Number of News and Recorded Natural Disasters

220241454

2023-05

```
options(tinytex.verbose = TRUE)
```

load libraries first.

```
library(here)
library(tidyverse)
library(dplyr)

#sort conflict packages in order to kint to pdf

library(conflicted)
conflict_prefer("filter","dplyr")
conflict_prefer("log","dplyr")
conflicts_prefer(plotly::layout)
```

Research Question

This project looks into the the relation between news report and disasters. It was argued that more "spectacular" disasters received more corresponding news coverage, and news reporter tend to prefer to report on these disasters than the more "common" ones. Thus this project aim to look into the relation between the number of disasters and the news coverage percentage, and whether if there is a tendency for some disaster types to receive more news coverage.

Data Origin

This project have two data sets:

- Number of reported natural disasters
- News coverage of disasters

Number of reported natural disasters

Data first published by *EM-DAT, CRED / UCLouvain, Brussels, Belgium*, 2022-11-27. Data was downloaded from *Our World in Data.* This data includes all categories classified as "natural disasters", which includes drought, earthquakes, extreme temperatures, extreme weather, floods, fogs, glacial lake outbursts, landslide, dry mass movements, volcanic activity, and wildfires.

```
#load data
nnd <- read.csv("data/number-of-natural-disaster-events.csv")

#check loading worked
head(nnd)
```

```
##             Entity Code Year Number.of.reported.natural.disasters
## 1 All disasters   NA 1900                                    6
## 2 All disasters   NA 1901                                    1
## 3 All disasters   NA 1902                                   10
## 4 All disasters   NA 1903                                   12
## 5 All disasters   NA 1904                                    4
## 6 All disasters   NA 1905                                    8
```

News coverage of disasters

Data first published by *EISENSEE AND STRÖMBERG 2007*, 2007. Data was downloaded from *Our World
in Data*. This data includes disaster's corresponding coverage in major US networks. It was suggested that it
was evident that "spectacular" disasters receive more coverage. This data includes natural disaster categories
such as drought and earthquakes.

```
#load data
newsnd <- read.csv("data/news-coverage-of-disasters.csv")

#check loading worked
head(newsnd)
```

```
##        Entity Code Year Share.in.news..Eisensee.and.Strömberg.2007.
## 1      Africa   NA 2002                                          4
## 2        Asia   NA 2002                                         13
## 3   Cold wave   NA 2002                                          6
## 4     Drought   NA 2002                                          4
## 5  Earthquake   NA 2002                                         33
## 6     Epidemic   NA 2002                                          2
```

First Plot

As the news coverage data frame only included date in Year 2002. I would like to show how the number of
recorded disasters have increased over the years. This is to demonstrate whether the system of recording
natural disasters was able to record most of the natural disasters that occurred in 2002. If some natural
disasters were not recorded, news reporters' source could be biased. And thus, it would be difficult to
interpret the news coverage of disasters data.

Data preparation

I selected a few disaster types and combined the rest into "other" because their number was small.

```r
#filter out data I want to focus on (all disasters, extreme weather, earthquake, flood, and drought)
#rest of the data are combined as "other"

#Group by year and entity, and sum the events

disasters_sum <-nnd %>%
  group_by(Year, Entity) %>%
  summarise(events_sum = sum(Number.of.reported.natural.disasters))
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `.groups` argument.
```

```r
#Sum "other" events

other <- disasters_sum %>%
  filter(Entity %in% c("Fog", "Glacial lake outburst",
                       "Dry mass movement", "Volcanic activity",
                       "Extreme temperature", "Landslide")) %>%
  group_by(Year) %>%
  summarise(events_sum = sum(events_sum)) %>%
  mutate(Entity = "Other")

#Combine data frames

binded <- bind_rows(disasters_sum, other)

#sort by yeaar

binded <- binded %>%
  arrange(Year)

#filter

nnd_filtered <- binded %>%
  filter(Entity %in% c("All disasters","Extreme weather",
                       "Earthquake","Flood","Drought","Other"))

#check the first few lines

head(nnd_filtered)
```

```
## # A tibble: 6 x 3
## # Groups:   Year [1]
##    Year Entity          events_sum
##    <int> <chr>               <int>
## 1  1900 All disasters           6
## 2  1900 Drought                 2
## 3  1900 Earthquake              1
## 4  1900 Extreme weather         1
## 5  1900 Flood                   1
## 6  1900 Other                   1
```

```
#save the data frame

write.csv(nnd_filtered, file = "processed/nnd_filtered.csv", row.names=FALSE)
```

Data visualisation

```
#setting the order of legand

nnd_filtered$Entity <-
  factor(nnd_filtered$Entity, levels = c("All disasters", "Flood",
                                         "Extreme weather", "Earthquake",
                                         "Drought","Other"))

#plot a graph

bestfit <- ggplot(nnd_filtered, aes (x = Year, y = events_sum,
                                     group = Entity)) +
  geom_smooth (method = "gam", se = F, span = 2,
               aes (col = Entity), alpha = 0.5, size = 0.5) +
  geom_line (alpha = 0.8, size = 0.8, aes (col = Entity))  +
  ggtitle("Number of Natural Disaster Events") +
  labs(title = "Number of Recorded Natural Disasters",
       subtitle = "1900 to 2022",
       x = "Year", y = "Number of events",
       fill = "natural disaster types", color = "Disaster Category") +
  theme(plot.title = element_text(face = "bold", hjust = 0.5),
        plot.subtitle = element_text(hjust = 0.5))

#let's see the plot

bestfit
```
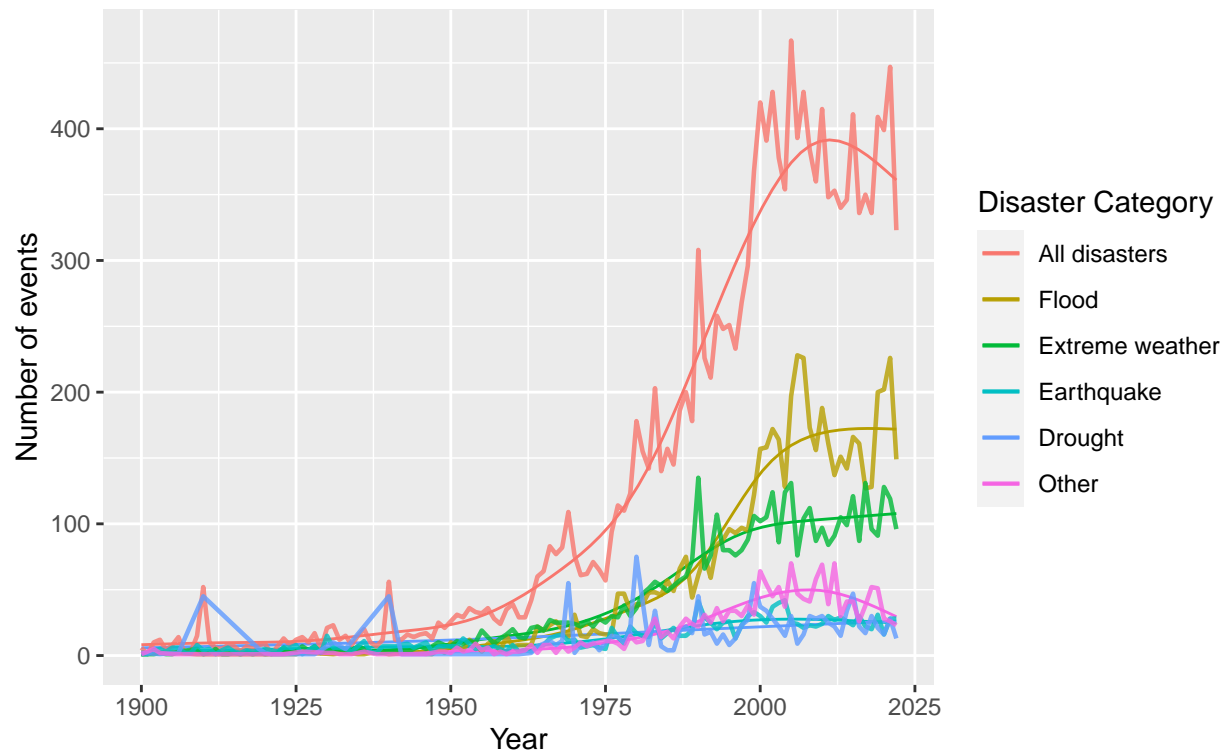
```
## `geom_smooth()` using formula = 'y ~ s(x, bs = "cs")'
```

Number of Recorded Natural Disasters
1900 to 2022

It could be suggested that since 2000, the number of recorded natural disasters have been relatively stable. Thus, suggests that news reporters should be able to access the source information with few bias.

```
#save the image

ggsave("nnd_bestfit.png", path = "outputs/", plot = bestfit,
       width=15, height=10,dpi=300)
```

```
## `geom_smooth()` using formula = 'y ~ s(x, bs = "cs")'
```

Second plot

Here will be a plot showing the number of reported natural disasters and the news coverage of neatural disasters in 2002.

Data preparation

```
#number of natural disaster in 2002
nnd2002 <- nnd %>% filter(Year == 2002)

#sort out nnd2002
```

```r
nnd2002 <- nnd2002 [-1, -c(2,3)]

DMM <- nnd2002 %>%
  filter(Entity %in% c("Landslide", "Dry mass movement")) %>%
  summarise(Number.of.reported.natural.disasters =
              sum(Number.of.reported.natural.disasters)) %>%
  mutate(Entity = "Dry mass movement")

nnd2002 <- bind_rows(DMM, nnd2002)
nnd2002 <- nnd2002 [-c(3,8),]

#sort out newsnd

#delete unwanted rows and columns

newsnd <- newsnd [-c(1,2,6,7,10,12,13), -c(2,3)]

#change names for better understanding

newsnd <- newsnd %>% rename(Share.in.news =
                              Share.in.news..Eisensee.and.Strömberg.2007.)

#uniform entity names, narrower terms are replaced by broader terms

newsnd <- newsnd %>%
  mutate (Entity = fct_recode(Entity,"Extreme weather"= "Storm",
                              "Extreme temperature" = "Cold wave",
                              "Dry mass movement" = "Landslide",
                              "Volcanic activity" = "Volcano",
                              "Wildfire" = "Fire"))

#combine data frames

news_n_nd <- merge(newsnd, nnd2002, by="Entity")
news_n_nd <- print(arrange(news_n_nd, -Number.of.reported.natural.disasters))
```

```
##                 Entity Share.in.news Number.of.reported.natural.disasters
## 1                Flood             9                                  172
## 2      Extreme weather            14                                  124
## 3           Earthquake            33                                   37
## 4              Drought             4                                   28
## 5             Wildfire            14                                   22
## 6    Dry mass movement             7                                   21
## 7  Extreme temperature             6                                   17
## 8    Volcanic activity            30                                    7
```

```r
#check the first few lines

news_n_nd
```

```
##                 Entity Share.in.news Number.of.reported.natural.disasters
## 1                Flood             9                                  172
## 2      Extreme weather            14                                  124
```

```
## 3          Earthquake                    33                                    37
## 4            Drought                     4                                     28
## 5           Wildfire                    14                                     22
## 6   Dry mass movement                    7                                     21
## 7 Extreme temperature                    6                                     17
## 8    Volcanic activity                   30                                     7
```

```r
#save data frame

write.csv(news_n_nd, file = "processed/news_n_nd.csv", row.names=FALSE)
```

## Data visualisation

```r
#plot

news_n_relation <- ggplot(news_n_nd) +
  geom_col(aes(x=Entity, y=Number.of.reported.natural.disasters),
           col = "dodgerblue4", fill= "dodgerblue4", alpha = 0.8) +
  geom_line(aes(x=Entity, y=Share.in.news, group=1), stat = "identity",
            col = "steelblue2", size = 1) +
  labs(title = "Number and News Percentage of Natural Disasters by Types",
       subtitle = "Year 2002, US") +
  scale_x_discrete("Type of natural disasters",
                   limits = c("Flood","Extreme weather","Earthquake",
                              "Drought","Wildfire",
                              "Dry mass movement","Extreme temperature",
                              "Volcanic activity")) +
  scale_y_continuous("Number of natural disasters",
                     sec.axis=sec_axis(~.*0.01,name="Percentage of News",
                                       labels=scales::percent)) +
  theme(plot.title = element_text(face = "bold", hjust = 0.5),
        plot.subtitle = element_text(hjust = 0.5),
        axis.title.y.left = element_text(face = "bold",color = "dodgerblue4"),
        axis.title.y.right = element_text(face = "bold",color = "steelblue2"))

#let's see the plot

news_n_relation
```
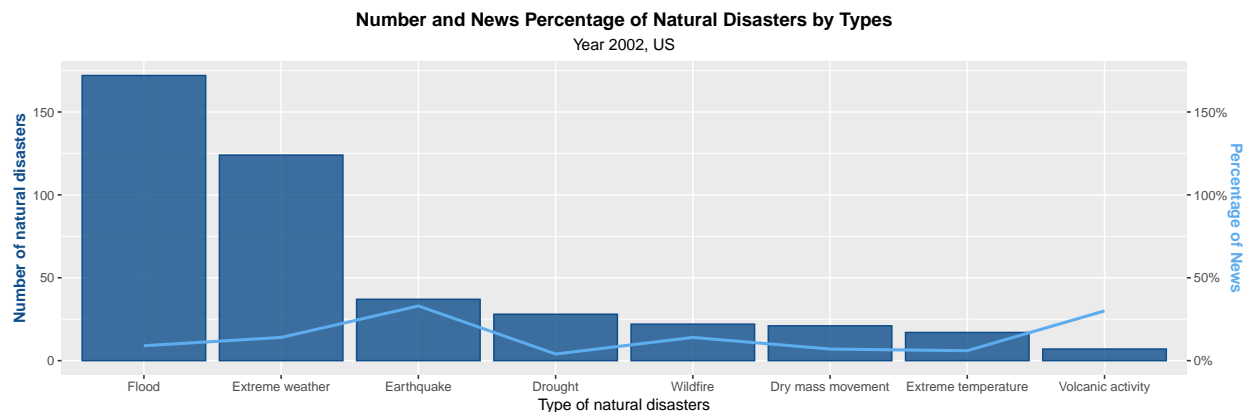
```
#save

ggsave("news_n_relation.png", path = "outputs/",
        plot = news_n_relation, width=15, height=10,dpi=300)
```

Summary

- Flood was the most common natural disaster, but the news coverage of flood-type disasters are low
- Earthquake and Volcanic activity as the more unusual natural disasters, had the top two news coverage among all the natural disasters covered in the data.

Thus, it could be suggested that there was a tendency that more 'spectacular' events such as earthquakes and volcanic activities had more news coverage than more 'common' ones such as Flood and Extreme weather.

Limitations

This project only included data from 2002 to produce the visualisation. The summary could be more valid if more news coverage data could be included (e.g. data from 2002 to 2022, or data from more counties).

Moreover, this project only included news coverage data from the United States, and it was compared to the number of reported natural disasters across the globe. This could also be biased as US news media would tend to report what happened locally. Thus, if the location of recorded natural disasters were available, natural disasters occurred in the US should also be included in the plot.

Colophon

- Eisensee, D., & Strömberg, T. (2007). News Droughts, News Floods, and U.S. Disaster Relief. The Quarterly Journal of Economics, 693-728. Retrieved from https://www.jstor.org/stable/25098856

- News coverage of disasters. (n.d.). Retrieved from Our World in Data: https://ourworldindata.org/grapher/news-coverage-of-disasters

- Number of recorded natural disaster events, 1900 to 2022. (n.d.). Retrieved from Our World in Data: https://ourworldindata.org/grapher/number-of-natural-disaster-events?time=earliest..latest

Repo for this project (the markdown file which generated this page, data and plot files) is here: *Y-AlpaCa/natural_disasters*