

# Analysis of Spatial wildfire occurrence in United States (1992-2018)

Lokesh Bharule

December 10, 2021

Loading of necessary packages.

```
library(tidyverse)
library(RSQLite)
library(maps)
library(lubridate)
library(ggthemes)
library(mapdata)
```

Create db connection.

```
usa_db_conn <- dbConnect(SQLite(), "F:/RCodes/DAB501/FPA_FOD_20210617.sqlite")
```

Show List of Tables.

```
as.data.frame(dbListTables(usa_db_conn))
```

```
##      dbListTables(usa_db_conn)
## 1                      Fires
## 2 NWCG_UnitIDActive_20200123
```

Get table

```
usa_fires_temp <-tbl(usa_db_conn, "Fires") %>% collect()
```

Write to a csv

```
write_csv(usa_fires_temp, 'F:/RCodes/DAB501/fires.csv')
```

Loading of data

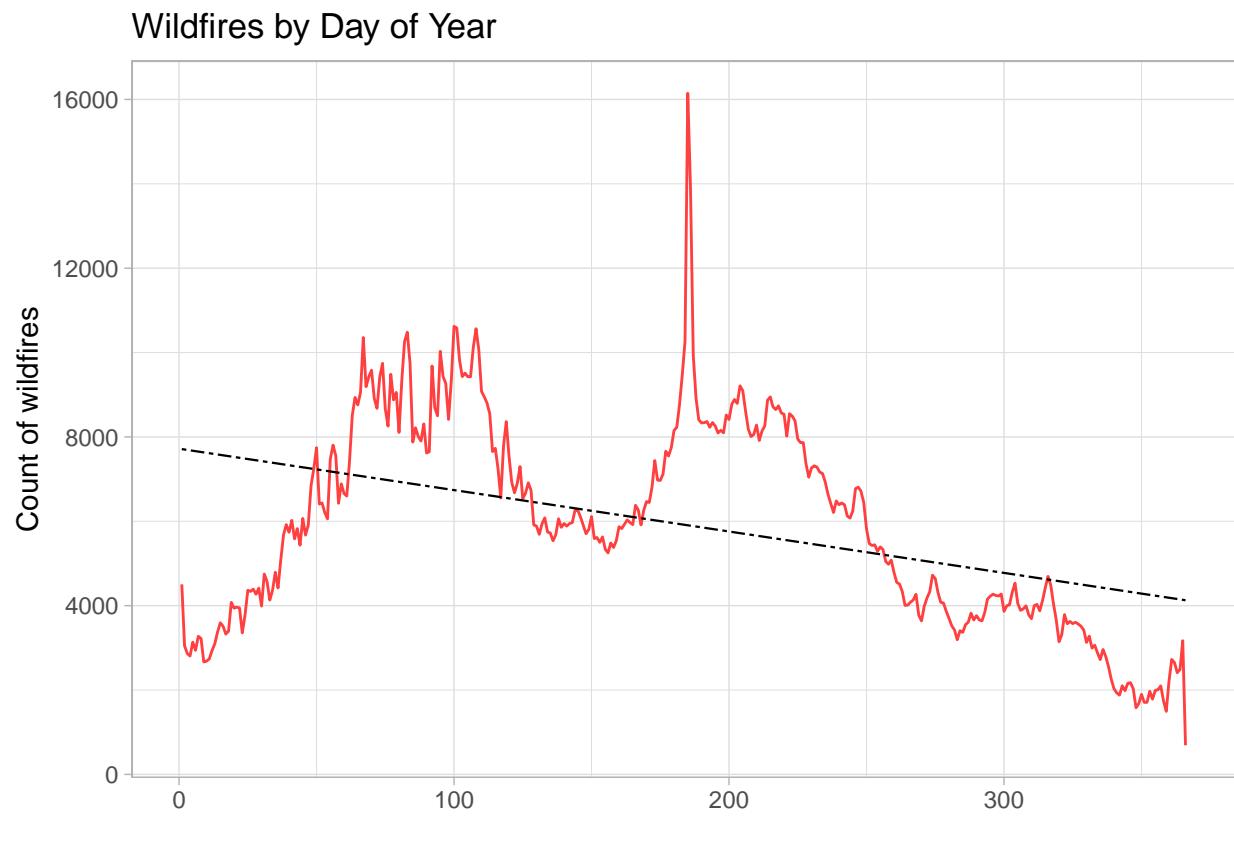
```
usa_fires <- read.csv('F:/RCodes/DAB501/fires.csv')
```

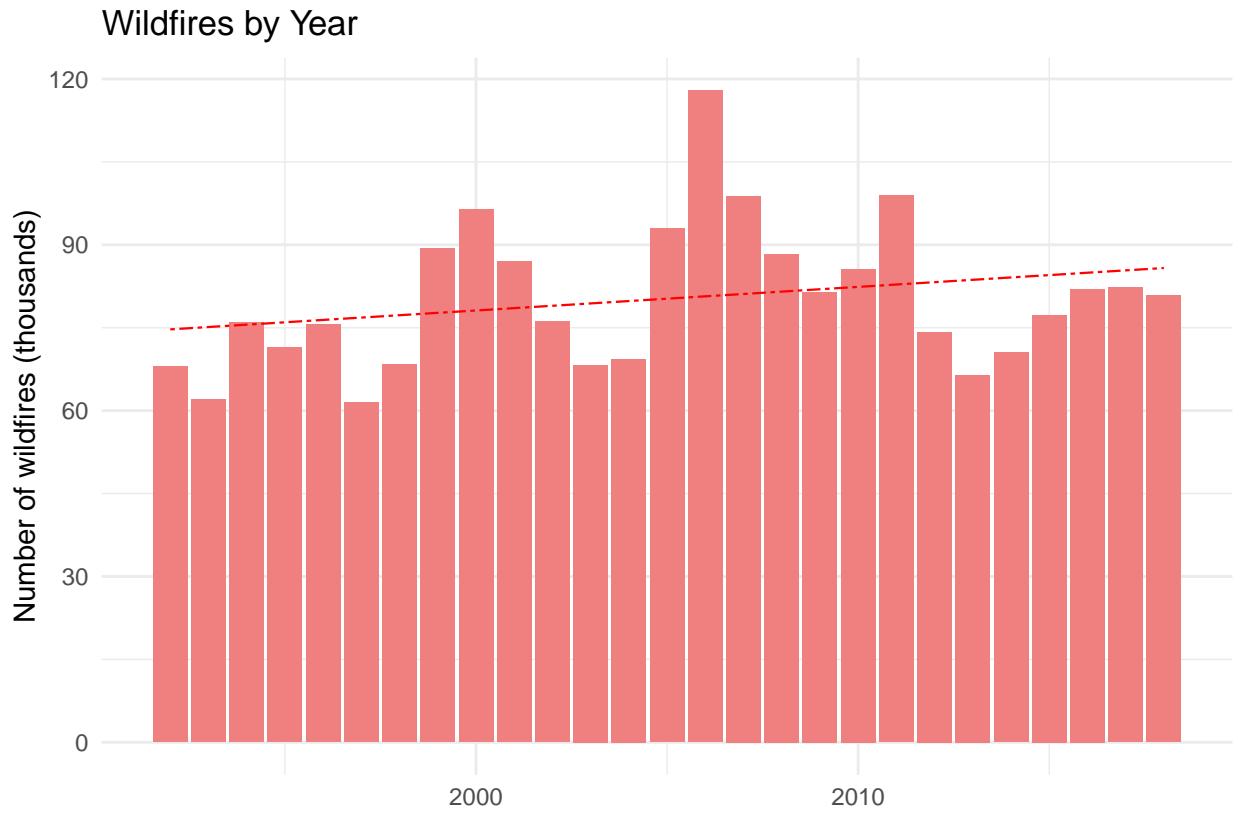
Disconnect from db

```
dbDisconnect(usa_db_conn)
```

## Timeseries

```
usa_fires %>%
  group_by(DISCOVERY_DOY) %>%
  summarize(n_fires = n()) %>%
  ggplot(aes(x = DISCOVERY_DOY, y = n_fires)) +
  geom_line(color = 'brown1') +
  geom_smooth(
    method = 'lm',
    se = FALSE,
    linetype = 'twodash',
    size = 0.4,
    color = 'black'
  ) +
  theme_light() +
  labs(x = '', y = 'Count of wildfires',
       title = 'Wildfires by Day of Year')
```





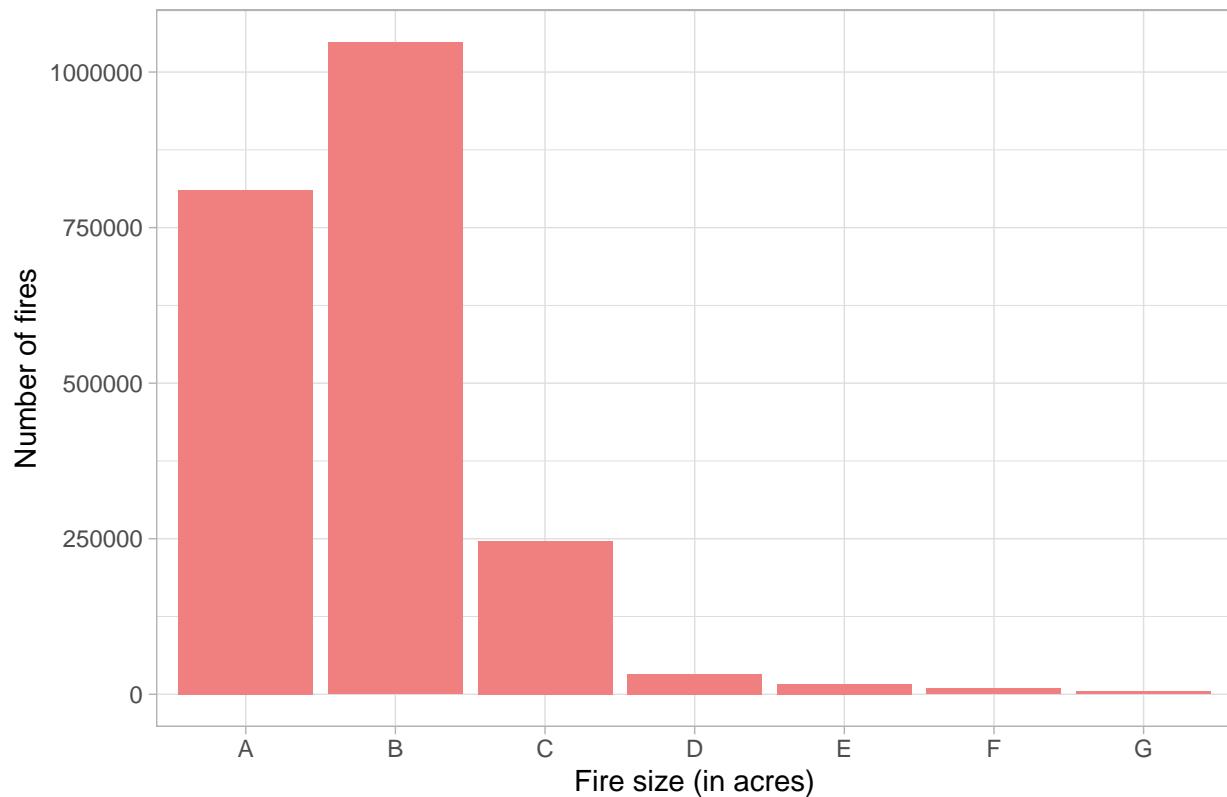
#### Annual

The number of fires per year were around 60,000 and 100,000 from 1992 to 2015. There was a spike in fires in 2006 to about 114,000. There is a small upward trend during this time period.

#### Fires size by class

```
usa_fires %>%
  group_by(FIRE_SIZE_CLASS) %>%
  summarize(n = n()) %>%
  ggplot(aes(x = FIRE_SIZE_CLASS, y = n)) +
  geom_bar(stat = 'identity', fill = 'lightcoral') +
  theme_light() +
  labs(x = 'Fire size (in acres)'
       , y = 'Number of fires'
       , title = 'Wildfires by Fire Size Class')
```

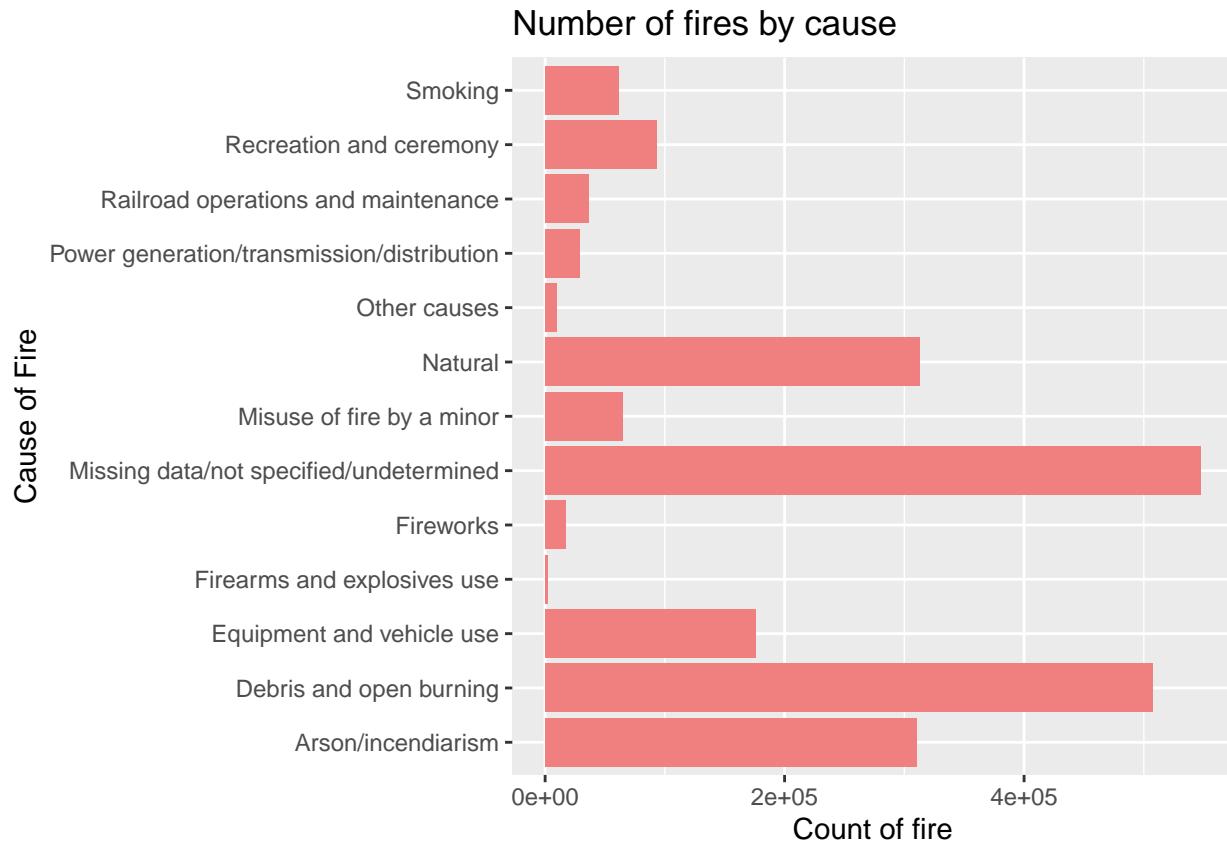
## Wildfires by Fire Size Class



What are the causes for wildfires and it's relationship with period and size?

### Count of fire by Causes

```
usa_fires %>%
  group_by(NWCG_GENERAL_CAUSE) %>%
  summarize(n = n()) %>%
  arrange(desc(n)) %>%
  ggplot(aes(x = NWCG_GENERAL_CAUSE, y = n)) +
  geom_bar(stat = 'identity', fill = 'lightcoral') +
  labs(x = 'Cause of Fire',
       y = 'Count of fire',
       title = 'Number of fires by cause') +
  coord_flip()
```



## Fire Period

```

#Date format conversion
date_conv2 <-
  as.Date(mdy_hm(usa_fires$DISCOVERY_DATE)) #yyyy-mm-dd
date_conv1 <-
  as.Date(mdy_hm(usa_fires$CONT_DATE))

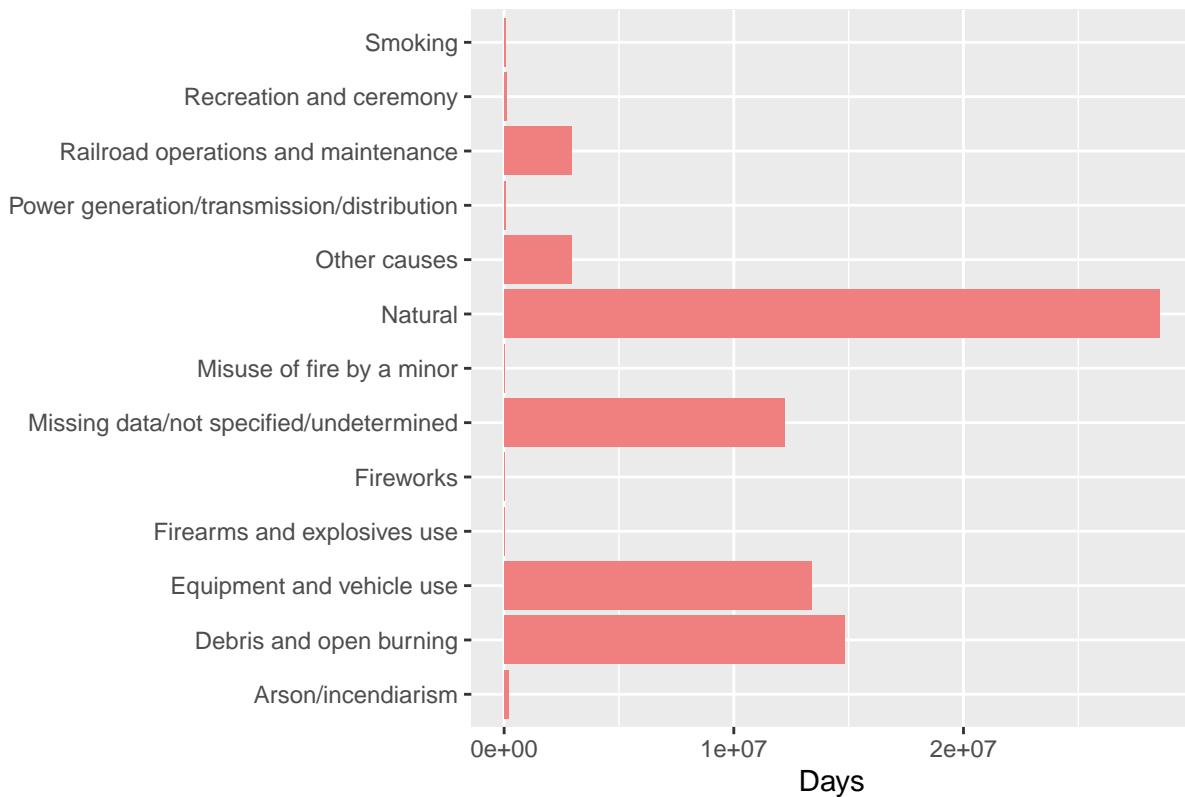
#Number of days fire was burning
usa_fires$Fire_Period <-
  difftime(date_conv1, date_conv2, units = 'days')

#Conversion of date format into numeric
usa_fires$Fire_Period <- as.numeric(usa_fires$Fire_Period)

usa_fires %>%
  group_by(NWCG_GENERAL_CAUSE) %>%
  ggplot(aes(x = reorder(NWCG_GENERAL_CAUSE, Fire_Period),
             y = Fire_Period)) +
  geom_bar(stat = 'identity', fill = 'lightcoral') +
  coord_flip() +
  labs(x = '',
       y = 'Days',
       title = 'Average Wildfire Fire Period by Cause')

```

## Average Wildfire Fire Period by Cause



## Geography of Wildfires

```

# Changing the state name to code
state.abb <- append(state.abb, c("DC", "PR"))
state.name <-
  append(state.name, c("District of Columbia", "Puerto Rico"))

usa_fires$region <-
  map_chr(usa_fires$STATE, function(x) {
    tolower(state.name[grep(x, state.abb)])
  })

# Pulling the US state map data
state_map <- map_data('state')

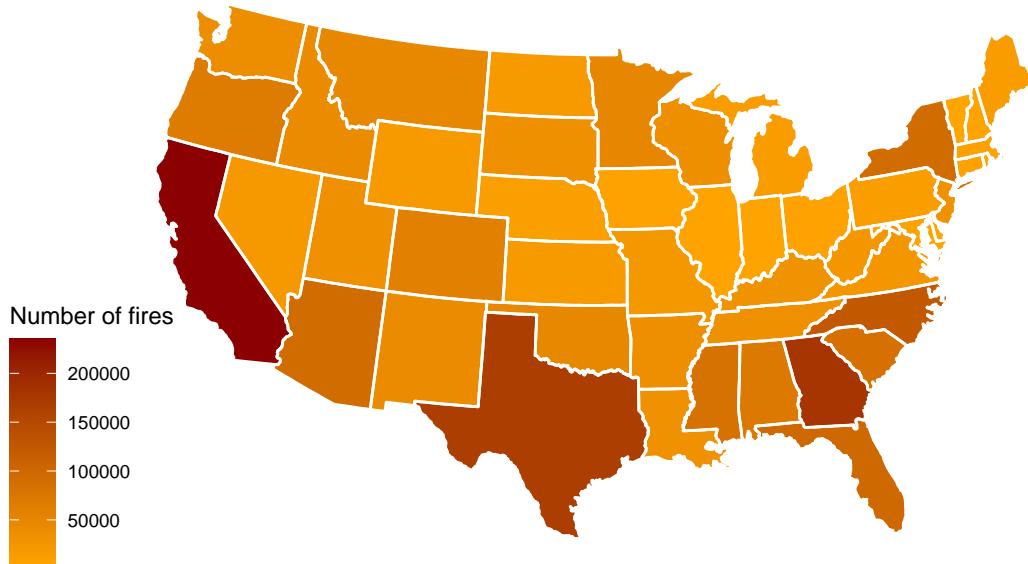
usa_fires %>%
  select(region) %>%
  group_by(region) %>%
  summarize(n = n()) %>%
  right_join(state_map, by = 'region') %>%
  ggplot(aes(
    x = long,
    y = lat,
    fill = region
  )) +
  geom_sf()
  
```

```

group = group,
fill = n
)) +
geom_polygon() +
geom_path(color = 'white') +
scale_fill_continuous(low = "orange",
                      high = "darkred",
                      name = 'Number of fires') +
theme_map() +
coord_map('albers', lat0 = 30, lat1 = 40) +
ggtitle("Wildfires in USA, 1992–2018") +
theme(plot.title = element_text(hjust = 0.5))

```

Wildfires in USA, 1992–2018



### Heatmap of Wildfire

```

world <- map_data("world")

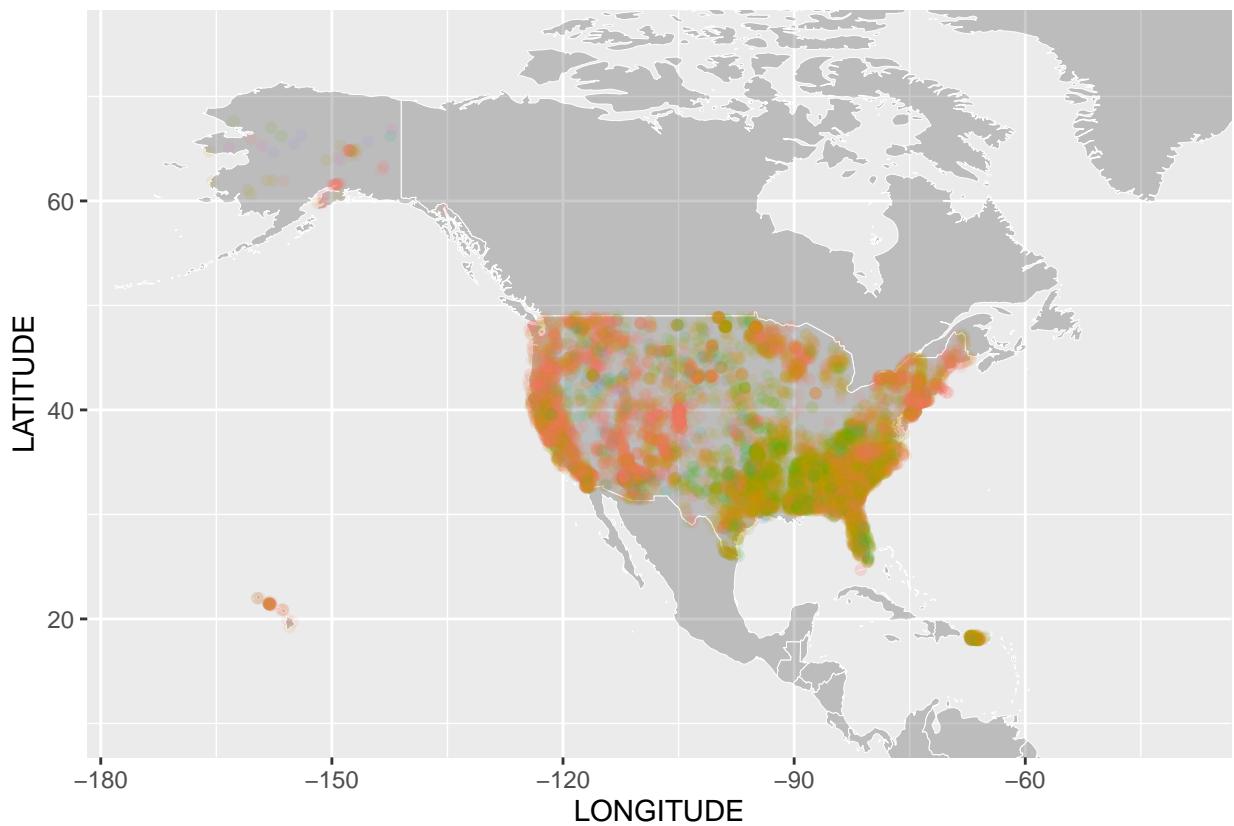
ggplot() +
  geom_map(
    data = world,
    map = world,
    aes( map_id = region),
    color = "white",

```

```

        size = 0.05,
        alpha = 1 / 4
    ) +
    geom_point(
        data = sample_n(usa_fires, 10000),
        aes(LONGITUDE, LATITUDE, color = FIRE_SIZE_CLASS),
        alpha = .1
    ) +
    coord_quickmap() +
    ylim(10, 75) + xlim(-175, -40) +
    guides(color = "none")

```



Map for each of the fire causes

```

state_para_func <- function(cause) {
  # Join dataframes and build map
  usa_fires %>%
    #filter(STAT_CAUSE_DESCR == cause) %>%
    select(region) %>%
    group_by(region) %>%
    summarize(n = n()) %>%
    right_join(state_map, by = 'region') %>%
    ggplot(aes(

```

```

    x = long,
    y = lat,
    group = group,
    fill = n
)) +
geom_polygon() +
geom_path(color = 'white') +
scale_fill_continuous(low = "orange",
                      high = "firebrick3",
                      name = 'Number of fires') +
theme_map() +
coord_map('albers', lat0 = 30, lat1 = 40) +
ggtitle(paste0("US Wildfires Caused by ", cause, ", 1992-2015")) +
theme(plot.title = element_text(hjust = 0.5))

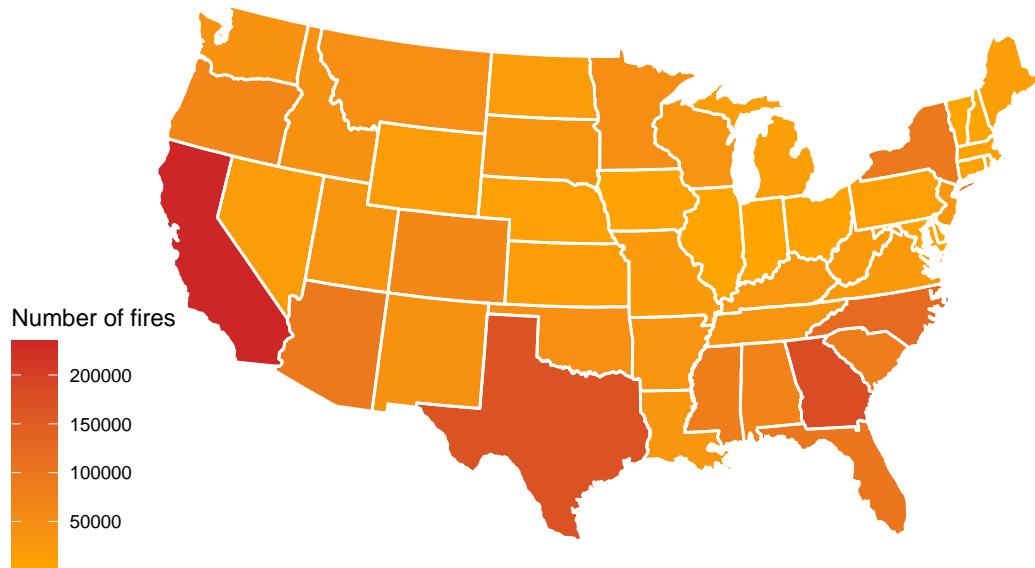
}

```

### Fires By State by causes

```
state_para_func(cause = 'Arson')
```

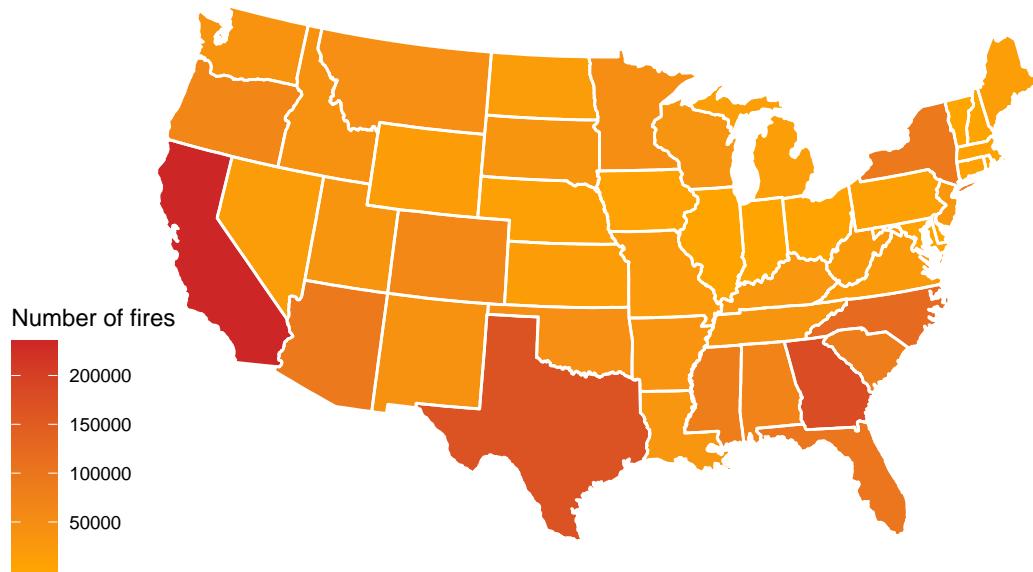
US Wildfires Caused by Arson, 1992–2015



Total

```
state_para_func(cause = 'Campfire')
```

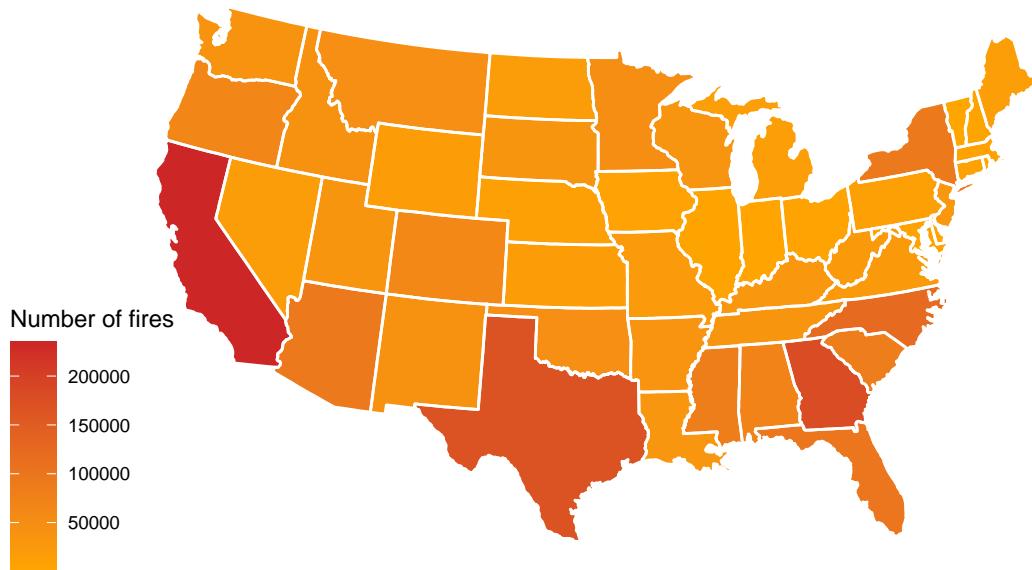
US Wildfires Caused by Campfire, 1992–2015



Campfire

```
state_para_func(cause = 'Children')
```

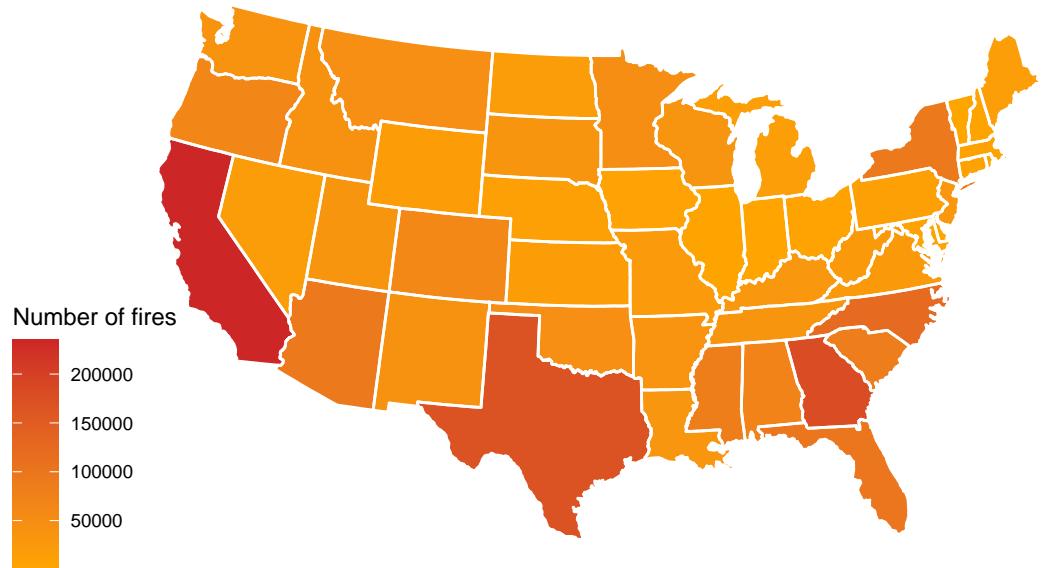
US Wildfires Caused by Children, 1992–2015



Children

```
state_para_func(cause = 'Debris Burning')
```

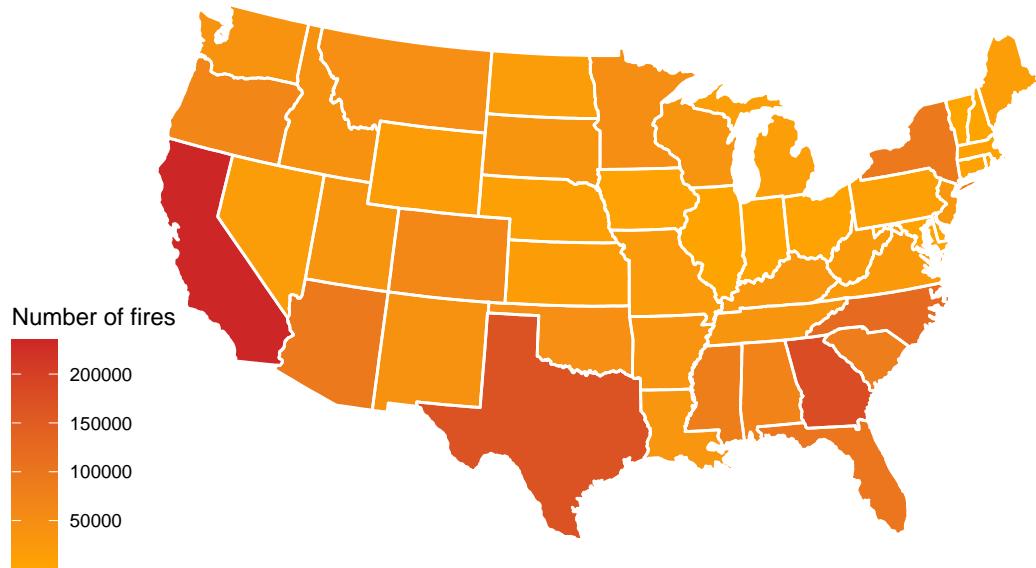
US Wildfires Caused by Debris Burning, 1992–2015



## Debris Burning

```
state_para_func(cause = 'Equipment Use')
```

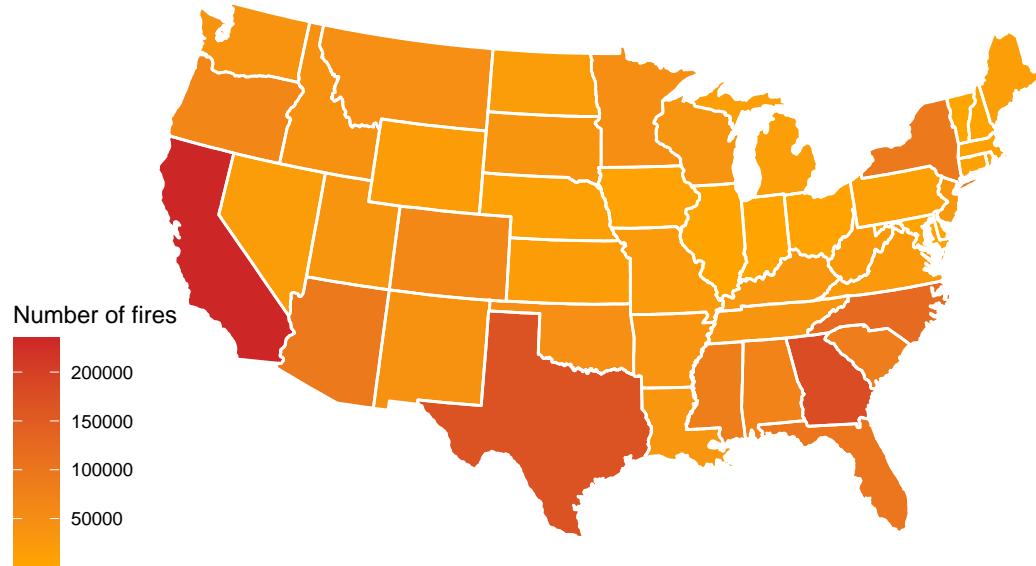
US Wildfires Caused by Equipment Use, 1992–2015



## Equipment Use

```
state_para_func(cause = 'Fireworks')
```

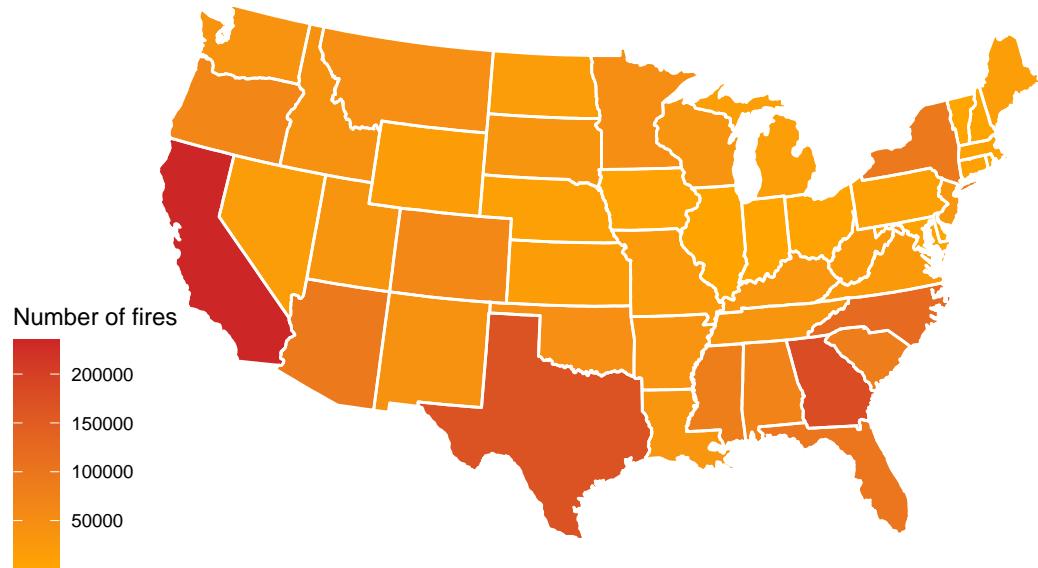
US Wildfires Caused by Fireworks, 1992–2015



## Fireworks

```
state_para_func(cause = 'Lightning')
```

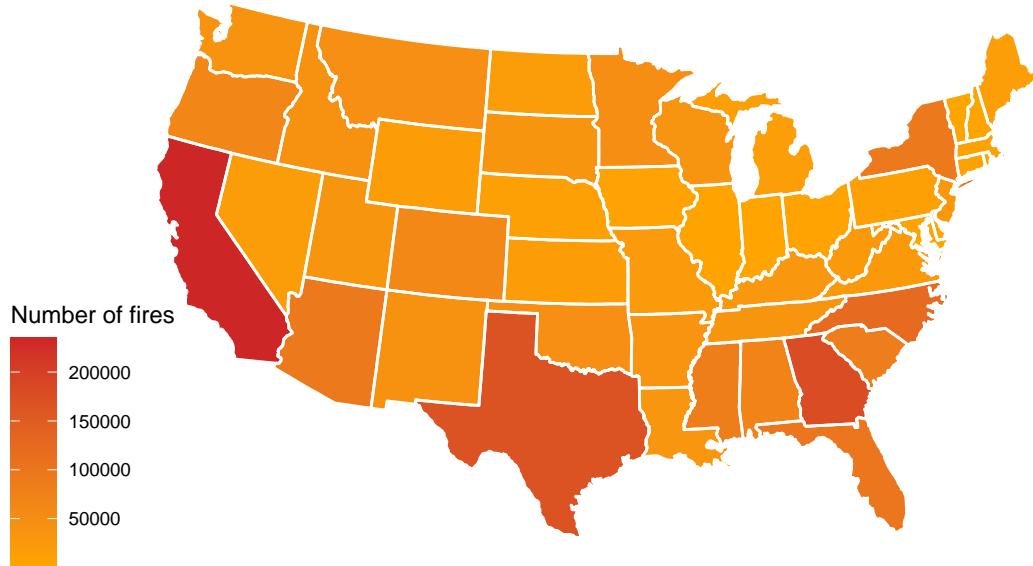
US Wildfires Caused by Lightning, 1992–2015



## Lightning

```
state_para_func(cause = 'Powerline')
```

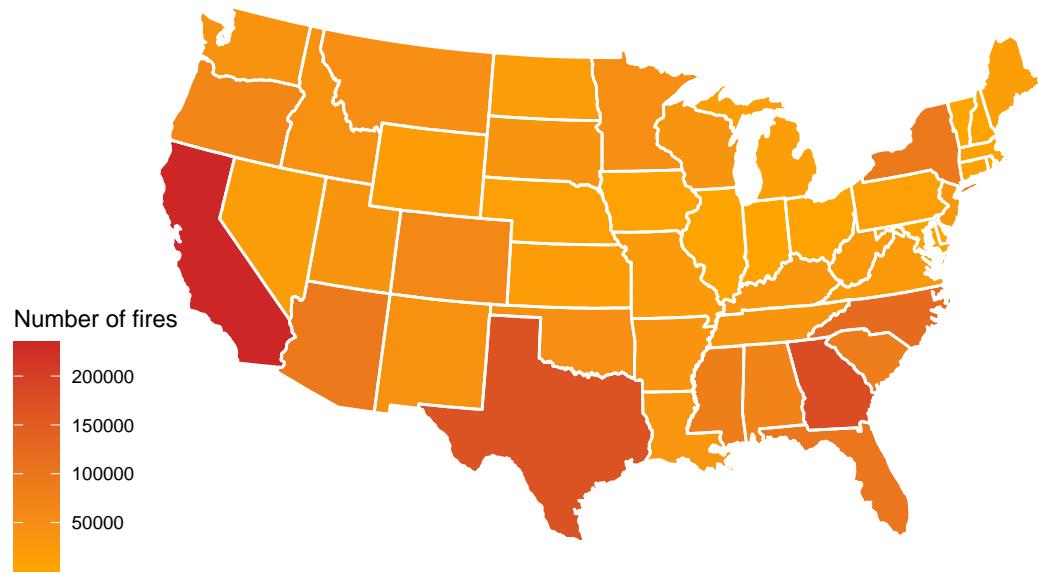
US Wildfires Caused by Powerline, 1992–2015



Powerline

```
state_para_func(cause = 'Railroad')
```

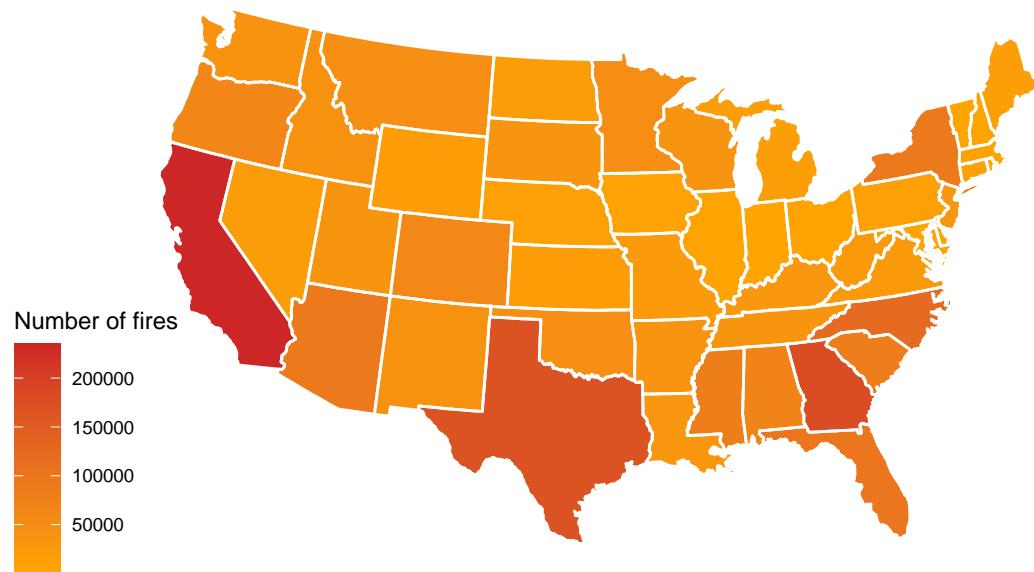
US Wildfires Caused by Railroad, 1992–2015



Railroad

```
state_para_func(cause = 'Smoking')
```

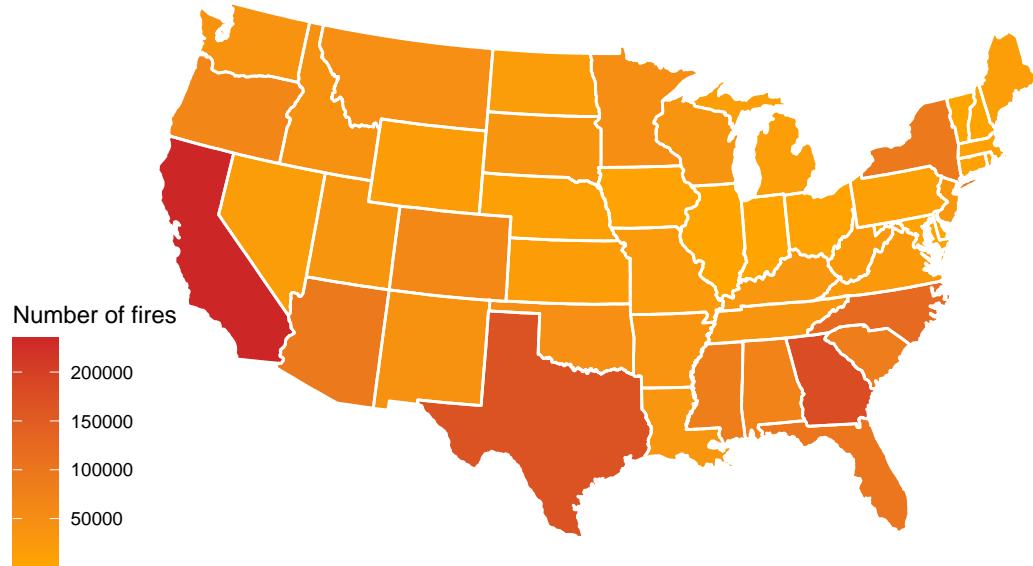
US Wildfires Caused by Smoking, 1992–2015



## Smoking

```
state_para_func(cause = 'Structure')
```

US Wildfires Caused by Structure, 1992–2015



## Structure

**Normalizing the data by state** If we normalize the data by state, we can see the number of wildfires per square mile. These maps will illustrate more because they will compare wildfire across states. Fortunately, R already has data on the area of US states in the state.x77 matrix.

```
# Create region column in state.x77 of lowercase state names
state.x77 <- state.x77 %>%
  as.data.frame() %>%
  mutate(region = tolower(rownames(state.x77)))

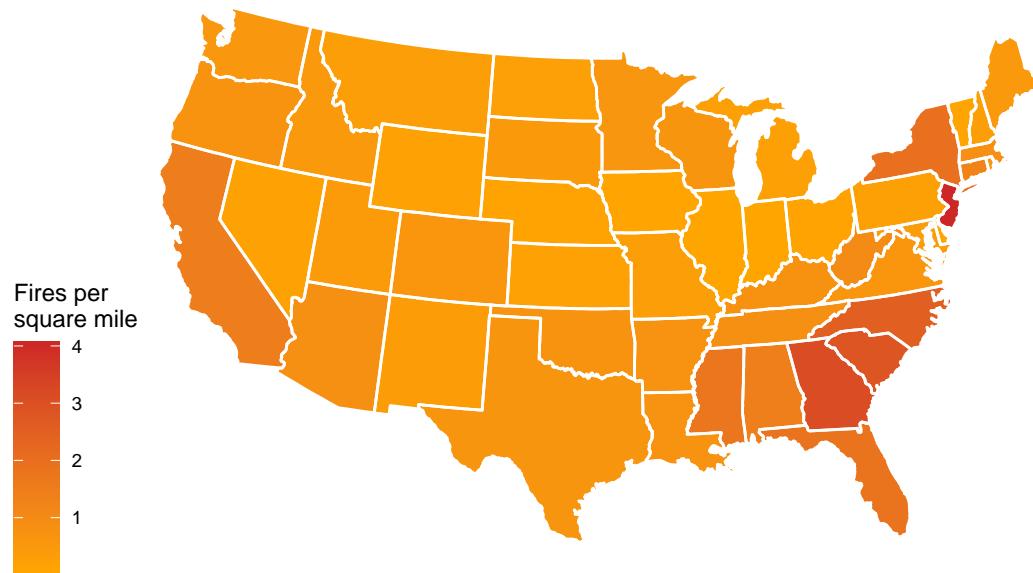
usa_fires %>%
  group_by(region) %>%
  summarize(n_fires = n()) %>%
  left_join(state.x77, by = 'region') %>%
  mutate(fires_per_sqm = n_fires / Area) %>%
  right_join(state_map, by = 'region') %>%
  ggplot(aes(
    x = long,
    y = lat,
    group = group,
    fill = fires_per_sqm
  )) +
  geom_polygon() +
  geom_path(color = 'white') +
  scale_fill_continuous(low = "orange",
                        high = "firebrick3",
```

```

            name = 'Fires per \nsquare mile') +
theme_map() +
coord_map('albers', lat0 = 30, lat1 = 40) +
ggtitle("Wildfires per Square Mile by 1992–2015") +
theme(plot.title = element_text(hjust = 0.5))

```

Wildfires per Square Mile by 1992–2015



#Once we adjust for size, it seems that New Jersey, New York, Georgia, South Carolina, and North Carolina have the most wildfires.

```

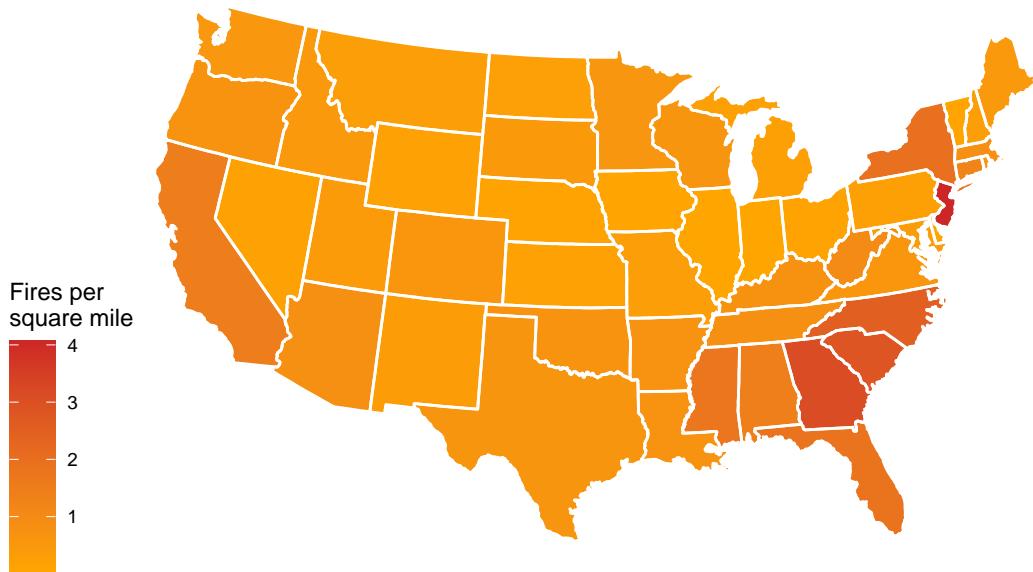
state_normalized_func <- function(cause) {
  usa_fires %>%
    #filter(STAT_CAUSE_DESCR == cause) %>%
    group_by(region) %>%
    summarize(n_fires = n()) %>%
    left_join(state.x77, by = 'region') %>%
    mutate(fires_per_sqm = n_fires / Area) %>%
    right_join(state_map, by = 'region') %>%
    ggplot(aes(
      x = long,
      y = lat,
      group = group,
      fill = fires_per_sqm
    )) +
    geom_polygon() +
    geom_path(color = 'white') +
    scale_fill_continuous(low = "orange",

```

```
        high = "firebrick3",
        name = 'Fires per \nsquare mile') +
theme_map() +
coord_map('albers', lat0 = 30, lat1 = 40) +
ggtitle(paste0("Wildfires Caused by ", cause, " per Square Mile 1992-2015")) +
theme(plot.title = element_text(hjust = 0.5))
}
```

```
state_normalized_func(cause = 'Arson')
```

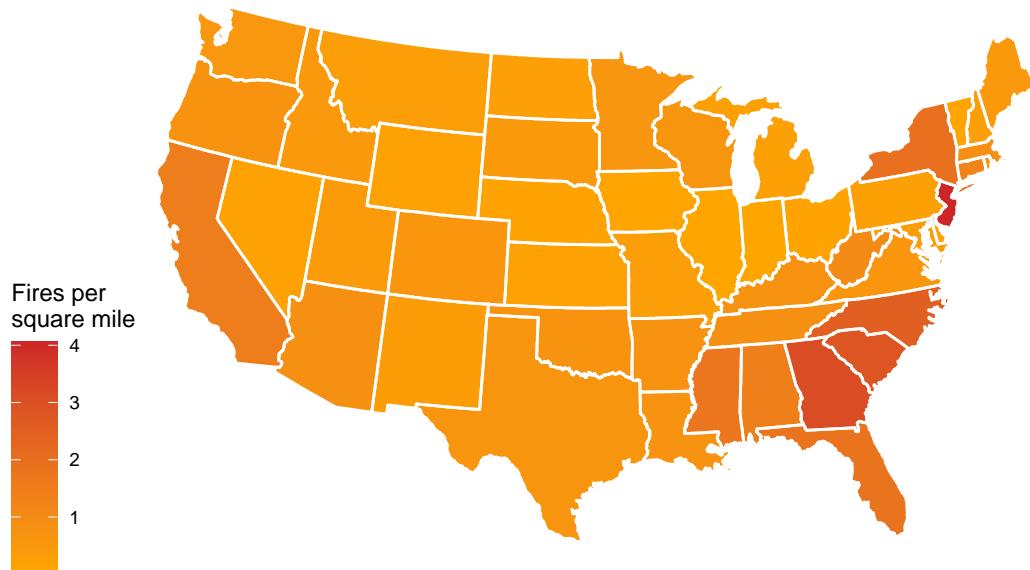
Wildfires Caused by Arson per Square Mile 1992–2015



## Arson

```
state_normalized_func(cause = 'Campfire')
```

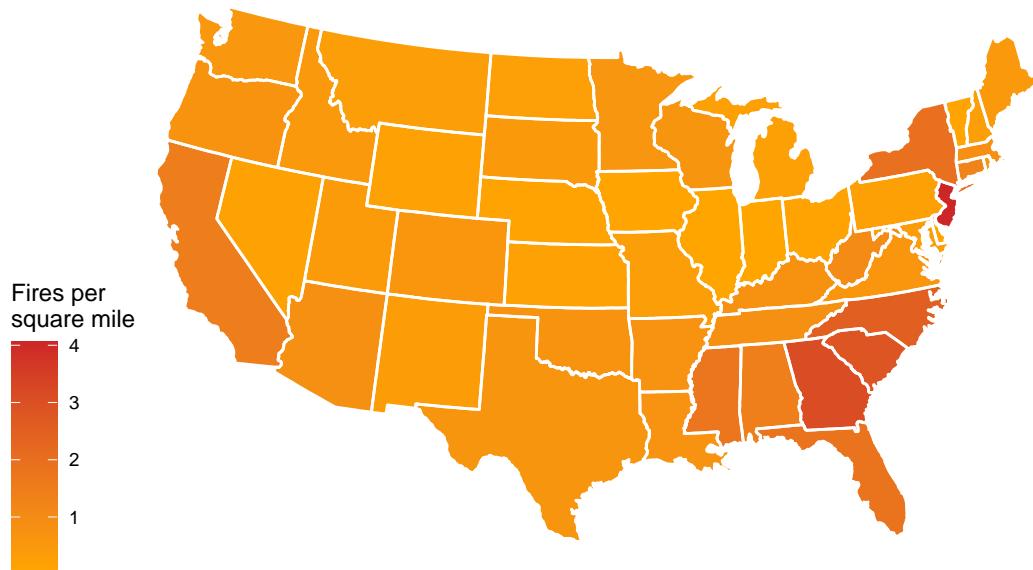
Wildfires Caused by Campfire per Square Mile 1992–2015



Campfire

```
state_normalized_func(cause = 'Children')
```

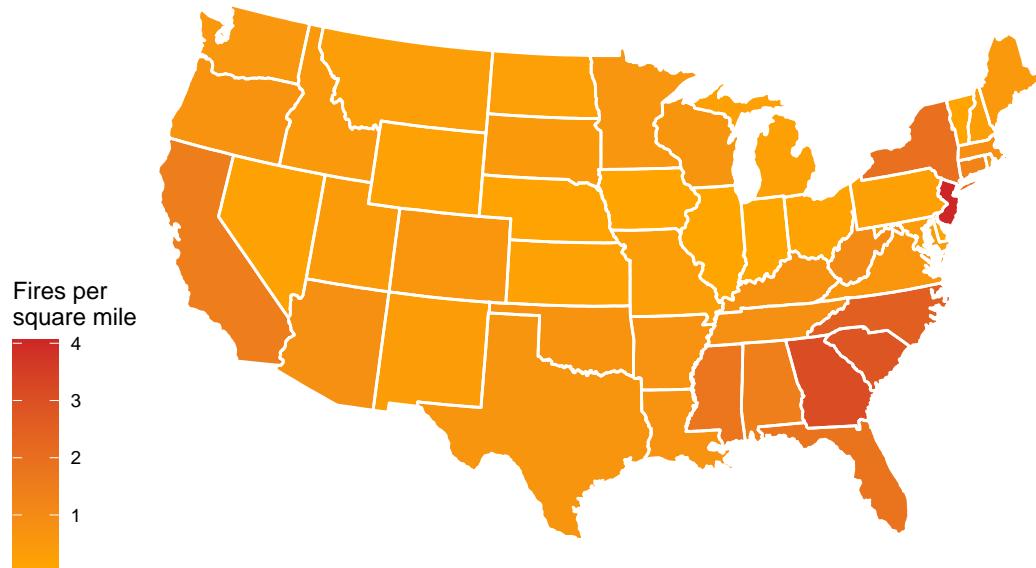
Wildfires Caused by Children per Square Mile 1992–2015



Children

```
state_normalized_func(cause = 'Debris Burning')
```

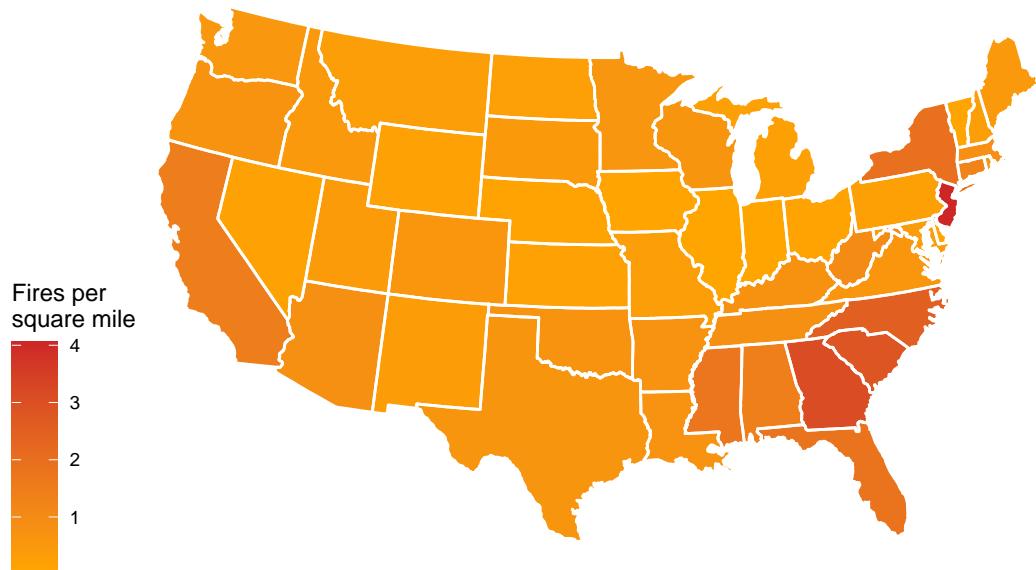
Wildfires Caused by Debris Burning per Square Mile 1992–2015



### Debris Burning

```
state_normalized_func(cause = 'Equipment Use')
```

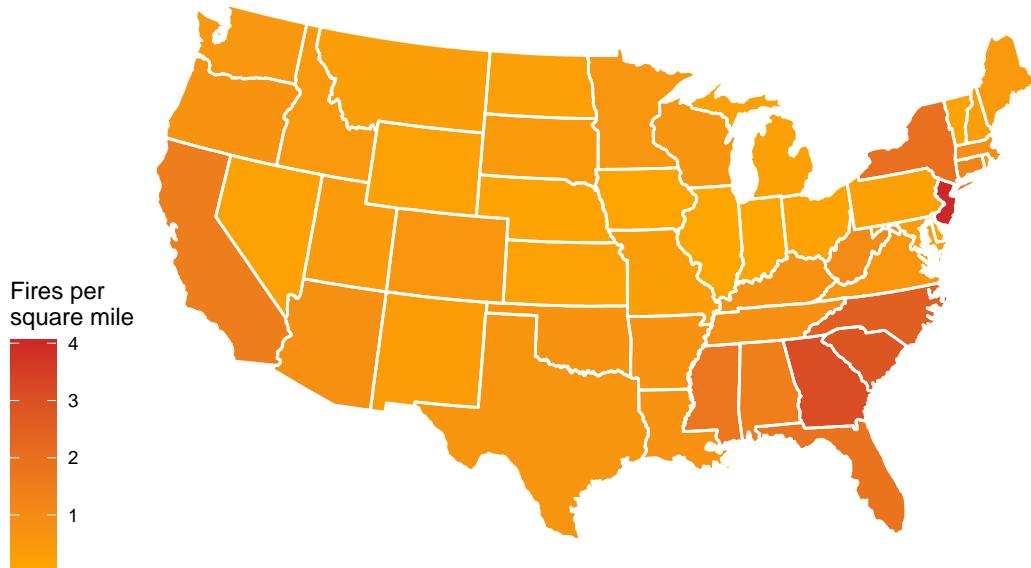
Wildfires Caused by Equipment Use per Square Mile 1992–2015



### Equipment Use

```
state_normalized_func(cause = 'Fireworks')
```

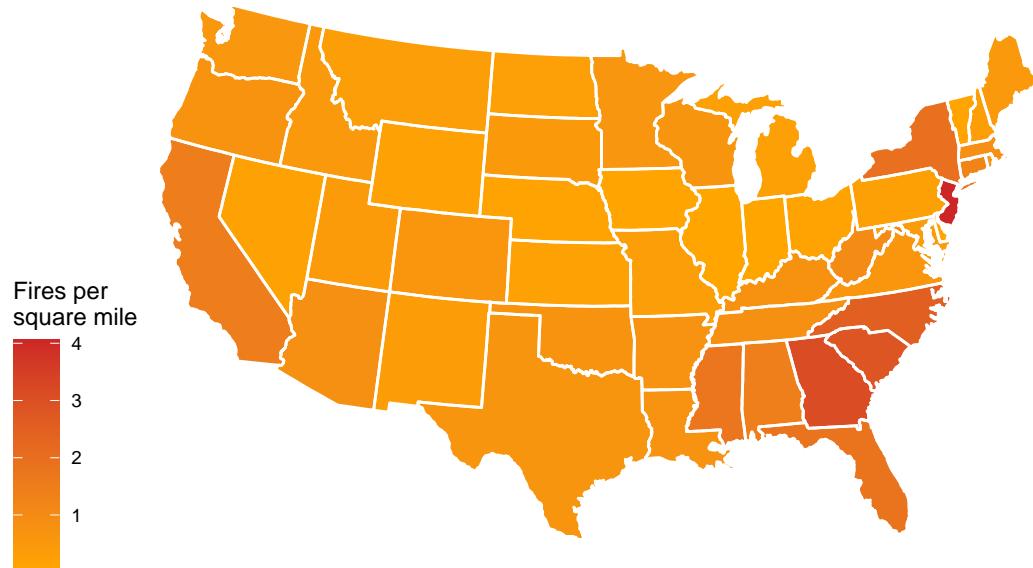
Wildfires Caused by Fireworks per Square Mile 1992–2015



Fireworks

```
state_normalized_func(cause = 'Lightning')
```

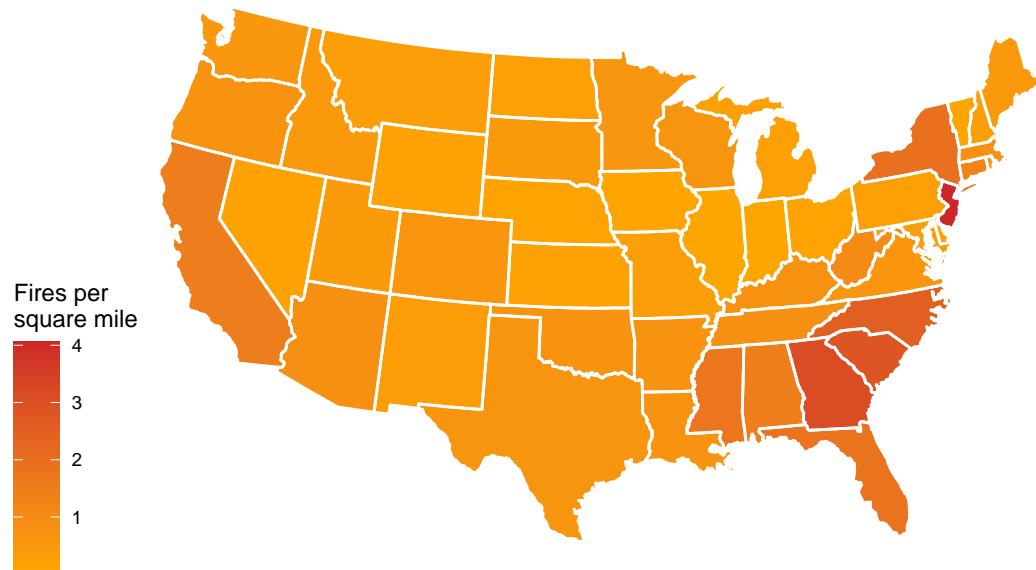
Wildfires Caused by Lightning per Square Mile 1992–2015



## Lightning

```
state_normalized_func(cause = 'Powerline')
```

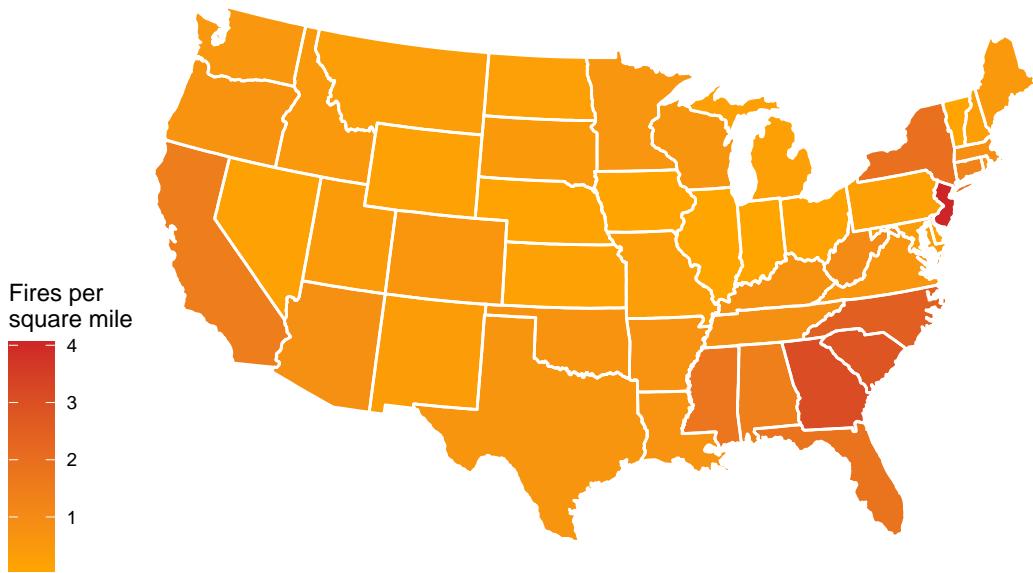
Wildfires Caused by Powerline per Square Mile 1992–2015



Powerline

```
state_normalized_func(cause = 'Railroad')
```

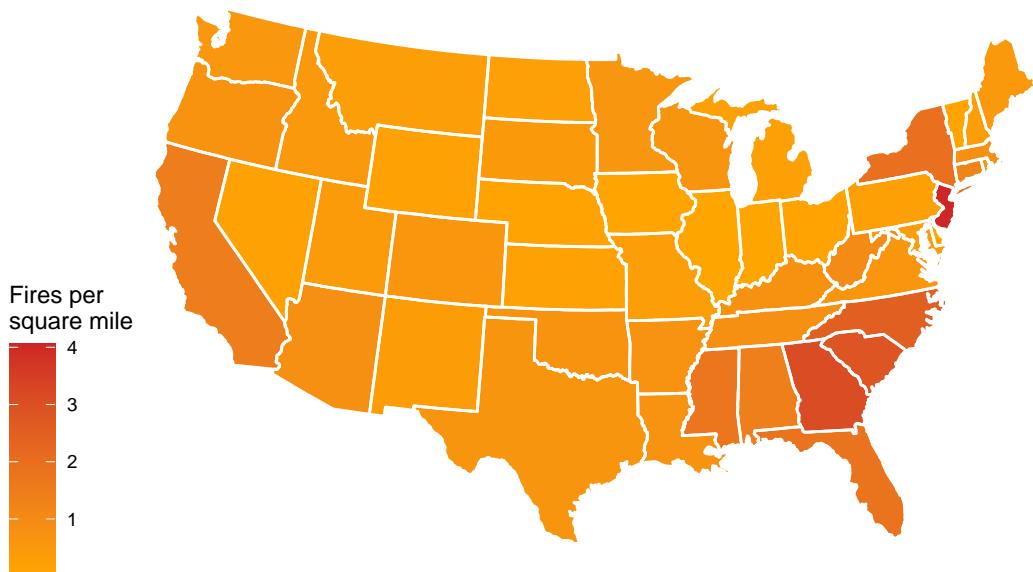
Wildfires Caused by Railroad per Square Mile 1992–2015



Railroad

```
state_normalized_func(cause = 'Smoking')
```

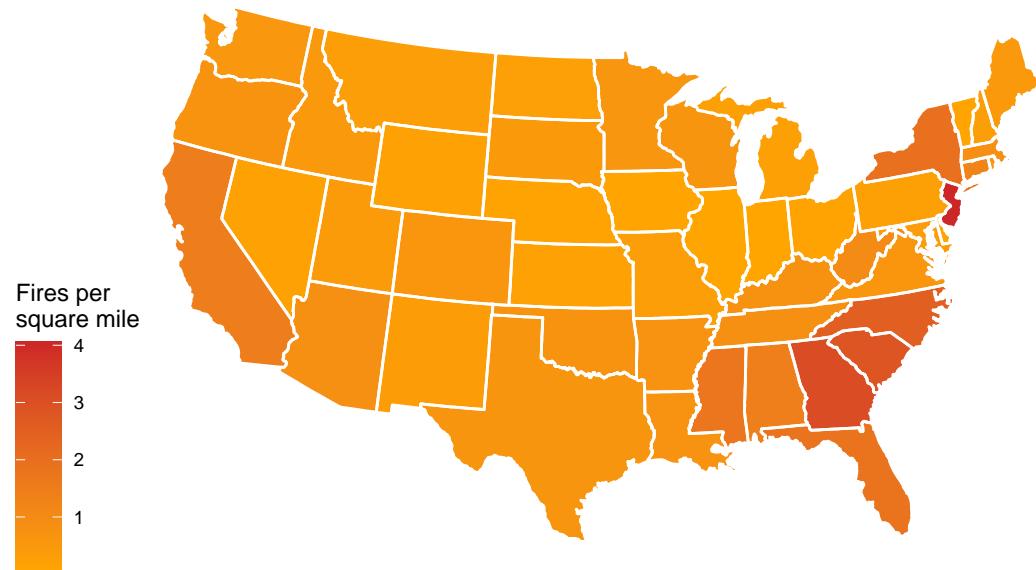
Wildfires Caused by Smoking per Square Mile 1992–2015



Smoking

```
state_normalized_func(cause = 'Structure')
```

## Wildfires Caused by Structure per Square Mile 1992–2015



## Structure

## Observation

Even though it is said that California has more number of wildfires, it is New Jersey that has more wildfires per square mile.