5-2014

The Failed State: Can it be Predicted? An Application of Random Forest for Variable Model Selection

Beth Ann Stewart

Utah State University

THE FAILED STATE: CAN IT BE PREDICTED? AN APPLICATION OF RANDOM FOREST FOR VARIABLE MODEL SELECTION

By

Beth Ann Stewart

A thesis submitted in partial fulfillment of the requirements for the degree

Of

MASTER OF ARTS

In

Political Science

Table of Contents

Introduction: The Failed State: Can It Be Predicted?

Research Question

Is it possible to identify the predictive characteristics for failed states using statistics? This research is an application of R[1] and Random Forest[2] with previously mined data[3] as a means to approach this political science question.

This research and the resulting paper is intended to do three things. **First**, give a working definition of the failed state for this project. **Second**, illustrate the rationale for using statistics, R[4] as the platform and the randomForest[5] package to create a model to identify predictor variables for the failed state. **Third**, evaluate multiple predictor variables arrived at through the methodology for the statistical model being created, which will distinguish strong correlations between the variables and a states' current stability status. This is an appraisal evaluated through

---

[1] http://cran.r-project.org/. R is the computer language platform used in this model. It is an open source language, meaning, anyone can use it without license and can create compatible software packages for others. Those who choose to use R or any of the packages associated with are encouraged to allow access to their package on one of the networks where CRan is maintained. R programming code resources were used to configure this statistical model and are cited as they are used in subsequent chapters of this work.

[2] Wiener, Matthew and Andy Liaw. "Classification and Regression by randomForest" R.News, 2002, Vol2 no.3. pgs 18-22. RandomForest is a downloadable software package from the CRan network site. It has been used for statistical purposes and has been touted as a reliable computer learning regressionary tree program. The predictability and reliability as such of this software has been expanding over more than just the hard sciences. (list uses of randomForest here)

[3] The data for this statistical model was obtained from the following sources that utilize data mining and other data collection techniques that is openly available for research and other uses- The political Terror Scale www.poitical terrorscale.org, The World Bank www.worldbank.org/data, The CIA Factbook www.cia.gov/library/publications/the-world-factbook, (CIRI) The Cingranelli-Richards Human Rights Dataset www.humanrightsdata.org, Freedom House www.freedomhouse.org, and the United Nations http://data.un.org Data mining is defined as any mass list of data collected. It can be 5 years of collecting information on telephone calls, or all of the diagnoses billed from Medicaid, Blue Cross and Kaiser insurance companies. The massiveness of the data collected is traditionally produced from collective computer sources and stored and can be with no particular use in mind. Due to the nature of the massiveness of the data, data mining is a very expensive endeavor and is usually performed by large institutions or companies. Theinformation age has enabled sharing of such data if there is a use. Privacy problems arise when people feel that certain shared information may be harmful to them in some way. For example, a diagnosis may prevent someone from getting a job if the information was automatically accessible without personal authorization.

[4] Ibid 1

[5] Ibid 2

R[6] and randomForest[7] within it as identified through the methodology. **Finally**, we will look at several failed states to help illustrate the usefulness of the statistical model created.

Data is everywhere with today's information age. Where we work, how we vote, our medical ailments, and where we live right down to a street view of our home address[8] is being tracked, filtered, stored and used to identify various things about people and their habits.[9] Information covering economics, government authority, civil societies, access to health care, authoritarian regimes, international disputes, human rights violations, economic solvency, unemployment rates and now even cellular phone records are being collected by individual databases through international institutions, governments, scholars, and NGO's.[10] Thousands of variables - bits of information, are collected into these databases, but almost no one is using them because of an inability to process it, resulting in a true "information overload"[11]. The challenge presented is how to utilize this mass of international data to address a real question thereby resulting in something concrete and constructive.

Data sets have become so large and unwieldy that even current technology cannot handle all the data in today's world of increasing speed and optimizing software programs.[12] The data is

[6] Ibid 1
[7] Ibid 2
[8] 'Google Earth' http://www.google.com/earth/index.html.
[9] "Big Data: The Next Frontier for Innovation, Competition, and Productivity", McKinsey Global Institute, Technology & Innovation, McKinsey & Company. http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation. and Brown, Brad, Michael Chui, and James Manyika. "Are you ready for the era of 'big data'." *McKinsey Quarterly* 4 (2011): 24-35.
[10] Ibid 8
[11] King, Gary, "Ensuring the Data-Rich Future of the Social Sciences", *Science*, 331 (2011), 719–721.
[12] Ibid 9

being collected faster and at greater proportions far surpassing processing capabilities.[13] In the words of John Naisbitt, “We are drowning in information, and starved for knowledge”[14].

So how do we filter through the cacophony of information in this data din? Most researchers who look at data indices are suffocated by their enormity. Rarely are these data monoliths put into practice because a majority of people really don’t know how to use the information to solve a problem or answer a question, rendering this data essentially useless. Conversely, however, this mass of information being stored presents endless possibilities for those willing to filter through it and harness it into a workable medium.[15] The desire to do just that has led me to attempt to use a quantitative approach in evaluating the failed state and to identify predictive variables contained therein. It is my hope that a quantitative approach to the failed state question will lend credibility to the field of political science, which has traditionally been considered a soft science, and create a platform and new direction for future research through the computer language of R[16] and the software package of randomForest[17] within in it.

We might as well start with one more question: Why address the question of the failed state? Since the 9/11 attacks in the United States against the World Trade Center and the Pentagon, the U.S. has considered failing states to be a growing concern.[18] Unstable regimes and porous borders create havens for human and drug trafficking, while weapons of mass destruction

---

[13] Ibid 9

[14] Naisbitt, John. *Megatrends 2000* (Avon, 1991).

[15] Jakulin, Aleks. "A Rant on the Virtues of Data Mining: Statistical Modeling, Causal Inference, and Social Science", *Statistical Modeling, causal Inference, and social Science*, 2012 http://andrewgelman.com/2007/08/a_rant_on_the_v/ [accessed 11 July 2012]. Aleks is not the first to discuss the problems of data mining. It has become a concern with our current administration regarding the collection phone records from Verizon users and whether the massive collection of private information is truly helpful in fighting domestic terrorism. For other source see ibid 3.

[16] Ibid 1

[17] Ibid 2

[18] Mallaby, Sebastian. *The World’s Banker: A Story of Failed States, Financial Crises, and the Wealth and Poverty of Nations* (Penguin Books, 2006).

undermine the stability of economic development and challenge international security.[19] Al Qaida terrorist groups are growing in numbers in Iraq, Syria, and Somalia.[20] The United States is interested in protecting their economic interest globally-a view that is supported by the National Security Strategy published by the Clinton Administration stipulating "selective U.S. engagement around the world".[21] There is a dichotomy and debate between intervention of a failed state and self-determination underpinned by conflicting economic and strategic interests. Each year the United States submits millions of dollars from its already tapped coffers to address these international concerns. Egypt is one such example.[22] Propping up countries, intervening with troops, funding favored rebel groups against authoritarian regimes from the objective of protecting the United States' interests have left the U.S., a country already in deficit, in an even greater financially challenged position.[23] It is agreed by pundits that it is far better to prop up a regime than to fix a failed/collapsed state thus making this research question so important.[24]

Although a majority of work in the political science field and the failed state has been qualitative, I believe it is possible to identify characteristics of failed states using statistical methods from information provided through data that has been mined.[25] This mined data coupled with the statistical model detailed in the methodology may make it more likely to predict failed

[19] Rotberg, Robert I. , "Failed States in a World of Terror", *Foreign Affairs*, 81 (2002), 127. And (Wyler, 2008)
[20] Williamson, Richard S. "Nation-Building: The dangers of Weak, failing and Failed States", The Whitehead Journal of Diplomacy and International Relations. Winter/Spring (2007). 14. www.journalofdiplomacy.org.
[21] Dorff, Robert H., "Democratization and Failed States: The Challenge of Ungovernability", *Parameters, US Army War College Quarterly - Summer 1996*, 1996, 17–31. The Clinton administration dealt with not only the tragedy of the Balkans but also the genocide in Rwanda in 1994. Dealing with the aftermath of problems created from failed states justified concluding that avoiding the failed state is a far more amenable prescription than recreation of state.
[22] There has been a developing foreign policy since the Truman administration giving aid to countries as part of the United States foreign policy goals. Since 9-11, the focus has altered somewhat in its emphasis to prop up states economically and gain a political chess piece in the future global negotiations. Radelet, Steven. "Bush and Foreign Aid", Foregin Affairs. Vol.82, No.5.(Sept-Oct, 2003) pp. 104-117. www.jstor.org/stable/20033686.
[23] Charles T. Call, "The Fallacy of the Failed State". Third World Quarterly, 29 (2008), 1491-1507.
[24] Ibid 20, Ibid 21.
[25] Ibid 3

or faltering states once these characteristics are determined. There has already been a great deal of work done on identifying characteristics of the failed state and its precursors, but again, the bulk of information on this topic is rendered in qualitative form.

Literature Review

The failed state dilemma is poised at the top of policy makers globally. This being said, it comes as no surprise the amount of research done to elucidate this topic. Exhaustive endeavors by preceding researchers have yet to come up with even a concise definition of the "failed state"[26]. Unfortunately most of these scholars hold in reserve a definitive statement for the failed state. Although a good number have suggestions for remediation, as well as descriptions for what can usually be found in the failed state, the vagueness or lack of definition leads the reader to assume that a country can be deemed a failed state if all of the identifiers illustrated by these respective scholars are present.

Rosa Brooks deems a failed state the opposite of what a state is. In short, basically the demise of the state constitutes a failed state. In her article *Failed States or the State as Failure*[27] she portrays the following to be present in a failed state: violence ensues and government loses control of its territories.[28] This oversimplifies things by implying that all that needs to be present for the state to be considered failed are these events. Brooks also mentions a subset of a definition for what constitutes weak or failing states and further suggests remediation for these states being best applied in the form of some international trusteeship such as the United

[26] Ibid 23.
[27] Brooks, Rosa, "Failed States or the State as Failure?" The University of Chicago Law Review. 72 (2005), pp 1159-1196.
[28] Ibid24

Nations.[29] This particular view of a trusteeship to oversee states when they have failed is shared with Helmand and Ratner. In their article, *Saving Failed States*, they view three possible faltering state scenarios: the failed, the failing, and the states too new to decide on, suggesting that the United Nations charters' primary focus is for the right to self-determination rather than a concern for a states potential longevity.[30] Helman and Ratner are of the few scholars who define the failed state, although ambiguous, as a failed nation which occurs when it is "utterly incapable of sustaining itself as a member of the international community"[31]. This particular definition falls short given that the researcher is suggesting a United Nations protectorate as a government stand-in for intervention in the event of a failed or failing state.[32]

Dorff approaches the definition of the failed state from the old adage "democracies do not fight other democracies"[33]. He also pinpoints large financial markets and global economies as a problem for the smaller failed state because the smaller state will have a harder time rebuilding itself in the event of collapse. Egypt's overthrow of their presidential elect Mohammed Morsi, Libya's lack of an official government after the ousting of Gaddafi are two counter examples to his argument. A final point is his mention that new democracies go to war and occasionally transition to authoritarian regimes.[34] None of these opinions approach a definition for a failed state, but instead suggest what he believes should constitute a democracy.

Probably the largest and most well written contributor to the qualitative method of evaluating the failed state is Robert Rotberg. Numerous articles share a constant thread

---

[29] Ibid 24
[30] Helman, Gerald B and Steven R. Ratner, "Saving failed States". *Foreign Policy*. No.89 (Winter, 1992- 1993). Pp. 3-20.
[31] Ibid 27
[32] Ibid 27
[33] Dorff, Robert H., "Democratization and Failed States: The Challenge of Ungovernability", *Parameters, US Army War College Quarterly - Summer 1996*, (1996): 17–31.
[34] Ibid 30

describing several factors present in the failed state, but, yet again, no real definition is offered for the failed state. He mentions several levels of the failed state. These are failed states, collapsed states, and weak states, with collapsed states equal to utter and complete disintegration of government using Somalia as an illustration.[35] Like others scholars, he mentions characteristics of the failed state which include but are not limited to: a lack of food for the populace, a decrease in education, a decrease in medical infrastructure, an increase in government and elitist corruption, and a block of civil societies.[36] It is unfortunate, but all of these pseudo-definitions of a failed state are too numerous to list.

The Fund for Peace describes the a failing state as "One that is losing legitimacy, maintains few or no functioning state institutions, offers few or no public services and is unable to contain or deliberately inspires social fragmentation"[37]. Furthermore they go on to say that a failed state is one in which it "forfeits the authority to make collective decision for the national population"[38]. They continue by adding that a failing state "fails to interact in formal relations with other states as a fully functioning member of the international community"[39]. I would argue that these two points should not define a failed state. They are better suited as identifiers of, but should not be included in the definition. Many states, regardless if it is positioned towards a division or collapse, still negotiate with other countries and trade goods. The Assad regime in Syria is one such example of this. The United Nations is in negotiations to decrease chemical weapon stores in that state.

---

[35] Rotberg, Robert I., "Failed States in a World of Terror", *Foreign Affairs*, 81 (2002), 127 and Rotberg, Robert I. "Failed states, collapsed states, weak states: Causes and indicators." *State failure and state weakness in a time of terror* (2003): 1-25. PDF from the Brookings institute http://www.brookings.edu

[36] Rotberg, Robert I. *When States Fail: Causes and Consequences*. Princeton University Press. New Jersey. 2004 pp.7.

[37] Fund For Peace, Failed State Index. http://library.fundforpeace.org/fsi

[38] Ibid 34

[39] Ibid 34

Although prior qualitative review of the failed state has given us a better knowledge base of these characteristics associated with the failed state such as lack of a government head, refugee influx, and a decrease in literacy as well as a decrease in life expectancy, the fact remains that throughout all of this previous research there is still an obvious lack of a concrete, working definition of the failed state by which to analyze this respective research against. It is disappointing that a majority of these pundits are giving advice on how to remedy a problem they have not really defined. Charles Call at least recognizes this issue when he says in his article, *The Fallacy of the Failed State*, that because of a lack of consensus for the definition, the term should just be thrown out.[40]

Failed State Definition

Taking all of the previous research into account and for the purposes of this paper, we define the failed state herein as failed when the following occurs-a total collapse of a government body and its respective institutions within defined and recognized geographic boundaries and/or a cessation of legitimacy in the government by its respective polity. A failed state can happen either by overthrow of a current regime either internally or externally. This definition allows civil war to be constituted as a failed state, as well as complete collapse of all infrastructures. Additionally, a failed state could be demonstrated by a secession of state, meaning state division of some sort as a result of civil war as was the case in Sudan and South Sudan[41] recently, or

[40] Ibid 23.
[41] "Republic of South Sudan Declares Independence". http://www.betterworldcampaign.org/un-peacekeeping/web-features/south-sudan-declares-independence.html [accessed 12 December 2012].

ethnic cleavages as was present in the case in the former Yugoslavia[42] as well as the former Soviet Union (U.S.S.R.).[43]

Previous work in the form of indices and statistics offers a starting point for this and future such research in the field of political science. This includes but is not limited to the index of State Weakness in the Developing World created by Susan Rice and the Brookings Institute[44], The Mo Ibrahim index[45] focusing on Africa which gives a complete score definition of the failed state as well as a methodology and sources for raw data, the CIFP or the Country indicators for Foreign Policy[46] which lists data about foreign aid, and finally the Fund For Peaces' Failed State Ranking that is released annually by Foreign Policy. All of these entities have done a great deal of work on the fragility of states by using quantitative methods.

The Brookings Institute has a rather concise index titled "Index of State Weakness in the Developing World"[47] in which Susan Rice and Patrick Stewart detail and categorize several lists of characteristics in which to gauge states fragility. These include economic, political, security and social welfare areas or 'baskets'[48]. Depending on the subcategories, all of these are weighed carefully, given a score and then the aggregate of that score goes into the basket and is tallied.[49]

---

[42] Laitin, David and Daniel Posner, "The Implications of Contructivism for Contructing Ethnic Fractionalization Indices", *The Comparative Politics Newsletter*, 12 http://todayinsci.com/QuotationsCategories/I_Cat/Information-Quotations.htm.
[43] Ibid 40.
[44]Rice, Susan and Stewart Patrick. "Index of State Weakness in the Developing World", *The Brookings Institution* http://www.brookings.edu/research/reports/2008/02/weak-states-index.
[45] Mo Ibrahim Foundation | Methodology, *Mo Ihabrim Foundation Index* http://www.moibrahimfoundation.org/IIAG-methodology [accessed 12 December 2012].
[46]CIFP - Country Indicators for Foreign Policy. http://www4.carleton.ca/cifp/ [accessed 12 December 2012].
[47] Ibid 42.
[48] Ibid 44.
[49] Ibid 42

A very detailed effort to the indices for failed state evaluation includes the Mo Ibrahim Foundations Index of African Governance.[50] They focus on Africa in specific, and as such, are able to be more precise in their findings, pin pointing areas of weakness and/or challenge, that larger western world indicees are not able to. Consistent with most indexes, there are the traditional categories that the school of political science has deemed relevant to failed states: rule of law, economic areas, and human development to name a few.

Country Indications for Foreign Policy or CIFP[51] is an online accessible publication for failed state ranking and index. Covering again the same sorts of indicators as others, this index also includes in its focus a human rights category and corruption figure. These details for human rights commonly have been lumped into rule of law for other indexes has previously not given full merit in its own category as it is here. Another different element this index has to offer is a category for corruption as well as political stability and violence. Produced by Carlton University in Canada, the raw data sources are listed so that other scholars may use it to further research.[52]

Monty Marshall and Benjamin Cole also have produced a “State Fragility Index and Matrix”[53]. Their index uses 0-5 as the basis for compiling scores. Political, security, economic and social indicators are scored and compiled into the ‘matrix’ which, again, basically is a table with the scores listed.[54]

Probably the most notable contributor toward failed state research is from the Fund for Peace. Its Failed State Index comes out each year in the Foreign Policy publication advertising

---

[50] Ibid 43
[51] Ibid 45
[52] Ibid 45
[53] Marshall, Monty and Benjamin Cole., "Global Report on Conflict, Governance and State Fragility 2008.", *Foreign Policy Bulletin*, 2008, 1–21.
[54] Ibid 45

the worst of the worst countries and failed states.[55] The Fund for Peace typically ranks 177 countries with a total ranking score of 120 where 1 is the lowest score to give a category with120 points possible for complete failure. Each sub category within the framework is given a 1-10 score with one being the best and 10 the worst possible score in that particular group.[56] Their method is detailed through the Conflict Assessment System Tool (CAST) which was created in 2001[57]. They describe this tool as an objective statistical assessment, implying that a statistical program was used to determine the final failed state ranking, or at least one was used to identify key areas projected as having relevancy for scoring the failed state. However, there is no mention in the methodology provided of a statistical program in CAST and/or how it works other than it separates the "relevant data from the irrelevant data"[58], and that they use "human analysis to ensure that the software has not misinterpreted the raw data"[59]. When reviewing the CAST manual, it appears to be a very detailed rubric by which their data figures are derived. They use three main categories called "clusters of leading societal indicators of state decay and internal collapse".[60] These are social, economic and political/military identifiers.[61] For each of the subsets, they suggest the researcher to begin with a baseline number of 1. Depending on the researchers evaluation of the identifiers given, and using the prescribed definitions as a guide, then they add a point if the characteristic is present. These numbers are tallied. With the 12 individual characteristics, the total points possible are 120 for a failed state.[62]

[55] Foreign Policy 2012
[56] Ibid 34
[57] Conflict Assessment System Tool (Cast) system, 2001. The manual can be located here. http://library.fundforpeace.org/fsi
[58] Ibid 34
[59] Ibid 34
[60] Ibid 54
[61] Ibid 54
[62] Ibid 54

Unfortunately, the Fund for Peace does not offer their raw data so that others may use it to further this research. You can copy the findings each year and convert the numbers to a workable excel table or a comma separated value (csv)[63] file. However they suggest in CAST to use their rubric for your own subjective analysis.[64]

Sadly, with the exception of the Mo Ibrahim Index[65], most of the producers of these failed state indices do lock away the raw data to protect their work. There is a plethora of raw data available electronically through NGO's and other institutions, but not all of it is in an accessible or workable standard. Some of it is offered in csv format, others provide only a chart with descriptions and numbers requiring it to be adapted and streamlined and into a functioning medium.

After review of all of the literature available and reviewing the mass of raw data, it became evident that out of all of this research, few if any had used statistics with their research and none of them were the product of a computer statistical model. The mention of statistical programs in past work is limited and altered depending on the interpretation, qualified, and altered if an analyzer (human form) deemed the statistical findings to be irrelevant. Wouldn't the question of the failed state would be more interesting if we could use the data and identify predictor variables for the failed state building a statistical model. How would we go about doing this? One of the considerations was to make the model in a format that anyone could use, not just someone with access to expensive supercomputer equipment, and it needed to be in a language that was probably not going to be considered extinct in the next few years like Fortran is now.

[63] CSV=comma separated file
[64] Ibid 54
[65] Ibid 43

Since there is a data surfeit, we have the ability to make more precise determinations about the failing state through more robust statistical methods. Although there are several different programs that are classification regression learning predictive tree models, we are using R[66] as the language platform with randomForest[67] as a software package in which to create the model. It was decided to use Random Forest and the predictor capabilities of the program for variable selection to create the model. Other options were SAS, SPSS and CART (Classification and Regression Trees), or IBM's InfoSphere Warehouse, however, both were unavailable due to cost as well as computer storage and ample processor speeds for which to run these complex programs. SAS and SPSS were omitted as statistical program possibilities due to a lack of availability. Both required either an expensive financial commitment that some universities are unwilling to carry, and/or a super processer that this size of dataset would require.

R[68] and randomForest[69] seemed to offer the potential to work through our question and circumvent the financial issues listed above. R[70] is a free language platform with open source downloadable from the CRan network. It has the ability to run smaller data sets or larger ones within certain parameters while offering the flexibility using a PC. All packages associated with R[71] are offered as free packages from the CRan network that gives global downloadable access.[72] These access locations are called mirrors. R is receiving increased recognition as a language

[66] Cran network. Ibid 1
[67] randomForest Ibid 2
[68] Ibid 62
[69] Ibid 63
[70] Ibid 62
[71] Ibid 62
[72] The Cran network uses several mirrors. These mirrors reflect a geographical point on the globe. The users of this network choose a mirror site in which to access software packages. The closer the proximity, the shorter download time is the thought.

which allows the user great adaptability for multiple purposes as noted from the variety of current software program packages available.[73]

Again, this paper and the method used to evaluate failed state predictors is meant to bring credit to a traditionally soft science field and provide a platform for further research. The best way to open a door is to create a model within a software program that anyone can re-create, through easy, accessible means. In this case the means are the accessible platform and the streamlined, downloadable, csv file raw data, further illustrating the rationale for choosing R.

Although 'R' itself has its challenges, namely that almost all its computer code has to be input manually — there are only a few graphical user interfaces for it. Many scholars are gravitating to it due to its exceptionally promising possibilities within the computer language of R and the benefits that its open source provides.[74]

Recursive partitioning statistics, tree modeling and Random Forest therein, have been consistently showing reliable predictive results.[75] Random Forest differs from other regression tree machine learning programs by the following: in addition to the regression bootstrap method to determine node split, it also takes a random sample[76] and runs this random sample against the root. What does this do exactly? Most statistical programs have some degree of error, due to biases of data supplied and the nature of the program itself looking for homogeneity in variables against the root. We build a forest of these like kind trees. With a random sample taken in the

[73] Ibid 1
[74] Open source is a term which means that it is not privately owned by a particular entity. There are no user rights given by anyone, and anyone can use the software. It is openly shared. However, proper citation for algorithms and software programs within the language are still required.
[75] Random Forest was used for gene selection applications. A paper on this describes how random forest does not over fit, can be used with large data sets and can still deal with noisy predictive variables. Diaz-Uriarte, Ramon and Sara Alvarex de Andres. "Gene Selection and Calssification of Microarray data using random forest". *BMC Bioinformatics*. 6 January 2006. http://www.biomedcentral.com/1471-2105/7/3. web access.
[76] Random sampling is also known as bagging or abbreviated as OOB meaning out of bag.

program, it allows the algorithm to be more accurate- fewer false positives, with a higher rate of node purity.[77] Large data sets can be very noisy and random forest has a lower error rate in the sample due to the random sampling. This is one of the few programs that use this factor programmed into the algorithm.[78]

Due to its predictability, flexibility and discerning capabilities, it should come as no surprise that the medical field has been using statistical tree programs to elucidate possible treatment scenarios for patients with acute illness. One such study was done by three researchers looking at patients suffering from renal cancer. In an effort to use the least invasive measures possible for the highest recovery rate, they evaluated cases based on applicable variables. It is because of the predictability of the recursive tree partitioning statistical model, that they have been able to make sufficient progress to apply the model's results in the exam room for cancer treatment.[79] Additionally, Furlanello has completed research using R and Random Forest evaluating the potential prevalence for tick-borne disease within the geographical boundaries of Trento, Italy where tick presence is common.[80]At our own institution, Dr. Richard Culter et al produced the following research paper and work "Random Forest for Classification in

[77] A node in tree modeling is defined as each place that a new tree is created. This node is a split from the parent tree. Node purity measures the accuracy of the split choice. For more information on nodes and tree modeling see explanation from Breiman and Cutler in randomForest software description. See ibid 2. For a graphed example of node purity in our research see table 3 and table 4.
[78] Ibid 2
[79] Three articles can be referred to for further information regarding the use of random forest algorithms, their predictability and use in cancer research that is now being applied to the exam room. Carter, Hannah, et al. "Cancer-specific high-throughput annotation of somatic mutations: computational prediction of driver missense mutations." *Cancer research* 69.16 (2009): 6660-6667. Ma, Yan, et al. "Predicting cancer drug response by proteomic profiling."*Clinical cancer research* 12.15 (2006): 4583-4589. https://clincancerres.aacrjournals.org/content/12/15/4583.full, and Kim, Hyung L., et al. "Using protein expressions to predict survival in clear cell renal carcinoma." *Clinical cancer research* 10.16 (2004): 5464-5471. http://clincancerres.aacrjournals.org/content/10/16/5464.web.
[80] Furlanello et. All. 2003.

Ecology"[81]. In it, he describes the accuracy of the predictability for the program even with some missing data.

Aside from the social sciences, machine learning is being developed to mimic the brain for adaptation in cell phones and Google search algorithms.[82] If your personal wireless device and search programs can mark and learn your behavior from previous inputs and inquiries, it can take less time to provide you with results that the computer program predicts you are looking for, thereby making the device more useful. In addition, Machine learning may even be coming to a TSA near you. A recent article released suggests the future use of self serve security kiosks that predict behavior based on your belongings and where you live. The device gets more accurate as more people go through the scanner based on false negatives (commits to memory), such as mistaking an electric razor for a bomb. [83]

Methodology

The methodology used for this research includes a platform language called R into which a software package application called randomForest is employed. Random Forest is so named because it computes (for a sample illustration) like an upside-down forest of ancestral/family (genealogical) trees. The last trees "grown" in this forest equal the most frequently manifest predictor variables (characteristics) from the computational set, in which 'set' comes from the

---

[81] Cutler, Richard, Thomas EdwardsJr., Karen Beard et all. "Random Forests for Classification in Ecology." *Ecology*, 88(11) 2007. 2783-2792.

[82]Some information on machine learning and smart phones can be found in the following article. Makeig, S.; Kothe, C., Mullen, T.,Bigdely-Shamlo et all. "Evolving Signal Processing for Brain–Computer Interfaces," *Proceedings of the IEEE* , vol.100, no.Special Centennial Issue, pp.1567,1584, May 13 2012 http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6169943&isnumber=6259910

For detailed information regarding search engine responses to individual inquiry and changes to those algorithms a good article can be referred to here. Thelwall, Mike. "The responsiveness of search engine indexes." *Cybermetrics*5.1 (2001): 8. http://cybermetrics.cindoc.csic.es/articles/v5i1p1.pdf

[83] http://travel.yahoo.com/blogs/compass/security-machine-might-replace-local-tsa-agent-212342236.html

raw data initially input. Along with random forest as the predictive software, Hmisc and Boruta are used to assist in selecting final variables for the model.

RandomForest[84] works by 'growing' trees from the raw data input in which the data can be likened to nodes and branches or seeds to saplings. Some data 'seeds/seedlings' turn into more robust saplings than others based on the homogeneity of the variables determined by the program. It is considered a kind of bootstrap program[85] that also uses out of bag (OOB) sampling to stabilize the data results. For this research, a forest was 'grown' in Random Forest by using the failed state ranking subjectively predetermined as the dependent variable in the index for the source set, or seed, to continue with the tree growth analogy, which then splits into subsets using the Gini impurity index[86] that determined the node or branch of the split. The Random Forest programs job is to build or grow homogenous trees in which the objective is to split off the trees that begin to look unlike the rest of the forest. This process is repeated on each derived subset and is considered complete when the splitting or "growth" that the program generates no longer adds value (per the law of diminishing returns)[87] to the predictive resultant variables. Other forms of statistical methods like this are called regressive partitioning,[88] Classification and

[84] For a detailed video explanation on randomForest you can look here for a five video series. https://www.youtube.com/watch?v=cQrvTYVN0ko

[85] The basic idea with a bootstrapping technique is to resample to data over and over. It is a way to determine the margin of error within a data set. This is employed in many R programs and in particular random Forest. For more information regarding the bootstrap method refer to Efron, Bradley, and Robert J. Tibshirani. *An introduction to the bootstrap*. Vol. 57. CRC press, 1994.pp.45-57.

[86] The Gini impurity is employed in random forest to determine when to split from the parent to grow another tree in the forest. Ibid 2.

[87] We use the term the law of diminishing returns to imply that adding another sampling does not yield any decreasing results. It is not meant to imply that adding further sampling will alter the results negatively after a certain number of samples taken.

[88] Recursive partitioning in regression models refers to the questions posed in the algorithm. It usually is a statement that determines the split in tree models for variable selection. See Torsten, Hothorn, Kurt Hornik and Achim Zeileis; "Unbiased Recursive Partitioning: A Conditional Interence Franework" *Journal of Computational and Graphical Statistics*. Vol. 15, Iss.3. 2006. http://www.tandfonline.com/doi/abs/10.1198/106186006X133933

regression trees (CART),[89] machine learning,[90] and employing bootstrapping and random sampling techniques.[91]

To explain regression tree programs in another way, the data set is asked a series of if then questions. Each question leads to the next question. In other words, it is a set of binary decisions. In theory, for illustration only, y equals 1 in this example or the failed state. We then ask a series of yes and no questions regarding the other variables. Labeling three variables as A, B, C, the question posed might be is A>.50. If the answer is yes, then the node split could be to the left, if the answer is no, then the node split is to the right and the next question is dependent on how the last one was answered. The next question might be on the left node split of the tree if A>.50 then is B >.75 and so on. From the answers of these yes and no questions, the probability the A=Y is a percentage of the probability from the question. In our case A, B and C represent specific characteristics in our failed state data set. The probability of the answer to the question is derived mathematically. In short, the higher the probability a variable is close to y, determines which variables are chosen.[92] Leo Breiman describes this classifier question and response split (our example is A>.50) as a binary tree.[93] Each tree split off makes up the forest. With each question that is answered, it develops pattern recognition. It is through this bootstrap algorithm that lends the predictability label commonly placed in these machine learning programs.

[89] Numerous articles discuss CART and its predictive characteristics. A good source can be found in the following article. Loh, Wei-Yin. "Classification and regression trees." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1.1 (2011): 14-23.
[90] Machine learning is another term for the computer creating predictive results based on input from the user over time. It is from information that the computer determines a pattern of behavior and predicts the users thinking.
[91] These techniques described in this paper are over simplified. For specific algorithm and distinctive mathematical equations and descriptions of all process within this statistical software programs employed refer to ibid2, Hmisc and Boruta
[92] Ibid 2
[93] Breiman, Leo, et al. *Classification and regression trees*. CRC press, 1984.

The end results in Random Forest are given with a percentage (%) of variance explained, as well as percentage (%) Inc MSE, also known as percentage increase over the Mean square error rate, and the Inc Node Purity, or increase of node purity. A percentage of variance explained is meant to tell the programmer how much variance in the variables could be explained. In other words, it is the probability percentage that was determined from the binary yes and no questions. If you look at the results from random forest on page (29) you can see that from one of our data sets, the percentage of variance explained was only 37.09%. So the program was able to predict accuracy 37 % of the time. This is not uncommon in very large noisy data sets that you are trying to whittle down. It is also because of this prediction accuracy rate that we used another program in conjunction with random forest to identify all relevant variables.

Percent increase mean square error as seen on table 3 and 4 illustrates a function of risk[94] or the percentage of the Gini impurity[95] being wrong corresponding to the "squared error loss"[96]. As the percentage increases, the mean square error loss measured is less. We can see the %IncMSE on Table 4. The higher the percentage, the higher the predictability of a particular variable Mean Square Error measures the average of the squares of the errors.[97]

The final item to evaluate in the random forest results are the Increase Node Purity as seen on Table 4. This increase just describes the probability of the decision made for the node split. The higher the node purity equals the higher the correlation to the failed state in our case.

---

[94] Grömping, Ulrike. "Variable importance assessment in regression: linear regression versus random forest." *The American Statistician* 63.4 (2009).
[95] Ibid 2
[96] Ibid2
[97] In regression analysis MSE is sometimes referred to the "unbiased estimate of error variance" .It can also mean squared prediction error.

Hmisc,[98] also referred to as Harrell Miscellaneous, is correlation matrix program. Using the spearman command within this software we can look for direct relationships between variables. When there is a large data set that you are trying to pare down, Hmisc[99] can identify close relationships between two variables. When you have extremely similar variables, the excess variable is considered redundant and should be dropped.

Boruta[100] is a software package written in R. It was designed to be run with Random Forest to specifically identify all relevant and non relevant variables. Random Forest gives us a classification predictor variable set from the forest creation - in our case, those variables most significant to the failed state. However, it does not tell us degrees of relevance for all of the variables. In some schools it is necessary to know all relevant variables, not just the most relevant ones. For example, in the case of medical research, a patient is worked up for cancer and presents with four symptoms, however, none of the four are the actual tumor itself. Even though the four are not a tumor, loss of appetite, low blood count, headaches and vomiting may still suggest presence of a brain tumor. Therefore all symptoms are relevant and should not be omitted. In the case of the failed state, high unemployment, exodus of the intelligence community (brain-drain), and rebel groups gaining power may not individually indicate a failed state, but collectively would be considered relevant and warrant further evaluation for it.

Boruta[101] is used in this methodology to pick out all relevant variables, not just the most relevant. We mentioned when describing random forest that there was a percentage of variance

---

[98] Ibid 1. http://cran.r-project.org/web/packages/Hmisc/index.html
[99] Ibid 87
[100] For a summary of the Boruta process see Kursa, Miron B. and Withold R. Rudnicki. "Feature Selection with the Boruta Package". *Journal of Statistical Software.* September 2010 Vol. 36. Issue 11. http://www.jstoresoft.org.
[101] Ibid 89

that could not be explained.[102] Boruta makes up for this by its algorithm which was designed to specifically wrap around the random forest program and its results. The program works by creating a shadow of all attributes, then calculates a Z score on this variable set. The Z score takes into "account the fluctuations of the mean accuracy loss in the trees from random forest".[103] Next, a minimum and a maximum Z score if derived from the shadow attributes in the program. Boruta assigns a hit to every attribute that scored better than the Mean Z score attribute or MZSA. Attributes which have importance significantly lower than MZSA are deemed unimportant and conversely those attributes scoring higher than MZSA are identified as being important. These values of importance are assigned for all the attributes. [104] When we look at table 5, we can see the mean Z, median Z, Min Z and Max Z as well as the decision. The range for decision and the Z score is different for each variable as we can see from our table. It is however the decision that we deem most important. Again, each variable needed to have a percentage of hits within the range of the Z score calculated for each to be deemed important. These scores are produced just to understand the range, not to imply that we use one particular point of importance over another. As long as the hits are within the range of min to max, it is considered valid in this program.

Although this method is very different from traditional analogies of the failed state, this research and the methodology used to evaluate the research question is intended to bring heightened credibility to a traditionally soft science field and provide accessible opportunity for future research.

[102] See results from random forest page 32
[103] Kursa, Miron B., and Witold R. Rudnicki. "Feature selection with the Boruta package." (2010).
[104] For a detailed account about the Boruta program, its authors describe in detail the full workings and equations for it. Kursa, Miron and Witold R. Rudnicki."Feature Selection with the Boruta Package" *Journal of Statistical Software* Sept. 2010, VOl 36, Issue 11.

Even though R has its own idiosyncrasies, namely the necessity of learning to code in R a fair amount due to the few graphical user interface options, the advantages of R outweighs the idiosyncrasy and scholars are gravitating to it because of the opportunities and flexibility it offers.

Many tree-based models have been used for data research in the applied sciences, including, medical purposes-arriving at prognosis and best treatment of cancer patients based on variables processed using such models. Since data sets can be altered with variable condition changes, the applied sciences make good use of the findings offered by these models. However, there is no good reason the social sciences cannot and should not be making equal use of such modeling. With regard to the failed state, we can make new data sets available from countries that were not as internationally transparent at earlier dates, thereby allowing them to be included in the data mix-adding to the comprehensiveness and cohesiveness of the data and the model.

Raw data for the purposes of this model was gleaned from multiple resources.[105] While collecting data, it was apparent that many data sets covered a specific area of government or function in the state, such as voting freedom, rule of law, women suffrage, or level of democracy. However, not all sources could be used because of the format of data storage.[106] Freedom House has an annual publication that describes states as one of the three: NF (not free), PF (partially free), and F (free). Although statistical computing has been employed, the data stores for their interpretation are not readily available. International indices evaluating life expectancy based on access to health care (amount the government pays per individual for health

[105] Ibid 3

[106] The data used fit a particular format. It needed to be numeric, and it had to be downloadable. If the data needed to be streamline and assigned values, other than freedom house, it was too time consuming and confusion to assign numeric values. Additionally, when assigning numeric values, there is no guarantee that it could be easily reproduced which would have been contrary to using accessible international data sources.

care), immunization rates, rate of curable diseases, drought conditions, access to education, and economic GDP were however used for our research purposes.[107] Additionally, raw resource data (oil, minerals), agriculture information, information technology and the following were added: information on human trafficking/sexual slavery, FGM (female genital mutilation), organ theft, ethnocide, and states' solvency. The United Nations has made an attempt to make this data publicly available.[108]

Incorporating as much of the good and useful bits of data derived from current indices and models, this research attempts to go even further by being more comprehensive and employing a unique methodology in this field to derive at a reliable and rather comprehensive predictive variable set, that can be resurrected or re-created for further research in the field through statistics.

After deciding on a software medium for which to create the failed state model, several other steps were essential in the process. In order to do this properly, it was necessary to list all the countries for which data was available. Several originating data sources came from countries other than the United States, in which a particular country's data was listed under a different name such as Republic of Korea, for example, which we will refer to as North Korea for our purposes. Individual countries were omitted from being included either because they lacked enough data or were missing sufficient data that could render them as useable/useful. These countries included but were not limited to Uzbekistan, Turkistan, among others. The reason(s) for a country demonstrating a dearth of data included the age of the country, (such as the above-mentioned, of which two were former members of The USSR, now Russia, whose geographic

[107] For a full list of original raw data resources see Ibid 3
[108] Ibid 3

boundaries have altered over recent years), resulting in a raw data gap. Also omitted were countries that had no data due to its provincial nature and/or its rule under others. Puerto Rico, for example, is a province of the United States and was therefore omitted. Scotland, Northern Ireland and the Falklands fall under the United Kingdom and were also left out of the study as was the Palestinian territories.[109] Given this, a resultant and still sizable 200 countries are included in this research.

An original failed state ranking for each of the 200 countries was created in which to run the collective data against. I used the CIA World Factbook[110] as the main source of information in which to determine each states present viability. This ranking would be used for this comprehensive model in which to run all of the other data against. The ranking is from 1 to 5, with 1 being the farthest removed from being a failed state, and a 5 considered a failed state. At the original time of data collection (2009)[111], the only two states that ranked a 5 were Iraq and Somalia, according to the definition previously presented in this paper. There have since been some states that have failed, which would be appropriate to include for further study. These 1-5 rankings were used so that anyone could replicate them according to the same definition and evaluation method used herein.

Again, the baseline is a number one, similar to the Fund for Peace. If a 1 was given, the country had a democracy in place, a strong economy, high age expectancy for its respective population, no sanctions posed against it and no alerts for travelers to the country. Airports, schools, and highway infrastructures were in place and the government spent a substantial

[109] See appendix 1 for country reconcile list.
[110] Ibid 3
[111] We began data collection in 2009 and as such, the results are based on the data collected during that this. World events however have not remained in 2009 and there are more examples and further data that was not available at the beginning of this research.

proportion of its revenue on its constituents. The rule of law was observed. The government was able to protect its own borders, and few if any refugees were exiting the country- spilling over into neighboring states or flowing into the country from bordering states.

The following key items pulled from the CIA World Factbook[112] were also considered for ranking: age of populous, type of government, and state of the economy that would warrant a change in the baseline score. These will be further detailed below.

Age and Condition of Populous

Consideration was given if over 70% of the population was under age 65. This would indicate that a state either has a short life-expectancy and/or a young populous was/is prominent. Resultant civil societies are fewer in the country due to the younger populous and government can also be weak with a predominantly younger populace. The birth rate, death rate, and migration rates significantly alter the average age of the residents. Any or all of these could be indicative of civil unrest which can create a state's inability to provide even for the basic needs of protecting its people. Additionally, health expenditures as a percentage of GDP was considered within this category, for example, the percentage of children under age five who were/are underweight. The literacy rate over age fifteen was also a factor. A number one was added to the base score accordingly given the considerations above.

Type of Government

The description for each respective state was evaluated from the CIA World Factbook.[113] We first looked to see if a government was identified. If no government was specifically

[112] Ibid 3
[113] Ibid 3

identified, such as democratic, autocratic etc., or if the government structure was unclear, then a one was added to the score. If there was, for example, an Islamic republic associated with it– we know that Islamic republics are considered more unstable than democracies or a republic, a one was added to the score. If it was not clear about government type, but documentation suggested that there was no constitution in place, if there was not a separation in the branches of government, or if it was defined as an authoritarian regime, a point was added to the failed state baseline. Note, most stable or defined strong regimes were clearly stated as such in the CIA World Factbook.

State of the Economy

A critical question when looking at the failed state is whether the economy is being propped up by international actors or is solvent. A country's unemployment rate, its percentage of GDP inflation, the type of products that were/are produced by the state, i.e., whether the state is natural-resource abundant, or what other resourc(es)/industry is used to sustain itself as its main sources of revenue were all considered. If there was a particularly high unemployment rate and it was known that government workers had not been paid in awhile, another point was added. We deemed arrear wages owed as an indication of either potential state insolvency, or misappropriation of state funds by its leaders.

In addition to the categories above, I looked to see if there was mass exodus of refugees or refugee pours instead of traditional migration patterns. If a transfer of population was due to civil unrest or loss of rule of law, another point was added. [114]

[114] See appendix 2

There were originally 600 plus variables from the data collected considered for 200 countries in this research, which needed to be condensed or simplified for working purposes. I divided the data into 20 separate files in which to read into the program R. R works by loading all of the data into ram at the same time.[115] Some computer languages and software platforms/programs will take the command that you give it, then the processor picks it up, puts it into ram, runs it, puts it back into memory and picks up the next command thus leaving ram and the processor relatively unencumbered, fast and free. R does not do this. Although it is quick, adaptable, and can handle large data sets, R still likes reasonable sized data sets. There are limits due to processor speed and memory space on any PC. These sets are still larger than other statistical programs would be able to handle on a laptop, but they still needed to be divided.

Each data set was run in the same fashion to determine predictor variables for the failed state prior to arriving at a final data set. To detail the methodology employed, I will work through just one of the data sets (BF_BO).[116]

I checked a tally to determine how many missing variables were in the data set.[117]

Our code

```
apply(BF_BO, 2, function(x) length(which(is.na(x))))
```

If you look at Table 1, you can see that GDPpercap_2005 is missing data for 22 different countries in the list. Since we have 200 countries in the data set, this is not considered a large percentage missing, and therefore, would be acceptable to keep.

---

[115] Because R loads all of the data into Ram at the same time, it can process all of the information there very quickly, however too much in a data set can be an issue for a regular PC. Chopping up the data therefore is helpful.

[116] See appendix 3 for R code used.

[117] See appendix 3 for full list of all R code employed.

Table 1

Country	Failing_Rank	LifeBirth_2005	AdultLit_2005
0	0	16	14
EnrolEduc_2005	GDPpercap_2005	LifeBirth_2006	AdultLit_2006
11	22	15	14
EnrolEduc_2006	GDPpercap_2006	HDI_2006 HDI_A1	Reconcil_Rank
11	19	21	21

After this, the data was run through a software package for R called Harrell Miscellaneous or Hmisc[118] which was able to help look at the relationships between the variables in a correlation matrix.[119]

We use the spearman command in Hmisc.

Our Code

```
library(Hmisc)
r.results <- rcorr(as.matrix(BF_BO[,2:ncol(BF_BO)]), type="spearman")
r.results$r
```

The matrix in Table 2 produced through Hmisc[120] allows us to look at the relationship between variables prior to running the data through random forest.[121] You can see the closer the

[118] Harrel, Franck E. Jr. Cran 2012-10-25. 14:00:08 Http://biostat.mc.vanerbuilt.edu
[119] See appendix 3 for all R code employed.

relationship between two variables is indicated by their numeric proximity to one another either positive or negative. The negative relationship indicates that as one variable goes up, the other goes down and vice verse. As a rule, any pair whose relationship is higher than .80, for the purpose of this paper, is deemed highly correlated and therefore one is redundant. In other words, both highly related variables predict in the same way usually necessitating the use of only one. For example GDP and GDP per Cap for the year 2006 could be highly correlated. We would therefore choose the one with a more complete data set for our purposes. Note is taken for all pairs bearing this relationship determined by Hmisc and are considered before final pairing prior to a Random Forest run.

For this particular data set you can see that HDI_A1Reconcil_Rank and HDI_2006 have a negative strong correlation of -0.999986410.[122] So these two variables are noted for final selection of variables to omit one due to this relationship.

Again, all of the variables are run through Hmisc[123] prior to randomForest[124] to locate directly related variables in an effort to either further solidify a relationship validated eventually through randomForest[125] and Boruta[126], or to extract it from the data set due to its redundancy. The end goal of this exercise is to end up with a model of about twenty or so variables from the hundreds that we started with which show a predictable relationship to the failed state statistically through random forest and Boruta.

---

[120] Ibid 87
[121] Ibid 2
[122] See table 2
[123] Ibid 87
[124] Ibid 2
[125] Ibid 2
[126] Ibid 89

Table 2

	Failing_Rank	LifeBirth_2005	AdultLit_2005	EnrolEduc_2005
Failing_Rank	1	0.06242993	0.02658016	0.076718941
LifeBirth_2005	0.06242993	1	0.05348156	-0.016884008
AdultLit_2005	0.02658016	0.05348156	1	0.046367958
EnrolEduc_2005	0.07671894	-0.01688401	0.04636796	1
GDPpercap_2005	-0.26392859	-0.06410334	-0.03690866	0.086575672
LifeBirth_2006	0.27416915	0.12889048	0.14461656	0.095038928
AdultLit_2006	0.09985825	0.05220463	0.62730968	0.017374132
EnrolEduc_2006	0.04383237	0.07998408	0.11584348	0.100154713
GDPpercap_2006	0.01886463	-0.0494695	-0.04241028	0.066649035
HDI_2006	-0.63302928	-0.03570063	-0.19248299	-0.005080359
HDI_A1Reconcil_Rank	0.63324142	0.03642339	0.19208394	0.005436544
	GDPpercap_2005	LifeBirth_2006	AdultLit_2006	EnrolEduc_2006
Failing_Rank	-0.263928592	0.27416915	0.09985825	0.043832373
LifeBirth_2005	-0.064103335	0.12889048	0.05220463	0.079984084
AdultLit_2005	-0.036908664	0.14461656	0.62730968	0.115843482
EnrolEduc_2005	0.086575672	0.09503893	0.01737413	0.100154713
GDPpercap_2005	1	-0.22368312	-0.07090337	0.003077174
LifeBirth_2006	-0.223683119	1	0.30828696	0.112480037
AdultLit_2006	-0.070903368	0.30828696	1	0.175718904
EnrolEduc_2006	0.003077174	0.11248004	0.1757189	1
GDPpercap_2006	-0.038675476	-0.05839336	-0.15373917	0.006740789
HDI_2006	0.430028647	-0.51835126	-0.38356614	-0.066614233
HDI_A1Reconcil_Rank	-0.429583192	0.51815039	0.38339147	0.066481166
	GDPpercap_2006	HDI_2006	HDI_A1Reconcil_Rank	
Failing_Rank	0.01886463	-0.633029282	0.633241415	
LifeBirth_2005	-0.0494695	-0.03570063	0.036423385	
AdultLit_2005	-0.04241028	-0.192482993	0.19208394	
EnrolEduc_2005	0.06664903	-0.005080359	0.005436544	
GDPpercap_2005	-0.03867548	0.430028647	-0.429583192	
LifeBirth_2006	-0.05839336	-0.518351257	0.518150389	
AdultLit_2006	-0.15373917	-0.383566141	0.38339147	
EnrolEduc_2006	0.00674079	-0.066614233	0.066481166	
GDPpercap_2006	1	0.028291663	-0.028308809	
HDI_2006	0.02829166	1	-0.99998641	
HDI_A1Reconcil_Rank	-0.02830881	-0.99998641	1	

Next we tell R to run Random Forest after looking for incomplete data sets that we may want to omit depending, or impute separately outside of random forest. Random Forest will impute values on its own in missing data sets if we tell it. Again, if there is a significant amount of missing data, those variables were dropped prior to running the program. This particular data set does not have enough missing data to warrant dropping the variables prior to running Random Forest. I did a trial run on this data set and several others to determine if there was a significant difference between dropping values, imputing separately, or just running it through Random Forest and deemed that just running the data through Random Forest at this point without imputation was sufficient for our purposes.[127]

Our Code

```
search()
library(randomForest)
str(BF_BO)
BF_BO$Failing_Rank
head(BF_BO)
BF_BO$FailingRank[BF_BO$FailingRank==NA,]
Incomplete <- BF_BO[!complete.cases(BF_BO),]
Incomplete
Incomplete$Country
set.seed(456)
BF_BO.rf=randomForest(Failing_Rank~.,data=BF_BO.data,importance=TRUE)
BF_BO.rf
```

[127] There is some degree of imputation automatically programmed into the software package of randomForest, however one has the ability to impute the data set prior to even running the program and beginning with a complete data set if desired. We omitted variables with large gaps in our data set so that we had a more clean result. For full list of R code employed See appendix 3.

The Results from randomForest

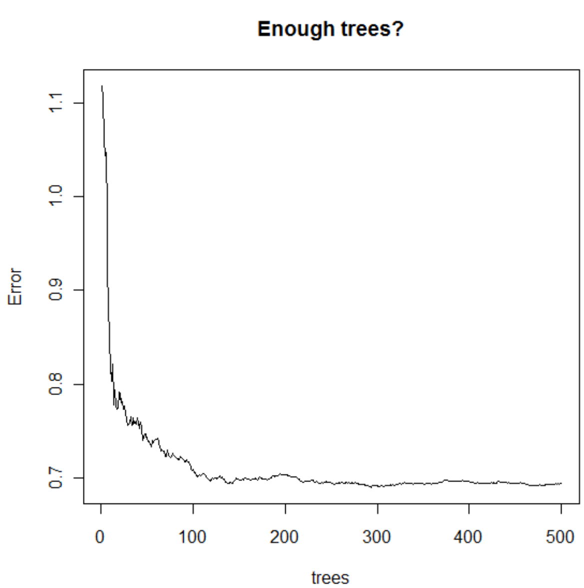
```
Call:
 randomForest(formula = Failing_Rank ~ ., data = BF_BO.data, importance = TRUE)
               Type of random forest: regression
                     Number of trees: 500
No. of variables tried at each split: 3

          Mean of squared residuals: 0.6947214
                    % Var explained: 37.09
```

Next we graph these results and ask the question, do we have enough trees? Would running the program more times yield different results? We grew 500 trees in our sample.

Our code

```
plot(BF_BO.rf, lty=1, main="Enough trees?")
```

Enough Trees Graph



We can see from our graph that there appears to be little change in the error rate, roughly, after 250 trees where our line begins to flatten out. So we know that with this data set, it is not necessary to increase the amount of sampling employed/trees grown.

Next we plot and make a graph of the Random Forest results according to variable importance.

Our code

```
round(importance(BF_BO.rf), 2)
varImpPlot(BF_BO.rf)
```

Table 3

	%IncMSE	IncNodePurity
LifeBirth_2005	-1.57	3.79
AdultLit_2005	10.87	11.65
EnrolEduc_2005	6.96	13.52
GDPpercap_2005	8.35	22.23
LifeBirth_2006	2.93	11.58
AdultLit_2006	5.74	6.94
EnrolEduc_2006	-0.94	9.14
GDPpercap_2006	3.45	11.22
HDI_2006	20.15	49.14

As we mentioned earlier, the higher the node purity and the higher the percentage of increase of the mean square error rate indicates a higher correlation to the failed state. These two things are the predictor measure of accuracy in randomForest.[128] "In regression trees, node impurity is measured by MSE, therefore the second measure that averages cumulative reduction in node impurity due to splits by a variable over all trees is labels as mean decrease in MSE. When you permute the value of a variable in out of bag data and make a prediction, the

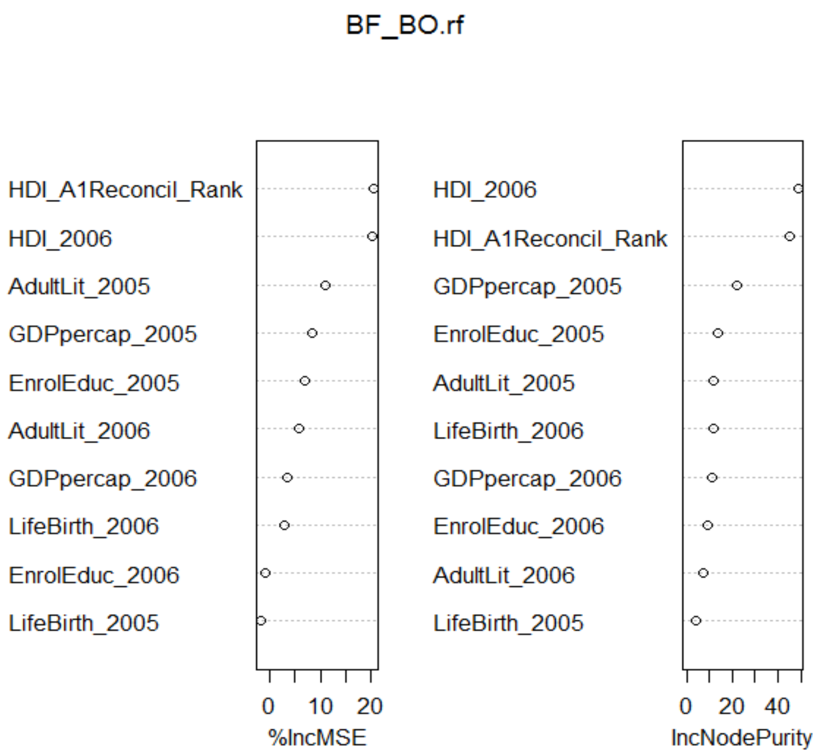
[128] Svetnik, Vladimir, et al. "Random forest: a classification and regression tool for compound classification and QSAR modeling." *Journal of chemical information and computer sciences* 43.6 (2003): 1947-1958.

expectation is that the MSE will increase, especially if the variable has some importance- hence the label % INC MSE or % increase MSE.[129]

Our graph on Table 4 shows the variables for this particular data set in descending order of importance. According to the graph, we see variables that seem to have the highest ranking on node purity, and mean square error explained (the variables or characteristics that seem to be prevalent regarding our failed state ranking). These would be HDI_2006 and HDI_A1Reconcil_Rank. We took note of these two variables when Hmisc[130] was run in R[131] which deemed these two variables to be almost identical in response to the failed state ranking. So, one of these will be dropped before determining the final model set. Next, it is always a good idea to look at partial dependence plots to see if there are any obvious relationships to take into account.[132] The results are on Table 4 on the following page.

[129] Ibid 128
[130] Ibid 87
[131] Ibid 1
[132] See appendix 3 for a complete list of all R code employed for this paper.

Table 4

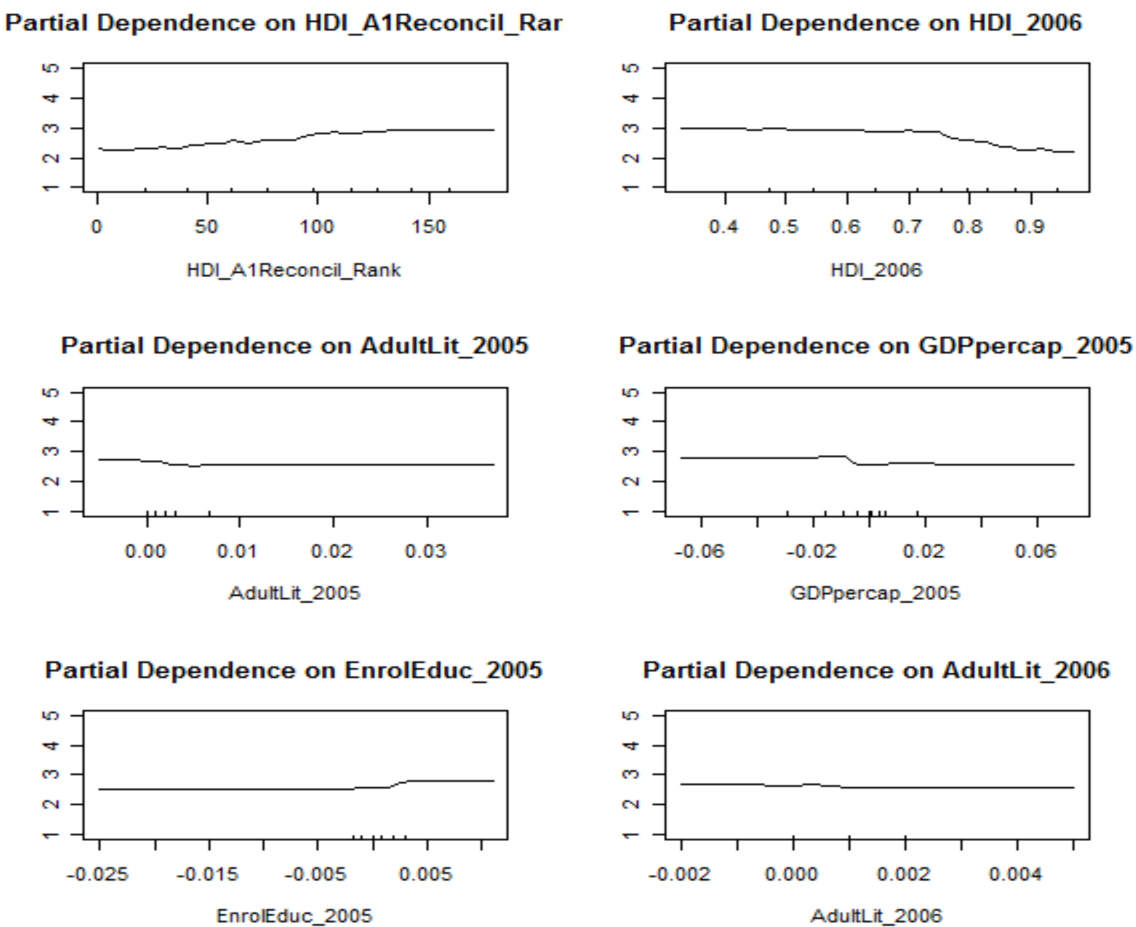


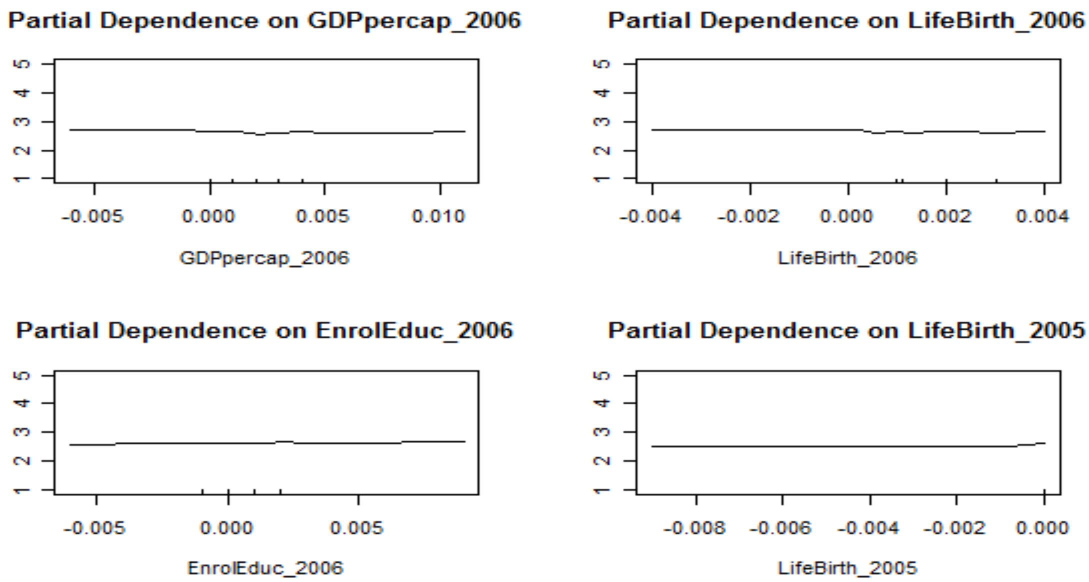
We graph all of the dependent variables using partial dependence plots ran through Random Forest with their results.

Our Code

```
imp<-importance(BF_BO.rf)
impvar<-rownames(imp)[order(imp[,1], decreasing=TRUE)]
par(mfrow=c(3,2))
for (i in seq_along(impvar)) {
  partialPlot(BF_BO.rf, BF_BO.data, impvar[i], xlab=impvar[i],
  main=paste("Partial Dependence on", impvar[i]), ylim=c(1,5))
  }
par(mfrow=c(1,1))
```

Table 5 Partial Dependence Plots





There seems to be little obvious responding variables from these graphs, so we employ the Boruta program to identify all relevant variables. As previously mentioned, Boruta identifies all relevant variables.

Here is our code…[133]

```
library(Boruta)
set.seed(498)
BF_BO.Boruta = Boruta(Failing_Rank~.,data=BF_BO.data,doTrace=2,maxRuns=400)
BF_BO.Boruta
attr = attStats(BF_BO.Boruta)
attr[order(-attr$medianZ),]
```

[133] For all code employed in the making of the model see appendix 3

Table 6

	Mean Z	Median Z	Min Z	Max Z	norm Hits	Decision
HDI_A1 Reconcil_Rank	19.9603965	19.976504	16.847558	23.020708	1	Confirmed
HDI_2006	18.932823	18.9644994	15.990646	21.516691	1	confirmed
AdultLit_2005	9.0000899	9.0545373	6.348779	12.607887	0.99302326	confirmed
GDP_percap_2005	8.693459	8.6262783	5.704974	11.813586	0.98837209	confirmed
AdultLit_2006	6.1957459	6.213406	2.813337	9.729706	0.92790698	confirmed
EnrolEduc_2005	5.9797098	5.9268323	2.311391	9.381075	0.88837209	confirmed
LifeBirth_2006	2.6465938	2.6576754	-1.351819	5.557084	0.43953488	tentative
GDPpercap_2006	1.7449177	1.8202295	-1.066098	4.13523	0.05116279	rejected
EnrolEduc_2006	-0.7323241	-0.6133068	-3.211732	1.076007	0	rejected
LifeBirth_2005	-0.9960767	-0.9855453	-3.39596	1.002132	0	rejected

We can see that the top two variables have a significant response to the ranking in table 5 and the range Boruta determined for each variable as well as hits associated with it. Note the top two variables that were previously identified as important through Random Forest, HDI_A1 Reconcil_rank and HDI_2006. In addition, we have other confirmed variables that we otherwise might have missed. Boruta determined the following variables to be relevant: HDI_A1Reconcil_Rank, HDI_2006, AdultLit_2005, GDPpercap_2005, AdultLit_2006, and EnrolEduc_2005. From this data set HDI_A1Reconcil_Rank will be dropped. The others are saved into another file.

This method was employed for all of the twenty different sub-data sets, from the 600 plus original raw data variables, employed to create the model. The most relevant variables pulled from the different data sets were then subject to Random Forest and Boruta once more before the final model and its variables were confirmed. This was a statistical manner of process and elimination to derive a final variable set.

So how does this model and method differ from previous research and why is this unique? First of all, most past research is offered in the form of an index, not a model created using statistics for variable selection, although it may be labeled as a model. The Fund for Peace Failed State Index that is published yearly in the Foreign Policy publication is an example.[134] The findings published are created through a CAST method[135] which is in part a statistical process as well an alteration of their statistical findings depending on the opinion of the analyst. Although qualitative interpretation is necessary in political science research, when creating any kind of index and rankings for a failed state, you can lose the opportunity to discover a variable and/or identifier that otherwise might be omitted from the data because it had not been considered before as being a key indicator of the failed state. If you are careful with what statistical program you choose and decrease variance error rates properly, you can use statistical programs for variable selection and then create a model useful for evaluating a states solvency and legitimacy. This is what we have done here as defined in the methodology. The CAST method referenced in the Fund for Peace is used to create an index. An index shows you where you are based on past events. CASTs' predictive capabilities are limited because the variables themselves were altered and the data used was altered based on the biases of the analyst. When the statistical program score is determined then altered or omitted by the researcher, it can bring all of the results into question, not to mention that you can miss something key to the research by assuming that it is a false positive.

[134] Ibid 55
[135] Ibid 57

To better illustrate this point of an index we can look at a very common index to Americans-the Standard and Poor's 500. This is a stock market index.[136] The index itself demonstrates at any given point the demand for a stock or option on a stock and the price that the last person was willing to pay for it. With a documented history of stock prices for all companies in the index, one can look at past analysis of a stock's performance through any given company's public trading history. The stock price and the index do not make inferences for future performance and it does not predict future solvency of any given company. It merely represents where it is today and gives documentation where it was. Millions have been spent to come up with a predictive stock analysis model.

Although previous researchers employ statistical methods to come with their rankings for the failed state, their indices still just tell you where a country was yesterday. The purposed outlined in the methodology is to take all such raw data from international indices available and after formulating a failed state ranking score, using a machine learning algorithm, in this case Random Forest, to create a model that can derive key characteristics/variables from the program. This program then illustrates variables that have strong correlations or a relationship to an increased failed state ranking. As complex and comprehensive as the data collection and processing was, this construct ends up being a basic but very functional and useful model. [137]

Although this research has a main quantitative focus, qualitative means were used to evaluate the justification of the independent variable ranking numbers (original failed state ranking, aka independent variable, previously described) against three known failed states and

[136] Kawaller, I.G., Koch, P.D. and Koch, T.W. (1987), "The Temporal Price Relationship between S&P 500 Futures and the S&P 500 Index". *The Journal of Finance*, 42: 1309–1329.
[137] See methodology section of this paper pp.

their conditions to demonstrate the validity of the model. Analysis of the results of this quantitative research and qualitative means of explanation is the subject of the next section.

Analysis of the data final model and failed state examples

After running all of the separate data clusters through the methodology described, what remains is a final variable set for our model. [138]

Again, the first step in the final analysis was to look at the data set through Hmisc. [139] By using the spearman command and the correlation matrix through Hmisc, we considered which variables may have a similar response, and could therefore be considered statistically redundant. Because Random Forest takes all of the variables into consideration when building the forest in the model, it is important to remove as many variables that are duplicates in an effort to explain the variation in the node selection, generating a more concise model.

We looked to make sure that there were no duplicates. Then we loaded up Random Forest and imputed any missing data. Imputation was done in the final model to make as complete of a data set as possible for the final run. Our goal was to have a final set of about twenty variables for this model.[140] We ran the data set through random forest and came up with a set of variables[141]then checked to make sure that we had enough trees and that the change over adding one more tree did not make a large difference.

[138] For a the raw data list used See Appendix 8
[139] See appendix 4 for Hmisc correlation matrix results table.
[140] Note, we began with over 600 variables and ended with approximately 20 for model selection.
[141] See appendix 7

Below is our results.[142]

Random Forest Results

Type of random forest: regression

Number of trees: 500

No. of variables tried at each split: 17

Mean of squared residuals: 0.5901746

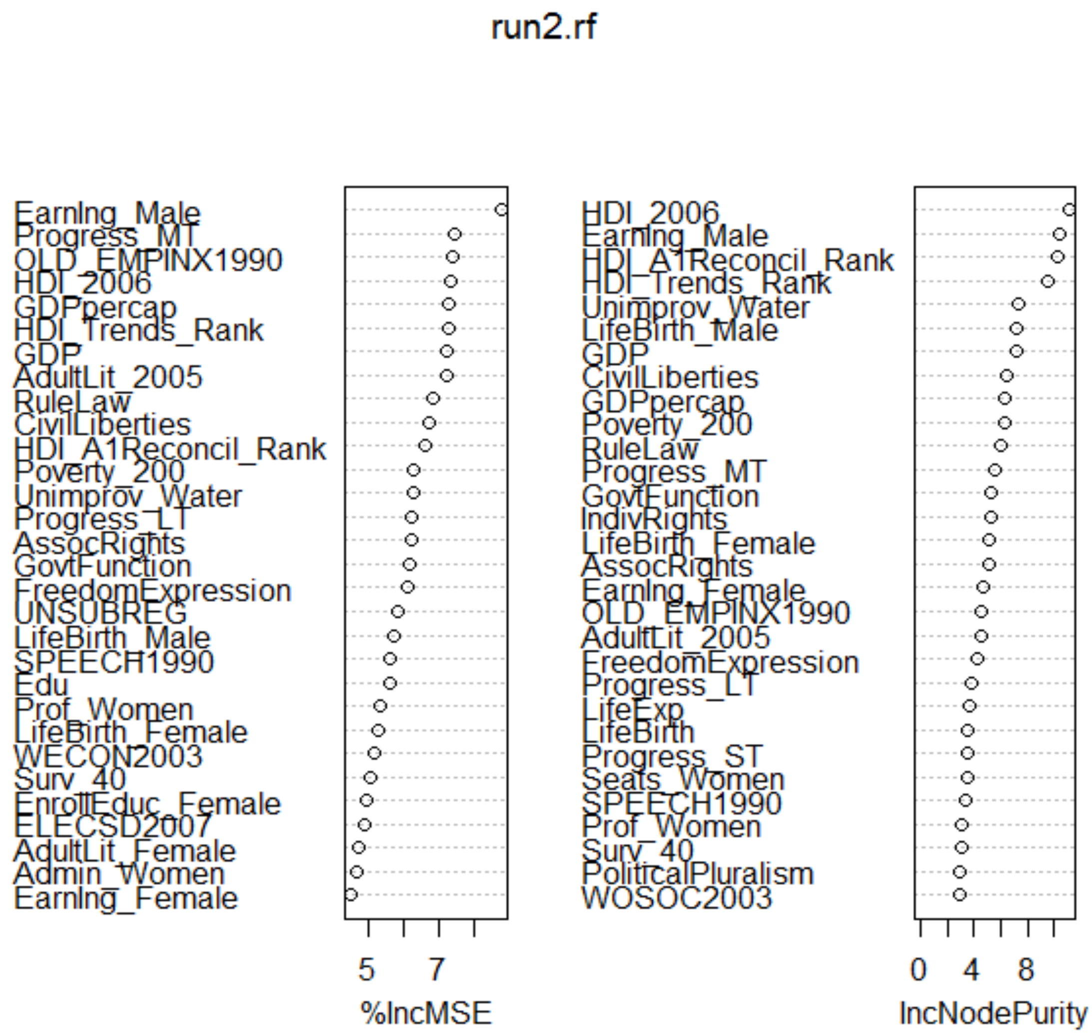
% Var explained: 45.91

A graph to view the results is found on Table 7.

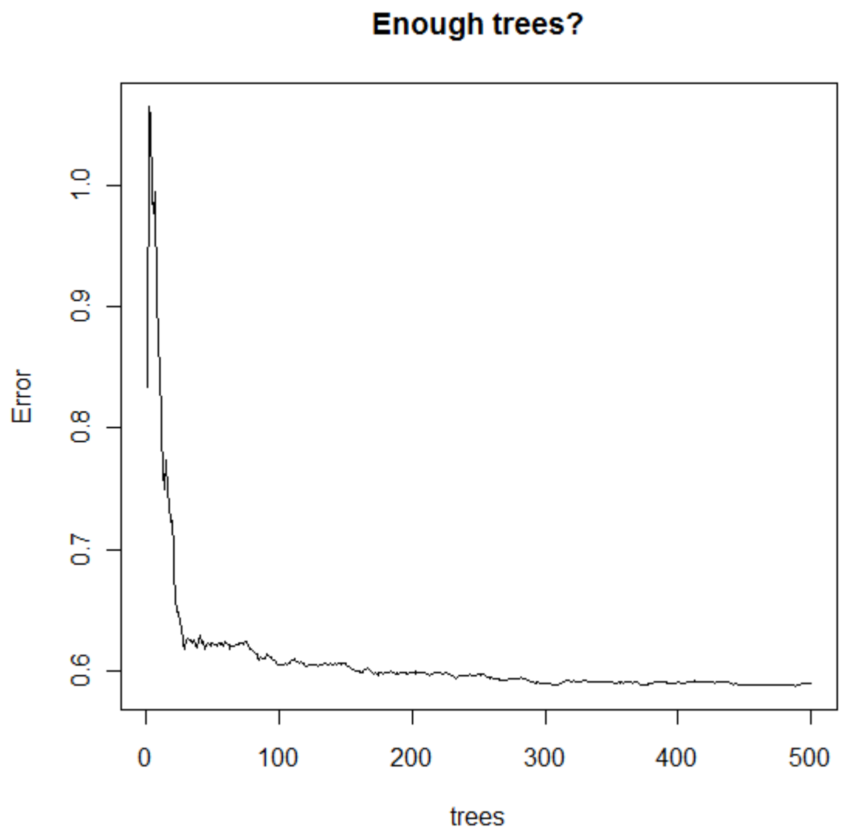
A few of the variables are rather obvious for showing a relationship. On our graph on Table 6 we see the male earnings, and HDI 2006 stand out at the top as having strong correlation to the failed state. Again, the higher the score indicates the higher the correlation to the failed state in both columns.

[142] For full list of R code employed see appendix 3

Table 7



Again we check to see if there were enough trees grown for the regression?



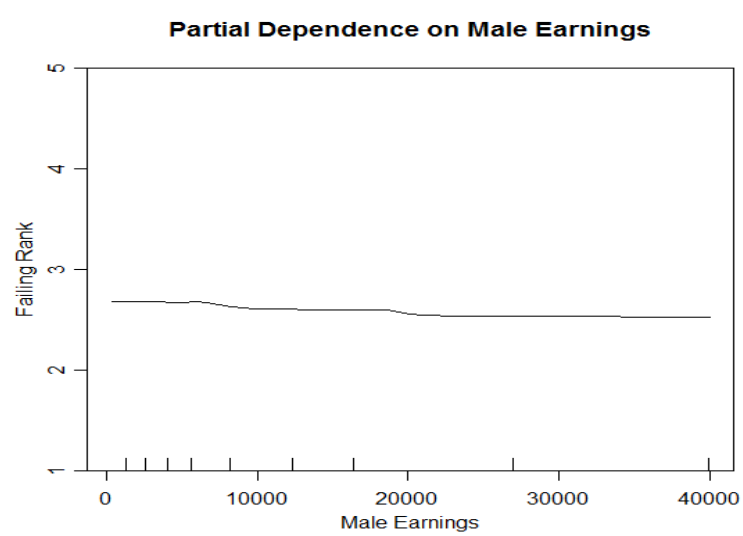
The next step was to run Boruta[143]. Although our Random Forest results graph gives us a nice look at the variables and their relationship to each other and the model, Boruta shows us all of the relevant variables after random forest has been run.[144] The final set of predictive variables for the failed state from this statistical practice are as follows under results in descending order of importance based on the statistical findings confirmed from the Random Forest regression paired with results from Boruta. [145]

[143] See appendix 3 for R code employed.
[144] See appendix 5 and 6
[145] See appendix 5 for Boruta results

Results

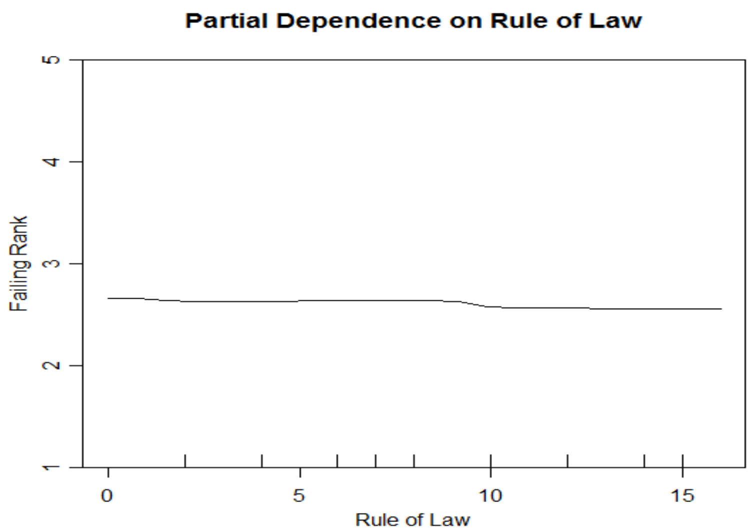
The results are Male earnings, rule of law, Human Development index, civil liberties, unimproved water, GDP, adult literacy, freedom of speech, associated rights, freedom of expression, life birth of males, poverty a percentage of the population, progress-mt, government function, followed by earnings per female, life birth female, individual rights, chance of survival over 40 years of age, political pluralism, progress-LT, Lifebirth, Life expectancy, WOSOC2003, Prof_Women, AdultLit_female, and education enrollment. [146]

Below are some individual partial plot graphs to show the relationships between the failed state ranking and variables confirmed in Boruta to have a statistical relationship.

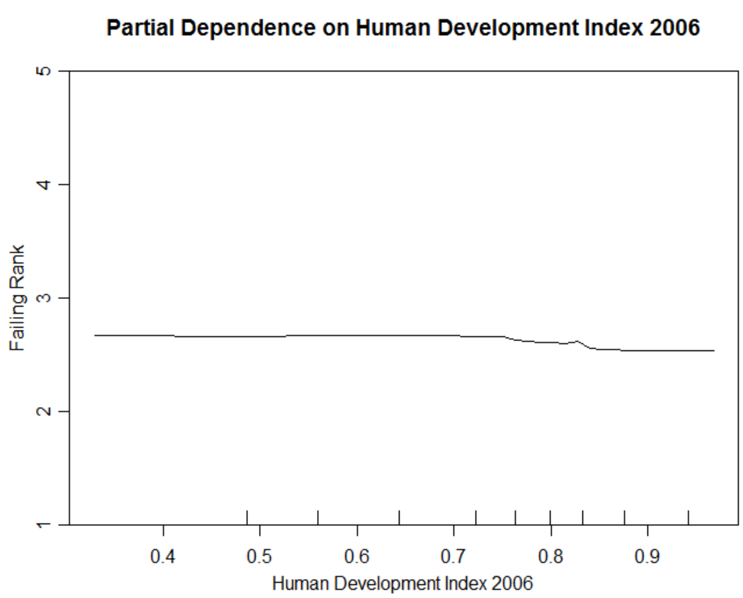


We can see in the male earnings partial plot that as earnings increase, the failed state ranking decreases. Male earnings is measured here as annual currency equal to U.S. dollars.

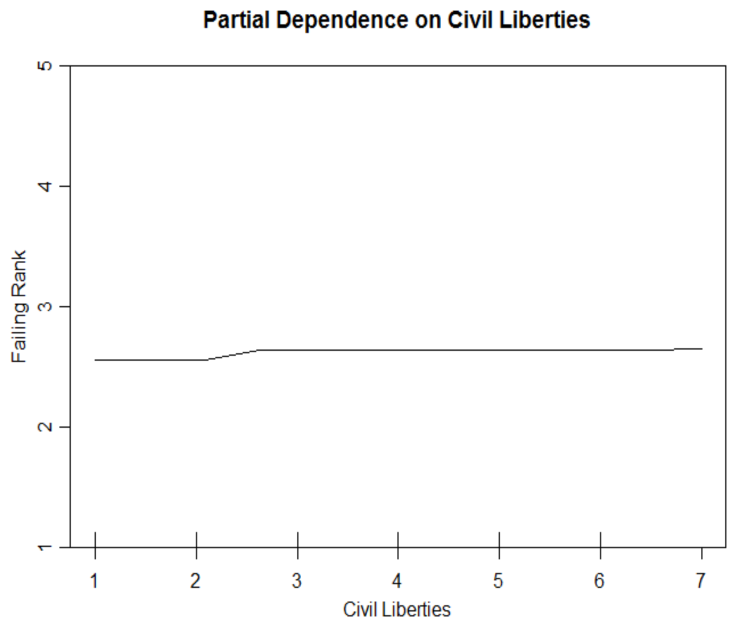
[146] See appendix 5 and 6 for the final model variables.



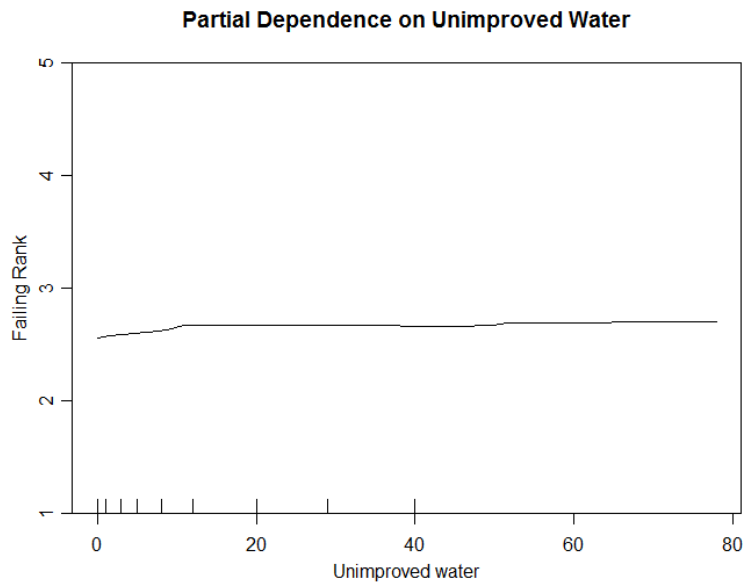
Rule of law is observed as a number scale based on several characteristics- Judicial process effectiveness and observance policing safeguards for civilians. You can see as the rule of law increases, the failed state ranking decreases.



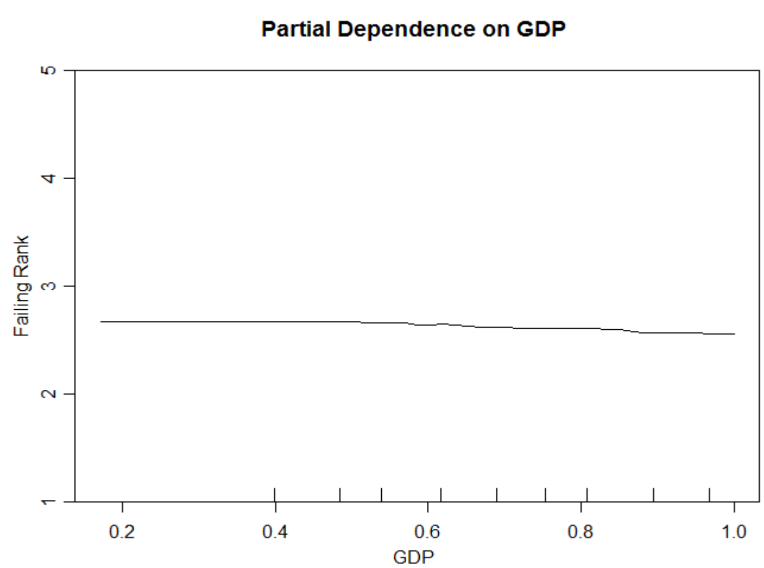
Here the human development index is measured on a percentage scale. It is meant to reflect a percentage of the population- .40 being 40 % and so on. As the failed state ranking goes down, the human development index increases.



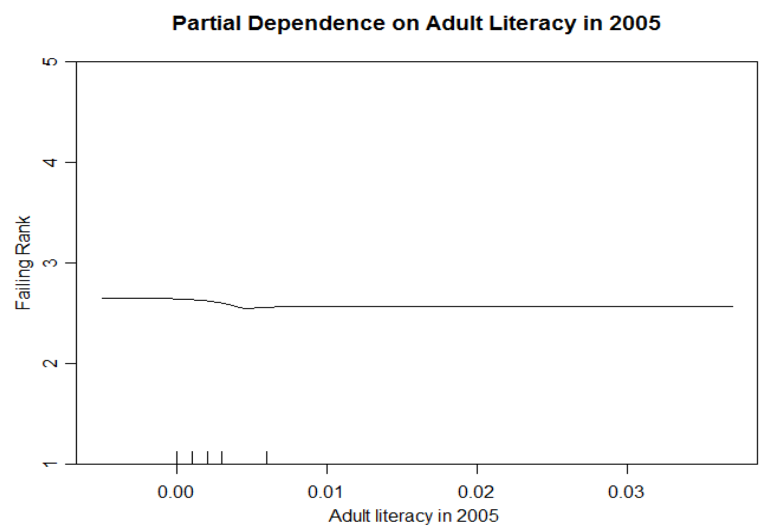
Civil liberties scale is a tally of overall civil liberties offered in a country. This relationship is a curious one. The curve is very slight. You can see that slowly as civil liberties increase, so does the failed state ranking. This finding was a surprise to me.



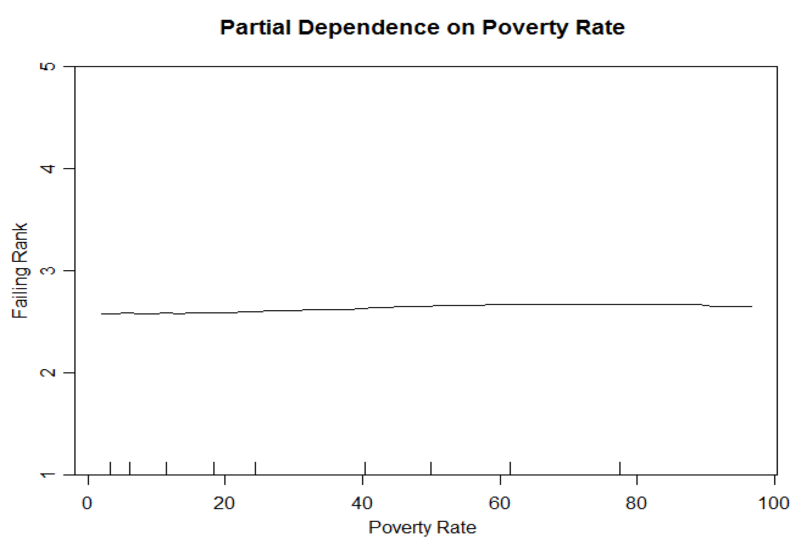
This partial plot shows a scale of the unimproved water in relationship to the percentage of the population that does not have access. It comes as no surprise that as the percentage of people without water improvements gets larger, so does the failed state ranking.



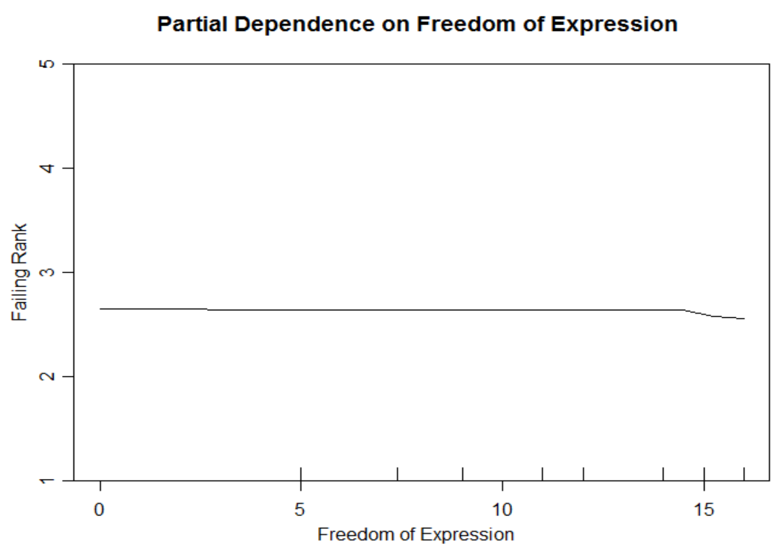
GDP is typically measured in U.S. dollars and is measured by the trillions. So a 1.0 would be one trillion and the percentages below that would be in the billions. Again we have a slightly discernible plot line that shows the failed state ranking going down as GDP increases.



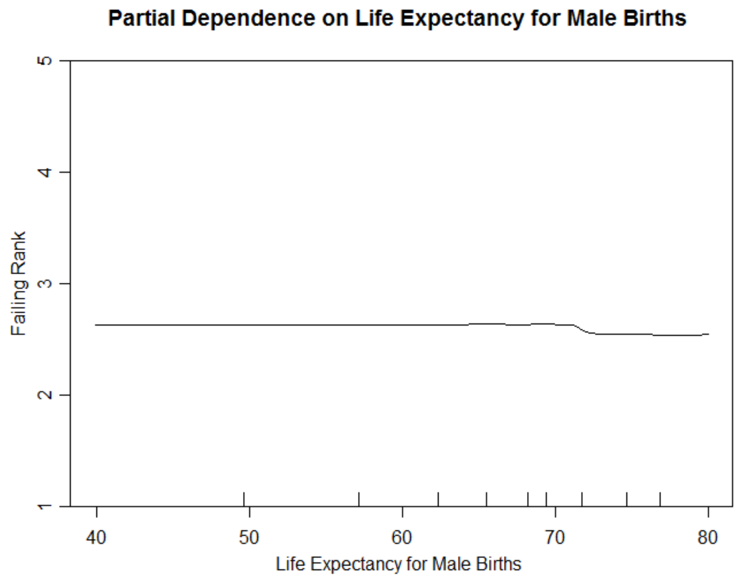
Adult literacy is measured as a percentage of the population that can read. As literacy goes up, failed state ranking goes down.



Poverty rate in the graph shows the complete spread of the population to 100%. Although a confirmed variable, this particular graph does not show an obvious relationship by itself. There is a slight decrease in the failed state at about the 90 % mark.

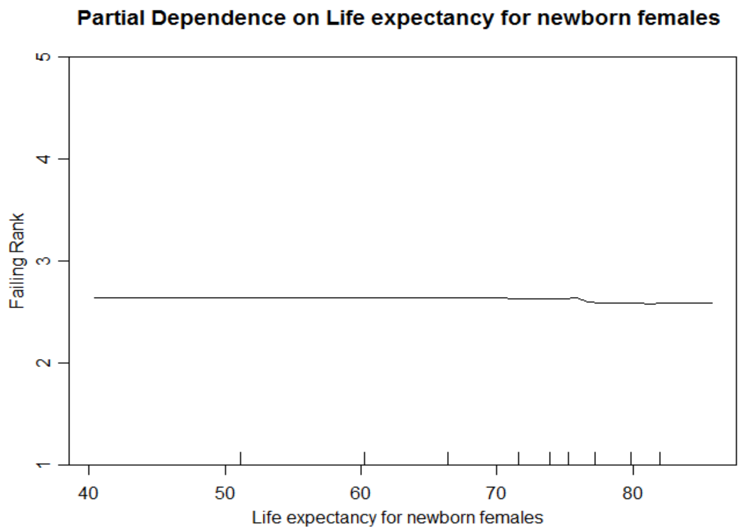


Freedom of expression is another variable measured in terms of score ranking from original data based on the amount of freedoms allowed. This encompasses freedom of speech, freedom of religion, and freedom of the press–the higher the measure of freedom of expression, the less chance of a failed state.

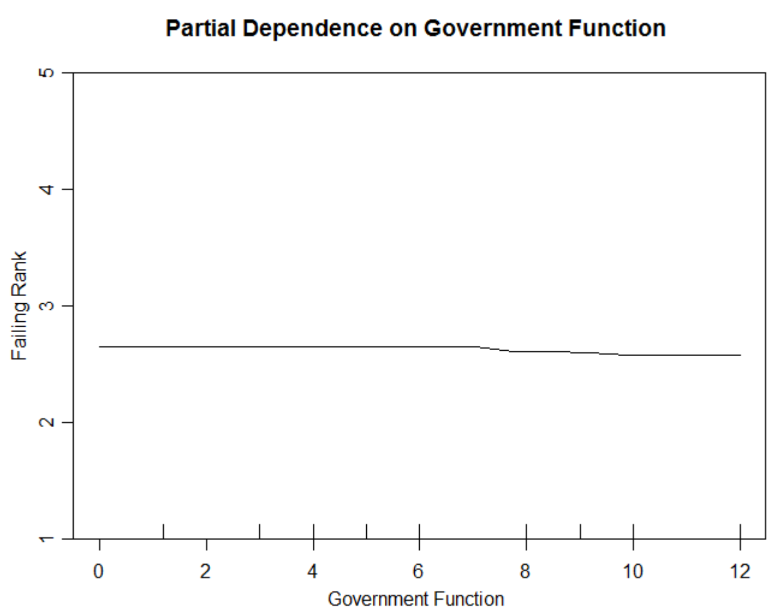


These numbers show the percentage of male boys born that live to age 5. The figure represents a percentage of the population of baby boys that survive to age 5. There are two different measures in the partial plots–life birth boys and a life expectancy for newborn females. It is widely known that baby boys do not fare as well as baby girls for survival particularly in food insecurity situations. Although the particular cause is unknown, regardless of medical research to determine otherwise, boys are more susceptible to respiratory failure and other problems associated with if proper care and nutrition is not available.[147]

[147] Elsmén, Emma, Mårten Steen, and Lena Hellström-Westas. "Sex and gender differences in newborn infants: why are boys at increased risk?." *Journal of Men's Health and Gender* 1.4 (2004): 303-311.



Similar to life expectancy for male births, life expectancy for newborn females is graphed in the overall percentage of babies' likelihood to survive to age 5. As this rate increases, we can see on the graph that the failing rank decreases.



Our last plotted example is government function. The numbers are a ranking of effectiveness for a government to function properly as a whole. Our plot shows a slight decrease of the failed state ranking as government function increases.

Table 8

Male earnings	Rule of law	Human development index	Civil liberties
Unimproved water	GDP	Adult literacy	Freedom of speech
Associated rights	Freedom of expression	Life birth males	Poverty
Government progress mid term	Government progress long term	Government function	Earnings per female
Life birth female	Individual rights	Chance of survival over 40	Political pluralism
Life births	Life expectancy	Womens social issues	Women professionals
Adult literacy female	Education enrollment		

Final model

Our final model employs the identifiers from table 7 for the failed state: Male earnings, rule of law, Human Development index, civil liberties, unimproved water, GDP, adult literacy, freedom of speech, associated rights, freedom of expression, life birth of males, poverty a percentage of the population, progress-mt, government function, followed by earnings per female, life birth female, individual rights, chance of survival over 40 years of age, political pluralism, progress-LT, Lifebirth, Life expectancy, WOSOC2003, Prof_Women, AdultLit_female, and education enrollment. [148] So how do we make use of these variables? We found the most straightforward way to use the model was to categorize the variables similar to

[148] See table 6

how the raw data was categorized from its original sources and then to look at the variables within failed state examples.

Final Variable Categories

The final variable set easily fits into four categories- ***Economy*** as a function of government, ***Health*** as a function of government, **Government overall**, and ***Social and Individuals rights*** wherein the government serves as a protector and supporter of these rights.

1. **Economy**- male earnings, female earnings, GDP (gross domestic product), and poverty a percentage of the population. Also included under this heading is government economic progress midterm and long term (listed in the data as progress MT and progress LT).

We know from studies conducted that states which employ economic solvency and growth have a higher ratio of well being and happiness than their non free nation counterparts including civil liberties and life expectancy.[149] Earnings of individuals, and more importantly, poverty, are significant factors in the failed state evaluation. It has been documented that the impoverished are more likely to become terrorists and other criminals.[150] People in general are malleable, and even more so when they are financially depressed. That being said, it is also important to point out that not all states with failed or collapsed economies become failed states. Iceland is one such example. Their economy completely collapsed, the government

[149] Dowd, Alan. "Freedom and Failed States." Fraser Forum Nov./Dec. 2012:15 print.
[150] Maleckova, Jitka. *Root causes of Terrorism: Myths, Reality and Ways Forward*. Routledge. New York. 2006. Pp. 33-43. ; and Bjorgo, Tore and Jitka Maleckova Root Causes of Terrorism: Myths, Reality and Ways Forward. Routledge New York. 2005. Pp. 33.

administration resigned, new officials were elected and implemented a course of policies to regain state solvency.[151]

2. Health- HDI (health development index) - This includes world health standards with access to vaccinations and other medical intervention available within the state, unimproved water (we know that bacteria disease processes such as tuberculosis, dysentery and the plague thrive in unclean areas where toilets are located close by the water source[152]), life births (the number of life births overall), life births male (the number of life male births), and life expectancy over the age of 40. [153]

3. Government- government functioning overall and the rule of law.

A stable government is one that can protect individual rights, provide social and other institutions that instill safety as well as a respected judicial process. A respected judicial process is necessary for rule of law to be recognized. Without it, vigilante groups and individual factions gain strength weakening a society.[154]

4. Social and Individual rights-civil liberties, freedom of speech, associated rights, freedom of expression, education enrollment, adult literacy, adult literacy in females and number of women in professional job positions as well as women's social rights for 2003 (WOSOC2003).

---

[151] The breakdown of the banking system in 2008 led to a transition of government in Iceland. For many countries, chaos would ensue. This was not the case with Iceland. See Ingimundarson, Valur, "A crisis of Affluence: The Politics of an Economic Breakdown in Iceland." *Irish Studies in International Affairs*. Vol 21, vol 21. 2010 pp. 57-69. http://ria.metapress.com/content/e41405167wvr06h3/

[152] See the World Health Organization for details about drinking water standards and waste management. http://www.who.int/topics/water/en/.

[153] For a list of all original data sources see ibid 3.

[154] The groups that rise up in the wake of a failed state or fill a role that a failing government no longer offers are not included in this list such as the civil society structures that grew when Somalia became a failed state. Lemarchand, Rene. "Uncivil states and civil societies: how illusion became reality." *Journal of Modern African Studies* 30.2 (1992): 177-191.

From the final variable set in table 7, we can evaluate former failed states to determine the value of the model. For this we use three examples, the former Yugoslavia, Rwanda and Syria.

Failed State examples

Yugoslavia

To many it comes as no surprise that the former Yugoslavia failed given its warring tumultuous past. WWI began with the assassination of the Archduke Ferdinand in Sarajevo during 1914.[155] There was a long standing animosity between those wanting to regain Serbian territory lost from Austrio-Hungarian wars of the past. To this day Serbs celebrate the battle of the blackbird where they actually lost as a defining moment of their strength and ability to persevere looking forward to the day that they can again be a united greater Serbia.[156] Following WWI, WWII brought border changes and divisions with part of the region siding with German coalition in Croatia, and part of the region siding with Slavic neighbor states of the Soviet Union. The treatment of Serbs in Croatia during the Nazi era was brutal and not forgotten. The Croat Nazi's (Ustache) were tortuous to Serbs in what could be described sadistic.[157]

[155] Long standing clashes in the Balkans and an inability to come to peaceful terms between the Austro Hungarian empire and the Serbian factions led to global powers taking sides. For more information about the Balkans, David

[156] Owen gives a concise account. Owen, David. *Balkan Odyssey*. Harvest Book. Harcourt Brace & Company. New York. 1995. 7-17.

[157] The atrocities of the Ustache during the second world war are well known in the Balkans. See De Figueiredo Jr, Rui JP, and Barry R. Weingast. "Rationality of fear: political opportunism and ethnic conflict." *Military Intervention in Civil Wars* (1997).

Because of the states violent history, many researchers immediately come to the conclusion that it fell again after Marshall Tito died and that only his rule was able to maintain Yugolsav nationalism that could not be maintained by another head of state.[158]

The entrenched problems are hard to ignore, however every failed state, and every non failed state has a history of which the United States is not exempt. We know that once the Soviet Union dropped back from support of Marshall Tito and Yugoslavia that he began a series of reforms in an effort to stabilize the region. Even though the country considered itself communist, each region was afforded some autonomy. Each region was allowed to have a central bank. Trade was encouraged. Travel across borders in the regions and internationally was allowed. Freedom of the press was observed and freedom of speech to some extent. Some factory workers were even known to strike for better working conditions. Taxes went to the larger collective. Marshall Tito became known as a benevolent dictator and a liberal communist. He ruled for 34 years until his death in 1980.[159]

For the most part, it appeared that this multicultural multi regional state was content due to the reforms set about. However if we look closely at the economy in 1978, something was happening in the state. The country lost about one million emigrants.[160] Inflation was on the rise in this weak economy and industry was deteriorating in Serbian regions. Yugoslavia was importing large numbers of goods and was unable to export with the same stamina that they were

---

[158] Denitch, Bogdan Denis. *Ethnic nationalism: The tragic death of Yugoslavia*. U of Minnesota Press, 1996.pp.51-70

[159] There has been a great deal of study regarding the decentralization in an effort to create nationality regarding Tito's Yugoslavia. Some good articles regarding this are Nicholas R. Lang (1975). The Dialectics of Decentralization: Economic Reform and Regional Inequality in Yugoslavia. World Politics, 27, pp 309-335. doi:10.2307/2010123.; Furubotn, Eirik G., and Svetozar Pejovich. "Property rights, economic decentralization, and the evolution of the Yugoslav firm, 1965-1972." *JL & Econ.*16 (1973): 275.; and Dunn, W. N. "Communal federalism: Dialectics of decentralization in socialist Yugoslavia." *Publius: The Journal of Federalism* 5.2 (1975): 127-150.

[160] Denitch, Boydan "Succession and Stability in Yugoslavia". *Journal of International Affairs*. Sept. 1978.

able to consume.[161] Everyone had a job, which was consistent with the communist ideal, however, work hours shrunk due to economic challenges. The former Austrio-Hungarian regions were more advanced and modern than their agrarian counterparts and as a result were sending much of their moneys to the central coffers to those underemployed provinces which created a source of resentment.[162]

Old Croat nationalist factions rose stirring secessionist sentiments.[163] Tito squashed this in its tracks and limited the education. By 1980 oil and gas prices were out of control at a 60 % increase from 1979-1980. Staples such as butter and milk could no longer be found in the country. The gas pumps ran dry.[164]

These events set the stage for the failed state. If we refer to our final failed state model, the following variables make sense given the events presented above. These are ***GDP, Male earnings, female earnings, education enrollment, poverty, government function, rule of law, life expectancy, and survival age over 40***. In our model, we see a direct correlation to the failed state from the following identifiers in the history of Yugoslavia. A shrinking Gross domestic product was due to industrial inefficiency and a negative trade deficit.[165] Male earnings and female earnings had been on the decline for some time. Education enrollment dropped after Croatia made a challenge to break up the federation during Tito's rule. Poverty was on the rise,

---

[161] Milanovic, Branko. "Poverty in Poland, Hungary, and Yugoslavia in the years of crisis*", 1978-87*. Vol. 507. *World Bank Publications*, 1990.

[162] Two good articles discuss these situations Jovic, Dejan. "The Disintegration of Yugoslavia A Critical Review of Explanatory Approaches." *European Journal of Social Theory* 4.1 (2001): 101-120 and this book source Woodward, Susan L. *Socialist unemployment: the political economy of Yugoslavia, 1945-1990*. Princeton University Press, 1995.pp. 193-202.

[163] Malesevic, Sinisa. *Ethnicity and Federalism in Communist Yugoslavia and its successor states. Autonomy and ethnicity: negotiating competing claims in multi-ethnic states*. Cambridge University Press(2000): 147.

[164] Mesa-Lago, Carmelo. "Unemployment in a Socialist Economy: Yugoslavia."*Industrial Relations: A Journal of Economy and Society* 10.1 (1971): pp. 49-69.

[165] Ibid 164

and with all of these, government function and its inability to stimulate the economy faltered.[166] Once Tito died, these factors came to the forefront which his successors were unable to maintain. Eventually the Rule of law could no longer be observed and the county fell once again into civil war thereby dropping its life expectancy rate well below that of any stable state and the chance of survival age over 40 years of age took a steep decline. [167]

Rwanda

A thousand hills mark the landscape surrounded by sinuous, graceful rivers pooling into turquoise lakes. Pastoral lands lie in the valley offering fine grazing areas for cattle while cultivated terraced hills rise up into the sky. White blankets of pyrethrum flowers sway in the landscape. Conical peaks from the Virunga Mountains create a backdrop to the West housing what was once one half of the worlds remaining mountain gorilla population. Beset by the Democratic Republic of Congo, Burundi, Uganda, and Tanzania, this is Rwanda.

For those that have seen this country, it is as picturesque as it sounds. However, in 1994, Rwanda's population had shrunk by approximately 3,800,000. The result of a massive genocide left 800,000 dead and displaced three million to the bordering states. Within three months, the Interahamwe militia almost completed its task of eradicating what they believed to be pests, cockroaches, *Inyenzi* from the earth.[168] It has been said that the murders were so numerous, blood from these killings caused Lake Victoria to run red. [169]

---

[166] Ibid 164

[167] A good resource for the effects of civil war and life expectancy can be seen here. Mackenbach, Johan P. "Political conditions and life expectancy in Europe, 1900–2008." *Social Science & Medicine* 82 (2013): 134-146.

[168] Keane, Fergal. *Season of Blood: A Rwandan Journey*. London WC2R ORL England: Penguin Books. 1995. Pp. 10.

[169] Neuffer, Elizabeth. *The Key to My Neighbors House: Seeking Justice in Bosnia and Rwanda*. New York. Picador.2002. pp. 84-86.

Rwanda is a fascinating case study. It is one of the smallest and most densely populated countries in Africa. In this restrained area, three different ethnic groups have coexisted for hundreds of years -Twa, Hutu and Tutsi.[170] As a result of a former Belgian and German protectorate, cultural distinctions and ethnic differences were defined by features and measurements of the skull by the Belgians. Identification cards were issued. Favored positions went to the Tutsi.[171] Eventually this economic class disparity led the Hutus to overthrow the government and place Hutu preferences in power. Regardless of identification cards and transition in power, there was a point where these groups had intermarried before the identity cards declared their differences. These ethnic classes did not formerly recognize physical distinctions amongst themselves.

So what happened in Rwanda? What is for certain is this country fell into civil war and genocide in 1994. A massive machete slaughter left indelible images in the minds of those who saw bodies of people along shorelines and littered throughout the landscape.[172] Due to the sheer systematic brutality of the genocide, most researchers focus on ethnic division as the reason for the failed state and miss perhaps even more significant problems that began long before ethnic foment gave rise to genocide.

We know that through the beginning of colonialism in 1897 and up to the time of the genocide in 1994, Rwanda basically subsisted in agriculture. During the Belgian colonial

---

[170] Smith, David Norman. "The psychocultural roots of genocide: Legitimacy and crisis in Rwanda." *American Psychologist* 53.7 (1998): 743.
[171] Mamdani, Mahmood. *When victims become killers: Colonialism, nativism, and the genocide in Rwanda*. Princeton University Press, 2001.
[172] An in depth analysis of the Rwandan genocide can be found in "Leave none to tell the Story: Genocide in Rwanda" written by Alison Des Forges. Human Rights Watch. New York. 1999.

occupation, roughly about 90 % of Rwandans practiced agriculture[173], with 75% of the farm land being used to cultivate coffee.[174] Cash cropping was introduced by Belgium and land tenure to foreign plantation owners was allowed with these limited land resources for pyrethrum production (an insecticide) and coffee. The transfer of precious land to foreigners left little for food propagation for the indigenous peoples. Rwandan struggled under these foreign landowner accommodations. It did not take long for economic hardship to cause negative sentiments amongst the people.

In 1959 Hutus overthrew the Tutsi ruling class that was originally instilled by the colonizers.[175] Tutsis fled to neighboring Uganda and Tanzania. They made an attempt to retake their position in 1963 through armed militia in the North, but were pushed back. All of the land of some 400,000 displaced Tutsis was available for Hutus, moreover, the elite Hutu groups. Abandoned land transferred to Hutus easily with no laws in place for repatriation of refugee flows.[176]

Since 1980 Rwanda fell to overpopulation and unequal land distribution. Elite groups sought to control land and mineral rich areas. The elite groups were called “abaryi- meaning eaters”,[177] by rural poor, and comprised of both Hutu and Tutsi. This exploitive group took any opportunity for advancement and comfort that they could to the extreme. At one point, 95% of the population covered 43% of the cultivated land area. The population growth had surpassed food production at this point. It was even more problematic because the choice of food

[173] Bigagaza, Jean, Carolyne Abong, and Cecile Mukarubuga. "Land scarcity, distribution and conflict in Rwanda." *Scarcity and Surfeit: The Ecology of Africa’s Conflicts, ACTS Press, Nairobi/Institute for Security Studies, Pretoria*(2002): 85-156.
[174] Schoenbrun, David L. “A Past Whose Time Has Come: Historical Context and History in Eastern Africa’s Great Lakes” *History and Theory* 32:4. Dec. 1993. Pp. 34.
[175] Ibid 171
[176] Van Der Meeren, Rachel. "Three decades in exile: Rwandan refugees 1960-1990." *J. Refugee Stud.* 9 (1996): 252.
[177] Ibid 171 pp.52.

consumption was a tuber akin to the sweet potato.[178]Tubers traditionally take up more land space. Education was offered to elite groups or those that could afford it. Extreme poverty ran among the rural population. To make matters worse, an economic crisis hit Rwanda in the late 1980's and coffee prices were beginning to spiral. At the time of the genocide, Rwanda had the highest population density within its state in the entire African continent.[179]

In the 1990's, high unemployment, land tenure programs benefitting elites and a massive drought left this overpopulated country and its people hungry.[180] The elite groups used propaganda to blame poverty on the Tutsis and instilled fear in the Hutus that Tutsis were going to kill them. Fueled by the belief that they would gain land and loot from those they killed, the genocide was swift.[181] It would be second generation Tutsis displaced from 1959 that made up the Rwandan Patriotic Front who eventually stopped the genocide.[182]

From our model variables, ***Male earnings, rule of law, adult literacy, poverty, government function, earnings per female, political pluralism, education enrollment and professional women in the population, GDP, HDI, and civil liberties,*** all are variables determined in our model to be closely linked with or are precursors for the failed state and can be seen from the historical event prior to the genocide.

The earnings overall had significantly decreased in this nation, rule of law did not protect everyone, only those with privileges. Poverty was everywhere. GDP was down due to a decline

---

[178] Grosse, Scott, Robert Ford, and Jennifer Olson. "MORE PEOPLE MORE TROUBLE: Population Growth and Agricultural Change in Rwanda (A Case Study of the Population-Agriculture-Environment Nexus)." 1994 USAID http://pdf.usaid.gov/pdf_docs/pnabw925.pdf
[179] Gros, Jean-Germain. "Towards a taxonomy of failed states in the New World Order: decaying Somalia, Liberia, Rwanda and Haiti." *Third World Quarterly*17.3 (1996): 455-472.
[180] Ibid 177.
[181] Gourevitch, Philip. *We wish to inform you that tomorrow we will be killed with our families: Stories from Rwanda*. Macmillan, 1998.
[182] Ibid 175

in coffee prices. Hutu elites ruled the government, and education was for a selected group. However, even being Hutu did not grant income and security. The country was failing economically and there was not enough food. It took little to use propaganda as a tool to fuel mass killings.

When the genocide broke out, all of the other infrastructures failed as well. There was no more rule of law. Women's social issues were not respected. It was not uncommon for rape as a weapon to be used during the genocide, and chance of survival over 40 was gone.

Syria

Finally we look at the state of Syria. Syria shares borders with the following countries: Turkey, Israel, Iraq, Lebanon and Jordan. At one point Syria enjoyed trade with several partners and brought in some of its money from its oil reserves.[183] Primarily a Muslim country, it allowed education, women in high positions and even women in the judicial system.[184] It was at one time, by all accounts, a rather transitioning country regarding education, civil liberties and freedoms.

Long considered terrorist sympathizers by the United States for allowing Hezbollah (considered terrorists of Israel) to reside within its borders without impunity, sanctions were put in place. Syria became a haven for many Palestinian refugees following the 1967 Three Day War, in which Israel claimed the Golan Heights.[185] The sanctions would mark the beginning of

[183] Kienle, Eberhard. "Syria, the Kuwait War, and the New World Order." *The Gulf War and the New World Order: International Relations of the Middle East*(1994): 383-398.
[184] Cardinal, Monique C. "Women and the judiciary in Syria: appointments process, training and career paths." *International Journal of the Legal Profession* 15.1-2 (2008): 123-139.
[185] Beres, Louis Rene, and Zalman Shoval. "On Demilitarizing a Palestinian Entity and the Golan Heights: An International Law Perspective." *Vand. J. Transnat'l L.* 28 (1995): 959.

an economic strangulation for Syria.[186] Other maneuvers by the United States following these sanctions included, but were not limited to, restricting the use of U.S. dollars in Syria and the imposition of high taxes on any U.S. company doing business in Syria. This proved to be enough of financial burden to cause complete withdraw of American foreign investment in the state.[187] Although laws were passed in Syria to entice foreign investment and protect property rights for those investments, the laws were poorly written leaving them, at times, open for ambiguous interpretation.[188] An inability to use the dollar for international business in Syria coupled with corrupt, public officiators, lead prospective investors to mistrust that Syria could provide protection of their future potential investment. Most foreign investors had pulled out of the country leaving an economic void.[189]

Even though economic prospects looked bleak, the regime continued to look for progressive ways in which to open up trade and receive financial assistance from other countries.[190] Strategically an opportunity presented itself in the first Gulf war. By siding itself with Saudi Arabia and Kuwait, Syria was able to trade that support in exchange for international aid. However, these funds and the opening of the oil trade did not do enough to entice its leaders to use the funding to create programs for its people. The government's main focus was on foreign investments and foreign business. [191] Now the country is lost in the chaos of civil war, with hundreds of thousands of its people killed, or displaced and starving.

---

[186] Sen, Kasturi, Waleed Al-Faisal, and Yaser AlSaleh. "Syria: effects of conflict and sanctions on public health." *Journal of Public Health* 35.2 (2013): 195-199.
[187] Perthes, Volker. "The Syrian economy in the 1980s." *The Middle East Journal*(1992): 37-58.
[188] Perthes, Volker. "The private sector, economic liberalization, and the prospects of democratization: The case of Syria and some other Arab countries."*Democracy without democrats* (1994): 243-269.
[189] Ibid 182
[190] Ibid 184
[191] Ibid 179

From our final model, the following variables seem to fit within the framework of events in Syria. These are as follows***; male earnings, rule of law, human development index, civil liberties, unimproved water, GDP, freedom of speech, associated rights, freedom of expression, life birth males and life birth females, poverty, government function, chance of survival over 40, life expectancy, professional women, womens social issues, and education enrollment.***

After foreign trade left Syria, unemployment rose. Women usually are the first groups of people to lose their jobs. According to our model, the following factors reside in Syria and were prevalent before the civil war. Male earnings were down, female earnings declining, GDP declining, government function was declining, and the rule of law (protection of property rights) was not being properly observed.[192] All of the other variables mentioned as indicators are obviously present due to the failed state. In civil war situations, life expectancy, life births, and babies surviving are diminished due to food insecurity and collapse of health care institutions. Muslim clerics have issued a fatwa granting people to eats dogs and cats because they are starving.[193] The Assad regime seems to be expulsing any persons who exercise and speak in defiance of the regime, so the associated rights, freedom of expression, and freedom of speech are no longer observed in this country. When civilian areas are bombed, so are the infrastructures including power and water sources. It does not take much to contaminate a well from leaking gas pipes, or irrigation systems.

[192] Moore, Pete W., and Bassel F. Salloukh. "Struggles under authoritarianism: Regimes, states, and professional associations in the Arab world." *International Journal of Middle East Studies* 39.01 (2007): 53-76.
[193] Hunter, Stuart. "Clerics in Syria Issue Fatwa Allowing Citizens to Eat Dogs, Cats toPprevent Starvation". The World Post. Oct 16, 2013. http://www.huffingtonpost.com/2013/10/15/syria-eating-dogs-cats-cleric-fatwa_n_4101821.html

The civil war and the World Health Organization announced a polio outbreak in Syria and in the refugee flow areas of Turkey and Iraq, which can only mean that its people were not vaccinated.[194] If it is only small children that have not been vaccinated, we can assume that the health development index was on the decline at least a couple of years ago to not have enough vaccine on hand. With a decline of the HDI and an outbreak of polio aside from the civil war, we can note that life expectancy has been declining in the state for some time. We know that a lack of access to proper health care diminishes lifespan. If Syria had been more transparent, we may have additional indicators consistent with our model prior to the opposition clashes and the outbreak of civil war in 2011.

Surprising findings

There are two rather surprising indicators that pop up in our final model. The first being women's rights and women's issues. Their prevalence in the model is unmistakable.[195]Adult literacy in females, the number of professional women, women's social rights overall and earnings per female all have a strong correlation to the failed state.

Originally it was my belief that only ethnic protections and equality between ethnicities were prudent for harmony in a state. Furthermore, although gender issues were important, they did not weigh with as much significance. After reviewing the results, I discovered that I was wrong about this. The United Nations has mandates for protection of women's rights. Women that are discriminated against, typically receive less health care attention, less education and

[194] More Polio Cases is Syria to be Confirmed this Week-WHO Reuters http://www.reuters.com/article/2013/11/25/syria-crisis-polio-idUSL5N0JA2X720131125
[195] See table

lower wages than men.[196] FGM (Female genital mutilation) is a cruel practice against girls in some African states and is a practice still continued. The importance to marry girls off in societies where they are not allowed to work outside of the home became the impetus for this brutal practice. Men could be very selective about their brides with purity ranking as a must. To ensure this, families practice FGM to show a girl's worthiness to marry, or a "proof of purity"[197]. Although this practice has been condemned, it still continues.[198] It is the individual rights that are not observed for women in the collective that lead to more harsh treatment and a harsher view of women. FGM is as abhorrent as male castration, which is a practice used to tame cattle bulls. After evaluations, it is apparent how these rights are so profoundly linked to the failed state. We only need to look to Afghanistan and Taliban rules imposed regarding women to see exactly how damaging it can be on a society.

The second surprise was regarding the civil liberties indicator in the model. The model showed that as civil liberties rise past a certain point in any given state, there was a positive relationship. In other words, as civil liberties increased past a particular threshold, so did the risk for a failed state. Malcolm Gladwell gave an example in *David and Goliath* that may shed some light on this. He suggests after some research, that there is a plateau regarding education, in which a person's value for it decreases~~-~~education is taken for granted, and a false sense of entitlement begins.[199] This entitlement problem comes not from the generation of people who

[196] See Ibid 3 www.undata.org. In 1980 the United Nations adopted a charter to protect women called CEDAW or Convenant on Women's Rights. Drinan, Robert Jr. *The Mobilization of Shame: A worldview of Human Rights.* Yale Univeristy Press. 2001. Pp. 35-44.
[197]Fauziya Kassindja entered the United State requesting asylum on December 17, 1994 for fear or subjugation of female circumcision. Her accounts of this practice the health problems as well as the economic burdens if a girl cannot marry in Togo because of an absence of this are well documented in her journey
Kassindja, Fauziya and Layli Miller. *Do they hear you when you Cry?* Random House. New York. 1998. Pp. 168.
[198] Ibid 188.
[199] Galdwell, Malcolm. *David and Goliath: Underdogs, Misfits, and the Art of Battling Giants.* Little Brown and Company. New York. 2013. 39-62. Print.

had to scrape to save so that their child could go to college, or those that had to work and do well in school for a scholarship. It seems to be the following generation that is the nexus for a false sense of entitlement. Expectations have changed for kids receiving more. We can look at the recent case of a girl that had moved out of her home at age 17 who was suing her parents for financial support for current and future academics.[200] Perhaps this is an example of why when civil liberties go up past a certain point the chance of a failed state does too. The entitlement expectations and proclivity against working for basics, sets people up for a fall if economies can no longer provide what those people understand as a given.

Another possible example of this is language interpreters. The United States encourages language individuality. Although English is the common language, electric companies offer two options, English and Spanish, in their automated system. Judicial and medical services must provide interpreters at the cost of the physician/medical community. In San Francisco, a huge melting pot area for immigration and tourism, there could be as many as 20 different languages- Spanish, Swahili, Persian, Arabic, Russian, Serbian, Portuguese, Italian, Chinese, Mandarin, Japanese, Korean, Vietnamese, Greek, French, and so on, although some research states that number to be even higher.[201] Written driver exams in California are currently being offered in the following languages in addition to English -- Amharic, Arabic, Armenian, Cambodian, Chinese, Croatian, French, German, Greek, Hebrew, Hindi, Hmong, Hungarian, Indonesian, Italian, Japanese, Korean, Laotian, Persian/Farsi, Polish, Portuguese, Punjabi, Romanian,

[200]Ly, Laura. "Student's Lawsuit Against Parnets for support loses first round in court". CNN Mar. 5, 2014. http://www.cnn.com/2014/03/04/justice/student-sues-parents-new-jersey/

[201]In this article from the Bay Area of San Francisco the figure touted for diverse languages was recorded at 112. Hendricks, Tyche."Bay Area/ Report: 112languages Spoken in Diverse Region". SF Gate http://www.sfgate.com/bayarea/article/BAY-AREA-Report-112-languages-spoken-in-2692403.php

Russian, Samoan, Spanish, Tagalog/Filipino, Thai, Tongan, Turkish, and Vietnamese. [202] When interpreters and language offerings go beyond, it could be considered an example of excessive civil liberties. Nationalism is no longer a goal or is even encouraged. Whether or not it would lead to a failed state is unknown.

Conclusion

Is it possible to identify the predictive characteristics for failed states using statistics with previously mined data? Furthermore, can the failed state be predicted?

In answer to the first question, we were able to construct a model that pinpoints indicators for the failed state. These markers appeared consistently and were prevalent in failed states from Random Forest and Boruta programs. Regression tree partitioning programs are being used in more fields of study than just the hard sciences. Through the process employed in the methodology, this paper has given a valid argument for the use of statistics and quantitative analysis in political science. Because of the nature of the statistical program and the methods employed, we were able to locate predictor variables/identifiers for the failed state by this model creation. However, this method would not have been possible without mined data sets.

There was a great deal of unavailable information regarding some countries due to lack of transparency or a short existence that would have given even more validity to the variables identified by their inclusion in the construct. Models created can only be as precise as the data available. It is possible that some variables would have had more importance if more countries could have been included. As it is, the model created was determined using no fewer than 200 countries, which was enough to provide sufficient data for our exercise. By assembling a model,

[202] http://www.dmv.ca.gov/dl/dl_info.htm#languages

the question of the failed state allows us to look at it through a different lens in the hopes of illustrating points we otherwise would not locate by qualitative means alone.

The variables identified for the model are Male earnings, rule of law, Human Development index, civil liberties, unimproved water, GDP, adult literacy, freedom of speech, associated rights, freedom of expression, life birth of males, poverty a percentage of the population, progress midterm, government function, followed by earnings per female, life birth female, individual rights, chance of survival over 40 years of age, political pluralism, progress long term, live births, life expectancy, WOSOC2003 (women's social issues for 2003) , women in professions, female adult literacy, and education enrollment.

In addition to using statistics to help answer our question, we also looked at three states qualitatively to help validate the variables chosen. Those states are the former Yugoslavia, Rwanda and Syria. Each of these states fit our definition of the failed state. There was enough history documented on all three to make reasonable qualitative illustrations for our variable model set and each state is distinctly different in its failure. Yugoslavia fell apart after the death of a long standing leader, was a former communist country, and is currently divided up into seven separate viable countries that are growing economically, (Serbia, Macedonia, Kosovo, Montenegro, Croatia, and Slovenia) with the exception of one that is under a United Nations protectorate–Bosnia-Herzegovina. Rwanda was a former colonial state that was ripped apart from a genocide that as well is growing economically. Syria is in the throes of civil war whose end is yet to be played out. All three are different and all had good documentation of past events and that allowed us to apply our variables for qualitative case analysis.

By using statistics as employed in the methodology, we hope to provide credit to the field of political science that is traditionally considered a soft science by enhancing the value of what we know through past research and opening the door to other methods of thinking

So, Can the failed state be predicted? Even after all of the research and deliberation taken, there are so many factors when dealing with failed states that there is no one size fits all model for predicting the failed state. The closest that I think we can come is to identify characteristics that show consistently to have a relationship with failed states. Along that vein, just because a state has several indicators present that are consistent with failed states, it is not necessarily a harbinger for a failed state.

States and whether or not they fail depends on so many factors. Society may be very tolerant, or conditioned to fill in the gaps where the government is lacking making it less likely to fail. In other situations it could take little to collapse the state. There are so many unknowns that it is very difficult to make a judgment about which states will fail and which will not.

My original belief was that states, which were propped up financially by wealthy nations, regardless of how unstable they were, would not fail due to external stability and support. However, Egypt proved this initial belief of mine to be wrong. Although this model is useful in identifying predictor variables, failed states themselves are too difficult to predict. There is no panacea for the failed state and there is no one predictor element for it. Although this model is very useful in identifying predictor variables, failed states themselves are too difficult to conclusively predict.

Bibliography

Beres, Louis Rene, and Zalman Shoval. "On Demilitarizing a Palestinian Entity and the Golan Heights: An International Law Perspective." *Vand. J. Transnat'l L.* 28 (1995): 959.

"Big Data: The Next Frontier for Innovation, Competition, and Productivity", McKinsey Global Institute, Technology & Innovation, McKinsey & Company. http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation.

Bigagaza, Jean, Carolyne Abong and Cecile Mukarubuga. "Land Scarcity, Distribution and Conflict in Africa. " www.issafrica.org/uploads/scarcitychapter2.pdf. 52-54. web accessed 10/31/13.

Bigagaza, Jean, Carolyne Abong, and Cecile Mukarubuga. "Land scarcity, distribution and conflict in Rwanda." *Scarcity and Surfeit: The Ecology of Africa's Conflicts, ACTS Press, Nairobi/Institute for Security Studies, Pretoria*(2002): 85-156.http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation [accessed 10 July 2012] Web

Bjorgo, Tore ed. *Root Causes of Terrorism: Myths, Reality and Ways Forward*. 2005. Routledge. New York. 2006. 33-43. Print.

Breiman, Leo, et al. *Classification and regression trees*. CRC press, 1984.

Breiman, Leo and Adele Cutler. Random Forest http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm

Brooks, R. E. "Failed States, or the State as Failure?" *The University of Chicago Law Review*, 72.4 (2005): 1159–1196. Print.

Brown, Brad, Michael Chui, and James Manyika. "Are you ready for the era of 'big data'." *McKinsey Quarterly* 4 (2011): 24-35.

Call, C. T. (2008). "The Fallacy of the Failed State." *Third World Quarterly*, 29.8 (2008): 1491–1507. Print.

Cardinal, Monique C. "Women and the judiciary in Syria: appointments process, training and career paths." *International Journal of the Legal Profession* 15.1-2 (2008): 123-139.

Carter, Hannah, et al. "Cancer-specific high-throughput annotation of somatic mutations: computational prediction of driver missense mutations." *Cancer research* 69.16 (2009): 6660-6667. http://cancerres.aacrjournals.org/content/69/16/6660.full web accessed.

Cingranelli, David L. and David Richards. *The Cingranelli-Richards Human Rights Dataset (CIRI)* www.humanrightsdata.org. web accessed 2013. 9.12.13.

'CIFP - Country Indicators for Foreign Policy' <http://www4.carleton.ca/cifp/> [accessed 12 December 2012] Web.

Cutler, Richard, Thoams Edwards Jr, Karen Beard et all. "Random Forests For Classification in Ecology. *Ecology*, 88(11) 2007. 2783-2792.

Dalgaard, Peter. *Introductory Statistics with R*. 2008. Springer. New York. 2nd edition. Print.

De Figueiredo Jr, Rui JP, and Barry R. Weingast. "Rationality of fear: political opportunism and ethnic conflict." *Military Intervention in Civil Wars* (1997).

De' ath, Glenn and Katharina E. Fabricius. "Classification and Regression Trees: A Powerful Yet Simple Technique for Ecological Data Analysis" *Ecology*, 81 (11). 2000. 3178-3192.

Denitch, Bogdan Denis. *Ethnic nationalism: The tragic death of Yugoslavia*. U of Minnesota Press, 1996.pp.51-70

DMV California. http://www.dmv.ca.gov/dl/dl_info.htm#languages

Des Forges, Alison. Leave None to Tell the Story: Genocide in Rwanda. Human Rights Watch. New York. 1999. Print.

Diaz-Uriarte, Ramon and Sara Alvarez de Andres. "Gene Selection and Classification of Microarray Data Using random Forest". BMC Bioinformatics. 6 January, 2006. http://www.biomedcentral.com/1471-2105/7/3. web accessed.

Dorff, Robert H., "Democratization and Failed States: The Challenge of Ungovernability", *Parameters, US Army War College Quarterly - Summer 1996*, (1996): 17–31. Print.

Dowd, Alan W. "Freedom and Failed States". *Fraser Forum*. Nov/Dec(2012): 15-19. www.fraserinstitute.org. Web

Dunn, W. N. "Communal federalism: Dialectics of decentralization in socialist Yugoslavia." *Publius: The Journal of Federalism* 5.2 (1975): 127-150.

Efron, Bradley, and Robert J. Tibshirani. *An introduction to the bootstrap*. Vol. 57. CRC press, 1994.Print.pp.45-57.

Elsmén, Emma, Mårten Steen, and Lena Hellström-Westas. "Sex and gender differences in newborn infants: why are boys at increased risk?." *Journal of Men's Health and Gender* 1.4 (2004): 303-311.

"The Failed States Index. Carnegie Endowment for International Peace. " Foreign Policy. No. 149. Jul.Aug 2005. 56-65. Print.

Fang, Janet. "Skip the TSA line and give yourself a security check" www.smartplanet.com/blog/bulletin/skip-the-tsa-line-and-give-yourself-a-security-check/32826. Web accessed 10/24/13.

Frank E Harrell Jr and with contributions from many other users. (2010). Hmisc: Harrell Miscellaneous. R package version 3.8-3. http://CRAN.R-project.org/package=Hmisc Web.

Fund For Peace, *Failed States Index.* www.fundforpeace. org. web. accessed 2008.

Furubotn, Eirik G., and Svetozar Pejovich. "Property rights, economic decentralization, and the evolution of the Yugoslav firm, 1965-1972." *JL & Econ.*16 (1973): 275.

Gellately, Robert ed. and Ben Kiernan ed. The Specter of Genocide: Mass Murder in Historical Perspective. Cambridge University Press. New York. 2003. 325-338. Print.

Gladwell, Malcolm. *David and Goliath: Underdogs, Misfits, and the Art of Battling Giants.* Little Brown and Company. New York. 2013. 39-62. Print.

Ghani, Ashraf and Clare Lockhart. *Fixing Failed States: A Framework for Rebuilding a Fractured World.* Oxford University Press. New York. 2008.

'Google Earth' <http://www.google.com/earth/index.html> [accessed 20 July 2012] Web.

Gourevitch, Philip. *We wish to inform you that tomorrow we will be killed with our families: Stories from Rwanda*. Macmillan, 1998.

Gow, James. International Diplomacy and the Yugoslav War: Triumph of the Lack of Will. Columbia University Press. New York. 1997. Print.

Grömping, Ulrike. "Variable importance assessment in regression: linear regression versus random forest." *The American Statistician* 63.4 (2009)

Gros, Jean-Germain. "Towards a taxonomy of failed states in the New World Order: decaying Somalia, Liberia, Rwanda and Haiti." *Third World Quarterly*17.3 (1996): 455-472.

Grosse, Scott, Robert Ford, and Jennifer Olson. "MORE PEOPLE MORE TROUBLE: Population Growth and Agricultural Change in Rwanda (A Case Study of the Population-Agriculture-Environment Nexus)." (1994). USAID PDF http://pdf.usaid.gov/pdf_docs/pnabw925.pdf

Hardin, Garret. "The Tragedy of the Commons". *Science*. 162:1243-48. 1968

Helman, Gerald B and Steven R. Ratner, “Saving failed States”. *Foreign Policy*. No.89 (Winter, 1992- 1993. Pp. 3-20

Hendricks, Tyche.”Bay Area/ Report: 112languages Spoken in Diverse Region”. SF Gate http://www.sfgate.com/bayarea/article/BAY-AREA-Report-112-languages-spoken-in-2692403.php

Hunter, Stuart. “Clerics in Syria Issue Fatwa Allowing Citizens to Eat Dogs, Cats toPprevent Starvation”. The World Post. Oct 16, 2013. http://www.huffingtonpost.com/2013/10/15/syria-eating-dogs-cats-cleric-fatwa_n_4101821.html

Huntington, Samuel P. *Political Order in Changing Societies.* Yale University. (1996). 216-218. Print.

Hursh, John. "Advancing Women's Rights Through Islamic Law: The Example of Morocco." *Berkeley. Journal of Gender, Law & Justice* 27.2 (2012): 252-305. Business Source Complete. Web.23 Oct. 2013.Web

‘Index of State Weakness in the Developing World’, *The Brookings Institution* http://www.brookings.edu/research/reports/2008/02/weak-states-index [accessed 1 August 2012] Web.

Ingimundarson, Valur, “A crisis of Affluence: The Politics of an Economic Breakdown in Iceland.” *Irish Studies in International Affairs*. Vol 21, vol 21. 2010 pp. 57-69. http://ria.metapress.com/content/e41405167wvr06h3/

Jakulin, Aleks, "A Rant on the Virtues of Data Mining: Statistical Modeling, Causal Inference, and Social Science", *Statistical Modeling, causal Inference, and social Science*, 2012 http://andrewgelman.com/2007/08/a_rant_on_the_v/ [accessed 11 July 2012] Web.

Jovic, Dejan. "The Disintegration of Yugoslavia A Critical Review of Explanatory Approaches." *European Journal of Social Theory* 4.1 (2001): 101-120.

Kassindja, Fauziya and Layli Miller. *Do they Hear You When You Cry?* Random House, inc. New York. 1998.

Kawaller, I.G., Koch, P.D. and Koch, T. W. (1987), "The Temporal Price Relationship between S&P 500 Futures and the S&P 500 Index". *The Journal of Finance*, 42: 1309–1329.

Keane, Fergal. *Season of Blood: A Rwandan Journey*. London WC2R ORL England: Penguin Books. 1995. Pp. 10.

Kienle, Eberhard. "Syria, the Kuwait War, and the New World Order." *The Gulf War and the New World Order: International Relations of the Middle East*(1994): 383-398.

Kim, Hyung L., et al. "Using protein expressions to predict survival in clear cell renal carcinoma." *Clinical cancer research* 10.16 (2004): 5464-5471. http://clincancerres.aacrjournals.org/content/10/16/5464.Web.

Kursa, Miron and Withold Rudnicki. " Feature Selection with the Boruta Package" . Journal of Statistical Software. Sept 2010 Vol 36 issue 11. www.jstatsoft.org. Print.

King, Gary, "Ensuring the Data-Rich Future of the Social Sciences", *Science*, 331 (2011), 719–721 doi:10.1126/science.1197872 Web.

Laitin, David, and Daniel Posner, 'The Implications of Contructivism for Contructing Ethnic Fractionalization Indices', *The Comparative Politics Newsletter*, 12 http://todayinsci.com/QuotationsCategories/I_Cat/Information-Quotations.htm [accessed 10 July 2012] Web.

Lang, Nicholas R. "The Dialectics of Decentralization: Economic Reform and Regional Inequality in Yugoslavia". *World Politics*, (1975). 27, pp 309-335. doi:10.2307/2010123.

Lemarchand, Rene. "Uncivil states and civil societies: how illusion became reality." *Journal of Modern African Studies* 30.2 (1992): 177-191.

Liaw, A. and M. Wiener, "Classification and Regression by randomForest". *R News* 2.3 (2002) : 18-22. Print.

Loh, Wei-Yin. "Classification and regression trees." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1.1 (2011): 14-23.

Ly, Laura. "Student's Lawsuit Against Parnets for support loses first round in court". CNN Mar. 5, 2014. http://www.cnn.com/2014/03/04/justice/student-sues-parents-new-jersey/

Ma, Yan, et al. "Predicting cancer drug response by proteomic profiling."*Clinical cancer research* 12.15 (2006): 4583-4589. https://clincancerres.aacrjournals.org/content/12/15/4583.full web.

Mackenbach, Johan P. "Political conditions and life expectancy in Europe, 1900–2008." *Social Science & Medicine* 82 (2013): 134-146.

Makeig, S.; Kothe, C.; Mullen, T.; Bigdely-Shamlo, N.; Zhilin Zhang; Kreutz-Delgado, K., "Evolving Signal Processing for Brain–Computer Interfaces," *Proceedings of the IEEE* , vol.100, no.Special Centennial Issue, pp.1567,1584, May 13 2012. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6169943&isnumber=6259910 web.

Malesevic, Sinisa. *Ethnicity and Federalism in Communist Yugoslavia and its successor states. Autonomy and ethnicity: negotiating competing claims in multi-ethnic states*. Cambridge University Press. (2000): 147.

Mallaby, Sebastian, *The World's Banker: A Story of Failed States, Financial Crises, and the Wealth and Poverty of Nations*. Penguin Books.New York. 2006. Print/

Mamdani, Mahmood. *When victims become killers: Colonialism, nativism, and the genocide in Rwanda*. Princeton University Press, 2001.

Marshall, Monty and Benjamin Cole., "Global Report on Conflict, Governance and State Fragility" *Foreign Policy Bulletin*, 2008: 1–21. Print.

Merkle Edgar C. and Victoria A. Shaffer. " Binary Recursive Partitioning: Background, Methods, and Application to Psychology" *British Journal of Mathematical and Statistical Psychology*. 2011, 64, 161- 181 Print.

Mesa-Lago, Carmelo. "Unemployment in a Socialist Economy: Yugoslavia."*Industrial Relations: A Journal of Economy and Society* 10.1 (1971): pp. 49-69.

Milanovic, Branko. *Poverty in Poland, Hungary, and Yugoslavia in the years of crisis, 1978-87*. Vol. 507. World Bank Publications, 1990.

'Mo Ibrahim Foundation | Methodology', *Mo Ihabrim Foundation Index* http://www.moibrahimfoundation.org/IIAG-methodology/ [accessed 12 December 2012]

Miron B. Kursa, Witold R. Rudnicki. "Feature Selection with the Boruta Package." *Journal of Statistical Software* 36.11 (2010) : 1-13. URL http://www.jstatsoft.org/v36/i11/. Web.

Moore, Pete W., and Bassel F. Salloukh. "Struggles under authoritarianism: Regimes, states, and professional associations in the Arab world." *International Journal of Middle East Studies* 39.01 (2007): 53-76.

More Polio Cases is Syria to be Confirmed this Week-WHO Reuters http://www.reuters.com/article/2013/11/25/syria-crisis-polio-idUSL5N0JA2X720131125

Naisbitt, John, *Megatrends 2000* (Avon, 1991) Web.

Neuffer, Elizabeth. *The Key to My Neighbors House: Seeking Justice in Bosnia and Rwanda*. New York. Picador.2002. pp. 84-86.

Owen, David. *Balkan Odyssey*. Harvest Book Harcourt Brace& Company. New York. 1995. Print.

Perthes, Volker. "The private sector, economic liberalization, and the prospects of democratization: The case of Syria and some other Arab countries."*Democracy without democrats* (1994): 243-269.

Perthes, Volker. "The Syrian economy in the 1980s." *The Middle East Journal*(1992): 37-58.

Radelet, Steven. "Bush and Foreign Aid". *Foreign Affairs.* Council on Foreign Relations. Vol. 82, No.5 (Sept-Oct, 2003) pp. 104-117.

Random Forest http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm

'Republic of South Sudan Declares Independence' http://www.betterworldcampaign.org/un-peacekeeping/web-features/south-sudan-declares-independence.html [accessed 12 December 2012] Web.

Rice, Susan. and Stewart Patrick. "Index of State Weakness in the Developing World". *Foreign Policy at Brookings.* The Brookings Institution, D.C. 2008. Print.

Rotberg, Robert I. "Failed states, collapsed states, weak states: Causes and indicators." *State failure and state weakness in a time of terror* (2003): 1-25.a PDF from the Brookings Institute http://www.brookings.edu

Rotberg, Robert I., "Failed States in a World of Terror", *Foreign Affairs*, 81 (2002), 127. Print.

Rotberg, Robert I. *When States Fail: Causes and Consequences*. Princeton University Press. New Jersey. 2004 pp.7.

Schultz, William F. *In Our Own Best Interest: How Defending Human Rights Benefits Us All* .Beacon Press. Boston. 2002.

Schoenbrun, David L. "A Past Whose Time Has Come: Historical Context and History in Eastern Africa's Great Lakes" *History and Theory* 32:4. Dec. 1993. Pp. 32- 56.

Sen, Kasturi, Waleed Al-Faisal, and Yaser AlSaleh. "Syria: effects of conflict and sanctions on public health." *Journal of Public Health* 35.2 (2013): 195-199.

Smith, David Norman. "The psychocultural roots of genocide: Legitimacy and crisis in Rwanda." *American Psychologist* 53.7 (1998): 743.

Spector, Phil. *Data Manipulation with R.* Springer. New York. 2008: 2-10,27,82. Print.

Svetnik, Vladimir, et al. "Random forest: a classification and regression tool for compound classification and QSAR modeling." *Journal of chemical information and computer sciences* 43.6 (2003): 1947-1958.

Takeuchi, Shinichi. "Conflict and Land tenure in Rwanda". Policy Dialogue. www.policydialogue.org. accessed 10/30/13. Web.

"The Continuing Challenge of Refugee Return in Bosnia Herzegovina". International Crisis Group Balkans Report 137. Dec 13 2002.

Thelwall, Mike. "The responsiveness of search engine indexes." *Cybermetrics*5.1 (2001): 8. http://cybermetrics.cindoc.csic.es/articles/v5i1p1.pdf. web

Torsten, Hothorn, Kurt Hornik and Achim Zeilieis. "Unbiased Recursive Partitioning: A Conditional Inference Framework". Jornal of Computational and Graphical Statistics. Vol. 15, Iss.3.,2006. http://www.tandfonline.com/doi/abs/10.1198/106186006X133933 web.

The United Nations data source. http://data.un.org

Van Der Meeren, Rachel. "Three decades in exile: Rwandan refugees 1960-1990." *J. Refugee Stud.* 9 (1996): 252.

Weeramantry, C.G. , *Islamic Jurisprudence: An International Prespective*. The Other Press. Kuala Lampur. 2001. 68-69. Print.

Williamson, Richard S. " Nation-Building: The Dangers of Weak, Failing, and Failed States". *The Whitehead Journal of Diplomacy and International Relations.* Winter/Spring (2007). 9-19. www.journalofdiplomacy.org Web.

Woodward, Susan L. *Socialist unemployment: the political economy of Yugoslavia, 1945-1990*. Princeton University Press, 1995.pp. 193-202.

World Health Organization. http://www.who.int/topics/water/en/

Zuur, Alain, Elena N. Ieno, & Erik H.W.G. Meesters. *A Beginner's Guide to R.* Springer. New York. 2009. Print.

http://www.smartplanet.com/blog/bulletin/skip-the-tsa-line-and-give-yourself-a-security-check/32826

Appendix 1

Country name reconciliation

Trend_Country	Rank_Country
	Afghanistan
Albania	Albania
Algeria	Algeria
Angola	Angola
Antigua and Barbuda	
Argentina	Argentina
Armenia	Armenia
Australia	Australia
Austria	Austria
Azerbaijan	Azerbaijan
Bahamas	Bahamas, The
Bahrain	Bahrain
Bangladesh	Bangladesh
Barbados	Barbados
Belarus	Belarus
Belgium	Belgium
Belize	Belize
Benin	Benin
Bhutan	Bhutan
Bolivia	Bolivia
Bosnia and Herzegovina	Bosnia and Herzegovina
Botswana	Botswana
Brazil	Brazil
Brunei Darussalam	
Bulgaria	Bulgaria
Burkina Faso	Burkina Faso
	Burma
Burundi	Burundi
Cambodia	Cambodia
Cameroon	Cameroon
Canada	Canada
Cape Verde	Cape Verde
Central African Republic	Central African Republic
Chad	Chad
Chile	Chile

China	China
Colombia	Colombia
Comoros	Comoros
Congo	Congo, Republic of
Congo (Democratic Republic of th	Congo, Democratic Republic of
Costa Rica	Costa Rica
Croatia	Croatia
Cuba	Cuba
Cyprus	Cyprus
Czech Republic	Czech Republic
Côte d'Ivoire	Côte d'Ivoire
Denmark	Denmark
Djibouti	Djibouti
Dominica	Dominica
Dominican Republic	Dominican Republic
Ecuador	Ecuador
Egypt	Egypt
El Salvador	El Salvador
Equatorial Guinea	Equatorial Guinea
Eritrea	Eritrea
Estonia	Estonia
Ethiopia	Ethiopia
Fiji	Fiji
Finland	Finland
France	France
Gabon	Gabon
Gambia	Gambia, The
Georgia	Georgia
Germany	Germany
Ghana	Ghana
Greece	Greece
Grenada	
Guatemala	Guatemala
Guinea	Guinea
Guinea-Bissau	Guinea-Bissau
Guyana	Guyana
Haiti	Haiti
Honduras	Honduras
Hong Kong, China (SAR)	Hong Kong
Hungary	Hungary
Iceland	Iceland
India	India

Indonesia	Indonesia
Iran (Islamic Republic of)	Iran
	Iraq
Ireland	Ireland
Israel	Israel
Italy	Italy
Jamaica	Jamaica
Japan	Japan
Jordan	Jordan
Kazakhstan	Kazakhstan
Kenya	Kenya
	Kiribati
	Korea, North
Korea (Republic of)	Korea, South
Kuwait	Kuwait
Kyrgyzstan	Kyrgyz Republic
Lao People's Democratic Republic	Laos
Latvia	Latvia
Lebanon	Lebanon
Lesotho	Lesotho
Liberia	Liberia
Libyan Arab Jamahiriya	Libya
	Liechtenstein
Lithuania	Lithuania
Luxembourg	Luxembourg
	Macau
Macedonia (TFYR)	Macedonia
Madagascar	Madagascar
Malawi	Malawi
Malaysia	Malaysia
Maldives	Maldives
Mali	Mali
Malta	Malta
Mauritania	Mauritania
Mauritius	Mauritius
Mexico	Mexico
	Micronesia, Fed St.
Moldova	Moldova
Mongolia	Mongolia
Montenegro	Montenegro, Republic of
Morocco	Morocco
Mozambique	Mozambique

Myanmar	
Namibia	Namibia
Nepal	Nepal
Netherlands	Netherlands
New Zealand	New Zealand
Nicaragua	Nicaragua
Niger	Niger
Nigeria	Nigeria
Norway	Norway
Occupied Palestinian Territories	
Oman	Oman
Pakistan	Pakistan
Panama	Panama
Papua New Guinea	Papua New Guinea
Paraguay	Paraguay
Peru	Peru
Philippines	Philippines
Poland	Poland
Portugal	Portugal
Qatar	Qatar
Romania	Romania
Russian Federation	Russia
Rwanda	Rwanda
Saint Kitts and Nevis	
Saint Lucia	Saint Lucia
Saint Vincent and the Grenadines	Saint Vincent and the Grenadines
Samoa	Samoa
Sao Tome and Principe	São Tomé and Príncipe
Saudi Arabia	Saudi Arabia
Senegal	Senegal
Serbia	Serbia
Seychelles	Seychelles
Sierra Leone	Sierra Leone
Singapore	Singapore
Slovakia	Slovak Republic
Slovenia	Slovenia
Solomon Islands	Solomon Islands
South Africa	South Africa
Spain	Spain
Sri Lanka	Sri Lanka
Sudan	Sudan
Suriname	Suriname

Swaziland	Swaziland
Sweden	Sweden
Switzerland	Switzerland
Syrian Arab Republic	Syrian Arab Republic
	Taiwan
Tajikistan	Tajikistan
Tanzania (United Republic of)	Tanzania
Thailand	Thailand
Timor-Leste	Timor-Leste
Togo	Togo
Tonga	Tonga
Trinidad and Tobago	Trinidad and Tobago
Tunisia	Tunisia
Turkey	Turkey
Turkmenistan	Turkmenistan
Uganda	Uganda
Ukraine	Ukraine
United Arab Emirates	United Arab Emirates
United Kingdom	United Kingdom
United States	United States
Uruguay	Uruguay
Uzbekistan	Uzbekistan
Vanuatu	Vanuatu
Venezuela (Bolivarian Republic o	Venezuela
Viet Nam	Vietnam
Yemen	Yemen
Zambia	Zambia
	Zimbabwe

Appendix 2

Original failed state ranking

Country	Failing Rank
Afghanistan	4.5
Albania	2
Algeria	3
Angola	3
Argentina	3
Armenia	3
Australia	1
Austria	1
Azerbaijan	3
Bahamas	1
Bahrain	1
Bangladesh	3
Barbados	1
Belarus	2
Belgium	1
Belize	1
Benin	3
Bhutan	3
Bolivia	3
Bosnia and Herzegovina	4
Botswana	2
Brazil	2
Bulgaria	2
Burkina Faso	3
Burma	2
Burundi	3
Cambodia	3
Cameroon	3
Canada	1
Cape Verde	2
Central African Republic	4
Chad	4
Chile	2
China	3
Colombia	4
Comoros	3
Congo (Democratic Republic of th	4
Congo	3
Costa Rica	1
Côte d'Ivoire	4
Croatia	2
Cuba	3
Cyprus	4

Czech Republic	1
Denmark	1
Djibouti	3
Dominica	2
Dominican Republic	3
Ecuador	2
Egypt	2
El Salvador	2
Equatorial Guinea	4
Eritrea	4
Estonia	1
Ethiopia	3
Fiji	4
Finland	1
France	2
Gabon	3
Gambia	3
Georgia	3
Germany	2
Ghana	2
Greece	1
Guatemala	3
Guinea	4.5
Guinea-Bissau	3
Guyana	3
Haiti	3
Honduras	2
Hong Kong, China (SAR)	2
Hungary	1
Iceland	2
India	3
Indonesia	3
Iran (Islamic Republic of)	3
Iraq	4
Ireland	2
Israel	2
Italy	2
Jamaica	3
Japan	2
Jordan	2
Kazakhstan	3
Kenya	3
Kiribati	3
Korea, North	3
Korea (Republic of)	2
Kuwait	2
Kyrgyzstan	3.5
Lao People's Democratic Republic	3
Latvia	2
Lebanon	3

Lesotho	4
Liberia	3
Libyan Arab Jamahiriya	3
Liechtenstein	1
Lithuania	3
Luxembourg	1
Macau	2
Macedonia (TFYR)	3
Madagascar	4
Malawi	3
Malaysia	3
Maldives	4
Mali	3
Malta	2
Mauritania	3
Mauritius	2
Mexico	3
Micronesia (Federated States of)	3
Moldova	4
Mongolia	2
Montenegro	2
Morocco	3
Mozambique	3
Namibia	2
Nepal	4
Netherlands	1
New Zealand	1
Nicaragua	4
Niger	4
Nigeria	4
Norway	1
Oman	2
Pakistan	4
Panama	2
Papua New Guinea	4
Paraguay	4
Peru	2
Philippines	4
Poland	2
Portugal	2
Qatar	2
Romania	5
Russian Federation	3
Rwanda	2
Saint Lucia	2
Saint Vincent and the Grenadines	2
Samoa	2
Sao Tome and Principe	2
Saudi Arabia	2
Senegal	3

Serbia	3
Seychelles	2
Sierra Leone	3
Singapore	1
Slovakia	2
Slovenia	1
Solomon Islands	4
South Africa	2
Spain	2
Sri Lanka	4
Sudan	4
Suriname	2
Swaziland	3
Sweden	1
Switzerland	1
Syrian Arab Republic	2
Taiwan	2
Tajikistan	4
Tanzania (United Republic of)	2
Thailand	3
Timor-Leste	4
Togo	4
Tonga	1
Trinidad and Tobago	1
Tunisia	2
Turkey	3
Turkmenistan	4
Uganda	3
Ukraine	2
United Arab Emirates	1
United Kingdom	1
United States	1
Uruguay	2
Uzbekistan	4
Vanuatu	1
Venezuela (Bolivarian Republic o	2
Viet Nam	2
Yemen	2
Zambia	3
Zimbabwe	4

Appendix 3

R code language employed

```
getwd()

## Read the data from csv file
# ---This command pertains only to my computer
#----You will need to name it something else depending on where you save
your files
AN_BF <- read.csv("C:\\Users\\Beth\\Documents\\R\\csv data files for
thesis\\AN_BF.csv", header=TRUE)


## Save the workspace as a name ex. AN_BF.Rdata
##--you will need to name it something else for your computer specific
#--this tells it where to save the workspace at so that I can email
results or view later
save.image("C:\\Users\\Beth\\Thesis data analysis\\AN_BF.Rdata")


## Print data
AN_BF

## Check "structure"
str(AN_BF)

## Tally the number of NAs in each column
apply(AN_BF, 2, function(x) length(which(is.na(x))))



## Graphically exploring the dataset, looking for expected relationships
#     with response
## and any strong correlations among predictor variables
# --> ncol(AN_BF) is the number of columns in AN_BF
#     Plot columns (i.e., variables)
#plot(AN_BF[,c(3:ncol(AN_BF),2)]) # scatterplot matrix
#pairs(AN_BF[,c(3:ncol(AN_BF),2)], panel = panel.smooth, lwd = 2, cex=
1.5, col="blue")# hmm...
pairs(AN_BF[,c(3:ncol(AN_BF),2)], panel = panel.smooth, lwd = 2, cex=
0.5, col="black")# hmm...

##  Compute all pairwise correlations if desired
# This was done on all computation sets
library(Hmisc)
r.results <- rcorr(as.matrix(AN_BF[,2:ncol(AN_BF)]), type="spearman")
r.results$r
```

```
# Create multiple plots for columns indexed by i (turn on Graphics window
History|Recording first)If you do not turn on the graphics window
#recording, you will need to re run the set. It will not be saved.

for (i in seq(from=3, to=ncol(AN_BF))){
   plot.data <- AN_BF[, c(i,2)]# Extract a pair of variables to plot
   plot.data <- na.omit(plot.data)# Delete observations for which either
variable is NA
   plot(plot.data, ylim=c(1,5))# Plot
   lines(smooth.spline(plot.data))# Augment with smoothed curve
   }
# --> PageUp/PageDown in Graphics window to see each plot


## Put object high in the list to determine comparison see attach in
intro book
search()

## Load randomForest package
library(randomForest)

## Look at structure of dataframe, and print values of Failing_Rank
str(AN_BF)
AN_BF$Failing_Rank

## Print data for first 5 countries
head(AN_BF)

## Print observations where failing rank is missing
AN_BF$FailingRank[AN_BF$FailingRank==NA,]

## Identify countries with any missing data
Incomplete <- AN_BF[!complete.cases(AN_BF),]

#  Print whole record for all countries with missing data
Incomplete

#  Print just Country for all countries with missing data
Incomplete$Country


## Copy the full dataset
## If rows (i.e., countries) or columns (i.e., variables) are omitted
using the code below,

## then this new dataset will be overwritten. If there are no edits, then
this dataset will

## move forward into random forest modeling
AN_BF.edited = AN_BF

## --> Do this step only if necessary or desired
## Make a dataset that omits selected countries, e.g.,
#  those with missing FailingRank
```

```
#  or those with too many variables with missing data
#  or those you want to drop for some other reason
#  Replace the numbers in the c( ) list with appropriate row numbers to
drop specific countries
#  Get a list of countries with row numbers
AN_BF.edited$Country
#  Drop countries in the list
AN_BF.edited = AN_BF.edited[-c(47,105,124,134,166,198),]
#  Check new structure; there should be fewer "obs.", consistent with
fewer countries
str(AN_BF.edited)
#  List countries in the new dataframe
AN_BF.edited$Country

## --> Do this step only if necessary or desired
## Make a dataset that omits selected variables, e.g.,
#  those with so many missing values that imputation does not seem
sensible,
#  or those you want to drop for some other reason (e.g., they are
redundant to other variables)
#  or to look at a small subset of variables that look "important"
## KEEP IN MIND that the total number of variables will be different if
some are dropped
#  Look at structure
str(AN_BF.edited)
#  Drop variables as desired by identifying columns in the c( ) list
AN_BF.edited = AN_BF.edited[,-c(4:11, 13:20)] # drop 4th through 11th,
and 13th through 20th
#  Check new structure to be sure that you dropped what you meant to drop
str(AN_BF.edited)

# --> Random forest modeling Choice #1:
#     Using an imputed dataset.
# --> For any computation that uses random numbers, like imputation or
random forest
#     modeling, you want to set a seed. Otherwise the computer will set
it by default
#     from the computer clock and you will not be able to exactly
duplicate the analysis.
set.seed(120)
# --> If there are no missing values, then an Error will be reported
#     and the AN_BF.imputed object will not be created. Hence any
subsequent command
#     that uses AN_BF.imputed will not run.
# --> If your dataset has no missing values (either by default or because
you deleted
#     all of the countries (rows) with missing data, then use
AN_BF.edited in subsequent
#     commands rather than AN_BF.imputed
# --> This process drops the Country variable from the AN_BF.imputed
object
#     Impute missing values
AN_BF.imputed <- rfImpute(Failing_Rank ~ .,
AN_BF.edited[,c(2:ncol(AN_BF.edited))])
```

```
# --> Check that all cases are now complete
complete.cases(AN_BF.imputed)
# --> Look at structure of new dataframe
str(AN_BF.imputed)
# --> Copy imputed dataset into dataset AN_BF.data to use in random
forest modeling
AN_BF.data = AN_BF.imputed
# --> Create a random forest object AN_BF.rf
set.seed(456)
AN_BF.rf=randomForest(Failing_Rank~.,data=AN_BF.data,importance=TRUE)
AN_BF.rf

# --> OR Random forest modeling Choice #2:
#     Using a complete dataset, either the original dataset if it had no
#     missing values or an edited dataset that was made by dropping all
rows (countries)
#     and/or columns (variables) with missing data.
# --> Copy imputed dataset into dataset AN_BF.data to use in random
forest modeling,
#     dropping the Country variable from the dataset object
AN_BF.data = AN_BF.edited[,c(2:ncol(AN_BF.edited))]
# --> Create a random forest object AN_BF.rf
set.seed(456)
AN_BF.rf=randomForest(Failing_Rank~.,data=AN_BF.data,importance=TRUE)
AN_BF.rf

# --> OR Random forest modeling Choice #3:
#     Using a dataset that may have missing values, dropping all
observations (countries)
#     with any missing data
# --> Copy imputed dataset into dataset AN_BF.data to use in random
forest modeling,
#     dropping the Country variable from the dataset object
AN_BF.data = na.omit(AN_BF.edited[,c(2:ncol(AN_BF.edited))])
# --> Check structure of new dataframe
str(AN_BF.data)
# --> Create a random forest object AN_BF.rf
set.seed(456)
AN_BF.rf=randomForest(Failing_Rank~.,data=AN_BF.data,importance=TRUE)
AN_BF.rf


## --> Continuing with the analysis, regardless of Choice
round(importance(AN_BF.rf), 2)
varImpPlot(AN_BF.rf)

## to automatically save the plot
setwd("C:\\Users\\Beth\\Documents\\R\\graphs for thesis\\")

#  change the file name to whatever the graph is
savePlot(filename="AN_BF Variable importance",type="emf")

#  multiple partial plots in 3 rows and 2 columns
imp<-importance(AN_BF.rf)
```

```
impvar<-rownames(imp)[order(imp[,1], decreasing=TRUE)]
par(mfrow=c(3,2))
for (i in seq_along(impvar)) {
  partialPlot(AN_BF.rf, AN_BF.data, impvar[i], xlab=impvar[i],
  main=paste("Partial Dependence on", impvar[i]), ylim=c(1,5))
  }
par(mfrow=c(1,1))

# for individual partial plot
# make graphic a large 1 by 1 matrix
par(mfrow=c(1,1))
partialPlot(AN_BF.rf, AN_BF.data, LifeBirth_Female,
xlab="LifeBirth_Female",
  ylab="Failing Rank",
  main=paste("Partial Dependence on", "LifeBirth_Female"), ylim=c(1,5))
savePlot(filename="C:\\Users\\Beth\\graphs for thesis\\AN_BF Partial Plot
LifeBirth_Female",
  type="emf")

##save new data objects in workspace
ls()
save.image("C:\\Users\\Beth\\Thesis data analysis\\AN_BF.Rdata")
```

Appendix 4 Hmisc results

```
> library(Hmisc)
> r.results <- rcorr(as.matrix(run2[,2:ncol(run2)]), type="spearman")
> r.results$r
```

	Failing_Rank	**AdultLit_2005**	**EnrolEduc_2005**	**GDPpercap_2005**
Failing_Rank	1.00000000	0.02697799	0.0851947516	-0.248745710
AdultLit_2005	0.02697799	1.00000000	0.0463679582	-0.036908664
EnrolEduc_2005	0.08519475	0.04636796	1.0000000000	0.086575672
GDPpercap_2005	-0.24874571	-0.03690866	0.0865756720	1.000000000
AdultLit_2006	0.11383769	0.62730968	0.0173741318	-0.070903368
HDI_2006	-0.63596439	-0.19248299	-0.0050803591	0.430028647
HDI_A1Reconcil_Rank	0.63617671	0.19208394	0.0054365443	-0.429583192
PoliticalRights	0.54707551	0.13005950	-0.0520302616	-0.006106872
CivilLiberties	0.57924598	0.15963461	-0.0294325501	-0.019184226
ElectoralProcess	-0.50577277	-0.10522731	0.0709697530	-0.028558489
PoliticalPluralism	-0.53421974	-0.10544204	0.0469114557	0.010303929
GovtFunction	-0.58038324	-0.13948955	0.0049100816	0.033694092
FreedomExpression	-0.56814373	-0.12557831	0.0154502923	0.034476887
AssocRights	-0.52849948	-0.12270474	0.0344730690	0.022420902
RuleLaw	-0.58514124	-0.16048244	0.0164689720	0.005314687
IndivRights	-0.60100901	-0.20106521	-0.0060573211	0.037086338
FreedomStatus	0.51773691	0.10652321	-0.0473107770	0.029706782
DISAP2005	-0.38758761	-0.03165168	0.1180226430	0.111951649
Progress_ST	0.18639658	0.21095045	0.2107104957	0.031337664

GDP	-0.61502534	-0.16067779	0.0063518756	0.536121130
HDI_Trends_Rank	0.63095719	0.19208394	0.0054365443	-0.429583192
GDPpercap	-0.61405259	-0.17096333	0.0003752598	0.536530316
LifeBirth	-0.60013753	-0.13539150	-0.0752368197	0.323006004
LifeExp	-0.59949005	-0.13514304	-0.0759409517	0.323116124
EarnIng_Female	-0.64474130	-0.19459738	0.0286328439	0.499794543
Seats_Women	-0.25070402	-0.06595705	0.0029300835	-0.067054778
Admin_Women	-0.17601755	-0.21298316	0.0094822142	0.011029314
Prof_Women	-0.10955402	-0.41417724	0.0997921675	-0.101032443
EarnIng_Male	-0.64237249	-0.14544363	0.0001138583	0.528035283
OLD_EMPINX1990	-0.44005004	-0.21366489	-0.0452510379	0.047151525
SPEECH1990	-0.44005945	-0.19672032	-0.0529744104	0.159773558
ELECSD2007	-0.43849507	-0.03614708	0.0361889787	-0.012454747
Progress_LT	0.37626520	0.38514042	0.0581945032	-0.015066378
Progress_MT	0.16695333	0.47412163	-0.0175123420	-0.285973549
ALR	-0.31345794	-0.26788428	0.0513533987	0.296765745
EdEnroll	-0.56001270	-0.21732248	0.0362962298	0.336800456
Edu	-0.53535765	-0.30806777	0.0282828324	0.305609882
WOSOC2003	-0.44952592	-0.19484411	0.0483556762	0.052507211
WECON2003	-0.46219319	-0.20606793	0.0634766221	0.097478129
WOPOL2006	-0.16770512	-0.15532932	-0.0731181279	-0.156219810
COW	0.24864329	0.12362177	-0.0537854023	-0.192376986
POLITY	0.25118226	0.12574111	-0.0534191951	-0.192376986
UNSUBREG	-0.24048531	-0.26868746	-0.0842922851	0.230426684
Unimprov_Water	0.49786022	0.07333523	0.0072254627	-0.285795748

LifeBirth_Female	-0.62169522	-0.16466987	-0.0333726257	0.324246198
LifeBirth_Male	-0.60210639	-0.11294315	-0.0735311434	0.324774683
AdultLit_Female	-0.34737045	-0.25754476	0.0343733840	0.258503467
AdultLit_Male	-0.30301630	-0.27829573	0.0422901288	0.319651425
EnrollEduc_Female	-0.58448786	-0.24265614	0.0049101929	0.333301306
EnrollEduc_Male	-0.53936839	-0.25156751	0.0162247103	0.334829837
Surv_40	0.45933023	0.04103788	0.0203386340	-0.326344043
Poverty_200	0.44574928	0.13835996	-0.1575469077	-0.355751991
POLPRIS1993	-0.02558983	0.10134294	-0.0436097234	-0.062632650

	AdultLit_2006	**HDI_2006**	**HDI_A1Reconcil_Rank**
Failing_Rank	0.11383769	-0.635964394	0.636176705
AdultLit_2005	0.62730968	-0.192482993	0.192083940
EnrolEduc_2005	0.01737413	-0.005080359	0.005436544
GDPpercap_2005	-0.07090337	0.430028647	-0.429583192
AdultLit_2006	1.00000000	-0.383566141	0.383391470
HDI_2006	-0.38356614	1.000000000	-0.999986410
HDI_A1Reconcil_Rank	0.38339147	-0.999986410	1.000000000
PoliticalRights	0.24357341	-0.537640929	0.537667453
CivilLiberties	0.27023625	-0.589725614	0.589622200
ElectoralProcess	-0.23420817	0.493383318	-0.493352205
PoliticalPluralism	-0.23018314	0.521081924	-0.520920813
GovtFunction	-0.25261465	0.560757399	-0.560876071
FreedomExpression	-0.22634208	0.519603014	-0.519470930
AssocRights	-0.22636390	0.507611513	-0.507318377

RuleLaw	-0.27063823	0.583854020	-0.583960295
IndivRights	-0.34219384	0.678043783	-0.677946806
FreedomStatus	0.19605599	-0.471143216	0.471388847
DISAP2005	-0.07118428	0.262089640	-0.261888802
Progress_ST	0.37397963	-0.369489759	0.369453013
GDP	-0.31272355	0.948860645	-0.948753536
HDI_Trends_Rank	0.38339147	-0.999986410	1.000000000
GDPpercap	-0.30870351	0.949518323	-0.949401975
LifeBirth	-0.32445711	0.933804154	-0.933910191
LifeExp	-0.32460445	0.933611095	-0.933715522
EarnIng_Female	-0.36146429	0.944190621	-0.944079757
Seats_Women	-0.17863925	0.234579623	-0.234277889
Admin_Women	-0.33580786	0.097385503	-0.096904181
Prof_Women	-0.52336985	0.183156326	-0.183604211
EarnIng_Male	-0.29610494	0.945450902	-0.945398688
OLD_EMPINX1990	-0.32388648	0.535366893	-0.534776509
SPEECH1990	-0.28245655	0.453809977	-0.453614801
ELECSD2007	-0.08909162	0.336264402	-0.336110085
Progress_LT	0.50145411	-0.329182148	0.328977227
Progress_MT	0.50059414	-0.301268607	0.301322252
ALR	-0.49664107	0.783320844	-0.783357978
EdEnroll	-0.35110492	0.869987845	-0.869750559
Edu	-0.48087588	0.878686428	-0.878678679
WOSOC2003	-0.29256171	0.582607567	-0.582167447
WECON2003	-0.25238499	0.535259247	-0.535613835

WOPOL2006	-0.23135683	0.163696811	-0.163358018
COW	0.24906459	-0.320214957	0.319751441
POLITY	0.25102791	-0.323124617	0.322670043
UNSUBREG	-0.26096317	0.535796463	-0.535898864
Unimprov_Water	0.25045758	-0.818641841	0.818749607
LifeBirth_Female	-0.36299199	0.944966912	-0.945028365
LifeBirth_Male	-0.28515211	0.907627106	-0.907730341
AdultLit_Female	-0.49747863	0.781164289	-0.781269133
AdultLit_Male	-0.47851905	0.774360538	-0.774483085
EnrollEduc_Female	-0.36894932	0.875091255	-0.874780774
EnrollEduc_Male	-0.36054954	0.855759263	-0.855338871
Surv_40	0.23032515	-0.912202239	0.912346423
Poverty_200	0.32868043	-0.900733054	0.900777519
POLPRIS1993	0.11815877	-0.040827919	0.040728055

	PoliticalRights	**CivilLiberties**	**ElectoralProcess**
Failing_Rank	0.547075510	0.57924598	-0.50577277
AdultLit_2005	0.130059496	0.15963461	-0.10522731
EnrolEduc_2005	-0.052030262	-0.02943255	0.07096975
GDPpercap_2005	-0.006106872	-0.01918423	-0.02855849
AdultLit_2006	0.243573412	0.27023625	-0.23420817
HDI_2006	-0.537640929	-0.58972561	0.49338332
HDI_A1Reconcil_Rank	0.537667453	0.58962220	-0.49335220
PoliticalRights	1.000000000	0.94438207	-0.96647555
CivilLiberties	0.944382071	1.00000000	-0.90561801

ElectoralProcess	-0.966475546	-0.90561801	1.00000000
PoliticalPluralism	-0.967481136	-0.94134063	0.93654853
GovtFunction	-0.954280257	-0.93319625	0.90849406
FreedomExpression	-0.930052221	-0.95643997	0.90285641
AssocRights	-0.932795167	-0.95032841	0.91222590
RuleLaw	-0.917431831	-0.96034533	0.86626369
IndivRights	-0.890894294	-0.94611591	0.85111308
FreedomStatus	0.937147439	0.91540259	-0.90272355
DISAP2005	-0.391567826	-0.40905383	0.37371707
Progress_ST	0.317109346	0.33829808	-0.28972384
GDP	-0.442463905	-0.49958798	0.40239230
HDI_Trends_Rank	0.539304614	0.59112281	-0.49542701
GDPpercap	-0.441896707	-0.49894869	0.40206948
LifeBirth	-0.507694662	-0.54861081	0.46441242
LifeExp	-0.507705510	-0.54860926	0.46435705
EarnIng_Female	-0.509087205	-0.57436627	0.47139797
Seats_Women	-0.169129208	-0.17236039	0.19086848
Admin_Women	-0.361237884	-0.40966564	0.37493256
Prof_Women	-0.295592576	-0.33614707	0.32507414
EarnIng_Male	-0.446632862	-0.50253177	0.40700874
OLD_EMPINX1990	-0.780963004	-0.82143790	0.76257253
SPEECH1990	-0.666586757	-0.69026029	0.64663094
ELECSD2007	-0.755801678	-0.70969433	0.77142590
Progress_LT	0.494528115	0.48357841	-0.47228754
Progress_MT	0.238795444	0.25357905	-0.23140550

ALR	-0.243760973	-0.32802901	0.21199600
EdEnroll	-0.533041298	-0.56538314	0.50561786
Edu	-0.527754903	-0.57763833	0.48754296
WOSOC2003	-0.518549263	-0.53657407	0.49574777
WECON2003	-0.400483310	-0.47706851	0.36419705
WOPOL2006	-0.283918798	-0.26337799	0.32338703
COW	0.431596041	0.40458068	-0.47683045
POLITY	0.432444811	0.40580153	-0.47756970
UNSUBREG	-0.158068240	-0.21357927	0.12505431
Unimprov_Water	0.351939023	0.41801113	-0.32900479
LifeBirth_Female	-0.551280260	-0.59265864	0.50950617
LifeBirth_Male	-0.503672481	-0.54320037	0.45627579
AdultLit_Female	-0.312993139	-0.40503189	0.27375188
AdultLit_Male	-0.208135977	-0.30488029	0.17642967
EnrollEduc_Female	-0.562310994	-0.59900844	0.53738141
EnrollEduc_Male	-0.556681752	-0.57583493	0.53276849
Surv_40	0.303058505	0.37749767	-0.25048339
Poverty_200	0.362890124	0.43846568	-0.37244534
POLPRIS1993	-0.081042729	-0.10071660	0.09601165

	PoliticalPluralism	**GovtFunction**	**FreedomExpression**
Failing_Rank	-0.53421974	-0.580383241	-0.56814373
AdultLit_2005	-0.10544204	-0.139489546	-0.12557831
EnrolEduc_2005	0.04691146	0.004910082	0.01545029
GDPpercap_2005	0.01030393	0.033694092	0.03447689

AdultLit_2006	-0.23018314	-0.252614647	-0.22634208
HDI_2006	0.52108192	0.560757399	0.51960301
HDI_A1Reconcil_Rank	-0.52092081	-0.560876071	-0.51947093
PoliticalRights	-0.96748114	-0.954280257	-0.93005222
CivilLiberties	-0.94134063	-0.933196247	-0.95643997
ElectoralProcess	0.93654853	0.908494055	0.90285641
PoliticalPluralism	1.00000000	0.923494518	0.93779808
GovtFunction	0.92349452	1.000000000	0.91280740
FreedomExpression	0.93779808	0.912807405	1.00000000
AssocRights	0.93330145	0.915465117	0.93549627
RuleLaw	0.91224945	0.933803260	0.91087741
IndivRights	0.89299595	0.905044854	0.90060008
FreedomStatus	-0.91974729	-0.902664602	-0.90618026
DISAP2005	0.37945792	0.383428782	0.40070885
Progress_ST	-0.33597210	-0.342695504	-0.33543631
GDP	0.42316464	0.483143061	0.44266516
HDI_Trends_Rank	-0.52247071	-0.562402070	-0.52160841
GDPpercap	0.42253974	0.482644200	0.44210896
LifeBirth	0.48506752	0.550043225	0.48420775
LifeExp	0.48538163	0.550045073	0.48435014
EarnIng_Female	0.49844387	0.535347044	0.52168965
Seats_Women	0.18328646	0.225743100	0.18597150
Admin_Women	0.35284200	0.359516889	0.44221571
Prof_Women	0.28336206	0.273163706	0.32873791
EarnIng_Male	0.43187645	0.477569699	0.44749025

OLD_EMPINX1990	0.78297788	0.757665157	0.80487406
SPEECH1990	0.67512572	0.626532197	0.69412965
ELECSD2007	0.74897516	0.716874361	0.73179197
Progress_LT	-0.48545143	-0.470533818	-0.48946443
Progress_MT	-0.22253641	-0.266235441	-0.26388118
ALR	0.22059275	0.251771897	0.24473839
EdEnroll	0.52428097	0.534959435	0.54515082
Edu	0.51863372	0.536301315	0.52927285
WOSOC2003	0.52192920	0.519593835	0.51619965
WECON2003	0.42172581	0.451772451	0.42449170
WOPOL2006	0.28983042	0.319381565	0.26854649
COW	-0.44038776	-0.400209546	-0.43980920
POLITY	-0.44002721	-0.400294334	-0.44047418
UNSUBREG	0.15186514	0.219719529	0.14879431
Unimprov_Water	-0.32313126	-0.388489008	-0.35152331
LifeBirth_Female	0.52711606	0.586949170	0.52089614
LifeBirth_Male	0.48775935	0.547114909	0.47604969
AdultLit_Female	0.29396018	0.312303603	0.33988979
AdultLit_Male	0.19503127	0.209660277	0.23070431
EnrollEduc_Female	0.55225074	0.570115864	0.58038723
EnrollEduc_Male	0.55314618	0.553605437	0.55801022
Surv_40	-0.28051412	-0.330858260	-0.27705276
Poverty_200	-0.35086536	-0.330064088	-0.36027229
POLPRIS1993	0.07560226	0.083667815	0.08423979

	AssocRights	RuleLaw	IndivRights	FreedomStatus
Failing_Rank	-0.52849948	-0.585141242	-0.601009011	0.51773691
AdultLit_2005	-0.12270474	-0.160482436	-0.201065212	0.10652321
EnrolEduc_2005	0.03447307	0.016468972	-0.006057321	-0.04731078
GDPpercap_2005	0.02242090	0.005314687	0.037086338	0.02970678
AdultLit_2006	-0.22636390	-0.270638227	-0.342193842	0.19605599
HDI_2006	0.50761151	0.583854020	0.678043783	-0.47114322
HDI_A1Reconcil_Rank	-0.50731838	-0.583960295	-0.677946806	0.47138885
PoliticalRights	-0.93279517	-0.917431831	-0.890894294	0.93714744
CivilLiberties	-0.95032841	-0.960345328	-0.946115911	0.91540259
ElectoralProcess	0.91222590	0.866263688	0.851113081	-0.90272355
PoliticalPluralism	0.93330145	0.912249446	0.892995954	-0.91974729
GovtFunction	0.91546512	0.933803260	0.905044854	-0.90266460
FreedomExpression	0.93549627	0.910877407	0.900600076	-0.90618026
AssocRights	1.00000000	0.903326452	0.885020554	-0.90136027
RuleLaw	0.90332645	1.000000000	0.931369185	-0.88168204
IndivRights	0.88502055	0.931369185	1.000000000	-0.85831940
FreedomStatus	-0.90136027	-0.881682038	-0.858319402	1.00000000
DISAP2005	0.38118166	0.432924509	0.369896233	-0.35271356
Progress_ST	-0.33471537	-0.355973721	-0.380730152	0.28981361
GDP	0.43337011	0.496219039	0.578128934	-0.39208645
HDI_Trends_Rank	-0.50939190	-0.585429370	-0.678846180	0.47366464
GDPpercap	0.43282714	0.495879650	0.577812135	-0.39143285
LifeBirth	0.48174530	0.544982791	0.658314168	-0.44294170
LifeExp	0.48177880	0.545050979	0.658454835	-0.44312918

EarnIng_Female	0.50724006	0.562206268	0.649822354	-0.45731312
Seats_Women	0.18437541	0.163147897	0.221796528	-0.16207905
Admin_Women	0.33993477	0.337255567	0.384615868	-0.40714043
Prof_Women	0.34134516	0.255117953	0.338858575	-0.35722876
EarnIng_Male	0.42999730	0.497813821	0.579064548	-0.39733073
OLD_EMPINX1990	0.78413832	0.781591654	0.797511578	-0.75761771
SPEECH1990	0.66621572	0.677234948	0.645810485	-0.62820321
ELECSD2007	0.73334098	0.674335003	0.657062292	-0.70252407
Progress_LT	-0.50133830	-0.495151192	-0.454295814	0.43065155
Progress_MT	-0.28944305	-0.267498165	-0.249489918	0.20248102
ALR	0.23055692	0.315209240	0.440868437	-0.23085994
EdEnroll	0.51493520	0.542042851	0.639029384	-0.49623731
Edu	0.51217872	0.567452431	0.653799772	-0.47999164
WOSOC2003	0.49893105	0.519890606	0.586038828	-0.47310460
WECON2003	0.41038567	0.497486442	0.529785454	-0.36614451
WOPOL2006	0.29613894	0.281660408	0.308042765	-0.25010797
COW	-0.44038177	-0.330050915	-0.399929792	0.40967098
POLITY	-0.44177458	-0.331039965	-0.401823014	0.40869999
UNSUBREG	0.16477765	0.235176206	0.281042755	-0.11539730
Unimprov_Water	-0.33562526	-0.392918140	-0.528129756	0.34662068
LifeBirth_Female	0.52316189	0.586472988	0.690726399	-0.47985685
LifeBirth_Male	0.46576345	0.547383904	0.650316358	-0.43232995
AdultLit_Female	0.29741171	0.396186680	0.511588335	-0.29893103
AdultLit_Male	0.20988682	0.297066927	0.434626371	-0.20098864
EnrollEduc_Female	0.54704469	0.569474757	0.663271666	-0.51802528

EnrollEduc_Male	0.54254824	0.554376602	0.654744983	-0.49900439
Surv_40	-0.26576987	-0.366238058	-0.509423077	0.27980503
Poverty_200	-0.32478002	-0.374444664	-0.577836871	0.34788749
POLPRIS1993	0.08413205	0.121352091	0.063788973	-0.11349798

	DISAP2005	**Progress_ST**	**GDP**	**HDI_Trends_Rank**
Failing_Rank	-0.38758761	0.186396584	-0.615025342	0.630957186
AdultLit_2005	-0.03165168	0.210950449	-0.160677791	0.192083940
EnrolEduc_2005	0.11802264	0.210710496	0.006351876	0.005436544
GDPpercap_2005	0.11195165	0.031337664	0.536121130	-0.429583192
AdultLit_2006	-0.07118428	0.373979628	-0.312723547	0.383391470
HDI_2006	0.26208964	-0.369489759	0.948860645	-0.999986410
HDI_A1Reconcil_Rank	-0.26188880	0.369453013	-0.948753536	1.000000000
PoliticalRights	-0.39156783	0.317109346	-0.442463905	0.539304614
CivilLiberties	-0.40905383	0.338298082	-0.499587983	0.591122806
ElectoralProcess	0.37371707	-0.289723843	0.402392298	-0.495427012
PoliticalPluralism	0.37945792	-0.335972100	0.423164636	-0.522470713
GovtFunction	0.38342878	-0.342695504	0.483143061	-0.562402070
FreedomExpression	0.40070885	-0.335436314	0.442665160	-0.521608412
AssocRights	0.38118166	-0.334715366	0.433370113	-0.509391904
RuleLaw	0.43292451	-0.355973721	0.496219039	-0.585429370
IndivRights	0.36989623	-0.380730152	0.578128934	-0.678846180
FreedomStatus	-0.35271356	0.289813608	-0.392086446	0.473664641
DISAP2005	1.00000000	-0.186140329	0.221342877	-0.259035915
Progress_ST	-0.18614033	1.000000000	-0.344958752	0.370078027

GDP	0.22134288	-0.344958752	1.000000000	-0.948297203
HDI_Trends_Rank	-0.25903592	0.370078027	-0.948297203	1.000000000
GDPpercap	0.22147235	-0.345790416	0.999931693	-0.948935449
LifeBirth	0.26513770	-0.352263957	0.833822966	-0.934383154
LifeExp	0.26473927	-0.352857381	0.833632469	-0.934191704
EarnIng_Female	0.22887747	-0.364341646	0.978015304	-0.944079757
Seats_Women	0.04473292	-0.317611843	0.202271461	-0.234277889
Admin_Women	0.18362819	-0.279127896	0.067554861	-0.096904181
Prof_Women	0.11569655	-0.123055354	0.100230381	-0.183604211
EarnIng_Male	0.22555541	-0.337840527	0.996558785	-0.945398688
OLD_EMPINX1990	0.27764863	-0.436269104	0.492423385	-0.535376728
SPEECH1990	0.26300561	-0.339644790	0.428210199	-0.455012769
ELECSD2007	0.28562614	-0.194780171	0.277367920	-0.339494526
Progress_LT	-0.33144596	0.705702841	-0.282749414	0.324108154
Progress_MT	-0.20802465	0.444662631	-0.363772452	0.309781939
ALR	0.17478155	-0.202458784	0.643701196	-0.781870544
EdEnroll	0.26815289	-0.370891184	0.785950124	-0.871269345
Edu	0.26020411	-0.382997453	0.770789921	-0.879901290
WOSOC2003	0.21412116	-0.237325519	0.462973744	-0.582167447
WECON2003	0.27278146	-0.096584536	0.458737671	-0.537212014
WOPOL2006	0.06807546	-0.316652805	0.157701313	-0.170307815
COW	-0.17106958	0.228187740	-0.298374653	0.322819710
POLITY	-0.17106958	0.228187740	-0.300163329	0.325753957
UNSUBREG	0.07847913	-0.114679649	0.499518812	-0.534262955
Unimprov_Water	-0.18781652	0.243267313	-0.753166914	0.819719911

LifeBirth_Female	0.27737620	-0.348840356	0.842526913	-0.945028365
LifeBirth_Male	0.27163446	-0.341296762	0.811933637	-0.907730341
AdultLit_Female	0.18938942	-0.245197400	0.650438786	-0.781269133
AdultLit_Male	0.15826266	-0.142696589	0.652076244	-0.774483085
EnrollEduc_Female	0.26854813	-0.363403320	0.785921633	-0.874780774
EnrollEduc_Male	0.22618480	-0.373896956	0.770073414	-0.855338871
Surv_40	-0.23300095	0.117368340	-0.777336240	0.912840903
Poverty_200	-0.16810924	0.108158514	-0.872796237	0.900777519
POLPRIS1993	0.16626303	0.004667582	-0.037407175	0.031312138

	GDPpercap	**LifeBirth**	**LifeExp**	**EarnIng_Female**
Failing_Rank	-0.6140525937	-0.60013753	-0.59949005	-0.64474130
AdultLit_2005	-0.1709633321	-0.13539150	-0.13514304	-0.19459738
EnrolEduc_2005	0.0003752598	-0.07523682	-0.07594095	0.02863284
GDPpercap_2005	0.5365303159	0.32300600	0.32311612	0.49979454
AdultLit_2006	-0.3087035120	-0.32445711	-0.32460445	-0.36146429
HDI_2006	0.9495183229	0.93380415	0.93361109	0.94419062
HDI_A1Reconcil_Rank	-0.9494019747	-0.93391019	-0.93371552	-0.94407976
PoliticalRights	-0.4418967068	-0.50769466	-0.50770551	-0.50908720
CivilLiberties	-0.4989486933	-0.54861081	-0.54860926	-0.57436627
ElectoralProcess	0.4020694792	0.46441242	0.46435705	0.47139797
PoliticalPluralism	0.4225397408	0.48506752	0.48538163	0.49844387
GovtFunction	0.4826442003	0.55004323	0.55004507	0.53534704
FreedomExpression	0.4421089590	0.48420775	0.48435014	0.52168965
AssocRights	0.4328271449	0.48174530	0.48177880	0.50724006
RuleLaw	0.4958796501	0.54498279	0.54505098	0.56220627

IndivRights	0.5778121352	0.65831417	0.65845484	0.64982235
FreedomStatus	-0.3914328516	-0.44294170	-0.44312918	-0.45731312
DISAP2005	0.2214723527	0.26513770	0.26473927	0.22887747
Progress_ST	-0.3457904160	-0.35226396	-0.35285738	-0.36434165
GDP	0.9999316931	0.83382297	0.83363247	0.97801530
HDI_Trends_Rank	-0.9489354491	-0.93438315	-0.93419170	-0.94407976
GDPpercap	1.0000000000	0.83794165	0.83773381	0.97929943
LifeBirth	0.8379416466	1.00000000	0.99998194	0.82415754
LifeExp	0.8377338052	0.99998194	1.00000000	0.82403356
EarnIng_Female	0.9792994261	0.82415754	0.82403356	1.00000000
Seats_Women	0.2025850564	0.21750504	0.21703789	0.25052324
Admin_Women	0.0503443778	0.01578032	0.01587339	0.19479311
Prof_Women	0.0810177699	0.06930289	0.07026724	0.23616332
EarnIng_Male	0.9964874387	0.83960348	0.83934879	0.96892852
OLD_EMPINX1990	0.4923361838	0.49735919	0.49798006	0.54505062
SPEECH1990	0.4281404912	0.42805210	0.42850965	0.45880964
ELECSD2007	0.2769346237	0.30917901	0.30930671	0.34058225
Progress_LT	-0.2833374441	-0.23367557	-0.23324271	-0.37816095
Progress_MT	-0.3646709621	-0.14368549	-0.14395040	-0.40452957
ALR	0.6467152834	0.63511461	0.63477045	0.69960511
EdEnroll	0.7916238904	0.76990497	0.76971710	0.81314903
Edu	0.7741427422	0.76786393	0.76755524	0.81711841
WOSOC2003	0.4629864693	0.50605863	0.50630742	0.53674096
WECON2003	0.4583230317	0.50723892	0.50647449	0.49472439
WOPOL2006	0.1584373116	0.13928635	0.13922371	0.19467112

COW	-0.2973598540	-0.33543274	-0.33605617	-0.32135794
POLITY	-0.2991590202	-0.33939192	-0.34001246	-0.32420799
UNSUBREG	0.5003862977	0.43119729	0.43050861	0.51639909
Unimprov_Water	-0.7535790801	-0.75996768	-0.75976670	-0.74140477
LifeBirth_Female	0.8444795609	0.97360003	0.97351211	0.84242189
LifeBirth_Male	0.8168677092	0.97839665	0.97844064	0.79425037
AdultLit_Female	0.6508321166	0.64332259	0.64333636	0.70126492
AdultLit_Male	0.6578247547	0.64284146	0.64261901	0.68865925
EnrollEduc_Female	0.7919008732	0.77614331	0.77576137	0.81747997
EnrollEduc_Male	0.7756068110	0.75986940	0.75976247	0.80654687
Surv_40	-0.7838840485	-0.97104955	-0.97090280	-0.74592066
Poverty_200	-0.8726763129	-0.81007999	-0.80958456	-0.85819232
POLPRIS1993	-0.0366780385	-0.01074270	-0.01153954	-0.07091042
	Seats_Women	**Admin_Women**	**Prof_Women**	**EarnIng_Male**
Failing_Rank	-0.250704020	-0.176017553	-0.10955402	-0.6423724890
AdultLit_2005	-0.065957054	-0.212983161	-0.41417724	-0.1454436332
EnrolEduc_2005	0.002930084	0.009482214	0.09979217	0.0001138583
GDPpercap_2005	-0.067054778	0.011029314	-0.10103244	0.5280352831
AdultLit_2006	-0.178639248	-0.335807860	-0.52336985	-0.2961049378
HDI_2006	0.234579623	0.097385503	0.18315633	0.9454509020
HDI_A1Reconcil_Rank	-0.234277889	-0.096904181	-0.18360421	-0.9453986883
PoliticalRights	-0.169129208	-0.361237884	-0.29559258	-0.4466328621
CivilLiberties	-0.172360390	-0.409665644	-0.33614707	-0.5025317669
ElectoralProcess	0.190868482	0.374932557	0.32507414	0.4070087373
PoliticalPluralism	0.183286458	0.352842003	0.28336206	0.4318764508

GovtFunction	0.225743100	0.359516889	0.27316371	0.4775696993
FreedomExpression	0.185971498	0.442215711	0.32873791	0.4474902451
AssocRights	0.184375405	0.339934766	0.34134516	0.4299972951
RuleLaw	0.163147897	0.337255567	0.25511795	0.4978138208
IndivRights	0.221796528	0.384615868	0.33885857	0.5790645480
FreedomStatus	-0.162079051	-0.407140434	-0.35722876	-0.3973307312
DISAP2005	0.044732921	0.183628187	0.11569655	0.2255554050
Progress_ST	-0.317611843	-0.279127896	-0.12305535	-0.3378405273
GDP	0.202271461	0.067554861	0.10023038	0.9965587854
HDI_Trends_Rank	-0.234277889	-0.096904181	-0.18360421	-0.9453986883
GDPpercap	0.202585056	0.050344378	0.08101777	0.9964874387
LifeBirth	0.217505038	0.015780324	0.06930289	0.8396034837
LifeExp	0.217037886	0.015873395	0.07026724	0.8393487930
EarnIng_Female	0.250523239	0.194793105	0.23616332	0.9689285159
Seats_Women	1.000000000	0.270869136	0.18667163	0.1784800887
Admin_Women	0.270869136	1.000000000	0.55485570	0.0420527458
Prof_Women	0.186671630	0.554855704	1.00000000	0.0667365342
EarnIng_Male	0.178480089	0.042052746	0.06673653	1.0000000000
OLD_EMPINX1990	0.236425236	0.469473720	0.48713535	0.4775406420
SPEECH1990	0.177913979	0.310264796	0.33900404	0.4133956730
ELECSD2007	0.157554716	0.363321334	0.25589523	0.2750814259
Progress_LT	-0.388169020	-0.506690502	-0.48862085	-0.2970980108
Progress_MT	-0.196966648	-0.426243305	-0.45877489	-0.3362294137
ALR	0.061867572	0.336560100	0.72111654	0.6393366456
EdEnroll	0.220086604	0.313123196	0.39774960	0.7925933599

Edu	0.252111644	0.284347355	0.53006250	0.7691794038
WOSOC2003	0.315631002	0.218239084	0.30404428	0.4715750217
WECON2003	0.272065282	0.181219324	0.19969313	0.4598045349
WOPOL2006	0.604778945	0.283548146	0.30742148	0.1138004288
COW	-0.297051758	-0.440053046	-0.35137358	-0.2823579907
POLITY	-0.294671714	-0.436909318	-0.34906891	-0.2845457792
UNSUBREG	0.024909241	-0.130753189	0.10268971	0.5175127387
Unimprov_Water	-0.079555154	-0.078664586	-0.34315640	-0.7614439130
LifeBirth_Female	0.211445138	0.071581215	0.14638385	0.8475594521
LifeBirth_Male	0.176071927	-0.048765160	-0.03498290	0.8216308951
AdultLit_Female	0.110621594	0.484909922	0.76123810	0.6312161088
AdultLit_Male	0.084503479	0.298121065	0.66249830	0.6393441558
EnrollEduc_Female	0.241279557	0.354479909	0.44256899	0.7923355699
EnrollEduc_Male	0.231503934	0.243843228	0.31778511	0.7793121338
Surv_40	0.058021832	0.119456396	-0.08131274	-0.7819010019
Poverty_200	0.033583030	-0.158697531	-0.44135115	-0.8726723790
POLPRIS1993	-0.197988003	0.162749872	-0.08437014	-0.0380039848

	OLD_EMPINX1990	**SPEECH1990**	**ELECSD2007**	**Progress_LT**
Failing_Rank	-0.44005004	-0.44005945	-0.43849507	0.3762651980
AdultLit_2005	-0.21366489	-0.19672032	-0.03614708	0.3851404190
EnrolEduc_2005	-0.04525104	-0.05297441	0.03618898	0.0581945032
GDPpercap_2005	0.04715152	0.15977356	-0.01245475	-0.0150663778
AdultLit_2006	-0.32388648	-0.28245655	-0.08909162	0.5014541149
HDI_2006	0.53536689	0.45380998	0.33626440	-0.3291821480

HDI_A1Reconcil_Rank	-0.53477651	-0.45361480	-0.33611009	0.3289772272
PoliticalRights	-0.78096300	-0.66658676	-0.75580168	0.4945281148
CivilLiberties	-0.82143790	-0.69026029	-0.70969433	0.4835784137
ElectoralProcess	0.76257253	0.64663094	0.77142590	-0.4722875357
PoliticalPluralism	0.78297788	0.67512572	0.74897516	-0.4854514301
GovtFunction	0.75766516	0.62653220	0.71687436	-0.4705338180
FreedomExpression	0.80487406	0.69412965	0.73179197	-0.4894644320
AssocRights	0.78413832	0.66621572	0.73334098	-0.5013383031
RuleLaw	0.78159165	0.67723495	0.67433500	-0.4951511919
IndivRights	0.79751158	0.64581048	0.65706229	-0.4542958140
FreedomStatus	-0.75761771	-0.62820321	-0.70252407	0.4306515455
DISAP2005	0.27764863	0.26300561	0.28562614	-0.3314459622
Progress_ST	-0.43626910	-0.33964479	-0.19478017	0.7057028413
GDP	0.49242339	0.42821020	0.27736792	-0.2827494144
HDI_Trends_Rank	-0.53537673	-0.45501277	-0.33949453	0.3241081536
GDPpercap	0.49233618	0.42814049	0.27693462	-0.2833374441
LifeBirth	0.49735919	0.42805210	0.30917901	-0.2336755693
LifeExp	0.49798006	0.42850965	0.30930671	-0.2332427055
EarnIng_Female	0.54505062	0.45880964	0.34058225	-0.3781609535
Seats_Women	0.23642524	0.17791398	0.15755472	-0.3881690204
Admin_Women	0.46947372	0.31026480	0.36332133	-0.5066905022
Prof_Women	0.48713535	0.33900404	0.25589523	-0.4886208475
EarnIng_Male	0.47754064	0.41339567	0.27508143	-0.2970980108
OLD_EMPINX1990	1.00000000	0.77460682	0.69534922	-0.5355617404
SPEECH1990	0.77460682	1.00000000	0.55443025	-0.4176353812

ELECSD2007	0.69534922	0.55443025	1.00000000	-0.3245204985
Progress_LT	-0.53556174	-0.41763538	-0.32452050	1.0000000000
Progress_MT	-0.40202385	-0.34352434	-0.06792083	0.7919880748
ALR	0.30970633	0.29745039	0.02979096	-0.2878564298
EdEnroll	0.56010413	0.48052365	0.35425991	-0.3103162348
Edu	0.54606318	0.48764977	0.31206605	-0.4099386334
WOSOC2003	0.49251539	0.37783498	0.44614226	-0.3801598251
WECON2003	0.33522117	0.21321660	0.23521544	-0.2600356638
WOPOL2006	0.34076020	0.28633839	0.28113094	-0.5105660558
COW	-0.45336005	-0.40182438	-0.38610086	0.3722806275
POLITY	-0.45336005	-0.40182438	-0.38870418	0.3722806275
UNSUBREG	0.14366330	0.14071129	0.01704218	-0.0358309671
Unimprov_Water	-0.30211017	-0.20691605	-0.13969818	-0.1467339396
LifeBirth_Female	0.52328551	0.44733515	0.35168314	-0.2554354370
LifeBirth_Male	0.46373397	0.40611520	0.30260348	-0.2067137361
AdultLit_Female	0.36739624	0.33679783	0.11260743	-0.3524624705
AdultLit_Male	0.22850463	0.26926067	0.02782055	-0.1308524609
EnrollEduc_Female	0.61640865	0.50627071	0.37255657	-0.3512610793
EnrollEduc_Male	0.59704596	0.52249694	0.37903893	-0.3196311891
Surv_40	-0.16995938	-0.14573340	-0.11494207	-0.1357613206
Poverty_200	-0.41974258	-0.35720897	-0.17851424	-0.0995636806
POLPRIS1993	0.16663888	0.16979913	0.11046769	-0.0008901446

	Progress_MT	ALR	EdEnroll	Edu
Failing_Rank	0.166953325	-0.31345794	-0.56001270	-0.53535765
AdultLit_2005	0.474121630	-0.26788428	-0.21732248	-0.30806777
EnrolEduc_2005	-0.017512342	0.05135340	0.03629623	0.02828283
GDPpercap_2005	-0.285973549	0.29676574	0.33680046	0.30560988
AdultLit_2006	0.500594139	-0.49664107	-0.35110492	-0.48087588
HDI_2006	-0.301268607	0.78332084	0.86998785	0.87868643
HDI_A1Reconcil_Rank	0.301322252	-0.78335798	-0.86975056	-0.87867868
PoliticalRights	0.238795444	-0.24376097	-0.53304130	-0.52775490
CivilLiberties	0.253579050	-0.32802901	-0.56538314	-0.57763833
ElectoralProcess	-0.231405497	0.21199600	0.50561786	0.48754296
PoliticalPluralism	-0.222536415	0.22059275	0.52428097	0.51863372
GovtFunction	-0.266235441	0.25177190	0.53495944	0.53630131
FreedomExpression	-0.263881177	0.24473839	0.54515082	0.52927285
AssocRights	-0.289443046	0.23055692	0.51493520	0.51217872
RuleLaw	-0.267498165	0.31520924	0.54204285	0.56745243
IndivRights	-0.249489918	0.44086844	0.63902938	0.65379977
FreedomStatus	0.202481017	-0.23085994	-0.49623731	-0.47999164
DISAP2005	-0.208024651	0.17478155	0.26815289	0.26020411
Progress_ST	0.444662631	-0.20245878	-0.37089118	-0.38299745
GDP	-0.363772452	0.64370120	0.78595012	0.77078992
HDI_Trends_Rank	0.309781939	-0.78187054	-0.87126935	-0.87990129
GDPpercap	-0.364670962	0.64671528	0.79162389	0.77414274
LifeBirth	-0.143685490	0.63511461	0.76990497	0.76786393
LifeExp	-0.143950403	0.63477045	0.76971710	0.76755524

EarnIng_Female	-0.404529572	0.69960511	0.81314903	0.81711841
Seats_Women	-0.196966648	0.06186757	0.22008660	0.25211164
Admin_Women	-0.426243305	0.33656010	0.31312320	0.28434736
Prof_Women	-0.458774894	0.72111654	0.39774960	0.53006250
EarnIng_Male	-0.336229414	0.63933665	0.79259336	0.76917940
OLD_EMPINX1990	-0.402023852	0.30970633	0.56010413	0.54606318
SPEECH1990	-0.343524337	0.29745039	0.48052365	0.48764977
ELECSD2007	-0.067920834	0.02979096	0.35425991	0.31206605
Progress_LT	0.791988075	-0.28785643	-0.31031623	-0.40993863
Progress_MT	1.000000000	-0.42781082	-0.32816473	-0.43873754
ALR	-0.427810818	1.00000000	0.76689011	0.95864218
EdEnroll	-0.328164726	0.76689011	1.00000000	0.92813212
Edu	-0.438737541	0.95864218	0.92813212	1.00000000
WOSOC2003	-0.207658142	0.47095063	0.58011156	0.62243956
WECON2003	-0.221221104	0.36734596	0.48800403	0.52185017
WOPOL2006	-0.279218525	0.03244329	0.21021147	0.24374177
COW	0.207228959	-0.22311851	-0.32824409	-0.31774911
POLITY	0.207228959	-0.22787155	-0.32987344	-0.31991744
UNSUBREG	-0.244513363	0.45981070	0.41477683	0.52354723
Unimprov_Water	0.103882909	-0.71180189	-0.68517154	-0.72342384
LifeBirth_Female	-0.173785850	0.68045771	0.77700460	0.79999632
LifeBirth_Male	-0.084527411	0.57825863	0.71929312	0.72394323
AdultLit_Female	-0.461882532	0.98795545	0.79562163	0.95333856
AdultLit_Male	-0.377919197	0.97792447	0.75787133	0.93429857
EnrollEduc_Female	-0.305663675	0.80426747	0.98913223	0.93424261

EnrollEduc_Male	-0.294358015	0.76212639	.97913605	0.91862005
Surv_40	0.007564762	-0.67562759	-0.72025543	-0.72516882
Poverty_200	0.081558116	0.74751902	-0.77429968	-0.77325255
POLPRIS1993	-0.043714061	-0.07480341	0.04188898	-0.02770853

	WOSOC2003	**WECON2003**	**WOPOL2006**	**COW**
Failing_Rank	-0.44952592	-0.46219319	-0.167705119	0.24864329
AdultLit_2005	-0.19484411	-0.20606793	-0.155329317	0.12362177
EnrolEduc_2005	0.04835568	0.06347662	-0.073118128	-0.05378540
GDPpercap_2005	0.05250721	0.09747813	-0.156219810	-0.19237699
AdultLit_2006	-0.29256171	-0.25238499	-0.231356829	0.24906459
HDI_2006	0.58260757	0.53525925	0.163696811	-0.32021496
HDI_A1Reconcil_Rank	-0.58216745	-0.53561383	-0.163358018	0.31975144
PoliticalRights	-0.51854926	-0.40048331	-0.283918798	0.43159604
CivilLiberties	-0.53657407	-0.47706851	-0.263377994	0.40458068
ElectoralProcess	0.49574777	0.36419705	0.323387027	-0.47683045
PoliticalPluralism	0.52192920	0.42172581	0.289830416	-0.44038776
GovtFunction	0.51959383	0.45177245	0.319381565	-0.40020955
FreedomExpression	0.51619965	0.42449170	0.268546492	-0.43980920
AssocRights	0.49893105	0.41038567	0.296138942	-0.44038177
RuleLaw	0.51989061	0.49748644	0.281660408	-0.33005092
IndivRights	0.58603883	0.52978545	0.308042765	-0.39992979
FreedomStatus	-0.47310460	-0.36614451	-0.250107974	0.40967098
DISAP2005	0.21412116	0.27278146	0.068075463	-0.17106958
Progress_ST	-0.23732552	-0.09658454	-0.316652805	0.22818774

GDP	0.46297374	0.45873767	0.157701313	-0.29837465
HDI_Trends_Rank	-0.58216745	-0.53721201	-0.170307815	0.32281971
GDPpercap	0.46298647	0.45832303	0.158437312	-0.29735985
LifeBirth	0.50605863	0.50723892	0.139286354	-0.33543274
LifeExp	0.50630742	0.50647449	0.139223710	-0.33605617
EarnIng_Female	0.53674096	0.49472439	0.194671124	-0.32135794
Seats_Women	0.31563100	0.27206528	0.604778945	-0.29705176
Admin_Women	0.21823908	0.18121932	0.283548146	-0.44005305
Prof_Women	0.30404428	0.19969313	0.307421476	-0.35137358
EarnIng_Male	0.47157502	0.45980453	0.113800429	-0.28235799
OLD_EMPINX1990	0.49251539	0.33522117	0.340760201	-0.45336005
SPEECH1990	0.37783498	0.21321660	0.286338389	-0.40182438
ELECSD2007	0.44614226	0.23521544	0.281130940	-0.38610086
Progress_LT	-0.38015983	-0.26003566	-0.510566056	0.37228063
Progress_MT	-0.20765814	-0.22122110	-0.279218525	0.20722896
ALR	0.47095063	0.36734596	0.032443289	-0.22311851
EdEnroll	0.58011156	0.48800403	0.210211471	-0.32824409
Edu	0.62243956	0.52185017	0.243741766	-0.31774911
WOSOC2003	1.00000000	0.58861929	0.266380310	-0.20251010
WECON2003	0.58861929	1.00000000	0.202627212	-0.11308328
WOPOL2006	0.26638031	0.20262721	1.000000000	-0.45789340
COW	-0.20251010	-0.11308328	-0.457893401	1.00000000
POLITY	-0.20237719	-0.11341916	-0.447453499	0.99999273
UNSUBREG	0.33823082	0.32748210	-0.088216223	0.17380078
Unimprov_Water	-0.36619890	-0.27642488	0.039763432	0.24633418

LifeBirth_Female	0.56308752	0.52017891	0.143488228	-0.33712566
LifeBirth_Male	0.49210614	0.51803166	0.074404337	-0.27554074
AdultLit_Female	0.47765532	0.41276705	0.075854979	-0.26410404
AdultLit_Male	0.46543166	0.43044057	-0.002795757	-0.19551815
EnrollEduc_Female	0.60667646	0.48877221	0.217017531	-0.35573554
EnrollEduc_Male	0.60346675	0.47171015	0.244185671	-0.34165320
Surv_40	-0.44860193	-0.39624077	0.160426244	0.14670704
Poverty_200	-0.47884870	-0.26578781	0.045208838	0.45346564
POLPRIS1993	-0.08513338	-0.07721172	-0.077233747	-0.01313319

	POLITY	**UNSUBREG**	**Unimprov_Water**	**LifeBirth_Female**
Failing_Rank	0.25118226	-0.24048531	0.497860223	-0.621695220
AdultLit_2005	0.12574111	-0.26868746	0.073335230	-0.164669871
EnrolEduc_2005	-0.05341920	-0.08429229	0.007225463	-0.033372626
GDPpercap_2005	-0.19237699	0.23042668	-0.285795748	0.324246198
AdultLit_2006	0.25102791	-0.26096317	0.250457585	-0.362991989
HDI_2006	-0.32312462	0.53579646	-0.818641841	0.944966912
HDI_A1Reconcil_Rank	0.32267004	-0.53589886	0.818749607	-0.945028365
PoliticalRights	0.43244481	-0.15806824	0.351939023	-0.551280260
CivilLiberties	0.40580153	-0.21357927	0.418011129	-0.592658639
ElectoralProcess	-0.47756970	0.12505431	-0.329004794	0.509506166
PoliticalPluralism	-0.44002721	0.15186514	-0.323131263	0.527116060
GovtFunction	-0.40029433	0.21971953	-0.388489008	0.586949170
FreedomExpression	-0.44047418	0.14879431	-0.351523310	0.520896137
AssocRights	-0.44177458	0.16477765	-0.335625261	0.523161888

RuleLaw	-0.33103997	0.23517621	-0.392918140	0.586472988
IndivRights	-0.40182301	0.28104275	-0.528129756	0.690726399
FreedomStatus	0.40869999	-0.11539730	0.346620679	-0.479856849
DISAP2005	-0.17106958	0.07847913	-0.187816516	0.277376205
Progress_ST	0.22818774	-0.11467965	0.243267313	-0.348840356
GDP	-0.30016333	0.49951881	-0.753166914	0.842526913
HDI_Trends_Rank	0.32575396	-0.53426296	0.819719911	-0.945028365
GDPpercap	-0.29915902	0.50038630	-0.753579080	0.844479561
LifeBirth	-0.33939192	0.43119729	-0.759967685	0.973600030
LifeExp	-0.34001246	0.43050861	-0.759766698	0.973512113
EarnIng_Female	-0.32420799	0.51639909	-0.741404772	0.842421889
Seats_Women	-0.29467171	0.02490924	-0.079555154	0.211445138
Admin_Women	-0.43690932	-0.13075319	-0.078664586	0.071581215
Prof_Women	-0.34906891	0.10268971	-0.343156397	0.146383852
EarnIng_Male	-0.28454578	0.51751274	-0.761443913	0.847559452
OLD_EMPINX1990	-0.45336005	0.14366330	-0.302110165	0.523285508
SPEECH1990	-0.40182438	0.14071129	-0.206916049	0.447335154
ELECSD2007	-0.38870418	0.01704218	-0.139698178	0.351683140
Progress_LT	0.37228063	-0.03583097	-0.146733940	-0.255435437
Progress_MT	0.20722896	-0.24451336	0.103882909	-0.173785850
ALR	-0.22787155	0.45981070	-0.711801887	0.680457711
EdEnroll	-0.32987344	0.41477683	-0.685171545	0.777004600
Edu	-0.31991744	0.52354723	-0.723423839	0.799996316
WOSOC2003	-0.20237719	0.33823082	-0.366198897	0.563087523
WECON2003	-0.11341916	0.32748210	-0.276424885	0.520178914

WOPOL2006	-0.44745350	-0.08821622	0.039763432	0.143488228
COW	0.99999273	0.17380078	0.246334180	-0.337125659
POLITY	1.00000000	0.17292447	0.251605004	-0.340739787
UNSUBREG	0.17292447	1.00000000	-0.413491279	0.456023246
Unimprov_Water	0.25160500	-0.41349128	1.000000000	-0.768605053
LifeBirth_Female	-0.34073979	0.45602325	-0.768605053	1.000000000
LifeBirth_Male	-0.27979946	0.43086550	-0.733549774	0.964120746
AdultLit_Female	-0.27008632	0.40519893	-0.686108232	0.686362803
AdultLit_Male	-0.20319155	0.45978534	-0.663447142	0.676071346
EnrollEduc_Female	-0.35707027	0.44728070	-0.715088308	0.784095526
EnrollEduc_Male	-0.34342533	0.45027903	-0.659576416	0.765674889
Surv_40	0.15333332	-0.43544862	0.772388458	-0.963847399
Poverty_200	0.45346564	-0.35442418	0.820479393	-0.837943852
POLPRIS1993	-0.01233003	-0.12073769	0.050677702	-0.008748074

	LifeBirth_Male	**AdultLit_Female**	**AdultLit_Male**
Failing_Rank	-0.60210639	-0.34737045	-0.303016305
AdultLit_2005	-0.11294315	-0.25754476	-0.278295726
EnrolEduc_2005	-0.07353114	0.03437338	0.042290129
GDPpercap_2005	0.32477468	0.25850347	0.319651425
AdultLit_2006	-0.28515211	-0.49747863	-0.478519052
HDI_2006	0.90762711	0.78116429	0.774360538
HDI_A1Reconcil_Rank	-0.90773034	-0.78126913	-0.774483085
PoliticalRights	-0.50367248	-0.31299314	-0.208135977
CivilLiberties	-0.54320037	-0.40503189	-0.304880291

ElectoralProcess	0.45627579	0.27375188	0.176429674
PoliticalPluralism	0.48775935	0.29396018	0.195031270
GovtFunction	0.54711491	0.31230360	0.209660277
FreedomExpression	0.47604969	0.33988979	0.230704308
AssocRights	0.46576345	0.29741171	0.209886819
RuleLaw	0.54738390	0.39618668	0.297066927
IndivRights	0.65031636	0.51158834	0.434626371
FreedomStatus	-0.43232995	-0.29893103	-0.200988635
DISAP2005	0.27163446	0.18938942	0.158262655
Progress_ST	-0.34129676	-0.24519740	-0.142696589
GDP	0.81193364	0.65043879	0.652076244
HDI_Trends_Rank	-0.90773034	-0.78126913	-0.774483085
GDPpercap	0.81686771	0.65083212	0.657824755
LifeBirth	0.97839665	0.64332259	0.642841458
LifeExp	0.97844064	0.64333636	0.642619014
EarnIng_Female	0.79425037	0.70126492	0.688659251
Seats_Women	0.17607193	0.11062159	0.084503479
Admin_Women	-0.04876516	0.48490992	0.298121065
Prof_Women	-0.03498290	0.76123810	0.662498295
EarnIng_Male	0.82163090	0.63121611	0.639344156
OLD_EMPINX1990	0.46373397	0.36739624	0.228504628
SPEECH1990	0.40611520	0.33679783	0.269260675
ELECSD2007	0.30260348	0.11260743	0.027820554
Progress_LT	-0.20671374	-0.35246247	-0.130852461
Progress_MT	-0.08452741	-0.46188253	-0.377919197

ALR	0.57825863	0.98795545	0.977924466
EdEnroll	0.71929312	0.79562163	0.757871330
Edu	0.72394323	0.95333856	0.934298575
WOSOC2003	0.49210614	0.47765532	0.465431660
WECON2003	0.51803166	0.41276705	0.430440575
WOPOL2006	0.07440434	0.07585498	-0.002795757
COW	-0.27554074	-0.26410404	-0.195518151
POLITY	-0.27979946	-0.27008632	-0.203191549
UNSUBREG	0.43086550	0.40519893	0.459785342
Unimprov_Water	-0.73354977	-0.68610823	-0.663447142
LifeBirth_Female	0.96412075	0.68636280	0.676071346
LifeBirth_Male	1.00000000	0.58399850	0.592186928
AdultLit_Female	0.58399850	1.00000000	0.944558859
AdultLit_Male	0.59218693	0.94455886	1.000000000
EnrollEduc_Female	0.72150993	0.83267605	0.777986944
EnrollEduc_Male	0.70876193	0.78295499	0.760698974
Surv_40	-0.95364803	-0.65647340	-0.678031206
Poverty_200	-0.76409930	-0.74054700	-0.753288865
POLPRIS1993	-0.03345465	-0.09169500	-0.127909467

	EnrollEduc_Female	**EnrollEduc_Male**	**Surv_40**	**Poverty_200**
Failing_Rank	-0.584487855	-0.539368391	0.459330231	0.44574928
AdultLit_2005	-0.242656142	-0.251567513	0.041037876	0.13835996
EnrolEduc_2005	0.004910193	0.016224710	0.020338634	-0.15754691
GDPpercap_2005	0.333301306	0.334829837	-0.326344043	-0.35575199

AdultLit_2006	-0.368949324	-0.360549539	0.230325148	0.32868043
HDI_2006	0.875091255	0.855759263	-0.912202239	-0.90073305
HDI_A1Reconcil_Rank	-0.874780774	-0.855338871	0.912346423	0.90077752
PoliticalRights	-0.562310994	-0.556681752	0.303058505	0.36289012
CivilLiberties	-0.599008441	-0.575834930	0.377497673	0.43846568
ElectoralProcess	0.537381411	0.532768488	-0.250483394	-0.37244534
PoliticalPluralism	0.552250743	0.553146183	-0.280514121	-0.35086536
GovtFunction	0.570115864	0.553605437	-0.330858260	-0.33006409
FreedomExpression	0.580387235	0.558010221	-0.277052760	-0.36027229
AssocRights	0.547044694	0.542548239	-0.265769869	-0.32478002
RuleLaw	0.569474757	0.554376602	-0.366238058	-0.37444466
IndivRights	0.663271666	0.654744983	-0.509423077	-0.57783687
FreedomStatus	-0.518025279	-0.499004394	0.279805034	0.34788749
DISAP2005	0.268548131	0.226184800	-0.233000949	-0.16810924
Progress_ST	-0.363403320	-0.373896956	0.117368340	0.10815851
GDP	0.785921633	0.770073414	-0.777336240	-0.87279624
HDI_Trends_Rank	-0.874780774	-0.855338871	0.912840903	0.90077752
GDPpercap	0.791900873	0.775606811	-0.783884048	-0.87267631
LifeBirth	0.776143312	0.759869397	-0.971049547	-0.81007999
LifeExp	0.775761366	0.759762466	-0.970902801	-0.80958456
EarnIng_Female	0.817479968	0.806546867	-0.745920658	-0.85819232
Seats_Women	0.241279557	0.231503934	0.058021832	0.03358303
Admin_Women	0.354479909	0.243843228	0.119456396	-0.15869753
Prof_Women	0.442568988	0.317785114	-0.081312738	-0.44135115
EarnIng_Male	0.792335570	0.779312134	-0.781901002	-0.87267238

OLD_EMPINX1990	0.616408646	0.597045958	-0.169959381	-0.41974258
SPEECH1990	0.506270707	0.522496939	-0.145733401	-0.35720897
ELECSD2007	0.372556567	0.379038930	-0.114942066	-0.17851424
Progress_LT	-0.351261079	-0.319631189	-0.135761321	-0.09956368
Progress_MT	-0.305663675	-0.294358015	0.007564762	0.08155812
ALR	0.804267466	0.762126386	-0.675627589	-0.74751902
EdEnroll	0.989132226	0.979136050	-0.720255435	-0.77429968
Edu	0.934242606	0.918620050	-0.725168824	-0.77325255
WOSOC2003	0.606676459	0.603466749	-0.448601931	-0.47884870
WECON2003	0.488772213	0.471710145	-0.396240771	-0.26578781
WOPOL2006	0.217017531	0.244185671	0.160426244	0.04520884
COW	-0.355735540	-0.341653198	0.146707043	0.45346564
POLITY	-0.357070267	-0.343425333	0.153333321	0.45346564
UNSUBREG	0.447280705	0.450279027	-0.435448617	-0.35442418
Unimprov_Water	-0.715088308	-0.659576416	0.772388458	0.82047939
LifeBirth_Female	0.784095526	0.765674889	-0.963847399	-0.83794385
LifeBirth_Male	0.721509933	0.708761930	-0.953648031	-0.76409930
AdultLit_Female	0.832676053	0.782954991	-0.656473398	-0.74054700
AdultLit_Male	0.777986944	0.760698974	-0.678031206	-0.75328887
EnrollEduc_Female	1.000000000	0.950588524	-0.734593928	-0.78947687
EnrollEduc_Male	0.950588524	1.000000000	-0.696624100	-0.77008772
Surv_40	-0.734593928	-0.696624100	1.000000000	0.84005624
Poverty_200	-0.789476871	-0.770087719	0.840056241	1.00000000
POLPRIS1993	0.002826840	-0.003617385	0.069253787	0.04562134

	POLPRIS1993
Failing_Rank	-0.0255898293
AdultLit_2005	0.1013429388
EnrolEduc_2005	-0.0436097234
GDPpercap_2005	-0.0626326501
AdultLit_2006	0.1181587651
HDI_2006	-0.0408279188
HDI_A1Reconcil_Rank	0.0407280549
PoliticalRights	-0.0810427293
CivilLiberties	-0.1007165983
ElectoralProcess	0.0960116461
PoliticalPluralism	0.0756022558
GovtFunction	0.0836678147
FreedomExpression	0.0842397884
AssocRights	0.0841320530
RuleLaw	0.1213520914
IndivRights	0.0637889728
FreedomStatus	-0.1134979799
DISAP2005	0.1662630290
Progress_ST	0.0046675815
GDP	-0.0374071747
HDI_Trends_Rank	0.0313121378
GDPpercap	-0.0366780385
LifeBirth	-0.0107427044
LifeExp	-0.0115395365

EarnIng_Female	-0.0709104240
Seats_Women	-0.1979880035
Admin_Women	0.1627498716
Prof_Women	-0.0843701437
EarnIng_Male	-0.0380039848
OLD_EMPINX1990	0.1666388810
SPEECH1990	0.1697991341
ELECSD2007	0.1104676872
Progress_LT	-0.0008901446
Progress_MT	-0.0437140614
ALR	-0.0748034120
EdEnroll	0.0418889821
Edu	-0.0277085323
WOSOC2003	-0.0851333812
WECON2003	-0.0772117227
WOPOL2006	-0.0772337466
COW	-0.0131331859
POLITY	-0.0123300264
UNSUBREG	-0.1207376868
Unimprov_Water	0.0506777018
LifeBirth_Female	-0.0087480741
LifeBirth_Male	-0.0334546492
AdultLit_Female	-0.0916950032
AdultLit_Male	-0.1279094666
EnrollEduc_Female	0.0028268399

EnrollEduc_Male	-0.0036173847
Surv_40	0.0692537874
Poverty_200	0.0456213392
POLPRIS1993	1.0000000000

Appendix 5

Boruta Results

	meanZ	medianZ	MinZ	maxZ	normHits	Decision
Earnings_Male	7.5653348	7.61574141	5.62706888	9.1626948	1	Confirmed
RuleLaw	7.2269491	7.24433719	5.52997138	8.8762238	0.988372093	Confirmed
HDI_2006	7.1866486	7.1866486	5.14013860	8.7759338	0.993023256	Confirmed
CivilLiberties	6.8686380	6.86319630	4.91997208	8.5750299	0.983720930	Confirmed
HDI_Trends_Rank	6.8536475	6.86042117	4.87929152	8.6574353	0.993023256	Confirmed
HDI_A1Reconcil_Rank	6.8452700	6.81835583	5.12679428	8.0973620	0.983720930	Confirmed
Unimprov_Water	6.6564019	6.67225343	4.51076947	8.8205913	0.986046512	Confirmed
GDP	6.6169581	6.62964347	4.46632684	8.3413279	0.990697674	Confirmed
GDPpercap	6.2931513	6.29969632	4.27561374	8.1365522	0.979069767	Confirmed
AdultLit_2005	6.2511449	6.25854794	3.09400413	8.7155516	0.972093023	Confirmed
OLD_EMPINX1990	6.2073087	6.19752575	4.38881210	8.1476327	0.981395349	Confirmed
SPEECH1990	6.0579632	6.08588915	4.08820884	8.1382888	0.981395349	Confirmed
AssocRights	5.9671259	5.99573419	4.27390405	7.9370283	0.974418605	Confirmed
FreedomExpression	5.7410892	5.76290319	3.49786690	7.3398607	0.951162791	Confirmed
LifeBirth_Male	5.6573027	5.68556470	3.98608757	7.3172314	0.960465116	Confirmed
Poverty_200	5.6490511	5.64273889	3.14818815	7.8794205	0.953488372	Confirmed
Progress_MT	5.5719758	5.57905495	2.32761442	8.8394059	0.925581395	Confirmed
GovtFunction	5.1842169	5.21516588	2.96044392	7.2070137	0.937209302	Confirmed
EarnIng_Female	5.1028746	5.11616863	3.02722566	6.6882073	0.923255814	Confirmed
LifeBirth_Female	4.9941838	4.98939272	3.30842873	6.5874214	0.918604651	Confirmed
IndivRights	4.6878620	4.69553060	2.87892855	6.0394965	0.879069767	Confirmed
Surv_40	4.6333056	4.68902042	2.44154473	6.5107425	0.853488372	Confirmed
PoliticalPluralism	4.5062036	4.56515515	2.24433515	6.3954653	0.858139535	Confirmed
Progress_LT	4.2859354	4.30972021	1.41009088	6.8003917	0.758139535	Confirmed
LifeBirth	4.2858806	4.28981423	2.32090565	5.7975827	0.788372093	Confirmed
LifeExp	4.3010242	4.28807632	2.43150986	5.8614181	0.811627907	Confirmed
WOSOC2003	4.1252955	4.15395539	1.77189964	6.4879093	0.748837209	Confirmed
Prof_Women	3.7511920	3.75191040	0.08118345	6.7700336	0.644186047	Confirmed
AdultLit_Female	3.6046336	3.63583989	0.55961701	5.9450258	0.595348837	Confirmed
EdEnroll	3.6333366	3.63400835	0.67176580	5.9917663	0.602325581	Confirmed
Edu	3.5193391	3.55918045	1.46595564	5.3803870	0.576744186	Confirmed
WECON2003	3.3732770	3.44016899	0.61099550	5.2298595	0.576744186	Confirmed
ALR	3.2495458	3.31780422	0.22254339	5.6346212	0.500000000	tentative
ELECSD2007	3.2311754	3.27815260	1.44269147	4.6636778	0.530232558	tentative
EnrollEduc_Female	3.0572417	3.08717811	-0.10882993	4.7724718	0.451162791	tentative
ElectoralProcess	3.0180520	3.02968905	0.61180720	4.6845675	0.441860465	tentative
UNSUBREG	2.9483912	2.97085340	-0.33628087	5.2695729	0.372093023	Rejected
EnrollEduc_Male	2.3966204	2.49947707	0.21855017	4.1032037	0.044186047	Rejected
Admin_Women	2.4125390	2.46138761	-0.03207636	5.0965314	0.046511628	Rejected
DISAP2005	2.3155786	2.42296135	-0.12588872	3.8365676	0.039534884	Rejected

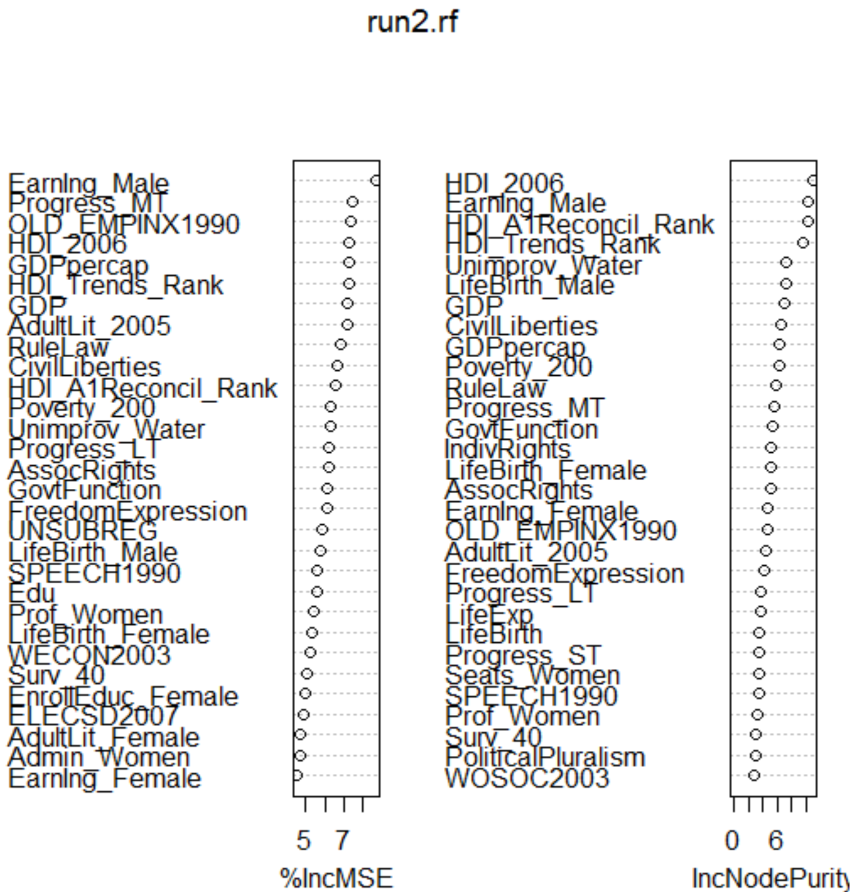
FreedomStatus	**2.3195119**	**2.39226313**	**1.34870207**	**3.4686590**	**0.023255814**	**Rejected**
PoliticalRights	**2.3078427**	**2.33610723**	**0.78141876**	**3.9903609**	**0.037209302**	**Rejected**
AdultLit_2006	**2.0403918**	**2.16197379**	**-0.49526028**	**3.8913776**	**0.030232558**	**Rejected**
Seats_Women	**2.0162893**	**2.15040620**	**-0.09587335**	**3.8431321**	**0.018604651**	**Rejected**
AdultLit_Male	**1.8612782**	**1.94466891**	**-0.69211584**	**3.6519175**	**0.020930233**	**Rejected**
GDPpercap_2005	**1.3826995**	**1.38390337**	**-0.33042109**	**2.7468396**	**0.000000000**	**Rejected**
POLITY	**1.3979751**	**1.36354068**	**-0.89501749**	**3.8497572**	**0.011627907**	**Rejected**
Progress_ST	**1.3538738**	**1.29000299**	**-0.39833326**	**3.2319258**	**0.009302326**	**Rejected**
COW	**0.9262780**	**0.98860094**	**-0.81726879**	**2.5140436**	**0.002325581**	**Rejected**
EnrolEduc_2005	**-0.1828905**	**-0.07425007**	**-2.61312323**	**1.5092043**	**0.000000000**	**Rejected**
POLPRIS1993	**-1.0636480**	**-0.92910357**	**-3.63456161**	**1.1545708**	**0.000000000**	**Rejected**

Appendix 6

Boruta results

EarnIng_Male	Confirmed
RuleLaw	Confirmed
HDI_2006	Confirmed
CivilLiberties	Confirmed
HDI_Trends_Rank	Confirmed
HDI_A1Reconcil_Rank	Confirmed
Unimprov_Water	Confirmed
GDP	Confirmed
GDPpercap	Confirmed
AdultLit_2005	Confirmed
OLD_EMPINX1990	Confirmed
SPEECH1990	Confirmed
AssocRights	Confirmed
FreedomExpression	Confirmed
LifeBirth_Male	Confirmed
Poverty_200	Confirmed
Progress_MT	Confirmed
GovtFunction	Confirmed
EarnIng_Female	Confirmed
LifeBirth_Female	Confirmed
IndivRights	Confirmed

Surv_40	**Confirmed**
PoliticalPluralism	**Confirmed**
Progress_LT	**Confirmed**
LifeBirth	**Confirmed**
LifeExp	**Confirmed**
WOSOC2003	**Confirmed**
Prof_Women	**Confirmed**
AdultLit_Female	**Confirmed**
EdEnroll	**Confirmed**


run2.rf
Earnlng_Male
Progress_MT
OLD_EMPINX1990
HDI_2006
GDPpercap
HDI_Trends_Rank
GDP
AdultLit_2005
RuleLaw
CivilLiberties
HDI_A1Reconcil_Rank
Poverty_200
Unimprov_Water
Progress_LT
AssocRights
GovtFunction
FreedomExpression
UNSUBREG
LifeBirth_Male
SPEECH1990
Edu
Prof_Women
LifeBirth_Female
WECON2003
Surv_40
EnrollEduc_Female
ELECSD2007
AdultLit_Female
Admin_Women
Earnlng_Female
5 7
%IncMSE
HDI_2006
Earnlng_Male
HDI_A1Reconcil_Rank
HDI_Trends_Rank
Unimprov_Water
LifeBirth_Male
GDP
CivilLiberties
GDPpercap
Poverty_200
RuleLaw
Progress_MT
GovtFunction
IndivRights
LifeBirth_Female
AssocRights
Earnlng_Female
OLD_EMPINX1990
AdultLit_2005
FreedomExpression
Progress_LT
LifeExp
LifeBirth
Progress_ST
Seats_Women
SPEECH1990
Prof_Women
Surv_40
PoliticalPluralism
WOSOC2003
0 6
IncNodePurity