Trusting the Government - the Asian Perspective

Introduction

We have selected Trust in National Government as the dependent variable that we attempt to explain using the social-demographic features based on the hypothesis that certain sections of society would be happy with the performance of the Government while few sections might not.

This information would help us to identify the sections of society which show distrust towards the government, and this could be used by groups concerned with democratic development, aligned with ABS' third main objective to disseminate survey results to academia and policy circles. One such use case would be for the government to identify the issues surrounding these groups and work upon policies towards their betterment.

We have come up with prediction models that use various variables, including the important variables given to us, and then identify the concerned citizen population. We have also run a Principal Component Analysis (PCA) to gauge which variables explain most of the data in each wave.

While we have the capability of running the model for each country, we have come up with a combined approach where we can see what factors influence the people in South-East Asia, which can be further used by different agencies to draw comparisons with the Rest of the Regions in the Global Barometer Survey.

Prior Work/Literature Review

While researching our Problem Statement, we came across many interesting articles, which talked about certain demographics and their trust in the Government. While these were informative, most of them focused on one particular country and one particular variable like Gender[1], so our target was to come up with a model that could combine socio-demographic variables and would work on multiple countries across different waves to provide a more accurate picture of the Trust in Government.

[1] Htun, M., Jensenius, F.R. Political Change, Women's Rights, and Public Opinion on Gender Equality in Myanmar. *Eur J Dev Res* 32, 457–481 (2020). https://doi.org/10.1057/s41287-020-00266-z

Proposed Methodology

Conducting EDA (see Appendix) for different variables in different countries, we got a basic understanding of the data and how people tend to vote and how satisfied they are with their governments.

Running a Principal Component Analysis is often helpful to understand the variables which account for most of the data variation. Hence we proposed running a PCA analysis before running our data models.

Data Modelling: Since our Target Variable is Trust in the Government, it is a categorical column with 4 values of Trust Level. Using this criterion, we have used a Multinomial Logistic Regression and a Random Forest Model to build a predictive model, since these are efficient for predicting our target variable. We have further used K-fold cross validation to check our OOS Accuracy. K-fold validation ensures that we aren't missing out on any data points for training and given the limited observations and implications of leaving out some training data, it is the best approach for validation in this scenario.

Analysis

After data cleaning steps and converting the variables to factors, we start with running a Principal Component Analysis (PCA) and try to establish the important variables which define most of our data. Here are the results of the top 3 loadings of the 2 waves:

PCA Results:

*Wave 1 (*Exhibit *1)*

- The **first loading** identified has high levels of trust in the government, parliament and courts. However, this group cohort does not trust the police.
- The **second loading** identified comprises poorer, older, non-single people with low trust in freedom of speech and their future economic condition.
- The **third loading** is made up of married individuals who believe freedom of speech and equality has improved.

Wave 2 (Exhibit 2)

- The **first loading** identified has high levels of trust in the government and courts however do not think that people are treated equally and have an equal say
- The **second loading** identified contains poor people with lower education levels who do not trust the government a lot
- The **third loading** identified contains protestors who voice their opinions via rallies and do not trust the government a lot.

Data Modelling Analysis:

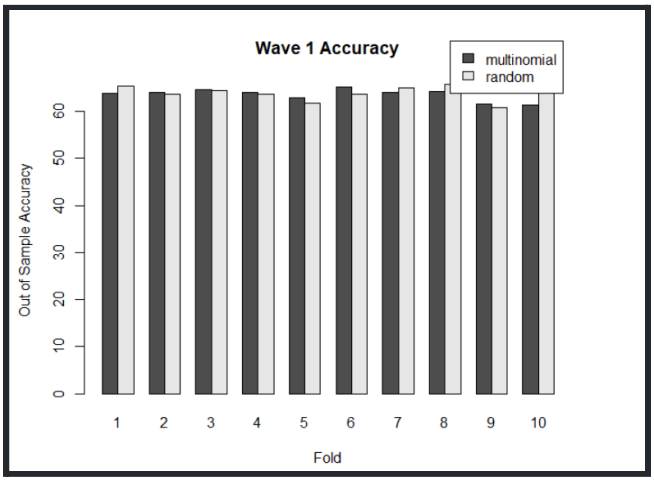
Models run:

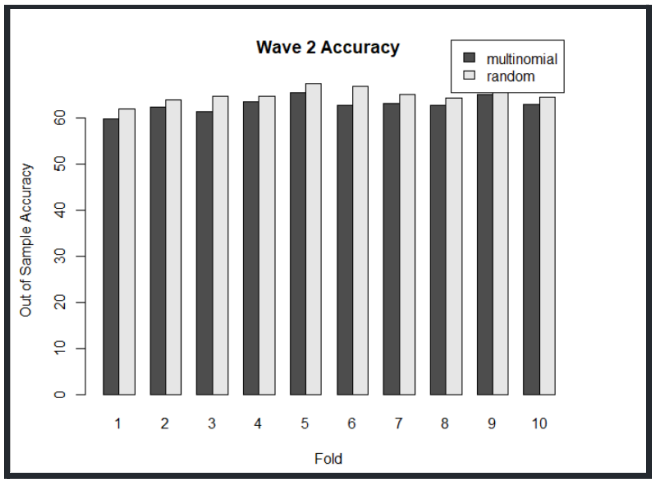
1. Multinomial Logistic Regression
2. Random Forest

Accuracy Validation: 10-fold Cross Validation

Model Performance, Out-Of-Sample Accuracy:

1. Wave 1
    a. Multinomial 63.58350
    b. Random Forest 63.88118
2. Wave 2
    a. Multinomial 62.74007
    b. Random Forest: 64.70992





Interpretation of results:

Wave 1 :

- As level of education increases, trust in governments reduce
- If people have faith and trust in the court systems and parliament, then that translates to an increased trust on the government
- If level of trust on the government is high, then freedom of speech can be an issue
- The level of trust on the government reduces when military presence in the government is high
- There is an inverse relationship between level of trust in police and level of trust in government
- In-sample accuracy for Wave 1 is 64.15551, hence we notice no overfitting.

Wave 2

- As proportion of males increase in a population, the level of trust on the government increases
- As a country becomes more educated, the level of trust on the government increases
- As income levels increase, the level of trust on the government increases
- Countries whose people have higher trust in the police, parliament and court system of a country tend to have a higher amount of trust on the government
- Countries with higher levels of freedom of speech generally trust the government
- The higher the number of people who believe everyone is treated equally, the more the trust in the government.
- In-sample accuracy for Wave 1 is 62.91142, hence we notice no overfitting.

We prefer Multinomial Logistic Regression because it gives us a more interpretable result, and hence a more actionable result. This would directly align with our goal to provide information which suggests what is causing people to lack trust in their Government.

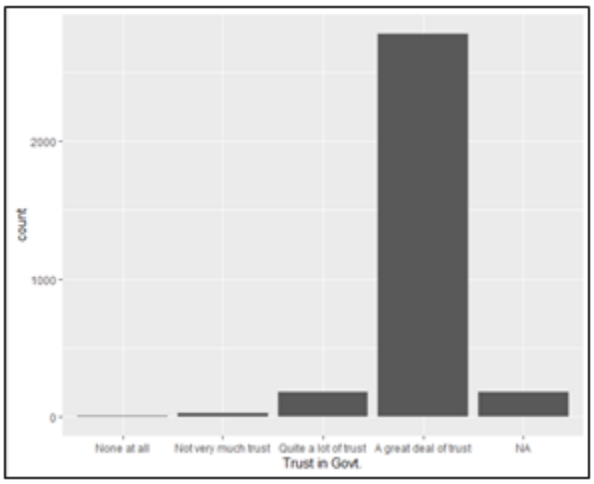
Conclusion

Through our analysis of the Asian Barometer datasets of Wave 1 and Wave 2, we have come up with functioning predictive models that have been run over all the countries for each of the waves. We have achieved approximately 64% accuracy with each of the models to predict our target variable, Trust in their Government. Using this model, we can predict whether a person might or might not be satisfied with the performance of their Government. Moreover, it gives us a starting point to discuss variation between other regions such as Europe and USA, and how their citizens' expectation of the Government differs from that of the Asian Countries. Using PCA we can also see where a large chunk of the population lies in terms of ideology and demographic features.

The models can also be run for individual countries, for governments to self-access and identify groups feeling distrust in the Government, and introduce policies to change this perception. Journalists and various NGOs and agencies can also shine light to these problems and influence Governments to recognize the problems faced by the citizens around them.

Appendix

Exploratory Data Analysis



Through Exploratory Data Analysis, we noticed that the data for Trust in Government in Mainland China was highly skewed, with a huge majority answering they have a ‘Great Deal of Trust’. Based on this observation we removed data of Mainland China.

We further used ggplot for univariate and bivariate plots and noticed some observations. In Thailand, women overall showed positive trust in the National Government, but we can observe an inverse trend between Trust and Education Level (higher the level of Education, lower the Trust in Govt.).

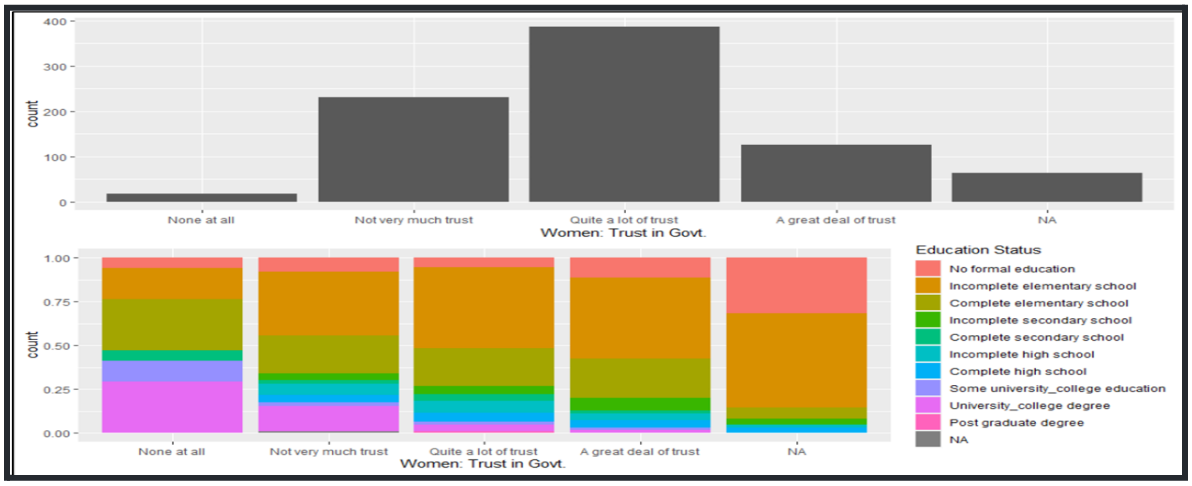


Exhibit 1 :

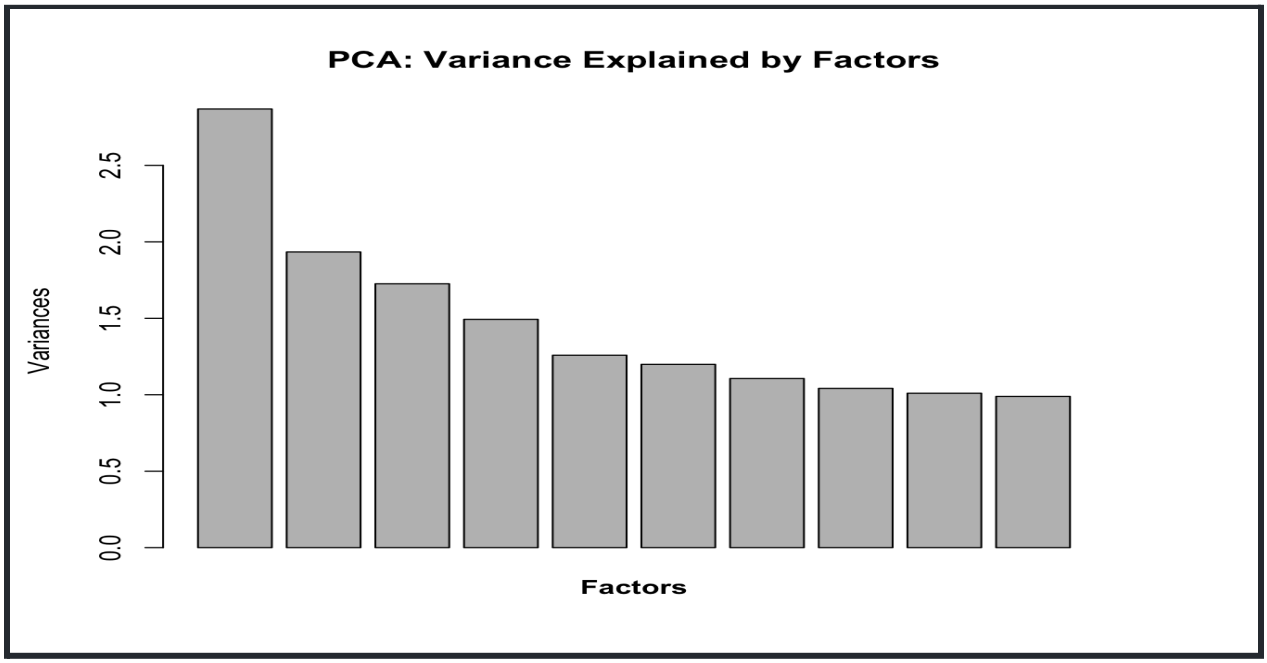


Exhibit 2: