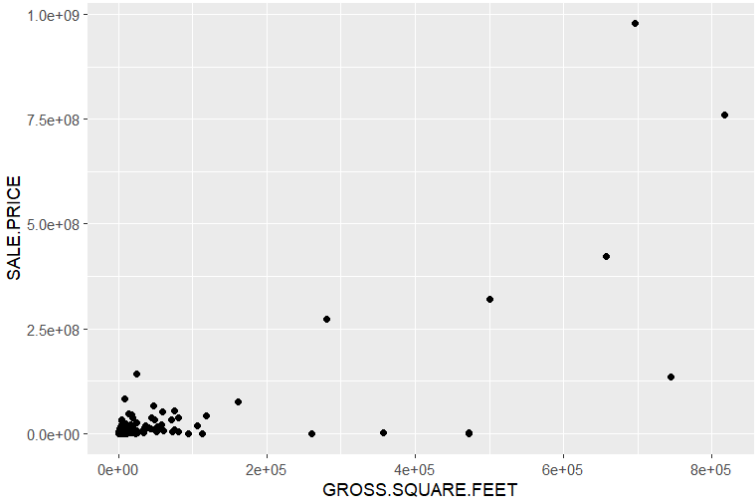
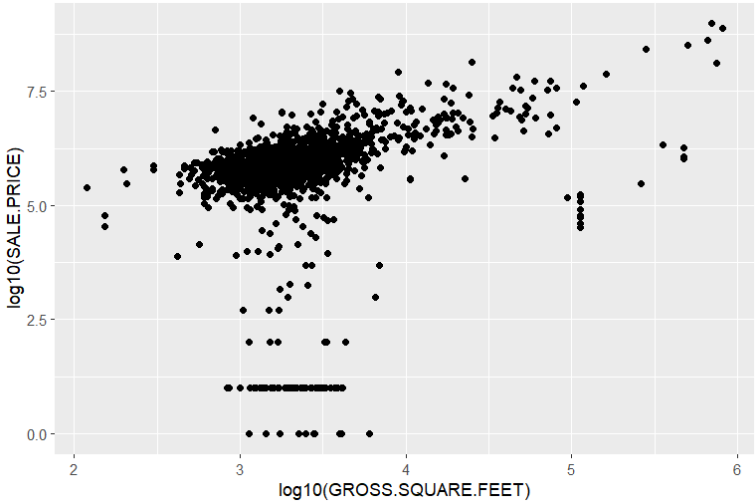
Line 26-28: Creating a scatterplot of SALE.PRICE vs GROSS.SQUARE.FEET

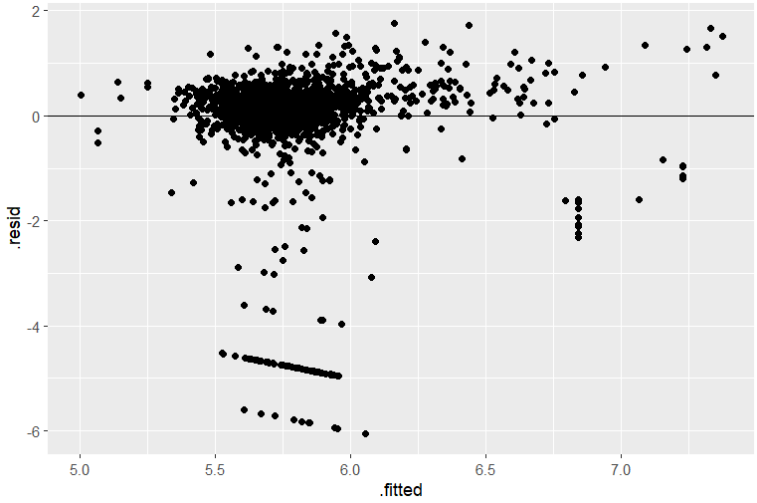


Line 30-31: Creating the same plot as before but on the log10 scale

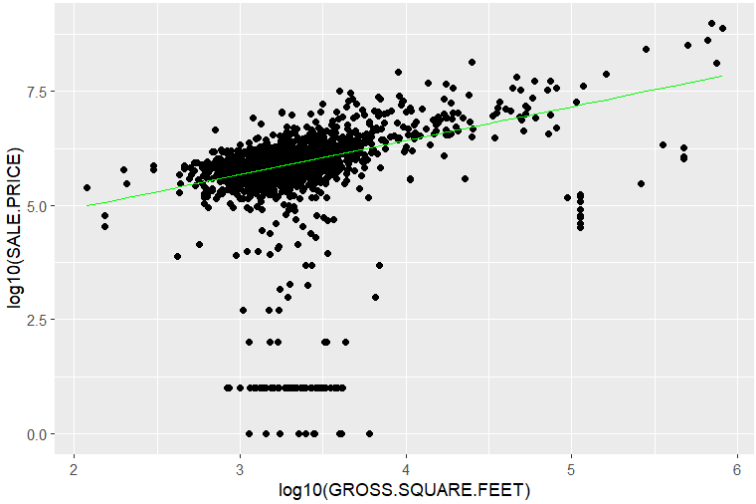


Line 34-37: Plotting the linear model trend line over previous plot

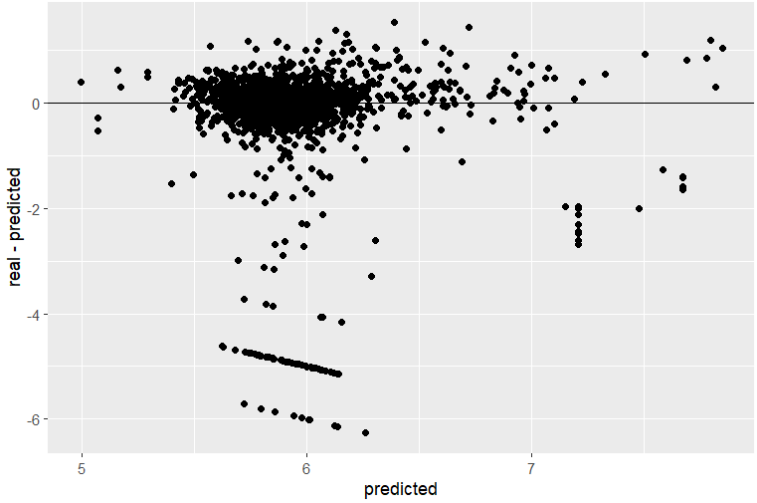
Line 39-41: Plotting the residuals of the linear model



Line 45-50: Plotting the SVM linear kernel model trend line over log10 plot

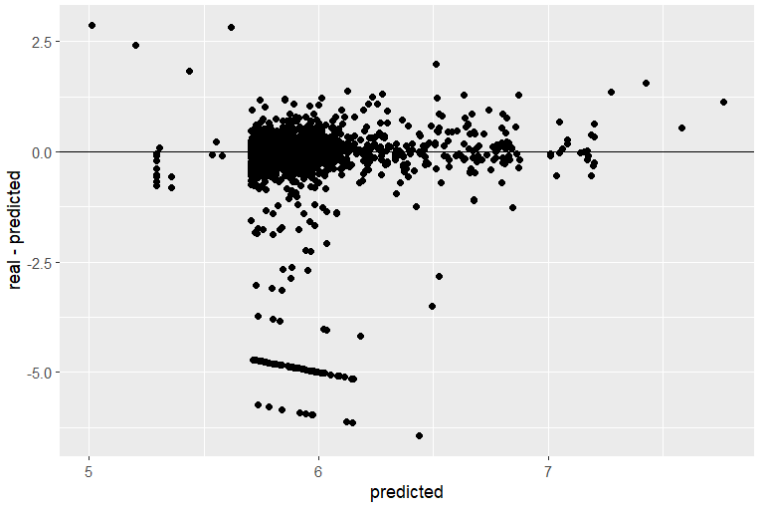


Line 52-54: Plotting residuals of SVM linear kernel



Line 57-62: Plotting the trend line of the SVM radial kernel model over the log10 scatterplot

Line 64-66: Plotting the residuals



Line 69-73: Creating the ranges for the gamma and cost used in the upcoming tuned SVM:

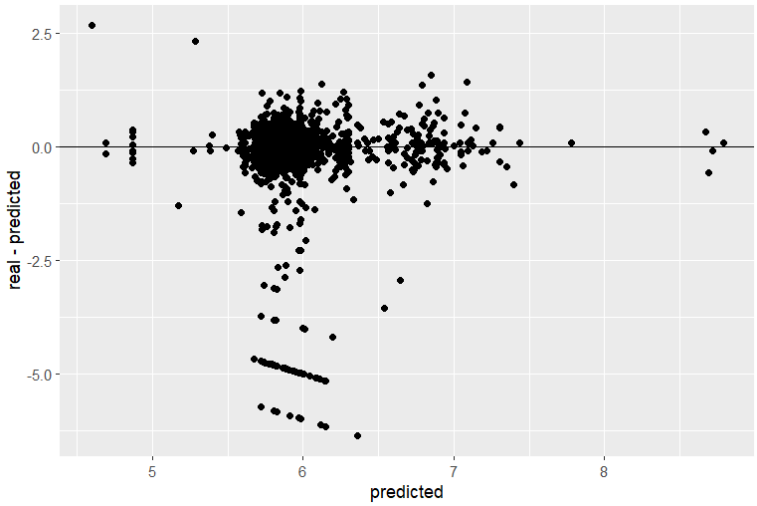
```
> gamma.range <- 10^seq(-3,2,1)
> gamma.range
[1] 1e-03 1e-02 1e-01 1e+00 1e+01 1e+02
> C.range <- 10^seq(-3,2,1)
> C.range
[1] 1e-03 1e-02 1e-01 1e+00 1e+01 1e+02
```

Line 76-79: Tuning the SVM

```
> tuned_dataset <- nyc_data_subset %>% slice_sample(n = 1000, replace = FALSE)
> tuned.svm <- tune.svm(log10(SALE.PRICE) ~ log10(GROSS.SQUARE.FEET), data=tuned_dataset, kernel="radial", g
amma = gamma.range, cost = C.range, tune.control=tune.control(cross = 5))
> tuned.svm

Parameter tuning of ‘svm’:

- sampling method: 10-fold cross validation

- best parameters:
 gamma cost
    10   10

- best performance: 0.9206542
```

Line 84-89: Plotting the trend line of the tuned SVM radial kernel over the log10 scatterplot

Line 91-93: Plotting the residuals



Line 107-219: Mean Absolute Error, Mean Square Error, and Root Mean Square Error over 100 k folds using Monte Carlo on the Models 1, 2, and 3 (Linear Model, SVM kernel=Linear, and SVM kernel=Radial respectively)

	mae	mse	rmse
1	0.4007624	0.7820272	0.8813665
2	0.3612353	0.8079171	0.8958912
3	0.3584261	0.8288916	0.9065581

Based on the metrics above, the SVM radial kernel seemed to perform the best, as despite having a higher over MSE and RMSE, it had a drastically lower MAE value. This means that it did the best at minimalizing the distance between errors. Conversely, the Linear Model performed the worst since it had overall the highest MAE by far.