Data Analysis Report: Annual FEMA Disaster Declarations

Michael V Cumbo

February 12, 2024

0.1 About this Data

Disaster Declarations Summaries is a summarized dataset describing all federally declared disasters. This dataset lists all official FEMA Disaster Declarations, beginning with the first disaster declaration in 1953 and features all three disaster declaration types: major disaster, emergency, and fire management assistance. The dataset includes declared recovery programs and geographic areas (county not available before 1964; Fire Management records are considered partial due to historical nature of the dataset).

```r
library(tidyverse)
library(lubridate)
library(RSQLite)
library(DBI)
library(ggplot2)
library(dplyr)
library(forcats)
library(GGally)
library(stringr)
library(magrittr)

setwd("~/workbook")
con <- dbConnect(RSQLite::SQLite(), "Disaster_Data.db")
dbListTables(con)


## [1] "US_Declarations_2023"     "sqlean_define"
## [3] "us_disaster_declarations"


declarations_2023 <- as_tibble(dbGetQuery(
  con,
  "SELECT
  disasterNumber,
  state,
  declarationType,
  incidentType,
  declarationDate
  FROM US_Declarations_2023
ORDER BY declarationDate;"
))
dbDisconnect(con)
```

```
# build an arrangment paired with mutaute in
# DPLYR library than pass it to ggplot and build the graph
# this reorders the factors in the
# graph so its easier to break down and look at
 v1 <- declarations_2023 %>%
  select("state", "incidentType", "Year") %>%
  group_by(incidentType) %>%
  summarise(n = n()) %>%
  arrange(n) %>%
  mutate(incidentType = factor(incidentType, levels = unique(incidentType))) %>%
  ggplot(aes(x = incidentType, y = n)) +
  geom_col() +
  theme_gray() +
  theme(axis.text.x = element_text(angle = 45, vjust = 0.5, hjust = 0.5))
print(v1)
```
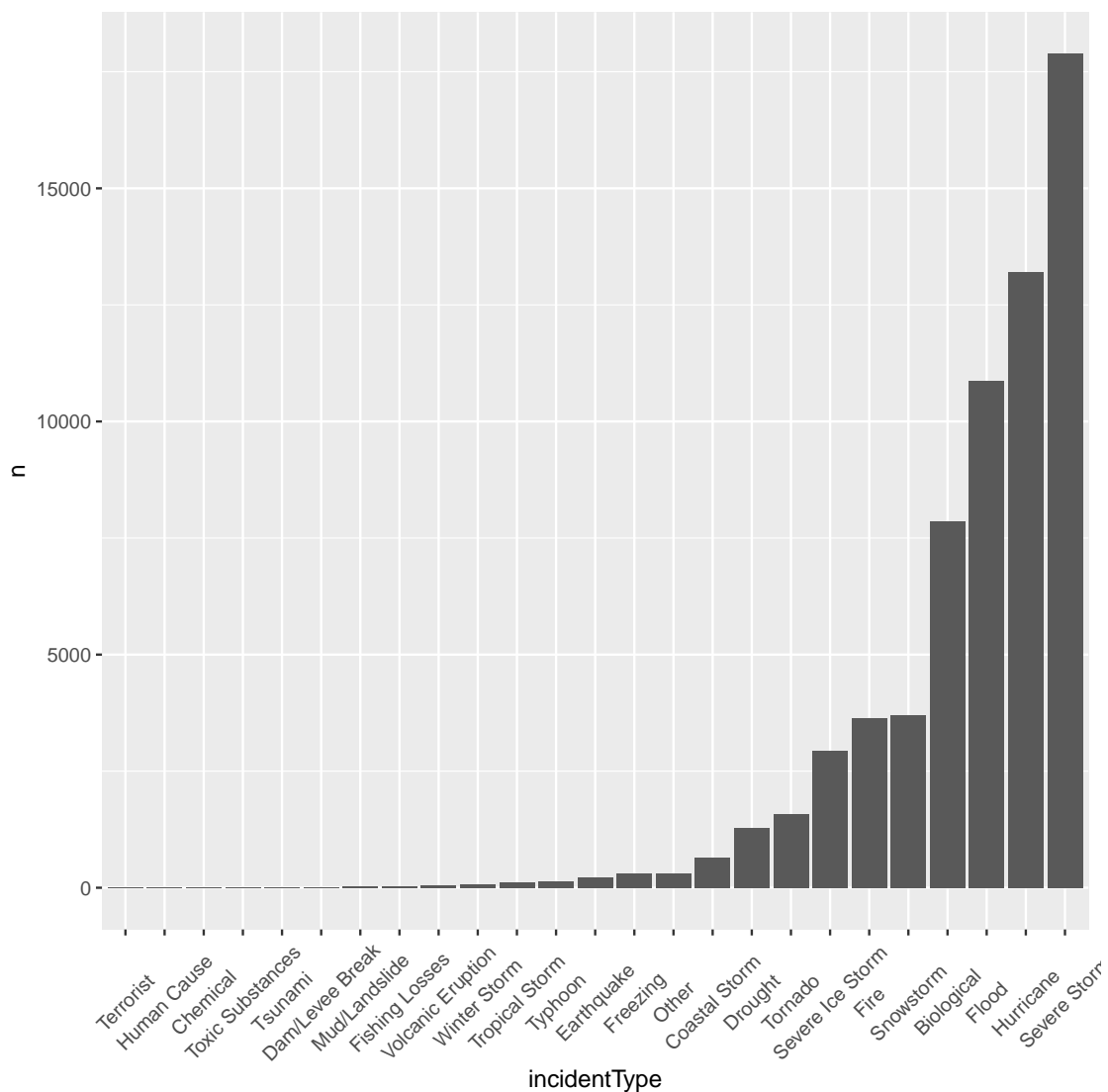

Figure 1: Histogram of the total number of disasters declared by the United States

Linear Regression Model Summary

Residuals

	Min	1Q	Median	3Q	Max
	-714.57	-312.30	-49.99	177.93	1361.83

Coefficients

	Estimate	Std. Error	t value	Pr($>$|t|)
(Intercept)	-49209.962	4821.925	-10.21	2.80×10^{-15} ***
Year	25.131	2.426	10.36	1.51×10^{-15} ***

Model Summary

Residual standard error	409.4 on 67 degrees of freedom
Multiple R-squared	0.6156
Adjusted R-squared	0.6098
F-statistic	107.3 on 1 and 67 DF
p-value	1.514×10^{-15}

$$y = \beta_0 + \beta_1 x + \epsilon \tag{1}$$

Coefficients:

- The intercept is -49209.962, implying the model's prediction for *DisasterCount* when *Year* is 0, which is not applicable in this context.

- The slope coefficient for *Year* is 25.131. This indicates an annual increase of approximately 25.131 in *DisasterCount*, as per the model.

- These predictions indicate an upward trend in *DisasterCount* over the years.

Statistical Significance:

- Both the intercept and the slope demonstrate statistical significance with a p-value < 0.001.

Model Fit:

- The R-squared value is 0.6156, signifying that approximately 61.56% of the variability in *DisasterCount* is explained by the year. However, a significant portion of variability remains unexplained.

- The Residual Standard Error (RSE) is 409.4, indicating the average deviation of data points from the fitted line.
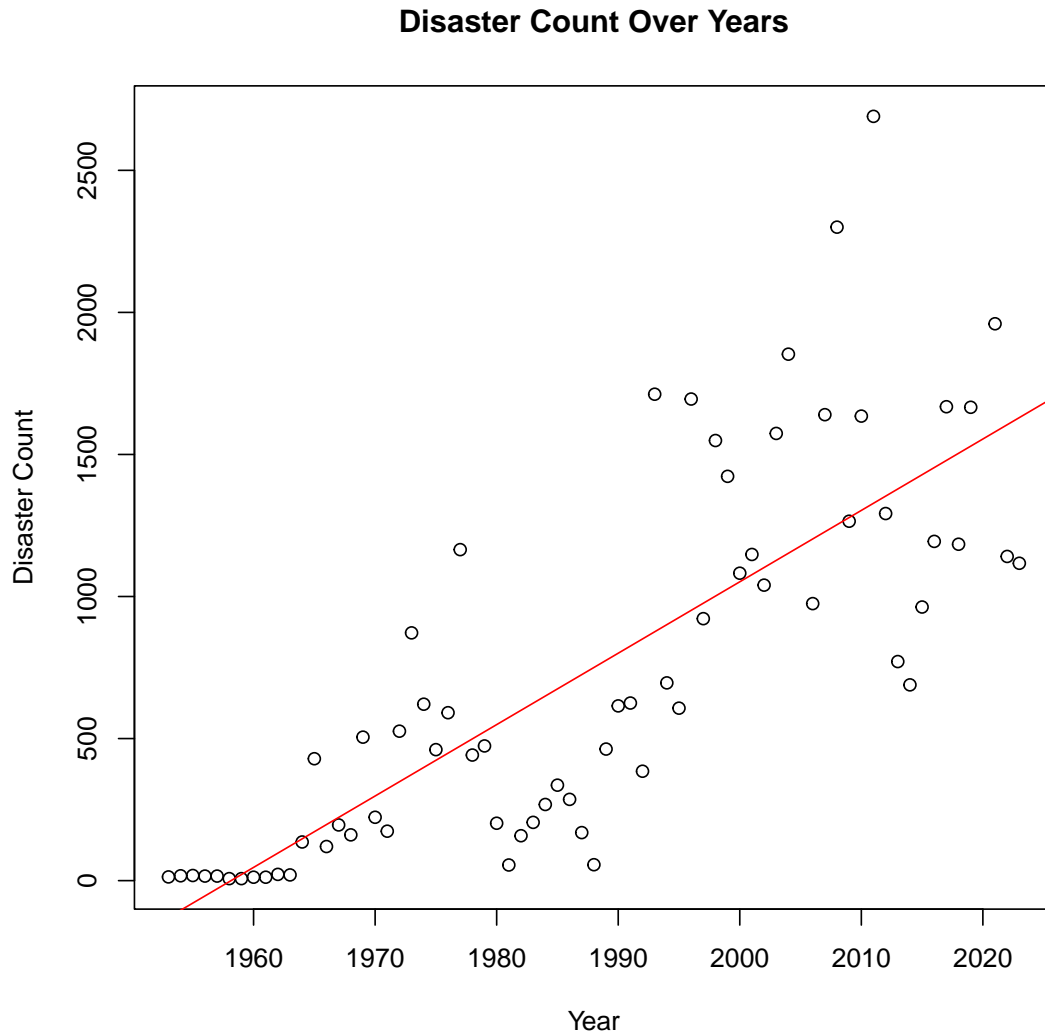
Figure 2: Simple Linear Rergression Model

0.2 Insights

- The model(see Figure: 2.) indicates a significant upward trend in disaster counts over the years.

- The presence of substantial residuals and an R-squared value of 0.6156 implies that, while there is a discernible trend, other unaccounted factors might also influence *Disaster Count* [2].

- Severe Storms, Hurricanes, Floods, and Biological make up the majority of declared disaster [1].

Considerations

- The Biological factor and the years 2020, 2005, and 2024 were filtered out of the linear model due to being significant outliers within the data set.

- Predictions for future years should be approached with caution due to the simplicity of the model and its exclusion of other potential predictive factors.

- When interpreting these results and making future decisions, the limitations of a simple linear regression model and the impact of external factors not included in the model should be considered.

- Further analysis of the top 3 populated factors present in the Histogram should be considered.

Analysis of Year-to-Year Change in Disaster Declarations

Table 1: Year-to-Year Change in number of Disaster Declarations by State

State	Year	YearToYearChange
TX	2005	693
ME	2020	609
TX	2020	567
LA	2020	497
PR	2020	351
TX	2008	349
TX	1998	315
GA	2020	308
GA	2017	266
AL	2020	265
MS	2020	263
OK	2007	245
VA	1996	227
LA	2005	222
FL	2004	220
MO	2011	220
TX	2011	215
KY	2020	206
AR	2020	204
FL	2022	203

0.3 Insights

- **High Fluctuations in Certain Years:** The years 2020, 2005, and 2008 stand out with significant changes in disaster declarations. This could indicate major disaster events or an increased number of smaller events that cumulatively led to a higher count.

- **Significant Increases in Texas (TX):** Texas shows the highest year-to-year increase in 2005 with 693 additional declarations compared to the previous year. Texas appears multiple times, indicating its vulnerability to disasters or its large size and varied climate contributing to different types of disasters.

- **Impact of 2020:** The year 2020 was significant for many states, including Maine (ME), Texas (TX), Louisiana (LA), Puerto Rico (PR), Georgia (GA), Alabama (AL), Mississippi (MS), Kentucky (KY), and Arkansas (AR), suggesting widespread disaster events or enhanced reporting mechanisms. Given the global pandemic's timeline, it's interesting to note such a high number of disaster declarations, suggesting that other disaster types (e.g., hurricanes, wildfires) also had a significant impact.

- **Variability Among States:** The list includes a mix of states from different regions, indicating that disasters affect a wide range of areas across the United States. Coastal states like Florida (FL) and Louisiana (LA) are featured alongside inland states like Missouri (MO) and Kentucky (KY), reflecting the diverse nature of disasters (natural, technological, or human-made) across the country.