wildfire Analysis in British Columbia (B.C.)

AUTHOR
Rakesh Das

PUBLISHED
November 9, 2024

Introduction

In this analysis, I have comprehensively addressed the task of examining wildfire occurrences in British Columbia (B.C.) over time, comparing historical data from 2012-2017 with current data from 2024. By scraping and analyzing the historical wildfire data, and integrating temperature and elevation data using the geodata package, I have visualized wildfire distributions over time. The analysis also includes a comparison of how temperature and elevation impact fire occurrences, using carefully crafted visualizations to illustrate these relationships. Additionally, I have explored the differences in wildfire distribution between the historical period and the current year, providing insights into the factors influencing wildfire behavior. All findings are supported by standalone visualizations and thorough written descriptions that discuss data and analysis limitations.

Installing and importing the required packages

```r
library(geodata)  # Geospatial data
```

```
Loading required package: terra
```

```
terra 1.7.83
```

```r
library(sf)       # Simple Features for R
```

```
Linking to GEOS 3.12.1, GDAL 3.8.4, PROJ 9.3.1; sf_use_s2() is TRUE
```

```r
library(terra)    # Spatial data analysis
library(raster)   # Raster data analysis
```

```
Loading required package: sp
```

```r
library(spData)   # Spatial data for R
library(rvest)    # Web scraping
library(ggplot2)  # Data visualization
library(patchwork) # Combine multiple ggplot2 plots
```

```

Attaching package: 'patchwork'

The following object is masked from 'package:raster':

    area

The following object is masked from 'package:terra':
```

```
    area
```

```r
library(gganimate) # Animated plots
```

```
Attaching package: 'gganimate'

The following object is masked from 'package:raster':

    animate

The following object is masked from 'package:terra':

    animate
```

```r
library(ggspatial) # Add spatial context to plots
library(rnaturalearth) # Natural Earth data
library(dplyr)    # Data manipulation
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:raster':

    intersect, select, union

The following objects are masked from 'package:terra':

    intersect, union

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```r
library(ggpubr)
```

```
Warning: package 'ggpubr' was built under R version 4.4.2


Attaching package: 'ggpubr'

The following object is masked from 'package:raster':

    rotate

The following object is masked from 'package:terra':

    rotate
```

```
library(cowplot)
```

```
Warning: package 'cowplot' was built under R version 4.4.2


Attaching package: 'cowplot'

The following object is masked from 'package:ggpubr':

    get_legend

The following object is masked from 'package:patchwork':

    align_plots
```

1) Web Scraping B.C. Wildfire Data (2012-2017)

Data Collection

- We are Scraping the British Columbia government's website to access historical wildfire data from 2012 to 2017 using rvest library.
- html_table function to extracts tables from the HTML content.The first table in the list contains the wildfire data, hence it is assigned to the wildfire_table variable.

```
url <- 'https://www2.gov.bc.ca/gov/content/safety/wildfire-status/about-bcws/wildfire-statisti
wildfire_html <- read_html(url)

tables <- wildfire_html %>% html_table()
wildfire_table <- tables[[1]]
head(wildfire_table)
```

```
# A tibble: 6 × 8
   Year `Fire Number` `Fire Centre` Latitude  Longitude  Geographic
  <int> <chr>         <chr>         <chr>     <chr>      <chr>
1  2017 C50011 (2017) Cariboo       51 56.952 123 08.113 Lee's Corner
2  2017 C50043 (2017) Cariboo       51 59.384 123 09.690 3 Km East Anaham Reser…
3  2017 C10248 (2017) Cariboo       53 10.077 123 02.022 Pantage Lake
4  2017 C40329 (2017) Cariboo       51 45.338 121 50.060 6km SE of Place Lake
5  2017 C20352 (2017) Cariboo       51 44.568 121 59.996 4 km W of Emerald Lake
6  2017 C20414 (2017) Cariboo       52 00.900 122 54.032 Raven Lake
# ℹ 2 more variables: `Discovery Date` <chr>, `Size (ha)` <dbl>
```

Data Pre-processing

checking for missing values

After extracting the data into a dataframe (wildfire_table), we need to identify any missing values.

```
sum(is.na(wildfire_table))
```

```
[1] 0
```

Renaming Columns

To improve the readability and understandability of the data, we are renaming the table columns using 'rename' function .

```
library(dplyr)
wildfire_table <- wildfire_table %>%
    rename(
        year = `Year`,
        fire_number = `Fire Number`,
        fire_centre = `Fire Centre`,
        latitude = `Latitude`,
        longitude = `Longitude`,
        geographic = `Geographic`,
        discovery_date = `Discovery Date`,
        size_ha = `Size (ha)`
    )
```

Convert Latitude and Longitude to Decimal Format

The mutate function is used to modify the latitude and longitude columns within the wildfire_table.

The convert_to_decimal function is applied to both columns, converting them to decimal format.

Since British Columbia is west of the Prime Meridian, a negative sign is assigned to the converted longitude values.

The current format for the discovery_date is getting converted it into a standard date format suitable for time-series analysis.

```
library(dplyr)

convert_to_decimal <- function(coord) {
  parts <- strsplit(as.character(coord), " ")
  degrees <- as.numeric(sapply(parts, `[`, 1))
  minutes <- as.numeric(sapply(parts, `[`, 2))
  decimal <- degrees + minutes / 60
  return(decimal)
}

wildfire_table <- wildfire_table %>%
  mutate(
    latitude = convert_to_decimal(latitude),
    longitude = -convert_to_decimal(longitude) # Negative for west longitude
  )
```

```
wildfire_table <- wildfire_table %>%
  mutate(discovery_date = as.Date(discovery_date, format = "%B %d, %Y"))
```

A new column named month is creared from the existing discovery_date column, which has been converted to a standard date format earlier.

Using the group_by and summarise function we are calculating total number of wildfires per month and per year.

From the counts, it is significant that no fire incidents were recorded in January and December, and only one incident occurred each in February and November.

In the next steps, we can proceed with the analysis excluding these months.

```
wildfire_table <- wildfire_table %>% mutate(month = format(as.Date(discovery_date, "%Y-%m-%d")
fire_counts_per_month <- wildfire_table %>% group_by(month) %>% summarise(count = n())
fire_counts_per_year <- wildfire_table %>% group_by(year) %>% summarise(count = n())
```

2) Pull temperature and elevation data for B.C using the geodata package. Crop and mask as needed.

Obtaining Elevation and Temperature Data:

The elevation_30s function from the geodata package is used to download a 30-second resolution elevation raster for Canada.

The worldclim_country function is used to download average temperature data for Canada.

With the gadm function from the geodata package, a shapefile of Canadian provinces and territories at level 1 is obtained and filtered for ‘British Columbia’.

st_transform function from the sf package is used to reproject the British Columbia shapefile to the same coordinate reference system (CRS) as the elevation data.

```
bc_elevation <- elevation_30s(country = "CAN", level = 1, path = tempdir())
bc_temp <- worldclim_country(country = "CAN", var = "tavg", path = tempdir())
bc_shape <- gadm(country = "CAN", level = 1, path = tempdir())
bc_shape <- bc_shape[bc_shape$NAME_1 == "British Columbia", ]
bc_shape_sf <- sf::st_as_sf(bc_shape)
bc_shape <- sf::st_transform(bc_shape_sf, st_crs(bc_elevation))
```

Cropping and Masking Elevation and Temperature Data:

crop function from the raster package is used to extract a portion of the bc_elevation and bc_temp raster that aligns with the extent of the bc_shape (British Columbia’s boundary).

The mask function from the raster package is employed to set the values of pixels outside the bc_shape to NA to isolating the elevation and temp data within British Columbia's boundaries.

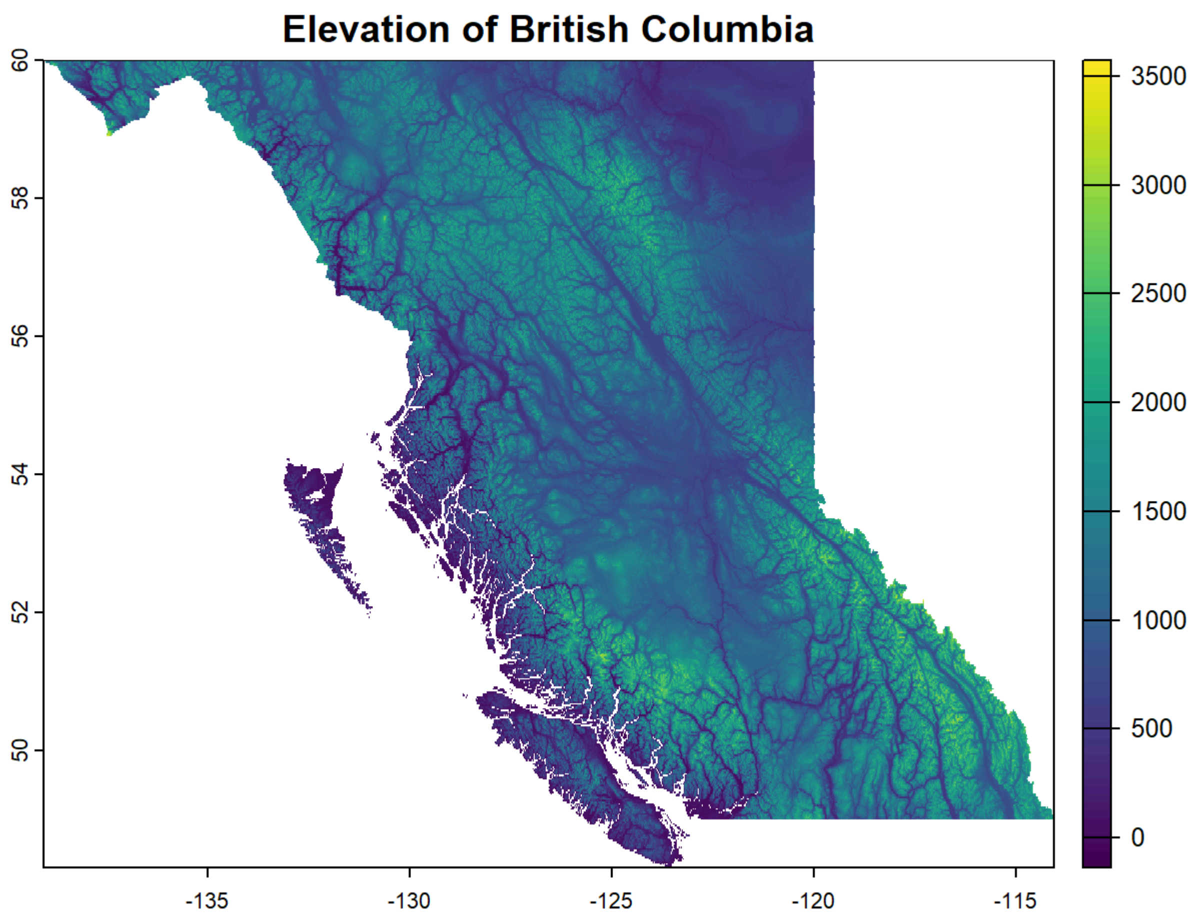
```
bc_elevation_cropped <- crop(bc_elevation, bc_shape)

bc_elevation_masked <- mask(bc_elevation_cropped, bc_shape)

bc_temp_cropped <- crop(bc_temp, bc_shape)
```

```
|---------|---------|---------|---------|
=========================================
```

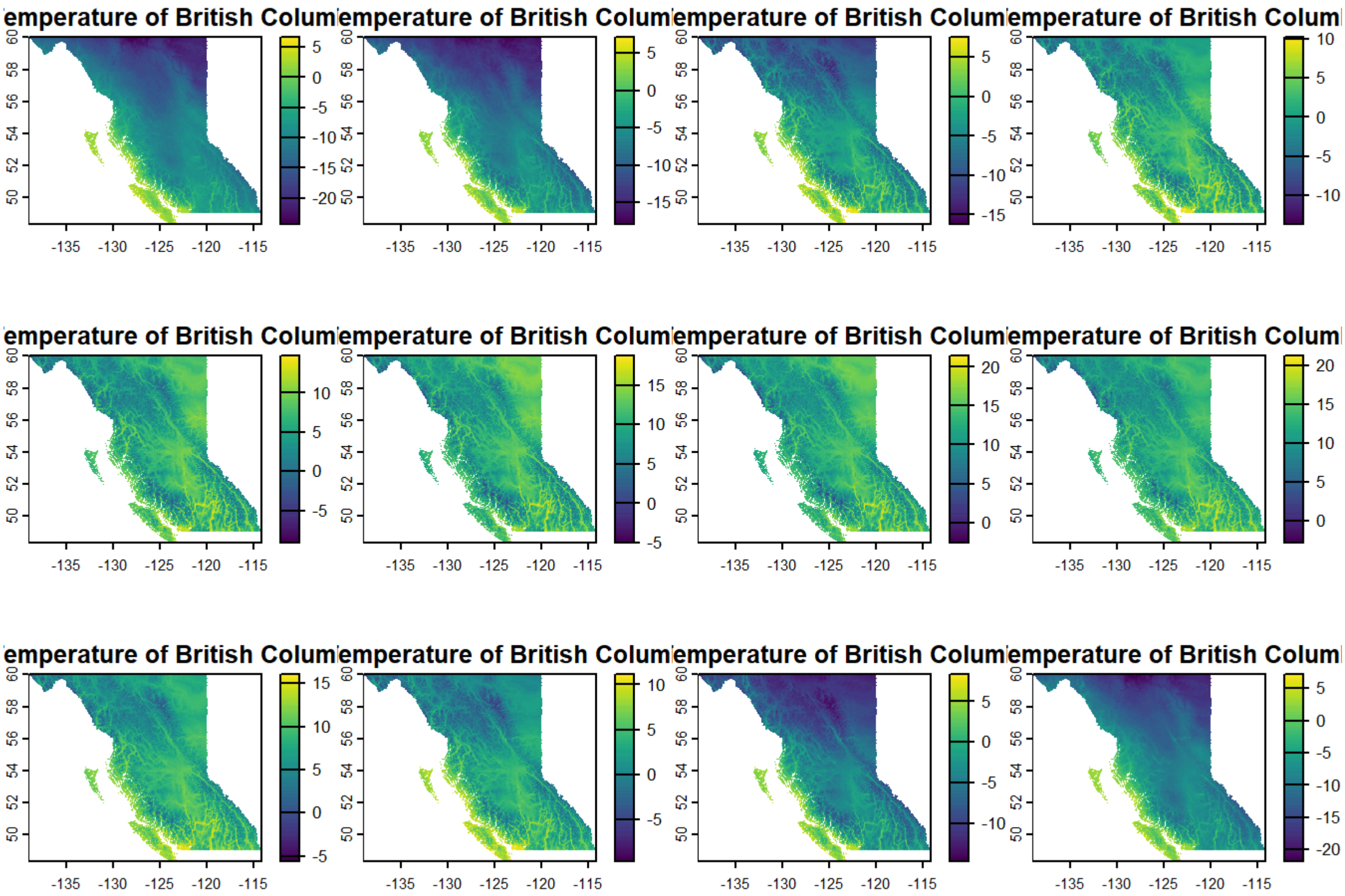
```
bc_temp_masked <- mask(bc_temp_cropped, bc_shape)
```

```
|---------|---------|---------|---------|
=========================================
```

```
# Plot the masked elevation data
plot(bc_elevation_masked, main = "Elevation of British Columbia")
```



```
# Plot the masked temperature data
plot(bc_temp_masked, main = "Mean Temperature of British Columbia")
```



The st_as_sf function from the sf package is used to convert the wildfire_table into a spatial points data frame (SF).

The st_transform function is used to reproject the wildfire points to the same coordinate reference system as the temperature and elevation rasters.

```
wildfire_sf <- wildfire_table %>%
  st_as_sf(coords = c("longitude", "latitude"),
           crs = 4326)

# Transforms CRS of wildfire data to match temperature raster
wildfire_sf_transformed <- st_transform(wildfire_sf, crs = st_crs(bc_temp_masked))
```

Using the function st_read function is used to read the current fire point, fire perimeter data from the downloaded data from the BC govt ‘Statistics & Geospatial Data’ website.

```
current_fire_perimeters_sf <- st_read(dsn="C:\\Users\\ASUS\\Documents\\Masters - Guelph\\Fall

current_fire_points_sf <- st_read(dsn="C:\\Users\\ASUS\\Documents\\Masters - Guelph\\Fall 2024
```

The st_transform function is used to convert the CRS of the wildfire_sf_transformed (which contains historical wildfire data) to match the coordinate reference system of the current fire perimeter/ fire point data.

```
wildfire_historical_sf <- st_transform(wildfire_sf_transformed , st_crs(current_fire_perimeter
```

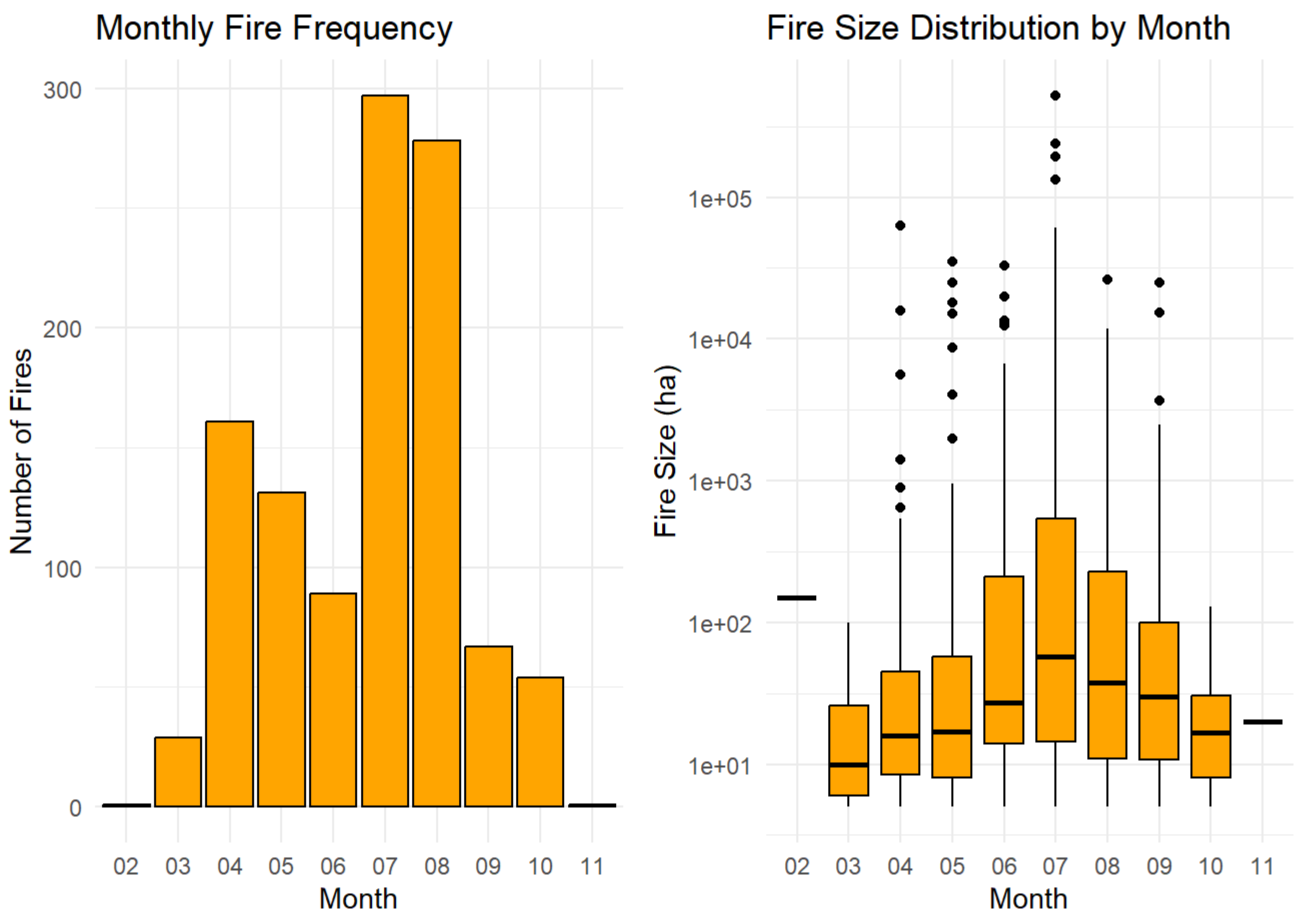
visualization of the fire frequency for each month

Monthly Fire Frequency Plot: - A bar plot is generatedto display the number of fires for each month.

Fire Size Distribution Plot: - A box plot is created to visualize the distribution of fire sizes for each month.

- The `patchwork` package is used to combine the two plots side-by-side, allowing for a comparative analysis of fire frequency and size distribution across different months.

```
# Count the number of fires for each month
monthly_fire_counts <- wildfire_table %>%
  group_by(month) %>%
  summarise(count = n())

# Create the Monthly Fire Frequency Plot
fire_frequency_plot <- ggplot(monthly_fire_counts, aes(x = month, y = count)) +
  geom_bar(stat = "identity", fill = "orange", color = "black") +
  labs(title = "Monthly Fire Frequency",
       x = "Month",
       y = "Number of Fires") +
  theme_minimal()

# Create the Fire Size Distribution Plot
fire_size_distribution_plot <- ggplot(wildfire_table, aes(x = month, y = size_ha)) +
  geom_boxplot(fill = "orange", color = "black") +
  scale_y_log10()+
  labs(title = "Fire Size Distribution by Month",
       x = "Month",
       y = "Fire Size (ha)") +
  theme_minimal()

# Combine the plots side by side using patchwork
combined_plot <- fire_frequency_plot + fire_size_distribution_plot

# Display the combined plot
print(combined_plot)
```



The analysis suggests that higher temperatures, particularly during the summer months of July and August, significantly impact fire occurrence.

These months not only see an increase in the number of fires but also larger and more variable fire sizes.

Conversely, during cooler months, fire occurrences and sizes are generally lower. This indicates a strong correlation between temperature and wildfire occurrences.

3) Visualize wildfires over time between 2012-2017 overtop a temperature raster.

The bc_temp_masked raster layer is converted into a data frame named temp_raster_df using as.data.frame function

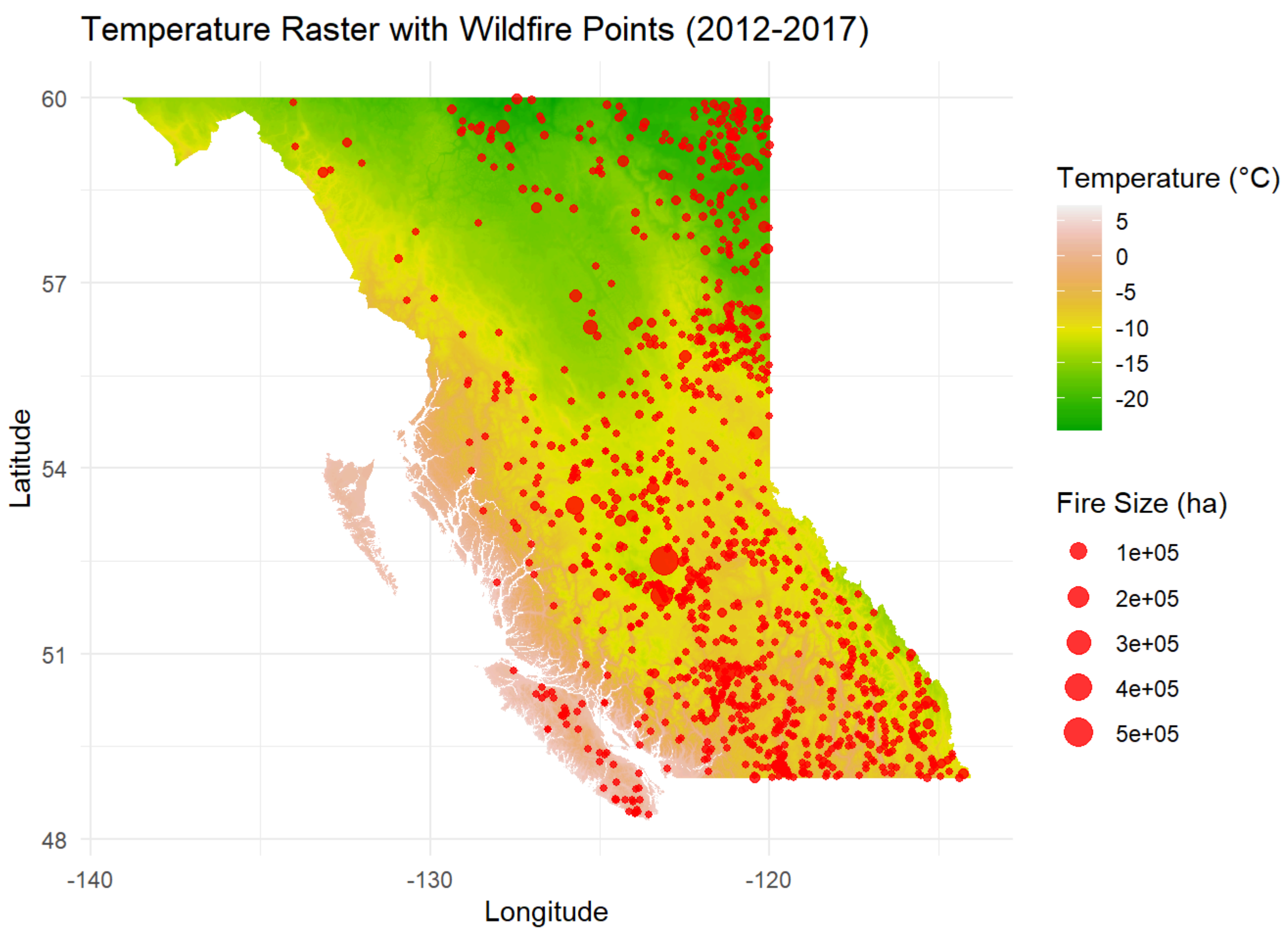
```
# Convert RasterLayer to data frame
temp_raster_df <- as.data.frame(bc_temp_masked, xy = TRUE)
colnames(temp_raster_df) <- c("x", "y", "temperature")
```

Visualization: Wildfires Overlaid on Temperature Raster

In the below section we are creating a refined plot that visualizes the relationship between temperature and wildfire occurrences in British Columbia.

-A detailed temperature gradient is craeted scale_fill_gradientn, terrain.colors(10). - geom_point layer adds the wildfire points to the temperature raster plot.

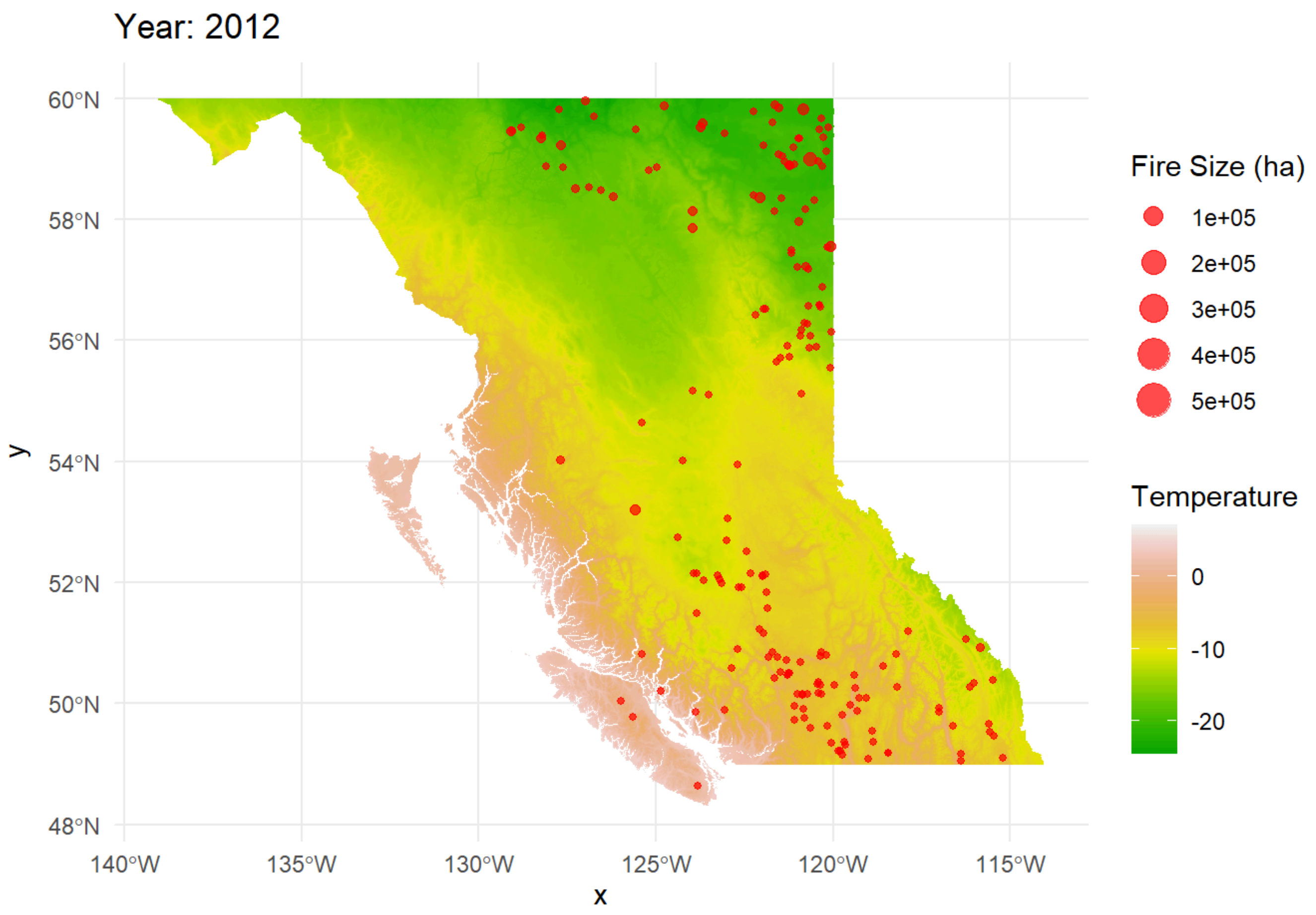
```
# Plot with refined temperature gradient
ggplot() +
  geom_raster(data = temp_raster_df,
              aes(x = x, y = y, fill = temperature)) +
  scale_fill_gradientn(
    colors = terrain.colors(10),
    limits = range(temp_raster_df$temperature, na.rm = TRUE),
    name = "Temperature (°C)",
    breaks = seq(-30, 30, by = 5)  # Adjust breaks for a more detailed scale
  ) +
  geom_point(data = wildfire_table,
             aes(x = longitude, y = latitude, size = size_ha),
             color = "red", alpha = 0.8) +
  scale_size_continuous(name = "Fire Size (ha)", range = c(1, 5)) +
  labs(
    title = "Temperature Raster with Wildfire Points (2012-2017)",
    x = "Longitude", y = "Latitude"
  ) +
  theme_minimal() +
  theme(legend.position = "right")
```



Based on observations from the visualization of wildfire occurrences and temperature gradients:

- The majority of wildfires are concentrated in regions with higher temperatures. This suggests a strong correlation between higher temperatures and increased wildfire activity.
- Larger wildfires, represented by bigger red dots, tend to occur more frequently in warmer areas. This indicates that higher temperatures may contribute to the intensity and spread of wildfires.
- These observations highlight the critical relationship between temperature and wildfire. With this information in mind, we are going to carry out the analysis further in this notebook.
- Due to huge amount of data, an animated map is created which will be visualizing the spatial distribution of wildfires in British Columbia over time, overlaid on a temperature raster.
- A base plot is created with a raster layer representing the temperature data is added using geom_raster and Wildfire locations are overlaid as points using geom_sf. The size of each point corresponds to the fire's size in hectares.The transition_states, animate function etc from the gganimate package are used to create the animation.

```
# Create the base plot
base_plot <- ggplot() +
  geom_raster(data = temp_raster_df,
              aes(x = x, y = y, fill = temperature)) +
  scale_fill_gradientn(colors = terrain.colors(10)) +
  geom_sf(data = wildfire_sf_transformed, aes(size = size_ha), color = "red", alpha = 0.7) +
  labs(title = "Temperature Raster with Wildfire Points (2012-2017)",
       fill = "Temperature",
       size = "Fire Size (ha)") +
  theme_minimal()

# Add animation transition
animated_plot <- base_plot +
  transition_states(year, transition_length = 2, state_length = 1) +
  ggtitle('Year: {closest_state}') +
  ease_aes('linear')

# Render the animation
animate(animated_plot, nframes = 6, fps = 1)
```



This animated map visually demonstrates how the spatial distribution of wildfires changes over the years, in relation to the temperature patterns.

4)how do temperature impact fire occurence?

Visualization 1:

By running this code, we obtain a visualization that allows us to visually assess potential associations between temperature and fire occurrences across different months in British Columbia.

plot_fires_for_month function is defined to create monthly plots.

The ggarrange function from the patchwork package is employed to arrange the eight individual plots (p1 to p8) into a 4x2 grid.

A common legend is displayed at the bottom of the grid for all plots.

The x-axis and y-axis tick labels are removed as this plot is primarily intended to showcase the effect of temperature on fire occurrences.

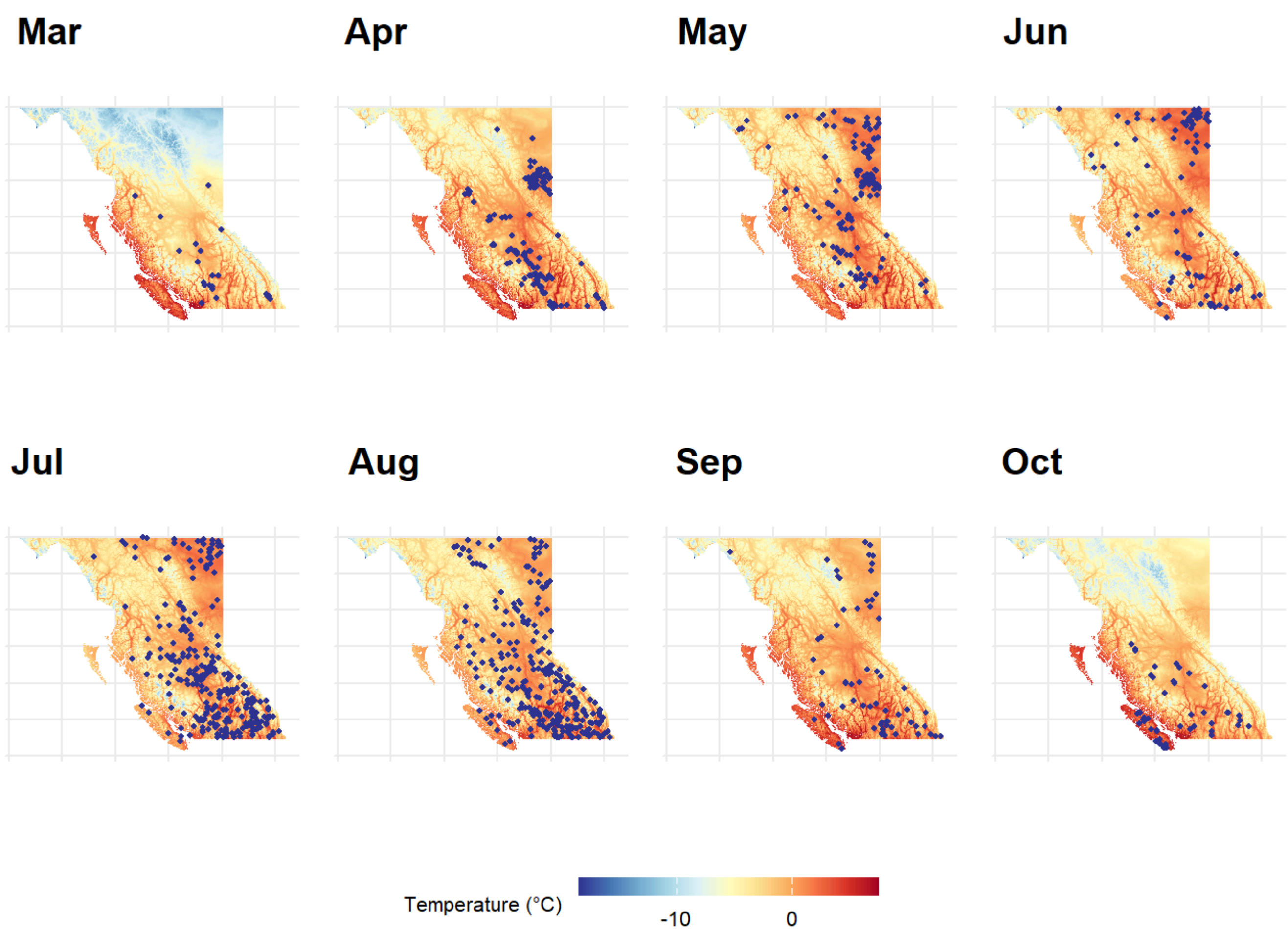
```
# Function to plot fires over temperature raster for a specific month
plot_fires_for_month <- function(month_no, wildfire_sf_transformed, bc_temp) {
  # Subset wildfire data for the given month
  wildfire_month <- wildfire_sf_transformed %>%
    filter(month == month_no)
```

```
  # Extract temperature layer for the given month
  temp_layer <- bc_temp_masked[[as.integer(month_no)]]

  # Convert raster to dataframe for ggplot
  temp_df <- as.data.frame(temp_layer, xy = TRUE)
  colnames(temp_df) <- c("x", "y", "temp")

  # Define a custom color palette for temperature
  temp_palette <- c("#313695", "#4575b4", "#74add1", "#abd9e9", "#e0f3f8", "#ffffbf", "#fee090

  # Plot using ggplot2
  ggplot() +
    geom_raster(data = temp_df, aes(x = x, y = y, fill = temp), interpolate = TRUE) +
    scale_fill_gradientn(colors = temp_palette, name = "Temperature (°C)", guide = guide_color
                         limits = range(temp_df$temp, na.rm = TRUE)
                         #oob = scales::squish
                         ) +
    geom_sf(data = wildfire_month, color = "#313695", alpha = 1, size = 0.8) +
    #scale_size_continuous(range = c(1, 10), limits = size_range) +  # Keep size scale static
    ggtitle(paste("Wildfires in Month:", month_no)) +
    theme_minimal() +
    theme(
  plot.title = element_blank(),
  axis.title.x = element_blank(),
  axis.title.y = element_blank(),
  axis.text.x = element_blank(),      # Remove x-axis tick labels
  axis.text.y = element_blank(),      # Remove y-axis tick labels
  axis.ticks.x = element_blank(),     # Remove x-axis ticks
  axis.ticks.y = element_blank(),     # Remove y-axis ticks
  legend.title = element_text(size = 8),
  legend.text = element_text(size = 8),
  legend.key.size = unit(0.5, "cm")
)
}
```

```
# Plot for a specific month, for example, July (month = "07")

p1 <- plot_fires_for_month("03", wildfire_sf_transformed, bc_temp)
p2 <- plot_fires_for_month("04", wildfire_sf_transformed, bc_temp)
p3 <- plot_fires_for_month("05", wildfire_sf_transformed, bc_temp)
p4 <- plot_fires_for_month("06", wildfire_sf_transformed, bc_temp)
p5 <- plot_fires_for_month("07", wildfire_sf_transformed, bc_temp)
p6 <- plot_fires_for_month("08", wildfire_sf_transformed, bc_temp)
p7 <- plot_fires_for_month("09", wildfire_sf_transformed, bc_temp)
p8 <- plot_fires_for_month("10", wildfire_sf_transformed, bc_temp)
```

```
ggarrange(
  p1,p2,p3,p4,p5,p6,p7,p8, labels = c("Mar","Apr","May","Jun","Jul","Aug","Sep","Oct"),
  ncol = 4, nrow = 2,
```

```
  common.legend = TRUE, legend = "bottom"
  )
```



The above plot shows the temperature distribution and fire occurrences across British Columbia for each month from March to October.

1. March to April:
   - The northern regions are predominantly cold (blue), while the southern regions are warmer (yellow to red). Few fire occurrences are observed, mostly in the southern warmer regions.
2. May to June:
   - The overall temperature increases, with more areas turning yellow to red, indicating warmer temperatures. There is a noticeable increase in fire occurrences, particularly in the central and southern regions where temperatures are higher.
3. July to August:
   - These months show the highest temperatures, with most areas in the red zone.
   - Fire occurrences peak during these months, with numerous fires spread across the region, especially in the warmer areas.
4. September to October:
   - Temperatures begin to decrease, with more areas turning yellow and some northern regions turning blue again.
   - The number of fire occurrences decreases significantly, correlating with the drop in temperature.

A strong correlation between temperature and fire occurrences can be observed. As temperatures rise from March to August, the number of fire occurrences increases, peaking in the hottest months (July and August). Conversely, as temperatures drop from September to October, fire occurrences decrease. This suggests that higher temperatures significantly impact the likelihood of fires occurring in this region.

Visualization 2: Wildfires Overlaid on Elevation Raster

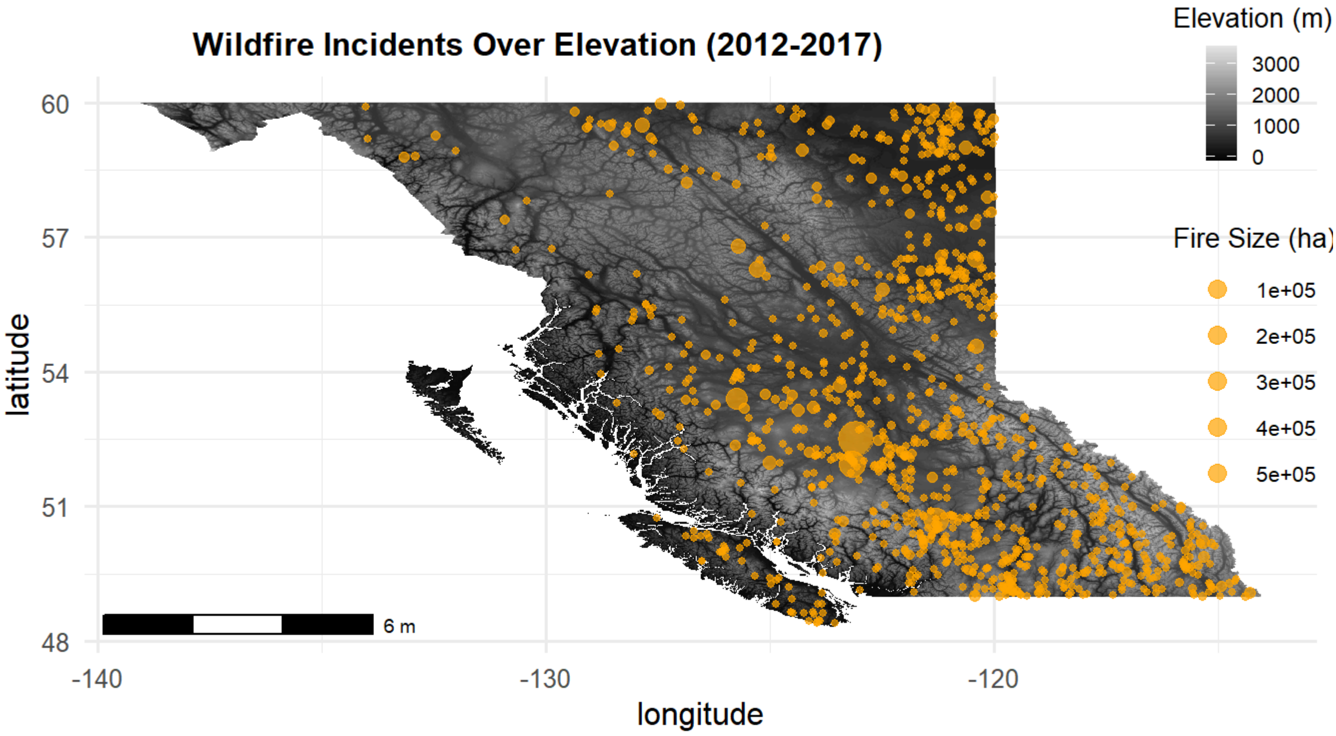
In order to assess whether there are any spatial patterns relating fire occurrences to elevation, we are creating the below plot using the elevation raster.

An elevation raster is displayed using geom_raster with the custom color palett and Wildfire locations are overlaid on the elevation map as points using geom_point.The size of each point corresponds to the fire size in hectors.

The plot is customized using annonation_scale, theme , guide etc.

```
library(RColorBrewer)

elevation_raster <- bc_elevation_masked

# Extract the elevation values
elevation_df <- as.data.frame(elevation_raster, xy = TRUE)
colnames(elevation_df) <- c("longitude", "latitude", "elevation")

# Custom color palette for elevation: grey color ramp
custom_palette <- grey(1:90 / 100)

# Plot elevation with wildfire points
ggplot() +
  geom_raster(data = elevation_df, aes(x = longitude, y = latitude, fill = elevation)) +
  scale_fill_gradientn(colors = custom_palette, name = "Elevation (m)", guide = guide_colorbar
  geom_point(data = wildfire_table,
             aes(x = longitude, y = latitude, size = size_ha),
             color = "orange", alpha = 0.7) +
  scale_size_continuous(name = "Fire Size (ha)", guide = guide_legend(override.aes = list(size
  labs(title = "Wildfire Incidents Over Elevation (2012-2017)") +
  coord_fixed() +
  theme_minimal(base_size = 12) +
  theme(
    legend.position = c(0.95, 0.7),  # Position legend inside plot (x, y) = (0.85, 0.15)
    legend.title = element_text(size = 10),
    legend.text = element_text(size = 8),
    plot.title = element_text(hjust = 0.2, size = 12, face = "bold")
  ) +
  annotation_scale(location = "bl", text_cex = 0.6)
```

```
Warning: A numeric `legend.position` argument in `theme()` was deprecated in ggplot2
3.5.0.
ℹ Please use the `legend.position.inside` argument of `theme()` instead.

Using plotunit = 'm'
```



The above plot shows wildfire incidents over different elevations , from 2012 to 2017.

The elevation is represented using a grayscale background, where darker shades indicate higher elevations.

Wildfire incidents are marked with orange dots, with the size of each dot representing the fire size in hectares (ha).

- Concentration at Lower Elevations:
  - The majority of wildfire incidents are concentrated in lower elevation areas, particularly in the northern and eastern parts of the region.
  - These areas appear lighter on the grayscale map.There are fewer wildfire incidents in higher elevation areas, which are indicated by darker shades on the map.
- Larger Fires at Lower Elevations:
  - Larger fires, depicted by bigger orange dots, are more prevalent at lower elevations.
  - These fires tend to occur in regions with flatter terrain.
- Smaller Fires at Higher Elevations:
  - Higher elevation areas have fewer and generally smaller wildfire incidents.

The plot suggests that elevation significantly impacts wildfire occurrence. Lower elevation areas tend to experience more frequent and larger wildfires, while higher elevation areas have fewer and smaller fires. This pattern could be attributed to various factors.

5) How does the historical distribution (2012-2017) of wildfires differ from this year's (2024) wildfires?

Visualization - Fire Size Comparison

To understand if fires in 2024 are generally larger or smaller than those in 2012–2017, we are using a boxplot.

The historical and current fire size data are combined into a single data frame (fire_size_comparison) with an additional column period to distinguish between the two time periods.

A logarithmic scale is applied to the y-axis using scale_y_log10() to better visualize the distribution of fire sizes.

read the current data

```
library(readr)
```

```
Attaching package: 'readr'

The following object is masked from 'package:rvest':

    guess_encoding
```

```
current_fire_points <- read_csv("C:\\Users\\ASUS\\Documents\\Masters - Guelph\\Fall 2024\\6200
```

```
Rows: 1685 Columns: 24

── Column specification ────────────────────────────────────────────────────────
Delimiter: ","
chr (11): FIRE_NO, RSPNS_TYPC, FIRESTATUS, FIRE_CAUSE, FIRE_TYPE, INCDNT_NM,...
dbl (12): FIRE_YEAR, IGN_DATE, FR_T_DTE, FIRE_CENTR, ZONE, FIRE_ID, LATITUDE...
lgl  (1): SHAPE

ℹ Use `spec()` to retrieve the full column specification for this data.
ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
current_fire_perimeters <- read_csv("C:\\Users\\ASUS\\Documents\\Masters - Guelph\\Fall 2024\\
```

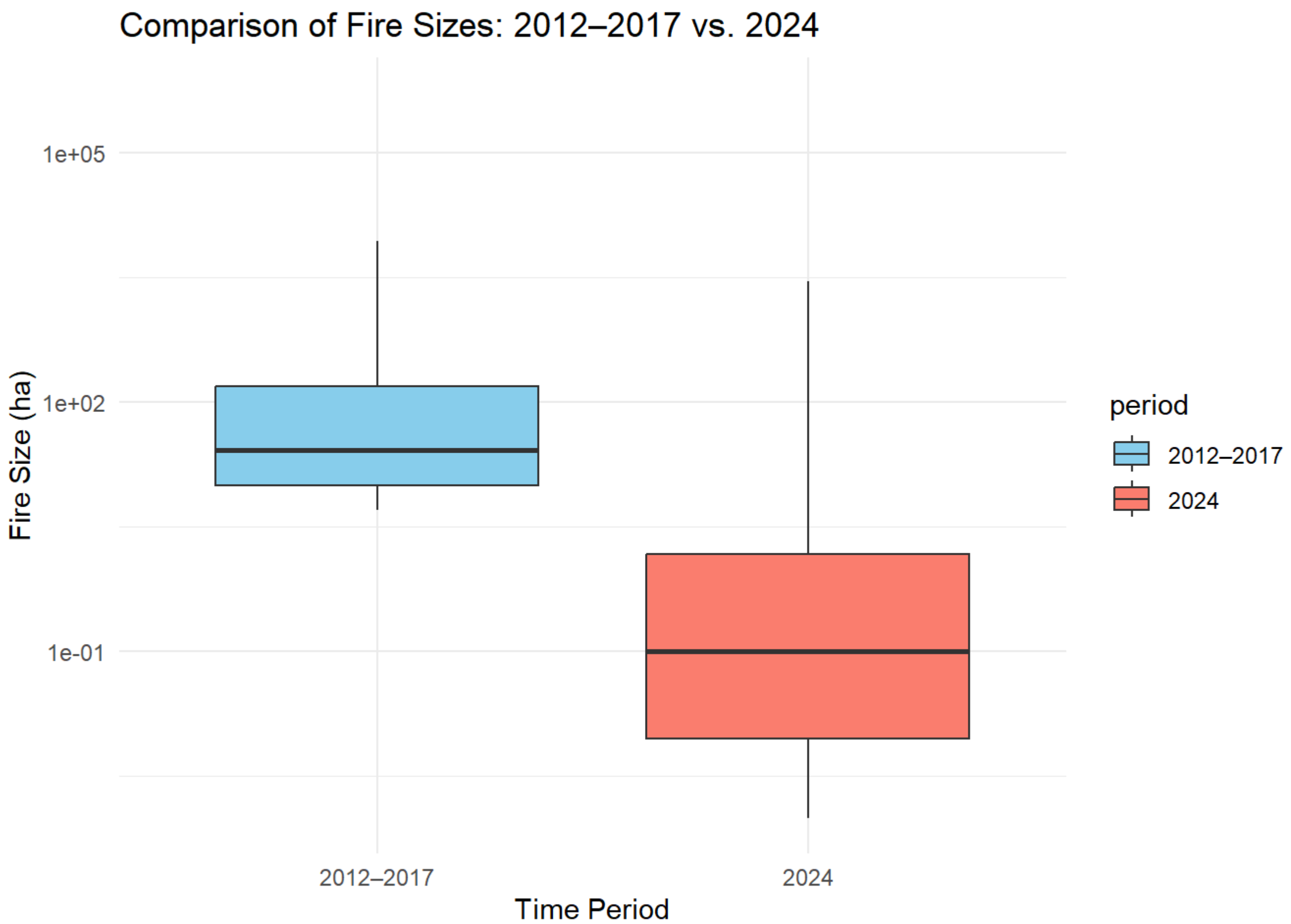
```
Rows: 454 Columns: 16
── Column specification ────────────────────────────────────────────────────────
Delimiter: ","
chr (5): FIRE_NO, SOURCE, FIRE_STAT, FIRE_URL, FCODE
dbl (8): VERSION_NO, FIRE_YEAR, SIZE_HA, TRACK_DATE, LOAD_DATE, AREA_SQM, FE...
lgl (3): SHAPE, X_COORDINATE, Y_COORDINATE
```

```
i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# Combine historical and current fire size data
historical_sizes <- wildfire_table %>% dplyr::select(size_ha) %>% mutate(period = "2012-2017")
current_sizes <- current_fire_points %>% dplyr::select(SIZE_HA) %>% rename(size_ha = SIZE_HA) 

fire_size_comparison <- bind_rows(historical_sizes, current_sizes)

# Plot boxplot
ggplot(fire_size_comparison, aes(x = period, y = size_ha, fill = period)) +
  geom_boxplot(outlier.shape = NA) +
  scale_y_log10() +  # Log scale for better visualization if fire sizes vary greatly
  labs(title = "Comparison of Fire Sizes: 2012-2017 vs. 2024",
       x = "Time Period", y = "Fire Size (ha)") +
  scale_fill_manual(values = c("2012-2017" = "skyblue", "2024" = "salmon")) +
  theme_minimal()
```

```
Warning in scale_y_log10(): log-10 transformation introduced infinite values.
```

```
Warning: Removed 33 rows containing non-finite outside the scale range
(`stat_boxplot()`).
```



The above box plot compares the distribution of fire sizes (in hectares) between two the time periods: 2012-2017 and 2024.

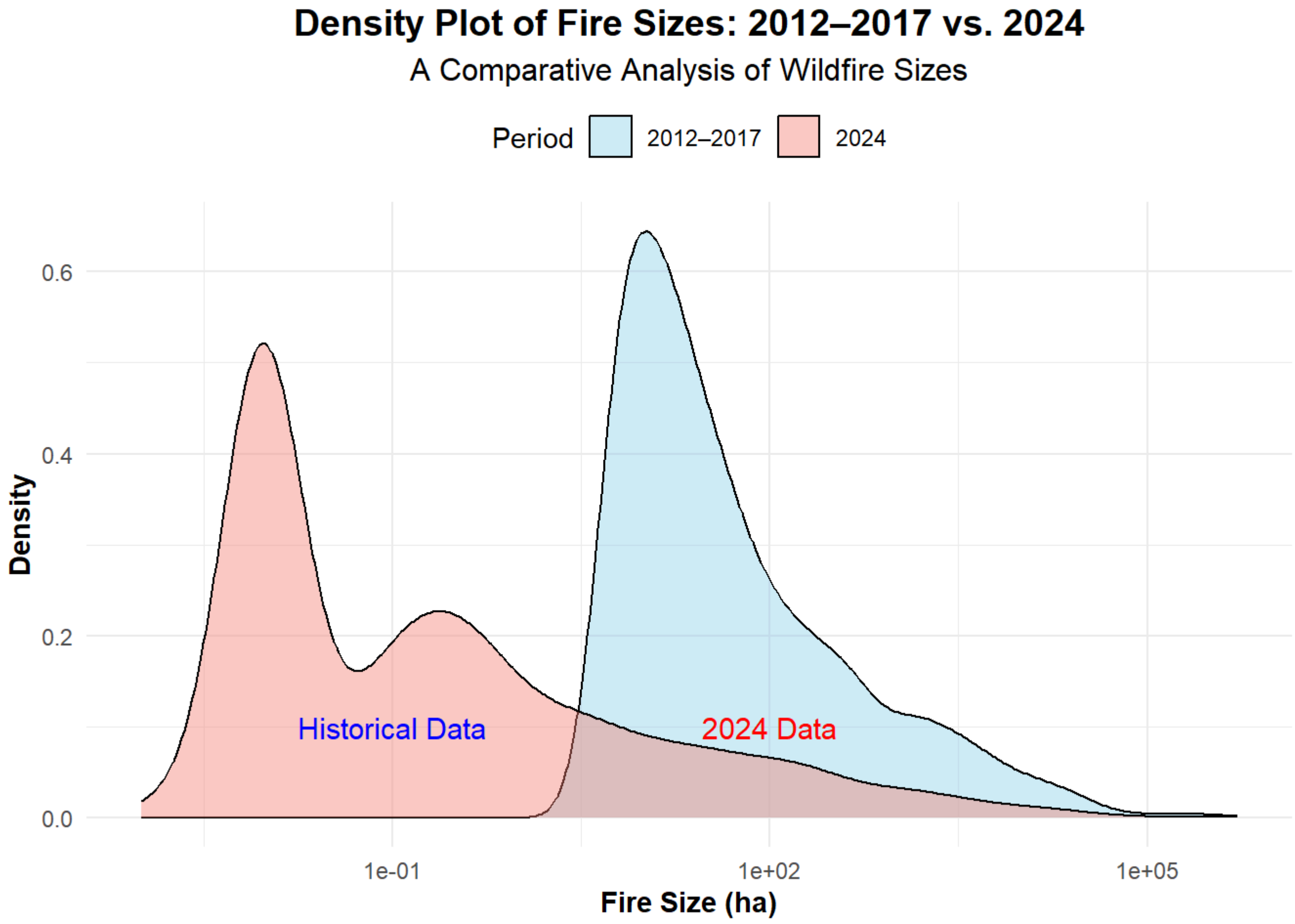
Increased Median Fire Size: The median fire size in 2024 is significantly higher than in the 2012-2017 period, indicating an increase in the typical fire size. Greater Variability: The IQR for 2024 is much larger than for 2012-2017, suggesting greater variability in fire sizes in 2024.

Fire sizes in 2024 are generally larger and more variable compared to the historical period of 2012-2017. This could be indicative of changing environmental conditions, increased fire activity, or other factors.

Density Plot of Fire Size to undestand the extent

```r
ggplot(fire_size_comparison, aes(x = size_ha, fill = period)) +
  geom_density(alpha = 0.4, color = "black") +
  scale_x_log10() +
  labs(
    title = "Density Plot of Fire Sizes: 2012-2017 vs. 2024",
    subtitle = "A Comparative Analysis of Wildfire Sizes",
    x = "Fire Size (ha)",
    y = "Density",
    fill = "Period"
  ) +
  scale_fill_manual(values = c(
    "2012-2017" = "skyblue",
    "2024" = "salmon"
  )) +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold", size = 14),
    plot.subtitle = element_text(hjust = 0.5, size = 12),
    axis.title = element_text(face = "bold"),
    legend.position = "top"
  ) +
  annotate("text", x = 0.1, y = 0.1, label = "Historical Data", color = "blue", size = 4) +
  annotate("text", x = 100, y = 0.1, label = "2024 Data", color = "red", size = 4)
```

```
Warning in scale_x_log10(): log-10 transformation introduced infinite values.

Warning: Removed 33 rows containing non-finite outside the scale range
(`stat_density()`).
```



Visualization

This map will show the spatial distribution of fires historically (2012–2017) compared to 2024.

wildfire_hist_points and wildfire_2024_points dataframes are created and transformed to a CRS to ensure spatial alignment.

The historical wildfire points are overlaid on the base plot using geom_sf.The current wildfire points are overlaid on a separate base plot.

The plot_grid function from the cowplot package is used to combine the historical and current wildfire plots side-by-side.

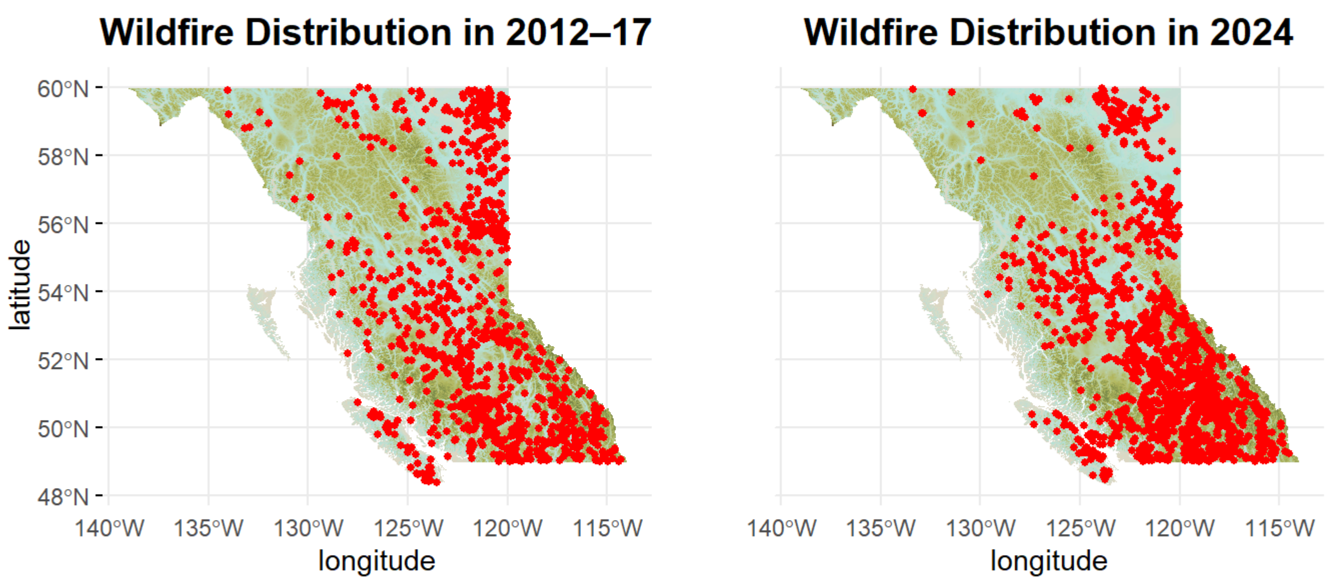
```
# Load elevation raster and transform wildfire points to match the CRS of the raster
elevation_raster <- bc_elevation_masked  # Replace with actual elevation raster if necessary
wildfire_hist_points <- st_transform(wildfire_sf_transformed, crs(elevation_raster))
wildfire_2024_points <- st_transform(current_fire_points_sf, crs(elevation_raster))

# Define the custom color palette
custom_palette <- c("#e5d9c2", "#b6e3db", "#b5ba61", "#7c8d4c", "#725428")

# Create a base plot with common elements
base_plot <- ggplot() +
  # Base elevation raster
  geom_raster(data = elevation_df, aes(x = longitude, y = latitude, fill = elevation)) +
```

```r
  scale_fill_gradientn(colors = custom_palette, name = "Elevation (m)") +
  # Add common theme
  theme_minimal() +
  theme(
    legend.position = "right",
    plot.title = element_text(hjust = 0.5, size = 14, face = "bold"),
    axis.title.y = element_blank()
  )

# Plot for Historical Wildfires (2012-2017)
plot_hist <- base_plot +
  # Overlay wildfire points
  geom_sf(data = wildfire_hist_points, color = "red", alpha = 1, size = 1) +
  # Add titles
  labs(title = "Wildfire Distribution in 2012-17") +
  theme(
    axis.title.y = element_text(angle = 90, vjust = 1.2),  # Adjust position of y-axis label
    axis.text.y = element_text(),
    axis.ticks.y = element_line()
  )

# Plot for Current Wildfires (2024) without y-axis label, ticks, and units
plot_2024 <- base_plot +
  # Overlay wildfire points
  geom_sf(data = wildfire_2024_points, color = "red", alpha = 1, size = 1) +
  # Add titles
  labs(title = "Wildfire Distribution in 2024") +
  theme(
    axis.title.y = element_blank(),
    axis.text.y = element_blank(),
    axis.ticks.y = element_blank()
  )

# Combine plots side by side with the same scale, increasing the plot size by 10%
plot_combined <- plot_grid(
  plot_hist + theme(legend.position = "none"),
  plot_2024 + theme(legend.position = "none"),
  ncol = 2,
  align = "hv",
  rel_widths = c(1.1, 1.1)  # Increase plot widths by 10%
)

# Extract the shared legend
shared_legend <- get_legend(base_plot + theme(legend.position = "bottom", legend.direction = "
```

```
Warning in get_plot_component(plot, "guide-box"): Multiple components found;
returning the first one. To return all, use `return_all = TRUE`.
```

```r
# Final plot with legend at the bottom
final_plot <- plot_grid(
  plot_combined,
  shared_legend,
```

```r
  ncol = 1,
  rel_heights = c(1, 0.1)
)

# Display the final plot
print(final_plot)
```



The above plots illustrate the distribution of wildfires in British Columbia (BC), Canada, over two distinct periods: 2012-2017 and the year 2024.

- Wildfires during 2012-2017 were relatively dispersed across various regions of BC. Clusters of wildfires are present, particularly in the northern and southeastern parts of the province. However, these clusters do not cover the entire region uniformly. The central and southwestern parts of BC seem to have fewer wildfire occurrences compared to these areas.
- The distribution pattern suggests that wildfires were somewhat sporadic, with certain areas experiencing higher frequencies than others. The density of wildfires in these regions points to recurring issues likely influenced by local climatic and environmental conditions.
- The 2024 map indicates a significant increase in both the number and density of wildfires compared to the historical period. Wildfires are more widespread, affecting a larger portion of the province.
- There is a particularly high concentration of wildfires in the southeastern region, indicating a dramatic escalation in this area. The northern region also shows an increased density of wildfires compared to the 2012-2017 period.

- The comparison between the two periods reveals a concerning escalation in wildfire activity in BC. The number and spread of wildfires have increased markedly in 2024, with wildfires becoming more widespread and densely concentrated in certain regions.

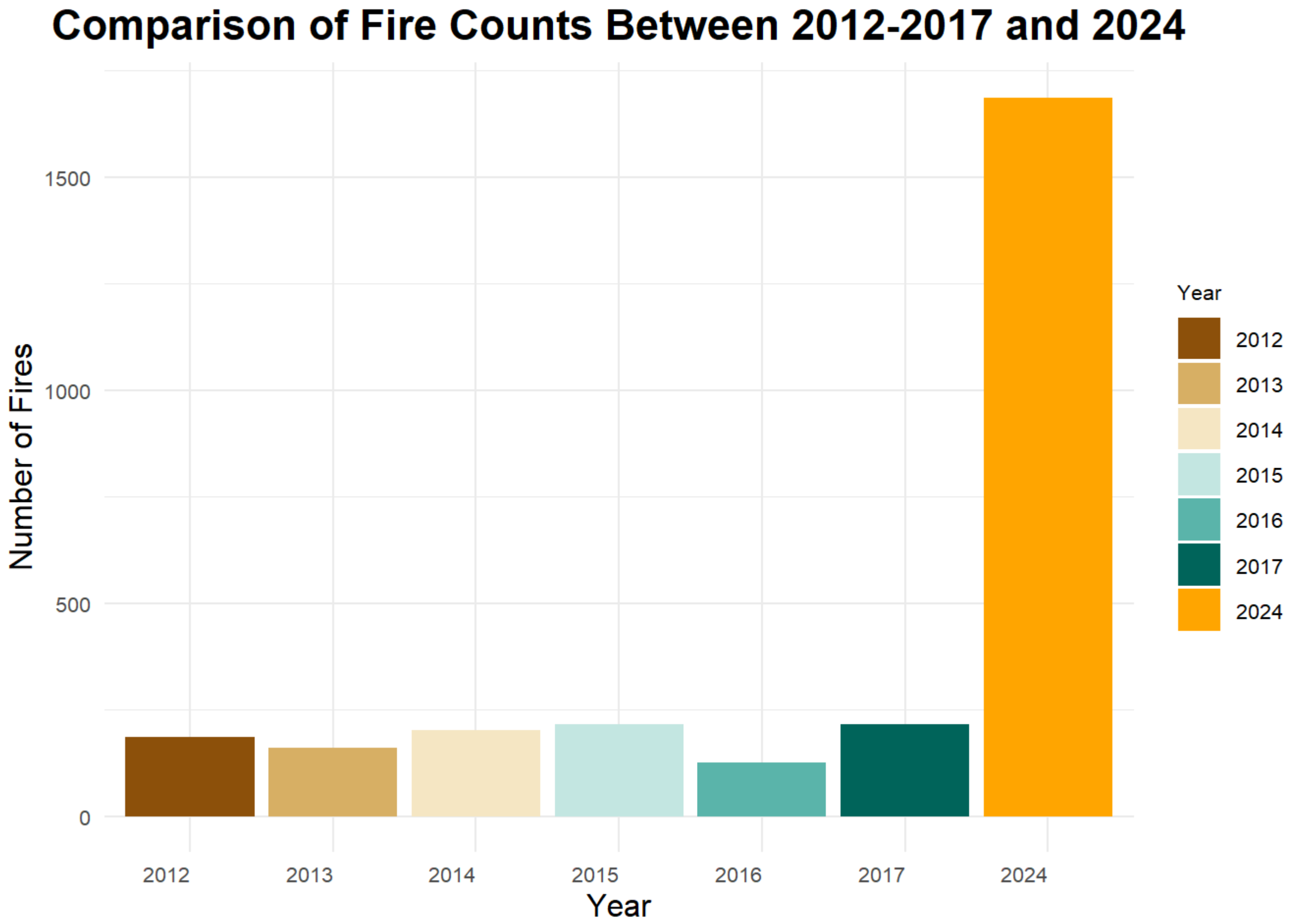
- The southeastern region of BC, in particular, has seen a dramatic rise in wildfire occurrences, indicating that this area might be experiencing more severe environmental stresses.

creating a bar chart to compare the number of fires between the historical period (2012-2017) and the current year (2024).

The historical and current wildfire data are combined into a single data frame and year column is added to the current fire data to match the format of the historical data.

combined_data is grouped by year, and the number of fires in each year is plotted with a bar chart.

```
# Combine the data
combined_data <- wildfire_table %>%
  bind_rows(
    current_fire_points_sf %>%
      mutate(year = 2024) %>%
      select(year)
  )

# Count the number of fires for each year
fire_counts <- combined_data %>%
  group_by(year) %>%
  summarise(num_fires = n())

# Create the bar chart
ggplot(fire_counts, aes(x = factor(year), y = num_fires, fill = factor(year))) +
  geom_bar(stat = "identity") +
  scale_fill_manual(values =  c("#8c510a", "#d8b365", "#f6e8c3", "#c7eae5", "#5ab4ac", "#01665
) +
  labs(
    title = "Comparison of Fire Counts Between 2012-2017 and 2024",
    x = "Year",
    y = "Number of Fires",
    fill = "Year"
  ) +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, size = 16, face = "bold"),
    axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12),
    axis.text.x = element_text(size = 8, hjust = 1),
    axis.text.y = element_text(size = 8),
    legend.title = element_text(size = 8),
    legend.text = element_text(size = 8)
  )
```



- From 2012 to 2017, the number of fires remains relatively low and stable, with each year seeing fewer than 500 fires. The slight fluctuations during these years suggest minor year-to-year variations, but overall, the fire occurrences were consistently low.
- The year 2024 shows a dramatic surge in wildfire counts, with the number of fires exceeding 1500. This represents a significant increase compared to the previous years