**Self-Assessment**

In this project, I took on the triangle role. I developed the initial version of our machine learning model, recommending the Random Forest Regressor as the main model to begin with. I researched and recommended the code for the Gradient Boosting model. I researched model optimization for the Forest regression using grid search cv. I wrote the code for the ml model preprocessing, feature engineering, splitting and scaling, and for the random forest regression modeling and the Hist Gradient Boosting model. I recoded the RF regressor to a classifier because I had made the error of using the stratify parameter in test-train-split, which caused the model to throw a quantity error. I binned the offending variable (our target variable) into discrete groupings for the classifier, an effort that proved superfluous once one team member determined the source of the error was using stratify. However, it was productive as an exercise in feature transformation. We returned to the regression model after removing the stratify parameter.

At the same time, I began the GitHub readme documentation. At the point where I had run the preliminary models, I suggested someone else on the team take on the ml modelling, and I focused on preliminary statistical analysis of the data set and writing the GitHub documentation.

Database architecture was mainly handled by two other team members. The third member focused on slide presentation and visuals. I participated in discussions about the data that we wished to include in the analysis beyond that which was contained in the main dataset. I advocated for inclusion of economic data and climate data, and after some research, supported the inclusion of crime data in the database (crime and economy is impossible to separate as a complete analysis). I disagreed with the decision to include weather data, and recommended substituting climate data. I reviewed the slides and the visuals, but to be honest, felt the work was of such high quality I found little to contribute other than to suggest red-gold palette for one visual and to request features be ordered by magnitude in another.

My biggest challenge is my very personal and characteristic fear of having my input and contributions be disregarded. The team thoroughly discussed source options for specific economic information and climate data as well as for crime data. Because we were working under tight time constraints, we all agreed to move the project forward with the sources we had identified at the beginning of the first segment. But I was concerned that some of the data decisions were somewhat arbitrary and did not serve a sound outcome, and that obstructions could have been overcome with less effort than generally acknowledged. My challenge in this process and subsequent discussions was to center myself, review my assumptions and perceptions, and re-equilibrize internally. It is an ongoing process.

**Team Assessment**

I have a great deal of respect and liking for my team, both as a whole and individually. In my view, we were a bunch of high achievers who were working on a project we were excited about and that was personally important to each of us. Inevitably, there were disagreements. The first big issue we faced was getting to a common understanding of the requirements for the project from the instructional side. The question we had to settle was which was more important: a sound statistical outcome to the model, or demonstrating our skills in specified techniques. I contacted the instructor to gain clarification, and understood the presentation to be the most critical. However, the instructional materials were very specific about grading criteria, and in the end, based on my team-mates’ judgment, we approached both equally. This was the proper approach, and I appreciated the team process.

Conflict and disagreements are part of human interaction. They are healthy processes. Conflict is dynamic and nurtures insight and creativity. It brings different perspectives to the fore and forces us to re-examine our ideas and assumptions. You cannot avoid disagreements, nor should you. The highest principle of human interaction is in managing ourselves, including conflict. That being said, there is such a thing as conflict based in pathology. That type cannot be managed in normal terms. This was not the case in my group. My team engaged in healthy, open disagreement. I may not have been completely happy with the outcome, but I trust my team and respect their abilities. I hope to continue association with each. I would counsel anyone going through this process to embrace team communication management, and in fact would recommend it be directly addressed as part of the initial strategy discussion with an agreed-upon protocol.

Lastly, I was lucky to be with people who were so engaged and excited about the project that we probably erred on the side of doing too much. No lack of abundance here! Our presentation ran a little long and our GitHub is deeply detailed, but I felt we supported one another throughout. Ours was a great team.

**The Project**

The housing prices project created a machine learning model to predict median house values in California based on the US government's Census data from 1990. The objective was to identify which features impact housing values and to train a model which can predict median house prices in any area with those features.

Following the sequence of data preparation, database creation, feature engineering and selection, the analysis proceeded comparing three machine learning models for best fit: Linear Regression, Random Forest Regressor, and Gradient Boosting Regression. After optimization and comparison, the optimized Gradient Boosting Regression demonstrated the most success in terms of accuracy and predictability.