Final Project Report
DSC 456
11/22/2021

Climate Change and Climate-Related Disasters

A. Introduction

As climate change continues to take its toll, scientists and scholars have argued for the case of protecting and advocating for the planet through conducting extensive research on the effects, causes, and outcomes of climate change. In result, a plethora of databases exist that encapsulate how the planet is continuing to decline by measuring toxic air pollution levels, rising temperatures, and sea levels.

The visualizations are based on the Climate Related Disasters Frequency database by the Center for Research on the Epidemiology of Disasters (CRED) which measures the frequency of climate disasters by country from the 1980's until 2020. There are 201 countries from both the Global North and Global South and 6 different types of disasters listed. In the analysis, I create visualizations to show the frequency of disasters by decade and country.

B. Exploratory Analysis

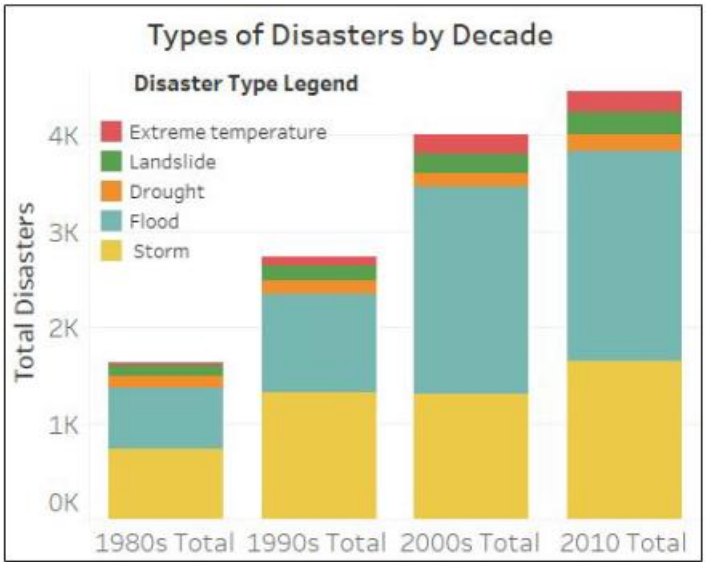
Before I discuss in detail about the exploratory analysis and other details of the dataset, I wanted to call out that this dataset is unique in a way that it had a wide format where every year had its own column. There could be numerous reasons why the source organization chose this sort of a format including using this as a summary table.

The dataset has records of Climate Related Disasters Frequency in a CSV file format YoY from 1980 to 2020. As mentioned earlier, the count of disasters is grouped by Type and By Year for all recorded countries. The dataset includes 948 rows and 48 columns, the dimensions included are the 1) **Object ID** – Unique Identifier of disaster 2) **Country** – Country Name 3) **Indicator** – Type of Disaster – Flood, Drought, Landslide etc. 4) **Code** – Single Code used across the dataset for all entries 5) **Unit** – Definition of metric count and 6) **Year** – 40 years spread across 40 columns, one column for each year from 1980 - 2020. The metric is the count of disasters by type by year.

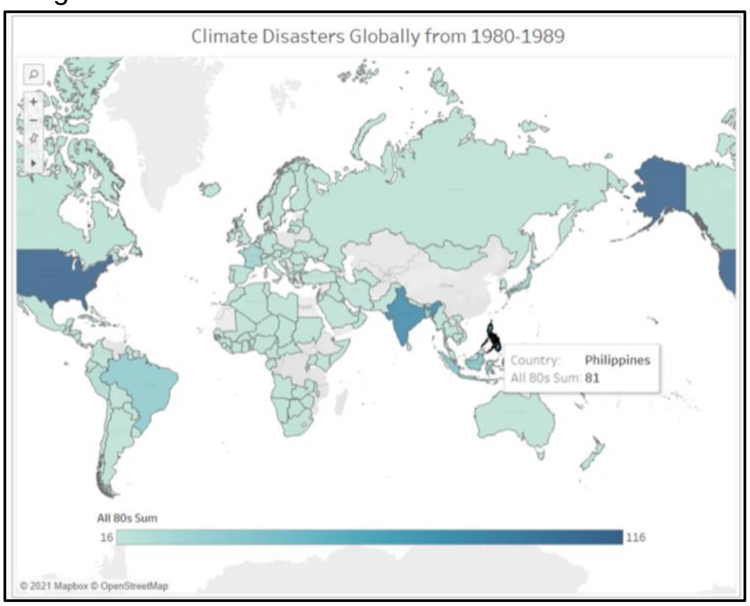
Source: International Disaster Database - https://www.emdat.be/

Final Project Report
DSC 456
11/22/2021

I deep dive into the data and see that I have year-wise data for 201 countries, 6 disaster types, and for the years 1980-2020. There is also a special dimension in the dataset, which contains the sum of all disasters for that country in that respective year. Looking at the data I can see from a first look that flood, and storms are major contributors to the total count of disasters. I observed an uptrend in the number of recorder Natural/Climate related disasters across the world Decade over Decade:



Similarly, on a quick look using tableau, I can see that in the 1980s, US, India, and the Philippines had approximately 10 times the disasters as compared to other countries in the globe.

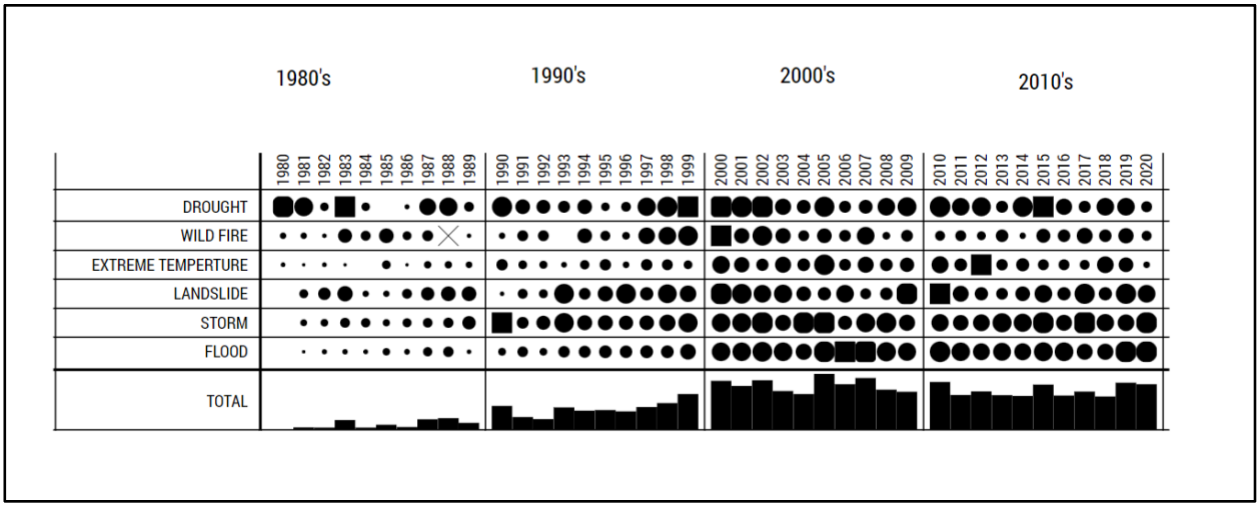


Final Project Report
DSC 456
11/22/2021

C. Visualizations

The following visualizations are created from the Climate Related Disasters Frequency database mentioned in the explanatory analysis.
Note: The following visualizations are to be graded as part of the final project.

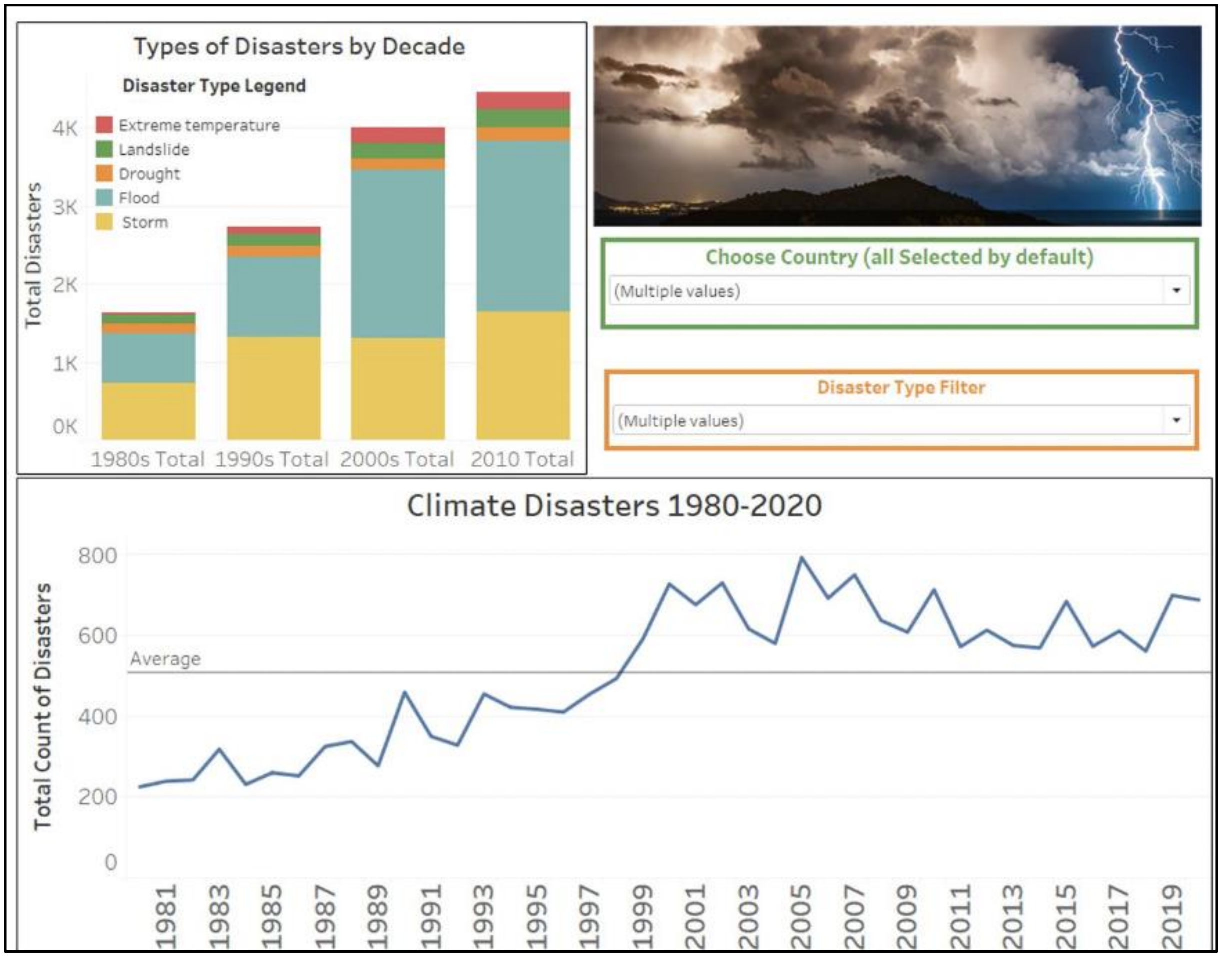
Bertin Matrix Displaying Disaster Frequency Over Time



The Bertin Matrix shown above, depicts the frequency of climate disasters by decade. The matrix includes the 6 types of natural disasters included in the dataset: drought, wildfires, extreme temperature, landslide, storms, and flood as well as the total number of climate disasters in each decade. Each decade is broken up by year, with a total of 40 years depicted. This visualization groups the frequency of similar disasters, however because each type of disaster is spread out throughout the decades, it is harder to see a more consistent and coherent grouping. What is apparent from this matrix is that the greatest frequency of climate change disasters was in the 2000's and 2010's with landslides, storms, and floods being the most frequent climate disasters. The following two visuals expand on this observation in two unique ways.

Final Project Report
DSC 456
11/22/2021

Interactive Dashboard for Visualizing Global Disasters in the last 40 years

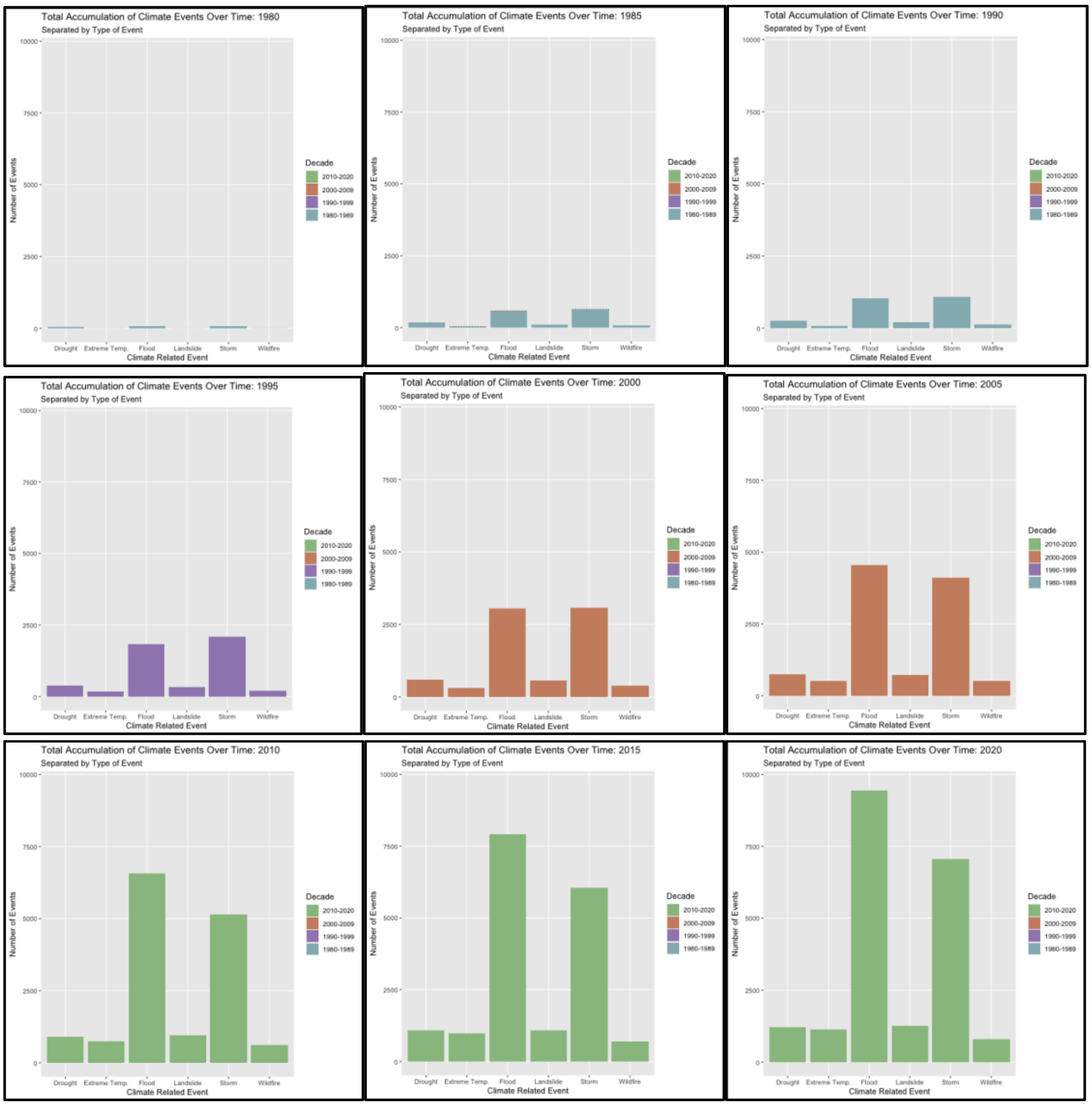


To further elaborate on changes in total amount of events and event type distribution, I developed the visualization above. This interactive dashboard was created for the audience, and it lets the audience select single or multiple countries and, single or multiple disasters. This helps the user easily deep dive into either a single country or multiple countries. However, if no selection is made, the dashboard delivers a global view without any filters for all disasters.

I had to be careful of a few key things while building this dashboard which were - Applying Filter to all worksheets on the dashboard *(as this is not the default setting on Tableau)* and choosing right visualizations to clearly show two different aspects of the data but also connect to tell a story. In the case, I am showing a YoY trend and, what is causing that trend (as in, which disasters).

This dashboard helps give the audience an overall picture and a trend, but also helps them understand what the factors behind this are. The two sheets work coherently to display an overall yet detailed story. Moreover, Interactivity provides flexibility to the users to play around the data and tweak the viz. as per their needs

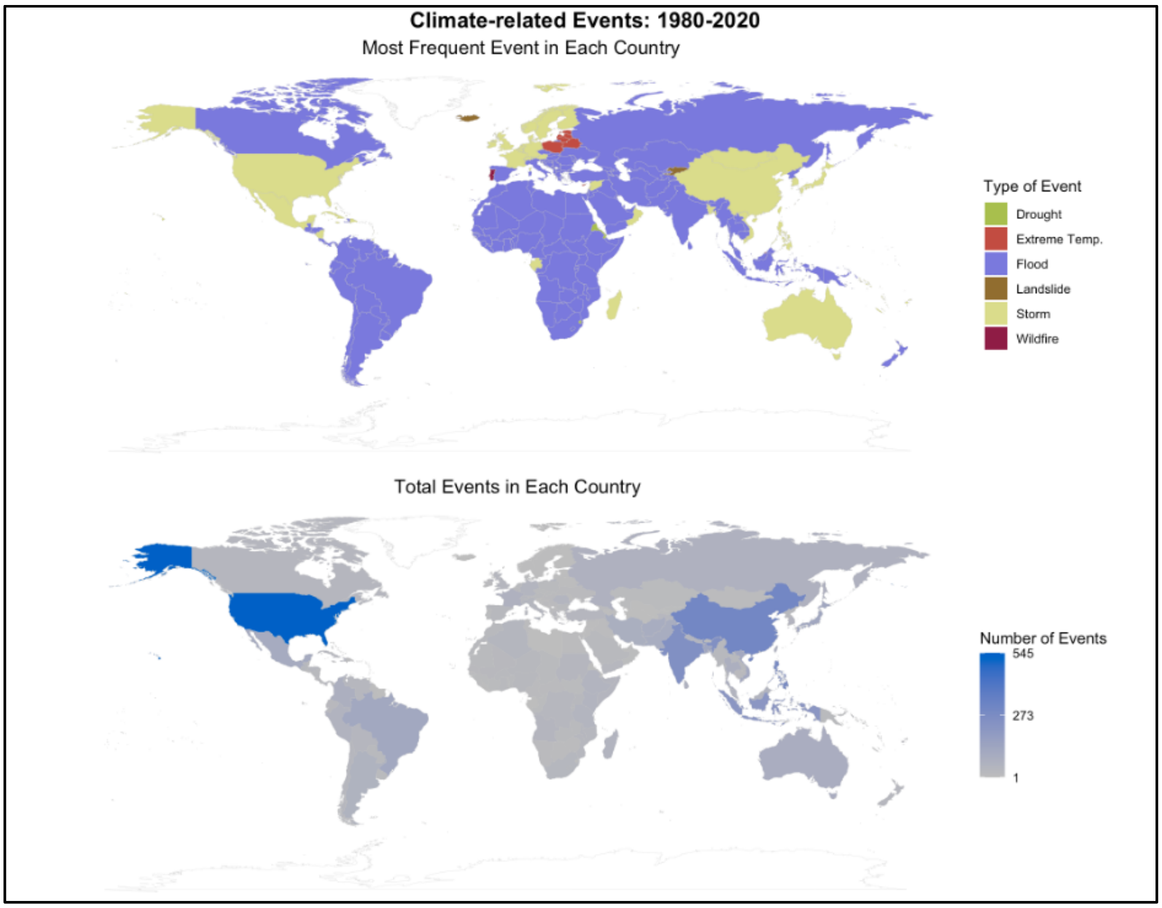
Animated Bar Graph Depicting Accumulation of Climate Events Over Time
Link: https://giphy.com/gifs/aHdK8TGPE54a7SBqBe



Although not depicted in the PDF, this graph is an animated bar chart that depicts the accumulation of climate-related events over the 40-year span of the data. The pictures above are 5-year snapshots extracted from the GIF, which can be opened through the link above.

This visualization expands on the dashboard's story, showing the actual growth piling up. The accumulation notably speeds up in the more recent decades. However, a slight lull can be noticed in the most recent years. This can also be seen in the plateau in the dashboard's line graph.

Choropleths: Most Frequent Event and Respective Counts



Finally, I move away from changes over time and instead focus on the four decades. Above, I see two stacked choropleths that look at all 40 years of data-—the top shows the most frequently occurring event by country, while the bottom shows the

total number of events for each country. By creating a new row-wise variable that sums the total events in each category and then grouping by countries' ISO3 codes, I can now look at the entirety of each countries' climate-related disaster history through the lens of a map. This allows us to compare the different counts of events for the separate countries and create a subset of the data that retains the most frequent.

I then use color to decode the difference in categories and amount. For event frequency, I use a discrete color scheme to distinguish the six most common types of events. For total events, I use a sequential color scheme with a neutral hue. After consideration, the blue scale was chosen to avoid any misleading use of colors, notably red and/or green, that could imply an improper measure, such as good vs. bad, hot vs. cold, etc.

While the choropleths were initially designed to be interpreted separately, the two maps also work together to identify countries with numerous events and of those countries, which events happen most frequently. For example, I can easily see that both China and the US both have higher event counts and both experience severe storms. However, because I also show the areas of the country, this allows us more insight on the smaller countries as well. I can see that Iceland does not have many events when compared to larger countries, but due to its relative size on the map, I can then give less weight to the lower choropleth and are left with the interesting observation that is has more landslides than other events, which is unique when considering the map.

D. Analysis and Discussion

The initial exploratory research into climate change shows that the frequency and total number of climate-related disasters appear to be increasing. Therefore, I took a closer look at these events over the previous four decades. Because the primary dataset addressed all global countries and 40 years of observations, I was able to look at the history of climate-related disasters with several different focuses: changes over time and by impact/frequency, as well as breaking these two foci down by location. Therefore, I undertook visualizing distribution of event type over time and total event over time. Additionally, I looked at these measures for the entire 40-year span and by country, where appropriate.

The Bertin Matrix encodes the increasing frequency in two separate ways—with shape and as a bar graph, allowing us to observe the constant rise but also the hopeful start of a plateau in the past decade. This is further supported by the interactive dashboard, which observes these points of interest as more detailed, separate but linked graphs. This allows multiple parties to investigate their area, event type, timeframe, or country of interest.

To look at the 40 years of data, I used choropleths to emphasize the frequency and amount of the most impactful events by country. I can see that most countries are impacted by floods and storms. Smaller countries, however, tend to have a wider variety of most frequently occurring events, such as landslides or extreme temperatures.

To summarize, the data investigated, and the visualizations created show that an impact of climate change is the increase in frequency of climate-related disasters. Over the past 40 years, the number of events has increased, especially during the last two decades. Of these events, storms and floods are most common, especially for larger countries. By continuing to track changes in climate-related events around the globe, I will gain additional insight into changes to the environment that would not be noticeable on a day-to-day level.

E. Appendix

Data Source:

IMF. (2021, February 27). Climate related disasters frequency. Climate Change Indicators Dashboard. Retrieved November 15, 2021, from https://climatedata.imf.org/datasets/b13b69ee0dde43a99c811f592af4e821_0/about.

RScript Code for Choropleth Visualization and Animation Visualization:

```
#DSC 465 Homework 4 | Prof. Brown
#Will Shoener | 11/14/2021

library(gganimate)
library(ggplot2)
library(tidyverse)
library(dplyr)
library(reshape)
library(gifski)
library(rworldmap)
library(maps)
library(gridExtra)
library(grid)
library(plyr)
library(reshape2)

##Load & preprocess data
setwd('/Users/wshoener/Downloads/')
#clean up data
eventsFull <- read.csv(file='Climate_related_disasters_frequency.csv', header=TRUE,sep=",")
events <- as.data.frame(subset(eventsFull, select = -c(Unit, Code)))
events$Indicator <- gsub("Climate related disasters frequency, Number of Disasters:
","",as.character(events$Indicator))
events <- events %>% replace(is.na(.), 0) %>% mutate(EventTotal = rowSums(.[6:46]))

#Fix Namibia
events[events$ObjectId == 587, "ISO2"] <- "NA"
events[events$ObjectId == 588, "ISO2"] <- "NA"
events[events$ObjectId == 589, "ISO2"] <- "NA"

##Prepare variations for Visualizations
eventsYearTypeGrouped <- events[5:46]
eventsYearTypeGrouped <- melt(eventsYearTypeGrouped, id=c("Indicator"))
eventsYearTypeGrouped <- eventsYearTypeGrouped %>% group_by(Indicator, variable) %>%
  summarise(across(everything(), list(sum = sum)))
names(eventsYearTypeGrouped) <- c('Event_Type','Year','Number_of_Events')
eventsYearTypeGrouped$Year <- as.integer(gsub("F","",as.character(eventsYearTypeGrouped$Year)))

#add in decade
eventsYearTypeGrouped$Decade <- ifelse(eventsYearTypeGrouped$Year %in% (1980:1989),'3',
                                       ifelse(eventsYearTypeGrouped$Year %in% (1990:1999),'2',
                                              ifelse(eventsYearTypeGrouped$Year %in% (2000:2009),'1',
                                                     ifelse(eventsYearTypeGrouped$Year %in% (2010:2020),'0',''))))

eventsBar <- filter(eventsYearTypeGrouped, Event_Type != "TOTAL")
eventsBar$Event_Type <- gsub("Extreme temperature","Extreme Temp.",as.character(eventsBar$Event_Type))

#Separate cumulative sums
eventHistoric1 <- eventsBar %>%
                  filter(Event_Type=='Extreme Temp.') %>%
```

```
                  group_by(Event_Type, Year) %>%
                  mutate(csum = cumsum(Number_of_Events))

eventHistoric2 <- eventsBar %>%
                  filter(Event_Type=='Drought') %>%
                  group_by(Event_Type, Year) %>%
                  mutate(csum = cumsum(Number_of_Events))

eventHistoric3 <- eventsBar %>%
                  filter(Event_Type=='Flood') %>%
                  group_by(Event_Type, Year) %>%
                  mutate(csum = cumsum(Number_of_Events))

eventHistoric4 <- eventsBar %>%
                  filter(Event_Type=='Landslide') %>%
                  group_by(Event_Type, Year) %>%
                  mutate(csum = cumsum(Number_of_Events))

eventHistoric5 <- eventsBar %>%
                  filter(Event_Type=='Storm') %>%
                  group_by(Event_Type, Year) %>%
                  mutate(csum = cumsum(Number_of_Events))

eventHistoric6 <- eventsBar %>%
                  filter(Event_Type=='Wildfire') %>%
                  group_by(Event_Type, Year) %>%
                  mutate(csum = cumsum(Number_of_Events))

eventHistoric <- rbind(eventHistoric1,eventHistoric2)
eventHistoric <- rbind(eventHistoric,eventHistoric3)
eventHistoric <- rbind(eventHistoric,eventHistoric4)
eventHistoric <- rbind(eventHistoric,eventHistoric5)
eventHistoric <- rbind(eventHistoric,eventHistoric6)

p <- ggplot(data = eventHistoric, aes(x=Event_Type,y=csum)) +
  geom_bar(stat="identity",aes(fill=factor(Decade))) + xlab('Climate Related Event') +
  ylab('Number of Events') + ggtitle('','Separated by Type of Event') +
  scale_fill_manual(values = c("#8bc288","#d68c69","#a686bf","#8ab8bf"), name = "Decade", labels = c("2010-2020","2000-
2009","1990-1999","1980-1989")) +
  theme(panel.border = element_blank(),panel.grid.minor = element_blank())

gifevent <- p + transition_time(Year) + ease_aes('linear', interval = 5) +
  labs(title = "Total Accumulation of Climate Events Over Time: {frame_time}")

gifevent

anim_save("climateevents_wshoener.gif", gifevent)

#World Map/Choropleths:
mapData <- filter(events, Indicator != "TOTAL")
mapData <- filter(mapData, Country != "All Countries and International Organizations")
mapData <- mapData %>% group_by(ISO3) %>% slice(which.max(EventTotal)) #get most common event per country
mapData$Indicator <- gsub("Extreme temperature","Extreme Temp.",as.character(mapData$Indicator))

world_map <- map_data(map = "world")
world_map$region <- iso.alpha(world_map$region, n=3)

m1 <- ggplot(mapData) +
  geom_map(aes(map_id = ISO3, fill = Indicator), map = world_map) +
  geom_polygon(data = world_map, aes(x = long, y = lat, group = group), colour = 'grey', fill = NA, size = 0.05) +
  expand_limits(x = world_map$long, y = world_map$lat) +
  scale_fill_manual(name = "Type of Event", values = c("#adbd5e", "#b55347", "#7a7cd6", "#8c6e3a", "#dbdb95",
"#852845")) +
  theme_void() + coord_fixed() + ggtitle("Most Frequent Event in Each Country") +
  theme(plot.title = element_text(hjust = 0.5))

m2 <- ggplot(mapData) +
  geom_map(aes(map_id = ISO3, fill = EventTotal), map = world_map) +
  geom_polygon(data = world_map, aes(x = long, y = lat, group = group), colour = 'grey', fill = NA, size = 0.05) +
  expand_limits(x = world_map$long, y = world_map$lat) +
  theme_void() + coord_fixed() + ggtitle("Total Events in Each Country") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_fill_gradient(low='#bdbdbd', high='#0462bf', name="Number of Events",
                      breaks=seq(min(mapData$EventTotal),max(mapData$EventTotal),
                                 (max(mapData$EventTotal)-min(mapData$EventTotal))/2))

grid.arrange(m1,m2, top = textGrob("Climate-related Events: 1980-2020",gp=gpar(fontsize=15,fontface="bold")))
```

Final Project Report
DSC 456
11/22/2021

Throughout this project, I focused on the data with a holistic approach, addressing all 40 years, rather than looking at changes over time. Due to the nature of the data, I immediately recognized the need for a map. After covering the geographic data portion of the class, it became clear that a choropleth made the most sense. For my initial version, and as part of the exploratory phase, I made a simple choropleth with a sequential color scheme depicting the number of total events. Later, I expanded this to contain the two stacked choropleths shown in the final report.

I also worked on a map chart showing the Global land temperature by country. The colours represent the average temperature in the region, and another way I did visualization is by doing bar chart animation for Global land temperature by states.
Further, I made a tree map displaying the average temperature from the 1800s to 2013 all over the world. I used tableau. Even if it is thick, I can deduce that temperature changes occurred throughout time. As I all know, winters have cold temperatures and summers have hot temperatures. I may deduce that lower to higher temperatures are represented by lighter to darker hues.

Discrete graph represents the average temperature between the years 2000 and 2013.
Heat map displays the average temperature between the years 2000 and 2013. The temperature field was calculated from kelvin scale to Fahrenheit scale for a more readable format. interpret that the lower to higher temperature are represented from lighter to darker shades.