

## Project Proposal – Megha Nanaki Parwani

### Missing Migrants Dataset

Accessed via: Kibble.com <https://www.kaggle.com/snocco/missing-migrants-project>

Methodology of data collection: <https://missingmigrants.iom.int/methodology>

#### Motivation for analyzing this data:

With violent conflicts erupting throughout the world in Yemen, Syria, Myanmar, Venezuela, and other places, there is an unprecedented rise in global migrations. Often the route to a destination country is dangerous, with thousands dying due to human and environmental obstacles. The number of deaths is not systematically recorded but the Missing Migrants dataset from Kaggle compiles many sources--from Coast Guards to NGOs-- to track migrant deaths when they occur. The dataset also tracks the gender of the victims above 18, the number of children killed, the continent/location of the death, the route along which the death occurred. While this data is grim, understanding the global spread, demographics and nature of these deaths could be a first step towards policy change or towards inspiring sympathy so that people can get more involved with refugee aid. Data visualization will make this data into more of a narrative that can move and educate people.

#### Data visualization options:

##### 1. Map of location of deaths of migrants (Heatmap)

This visualization will identify trends in which groups of migrants or which locations end up suffering the most fatalities in the course of migration. The different regions in the dataset are: North Africa, Europe, US-Canada, US-Mexico, Sub-Saharan Africa, Mediterranean. The number of deaths will be represented by colors of a gradient. I will follow the model of the 'World of Happiness' (Clementine Chou) map.

##### 2. 6 box charts of different migration routes, showing spread of fatality data for the different migration routes (Mediterranean, Eastern Africa, Western Africa, Northern Africa, Northern American, Central America (not same as locations from Pt1))

This visualization might help us analyze whether some routes are deadlier than others. Number of deaths will be on the y-axis and migration routes will be across the x-axis.

##### 3. Histogram: Genders of victims along different migration route: helps discern whether gender is more likely to suffer fatalities on certain routes.

This visualization will help us notice if there is a gender-based dimension to deaths. I will use two colors to represent the two genders, male and female. Number of deaths will be on the x-axis and migration routes will be across the y-axis.

**Commented [1]:** For the project proposal, submit a 1-2 paragraph description of your final project. You should provide details on the problem you wish to explore, any datasets that you will use, and potential visualizations. The proposal is a way for me to make sure that you have enough structure and data to complete the project. The project proposal should be submitted as a text file (any format will do). In implementing your final project, you should try to stick as closely to your proposal as possible. If you are unable to do so, please provide a paragraph in your final project to explain the deviations from the proposal.