ANA 515 Week 06, Assignment 3-Storm Event Analysis

SACHINSINGH TOMAR

2024-07-01

#code for to read local file and assign for further process.  
stormevent\_data <- "/Users/sachintomar/Desktop/McDaniel/ANA 515/Week 06//StormEventsDetails.csv" # Code for if your want to read file from local folder.  
  
# Read the dataset from the URL  
storm\_data <- read\_csv(stormevent\_data,show\_col\_types = FALSE) #read\_csv is part of the readr package, which is included in the tidyverse collection of R packages  
  
#display the first few rows of the dataframe  
head(storm\_data)

## # A tibble: 6 × 51  
## BEGIN\_YEARMONTH BEGIN\_DAY BEGIN\_TIME END\_YEARMONTH END\_DAY END\_TIME EPISODE\_ID  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 199704 21 1727 199704 21 1727 2402786  
## 2 199704 21 1730 199704 21 1730 2402786  
## 3 199704 21 1630 199704 21 1630 2402786  
## 4 199707 17 500 199707 17 500 2402790  
## 5 199707 16 2354 199707 16 2354 2402790  
## 6 199707 17 515 199707 17 515 2402790  
## # ℹ 44 more variables: EVENT\_ID <dbl>, STATE <chr>, STATE\_FIPS <dbl>,  
## # YEAR <dbl>, MONTH\_NAME <chr>, EVENT\_TYPE <chr>, CZ\_TYPE <chr>,  
## # CZ\_FIPS <dbl>, CZ\_NAME <chr>, WFO <chr>, BEGIN\_DATE\_TIME <chr>,  
## # CZ\_TIMEZONE <chr>, END\_DATE\_TIME <chr>, INJURIES\_DIRECT <dbl>,  
## # INJURIES\_INDIRECT <dbl>, DEATHS\_DIRECT <dbl>, DEATHS\_INDIRECT <dbl>,  
## # DAMAGE\_PROPERTY <chr>, DAMAGE\_CROPS <chr>, SOURCE <chr>, MAGNITUDE <dbl>,  
## # MAGNITUDE\_TYPE <lgl>, FLOOD\_CAUSE <lgl>, CATEGORY <lgl>, …

# Select the desired columns  
selected\_columns <- c(  
 'BEGIN\_YEARMONTH',  
 'EPISODE\_ID',  
 'STATE',  
 'STATE\_FIPS',  
 'CZ\_NAME', # County name  
 'CZ\_TYPE',  
 'CZ\_FIPS',  
 'EVENT\_TYPE'  
)  
  
# Limit the dataframe to the selected columns  
limited\_stormdata <- storm\_data %>% select(all\_of(selected\_columns))  
  
# Display the first few rows of the limited dataframe  
head(limited\_stormdata)

## # A tibble: 6 × 8  
## BEGIN\_YEARMONTH EPISODE\_ID STATE STATE\_FIPS CZ\_NAME CZ\_TYPE CZ\_FIPS EVENT\_TYPE  
## <dbl> <dbl> <chr> <dbl> <chr> <chr> <dbl> <chr>   
## 1 199704 2402786 TENN… 47 LAWREN… C 99 Hail   
## 2 199704 2402786 TENN… 47 LAWREN… C 99 Hail   
## 3 199704 2402786 TENN… 47 GILES C 55 Hail   
## 4 199707 2402790 KANS… 20 SEDGWI… C 173 Hail   
## 5 199707 2402790 KANS… 20 SEDGWI… C 173 Hail   
## 6 199707 2402790 KANS… 20 SEDGWI… C 173 Thunderst…

# Arrange the dataframe by the state name  
arranged\_stormdata <- limited\_stormdata %>% arrange(STATE)  
  
# Display the first few rows of the arranged dataframe  
head(arranged\_stormdata)

## # A tibble: 6 × 8  
## BEGIN\_YEARMONTH EPISODE\_ID STATE STATE\_FIPS CZ\_NAME CZ\_TYPE CZ\_FIPS EVENT\_TYPE  
## <dbl> <dbl> <chr> <dbl> <chr> <chr> <dbl> <chr>   
## 1 199707 2066828 ALAB… 1 COLBERT C 33 Hail   
## 2 199708 2066984 ALAB… 1 CALHOUN Z 19 Cold/Wind…  
## 3 199707 2066827 ALAB… 1 LAUDER… C 77 Hail   
## 4 199708 2066992 ALAB… 1 MONTGO… C 101 Flash Flo…  
## 5 199707 2066242 ALAB… 1 WASHIN… C 129 Thunderst…  
## 6 199707 2066243 ALAB… 1 MOBILE C 97 Thunderst…

arranged\_stormdata <- arranged\_stormdata %>%  
 mutate(  
 STATE = str\_to\_title(STATE),  
 CZ\_NAME = str\_to\_title(CZ\_NAME)  
 )  
  
# Display the first few rows of the modified data  
head(arranged\_stormdata)

## # A tibble: 6 × 8  
## BEGIN\_YEARMONTH EPISODE\_ID STATE STATE\_FIPS CZ\_NAME CZ\_TYPE CZ\_FIPS EVENT\_TYPE  
## <dbl> <dbl> <chr> <dbl> <chr> <chr> <dbl> <chr>   
## 1 199707 2066828 Alab… 1 Colbert C 33 Hail   
## 2 199708 2066984 Alab… 1 Calhoun Z 19 Cold/Wind…  
## 3 199707 2066827 Alab… 1 Lauder… C 77 Hail   
## 4 199708 2066992 Alab… 1 Montgo… C 101 Flash Flo…  
## 5 199707 2066242 Alab… 1 Washin… C 129 Thunderst…  
## 6 199707 2066243 Alab… 1 Mobile C 97 Thunderst…

rows\_arranged = nrow(arranged\_stormdata)  
rows\_arranged

## [1] 41991

# Filter the events to include only those with CZ\_TYPE of 'C'  
filtered\_storm\_events <- arranged\_stormdata %>%  
 filter(CZ\_TYPE == "C") %>%  
 select(-CZ\_TYPE)  
  
# Display the first few rows of the filtered dataframe  
head(filtered\_storm\_events)

## # A tibble: 6 × 7  
## BEGIN\_YEARMONTH EPISODE\_ID STATE STATE\_FIPS CZ\_NAME CZ\_FIPS EVENT\_TYPE   
## <dbl> <dbl> <chr> <dbl> <chr> <dbl> <chr>   
## 1 199707 2066828 Alabama 1 Colbert 33 Hail   
## 2 199707 2066827 Alabama 1 Lauderdale 77 Hail   
## 3 199708 2066992 Alabama 1 Montgomery 101 Flash Flood   
## 4 199707 2066242 Alabama 1 Washington 129 Thunderstorm…  
## 5 199707 2066243 Alabama 1 Mobile 97 Thunderstorm…  
## 6 199707 2066244 Alabama 1 Baldwin 3 Tornado

rows\_filtered = nrow(filtered\_storm\_events)  
rows\_filtered

## [1] 25078

# Pad the state and county FIPS with a "0" at the beginning  
filtered\_storm\_events <- filtered\_storm\_events %>%  
 mutate(  
 STATE\_FIPS = str\_pad(STATE\_FIPS, width = 3, side = "left", pad = "0"),  
 CZ\_FIPS = str\_pad(CZ\_FIPS, width = 3, side = "left", pad = "0")  
 )  
  
# Unite the state and county FIPS into one FIPS column  
filtered\_storm\_events <- filtered\_storm\_events %>%  
 unite("FIPS", STATE\_FIPS, CZ\_FIPS, sep = "")  
  
# Display the first few rows  
head(filtered\_storm\_events)

## # A tibble: 6 × 6  
## BEGIN\_YEARMONTH EPISODE\_ID STATE FIPS CZ\_NAME EVENT\_TYPE   
## <dbl> <dbl> <chr> <chr> <chr> <chr>   
## 1 199707 2066828 Alabama 001033 Colbert Hail   
## 2 199707 2066827 Alabama 001077 Lauderdale Hail   
## 3 199708 2066992 Alabama 001101 Montgomery Flash Flood   
## 4 199707 2066242 Alabama 001129 Washington Thunderstorm Wind  
## 5 199707 2066243 Alabama 001097 Mobile Thunderstorm Wind  
## 6 199707 2066244 Alabama 001003 Baldwin Tornado

# Change all column names to lower case  
filtered\_storm\_events <- filtered\_storm\_events %>%  
 rename\_all(tolower)  
  
# Display the first few rows of the dataframe with lower case column names  
head(filtered\_storm\_events)

## # A tibble: 6 × 6  
## begin\_yearmonth episode\_id state fips cz\_name event\_type   
## <dbl> <dbl> <chr> <chr> <chr> <chr>   
## 1 199707 2066828 Alabama 001033 Colbert Hail   
## 2 199707 2066827 Alabama 001077 Lauderdale Hail   
## 3 199708 2066992 Alabama 001101 Montgomery Flash Flood   
## 4 199707 2066242 Alabama 001129 Washington Thunderstorm Wind  
## 5 199707 2066243 Alabama 001097 Mobile Thunderstorm Wind  
## 6 199707 2066244 Alabama 001003 Baldwin Tornado

# Load the data on U.S. states  
data("state")  
  
# Create a dataframe with state name, area, and region  
US\_states <- data.frame(  
 US\_State = state.name,  
 Area = state.area,  
 Region = state.region  
)  
  
# Display the first few rows of the states dataframe  
head(US\_states)

## US\_State Area Region  
## 1 Alabama 51609 South  
## 2 Alaska 589757 West  
## 3 Arizona 113909 West  
## 4 Arkansas 53104 South  
## 5 California 158693 West  
## 6 Colorado 104247 West

# Create a dataframe with the number of events per state  
events\_per\_state <- filtered\_storm\_events %>%  
 group\_by(state) %>%  
 summarise(number\_of\_events = n())  
  
# Display the first few rows of the events per state dataframe  
head(events\_per\_state)

## # A tibble: 6 × 2  
## state number\_of\_events  
## <chr> <int>  
## 1 Alabama 652  
## 2 Alaska 12  
## 3 Arizona 248  
## 4 Arkansas 628  
## 5 California 283  
## 6 Colorado 586

# Merge the events\_per\_state dataframe with the states\_data dataframe  
merged\_data <- merge(events\_per\_state, US\_states, by.x = "state", by.y = "US\_State")  
  
# Display the first few rows of the merged dataframe  
head(merged\_data)

## state number\_of\_events Area Region  
## 1 Alabama 652 51609 South  
## 2 Alaska 12 589757 West  
## 3 Arizona 248 113909 West  
## 4 Arkansas 628 53104 South  
## 5 California 283 158693 West  
## 6 Colorado 586 104247 West

# Create the scatter plot  
ggplot(merged\_data, aes(x = Area, y = number\_of\_events, color = Region)) +  
 geom\_point(size = 3) +  
 scale\_color\_manual(values = c(  
 "Northeast" = "red",  
 "South" = "green",  
 "North Central" = "blue",  
 "West" = "purple"  
 )) +  
 labs(  
 title = "Number of Storm Events vs Land Area",  
 x = "Land Area (square miles)",  
 y = "Number of Storm Events"  
 ) +  
 theme\_minimal()

