Brooke Ann Coco
Project McNulty Proposal
To Coup Or Not to Coup?
April 24, 2019

Motivation:
In their 2011 book, *The Dictators Handbook: Why Bad Behavior Is Almost Always Good Politics*, authors Bruce Bueno de Mesquita and Alastair Smith argue the underlying motivation of any ruler, no matter how good their intentions, is to remain in power. Thus, all rulers, whether they be democratic or authoritarian, are bound by the same sets of political rules, which the authors outline throughout the remainder of the book. To that end, this project aims to utilize supervised learning techniques to classify whether a sovereign nation will experience a coup attempt within a given year.

Data:
The data for this project will be obtained from the following datasets:

1. OEF Research's **REIGN (Rulers, Elections, and Irregular Governance)** dataset contains information about the political conditions in over 200 countries for each month for each month they were independent spanning from 1950 to the present.
2. The **World Bank** has a dataset of Gini indexes for each country spanning back to the 1950s, though there are quite a few missing data points. I am still in the process of searching for a more complete data set.
3. The **US State Department** manages a database that records all US foreign assistance against nearly 300 indicators.
4. The **Penn World Table** contains information on relative levels of income, output, and productivity covering 182 countries between 1950 and 2017.
5. And MORE!

Features:
Throughout their book, Bueno de Mesquita and Smith outline several methods through which political leaders maintain power. The features of my model attempt to capture some of these methods. Information in red is subject to change.

Feature Name	Variable Type	Description	Source
coups[*]	Categorical	Identifies whether a coup attempt occurred	REIGN
regime_tenure	Continuous	Number of years a specific political regime has been in place	REIGN
military_career	Categorical	Identifies whether the primary career and/or source of legitimacy of a nation's leader stems from a career in the military, police force, or defense ministry	REIGN

[*] Target variable

regime_ type	Categorical	Type of political regime (out of 16 regime types)	REIGN
prev_conflict	Categorical	Identifies if a violent civil conflict occurred in the country that year	REIGN
gini	Continuous	The Gini index measures wealth distribution and ranges from 0 (perfect equality) to 1 (perfect inequality)	World Bank
aid_received	Continuous	**Currently:** Amount of foreign aid received from the US per year **If Possible**: Total amount of foreign aid received	U.S. Department of State
aid_given	Continuous	**If Possible**: Total amount of foreign aid given	TBD
living_standard	TBD	**Looking into dataset to determine best measure of relative living standard**	Penn World Table
gdp_ per_capita	Continuous	Nation's GDP per capita by year	Penn World Table
nat_resources	Continuous	Percent of national revenue gained from natural resources	TBD

I am also tossing around a few more feature ideas that I have yet to explore. They will be solidified by tonight.

Known Unknowns:

I realize that my data is going to be messy and that there is the potential for a lot of information to be missing. However, if my goal is to work with nonprofits or NGOs, I need to get used to working with sparse or incomplete data to tackle meaningful issues. This may mean that I will need to be creative in the way that I account for missing data points.

Also common among nonprofits and NGOs is the tendency to tackle complex issues with no clear solution. This is certainly the case within this project. Thus, it is likely that my final model will not have much, if any, predictive power. The point of this project is not to perfectly predict when a coup will occur in a given country. Instead, the goal is to discover which features are better indicators of a coup attempts as well as to better understand how certain feature variables interact with one another to make coup attempts more likely. To this end, I would be happy with a model that has the ability to generate any nuggets of insight. I see this project not only as a way to practice supervised learning techniques and workflow, but also an exercise in working with complex problems.