Self- Assessment

My role in the team project was rotating every week – database, github and ML model. For Database – my role was to find the good data and then designing and data processing. We used pgadmin for our database. For Github I was involved in setting up branch and then merging the final files. For ML model, I was involved in initial running model with sample database. Then, creating random forest and Logistic regression model on the full database.

I have used resources help from instructor, and TAs which was great to develop deeper understanding of data processing and ML model. Team discussions were greatly helpful as well.

The greatest personal challenge was collecting the data and then troubleshooting ML model. Once that was fixed, model ran ok.

Team Assessment

Maria, Sam and I worked together for the entire project on resolving, troubleshooting and creating an action plan for each member each Tuesday and Thursday when we met. Then we continued the discussions through slack.

Summary of Project

Our topic: Predict housing prices in New York City with additional features around broad economic factors.

We started out with Logistic regression model and finalized with Random Forest Model. Our model used unemployment, income, # of bedrooms, # of bathrooms, Square Feet by Zip codes to predict the housing prices in NYC. We used flask web application for front end Forecasting and backend was Random Forest Model. The model was 74 % accurate.