# Homework 3:

# Advanced Data Analysis in Python

Due before class Wednesday, April 15, 2020

The purpose of this homework is to familiarize you with hierarchical Bayesian analysis as implemented through the Stan programming language. There is a file on GitHub titled trend2.csv. Read this data into Python. The variables of interest are as follows:

- Outcome: church2 — this is a measure of religiosity in a country.
- Explanatory: gini_net — this is a measure of inequality in a country.
- Control: rgdpl

Run a hierarchical linear model with the above variables, and include random intercepts for country and year. For the model, use diffuse priors for all parameters. Next, change the $\beta$ prior for the main explanatory variable to be highly informative, but relatively far away from the posterior estimate obtained from before. How is the posterior different between the two models?

These models should be run in Stan through Python. Submit your code, and just briefly answer the question in a comment at the end of the file. I only need the one file for this homework, no test file is needed. IF YOU CANNOT GET STAN RUNNING AND YOU HAVE SPOKEN WITH ME ABOUT IT, write the Stan code and submit it, without actually running the models, or collaborate with a classmate who has Stan working.