RELATIONSHIP BETWEEN THE FILES

- **Model_10_10_2016.ipynb**

  is used to create the models and make the necessary predictions.

- **Sosumi.aiff**

  is a sound file used by the programs in
  Model_10_10_2016.ipynb as a reminder for when the model is done fitting the data

- **urban_poverty.R**

  is an R file used to add a couple of variables to the data. See the file for details of the program. I left comment in them to help a curious mind figure out what the program does.

- **Model_files**

  is a folder of all actual model that has been written out onto the disk. Please none of the files in the folder should be tempered with.

How to open the Model_10_10_2016.ipynb

To open the file

1. Have anaconda installed. In windows, it comes with the anaconda navigator but in linux it doesn't and one would need

to visit [anaconda webpage](#) to follow their instruction to setup.

2. using the command prompt in windows or the terminal in linux, **cd** into the directory that houses that files in this folder(Isaac).

3. Once you are in the folder containing the files of the models, type:

   `jupyter notebook`

   This will start up the jupyter notebook.

4. Feel free to chane the directory to point to where the files sit on your computer

Running the Models/ Making predictions

Each cell in the notebook must be executed in a linear function starting from the top of the file. If the already fitted model doesn't work on your computer then it's likely to be a mismatch between the version of the jupyter notebook I used in fitting the model and the one you are using. As at the time of creation of this document[11/28/2016], files dumped to `"model_files"` needs to be loaded with the current version of jupyter notebook. In the unlikely event that you are unable to open the files, I will advice to fit the model again. Since the work is reproducible, you have the same results.

To fit the model

Run the cells in the section `"Building models"`. The first cell builds with the Random forest, the second cell builds with Gradient Boosting

and the third cell builds the Logistic model. Please note that you have to build the model one after the other, that is you have to allow each cell to finish executing the program in them before running the next cell.

NOTE: you need to change the directory where you want to models to be written to before running each cell.

Making predictions

Run the codes in the section `"making predictions"`. Once again change the file path to reflect where you saved the models in the previous sections. Now continue executing the cells in a linear fashion.

Please me at `isaac.oduro@case.edu` if you have any question.