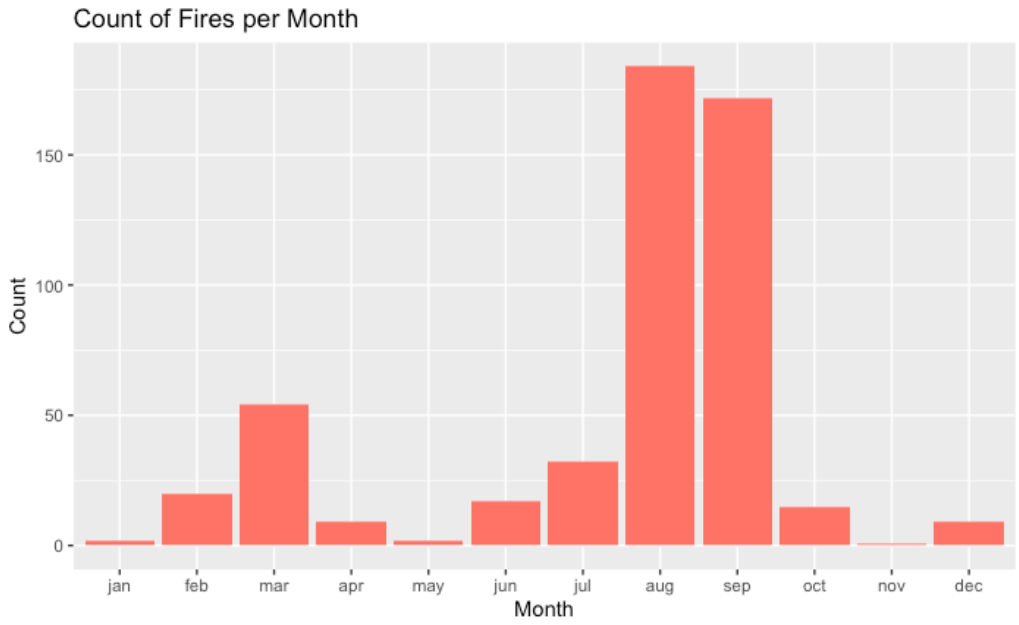
Portugal Forest Fire Report

Data is taken form Montesinho Park in Portugal. There are 13 separate variables found in the datasets which include: X-spatial coordinate, Y-spatial coordinate, Month, Day, FFMC, DMC, DC, ISI, Temp, RH, Wind, Rain, Area. This dataset comes from Paulo Cortez and Anibal Morais from the Department of Information Systems/R&D Algoritmi Centre at the University of Minho in Guimaraes, Portugal. The data is from the Montesinho natural park which is located in the Northeast region of Portugal and it the data collected is from January 2000 to December 2003.

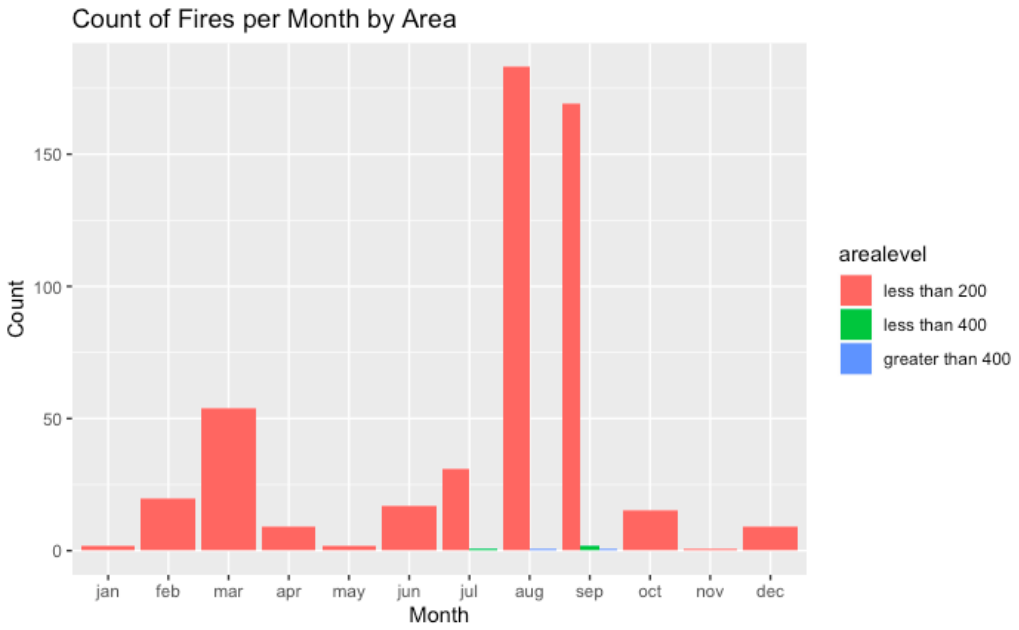
With this data set we will be looking to see what time of year do most forest fires happen Montesinho National Park? What time of year has the largest forest fires Montesinho National Park? Is there a significant linear relationship between temperature, rain, wind speed or relative humidity and the amount of burned area? Do more fires happen in a specific area of Montesinho National Park?

The data was downloaded through as a csv. The data was cleaned by re-ordering the month and day variables. Added a factor variable for area with three levels which included less than 200 hectares, less than 400 hectare and greater than 400 hectares. The data had to have subsets of area level and time of year. The data was checked to see if there were any missing values.

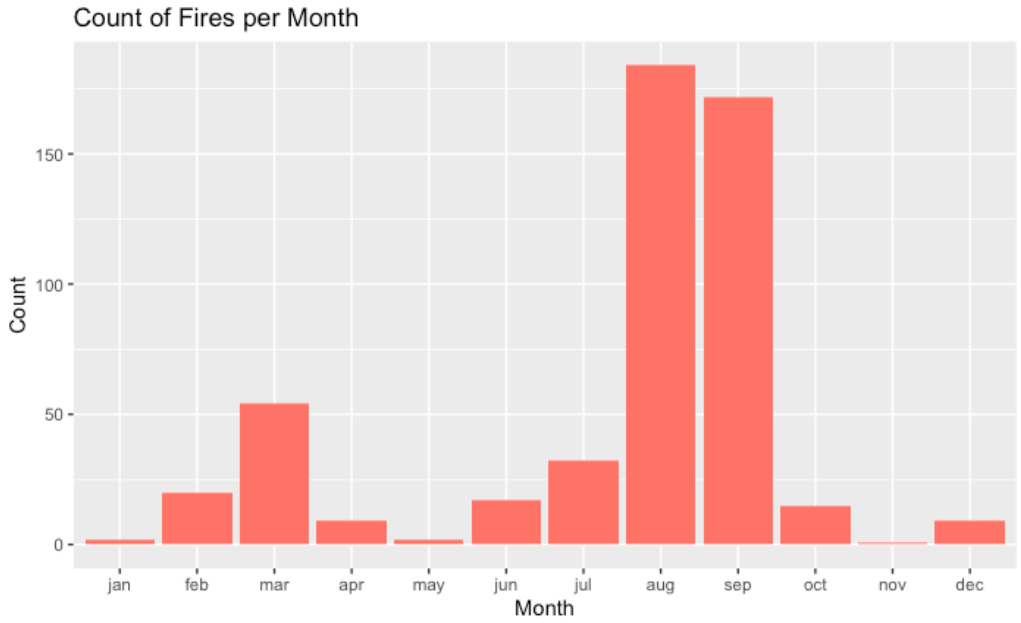
First, we checked which months had the most fires. We used the count of fires that happened each month. The data showed March, August, September had to most fires per month.

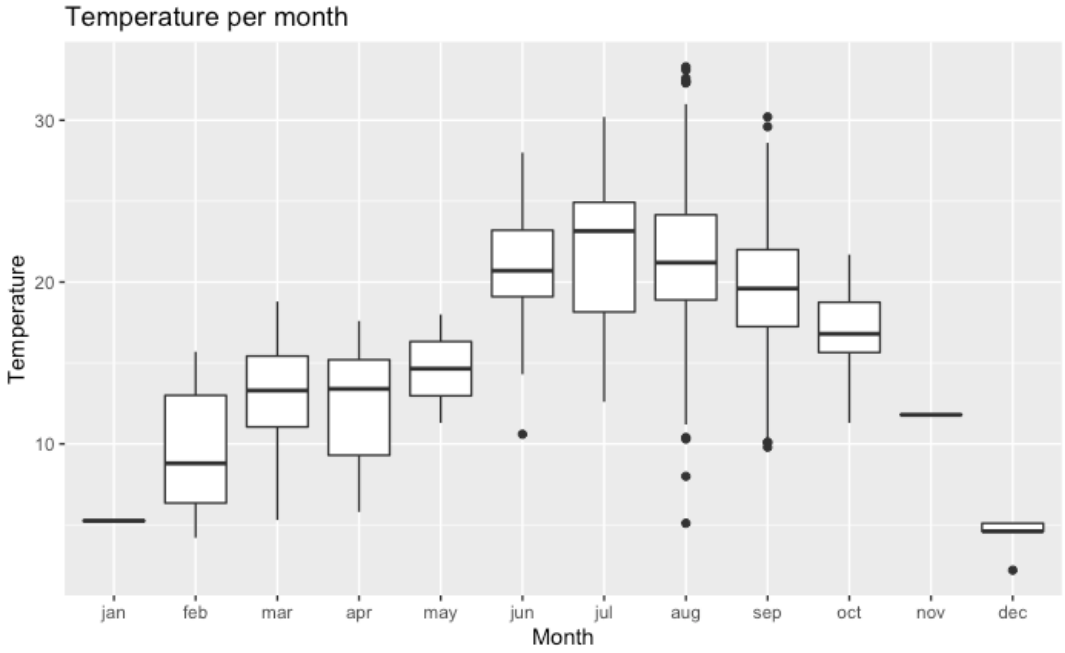


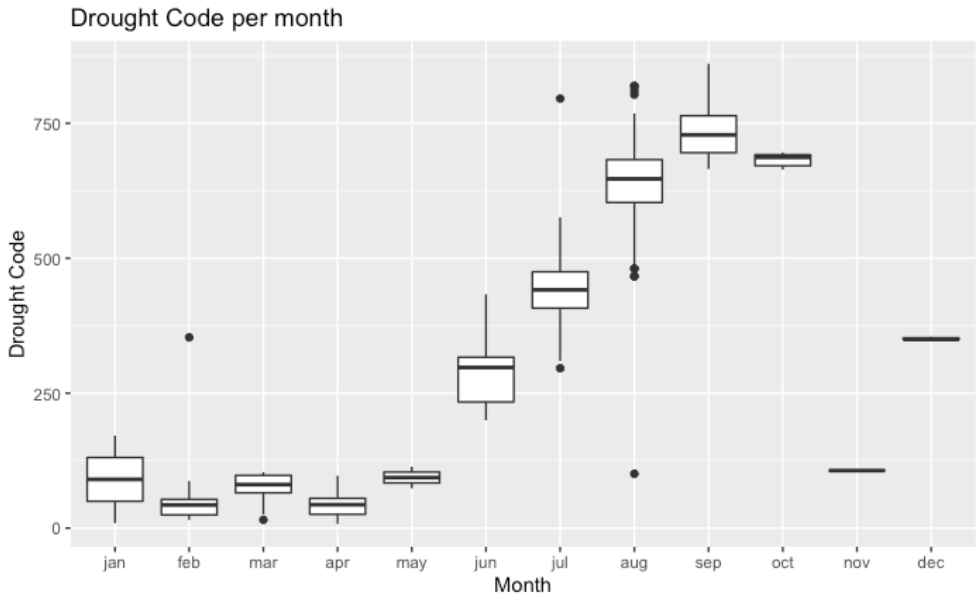
Going more in depth we used our three levels of area to show which months had the greatest amount of area the fire covered. The data showed that July, August, September had the greatest areas of fire. It also shows that there are more smaller fires than there are bigger fires.

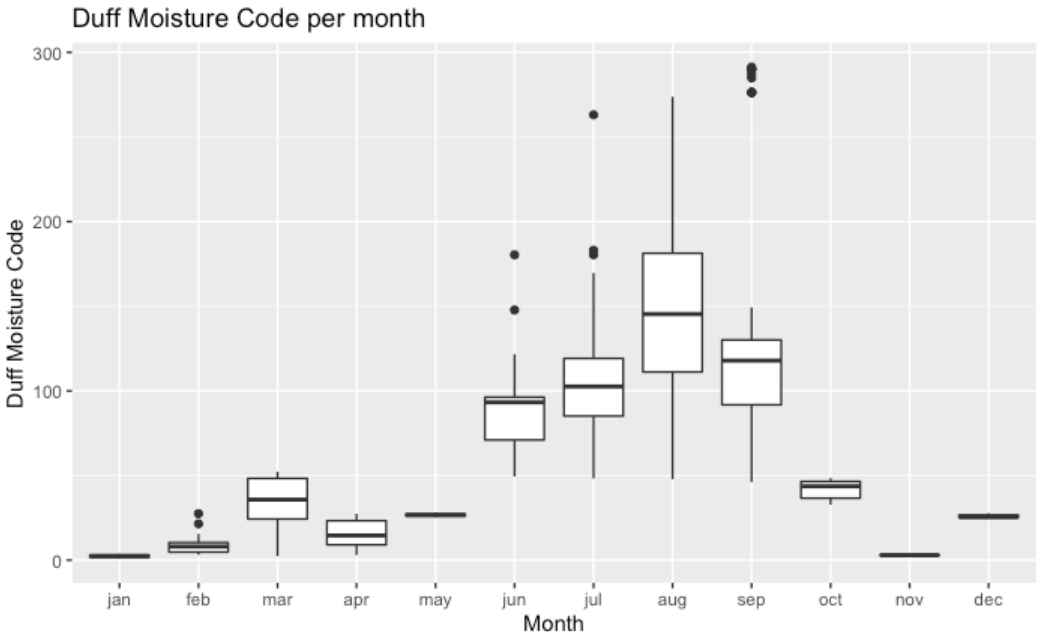


There has been some correlation with our data. Temperature, Drought code and duff moisture all correlate with count of fires per month. They all have higher averages around June through September. We used these variables because they were the only variable that had any correlation. Also, Drought code and duff code are variables that contribute to fire intensity. They measure the moisture of shallow and deep organic layers.

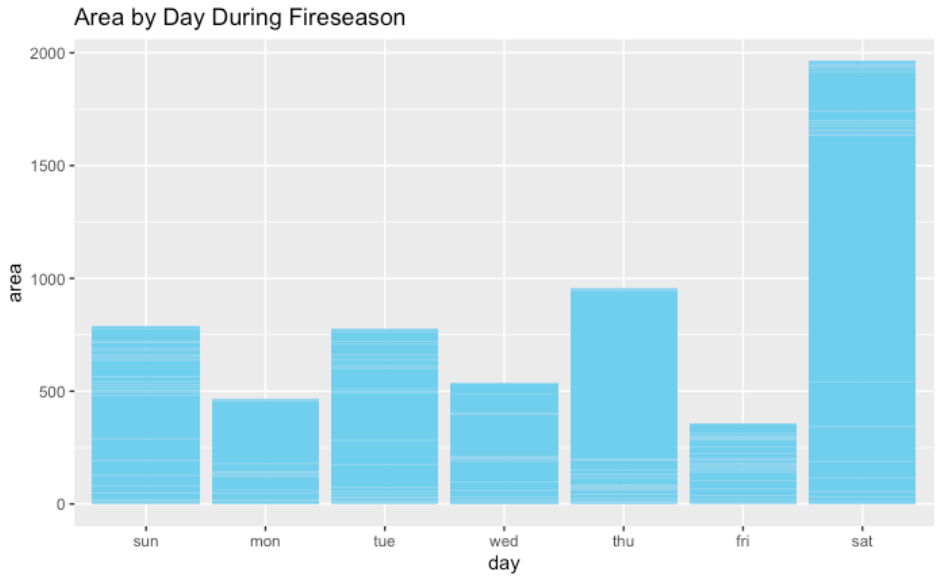


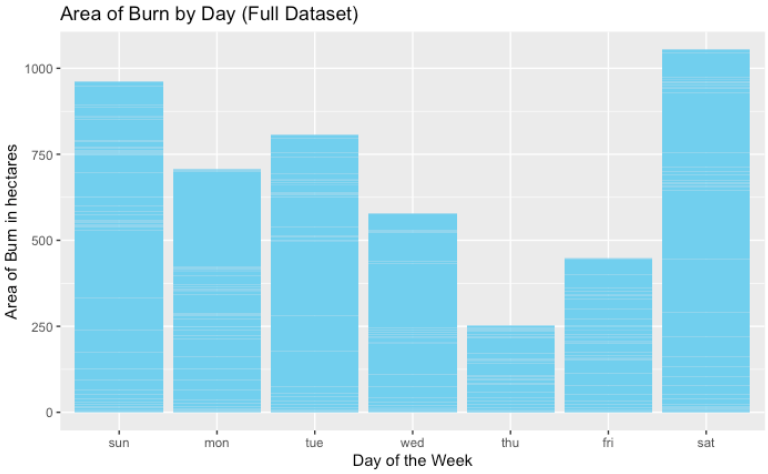


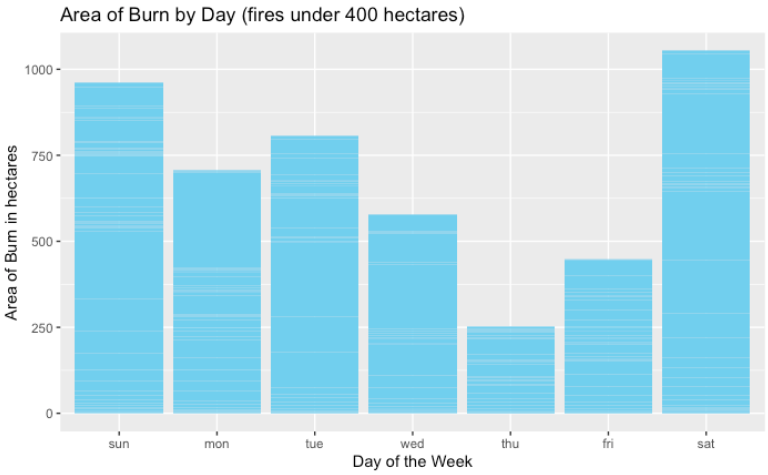




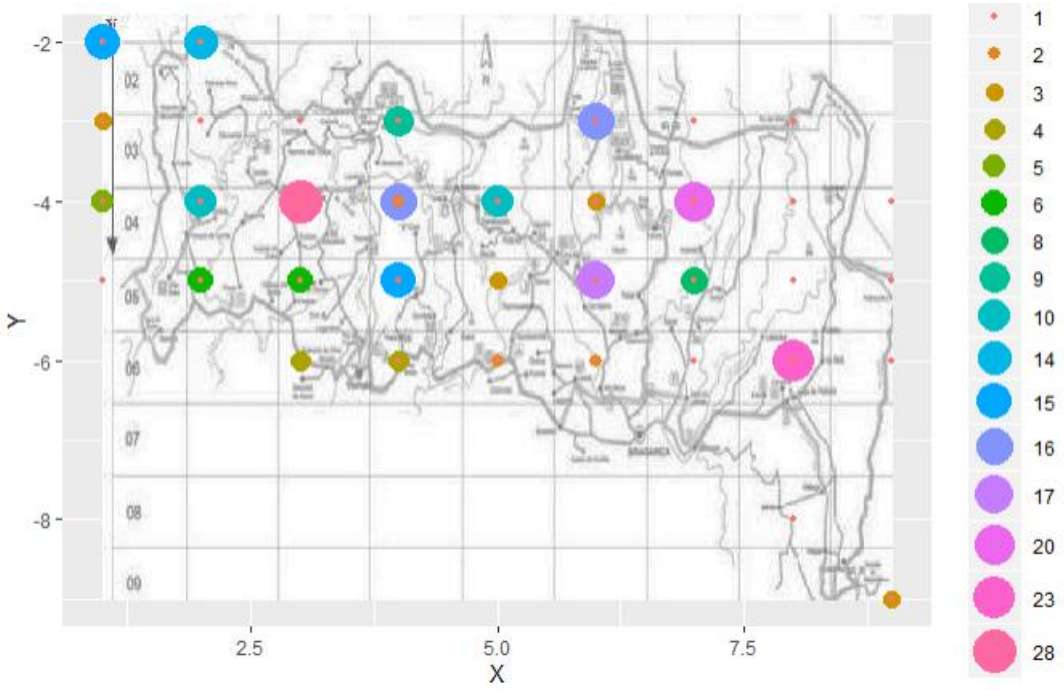
We went even further and tried to see if there were any correlation with weekdays vs area of fire. The data showed that the Saturday has the most area burn of fire. We also saw that the weekends had the largest amount of smaller fires. Overall the weekends had the most area of burn in hectares.

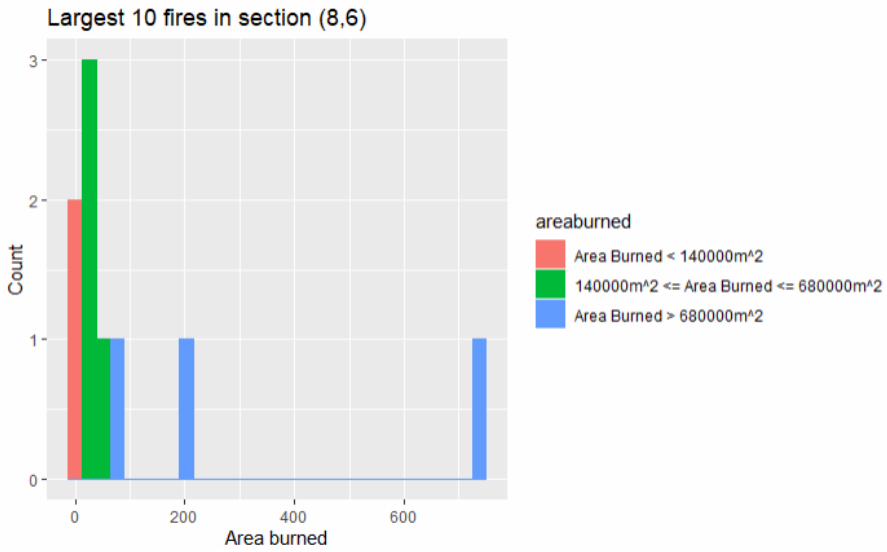

Area by Day During Fireseason
area
2000
1500
1000
500
0
sun
mon
tue
wed
thu
fri
sat
day

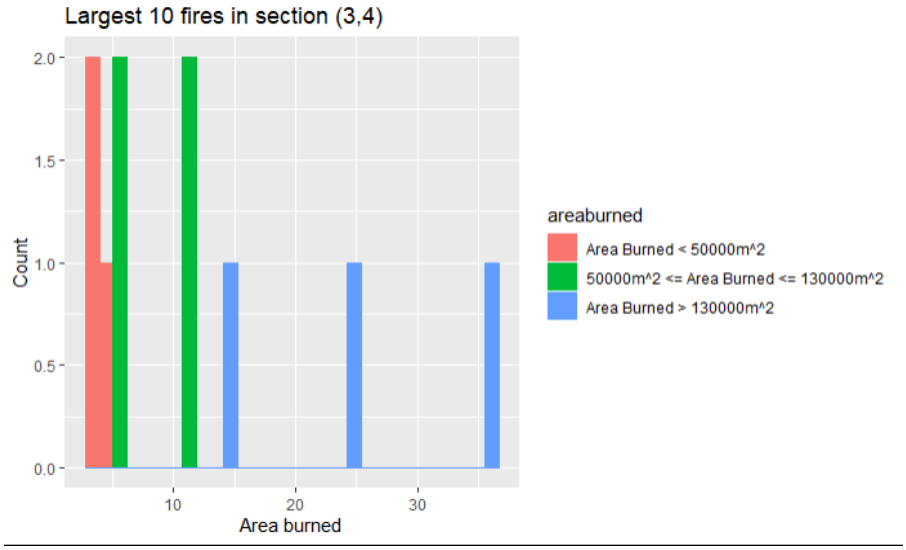

Area of Burn by Day (Full Dataset)
Area of Burn in hectares
1000
750
500
250
0
sun
mon
tue
wed
thu
fri
sat
Day of the Week


Area of Burn by Day (fires under 400 hectares)
Area of Burn in hectares
1000
750
500
250
0
sun
mon
tue
wed
thu
fri
sat
Day of the Week

Last we checked to see where the higher density of fires was. We took a map and saw that it was scattered and there was no significant spot where there were any fires. We as well determined which section had the largest fires. It showed the section







Some of the difficulties we had with our data was that it was difficult to correlate the data and it was difficult to make a model. There was no data that told how many fires happened each year so we could not see if fires would increase as time went by. The variables that were codes (FFMC, DMC, ISI) were difficult to understand what each was telling you and what formulas they used to come up with these variables.

In conclusion it is difficult to predict how large a forest fire will become since there are many other external variables not included in data set such as: Type of vegetation in a certain area, Firefighting intervention. There was also a trend that September, August, and July had the most fires out of the year. Those months also had the highest hectares of area burned. A fifth of fires occurred in sections (3,4) and (8.6)