Capstone Project 2: Project Proposal

Of the four proposed capstone ideas I have decided to go with the Rulers, Elections, and Irregular Governance Dataset (REIGN) dataset.

I subscribe to a bi-weekly, emailed newsletter containing a curated list of new and interesting datasets called "Data is Plural", edited and maintained by Jeremy Singer-Vine. I recently received one such newsletter containing several interesting datasets, but the REIGN dataset, in particular, sparked my curiosity.

The REIGN dataset describes political conditions in every country each and every month. These conditions include the tenures and personal characteristics of world leaders, the types of political institutions and political regimes in effect, election outcomes and election announcements, and irregular events like coups, coup attempts and other violent conflicts.

The dataset covers more than 200 countries for each month they were independent, January 1950 to the present. REIGN was created by gathering original data, compiling other datasets on political conditions, reviewing their coding rules, and updating all information to the present.

I would like to use this dataset to make predictions about coup attempts. Specifically, I would like to analyze leader characteristics to determine which features are the most important in explaining the likelihood of coup attempts utilizing a Random Forest Classifier.

I then plan on using this Random Forest model to determine country by country coup probabilities and to create a ranked list of countries ordered from most likely to have a coup attempt to least likely.

I would also like to explore other questions, such as which features contribute to successful versus unsuccessful coups. I plan on doing this by subsetting the data to analyze only coup attempts and to rerun my model on this subsetted data to determine which features contribute to successful coup attempts.