Checkpoint 1

1. **What did you do this week?**
   - Setup GitHub repo.
   - Combined several datasets from state level economic, environmental, and health factors. Also added a SARS dataset.
   - Commit to repo with linear regression using R.
   - Ran a time series analysis using Tensorflow and received an initial score of 18.5%.
   - Using AutoML to explore potential deep learning, machine learning, and time series models then will export the model via an API.
     - Exploring different ways to train and test the model against SARS data.
   - Standardized data from the initial dataset posted by John Hopkins University Github into 1 dataset based on Date, Province State, Country Region, Confirmed, Deaths, Recovered.
   - Developed Task List.
   - Attempting to configure SparkR, RStudio, and the Grid with the HPC team.

2. **What are you planning to do next week?**
   - Finalize datasets and data sources.
   - Make a commit to GitHub to complete linear regression from Regression_R file in GitHub.
   - Complete a time series model expressing a metric like Mean Absolute Error using AutoML as well as the increase the Tensorflow notebook up from 18.5%
   - Configure SparkR, shinyjs, Shiny, and the Grid correctly.
   - Explore dashboard layouts and style.
   - Prepare the data and setup Spark Stream using the Grid.

2. **What problems are you having?**
   - Initially, my biggest problem was overall project organization and understanding what tools could be used.
   - Lack of complete data for factors at the city/county/zip code level.
   - Cleansing the data for a standardized dataset.
   - X11 Forwarding gave me issues for display. Had to export dbus, first.
   - Time needed for model training, development, and selection.
   - SparkR was not an installed package on my node through the Grid.
   - Issues launching RStudio because of X11 Forwarding errors.