

## The following object is masked from 'package:boot':  
##   
## logit

## lag Autocorrelation D-W Statistic p-value  
## 1 0.6422763 0.7154469 0  
## Alternative hypothesis: rho != 0

From the output, we can see that the test statistic is 0.7154469 and the corresponding p-value is 0 which is less than 0.05, we can reject the null hypothesis and conclude that the residuals in this model are auto correlated.

## Perform the necessary calculations to assess the assumption of no multicollinearity and state if the condition is met or not.

## `Sale Date` sq\_ft\_lot bedrooms bath\_full\_count year\_built   
## 1.003512 1.034197 1.130426 1.435768 1.324972

The VIF values are less than 2, hence condition is met for the assumption of no multicollinearity.

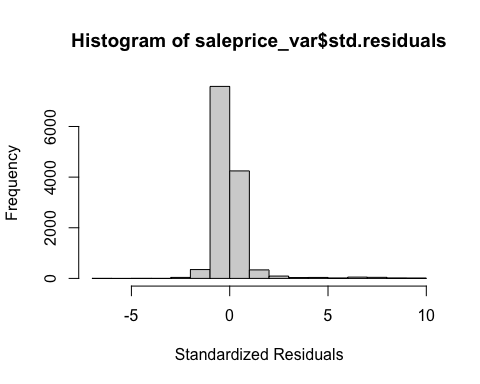
## Visually check the assumptions related to the residuals using the plot() and hist() functions. Summarize what each graph is informing you of and if any anomalies are present.

## Warning in plot.window(xlim, ylim, "", ...): "scale" is not a graphical  
## parameter

## Warning in title(main = main, sub = sub, xlab = xlab, ylab = ylab, ...): "scale"  
## is not a graphical parameter

## Warning in axis(1, ...): "scale" is not a graphical parameter

## Warning in axis(2, ...): "scale" is not a graphical parameter

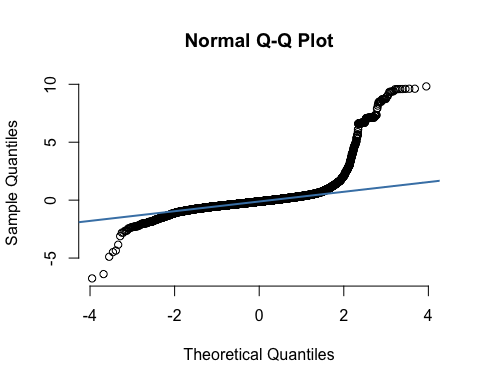
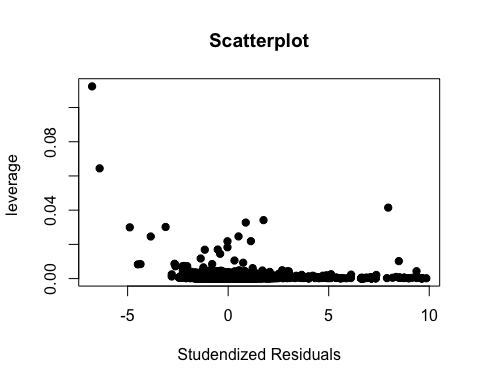
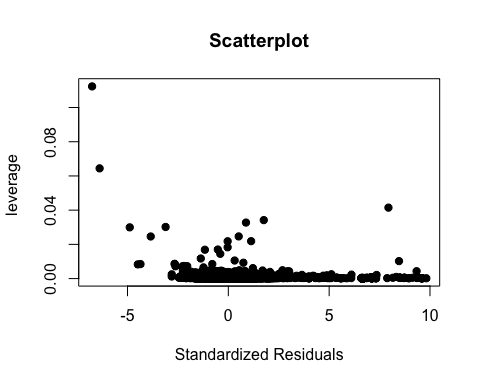
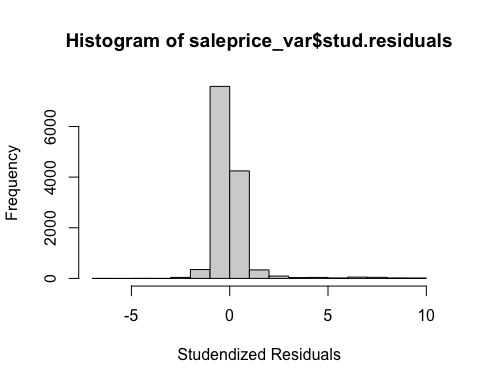


## Warning in plot.window(xlim, ylim, "", ...): "scale" is not a graphical  
## parameter

## Warning in title(main = main, sub = sub, xlab = xlab, ylab = ylab, ...): "scale"  
## is not a graphical parameter

## Warning in axis(1, ...): "scale" is not a graphical parameter

## Warning in axis(2, ...): "scale" is not a graphical parameter



As per the plots, no anomalies are observed.

## Overall, is this regression model unbiased? If an unbiased regression model, what does this tell us about the sample vs. the entire population model?

This regression model appears unbaised.