

Notes to Price Distribution

One key aspect of exploratory data analysis is investigating the *distribution* of the outcome variable. In our case the price distribution is highly right-skewed with a few very expensive listings pulling the mean and median of the price distribution further away from each other.

Some statistical models such as *Linear Regression* tend to perform better when the outcome distribution is symmetric and approximately normal, whereas some very flexible algorithms like *Neural Networks* do not make any distributional assumptions and are capable of modeling any kind of distribution accurately. Figure 1 illustrates an approximate normal distribution can be achieved with a simple logarithmic distribution.

Whereas *all* of the classical models benefitted from the log-transformation resulting in lower error metrics, this was not the case for the Neural Network we used.

In fact, training turned out to be more challenging, since the *magnitude* of the losses by comparing true and predicted price on the logarithmic scale was drastically reduced, leading to smaller gradients and thus smaller weight updates. This issue could be mitigated to some extent with a larger learning rate. However, in contrast to the untransformed version, the network still suffered from *vanishing gradients* and *loss plateaus*.

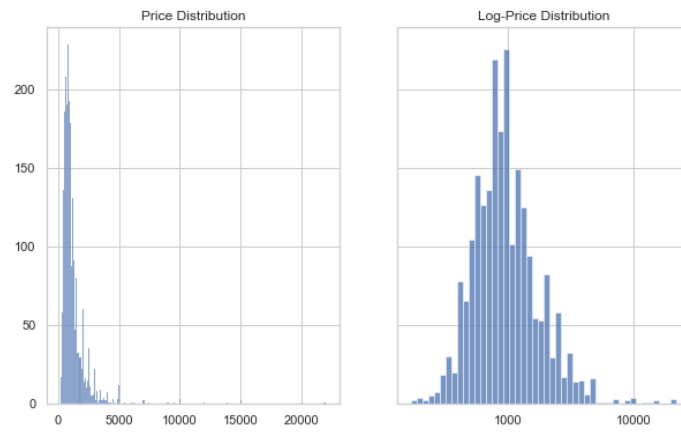


Figure 1: Distribution of Apartment Prices on original and logarithmic Scale