**Stat 502 Class Project Weekly Updates**

**Competition:**

**Team Kaggle Name (if Relevant): \_\_\_\_\_\_\_\_\_\_\_Y\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Team Members: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Eryn Blagg,\_\_Yiqun Jiang\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Week 10 (March 23-27) Summary of Activity and Progress**

Preprocessing: we transformed waterfront and zip code to be factors and delete date and property.

Model: This week we tried to use a simple linear regression model to predict house price but one problem we meet is with this method, we may get negative predictions. It’s commonsense that house price should be positive so we modify our model a little bit to be absolute linear regression model. We forced the prediction to be positive to get more reasonable result.

Future plan: one thing we can do to improve our results I think is input engineering. Besides we would try more models.

**Week 11 (March 30-April 3) Summary of Activity and Progress**

**Week 12 (April 6-10) Summary of Activity and Progress**

**Week 13 (April 13-17) Summary of Activity and Progress**

**Week 14 (April 20-24) Summary of Activity and Progress**

**Week 15 (April 27-May 1) Summary of Activity and Progress**

**Week 16 (May 4-May 8) (Finals Week) Summary of Activity and Progress**