Smoothing Model Building

Maya Le

Summary

The model produces a smoothing model that predict property prices based on a set of training data. First, we prepare the data by filling NA values by simple imputation methods. Then, we analyze the variables histogram and scatterplot to determine whether they need transformation. Lastly, we build the smoothing model.

The prediction error will be evaluted using RMLSE (Root Mean Squared Logarithmic Error).

Preprocessing

There are a few observations with no rooms. This can be a sign that this data point was recorded incorrectly. We will drop outliers from the training data set, which are the observations with no rooms, bedrooms, or kitchens.

Transformation (if any, delete if none)

- price: log-transformed.

New Variables (if any, delete if none)

- sinceRmdl: how many years since the property was remodelled.
- invest: indicates whether the property was bought before it was built.

Missing data handling

- yr_rmdl: replace the missing values in yr_rmdl with median.

Model Building

Main package used: `magicSpline`

Forward selection and adding terms using thin plate spline.

Final Model

- The final model is $log(price) \sim$ bathrm + ac +bedrm + invest + s(sinceRmdl) + s(saledate) + style + grade + cndtn + intwall + s(rooms) + fireplaces + s(latitude) + s(longitude) + nbhd + quadrant +
  te(gba,landarea) + te(saledate,fireplaces) + te(longitude,gba) + te(gba,latitude) + te(longitude,eyb)

+ te(saledate,gba) + te(saledate,eyb,ayb) + te(gba,latitude,landarea) + te(bathrm,gba,eyb) + te(gba,landarea,ayb)

1.Preprocessing

1.1 Loading data

```
setwd("C:/Users/maian/OneDrive - University of Waterloo/Documents/UW/W24/Stat444/Project Smooth")
load("smooth.Rdata")
```

1.2 Missing data handling

First, checking for outliers, we see that there are properties with zero rooms, bathrooms, or kitchens. This could indicate errors when the data is recorded. Therefore, we will remove these observations.

```
dtrain <- subset(dtrain, rooms != 0 & bedrm != 0 & kitchens != 0 & bathrm != 0)
```

Next, we see that there are some missing values:

```
colSums(is.na(dtrain))
```

```
##     bathrm  hf_bathrm       heat         ac      rooms      bedrm        ayb
##          0          0          0          0          0          0         14
##    yr_rmdl        eyb    stories   saledate      price        gba      style
##       1995          0          4          0          0          0          0
##      grade      cndtn   extwall       roof    intwall   kitchens fireplaces
##          0          0          0          0          0          0          0
##   landarea   latitude  longitude       nbhd       ward   quadrant
##          0          0          0          0          0         27
```

As for saledate, we will take the year only and calculate the age of the property when it was sold.

```
dtrain$saledate <- year(dtrain$saledate)
```

Ayb is the year that the property was first built. To fill this value, we will use the average age of the other properties, then deduct the age from saledate.

```
avgAge <- mean(dtrain$saledate - dtrain$ayb, na.rm = TRUE)
dtrain$ayb <- ifelse(is.na(dtrain$ayb), dtrain$saledate - avgAge, dtrain$ayb)
```

yr_rmdl is the year that the propety was remodelled. If the value is NA, then we can assume that the property has never been remodelled, and we can use ayb value instead.

```
dtrain$yr_rmdl <- ifelse(is.na(dtrain$yr_rmdl), dtrain$ayb, dtrain$yr_rmdl)
```

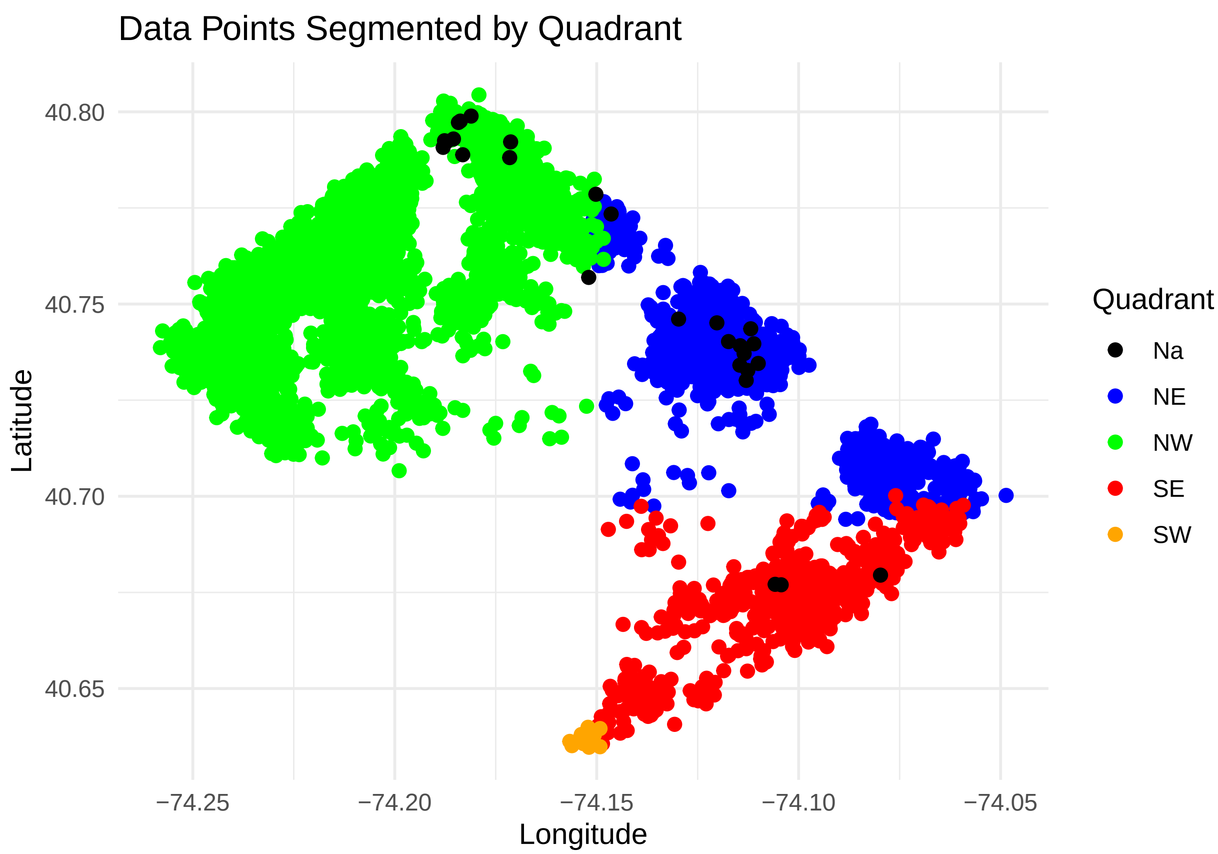
The number of stories that were missing can potentially be extracted from style.

```r
extract_stories <- function(style) {
  if (grepl("Story", style)) {
    # If the style contains "story", extract the numeric part
    as.numeric(strsplit(style, " ")[[1]][1])
  } else if (grepl("Bi-level", style)) {
    2
  } else if (grepl("Split", style)) {
    1.5
  } else {
    1
  }
}

dtrain$stories <- ifelse(is.na(dtrain$stories),
                         sapply(dtrain$style, extract_stories),dtrain$stories)
```

Based on this plot, we can see the missing quadrant values, where the properties with quadrant values of NA is presented as black dots:

```r
ggplot(data = dtrain, aes(x = longitude, y = latitude)) +
  labs(title = "Data Points Segmented by Quadrant", x = "Longitude", y = "Latitude") +
  theme_minimal() +
  geom_point(data = subset(dtrain, quadrant == "NE"), aes(color = "NE"), size = 2) +
  geom_point(data = subset(dtrain, quadrant == "NW"), aes(color = "NW"), size = 2) +
  geom_point(data = subset(dtrain, quadrant == "SE"), aes(color = "SE"), size = 2) +
  geom_point(data = subset(dtrain, quadrant == "SW"), aes(color = "SW"), size = 2) +
  geom_point(data = subset(dtrain, is.na(quadrant)), aes(color = "Na"), size = 2)+
  scale_color_manual(name = "Quadrant",
                     values = c(NE = "blue", NW = "green", SE = "red", SW = "orange",
                                Na = "black"))
```



Therefore, we can use k-nearest-neighbor method to fill out missing quadrant values using their coordinates.

```
fill_quadrant_knn <- function(data, k) {
  complete_data <- data[!is.na(data$quadrant), c("latitude", "longitude")]
  complete_labels <- data$quadrant[!is.na(data$quadrant)]
  missing_data <- data[is.na(data$quadrant), c("latitude", "longitude")]

  knn_result <- knn(train = complete_data, test = missing_data,
                    cl = complete_labels, k = k)

  data$quadrant[is.na(data$quadrant)] <- knn_result
  return(data)
}

dtrain <- fill_quadrant_knn(dtrain, 5)

labels <- c("NE","NW","SE")
for (i in seq(1,3)){
  dtrain$quadrant <- ifelse(dtrain$quadrant == i, labels[i],dtrain$quadrant)
}
```

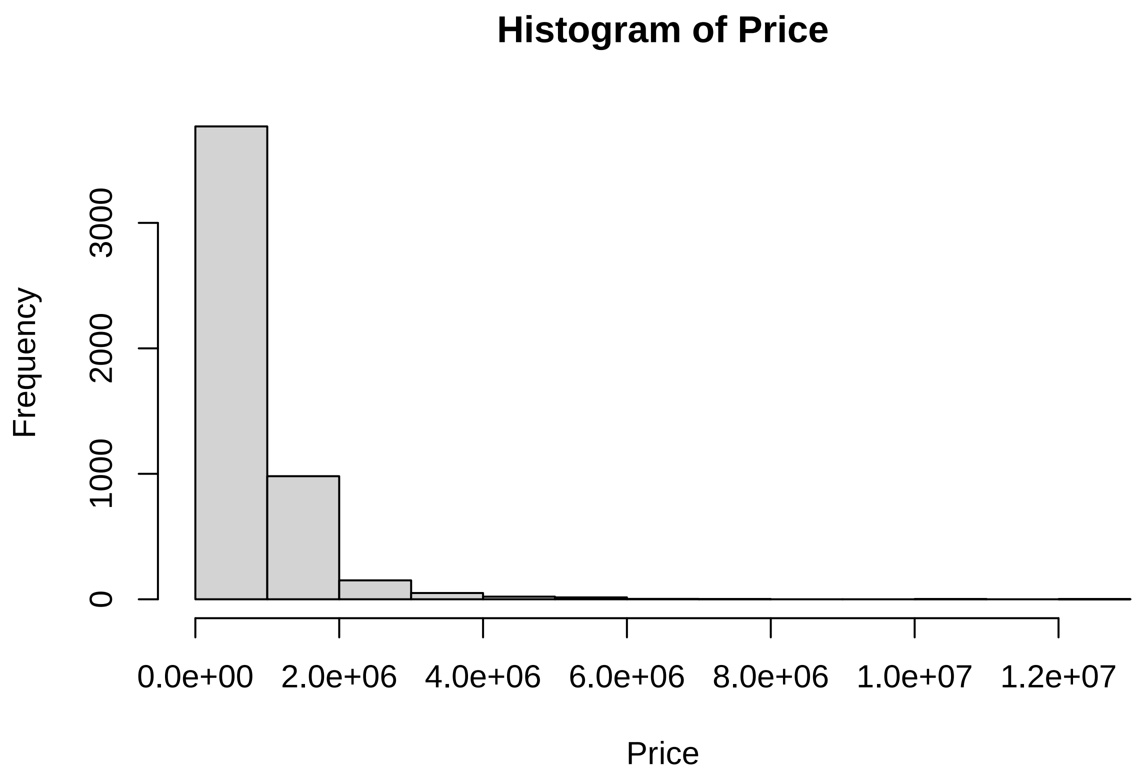
As for the testing data set, there are some levels in categorical variables that weren't in the original training data set. To run the evaluation file, we will replace these values using the most common values.

for (col in names(dtest)[sapply(dtest, is.character)]) { new_levels <- setdiff(unique(dtest[[col]]), unique(dtrain[[col]])) if (length(new_levels) > 0) { most_common_level <- names(sort(table(dtrain[[col]]), decreasing = TRUE))[1] dtest[[col]][dtest[[col]] %in% new_levels] <- most_common_level } }

2. Model building

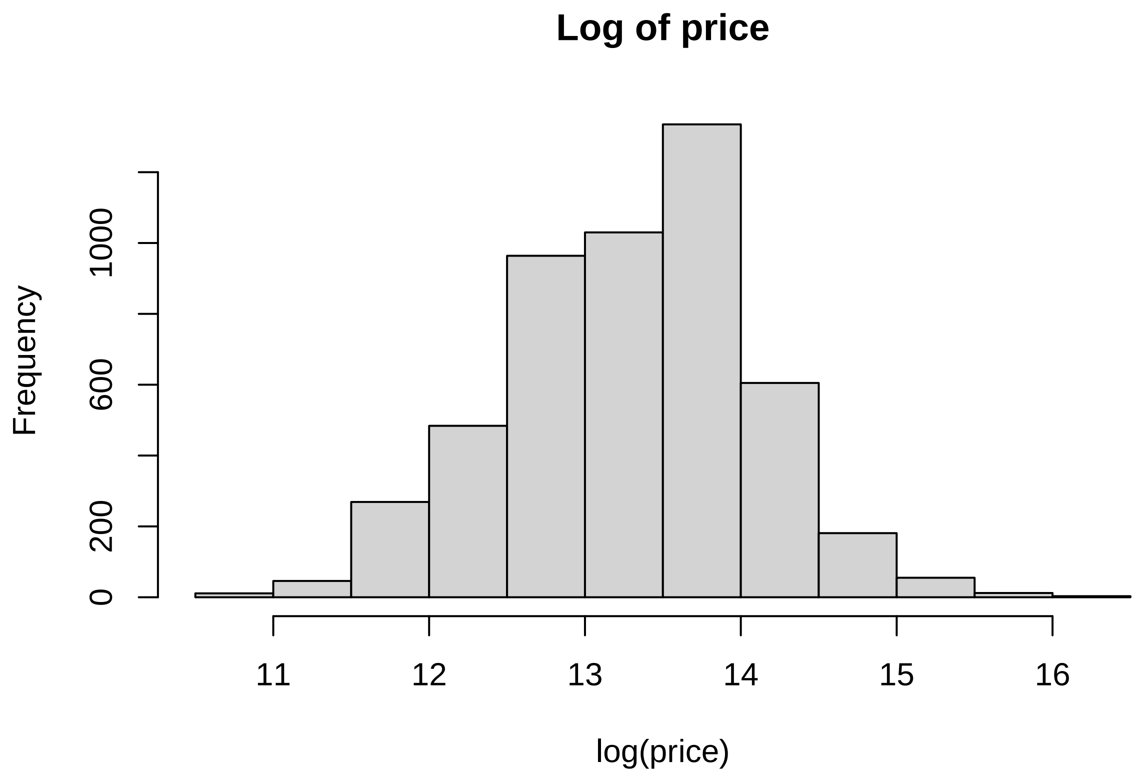
First, we need to transform the variables.

```
hist(dtrain$price, main = "Histogram of Price", xlab = "Price")
```



The histogram is right skewed. Therefore, we will apply log transform to get a normal distributed histogram.

```
hist(log(dtrain$price), main = "Log of price", xlab = "log(price)")
```



We will treat a half bathroom as 0.5 bathroom, and combine the two variables.

```
dtrain$bathrm <- dtrain$bathrm + (dtrain$hf_bathrm * 0.5)
```

We will create a new variable that shows how long the property was remodeled.

```
dtrain$sinceRmdl <- ifelse(is.na(dtrain$yr_rmdl)|dtrain$yr_rmdl > dtrain$saledate,
                           dtrain$saledate - dtrain$ayb,
                           dtrain$saledate - dtrain$yr_rmdl)
```

Also, notice that there are some property that was bought before it was built. That is, ayb is larger than saledate.

```
sum(dtrain$ayb > dtrain$saledate)
```

```
## [1] 35
```

Therefore, we will create a variable called invest, that indicates "Y" if the property was bought before it was built.

```
dtrain$invest <- ifelse(dtrain$ayb > dtrain$saledate, "Y", "N")
```

Next, we will choose a model by AIC in a Stepwise Algorithm. Starting with the basic model log(price) ~ 1.

```r
para <- c(
  "bathrm",  "ac", "rooms", "bedrm", "ayb", "eyb",
  "stories", "saledate", "price", "gba", "kitchens", "fireplaces", "landarea",
  "latitude", "longitude", "quadrant","sinceRmdl", "invest")

match_index <- match(para, names(dtrain))
train <- dtrain[,match_index]

basic <- lm(log(price) ~ 1, data=train)
full <- lm(log(price) ~ (.)^2, data=train)

mod <- step(basic,scope=list(lower=basic, upper=full),
                               direction="forward", trace=0)
summary(mod)
```

```
##
## Call:
## lm(formula = log(price) ~ bathrm + longitude + saledate + gba +
##     invest + fireplaces + quadrant + eyb + ayb + landarea + ac +
##     latitude + sinceRmdl + stories + rooms + longitude:quadrant +
##     fireplaces:landarea + fireplaces:latitude + saledate:sinceRmdl +
##     saledate:quadrant + invest:quadrant + quadrant:latitude +
##     longitude:latitude + gba:landarea + quadrant:landarea + bathrm:latitude +
##     bathrm:saledate + gba:quadrant + longitude:gba + gba:latitude +
##     bathrm:gba + bathrm:eyb + ayb:ac + quadrant:ac + eyb:ayb +
##     gba:invest + longitude:saledate + saledate:ac + saledate:fireplaces +
##     invest:landarea + longitude:ayb + quadrant:ayb + gba:eyb +
##     quadrant:sinceRmdl + ayb:sinceRmdl + gba:fireplaces + fireplaces:eyb +
##     stories:rooms + longitude:stories + sinceRmdl:rooms + quadrant:stories +
##     saledate:gba + longitude:fireplaces + fireplaces:quadrant +
##     quadrant:eyb + longitude:eyb + latitude:stories + landarea:rooms +
##     invest:latitude + invest:sinceRmdl + invest:stories + eyb:latitude +
##     longitude:rooms + bathrm:ac + saledate:latitude + longitude:invest +
##     invest:fireplaces + fireplaces:sinceRmdl + latitude:rooms +
##     eyb:stories + sinceRmdl:stories + gba:sinceRmdl + saledate:eyb +
##     invest:rooms + latitude:sinceRmdl + landarea:latitude, data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23769 -0.13227 -0.00667  0.12979  0.89147
##
## Coefficients: (2 not defined because of singularities)
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             2.962e+05  2.513e+04  11.789  < 2e-16 ***
## bathrm                  1.789e+01  6.861e+00   2.608 0.009141 **
## longitude               3.999e+03  3.384e+02  11.817  < 2e-16 ***
## saledate               -6.145e+00  2.048e+00  -3.000 0.002711 **
## gba                     1.673e-01  2.390e-02   6.999 2.94e-12 ***
## investY                -9.130e+02  2.311e+02  -3.951 7.89e-05 ***
## fireplaces              9.528e+01  2.109e+01   4.518 6.38e-06 ***
## quadrantNW              7.983e+02  6.688e+01  11.937  < 2e-16 ***
## quadrantSE             -1.682e+02  1.029e+02  -1.635 0.102212
## quadrantSW              4.486e+04  2.915e+04   1.539 0.123863
```

```
## eyb                    -2.218e+00  1.294e+00  -1.714 0.086666 .
## ayb                     5.236e+00  6.309e-01   8.299  < 2e-16 ***
## landarea               -3.863e-03  2.659e-03  -1.453 0.146337
## acY                     1.030e+01  3.112e+00   3.310 0.000939 ***
## latitude               -7.127e+03  6.099e+02 -11.685  < 2e-16 ***
## sinceRmdl               2.451e-01  2.585e-01   0.948 0.343090
## stories                -1.484e+02  3.496e+01  -4.244 2.24e-05 ***
## rooms                  -4.064e+00  3.645e+00  -1.115 0.264910
## longitude:quadrantNW    4.545e+00  7.608e-01   5.974 2.47e-09 ***
## longitude:quadrantSE   -8.759e-01  1.036e+00  -0.845 0.398044
## longitude:quadrantSW    3.700e+02  2.367e+02   1.564 0.117991
## fireplaces:landarea    -1.230e-06  9.118e-07  -1.349 0.177428
## fireplaces:latitude    -7.298e-01  2.682e-01  -2.721 0.006524 **
## saledate:sinceRmdl      8.757e-05  1.724e-05   5.080 3.91e-07 ***
## saledate:quadrantNW    -1.077e-02  2.579e-03  -4.175 3.03e-05 ***
## saledate:quadrantSE    -5.841e-03  2.267e-03  -2.576 0.010013 *
## saledate:quadrantSW    -5.117e-02  3.477e-02  -1.472 0.141143
## investY:quadrantNW     -4.325e-01  3.632e-01  -1.191 0.233788
## investY:quadrantSE      2.621e-01  1.760e-01   1.490 0.136398
## investY:quadrantSW             NA         NA      NA       NA
## quadrantNW:latitude    -1.081e+01  1.081e+00  -9.995  < 2e-16 ***
## quadrantSE:latitude     2.935e+00  1.366e+00   2.149 0.031699 *
## quadrantSW:latitude    -4.252e+02  3.077e+02  -1.382 0.167041
## longitude:latitude     -9.600e+01  8.235e+00 -11.658  < 2e-16 ***
## gba:landarea           -1.958e-09  5.131e-10  -3.816 0.000137 ***
## quadrantNW:landarea     1.118e-05  4.305e-06   2.597 0.009433 **
## quadrantSE:landarea     6.104e-06  6.546e-06   0.932 0.351197
## quadrantSW:landarea    -5.338e-04  5.042e-04  -1.059 0.289714
## bathrm:latitude        -3.604e-01  1.598e-01  -2.256 0.024142 *
## bathrm:saledate        -2.397e-03  6.620e-04  -3.620 0.000297 ***
## gba:quadrantNW          1.932e-04  3.275e-05   5.898 3.92e-09 ***
## gba:quadrantSE         -1.472e-04  3.548e-05  -4.148 3.41e-05 ***
## gba:quadrantSW          6.655e-06  1.086e-03   0.006 0.995110
## longitude:gba           1.451e-03  2.286e-04   6.346 2.41e-10 ***
## gba:latitude           -1.590e-03  3.322e-04  -4.787 1.74e-06 ***
## bathrm:gba             -1.194e-05  2.590e-06  -4.610 4.12e-06 ***
## bathrm:eyb              8.709e-04  3.326e-04   2.618 0.008869 **
## ayb:acY                -1.799e-03  7.413e-04  -2.426 0.015283 *
## quadrantNW:acY          1.113e-01  2.399e-02   4.638 3.62e-06 ***
## quadrantSE:acY          1.283e-01  3.369e-02   3.808 0.000142 ***
## quadrantSW:acY         -1.568e+00  1.205e+00  -1.301 0.193177
## eyb:ayb                -6.228e-05  1.219e-05  -5.109 3.35e-07 ***
## gba:investY            -1.292e-04  5.257e-05  -2.458 0.014024 *
## longitude:saledate     -5.776e-02  2.040e-02  -2.832 0.004649 **
## saledate:acY           -3.383e-03  1.332e-03  -2.539 0.011138 *
## saledate:fireplaces    -2.764e-03  6.784e-04  -4.075 4.68e-05 ***
## investY:landarea        4.235e-05  1.024e-05   4.134 3.63e-05 ***
## longitude:ayb           6.902e-02  8.518e-03   8.104 6.68e-16 ***
## quadrantNW:ayb          7.894e-03  9.697e-04   8.141 4.94e-16 ***
## quadrantSE:ayb          3.595e-03  9.041e-04   3.976 7.11e-05 ***
## quadrantSW:ayb         -1.877e-02  2.069e-02  -0.908 0.364178
## gba:eyb                 6.132e-07  4.041e-07   1.517 0.129256
## quadrantNW:sinceRmdl    1.140e-03  3.662e-04   3.113 0.001861 **
## quadrantSE:sinceRmdl    1.515e-04  5.803e-04   0.261 0.794066
```

```
## quadrantSW:sinceRmdl  -1.689e-02  2.314e-02  -0.730 0.465418
## ayb:sinceRmdl         -1.512e-05  6.706e-06  -2.255 0.024168 *
## gba:fireplaces        -1.498e-05  3.592e-06  -4.171 3.09e-05 ***
## fireplaces:eyb         1.688e-03  3.811e-04   4.430 9.62e-06 ***
## stories:rooms         -1.019e-02  3.122e-03  -3.264 0.001106 **
## longitude:stories     -1.288e+00  3.514e-01  -3.665 0.000250 ***
## sinceRmdl:rooms       -1.268e-04  6.797e-05  -1.866 0.062155 .
## quadrantNW:stories    -1.463e-01  4.618e-02  -3.167 0.001549 **
## quadrantSE:stories     6.451e-02  3.813e-02   1.692 0.090743 .
## quadrantSW:stories    -1.688e+00  1.627e+00  -1.037 0.299659
## saledate:gba           2.002e-06  7.996e-07   2.504 0.012327 *
## longitude:fireplaces   8.534e-01  2.073e-01   4.117 3.90e-05 ***
## fireplaces:quadrantNW  8.681e-02  2.725e-02   3.186 0.001450 **
## fireplaces:quadrantSE  2.114e-02  2.497e-02   0.847 0.397287
## fireplaces:quadrantSW  1.117e-01  3.427e-01   0.326 0.744430
## quadrantNW:eyb        -7.534e-03  1.778e-03  -4.236 2.31e-05 ***
## quadrantSE:eyb        -5.819e-03  1.776e-03  -3.276 0.001061 **
## quadrantSW:eyb                NA         NA      NA       NA
## longitude:eyb         -5.781e-02  1.464e-02  -3.947 8.01e-05 ***
## latitude:stories       1.223e+00  4.254e-01   2.876 0.004049 **
## landarea:rooms        -9.777e-07  3.777e-07  -2.589 0.009657 **
## investY:latitude       1.262e+01  2.702e+00   4.673 3.05e-06 ***
## investY:sinceRmdl      7.650e-02  1.899e-02   4.029 5.69e-05 ***
## investY:stories        3.090e-01  1.284e-01   2.407 0.016136 *
## eyb:latitude          -4.405e-02  1.463e-02  -3.012 0.002612 **
## longitude:rooms       -1.484e-01  5.158e-02  -2.877 0.004028 **
## bathrm:acY            -2.615e-02  1.174e-02  -2.227 0.025973 *
## saledate:latitude      5.079e-02  2.609e-02   1.947 0.051593 .
## longitude:investY     -5.374e+00  2.351e+00  -2.286 0.022318 *
## investY:fireplaces    -1.729e-01  8.128e-02  -2.127 0.033429 *
## fireplaces:sinceRmdl   3.611e-04  1.580e-04   2.285 0.022349 *
## latitude:rooms        -1.695e-01  7.640e-02  -2.219 0.026543 *
## eyb:stories            1.622e-03  5.576e-04   2.909 0.003639 **
## sinceRmdl:stories      9.183e-04  3.025e-04   3.035 0.002415 **
## gba:sinceRmdl         -4.380e-07  1.870e-07  -2.343 0.019180 *
## saledate:eyb          -7.465e-05  4.297e-05  -1.737 0.082415 .
## investY:rooms         -4.452e-02  2.742e-02  -1.624 0.104511
## latitude:sinceRmdl    -9.675e-03  6.366e-03  -1.520 0.128601
## landarea:latitude      9.547e-05  6.526e-05   1.463 0.143585
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2228 on 4894 degrees of freedom
## Multiple R-squared:  0.922,  Adjusted R-squared:  0.9204
## F-statistic: 578.8 on 100 and 4894 DF,  p-value: < 2.2e-16
```

Based on this result, we can also see the interaction terms. After performing forward selection and adding terms using thin plate spline, we arrive at our final model:

```
fit <- bam(log(price) ~ bathrm + ac +bedrm + invest +
             s(sinceRmdl) +  s(saledate) + style + grade +
             cndtn + intwall + s(rooms) + fireplaces  +
             s(latitude) + s(longitude) + nbhd + quadrant +
             te(gba,landarea) + te(saledate,fireplaces) + te(longitude,gba) +
```

```r
             te(gba,latitude) +
             te(longitude,eyb) + te(saledate,gba) + te(saledate,eyb,ayb) +
             te(gba,latitude,landarea) +
             te(bathrm,gba,eyb) + te(gba,landarea,ayb),
           data=dtrain)

summary(fit)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## log(price) ~ bathrm + ac + bedrm + invest + s(sinceRmdl) + s(saledate) +
##     style + grade + cndtn + intwall + s(rooms) + fireplaces +
##     s(latitude) + s(longitude) + nbhd + quadrant + te(gba, landarea) +
##     te(saledate, fireplaces) + te(longitude, gba) + te(gba, latitude) +
##     te(longitude, eyb) + te(saledate, gba) + te(saledate, eyb,
##     ayb) + te(gba, latitude, landarea) + te(bathrm, gba, eyb) +
##     te(gba, landarea, ayb)
##
## Parametric coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)          12.9278726  0.0925013 139.759  < 2e-16 ***
## bathrm                0.0239947  0.0269159   0.891 0.372722
## acY                   0.0285510  0.0064310   4.440 9.22e-06 ***
## bedrm                 0.0099412  0.0027602   3.602 0.000320 ***
## investY              -0.9235852  0.0270980 -34.083  < 2e-16 ***
## style1.5 Story Fin    0.0081674  0.0104152   0.784 0.432969
## style1.5 Story Unfin -0.0323294  0.0374459  -0.863 0.387981
## style2 Story         -0.0082197  0.0083270  -0.987 0.323634
## style2.5 Story Fin    0.0152017  0.0100174   1.518 0.129199
## style2.5 Story Unfin  0.0061256  0.0160012   0.383 0.701870
## style3 Story          0.0084231  0.0154758   0.544 0.586276
## style4 Story         -0.0001661  0.0797332  -0.002 0.998338
## styleBi-Level         0.0298798  0.1328727   0.225 0.822086
## styleSplit Foyer      0.0770393  0.0249907   3.083 0.002063 **
## styleSplit Level     -0.0013625  0.0189519  -0.072 0.942691
## gradeAverage         -0.0516240  0.0084648  -6.099 1.15e-09 ***
## gradeExcellent        0.1397716  0.0138784  10.071  < 2e-16 ***
## gradeExceptional-A    0.2864517  0.0249685  11.473  < 2e-16 ***
## gradeExceptional-B    0.4032510  0.0299799  13.451  < 2e-16 ***
## gradeExceptional-C    0.5009276  0.0898915   5.573 2.65e-08 ***
## gradeExceptional-D    0.6311217  0.0632205   9.983  < 2e-16 ***
## gradeFair Quality    -0.0423178  0.0410534  -1.031 0.302687
## gradeGood Quality     0.0263710  0.0088195   2.990 0.002803 **
## gradeLow Quality     -0.6079385  0.1612053  -3.771 0.000164 ***
## gradeSuperior         0.2145643  0.0171991  12.475  < 2e-16 ***
## gradeVery Good        0.0631011  0.0104114   6.061 1.46e-09 ***
## cndtnExcellent        0.3052602  0.0232226  13.145  < 2e-16 ***
## cndtnFair            -0.1128771  0.0302397  -3.733 0.000192 ***
## cndtnGood             0.0684622  0.0050045  13.680  < 2e-16 ***
## cndtnPoor            -0.1545347  0.0688209  -2.245 0.024784 *
```

```
## cndtnVery Good        0.1790029  0.0082571  21.679  < 2e-16 ***
## intwallCeramic Tile  -0.0142504  0.0897892  -0.159 0.873905
## intwallDefault        0.0605269  0.0998052   0.606 0.544245
## intwallHardwood       0.0339133  0.0137792   2.461 0.013883 *
## intwallHardwood/Carp  0.0102087  0.0144208   0.708 0.479032
## intwallLt Concrete    0.3437221  0.0644854   5.330 1.03e-07 ***
## intwallParquet        0.0001861  0.1323344   0.001 0.998878
## intwallWood Floor     0.0022143  0.0173922   0.127 0.898696
## fireplaces            0.0080053  0.0124555   0.643 0.520442
## nbhdA2                0.1317892  0.0438424   3.006 0.002661 **
## nbhdA3               -0.5597011  0.0808789  -6.920 5.11e-12 ***
## nbhdA4               -0.7588760  0.0885102  -8.574  < 2e-16 ***
## nbhdA5                0.1964463  0.0515051   3.814 0.000138 ***
## nbhdA6               -0.0966487  0.0790242  -1.223 0.221380
## nbhdA7               -0.0328512  0.0232976  -1.410 0.158584
## nbhdA8                0.0055874  0.0555563   0.101 0.919895
## nbhdA9                0.3706944  0.0811870   4.566 5.10e-06 ***
## nbhdB1                0.6505498  0.0764967   8.504  < 2e-16 ***
## nbhdB2                0.1331306  0.0405957   3.279 0.001048 **
## nbhdB3               -0.0854327  0.0339871  -2.514 0.011980 *
## nbhdB4                0.3876220  0.0449754   8.619  < 2e-16 ***
## nbhdB5                0.1060593  0.0359665   2.949 0.003205 **
## nbhdB6               -0.1893561  0.0772691  -2.451 0.014297 *
## nbhdB7               -0.5812177  0.0826634  -7.031 2.34e-12 ***
## nbhdB8                0.1082066  0.0237556   4.555 5.37e-06 ***
## nbhdB9               -0.3662756  0.0779103  -4.701 2.66e-06 ***
## nbhdC1                0.0091419  0.0754593   0.121 0.903577
## nbhdC2                0.6252085  0.1461493   4.278 1.92e-05 ***
## nbhdC3                0.1850442  0.0412523   4.486 7.44e-06 ***
## nbhdC4               -0.3880513  0.0823931  -4.710 2.55e-06 ***
## nbhdC5               -0.1641621  0.0914592  -1.795 0.072729 .
## nbhdC6                0.2378111  0.0669087   3.554 0.000383 ***
## nbhdC7                0.3776979  0.0790950   4.775 1.85e-06 ***
## nbhdC8                0.6291550  0.0582048  10.809  < 2e-16 ***
## nbhdC9                0.3730017  0.1038862   3.590 0.000333 ***
## nbhdD1                0.0549849  0.0441169   1.246 0.212699
## nbhdD2               -0.2425476  0.0808675  -2.999 0.002720 **
## nbhdD3                0.5403776  0.0513007  10.534  < 2e-16 ***
## nbhdD4                0.1790063  0.0511033   3.503 0.000465 ***
## nbhdD5                0.4206325  0.0997196   4.218 2.51e-05 ***
## nbhdD6               -0.2912549  0.0748413  -3.892 0.000101 ***
## nbhdD7               -0.4833020  0.0858321  -5.631 1.90e-08 ***
## nbhdD8                0.3064015  0.0516085   5.937 3.11e-09 ***
## nbhdD9               -0.1199197  0.0588514  -2.038 0.041638 *
## nbhdE1                0.2460622  0.0516878   4.761 1.99e-06 ***
## nbhdE2                0.2362451  0.0442184   5.343 9.58e-08 ***
## nbhdE3                0.3050080  0.0508457   5.999 2.14e-09 ***
## nbhdE4                0.2006658  0.0774664   2.590 0.009617 **
## nbhdE5                0.5033894  0.0558437   9.014  < 2e-16 ***
## nbhdE6                0.2148343  0.0520513   4.127 3.73e-05 ***
## nbhdE7               -0.0721630  0.0332439  -2.171 0.030002 *
## nbhdE8               -0.4394392  0.0827712  -5.309 1.15e-07 ***
## nbhdE9               -0.0699163  0.0517035  -1.352 0.176358
## nbhdF1                0.0757878  0.0262046   2.892 0.003843 **
```

```
## nbhdF2               0.1790168  0.0484841   3.692 0.000225 ***
## nbhdF3              -0.0263317  0.0439705  -0.599 0.549301
## nbhdF4               0.2023651  0.1474483   1.372 0.169988
## nbhdF5               0.1788995  0.0444920   4.021 5.89e-05 ***
## nbhdF6               0.2633716  0.0495051   5.320 1.08e-07 ***
## nbhdF7               0.3657911  0.0497831   7.348 2.36e-13 ***
## nbhdF8              -0.1417498  0.0605073  -2.343 0.019186 *
## quadrantNW           0.1042935  0.0289311   3.605 0.000315 ***
## quadrantSE           0.0447303  0.0342636   1.305 0.191794
## quadrantSW           0.0171712  0.0646673   0.266 0.790612
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                               edf  Ref.df      F  p-value
## s(sinceRmdl)                5.630   6.769 47.261  < 2e-16 ***
## s(saledate)                 8.433   8.809 39.038  < 2e-16 ***
## s(rooms)                    3.382   4.281  3.411 0.007501 **
## s(latitude)                 7.876   8.515  7.453 2.59e-05 ***
## s(longitude)                8.056   8.759  8.573  < 2e-16 ***
## te(gba,landarea)            7.592   8.630 17.413  < 2e-16 ***
## te(saledate,fireplaces)     5.917   7.685  3.781 0.000179 ***
## te(longitude,gba)           6.183  19.000  3.097  < 2e-16 ***
## te(gba,latitude)            8.092  19.000  2.593  < 2e-16 ***
## te(longitude,eyb)           6.853   8.928  1.248 0.288762
## te(saledate,gba)            2.726  13.000  2.666  < 2e-16 ***
## te(saledate,eyb,ayb)       22.600  30.166  6.961  < 2e-16 ***
## te(gba,latitude,landarea)   8.404  74.000  0.511 7.59e-07 ***
## te(bathrm,gba,eyb)         24.807 100.000  1.135  < 2e-16 ***
## te(gba,landarea,ayb)        2.096  77.000  0.100 0.002742 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 643/644
## R-sq.(adj) =  0.973   Deviance explained = 97.4%
## fREML = -2663.8  Scale est. = 0.017033  n = 4995
```

```r
AIC(fit)
```

```
## [1] -5917.177
```

```r
BIC(fit)
```

```
## [1] -4359.465
```