EM-DAT Data Visualization Tutorial

May 12, 2021

## Load libraries

First, we need to load the following libraries:

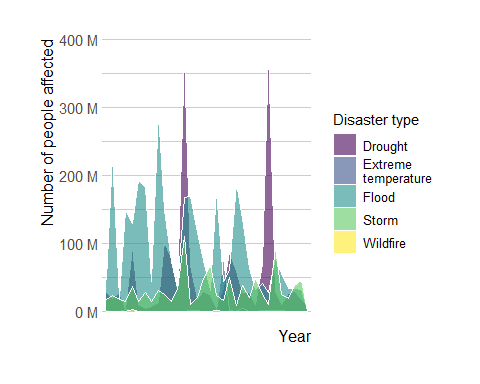
library(here)  
library(readxl)  
library(janitor)  
library(dplyr)  
library(stringr)  
library(ggplot2)  
library(hrbrthemes)  
library(viridis)  
library(scales)  
library(extrafont)  
loadfonts(device = "win")

## Import dataset

em\_dat <- utils::read.csv(here::here("scripts/cleaning/em-dat",   
 "em-dat-clean.csv"), stringsAsFactors = TRUE)

## Create visualization

em\_dat %>%  
 dplyr::select(-country, -iso) %>%  
 dplyr::filter(as.numeric(year) >= 1990) %>%  
 dplyr::group\_by(year, disaster\_type) %>%  
 dplyr::summarise\_all(funs(sum), na.rm = TRUE) %>%  
 ggplot2::ggplot(aes(x = year, y = total\_affected,   
 fill = str\_wrap(disaster\_type, 15),   
 group = disaster\_type)) + ggplot2::geom\_area(alpha = 0.6,   
 size = 0.1, colour = "white", position = "identity") +   
 viridis::scale\_fill\_viridis(name = "Disaster type",   
 discrete = TRUE) + hrbrthemes::theme\_ipsum(base\_size = 11) +   
 ggplot2::scale\_y\_continuous(labels = scales::unit\_format(unit = "M",   
 scale = 1e-06)) + ggplot2::ylab("Number of people affected") +   
 ggplot2::scale\_x\_discrete(name = "Year",   
 breaks = seq(1990, 2021, 5)) + ggplot2::theme(text = element\_text(family = "Arial",   
 size = 11), axis.title.x = element\_text(family = "Arial",   
 size = 12), axis.title.y = element\_text(family = "Arial",   
 size = 12), legend.text = element\_text(family = "Arial",   
 size = 10))



ggplot2::ggsave(here("images", "vulnerability-01.svg"),   
 device = "svg", dpi = 300)  
  
rm(em\_dat)

# Export as an R script for future use

Only run this chunk manually once within the .Rmd file. It produces an error when knitting it as a whole because of chunk label duplicates. As of May 12, 2021, there hasn’t been a viable solution to run the code below when as part of the knitting process.

knitr::purl("em-dat-visualizations.Rmd",   
 "em-dat-visualizations.R")  
knitr::write\_bib(.packages(), "packages.bib")

# Software used

Firke, Sam. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*, 2021. <https://github.com/sfirke/janitor>.

Garnier, Simon. *Viridis: Default Color Maps from Matplotlib*, 2021. <https://CRAN.R-project.org/package=viridis>.

———. *viridisLite: Colorblind-Friendly Color Maps (Lite Version)*, 2021. <https://CRAN.R-project.org/package=viridisLite>.

Müller, Kirill. *Here: A Simpler Way to Find Your Files*, 2020. <https://CRAN.R-project.org/package=here>.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2021. <https://www.R-project.org/>.

Rudis, Bob. *Hrbrthemes: Additional Themes, Theme Components and Utilities for Ggplot2*, 2020. <http://github.com/hrbrmstr/hrbrthemes>.

Wickham, Hadley. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. <https://ggplot2.tidyverse.org>.

———. *Stringr: Simple, Consistent Wrappers for Common String Operations*, 2019. <https://CRAN.R-project.org/package=stringr>.

Wickham, Hadley, and Jennifer Bryan. *Readxl: Read Excel Files*, 2019. <https://CRAN.R-project.org/package=readxl>.

Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*, 2020. <https://CRAN.R-project.org/package=ggplot2>.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. *Dplyr: A Grammar of Data Manipulation*, 2021. <https://CRAN.R-project.org/package=dplyr>.

Wickham, Hadley, and Dana Seidel. *Scales: Scale Functions for Visualization*, 2020. <https://CRAN.R-project.org/package=scales>.

Winston Chang. *Extrafont: Tools for Using Fonts*, 2014. <https://github.com/wch/extrafont>.