Refugee Data Analysis

# Introduction

This endeavor dives deep into the intricacies of refugee data, emphasizing the key countries that contribute to refugee numbers. By tapping into the power of data visualization tools and advanced algorithms, the study unravels nuanced patterns of refugee movements and their associated decisions.

# Data Overview

The dataset we're working with paints a vivid picture of the countries refugees come from, the sheer numbers, and the verdicts associated with their asylum requests. This data stretches from the dawn of the new millennium in 2000 up to 2016. With this rich dataset, we're able to craft visual tales, using maps to pinpoint application origins and the nations granting refuge.

# Backdrop

The globe grapples with the ever-growing refugee crisis. As regions face turmoil and strife, an increasing number of souls seek solace in safer shores. To grasp the contours of these migratory waves, a data-driven lens becomes indispensable. This analysis endeavors to shine a light on these migratory paths, harnessing the might of data and visual narratives.

# Exploratory Data Analysis (EDA)

1. Nations Leading the Refugee Count: A visual exploration probably lays out the top nations producing the highest refugee counts.  
2. Seventeen-Year Decision Distribution for Elite Five: This segment delves into the decision-making trends concerning asylum applications for the top quintet over a span of 17 years.  
3. Mapping the Migrations: A geographical spectacle, this section traces the roots of the asylum seekers and the nations that embrace them.

# Crafting the Data

The art of data transformation can elevate algorithmic performance. In this exploration:  
  
- Label Encoding: A method that grants numeric identities to categorical data, ensuring algorithms digest the data without losing categorical nuances.  
  
- Target variable construction: We summed 2 columns - ‘Total pending start-year’, ‘Applied during year’ into one column called ‘E+G’. We then proceeded to calculate ‘acceptance\_rate‘ column with ‘decisions\_recognised’ divided by the ‘E+G’ column. It was this column that was used as our target variable.

# Algorithmic Arsenal

Machine learning models stand as the pillars of our predictions using intel sklearnex:  
  
1. Polynomial Regression: Venturing beyond linear boundaries, this model crafts curves to fit data points, capturing the essence of non-linear data dynamics.  
  
2. Random Forest Regressor: A symphony of decision trees, this ensemble captures diverse data nuances. Its inherent design fends off overfitting.  
  
3. XGBoost Regressor: A shining star in the gradient boosting realm, XGBoost is revered for its swift executions and stellar model outcomes.

4.) Artificial neural network – The constructed NN consisted of 3 layers (including the output layer). This is made using the Intel OneAPI OneDNN library.

# Performance Metrics

A trio of metrics provide a window into model prowess:  
  
- RMSE: A metric offering insights into prediction accuracy by quantifying the average squared deviations between predictions and reality.  
  
- MAE: A direct measure, MAE lays out the average prediction deviations, presenting the error magnitude plainly.  
  
- R^2: A metric that encapsulates the data variance decoded by the model. A higher value indicates a model that's in sync with the data's rhythms.

# Proposed model: Our proposed model is the artificial network with a validation loss of 0.1382.

# OneAPI's Strengths

OneAPI, a unified programming beacon, offers several merits:  
  
1. Harmonized Programming: OneAPI bridges diverse hardware architectures under one umbrella.  
  
2. Performance Boost: Leveraging hardware accelerators, OneAPI amplifies execution speeds, crucial for voluminous datasets.  
  
3. Evolving Flexibility: Future-proofed by design, OneAPI welcomes diverse accelerators.  
  
4. Tailored Libraries: Housing optimized libraries spanning various domains, OneAPI enhances performance seamlessly.  
  
5. Everywhere and Anywhere: OneAPI's code remains versatile, ready to run on any compliant hardware.

# Wrapping Up

Navigating the complex web of refugee movements, this analysis brings to light patterns and insights that were previously obscured. By sculpting the data and deploying potent algorithms, not only have we unraveled past trends, but we've also charted potential future trajectories. This knowledge holds the power to inform and guide global strategies, ensuring resources reach those who need them most.