

Single Family Residential Regression Results

For ECONorthwest

Salma Huque

04/01/20

Import Packages + Read in Data

```
library(tidyverse)
library(magrittr)
library(here)
library(data.table)
library(MASS)
library(xtable)
library(knitr)
library(stargazer)
library(pander)
library(kableExtra)
library(lubridate)

# read in csv
thesis_data <- read_csv(here::here("DATA", "thesis-data.csv"))
```

Cleaning

This chunk cleans the single-family residential data. It involves selecting the appropriate variables, and flagging certain variables in order to determine which transactions can be considered “arms length”. This chunk also summarizes the multiple variables originally created for basements, attics, garages, etc. (the original dataset had multiple types of each) and creates a numeric and dummy version of each.

```
# Subset Data by Single Family Residential
SFR <- thesis_data %>%
  filter(PRPCD_DESC == "RESIDENTIAL IMPROVED", prop_type == "Single-family")

# Clean SFR Dataframe
test1 <- SFR %>%
  dplyr::select(-c(OWNER2, OWNER3, OWNERZIP,
    MKTVALYR1, MKTVALYR2,
    BLDGVAL1, BLDGVAL2,
    LANDVAL1, LANDVAL2,
    TOTALVAL1, TOTALVAL2,
    MS_GRADE, ES_GRADE,
    LEGAL_DESC, TAXCODE, PROP_CODE, LANDUSE, BEDROOMS, ACC_STATUS, NAME, COMMPLAN, SHARED, COAL),
  mutate(top_1 = SALEPRICE > quantile(SALEPRICE, .99),
```

```

MKTVALYR3 = case_when(MKTVALYR3 != 2018 ~ 2017,
                      TRUE ~ 2018),

price_diff = SALEPRICE - LANDVAL3,
price_ratio = SALEPRICE/LANDVAL3 * 100,
vacant_dummy = PRPCD_DESC == "VACANT LAND",
llc_flag = grepl("LLC", OWNER1),
proud_flag = grepl("PROUD", OWNER1),
trust_flag = grepl("TRUST", OWNER1) & !grepl("FAMILY", OWNER1) & !grepl("LIVING", OWNER1)) %>%
mutate(arms_length = price_ratio > 20,
       na_if(yearbuilt, 0))

# switch the NAs in the constraints to 0s
to0 <- function(x){ifelse(is.na(x), 0, x)}

# Make a string of constraints and percent canopy names called "con_names"
clean <- SFR %>%
  dplyr::select(contains("con"), contains("pct")) %>%
  dplyr::select(-c("CONCRETE", "FIN SECOND", "PAVING/CONCRETE ONLY", "UNF SECOND", "conFldway", "pct_con"))
clean_names <- names(clean)
con_names <- paste(clean_names, collapse = " + ")

# removing observations
dat1 <- test1 %>% #25,266
  mutate_at(vars(clean_names), to0) %>%
  filter(top_1 == FALSE, #25,013, #253
         arms_length == TRUE, #24,495 #472
         vacant_dummy == FALSE, #24,495 #0
         proud_flag == FALSE, #24,480 #15
         llc_flag == FALSE, #24,128 #413
         trust_flag == FALSE #24,081 #52
  )

#garage sqft
gar_sqft_sum <- dat1 %>%
  dplyr::select(matches("gar")) %>%
  rowSums()

# attic sqft
attic_sqft_sum <- dat1 %>%
  dplyr::select(matches("attic")) %>%
  rowSums()

# basement sqft
bsmt_sqft_sum <- dat1 %>%
  dplyr::select(matches("bsmt")) %>%
  dplyr::select(-c("BSMT PARKING", "BSMT GAR")) %>%
  rowSums()

dat1 %<>%
  mutate(garage_sqft = gar_sqft_sum,
         garage_dum = garage_sqft > 0,
         attic_sqft = attic_sqft_sum,
         attic_dum = attic_sqft > 0,

```

```

    bsmt_sqft = bsmt_sqft_sum,
    bsmt_dum = bsmt_sqft > 0,
    saledate = mdy(saledate),
    year_sold = year(saledate)) %>%
filter(between(totalsqft, 1, 7500),
       yearbuilt > 1500)

```

Exploration

This chunk looks at some of the control variables in order to determine any necessary changes to functional form, and to further inform cleaning.

```

# A Look at the Dataset-----

salma <- summary(dat1) # notes in thesis notebook

controls <- dat1 %>%
  dplyr::select(price_diff, pct_canopy_cov, taxlot_area, yearbuilt, totalsqft, dist_ugb, dist_cityhall,
    SALEPRICE)
nm <- names(controls)

# Correlation-----

correlation <- cor(controls, use = "pairwise.complete.obs")
#takes only obs for which both vars are not NA

# Scatterplots-----

for (i in seq_along(nm)) {
  print(names(controls[i]))
  print(ggplot(controls, aes_string(y = nm[length(nm)], x = nm[i])) +
    geom_point(alpha = 0.1) +
    geom_smooth())
}

ggplot(dat1, aes(y = SALEPRICE, x = TOTALVAL3)) +
  geom_point(alpha = 0.1)

ggplot(dat1, aes(y = log(SALEPRICE), x = taxlot_area)) +
  geom_point(alpha = 0.1)

# Outliers -- Box Plots-----

# best for discrete x

for (i in seq_along(nm)) {
  print(names(controls[i]))
  print(ggplot(controls, aes_string(y = nm[length(nm)], x = nm[i])) +
    geom_boxplot(alpha = 0.1))
}

# Distribution -- Density Plots-----

```

```
# Now to check the linearity of the variables it is a good practice to plot distribution graph and look

for (i in seq_along(nm)) {
  print(names(controls[i]))
  print(ggplot(controls, aes_string(x = nm[i])) +
    geom_histogram(bins = 50))
}
```

Regressions

I have included two of my preliminary regressions here: a linear regression and a Box-Cox regression. Apologies that the results tables are so long! I had both neighborhood and high school districting fixed effects.

```
# LINEAR REGRESSION-----

# formula: only dummies, no percent except for percent canopy cover
sfr_lin = formula(paste0("SALEPRICE ~
                          CN_score + percent_vacant + dist_cityhall + dist_ugb + totalsqft + yearbuilt"))

# reg
sfr_lin_mod <- lm(sfr_lin, dat1)
summary(sfr_lin_mod)
```

```
##
## Call:
## lm(formula = sfr_lin, data = dat1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -659207  -49654    1771   49929   648721
##
## Coefficients: (6 not defined because of singularities)
##                                     Estimate Std. Error
## (Intercept)                    -5.488e+07   1.418e+06
## CN_score                       -3.266e+02   1.058e+02
## percent_vacant                   1.577e+02   3.277e+03
## dist_cityhall                   -5.649e+00   4.095e-01
## dist_ugb                        -1.235e+00   5.077e-01
## totalsqft                       6.042e+01   1.339e+00
## yearbuilt                       -4.210e+02   4.311e+01
## f_baths                         6.134e+04   1.382e+03
## h_baths                         4.399e+04   1.746e+03
## n_fireplaces                    -5.666e+03   4.249e+03
## ADUdummy                       -3.565e+04   7.083e+03
## pct_canopy_cov                  -1.389e+04   4.830e+03
## taxlot_area                     1.754e+00   1.230e-01
## garage_dumTRUE                   1.264e+04   2.294e+03
## attic_dumTRUE                    -5.226e+03   2.097e+03
## bsmt_dumTRUE                     -8.739e+03   2.175e+03
## year_sold                       2.788e+04   7.009e+02
## MapLabelArbor Lodge             -1.676e+04   2.202e+04
```

## MapLabelArdenwald-Johnson Creek	-1.337e+05	2.500e+04
## MapLabelArgay Terrace	-1.944e+05	2.087e+04
## MapLabelArlington Heights	-3.196e+04	3.007e+04
## MapLabelArnold Creek	-1.220e+05	1.946e+04
## MapLabelAshcreek	-1.599e+05	1.890e+04
## MapLabelBeaumont-Wilshire	-1.055e+04	9.858e+03
## MapLabelBoise	-1.848e+04	3.079e+04
## MapLabelBrentwood-Darlington	-1.643e+05	1.570e+04
## MapLabelBridlemile	-1.783e+05	2.528e+04
## MapLabelBrooklyn Action Corps	-1.338e+05	1.854e+04
## MapLabelBuckman Community Association	-5.134e+04	2.172e+04
## MapLabelCathedral Park	1.545e+04	2.852e+04
## MapLabelCentennial Community Association	-1.661e+05	1.920e+04
## MapLabelCollins View	-1.605e+05	2.016e+04
## MapLabelConcordia	-3.793e+04	1.303e+04
## MapLabelCreston-Kenilworth	-1.152e+05	1.572e+04
## MapLabelCrestwood	-1.804e+05	1.988e+04
## MapLabelCully Association of Neighbors	-1.165e+05	1.225e+04
## MapLabelEast Columbia	-1.628e+05	2.684e+04
## MapLabelEastmoreland	7.211e+04	1.622e+04
## MapLabelEliot	-1.547e+05	4.740e+04
## MapLabelFar Southwest	-1.264e+05	2.058e+04
## MapLabelForest Park	-6.744e+04	2.800e+04
## MapLabelFoster-Powell	-1.174e+05	1.610e+04
## MapLabelGlenfair	-1.772e+05	2.499e+04
## MapLabelGoose Hollow Foothills League	-8.679e+04	3.414e+04
## MapLabelGrant Park	4.908e+04	1.136e+04
## MapLabelHayden Island Neighborhood Network	5.063e+04	3.602e+04
## MapLabelHayhurst	-1.664e+05	1.892e+04
## MapLabelHazelwood	-1.622e+05	1.645e+04
## MapLabelHealy Heights	-7.397e+03	4.050e+04
## MapLabelHillsdale	-1.412e+05	1.905e+04
## MapLabelHillside	-3.943e+04	2.810e+04
## MapLabelHollywood	3.777e+04	2.416e+04
## MapLabelHomestead	-2.133e+05	3.315e+04
## MapLabelHosford-Abernethy Neighborhood District Assn.	-2.634e+04	1.586e+04
## MapLabelHumboldt	-1.489e+04	2.190e+04
## MapLabelIrvington Community Association	2.725e+04	1.185e+04
## MapLabelKenton	-3.885e+04	2.472e+04
## MapLabelKerns	-2.876e+04	2.692e+04
## MapLabelKing	-8.181e+04	1.740e+04
## MapLabelLaurelhurst	4.035e+04	1.009e+04
## MapLabelLents	-1.639e+05	1.518e+04
## MapLabelLinnton	-1.812e+05	3.228e+04
## MapLabelMadison South	-1.197e+05	1.311e+04
## MapLabelMaplewood	-1.355e+05	1.979e+04
## MapLabelMarkham	-1.907e+05	1.946e+04
## MapLabelMarshall Park	-1.451e+05	2.106e+04
## MapLabelMC Unclaimed #13	-1.428e+05	9.326e+04
## MapLabelMill Park	-1.683e+05	1.722e+04
## MapLabelMontavilla	-9.357e+04	1.255e+04
## MapLabelMt. Scott-Arleta	-1.126e+05	1.602e+04
## MapLabelMt. Tabor	1.261e+04	1.358e+04
## MapLabelMultnomah	-1.445e+05	1.894e+04

## MapLabelNorth Tabor	-6.668e+04	1.442e+04
## MapLabelNorthwest District Association	7.639e+04	2.924e+04
## MapLabelNorthwest Heights	-4.392e+04	2.592e+04
## MapLabelOverlook	3.727e+03	2.274e+04
## MapLabelParkrose	-1.974e+05	2.131e+04
## MapLabelParkrose Heights Association of Neighbors	-1.934e+05	2.064e+04
## MapLabelPiedmont	-3.993e+04	2.165e+04
## MapLabelPleasant Valley	-1.509e+05	1.834e+04
## MapLabelPortsmouth	-3.694e+04	2.596e+04
## MapLabelPowellhurst-Gilbert	-1.613e+05	1.756e+04
## MapLabelReed	-6.923e+04	1.810e+04
## MapLabelRichmond	-2.771e+04	1.418e+04
## MapLabelRose City Park	-4.419e+04	9.576e+03
## MapLabelRoseway	-7.216e+04	1.216e+04
## MapLabelRussell	-1.724e+05	2.078e+04
## MapLabelSabin Community Association	-8.659e+03	1.080e+04
## MapLabelSellwood-Moreland Improvement League	-2.197e+04	1.541e+04
## MapLabelSouth Burlingame	-1.626e+05	2.025e+04
## MapLabelSouth Portland	-7.141e+04	2.218e+04
## MapLabelSouth Tabor	-7.214e+04	1.503e+04
## MapLabelSouthwest Hills Residential League	-5.751e+04	2.535e+04
## MapLabelSt. Johns	-2.569e+04	2.600e+04
## MapLabelSullivan's Gulch	-1.070e+05	2.314e+04
## MapLabelSumner Association of Neighbors	-1.870e+05	2.399e+04
## MapLabelSunderland Association of Neighbors	-1.420e+05	6.142e+04
## MapLabelSunnyside	-1.677e+04	1.598e+04
## MapLabelSylvan-Highlands	-1.484e+05	2.833e+04
## MapLabelUniversity Park	-1.635e+03	2.603e+04
## MapLabelVernon	-4.451e+03	1.918e+04
## MapLabelWest Portland Park	-1.794e+05	1.936e+04
## MapLabelWilkes Community Group	-1.801e+05	2.183e+04
## MapLabelWoodland Park	-1.869e+05	2.867e+04
## MapLabelWoodlawn	-4.536e+04	1.974e+04
## MapLabelWoodstock	-7.450e+04	1.627e+04
## HIGH_SCHCLEVELAND HS	-4.179e+04	1.493e+04
## HIGH_SCHDAVID DOUGLAS HS	-3.906e+04	9.155e+03
## HIGH_SCHFRANKLIN HS	-4.744e+04	1.354e+04
## HIGH_SCHGRANT HS	-1.107e+04	1.597e+04
## HIGH_SCHGRANT-JEFF	-7.740e+04	2.373e+04
## HIGH_SCHLINCOLN HS	7.010e+04	1.687e+04
## HIGH_SCHMADISON HS	-8.511e+04	1.477e+04
## HIGH_SCHMADISON-JEFF	-9.257e+04	1.762e+04
## HIGH_SCHPARKROSE HS	-2.820e+04	1.385e+04
## HIGH_SCHREYNOLDS HS	-1.431e+04	1.366e+04
## HIGH_SCHRIVERDALE HS	4.306e+05	7.139e+04
## HIGH_SCHROOS-JEFF	-1.223e+05	2.424e+04
## HIGH_SCHROOSEVELT CAMPUS	-1.343e+05	2.757e+04
## HIGH_SCHSUNSET HS	7.930e+04	3.313e+04
## HIGH_SCHWILSON HS	NA	NA
## conECSI	9.076e+04	1.000e+05
## conLUST	-4.625e+04	7.172e+04
## conHist	3.315e+04	6.568e+03
## conHistLdm	2.155e+04	3.067e+04
## conNatAm	NA	NA

## conAirHgt	1.743e+04	4.247e+03
## conHeliprt	NA	NA
## conNoise	-2.359e+03	9.408e+03
## conGW	4.677e+04	5.339e+04
## conLSHA	1.114e+04	4.037e+03
## conSLID0	-1.874e+04	5.978e+03
## conSlp25	1.047e+04	4.101e+03
## conInstit	NA	NA
## conPrvCom	2.754e+05	1.008e+05
## conPubOwn	-1.241e+04	4.084e+04
## conSewer	1.234e+04	1.301e+04
## conStorm	-3.228e+03	3.177e+03
## conTranCap	-6.411e+03	3.448e+03
## conTranInt	-1.031e+04	3.705e+03
## conTranSub	-9.918e+03	2.174e+03
## conWater	-7.123e+03	4.155e+03
## conView	NA	NA
## conCovrly	-2.129e+03	4.518e+03
## conPovrly	-2.643e+04	6.268e+03
## conWetland	-6.517e+04	2.420e+04
## conFld100	-1.304e+04	1.244e+04
## pct_conFld100	1.876e+04	1.952e+04
## pct_conPovrly	1.750e+04	1.012e+05
## pct_conCovrly	2.937e+04	8.734e+03
## pct_conWetland	NA	NA
##	t value	Pr(> t)
## (Intercept)	-38.708	< 2e-16 ***
## CN_score	-3.086	0.002033 **
## percent_vacant	0.048	0.961626
## dist_cityhall	-13.793	< 2e-16 ***
## dist_ugb	-2.432	0.015032 *
## totalsqft	45.128	< 2e-16 ***
## yearbuilt	-9.765	< 2e-16 ***
## f_baths	44.385	< 2e-16 ***
## h_baths	25.193	< 2e-16 ***
## n_fireplaces	-1.334	0.182331
## ADUdummy	-5.033	4.89e-07 ***
## pct_canopy_cov	-2.875	0.004046 **
## taxlot_area	14.260	< 2e-16 ***
## garage_dumTRUE	5.511	3.62e-08 ***
## attic_dumTRUE	-2.493	0.012690 *
## bsmt_dumTRUE	-4.018	5.90e-05 ***
## year_sold	39.784	< 2e-16 ***
## MapLabelArbor Lodge	-0.761	0.446371
## MapLabelArdenwald-Johnson Creek	-5.350	8.91e-08 ***
## MapLabelArgay Terrace	-9.312	< 2e-16 ***
## MapLabelArlington Heights	-1.063	0.287940
## MapLabelArnold Creek	-6.267	3.76e-10 ***
## MapLabelAshcreek	-8.460	< 2e-16 ***
## MapLabelBeaumont-Wilshire	-1.070	0.284693
## MapLabelBoise	-0.600	0.548356
## MapLabelBrentwood-Darlington	-10.463	< 2e-16 ***
## MapLabelBridlemile	-7.055	1.80e-12 ***
## MapLabelBrooklyn Action Corps	-7.214	5.67e-13 ***

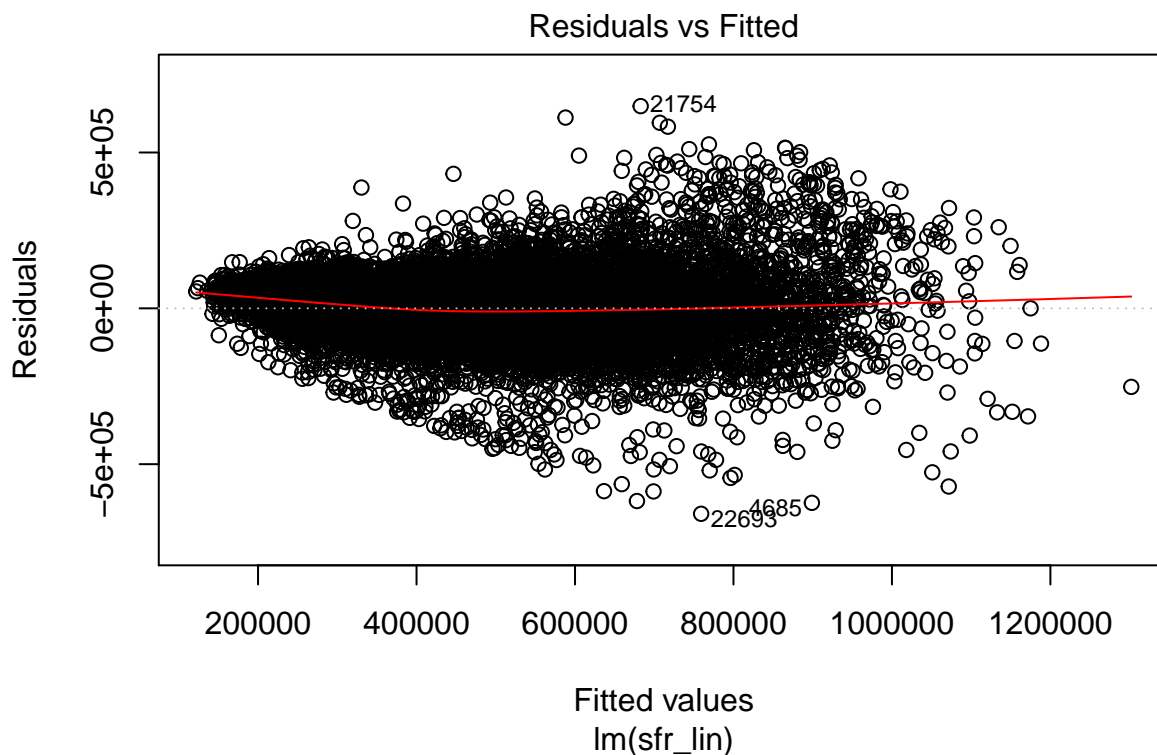
## MapLabelBuckman Community Association	-2.364	0.018101	*
## MapLabelCathedral Park	0.542	0.588163	
## MapLabelCentennial Community Association	-8.653	< 2e-16	***
## MapLabelCollins View	-7.962	1.80e-15	***
## MapLabelConcordia	-2.911	0.003612	**
## MapLabelCreston-Kenilworth	-7.330	2.41e-13	***
## MapLabelCrestwood	-9.071	< 2e-16	***
## MapLabelCully Association of Neighbors	-9.510	< 2e-16	***
## MapLabelEast Columbia	-6.067	1.34e-09	***
## MapLabelEastmoreland	4.445	8.85e-06	***
## MapLabelEliot	-3.264	0.001100	**
## MapLabelFar Southwest	-6.144	8.24e-10	***
## MapLabelForest Park	-2.409	0.016022	*
## MapLabelFoster-Powell	-7.290	3.25e-13	***
## MapLabelGlenfair	-7.092	1.37e-12	***
## MapLabelGoose Hollow Foothills League	-2.542	0.011035	*
## MapLabelGrant Park	4.322	1.56e-05	***
## MapLabelHayden Island Neighborhood Network	1.406	0.159845	
## MapLabelHayhurst	-8.795	< 2e-16	***
## MapLabelHazelwood	-9.860	< 2e-16	***
## MapLabelHealy Heights	-0.183	0.855083	
## MapLabelHillsdale	-7.412	1.31e-13	***
## MapLabelHillside	-1.403	0.160617	
## MapLabelHollywood	1.563	0.118049	
## MapLabelHomestead	-6.434	1.28e-10	***
## MapLabelHosford-Abernethy Neighborhood District Assn.	-1.662	0.096632	.
## MapLabelHumboldt	-0.680	0.496472	
## MapLabelIrvington Community Association	2.300	0.021482	*
## MapLabelKenton	-1.572	0.115974	
## MapLabelKerns	-1.068	0.285508	
## MapLabelKing	-4.701	2.62e-06	***
## MapLabelLaurelhurst	3.998	6.42e-05	***
## MapLabelLents	-10.803	< 2e-16	***
## MapLabelLinnton	-5.612	2.03e-08	***
## MapLabelMadison South	-9.128	< 2e-16	***
## MapLabelMaplewood	-6.846	7.86e-12	***
## MapLabelMarkham	-9.801	< 2e-16	***
## MapLabelMarshall Park	-6.891	5.74e-12	***
## MapLabelMC Unclaimed #13	-1.531	0.125792	
## MapLabelMill Park	-9.774	< 2e-16	***
## MapLabelMontavilla	-7.457	9.27e-14	***
## MapLabelMt. Scott-Arleta	-7.029	2.16e-12	***
## MapLabelMt. Tabor	0.929	0.353136	
## MapLabelMultnomah	-7.628	2.52e-14	***
## MapLabelNorth Tabor	-4.624	3.79e-06	***
## MapLabelNorthwest District Association	2.613	0.008988	**
## MapLabelNorthwest Heights	-1.694	0.090230	.
## MapLabelOverlook	0.164	0.869790	
## MapLabelParkrose	-9.260	< 2e-16	***
## MapLabelParkrose Heights Association of Neighbors	-9.373	< 2e-16	***
## MapLabelPiedmont	-1.845	0.065104	.
## MapLabelPleasant Valley	-8.230	< 2e-16	***
## MapLabelPortsmouth	-1.423	0.154791	
## MapLabelPowellhurst-Gilbert	-9.184	< 2e-16	***

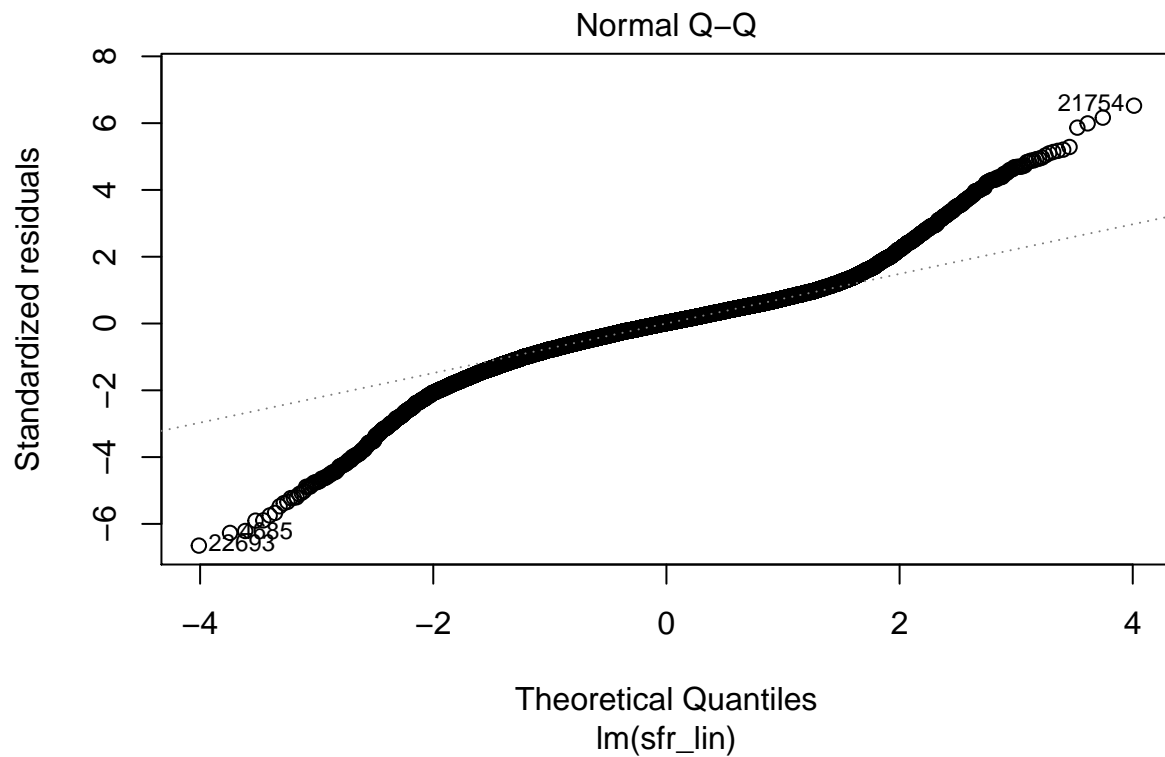
## MapLabelReed	-3.825	0.000131	***
## MapLabelRichmond	-1.955	0.050656	.
## MapLabelRose City Park	-4.615	3.96e-06	***
## MapLabelRoseway	-5.935	3.00e-09	***
## MapLabelRussell	-8.292	< 2e-16	***
## MapLabelSabin Community Association	-0.802	0.422749	
## MapLabelSellwood-Moreland Improvement League	-1.426	0.153888	
## MapLabelSouth Burlingame	-8.027	1.07e-15	***
## MapLabelSouth Portland	-3.219	0.001287	**
## MapLabelSouth Tabor	-4.800	1.60e-06	***
## MapLabelSouthwest Hills Residential League	-2.268	0.023328	*
## MapLabelSt. Johns	-0.988	0.323243	
## MapLabelSullivan's Gulch	-4.623	3.81e-06	***
## MapLabelSumner Association of Neighbors	-7.796	6.79e-15	***
## MapLabelSunderland Association of Neighbors	-2.312	0.020766	*
## MapLabelSunnyside	-1.050	0.293901	
## MapLabelSylvan-Highlands	-5.239	1.63e-07	***
## MapLabelUniversity Park	-0.063	0.949933	
## MapLabelVernon	-0.232	0.816489	
## MapLabelWest Portland Park	-9.264	< 2e-16	***
## MapLabelWilkes Community Group	-8.251	< 2e-16	***
## MapLabelWoodland Park	-6.519	7.31e-11	***
## MapLabelWoodlawn	-2.298	0.021603	*
## MapLabelWoodstock	-4.578	4.73e-06	***
## HIGH_SCHCLEVELAND HS	-2.799	0.005126	**
## HIGH_SCHDAVID DOUGLAS HS	-4.267	1.99e-05	***
## HIGH_SCHFRANKLIN HS	-3.504	0.000460	***
## HIGH_SCHGRANT HS	-0.693	0.488442	
## HIGH_SCHGRANT-JEFF	-3.261	0.001112	**
## HIGH_SCHLINCOLN HS	4.155	3.27e-05	***
## HIGH_SCHMADISON HS	-5.764	8.39e-09	***
## HIGH_SCHMADISON-JEFF	-5.252	1.52e-07	***
## HIGH_SCHPARKROSE HS	-2.036	0.041741	*
## HIGH_SCHREYNOLDS HS	-1.048	0.294866	
## HIGH_SCHRIVERDALE HS	6.031	1.66e-09	***
## HIGH_SCHROOS-JEFF	-5.046	4.56e-07	***
## HIGH_SCHROOSEVELT CAMPUS	-4.869	1.13e-06	***
## HIGH_SCHSUNSET HS	2.394	0.016677	*
## HIGH_SCHWILSON HS	NA	NA	
## conECSEI	0.907	0.364208	
## conLUST	-0.645	0.519000	
## conHist	5.048	4.52e-07	***
## conHistLdm	0.703	0.482252	
## conNatAm	NA	NA	
## conAirHgt	4.104	4.08e-05	***
## conHeliprt	NA	NA	
## conNoise	-0.251	0.802048	
## conGW	0.876	0.381100	
## conLSHA	2.759	0.005807	**
## conSLID0	-3.134	0.001725	**
## conSlp25	2.552	0.010710	*
## conInstit	NA	NA	
## conPrvCom	2.733	0.006288	**
## conPubOwn	-0.304	0.761312	

```
## conSewer                0.948 0.342941
## conStorm                -1.016 0.309687
## conTranCap              -1.859 0.063019 .
## conTranInt              -2.782 0.005413 **
## conTranSub              -4.561 5.12e-06 ***
## conWater                -1.714 0.086468 .
## conView                 NA      NA
## conCovrly               -0.471 0.637452
## conPovrly               -4.217 2.48e-05 ***
## conWetland              -2.693 0.007097 **
## conFld100               -1.049 0.294256
## pct_conFld100           0.961 0.336493
## pct_conPovrly           0.173 0.862766
## pct_conCovrly           3.363 0.000772 ***
## pct_conWetland          NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 99780 on 16383 degrees of freedom
## (6539 observations deleted due to missingness)
## Multiple R-squared:  0.7585, Adjusted R-squared:  0.7563
## F-statistic: 357.3 on 144 and 16383 DF, p-value: < 2.2e-16
```

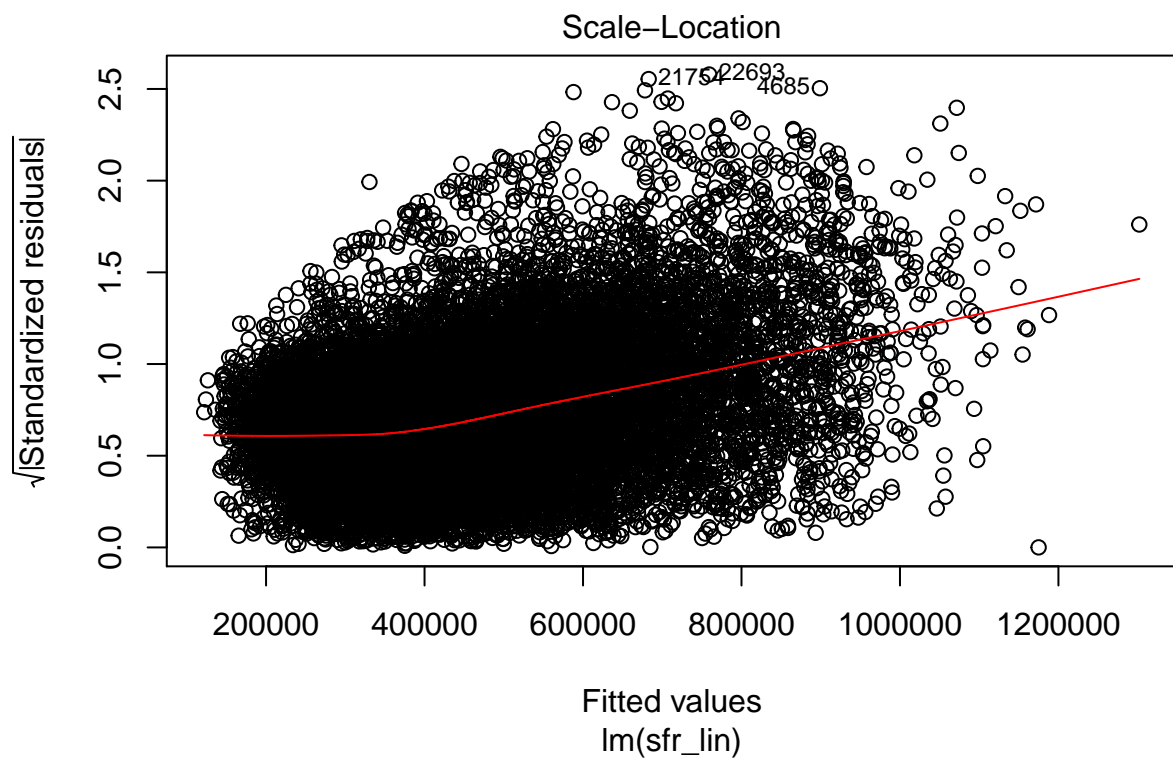
```
plot(sfr_lin_mod)
```

```
## Warning: not plotting observations with leverage one:
## 12269, 14945
```



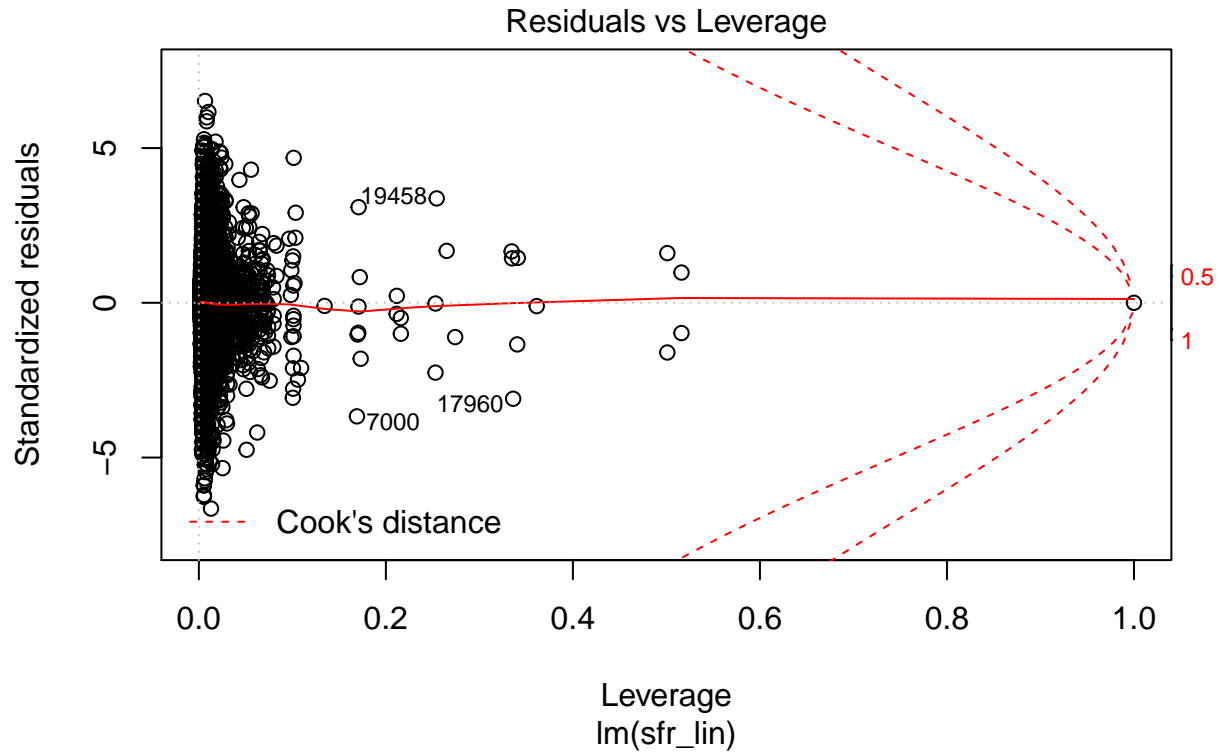


```
## Warning: not plotting observations with leverage one:
## 12269, 14945
```



```
## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```

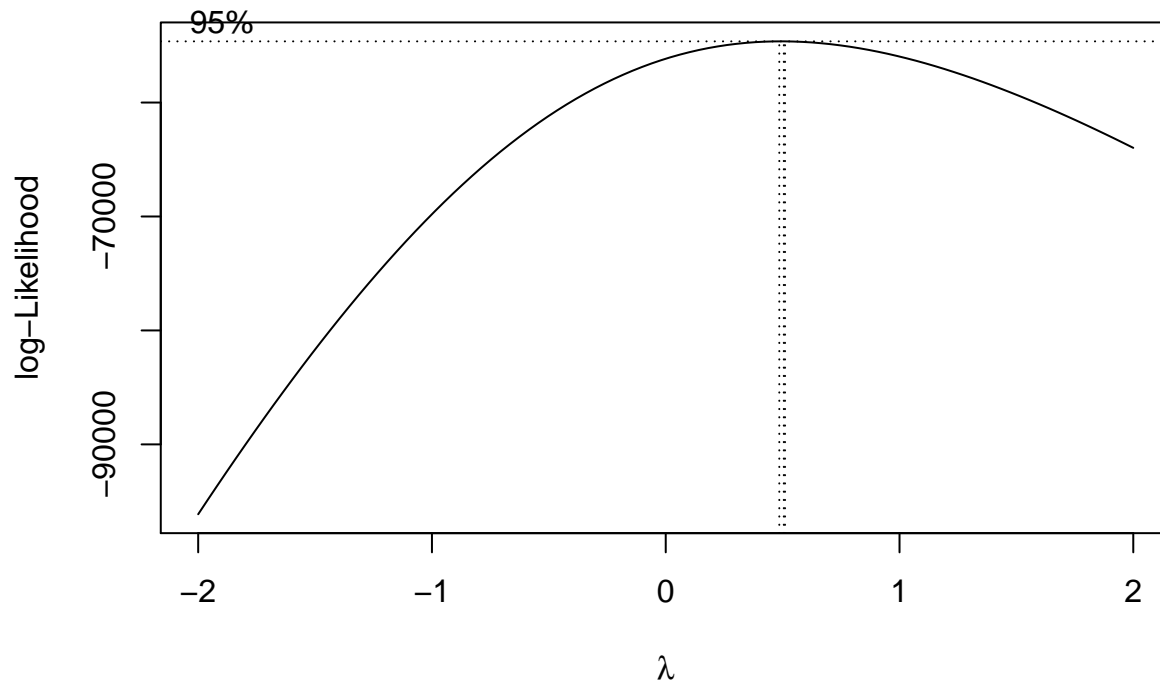
```
## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```



```
# BOX-COX (LINEAR) REGRESSION-----
```

```
bc = boxcox(sfr_lin_mod, lambda = seq(-3, 3, 1))
```

```
## Warning: In lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :  
## extra argument 'lambda' will be disregarded
```



```
best_lam = bc$x[which(bc$y == max(bc$y))]
dat1 %<>%
  mutate(SALEPRICEbc = SALEPRICE^best_lam)

# pre-analysis formula
sfr_bc = formula(paste0("SALEPRICEbc ~
                        CN_score + percent_vacant + dist_cityhall + dist_ugb + totalsqft + yearbuilt"))

# regression
sfr_mod_bc <- lm(sfr_bc, dat1)
summary(sfr_mod_bc)
```

```
##
## Call:
## lm(formula = sfr_bc, data = dat1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -624.99  -35.17    4.82   41.47  366.67
##
## Coefficients: (6 not defined because of singularities)
##
##              Estimate Std. Error
## (Intercept)  -4.566e+04  1.073e+03
## CN_score      -2.691e-01  8.009e-02
## percent_vacant  1.228e+00  2.480e+00
## dist_cityhall  -4.254e-03  3.099e-04
## dist_ugb       -1.108e-03  3.841e-04
## totalsqft      4.242e-02  1.013e-03
## yearbuilt     -2.229e-01  3.262e-02
## f_baths        4.548e+01  1.046e+00
## h_baths        3.212e+01  1.321e+00
```

## n_fireplaces	-3.118e+00	3.215e+00
## ADUdummy	-2.648e+01	5.360e+00
## pct_canopy_cov	-9.646e+00	3.655e+00
## taxlot_area	1.161e-03	9.308e-05
## garage_dumTRUE	1.100e+01	1.736e+00
## attic_dumTRUE	-7.605e-02	1.586e+00
## bsmt_dumTRUE	-5.838e-01	1.646e+00
## year_sold	2.325e+01	5.304e-01
## MapLabelArbor Lodge	-4.298e+00	1.666e+01
## MapLabelArdenwald-Johnson Creek	-9.670e+01	1.891e+01
## MapLabelArgay Terrace	-1.539e+02	1.579e+01
## MapLabelArlington Heights	-5.429e+01	2.276e+01
## MapLabelArnold Creek	-1.009e+02	1.472e+01
## MapLabelAshcreek	-1.268e+02	1.430e+01
## MapLabelBeaumont-Wilshire	-8.994e+00	7.459e+00
## MapLabelBoise	6.177e+00	2.330e+01
## MapLabelBrentwood-Darlington	-1.291e+02	1.188e+01
## MapLabelBridlemile	-1.391e+02	1.913e+01
## MapLabelBrooklyn Action Corps	-9.097e+01	1.403e+01
## MapLabelBuckman Community Association	-3.049e+01	1.643e+01
## MapLabelCathedral Park	2.263e+01	2.158e+01
## MapLabelCentennial Community Association	-1.521e+02	1.453e+01
## MapLabelCollins View	-1.281e+02	1.526e+01
## MapLabelConcordia	-2.588e+01	9.860e+00
## MapLabelCreston-Kenilworth	-8.197e+01	1.189e+01
## MapLabelCrestwood	-1.441e+02	1.504e+01
## MapLabelCully Association of Neighbors	-8.770e+01	9.271e+00
## MapLabelEast Columbia	-1.267e+02	2.031e+01
## MapLabelEastmoreland	5.055e+01	1.228e+01
## MapLabelEliot	-9.494e+01	3.586e+01
## MapLabelFar Southwest	-9.995e+01	1.557e+01
## MapLabelForest Park	-7.103e+01	2.119e+01
## MapLabelFoster-Powell	-8.539e+01	1.218e+01
## MapLabelGlenfair	-1.511e+02	1.891e+01
## MapLabelGoose Hollow Foothills League	-9.009e+01	2.584e+01
## MapLabelGrant Park	3.418e+01	8.593e+00
## MapLabelHayden Island Neighborhood Network	4.854e+01	2.726e+01
## MapLabelHayhurst	-1.312e+02	1.432e+01
## MapLabelHazelwood	-1.344e+02	1.244e+01
## MapLabelHealy Heights	-4.531e+01	3.064e+01
## MapLabelHillsdale	-1.164e+02	1.442e+01
## MapLabelHillside	-6.281e+01	2.126e+01
## MapLabelHollywood	3.287e+01	1.828e+01
## MapLabelHomestead	-1.684e+02	2.508e+01
## MapLabelHosford-Abernethy Neighborhood District Assn.	-1.747e+01	1.200e+01
## MapLabelHumboldt	-3.869e+00	1.657e+01
## MapLabelIrvington Community Association	1.199e+01	8.965e+00
## MapLabelKenton	-2.079e+01	1.870e+01
## MapLabelKerns	-1.819e+01	2.037e+01
## MapLabelKing	-5.333e+01	1.317e+01
## MapLabelLaurelhurst	3.027e+01	7.637e+00
## MapLabelLents	-1.368e+02	1.148e+01
## MapLabelLinnton	-1.381e+02	2.443e+01
## MapLabelMadison South	-9.536e+01	9.920e+00

## MapLabelMaplewood	-1.062e+02	1.498e+01
## MapLabelMarkham	-1.496e+02	1.472e+01
## MapLabelMarshall Park	-1.153e+02	1.594e+01
## MapLabelMC Unclaimed #13	-1.250e+02	7.056e+01
## MapLabelMill Park	-1.432e+02	1.303e+01
## MapLabelMontavilla	-7.137e+01	9.494e+00
## MapLabelMt. Scott-Arleta	-8.079e+01	1.212e+01
## MapLabelMt. Tabor	1.464e+01	1.028e+01
## MapLabelMultnomah	-1.152e+02	1.433e+01
## MapLabelNorth Tabor	-4.313e+01	1.091e+01
## MapLabelNorthwest District Association	1.629e+01	2.212e+01
## MapLabelNorthwest Heights	-5.587e+01	1.962e+01
## MapLabelOverlook	1.027e+01	1.721e+01
## MapLabelParkrose	-1.666e+02	1.613e+01
## MapLabelParkrose Heights Association of Neighbors	-1.620e+02	1.562e+01
## MapLabelPiedmont	-2.270e+01	1.638e+01
## MapLabelPleasant Valley	-1.179e+02	1.387e+01
## MapLabelPortsmouth	-2.119e+01	1.964e+01
## MapLabelPowellhurst-Gilbert	-1.402e+02	1.329e+01
## MapLabelReed	-4.096e+01	1.370e+01
## MapLabelRichmond	-1.319e+01	1.073e+01
## MapLabelRose City Park	-2.292e+01	7.246e+00
## MapLabelRoseway	-4.998e+01	9.200e+00
## MapLabelRussell	-1.426e+02	1.573e+01
## MapLabelSabin Community Association	-5.684e+00	8.173e+00
## MapLabelSellwood-Moreland Improvement League	-1.022e+01	1.166e+01
## MapLabelSouth Burlingame	-1.297e+02	1.533e+01
## MapLabelSouth Portland	-6.655e+01	1.679e+01
## MapLabelSouth Tabor	-4.492e+01	1.137e+01
## MapLabelSouthwest Hills Residential League	-7.143e+01	1.918e+01
## MapLabelSt. Johns	-1.581e+01	1.967e+01
## MapLabelSullivan's Gulch	-6.369e+01	1.751e+01
## MapLabelSumner Association of Neighbors	-1.596e+02	1.815e+01
## MapLabelSunderland Association of Neighbors	-1.316e+02	4.647e+01
## MapLabelSunnyside	-2.681e+00	1.209e+01
## MapLabelSylvan-Highlands	-1.219e+02	2.144e+01
## MapLabelUniversity Park	1.116e+01	1.970e+01
## MapLabelVernon	-8.856e-01	1.451e+01
## MapLabelWest Portland Park	-1.437e+02	1.465e+01
## MapLabelWilkes Community Group	-1.454e+02	1.652e+01
## MapLabelWoodland Park	-1.639e+02	2.169e+01
## MapLabelWoodlawn	-2.810e+01	1.494e+01
## MapLabelWoodstock	-4.820e+01	1.231e+01
## HIGH_SCHCLEVELAND HS	-4.322e+01	1.130e+01
## HIGH_SCHDAVID DOUGLAS HS	-3.667e+01	6.927e+00
## HIGH_SCHFRANKLIN HS	-4.724e+01	1.024e+01
## HIGH_SCHGRANT HS	-2.323e+01	1.209e+01
## HIGH_SCHGRANT-JEFF	-7.567e+01	1.796e+01
## HIGH_SCHLINCOLN HS	4.389e+01	1.277e+01
## HIGH_SCHMADISON HS	-7.691e+01	1.117e+01
## HIGH_SCHMADISON-JEFF	-7.745e+01	1.334e+01
## HIGH_SCHPARKROSE HS	-2.617e+01	1.048e+01
## HIGH_SCHREYNOLDS HS	-1.650e+01	1.034e+01
## HIGH_SCHRIVERDALE HS	2.483e+02	5.402e+01

## HIGH_SCHROOS-JEFF	-1.053e+02	1.834e+01
## HIGH_SCHROOSEVELT CAMPUS	-1.195e+02	2.086e+01
## HIGH_SCHSUNSET HS	4.709e+01	2.507e+01
## HIGH_SCHWILSON HS	NA	NA
## conECSEI	6.525e+01	7.569e+01
## conLUST	-2.702e+01	5.427e+01
## conHist	2.327e+01	4.970e+00
## conHistLdm	3.145e+00	2.321e+01
## conNatAm	NA	NA
## conAirHgt	1.334e+01	3.213e+00
## conHeliprt	NA	NA
## conNoise	-2.288e+00	7.119e+00
## conGW	4.154e+01	4.040e+01
## conLSHA	1.112e+01	3.055e+00
## conSLIDO	-1.043e+01	4.524e+00
## conSlp25	4.974e+00	3.103e+00
## conInstit	NA	NA
## conPrvCom	1.338e+02	7.626e+01
## conPubOwn	-2.648e+01	3.090e+01
## conSewer	1.059e+01	9.847e+00
## conStorm	-7.636e-01	2.404e+00
## conTranCap	-5.001e+00	2.609e+00
## conTranInt	-9.466e+00	2.803e+00
## conTranSub	-7.727e+00	1.645e+00
## conWater	-8.233e+00	3.144e+00
## conView	NA	NA
## conCovrly	3.324e-01	3.419e+00
## conPovrly	-1.478e+01	4.743e+00
## conWetland	-4.194e+01	1.831e+01
## conFld100	-8.352e+00	9.410e+00
## pct_conFld100	1.099e+01	1.477e+01
## pct_conPovrly	1.395e+01	7.659e+01
## pct_conCovrly	2.685e+01	6.609e+00
## pct_conWetland	NA	NA
##	t value	Pr(> t)
## (Intercept)	-42.567	< 2e-16 ***
## CN_score	-3.360	0.000780 ***
## percent_vacant	0.495	0.620417
## dist_cityhall	-13.728	< 2e-16 ***
## dist_ugb	-2.885	0.003921 **
## totalsqft	41.873	< 2e-16 ***
## yearbuilt	-6.834	8.57e-12 ***
## f_baths	43.493	< 2e-16 ***
## h_baths	24.312	< 2e-16 ***
## n_fireplaces	-0.970	0.332192
## ADUdummy	-4.940	7.87e-07 ***
## pct_canopy_cov	-2.639	0.008315 **
## taxlot_area	12.473	< 2e-16 ***
## garage_dumTRUE	6.335	2.43e-10 ***
## attic_dumTRUE	-0.048	0.961767
## bsmt_dumTRUE	-0.355	0.722821
## year_sold	43.835	< 2e-16 ***
## MapLabelArbor Lodge	-0.258	0.796401
## MapLabelArdenwald-Johnson Creek	-5.113	3.21e-07 ***

## MapLabelArgay Terrace	-9.744	< 2e-16	***
## MapLabelArlington Heights	-2.386	0.017058	*
## MapLabelArnold Creek	-6.851	7.57e-12	***
## MapLabelAshcreek	-8.870	< 2e-16	***
## MapLabelBeaumont-Wilshire	-1.206	0.227931	
## MapLabelBoise	0.265	0.790918	
## MapLabelBrentwood-Darlington	-10.866	< 2e-16	***
## MapLabelBridlemile	-7.273	3.67e-13	***
## MapLabelBrooklyn Action Corps	-6.484	9.19e-11	***
## MapLabelBuckman Community Association	-1.855	0.063592	.
## MapLabelCathedral Park	1.048	0.294444	
## MapLabelCentennial Community Association	-10.467	< 2e-16	***
## MapLabelCollins View	-8.393	< 2e-16	***
## MapLabelConcordia	-2.624	0.008686	**
## MapLabelCreston-Kenilworth	-6.894	5.64e-12	***
## MapLabelCrestwood	-9.579	< 2e-16	***
## MapLabelCully Association of Neighbors	-9.459	< 2e-16	***
## MapLabelEast Columbia	-6.238	4.56e-10	***
## MapLabelEastmoreland	4.118	3.84e-05	***
## MapLabelEliot	-2.647	0.008125	**
## MapLabelFar Southwest	-6.418	1.41e-10	***
## MapLabelForest Park	-3.353	0.000803	***
## MapLabelFoster-Powell	-7.010	2.48e-12	***
## MapLabelGlenfair	-7.992	1.41e-15	***
## MapLabelGoose Hollow Foothills League	-3.487	0.000490	***
## MapLabelGrant Park	3.978	6.99e-05	***
## MapLabelHayden Island Neighborhood Network	1.781	0.074913	.
## MapLabelHayhurst	-9.165	< 2e-16	***
## MapLabelHazelwood	-10.799	< 2e-16	***
## MapLabelHealy Heights	-1.479	0.139258	
## MapLabelHillsdale	-8.072	7.40e-16	***
## MapLabelHillside	-2.954	0.003143	**
## MapLabelHollywood	1.797	0.072292	.
## MapLabelHomestead	-6.716	1.93e-11	***
## MapLabelHosford-Abernethy Neighborhood District Assn.	-1.456	0.145488	
## MapLabelHumboldt	-0.233	0.815415	
## MapLabelIrvington Community Association	1.337	0.181250	
## MapLabelKenton	-1.112	0.266290	
## MapLabelKerns	-0.893	0.371984	
## MapLabelKing	-4.050	5.15e-05	***
## MapLabelLaurelhurst	3.964	7.40e-05	***
## MapLabelLents	-11.917	< 2e-16	***
## MapLabelLinnton	-5.653	1.60e-08	***
## MapLabelMadison South	-9.613	< 2e-16	***
## MapLabelMaplewood	-7.089	1.41e-12	***
## MapLabelMarkham	-10.160	< 2e-16	***
## MapLabelMarshall Park	-7.236	4.81e-13	***
## MapLabelMC Unclaimed #13	-1.771	0.076616	.
## MapLabelMill Park	-10.989	< 2e-16	***
## MapLabelMontavilla	-7.518	5.86e-14	***
## MapLabelMt. Scott-Arleta	-6.663	2.77e-11	***
## MapLabelMt. Tabor	1.425	0.154279	
## MapLabelMultnomah	-8.041	9.51e-16	***
## MapLabelNorth Tabor	-3.953	7.75e-05	***

## MapLabelNorthwest District Association	0.736	0.461503	
## MapLabelNorthwest Heights	-2.848	0.004402	**
## MapLabelOverlook	0.597	0.550764	
## MapLabelParkrose	-10.332	< 2e-16	***
## MapLabelParkrose Heights Association of Neighbors	-10.372	< 2e-16	***
## MapLabelPiedmont	-1.386	0.165773	
## MapLabelPleasant Valley	-8.496	< 2e-16	***
## MapLabelPortsmouth	-1.079	0.280669	
## MapLabelPowellhurst-Gilbert	-10.549	< 2e-16	***
## MapLabelReed	-2.991	0.002789	**
## MapLabelRichmond	-1.229	0.218928	
## MapLabelRose City Park	-3.163	0.001566	**
## MapLabelRoseway	-5.433	5.62e-08	***
## MapLabelRussell	-9.067	< 2e-16	***
## MapLabelSabin Community Association	-0.695	0.486762	
## MapLabelSellwood-Moreland Improvement League	-0.877	0.380585	
## MapLabelSouth Burlingame	-8.461	< 2e-16	***
## MapLabelSouth Portland	-3.965	7.38e-05	***
## MapLabelSouth Tabor	-3.950	7.84e-05	***
## MapLabelSouthwest Hills Residential League	-3.723	0.000197	***
## MapLabelSt. Johns	-0.803	0.421703	
## MapLabelSullivan's Gulch	-3.637	0.000277	***
## MapLabelSumner Association of Neighbors	-8.791	< 2e-16	***
## MapLabelSunderland Association of Neighbors	-2.833	0.004620	**
## MapLabelSunnyside	-0.222	0.824520	
## MapLabelSylvan-Highlands	-5.687	1.31e-08	***
## MapLabelUniversity Park	0.566	0.571190	
## MapLabelVernon	-0.061	0.951340	
## MapLabelWest Portland Park	-9.809	< 2e-16	***
## MapLabelWilkes Community Group	-8.803	< 2e-16	***
## MapLabelWoodland Park	-7.555	4.41e-14	***
## MapLabelWoodlawn	-1.881	0.059983	.
## MapLabelWoodstock	-3.914	9.11e-05	***
## HIGH_SCHCLEVELAND HS	-3.826	0.000131	***
## HIGH_SCHDAVID DOUGLAS HS	-5.293	1.22e-07	***
## HIGH_SCHFRANKLIN HS	-4.611	4.04e-06	***
## HIGH_SCHGRANT HS	-1.922	0.054653	.
## HIGH_SCHGRANT-JEFF	-4.214	2.53e-05	***
## HIGH_SCHLINCOLN HS	3.438	0.000588	***
## HIGH_SCHMADISON HS	-6.883	6.06e-12	***
## HIGH_SCHMADISON-JEFF	-5.808	6.44e-09	***
## HIGH_SCHPARKROSE HS	-2.497	0.012526	*
## HIGH_SCHREYNOLDS HS	-1.597	0.110375	
## HIGH_SCHRIVERDALE HS	4.596	4.35e-06	***
## HIGH_SCHROOS-JEFF	-5.740	9.65e-09	***
## HIGH_SCHROOSEVELT CAMPUS	-5.730	1.02e-08	***
## HIGH_SCHSUNSET HS	1.879	0.060323	.
## HIGH_SCHWILSON HS	NA	NA	
## conEC SI	0.862	0.388673	
## conLUST	-0.498	0.618569	
## conHist	4.683	2.86e-06	***
## conHistLdm	0.136	0.892202	
## conNatAm	NA	NA	
## conAirHgt	4.152	3.31e-05	***

```

## conHeliprt          NA      NA
## conNoise            -0.321 0.747945
## conGW               1.028 0.303921
## conLSHA             3.639 0.000275 ***
## conSLID0           -2.306 0.021148 *
## conSlp25            1.603 0.109022
## conInstit           NA      NA
## conPrvCom           1.755 0.079284 .
## conPubOwn          -0.857 0.391572
## conSewer            1.076 0.282024
## conStorm           -0.318 0.750782
## conTranCap         -1.917 0.055291 .
## conTranInt         -3.377 0.000735 ***
## conTranSub         -4.696 2.67e-06 ***
## conWater           -2.619 0.008838 **
## conView            NA      NA
## conCovrly           0.097 0.922544
## conPovrly          -3.117 0.001829 **
## conWetland         -2.290 0.022026 *
## conFld100          -0.888 0.374807
## pct_conFld100       0.744 0.456608
## pct_conPovrly       0.182 0.855426
## pct_conCovrly       4.063 4.87e-05 ***
## pct_conWetland      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 75.5 on 16383 degrees of freedom
## (6539 observations deleted due to missingness)
## Multiple R-squared:  0.7555, Adjusted R-squared:  0.7533
## F-statistic: 351.5 on 144 and 16383 DF, p-value: < 2.2e-16

```

```
#plot(sfr_mod_bc)
```