

Comparing NRI Scores from California Counties and the Rest of the United States Counties

Melannie Moreno Rolón

2026-01-28

How do FEMA National Risk Index scores for counties in California compare to those in other states?

```
# Load packages
library(tidyverse)
library(here)

# Read in NRI county-level data
nri_raw <- read_csv(here("data", "National_Risk_Index_Counties_807384124455672111.csv"))

# Remove US territories
nri_states <- nri_raw %>%
  janitor::clean_names() %>%
  filter(!state_name_abbreviation %in% c("PR", "GU", "VI", "AS", "MP"))

# Select and rename variables + create a comparative variable
nri_compare <- nri_states %>%
  # Select variables of interest
  select(state_name, state_name_abbreviation, county_name,
         national_risk_index_score_composite) %>%

  # Rename variables
  rename(state_abbr = state_name_abbreviation,
         nri_score = national_risk_index_score_composite) %>%
  mutate(

    # Create a comparative variable
    region = if_else(condition = state_abbr == "CA",
```

```

    "California", "Other States")
)

#|fig-alt: "Density plot comparing county-level National Risk Index composite scores for Cal.

# Plot the density of NRI scores
ggplot(nri_compare, aes(x = nri_score, fill = region)) +

  # Create density plot
  geom_density(na.rm = TRUE, alpha = 0.7, color = "white") +

  scale_fill_manual(values = c("California" = "gold",
                                "Other States" = "lightblue1")) +
  scale_alpha_manual(values = c("California" = 0.75,
                                "Other States" = 0.35)) +

  scale_y_continuous(expand = c(0,NA)) +
  scale_x_continuous(expand = c(0,NA)) +

##~~~~~
##                                     Labeling                                     -----
##~~~~~

labs(title = "How do FEMA National Risk Index scores for \ncounties in
      California compare to those in other \nstates?",

      subtitle = "California counties are heavily concentrated at the highest
      NRI scores, \nwhile counties in other states span a much wider range of
      risk levels.",

      caption = "Source: FEMA Resilience Analysis & Planning Tool (2025)",

      alt_insight = "Density plot comparing county-level National Risk Index
      composite scores for California and other U.S. states. California's
      distribution is concentrated at scores around 80-100, while counties in
      other states span a wider range across the full 0-100 scale.",

      x = "National Risk Index Score Composite",
      y = "Density") +

##~~~~~
##                                     Theme Modifications                                     -----

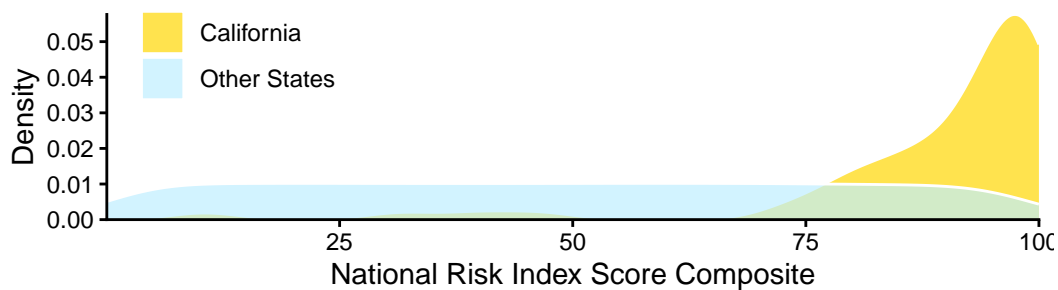
```

```
##~~~~~
theme_classic() +
theme(
  plot.title = element_text(size = 16),
  legend.position = c(0.14, 0.80),
  legend.title = element_blank()
) +

guides(alpha = "none")
```

How do FEMA National Risk Index scores for counties in California compare to those in other states?

California counties are heavily concentrated at the highest NRI scores, while counties in other states span a much wider range of risk levels.



Source: FEMA Resilience Analysis & Planning Tool (2025)

V. 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)?

- National Risk Index (NRI) Score – Composite

This is a numeric, continuous variable. It represents a county-level composite score summarizing overall natural hazard risk.

- County Name

This variable is categorical (nominal) and identifies individual counties.

- State Name

This is another categorical (nominal) variable that indicates the U.S. state each county belongs to.

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?

I selected a density chart because the research question focuses on comparing the distributions of county-level National Risk Index scores between California and other states. Density plots allow for direct comparison of a distribution's spread and overlap, which is I found could be useful for understanding how risk levels differ across groups. Alternative graphic forms considered included boxplots, violin plots, and ridgeline density plots. A two-group density plot was ultimately chosen because it clearly communicates distributional differences while remaining accessible to a non-technical audience.

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

This graphic found that California NRI scores are distributed at higher levels, while counties in other states have a wider range of NRI scores.

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

I wanted to avoid stacked shapes competing for attention and prevent scanning the graphic vertically, so the modification I executed was to create a **region** parameter that classifies all counties from California as "California" and all other counties as "Other States". This avoids plotting all 50 states individually.

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

I wanted to implement a fixed aspect ratio so that the rendered plot would be visible without overlaps or scaling issues. I tried using `coord_fixed()`, but did not find the right way to implement it.