# Visualizing FEMA NRI Data

Aakriti Poudel

2026-02-22

## Table of contents

## FEMA National Risk Index Scores for Counties in California

Create a data viz that helps to answer the question, How do FEMA National Risk Index scores for counties in California compare to those in other states?

### Set up

```
# Load all necessary libraries
library(tidyverse)
library(here)
library(janitor)
library(ggplot2)
```

### Import and read data

```
# Read data set
nri <- read.csv(here("data", "National_Risk_Index_Counties_807384124455672111.csv")) %>%
  clean_names()
```

**Data cleaning and wrangling**

```
# Extract the 50 states
us_states <- nri %>%
  filter(state_name %in% state.name)

# Aggregate county to the state level
states <- us_states %>%
  group_by(state_name) %>%
  # Calculate median value as it is robust to outliers
  summarize(median_risk_score = median(national_risk_index_score_composite, na.rm = TRUE),
            # Calculate average national risk score
            mean_risk_score = mean(national_risk_index_score_composite, na.rm = TRUE),
            # Calculate average annual expected loss loss score
            mean_annual_loss = mean(expected_annual_loss_score_composite, na.rm = TRUE),
            total_population = sum((population_2020), na.rm = TRUE),
            n_counties = n(),
            # Calculate range of risk, minimum to maximum
            min_risk = min(national_risk_index_score_composite, na.rm = TRUE),
            max_risk = max(national_risk_index_score_composite, na.rm = TRUE))

# Calculate for California
california <- states %>%
  mutate(is_california = ifelse(state_name == "California", "California", "Other state"),
    # Create a risk categories
    risk_category = case_when(
      median_risk_score < 25 ~ "Very low",
      median_risk_score < 50 ~ "Low",
      median_risk_score < 75 ~ "Moderate",
      median_risk_score < 90 ~ "High",
      TRUE ~ "Very High")) %>%
  # Sort by median risk score
  arrange(desc(median_risk_score))
```
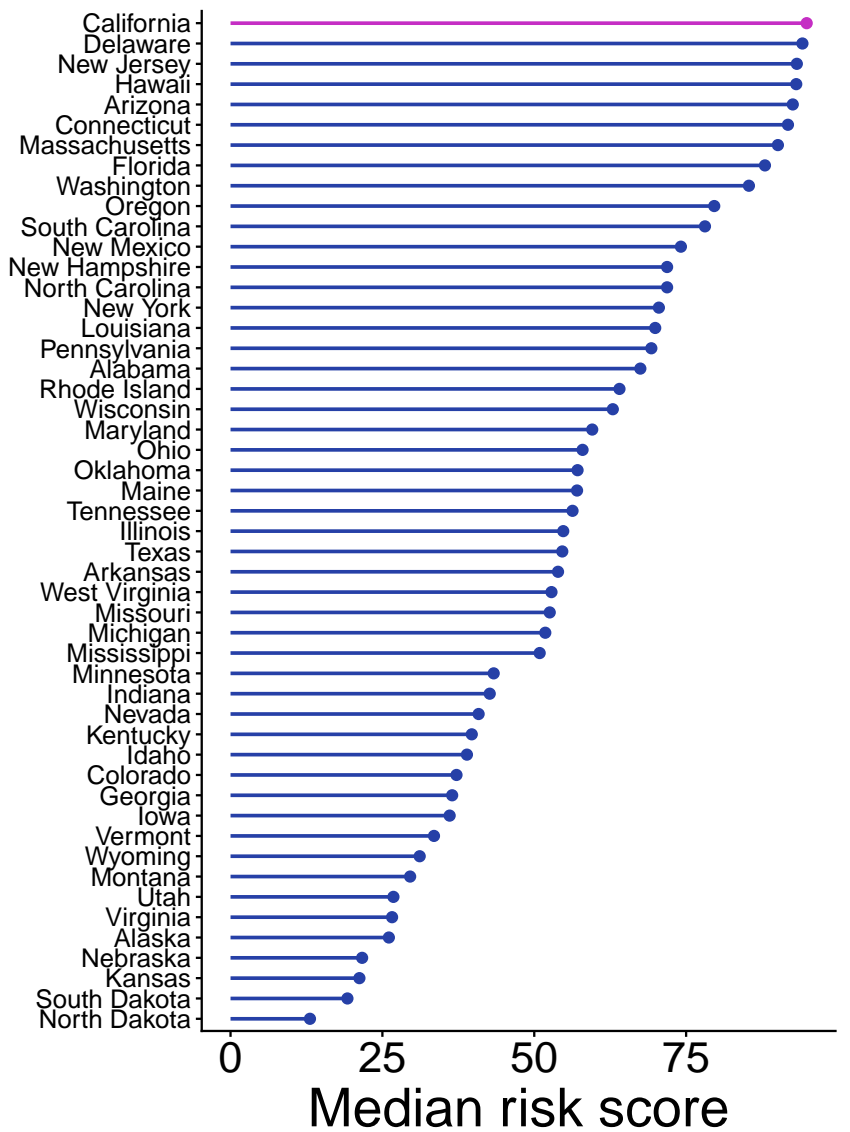
**Data visualization**

```r
ggplot(california, aes(x = reorder(state_name, median_risk_score),
                       y = median_risk_score,
                       color = is_california)) +
  geom_segment(aes(xend = state_name, y = 0, yend = median_risk_score), linewidth = 0.8) +
  geom_point(size = 2) +
  coord_flip() +
  scale_color_manual(values = c("California" = "#c52ec4", "Other state" = "#2640a7")) +
  labs(x = NULL,
       y = "Median risk score",
       title = "FEMA's National Risk Index (NRI) Score for Natural Hazards in US States",
       subtitle = "The median NRI score for US states",
       caption = "Data Source: FEMA National Risk Index, 2025",
       color = NULL) +
  theme_classic() +
  theme(legend.position = "bottom",
        legend.text = element_text(size = 20, face = "bold"),
        legend.key.size = unit(3, "cm"),
        plot.title = element_text(face = "bold", size = 25),
        plot.subtitle = element_text(size = 20),
        plot.caption = element_text(hjust = 0, color = "black", size = 20 ),
        axis.title.x = element_text(size = 25),
        axis.text.x = element_text(size = 20),
        axis.text.y = element_text(size = 12))
```

# FEMA's National Risk Ind

The median NRI score for US stat



Median risk score

—•— **California**    —•— **Other s**

Data Source: FEMA National Risk

**Answer some questions**

**1. What are your variables of interest and what kinds of data (e.g. numeric, categorical, ordered, etc.) are they (a bullet point list is fine)?**

| Variables of interest | Type of data |
| --- | --- |
| State name | Categorical |
| County name | Categorical |
| National Risk Index Score Composite | Numeric |
| National Risk Index Rating Composite | Categorical |
| Expected Annual Loss Score Composite | Numeric |

**2. How did you decide which type of graphic form was best suited for answering the question? What alternative graphic forms could you have used instead? Why did you settle on this particular graphic form?**

**Answer:** I chose a lollipop chart because it shows rankings for all 50 states in a clear, uncluttered way. The horizontal layout also makes the state names easy to read. The alternative graphic forms I could have used are bar chart or a box plot, but I feel like with 50 states these options can be too compact and visually heavy. I ultimately selected the lollipop chart because it keeps the visualization clean, makes rankings easy to interpret, and allows California to stand out relative to other states.

**3. Summarize your main finding in no more than two sentences.**

**Answer:** Based on FEMA National Risk Index scores, California ranks among the highest-risk states for natural hazards, along with states such as Delaware and New Jersey. In contrast, states like North Dakota, South Dakota and Kansas have the lowest median risk levels.

**4. What modifications did you make to this visualization to make it more easily readable?**

**Answer:** I made several changes to improve readability. The states were ordered from highest to lowest median risk score so comparisons are easy to follow. California is highlighted in a different color to make it stand out. The chart was flipped so state names are easier to read. The figure height and text sizes were adjusted to reduce crowding, and a clean theme was used.

**5. Is there anything you wanted to implement, but didn't know how? If so, please describe.**

**Answer:** I wanted to add a vertical reference line to show the national median risk score, but I am not sure how to do it in this plot. I also thought about using state abbreviations instead of full names to save space, but didn't know how to do it in a proper way.