G.L.A.P.S Progress Report:

We currently have a model that predicts the median home values of homes in any county within the United States for the year of 2011 with an absolute percent error of approximately 4.2%. To obtain the prediction the normalized values for all of the attributes must be provided to the model.

Additionally, we also currently know how to use TensorBoard to display the machine learning process that our model is completing, as well as, varying graphs that display the effectiveness of our model to predict the desired output.

We were able to obtain our data using the census APIs and then visualize it with Facets. Visualizing our data helped us to notice any places where a default number may have been set that could skew the normalization and in turn our predictions.

Currently, we are working on moving forward to the creation of a time series forecast model. We will require a model that will have memory layers for the data each year to help the model determine trends which it will use in its prediction. We are focusing on LSTM models to complete this task.

Our issues currently are based on needing a deeper understanding of the subject and how to complete the model, which we will attempt to resolve through research and trial and error. The other issues we are facing are finding the most effective way to train the model. Our current thought is to focus solely on one state so that there aren’t as many variables for with which the machine has to deal.

In focusing on the one state (in this case NC) we then faced issues with how to format our data once again so that it is split into arrays that hold the data for all the years for each county within its own array. We were able to complete this by sorting the data with pandas and then reshaping the array with numpy.

From here we will now continue to research and focus on the creation of a model that can create a future prediction regardless of accuracy. We will apply our knowledge of TensorBoard and saving models to help us visualize which factors are helping or hindering our model.