- **DataCleaning.py** in Data Cleaning file used the original train and test file from Kaggle to generate cleaned data, it includes log transformation to

- 15 models were stacked as the first layer, and 2 models were stacked as the second layer, the Kernel Ridge and Lasso were both tested as the final blending model, and have equivalent performances

- The whole stacking and blending process will be repeated twice, the output of the first process will be used for outlier detection based on outlier test function in statsmodel. The second process first excluded all the outliers then rerun the whole process again. This help the model improve the log RMSE by 3%

- The files in the parameter tuning file were used to tune the parameter of each individual models