Project Draft

Due: October 30, 11:59pm

```
library(tidyverse)
climate <- read_csv("data/climate_update.csv")

#climate2 <- data.frame(t(climate[-1]))
#colnames(climate2) <- climate[, -1]

#glimpse(climate2)
```

Introduction

As the world struggles to convince so many people about the urgency of climate change it needs to be known that the threat we face is not just that of rising sea levels or CO2 levels—it is that of losing our homes. It is that of entire cities having to uproot and move elsewhere because they can no longer sustain themselves. Far from just a small increase in temperature, but a disruption of our lives as we know it.

Each year, tens of millions of people are driven from their homes by floods, storms, and droughts. The Ecological Threat Register, conducted by The Sydney-based Institute for Economics and Peace (IEP), measures ecological threats over 157 independent states and territories. The report projects that as many as 1.2 billion people around the world could be displaced by 2050 (Institute for Economics and Peace, 2020). Moreover, adverse effects of global climate change will induce more extreme weather, growing food and water insecurity, and rising sea levels which will cause the number of displaced people to rise (UNHCR, 2019). The report additionally identifies three clusters of ecological hotspots: the Sahel-Horn belt of Africa, from Mauritania to Somalia; the Southern African belt, from Angola to Madagascar, and the Middle East and Central Asian belt, from Syria to Pakistan.

The intersection of climate change and migration requires comprehensive data analysis and solutions to the multidimensional challenges it creates (Podesta, 2019). Therefore, analyzing the dynamics between climate change indicators and displaced people not due to conflict can reveal opportunities for interventions.

Our primary goal in this project is to understand the correlation between climate change indicators and socioeconomic factors, and the resulting displacement. We would like to look at the environmental and socioeconomic data that might predict displacement. For instance, previous literature has indicated that variables such as rainfall, agricultural yield, and low-lying areas may be associated with internal displacement, and we would like to see if the data from the World Development Bank are consistent with these claims.

We hypothesize that the percentage of total population below 5 meters of elevation in island nations, the percentage of total population living in slums, poverty, and falling agricultural output will be the strongest correlates and predictors of displacement of people due to climate in following years in regions identified by The Ecological Threat Register.

Methodology

In order to analyze our data, we will run multiple linear regressions on internally displaced people associated with disasters and other variables that have been associated with displacement. We will look at regions of countries in the world in order to make the data analysis easier for us. We would like to see which variables are predictors of internal displacement. Seeing these predictors will allow us to easily see which variables are associated with high amounts of displacement due to disaster. Some of the variables that we will focus on

include annual rainfall, agricultural yield, and the percentage of land that is low-lying in the country. For some of the variables with fewer data points but are unlikely to change over the past 20 years, such as the percentage of land that is low-lying, we will turn these variables into binary categorical variables. This will allow us to use these variables in our data analysis.

Once we identify some of the better predictors of displacement, we would like to run two sample hypothesis tests to see if there are differences in certain variables in countries that are prone to displacement. Specifically, we could run two sample hypothesis tests between high- and low-displacement regions as mentioned in the introduction with relation to certain predictor variables. This way, we could see the extent to how different these variables are for these two classifications. Additionally, we would like to run a Chi-Squared test on the displacement due to disasters and displacement due to conflict. We would like to run this test because it can help us elucidate whether these displacements are related. We would run this test because, if there is a high amount of displacement due to conflict, this could affect the displacement due to disaster, so we would want to see if these variables are independent.

```
#displacement <- climate %>%
  #filter(`Series Name` == "Internally displaced persons, new displacement associated with disasters (n
  #filter(`Country Name` == "Vietnam") %>%
  #select(!(`Country Name` | `Country Code` | `Series Code`))

#climate_update = read.csv("data/climate_update.csv")


#climate2 <- data.frame(t(displacement[-1]))
#colnames(climate2) <- displacement[, -1]

#climate2 <- climate2 %>%
  #filter(X1 != ".." & X2 != "..")

#disp_lm <- lm(X1 ~ X2, data = climate2)

#tidy(disp_lm)
```

```
climateData <- climate %>%
  mutate(MiddleEastAndCentralAsianBelt = ifelse(`Country Name` == "Syrian Arab Republic" |
        `Country Name` == "Iraq" | `Country Name` == "Afghanistan" | `Country Name` == "Pakistan" |
        `Country Name` == "Iran, Islamic Rep.", 1, 0)) %>%
  mutate(SouthernAfricanBelt = ifelse(`Country Name` == "Angola" | `Country Name` == "Zambia" |
        `Country Name` == "Zimbabwe" | `Country Name` == "Mozambique" | `Country Name` == "Malawi" |
        `Country Name` == "Madagascar", 1, 0)) %>%
  mutate(SahelHornBelt = ifelse(`Country Name` == "Senegal" | `Country Name` == "Mauritania" |
        `Country Name` == "Mali" |`Country Name` == "Burkina Faso" | `Country Name` == "Niger" |
        `Country Name` == "Chad" | `Country Name` == "Sudan" | `Country Name` == "Eritrea", 1, 0))
```

Results

Discussion