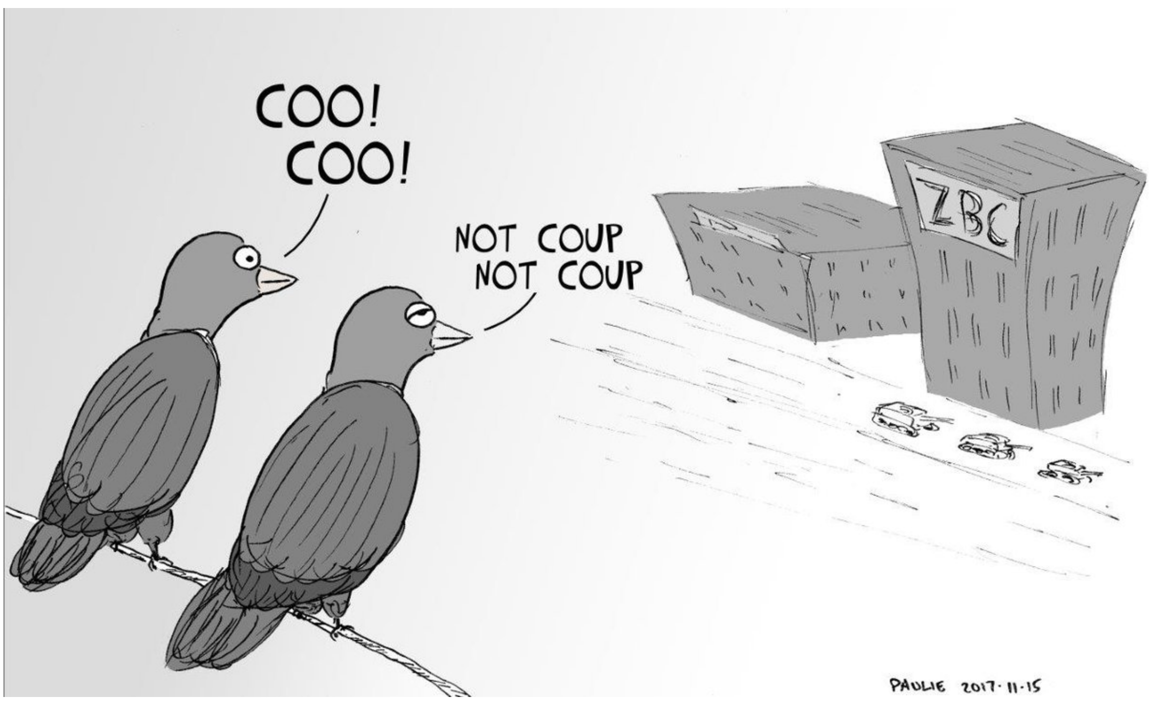
CAPSTONE 2

A machine learning analysis of coup attempts in Python using a variable threshold Random Forest Regressor model

5/23/19

RYNE SCHULTZ



Part I – Data Wrangling

Problem Statement

What factors contribute to coup attempts? A coup – or coup d'état (literally meaning a "stroke of state" or "blow against the state") – is an overthrow of an existing government. Typically, it refers to an illegal, unconstitutional seizure of power by a dictator, the military, or a political faction.

Many political scientists have analyzed coups – particularly the rise in coup attempts during the Cold War and post-Cold War period. Bruce Bueno de Mesquita and Alastair Smith in their study “The Logic of Political Survival”[1] asked a simple question: why do autocratic leaders who produce corruption, war, and misery often endure for decades in office?

In their estimation, two factors govern the longevity of autocratic regimes: the “selectorate” – the set of people with a say in choosing leaders and with a prospect of gaining access to special privileges doled out by leaders, and the “winning coalition” – a subgroup of the selectorate who maintain incumbents in office and in exchange receive special privileges.

Their theory is simple: every dictator answers to some group that retains them in power – their winning coalition. If the leader loses the loyalty of a sufficient number of members of the winning coalition, a challenger can remove and replace them in office. Coalition members, themselves, are drawn from the broader selectorate. The incentive to defect from the incumbent to a challenger depends on the prospects of being included in the challenger’s winning coalition if they should replace the incumbent.

The larger the selectorate relative to the winning coalition, the smaller the chance that a given member of the current leader’s coalition will be included in the challenger’s new winning coalition. Therefore, in political systems characterized by small winning coalitions and large selectorates – as is common in many types of autocracies – supporters of the leader are particularly loyal because the risk and

[1] Bruce Bueno de Mesquita & Alastair Smith & Randolph M. Siverson & James D. Morrow, 2005. "The Logic of Political Survival," MIT Press Books, The MIT Press, edition 1, volume 1, number 0262524406, January.

cost of exclusion if the challenger comes to power is high. Thus, we should expect these types of regimes to persist, and for coup attempts to be rare.

However, when we look at the data, we see that coup attempts are hardly infrequent. Indeed, since 1950 there have been 464 coup attempts, of which half have been successful. And even more puzzling, of the 464 total coup attempts, more than three-quarters took place under *autocratic* regimes.

Brett Casper and Scott Tyson analyzed this paradox in their study "Popular Protest and Elite Coordination in a Coup d'etat"[2] In their analysis, Casper and Tyson recognize that "elites" (similar in definition to Mesquita and Smith's "selectorate") face a daunting coordination problem when contemplating a coup. Citizens, who desire political reform, face a similar coordination problem when contemplating protest. Since elites and citizens interact with the same regime, these coordination problems are invariably linked. Specifically, Casper and Tyson posit that *protests* aggregate citizen information and provide elites with a public signal which helps them coordinate in a coup.

Therefore, those factors that contribute to public protest in turn signal to elites the potential weakness of the current regime. Thus, if we better understand those factors that lead to public discontent, we may better understand those factors that lead to coup attempts.

Data Description

Towards this end, Curtis Bell assembled the Rulers, Elections, and Irregular Governance (REIGN) dataset[3], a dataset that describes political conditions in every country each and every month. These conditions include the tenures and personal characteristics of leaders, the types of political institutions in effect, election-related outcomes and announcements, and irregular events like coup attempts and other violent conflicts. By analyzing this varied set of factors, we may better understand which factors contribute to coup attempts.

---

[2] Casper, Brett and Tyson, Scott, Popular Protest and Elite Coordination in a Coup d'etat (July 17, 2013). Available at SSRN: https://ssrn.com/abstract=2105409 or http://dx.doi.org/10.2139/ssrn.2105409

[3] Bell, Curtis. 2016. The Rulers, Elections, and Irregular Governance (REIGN) Dataset. Broomfield, CO: OEF Research.

The dataset contains 38 columns, which can be broken down into 5 main categories:

1. Observation Identifiers
    a. **ccode –** a one to three-digit number that uniquely identifies each of the 201 countries included in the dataset.
    b. country – a three-letter abbreviation for each country in the dataset.
    c. **leader –** the de facto leader's name.
    d. **year –** the calendar year, ranging from 1950 to the present year.
    e. **month –** the month number, ranging from 1 (January) to 12 (December).
2. Leader Characteristics
    a. **elected –** 0 for months when the incumbent leader has never been elected to the highest office and 1 if the leader has been elected to that office. Note, elections need not be democratic or competitive. The variable is equal to 1 whenever a leader's rule has been legitimized through some form of election.
    b. **age –** an approximation of the leader's age calculated by subtracting the leader's birth year from year. It takes the same value for each month in a calendar year.
    c. **male –** equals 1 if the leader is male and 0 if the leader is female.
    d. **militarycareer –** equal to 1 if the leader's primary career and/or source of legitimacy or authority comes from his or her career in the military, police force, or defense ministry. Previous service is not sufficient. Rather, this designation is reserved for those whose primary affiliation prior to taking power can be described by these specific career categories.
    e. **tenure_months** - counts the number of months the leader has been in power in his or her present term. The inauguration month takes a value of 1. **tenure_months** resets every time a leader leaves and reenters office.
3. Regime Characteristics
    a. **government** - places each month into one of 16 regime types:
        i. Democracies
            1. *Presidential Democracy* – Democracy in which the executive is distinct from the legislative branch and

considerable decision-making authority is granted to the executive. Presidential systems have presidents who serve as chief executives rather than figureheads.

2. *Parliamentary Democracy* – Democracy in which legislatures are more powerful and executives are less autonomous. Generally speaking, countries with powerful prime ministers and general elections are parliamentary democracies. Hybrid semi-presidential systems are classified case-by-case but are usually grouped with parliamentary democracies.

ii. Non-Democracies

1. *Personalist Systems* – Power is highly concentrated in the hands of a non-monarch dictator who is relatively unconstrained by a military or political party. Contemporary examples include Russia, Sudan, and Chad.
2. *Monarchies* – Power is highly concentrated in the hands of a monarch who is much more than just a figurehead. Contemporary examples include Swaziland, Kuwait, and Morocco.
3. *Single-Party Systems* – Power is held by the head of a party. Executive power is effectively checked by the party or ruling committee. Contemporary examples include China, Angola, and Ethiopia.
4. *Oligarchies* – Power is held by the head of party, but unlike other single-party systems this party explic- itly represents the interests of one elite segment of society. Past examples include apartheid-era South Africa and Rhodesia under Ian Smith.
5. *Party-Personalist Hybrids* – An intermediate hybrid where a party apparatus supports a dictator, yet the party's identity is concentrated around the person in power and it has few meaningful checks on executive power. Examples include Eritrea, and North Korea.

6. *Military Juntas* – A military committee runs the country. One officer typically serves as head, but this head serves the interests of the committee and his power is checked by other members of the military. Recent examples include Thailand and Algeria.
7. *Indirect Military Juntas* – The military has de facto power, but rules behind a civilian puppet.
8. *Personalist-Military Hybrids* – A hybrid of military and personalist institutions in which a dictator holds most power and is relatively unchecked, yet the dictator's authority is rooted in military support. These systems often evolve from juntas when power is consolidated around a single individual. Examples include Chile under Pinochet, Pakistan under Zia and Musharraf, and Fiji under Bainimarama.
9. *Party-Military Hybrids* – Militarized single-party states in which most or all members of the ruling party are military elites.
10. *Party-Personalist-Military Hybrids* – A dictator rules with the support of a militarized single-party state but is relatively unchecked by these institutions. Examples include Egypt after 1952, Indonesia under Suharto, and Syria under the Assads.

iii. Interim Periods

1. *Warlordism* – occurs only in countries that are torn apart by conflict to the extent that they do not have a functional government.
2. *Foreign-Occupied* – occur where foreign politicians or militaries hold de facto power over a government.
3. *Civilian Provisional* – explicitly temporary arrangement that usually proceeds completed transitions to democracy or follow coups and constitutional crises.
4. *Military Provisional* – Interim regimes are only called "military provisional" if the military is holding power

until an election or some other formalized legitimizing event can occur.

4. Election Events
    a. **anticipation** – equal to 1 if a qualifying election is anticipated within the next six months.
    b. **ref_ant** – equal to 1 if a qualifying referendum is anticipated within the next six months.
    c. **leg_ant** – equal to 1 if a qualifying legislative election or general election is anticipated within the next six months.
    d. **exec_ant** – equal to 1 if a qualifying executive election is anticipated within the next six months.
    e. **irreg_lead_ant** – equal to 1 if a qualifying irregular leader election is scheduled within the next few months. Elections are irregular if they are not occurring as part of an established pattern or norm for executive selection. Elections are leader elections if they will directly choose the next leader via presidential or general elections.
    f. **election_now** – equal to 1 if a qualifying election occurs during the month observed.
    g. **election_recent** – equal to 1 in the six months following a qualifying election.
    h. **leg_recent** – equal to 1 in the six months following a qualifying legislative or general election.
    i. **exec_recent** – equal to 1 in the six months following a qualifying executive or presidential election.
    j. **lead_recent** – equal to 1 in the six months following a qualifying non-referendum election.
    k. **ref_recent** – equal to 1 in the six months following a qualifying referendum.
    l. **direct_recent** – equal to 1 in the six months following a qualifying direct (popular) election.
    m. **indirect_recent** – equal to 1 in the six months following a qualifying indirect (elite) election.
    n. **victory_recent** – equal to 1 in the six months following a victory for the incumbent leader or incumbent political party.

o. **defeat_recent** – equal to 1 in the six months following a victory for the incumbent leader or incumbent political party.
p. **change_recent** – equal to 1 in the six months following an election that forced a change in leadership (individual, not party).
q. **nochange_recent** – equal to 1 in the six months following an election that did not force a change in leader- ship (individual, not party).
r. **delayed** – equal to 1 in the six months following an election cancellation or delay of more than two weeks.
s. **lastelection** – the number of months since the last qualifying election (election for highest office or referendum that would expand executive power), or, in the absence of previous elections, the number of months since the political system last changed.
t. **loss** – the number of months since the incumbent or incumbent political party last lost a competitive election, or, in the absence of previous losses, the number of months since the political system last changed.

5. Irregular Events:
    a. **irregular** – unknown definition (will be removed)
    b. **prev_conflict** – equal to 1 if a qualifying violent civil conflict occurred somewhere in the country in the previous month.
    c. **pt_suc** – equal to 1 if a successful coup occurred against the leader that month.
    d. **pt_attempt** – equal to 1 if a coup attempt occurred against the leader that month.
    e. **precip** – precipitation (SPI) estimates using NOAA's current month Precipitation Reconstruction over Land (PREC/L) release.
    f. **couprisk** – unknown definition (will be removed).
    g. **pctile_risk** – unknown definition (will be removed).

Removing Undefined Features

There are three features in our dataset that were left undefined by the curator of the dataset. As such, we cannot include these features – **irregular**, **couprisk**, and **pctile_risk** – since we cannot accurately define them. As a result, we reduce the number of features from 38 to 35, with 132,866 total observations.

Converting ‘government’ to Dummy Variables

The next step in making the data usable in our model was to convert the **government** variable into a series of dummy variables using one-hot-encoding. By splitting this single variable into 16 new variables – one for each of the different government regime types – we are able to determine whether regime-type has a significant effect on the likelihood of coup attempts.

Thus, the single variable, **government**, was converted into **government_Dominant Party**, **government_Foreign/Occupied**, **government_Indirect Military**, and so on. As a result, the number of features in our model increased from 35 to 50, with a total of 132,866 observations.

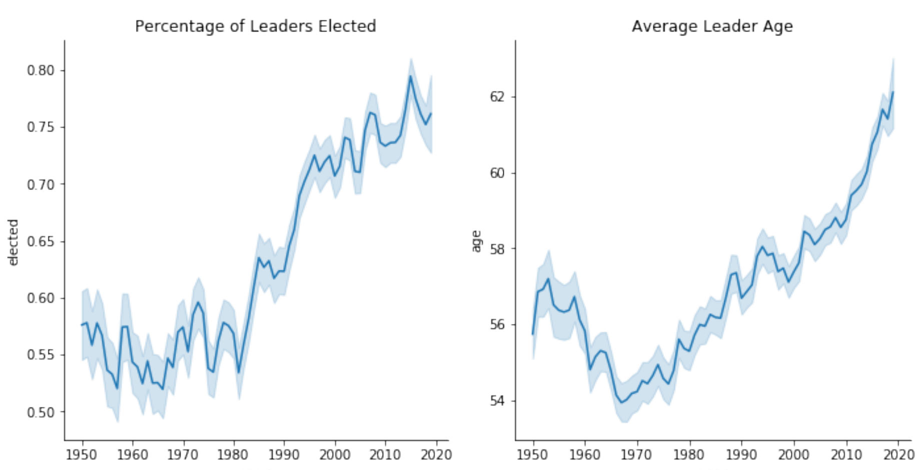
Removing Missing Values

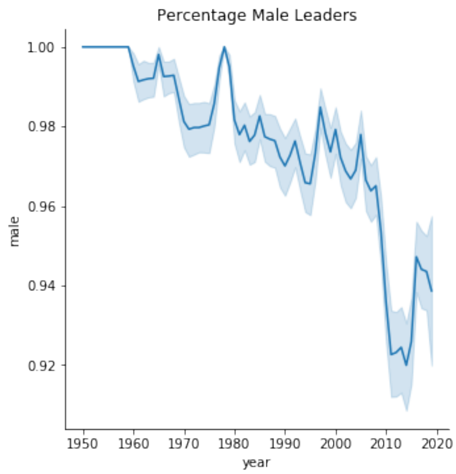
An ideal observation should contain values for *all* of the features in our dataset. As such, I aimed to remove any observations with missing values. The original dataset contained 132,866 observations, and after removing any observations with missing values the number of observations remained the same, indicating no missing values in the dataset.

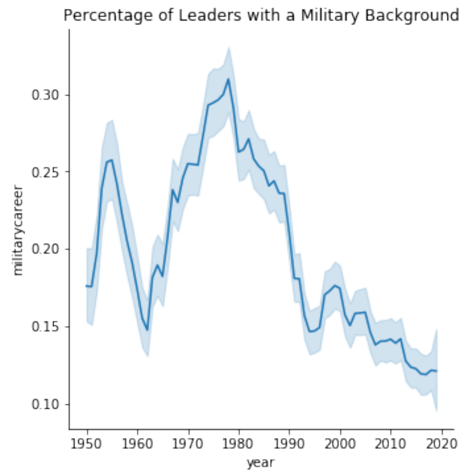
Part II – Exploratory Data Analysis

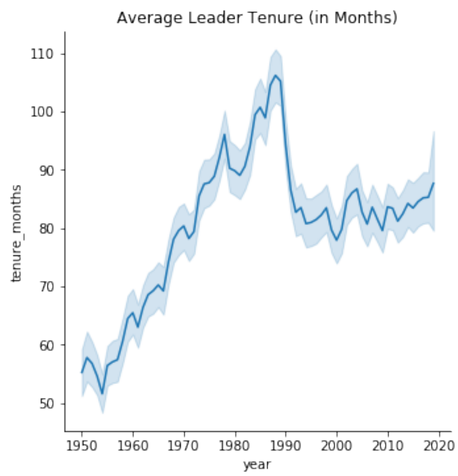
Graphical Analysis

I first graphically analyzed leader characteristics over time:









After graphically analyzing these leader characteristics we see some obvious trends. First, the percentage of elected leaders has steadily increased, from a little over half in the early 1970s to more than three-quarters today.

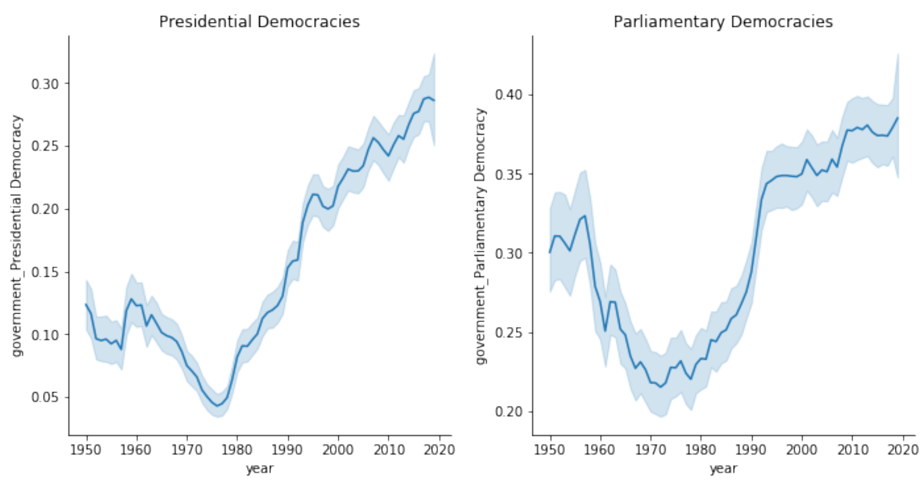
Second, there has been a persistent increase in the average leader age, from about 55 years in the 1950s to around 62 years today. This may either be an effect of greater democratization (for example, experience may be valued more in democracies than in autocracies) or a more general aging of the global population.

Third, though the male percentage of global leaders is still overwhelmingly lopsided, there has been some steady improvement in the percentage of leaders that are female, from 0% at its nadir to nearly 8% at its peak.

Fourth, the percentage of leaders with a military background has persistently declined, from more than 30% in the late 1970s to about 12% today. It is important to reiterate that previous service is not sufficient for a leader to be designated as having a “military background”. Rather, this designation is reserved for those whose *primary* affiliation prior to taking power can be described by the military, police, or defense ministry.

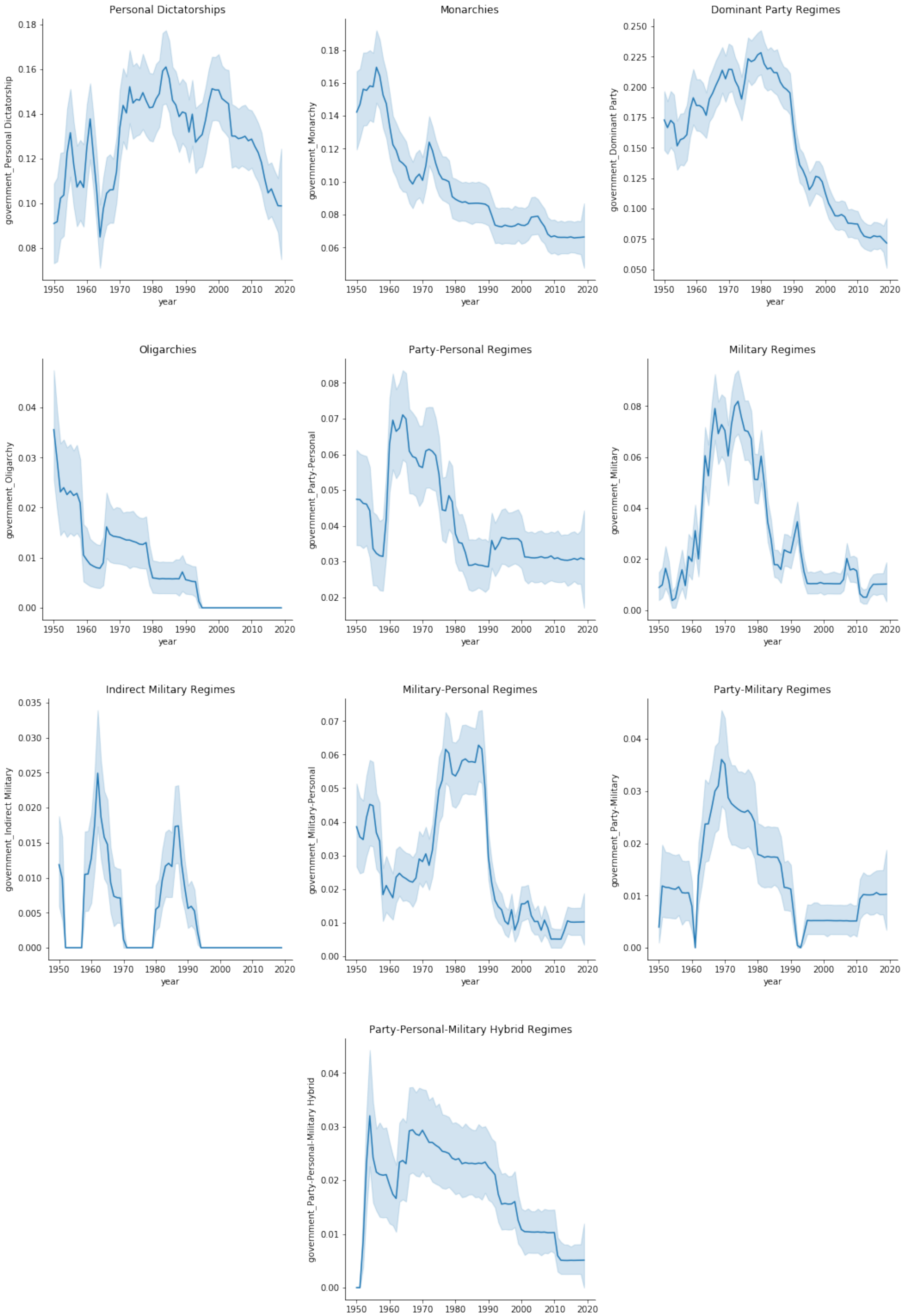
Fifth, and finally, it appears that average leader tenure initially rose from the early 1970s through the late 1980s – rising to a peak of more than 105 month (a little less than 9 years), but began to fall precipitously in the early 90s, where it has since hovered between 80 to 85 months (around 7 years). This is likely due to increased democratization as democracies are characterized by large coalitions and large selectorates. Thus, supporters of the leader have weak bonds of special privileges and so are more willing to defect, resulting in shorter average tenures.

I then graphically analyzed government regime trends. I first analyzed the democratic regimes:



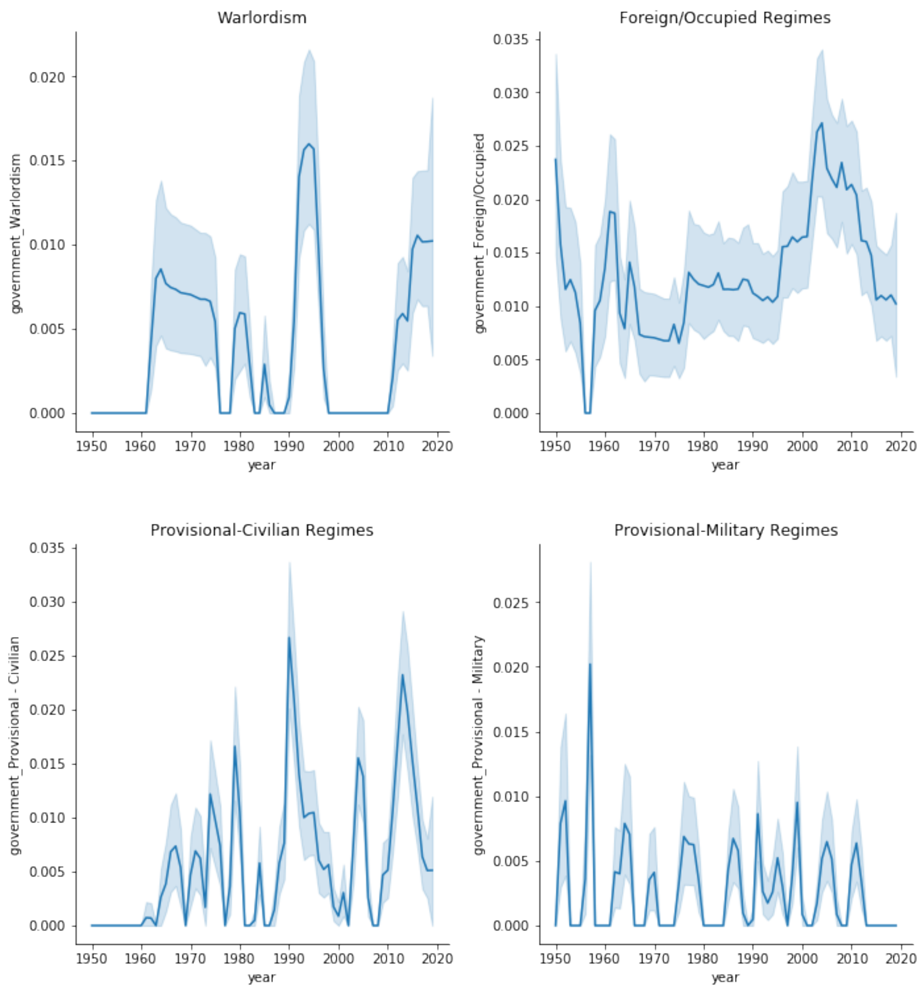
We do, indeed, see steady increases in democratization, with *Presidential Democracies* increasing from a low of ~5% of total regimes in the early 1970s to more than 25% today, and *Parliamentary Democracies* initially decreasing to a low of ~22% in the mid-1970s to nearly 40% today.

When we look at Non-Democratic regime types:


Personal Dictatorships
government_Personal Dictatorship
year
Monarchies
government_Monarchy
year
Dominant Party Regimes
government_Dominant Party
year
Oligarchies
government_Oligarchy
year
Party-Personal Regimes
government_Party-Personal
year
Military Regimes
government_Military
year
Indirect Military Regimes
government_Indirect Military
year
Military-Personal Regimes
government_Military-Personal
year
Party-Military Regimes
government_Party-Military
year
Party-Personal-Military Hybrid Regimes
government_Party-Personal-Military Hybrid
year

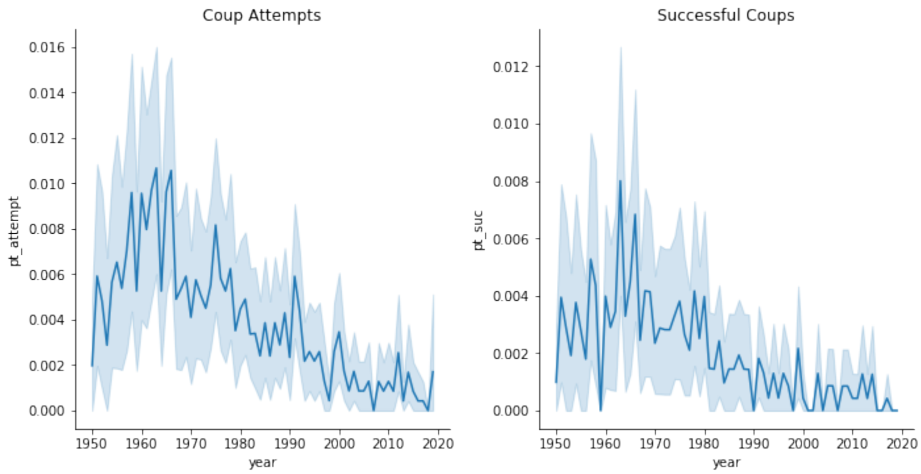
we see obvious declines across the board. Some regime types have arrested their declines (like *party-personal* regimes), but the overall trend is definitively downward.

The same is true when we analyze the interim regimes:

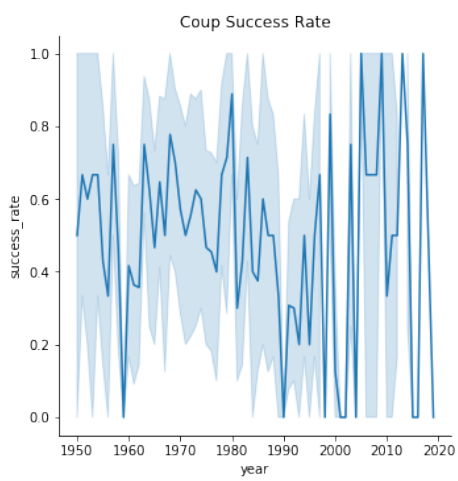


Either the interim regime is highly infrequent (*Warlordism*, *Provisional-Civilian*, and *Provisional-Military*) or it has experienced recent and pronounced declines (*Foreign/Occupied Regimes*).

In all, there is a clear trend towards greater democratization, with some persistent minority regime types – interspersed with infrequent provisional governments. Finally, when we look at coup attempts and coup successes:



we see, clearly, their relative infrequence and variability. However, there does seems to be a clear, asymptotic decline towards 0. This suggests that coup attempts are becoming ever more infrequent – perhaps a consequence of the greater stability of democratic regimes. The same holds true for coup successes: coup successes have declined alongside coup attempts, and even we look at coup success rates over time:



we cannot make out any reliable trends.

Part III – Running the Model

In order to create a model that accurately predicts the likelihood of a coup attempt, and which also provides us insight into which factors contribute the most to the likelihood of a coup attempt, I used a Random Forest model.

However, I did not use a Random Forest *Classifier* model due the nuances regarding how we measure model success. Perhaps the simplest method would have been to establish the X and y values, split the data into train and test sets, run the Random Forest Classifier model on training set, and check the accuracy of the model with regards to the test set.

However, the nuances of the underlying data preclude us from adopting this more straightforward procedure. The underlying dataset is extremely unbalanced; indeed, of the 132,866 complete observations in our data, only 464 (~0.35%) are of the positive class (i.e. coup attempts). Therefore, a naïve use of the Random Forest Classifier model using accuracy (i.e. # of correct predictions / total predictions) as our measure of success would have resulted in a model that would, understandably, seek to classify *all* observations as being part of the negative class. In other words, to maximize accuracy, the model would simply classify every observation as *not* being a coup attempt. This inherently undermines the utility of the model and so we must utilize a different methodology.

The benefit of Random Forest Regressor is that it provides *probabilities*, not predictions. When Random Forest Classifier arrives at a prediction, it does so by using a default threshold of 0.5. Observations with a probability of being part of the positive class (i.e. 1) greater than 0.5 are predicted to be part of that positive class, while those observation with a probability less than 0.5 are predicted to part of the negative class (i.e. 0). Therefore, by explicitly deriving the probabilities rather than the predictions, we can vary this threshold in order to better determine the utility of our model.

However, we must also determine our definition of model success. As previously stated, in cases such as this – where the underlying data is so dramatically unbalanced – accuracy is poor measure of success. When predicting rare events – for example, testing for the incidence of a rare disease – we want a model that captures as many of the *positive* cases as possible, a measure known as "recall": the fraction of relevant instances that have been retrieved over the total amount of relevant instances.

In other words, we want a test that is capable of determining every positive case in a given population. For example, if 100 people out of 10,000 have a rare disease, a test with perfect recall would determine all 100 positive cases.

Of course, such a test may also result in a large number of false positives: the test may perfectly determine the number of true positive cases, but it may do so at the expense of testing positive for 1,000 people. Thus, the test may be said to have poor "precision": the fraction of relevant instances among the retrieved instances.

In the case of our hypothetical rare disease test, the fraction of relevant instances (100) over the total number of relevant instances (100) would result in a recall score of 1.0. However, the fraction of relevant instances (100) among the retrieved instances (1000), would result in a precision score of only 0.1. The difficulty is in determining the appropriate tradeoff between recall and precision.

One attempt at balancing this tradeoff is known as the "F1" score: the harmonic average of the precision and recall, where an F1 score reaches its best value at 1 (perfect precision and recall) and worst at 0. However, the F1 score isn't perfect, as it gives equal importance to precision and recall. In practice, different types of mis-classifications incur different costs. In other words, the relative importance of precision and recall is an *aspect* of the problem at hand.

Therefore, my model determines *three* optimal thresholds: one for precision, one for recall, and one for the F1 score (as well as a default threshold of 0.5 for reference). Thus, the end-user of the model can ultimately determine which threshold is right given the circumstances.

The model proceeds first by establishing our X and y values and splits the data into training and test sets. I then established an instance of the Random Forest Regressor model using 10,000 estimators (i.e. 10,000 decision trees). This significantly increases the run-time of the model, but what we lose in speed we gain in model sensitivity.

I then fit the model to the training data, established the model's predictions on the test set, and created a new Data Frame to analyze the results. Normally, one would cross-validate the results, but since there are so many estimators being used in the model (10,000) relative to the total number of observations, the Out of Bag (OOB) estimate should asymptotically converge towards the best OOB estimate value, especially since the model's oob_score (i.e. whether to use out-of-bag samples to estimate the R^2 on unseen data) is set to "True".

In order determine the optimal thresholds for precision, recall, and the F1 score, I ran a for-loop that iterated over 1,000 possible thresholds from 0.0 to 1.0. I then separated out the false negatives, true positives, true negatives, and false positives. I then calculated the model's precision, recall, and F1 score.

Next, I determined the optimal thresholds to maximize the model's precision, recall, and F1 score (as well as including a default threshold of 0.5 as reference threshold).

```
### Optimal Precision Threshold ###
threshold    0.222222
precision    0.054054
recall       0.026490
f1_score     0.035556

### Optimal Recall Threshold ###
threshold    0.000000
precision    0.005366
recall       0.814570
f1_score     0.010662

### Optimal F1 Threshold ###
threshold    0.127127
precision    0.040000
recall       0.046358
f1_score     0.042945

### Default Threshold ###
threshold    0.500501
precision    0.000000
recall       0.000000
f1_score     0.010101
```

Finally, I printed a confusion matrix for each of the optimal thresholds to determine the number of false negatives, true positives, true negatives, and false positives:

```
### Optimal Precision Confusion Matrix ###
False Negatives:  147
True Positives:  4
True Negatives:  39637
False Positives:   70

### Optimal Recall Confusion Matrix ###
False Negatives:  28
True Positives:  123
True Negatives:  16909
False Positives:   22798

### Optimal F1 Confusion Matrix ###
False Negatives:  144
True Positives:  7
True Negatives:  39539
False Positives:   168
```

```
### Default Confusion Matrix ###
False Negatives:  151
True Positives:  0
True Negatives:  39699
False Positives:   8
```

Part IV – Analyzing the Results

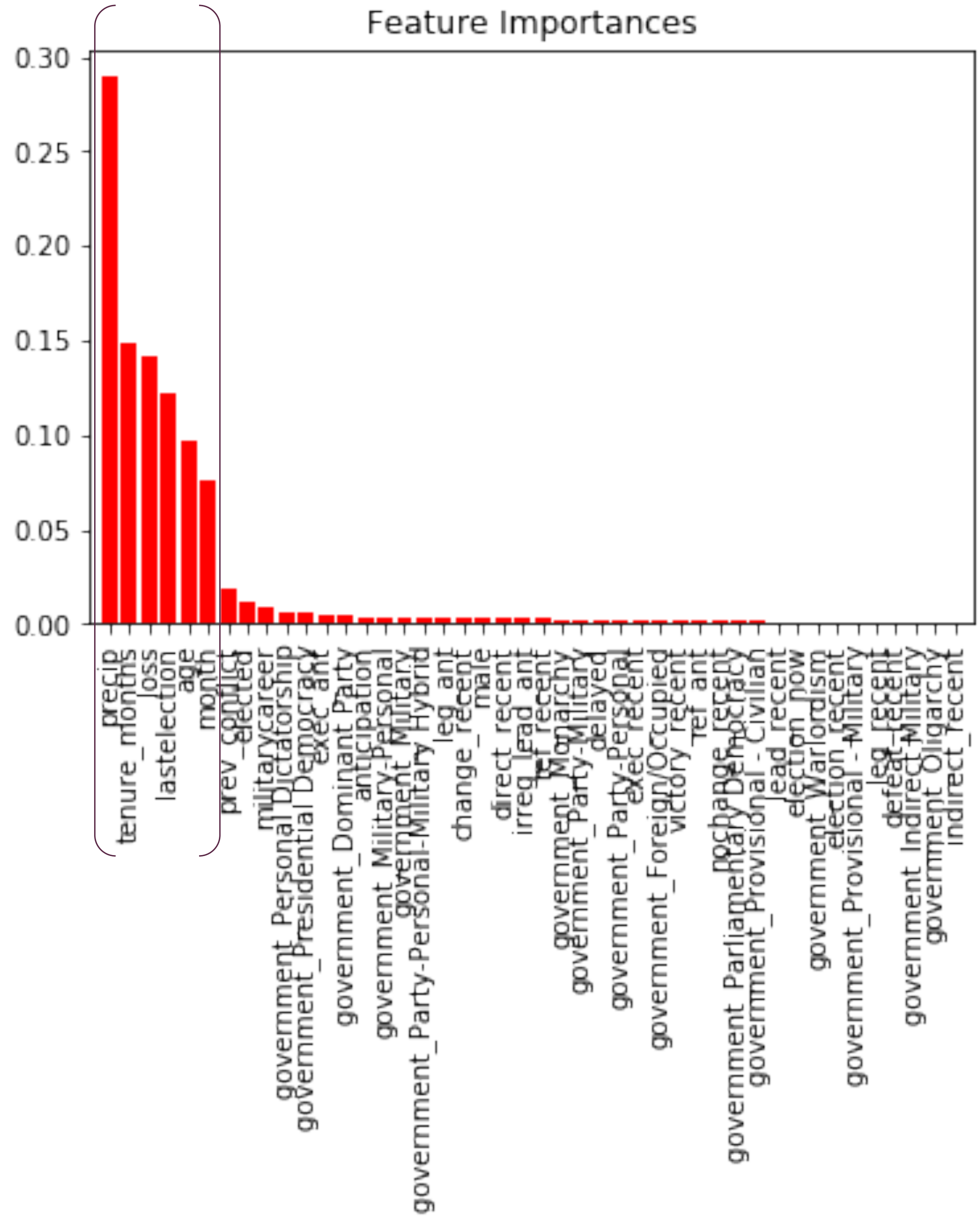
When we analyze the results of our model, we see that the optimal threshold for precision stands at ~0.22, recall at 0, and F1 at ~0.13.

An optimal recall threshold of 0 make sense given the severe imbalances in the underlying data. The relative incidence of true positives is extremely rare; therefore, the model must be extremely sensitive in order to capture those true positives. By setting the threshold to 0 we are able to capture as many true positives as possible given the number of estimators utilized in the model; and by increasing the number of estimators we can increase the sensitivity of our model to such an extent that it will capture *all* true positives.

However, there is an obvious cost to increasing our model's recall: as recall improves our false positive rate increases dramatically. In fact, if it weren't for the limitations of the model's run time, we could attain perfect recall, and as a result it is likely that the false positive rate would increase to the total number of remaining observations. As a result, we will have attained perfect recall at the expense of attaining the highest possible false positive rate.

As such, the optimal tradeoff between precision and recall – the F1 score – indicates that our model was able to correctly call 7 out of 151 (~5%) actual positives and 39,539 out of 39,707 (~99%) actual negatives.

When we analyze which features are the most important in explaining the variation in our dependent variable (coup attempts):



six features stand out from the rest (in brackets above):

1. **precip –** 0.289332
2. **tenure_months –** 0.149101
3. **loss –** 0.142232
4. **lastelection** – 0.122175
5. **age** – 0.097177
6. **month** – 0.076125

By the far the single most important feature is **precip** or precipitation. This makes some intuitive sense since too much precipitation (flooding) or too little precipitation (drought) can adversely affect crop outputs. As a result, food prices rise, and household incomes decline. The populace becomes ever more desperate and dissention increases. As a result, either the population or the selectorate rise up against the current regime. This may also explain why month of the year has such a large explanatory effect: rainy or dry seasons corresponds with certain months of the year, so **month** may be capturing this cyclical weather effect.

Both **tenure_months** (the number of months the leader has been in power in his or her present term) and **age** (an approximation of the leader's age) seem to be related as well, as long tenures and old age both expose a leader to increased likelihood of coupe attempts. Long tenures may stoke resentment amongst the selectorate, and old age may itself signal weakness due to the health effects of ageing.

Finally, **loss** (the number of months since the incumbent or incumbent political party last lost a competitive election) and **last_election** (the number of months since the last qualifying election) appear to be work together as well. When the incumbent or incumbent political party loses a competitive election, this is a direct signal from the population to the selectorate regarding the incumbent regime's weakness and/or illegitimacy. And as time passes since that election (as captured by **last_election**) the expectation for change intensifies.

If we subset the data to examine only the most recent year and month (March 2019) and look at the model's predicted probability of a coup attempt by country, we can create a league table of countries ordered by likelihood of coup attempt. For March 2019, the 5 countries with the highest likelihood of coup attempt are:

1. Gabon – 0.0427
2. Turkey – 0.0318
3. Belgium – 0.0251
4. Congo/Zaire – 0.0228
5. Yemen – 0.0192

If we refer back to our thresholds, we see that none of these top five countries meet our F1 score threshold of 0.127127, and so we would not predict a coup attempt in the current period.

If we again subset the data to analyze only those observations in which a coup attempt occurred, we can further analyze our thresholds to see if we can predict whether a given coup will be successful or not, based on the relative likelihood of the coup occurring at all.

When we look at our model's predicted probabilities for this subset and run a for-loop to determine the optimal thresholds for precision, recall, and F1 we find the following results:

```
### Optimal Precision Threshold ###
threshold    0.000000
precision    0.502294
recall       0.943966
f1_score     0.655689

### Optimal Recall Threshold ###
threshold    0.000000
precision    0.502294
recall       0.943966
f1_score     0.655689

### Optimal F1 Threshold ###
threshold    0.000000
precision    0.502294
recall       0.943966
f1_score     0.655689

### Default Threshold ###
threshold    0.500501
precision    0.498403
recall       0.672414
```

It appears that the optimal threshold – as determined by all three measures – is 0. In other words, if the predicted probability of a coup attempt meets our original F1 threshold of 0.127127, and if we also assume all such coup attempts will also be successful, we will maximize our precision, recall, and F1 scores.

We can analyze the resulting confusion matrices as well:

```
### Optimal Precision Confusion Matrix ###
False Negatives:  13
True Positives:  219
True Negatives:  15
False Positives:  217

### Optimal Recall Confusion Matrix ###
False Negatives:  13
True Positives:  219
True Negatives:  15
```

```
False Positives:  217

### Optimal F1 Confusion Matrix ###
False Negatives:  7
True Positives:  103
True Negatives:  7
False Positives:  123

### Default Confusion Matrix ###
False Negatives:  76
True Positives:  156
True Negatives:  75
False Positives:  157
```

We can see that if the main goal is to maximize precision, recall, or to find the optimal tradeoff between precision and recall our optimal threshold is 0 (i.e. we assume that all coup attempts will be successful).

However, this inelegant method of simply predicting that all coup attempts will be successful is hardly a prediction at all. Indeed, it makes much more intuitive sense to say that a coup attempt with a higher likelihood of occurring is also one that is more likely to succeed. However, given the paucity of positive cases we cannot say anything more definitive.

Part V – Conclusions

Through this analysis we have come to understand, a little bit better, the historical trends that have driven the decline in coup attempts over the past half-century, which factors contribute to civil unrest, and, thus, which factors help to overcome the coordination problems that the selectorate faces when contemplating a coup attempt.

The results seem to indicate that greater democratization of the globe over the past half century appears to, at least in part, drive the secular decline in coup attempts. How?

Mesquita and Smith note in their paper that leaders, all of whom face challengers who wish to depose them, maintain their coalition of supporters by taxing and spending in ways that allocate mixes of public and private goods. The nature of the mix depends on the size of the winning coalition.

More specifically, leaders who depend on only a select few to keep them in office, especially when they are drawn from a relatively large pool of potential supporters, engender loyalty among their backers by providing them with access to ample personal, private benefits they would not otherwise have if they were not in the coalition.

By contrast, with many supporters demanding rewards, the cost of personal benefits required to keep their loyalty becomes prohibitive. Instead, regardless of how altruistic or civic-minded the leader is, leaders who rely on a large coalition emphasize the production of goods that benefit everyone in society.

Democracies are characterized by these large coalitions. As such, Democracies spend a greater percentage of their tax revenues on public goods like dykes, dams, levees, aqueducts, and other forms of infrastructure that smooth out extreme precipitation events.

As the model shows, the single most important factor in determining the likelihood of a coup is precipitation. As more and more societies democratized over the past half-century, greater and greater percentages of global tax revenues were spent on public infrastructure – particularly on infrastructure that blunted the effects of extreme precipitation events.

It is my hypothesis that through this mechanism of increased democratization – leading to greater infrastructure spending, which in turn results in less detrimental effects of extreme precipitation events – that coup attempts have declined.

Indeed, one particularly interesting area of further study would be to see if the contribution of precipitation in explaining the likelihood of coup attempts has declined along with the increased democratization. If my hypothesis is correct, this should be the case.