

# DATA\*6200 Assignment 2

AUTHOR

Atharva Vichare

## Introduction

Wildfires are a growing concern in Canada, particularly in British Columbia (B.C) due to climate and environmental factors. In this assignment we will perform analysis on wildfire distribution in B.C. from 2012-2017 and also compare it with recent data from 2024 to identify the trends and how factors like temperature as well as elevation contribute to wildfire occurrence.

## Importing important libraries

```
# Load necessary libraries
library(rvest)      # For web scraping
library(sf)         # For handling spatial data
```

Linking to GEOS 3.12.1, GDAL 3.8.4, PROJ 9.3.1; sf\_use\_s2() is TRUE

```
library(raster)      # For raster data manipulation
```

Loading required package: sp

```
library(ggplot2)    # For plotting
library(dplyr)       # For data manipulation
```

Attaching package: 'dplyr'

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

intersect, select, union

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

filter, lag

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

intersect, setdiff, setequal, union

```
library(ggspatial)   # For map overlays
library(terra)        # Alternative for raster manipulation if needed
```

terra 1.7.83

```
library(geodata)      # Pull the temperature and elevation data
library(osmdata)      # To get BC boundary
```

Data (c) OpenStreetMap contributors, ODbL 1.0. <https://www.openstreetmap.org/copyright>

```
library(ggplot2)      # For plotting maps
library(dplyr)         # For data manipulation
library(viridis)       # For importing color for raster and maps
```

Loading required package: viridisLite

```
library(leaflet)       # For plotting interactive maps
library(htmltools)     # For customizing UI of interactive maps
library(readxl)        # For importing excel file
library(magick)        # To view plots saved as images due to memory issue
```

Linking to ImageMagick 6.9.12.98

Enabled features: cairo, freetype, fftw, ghostscript, heic, lcms, pango, raw, rsvg, webp

Disabled features: fontconfig, x11

```
library(patchwork)    # To combine the plots
```

Attaching package: 'patchwork'

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

area

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

area

```
library(knitr)         # To display saved image in quarto
```

Attaching package: 'knitr'

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

spin

## Question 1

## Scraping 2012-2017 wildfire data in BC from gov.bc.ca site

To analyze the historical data we will first use data scrapping techniques taught in class to scrape wildfire data from 2012-2017 present on the gov.bc.ca site.

```
# Step 1: Read the html
wildfire_html <- read_html("https://www2.gov.bc.ca/gov/content/safety/wildfire-status/about-wildfires")
wildfire_html
```

```
{html_document}
<html lang="en">
[1] <head>\n<meta http-equiv="Content-Type" content="text/html; charset=UTF-8 ...
[2] <body>\n<div id="__next" data-reactroot="">\n<script src="/env-config.js" ...
```

Note : Run the cells individually because when we just hit run all cells the cells wont run or mess up the visualizations for some reason. Apologies for the inconvenience.

```
# Step 2 : Retrive data from the html using table fuction
tables <- wildfire_html |>
  html_nodes("table") |>
  html_nodes(xpath = "//table[caption/p/strong[contains(text(), 'Locations of all Wildfires')]]") |>
  html_table(fill = TRUE)
# Replace 1 with the correct index if the first table is not the one you need
wildfire_data <- as_tibble(tables[[1]])
# Display the first few data
head(wildfire_data)
```

```
# 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>
```

After successfully scrapping the historical data of wildfire (2012-2017) I noticed the longitudes and latitudes are in degrees & minutes format so first thing we will do is to convert them into decimal degrees format by creating a function called convert\_to\_decimal () which will split them by spaces and convert them into numeric type. Further they will be converted into decimal degrees by the formula.

```
# Convert longitude and latitude into degrees from minutes format
convert_to_decimal <- function(coord) {
  parts <- strsplit(coord, " ")[[1]]
```

```

degrees <- as.numeric(parts[1])
minutes <- as.numeric(parts[2])
decimal_degrees <- degrees + (minutes / 60)
return(decimal_degrees)
}
# Apply the function to latitude and longitude columns
wildfire_data$Latitude <- sapply(wildfire_data$Latitude, convert_to_decimal)
wildfire_data$Longitude <- sapply(wildfire_data$Longitude, convert_to_decimal)
head(wildfire_data)

```

```

# A tibble: 6 × 8
  Year `Fire Number` `Fire Centre` Latitude Longitude Geographic
  <int> <chr>        <chr>       <dbl>    <dbl> <chr>
1 2017 C50011 (2017) Cariboo      51.9     123. Lee's Corner
2 2017 C50043 (2017) Cariboo      52.0     123. 3 Km East Anaham Reserve
3 2017 C10248 (2017) Cariboo      53.2     123. Pantage Lake
4 2017 C40329 (2017) Cariboo      51.8     122. 6km SE of Place Lake
5 2017 C20352 (2017) Cariboo      51.7     122. 4 km W of Emerald Lake
6 2017 C20414 (2017) Cariboo      52.0     123. Raven Lake
# i 2 more variables: `Discovery Date` <chr>, `Size (ha)` <dbl>

```

Next I noticed a column named as "Size(ha)" , "Discovery Date" , "Fire Centre" & "Fire Number" which can cause us problem due to its structure in future. So we will use rename() to rename that particular column

```

#Rename the column
wildfire_data <- wildfire_data |>
  rename(
    Size_ha =`Size (ha)` ,
    Fire_Centre = `Fire Centre` ,
    Fire_Number = `Fire Number` ,
    Discovery_Date = `Discovery Date`
  )
head(wildfire_data)

```

```

# A tibble: 6 × 8
  Year Fire_Number Fire_Centre Latitude Longitude Geographic Discovery_Date
  <int> <chr>        <chr>       <dbl>    <dbl> <chr>          <chr>
1 2017 C50011 (2017) Cariboo      51.9     123. Lee's Corner April 6, 2017
2 2017 C50043 (2017) Cariboo      52.0     123. 3 Km East A... April 21, 2017
3 2017 C10248 (2017) Cariboo      53.2     123. Pantage Lake June 4, 2017
4 2017 C40329 (2017) Cariboo      51.8     122. 6km SE of P... June 18, 2017
5 2017 C20352 (2017) Cariboo      51.7     122. 4 km W of E... June 23, 2017
6 2017 C20414 (2017) Cariboo      52.0     123. Raven Lake   June 26, 2017
# i 1 more variable: Size_ha <dbl>

```

After carefully observing the latitudes and longitudes we can notice the longitudes of British Columbia area in the dataset are positive but according to our knowledge since B.C. is the west of Prime Meridian longitudes should be in negative format. We will convert them into negative form using mutate() so that while plotting maps in future we wont have to do

```
# Convert Longitudes into negative format since BC has negative longitutde
wildfire_data <- wildfire_data |>
  mutate(Longitude = -abs(Longitude))
head(wildfire_data)
```

```
# A tibble: 6 × 8
  Year Fire_Number Fire_Centre Latitude Longitude Geographic Discovery_Date
  <int> <chr>      <chr>     <dbl>    <dbl> <chr>        <chr>
1 2017 C50011 (2017) Cariboo      51.9     -123. Lee's Corner April 6, 2017
2 2017 C50043 (2017) Cariboo      52.0     -123. 3 Km East A... April 21, 2017
3 2017 C10248 (2017) Cariboo      53.2     -123. Pantage Lake June 4, 2017
4 2017 C40329 (2017) Cariboo      51.8     -122. 6km SE of P... June 18, 2017
5 2017 C20352 (2017) Cariboo      51.7     -122. 4 km W of E... June 23, 2017
6 2017 C20414 (2017) Cariboo      52.0     -123. Raven Lake   June 26, 2017
# i 1 more variable: Size_ha <dbl>
```

## Question 2

### Pulling the temperature data using 'geodata' package

Following the documentation of geodata package we will use worldclim\_country to fetch country wise Canada's temperature first.

```
can <- worldclim_country("Canada", var="tavg", res=10, path=tempdir())
```

The major problem now occurs is that we get 12 different raster data for 12 months . But since December, January & February are the winter season months when temperature shoots below 0 degrees it is highly unlikely there will be any wildfire during that period and so we will only take raster of all the months excluding the winter season.

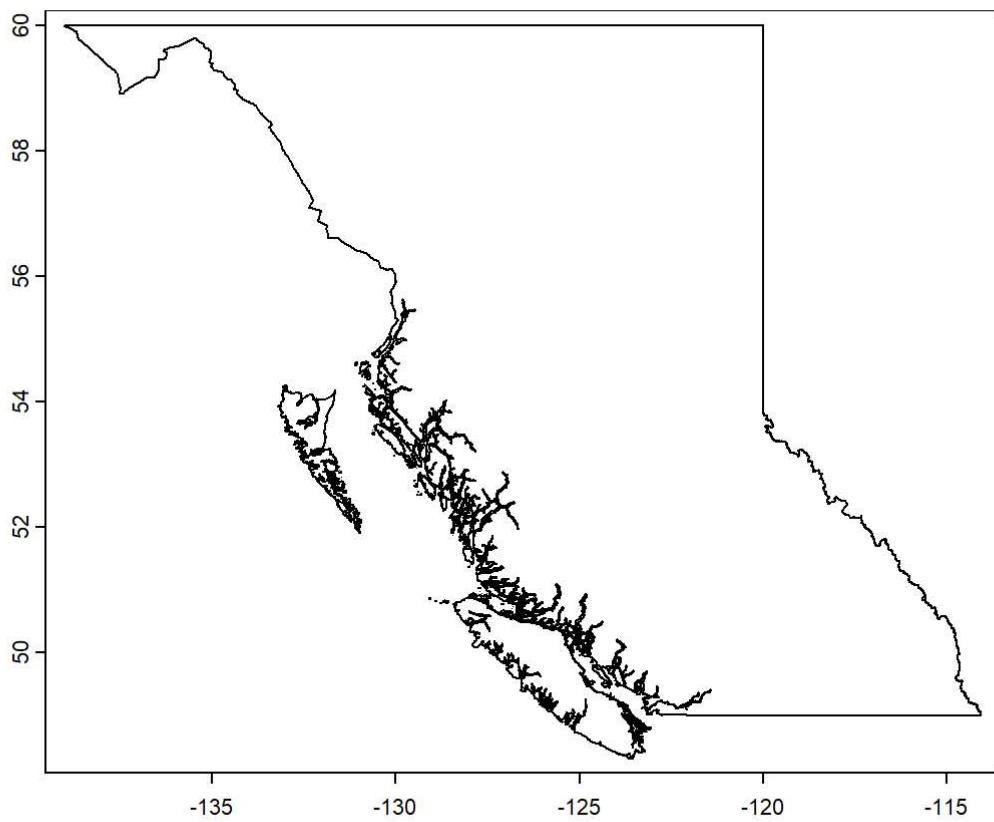
```
# Taking temperature data from March to November
can_no_winter <- can[[c(3:11)]]
```

After getting the country wise temperature data we will retrive boundaries of different provinces using the '**gadmTools**' or '**GADMTools**' (**Global Administrative Areas**) database.

```
# Extract boundaries of different provinces
province_boundary <- gadm("Canada", level = 1, path = tempdir())
```

Now we extract British Columbia boundary from province\_boundary .

```
#Extract only BC boundary
bc_boundary <- province_boundary[province_boundary$NAME_1 == "British Columbia", ]
plot(bc_boundary)
```

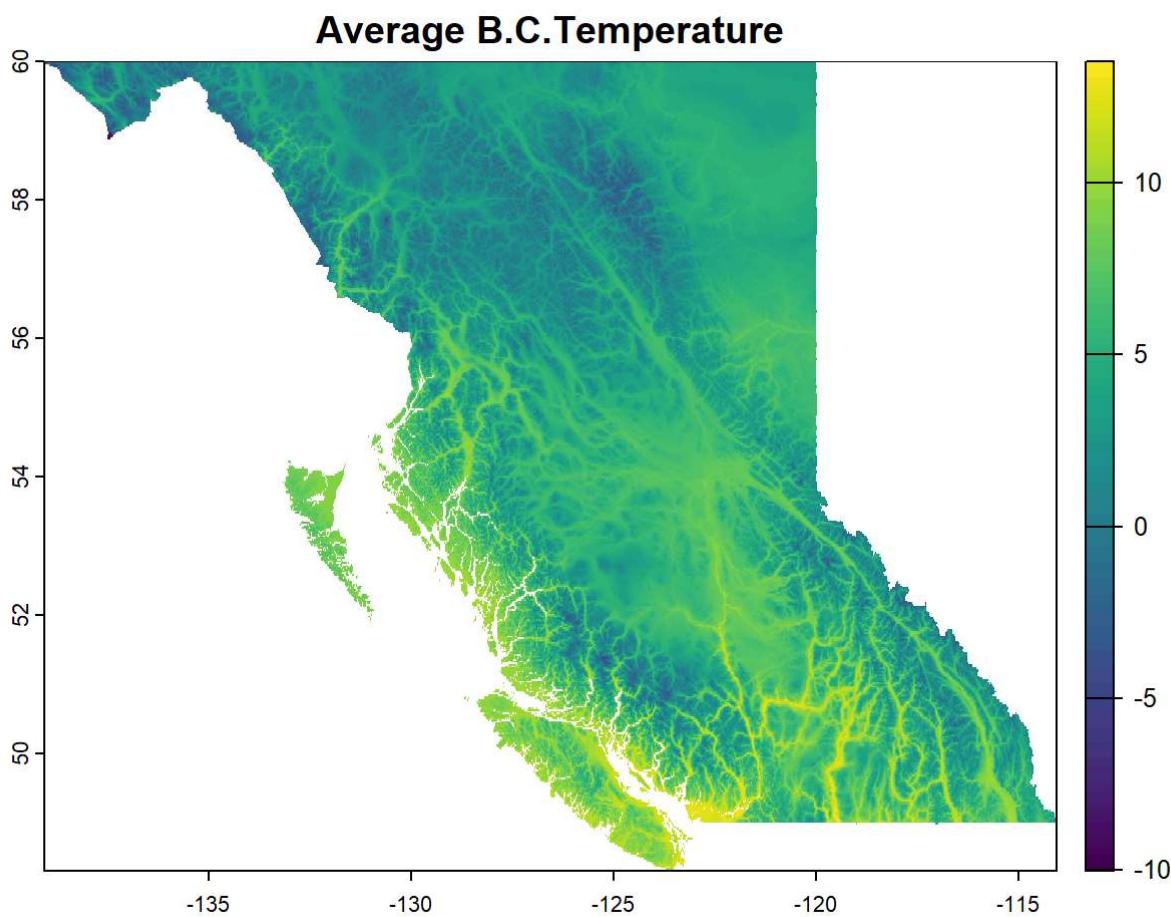


Now since our British Columbia boundary ready to use we will crop and mask the temperature raster we extracted of Canada to get B.C. temperature raster .

```
bc_temp <- crop(can_no_winter, bc_boundary) # Crop to the extent of BC  
bc_temp <- mask(bc_temp, bc_boundary) # Mask using the exact BC boundary
```

Now after masking we got temperature raster for B.C. from March to November with 9 plots in total but we need to show only one temperature raster so will take average of all the 9 months to get just on average B.C. temperature raster.

```
# Calculate the average BC temperature for the raster  
avg_temp_bc <- mean(bc_temp)  
plot(avg_temp_bc, main = "Average B.C.Temperature")
```



## Pulling the elevation data using geodata package

Following the documentation of geodata package we will use worldclim\_country to fetch country wise Canada's elevation data first.

```
# Pull the elevation data for Canada
can_elev <- elevation_30s(country="CAN", path=tempdir(), mask=TRUE)
```

Here we don't have to deal with any problem of 12 plots like in the temperature data. We just get one plot as we want

After getting the country wise elevation data we will get boundaries again of different provinces again to avoid any errors using the '**gadmTools**' or '**GADMTools**' (**Global Administrative Areas**) database

```
# Extract province boundaries
province_boundary2 <- gadm("Canada", level = 1, path = tempdir())
```

Now we again extract British Columbia boundary from province\_boundary .

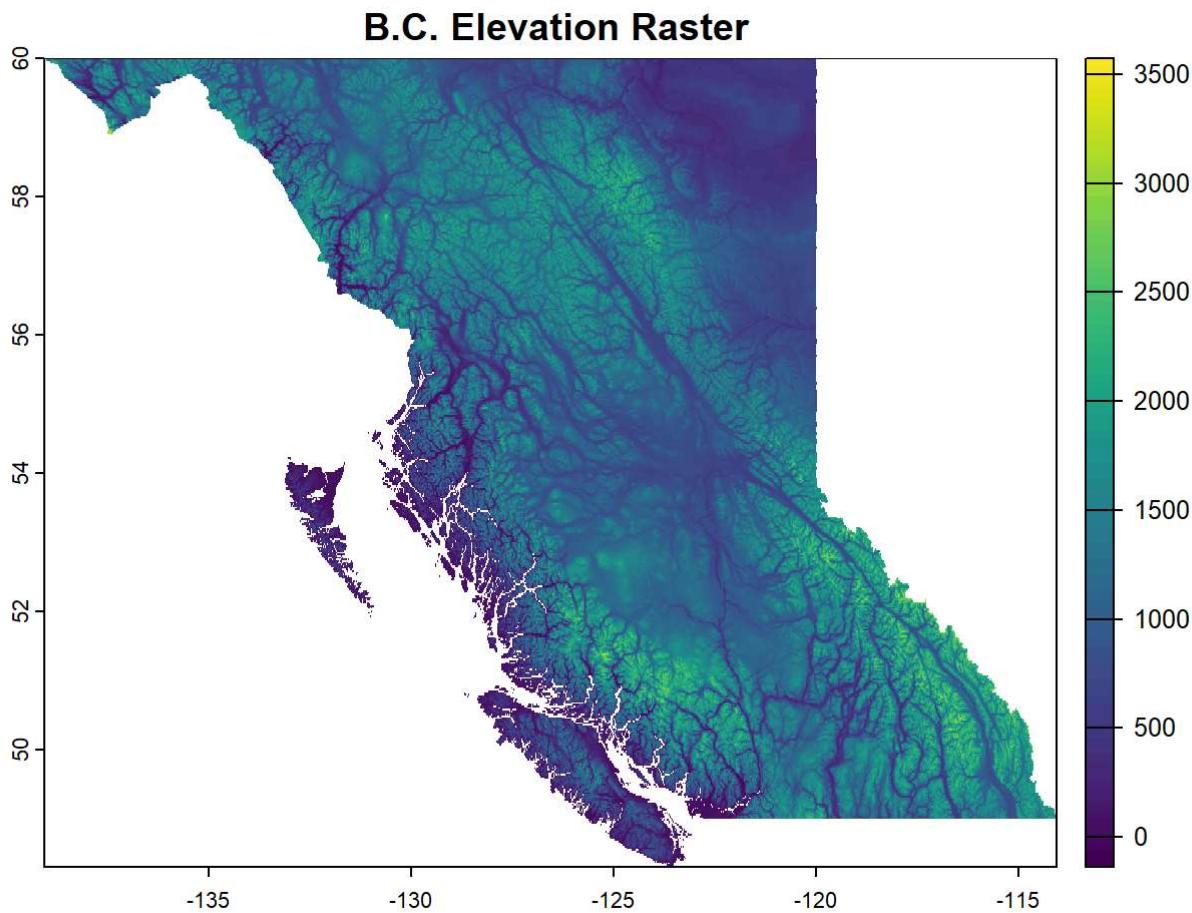
```
#Extract only BC boundary
```

```
bc_boundary2 <- province_boundary2[province_boundary2$NAME_1 == "British Columbia", ]
```

Now we will perform cropping and masking to fit the elevation raster data inside the BC boundary same like we did above for temperature.

```
bc_elev <- crop(can_elev, bc_boundary2) # Crop to the extent of BC
bc_elev <- mask(bc_elev, bc_boundary2) # Mask using the exact BC boundary
```

```
#Plot the elevation raster
plot(bc_elev, main = " B.C. Elevation Raster")
```



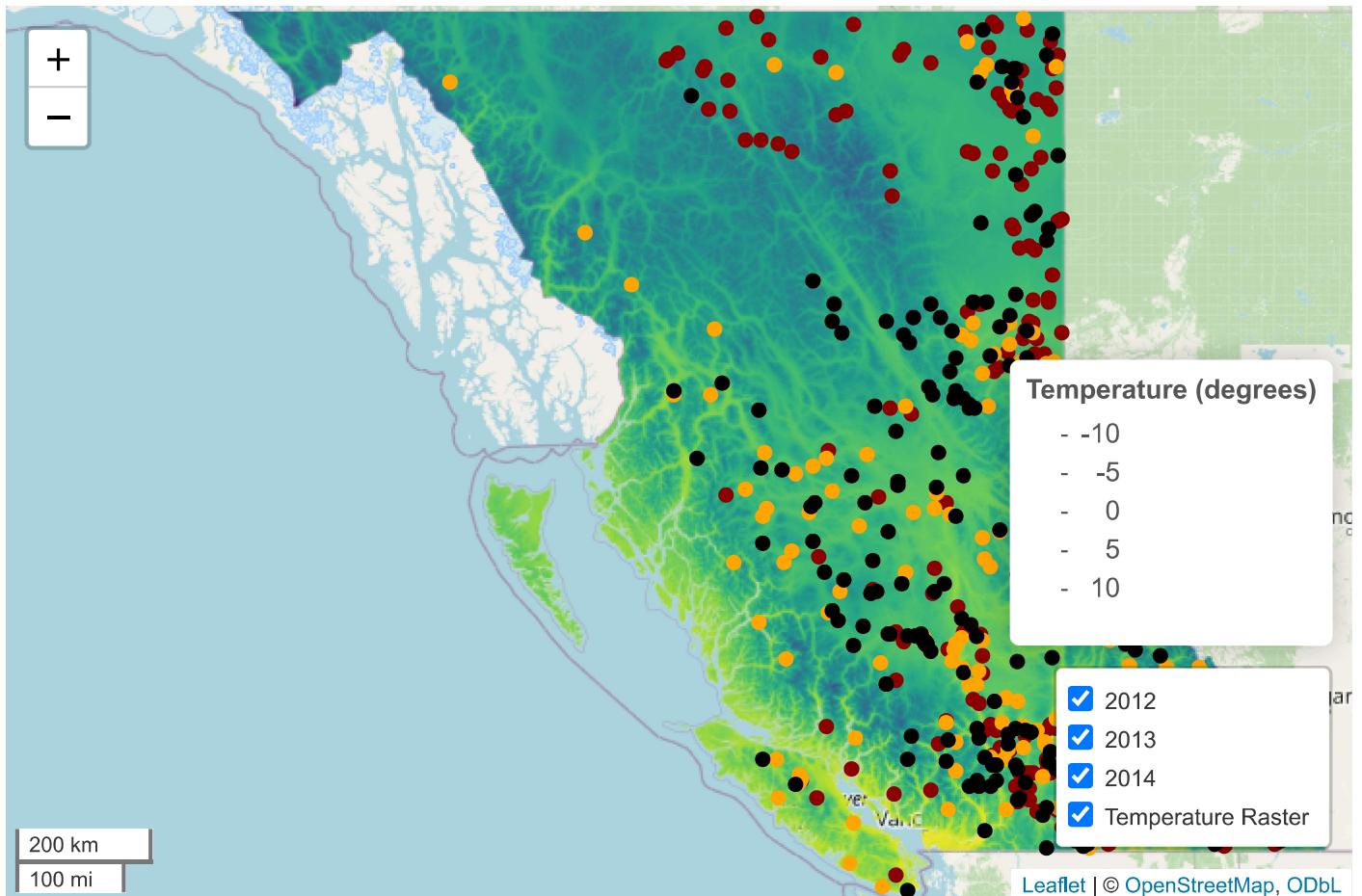
## Question 3

**Visualizing overtime between 2012-2017 overtop a temperature raster**

### 2012, 2013 & 2014 wildfire overtop temperature raster

I have used leaflet to make this interactive graph since we are supposed to visualize wildfire between 2012-2017 over the temperature raster and plotting all the points in one graph makes the graph crowded also

making one graph for each year is repetitive work so I rather choose to make them in a group of 3 (2012, 2013 & 2014) where we can visualize wildfire separately or in combination to get better understanding.



1. The red dots represent wildfire locations for 2012
2. The orange dots represent wildfire locations for 2013
3. The black dots represent wildfire location for 2014

The background is the Average temperature raster for B.C. between March and November.

### **2012 Wildfires :**

The wildfires appear to be concentrated in the southern and northern regions of BC. The areas with higher temperature seem to be hot-spot for wildfire suggesting the contribution of high temperature in wildfire.

1. The southern region, appears to be the hot-spot for wildfires. This area has a high number of fire points and overlaps with the warmer temperature regions.
2. The coastal areas, particularly along the Pacific coast, have fewer fire points. This aligns with the cooler temperatures indicated by the raster in these regions.
3. The northern parts of BC also show a significant number of fires, especially in the central and eastern regions. This could be due to a combination of factors, including higher temperature.

### **2013 Wildfires :**

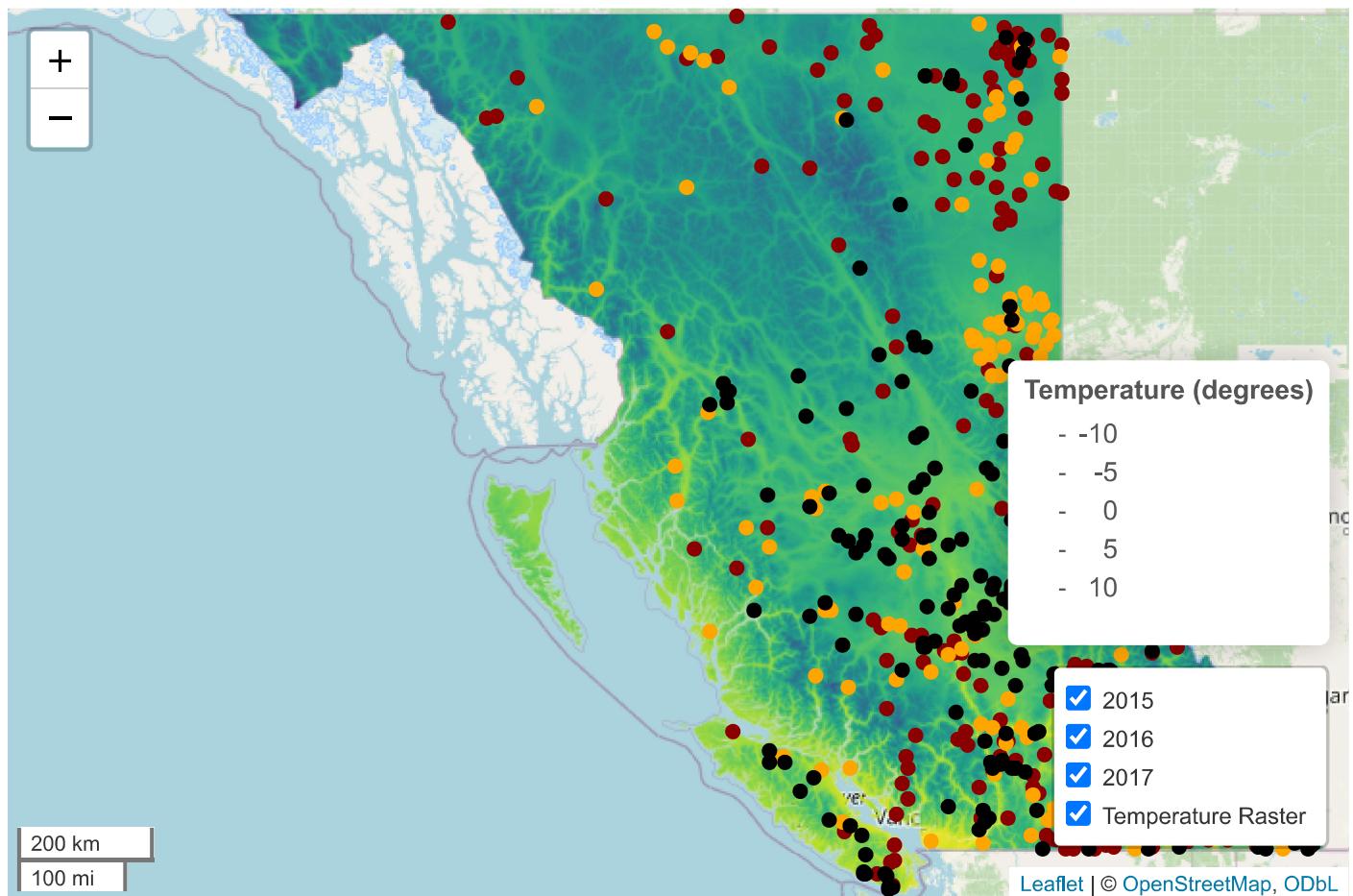
In 2013 the wildfires also seem to be concentrated more in the central and southern regions of BC as compared to 2012.

1. The central & southern region, appears to be the hotspot for wildfires as compared to the southern region from 2012.
2. The coastal areas have fewer wildfires same as 2012.
3. Although there seem to be some wildfires in northern region but they are much fewer as compared to 2012 . This can be due to various reasons like better fire management from authorities, etc.

### **2014 Wildfires :**

In 2014, as compared to previous years of 2012 & 2013 the wildfire distribution has shifted more towards the southern region.

### **2015, 2016 & 2017 wildfire overtop temperature raster**



1. The red dots represent wildfire locations for 2015.
2. The orange dots represent wildfire locations for 2016.
3. The black dots represent wildfire location for 2017.

The background is the Average temperature raster for B.C. between March and November.

### **2015 Wildfires :**

The wildfires appear to be concentrated in the southeastern and northern regions of BC same as 2012.

1. The southeastern region contains notable cluster of wildfires.
2. The central region seems to have fewer wildfires just like the coastal regions near the Pacific Ocean
3. The northern parts of BC also show a significant number of fires, especially in the eastern region.

### **2016 Wildfires :**

As compared to 2015 the wildfires appear to be less in number.

1. Most of the wildfires are concentrated near the southern and central region .
2. The northern region of BC has fewer wildfires than 2015 .

### **2017 Wildfires :**

The wildfires again increased in number with most of the concentration in southern and central region .

1. The southeastern region contains notable cluster of wildfires.
2. The norther region appears to be empty with a few exceptions on its eastern border .

### **Overall Comparision :**

1. In both maps, wildfire points are scattered across the entire province, but the second map shows more clustering in the interior and southern parts of British Columbia, while the first map has a wider spread with more points in the southern region as well.
2. The second map has more intense fire points in certain areas, indicating more fires in the 2015-2017 period.
3. The first map (2012-2014) shows a more even distribution across BC, with a slightly lower density compared to the second map (2015-2017)
4. The second map (2015-2017) has a higher density of points, especially concentrated in the central and southern regions of BC, indicating more wildfires those areas.

## **Question 4**

### **How do temperature impact fire occurrence?**

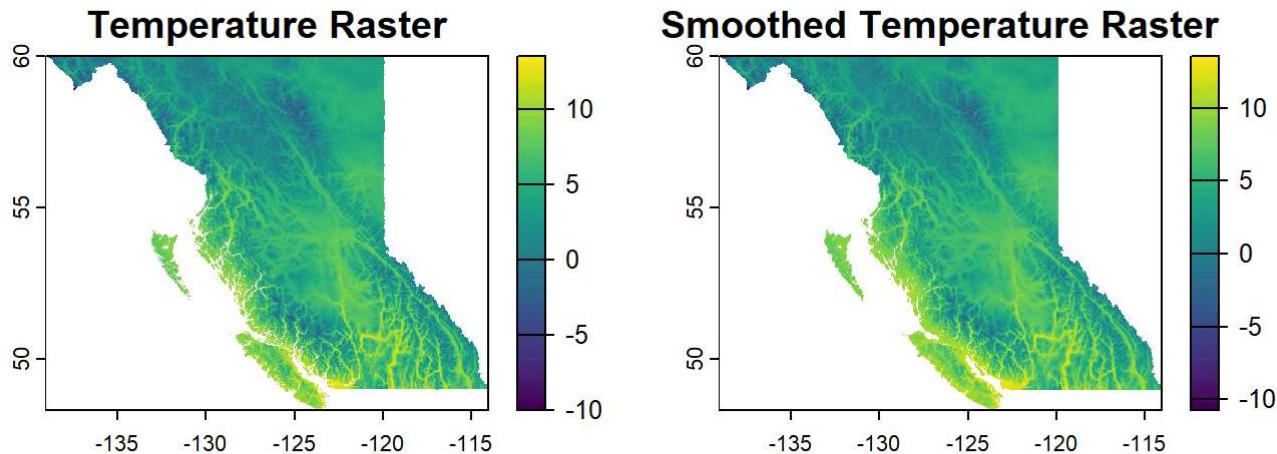
We can answer this question in many different ways if we had data frame with temperature and elevation data along with wildfire data. But instead we have temperature and elevation raster's not as data frame. So the best way to answer this question is by using focal and zonal operations on these raster's taught in class.

Firstly we apply focal and zonal operations to process the temperature raster(data) and see its relationship with wildfire occurrences

```
# Convert the wildfire data into spatial points
wildfire_points <- st_as_sf(wildfire_data, coords = c("Longitude", "Latitude"), crs = st_crs(wildfire_data))
```

The next step is to apply focal operation to smooth the temperature raster by calculating average temperature around each pixel this helps to reduce noise and show temperature pattern more clearly.

```
# Focal operation to smooth the temperature data
focal_temp <- focal(avg_temp_bc, w = matrix(1,nrow= 3,ncol = 3), fun = mean, na.rm = TRUE)
```



Now zonal operation starts by adding buffer zone around each wildfire and calculate average temperature within that zones which temperature indications for that particular wildfire

```
# Zonal operation to create buffer around each wildfire
wildfire_buffer <- st_buffer(wildfire_points, dist = 5000)
```

Next is to extract that temperature into a new data frame.

```
# Extract buffered temperature to new df and add it to the wildfire_points
buffered_temp <- extract(focal_temp, wildfire_buffer, fun = mean, na.rm = TRUE)
```

```
wildfire_points$Avg_Temperature <- buffered_temp
head(wildfire_points)
```

Simple feature collection with 6 features and 7 fields

Geometry type: POINT

Dimension: XY

Bounding box: xmin: -123.1615 ymin: 51.7428 xmax: -121.8343 ymax: 53.16795

Geodetic CRS: WGS 84

# A tibble: 6 × 8

	Year	Fire_Number	Fire_Centre	Geographic	Discovery_Date	Size_ha	<dbl>
1	2017	C50011 (2017)	Cariboo	Lee's Corner	April 6, 2017	8.7	
2	2017	C50043 (2017)	Cariboo	3 Km East Anaham Reser...	April 21, 2017	8	
3	2017	C10248 (2017)	Cariboo	Pantage Lake	June 4, 2017	132.	
4	2017	C40329 (2017)	Cariboo	6km SE of Place Lake	June 18, 2017	27.4	
5	2017	C20352 (2017)	Cariboo	4 km W of Emerald Lake	June 23, 2017	25.9	
6	2017	C20414 (2017)	Cariboo	Raven Lake	June 26, 2017	41.8	

# i 2 more variables: geometry <POINT [°]>, Avg\_Temperature <dbl>

Now when we do this the Avg\_temperature gets attached to the wildfire\_points data as a data frame list with two columns called 'id' & 'focal\_mean'. We will remove this type of list and only extract Avg\_Temperature from there and store it as a new column.

```
# Extract mean from the Temperature data frame
wildfire_points <- wildfire_points |>
  mutate(
    Avg_Temperature = Avg_Temperature$focal_mean
  )
head(wildfire_points)
```

Simple feature collection with 6 features and 7 fields

Geometry type: POINT

Dimension: XY

Bounding box: xmin: -123.1615 ymin: 51.7428 xmax: -121.8343 ymax: 53.16795

Geodetic CRS: WGS 84

# A tibble: 6 × 8

	Year	Fire_Number	Fire_Centre	Geographic	Discovery_Date	Size_ha	<dbl>
1	2017	C50011 (2017)	Cariboo	Lee's Corner	April 6, 2017	8.7	
2	2017	C50043 (2017)	Cariboo	3 Km East Anaham Reser...	April 21, 2017	8	
3	2017	C10248 (2017)	Cariboo	Pantage Lake	June 4, 2017	132.	
4	2017	C40329 (2017)	Cariboo	6km SE of Place Lake	June 18, 2017	27.4	
5	2017	C20352 (2017)	Cariboo	4 km W of Emerald Lake	June 23, 2017	25.9	
6	2017	C20414 (2017)	Cariboo	Raven Lake	June 26, 2017	41.8	

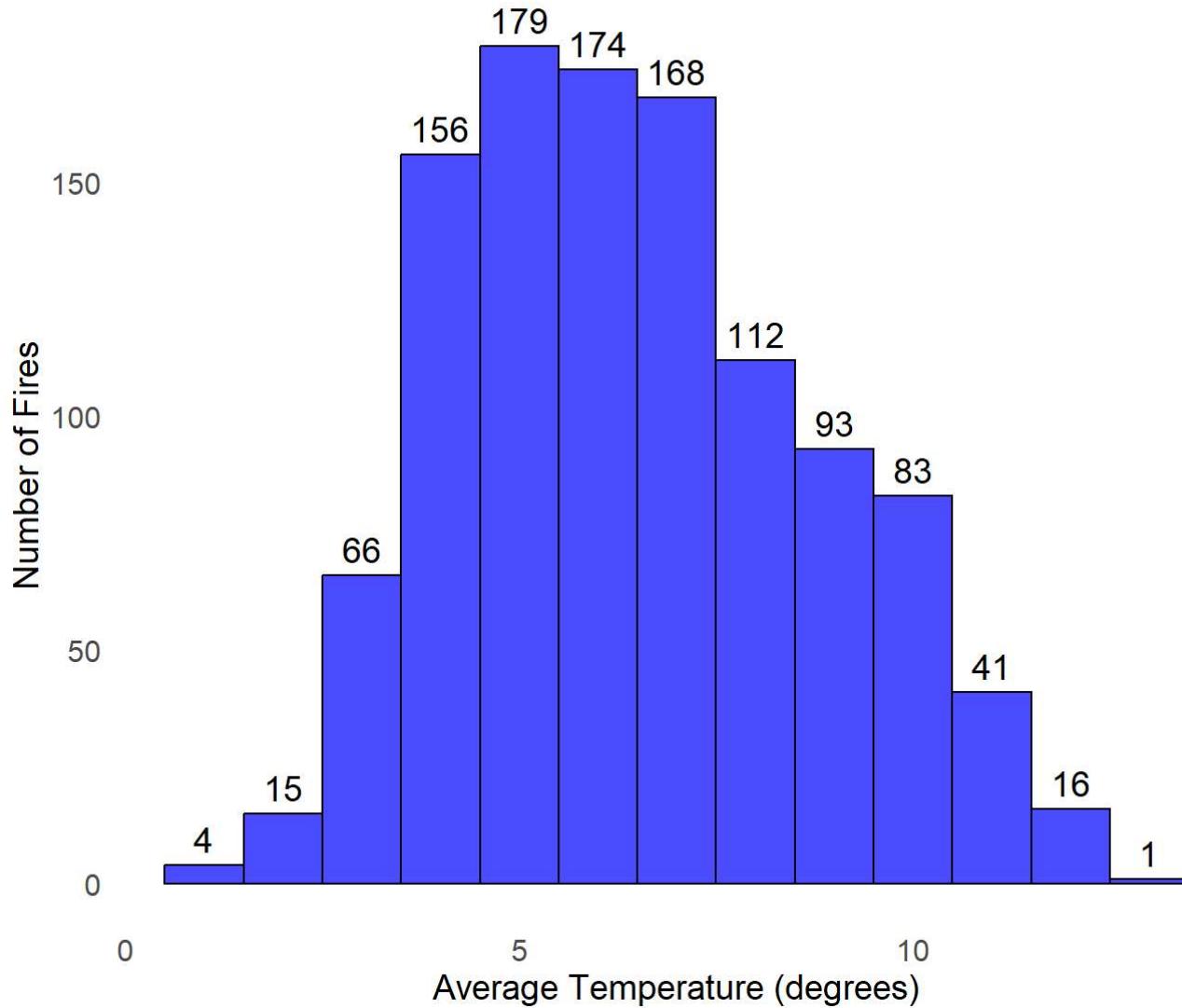
# i 2 more variables: geometry <POINT [°]>, Avg\_Temperature <dbl>

Now our data is ready so we will plot a histogram to see the impact of temperature on wildfire occurrence .

Warning: The dot-dot notation (`..count..`) was deprecated in ggplot2 3.4.0.

i Please use `after\_stat(count)` instead.

## Impact of Temperature on Wildfire Occurrence

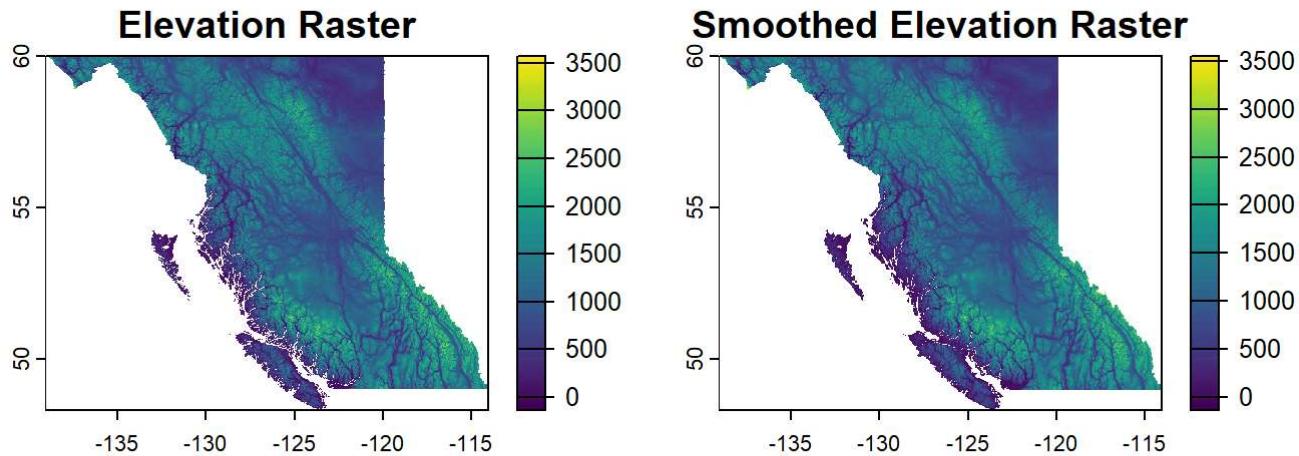


From the graph we interpret that highest number of wildfires occurred in the 5-9 degrees temperature range with most occurring around 7-8 degrees temperature with count over 160 wildfires. This pattern suggest that wildfires have mainly occurred between the range of 5-9 degrees. Apart from this range the wildfires start decreasing in number as temperature rise to 10 degrees and as temperature starts decreasing to 3 degrees. This can happen because in extreme cold fires cant occur due to cold conditions and at high temperature they don't occur because there might be no vegetation to ignite the wildfires.

## How do elevation impact fire occurrence?

We apply the same techniques of focal and zonal operation here as we did for temperature.

```
# Focal operation smooth the elevation data
focal_elev <- focal(bc_elev, w = matrix(1,nrow = 3,ncol = 3), fun = mean, na.rm = TRUE)
```



```
# Zonal operation to create buffer around each wildfire
wildfire_buffer_elev <- st_buffer(wildfire_points, dist = 5000)
# Extract buffered elevation to new df and add it to the wildfire_points
buffered_elev <- extract(focal_elev, wildfire_buffer_elev, fun = mean, na.rm = TRUE)
wildfire_points$Avg_Elevation <- buffered_elev
# Extract mean from the Temperature data frame
wildfire_points <- wildfire_points |>
  mutate(
    Avg_Elevation = Avg_Elevation$focal_mean
  )
head(wildfire_points)
```

Simple feature collection with 6 features and 8 fields  
 Geometry type: POINT  
 Dimension: XY  
 Bounding box: xmin: -123.1615 ymin: 51.7428 xmax: -121.8343 ymax: 53.16795

Geodetic CRS: WGS 84

# A tibble: 6 × 9

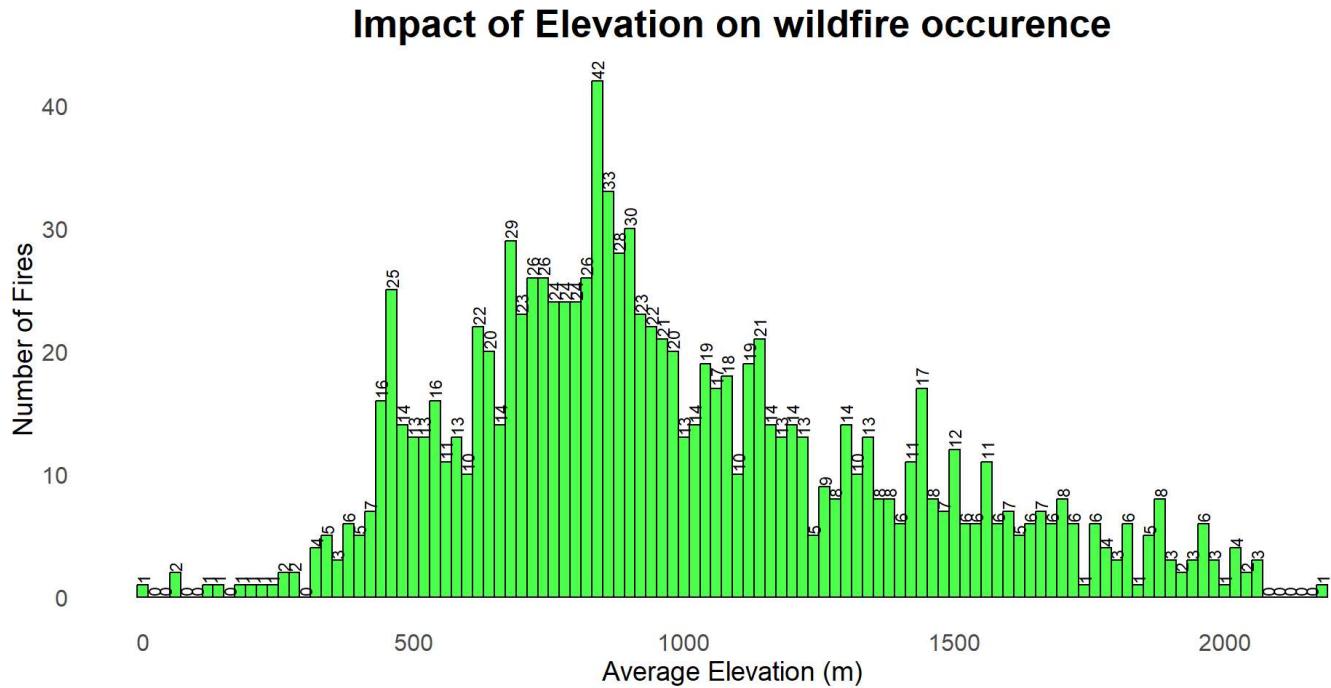
	Year	Fire_Number	Fire_Centre	Geographic	Discovery_Date	Size_ha
	<int>	<chr>	<chr>	<chr>	<chr>	<dbl>
1	2017	C50011 (2017)	Cariboo	Lee's Corner	April 6, 2017	8.7
2	2017	C50043 (2017)	Cariboo	3 Km East Anaham Reser...	April 21, 2017	8
3	2017	C10248 (2017)	Cariboo	Pantage Lake	June 4, 2017	132.

```

4 2017 C40329 (2017) Cariboo      6km SE of Place Lake    June 18, 2017    27.4
5 2017 C20352 (2017) Cariboo      4 km W of Emerald Lake June 23, 2017    25.9
6 2017 C20414 (2017) Cariboo      Raven Lake             June 26, 2017    41.8
# i 3 more variables: geometry <POINT [°]>, Avg_Temperature <dbl>,
#   Avg_Elevation <dbl>

```

Now our data is ready so we will plot a histogram to see the impact of elevation on wildfire occurrence .



From this graph we can interpret that, the number of fire occurrence increases as elevation increases going max around the range of 500-1000 meters which suggest fire occurrence is more common in these elevation range. As elevation decreases from 500m the number of fires also decreasing indicating fire rarely occurs in areas close to sea level. As the elevation increases above 1000m the number of fires decrease with very high elevation even show no fire . This might be again due to cold temperature at these elevations and also no vegetation to ignite the wildfires.

# Question 5

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

For this we have to first download the current data from [gov.bc.ca](http://gov.bc.ca) site and load it inside our R environment . Since it is an excel file we will use `read_excel()` function to load the current wildfire data

```
# Read the current data )
wildfire_current <- read_excel("C:/Users/avich/Downloads/R/Current Wildfire.xlsx")
head(wildfire_current)
```

```
# A tibble: 6 × 24
  FIRE_NO FIRE_YEAR RSPNS_TYPIC IGN_DATE FR_T_DTE FIRESTATUS FIRE_CAUSE
  <chr>     <dbl> <chr>       <dbl>   <dbl> <chr>      <chr>
1 N52580      2024 Full        2.02e13  2.02e13 Out    Lightning
2 G70039      2024 Full        2.02e13  2.02e13 Out    Person
3 G30068      2024 Full        2.02e13  2.02e13 Out    Person
4 G70040      2024 Full        2.02e13  2.02e13 Out    Person
5 G40102      2024 Full        2.02e13  2.02e13 Out    Person
6 G80426      2024 Full        2.02e13  2.02e13 Out    Person
# i 17 more variables: FIRE_CENTR <dbl>, ZONE <dbl>, FIRE_ID <dbl>,
#   FIRE_TYPE <chr>, INCDNT_NM <chr>, GEO_DESC <chr>, LATITUDE <dbl>,
#   LONGITUDE <dbl>, SIZE_HA <dbl>, FIRE_URL <chr>, FR_F_NT_ND <chr>,
#   WS_FR_F_ND <chr>, FCODE <chr>, SHAPE <lgl>, OBJECTID <dbl>,
#   X_COORDINATE <chr>, Y_COORDINATE <chr>
```

After loading the current dataset I noticed there are lot of columns which are really not required for our analysis so we will drop them .

```
# Drop columns that are unnecessary
clean_wildfire <- wildfire_current |>
  select(-RSPNS_TYPIC, -FR_T_DTE, -FIRESTATUS, -FIRE_CAUSE, -ZONE, -FIRE_ID, -FIRE_TYPE, -INCDNT,
         -FIRE_URL, -FR_F_NT_ND, -WS_FR_F_ND, -FCODE, -SHAPE, -OBJECTID, -X_COORDINATE, -Y_COORDINATE)
head(clean_wildfire)
```

```
# A tibble: 6 × 8
  FIRE_NO FIRE_YEAR IGN_DATE FIRE_CENTR GEO_DESC          LATITUDE LONGITUDE SIZE_HA
  <chr>     <dbl> <dbl>       <dbl> <chr>           <dbl>     <dbl>   <dbl>
1 N52580      2024  2.02e13      6 Renata Creek      49.4    -118.     1
2 G70039      2024  2.02e13      4 Radar Lake       55.7    -120.     0.5
3 G30068      2024  2.02e13      4 Murphy Creek      53.2    -120.     0.25
4 G70040      2024  2.02e13      4 Briar Ridge       55.7    -120.    19.8
5 G40102      2024  2.02e13      4 Tachick Lake      54.0    -124.     0.5
6 G80426      2024  2.02e13      4 Tommy Lakes ...     56.9    -122.    0.009
```

Befor changing the column names we can see the column "IGN\_DATE" is in 'YYYYMMDDHHMMSS'. This column matches with the column called "Discovery Date" in our historical data so we will first convert "IGN\_DATE" into 'Month,Date Year' format and then proceed with the change in column names .

```
#Convert IGN_DATE into proper format
clean_wildfire <- clean_wildfire |>
  mutate(
    IGN_DATE = as.character(IGN_DATE),
    IGN_DATE = as.POSIXct(IGN_DATE, format="%Y%m%d%H%M%S", tz="UTC"),
    IGN_DATE = format(IGN_DATE, "%B %d, %Y") # Format as "Month Day, Year"
  )
# View the updated data
head(clean_wildfire)
```

```
# A tibble: 6 × 8
  FIRE_NO FIRE_YEAR IGN_DATE      FIRE_CENTR GEO_DESC LATITUDE LONGITUDE SIZE_HA
  <chr>     <dbl> <chr>          <dbl> <chr>     <dbl>     <dbl>    <dbl>
1 N52580     2024 September 25...       6 Renata ...   49.4    -118.     1
2 G70039     2024 April 07, 20...       4 Radar L...   55.7    -120.     0.5
3 G30068     2024 April 14, 20...       4 Murphy ...   53.2    -120.     0.25
4 G70040     2024 April 07, 20...       4 Briar R...   55.7    -120.    19.8
5 G40102     2024 April 19, 20...       4 Tachick...   54.0    -124.     0.5
6 G80426     2024 May 31, 2024        4 Tommy L...   56.9    -122.    0.009
```

We want to join this dataset to our historical data row wise so we will change column names to the one matching in our historical dataset

```
#Renaming the columns in wildfire_new
clean_wildfire <- clean_wildfire |>
  rename(
    Fire_Number = `FIRE_NO`,
    Discovery_Date = `IGN_DATE`,
    Year = `FIRE_YEAR`,
    Fire_Centre = `FIRE_CENTR`,
    Geographic = `GEO_DESC`,
    Latitude = `LATITUDE`,
    Longitude = `LONGITUDE`,
    Size_ha = `SIZE_HA`)
head(clean_wildfire)
```

```
# A tibble: 6 × 8
  Fire_Number Year Discovery_Date      Fire_Centre Geographic Latitude Longitude
  <chr>        <dbl> <chr>          <dbl> <chr>     <dbl>     <dbl>
1 N52580      2024 September 25, 2024       6 Renata Cr...   49.4    -118.
2 G70039      2024 April 07, 2024           4 Radar Lake    55.7    -120.
3 G30068      2024 April 14, 2024           4 Murphy Cr...   53.2    -120.
4 G70040      2024 April 07, 2024           4 Briar Rid...   55.7    -120.
5 G40102      2024 April 19, 2024           4 Tachick L...   54.0    -124.
6 G80426      2024 May 31, 2024            4 Tommy Lak...   56.9    -122.
# i 1 more variable: Size_ha <dbl>
```

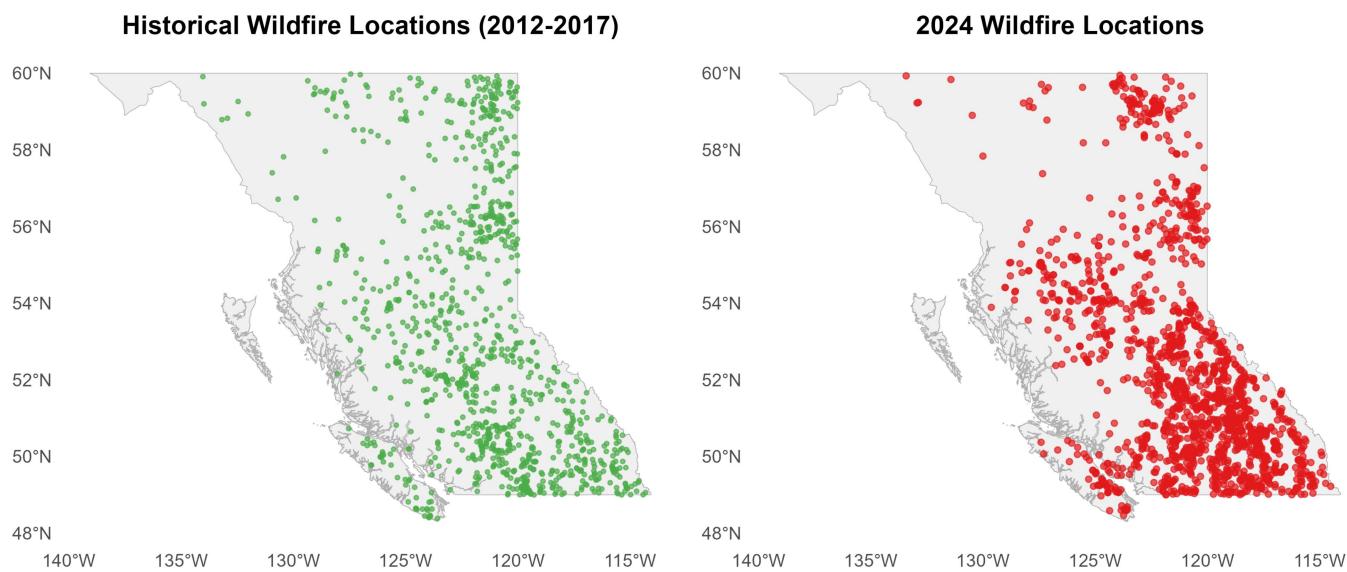
Now we will combine both the historical as well as current wildfire dataset row wise .

```
#Make this column character type because we dont have Fire Centre names in the current w
clean_wildfire$Fire_Centre <- as.character(clean_wildfire$Fire_Centre)
# Combine both the datasets
combined_data <- bind_rows(wildfire_data, clean_wildfire)
head(combined_data)
```

```
# A tibble: 6 × 8
  Year Fire_Number Fire_Centre Latitude Longitude Geographic Discovery_Date
  <dbl> <chr>       <chr>     <dbl>     <dbl> <chr>      <chr>
1 2017 C50011 (2017) Cariboo      51.9     -123. Lee's Corner April 6, 2017
```

2	2017 C50043 (2017) Cariboo	52.0	-123. 3 Km East A... April 21, 2017
3	2017 C10248 (2017) Cariboo	53.2	-123. Pantage Lake June 4, 2017
4	2017 C40329 (2017) Cariboo	51.8	-122. 6km SE of P... June 18, 2017
5	2017 C20352 (2017) Cariboo	51.7	-122. 4 km W of E... June 23, 2017
6	2017 C20414 (2017) Cariboo	52.0	-123. Raven Lake June 26, 2017
# i 1 more variable: Size_ha <dbl>			

Now our dataset is ready to answer the question of how historical distribution (2012-2017) differ from this year's (2024) plot. We will plot both the maps together by using combine <- p1+p2 method taught in class so we can take look at both the plots together.



From the above graph we can interpret the following things :

- The historical wildfires (2012-2017), shown on the left in green, are spread across various regions in B.C., with higher concentrations in the southern and eastern areas. The 2024 wildfires, shown on the right in red, are concentrated, particularly in the central and southern regions of B.C., suggesting a possible increase in wildfire occurrences in these areas.
- Northern regions had fewer fires historically, but in 2024, some wildfire activity is visible there as well, suggesting that fires are now spreading to areas less affected in historical years.
- The change in the pattern of wildfires may reflect differences in climate conditions, vegetation, etc. The increase in density in specific areas can be due to higher temperatures, making certain regions more prone to fires.

The graph suggests that wildfires in 2024 are more intense and more spread, especially in central and southern B.C., compared to the distribution seen in historical 2012-2017 data. This may indicate increasing risk of wildfires in B.C. in the coming years.