

PROJECT PORTFOLIO DATA ANALYSIS ON TRAFFICKING IN HUMAN BEINGS

Capstone Project Summary

Submitted by: Hanna Asipovich
Supervisor: Zoltan Toth

Submitted to Central European University
Department of Economics and Business

In partial fulfilment of the requirement for the degree of Master of Science in Business Analytics

This is the summary of the capstone project completed for a non-profit organization (*the Client*) to carry out data research and data analysis for the project portfolio. The subject matter is related to the trafficking in human beings, covers a specific region and certain issues of interest.

1. Scope of Work

I completed the assignment by carrying out data research, data cleaning and transformation, undertaking an initial analysis, identifying additional key data points, and preparing visualisations in Tableau to highlight regional trends and patterns. Based on policy work experience and to address an emerging request by the Client, I developed an additional product, a monitoring tool with both qualitative and quantitative analytical sections. In addition, I carried out tailored training on data use in policy work for the staff to present initial findings and to encourage further data exploration.

2. Data Sources and Characteristics

All data came from open sources - UNODC, US Department of State, World Bank, and other secondary sources. All data is freely available and based on the submitted reports by public authorities. Data is oftentimes not cleaned or transformed for the purposes of the research. Cleaning and enriching the data would be useful for the purpose of research and further work on eliminating trafficking in human beings. To tailor the analysis to the Client's needs I worked directly with domain experts, which also allowed us to identify some faulty information and even fix coding of an original database. The corrected version is now available for other researchers in the field.

3. Data Processing

The original database included 46, 556 observations, which were further filtered to 11,574 observations to correspond to the geographic area and timeseries span of 2017-2021. Data processing pipeline codes are available on Github. For example, countries of the region were filtered and their names were fixed, the coded "<5" data was averaged, non-reporting countries values were included at "0", additional data by the World Bank was included to normalise the number of victims by the million of population. All data was presented in the long format for it to be later used in the draft monitoring tool which is Excel based.

4. Data Exploration and Visualisations

For data visualizations Tableau was used to prepare visuals around high level statistics for easier perception and manipulation for the show case for the Client. A variety of graphs were used: histograms, trend and pie charts, maps. The charts were organized in dashboards to address specific issues, e.g. trends overtime, specific data on number of victims and normalized numbers, gender specific information.

5. Findings

Based on descriptive statistics and analysis of the processed data the following findings were made and shared with the Client:

- **There has been a continuous increase in the number of offences and the number of detected victims of trafficking** since 2017 till 2021, the timeframe of the analysis.
- At the same time there was **a clear drop in the overall number of prosecutions and convictions**, which can be explained by more prolonged trial procedures than victim detection

and the fact that in 2020 access to justice was affected by COVID 19 health emergency. E.g. especially looking at the significantly lower numbers in this dimension in the time series provided by USA, which affected the total count.

- **In prosecutions and convictions France leads the way**, however, if excluded than one can see that USA, Romania, Belgium and Uzbekistan had quite good standing on judicial recourse.
- The United Kingdom, the United States, Ukraine, France and the Netherlands **are the top five in the total number of victims**. However, with normalisation, one can see that Montenegro, Bulgaria, the Netherlands, Norway and Finland have high numbers of **victims by one million of population**.
- **The countries, outliers, both in the high total number and the number of victims by a million of population are:** Ukraine, France, the Netherlands, Romania and Bulgaria. These are the countries most susceptible to trafficking and should be the priority in program interventions.

6. Follow Up Questions

Upon presenting the results to the Client, further questions were developed by the staff to be addressed in future research. More granulated data would be needed to, for example, identify recruitment practices online, transit routes, composite variables influencing victims and perpetrators of THB, correlations identified through data obtained in counter-terrorism/anti-money laundering law enforcement efforts.

7. Summary

Complete elimination of trafficking in human beings is one of the UN sustainable development goals. As countries of the region are moving forward to counter this criminal activity, statistics and data are necessary to analyse the trends and address this problem based on geographic and thematic evidence. This analysis contributed by extracting relevant data and providing high level data analytics to the Client based on their request, finally, building a tailored inhouse tool for further monitoring. I used freely available databases and raw data available there, built data processing pipelines, cleaned and transformed the data, carried out high level descriptive statistical analysis and developed visualisations in Tableau.

The data analysis identified an upward trend in the number of trafficking victims, looked into the countries with the highest numbers of victims both in absolute and normalised numbers, and countries which grant the best judicial recourse. Research can further benefit from more granulated data in order to address nuisances of the trafficking, e.g. geospatial data on transit routes and heatmaps for origin/destination, THB hotspots and proximity analysis, online recruitment and text analysis, qualitative analysis on presence and reform of THB laws, financial investigations and possible correlation with other criminal activities.