DOPP 2019W Exercise 3 - Group 32

Analysis of flows of refugees between countries

Contributors

- Eszter Katalin Bognar - 11931695
- Luis Kolb - 01622731
- Alexander Leitner - 01525882

Q1: What is the most accurate overview of flows of refugees between countries that can be obtained?

We selected 4 datasets to use:

- OECD International Migration Database data
- Gross Domestic Product per Capita data
- Human Development Index data (http://hdr.undp.org/en/data)
- World Governance Index data

First of all we clarified the term refugee: displaced person who has been forced to cross national boundaries and who cannot return home safely. Such a person may be called an asylum seeker until granted refugee status by the contracting state or the UNHCR.

According to that we decided to use the data from the OECD database, and collected the data for asylum seeker flows between countries.

Refugee flows are facilitated mostly by crime and hunger so we decided to add **GDP per capita** and **human development index**es to assess the wealth of the origin and source countries. Since crime and political instability also influence migration, we added the **world governance index** data to our dataset with the following values:

- Control of Corruption: Estimate
- Government Effectiveness: Estimate
- Political Stability and Absence of Violence/Terrorism: Estimate
- Rule of Law: Estimate
- Regulatory Quality: Estimate
- Voice and Accountability: Estimate

Putting all together we had a comprehensive overview of refugee flows.

Q2: Are there typical characteristics of refugee origin and destination countries?

The World Governance Indicators of origin countries negatively correlate with the number of asylum seekers originating from there - when some Indicators decrease, the number of refugees originating from that country tends to increase. Important to notice, HDI and GDP seemingly have no correlation with the number of refugees fleeing a country or choosing a country to flee to, according to these graphs.

Q3: Are there typical characteristics of large flows of refugees?

This question is similar to the one before it, and parts of this analysis here are relevant for the other question as well. Typically, large flows of refugees originate from countries with rather negative Indicators from the WGI, as well as lower HDI and GDP. Destination countries typically register a much higher GDP and HDI, as well as much better/higher Indicators from the WGI.

Another (very important) aspect to notice is the historical and geographical background. Iran and Afghanistan have recorded some of the highest numbers of refugees fleeing these countries between 2000-2017, and Turkey has the highest influx of refugees in the world in this time period. These countries are, first and foremost, very close geographically, Iran bordering on Turkey and Afghanistan bordering on Iran. It stands to reason that refugees will flee to the nearest, safe country, even if its WGI and other metrics are not as good or as high as some other, far away countries.

The same pattern is visible when looking at Serbia, producing about 500.000 refugees from 2000 onwards, and Hungary and neighboring states being among the countries with the highest refugee influx. Since we only use data from 2000 onwards, only a small part of the Yugoslav War (1991-2001) is represented here. Still, it is enough to push Serbia to be the country producing the third-most refugees in this time period.

Q4: Can countries that will produce large numbers of refugees be predicted? Can refugee flows be predicted?

First we have to prepare the data set “merge” to separate the columns that we need for the prediction. Then we could choose the prediction for any country that we want, for example Afganistan. After some closer observation the number of refugees from that country are the highest ones from the data set, coincidentally.
We used a linear regression to predict the data for 2017. For that we use the previous data from 2000-2016. At the end it shows that the prediction is not that accurate
as expected. The accuressy is only about 10%. The reason for that is that the flow of refugees is very hard to predict. It could be at every time possible that a war could break out and then the number of refugees will rise immediately.