**Background**

<https://www.unhcr.org/en-us/the-global-compact-on-refugees.html> - global compact overview

<https://www.unhcr.org/gcr/GCR_English.pdf> - global compact text

<https://www.unhcr.org/en-us/news/press/2020/4/5eaa991c4/unhcr-warns-looming-humanitarian-crisis-south-sudan-amidst-ongoing-fighting.html> - cv19

<https://www.unhcr.org/en-us/news/briefing/2020/4/5ea7ecd84/unhcr-shortfall-threatens-critical-aid-nearly-million-displaced-yemenis.html> - yemen

<https://jnslp.com/wp-content/uploads/2018/11/The_Border_and_Beyond_3.pdf>

**Introduction**

The complex global refugee crisis presents challenges for developing and developed countries alike. As persecution, climate change, conflict and poverty become more acute, millions of refugees are being forced to flee their homes and seek extended residence in peripheral countries. Some of these countries, facing their own economic and developmental issues, do not have the capacity or public will to support these refugees for significant periods of time, increasing instability and threatening a disjointed and potentially xenophobic response.

Natural disasters and conflict in countries such as South Sudan, Myanmar, Venezuela and Syria have produced millions of refugees that are left to live in austere conditions. With the advent of the Coronavirus, ravaged medical infrastructure and a decreasing availability of basic necessities, these refugees face a grim outlook as countries look inward to ameliorate their own public health crises. Aid organizations, which have historically stood up to provide humanitarian assistance, must navigate restricted transportation networks and a lack of funding, hamstringing their life-saving capacity.

International organizations, such as the United Nation’s High Commissioner for Refugees (UNHCR), were created for the sole purpose of responding to this refugee crisis. Through private donations and UN member nation burden-sharing, UNHCR seeks to: ease pressure on host countries, enhance refugee self-reliance, expand access to third country solutions and support setting conditions conducive for a safe return to host countries. As funding and involvement from member states decreases, UNHCR must effectively prioritize flashpoint locations to maximize their response options amidst increasing uncertainty.

**Analysis and Models**

Once they had the data from the UNCHR and Freedom House they needed to combine the two data sets together. In order to properly do this, they had to match them on a key. The key was country in this case and the data was merged in Excel. The Freedom House Index values were appended to the original data of the UNHCR. Once the tables were properly merged in Excel it needed to be cleaned in order to analyze for future modeling.

The rows that were either totals or did not have a country related were removed from the tables. Then the data was given proper headers to then be able to easily fix column types. All data initially was the variable type of character and needed to be changed to either factor or numerical. Country, Region, Free, Political Rights, and Civil Liberties were converted to factors while all other variables were made to be numerical.

Once all columns were associated to their proper type another data frame was created to make cuts within the data either to be binned. Binning was done with three to four bins and was based on the range of the data in each of the data. These separate data frames allowed for flexibility for modeling that will be shown in the below analysis.

**Results**

For the Association Rule Mining they initially looked to run rules on the tables that involved the information of where the asylum seekers were sending their applications.

lhs rhs support confidence lift

{2010=small, Rule of Law=mid-low} => {2014=small}    0.2 0.8 1.10

This rule shows that if the country had a small 2010 of applications submitted and low rule of law then the outcome would usually result in a small number of applications in 2014. This rule shows decent support and confidence. The lift is promising at 1.10 and shows that it can be generally dependent.

lhs rhs support confidence lift

{Free=F,Governance=mid-high}     => {2013=small}       0.3    0.8 1.07

The above rule shows that if the Freedom Index has indicated that the country is Free and their Governance is mid-high then they would have a small number of applications in 2013. The confidence is fairly high at 80 percent, but the lift again is showing that this is a rule that can be trusted.

lhs rhs support confidence lift

{2013=small-high,Liberties Score=F}   => {2012=small-high} 0.1    1    5

This rule shows that if the country had a small-high number of applications in 2013 and a Liberties score of Free then their 2012 was likely small-high. The small support and confidence of 1 is somewhat questionable in this case. The high lift at 5 too also can point that this rule is important, but should be taken with a grain of salt as there may have only be a handful of countries that follow this.

Next, they wanted to see if they looked at where the asylum seekers were coming from and if the results would differ from the above rules that were generated from where the applications were being put in.

lhs rhs support confidence lift

{2013=highest,PolPluralism=mid} =>  {PolRights=mid-low}    0.36  0.8    1.9

This rule is showing that if the PolRights is mid-low then it is likely that the country had a highest 2013 and mid PolPluralism. This is interesting to show what can make up some of the Freedom Index values from both Freedom Index and UNHCR variables. The support at 36 percent and confidence of 80 percent are good. The lift of 1.9 is also a good indicator that this rule can be trusted.

lhs rhs support confidence lift

{2014=highest,Expression=low}    =>    {Free=NF}        0.36  0.86    2.2

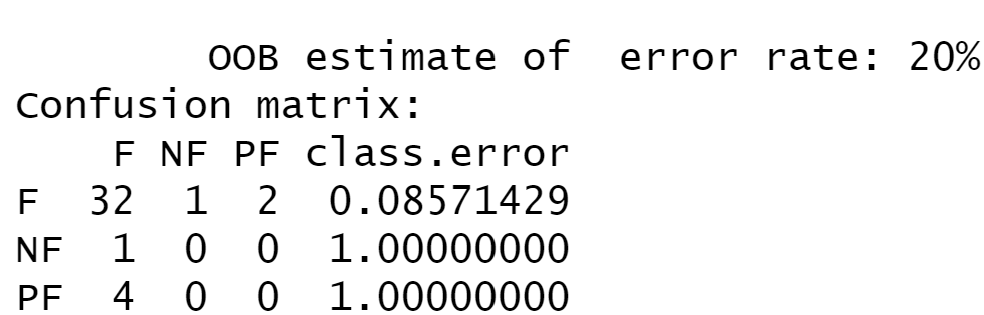
This rule is showing that if 2014 had a highest number of seekers from the country and a low Expression then the country would be considered Free. The support and confidence show strong reason to believe in this rule. The lift also supports both the support and confidence and show that the rule can be valid.

lhs rhs support confidence lift

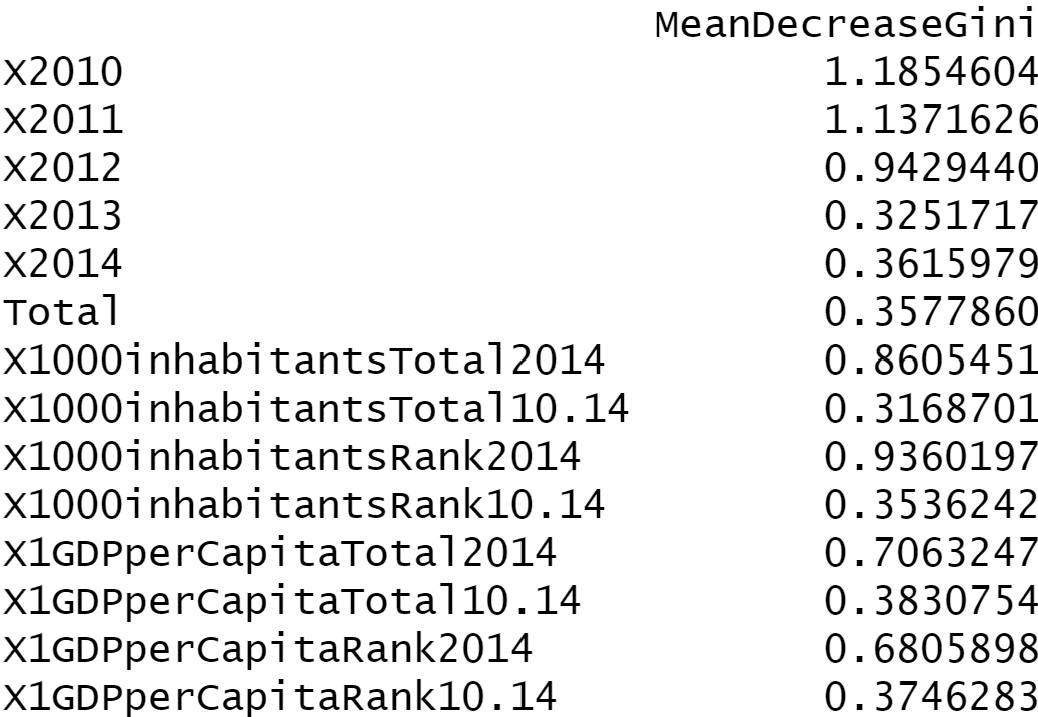
{Expression=low,Liberties Score=NF}    =>             {Governance=low}    0.36  0.86    1.9

This final rule is saying that if the Governance of the country is considered low then they have a strong chance of having a low Expression and a Not Free Liberties value. The support and confidence match the rule above and as previously discussed show good value to the rule. Then the lift is a bit lower than the previous but is still valid at 1.9 and shows strong promise to mark this rule as reliable.

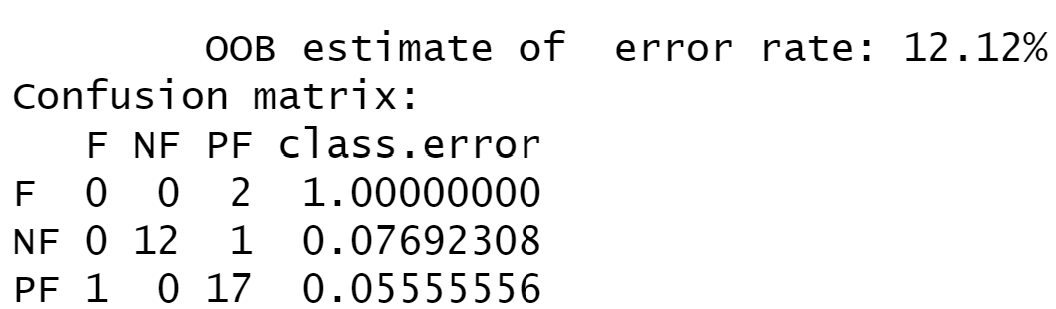
The final models that were run were RandomForest on both tables for where asylum seekers were seeking that asylum and also where thy originated from. The first model below is showing the RandomForest run on the first table of where asylum seekers were seeking the asylum. It was run to predict whether a country was labeled as Free. In order for the Freedom Index values not to show a perfect prediction those values were removed for the model.



The error rate of 20% is promising and can be taken that defining a Free country off of the first table is somewhat reliable. The importance of the variables below show that the asylum numbers from the years showed strong importance in the model. They would recommend to look into collecting more data from these countries by territories in them.



The next model was run on the table of where the applicants were coming from. They were more interested in running the model on this table as it could produce a better path forward when determining if a country is Free. As with the previous the Freedom Index values were taken out except for Political Rights and Civil Liberties. They determined these values important and good values for the UNCHR to populate.



This model shows a better prediction model than the model built on the previous data and used even less variables. The importance of the variables shows that the variables that come from the Freedom House are showing strong importance. This makes sense but the other variables are also encouraging as these are values that the UNHCR already have.



Draft – something like this will go in the final draft – we can do simple bullet points for the practice presentation.

**Motivation**

<https://www.humanrightsfirst.org/resource/refugees-and-national-security> - national security

An exacerbated refugee crisis is global security concern

Regional stability supports global prosperity

Reduced forced migration can stop the spread of contagious disease

I take motivation as “why do we care about doing this?”

**Problem Statement**

How can the team advise future humanitarian assistance operations by predicting/modeling/whatevering future refugee migration trends based off of historical data in support of minimizing global societal impacts?

Or

???

Kind of want to see what EDA looks like before we sign ourselves up for something crazy.