Following Apapiu's notebook, the sales price data was a normal distribution skewed right as pointed out. Following his notebook, I shifted the sale price data right to normalize it. Other than that, I one-hot encoded all the categorical variables and I replaced all empty entries with the mean of their respective column. After preprocessing the data, I decided to compare a ridge regression model with a lasso regression model. After optimizing the alphas of each through cross-validation, the ridge regression model yielded a score of 0.12479 with an alpha of 10 while the lasso regression model yielded a 0.12455. I decided to also try a XGBRegressor model (parameters: n_estimators=360, max_depth=2, learning_rate=0.1) which yielded a worst score of 0.13278. In an attempt to improve my score, I decided to follow Apapiu's suggestion and take a weighted average of the lasso model and the XGBRegressor model. Taking 30% of the XGBRegressor predictions and 70% of the Lasso predictions, the score of the prediction came out to 0.12299. Looking back to what worked and what didn't, optimizing the alphas for each model through cross-validation greatly improved the score and model's performance. Furthermore, taking a weighted combination of the Lasso predictions and the XGBRegressor model yielded the best results. On the contrary, simply using an XGBRegressor model did not further improve performance when compared to the Lasso and Ridge models.