

SalePrice ~ Neighborhood

SalePrice ~ Neighborhood + GrLivArea

SalePrice ~ Neighborhood + GrLivArea + ExterQual

SalePrice ~ Neighborhood + GrLivArea + ExterQual + MSSubClass

SalePrice ~ Neighborhood + GrLivArea + ExterQual + MSSubClass +
GarageCars

SalePrice ~ Neighborhood + GrLivArea + ExterQual + MSSubClass +
GarageCars + Fireplaces

SalePrice ~ Neighborhood + GrLivArea + ExterQual + MSSubClass +
GarageCars + Fireplaces + YearRemodAdd

SalePrice ~ Neighborhood + GrLivArea + ExterQual + MSSubClass +
GarageCars + Fireplaces + YearRemodAdd + HouseStyle

SalePrice ~ Neighborhood + GrLivArea + ExterQual + MSSubClass +
GarageCars + Fireplaces + YearRemodAdd + HouseStyle + Foundation

SalePrice ~ Neighborhood + GrLivArea + ExterQual + MSSubClass +
GarageCars + Fireplaces + YearRemodAdd + HouseStyle + Foundation +
YearBuilt

SalePrice ~ Neighborhood + GrLivArea + ExterQual + MSSubClass +
GarageCars + Fireplaces + YearRemodAdd + HouseStyle + Foundation +
YearBuilt + BedroomAbvGr

Lab 3 was honestly a mess. I learned a lot but was set, by my own hand and by circumstances, to fail. For what it's worth I, Kevin Elkin, did all the work for this assignment with absolutely no assistance from my other group members. I understand this does not excuse the quality of the work provided, do as you will.

Above are the models created by the forward selection method. Based on the results depicted in the output for the set function the model gets notably better with each iteration. The final step to the full function is only marginally better I would still consider this more complete model to be the best. The reason for this is that the added predictor variables intuitively make sense in contributing to greater

sales prices. Though some variables are more intuitions relating to other overarching considerations of value, such as fireplaces. Fireplaces are not likely to yield a greater sale price by the merit of their existence but they are tied to the nature and style to the house in a way that may not be entirely encompassed by other variables. In this way I would consider the final output better than the lesser models.

I struggled with null values. This was by the far the most challenging part of creating the models. I do not believe I utilized the scope input of the function correctly. I tried to avoid variables with null values. Many of which were seemingly unimportant anyway. I still do not fully understand how one can perform the tasks at hand without replacing these values, and even still I would assume their to be other issues associated with doing so. I ran out of time to properly organize the outputs in a plot. This would have easily been something that could have been performed with assistance. That being said I do understand the intuitions of the how the RMSE works and how it can be displayed in the manner requested. Similarly, when working towards the final model I would have run the set function to run a backward selection method as well a simultaneous direction selection. I also would have worked to increase the scope and manage the null values by either replacing them with 0's or working to find a better way for them to be implemented. I should have started the assignment earlier and coaxed my group more vigorously to assist, with this hindsight the task could have been better accomplished. I will not fall victim to this again.