Project Progress Report

Author: Zifeng Chen

Group members: Ailong Chen,

Haowei Cheng

- Our group project is using several factors such as the number of bedrooms, the area of the pool, etc. to predict the house price.

- My role in this group is doing a literature search with Ailong to find the possible variables that significantly affect the house price. Then, I would tell Haowei who would use the linear regression model to train the data. Finally, we would use what we learn in class to make our model better.

- So far, I have found some significant numerical variables: The building class, Overall material, finish quality, Overall condition rating, Original construction date, etc. I am working on getting the correlation between the variable so that I can remove some of the variables.

- **The challenges:** we have a total of 80+ variables including both categorical and numerical variables, and when we use the linear regression model to train the data, some variables would cause the error. It is hard to figure out which variables are good for us to use. The other challenge is that it is the

first time for me to do the feature selection when the target variable is

numerical (house price). I need more time to do the research.