I have provided two data sets: a training data set and a predictions data set. The training data set (full description below) contains ozone measurements as well as other factors. The prediction data set contains all of the factors but is missing the ozone measurement.

1. Use the training data set to build a regression tree with ozone (O3) as the outcome.

2. Use the tree you built to predict the missing values of ozone based on the observed factors.

Whoever's prediction minimize mean squared error (MSE) wins a prize or something.

Training data: http://bit.ly/1tfXc1t

Prediction Data: http://bit.ly/1DYRa8R

- O3: Ozone concentration (parts per million)

- vh: Vandenburg 500 millibar height (in)

- wind: wind speed (mph)

- humidity: humidty (Percent)

- temp: temperature (Celsius)

- dpg: Pressure gradient (mmhg)

- ibt: Inversion base temperature (Fahrenheit)

- vis: visibility (miles)

- doy: day of year