Final Project Executive Summary

We are interested in predicting the total rent for apartments with different attributes. Utilizing 138 observations, we use the following variables to help predict log(total rent): number of rooms, price trend, living space in km, whether there is a balcony, and the service charge. Collectively, these variables work to explain over 77% of the variation in its linear relationship with total rent charge.

To come to this final model, we first started with a larger model that included along with the variables mentioned above, also included both firing type, a categorical variable, and the interaction between the price trend and the number of km of living space. Additionally, this model originally predicted total rent instead of logged total rent. We decided to log total rent because after creating a box cox plot, we found that it would be most ideal to log our outcome variable. From here, we used a backward stepwise search to find the model with the lowest AIC and then compared its adjusted R^2 value with that of the larger, original, nested model.

This model would allow a realtor or anyone working with apartment prices make better predictions regarding how much an apartment costs to rent or how much an apartments rent should be set at. The model has good overall significance and practical predictor variables that customers/sellers would consider when renting/renting-out an apartment. This could also allow customers to see if they are being over/under charged for an apartment based on its attributes.

We see from the output of our model that all the variables are significant at the 5% or smaller level, with the largest coefficients being for whether there is a balcony and the current price trend. Because whether an apartment has a balcony can at most be equal to 1, we know it will not affect the total rent greatly because at most it adds 10 cents to total rent. On the other hand, the current price trend has a much greater ability to change the overall rent because the price trend is continuous, allowing it the possibility of having a rather large effect on total rent. I would also take a close look at the livingspace variable because it is likely that the number of km of living space in an apartment is rather large, increasing the rent greatly without any way of reducing rent as we will not see negative numbers for km. Laslty, I would advise to be cautious about the coefficient for the number of rooms because although it is negative, we see somewhat of a parabola when number of rooms is plotted against total rent. Moving forward, I would consider a model that possibly has a squared term for number of rooms.