Contents

[Load packages. 5](#_Toc36457850)

[Data Preparation/Clean-up/Cleaning 10](#_Toc36457851)

[Feature selection using FSelector Use information gain to find best attributes. 27](#_Toc36457852)

[Exploratory Data Analysis 28](#_Toc36457853)

[Outlier detection 104](#_Toc36457854)

[Comparing stats with “Number” of forest fires before and after reducing outliers 108](#_Toc36457855)

[Preparing data frame for CAPPED OUTLIERS 111](#_Toc36457856)

[Discretization test 115](#_Toc36457857)

[17. Prepare packages for models 117](#_Toc36457858)

[18. Split the data into training and test set 118](#_Toc36457859)

[Set the seed to make your partition reproducible 118](#_Toc36457860)

[We want to make sure that the training set and the test set do not have any common data points. 118](#_Toc36457861)

[R built in function “Sample”" randomly selects samples 118](#_Toc36457862)

[19.Set up model with K-fold cross-validation 119](#_Toc36457863)

[Defining the training controls for multiple models 119](#_Toc36457864)

[20.Build the models with start time and end time for each model 120](#_Toc36457865)

[Comment out different models types for variable change. 120](#_Toc36457866)

[21.View the total amount of time taken to run each model. 151](#_Toc36457867)

[22.View summaries of the models. 152](#_Toc36457868)

[23.Evaluation of techniques 172](#_Toc36457869)

[View the results of the models 172](#_Toc36457870)

[24.Find the best results for each model. 176](#_Toc36457871)

[25. Predict on test set 184](#_Toc36457872)

[26. Time taken for prediction 198](#_Toc36457873)

[27. Compare correlation between actual and predicted 199](#_Toc36457874)

[27. RMSE between actual and predicted 200](#_Toc36457875)

[28. MAE between actual and predicted 202](#_Toc36457876)

[29. r squared between actual and predicted 203](#_Toc36457877)

[30. Combine predicted RMSE, MAE R squared, time 206](#_Toc36457878)

[31. Compare the two sets of RMSE, MAE, r squared, time 207](#_Toc36457879)

[30. RMSE, Rsquared MAE on training set 208](#_Toc36457880)

[31.Model below is separate from models above 210](#_Toc36457881)

[Additional for testing: Decision Tree 210](#_Toc36457882)

[32. 250](#_Toc36457883)

[Model below is separate from models above 250](#_Toc36457884)

[Additional for testing: Random Forest 250](#_Toc36457885)

[33.Model below is separate from models above 258](#_Toc36457886)

[Additional for testing: Linear Regression for Predicting Forest Fires. 258](#_Toc36457887)

[34.Model below is separate from models above 264](#_Toc36457888)

[Additional for testing: prediction 264](#_Toc36457889)

[35.RMSE Prediction for Years 269](#_Toc36457890)

[36.Prepare Time Series Analysis. 272](#_Toc36457891)

[37.Model: Time Series Analysis. 277](#_Toc36457892)

[38.Model: Forecasting Time Series ARIMA Analysis. 281](#_Toc36457893)

[39.Analysis of Stationarity 285](#_Toc36457894)

[Autocorrelation function (ACF) of the time series 285](#_Toc36457895)

[40.Plot of the differenced time series 286](#_Toc36457896)

[41.Autocorrelation function (ACF) of the differenced time series 287](#_Toc36457897)

[42.Partial autocorrelation function (PACF) of the differenced time series 288](#_Toc36457898)

[43.Ljung-Box test of autocorrelation 289](#_Toc36457899)

[44. 289](#_Toc36457900)

[Augmented Dickey-Fuller test of stationarity 289](#_Toc36457901)

[40.Additional testing 290](#_Toc36457902)

[Variable Selection for Multiple Linear Regression in R 290](#_Toc36457903)

[41.Additional testing 293](#_Toc36457904)

[Forward and Backward selection algorithm 293](#_Toc36457905)

[42.Additional testing 299](#_Toc36457906)

[Variable selection using automatic methods 299](#_Toc36457907)

[43. 303](#_Toc36457908)

[Additional testing 303](#_Toc36457909)

[Prediction using k Nearest Neighbor Regression 303](#_Toc36457910)

[Additional time series data 310](#_Toc36457911)

[OUTLIERS CAPPED 317](#_Toc36457912)

[OUTLIERS CAPPED 319](#_Toc36457913)

[Prepare packages for models with outliers removed 319](#_Toc36457914)

[OUTLIERS CAPPED 319](#_Toc36457915)

[Split the data into training and test set 319](#_Toc36457916)

[Set the seed to make your partition reproducible 319](#_Toc36457917)

[We want to make sure that the training set and the test set do not have any common data points. 319](#_Toc36457918)

[R built in function “Sample”" randomly selects samples 319](#_Toc36457919)

[OUTLIERS CAPPED 319](#_Toc36457920)

[Set up model 319](#_Toc36457921)

[Set up K-fold cross-validation 319](#_Toc36457922)

[Defining the training controls for multiple models 319](#_Toc36457923)

[OUTLIERS CAPPED 319](#_Toc36457924)

[Build the models with start time and end time for each model 319](#_Toc36457925)

[OUTLIERS CAPPED 346](#_Toc36457926)

[View the total amount of time taken to run each model. 346](#_Toc36457927)

[OUTLIERS CAPPED 346](#_Toc36457928)

[View summaries of the models. 346](#_Toc36457929)

[OUTLIERS CAPPED 367](#_Toc36457930)

[Evaluation of techniques 367](#_Toc36457931)

[View the results of the models 367](#_Toc36457932)

[OUTLIERS CAPPED 369](#_Toc36457933)

[Find the best results for each model. 369](#_Toc36457934)

[OUTLIERS CAPPED 377](#_Toc36457935)

[Predict on test set 377](#_Toc36457936)

[OUTLIERS CAPPED 390](#_Toc36457937)

[Time taken for prediction 390](#_Toc36457938)

[OUTLIERS CAPPED 391](#_Toc36457939)

[Compare correlation between actual and predicted 391](#_Toc36457940)

[OUTLIERS CAPPED 392](#_Toc36457941)

[RMSE between actual and predicted 392](#_Toc36457942)

[OUTLIERS CAPPED 393](#_Toc36457943)

[MAE between actual and predicted 393](#_Toc36457944)

[OUTLIERS CAPPED 394](#_Toc36457945)

[r squared between actual and predicted 394](#_Toc36457946)

[OUTLIERS CAPPED 396](#_Toc36457947)

[Combine predicted RMSE, MAE R squared, time 396](#_Toc36457948)

[OUTLIERS CAPPED 396](#_Toc36457949)

[Compare the two sets of RMSE, MAE, r squared, time 396](#_Toc36457950)

[OUTLIERS CAPPED 397](#_Toc36457951)

[RMSE, Rsquared MAE on training set 397](#_Toc36457952)

CKME 136 Capstone Forest Fires

## Load packages.

Link for “Initial Results and Code” on Github <https://github.com/ed209robo/Ryerson>

#Make sure to use libraries below  
library(dplyr)

## Warning: package 'dplyr' was built under R version 3.5.3

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

#install.packages("ggplot2")  
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.5.3

#install.packages("ggpubr")  
library(ggpubr)

## Loading required package: magrittr

#install.packages("tidyr")  
library(tidyr)

## Warning: package 'tidyr' was built under R version 3.5.3

##   
## Attaching package: 'tidyr'

## The following object is masked from 'package:magrittr':  
##   
## extract

#install.packages("scales")  
library(scales)

## Warning: package 'scales' was built under R version 3.5.3

#install.packages("wesanderson")  
#library(wesanderson)  
#install.packages("viridis") # Install  
library(viridis) # Load

## Warning: package 'viridis' was built under R version 3.5.3

## Loading required package: viridisLite

##   
## Attaching package: 'viridis'

## The following object is masked from 'package:scales':  
##   
## viridis\_pal

#theme\_set(theme\_pubclean())  
library(ggrepel)

## Warning: package 'ggrepel' was built under R version 3.5.3

#install.packages("janitor")  
library(janitor)

## Warning: package 'janitor' was built under R version 3.5.3

##   
## Attaching package: 'janitor'

## The following objects are masked from 'package:stats':  
##   
## chisq.test, fisher.test

#install.packages("PerformanceAnalytics")  
library(PerformanceAnalytics)

## Warning: package 'PerformanceAnalytics' was built under R version 3.5.3

## Loading required package: xts

## Warning: package 'xts' was built under R version 3.5.3

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 3.5.3

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

##   
## Attaching package: 'xts'

## The following objects are masked from 'package:dplyr':  
##   
## first, last

##   
## Attaching package: 'PerformanceAnalytics'

## The following object is masked from 'package:graphics':  
##   
## legend

library(randomForest)

## Warning: package 'randomForest' was built under R version 3.5.3

## randomForest 4.6-14

## Type rfNews() to see new features/changes/bug fixes.

##   
## Attaching package: 'randomForest'

## The following object is masked from 'package:ggplot2':  
##   
## margin

## The following object is masked from 'package:dplyr':  
##   
## combine

#install.packages("e1071")  
library(e1071)

## Warning: package 'e1071' was built under R version 3.5.3

##   
## Attaching package: 'e1071'

## The following objects are masked from 'package:PerformanceAnalytics':  
##   
## kurtosis, skewness

library(caret)

## Warning: package 'caret' was built under R version 3.5.3

## Loading required package: lattice

#install.packages("caretEnsemble")  
library(caretEnsemble)

## Warning: package 'caretEnsemble' was built under R version 3.5.3

##   
## Attaching package: 'caretEnsemble'

## The following object is masked from 'package:ggplot2':  
##   
## autoplot

# install.packages("party")  
library(party)

## Warning: package 'party' was built under R version 3.5.3

## Loading required package: grid

## Loading required package: mvtnorm

## Loading required package: modeltools

## Loading required package: stats4

## Loading required package: strucchange

## Warning: package 'strucchange' was built under R version 3.5.3

## Loading required package: sandwich

## Warning: package 'sandwich' was built under R version 3.5.3

library(rpart)  
#install.packages("formattable")  
#detach(package:MASS, unload=TRUE)  
#library(formattable)  
#install.packages("DT")  
#library(DT)  
#install.packages("FSelector")  
library(FSelector)

## Warning: package 'FSelector' was built under R version 3.5.3

#??FSelector

## Data Preparation/Clean-up/Cleaning

1. Read the “fores fire .csv” file from the following website.

fires <-read.csv('CKME 136 Forest Fire Data.csv',header=T)

1. Have a look at the data set. View(train.data)

head(fires)

## Année Cause Data.qualifier Juridiction Jurisdiction  
## 1 1990 Forest industry a Alberta Alberta  
## 2 1991 Forest industry a Alberta Alberta  
## 3 1992 Forest industry a Alberta Alberta  
## 4 1993 Forest industry a Alberta Alberta  
## 5 1994 Forest industry a Alberta Alberta  
## 6 1995 Forest industry a Alberta Alberta  
## Niveau.d.intervention Number Origine Protection.zone  
## 1 Normale 22 Industrie forestiere Intensive  
## 2 Normale 14 Industrie forestiere Intensive  
## 3 Normale 12 Industrie forestiere Intensive  
## 4 Normale 11 Industrie forestiere Intensive  
## 5 Normale 13 Industrie forestiere Intensive  
## 6 Normale 14 Industrie forestiere Intensive  
## Response.category Year Zone.de.protection  
## 1 Full 1990 Intensive  
## 2 Full 1991 Intensive  
## 3 Full 1992 Intensive  
## 4 Full 1993 Intensive  
## 5 Full 1994 Intensive  
## 6 Full 1995 Intensive

tail(fires)

## Année Cause Data.qualifier  
## 19866 2018 Unspecified human activities p  
## 19867 2018 Unspecified human activities p  
## 19868 2018 Unspecified human activities p  
## 19869 2018 Unspecified human activities p  
## 19870 2018 Unspecified human activities p  
## 19871 2018 Unspecified human activities p  
## Juridiction Jurisdiction Niveau.d.intervention  
## 19866 Nouvelle-Écosse Nova Scotia Modulée  
## 19867 Ontario Ontario Modulée  
## 19868 Île-du-Prince-Édouard Prince Edward Island Modulée  
## 19869 Québec Quebec Modulée  
## 19870 Saskatchewan Saskatchewan Modulée  
## 19871 Yukon Yukon Modulée  
## Number Origine Protection.zone  
## 19866 0 Activités humaines indéterminées Unspecified  
## 19867 11 Activités humaines indéterminées Unspecified  
## 19868 0 Activités humaines indéterminées Unspecified  
## 19869 6 Activités humaines indéterminées Unspecified  
## 19870 13 Activités humaines indéterminées Unspecified  
## 19871 6 Activités humaines indéterminées Unspecified  
## Response.category Year Zone.de.protection  
## 19866 Modified 2018 Indéterminée  
## 19867 Modified 2018 Indéterminée  
## 19868 Modified 2018 Indéterminée  
## 19869 Modified 2018 Indéterminée  
## 19870 Modified 2018 Indéterminée  
## 19871 Modified 2018 Indéterminée

str(fires) #properties and elements of (fires)

## 'data.frame': 19871 obs. of 12 variables:  
## $ Année : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Data.qualifier : Factor w/ 8 levels "a","e","E","n",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Juridiction : Factor w/ 13 levels "Alberta","Colombie-Britannique",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Niveau.d.intervention: Factor w/ 4 levels "Aucune","Indéterminée",..: 4 4 4 4 4 4 4 4 4 4 ...  
## $ Number : int 22 14 12 11 13 14 8 29 10 20 ...  
## $ Origine : Factor w/ 10 levels "Activités humaines indéterminées",..: 8 8 8 8 8 8 8 8 8 8 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category : Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Zone.de.protection : Factor w/ 3 levels "Indéterminée",..: 2 2 2 2 2 2 2 2 2 2 ...

# Check data types of attributes  
sapply(fires, class)

## Année Cause Data.qualifier   
## "integer" "factor" "factor"   
## Juridiction Jurisdiction Niveau.d.intervention   
## "factor" "factor" "factor"   
## Number Origine Protection.zone   
## "integer" "factor" "factor"   
## Response.category Year Zone.de.protection   
## "factor" "integer" "factor"

sapply(fires, typeof)

## Année Cause Data.qualifier   
## "integer" "integer" "integer"   
## Juridiction Jurisdiction Niveau.d.intervention   
## "integer" "integer" "integer"   
## Number Origine Protection.zone   
## "integer" "integer" "integer"   
## Response.category Year Zone.de.protection   
## "integer" "integer" "integer"

1. Extract relevant columns.

new\_fire <- fires[, c("Cause", "Jurisdiction", "Number", "Protection.zone", "Response.category", "Year")]

1. Check for missing values.

sum(is.na(new\_fire$Cause) == TRUE) # 0 Missing values.

## [1] 0

length(new\_fire$Cause)

## [1] 19871

sum(is.na(new\_fire$Jurisdiction) == TRUE) # 0 Missing values.

## [1] 0

length(new\_fire$Jurisdiction)

## [1] 19871

sum(is.na(new\_fire$Number) == TRUE) # 8352 initial missing values for "Number" field.

## [1] 8352

length(new\_fire$Number)

## [1] 19871

sum(is.na(new\_fire$Protection.zone) == TRUE) # 0 Missing values.

## [1] 0

length(new\_fire$Protection.zone)

## [1] 19871

sum(is.na(new\_fire$Response.category) == TRUE) # 0 Missing values.

## [1] 0

length(new\_fire$Response.category)

## [1] 19871

sum(is.na(new\_fire$Year) == TRUE) # 0 Missing values.

## [1] 0

length(new\_fire$Year)

## [1] 19871

1. Only “Number” has missing rows. Remove all rows with missing values.

# Remove remaining records with missing values.  
FireClean <- new\_fire[complete.cases(new\_fire),]  
  
nrow(FireClean) #11519 rows remaining

## [1] 11519

1. Check attributes after missing rows have been removed.

#attach(FireClean)  
  
head(FireClean)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1 Forest industry Alberta 22 Intensive Full  
## 2 Forest industry Alberta 14 Intensive Full  
## 3 Forest industry Alberta 12 Intensive Full  
## 4 Forest industry Alberta 11 Intensive Full  
## 5 Forest industry Alberta 13 Intensive Full  
## 6 Forest industry Alberta 14 Intensive Full  
## Year  
## 1 1990  
## 2 1991  
## 3 1992  
## 4 1993  
## 5 1994  
## 6 1995

tail(FireClean)

## Cause Jurisdiction Number  
## 19866 Unspecified human activities Nova Scotia 0  
## 19867 Unspecified human activities Ontario 11  
## 19868 Unspecified human activities Prince Edward Island 0  
## 19869 Unspecified human activities Quebec 6  
## 19870 Unspecified human activities Saskatchewan 13  
## 19871 Unspecified human activities Yukon 6  
## Protection.zone Response.category Year  
## 19866 Unspecified Modified 2018  
## 19867 Unspecified Modified 2018  
## 19868 Unspecified Modified 2018  
## 19869 Unspecified Modified 2018  
## 19870 Unspecified Modified 2018  
## 19871 Unspecified Modified 2018

str(FireClean)

## 'data.frame': 11519 obs. of 6 variables:  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Number : int 22 14 12 11 13 14 8 29 10 20 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...

dim(FireClean) # 11519 rows, 6 columns

## [1] 11519 6

# Check data types of attributes  
sapply(FireClean, class)

## Cause Jurisdiction Number Protection.zone   
## "factor" "factor" "integer" "factor"   
## Response.category Year   
## "factor" "integer"

levels(FireClean$Cause)

## [1] "Forest industry" "Incendiary"   
## [3] "Lightning" "Miscellaneous known causes"   
## [5] "Other industry" "Railways"   
## [7] "Recreation" "Residents"   
## [9] "Unspecified" "Unspecified human activities"

levels(FireClean$Jurisdiction)

## [1] "Alberta" "British Columbia"   
## [3] "Manitoba" "National parks"   
## [5] "New Brunswick" "Newfoundland and Labrador"  
## [7] "Northwest Territories" "Nova Scotia"   
## [9] "Ontario" "Prince Edward Island"   
## [11] "Quebec" "Saskatchewan"   
## [13] "Yukon"

levels(FireClean$Protection.zone)

## [1] "Intensive" "Limited" "Unspecified"

levels(FireClean$Response.category)

## [1] "Full" "Modified" "None" "Unspecified"

levels(FireClean$Year)

## NULL

summary(FireClean) # Only the "number" attribute maybe usefull with the summary

## Cause Jurisdiction   
## Lightning :1325 Quebec :1516   
## Unspecified :1290 Newfoundland and Labrador:1486   
## Miscellaneous known causes:1279 Ontario :1368   
## Recreation :1274 Manitoba :1304   
## Incendiary :1271 Yukon :1059   
## Residents :1268 Northwest Territories : 773   
## (Other) :3812 (Other) :4013   
## Number Protection.zone Response.category Year   
## Min. : 0.00 Intensive :8190 Full :4204 Min. :1990   
## 1st Qu.: 0.00 Limited :3281 Modified :3521 1st Qu.:1997   
## Median : 0.00 Unspecified: 48 None :3681 Median :2004   
## Mean : 18.64 Unspecified: 113 Mean :2004   
## 3rd Qu.: 4.00 3rd Qu.:2011   
## Max. :2913.00 Max. :2018   
##

1. Shorten “Jurisdiction” name

FireClean$Juris\_Long <- FireClean$Jurisdiction # Duplicate Jurisdiction column  
  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "British Columbia"] <- "BC"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Alberta"] <- "AB"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "National parks"] <- "NP"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Northwest Territories"] <- "NT"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Prince Edward Island"] <- "PE"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Yukon"] <- "YT"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "New Brunswick"] <- "NB"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Nova Scotia"] <- "NS"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Quebec"] <- "QC"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Manitoba"] <- "MB"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Newfoundland and Labrador"] <- "NL"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Ontario"] <- "ON"  
levels(FireClean$Jurisdiction)[levels(FireClean$Jurisdiction) == "Saskatchewan"] <- "SK"  
levels(FireClean$Jurisdiction)

## [1] "AB" "BC" "MB" "NP" "NB" "NL" "NT" "NS" "ON" "PE" "QC" "SK" "YT"

1. create new column: Cause\_Grouped People vs Lightning

# Group Human causes together  
# Lightning is a stand alone cause  
FireClean <- FireClean %>%  
 mutate(Cause\_Grouped = case\_when(  
 Cause == "Lightning" ~ "Lightning",  
 TRUE ~ "People"  
 )  
 )

1. Create new column: Time1 Group Years

# Create new column "Time1" and group years into time periods  
FireClean <- FireClean %>%  
 mutate(Time1 = case\_when(  
 Year <= 1995 ~ "Early 90s",  
 Year >= 1996 & Year <= 2000 ~ "Late 90s",  
 Year >= 2001 & Year <= 2005 ~ "Early 10s",  
 Year >= 2006 & Year <= 2010 ~ "Late 10s",  
 Year >= 2011 & Year <= 2015 ~ "Early 20s",  
 Year >= 2016 ~ "Late 20s"  
 )  
 )

1. Create new column: Time2 Group Years

# Create new column "Time2" and group years into time periods  
FireClean <- FireClean %>%  
 mutate(Time2 = case\_when(  
 Year >= 1990 & Year <= 1999 ~ "1990s",  
 Year >= 2000 & Year <= 2009 ~ "2000s",  
 Year >= 2010 & Year <= 2018 ~ "2010s"  
 )  
 )

1. Group Provinces into regions

# Create new column "Region" and group provinces into regions  
FireClean <- FireClean %>%   
 mutate(Region = case\_when(  
 Jurisdiction %in% c("AB", "MB", "SK") ~ "Prairie Region",  
 Jurisdiction %in% c("BC") ~ "Pacific Region",  
 Jurisdiction %in% c("NP") ~ "National Parks",  
 Jurisdiction %in% c("NB", "NL", "NS", "PE") ~ "Atlantic Region",  
 Jurisdiction %in% c("ON", "QC") ~ "Central Region",  
 Jurisdiction %in% c("YT", "NT") ~ "North Region"  
 )  
 )

1. Check structure of attributes again.

#attach(FireClean)  
  
head(FireClean)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1 Forest industry AB 22 Intensive Full  
## 2 Forest industry AB 14 Intensive Full  
## 3 Forest industry AB 12 Intensive Full  
## 4 Forest industry AB 11 Intensive Full  
## 5 Forest industry AB 13 Intensive Full  
## 6 Forest industry AB 14 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1 1990 Alberta People Early 90s 1990s Prairie Region  
## 2 1991 Alberta People Early 90s 1990s Prairie Region  
## 3 1992 Alberta People Early 90s 1990s Prairie Region  
## 4 1993 Alberta People Early 90s 1990s Prairie Region  
## 5 1994 Alberta People Early 90s 1990s Prairie Region  
## 6 1995 Alberta People Early 90s 1990s Prairie Region

tail(FireClean)

## Cause Jurisdiction Number Protection.zone  
## 11514 Unspecified human activities NS 0 Unspecified  
## 11515 Unspecified human activities ON 11 Unspecified  
## 11516 Unspecified human activities PE 0 Unspecified  
## 11517 Unspecified human activities QC 6 Unspecified  
## 11518 Unspecified human activities SK 13 Unspecified  
## 11519 Unspecified human activities YT 6 Unspecified  
## Response.category Year Juris\_Long Cause\_Grouped Time1  
## 11514 Modified 2018 Nova Scotia People Late 20s  
## 11515 Modified 2018 Ontario People Late 20s  
## 11516 Modified 2018 Prince Edward Island People Late 20s  
## 11517 Modified 2018 Quebec People Late 20s  
## 11518 Modified 2018 Saskatchewan People Late 20s  
## 11519 Modified 2018 Yukon People Late 20s  
## Time2 Region  
## 11514 2010s Atlantic Region  
## 11515 2010s Central Region  
## 11516 2010s Atlantic Region  
## 11517 2010s Central Region  
## 11518 2010s Prairie Region  
## 11519 2010s North Region

str(FireClean)

## 'data.frame': 11519 obs. of 11 variables:  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "AB","BC","MB",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Number : int 22 14 12 11 13 14 8 29 10 20 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Cause\_Grouped : chr "People" "People" "People" "People" ...  
## $ Time1 : chr "Early 90s" "Early 90s" "Early 90s" "Early 90s" ...  
## $ Time2 : chr "1990s" "1990s" "1990s" "1990s" ...  
## $ Region : chr "Prairie Region" "Prairie Region" "Prairie Region" "Prairie Region" ...

dim(FireClean) # 11519 rows, 6 columns

## [1] 11519 11

# Check data types of attributes  
sapply(FireClean, class)

## Cause Jurisdiction Number Protection.zone   
## "factor" "factor" "integer" "factor"   
## Response.category Year Juris\_Long Cause\_Grouped   
## "factor" "integer" "factor" "character"   
## Time1 Time2 Region   
## "character" "character" "character"

levels(FireClean$Cause)

## [1] "Forest industry" "Incendiary"   
## [3] "Lightning" "Miscellaneous known causes"   
## [5] "Other industry" "Railways"   
## [7] "Recreation" "Residents"   
## [9] "Unspecified" "Unspecified human activities"

levels(FireClean$Jurisdiction)

## [1] "AB" "BC" "MB" "NP" "NB" "NL" "NT" "NS" "ON" "PE" "QC" "SK" "YT"

levels(FireClean$Protection.zone)

## [1] "Intensive" "Limited" "Unspecified"

levels(FireClean$Response.category)

## [1] "Full" "Modified" "None" "Unspecified"

levels(FireClean$Year)

## NULL

summary(FireClean) # Only the "number" attribute maybe usefull with the summary

## Cause Jurisdiction Number   
## Lightning :1325 QC :1516 Min. : 0.00   
## Unspecified :1290 NL :1486 1st Qu.: 0.00   
## Miscellaneous known causes:1279 ON :1368 Median : 0.00   
## Recreation :1274 MB :1304 Mean : 18.64   
## Incendiary :1271 YT :1059 3rd Qu.: 4.00   
## Residents :1268 NT : 773 Max. :2913.00   
## (Other) :3812 (Other):4013   
## Protection.zone Response.category Year   
## Intensive :8190 Full :4204 Min. :1990   
## Limited :3281 Modified :3521 1st Qu.:1997   
## Unspecified: 48 None :3681 Median :2004   
## Unspecified: 113 Mean :2004   
## 3rd Qu.:2011   
## Max. :2018   
##   
## Juris\_Long Cause\_Grouped Time1   
## Quebec :1516 Length:11519 Length:11519   
## Newfoundland and Labrador:1486 Class :character Class :character   
## Ontario :1368 Mode :character Mode :character   
## Manitoba :1304   
## Yukon :1059   
## Northwest Territories : 773   
## (Other) :4013   
## Time2 Region   
## Length:11519 Length:11519   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##   
##

summary(FireClean$Number)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 0.00 0.00 18.64 4.00 2913.00

summary(FireClean$Year)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1990 1997 2004 2004 2011 2018

## Feature selection using FSelector Use information gain to find best attributes.

data <- tbl\_df(FireClean)  
  
head(data)

## # A tibble: 6 x 11  
## Cause Jurisdiction Number Protection.zone Response.catego~ Year  
## <fct> <fct> <int> <fct> <fct> <int>  
## 1 Fore~ AB 22 Intensive Full 1990  
## 2 Fore~ AB 14 Intensive Full 1991  
## 3 Fore~ AB 12 Intensive Full 1992  
## 4 Fore~ AB 11 Intensive Full 1993  
## 5 Fore~ AB 13 Intensive Full 1994  
## 6 Fore~ AB 14 Intensive Full 1995  
## # ... with 5 more variables: Juris\_Long <fct>, Cause\_Grouped <chr>,  
## # Time1 <chr>, Time2 <chr>, Region <chr>

weights <- information.gain(Number~., data)  
print(weights)

## attr\_importance  
## Cause 0.062070964  
## Jurisdiction 0.150794057  
## Protection.zone 0.308965544  
## Response.category 0.392000009  
## Year 0.006340241  
## Juris\_Long 0.150794057  
## Cause\_Grouped 0.023934581  
## Time1 0.005551252  
## Time2 0.004146064  
## Region 0.056031029

subset <- cutoff.k(weights, 12)  
  
f <- as.simple.formula(subset, "Number")  
print(f)

## Number ~ Response.category + Protection.zone + Jurisdiction +   
## Juris\_Long + Cause + Region + Cause\_Grouped + Year + Time1 +   
## Time2  
## <environment: 0x000000001e713138>

## Exploratory Data Analysis

1. Check structure of data

head(data)

## # A tibble: 6 x 11  
## Cause Jurisdiction Number Protection.zone Response.catego~ Year  
## <fct> <fct> <int> <fct> <fct> <int>  
## 1 Fore~ AB 22 Intensive Full 1990  
## 2 Fore~ AB 14 Intensive Full 1991  
## 3 Fore~ AB 12 Intensive Full 1992  
## 4 Fore~ AB 11 Intensive Full 1993  
## 5 Fore~ AB 13 Intensive Full 1994  
## 6 Fore~ AB 14 Intensive Full 1995  
## # ... with 5 more variables: Juris\_Long <fct>, Cause\_Grouped <chr>,  
## # Time1 <chr>, Time2 <chr>, Region <chr>

tail(data)

## # A tibble: 6 x 11  
## Cause Jurisdiction Number Protection.zone Response.catego~ Year  
## <fct> <fct> <int> <fct> <fct> <int>  
## 1 Unsp~ NS 0 Unspecified Modified 2018  
## 2 Unsp~ ON 11 Unspecified Modified 2018  
## 3 Unsp~ PE 0 Unspecified Modified 2018  
## 4 Unsp~ QC 6 Unspecified Modified 2018  
## 5 Unsp~ SK 13 Unspecified Modified 2018  
## 6 Unsp~ YT 6 Unspecified Modified 2018  
## # ... with 5 more variables: Juris\_Long <fct>, Cause\_Grouped <chr>,  
## # Time1 <chr>, Time2 <chr>, Region <chr>

str(data)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 11519 obs. of 11 variables:  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "AB","BC","MB",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Number : int 22 14 12 11 13 14 8 29 10 20 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Cause\_Grouped : chr "People" "People" "People" "People" ...  
## $ Time1 : chr "Early 90s" "Early 90s" "Early 90s" "Early 90s" ...  
## $ Time2 : chr "1990s" "1990s" "1990s" "1990s" ...  
## $ Region : chr "Prairie Region" "Prairie Region" "Prairie Region" "Prairie Region" ...

dim(data) # 11519 rows, 6 columns

## [1] 11519 11

# Check data types of attributes  
sapply(data, class)

## Cause Jurisdiction Number Protection.zone   
## "factor" "factor" "integer" "factor"   
## Response.category Year Juris\_Long Cause\_Grouped   
## "factor" "integer" "factor" "character"   
## Time1 Time2 Region   
## "character" "character" "character"

levels(data$Cause)

## [1] "Forest industry" "Incendiary"   
## [3] "Lightning" "Miscellaneous known causes"   
## [5] "Other industry" "Railways"   
## [7] "Recreation" "Residents"   
## [9] "Unspecified" "Unspecified human activities"

levels(data$Jurisdiction)

## [1] "AB" "BC" "MB" "NP" "NB" "NL" "NT" "NS" "ON" "PE" "QC" "SK" "YT"

levels(data$Protection.zone)

## [1] "Intensive" "Limited" "Unspecified"

levels(data$Response.category)

## [1] "Full" "Modified" "None" "Unspecified"

levels(data$Year)

## NULL

summary(data) # Only the "number" attribute maybe usefull with the summary

## Cause Jurisdiction Number   
## Lightning :1325 QC :1516 Min. : 0.00   
## Unspecified :1290 NL :1486 1st Qu.: 0.00   
## Miscellaneous known causes:1279 ON :1368 Median : 0.00   
## Recreation :1274 MB :1304 Mean : 18.64   
## Incendiary :1271 YT :1059 3rd Qu.: 4.00   
## Residents :1268 NT : 773 Max. :2913.00   
## (Other) :3812 (Other):4013   
## Protection.zone Response.category Year   
## Intensive :8190 Full :4204 Min. :1990   
## Limited :3281 Modified :3521 1st Qu.:1997   
## Unspecified: 48 None :3681 Median :2004   
## Unspecified: 113 Mean :2004   
## 3rd Qu.:2011   
## Max. :2018   
##   
## Juris\_Long Cause\_Grouped Time1   
## Quebec :1516 Length:11519 Length:11519   
## Newfoundland and Labrador:1486 Class :character Class :character   
## Ontario :1368 Mode :character Mode :character   
## Manitoba :1304   
## Yukon :1059   
## Northwest Territories : 773   
## (Other) :4013   
## Time2 Region   
## Length:11519 Length:11519   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##   
##

summary(data$Number)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 0.00 0.00 18.64 4.00 2913.00

summary(data$Year)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1990 1997 2004 2004 2011 2018

1. Visulization, Barplot, Boxplot, Exploratory Data Analysis, Outliers

########################################################################################################################################  
#11 STACKED BAR CHART (STACKED BY CAUSE) FOR FOREST FIRES IN BC 1990 - 2018  
pivot3 <- data %>%  
dplyr::select(Year, Number, Cause\_Grouped)  
head(pivot3)

## # A tibble: 6 x 3  
## Year Number Cause\_Grouped  
## <int> <int> <chr>   
## 1 1990 22 People   
## 2 1991 14 People   
## 3 1992 12 People   
## 4 1993 11 People   
## 5 1994 13 People   
## 6 1995 14 People

pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Cause\_Grouped, Number, Year, Jurisdiction) %>%   
 filter(Jurisdiction == "BC") %>%   
 group\_by(Year, Cause\_Grouped) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
  
# Matrix with Cause-Grouped (Lightning & Human), Total number of fires for each year.  
pivot3 %>%   
 spread(Year, sum\_Number)

## # A tibble: 2 x 30  
## Cause\_Grouped `1990` `1991` `1992` `1993` `1994` `1995` `1996` `1997`  
## <chr> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 Lightning 2015 759 2344 609 2913 342 723 675  
## 2 People 1240 1254 1461 888 1175 1132 623 486  
## # ... with 21 more variables: `1998` <int>, `1999` <int>, `2000` <int>,  
## # `2001` <int>, `2002` <int>, `2003` <int>, `2004` <int>, `2005` <int>,  
## # `2006` <int>, `2007` <int>, `2008` <int>, `2009` <int>, `2010` <int>,  
## # `2011` <int>, `2012` <int>, `2013` <int>, `2014` <int>, `2015` <int>,  
## # `2016` <int>, `2017` <int>, `2018` <int>

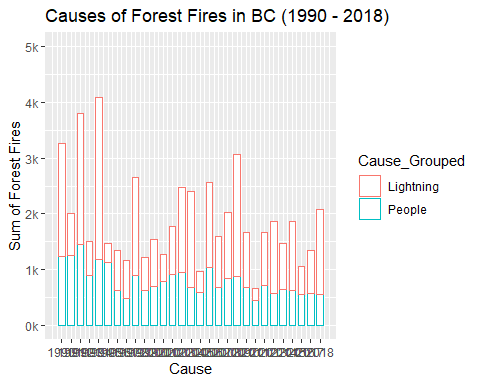
############  
#SUMMING COLUMNS AND ROWS  
pivot4 <- pivot3 %>%   
 spread(Year, sum\_Number)  
  
pivot4 # Matrix with Cause-Grouped (Lightning & Human), Total number of fires for each year.

## # A tibble: 2 x 30  
## Cause\_Grouped `1990` `1991` `1992` `1993` `1994` `1995` `1996` `1997`  
## <chr> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 Lightning 2015 759 2344 609 2913 342 723 675  
## 2 People 1240 1254 1461 888 1175 1132 623 486  
## # ... with 21 more variables: `1998` <int>, `1999` <int>, `2000` <int>,  
## # `2001` <int>, `2002` <int>, `2003` <int>, `2004` <int>, `2005` <int>,  
## # `2006` <int>, `2007` <int>, `2008` <int>, `2009` <int>, `2010` <int>,  
## # `2011` <int>, `2012` <int>, `2013` <int>, `2014` <int>, `2015` <int>,  
## # `2016` <int>, `2017` <int>, `2018` <int>

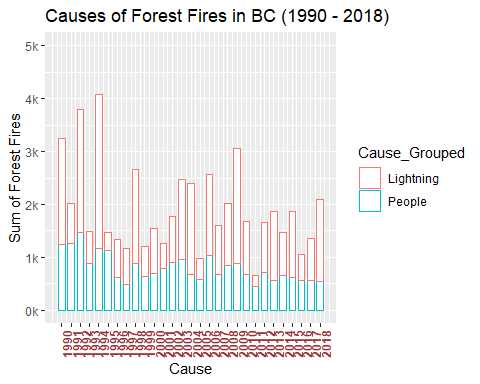
pivot5 <- pivot4 %>%  
 adorn\_totals("row") %>%   
 adorn\_totals("col")   
  
pivot5 # Matrix WITH TOTALS for Cause-Grouped (Lightning & Human), Total number of fires for each year.

## Cause\_Grouped 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001  
## Lightning 2015 759 2344 609 2913 342 723 675 1773 585 842 479  
## People 1240 1254 1461 888 1175 1132 623 486 889 629 697 785  
## Total 3255 2013 3805 1497 4088 1474 1346 1161 2662 1214 1539 1264  
## 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015  
## 870 1513 1716 384 1536 912 1175 2184 992 209 944 1296 816 1237  
## 911 959 682 587 1033 682 848 880 681 446 718 569 649 621  
## 1781 2472 2398 971 2569 1594 2023 3064 1673 655 1662 1865 1465 1858  
## 2016 2017 2018 Total  
## 486 782 1537 32648  
## 563 569 549 23206  
## 1049 1351 2086 55854

##################################################  
# Function to make number scale easier to read on x-axis  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
p <- ggplot(data = pivot3, aes(x = Year,  
 y = sum\_Number,  
 color = Cause\_Grouped)) +  
 geom\_bar(stat="identity", width = 0.7, fill="white") +  
 # geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 # vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 scale\_x\_continuous(breaks=1990:2018)+  
 #stat\_summary(fun.y = sum, aes(label = ..y.., group = Year), geom = "text", vjust= -1.5, show\_guide = F)+  
 xlab("Cause") +  
 ylab("Causes of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 5000))+  
 labs(title = "Causes of Forest Fires in BC (1990 - 2018)")  
  
p



p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 0 ,size = 9, angle = 90))



########################################################################################################################################  
#10B Matrix WITH TOTALS for Regions and total number of fires for each cause (Lightning & Human) in the region.  
pivot3 <- data %>%  
dplyr::select(Region, Number, Cause\_Grouped)  
head(pivot3)

## # A tibble: 6 x 3  
## Region Number Cause\_Grouped  
## <chr> <int> <chr>   
## 1 Prairie Region 22 People   
## 2 Prairie Region 14 People   
## 3 Prairie Region 12 People   
## 4 Prairie Region 11 People   
## 5 Prairie Region 13 People   
## 6 Prairie Region 14 People

pivot3 <- data %>% #Groups Cause and Regions together, sums Number  
 dplyr::select(Region, Number, Cause\_Grouped) %>%   
 group\_by(Cause\_Grouped, Region) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))  
   
# Matrix WITH TOTALS for Regions and total number of fires for each cause (Lightning & Human) in the region.  
pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)

## # A tibble: 6 x 3  
## Region Lightning People  
## <chr> <int> <int>  
## 1 Atlantic Region 2120 21073  
## 2 Central Region 25023 30116  
## 3 National Parks 1421 1096  
## 4 North Region 8549 2163  
## 5 Pacific Region 32648 23206  
## 6 Prairie Region 31234 36116

############  
#SUMMING COLUMNS AND ROWS  
pivot4 <- pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)  
  
pivot4

## # A tibble: 6 x 3  
## Region Lightning People  
## <chr> <int> <int>  
## 1 Atlantic Region 2120 21073  
## 2 Central Region 25023 30116  
## 3 National Parks 1421 1096  
## 4 North Region 8549 2163  
## 5 Pacific Region 32648 23206  
## 6 Prairie Region 31234 36116

pivot5 <- pivot4 %>%  
 adorn\_totals("row") %>%   
 adorn\_totals("col")   
  
pivot5

## Region Lightning People Total  
## Atlantic Region 2120 21073 23193  
## Central Region 25023 30116 55139  
## National Parks 1421 1096 2517  
## North Region 8549 2163 10712  
## Pacific Region 32648 23206 55854  
## Prairie Region 31234 36116 67350  
## Total 100995 113770 214765

##################################################  
#CORRELATION USE FOR pivot3 ONLY  
CorDataFrame <- pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)  
  
CorDataFrame

## # A tibble: 6 x 3  
## Region Lightning People  
## <chr> <int> <int>  
## 1 Atlantic Region 2120 21073  
## 2 Central Region 25023 30116  
## 3 National Parks 1421 1096  
## 4 North Region 8549 2163  
## 5 Pacific Region 32648 23206  
## 6 Prairie Region 31234 36116

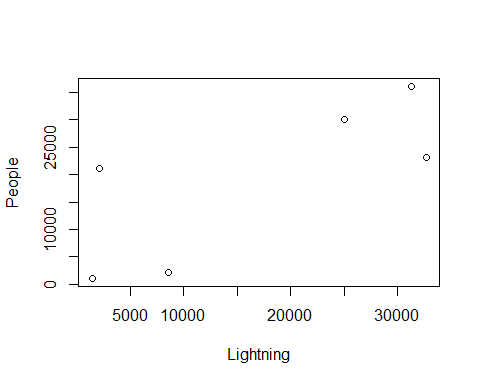
sapply(CorDataFrame, is.numeric) # Which columns are numeric?

## Region Lightning People   
## FALSE TRUE TRUE

my\_num\_data <- CorDataFrame[, sapply(CorDataFrame, is.numeric)] # Subset numeric columns  
my\_num\_data

## # A tibble: 6 x 2  
## Lightning People  
## <int> <int>  
## 1 2120 21073  
## 2 25023 30116  
## 3 1421 1096  
## 4 8549 2163  
## 5 32648 23206  
## 6 31234 36116

plot(my\_num\_data) # Works



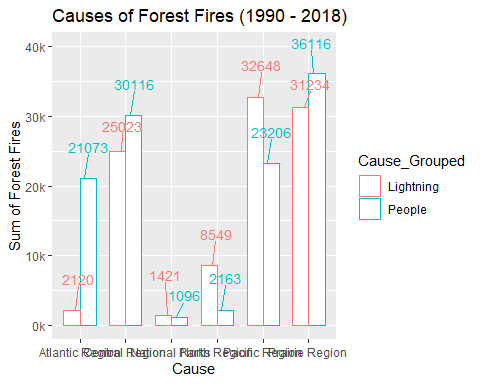
cor(my\_num\_data)

## Lightning People  
## Lightning 1.0000000 0.7550789  
## People 0.7550789 1.0000000

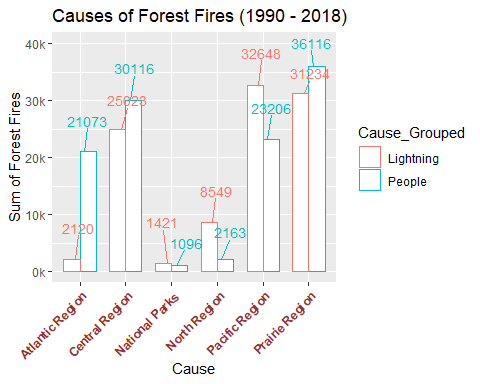
##################################################  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
p <- ggplot(data = pivot3,  
 aes(x = Region, y = sum\_Number, color=Cause\_Grouped)) +  
 geom\_bar(position = "dodge", stat="identity", width = 0.7, fill="white") +  
 geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 xlab("Cause") +  
 ylab("Causes of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 40000))+  
 labs(title = "Causes of Forest Fires (1990 - 2018)")

## Warning: `show\_guide` has been deprecated. Please use `show.legend`  
## instead.

p



p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 1,size = 9, angle = 45))



########################################################################################################################################  
#10 Matrix WITH TOTALS for Regions and total number of fires for each cause (Lightning & Human) in the region.  
pivot3 <- data %>%  
dplyr::select(Region, Number, Cause\_Grouped)  
head(pivot3)

## # A tibble: 6 x 3  
## Region Number Cause\_Grouped  
## <chr> <int> <chr>   
## 1 Prairie Region 22 People   
## 2 Prairie Region 14 People   
## 3 Prairie Region 12 People   
## 4 Prairie Region 11 People   
## 5 Prairie Region 13 People   
## 6 Prairie Region 14 People

pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Region, Number, Cause\_Grouped) %>%   
 group\_by(Cause\_Grouped, Region) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
  
pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)

## # A tibble: 6 x 3  
## Region Lightning People  
## <chr> <int> <int>  
## 1 Atlantic Region 2120 21073  
## 2 Central Region 25023 30116  
## 3 National Parks 1421 1096  
## 4 North Region 8549 2163  
## 5 Pacific Region 32648 23206  
## 6 Prairie Region 31234 36116

##################################################  
#CORRELATION  
CorDataFrame <- pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)  
  
CorDataFrame

## # A tibble: 6 x 3  
## Region Lightning People  
## <chr> <int> <int>  
## 1 Atlantic Region 2120 21073  
## 2 Central Region 25023 30116  
## 3 National Parks 1421 1096  
## 4 North Region 8549 2163  
## 5 Pacific Region 32648 23206  
## 6 Prairie Region 31234 36116

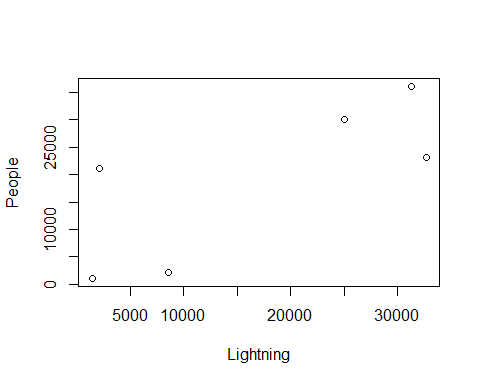
sapply(CorDataFrame, is.numeric) # Which columns are numeric?

## Region Lightning People   
## FALSE TRUE TRUE

my\_num\_data <- CorDataFrame[, sapply(CorDataFrame, is.numeric)] # Subset numeric columns  
my\_num\_data

## # A tibble: 6 x 2  
## Lightning People  
## <int> <int>  
## 1 2120 21073  
## 2 25023 30116  
## 3 1421 1096  
## 4 8549 2163  
## 5 32648 23206  
## 6 31234 36116

plot(my\_num\_data) # Works



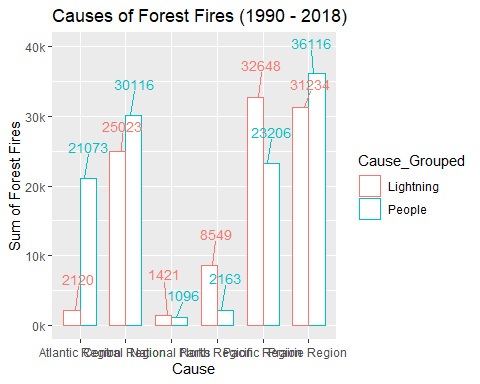
cor(my\_num\_data)

## Lightning People  
## Lightning 1.0000000 0.7550789  
## People 0.7550789 1.0000000

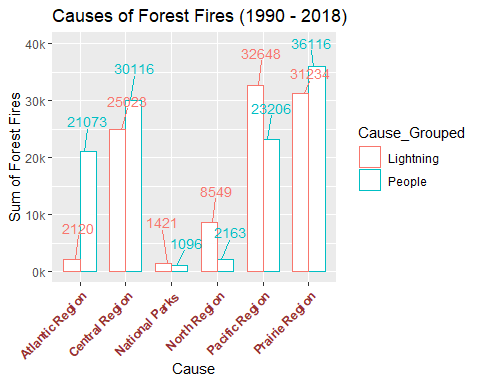
##################################################  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
p <- ggplot(data = pivot3,  
 aes(x = Region, y = sum\_Number, color=Cause\_Grouped)) +  
 geom\_bar(position = "dodge", stat="identity", width = 0.7, fill="white") +  
 geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 xlab("Cause") +  
 ylab("Causes of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 40000))+  
 labs(title = "Causes of Forest Fires (1990 - 2018)")

## Warning: `show\_guide` has been deprecated. Please use `show.legend`  
## instead.

p



p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 1,size = 9, angle = 45))



########################################################################################################################################  
#9 Time Series chart for number of fires in each province (1990-2018)  
pivot3 <- data %>%  
dplyr::select(Year, Number, Jurisdiction)  
head(pivot3)

## # A tibble: 6 x 3  
## Year Number Jurisdiction  
## <int> <int> <fct>   
## 1 1990 22 AB   
## 2 1991 14 AB   
## 3 1992 12 AB   
## 4 1993 11 AB   
## 5 1994 13 AB   
## 6 1995 14 AB

pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Jurisdiction, Number, Year) %>%   
 group\_by(Year, Jurisdiction) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
  
pivot3 %>%   
 spread(Year, sum\_Number)

## # A tibble: 13 x 30  
## Jurisdiction `1990` `1991` `1992` `1993` `1994` `1995` `1996` `1997`  
## <fct> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 AB 1296 923 1055 848 872 803 376 456  
## 2 BC 3255 2013 3805 1497 4088 1474 1346 1161  
## 3 MB 570 676 298 239 555 660 424 373  
## 4 NP 128 56 72 63 166 62 NA 55  
## 5 NB 377 656 576 430 518 547 367 368  
## 6 NL 197 166 109 83 143 103 148 110  
## 7 NT 236 331 285 469 627 215 350 105  
## 8 NS 498 733 299 317 245 408 272 371  
## 9 ON 1614 2560 960 743 1053 2122 1245 1636  
## 10 PE 38 48 27 29 43 29 0 34  
## 11 QC 851 1216 765 543 499 1265 1250 876  
## 12 SK 897 762 701 646 699 650 422 491  
## 13 YT 154 187 116 136 255 148 149 112  
## # ... with 21 more variables: `1998` <int>, `1999` <int>, `2000` <int>,  
## # `2001` <int>, `2002` <int>, `2003` <int>, `2004` <int>, `2005` <int>,  
## # `2006` <int>, `2007` <int>, `2008` <int>, `2009` <int>, `2010` <int>,  
## # `2011` <int>, `2012` <int>, `2013` <int>, `2014` <int>, `2015` <int>,  
## # `2016` <int>, `2017` <int>, `2018` <int>

#Produces matrix with zero  
pivot4 <- pivot3 %>%   
 spread(Year, sum\_Number)  
  
pivot4

## # A tibble: 13 x 30  
## Jurisdiction `1990` `1991` `1992` `1993` `1994` `1995` `1996` `1997`  
## <fct> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 AB 1296 923 1055 848 872 803 376 456  
## 2 BC 3255 2013 3805 1497 4088 1474 1346 1161  
## 3 MB 570 676 298 239 555 660 424 373  
## 4 NP 128 56 72 63 166 62 NA 55  
## 5 NB 377 656 576 430 518 547 367 368  
## 6 NL 197 166 109 83 143 103 148 110  
## 7 NT 236 331 285 469 627 215 350 105  
## 8 NS 498 733 299 317 245 408 272 371  
## 9 ON 1614 2560 960 743 1053 2122 1245 1636  
## 10 PE 38 48 27 29 43 29 0 34  
## 11 QC 851 1216 765 543 499 1265 1250 876  
## 12 SK 897 762 701 646 699 650 422 491  
## 13 YT 154 187 116 136 255 148 149 112  
## # ... with 21 more variables: `1998` <int>, `1999` <int>, `2000` <int>,  
## # `2001` <int>, `2002` <int>, `2003` <int>, `2004` <int>, `2005` <int>,  
## # `2006` <int>, `2007` <int>, `2008` <int>, `2009` <int>, `2010` <int>,  
## # `2011` <int>, `2012` <int>, `2013` <int>, `2014` <int>, `2015` <int>,  
## # `2016` <int>, `2017` <int>, `2018` <int>

#Remove NAs from matrix, replace with zero  
pivot4[is.na(pivot4)] <- 0  
#Matrix now has NAs removed, now replaced with zero  
pivot4

## # A tibble: 13 x 30  
## Jurisdiction `1990` `1991` `1992` `1993` `1994` `1995` `1996` `1997`  
## <fct> <int> <int> <int> <int> <int> <int> <dbl> <int>  
## 1 AB 1296 923 1055 848 872 803 376 456  
## 2 BC 3255 2013 3805 1497 4088 1474 1346 1161  
## 3 MB 570 676 298 239 555 660 424 373  
## 4 NP 128 56 72 63 166 62 0 55  
## 5 NB 377 656 576 430 518 547 367 368  
## 6 NL 197 166 109 83 143 103 148 110  
## 7 NT 236 331 285 469 627 215 350 105  
## 8 NS 498 733 299 317 245 408 272 371  
## 9 ON 1614 2560 960 743 1053 2122 1245 1636  
## 10 PE 38 48 27 29 43 29 0 34  
## 11 QC 851 1216 765 543 499 1265 1250 876  
## 12 SK 897 762 701 646 699 650 422 491  
## 13 YT 154 187 116 136 255 148 149 112  
## # ... with 21 more variables: `1998` <dbl>, `1999` <dbl>, `2000` <dbl>,  
## # `2001` <int>, `2002` <int>, `2003` <int>, `2004` <int>, `2005` <int>,  
## # `2006` <int>, `2007` <int>, `2008` <int>, `2009` <int>, `2010` <int>,  
## # `2011` <int>, `2012` <int>, `2013` <int>, `2014` <int>, `2015` <int>,  
## # `2016` <int>, `2017` <int>, `2018` <int>

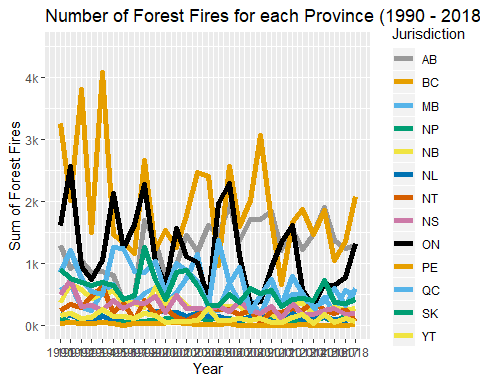
############  
#SUMMING COLUMNS AND ROWS  
  
pivot5 <- pivot4 %>%  
 adorn\_totals("row") %>%   
 adorn\_totals("col")   
  
pivot5

## Jurisdiction 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000  
## AB 1296 923 1055 848 872 803 376 456 1698 1355 783  
## BC 3255 2013 3805 1497 4088 1474 1346 1161 2662 1214 1539  
## MB 570 676 298 239 555 660 424 373 515 613 354  
## NP 128 56 72 63 166 62 0 55 0 0 0  
## NB 377 656 576 430 518 547 367 368 286 607 333  
## NL 197 166 109 83 143 103 148 110 192 228 219  
## NT 236 331 285 469 627 215 350 105 399 170 275  
## NS 498 733 299 317 245 408 272 371 348 464 212  
## ON 1614 2560 960 743 1053 2122 1245 1636 2278 1016 644  
## PE 38 48 27 29 43 29 0 34 27 34 0  
## QC 851 1216 765 543 499 1265 1250 876 854 1037 516  
## SK 897 762 701 646 699 650 422 491 1266 735 419  
## YT 154 187 116 136 255 148 149 112 198 160 55  
## Total 10111 10327 9068 6043 9763 8486 6349 6148 10723 7633 5349  
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014  
## 989 1447 1188 1612 1448 1954 1349 1712 1710 1840 1218 1568 1226 1470  
## 1264 1781 2472 2398 971 2569 1594 2023 3064 1673 655 1662 1865 1465  
## 537 754 1148 234 248 682 382 397 184 581 312 497 494 245  
## 128 86 112 101 107 118 65 108 134 121 76 86 93 84  
## 490 317 228 253 306 308 316 168 197 179 81 344 354 178  
## 202 143 191 153 145 96 87 139 176 61 53 198 101 124  
## 127 85 160 297 261 261 184 241 42 224 207 279 248 385  
## 486 267 272 258 302 234 393 248 198 313 116 352 171 171  
## 1561 1111 1012 424 1961 2298 1124 341 385 931 1334 1619 580 303  
## 46 29 14 20 13 36 8 4 8 4 4 8 10 4  
## 1003 895 716 319 1374 683 935 222 483 737 329 795 515 292  
## 857 880 640 329 323 501 370 599 511 571 302 422 430 403  
## 63 66 77 282 83 80 110 76 118 56 56 126 177 34  
## 7753 7861 8230 6680 7542 9820 6917 6278 7210 7291 4743 7956 6264 5158  
## 2015 2016 2017 2018 Total  
## 1898 1376 1244 1288 37002  
## 1858 1049 1351 2086 55854  
## 459 202 559 477 13669  
## 122 65 169 140 2517  
## 221 285 245 282 9817  
## 128 91 80 132 3998  
## 245 189 262 59 7218  
## 247 274 175 190 8834  
## 668 645 776 1327 34271  
## 5 8 4 10 544  
## 384 602 319 593 20868  
## 720 364 353 416 16679  
## 185 53 115 67 3494  
## 7140 5203 5652 7067 214765

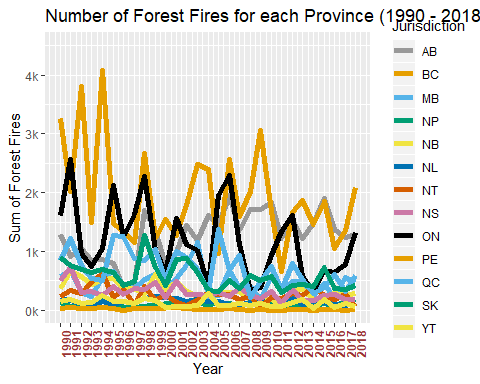
##################################################  
# Test 4 - Time Series chart for number of fires in each province (1990-2018) with thicker lines  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
  
p <- ggplot(data = pivot3, aes(x = Year,  
 y = sum\_Number,  
 color = Jurisdiction)  
 ) +  
 geom\_line(stat="identity", width = 0.7, fill="white", size = 2) +  
 # geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 # vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 scale\_color\_manual(values = c("#999999", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2", "#D55E00", "#CC79A7",  
 "#000000", "#E69F00", "#56B4E9", "#009E73", "#F0E442")) +  
 scale\_x\_continuous(breaks=1990:2018)+  
 #stat\_summary(fun.y = sum, aes(label = ..y.., group = Year), geom = "text", vjust= -1.5, show\_guide = F)+  
 xlab("Year") +  
 ylab("Number of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 4500))+  
 labs(title = "Number of Forest Fires for each Province (1990 - 2018)")

## Warning: Ignoring unknown parameters: width, fill

p



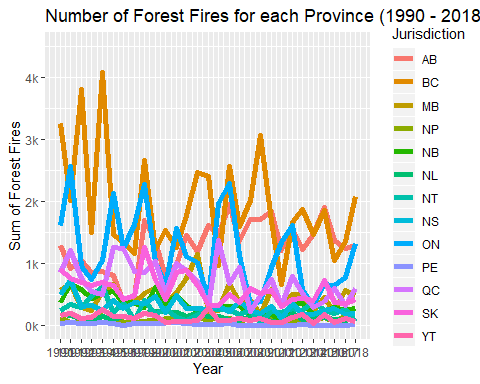
p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 0 ,size = 9, angle = 90))



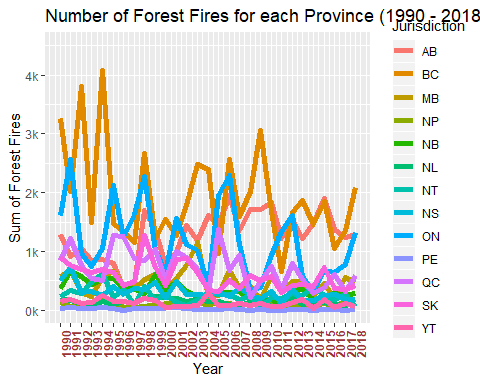
##################################################  
# Test 3 - Time Series chart for number of fires in each province (1990-2018) with thicker lines  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
  
p <- ggplot(data = pivot3, aes(x = Year,  
 y = sum\_Number,  
 color = Jurisdiction)  
 ) +  
 geom\_line(stat="identity", width = 0.7, fill="white", size = 2) +  
 # geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 # vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 scale\_x\_continuous(breaks=1990:2018)+  
 #stat\_summary(fun.y = sum, aes(label = ..y.., group = Year), geom = "text", vjust= -1.5, show\_guide = F)+  
 xlab("Year") +  
 ylab("Number of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 4500))+  
 labs(title = "Number of Forest Fires for each Province (1990 - 2018)")

## Warning: Ignoring unknown parameters: width, fill

p



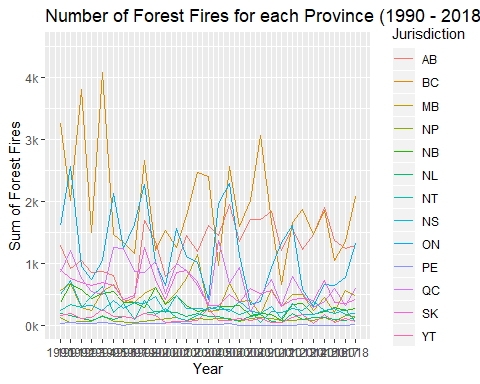
p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 0 ,size = 9, angle = 90))



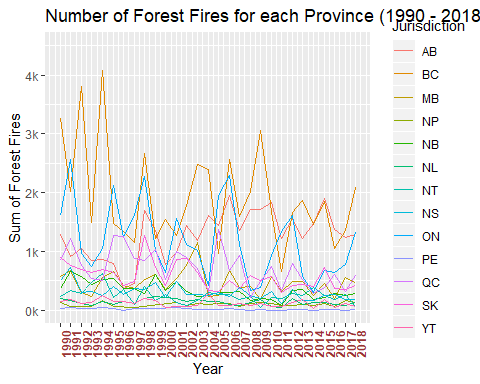
##################################################  
# Test 2 - Time Series chart for number of fires in each province (1990-2018) with thin lines  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
  
p <- ggplot(data = pivot3, aes(x = Year,  
 y = sum\_Number,  
 color = Jurisdiction)  
 ) +  
 geom\_line(stat="identity", width = 0.7, fill="white") +  
 # geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 # vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 scale\_x\_continuous(breaks=1990:2018)+  
 #stat\_summary(fun.y = sum, aes(label = ..y.., group = Year), geom = "text", vjust= -1.5, show\_guide = F)+  
 xlab("Year") +  
 ylab("Number of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 4500))+  
 labs(title = "Number of Forest Fires for each Province (1990 - 2018)")

## Warning: Ignoring unknown parameters: width, fill

p



p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 0 ,size = 9, angle = 90))



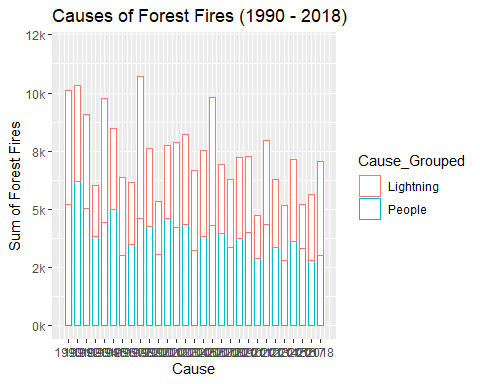
########################################################################################################################################  
#8 - Stacked bar (cause of fire, Lightning vs Human) chart for all provinces by (1990-2018)  
pivot3 <- data %>%  
dplyr::select(Year, Number, Cause\_Grouped)  
head(pivot3)

## # A tibble: 6 x 3  
## Year Number Cause\_Grouped  
## <int> <int> <chr>   
## 1 1990 22 People   
## 2 1991 14 People   
## 3 1992 12 People   
## 4 1993 11 People   
## 5 1994 13 People   
## 6 1995 14 People

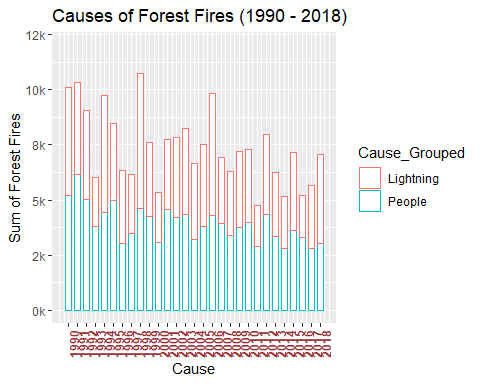
pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Cause\_Grouped, Number, Year) %>%   
 group\_by(Year, Cause\_Grouped) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
  
pivot3 %>%   
 spread(Year, sum\_Number)

## # A tibble: 2 x 30  
## Cause\_Grouped `1990` `1991` `1992` `1993` `1994` `1995` `1996` `1997`  
## <chr> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 Lightning 4895 4146 4021 2229 5324 3504 3317 2652  
## 2 People 5216 6181 5047 3814 4439 4982 3032 3496  
## # ... with 21 more variables: `1998` <int>, `1999` <int>, `2000` <int>,  
## # `2001` <int>, `2002` <int>, `2003` <int>, `2004` <int>, `2005` <int>,  
## # `2006` <int>, `2007` <int>, `2008` <int>, `2009` <int>, `2010` <int>,  
## # `2011` <int>, `2012` <int>, `2013` <int>, `2014` <int>, `2015` <int>,  
## # `2016` <int>, `2017` <int>, `2018` <int>

##################################################  
# Test 8 Stacked bar (cause of fire, Lightning vs Human) chart for all provinces by (1990-2018)  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
p <- ggplot(data = pivot3, aes(x = Year,  
 y = sum\_Number,  
 color = Cause\_Grouped)  
 ) +  
 geom\_bar(stat="identity", width = 0.7, fill="white") +  
 # geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 # vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 scale\_x\_continuous(breaks=1990:2018)+  
 #stat\_summary(fun.y = sum, aes(label = ..y.., group = Year), geom = "text", vjust= -1.5, show\_guide = F)+  
 xlab("Cause") +  
 ylab("Causes of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 12000))+  
 labs(title = "Causes of Forest Fires (1990 - 2018)")  
  
p



p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 0 ,size = 9, angle = 90))



########################################################################################################################################  
#7 Boxplot outlier detection  
pivot3 <- data %>%  
dplyr::select(Jurisdiction, Number, Year)  
head(pivot3)

## # A tibble: 6 x 3  
## Jurisdiction Number Year  
## <fct> <int> <int>  
## 1 AB 22 1990  
## 2 AB 14 1991  
## 3 AB 12 1992  
## 4 AB 11 1993  
## 5 AB 13 1994  
## 6 AB 14 1995

pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Jurisdiction, Number, Year) %>%   
 group\_by(Year, Jurisdiction) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
   
#Produces matrix with zero  
pivot4 <- pivot3 %>%   
 spread(Year, sum\_Number)  
  
pivot4

## # A tibble: 13 x 30  
## Jurisdiction `1990` `1991` `1992` `1993` `1994` `1995` `1996` `1997`  
## <fct> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 AB 1296 923 1055 848 872 803 376 456  
## 2 BC 3255 2013 3805 1497 4088 1474 1346 1161  
## 3 MB 570 676 298 239 555 660 424 373  
## 4 NP 128 56 72 63 166 62 NA 55  
## 5 NB 377 656 576 430 518 547 367 368  
## 6 NL 197 166 109 83 143 103 148 110  
## 7 NT 236 331 285 469 627 215 350 105  
## 8 NS 498 733 299 317 245 408 272 371  
## 9 ON 1614 2560 960 743 1053 2122 1245 1636  
## 10 PE 38 48 27 29 43 29 0 34  
## 11 QC 851 1216 765 543 499 1265 1250 876  
## 12 SK 897 762 701 646 699 650 422 491  
## 13 YT 154 187 116 136 255 148 149 112  
## # ... with 21 more variables: `1998` <int>, `1999` <int>, `2000` <int>,  
## # `2001` <int>, `2002` <int>, `2003` <int>, `2004` <int>, `2005` <int>,  
## # `2006` <int>, `2007` <int>, `2008` <int>, `2009` <int>, `2010` <int>,  
## # `2011` <int>, `2012` <int>, `2013` <int>, `2014` <int>, `2015` <int>,  
## # `2016` <int>, `2017` <int>, `2018` <int>

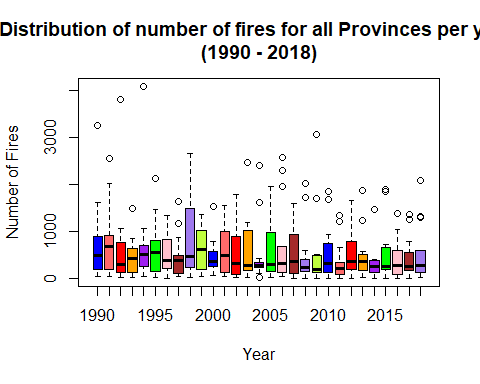
#Remove NAs from matrix, replace with zero  
pivot4[is.na(pivot4)] <- 0  
#Matrix now has NAs removed, now replaced with zero  
pivot4

## # A tibble: 13 x 30  
## Jurisdiction `1990` `1991` `1992` `1993` `1994` `1995` `1996` `1997`  
## <fct> <int> <int> <int> <int> <int> <int> <dbl> <int>  
## 1 AB 1296 923 1055 848 872 803 376 456  
## 2 BC 3255 2013 3805 1497 4088 1474 1346 1161  
## 3 MB 570 676 298 239 555 660 424 373  
## 4 NP 128 56 72 63 166 62 0 55  
## 5 NB 377 656 576 430 518 547 367 368  
## 6 NL 197 166 109 83 143 103 148 110  
## 7 NT 236 331 285 469 627 215 350 105  
## 8 NS 498 733 299 317 245 408 272 371  
## 9 ON 1614 2560 960 743 1053 2122 1245 1636  
## 10 PE 38 48 27 29 43 29 0 34  
## 11 QC 851 1216 765 543 499 1265 1250 876  
## 12 SK 897 762 701 646 699 650 422 491  
## 13 YT 154 187 116 136 255 148 149 112  
## # ... with 21 more variables: `1998` <dbl>, `1999` <dbl>, `2000` <dbl>,  
## # `2001` <int>, `2002` <int>, `2003` <int>, `2004` <int>, `2005` <int>,  
## # `2006` <int>, `2007` <int>, `2008` <int>, `2009` <int>, `2010` <int>,  
## # `2011` <int>, `2012` <int>, `2013` <int>, `2014` <int>, `2015` <int>,  
## # `2016` <int>, `2017` <int>, `2018` <int>

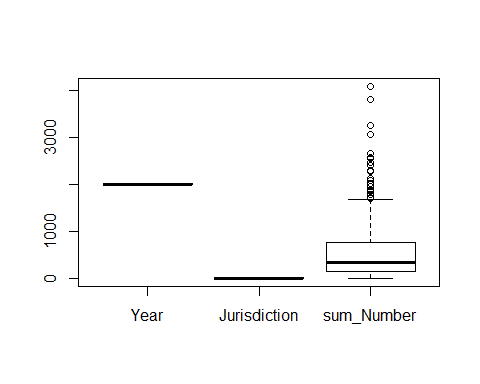
############  
#SUMMING COLUMNS AND ROWS  
  
pivot5 <- pivot4 %>%  
 adorn\_totals("row") %>%   
 adorn\_totals("col")   
  
pivot5

## Jurisdiction 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000  
## AB 1296 923 1055 848 872 803 376 456 1698 1355 783  
## BC 3255 2013 3805 1497 4088 1474 1346 1161 2662 1214 1539  
## MB 570 676 298 239 555 660 424 373 515 613 354  
## NP 128 56 72 63 166 62 0 55 0 0 0  
## NB 377 656 576 430 518 547 367 368 286 607 333  
## NL 197 166 109 83 143 103 148 110 192 228 219  
## NT 236 331 285 469 627 215 350 105 399 170 275  
## NS 498 733 299 317 245 408 272 371 348 464 212  
## ON 1614 2560 960 743 1053 2122 1245 1636 2278 1016 644  
## PE 38 48 27 29 43 29 0 34 27 34 0  
## QC 851 1216 765 543 499 1265 1250 876 854 1037 516  
## SK 897 762 701 646 699 650 422 491 1266 735 419  
## YT 154 187 116 136 255 148 149 112 198 160 55  
## Total 10111 10327 9068 6043 9763 8486 6349 6148 10723 7633 5349  
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014  
## 989 1447 1188 1612 1448 1954 1349 1712 1710 1840 1218 1568 1226 1470  
## 1264 1781 2472 2398 971 2569 1594 2023 3064 1673 655 1662 1865 1465  
## 537 754 1148 234 248 682 382 397 184 581 312 497 494 245  
## 128 86 112 101 107 118 65 108 134 121 76 86 93 84  
## 490 317 228 253 306 308 316 168 197 179 81 344 354 178  
## 202 143 191 153 145 96 87 139 176 61 53 198 101 124  
## 127 85 160 297 261 261 184 241 42 224 207 279 248 385  
## 486 267 272 258 302 234 393 248 198 313 116 352 171 171  
## 1561 1111 1012 424 1961 2298 1124 341 385 931 1334 1619 580 303  
## 46 29 14 20 13 36 8 4 8 4 4 8 10 4  
## 1003 895 716 319 1374 683 935 222 483 737 329 795 515 292  
## 857 880 640 329 323 501 370 599 511 571 302 422 430 403  
## 63 66 77 282 83 80 110 76 118 56 56 126 177 34  
## 7753 7861 8230 6680 7542 9820 6917 6278 7210 7291 4743 7956 6264 5158  
## 2015 2016 2017 2018 Total  
## 1898 1376 1244 1288 37002  
## 1858 1049 1351 2086 55854  
## 459 202 559 477 13669  
## 122 65 169 140 2517  
## 221 285 245 282 9817  
## 128 91 80 132 3998  
## 245 189 262 59 7218  
## 247 274 175 190 8834  
## 668 645 776 1327 34271  
## 5 8 4 10 544  
## 384 602 319 593 20868  
## 720 364 353 416 16679  
## 185 53 115 67 3494  
## 7140 5203 5652 7067 214765

##################################################  
boxplot(sum\_Number~Year,  
 data=pivot3,  
 main="Distribution of number of fires for all Provinces per year \n(1990 - 2018)",  
 xlab="Year",  
 ylab="Number of Fires",  
 col=c("blue", "indianred1","red","orange","purple","green",  
 "pink","brown", "mediumpurple2","olivedrab1"),  
 border="black"  
)



# Outlier detection  
OutVals = boxplot(pivot3)$out



OutVals

## [1] 3255 2013 2560 3805 4088 2122 1698 2662 2278 1781 2472 2398 1961 1954  
## [15] 2569 2298 1712 2023 1710 3064 1840 1865 1898 1858 2086

# Find Outlier index position  
#which(pivot3 )  
########################################################################################################################################  
#6 Boxplot - outlier detection  
pivot3 <- data %>%  
dplyr::select(Jurisdiction, Number, Year)  
head(pivot3)

## # A tibble: 6 x 3  
## Jurisdiction Number Year  
## <fct> <int> <int>  
## 1 AB 22 1990  
## 2 AB 14 1991  
## 3 AB 12 1992  
## 4 AB 11 1993  
## 5 AB 13 1994  
## 6 AB 14 1995

pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Jurisdiction, Number, Year) %>%   
 group\_by(Year, Jurisdiction) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
  
pivot3 %>%   
 spread(Jurisdiction, sum\_Number)

## # A tibble: 29 x 14  
## # Groups: Year [29]  
## Year AB BC MB NP NB NL NT NS ON PE QC  
## <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 1990 1296 3255 570 128 377 197 236 498 1614 38 851  
## 2 1991 923 2013 676 56 656 166 331 733 2560 48 1216  
## 3 1992 1055 3805 298 72 576 109 285 299 960 27 765  
## 4 1993 848 1497 239 63 430 83 469 317 743 29 543  
## 5 1994 872 4088 555 166 518 143 627 245 1053 43 499  
## 6 1995 803 1474 660 62 547 103 215 408 2122 29 1265  
## 7 1996 376 1346 424 NA 367 148 350 272 1245 0 1250  
## 8 1997 456 1161 373 55 368 110 105 371 1636 34 876  
## 9 1998 1698 2662 515 NA 286 192 399 348 2278 27 854  
## 10 1999 1355 1214 613 NA 607 228 170 464 1016 34 1037  
## # ... with 19 more rows, and 2 more variables: SK <int>, YT <int>

#Produces matrix with zero  
pivot4 <- pivot3 %>%   
 spread(Jurisdiction, sum\_Number)  
  
pivot4

## # A tibble: 29 x 14  
## # Groups: Year [29]  
## Year AB BC MB NP NB NL NT NS ON PE QC  
## <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 1990 1296 3255 570 128 377 197 236 498 1614 38 851  
## 2 1991 923 2013 676 56 656 166 331 733 2560 48 1216  
## 3 1992 1055 3805 298 72 576 109 285 299 960 27 765  
## 4 1993 848 1497 239 63 430 83 469 317 743 29 543  
## 5 1994 872 4088 555 166 518 143 627 245 1053 43 499  
## 6 1995 803 1474 660 62 547 103 215 408 2122 29 1265  
## 7 1996 376 1346 424 NA 367 148 350 272 1245 0 1250  
## 8 1997 456 1161 373 55 368 110 105 371 1636 34 876  
## 9 1998 1698 2662 515 NA 286 192 399 348 2278 27 854  
## 10 1999 1355 1214 613 NA 607 228 170 464 1016 34 1037  
## # ... with 19 more rows, and 2 more variables: SK <int>, YT <int>

#Remove NAs from matrix, replace with zero  
pivot4[is.na(pivot4)] <- 0  
#Matrix now has NAs removed, now replaced with zero  
pivot4

## # A tibble: 29 x 14  
## # Groups: Year [29]  
## Year AB BC MB NP NB NL NT NS ON PE QC  
## <int> <int> <int> <int> <dbl> <int> <int> <int> <int> <int> <dbl> <int>  
## 1 1990 1296 3255 570 128 377 197 236 498 1614 38 851  
## 2 1991 923 2013 676 56 656 166 331 733 2560 48 1216  
## 3 1992 1055 3805 298 72 576 109 285 299 960 27 765  
## 4 1993 848 1497 239 63 430 83 469 317 743 29 543  
## 5 1994 872 4088 555 166 518 143 627 245 1053 43 499  
## 6 1995 803 1474 660 62 547 103 215 408 2122 29 1265  
## 7 1996 376 1346 424 0 367 148 350 272 1245 0 1250  
## 8 1997 456 1161 373 55 368 110 105 371 1636 34 876  
## 9 1998 1698 2662 515 0 286 192 399 348 2278 27 854  
## 10 1999 1355 1214 613 0 607 228 170 464 1016 34 1037  
## # ... with 19 more rows, and 2 more variables: SK <int>, YT <int>

############  
#SUMMING COLUMNS AND ROWS  
  
pivot5 <- pivot4 %>%  
 adorn\_totals("row") %>%   
 adorn\_totals("col")   
  
pivot5

## Year AB BC MB NP NB NL NT NS ON PE QC SK  
## 1990 1296 3255 570 128 377 197 236 498 1614 38 851 897  
## 1991 923 2013 676 56 656 166 331 733 2560 48 1216 762  
## 1992 1055 3805 298 72 576 109 285 299 960 27 765 701  
## 1993 848 1497 239 63 430 83 469 317 743 29 543 646  
## 1994 872 4088 555 166 518 143 627 245 1053 43 499 699  
## 1995 803 1474 660 62 547 103 215 408 2122 29 1265 650  
## 1996 376 1346 424 0 367 148 350 272 1245 0 1250 422  
## 1997 456 1161 373 55 368 110 105 371 1636 34 876 491  
## 1998 1698 2662 515 0 286 192 399 348 2278 27 854 1266  
## 1999 1355 1214 613 0 607 228 170 464 1016 34 1037 735  
## 2000 783 1539 354 0 333 219 275 212 644 0 516 419  
## 2001 989 1264 537 128 490 202 127 486 1561 46 1003 857  
## 2002 1447 1781 754 86 317 143 85 267 1111 29 895 880  
## 2003 1188 2472 1148 112 228 191 160 272 1012 14 716 640  
## 2004 1612 2398 234 101 253 153 297 258 424 20 319 329  
## 2005 1448 971 248 107 306 145 261 302 1961 13 1374 323  
## 2006 1954 2569 682 118 308 96 261 234 2298 36 683 501  
## 2007 1349 1594 382 65 316 87 184 393 1124 8 935 370  
## 2008 1712 2023 397 108 168 139 241 248 341 4 222 599  
## 2009 1710 3064 184 134 197 176 42 198 385 8 483 511  
## 2010 1840 1673 581 121 179 61 224 313 931 4 737 571  
## 2011 1218 655 312 76 81 53 207 116 1334 4 329 302  
## 2012 1568 1662 497 86 344 198 279 352 1619 8 795 422  
## 2013 1226 1865 494 93 354 101 248 171 580 10 515 430  
## 2014 1470 1465 245 84 178 124 385 171 303 4 292 403  
## 2015 1898 1858 459 122 221 128 245 247 668 5 384 720  
## 2016 1376 1049 202 65 285 91 189 274 645 8 602 364  
## 2017 1244 1351 559 169 245 80 262 175 776 4 319 353  
## 2018 1288 2086 477 140 282 132 59 190 1327 10 593 416  
## Total 37002 55854 13669 2517 9817 3998 7218 8834 34271 544 20868 16679  
## YT Total  
## 154 10111  
## 187 10327  
## 116 9068  
## 136 6043  
## 255 9763  
## 148 8486  
## 149 6349  
## 112 6148  
## 198 10723  
## 160 7633  
## 55 5349  
## 63 7753  
## 66 7861  
## 77 8230  
## 282 6680  
## 83 7542  
## 80 9820  
## 110 6917  
## 76 6278  
## 118 7210  
## 56 7291  
## 56 4743  
## 126 7956  
## 177 6264  
## 34 5158  
## 185 7140  
## 53 5203  
## 115 5652  
## 67 7067  
## 3494 214765

##################################################  
#CORRELATION  
CorDataFrame <- pivot3 %>%   
 spread(Jurisdiction, sum\_Number)  
  
CorDataFrame

## # A tibble: 29 x 14  
## # Groups: Year [29]  
## Year AB BC MB NP NB NL NT NS ON PE QC  
## <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 1990 1296 3255 570 128 377 197 236 498 1614 38 851  
## 2 1991 923 2013 676 56 656 166 331 733 2560 48 1216  
## 3 1992 1055 3805 298 72 576 109 285 299 960 27 765  
## 4 1993 848 1497 239 63 430 83 469 317 743 29 543  
## 5 1994 872 4088 555 166 518 143 627 245 1053 43 499  
## 6 1995 803 1474 660 62 547 103 215 408 2122 29 1265  
## 7 1996 376 1346 424 NA 367 148 350 272 1245 0 1250  
## 8 1997 456 1161 373 55 368 110 105 371 1636 34 876  
## 9 1998 1698 2662 515 NA 286 192 399 348 2278 27 854  
## 10 1999 1355 1214 613 NA 607 228 170 464 1016 34 1037  
## # ... with 19 more rows, and 2 more variables: SK <int>, YT <int>

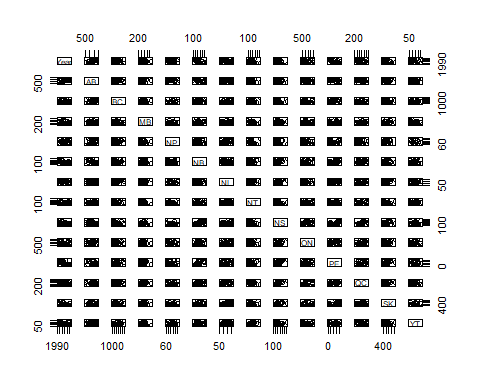
sapply(CorDataFrame, is.numeric) # Which columns are numeric?

## Year AB BC MB NP NB NL NT NS ON PE QC SK YT   
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

my\_num\_data <- CorDataFrame[, sapply(CorDataFrame, is.numeric)] # Subset numeric columns  
my\_num\_data

## # A tibble: 29 x 14  
## # Groups: Year [29]  
## Year AB BC MB NP NB NL NT NS ON PE QC  
## <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>  
## 1 1990 1296 3255 570 128 377 197 236 498 1614 38 851  
## 2 1991 923 2013 676 56 656 166 331 733 2560 48 1216  
## 3 1992 1055 3805 298 72 576 109 285 299 960 27 765  
## 4 1993 848 1497 239 63 430 83 469 317 743 29 543  
## 5 1994 872 4088 555 166 518 143 627 245 1053 43 499  
## 6 1995 803 1474 660 62 547 103 215 408 2122 29 1265  
## 7 1996 376 1346 424 NA 367 148 350 272 1245 0 1250  
## 8 1997 456 1161 373 55 368 110 105 371 1636 34 876  
## 9 1998 1698 2662 515 NA 286 192 399 348 2278 27 854  
## 10 1999 1355 1214 613 NA 607 228 170 464 1016 34 1037  
## # ... with 19 more rows, and 2 more variables: SK <int>, YT <int>

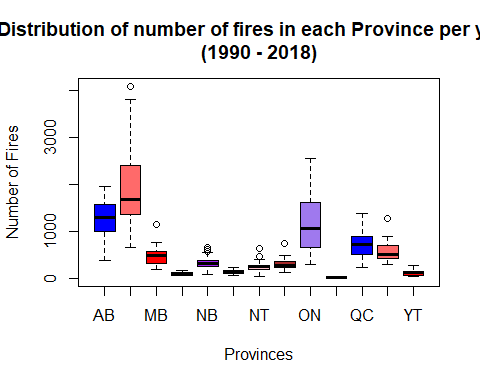
plot(my\_num\_data) # Works



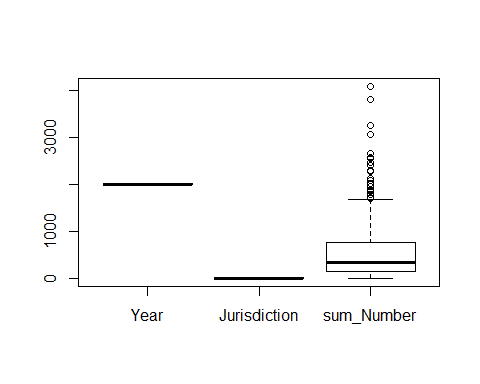
cor(my\_num\_data)

## Year AB BC MB NP NB  
## Year 1.0000000 0.546335923 -0.36110450 -0.185100904 NA -0.7000538  
## AB 0.5463359 1.000000000 0.15023454 -0.007571187 NA -0.5189530  
## BC -0.3611045 0.150234537 1.00000000 0.161751543 NA 0.2114017  
## MB -0.1851009 -0.007571187 0.16175154 1.000000000 NA 0.2038883  
## NP NA NA NA NA 1 NA  
## NB -0.7000538 -0.518953042 0.21140166 0.203888335 NA 1.0000000  
## NL -0.3317171 -0.041129119 0.22401619 0.247988400 NA 0.3037542  
## NT -0.3474347 -0.172973137 0.30415171 -0.120227227 NA 0.2320419  
## NS -0.6080368 -0.232706662 0.01817127 0.286415878 NA 0.6993488  
## ON -0.4188543 -0.148060585 -0.01483963 0.375743689 NA 0.4197618  
## PE NA NA NA NA NA NA  
## QC -0.5246978 -0.367004187 -0.16781797 0.281679330 NA 0.5951643  
## SK -0.5412805 0.062870086 0.41185714 0.418406166 NA 0.3824341  
## YT -0.3859273 -0.076749198 0.44794406 0.015002077 NA 0.3890247  
## NL NT NS ON PE QC  
## Year -0.33171713 -0.347434654 -0.608036773 -0.41885428 NA -0.52469782  
## AB -0.04112912 -0.172973137 -0.232706662 -0.14806059 NA -0.36700419  
## BC 0.22401619 0.304151713 0.018171267 -0.01483963 NA -0.16781797  
## MB 0.24798840 -0.120227227 0.286415878 0.37574369 NA 0.28167933  
## NP NA NA NA NA NA NA  
## NB 0.30375423 0.232041878 0.699348814 0.41976184 NA 0.59516430  
## NL 1.00000000 -0.056363050 0.378076786 0.12265643 NA 0.25919831  
## NT -0.05636305 1.000000000 0.003673412 0.02984490 NA -0.11646900  
## NS 0.37807679 0.003673412 1.000000000 0.58423893 NA 0.67343110  
## ON 0.12265643 0.029844896 0.584238930 1.00000000 NA 0.70744952  
## PE NA NA NA NA 1 NA  
## QC 0.25919831 -0.116469001 0.673431099 0.70744952 NA 1.00000000  
## SK 0.42672134 0.134702176 0.490365233 0.36364534 NA 0.26994050  
## YT 0.20183017 0.487585257 0.253303187 0.09305185 NA 0.03992442  
## SK YT  
## Year -0.54128047 -0.38592735  
## AB 0.06287009 -0.07674920  
## BC 0.41185714 0.44794406  
## MB 0.41840617 0.01500208  
## NP NA NA  
## NB 0.38243408 0.38902468  
## NL 0.42672134 0.20183017  
## NT 0.13470218 0.48758526  
## NS 0.49036523 0.25330319  
## ON 0.36364534 0.09305185  
## PE NA NA  
## QC 0.26994050 0.03992442  
## SK 1.00000000 0.27080987  
## YT 0.27080987 1.00000000

##################################################  
# Test 1  
boxplot(sum\_Number~Jurisdiction,  
 data=pivot3,  
 main="Distribution of number of fires in each Province per year \n(1990 - 2018)",  
 xlab="Provinces",  
 ylab="Number of Fires",  
 col=c("blue", "indianred1","red","orange","purple","green",  
 "pink","brown", "mediumpurple2","olivedrab1"),  
 border="black"  
)



# Outlier detection  
OutVals = boxplot(pivot3)$out



OutVals

## [1] 3255 2013 2560 3805 4088 2122 1698 2662 2278 1781 2472 2398 1961 1954  
## [15] 2569 2298 1712 2023 1710 3064 1840 1865 1898 1858 2086

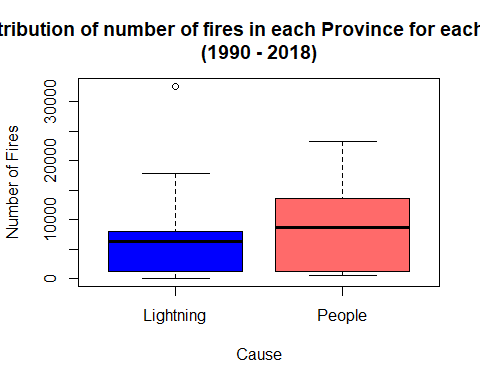
########################################################################################################################################  
#5 Boxplot - outlier detection  
pivot3 <- data %>%  
dplyr::select(Jurisdiction, Number, Cause\_Grouped)  
head(pivot3)

## # A tibble: 6 x 3  
## Jurisdiction Number Cause\_Grouped  
## <fct> <int> <chr>   
## 1 AB 22 People   
## 2 AB 14 People   
## 3 AB 12 People   
## 4 AB 11 People   
## 5 AB 13 People   
## 6 AB 14 People

pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Jurisdiction, Number, Cause\_Grouped) %>%   
 group\_by(Cause\_Grouped, Jurisdiction) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
   
pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)

## # A tibble: 13 x 3  
## Jurisdiction Lightning People  
## <fct> <int> <int>  
## 1 AB 16457 20545  
## 2 BC 32648 23206  
## 3 MB 6838 6831  
## 4 NP 1421 1096  
## 5 NB 1156 8661  
## 6 NL 765 3233  
## 7 NT 6274 944  
## 8 NS 198 8636  
## 9 ON 17782 16489  
## 10 PE 1 543  
## 11 QC 7241 13627  
## 12 SK 7939 8740  
## 13 YT 2275 1219

###############################################  
# Test 2  
boxplot(sum\_Number~Cause\_Grouped,  
 data=pivot3,  
 main="Distribution of number of fires in each Province for each cause \n(1990 - 2018)",  
 xlab="Cause",  
 ylab="Number of Fires",  
 col=c("blue", "indianred1","red","orange","purple","green",  
 "pink","brown", "mediumpurple2","olivedrab1"),  
 border="black"  
)



# Outlier detection  
#OutVals = boxplot(pivot3)$out  
#OutVals  
########################################################################################################################################  
#4  
pivot3 <- data %>%  
dplyr::select(Jurisdiction, Number, Cause\_Grouped)  
head(pivot3)

## # A tibble: 6 x 3  
## Jurisdiction Number Cause\_Grouped  
## <fct> <int> <chr>   
## 1 AB 22 People   
## 2 AB 14 People   
## 3 AB 12 People   
## 4 AB 11 People   
## 5 AB 13 People   
## 6 AB 14 People

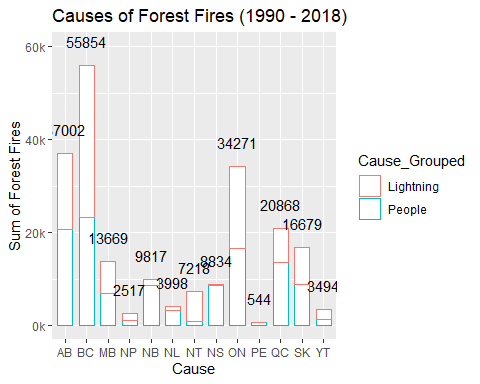
pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Jurisdiction, Number, Cause\_Grouped) %>%   
 group\_by(Cause\_Grouped, Jurisdiction) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
   
pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)

## # A tibble: 13 x 3  
## Jurisdiction Lightning People  
## <fct> <int> <int>  
## 1 AB 16457 20545  
## 2 BC 32648 23206  
## 3 MB 6838 6831  
## 4 NP 1421 1096  
## 5 NB 1156 8661  
## 6 NL 765 3233  
## 7 NT 6274 944  
## 8 NS 198 8636  
## 9 ON 17782 16489  
## 10 PE 1 543  
## 11 QC 7241 13627  
## 12 SK 7939 8740  
## 13 YT 2275 1219

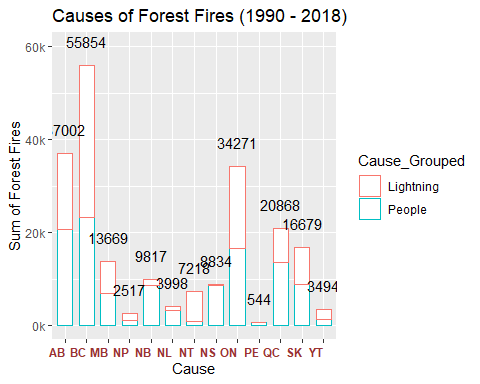
##################################################  
# Test 5  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
p <- ggplot(data = pivot3, aes(x = Jurisdiction,  
 y = sum\_Number,  
 color = Cause\_Grouped)  
 ) +  
 geom\_bar(stat="identity", width = 0.7, fill="white") +  
 # geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 # vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 stat\_summary(fun.y = sum, aes(label = ..y.., group = Jurisdiction), geom = "text", vjust= -1.5, show\_guide = F)+  
 xlab("Cause") +  
 ylab("Causes of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 60000))+  
 labs(title = "Causes of Forest Fires (1990 - 2018)")

## Warning: `show\_guide` has been deprecated. Please use `show.legend`  
## instead.

p



p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 1,size = 9, angle = 0))



########################################################################################################################################  
#3  
pivot3 <- data %>%  
dplyr::select(Jurisdiction, Number, Cause\_Grouped)  
head(pivot3)

## # A tibble: 6 x 3  
## Jurisdiction Number Cause\_Grouped  
## <fct> <int> <chr>   
## 1 AB 22 People   
## 2 AB 14 People   
## 3 AB 12 People   
## 4 AB 11 People   
## 5 AB 13 People   
## 6 AB 14 People

# attach(pivot3)  
# detach(pivot3)  
  
pivot3 <- data %>% #Groups Cause together and sums Number  
 dplyr::select(Jurisdiction, Number, Cause\_Grouped) %>%   
 group\_by(Cause\_Grouped, Jurisdiction) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))   
   
pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)

## # A tibble: 13 x 3  
## Jurisdiction Lightning People  
## <fct> <int> <int>  
## 1 AB 16457 20545  
## 2 BC 32648 23206  
## 3 MB 6838 6831  
## 4 NP 1421 1096  
## 5 NB 1156 8661  
## 6 NL 765 3233  
## 7 NT 6274 944  
## 8 NS 198 8636  
## 9 ON 17782 16489  
## 10 PE 1 543  
## 11 QC 7241 13627  
## 12 SK 7939 8740  
## 13 YT 2275 1219

##################################################  
#CORRELATION  
CorDataFrame <- pivot3 %>%   
 spread(Cause\_Grouped, sum\_Number)  
  
CorDataFrame

## # A tibble: 13 x 3  
## Jurisdiction Lightning People  
## <fct> <int> <int>  
## 1 AB 16457 20545  
## 2 BC 32648 23206  
## 3 MB 6838 6831  
## 4 NP 1421 1096  
## 5 NB 1156 8661  
## 6 NL 765 3233  
## 7 NT 6274 944  
## 8 NS 198 8636  
## 9 ON 17782 16489  
## 10 PE 1 543  
## 11 QC 7241 13627  
## 12 SK 7939 8740  
## 13 YT 2275 1219

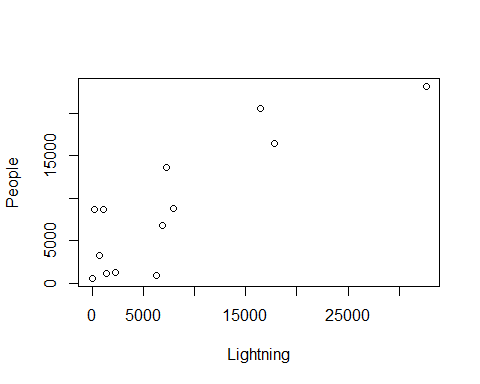
sapply(CorDataFrame, is.numeric) # Which columns are numeric?

## Jurisdiction Lightning People   
## FALSE TRUE TRUE

my\_num\_data <- CorDataFrame[, sapply(CorDataFrame, is.numeric)] # Subset numeric columns  
my\_num\_data

## # A tibble: 13 x 2  
## Lightning People  
## <int> <int>  
## 1 16457 20545  
## 2 32648 23206  
## 3 6838 6831  
## 4 1421 1096  
## 5 1156 8661  
## 6 765 3233  
## 7 6274 944  
## 8 198 8636  
## 9 17782 16489  
## 10 1 543  
## 11 7241 13627  
## 12 7939 8740  
## 13 2275 1219

plot(my\_num\_data) # Works



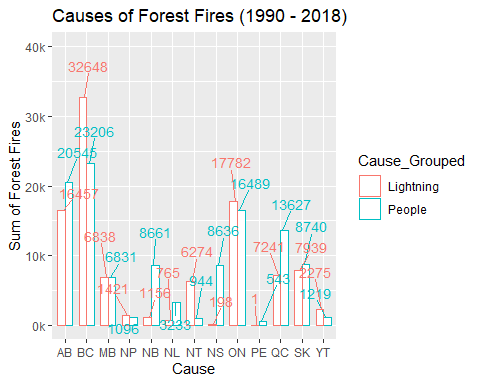
cor(my\_num\_data)

## Lightning People  
## Lightning 1.0000000 0.8506535  
## People 0.8506535 1.0000000

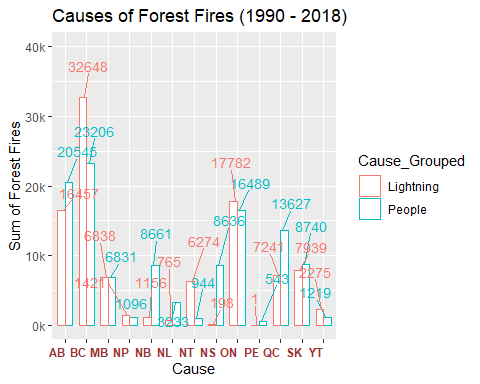
##################################################  
# Test 11  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
  
p <- ggplot(data = pivot3,  
 aes(x = Jurisdiction, y = sum\_Number, color=Cause\_Grouped)) +  
 geom\_bar(position = "dodge", stat="identity", width = 0.7, fill="white") +  
 geom\_text\_repel(aes(label=sum\_Number), show\_guide = F, position=position\_dodge(width=0.4),  
 vjust= -2.4, hjust = 0.4, size = 3.8, angle = 0)+  
 xlab("Cause") +  
 ylab("Causes of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 40000))+  
 labs(title = "Causes of Forest Fires (1990 - 2018)")

## Warning: `show\_guide` has been deprecated. Please use `show.legend`  
## instead.

p



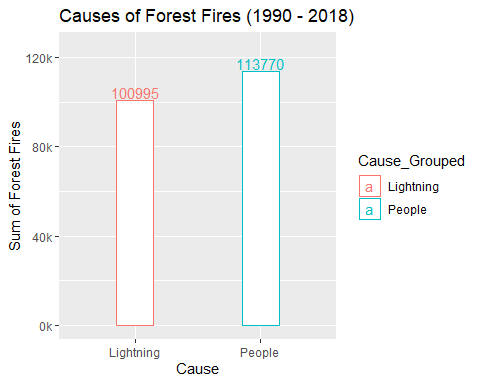
p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 1,size = 9, angle = 0))



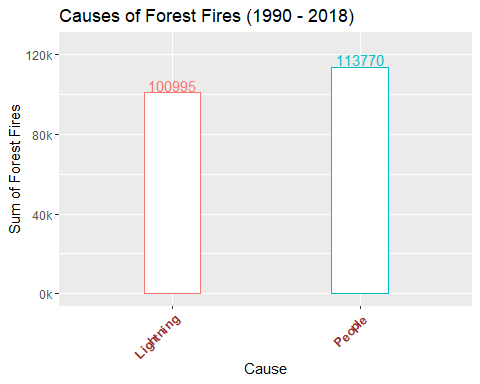
########################################################################################################################################  
#2  
#Below works with percentage totals  
GCause\_HvsL <- data %>% #Groups Cause together and sums Number  
 group\_by(Cause\_Grouped) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE)) %>%   
 mutate(rel.freq = paste0(round(100 \* sum\_Number/sum(sum\_Number),0),"%"))  
  
GCause\_HvsL

## # A tibble: 2 x 3  
## Cause\_Grouped sum\_Number rel.freq  
## <chr> <int> <chr>   
## 1 Lightning 100995 47%   
## 2 People 113770 53%

##################################################  
# Boxplot GCause 11, testing  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
p <- ggplot(data = GCause\_HvsL,  
 aes(x = Cause\_Grouped, y = sum\_Number, color=Cause\_Grouped)) +  
 geom\_bar(stat="identity", width = 0.3, fill="white") +  
 geom\_text(aes(label=sum\_Number), position=position\_dodge(width=0.9), vjust=-0.15)+  
 xlab("Cause") +  
 ylab("Causes of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 125000))+  
 labs(title = "Causes of Forest Fires (1990 - 2018)")  
  
p



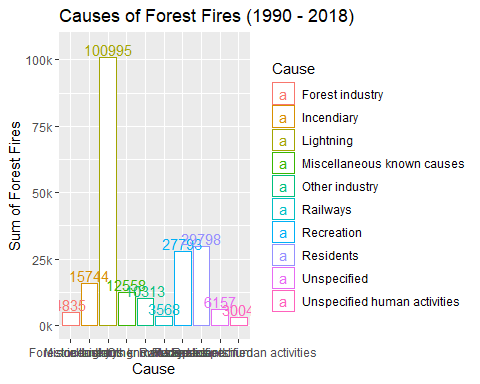
p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 1,size = 9, angle = 45), legend.position = "none")



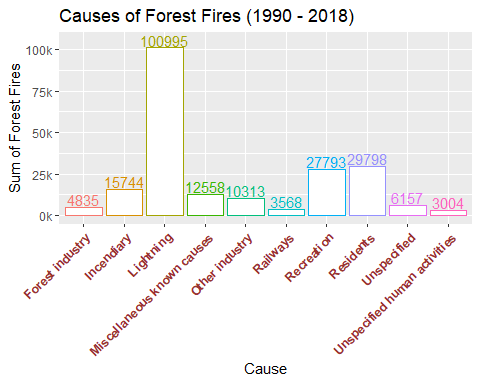
########################################################################################################################################  
#1  
#Below works with percentage totals  
GCause <- data %>% #Groups Cause together and sums Number  
 group\_by(Cause) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE)) %>%   
 mutate(rel.freq = paste0(round(100 \* sum\_Number/sum(sum\_Number),0),"%"))  
  
GCause

## # A tibble: 10 x 3  
## Cause sum\_Number rel.freq  
## <fct> <int> <chr>   
## 1 Forest industry 4835 2%   
## 2 Incendiary 15744 7%   
## 3 Lightning 100995 47%   
## 4 Miscellaneous known causes 12558 6%   
## 5 Other industry 10313 5%   
## 6 Railways 3568 2%   
## 7 Recreation 27793 13%   
## 8 Residents 29798 14%   
## 9 Unspecified 6157 3%   
## 10 Unspecified human activities 3004 1%

##################################################  
ks <- function (x) { number\_format(accuracy = 1,  
 scale = 1/1000,  
 suffix = "k",  
 big.mark = ",")(x) }  
  
p <- ggplot(data = GCause,  
 aes(x = Cause, y = sum\_Number, color=Cause)) +  
 geom\_bar(stat="identity", fill="white") +  
 geom\_text(aes(label=sum\_Number), position=position\_dodge(width=0.9), vjust=-0.15)+  
 xlab("Cause") +  
 ylab("Causes of Forest Fires (1990 - 2018)") +  
 scale\_y\_continuous(name="Sum of Forest Fires", labels = ks)+  
 coord\_cartesian(ylim = c(0, 105000))+  
 labs(title = "Causes of Forest Fires (1990 - 2018)")  
  
p



p + theme(  
 axis.text.x = element\_text(face = "bold", color = "#993333", hjust = 1,size = 9, angle = 45), legend.position = "none")



########################################################################################################################################  
#Sample code for various statistics  
data %>% #Groups Cause, Year together and sums Number  
 group\_by(Cause, Year) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 262 x 3  
## # Groups: Cause [10]  
## Cause Year sum\_Number  
## <fct> <int> <int>  
## 1 Forest industry 1990 355  
## 2 Forest industry 1991 299  
## 3 Forest industry 1992 258  
## 4 Forest industry 1993 240  
## 5 Forest industry 1994 222  
## 6 Forest industry 1995 309  
## 7 Forest industry 1996 282  
## 8 Forest industry 1997 276  
## 9 Forest industry 1998 151  
## 10 Forest industry 1999 239  
## # ... with 252 more rows

#Below code works  
data %>% #Groups Cause, Year together and sums Number  
 group\_by(Cause, Jurisdiction, Year) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 3,219 x 4  
## # Groups: Cause, Jurisdiction [129]  
## Cause Jurisdiction Year sum\_Number  
## <fct> <fct> <int> <int>  
## 1 Forest industry AB 1990 22  
## 2 Forest industry AB 1991 14  
## 3 Forest industry AB 1992 12  
## 4 Forest industry AB 1993 11  
## 5 Forest industry AB 1994 13  
## 6 Forest industry AB 1995 14  
## 7 Forest industry AB 1996 8  
## 8 Forest industry AB 1997 29  
## 9 Forest industry AB 1998 10  
## 10 Forest industry AB 1999 20  
## # ... with 3,209 more rows

#Below code works  
data %>% #Groups Cause, Year together and sums Number  
 group\_by(Jurisdiction, Year) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 372 x 3  
## # Groups: Jurisdiction [13]  
## Jurisdiction Year sum\_Number  
## <fct> <int> <int>  
## 1 AB 1990 1296  
## 2 AB 1991 923  
## 3 AB 1992 1055  
## 4 AB 1993 848  
## 5 AB 1994 872  
## 6 AB 1995 803  
## 7 AB 1996 376  
## 8 AB 1997 456  
## 9 AB 1998 1698  
## 10 AB 1999 1355  
## # ... with 362 more rows

#Below code works  
data %>% #Groups Cause, Year together and sums Number  
 group\_by(Year, Jurisdiction) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 372 x 3  
## # Groups: Year [29]  
## Year Jurisdiction sum\_Number  
## <int> <fct> <int>  
## 1 1990 AB 1296  
## 2 1990 BC 3255  
## 3 1990 MB 570  
## 4 1990 NP 128  
## 5 1990 NB 377  
## 6 1990 NL 197  
## 7 1990 NT 236  
## 8 1990 NS 498  
## 9 1990 ON 1614  
## 10 1990 PE 38  
## # ... with 362 more rows

#Below code works  
data %>% #Groups Cause, Year together and sums Number  
 group\_by(Jurisdiction) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 13 x 2  
## Jurisdiction sum\_Number  
## <fct> <int>  
## 1 AB 37002  
## 2 BC 55854  
## 3 MB 13669  
## 4 NP 2517  
## 5 NB 9817  
## 6 NL 3998  
## 7 NT 7218  
## 8 NS 8834  
## 9 ON 34271  
## 10 PE 544  
## 11 QC 20868  
## 12 SK 16679  
## 13 YT 3494

#Below code works  
data %>% #Groups Cause, Year together and sums Number  
 group\_by(Jurisdiction) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 13 x 2  
## Jurisdiction sum\_Number  
## <fct> <int>  
## 1 AB 37002  
## 2 BC 55854  
## 3 MB 13669  
## 4 NP 2517  
## 5 NB 9817  
## 6 NL 3998  
## 7 NT 7218  
## 8 NS 8834  
## 9 ON 34271  
## 10 PE 544  
## 11 QC 20868  
## 12 SK 16679  
## 13 YT 3494

#######################################################  
#Below code works  
data %>% #Groups Cause, Year, Jurisdiction and sums Number  
 group\_by(Cause, Year, Jurisdiction) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE)) %>%   
 filter(Jurisdiction == "BC") %>%  
 filter(Cause == "Lightning")

## # A tibble: 29 x 4  
## # Groups: Cause, Year [38]  
## Cause Year Jurisdiction sum\_Number  
## <fct> <int> <fct> <int>  
## 1 Lightning 1990 BC 2015  
## 2 Lightning 1991 BC 759  
## 3 Lightning 1992 BC 2344  
## 4 Lightning 1993 BC 609  
## 5 Lightning 1994 BC 2913  
## 6 Lightning 1995 BC 342  
## 7 Lightning 1996 BC 723  
## 8 Lightning 1997 BC 675  
## 9 Lightning 1998 BC 1773  
## 10 Lightning 1999 BC 585  
## # ... with 19 more rows

#Below code works  
data %>% #Groups Cause together and sums Number  
 group\_by(Cause, Year, Jurisdiction) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE)) %>%   
 filter(Jurisdiction == "AB") %>%  
 filter(Cause == "Lightning")

## # A tibble: 29 x 4  
## # Groups: Cause, Year [38]  
## Cause Year Jurisdiction sum\_Number  
## <fct> <int> <fct> <int>  
## 1 Lightning 1990 AB 914  
## 2 Lightning 1991 AB 466  
## 3 Lightning 1992 AB 626  
## 4 Lightning 1993 AB 517  
## 5 Lightning 1994 AB 499  
## 6 Lightning 1995 AB 357  
## 7 Lightning 1996 AB 217  
## 8 Lightning 1997 AB 242  
## 9 Lightning 1998 AB 1192  
## 10 Lightning 1999 AB 890  
## # ... with 19 more rows

#Below works great, show max number of fires and year for Alberta  
data %>% #Groups Cause together and sums Number  
 group\_by(Jurisdiction, Cause, Year) %>%   
 summarize(max\_Number = max(Number, na.rm = TRUE)) %>%   
 top\_n(n=1) %>%   
 filter(Jurisdiction == "AB") %>%  
 filter(Cause == "Lightning")

## Selecting by max\_Number

## # A tibble: 1 x 4  
## # Groups: Jurisdiction, Cause [130]  
## Jurisdiction Cause Year max\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Lightning 1998 1192

#Below works great, Top 10 years for Lightning in Alberta  
data %>% #Groups Cause together and sums Number  
 group\_by(Jurisdiction, Cause, Year) %>%   
 summarize(max\_Number = max(Number, na.rm = TRUE)) %>%   
 top\_n(n=10) %>%   
 filter(Jurisdiction == "AB") %>%  
 filter(Cause == "Lightning") %>%   
 arrange(desc(max\_Number))

## Selecting by max\_Number

## # A tibble: 10 x 4  
## # Groups: Jurisdiction, Cause [1]  
## Jurisdiction Cause Year max\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Lightning 1998 1192  
## 2 AB Lightning 1990 914  
## 3 AB Lightning 1999 890  
## 4 AB Lightning 2002 868  
## 5 AB Lightning 2008 779  
## 6 AB Lightning 2015 772  
## 7 AB Lightning 2006 746  
## 8 AB Lightning 2010 741  
## 9 AB Lightning 2004 732  
## 10 AB Lightning 1992 626

#Below works great, Top 10 years for Lightning in Alberta  
data %>% #Groups Cause together and sums Number  
 group\_by(Jurisdiction, Cause, Year) %>%   
 summarize(max\_Number = max(Number, na.rm = TRUE)) %>%   
 top\_n(n=10) %>%   
 filter(Jurisdiction == "AB") %>%  
 filter(Cause == "Lightning")

## Selecting by max\_Number

## # A tibble: 10 x 4  
## # Groups: Jurisdiction, Cause [130]  
## Jurisdiction Cause Year max\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Lightning 1990 914  
## 2 AB Lightning 1992 626  
## 3 AB Lightning 1998 1192  
## 4 AB Lightning 1999 890  
## 5 AB Lightning 2002 868  
## 6 AB Lightning 2004 732  
## 7 AB Lightning 2006 746  
## 8 AB Lightning 2008 779  
## 9 AB Lightning 2010 741  
## 10 AB Lightning 2015 772

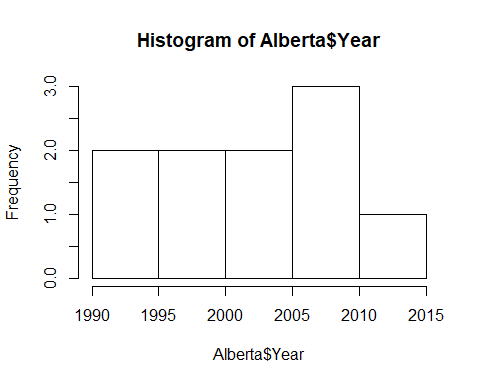
#Below works great, Top 10 years for Lightning in Alberta  
Alberta <- data %>% #Groups Cause together and sums Number  
 group\_by(Jurisdiction, Cause, Year) %>%   
 summarize(max\_Number = max(Number, na.rm = TRUE)) %>%   
 top\_n(n=10) %>%   
 filter(Jurisdiction == "AB") %>%  
 filter(Cause == "Lightning")

## Selecting by max\_Number

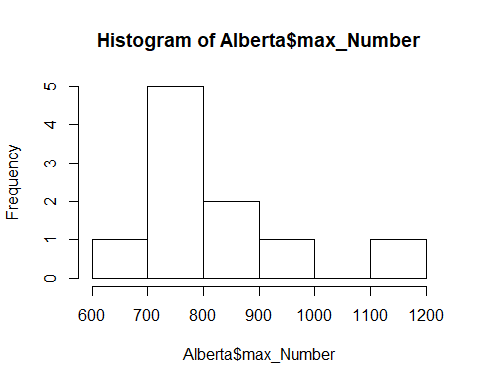
Alberta

## # A tibble: 10 x 4  
## # Groups: Jurisdiction, Cause [130]  
## Jurisdiction Cause Year max\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Lightning 1990 914  
## 2 AB Lightning 1992 626  
## 3 AB Lightning 1998 1192  
## 4 AB Lightning 1999 890  
## 5 AB Lightning 2002 868  
## 6 AB Lightning 2004 732  
## 7 AB Lightning 2006 746  
## 8 AB Lightning 2008 779  
## 9 AB Lightning 2010 741  
## 10 AB Lightning 2015 772

hist(Alberta$Year)



hist(Alberta$max\_Number)



#Below works  
data %>% #Groups Cause together and sums Number  
 group\_by(Jurisdiction, Cause, Year) %>%   
 summarize(max\_Number = max(Number, na.rm = TRUE)) %>%   
 filter(Jurisdiction == "AB") %>%  
 filter(Cause == "Lightning") %>%   
 filter(Year > "2003")

## # A tibble: 15 x 4  
## # Groups: Jurisdiction, Cause [130]  
## Jurisdiction Cause Year max\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Lightning 2004 732  
## 2 AB Lightning 2005 432  
## 3 AB Lightning 2006 746  
## 4 AB Lightning 2007 513  
## 5 AB Lightning 2008 779  
## 6 AB Lightning 2009 566  
## 7 AB Lightning 2010 741  
## 8 AB Lightning 2011 215  
## 9 AB Lightning 2012 436  
## 10 AB Lightning 2013 287  
## 11 AB Lightning 2014 563  
## 12 AB Lightning 2015 772  
## 13 AB Lightning 2016 514  
## 14 AB Lightning 2017 416  
## 15 AB Lightning 2018 511

#Below works, Top 10 after 2003  
data %>% #Groups Cause together and sums Number  
 group\_by(Jurisdiction, Cause, Year) %>%   
 summarize(max\_Number = max(Number, na.rm = TRUE)) %>%   
 filter(Jurisdiction == "AB") %>%  
 filter(Cause == "Lightning") %>%   
 filter(Year > "2003") %>%   
 arrange(desc(max\_Number))

## # A tibble: 15 x 4  
## # Groups: Jurisdiction, Cause [1]  
## Jurisdiction Cause Year max\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Lightning 2008 779  
## 2 AB Lightning 2015 772  
## 3 AB Lightning 2006 746  
## 4 AB Lightning 2010 741  
## 5 AB Lightning 2004 732  
## 6 AB Lightning 2009 566  
## 7 AB Lightning 2014 563  
## 8 AB Lightning 2016 514  
## 9 AB Lightning 2007 513  
## 10 AB Lightning 2018 511  
## 11 AB Lightning 2012 436  
## 12 AB Lightning 2005 432  
## 13 AB Lightning 2017 416  
## 14 AB Lightning 2013 287  
## 15 AB Lightning 2011 215

data %>%   
 group\_by(Jurisdiction, Cause, Year) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 3,219 x 4  
## # Groups: Jurisdiction, Cause [129]  
## Jurisdiction Cause Year sum\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Forest industry 1990 22  
## 2 AB Forest industry 1991 14  
## 3 AB Forest industry 1992 12  
## 4 AB Forest industry 1993 11  
## 5 AB Forest industry 1994 13  
## 6 AB Forest industry 1995 14  
## 7 AB Forest industry 1996 8  
## 8 AB Forest industry 1997 29  
## 9 AB Forest industry 1998 10  
## 10 AB Forest industry 1999 20  
## # ... with 3,209 more rows

#Below works, Top 10 after 2003  
data %>% #Groups Cause together and sums Number  
 group\_by(Jurisdiction, Cause, Year) %>%   
 summarize(max\_Number = max(Number, na.rm = TRUE)) %>%   
 filter(Jurisdiction == "AB") %>%  
 filter(Cause == "Lightning") %>%   
 filter(Year > "2003") %>%   
 arrange(desc(max\_Number))

## # A tibble: 15 x 4  
## # Groups: Jurisdiction, Cause [1]  
## Jurisdiction Cause Year max\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Lightning 2008 779  
## 2 AB Lightning 2015 772  
## 3 AB Lightning 2006 746  
## 4 AB Lightning 2010 741  
## 5 AB Lightning 2004 732  
## 6 AB Lightning 2009 566  
## 7 AB Lightning 2014 563  
## 8 AB Lightning 2016 514  
## 9 AB Lightning 2007 513  
## 10 AB Lightning 2018 511  
## 11 AB Lightning 2012 436  
## 12 AB Lightning 2005 432  
## 13 AB Lightning 2017 416  
## 14 AB Lightning 2013 287  
## 15 AB Lightning 2011 215

data %>% #Groups Cause together and sums Number. Also provides avg for Number of fires.  
 group\_by(Cause) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE),  
 count = n(),  
 average\_fire = mean(Number, na.rm = TRUE))

## # A tibble: 10 x 4  
## Cause sum\_Number count average\_fire  
## <fct> <int> <int> <dbl>  
## 1 Forest industry 4835 1262 3.83  
## 2 Incendiary 15744 1271 12.4   
## 3 Lightning 100995 1325 76.2   
## 4 Miscellaneous known causes 12558 1279 9.82  
## 5 Other industry 10313 1261 8.18  
## 6 Railways 3568 1265 2.82  
## 7 Recreation 27793 1274 21.8   
## 8 Residents 29798 1268 23.5   
## 9 Unspecified 6157 1290 4.77  
## 10 Unspecified human activities 3004 24 125.

data %>% #Groups Cause together and sums Number. Also provides avg for Number of fires.  
 group\_by(Cause) %>%   
 summarize(  
 sum\_Number = sum(Number, na.rm = TRUE),  
 count = n(),  
 average\_fire = mean(Number, na.rm = TRUE),  
 total = sum(sum\_Number)  
 )

## # A tibble: 10 x 5  
## Cause sum\_Number count average\_fire total  
## <fct> <int> <int> <dbl> <int>  
## 1 Forest industry 4835 1262 3.83 4835  
## 2 Incendiary 15744 1271 12.4 15744  
## 3 Lightning 100995 1325 76.2 100995  
## 4 Miscellaneous known causes 12558 1279 9.82 12558  
## 5 Other industry 10313 1261 8.18 10313  
## 6 Railways 3568 1265 2.82 3568  
## 7 Recreation 27793 1274 21.8 27793  
## 8 Residents 29798 1268 23.5 29798  
## 9 Unspecified 6157 1290 4.77 6157  
## 10 Unspecified human activities 3004 24 125. 3004

data %>% #Groups Year and sums Number  
 group\_by(Year) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 29 x 2  
## Year sum\_Number  
## <int> <int>  
## 1 1990 10111  
## 2 1991 10327  
## 3 1992 9068  
## 4 1993 6043  
## 5 1994 9763  
## 6 1995 8486  
## 7 1996 6349  
## 8 1997 6148  
## 9 1998 10723  
## 10 1999 7633  
## # ... with 19 more rows

data %>% #Groups Jurisdiction and sums Number  
 group\_by(Jurisdiction) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 13 x 2  
## Jurisdiction sum\_Number  
## <fct> <int>  
## 1 AB 37002  
## 2 BC 55854  
## 3 MB 13669  
## 4 NP 2517  
## 5 NB 9817  
## 6 NL 3998  
## 7 NT 7218  
## 8 NS 8834  
## 9 ON 34271  
## 10 PE 544  
## 11 QC 20868  
## 12 SK 16679  
## 13 YT 3494

data %>%   
 group\_by(Year) %>%  
 group\_by(Jurisdiction) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 13 x 2  
## Jurisdiction sum\_Number  
## <fct> <int>  
## 1 AB 37002  
## 2 BC 55854  
## 3 MB 13669  
## 4 NP 2517  
## 5 NB 9817  
## 6 NL 3998  
## 7 NT 7218  
## 8 NS 8834  
## 9 ON 34271  
## 10 PE 544  
## 11 QC 20868  
## 12 SK 16679  
## 13 YT 3494

data %>%   
 group\_by(Jurisdiction) %>%  
 group\_by(Year) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 29 x 2  
## Year sum\_Number  
## <int> <int>  
## 1 1990 10111  
## 2 1991 10327  
## 3 1992 9068  
## 4 1993 6043  
## 5 1994 9763  
## 6 1995 8486  
## 7 1996 6349  
## 8 1997 6148  
## 9 1998 10723  
## 10 1999 7633  
## # ... with 19 more rows

data %>%   
 group\_by(Jurisdiction) %>%  
 group\_by(Year) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 29 x 2  
## Year sum\_Number  
## <int> <int>  
## 1 1990 10111  
## 2 1991 10327  
## 3 1992 9068  
## 4 1993 6043  
## 5 1994 9763  
## 6 1995 8486  
## 7 1996 6349  
## 8 1997 6148  
## 9 1998 10723  
## 10 1999 7633  
## # ... with 19 more rows

data %>%   
 group\_by(Jurisdiction) %>%  
 group\_by(Year) %>%   
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 29 x 2  
## Year sum\_Number  
## <int> <int>  
## 1 1990 10111  
## 2 1991 10327  
## 3 1992 9068  
## 4 1993 6043  
## 5 1994 9763  
## 6 1995 8486  
## 7 1996 6349  
## 8 1997 6148  
## 9 1998 10723  
## 10 1999 7633  
## # ... with 19 more rows

data %>%   
 group\_by(Jurisdiction, Cause, Year) %>%  
 summarize(sum\_Number = sum(Number, na.rm = TRUE))

## # A tibble: 3,219 x 4  
## # Groups: Jurisdiction, Cause [129]  
## Jurisdiction Cause Year sum\_Number  
## <fct> <fct> <int> <int>  
## 1 AB Forest industry 1990 22  
## 2 AB Forest industry 1991 14  
## 3 AB Forest industry 1992 12  
## 4 AB Forest industry 1993 11  
## 5 AB Forest industry 1994 13  
## 6 AB Forest industry 1995 14  
## 7 AB Forest industry 1996 8  
## 8 AB Forest industry 1997 29  
## 9 AB Forest industry 1998 10  
## 10 AB Forest industry 1999 20  
## # ... with 3,209 more rows

data$Fire\_Cause\_Human = as.character(data$Cause)  
  
data$Fire\_Cause\_Human[ data$Fire\_Cause\_Human != "Lightning" & data$Fire\_Cause\_Human != "Unspecified" ] = "Human"  
# There are three types of fire causes, Lightning, Human and Unspecified.   
  
unique(data$Fire\_Cause\_Human)

## [1] "Human" "Lightning" "Unspecified"

#[1] "Human" "Lightning" "Unspecified"  
  
names(data)

## [1] "Cause" "Jurisdiction" "Number"   
## [4] "Protection.zone" "Response.category" "Year"   
## [7] "Juris\_Long" "Cause\_Grouped" "Time1"   
## [10] "Time2" "Region" "Fire\_Cause\_Human"

# [1] "Cause" "Jurisdiction" "Number" "Protection.zone" "Response.category" "Year" "Juris\_Long" "Cause\_Grouped" "Time1" "Time2"   
# [11] "Region" "Fire\_Cause\_Human"   
  
unique(data$Cause\_Grouped)

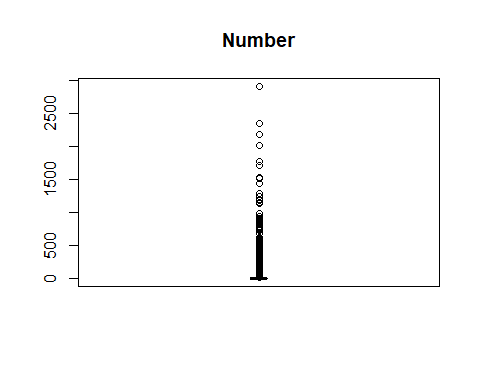
## [1] "People" "Lightning"

#[1] "People" "Lightning"  
  
unique(data$Year)

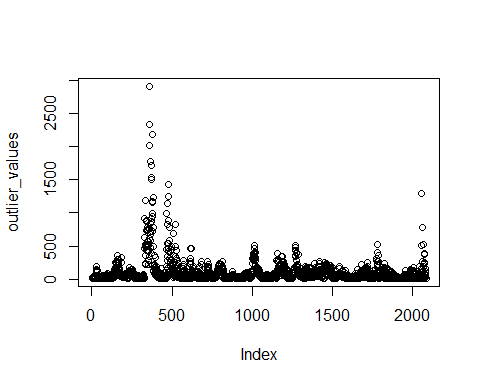
## [1] 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003  
## [15] 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017  
## [29] 2018

## Outlier detection

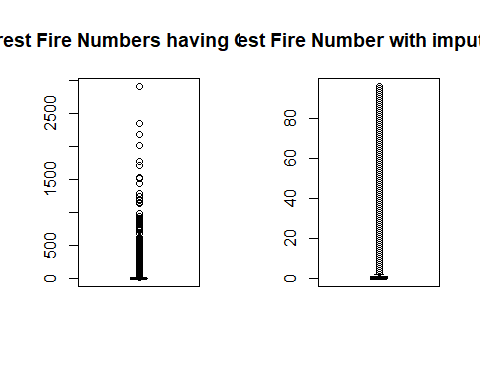
# Treating the outliers with mean/median imputation  
# We can handle outliers with mean or median imputation by replacing the observations lower than the 5th percentile with mean   
# and those higher than 95th percentile with median.   
# We can use the same statistics, mean or median, to impute outliers in both directions:  
  
# Outlier detection  
outlier\_values <- boxplot.stats(data$Number)$out  
boxplot(data$Number, main="Number", boxwex=0.1)



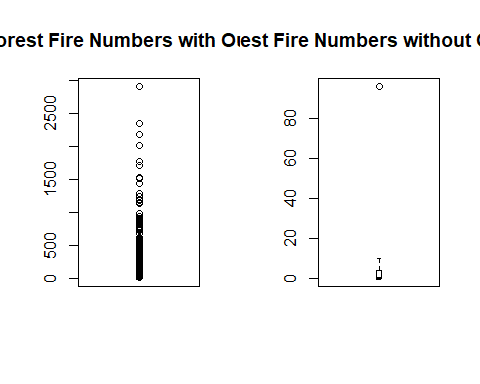
plot(outlier\_values)



impute\_outliers <- function(x,removeNA = TRUE){  
 quantiles <- quantile( x, c(.05, .95 ),na.rm = removeNA )  
 x[ x < quantiles[1] ] <- mean(x,na.rm = removeNA )  
 x[ x > quantiles[2] ] <- median(x,na.rm = removeNA )  
 x  
}  
  
imputed\_data <- impute\_outliers(data$Number)  
  
par(mfrow = c(1, 2))  
  
boxplot(data$Number, main="Forest Fire Numbers having Outliers", boxwex=0.3)  
  
boxplot(imputed\_data, main="Forest Fire Number with imputed data", boxwex=0.3)



########################################################################################################################################  
# Handling extreme values with capping  
# To handle extreme values that lie outside the 1.5 \* IQR(Inter Quartile Range) limits,   
# we could cap them by replacing those observations that lie below the lower limit,   
# with the value of 5th percentile and those that lie above the upper limit,   
# with the value of 95th percentile,   
  
replace\_outliers <- function(x, removeNA = TRUE) {  
 Number <- x  
 qnt <- quantile(Number, probs=c(.25, .75), na.rm = removeNA)  
 caps <- quantile(Number, probs=c(.05, .95), na.rm = removeNA)  
 H <- 1.5 \* IQR(Number, na.rm = removeNA)  
 Number[Number < (qnt[1] - H)] <- caps[1]  
 Number[Number > (qnt[2] + H)] <- caps[2]  
 Number  
 }  
  
capped\_Number <- replace\_outliers(data$Number)  
  
  
par(mfrow = c(1, 2))  
  
boxplot(data$Number, main="Forest Fire Numbers with Outliers", boxwex=0.1)  
  
boxplot(capped\_Number, main="Forest Fire Numbers without Outliers", boxwex=0.1)



## Comparing stats with “Number” of forest fires before and after reducing outliers

length(capped\_Number) # [1] 11519

## [1] 11519

max(capped\_Number) # [1] 96

## [1] 96

min(capped\_Number) # [1] 0

## [1] 0

mean(capped\_Number) # [1] 18.02405

## [1] 18.02405

median(capped\_Number) # [1] 0

## [1] 0

class(capped\_Number) # [1] "numeric"

## [1] "numeric"

summary(capped\_Number)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 0.00 0.00 18.02 4.00 96.00

# Min. 1st Qu. Median Mean 3rd Qu. Max.   
 # 0.00 0.00 0.00 18.02 4.00 96.00  
  
length(data$Number) # [1] 11519

## [1] 11519

max(data$Number) # [1] 2913

## [1] 2913

min(data$Number) # [1] 0

## [1] 0

mean(data$Number) # [1] 18.64441

## [1] 18.64441

median(data$Number) # [1] 0

## [1] 0

class(data$Number) # [1] "integer"

## [1] "integer"

summary(data$Number)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 0.00 0.00 18.64 4.00 2913.00

# Min. 1st Qu. Median Mean 3rd Qu. Max.   
 # 0.00 0.00 0.00 18.64 4.00 2913.00

## Preparing data frame for CAPPED OUTLIERS

outlier\_data\_frame <- data.frame(data)   
  
head(outlier\_data\_frame)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1 Forest industry AB 22 Intensive Full  
## 2 Forest industry AB 14 Intensive Full  
## 3 Forest industry AB 12 Intensive Full  
## 4 Forest industry AB 11 Intensive Full  
## 5 Forest industry AB 13 Intensive Full  
## 6 Forest industry AB 14 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1 1990 Alberta People Early 90s 1990s Prairie Region  
## 2 1991 Alberta People Early 90s 1990s Prairie Region  
## 3 1992 Alberta People Early 90s 1990s Prairie Region  
## 4 1993 Alberta People Early 90s 1990s Prairie Region  
## 5 1994 Alberta People Early 90s 1990s Prairie Region  
## 6 1995 Alberta People Early 90s 1990s Prairie Region  
## Fire\_Cause\_Human  
## 1 Human  
## 2 Human  
## 3 Human  
## 4 Human  
## 5 Human  
## 6 Human

tail(outlier\_data\_frame)

## Cause Jurisdiction Number Protection.zone  
## 11514 Unspecified human activities NS 0 Unspecified  
## 11515 Unspecified human activities ON 11 Unspecified  
## 11516 Unspecified human activities PE 0 Unspecified  
## 11517 Unspecified human activities QC 6 Unspecified  
## 11518 Unspecified human activities SK 13 Unspecified  
## 11519 Unspecified human activities YT 6 Unspecified  
## Response.category Year Juris\_Long Cause\_Grouped Time1  
## 11514 Modified 2018 Nova Scotia People Late 20s  
## 11515 Modified 2018 Ontario People Late 20s  
## 11516 Modified 2018 Prince Edward Island People Late 20s  
## 11517 Modified 2018 Quebec People Late 20s  
## 11518 Modified 2018 Saskatchewan People Late 20s  
## 11519 Modified 2018 Yukon People Late 20s  
## Time2 Region Fire\_Cause\_Human  
## 11514 2010s Atlantic Region Human  
## 11515 2010s Central Region Human  
## 11516 2010s Atlantic Region Human  
## 11517 2010s Central Region Human  
## 11518 2010s Prairie Region Human  
## 11519 2010s North Region Human

str(outlier\_data\_frame)

## 'data.frame': 11519 obs. of 12 variables:  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "AB","BC","MB",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Number : int 22 14 12 11 13 14 8 29 10 20 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Cause\_Grouped : chr "People" "People" "People" "People" ...  
## $ Time1 : chr "Early 90s" "Early 90s" "Early 90s" "Early 90s" ...  
## $ Time2 : chr "1990s" "1990s" "1990s" "1990s" ...  
## $ Region : chr "Prairie Region" "Prairie Region" "Prairie Region" "Prairie Region" ...  
## $ Fire\_Cause\_Human : chr "Human" "Human" "Human" "Human" ...

dim(outlier\_data\_frame) # 11519 rows, 12 columns

## [1] 11519 12

outlier\_data\_frame1 <- data.frame(data) %>%   
 dplyr::select(-3)  
  
outlier\_data\_frame2 <- data.frame(  
 outlier\_data\_frame1,  
 capped\_Number)   
# Reorder columns  
outlier\_data\_frame3 <- outlier\_data\_frame2[,c(1,2,12,3,4,5,6,7,8,9,10,11)]  
  
head(outlier\_data\_frame3)

## Cause Jurisdiction capped\_Number Protection.zone  
## 1 Forest industry AB 96 Intensive  
## 2 Forest industry AB 96 Intensive  
## 3 Forest industry AB 96 Intensive  
## 4 Forest industry AB 96 Intensive  
## 5 Forest industry AB 96 Intensive  
## 6 Forest industry AB 96 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 1 Full 1990 Alberta People Early 90s 1990s  
## 2 Full 1991 Alberta People Early 90s 1990s  
## 3 Full 1992 Alberta People Early 90s 1990s  
## 4 Full 1993 Alberta People Early 90s 1990s  
## 5 Full 1994 Alberta People Early 90s 1990s  
## 6 Full 1995 Alberta People Early 90s 1990s  
## Region Fire\_Cause\_Human  
## 1 Prairie Region Human  
## 2 Prairie Region Human  
## 3 Prairie Region Human  
## 4 Prairie Region Human  
## 5 Prairie Region Human  
## 6 Prairie Region Human

tail(outlier\_data\_frame3)

## Cause Jurisdiction capped\_Number  
## 11514 Unspecified human activities NS 0  
## 11515 Unspecified human activities ON 96  
## 11516 Unspecified human activities PE 0  
## 11517 Unspecified human activities QC 6  
## 11518 Unspecified human activities SK 96  
## 11519 Unspecified human activities YT 6  
## Protection.zone Response.category Year Juris\_Long  
## 11514 Unspecified Modified 2018 Nova Scotia  
## 11515 Unspecified Modified 2018 Ontario  
## 11516 Unspecified Modified 2018 Prince Edward Island  
## 11517 Unspecified Modified 2018 Quebec  
## 11518 Unspecified Modified 2018 Saskatchewan  
## 11519 Unspecified Modified 2018 Yukon  
## Cause\_Grouped Time1 Time2 Region Fire\_Cause\_Human  
## 11514 People Late 20s 2010s Atlantic Region Human  
## 11515 People Late 20s 2010s Central Region Human  
## 11516 People Late 20s 2010s Atlantic Region Human  
## 11517 People Late 20s 2010s Central Region Human  
## 11518 People Late 20s 2010s Prairie Region Human  
## 11519 People Late 20s 2010s North Region Human

str(outlier\_data\_frame3)

## 'data.frame': 11519 obs. of 12 variables:  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "AB","BC","MB",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ capped\_Number : num 96 96 96 96 96 96 8 96 10 96 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Cause\_Grouped : chr "People" "People" "People" "People" ...  
## $ Time1 : chr "Early 90s" "Early 90s" "Early 90s" "Early 90s" ...  
## $ Time2 : chr "1990s" "1990s" "1990s" "1990s" ...  
## $ Region : chr "Prairie Region" "Prairie Region" "Prairie Region" "Prairie Region" ...  
## $ Fire\_Cause\_Human : chr "Human" "Human" "Human" "Human" ...

dim(outlier\_data\_frame3) # 11519 rows, 12 columns

## [1] 11519 12

# outlier\_data\_frame3$capped\_Number <- as.integer (outlier\_data\_frame3$capped\_Number) #Change to factor  
# class(capped\_Number) # [1] "numeric"  
  
#data$Fire\_Cause\_Human <- as.factor(data$Fire\_Cause\_Human) #Change to factor

## Discretization test

#install.packages("dataPreparation")  
library(dataPreparation)

## Warning: package 'dataPreparation' was built under R version 3.5.3

## Loading required package: lubridate

## Warning: package 'lubridate' was built under R version 3.5.3

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':  
##   
## date

## Loading required package: stringr

## Warning: package 'stringr' was built under R version 3.5.3

##   
## Attaching package: 'stringr'

## The following object is masked from 'package:strucchange':  
##   
## boundary

## Loading required package: Matrix

##   
## Attaching package: 'Matrix'

## The following objects are masked from 'package:tidyr':  
##   
## expand, pack, unpack

## Loading required package: progress

## Warning: package 'progress' was built under R version 3.5.3

## dataPreparation 0.4.3

## Type dataPrepNews() to see new features/changes/bug fixes.

bins <- build\_bins(dataSet = data, cols = "Number", n\_bins = 10, type = "equal\_freq")

## [1] "fastDiscretization: I will build splits for 1 numeric columns using, equal\_freq method."  
## [1] "equal\_freq\_splits: Number can't provide 10 equal freq bins; instead you will have 5 bins."  
## [1] "fastDiscretization: it took me: 0s to build splits for 1 numeric columns."

print(bins)

## $Number  
## [1] -Inf 0 1 8 38 Inf

# $Number  
# [1] -Inf 0 1 8 38 Inf

## 17. Prepare packages for models

library(caret)  
library(lars)

## Loaded lars 1.2

library(elasticnet)

## Split the data into training and test set

## Set the seed to make your partition reproducible

## We want to make sure that the training set and the test set do not have any common data points.

## R built in function “Sample”" randomly selects samples

set.seed(1235)   
train\_index <- sample(1:nrow(data), 0.80 \* nrow(data))  
train.set <- data[train\_index,]  
test.set <- data[-train\_index,]

## 19.Set up model with K-fold cross-validation

## Defining the training controls for multiple models

train.control <- trainControl(method = "repeatedcv",   
 number = 3, repeats = 5)

## 20.Build the models with start time and end time for each model

## Comment out different models types for variable change.

#############################################################################################################################################  
# Model 1: lm model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
lm\_model <- train(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lm",  
 trControl = train.control)  
  
# lm\_model <- train(Number ~ Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lm",  
# trControl = train.control)  
#   
# lm\_model <- train(Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lm",  
# trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
lm\_model\_time.taken <- end.time - start.time  
#lm\_model\_time.taken  
#############################################################################################################################################  
# Model 2: glm model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
glm\_model <- train(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "glm",  
 trControl = train.control)  
  
# glm\_model <- train(Number ~ Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "glm",  
# trControl = train.control)  
#   
# glm\_model <- train(Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "glm",  
# trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
glm\_model\_time.taken <- end.time - start.time  
#glm\_model\_time.taken  
#############################################################################################################################################  
# Model 3: lasso model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
lasso\_Mod <- train(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lasso",  
 trControl = train.control)  
  
# lasso\_Mod <- train(Number ~ Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lasso",  
# trControl = train.control)  
#   
# lasso\_Mod <- train(Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lasso",  
# trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
lasso\_Mod\_time.taken <- end.time - start.time  
#lasso\_Mod\_time.taken  
#############################################################################################################################################  
# Model 4: knn model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
knn\_model <- train(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "knn",  
 trControl = train.control)  
  
# knn\_model <- train(Number ~ Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "knn",  
# trControl = train.control)  
#   
# knn\_model <- train(Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "knn",  
# trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
knn\_model\_time.taken <- end.time - start.time  
#knn\_model\_time.taken  
#############################################################################################################################################  
# Model 5: leapForward model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
LF\_model <- train(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "leapForward",  
 trControl = train.control)  
  
# LF\_model <- train(Number ~ Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "leapForward",  
# trControl = train.control)  
#   
# LF\_model <- train(Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "leapForward",  
# trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
LF\_model\_time.taken <- end.time - start.time  
#LF\_model\_time.taken  
#############################################################################################################################################  
# Model 6: leapBackward model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
LB\_model <- train(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "leapBackward",  
 trControl = train.control)  
  
# LB\_model <- train(Number ~ Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "leapBackward",  
# trControl = train.control)  
#   
# LB\_model <- train(Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "leapBackward",  
# trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
LB\_model\_time.taken <- end.time - start.time  
#LB\_model\_time.taken  
#############################################################################################################################################  
# Model 7: lmStepAIC model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
lmStepAIC\_Mod <- train(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lmStepAIC",  
 trControl = train.control)

## Start: AIC=53049.47  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 7873 34353677 53049  
## <none> 34345803 53049  
## - Response.categoryUnspecified 1 21595 34367398 53051  
## - Fire\_Cause\_HumanUnspecified 1 30214 34376017 53053  
## - Juris\_LongOntario 1 78498 34424301 53061  
## - Juris\_LongSaskatchewan 1 84654 34430457 53063  
## - `Juris\_LongNew Brunswick` 1 105214 34451017 53066  
## - Juris\_LongQuebec 1 122471 34468274 53069  
## - Protection.zoneLimited 1 129255 34475058 53071  
## - Protection.zoneUnspecified 1 157986 34503789 53076  
## - Juris\_LongManitoba 1 173576 34519379 53078  
## - `Juris\_LongNova Scotia` 1 178878 34524681 53079  
## - `Juris\_LongNorthwest Territories` 1 210584 34556387 53085  
## - `Juris\_LongBritish Columbia` 1 217307 34563110 53086  
## - Juris\_LongYukon 1 288330 34634134 53099  
## - `Juris\_LongNewfoundland and Labrador` 1 291362 34637165 53099  
## - `Juris\_LongNational parks` 1 292935 34638738 53100  
## - `Juris\_LongPrince Edward Island` 1 344269 34690072 53109  
## - Response.categoryModified 1 1583931 35929735 53324  
## - Response.categoryNone 1 1619108 35964912 53330  
## - Fire\_Cause\_HumanLightning 1 1902141 36247945 53379  
##   
## Step: AIC=53048.88  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 34353677 53049  
## - Response.categoryUnspecified 1 26583 34380260 53052  
## - Fire\_Cause\_HumanUnspecified 1 29699 34383376 53052  
## - Juris\_LongOntario 1 76315 34429991 53061  
## - Juris\_LongSaskatchewan 1 86556 34440233 53062  
## - `Juris\_LongNew Brunswick` 1 105096 34458772 53066  
## - Juris\_LongQuebec 1 121568 34475245 53069  
## - Protection.zoneLimited 1 132962 34486638 53071  
## - Protection.zoneUnspecified 1 152173 34505850 53074  
## - Juris\_LongManitoba 1 178013 34531690 53079  
## - `Juris\_LongNova Scotia` 1 179273 34532949 53079  
## - `Juris\_LongNorthwest Territories` 1 213980 34567656 53085  
## - `Juris\_LongBritish Columbia` 1 221162 34574839 53086  
## - `Juris\_LongNewfoundland and Labrador` 1 289292 34642968 53098  
## - Juris\_LongYukon 1 293849 34647526 53099  
## - `Juris\_LongNational parks` 1 296013 34649690 53100  
## - `Juris\_LongPrince Edward Island` 1 338802 34692479 53107  
## - Response.categoryModified 1 1581304 35934981 53323  
## - Response.categoryNone 1 1612306 35965983 53329  
## - Fire\_Cause\_HumanLightning 1 1904544 36258221 53378  
## Start: AIC=53443.52  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 3980 36579347 53442  
## <none> 36575367 53444  
## - Response.categoryUnspecified 1 17517 36592884 53444  
## - Fire\_Cause\_HumanUnspecified 1 39763 36615130 53448  
## - Protection.zoneUnspecified 1 52881 36628249 53450  
## - Juris\_LongOntario 1 91975 36667342 53457  
## - Juris\_LongSaskatchewan 1 129413 36704780 53463  
## - `Juris\_LongBritish Columbia` 1 169027 36744394 53470  
## - Juris\_LongQuebec 1 196891 36772258 53475  
## - Protection.zoneLimited 1 207869 36783237 53476  
## - `Juris\_LongNew Brunswick` 1 212977 36788344 53477  
## - Juris\_LongManitoba 1 270617 36845985 53487  
## - `Juris\_LongNova Scotia` 1 300416 36875783 53492  
## - `Juris\_LongNorthwest Territories` 1 334795 36910163 53498  
## - Juris\_LongYukon 1 376960 36952327 53505  
## - `Juris\_LongNational parks` 1 377133 36952500 53505  
## - `Juris\_LongNewfoundland and Labrador` 1 381429 36956796 53505  
## - `Juris\_LongPrince Edward Island` 1 458529 37033896 53518  
## - Response.categoryModified 1 1721211 38296579 53724  
## - Response.categoryNone 1 1835769 38411136 53742  
## - Fire\_Cause\_HumanLightning 1 2299217 38874584 53816  
##   
## Step: AIC=53442.19  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 36579347 53442  
## - Response.categoryUnspecified 1 20358 36599705 53444  
## - Fire\_Cause\_HumanUnspecified 1 39501 36618848 53447  
## - Protection.zoneUnspecified 1 50172 36629519 53449  
## - Juris\_LongOntario 1 90364 36669712 53455  
## - Juris\_LongSaskatchewan 1 130382 36709729 53462  
## - `Juris\_LongBritish Columbia` 1 171458 36750806 53469  
## - Juris\_LongQuebec 1 196325 36775673 53473  
## - Protection.zoneLimited 1 211993 36791340 53476  
## - `Juris\_LongNew Brunswick` 1 214379 36793726 53476  
## - Juris\_LongManitoba 1 274173 36853520 53486  
## - `Juris\_LongNova Scotia` 1 300473 36879820 53490  
## - `Juris\_LongNorthwest Territories` 1 336899 36916246 53497  
## - `Juris\_LongNational parks` 1 379141 36958489 53504  
## - `Juris\_LongNewfoundland and Labrador` 1 380714 36960061 53504  
## - Juris\_LongYukon 1 381460 36960807 53504  
## - `Juris\_LongPrince Edward Island` 1 455589 37034936 53516  
## - Response.categoryModified 1 1721115 38300462 53723  
## - Response.categoryNone 1 1831924 38411272 53740  
## - Fire\_Cause\_HumanLightning 1 2302177 38881525 53815  
## Start: AIC=55333.26  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Response.categoryUnspecified 1 9948 49821725 55332  
## - Year 1 13415 49825192 55333  
## <none> 49811777 55333  
## - Fire\_Cause\_HumanUnspecified 1 30025 49841802 55335  
## - Protection.zoneUnspecified 1 95398 49907175 55343  
## - Protection.zoneLimited 1 139109 49950885 55348  
## - Juris\_LongOntario 1 149003 49960780 55350  
## - Juris\_LongSaskatchewan 1 159745 49971521 55351  
## - `Juris\_LongNew Brunswick` 1 246046 50057823 55362  
## - Juris\_LongQuebec 1 275703 50087480 55365  
## - `Juris\_LongNova Scotia` 1 303162 50114938 55369  
## - `Juris\_LongBritish Columbia` 1 346562 50158339 55374  
## - Juris\_LongManitoba 1 352679 50164456 55375  
## - `Juris\_LongNorthwest Territories` 1 394759 50206535 55380  
## - `Juris\_LongNational parks` 1 423900 50235677 55383  
## - `Juris\_LongPrince Edward Island` 1 451481 50263258 55387  
## - Juris\_LongYukon 1 484192 50295969 55391  
## - `Juris\_LongNewfoundland and Labrador` 1 493711 50305488 55392  
## - Response.categoryModified 1 2048637 51860414 55579  
## - Response.categoryNone 1 2052050 51863826 55579  
## - Fire\_Cause\_HumanLightning 1 3473224 53285000 55745  
##   
## Step: AIC=55332.49  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone  
##   
## Df Sum of Sq RSS AIC  
## <none> 49821725 55332  
## - Year 1 17671 49839396 55333  
## - Fire\_Cause\_HumanUnspecified 1 30319 49852044 55334  
## - Protection.zoneUnspecified 1 96209 49917934 55342  
## - Protection.zoneLimited 1 143276 49965001 55348  
## - Juris\_LongOntario 1 148350 49970075 55349  
## - Juris\_LongSaskatchewan 1 152494 49974219 55349  
## - `Juris\_LongNew Brunswick` 1 248084 50069809 55361  
## - Juris\_LongQuebec 1 274055 50095780 55364  
## - `Juris\_LongNova Scotia` 1 302946 50124671 55368  
## - Juris\_LongManitoba 1 344049 50165774 55373  
## - `Juris\_LongBritish Columbia` 1 345663 50167388 55373  
## - `Juris\_LongNorthwest Territories` 1 388991 50210716 55378  
## - `Juris\_LongNational parks` 1 423202 50244927 55382  
## - `Juris\_LongPrince Edward Island` 1 453425 50275150 55386  
## - Juris\_LongYukon 1 481544 50303269 55390  
## - `Juris\_LongNewfoundland and Labrador` 1 491370 50313095 55391  
## - Response.categoryModified 1 2107677 51929402 55585  
## - Response.categoryNone 1 2124606 51946331 55587  
## - Fire\_Cause\_HumanLightning 1 3470983 53292708 55744  
## Start: AIC=54557.16  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 43843722 54557  
## - Response.categoryUnspecified 1 20909 43864631 54558  
## - Year 1 22157 43865879 54558  
## - Fire\_Cause\_HumanUnspecified 1 33352 43877074 54560  
## - Juris\_LongOntario 1 80093 43923815 54566  
## - Juris\_LongSaskatchewan 1 108366 43952089 54570  
## - Protection.zoneLimited 1 142404 43986126 54575  
## - Protection.zoneUnspecified 1 173871 44017593 54579  
## - Juris\_LongQuebec 1 179406 44023128 54580  
## - `Juris\_LongNew Brunswick` 1 182369 44026092 54581  
## - Juris\_LongManitoba 1 236454 44080176 54588  
## - `Juris\_LongNova Scotia` 1 247267 44090989 54590  
## - `Juris\_LongNorthwest Territories` 1 296306 44140028 54597  
## - `Juris\_LongBritish Columbia` 1 328382 44172105 54601  
## - `Juris\_LongNational parks` 1 328853 44172575 54601  
## - `Juris\_LongNewfoundland and Labrador` 1 358695 44202417 54605  
## - Juris\_LongYukon 1 364105 44207827 54606  
## - `Juris\_LongPrince Edward Island` 1 393315 44237037 54610  
## - Response.categoryModified 1 1816137 45659859 54805  
## - Response.categoryNone 1 1892605 45736327 54815  
## - Fire\_Cause\_HumanLightning 1 2693351 46537073 54921  
## Start: AIC=54129.34  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 649 40947288 54127  
## <none> 40946639 54129  
## - Protection.zoneUnspecified 1 15764 40962403 54130  
## - Response.categoryUnspecified 1 24032 40970671 54131  
## - Fire\_Cause\_HumanUnspecified 1 29096 40975735 54132  
## - Juris\_LongOntario 1 97882 41044520 54142  
## - Juris\_LongSaskatchewan 1 126510 41073149 54146  
## - Juris\_LongQuebec 1 152154 41098793 54150  
## - Protection.zoneLimited 1 193577 41140216 54156  
## - `Juris\_LongNew Brunswick` 1 194144 41140782 54156  
## - `Juris\_LongNova Scotia` 1 233896 41180535 54162  
## - Juris\_LongManitoba 1 248454 41195093 54165  
## - `Juris\_LongNorthwest Territories` 1 273481 41220119 54168  
## - `Juris\_LongBritish Columbia` 1 329871 41276510 54177  
## - Juris\_LongYukon 1 344655 41291294 54179  
## - `Juris\_LongNewfoundland and Labrador` 1 353044 41299683 54180  
## - `Juris\_LongNational parks` 1 355780 41302419 54180  
## - `Juris\_LongPrince Edward Island` 1 361960 41308599 54181  
## - Response.categoryModified 1 1739751 42686390 54383  
## - Response.categoryNone 1 1802332 42748971 54392  
## - Fire\_Cause\_HumanLightning 1 2614442 43561081 54508  
##   
## Step: AIC=54127.44  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 40947288 54127  
## - Protection.zoneUnspecified 1 15241 40962528 54128  
## - Response.categoryUnspecified 1 25799 40973086 54129  
## - Fire\_Cause\_HumanUnspecified 1 29058 40976346 54130  
## - Juris\_LongOntario 1 97410 41044697 54140  
## - Juris\_LongSaskatchewan 1 127045 41074333 54144  
## - Juris\_LongQuebec 1 152044 41099331 54148  
## - `Juris\_LongNew Brunswick` 1 194380 41141668 54155  
## - Protection.zoneLimited 1 195216 41142503 54155  
## - `Juris\_LongNova Scotia` 1 234033 41181320 54160  
## - Juris\_LongManitoba 1 250297 41197585 54163  
## - `Juris\_LongNorthwest Territories` 1 274951 41222239 54167  
## - `Juris\_LongBritish Columbia` 1 331618 41278905 54175  
## - Juris\_LongYukon 1 347062 41294349 54177  
## - `Juris\_LongNewfoundland and Labrador` 1 353019 41300306 54178  
## - `Juris\_LongNational parks` 1 357683 41304971 54179  
## - `Juris\_LongPrince Edward Island` 1 361313 41308601 54179  
## - Response.categoryModified 1 1739363 42686651 54381  
## - Response.categoryNone 1 1803625 42750913 54390  
## - Fire\_Cause\_HumanLightning 1 2614348 43561635 54506  
## Start: AIC=53357.26  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Response.categoryUnspecified 1 5934 36116468 53356  
## - Year 1 8560 36119094 53357  
## <none> 36110534 53357  
## - Fire\_Cause\_HumanUnspecified 1 34842 36145376 53361  
## - `Juris\_LongBritish Columbia` 1 100314 36210849 53372  
## - Juris\_LongSaskatchewan 1 134422 36244956 53378  
## - Protection.zoneLimited 1 136512 36247046 53378  
## - Juris\_LongOntario 1 145867 36256401 53380  
## - Protection.zoneUnspecified 1 148364 36258899 53380  
## - `Juris\_LongNew Brunswick` 1 187373 36297908 53387  
## - Juris\_LongQuebec 1 259493 36370027 53399  
## - `Juris\_LongNova Scotia` 1 303683 36414217 53407  
## - Juris\_LongManitoba 1 304469 36415003 53407  
## - `Juris\_LongNorthwest Territories` 1 366560 36477094 53417  
## - `Juris\_LongNational parks` 1 414511 36525046 53425  
## - Juris\_LongYukon 1 439531 36550065 53430  
## - `Juris\_LongNewfoundland and Labrador` 1 450618 36561152 53431  
## - `Juris\_LongPrince Edward Island` 1 512978 36623512 53442  
## - Response.categoryModified 1 1786740 37897274 53652  
## - Response.categoryNone 1 1804104 37914639 53655  
## - Fire\_Cause\_HumanLightning 1 2267260 38377795 53729  
##   
## Step: AIC=53356.27  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 11214 36127683 53356  
## <none> 36116468 53356  
## - Fire\_Cause\_HumanUnspecified 1 35149 36151618 53360  
## - `Juris\_LongBritish Columbia` 1 99902 36216371 53371  
## - Juris\_LongSaskatchewan 1 129149 36245618 53376  
## - Protection.zoneLimited 1 139667 36256136 53378  
## - Juris\_LongOntario 1 145607 36262076 53379  
## - Protection.zoneUnspecified 1 148865 36265333 53380  
## - `Juris\_LongNew Brunswick` 1 188593 36305061 53386  
## - Juris\_LongQuebec 1 258741 36375210 53398  
## - Juris\_LongManitoba 1 299165 36415633 53405  
## - `Juris\_LongNova Scotia` 1 303680 36420148 53406  
## - `Juris\_LongNorthwest Territories` 1 363024 36479493 53416  
## - `Juris\_LongNational parks` 1 414360 36530828 53424  
## - Juris\_LongYukon 1 437897 36554365 53428  
## - `Juris\_LongNewfoundland and Labrador` 1 449580 36566048 53430  
## - `Juris\_LongPrince Edward Island` 1 515158 36631626 53441  
## - Response.categoryModified 1 1832516 37948985 53658  
## - Response.categoryNone 1 1858883 37975352 53663  
## - Fire\_Cause\_HumanLightning 1 2264615 38381083 53728  
##   
## Step: AIC=53356.18  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone  
##   
## Df Sum of Sq RSS AIC  
## <none> 36127683 53356  
## - Fire\_Cause\_HumanUnspecified 1 34831 36162514 53360  
## - `Juris\_LongBritish Columbia` 1 103302 36230985 53372  
## - Juris\_LongSaskatchewan 1 128757 36256440 53376  
## - Protection.zoneUnspecified 1 141286 36268969 53378  
## - Juris\_LongOntario 1 142117 36269800 53378  
## - Protection.zoneLimited 1 143581 36271263 53379  
## - `Juris\_LongNew Brunswick` 1 191150 36318832 53387  
## - Juris\_LongQuebec 1 256213 36383895 53398  
## - Juris\_LongManitoba 1 303293 36430976 53406  
## - `Juris\_LongNova Scotia` 1 303368 36431050 53406  
## - `Juris\_LongNorthwest Territories` 1 365180 36492862 53416  
## - `Juris\_LongNational parks` 1 416452 36544135 53425  
## - Juris\_LongYukon 1 445759 36573441 53430  
## - `Juris\_LongNewfoundland and Labrador` 1 446291 36573974 53430  
## - `Juris\_LongPrince Edward Island` 1 508213 36635895 53440  
## - Response.categoryModified 1 1834053 37961736 53658  
## - Response.categoryNone 1 1854232 37981914 53662  
## - Fire\_Cause\_HumanLightning 1 2268008 38395691 53728  
## Start: AIC=53748.08  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 5469 38488008 53747  
## <none> 38482539 53748  
## - Response.categoryUnspecified 1 22243 38504782 53750  
## - Fire\_Cause\_HumanUnspecified 1 33075 38515614 53751  
## - Juris\_LongSaskatchewan 1 93967 38576506 53761  
## - Protection.zoneUnspecified 1 125155 38607694 53766  
## - Juris\_LongOntario 1 131262 38613801 53767  
## - Protection.zoneLimited 1 134877 38617416 53768  
## - `Juris\_LongNew Brunswick` 1 187029 38669568 53776  
## - Juris\_LongQuebec 1 201525 38684064 53778  
## - `Juris\_LongNova Scotia` 1 272217 38754756 53789  
## - Juris\_LongManitoba 1 274068 38756607 53790  
## - `Juris\_LongBritish Columbia` 1 298690 38781229 53794  
## - `Juris\_LongNorthwest Territories` 1 307952 38790491 53795  
## - `Juris\_LongNational parks` 1 347869 38830408 53801  
## - `Juris\_LongNewfoundland and Labrador` 1 376457 38858996 53806  
## - Juris\_LongYukon 1 380333 38862872 53806  
## - `Juris\_LongPrince Edward Island` 1 435201 38917740 53815  
## - Response.categoryNone 1 1714144 40196683 54014  
## - Response.categoryModified 1 1752530 40235069 54020  
## - Fire\_Cause\_HumanLightning 1 2198980 40681519 54087  
##   
## Step: AIC=53746.95  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 38488008 53747  
## - Response.categoryUnspecified 1 26486 38514495 53749  
## - Fire\_Cause\_HumanUnspecified 1 33138 38521146 53750  
## - Juris\_LongSaskatchewan 1 95839 38583848 53760  
## - Protection.zoneUnspecified 1 120610 38608618 53764  
## - Juris\_LongOntario 1 129538 38617546 53766  
## - Protection.zoneLimited 1 137189 38625198 53767  
## - `Juris\_LongNew Brunswick` 1 187702 38675710 53775  
## - Juris\_LongQuebec 1 201331 38689340 53777  
## - `Juris\_LongNova Scotia` 1 273263 38761271 53788  
## - Juris\_LongManitoba 1 279525 38767534 53789  
## - `Juris\_LongBritish Columbia` 1 301801 38789809 53793  
## - `Juris\_LongNorthwest Territories` 1 312808 38800817 53795  
## - `Juris\_LongNational parks` 1 351492 38839500 53801  
## - `Juris\_LongNewfoundland and Labrador` 1 376510 38864518 53805  
## - Juris\_LongYukon 1 387502 38875510 53806  
## - `Juris\_LongPrince Edward Island` 1 431411 38919420 53813  
## - Response.categoryNone 1 1708874 40196883 54012  
## - Response.categoryModified 1 1750484 40238493 54018  
## - Fire\_Cause\_HumanLightning 1 2202865 40690874 54087  
## Start: AIC=54929.98  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 3399 46650122 54928  
## <none> 46646723 54930  
## - Protection.zoneUnspecified 1 24055 46670778 54931  
## - Response.categoryUnspecified 1 40744 46687467 54933  
## - Fire\_Cause\_HumanUnspecified 1 40898 46687621 54933  
## - Juris\_LongOntario 1 121170 46767893 54944  
## - Protection.zoneLimited 1 155528 46802251 54948  
## - Juris\_LongSaskatchewan 1 158299 46805022 54949  
## - `Juris\_LongNew Brunswick` 1 209060 46855783 54955  
## - Juris\_LongQuebec 1 231696 46878419 54958  
## - `Juris\_LongNova Scotia` 1 276653 46923377 54964  
## - Juris\_LongManitoba 1 303061 46949784 54968  
## - `Juris\_LongBritish Columbia` 1 326567 46973290 54971  
## - `Juris\_LongNorthwest Territories` 1 351696 46998419 54974  
## - `Juris\_LongNational parks` 1 397153 47043876 54980  
## - `Juris\_LongPrince Edward Island` 1 437753 47084476 54985  
## - `Juris\_LongNewfoundland and Labrador` 1 442147 47088870 54986  
## - Juris\_LongYukon 1 453595 47100318 54987  
## - Response.categoryModified 1 1955087 48601810 55180  
## - Response.categoryNone 1 2031087 48677810 55190  
## - Fire\_Cause\_HumanLightning 1 3206707 49853430 55336  
##   
## Step: AIC=54928.43  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 46650122 54928  
## - Protection.zoneUnspecified 1 22416 46672539 54929  
## - Fire\_Cause\_HumanUnspecified 1 40645 46690767 54932  
## - Response.categoryUnspecified 1 45011 46695133 54932  
## - Juris\_LongOntario 1 119643 46769765 54942  
## - Juris\_LongSaskatchewan 1 158369 46808492 54947  
## - Protection.zoneLimited 1 158587 46808710 54947  
## - `Juris\_LongNew Brunswick` 1 209430 46859552 54954  
## - Juris\_LongQuebec 1 230659 46880781 54957  
## - `Juris\_LongNova Scotia` 1 276132 46926254 54963  
## - Juris\_LongManitoba 1 306429 46956551 54967  
## - `Juris\_LongBritish Columbia` 1 329572 46979695 54970  
## - `Juris\_LongNorthwest Territories` 1 353800 47003923 54973  
## - `Juris\_LongNational parks` 1 399495 47049617 54979  
## - `Juris\_LongPrince Edward Island` 1 434717 47084839 54983  
## - `Juris\_LongNewfoundland and Labrador` 1 440929 47091052 54984  
## - Juris\_LongYukon 1 457770 47107892 54986  
## - Response.categoryModified 1 1953326 48603449 55178  
## - Response.categoryNone 1 2027703 48677825 55188  
## - Fire\_Cause\_HumanLightning 1 3205016 49855138 55335  
## Start: AIC=53290.57  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Response.categoryUnspecified 1 739 35676793 53289  
## <none> 35676054 53291  
## - Year 1 19871 35695924 53292  
## - Fire\_Cause\_HumanUnspecified 1 26893 35702947 53293  
## - Juris\_LongOntario 1 71534 35747587 53301  
## - Juris\_LongSaskatchewan 1 117296 35793349 53309  
## - `Juris\_LongBritish Columbia` 1 123668 35799722 53310  
## - Juris\_LongQuebec 1 153349 35829403 53315  
## - `Juris\_LongNew Brunswick` 1 157954 35834007 53316  
## - Protection.zoneUnspecified 1 180090 35856144 53320  
## - Protection.zoneLimited 1 180422 35856476 53320  
## - Juris\_LongManitoba 1 208858 35884912 53324  
## - `Juris\_LongNova Scotia` 1 232594 35908648 53328  
## - `Juris\_LongNorthwest Territories` 1 270025 35946078 53335  
## - Juris\_LongYukon 1 311985 35988039 53342  
## - `Juris\_LongNewfoundland and Labrador` 1 340986 36017039 53347  
## - `Juris\_LongNational parks` 1 348338 36024391 53348  
## - `Juris\_LongPrince Edward Island` 1 395599 36071653 53356  
## - Response.categoryModified 1 1647312 37323365 53566  
## - Response.categoryNone 1 1763685 37439739 53585  
## - Fire\_Cause\_HumanLightning 1 2195946 37872000 53656  
##   
## Step: AIC=53288.69  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone  
##   
## Df Sum of Sq RSS AIC  
## <none> 35676793 53289  
## - Year 1 21625 35698418 53290  
## - Fire\_Cause\_HumanUnspecified 1 27020 35703813 53291  
## - Juris\_LongOntario 1 71483 35748276 53299  
## - Juris\_LongSaskatchewan 1 116924 35793717 53307  
## - `Juris\_LongBritish Columbia` 1 123496 35800289 53308  
## - Juris\_LongQuebec 1 153119 35829912 53313  
## - `Juris\_LongNew Brunswick` 1 158433 35835226 53314  
## - Protection.zoneUnspecified 1 180413 35857206 53318  
## - Protection.zoneLimited 1 181690 35858483 53318  
## - Juris\_LongManitoba 1 208297 35885090 53322  
## - `Juris\_LongNova Scotia` 1 232611 35909404 53327  
## - `Juris\_LongNorthwest Territories` 1 269323 35946116 53333  
## - Juris\_LongYukon 1 311561 35988354 53340  
## - `Juris\_LongNewfoundland and Labrador` 1 340702 36017495 53345  
## - `Juris\_LongNational parks` 1 348313 36025106 53346  
## - `Juris\_LongPrince Edward Island` 1 396259 36073052 53355  
## - Response.categoryModified 1 1674787 37351580 53569  
## - Response.categoryNone 1 1802052 37478845 53589  
## - Fire\_Cause\_HumanLightning 1 2195263 37872056 53654  
## Start: AIC=52839.63  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 254 33192637 52838  
## - Protection.zoneUnspecified 1 828 33193211 52838  
## <none> 33192383 52840  
## - Response.categoryUnspecified 1 11946 33204329 52840  
## - Fire\_Cause\_HumanUnspecified 1 35443 33227825 52844  
## - Juris\_LongOntario 1 70789 33263172 52851  
## - Juris\_LongSaskatchewan 1 110050 33302433 52858  
## - `Juris\_LongNew Brunswick` 1 158795 33351178 52867  
## - Juris\_LongQuebec 1 173034 33365417 52870  
## - Protection.zoneLimited 1 175362 33367745 52870  
## - `Juris\_LongBritish Columbia` 1 218244 33410627 52878  
## - Juris\_LongManitoba 1 231290 33423673 52880  
## - `Juris\_LongNova Scotia` 1 249538 33441921 52884  
## - `Juris\_LongNorthwest Territories` 1 260185 33452568 52886  
## - `Juris\_LongNewfoundland and Labrador` 1 330203 33522586 52898  
## - `Juris\_LongNational parks` 1 344663 33537046 52901  
## - `Juris\_LongPrince Edward Island` 1 350157 33542540 52902  
## - Juris\_LongYukon 1 359571 33551954 52904  
## - Response.categoryModified 1 1602605 34794988 53127  
## - Response.categoryNone 1 1738814 34931197 53151  
## - Fire\_Cause\_HumanLightning 1 2499749 35692132 53284  
##   
## Step: AIC=52837.67  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Protection.zoneUnspecified 1 961 33193598 52836  
## <none> 33192637 52838  
## - Response.categoryUnspecified 1 11692 33204329 52838  
## - Fire\_Cause\_HumanUnspecified 1 35518 33228154 52842  
## - Juris\_LongOntario 1 71346 33263983 52849  
## - Juris\_LongSaskatchewan 1 109939 33302576 52856  
## - `Juris\_LongNew Brunswick` 1 158667 33351304 52865  
## - Juris\_LongQuebec 1 173210 33365847 52868  
## - Protection.zoneLimited 1 175221 33367858 52868  
## - `Juris\_LongBritish Columbia` 1 217990 33410627 52876  
## - Juris\_LongManitoba 1 231086 33423722 52878  
## - `Juris\_LongNova Scotia` 1 249643 33442280 52882  
## - `Juris\_LongNorthwest Territories` 1 259939 33452576 52884  
## - `Juris\_LongNewfoundland and Labrador` 1 330370 33523007 52897  
## - `Juris\_LongNational parks` 1 344409 33537046 52899  
## - `Juris\_LongPrince Edward Island` 1 352528 33545165 52901  
## - Juris\_LongYukon 1 359471 33552108 52902  
## - Response.categoryModified 1 1603592 34796228 53126  
## - Response.categoryNone 1 1744476 34937113 53150  
## - Fire\_Cause\_HumanLightning 1 2499613 35692250 53282  
##   
## Step: AIC=52835.85  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Response.categoryModified + Response.categoryNone +   
## Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 33193598 52836  
## - Response.categoryUnspecified 1 11648 33205246 52836  
## - Fire\_Cause\_HumanUnspecified 1 35669 33229267 52840  
## - Juris\_LongOntario 1 71338 33264936 52847  
## - Juris\_LongSaskatchewan 1 109636 33303234 52854  
## - `Juris\_LongNew Brunswick` 1 158047 33351645 52863  
## - Juris\_LongQuebec 1 173321 33366919 52866  
## - Protection.zoneLimited 1 176170 33369768 52866  
## - `Juris\_LongBritish Columbia` 1 218075 33411672 52874  
## - Juris\_LongManitoba 1 231194 33424791 52876  
## - `Juris\_LongNova Scotia` 1 249426 33443024 52880  
## - `Juris\_LongNorthwest Territories` 1 260422 33454020 52882  
## - `Juris\_LongNewfoundland and Labrador` 1 330230 33523828 52895  
## - `Juris\_LongNational parks` 1 344369 33537967 52897  
## - `Juris\_LongPrince Edward Island` 1 351726 33545324 52899  
## - Juris\_LongYukon 1 359384 33552982 52900  
## - Response.categoryModified 1 1602969 34796567 53124  
## - Response.categoryNone 1 1746464 34940062 53149  
## - Fire\_Cause\_HumanLightning 1 2527787 35721385 53285  
## Start: AIC=54924.3  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 8008 46551486 54923  
## <none> 46543478 54924  
## - Fire\_Cause\_HumanUnspecified 1 27079 46570557 54926  
## - Response.categoryUnspecified 1 28452 46571930 54926  
## - Juris\_LongOntario 1 123981 46667460 54939  
## - Juris\_LongSaskatchewan 1 144042 46687520 54941  
## - Protection.zoneLimited 1 164688 46708166 54944  
## - Protection.zoneUnspecified 1 197848 46741326 54948  
## - `Juris\_LongBritish Columbia` 1 230421 46773900 54953  
## - Juris\_LongQuebec 1 232439 46775917 54953  
## - `Juris\_LongNew Brunswick` 1 236987 46780466 54954  
## - Juris\_LongManitoba 1 298889 46842367 54962  
## - `Juris\_LongNova Scotia` 1 299167 46842646 54962  
## - `Juris\_LongNorthwest Territories` 1 359442 46902920 54970  
## - `Juris\_LongNational parks` 1 400198 46943677 54975  
## - Juris\_LongYukon 1 414761 46958240 54977  
## - `Juris\_LongNewfoundland and Labrador` 1 443404 46986882 54981  
## - `Juris\_LongPrince Edward Island` 1 509357 47052836 54989  
## - Response.categoryModified 1 2013590 48557069 55183  
## - Response.categoryNone 1 2028889 48572367 55184  
## - Fire\_Cause\_HumanLightning 1 2841081 49384559 55286  
##   
## Step: AIC=54923.36  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 46551486 54923  
## - Fire\_Cause\_HumanUnspecified 1 26731 46578217 54925  
## - Response.categoryUnspecified 1 34164 46585650 54926  
## - Juris\_LongOntario 1 121960 46673447 54937  
## - Juris\_LongSaskatchewan 1 146263 46697749 54941  
## - Protection.zoneLimited 1 167296 46718782 54943  
## - Protection.zoneUnspecified 1 191217 46742703 54947  
## - Juris\_LongQuebec 1 232041 46783528 54952  
## - `Juris\_LongBritish Columbia` 1 234727 46786213 54952  
## - `Juris\_LongNew Brunswick` 1 238482 46789968 54953  
## - `Juris\_LongNova Scotia` 1 299709 46851196 54961  
## - Juris\_LongManitoba 1 305395 46856881 54962  
## - `Juris\_LongNorthwest Territories` 1 363900 46915386 54969  
## - `Juris\_LongNational parks` 1 402447 46953934 54974  
## - Juris\_LongYukon 1 422308 46973794 54977  
## - `Juris\_LongNewfoundland and Labrador` 1 443212 46994698 54980  
## - `Juris\_LongPrince Edward Island` 1 504193 47055679 54988  
## - Response.categoryModified 1 2012385 48563872 55181  
## - Response.categoryNone 1 2022303 48573789 55183  
## - Fire\_Cause\_HumanLightning 1 2840580 49392067 55285  
## Start: AIC=54150.84  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Response.categoryUnspecified 1 9327 41099526 54150  
## <none> 41090200 54151  
## - Fire\_Cause\_HumanUnspecified 1 34230 41124429 54154  
## - Year 1 35533 41125733 54154  
## - Juris\_LongSaskatchewan 1 114588 41204788 54166  
## - Juris\_LongOntario 1 123710 41213910 54167  
## - Protection.zoneLimited 1 131668 41221867 54168  
## - `Juris\_LongNew Brunswick` 1 156369 41246568 54172  
## - Juris\_LongQuebec 1 178267 41268467 54175  
## - Protection.zoneUnspecified 1 227482 41317682 54183  
## - `Juris\_LongNova Scotia` 1 228720 41318920 54183  
## - Juris\_LongManitoba 1 254035 41344235 54187  
## - `Juris\_LongBritish Columbia` 1 275335 41365535 54190  
## - `Juris\_LongNorthwest Territories` 1 309963 41400162 54195  
## - `Juris\_LongNational parks` 1 346376 41436576 54200  
## - Juris\_LongYukon 1 363018 41453217 54203  
## - `Juris\_LongNewfoundland and Labrador` 1 381592 41471792 54206  
## - `Juris\_LongPrince Edward Island` 1 402355 41492554 54209  
## - Response.categoryNone 1 1734748 42824948 54403  
## - Response.categoryModified 1 1746933 42837133 54405  
## - Fire\_Cause\_HumanLightning 1 2267597 43357796 54479  
##   
## Step: AIC=54150.24  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone  
##   
## Df Sum of Sq RSS AIC  
## <none> 41099526 54150  
## - Fire\_Cause\_HumanUnspecified 1 34865 41134391 54153  
## - Year 1 42322 41141848 54155  
## - Juris\_LongSaskatchewan 1 107840 41207367 54164  
## - Juris\_LongOntario 1 123185 41222712 54167  
## - Protection.zoneLimited 1 135234 41234760 54168  
## - `Juris\_LongNew Brunswick` 1 157959 41257485 54172  
## - Juris\_LongQuebec 1 177190 41276716 54175  
## - Protection.zoneUnspecified 1 228053 41327579 54182  
## - `Juris\_LongNova Scotia` 1 228590 41328117 54182  
## - Juris\_LongManitoba 1 247205 41346731 54185  
## - `Juris\_LongBritish Columbia` 1 274404 41373930 54189  
## - `Juris\_LongNorthwest Territories` 1 305544 41405071 54194  
## - `Juris\_LongNational parks` 1 345871 41445398 54200  
## - Juris\_LongYukon 1 361059 41460585 54202  
## - `Juris\_LongNewfoundland and Labrador` 1 380236 41479762 54205  
## - `Juris\_LongPrince Edward Island` 1 404114 41503641 54208  
## - Response.categoryNone 1 1794259 42893785 54411  
## - Response.categoryModified 1 1795703 42895229 54411  
## - Fire\_Cause\_HumanLightning 1 2266424 43365950 54478  
## Start: AIC=53342.01  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 11235 36032235 53342  
## - Response.categoryUnspecified 1 11522 36032522 53342  
## <none> 36021000 53342  
## - Fire\_Cause\_HumanUnspecified 1 32432 36053432 53346  
## - `Juris\_LongBritish Columbia` 1 78126 36099126 53353  
## - Protection.zoneUnspecified 1 142720 36163719 53364  
## - Protection.zoneLimited 1 147751 36168751 53365  
## - Juris\_LongOntario 1 156086 36177086 53367  
## - Juris\_LongSaskatchewan 1 204927 36225927 53375  
## - `Juris\_LongNew Brunswick` 1 262113 36283113 53385  
## - Juris\_LongQuebec 1 284019 36305019 53388  
## - Juris\_LongManitoba 1 350447 36371447 53399  
## - `Juris\_LongNova Scotia` 1 364163 36385163 53402  
## - `Juris\_LongNorthwest Territories` 1 424171 36445171 53412  
## - `Juris\_LongNational parks` 1 457001 36478001 53417  
## - `Juris\_LongPrince Edward Island` 1 498512 36519512 53424  
## - `Juris\_LongNewfoundland and Labrador` 1 505759 36526759 53426  
## - Juris\_LongYukon 1 531695 36552695 53430  
## - Response.categoryNone 1 1725222 37746222 53627  
## - Response.categoryModified 1 1758242 37779241 53633  
## - Fire\_Cause\_HumanLightning 1 2166743 38187743 53699  
##   
## Step: AIC=53341.93  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 36032235 53342  
## - Response.categoryUnspecified 1 15882 36048117 53343  
## - Fire\_Cause\_HumanUnspecified 1 32107 36064342 53345  
## - `Juris\_LongBritish Columbia` 1 80657 36112892 53354  
## - Protection.zoneUnspecified 1 135424 36167660 53363  
## - Juris\_LongOntario 1 152891 36185126 53366  
## - Protection.zoneLimited 1 154130 36186365 53366  
## - Juris\_LongSaskatchewan 1 209042 36241277 53375  
## - `Juris\_LongNew Brunswick` 1 265569 36297804 53385  
## - Juris\_LongQuebec 1 283402 36315637 53388  
## - Juris\_LongManitoba 1 357540 36389775 53401  
## - `Juris\_LongNova Scotia` 1 364577 36396812 53402  
## - `Juris\_LongNorthwest Territories` 1 431605 36463840 53413  
## - `Juris\_LongNational parks` 1 462532 36494767 53418  
## - `Juris\_LongPrince Edward Island` 1 491558 36523793 53423  
## - `Juris\_LongNewfoundland and Labrador` 1 504670 36536905 53425  
## - Juris\_LongYukon 1 540522 36572757 53431  
## - Response.categoryNone 1 1716150 37748385 53626  
## - Response.categoryModified 1 1755167 37787402 53632  
## - Fire\_Cause\_HumanLightning 1 2168754 38200989 53699  
## Start: AIC=54317.5  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 1401 42221589 54316  
## <none> 42220189 54317  
## - Response.categoryUnspecified 1 20517 42240706 54318  
## - Fire\_Cause\_HumanUnspecified 1 35748 42255937 54321  
## - Juris\_LongOntario 1 88411 42308600 54328  
## - Juris\_LongSaskatchewan 1 109615 42329804 54331  
## - Juris\_LongQuebec 1 144970 42365159 54337  
## - Protection.zoneLimited 1 149465 42369654 54337  
## - Protection.zoneUnspecified 1 153444 42373633 54338  
## - `Juris\_LongNew Brunswick` 1 158131 42378320 54338  
## - Juris\_LongManitoba 1 208865 42429053 54346  
## - `Juris\_LongNova Scotia` 1 225192 42445381 54348  
## - `Juris\_LongNorthwest Territories` 1 248454 42468643 54352  
## - Juris\_LongYukon 1 311778 42531966 54361  
## - `Juris\_LongNewfoundland and Labrador` 1 325171 42545360 54363  
## - `Juris\_LongNational parks` 1 347972 42568161 54366  
## - `Juris\_LongPrince Edward Island` 1 380262 42600451 54371  
## - `Juris\_LongBritish Columbia` 1 408185 42628374 54375  
## - Response.categoryModified 1 1762770 43982958 54567  
## - Response.categoryNone 1 1848951 44069140 54579  
## - Fire\_Cause\_HumanLightning 1 2864070 45084259 54719  
##   
## Step: AIC=54315.7  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 42221589 54316  
## - Response.categoryUnspecified 1 22573 42244162 54317  
## - Fire\_Cause\_HumanUnspecified 1 35585 42257175 54319  
## - Juris\_LongOntario 1 87739 42309329 54326  
## - Juris\_LongSaskatchewan 1 110490 42332079 54330  
## - Juris\_LongQuebec 1 144705 42366295 54335  
## - Protection.zoneLimited 1 150831 42372420 54336  
## - Protection.zoneUnspecified 1 152139 42373729 54336  
## - `Juris\_LongNew Brunswick` 1 158161 42379751 54337  
## - Juris\_LongManitoba 1 211400 42432990 54344  
## - `Juris\_LongNova Scotia` 1 225147 42446737 54346  
## - `Juris\_LongNorthwest Territories` 1 249612 42471202 54350  
## - Juris\_LongYukon 1 315007 42536597 54359  
## - `Juris\_LongNewfoundland and Labrador` 1 324948 42546538 54361  
## - `Juris\_LongNational parks` 1 349139 42570728 54364  
## - `Juris\_LongPrince Edward Island` 1 378862 42600452 54369  
## - `Juris\_LongBritish Columbia` 1 410573 42632162 54373  
## - Response.categoryModified 1 1761661 43983251 54565  
## - Response.categoryNone 1 1849832 44071421 54577  
## - Fire\_Cause\_HumanLightning 1 2865425 45087014 54717  
## Start: AIC=54386.41  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 13758 42655764 54386  
## <none> 42642006 54386  
## - Response.categoryUnspecified 1 16221 42658227 54387  
## - Protection.zoneUnspecified 1 28013 42670019 54388  
## - Fire\_Cause\_HumanUnspecified 1 29881 42671887 54389  
## - Juris\_LongSaskatchewan 1 67249 42709255 54394  
## - Juris\_LongOntario 1 76697 42718703 54395  
## - `Juris\_LongNew Brunswick` 1 136910 42778916 54404  
## - Juris\_LongQuebec 1 162605 42804611 54408  
## - Protection.zoneLimited 1 173714 42815721 54409  
## - `Juris\_LongNova Scotia` 1 202462 42844468 54414  
## - Juris\_LongManitoba 1 232561 42874567 54418  
## - `Juris\_LongNorthwest Territories` 1 267511 42909517 54423  
## - `Juris\_LongNational parks` 1 296085 42938091 54427  
## - `Juris\_LongBritish Columbia` 1 296221 42938227 54427  
## - Juris\_LongYukon 1 311148 42953154 54429  
## - `Juris\_LongNewfoundland and Labrador` 1 333230 42975236 54432  
## - `Juris\_LongPrince Edward Island` 1 383188 43025194 54439  
## - Response.categoryModified 1 1834432 44476438 54643  
## - Response.categoryNone 1 1929557 44571564 54656  
## - Fire\_Cause\_HumanLightning 1 2568608 45210614 54744  
##   
## Step: AIC=54386.39  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 42655764 54386  
## - Response.categoryUnspecified 1 21423 42677187 54387  
## - Protection.zoneUnspecified 1 23871 42679635 54388  
## - Fire\_Cause\_HumanUnspecified 1 29303 42685068 54389  
## - Juris\_LongSaskatchewan 1 67265 42723029 54394  
## - Juris\_LongOntario 1 73991 42729755 54395  
## - `Juris\_LongNew Brunswick` 1 137342 42793107 54404  
## - Juris\_LongQuebec 1 161445 42817209 54408  
## - Protection.zoneLimited 1 178600 42834364 54410  
## - `Juris\_LongNova Scotia` 1 203010 42858775 54414  
## - Juris\_LongManitoba 1 238022 42893786 54419  
## - `Juris\_LongNorthwest Territories` 1 272205 42927970 54423  
## - `Juris\_LongNational parks` 1 299382 42955147 54427  
## - `Juris\_LongBritish Columbia` 1 301704 42957468 54428  
## - Juris\_LongYukon 1 318605 42974370 54430  
## - `Juris\_LongNewfoundland and Labrador` 1 331617 42987381 54432  
## - `Juris\_LongPrince Edward Island` 1 377247 43033011 54438  
## - Response.categoryModified 1 1835592 44491356 54643  
## - Response.categoryNone 1 1923301 44579065 54655  
## - Fire\_Cause\_HumanLightning 1 2569980 45225744 54744  
## Start: AIC=81041.07  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 11579 60532693 81041  
## <none> 60521114 81041  
## - Response.categoryUnspecified 1 23579 60544694 81043  
## - Fire\_Cause\_HumanUnspecified 1 48906 60570020 81047  
## - Protection.zoneUnspecified 1 146151 60667265 81061  
## - Juris\_LongOntario 1 157841 60678955 81063  
## - Juris\_LongSaskatchewan 1 183755 60704869 81067  
## - Protection.zoneLimited 1 236177 60757291 81075  
## - `Juris\_LongNew Brunswick` 1 276479 60797593 81081  
## - Juris\_LongQuebec 1 289853 60810967 81083  
## - `Juris\_LongBritish Columbia` 1 360728 60881843 81094  
## - `Juris\_LongNova Scotia` 1 389243 60910357 81098  
## - Juris\_LongManitoba 1 390451 60911566 81098  
## - `Juris\_LongNorthwest Territories` 1 464164 60985278 81109  
## - `Juris\_LongNational parks` 1 547198 61068313 81122  
## - Juris\_LongYukon 1 568954 61090069 81125  
## - `Juris\_LongNewfoundland and Labrador` 1 576201 61097315 81126  
## - `Juris\_LongPrince Edward Island` 1 630447 61151561 81135  
## - Response.categoryModified 1 2685813 63206928 81439  
## - Response.categoryNone 1 2759038 63280152 81450  
## - Fire\_Cause\_HumanLightning 1 3788701 64309815 81599  
##   
## Step: AIC=81040.83  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 60532693 81041  
## - Response.categoryUnspecified 1 29530 60562223 81043  
## - Fire\_Cause\_HumanUnspecified 1 48363 60581056 81046  
## - Protection.zoneUnspecified 1 138536 60671229 81060  
## - Juris\_LongOntario 1 154665 60687358 81062  
## - Juris\_LongSaskatchewan 1 186033 60718727 81067  
## - Protection.zoneLimited 1 242169 60774862 81076  
## - `Juris\_LongNew Brunswick` 1 277878 60810572 81081  
## - Juris\_LongQuebec 1 288795 60821488 81083  
## - `Juris\_LongBritish Columbia` 1 366327 60899020 81094  
## - `Juris\_LongNova Scotia` 1 389527 60922220 81098  
## - Juris\_LongManitoba 1 398114 60930808 81099  
## - `Juris\_LongNorthwest Territories` 1 470080 61002773 81110  
## - `Juris\_LongNational parks` 1 551875 61084568 81122  
## - `Juris\_LongNewfoundland and Labrador` 1 574878 61107571 81126  
## - Juris\_LongYukon 1 578787 61111480 81127  
## - `Juris\_LongPrince Edward Island` 1 623000 61155693 81133  
## - Response.categoryModified 1 2683280 63215974 81439  
## - Response.categoryNone 1 2748751 63281444 81448  
## - Fire\_Cause\_HumanLightning 1 3791371 64324065 81599

# lmStepAIC\_Mod <- train(Number ~ Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lmStepAIC",  
# trControl = train.control)  
#   
# lmStepAIC\_Mod <- train(Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lmStepAIC",  
# trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
lmStepAIC\_Mod\_time.taken <- end.time - start.time  
#lmStepAIC\_Mod\_time.taken

## 21.View the total amount of time taken to run each model.

Model\_Time <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 Time\_train=c(lm\_model\_time.taken, glm\_model\_time.taken, lasso\_Mod\_time.taken, knn\_model\_time.taken, LF\_model\_time.taken, LB\_model\_time.taken, lmStepAIC\_Mod\_time.taken))  
Model\_Time

## Name Time\_train  
## 1 lm\_model 1.9785759 secs  
## 2 glm\_model 1.6067028 secs  
## 3 lasso\_Mod 1.8617730 secs  
## 4 knn\_model 22.7727718 secs  
## 5 LF\_model 0.9035399 secs  
## 6 LB\_model 0.8976009 secs  
## 7 lmStepAIC\_Mod 5.1023571 secs

## 22.View summaries of the models.

############################################  
# Model 1  
# Summarize the results for lm\_model  
print(lm\_model)

## Linear Regression   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 80.77566 0.1693413 31.68217  
##   
## Tuning parameter 'intercept' was held constant at a value of TRUE

summary(lm\_model)

##   
## Call:  
## lm(formula = .outcome ~ ., data = dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -154.24 -26.96 0.70 14.48 2740.33   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 357.4664 216.5346 1.651 0.09880  
## `Juris\_LongBritish Columbia` 36.3759 4.9139 7.403 1.45e-13  
## Juris\_LongManitoba -33.2714 4.3200 -7.702 1.48e-14  
## `Juris\_LongNational parks` -45.1823 4.9556 -9.117 < 2e-16  
## `Juris\_LongNew Brunswick` -42.6220 6.5766 -6.481 9.59e-11  
## `Juris\_LongNewfoundland and Labrador` -39.5928 4.2319 -9.356 < 2e-16  
## `Juris\_LongNorthwest Territories` -39.0974 4.6560 -8.397 < 2e-16  
## `Juris\_LongNova Scotia` -36.7600 4.7804 -7.690 1.63e-14  
## Juris\_LongOntario -21.0458 4.2979 -4.897 9.91e-07  
## `Juris\_LongPrince Edward Island` -54.8988 5.6097 -9.786 < 2e-16  
## Juris\_LongQuebec -28.0000 4.2196 -6.636 3.41e-11  
## Juris\_LongSaskatchewan -26.4579 5.0077 -5.283 1.30e-07  
## Juris\_LongYukon -41.5927 4.4738 -9.297 < 2e-16  
## Year -0.1433 0.1081 -1.326 0.18478  
## Fire\_Cause\_HumanLightning 64.5798 2.6919 23.991 < 2e-16  
## Fire\_Cause\_HumanUnspecified -7.3782 2.7069 -2.726 0.00643  
## Protection.zoneLimited -13.0634 2.1809 -5.990 2.18e-09  
## Protection.zoneUnspecified 62.1442 13.1887 4.712 2.49e-06  
## Response.categoryModified -42.5181 2.1049 -20.199 < 2e-16  
## Response.categoryNone -42.5989 2.0808 -20.473 < 2e-16  
## Response.categoryUnspecified 16.7966 8.8748 1.893 0.05844  
##

## (Intercept) .   
## `Juris\_LongBritish Columbia` \*\*\*  
## Juris\_LongManitoba \*\*\*  
## `Juris\_LongNational parks` \*\*\*  
## `Juris\_LongNew Brunswick` \*\*\*  
## `Juris\_LongNewfoundland and Labrador` \*\*\*  
## `Juris\_LongNorthwest Territories` \*\*\*  
## `Juris\_LongNova Scotia` \*\*\*  
## Juris\_LongOntario \*\*\*  
## `Juris\_LongPrince Edward Island` \*\*\*  
## Juris\_LongQuebec \*\*\*  
## Juris\_LongSaskatchewan \*\*\*  
## Juris\_LongYukon \*\*\*  
## Year   
## Fire\_Cause\_HumanLightning \*\*\*  
## Fire\_Cause\_HumanUnspecified \*\*   
## Protection.zoneLimited \*\*\*  
## Protection.zoneUnspecified \*\*\*  
## Response.categoryModified \*\*\*  
## Response.categoryNone \*\*\*  
## Response.categoryUnspecified .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 81.13 on 9194 degrees of freedom  
## Multiple R-squared: 0.1705, Adjusted R-squared: 0.1687   
## F-statistic: 94.51 on 20 and 9194 DF, p-value: < 2.2e-16

lm\_model$finalModel

##   
## Call:  
## lm(formula = .outcome ~ ., data = dat)  
##   
## Coefficients:  
## (Intercept)   
## 357.4664   
## `Juris\_LongBritish Columbia`   
## 36.3759   
## Juris\_LongManitoba   
## -33.2714   
## `Juris\_LongNational parks`   
## -45.1823   
## `Juris\_LongNew Brunswick`   
## -42.6220   
## `Juris\_LongNewfoundland and Labrador`   
## -39.5928   
## `Juris\_LongNorthwest Territories`   
## -39.0974   
## `Juris\_LongNova Scotia`   
## -36.7600   
## Juris\_LongOntario   
## -21.0458   
## `Juris\_LongPrince Edward Island`   
## -54.8988   
## Juris\_LongQuebec   
## -28.0000   
## Juris\_LongSaskatchewan   
## -26.4579   
## Juris\_LongYukon   
## -41.5927   
## Year   
## -0.1433   
## Fire\_Cause\_HumanLightning   
## 64.5798   
## Fire\_Cause\_HumanUnspecified   
## -7.3782   
## Protection.zoneLimited   
## -13.0634   
## Protection.zoneUnspecified   
## 62.1442   
## Response.categoryModified   
## -42.5181   
## Response.categoryNone   
## -42.5989   
## Response.categoryUnspecified   
## 16.7966

lm\_model$modelType

## [1] "Regression"

############################################  
# Model 2  
# Summarize the results for glm model  
print(glm\_model)

## Generalized Linear Model   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 80.77566 0.1693413 31.68217

summary(glm\_model)

##   
## Call:  
## NULL  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -154.24 -26.96 0.70 14.48 2740.33   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 357.4664 216.5346 1.651 0.09880  
## `Juris\_LongBritish Columbia` 36.3759 4.9139 7.403 1.45e-13  
## Juris\_LongManitoba -33.2714 4.3200 -7.702 1.48e-14  
## `Juris\_LongNational parks` -45.1823 4.9556 -9.117 < 2e-16  
## `Juris\_LongNew Brunswick` -42.6220 6.5766 -6.481 9.59e-11  
## `Juris\_LongNewfoundland and Labrador` -39.5928 4.2319 -9.356 < 2e-16  
## `Juris\_LongNorthwest Territories` -39.0974 4.6560 -8.397 < 2e-16  
## `Juris\_LongNova Scotia` -36.7600 4.7804 -7.690 1.63e-14  
## Juris\_LongOntario -21.0458 4.2979 -4.897 9.91e-07  
## `Juris\_LongPrince Edward Island` -54.8988 5.6097 -9.786 < 2e-16  
## Juris\_LongQuebec -28.0000 4.2196 -6.636 3.41e-11  
## Juris\_LongSaskatchewan -26.4579 5.0077 -5.283 1.30e-07  
## Juris\_LongYukon -41.5927 4.4738 -9.297 < 2e-16  
## Year -0.1433 0.1081 -1.326 0.18478  
## Fire\_Cause\_HumanLightning 64.5798 2.6919 23.991 < 2e-16  
## Fire\_Cause\_HumanUnspecified -7.3782 2.7069 -2.726 0.00643  
## Protection.zoneLimited -13.0634 2.1809 -5.990 2.18e-09  
## Protection.zoneUnspecified 62.1442 13.1887 4.712 2.49e-06  
## Response.categoryModified -42.5181 2.1049 -20.199 < 2e-16  
## Response.categoryNone -42.5989 2.0808 -20.473 < 2e-16  
## Response.categoryUnspecified 16.7966 8.8748 1.893 0.05844  
##   
## (Intercept) .   
## `Juris\_LongBritish Columbia` \*\*\*  
## Juris\_LongManitoba \*\*\*  
## `Juris\_LongNational parks` \*\*\*  
## `Juris\_LongNew Brunswick` \*\*\*  
## `Juris\_LongNewfoundland and Labrador` \*\*\*  
## `Juris\_LongNorthwest Territories` \*\*\*  
## `Juris\_LongNova Scotia` \*\*\*  
## Juris\_LongOntario \*\*\*  
## `Juris\_LongPrince Edward Island` \*\*\*  
## Juris\_LongQuebec \*\*\*  
## Juris\_LongSaskatchewan \*\*\*  
## Juris\_LongYukon \*\*\*  
## Year   
## Fire\_Cause\_HumanLightning \*\*\*  
## Fire\_Cause\_HumanUnspecified \*\*   
## Protection.zoneLimited \*\*\*  
## Protection.zoneUnspecified \*\*\*  
## Response.categoryModified \*\*\*  
## Response.categoryNone \*\*\*  
## Response.categoryUnspecified .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 6582.675)  
##   
## Null deviance: 72963124 on 9214 degrees of freedom  
## Residual deviance: 60521114 on 9194 degrees of freedom  
## AIC: 107194  
##   
## Number of Fisher Scoring iterations: 2

glm\_model$finalModel

##   
## Call: NULL  
##   
## Coefficients:  
## (Intercept)   
## 357.4664   
## `Juris\_LongBritish Columbia`   
## 36.3759   
## Juris\_LongManitoba   
## -33.2714   
## `Juris\_LongNational parks`   
## -45.1823   
## `Juris\_LongNew Brunswick`   
## -42.6220   
## `Juris\_LongNewfoundland and Labrador`   
## -39.5928   
## `Juris\_LongNorthwest Territories`   
## -39.0974   
## `Juris\_LongNova Scotia`   
## -36.7600   
## Juris\_LongOntario   
## -21.0458   
## `Juris\_LongPrince Edward Island`   
## -54.8988   
## Juris\_LongQuebec   
## -28.0000   
## Juris\_LongSaskatchewan   
## -26.4579   
## Juris\_LongYukon   
## -41.5927   
## Year   
## -0.1433   
## Fire\_Cause\_HumanLightning   
## 64.5798   
## Fire\_Cause\_HumanUnspecified   
## -7.3782   
## Protection.zoneLimited   
## -13.0634   
## Protection.zoneUnspecified   
## 62.1442   
## Response.categoryModified   
## -42.5181   
## Response.categoryNone   
## -42.5989   
## Response.categoryUnspecified   
## 16.7966   
##   
## Degrees of Freedom: 9214 Total (i.e. Null); 9194 Residual  
## Null Deviance: 72960000   
## Residual Deviance: 60520000 AIC: 107200

glm\_model$modelType

## [1] "Regression"

############################################  
# Model 3  
# Summarize the results  
print(lasso\_Mod)

## The lasso   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results across tuning parameters:  
##   
## fraction RMSE Rsquared MAE   
## 0.1 85.28627 0.09964584 27.00541  
## 0.5 81.22124 0.16058728 28.09510  
## 0.9 80.78302 0.16898311 31.18791  
##   
## RMSE was used to select the optimal model using the smallest value.  
## The final value used for the model was fraction = 0.9.

summary(lasso\_Mod)

## Length Class Mode   
## call 4 -none- call   
## actions 23 -none- list   
## allset 20 -none- numeric   
## beta.pure 460 -none- numeric   
## vn 20 -none- character  
## mu 1 -none- numeric   
## normx 20 -none- numeric   
## meanx 20 -none- numeric   
## lambda 1 -none- numeric   
## L1norm 23 -none- numeric   
## penalty 23 -none- numeric   
## df 23 -none- numeric   
## Cp 23 -none- numeric   
## sigma2 1 -none- numeric   
## xNames 20 -none- character  
## problemType 1 -none- character  
## tuneValue 1 data.frame list   
## obsLevels 1 -none- logical   
## param 0 -none- list

lasso\_Mod$finalModel

##   
## Call:  
## elasticnet::enet(x = as.matrix(x), y = y, lambda = 0)  
## Cp statistics of the Lasso fit   
## Cp: 1871.115 1661.125 1210.760 1162.009 753.090 374.862 347.432 329.418 324.540 295.076 281.463 232.746 223.123 197.905 161.413 149.090 149.548 145.647 97.212 77.744 76.761 57.844 21.000   
## DF: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 19 19 20 21   
## Sequence of moves:  
## Fire\_Cause\_HumanLightning Juris\_LongBritish Columbia  
## Var 14 1  
## Step 1 2  
## Response.categoryNone Response.categoryModified  
## Var 19 18  
## Step 3 4  
## Protection.zoneLimited Protection.zoneUnspecified  
## Var 16 17  
## Step 5 6  
## Juris\_LongNewfoundland and Labrador Juris\_LongPrince Edward Island  
## Var 5 9  
## Step 7 8  
## Juris\_LongYukon Juris\_LongOntario Juris\_LongNational parks  
## Var 12 8 3  
## Step 9 10 11  
## Fire\_Cause\_HumanUnspecified Response.categoryUnspecified  
## Var 15 20  
## Step 12 13  
## Juris\_LongNorthwest Territories Juris\_LongNova Scotia Year  
## Var 6 7 13  
## Step 14 15 16  
## Juris\_LongManitoba Juris\_LongNew Brunswick Juris\_LongQuebec  
## Var 2 4 10  
## Step 17 18 19  
## Juris\_LongOntario Juris\_LongSaskatchewan Juris\_LongOntario   
## Var -8 11 8 23  
## Step 20 21 22 23

lasso\_Mod$modelType

## [1] "Regression"

############################################  
# Model 4  
# Summarize the results for knn model  
print(knn\_model)

## k-Nearest Neighbors   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results across tuning parameters:  
##   
## k RMSE Rsquared MAE   
## 5 76.46193 0.2743102 19.88133  
## 7 77.90343 0.2502236 20.71882  
## 9 79.47114 0.2126232 21.15818  
##   
## RMSE was used to select the optimal model using the smallest value.  
## The final value used for the model was k = 5.

summary(knn\_model)

## Length Class Mode   
## learn 2 -none- list   
## k 1 -none- numeric   
## theDots 0 -none- list   
## xNames 20 -none- character  
## problemType 1 -none- character  
## tuneValue 1 data.frame list   
## obsLevels 1 -none- logical   
## param 0 -none- list

knn\_model$finalModel

## 5-nearest neighbor regression model

knn\_model$modelType

## [1] "Regression"

############################################  
# Model 5  
# Summarize the results for LF model  
print(LF\_model)

## Linear Regression with Forward Selection   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results across tuning parameters:  
##   
## nvmax RMSE Rsquared MAE   
## 2 84.39884 0.0922847 26.34798  
## 3 83.68595 0.1076692 29.48269  
## 4 81.83715 0.1471318 30.66423  
##   
## RMSE was used to select the optimal model using the smallest value.  
## The final value used for the model was nvmax = 4.

summary(LF\_model)

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: forward  
## Juris\_LongBritish Columbia Juris\_LongManitoba  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) "\*" " "   
## 3 ( 1 ) "\*" " "   
## 4 ( 1 ) "\*" " "   
## Juris\_LongNational parks Juris\_LongNew Brunswick  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongNewfoundland and Labrador  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "   
## Juris\_LongNorthwest Territories Juris\_LongNova Scotia  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongOntario Juris\_LongPrince Edward Island Juris\_LongQuebec  
## 1 ( 1 ) " " " " " "   
## 2 ( 1 ) " " " " " "   
## 3 ( 1 ) " " " " " "   
## 4 ( 1 ) " " " " " "   
## Juris\_LongSaskatchewan Juris\_LongYukon Year  
## 1 ( 1 ) " " " " " "   
## 2 ( 1 ) " " " " " "   
## 3 ( 1 ) " " " " " "   
## 4 ( 1 ) " " " " " "   
## Fire\_Cause\_HumanLightning Fire\_Cause\_HumanUnspecified  
## 1 ( 1 ) "\*" " "   
## 2 ( 1 ) "\*" " "   
## 3 ( 1 ) "\*" " "   
## 4 ( 1 ) "\*" " "   
## Protection.zoneLimited Protection.zoneUnspecified  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Response.categoryModified Response.categoryNone  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " "\*"   
## 4 ( 1 ) "\*" "\*"   
## Response.categoryUnspecified  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "

LF\_model$finalModel

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: forward

LF\_model$modelType

## [1] "Regression"

############################################  
# Model 6  
# Summarize the results for LB Modle  
print(LB\_model)

## Linear Regression with Backwards Selection   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results across tuning parameters:  
##   
## nvmax RMSE Rsquared MAE   
## 2 85.12492 0.07581972 27.81353  
## 3 83.74358 0.10550111 28.73342  
## 4 82.11647 0.14152807 30.34153  
##   
## RMSE was used to select the optimal model using the smallest value.  
## The final value used for the model was nvmax = 4.

summary(LB\_model)

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: backward  
## Juris\_LongBritish Columbia Juris\_LongManitoba  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) "\*" " "   
## Juris\_LongNational parks Juris\_LongNew Brunswick  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongNewfoundland and Labrador  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "   
## Juris\_LongNorthwest Territories Juris\_LongNova Scotia  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongOntario Juris\_LongPrince Edward Island Juris\_LongQuebec  
## 1 ( 1 ) " " " " " "   
## 2 ( 1 ) " " " " " "   
## 3 ( 1 ) " " " " " "   
## 4 ( 1 ) " " " " " "   
## Juris\_LongSaskatchewan Juris\_LongYukon Year  
## 1 ( 1 ) " " " " " "   
## 2 ( 1 ) " " " " " "   
## 3 ( 1 ) " " " " " "   
## 4 ( 1 ) " " " " " "   
## Fire\_Cause\_HumanLightning Fire\_Cause\_HumanUnspecified  
## 1 ( 1 ) "\*" " "   
## 2 ( 1 ) "\*" " "   
## 3 ( 1 ) "\*" " "   
## 4 ( 1 ) "\*" " "   
## Protection.zoneLimited Protection.zoneUnspecified  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Response.categoryModified Response.categoryNone  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " "\*"   
## 3 ( 1 ) "\*" "\*"   
## 4 ( 1 ) "\*" "\*"   
## Response.categoryUnspecified  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "

LB\_model$finalModel

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: backward

LB\_model$modelType

## [1] "Regression"

############################################  
# Model 7  
# Summarize the results for lmStepAIC  
print(lmStepAIC\_Mod)

## Linear Regression with Stepwise Selection   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 80.79002 0.1690438 31.69457

summary(lmStepAIC\_Mod)

##   
## Call:  
## lm(formula = .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified, data = dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -154.23 -27.83 0.12 14.05 2741.45   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 70.323 3.576 19.663 < 2e-16  
## `Juris\_LongBritish Columbia` 36.629 4.910 7.460 9.47e-14  
## Juris\_LongManitoba -33.555 4.315 -7.777 8.26e-15  
## `Juris\_LongNational parks` -45.359 4.954 -9.156 < 2e-16  
## `Juris\_LongNew Brunswick` -42.727 6.576 -6.497 8.62e-11  
## `Juris\_LongNewfoundland and Labrador` -39.546 4.232 -9.345 < 2e-16  
## `Juris\_LongNorthwest Territories` -39.320 4.653 -8.450 < 2e-16  
## `Juris\_LongNova Scotia` -36.773 4.781 -7.692 1.60e-14  
## Juris\_LongOntario -20.816 4.295 -4.847 1.27e-06  
## `Juris\_LongPrince Edward Island` -54.492 5.602 -9.728 < 2e-16  
## Juris\_LongQuebec -27.948 4.220 -6.623 3.71e-11  
## Juris\_LongSaskatchewan -26.614 5.007 -5.316 1.09e-07  
## Juris\_LongYukon -41.896 4.468 -9.376 < 2e-16  
## Fire\_Cause\_HumanLightning 64.601 2.692 23.998 < 2e-16  
## Fire\_Cause\_HumanUnspecified -7.337 2.707 -2.710 0.00673  
## Protection.zoneLimited -13.211 2.178 -6.065 1.37e-09  
## Protection.zoneUnspecified 60.081 13.097 4.587 4.55e-06  
## Response.categoryModified -42.497 2.105 -20.189 < 2e-16  
## Response.categoryNone -42.480 2.079 -20.434 < 2e-16  
## Response.categoryUnspecified 18.580 8.773 2.118 0.03421  
##   
## (Intercept) \*\*\*  
## `Juris\_LongBritish Columbia` \*\*\*  
## Juris\_LongManitoba \*\*\*  
## `Juris\_LongNational parks` \*\*\*  
## `Juris\_LongNew Brunswick` \*\*\*  
## `Juris\_LongNewfoundland and Labrador` \*\*\*  
## `Juris\_LongNorthwest Territories` \*\*\*  
## `Juris\_LongNova Scotia` \*\*\*  
## Juris\_LongOntario \*\*\*  
## `Juris\_LongPrince Edward Island` \*\*\*  
## Juris\_LongQuebec \*\*\*  
## Juris\_LongSaskatchewan \*\*\*  
## Juris\_LongYukon \*\*\*  
## Fire\_Cause\_HumanLightning \*\*\*  
## Fire\_Cause\_HumanUnspecified \*\*   
## Protection.zoneLimited \*\*\*  
## Protection.zoneUnspecified \*\*\*  
## Response.categoryModified \*\*\*  
## Response.categoryNone \*\*\*  
## Response.categoryUnspecified \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 81.14 on 9195 degrees of freedom  
## Multiple R-squared: 0.1704, Adjusted R-squared: 0.1687   
## F-statistic: 99.38 on 19 and 9195 DF, p-value: < 2.2e-16

lmStepAIC\_Mod$finalModel

##   
## Call:  
## lm(formula = .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified, data = dat)  
##   
## Coefficients:  
## (Intercept)   
## 70.323   
## `Juris\_LongBritish Columbia`   
## 36.629   
## Juris\_LongManitoba   
## -33.555   
## `Juris\_LongNational parks`   
## -45.359   
## `Juris\_LongNew Brunswick`   
## -42.727   
## `Juris\_LongNewfoundland and Labrador`   
## -39.546   
## `Juris\_LongNorthwest Territories`   
## -39.320   
## `Juris\_LongNova Scotia`   
## -36.773   
## Juris\_LongOntario   
## -20.816   
## `Juris\_LongPrince Edward Island`   
## -54.492   
## Juris\_LongQuebec   
## -27.948   
## Juris\_LongSaskatchewan   
## -26.614   
## Juris\_LongYukon   
## -41.896   
## Fire\_Cause\_HumanLightning   
## 64.601   
## Fire\_Cause\_HumanUnspecified   
## -7.337   
## Protection.zoneLimited   
## -13.211   
## Protection.zoneUnspecified   
## 60.081   
## Response.categoryModified   
## -42.497   
## Response.categoryNone   
## -42.480   
## Response.categoryUnspecified   
## 18.580

lmStepAIC\_Mod$modelType

## [1] "Regression"

############################################

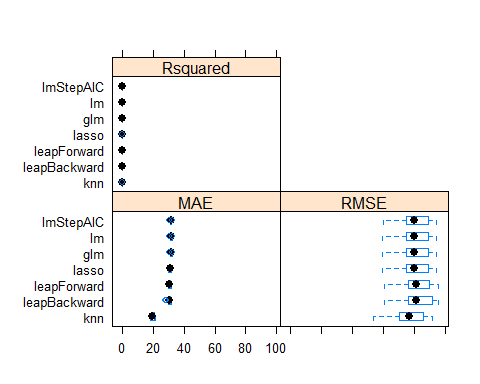
## 23.Evaluation of techniques

## View the results of the models

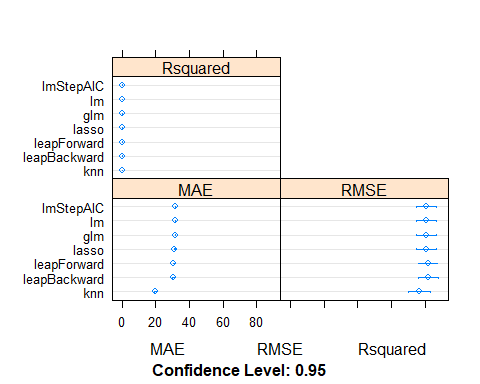
# Evaluation of techniques  
results <- resamples(list(lm=lm\_model, glm=glm\_model,lasso=lasso\_Mod, knn=knn\_model, leapForward=LF\_model,leapBackward=LB\_model, lmStepAIC=lmStepAIC\_Mod))  
summary(results)

##   
## Call:  
## summary.resamples(object = results)  
##   
## Models: lm, glm, lasso, knn, leapForward, leapBackward, lmStepAIC   
## Number of resamples: 15   
##   
## MAE   
## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's  
## lm 30.57488 31.40648 31.66158 31.68217 31.81701 32.51773 0  
## glm 30.57488 31.40648 31.66158 31.68217 31.81701 32.51773 0  
## lasso 30.13162 30.87034 31.15810 31.18791 31.37390 32.10633 0  
## knn 18.46061 19.18022 19.86026 19.88133 20.55260 21.45344 0  
## leapForward 29.89270 30.25847 30.76301 30.66423 30.95333 31.63156 0  
## leapBackward 27.69117 29.96655 30.56175 30.34153 30.95333 31.63156 0  
## lmStepAIC 30.60800 31.44798 31.67075 31.69457 31.83015 32.53927 0  
##   
## RMSE   
## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's  
## lm 59.73223 75.11455 79.98791 80.77566 89.39661 94.59078 0  
## glm 59.73223 75.11455 79.98791 80.77566 89.39661 94.59078 0  
## lasso 59.59428 75.15148 80.03397 80.78302 89.37098 94.64862 0  
## knn 53.74013 70.10013 77.11369 76.46193 85.88831 91.94024 0  
## leapForward 60.96300 76.35707 81.26035 81.83715 90.10764 95.67895 0  
## leapBackward 60.96300 76.35707 81.26035 82.11647 91.63925 95.67895 0  
## lmStepAIC 59.76533 75.11189 79.99824 80.79002 89.41443 94.62547 0  
##   
## Rsquared   
## Min. 1st Qu. Median Mean 3rd Qu. Max.  
## lm 0.1468470 0.1637254 0.1695050 0.1693413 0.1756397 0.1905462  
## glm 0.1468470 0.1637254 0.1695050 0.1693413 0.1756397 0.1905462  
## lasso 0.1462386 0.1641531 0.1685045 0.1689831 0.1751686 0.1906132  
## knn 0.2152862 0.2468016 0.2582658 0.2743102 0.3006185 0.3931389  
## leapForward 0.1232120 0.1420648 0.1442352 0.1471318 0.1559825 0.1704779  
## leapBackward 0.1085446 0.1320164 0.1433190 0.1415281 0.1544986 0.1704779  
## lmStepAIC 0.1461491 0.1632595 0.1692931 0.1690438 0.1751376 0.1903545  
## NA's  
## lm 0  
## glm 0  
## lasso 0  
## knn 0  
## leapForward 0  
## leapBackward 0  
## lmStepAIC 0

# Compare results with boxplots   
bwplot(results)



# Compare results with dot plots   
dotplot(results)



## 24.Find the best results for each model.

get\_best\_result = function(caret\_fit) {  
 best = which(rownames(caret\_fit$results) == rownames(caret\_fit$bestTune))  
 best\_result = caret\_fit$results[best, ]  
 rownames(best\_result) = NULL  
 best\_result  
}  
  
###################################################################################  
lm\_best <- data.frame(get\_best\_result(lm\_model) %>%  
 dplyr::select(2:4))  
lm\_model$finalModel

##   
## Call:  
## lm(formula = .outcome ~ ., data = dat)  
##   
## Coefficients:  
## (Intercept)   
## 357.4664   
## `Juris\_LongBritish Columbia`   
## 36.3759   
## Juris\_LongManitoba   
## -33.2714   
## `Juris\_LongNational parks`   
## -45.1823   
## `Juris\_LongNew Brunswick`   
## -42.6220   
## `Juris\_LongNewfoundland and Labrador`   
## -39.5928   
## `Juris\_LongNorthwest Territories`   
## -39.0974   
## `Juris\_LongNova Scotia`   
## -36.7600   
## Juris\_LongOntario   
## -21.0458   
## `Juris\_LongPrince Edward Island`   
## -54.8988   
## Juris\_LongQuebec   
## -28.0000   
## Juris\_LongSaskatchewan   
## -26.4579   
## Juris\_LongYukon   
## -41.5927   
## Year   
## -0.1433   
## Fire\_Cause\_HumanLightning   
## 64.5798   
## Fire\_Cause\_HumanUnspecified   
## -7.3782   
## Protection.zoneLimited   
## -13.0634   
## Protection.zoneUnspecified   
## 62.1442   
## Response.categoryModified   
## -42.5181   
## Response.categoryNone   
## -42.5989   
## Response.categoryUnspecified   
## 16.7966

###################################################################################  
glm\_best <- data.frame(get\_best\_result(glm\_model) %>%   
 dplyr::select(2:4))  
glm\_model$finalModel

##   
## Call: NULL  
##   
## Coefficients:  
## (Intercept)   
## 357.4664   
## `Juris\_LongBritish Columbia`   
## 36.3759   
## Juris\_LongManitoba   
## -33.2714   
## `Juris\_LongNational parks`   
## -45.1823   
## `Juris\_LongNew Brunswick`   
## -42.6220   
## `Juris\_LongNewfoundland and Labrador`   
## -39.5928   
## `Juris\_LongNorthwest Territories`   
## -39.0974   
## `Juris\_LongNova Scotia`   
## -36.7600   
## Juris\_LongOntario   
## -21.0458   
## `Juris\_LongPrince Edward Island`   
## -54.8988   
## Juris\_LongQuebec   
## -28.0000   
## Juris\_LongSaskatchewan   
## -26.4579   
## Juris\_LongYukon   
## -41.5927   
## Year   
## -0.1433   
## Fire\_Cause\_HumanLightning   
## 64.5798   
## Fire\_Cause\_HumanUnspecified   
## -7.3782   
## Protection.zoneLimited   
## -13.0634   
## Protection.zoneUnspecified   
## 62.1442   
## Response.categoryModified   
## -42.5181   
## Response.categoryNone   
## -42.5989   
## Response.categoryUnspecified   
## 16.7966   
##   
## Degrees of Freedom: 9214 Total (i.e. Null); 9194 Residual  
## Null Deviance: 72960000   
## Residual Deviance: 60520000 AIC: 107200

###################################################################################  
lasso\_best <- data.frame(get\_best\_result(lasso\_Mod) %>%   
 dplyr::select(2:4))  
lasso\_Mod$finalModel

##   
## Call:  
## elasticnet::enet(x = as.matrix(x), y = y, lambda = 0)  
## Cp statistics of the Lasso fit   
## Cp: 1871.115 1661.125 1210.760 1162.009 753.090 374.862 347.432 329.418 324.540 295.076 281.463 232.746 223.123 197.905 161.413 149.090 149.548 145.647 97.212 77.744 76.761 57.844 21.000   
## DF: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 19 19 20 21   
## Sequence of moves:  
## Fire\_Cause\_HumanLightning Juris\_LongBritish Columbia  
## Var 14 1  
## Step 1 2  
## Response.categoryNone Response.categoryModified  
## Var 19 18  
## Step 3 4  
## Protection.zoneLimited Protection.zoneUnspecified  
## Var 16 17  
## Step 5 6  
## Juris\_LongNewfoundland and Labrador Juris\_LongPrince Edward Island  
## Var 5 9  
## Step 7 8  
## Juris\_LongYukon Juris\_LongOntario Juris\_LongNational parks  
## Var 12 8 3  
## Step 9 10 11  
## Fire\_Cause\_HumanUnspecified Response.categoryUnspecified  
## Var 15 20  
## Step 12 13  
## Juris\_LongNorthwest Territories Juris\_LongNova Scotia Year  
## Var 6 7 13  
## Step 14 15 16  
## Juris\_LongManitoba Juris\_LongNew Brunswick Juris\_LongQuebec  
## Var 2 4 10  
## Step 17 18 19  
## Juris\_LongOntario Juris\_LongSaskatchewan Juris\_LongOntario   
## Var -8 11 8 23  
## Step 20 21 22 23

###################################################################################  
knn\_best <- data.frame(get\_best\_result(knn\_model) %>%   
 dplyr::select(2:4))  
knn\_model$finalModel

## 5-nearest neighbor regression model

###################################################################################  
LF\_best <- data.frame(get\_best\_result(LF\_model) %>%   
 dplyr::select(2:4))  
LF\_model$finalModel

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: forward

###################################################################################  
LB\_best <- data.frame(get\_best\_result(LB\_model) %>%   
 dplyr::select(2:4))  
LB\_model$finalModel

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: backward

###################################################################################  
lmStep\_best <- data.frame(get\_best\_result(lmStepAIC\_Mod) %>%   
 dplyr::select(2:4))  
lmStepAIC\_Mod$finalModel

##   
## Call:  
## lm(formula = .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified, data = dat)  
##   
## Coefficients:  
## (Intercept)   
## 70.323   
## `Juris\_LongBritish Columbia`   
## 36.629   
## Juris\_LongManitoba   
## -33.555   
## `Juris\_LongNational parks`   
## -45.359   
## `Juris\_LongNew Brunswick`   
## -42.727   
## `Juris\_LongNewfoundland and Labrador`   
## -39.546   
## `Juris\_LongNorthwest Territories`   
## -39.320   
## `Juris\_LongNova Scotia`   
## -36.773   
## Juris\_LongOntario   
## -20.816   
## `Juris\_LongPrince Edward Island`   
## -54.492   
## Juris\_LongQuebec   
## -27.948   
## Juris\_LongSaskatchewan   
## -26.614   
## Juris\_LongYukon   
## -41.896   
## Fire\_Cause\_HumanLightning   
## 64.601   
## Fire\_Cause\_HumanUnspecified   
## -7.337   
## Protection.zoneLimited   
## -13.211   
## Protection.zoneUnspecified   
## 60.081   
## Response.categoryModified   
## -42.497   
## Response.categoryNone   
## -42.480   
## Response.categoryUnspecified   
## 18.580

###################################################################################  
total <- rbind(lm\_best, glm\_best, lasso\_best, knn\_best, LF\_best, LB\_best, lmStep\_best)  
total\_best\_train <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 total,  
 Model\_Time)%>%   
 dplyr::select(-5)  
total\_best\_train

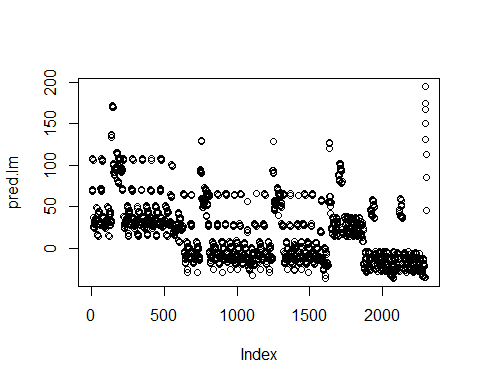
## Name RMSE Rsquared MAE Time\_train  
## 1 lm\_model 80.77566 0.1693413 31.68217 1.9785759 secs  
## 2 glm\_model 80.77566 0.1693413 31.68217 1.6067028 secs  
## 3 lasso\_Mod 80.78302 0.1689831 31.18791 1.8617730 secs  
## 4 knn\_model 76.46193 0.2743102 19.88133 22.7727718 secs  
## 5 LF\_model 81.83715 0.1471318 30.66423 0.9035399 secs  
## 6 LB\_model 82.11647 0.1415281 30.34153 0.8976009 secs  
## 7 lmStepAIC\_Mod 80.79002 0.1690438 31.69457 5.1023571 secs

## 25. Predict on test set

#############################################################################################################################################  
# Model 1: lm model  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.lm = predict(lm\_model, newdata = test.set)  
output <- cbind(test.set, pred.lm)  
head(output)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry AB 11 Intensive Full  
## 2: Forest industry AB 26 Intensive Full  
## 3: Forest industry AB 19 Intensive Full  
## 4: Forest industry AB 41 Intensive Full  
## 5: Forest industry AB 41 Intensive Full  
## 6: Forest industry BC 3 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 2002 Alberta People Early 10s 2000s Prairie Region  
## 2: 2006 Alberta People Late 10s 2000s Prairie Region  
## 3: 2007 Alberta People Late 10s 2000s Prairie Region  
## 4: 2008 Alberta People Late 10s 2000s Prairie Region  
## 5: 2015 Alberta People Early 20s 2010s Prairie Region  
## 6: 1998 British Columbia People Late 90s 1990s Pacific Region  
## Fire\_Cause\_Human pred.lm  
## 1: Human 70.56723  
## 2: Human 69.99401  
## 3: Human 69.85070  
## 4: Human 69.70739  
## 5: Human 68.70425  
## 6: Human 107.51632

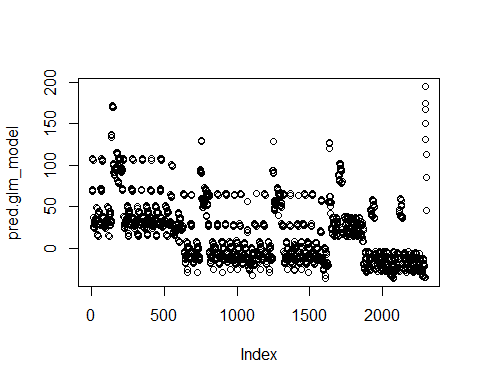
plot(pred.lm)



# Stop the clock  
end.time <- Sys.time()  
  
pre.lm\_model\_time.taken <- end.time - start.time  
#############################################################################################################################################  
# Model 2: glm model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.glm\_model = predict(glm\_model, newdata = test.set)  
output <- cbind(test.set, pred.glm\_model)  
head(output)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry AB 11 Intensive Full  
## 2: Forest industry AB 26 Intensive Full  
## 3: Forest industry AB 19 Intensive Full  
## 4: Forest industry AB 41 Intensive Full  
## 5: Forest industry AB 41 Intensive Full  
## 6: Forest industry BC 3 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 2002 Alberta People Early 10s 2000s Prairie Region  
## 2: 2006 Alberta People Late 10s 2000s Prairie Region  
## 3: 2007 Alberta People Late 10s 2000s Prairie Region  
## 4: 2008 Alberta People Late 10s 2000s Prairie Region  
## 5: 2015 Alberta People Early 20s 2010s Prairie Region  
## 6: 1998 British Columbia People Late 90s 1990s Pacific Region  
## Fire\_Cause\_Human pred.glm\_model  
## 1: Human 70.56723  
## 2: Human 69.99401  
## 3: Human 69.85070  
## 4: Human 69.70739  
## 5: Human 68.70425  
## 6: Human 107.51632

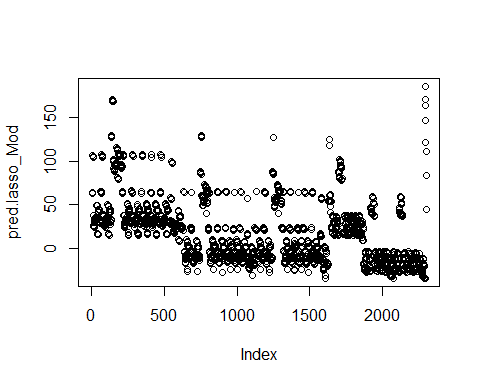
plot(pred.glm\_model)



# Stop the clock  
end.time <- Sys.time()  
  
pre.glm\_model\_time.taken <- end.time - start.time  
#glm\_model\_time.taken  
  
#############################################################################################################################################  
# Model 3: lasso model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.lasso\_Mod = predict(lasso\_Mod, newdata = test.set)  
output <- cbind(test.set, pred.lasso\_Mod)  
head(output)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry AB 11 Intensive Full  
## 2: Forest industry AB 26 Intensive Full  
## 3: Forest industry AB 19 Intensive Full  
## 4: Forest industry AB 41 Intensive Full  
## 5: Forest industry AB 41 Intensive Full  
## 6: Forest industry BC 3 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 2002 Alberta People Early 10s 2000s Prairie Region  
## 2: 2006 Alberta People Late 10s 2000s Prairie Region  
## 3: 2007 Alberta People Late 10s 2000s Prairie Region  
## 4: 2008 Alberta People Late 10s 2000s Prairie Region  
## 5: 2015 Alberta People Early 20s 2010s Prairie Region  
## 6: 1998 British Columbia People Late 90s 1990s Pacific Region  
## Fire\_Cause\_Human pred.lasso\_Mod  
## 1: Human 64.25467  
## 2: Human 63.74066  
## 3: Human 63.61216  
## 4: Human 63.48366  
## 5: Human 62.58415  
## 6: Human 106.51361

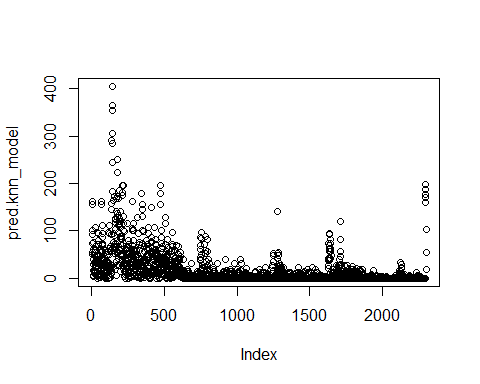
plot(pred.lasso\_Mod)



# Stop the clock  
end.time <- Sys.time()  
  
pre.lasso\_Mod\_time.taken <- end.time - start.time  
#lasso\_Mod\_time.taken  
#############################################################################################################################################  
# Model 4: knn model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.knn\_model = predict(knn\_model, newdata = test.set)  
output <- cbind(test.set, pred.knn\_model)  
head(output)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry AB 11 Intensive Full  
## 2: Forest industry AB 26 Intensive Full  
## 3: Forest industry AB 19 Intensive Full  
## 4: Forest industry AB 41 Intensive Full  
## 5: Forest industry AB 41 Intensive Full  
## 6: Forest industry BC 3 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 2002 Alberta People Early 10s 2000s Prairie Region  
## 2: 2006 Alberta People Late 10s 2000s Prairie Region  
## 3: 2007 Alberta People Late 10s 2000s Prairie Region  
## 4: 2008 Alberta People Late 10s 2000s Prairie Region  
## 5: 2015 Alberta People Early 20s 2010s Prairie Region  
## 6: 1998 British Columbia People Late 90s 1990s Pacific Region  
## Fire\_Cause\_Human pred.knn\_model  
## 1: Human 92.83333  
## 2: Human 53.71591  
## 3: Human 49.75824  
## 4: Human 162.60000  
## 5: Human 154.80000  
## 6: Human 101.08571

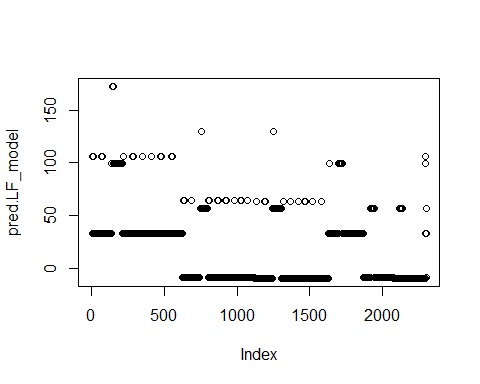
plot(pred.knn\_model)



# Stop the clock  
end.time <- Sys.time()  
  
pre.knn\_model\_time.taken <- end.time - start.time  
#knn\_model\_time.taken  
#############################################################################################################################################  
# Model 5: leapForward model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.LF\_model = predict(LF\_model, newdata = test.set)  
output <- cbind(test.set, pred.LF\_model)  
head(output)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry AB 11 Intensive Full  
## 2: Forest industry AB 26 Intensive Full  
## 3: Forest industry AB 19 Intensive Full  
## 4: Forest industry AB 41 Intensive Full  
## 5: Forest industry AB 41 Intensive Full  
## 6: Forest industry BC 3 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 2002 Alberta People Early 10s 2000s Prairie Region  
## 2: 2006 Alberta People Late 10s 2000s Prairie Region  
## 3: 2007 Alberta People Late 10s 2000s Prairie Region  
## 4: 2008 Alberta People Late 10s 2000s Prairie Region  
## 5: 2015 Alberta People Early 20s 2010s Prairie Region  
## 6: 1998 British Columbia People Late 90s 1990s Pacific Region  
## Fire\_Cause\_Human pred.LF\_model  
## 1: Human 33.46592  
## 2: Human 33.46592  
## 3: Human 33.46592  
## 4: Human 33.46592  
## 5: Human 33.46592  
## 6: Human 106.43868

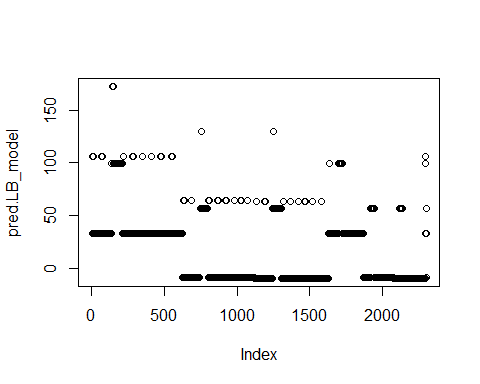
plot(pred.LF\_model)



# Stop the clock  
end.time <- Sys.time()  
  
pre.LF\_model\_time.taken <- end.time - start.time  
#LF\_model\_time.taken  
#############################################################################################################################################  
# Model 6: leapBackward model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.LB\_model = predict(LB\_model, newdata = test.set)  
output <- cbind(test.set, pred.LB\_model)  
head(output)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry AB 11 Intensive Full  
## 2: Forest industry AB 26 Intensive Full  
## 3: Forest industry AB 19 Intensive Full  
## 4: Forest industry AB 41 Intensive Full  
## 5: Forest industry AB 41 Intensive Full  
## 6: Forest industry BC 3 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 2002 Alberta People Early 10s 2000s Prairie Region  
## 2: 2006 Alberta People Late 10s 2000s Prairie Region  
## 3: 2007 Alberta People Late 10s 2000s Prairie Region  
## 4: 2008 Alberta People Late 10s 2000s Prairie Region  
## 5: 2015 Alberta People Early 20s 2010s Prairie Region  
## 6: 1998 British Columbia People Late 90s 1990s Pacific Region  
## Fire\_Cause\_Human pred.LB\_model  
## 1: Human 33.46592  
## 2: Human 33.46592  
## 3: Human 33.46592  
## 4: Human 33.46592  
## 5: Human 33.46592  
## 6: Human 106.43868

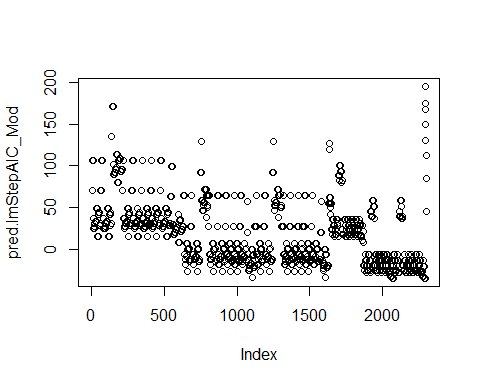
plot(pred.LB\_model)



# Stop the clock  
end.time <- Sys.time()  
  
pre.LB\_model\_time.taken <- end.time - start.time  
#LB\_model\_time.taken  
#############################################################################################################################################  
# Model 7: lmStepAIC model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.lmStepAIC\_Mod = predict(lmStepAIC\_Mod, newdata = test.set)  
output <- cbind(test.set, pred.lmStepAIC\_Mod)  
head(output)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry AB 11 Intensive Full  
## 2: Forest industry AB 26 Intensive Full  
## 3: Forest industry AB 19 Intensive Full  
## 4: Forest industry AB 41 Intensive Full  
## 5: Forest industry AB 41 Intensive Full  
## 6: Forest industry BC 3 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 2002 Alberta People Early 10s 2000s Prairie Region  
## 2: 2006 Alberta People Late 10s 2000s Prairie Region  
## 3: 2007 Alberta People Late 10s 2000s Prairie Region  
## 4: 2008 Alberta People Late 10s 2000s Prairie Region  
## 5: 2015 Alberta People Early 20s 2010s Prairie Region  
## 6: 1998 British Columbia People Late 90s 1990s Pacific Region  
## Fire\_Cause\_Human pred.lmStepAIC\_Mod  
## 1: Human 70.32272  
## 2: Human 70.32272  
## 3: Human 70.32272  
## 4: Human 70.32272  
## 5: Human 70.32272  
## 6: Human 106.95205

plot(pred.lmStepAIC\_Mod)



# Stop the clock  
end.time <- Sys.time()  
  
pre.lmStepAIC\_Mod\_time.taken <- end.time - start.time  
#lmStepAIC\_Mod\_time.taken  
#############################################################################################################################################

## 26. Time taken for prediction

Model\_TimePred <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 Predict\_Time=c(pre.lm\_model\_time.taken, pre.glm\_model\_time.taken, pre.lasso\_Mod\_time.taken, pre.knn\_model\_time.taken, pre.LF\_model\_time.taken, pre.LB\_model\_time.taken, pre.lmStepAIC\_Mod\_time.taken))  
Model\_TimePred

## Name Predict\_Time  
## 1 lm\_model 0.1805122 secs  
## 2 glm\_model 0.1366329 secs  
## 3 lasso\_Mod 0.1117060 secs  
## 4 knn\_model 0.7819109 secs  
## 5 LF\_model 0.1077108 secs  
## 6 LB\_model 0.1416218 secs  
## 7 lmStepAIC\_Mod 0.2054460 secs

## 27. Compare correlation between actual and predicted

#############################################################################################################################################  
# Formula to calculate correlation   
corr\_lm\_model <- round(cor(test.set$Number, predict(lm\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_glm\_model <- round(cor(test.set$Number, predict(glm\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_lasso\_Mod <- round(cor(test.set$Number, predict(lasso\_Mod, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_knn\_model <- round(cor(test.set$Number, predict(knn\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_LF\_model <- round(cor(test.set$Number, predict(LF\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_LB\_model <- round(cor(test.set$Number, predict(LB\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_lmStepAIC\_Mod <- round(cor(test.set$Number, predict(lmStepAIC\_Mod, test.set), method = c("pearson", "kendall", "spearman")),2)  
  
total\_corr <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 corr\_pred=c(corr\_lm\_model, corr\_glm\_model, corr\_lasso\_Mod, corr\_knn\_model, corr\_LF\_model, corr\_LB\_model, corr\_lmStepAIC\_Mod))  
total\_corr

## Name corr\_pred  
## 1 lm\_model 0.47  
## 2 glm\_model 0.47  
## 3 lasso\_Mod 0.47  
## 4 knn\_model 0.69  
## 5 LF\_model 0.41  
## 6 LB\_model 0.41  
## 7 lmStepAIC\_Mod 0.47

## 27. RMSE between actual and predicted

#############################################################################################################################################  
# Formula to calculate RMSE on test set  
calc\_rmse = function(actual, predicted) {  
 sqrt(mean((actual - predicted) ^ 2))  
}  
  
# RMSE value on test set  
rmse\_lm\_model <- calc\_rmse(actual = test.set$Number,  
 predicted = predict(lm\_model, test.set))  
  
# RMSE value on test set  
rmse\_glm\_model <- calc\_rmse(actual = test.set$Number,  
 predicted = predict(glm\_model, test.set))  
  
# RMSE value on test set  
rmse\_lasso\_Mod <- calc\_rmse(actual = test.set$Number,  
 predicted = predict(lasso\_Mod, test.set))  
  
# RMSE value on test set  
rmse\_knn\_model <- calc\_rmse(actual = test.set$Number,  
 predicted = predict(knn\_model, test.set))  
  
# RMSE value on test set  
rmse\_LF\_model <- calc\_rmse(actual = test.set$Number,  
 predicted = predict(LF\_model, test.set))  
  
# RMSE value on test set  
rmse\_LB\_model <- calc\_rmse(actual = test.set$Number,  
 predicted = predict(LB\_model, test.set))  
  
# RMSE value on test set  
rmse\_lmStepAIC\_Mod <- calc\_rmse(actual = test.set$Number,  
 predicted = predict(lmStepAIC\_Mod, test.set))  
  
total\_RMSE <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 RMSE\_Pred=c(rmse\_lm\_model, rmse\_glm\_model, rmse\_lasso\_Mod, rmse\_knn\_model, rmse\_LF\_model, rmse\_LB\_model, rmse\_lmStepAIC\_Mod))  
total\_RMSE

## Name RMSE\_Pred  
## 1 lm\_model 74.75909  
## 2 glm\_model 74.75909  
## 3 lasso\_Mod 74.78981  
## 4 knn\_model 65.58049  
## 5 LF\_model 76.98764  
## 6 LB\_model 76.98764  
## 7 lmStepAIC\_Mod 74.75997

## 28. MAE between actual and predicted

#############################################################################################################################################  
  
MAE\_test\_lm\_model <- MAE(test.set$Number, predict(lm\_model, test.set))  
MAE\_test\_glm\_model <- MAE(test.set$Number, predict(glm\_model, test.set))  
MAE\_test\_lasso\_Mod <- MAE(test.set$Number, predict(lasso\_Mod, test.set))  
MAE\_test\_knn\_model <- MAE(test.set$Number, predict(knn\_model, test.set))  
MAE\_test\_LF\_model <- MAE(test.set$Number, predict(LF\_model, test.set))  
MAE\_test\_LB\_model <- MAE(test.set$Number, predict(LB\_model, test.set))  
MAE\_test\_lmStepAIC\_Mod <- MAE(test.set$Number, predict(lmStepAIC\_Mod, test.set))  
  
total\_MAE <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 MAE\_Pred=c(MAE\_test\_lm\_model, MAE\_test\_glm\_model, MAE\_test\_lasso\_Mod, MAE\_test\_knn\_model, MAE\_test\_LF\_model, MAE\_test\_LB\_model, MAE\_test\_lmStepAIC\_Mod))  
total\_MAE

## Name MAE\_Pred  
## 1 lm\_model 33.32379  
## 2 glm\_model 33.32379  
## 3 lasso\_Mod 32.78662  
## 4 knn\_model 18.63506  
## 5 LF\_model 32.49316  
## 6 LB\_model 32.49316  
## 7 lmStepAIC\_Mod 33.33735

## 29. r squared between actual and predicted

calc\_rss = function(actual, predicted) {  
 sum((predicted - actual) ^ 2) ## residual sum of squares  
}  
  
calc\_tss = function(actual, predicted) {  
 sum((actual - mean(actual)) ^ 2) ## total sum of squares  
}  
  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$Number,  
 predicted = predict(lm\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$Number,  
 predicted = predict(lm\_model, test.set))  
  
rsq\_lm\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$Number,  
 predicted = predict(glm\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$Number,  
 predicted = predict(glm\_model, test.set))  
  
rsq\_glm\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$Number,  
 predicted = predict(lasso\_Mod, test.set))  
  
caltss <- calc\_tss(actual = test.set$Number,  
 predicted = predict(lasso\_Mod, test.set))  
  
rsq\_lasso\_Mod <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$Number,  
 predicted = predict(knn\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$Number,  
 predicted = predict(knn\_model, test.set))  
  
rsq\_knn\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$Number,  
 predicted = predict(LF\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$Number,  
 predicted = predict(LF\_model, test.set))  
  
rsq\_LF\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$Number,  
 predicted = predict(LB\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$Number,  
 predicted = predict(LB\_model, test.set))  
  
rsq\_LB\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$Number,  
 predicted = predict(lmStepAIC\_Mod, test.set))  
  
caltss <- calc\_tss(actual = test.set$Number,  
 predicted = predict(lmStepAIC\_Mod, test.set))  
  
rsq\_lmStepAIC\_Mod <- 1 - calrss/caltss  
#######################################################  
  
total\_rsq <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 R\_squared\_Pred=c(rsq\_lm\_model, rsq\_glm\_model, rsq\_lasso\_Mod, rsq\_knn\_model, rsq\_LF\_model, rsq\_LB\_model, rsq\_lmStepAIC\_Mod))  
total\_rsq

## Name R\_squared\_Pred  
## 1 lm\_model 0.2161545  
## 2 glm\_model 0.2161545  
## 3 lasso\_Mod 0.2155102  
## 4 knn\_model 0.3968132  
## 5 LF\_model 0.1687255  
## 6 LB\_model 0.1687255  
## 7 lmStepAIC\_Mod 0.2161360

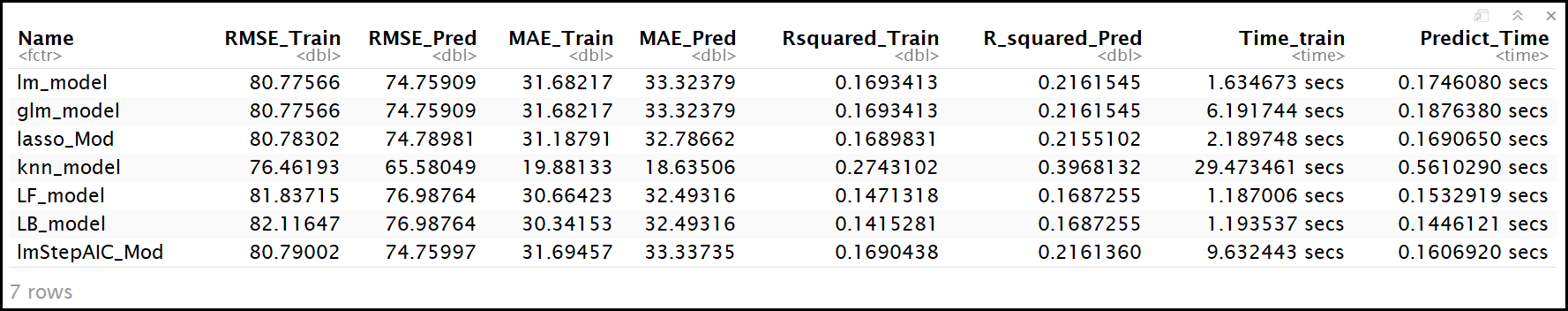
## 30. Combine predicted RMSE, MAE R squared, time

total\_combALLpred <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 Model\_TimePred,  
 total\_RMSE,  
 total\_MAE,  
 total\_rsq) %>%   
 dplyr::select(-2,-4,-6,-8)  
total\_combALLpred

## Name Predict\_Time RMSE\_Pred MAE\_Pred R\_squared\_Pred  
## 1 lm\_model 0.1805122 secs 74.75909 33.32379 0.2161545  
## 2 glm\_model 0.1366329 secs 74.75909 33.32379 0.2161545  
## 3 lasso\_Mod 0.1117060 secs 74.78981 32.78662 0.2155102  
## 4 knn\_model 0.7819109 secs 65.58049 18.63506 0.3968132  
## 5 LF\_model 0.1077108 secs 76.98764 32.49316 0.1687255  
## 6 LB\_model 0.1416218 secs 76.98764 32.49316 0.1687255  
## 7 lmStepAIC\_Mod 0.2054460 secs 74.75997 33.33735 0.2161360

## 31. Compare the two sets of RMSE, MAE, r squared, time

total\_setscomp <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 total\_combALLpred,  
 total\_best\_train) %>%   
 dplyr::select(-2,-7)  
# Reorder columns  
total\_setscomp2 <- total\_setscomp[,c(1,6,3,8,4,7,5,9,2)]  
  
# Change name of column in data.frame  
total\_setscomp2 %>%   
 rename(  
 RMSE\_Train = RMSE,  
 MAE\_Train = MAE,  
 Rsquared\_Train = Rsquared  
 )

## Name RMSE\_Train RMSE\_Pred MAE\_Train MAE\_Pred Rsquared\_Train  
## 1 lm\_model 80.77566 74.75909 31.68217 33.32379 0.1693413  
## 2 glm\_model 80.77566 74.75909 31.68217 33.32379 0.1693413  
## 3 lasso\_Mod 80.78302 74.78981 31.18791 32.78662 0.1689831  
## 4 knn\_model 76.46193 65.58049 19.88133 18.63506 0.2743102  
## 5 LF\_model 81.83715 76.98764 30.66423 32.49316 0.1471318  
## 6 LB\_model 82.11647 76.98764 30.34153 32.49316 0.1415281  
## 7 lmStepAIC\_Mod 80.79002 74.75997 31.69457 33.33735 0.1690438  
## R\_squared\_Pred Time\_train Predict\_Time  
## 1 0.2161545 1.9785759 secs 0.1805122 secs  
## 2 0.2161545 1.6067028 secs 0.1366329 secs  
## 3 0.2155102 1.8617730 secs 0.1117060 secs  
## 4 0.3968132 22.7727718 secs 0.7819109 secs  
## 5 0.1687255 0.9035399 secs 0.1077108 secs  
## 6 0.1687255 0.8976009 secs 0.1416218 secs  
## 7 0.2161360 5.1023571 secs 0.2054460 secs

## 30. RMSE, Rsquared MAE on training set

# RMSE, Rsquared MAE on training set  
lm\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 80.78 0.17 31.68

# RMSE, Rsquared MAE on training set  
glm\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 80.78 0.17 31.68

# RMSE, Rsquared MAE on training set  
lasso\_Mod$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 85.29 0.10 27.01  
## 2 81.22 0.16 28.10  
## 3 80.78 0.17 31.19

# RMSE, Rsquared MAE on training set  
knn\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 76.46 0.27 19.88  
## 2 77.90 0.25 20.72  
## 3 79.47 0.21 21.16

# RMSE, Rsquared MAE on training set  
LF\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 84.40 0.09 26.35  
## 2 83.69 0.11 29.48  
## 3 81.84 0.15 30.66

# RMSE, Rsquared MAE on training set  
LB\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 85.12 0.08 27.81  
## 2 83.74 0.11 28.73  
## 3 82.12 0.14 30.34

# RMSE, Rsquared MAE on training set  
lmStepAIC\_Mod$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 80.79 0.17 31.69

# 31.Model below is separate from models above

## Additional for testing: Decision Tree

summary(data)

## Cause Jurisdiction Number   
## Lightning :1325 QC :1516 Min. : 0.00   
## Unspecified :1290 NL :1486 1st Qu.: 0.00   
## Miscellaneous known causes:1279 ON :1368 Median : 0.00   
## Recreation :1274 MB :1304 Mean : 18.64   
## Incendiary :1271 YT :1059 3rd Qu.: 4.00   
## Residents :1268 NT : 773 Max. :2913.00   
## (Other) :3812 (Other):4013   
## Protection.zone Response.category Year   
## Intensive :8190 Full :4204 Min. :1990   
## Limited :3281 Modified :3521 1st Qu.:1997   
## Unspecified: 48 None :3681 Median :2004   
## Unspecified: 113 Mean :2004   
## 3rd Qu.:2011   
## Max. :2018   
##   
## Juris\_Long Cause\_Grouped Time1   
## Quebec :1516 Length:11519 Length:11519   
## Newfoundland and Labrador:1486 Class :character Class :character   
## Ontario :1368 Mode :character Mode :character   
## Manitoba :1304   
## Yukon :1059   
## Northwest Territories : 773   
## (Other) :4013   
## Time2 Region Fire\_Cause\_Human   
## Length:11519 Length:11519 Length:11519   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##

class(data$Fire\_Cause\_Human) # [1] "factor"

## [1] "character"

unique(data$Fire\_Cause\_Human)

## [1] "Human" "Lightning" "Unspecified"

data$Fire\_Cause\_Human <- as.factor(data$Fire\_Cause\_Human) #Change to factor  
class(data$Fire\_Cause\_Human)

## [1] "factor"

unique(data$Fire\_Cause\_Human)

## [1] Human Lightning Unspecified  
## Levels: Human Lightning Unspecified

names(data)

## [1] "Cause" "Jurisdiction" "Number"   
## [4] "Protection.zone" "Response.category" "Year"   
## [7] "Juris\_Long" "Cause\_Grouped" "Time1"   
## [10] "Time2" "Region" "Fire\_Cause\_Human"

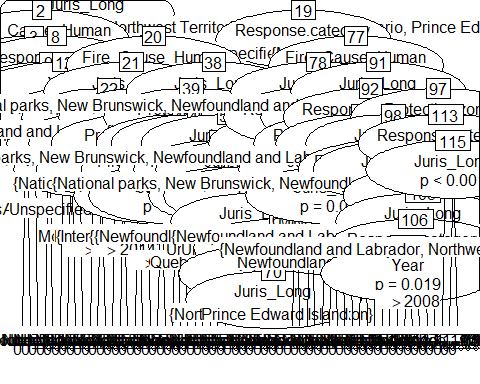
str(data)

## Classes 'data.table' and 'data.frame': 11519 obs. of 12 variables:  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "AB","BC","MB",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Number : int 22 14 12 11 13 14 8 29 10 20 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Cause\_Grouped : chr "People" "People" "People" "People" ...  
## $ Time1 : chr "Early 90s" "Early 90s" "Early 90s" "Early 90s" ...  
## $ Time2 : chr "1990s" "1990s" "1990s" "1990s" ...  
## $ Region : chr "Prairie Region" "Prairie Region" "Prairie Region" "Prairie Region" ...  
## $ Fire\_Cause\_Human : Factor w/ 3 levels "Human","Lightning",..: 1 1 1 1 1 1 1 1 1 1 ...  
## - attr(\*, ".internal.selfref")=<externalptr>

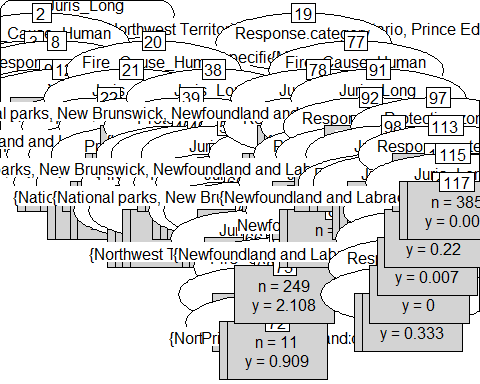
# Partition data into Training and Validation datasets  
set.seed(1235)  
pd <- sample(2,nrow(data),replace = TRUE, prob = c(0.7,0.3))  
train <- data[pd==1,]  
validate <- data[pd==2,]  
  
# Decision Tree with party  
tree <- ctree(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone +   
 Response.category, data = train, controls = ctree\_control(mincriterion = 0.99, minsplit=900))  
  
tree <- ctree(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train)  
tree

##   
## Conditional inference tree with 59 terminal nodes  
##   
## Response: Number   
## Inputs: Juris\_Long, Year, Fire\_Cause\_Human, Protection.zone, Response.category   
## Number of observations: 7986   
##   
## 1) Juris\_Long == {Alberta, British Columbia}; criterion = 1, statistic = 472.214  
## 2) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 196.933  
## 3) Response.category == {Full}; criterion = 1, statistic = 61.204  
## 4) Juris\_Long == {British Columbia}; criterion = 0.983, statistic = 8.537  
## 5)\* weights = 24   
## 4) Juris\_Long == {Alberta}  
## 6)\* weights = 21   
## 3) Response.category == {Modified, None}  
## 7)\* weights = 65   
## 2) Fire\_Cause\_Human == {Human, Unspecified}  
## 8) Response.category == {Full}; criterion = 1, statistic = 292.308  
## 9) Fire\_Cause\_Human == {Human}; criterion = 0.996, statistic = 42.233  
## 10)\* weights = 275   
## 9) Fire\_Cause\_Human == {Unspecified}  
## 11)\* weights = 38   
## 8) Response.category == {Modified, None}  
## 12) Year <= 2012; criterion = 1, statistic = 17.345  
## 13)\* weights = 469   
## 12) Year > 2012  
## 14) Juris\_Long == {British Columbia}; criterion = 1, statistic = 21.522  
## 15) Response.category == {Modified}; criterion = 0.986, statistic = 8.893  
## 16)\* weights = 13   
## 15) Response.category == {None}  
## 17)\* weights = 10   
## 14) Juris\_Long == {Alberta}  
## 18)\* weights = 57   
## 1) Juris\_Long == {Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Ontario, Prince Edward Island, Quebec, Saskatchewan, Yukon}  
## 19) Response.category == {Full, Unspecified}; criterion = 1, statistic = 610.356  
## 20) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 251.098  
## 21) Juris\_Long == {Manitoba, Northwest Territories, Ontario, Quebec, Saskatchewan}; criterion = 1, statistic = 66.847  
## 22) Protection.zone == {Intensive}; criterion = 1, statistic = 36.788  
## 23) Juris\_Long == {Ontario}; criterion = 0.98, statistic = 15.378  
## 24)\* weights = 20   
## 23) Juris\_Long == {Manitoba, Northwest Territories, Quebec, Saskatchewan}  
## 25) Year <= 2006; criterion = 0.999, statistic = 13.762  
## 26)\* weights = 48   
## 25) Year > 2006  
## 27)\* weights = 28   
## 22) Protection.zone == {Limited, Unspecified}  
## 28) Juris\_Long == {Ontario}; criterion = 0.964, statistic = 13.966  
## 29)\* weights = 24   
## 28) Juris\_Long == {Manitoba, Northwest Territories, Quebec, Saskatchewan}  
## 30) Year <= 2014; criterion = 0.95, statistic = 29.574  
## 31)\* weights = 32   
## 30) Year > 2014  
## 32)\* weights = 8   
## 21) Juris\_Long == {National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Yukon}  
## 33) Juris\_Long == {National parks, New Brunswick, Yukon}; criterion = 1, statistic = 36.341  
## 34) Protection.zone == {Intensive, Unspecified}; criterion = 0.995, statistic = 13.988  
## 35)\* weights = 61   
## 34) Protection.zone == {Limited}  
## 36)\* weights = 8   
## 33) Juris\_Long == {Newfoundland and Labrador, Nova Scotia, Prince Edward Island}  
## 37)\* weights = 70   
## 20) Fire\_Cause\_Human == {Human, Unspecified}  
## 38) Juris\_Long == {New Brunswick, Nova Scotia, Ontario, Quebec, Saskatchewan}; criterion = 1, statistic = 290.184  
## 39) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 202.195  
## 40) Juris\_Long == {Ontario, Quebec}; criterion = 1, statistic = 50.044  
## 41) Fire\_Cause\_Human == {Human}; criterion = 1, statistic = 36.648  
## 42) Year <= 2003; criterion = 0.998, statistic = 33.772  
## 43)\* weights = 146   
## 42) Year > 2003  
## 44)\* weights = 147   
## 41) Fire\_Cause\_Human == {Unspecified}  
## 45) Juris\_Long == {Ontario}; criterion = 1, statistic = 24.137  
## 46)\* weights = 13   
## 45) Juris\_Long == {Quebec}  
## 47)\* weights = 22   
## 40) Juris\_Long == {New Brunswick, Nova Scotia, Saskatchewan}  
## 48)\* weights = 464   
## 39) Protection.zone == {Limited}  
## 49) Juris\_Long == {Ontario}; criterion = 1, statistic = 42.246  
## 50)\* weights = 147   
## 49) Juris\_Long == {Quebec, Saskatchewan}  
## 51) Fire\_Cause\_Human == {Human}; criterion = 0.969, statistic = 7.465  
## 52)\* weights = 149   
## 51) Fire\_Cause\_Human == {Unspecified}  
## 53)\* weights = 19   
## 38) Juris\_Long == {Manitoba, National parks, Newfoundland and Labrador, Northwest Territories, Prince Edward Island, Yukon}  
## 54) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 139.117  
## 55) Juris\_Long == {Manitoba}; criterion = 1, statistic = 139.607  
## 56) Fire\_Cause\_Human == {Human}; criterion = 0.977, statistic = 24.222  
## 57)\* weights = 135   
## 56) Fire\_Cause\_Human == {Unspecified}  
## 58)\* weights = 17   
## 55) Juris\_Long == {National parks, Newfoundland and Labrador, Northwest Territories, Prince Edward Island, Yukon}  
## 59) Juris\_Long == {Newfoundland and Labrador}; criterion = 1, statistic = 64.472  
## 60) Fire\_Cause\_Human == {Human}; criterion = 0.994, statistic = 22.477  
## 61)\* weights = 136   
## 60) Fire\_Cause\_Human == {Unspecified}  
## 62)\* weights = 23   
## 59) Juris\_Long == {National parks, Northwest Territories, Prince Edward Island, Yukon}  
## 63) Year <= 1998; criterion = 0.999, statistic = 14.263  
## 64)\* weights = 189   
## 63) Year > 1998  
## 65) Juris\_Long == {National parks}; criterion = 1, statistic = 24.797  
## 66)\* weights = 101   
## 65) Juris\_Long == {Northwest Territories, Prince Edward Island, Yukon}  
## 67) Fire\_Cause\_Human == {Unspecified}; criterion = 0.98, statistic = 8.247  
## 68) Year <= 2002; criterion = 0.982, statistic = 10.469  
## 69)\* weights = 8   
## 68) Year > 2002  
## 70) Juris\_Long == {Northwest Territories, Yukon}; criterion = 0.993, statistic = 13.022  
## 71)\* weights = 16   
## 70) Juris\_Long == {Prince Edward Island}  
## 72)\* weights = 11   
## 67) Fire\_Cause\_Human == {Human}  
## 73)\* weights = 249   
## 54) Protection.zone == {Limited}  
## 74) Juris\_Long == {Manitoba}; criterion = 1, statistic = 76.376  
## 75)\* weights = 127   
## 74) Juris\_Long == {Newfoundland and Labrador, Northwest Territories, Yukon}  
## 76)\* weights = 265   
## 19) Response.category == {Modified, None}  
## 77) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 698.129  
## 78) Juris\_Long == {Northwest Territories, Saskatchewan}; criterion = 1, statistic = 87.131  
## 79) Response.category == {Modified}; criterion = 1, statistic = 43.768  
## 80)\* weights = 34   
## 79) Response.category == {None}  
## 81)\* weights = 37   
## 78) Juris\_Long == {Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 82) Protection.zone == {Limited, Unspecified}; criterion = 1, statistic = 48.536  
## 83) Response.category == {None}; criterion = 1, statistic = 42.523  
## 84) Juris\_Long == {Ontario}; criterion = 1, statistic = 34.329  
## 85)\* weights = 16   
## 84) Juris\_Long == {Manitoba, Newfoundland and Labrador, Quebec, Yukon}  
## 86) Juris\_Long == {Manitoba, Quebec, Yukon}; criterion = 0.992, statistic = 15.33  
## 87)\* weights = 47   
## 86) Juris\_Long == {Newfoundland and Labrador}  
## 88)\* weights = 21   
## 83) Response.category == {Modified}  
## 89)\* weights = 94   
## 82) Protection.zone == {Intensive}  
## 90)\* weights = 263   
## 77) Fire\_Cause\_Human == {Human, Unspecified}  
## 91) Juris\_Long == {Saskatchewan}; criterion = 1, statistic = 280.785  
## 92) Response.category == {Modified}; criterion = 1, statistic = 22.967  
## 93) Year <= 2015; criterion = 0.997, statistic = 11.869  
## 94)\* weights = 81   
## 93) Year > 2015  
## 95)\* weights = 11   
## 92) Response.category == {None}  
## 96)\* weights = 105   
## 91) Juris\_Long == {Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 97) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 359.436  
## 98) Juris\_Long == {Manitoba, Ontario}; criterion = 1, statistic = 323.403  
## 99) Fire\_Cause\_Human == {Human}; criterion = 0.999, statistic = 118.705  
## 100)\* weights = 465   
## 99) Fire\_Cause\_Human == {Unspecified}  
## 101)\* weights = 73   
## 98) Juris\_Long == {National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Prince Edward Island, Quebec, Yukon}  
## 102) Juris\_Long == {Northwest Territories, Quebec}; criterion = 1, statistic = 213.276  
## 103)\* weights = 593   
## 102) Juris\_Long == {National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Yukon}  
## 104) Year <= 2016; criterion = 0.982, statistic = 420.609  
## 105) Juris\_Long == {National parks}; criterion = 1, statistic = 30.174  
## 106) Response.category == {Modified}; criterion = 0.992, statistic = 10.069  
## 107) Year <= 2008; criterion = 0.981, statistic = 8.351  
## 108)\* weights = 82   
## 107) Year > 2008  
## 109)\* weights = 39   
## 106) Response.category == {None}  
## 110)\* weights = 109   
## 105) Juris\_Long == {Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Yukon}  
## 111)\* weights = 914   
## 104) Year > 2016  
## 112)\* weights = 59   
## 97) Protection.zone == {Limited}  
## 113) Response.category == {None}; criterion = 1, statistic = 25.025  
## 114)\* weights = 652   
## 113) Response.category == {Modified}  
## 115) Juris\_Long == {Manitoba, Ontario}; criterion = 1, statistic = 31.605  
## 116)\* weights = 251   
## 115) Juris\_Long == {Newfoundland and Labrador, Northwest Territories, Quebec, Yukon}  
## 117)\* weights = 385

plot(tree)



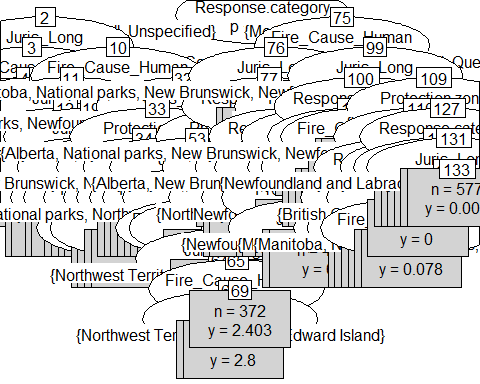
plot(tree, type="simple")



fire\_ctree <- ctree(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = data)  
  
print(fire\_ctree)

##   
## Conditional inference tree with 67 terminal nodes  
##   
## Response: Number   
## Inputs: Juris\_Long, Year, Fire\_Cause\_Human, Protection.zone, Response.category   
## Number of observations: 11519   
##   
## 1) Response.category == {Full, Unspecified}; criterion = 1, statistic = 688.349  
## 2) Juris\_Long == {Alberta, British Columbia}; criterion = 1, statistic = 665.926  
## 3) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 281.612  
## 4) Juris\_Long == {British Columbia}; criterion = 0.999, statistic = 13.588  
## 5)\* weights = 29   
## 4) Juris\_Long == {Alberta}  
## 6)\* weights = 29   
## 3) Fire\_Cause\_Human == {Human, Unspecified}  
## 7) Fire\_Cause\_Human == {Human}; criterion = 1, statistic = 53.383  
## 8)\* weights = 394   
## 7) Fire\_Cause\_Human == {Unspecified}  
## 9)\* weights = 56   
## 2) Juris\_Long == {Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Ontario, Prince Edward Island, Quebec, Saskatchewan, Yukon}  
## 10) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 356.913  
## 11) Juris\_Long == {Ontario, Saskatchewan}; criterion = 1, statistic = 113.941  
## 12) Protection.zone == {Intensive, Unspecified}; criterion = 0.999, statistic = 18.185  
## 13) Juris\_Long == {Ontario}; criterion = 0.987, statistic = 9.036  
## 14)\* weights = 29   
## 13) Juris\_Long == {Saskatchewan}  
## 15) Year <= 2003; criterion = 0.981, statistic = 8.341  
## 16)\* weights = 14   
## 15) Year > 2003  
## 17)\* weights = 15   
## 12) Protection.zone == {Limited}  
## 18)\* weights = 29   
## 11) Juris\_Long == {Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Prince Edward Island, Quebec, Yukon}  
## 19) Juris\_Long == {Manitoba, Northwest Territories, Quebec}; criterion = 1, statistic = 80.329  
## 20) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 46.103  
## 21) Year <= 2007; criterion = 0.998, statistic = 13.069  
## 22)\* weights = 54   
## 21) Year > 2007  
## 23)\* weights = 33   
## 20) Protection.zone == {Limited}  
## 24)\* weights = 51   
## 19) Juris\_Long == {National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Yukon}  
## 25) Juris\_Long == {National parks, New Brunswick, Yukon}; criterion = 1, statistic = 51.811  
## 26) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 23.643  
## 27)\* weights = 83   
## 26) Protection.zone == {Limited}  
## 28)\* weights = 16   
## 25) Juris\_Long == {Newfoundland and Labrador, Nova Scotia, Prince Edward Island}  
## 29) Juris\_Long == {Newfoundland and Labrador, Nova Scotia}; criterion = 0.996, statistic = 14.45  
## 30)\* weights = 86   
## 29) Juris\_Long == {Prince Edward Island}  
## 31)\* weights = 25   
## 10) Fire\_Cause\_Human == {Human, Unspecified}  
## 32) Juris\_Long == {New Brunswick, Nova Scotia, Ontario, Quebec, Saskatchewan}; criterion = 1, statistic = 403.349  
## 33) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 257.364  
## 34) Juris\_Long == {Ontario, Quebec}; criterion = 1, statistic = 51.242  
## 35) Fire\_Cause\_Human == {Human}; criterion = 1, statistic = 37.708  
## 36) Year <= 2001; criterion = 1, statistic = 34.011  
## 37)\* weights = 168   
## 36) Year > 2001  
## 38)\* weights = 226   
## 35) Fire\_Cause\_Human == {Unspecified}  
## 39) Juris\_Long == {Ontario}; criterion = 1, statistic = 39.384  
## 40)\* weights = 28   
## 39) Juris\_Long == {Quebec}  
## 41)\* weights = 28   
## 34) Juris\_Long == {New Brunswick, Nova Scotia, Saskatchewan}  
## 42) Year <= 2002; criterion = 1, statistic = 17.017  
## 43)\* weights = 318   
## 42) Year > 2002  
## 44)\* weights = 350   
## 33) Protection.zone == {Limited}  
## 45) Juris\_Long == {Ontario}; criterion = 1, statistic = 58.864  
## 46) Fire\_Cause\_Human == {Human}; criterion = 0.957, statistic = 6.859  
## 47)\* weights = 181   
## 46) Fire\_Cause\_Human == {Unspecified}  
## 48)\* weights = 27   
## 45) Juris\_Long == {Quebec, Saskatchewan}  
## 49) Fire\_Cause\_Human == {Human}; criterion = 0.996, statistic = 11.431  
## 50)\* weights = 203   
## 49) Fire\_Cause\_Human == {Unspecified}  
## 51)\* weights = 29   
## 32) Juris\_Long == {Manitoba, National parks, Newfoundland and Labrador, Northwest Territories, Prince Edward Island, Yukon}  
## 52) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 164.258  
## 53) Juris\_Long == {Manitoba}; criterion = 1, statistic = 216.552  
## 54) Fire\_Cause\_Human == {Human}; criterion = 0.999, statistic = 26.419  
## 55)\* weights = 197   
## 54) Fire\_Cause\_Human == {Unspecified}  
## 56)\* weights = 27   
## 53) Juris\_Long == {National parks, Newfoundland and Labrador, Northwest Territories, Prince Edward Island, Yukon}  
## 57) Juris\_Long == {Newfoundland and Labrador}; criterion = 1, statistic = 101.806  
## 58) Fire\_Cause\_Human == {Human}; criterion = 0.998, statistic = 23.186  
## 59)\* weights = 197   
## 58) Fire\_Cause\_Human == {Unspecified}  
## 60)\* weights = 28   
## 57) Juris\_Long == {National parks, Northwest Territories, Prince Edward Island, Yukon}  
## 61) Year <= 1998; criterion = 1, statistic = 18.349  
## 62)\* weights = 267   
## 61) Year > 1998  
## 63) Juris\_Long == {National parks}; criterion = 1, statistic = 26.485  
## 64)\* weights = 137   
## 63) Juris\_Long == {Northwest Territories, Prince Edward Island, Yukon}  
## 65) Fire\_Cause\_Human == {Unspecified}; criterion = 0.996, statistic = 33.371  
## 66) Juris\_Long == {Yukon}; criterion = 0.997, statistic = 14.959  
## 67)\* weights = 19   
## 66) Juris\_Long == {Northwest Territories, Prince Edward Island}  
## 68)\* weights = 35   
## 65) Fire\_Cause\_Human == {Human}  
## 69)\* weights = 372   
## 52) Protection.zone == {Limited}  
## 70) Juris\_Long == {Manitoba}; criterion = 1, statistic = 89.964  
## 71)\* weights = 177   
## 70) Juris\_Long == {Newfoundland and Labrador, Northwest Territories, Yukon}  
## 72) Juris\_Long == {Newfoundland and Labrador}; criterion = 0.971, statistic = 10.271  
## 73)\* weights = 224   
## 72) Juris\_Long == {Northwest Territories, Yukon}  
## 74)\* weights = 136   
## 1) Response.category == {Modified, None}  
## 75) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 1031.941  
## 76) Juris\_Long == {Northwest Territories, Saskatchewan}; criterion = 1, statistic = 132.153  
## 77) Response.category == {Modified}; criterion = 1, statistic = 57.155  
## 78)\* weights = 47   
## 77) Response.category == {None}  
## 79)\* weights = 50   
## 76) Juris\_Long == {Alberta, British Columbia, Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 80) Protection.zone == {Limited, Unspecified}; criterion = 1, statistic = 102.629  
## 81) Response.category == {None}; criterion = 1, statistic = 66.737  
## 82) Juris\_Long == {Ontario, Yukon}; criterion = 1, statistic = 34.416  
## 83)\* weights = 50   
## 82) Juris\_Long == {Manitoba, Newfoundland and Labrador, Quebec}  
## 84) Juris\_Long == {Manitoba, Quebec}; criterion = 1, statistic = 19.943  
## 85)\* weights = 50   
## 84) Juris\_Long == {Newfoundland and Labrador}  
## 86)\* weights = 28   
## 81) Response.category == {Modified}  
## 87) Juris\_Long == {British Columbia, Manitoba, National parks}; criterion = 1, statistic = 83.653  
## 88)\* weights = 25   
## 87) Juris\_Long == {Alberta, New Brunswick, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 89) Protection.zone == {Unspecified}; criterion = 1, statistic = 19.388  
## 90)\* weights = 8   
## 89) Protection.zone == {Limited}  
## 91) Juris\_Long == {Newfoundland and Labrador, Ontario}; criterion = 1, statistic = 21.72  
## 92)\* weights = 55   
## 91) Juris\_Long == {Quebec, Yukon}  
## 93)\* weights = 44   
## 80) Protection.zone == {Intensive}  
## 94) Juris\_Long == {British Columbia, Manitoba, National parks, Ontario, Quebec, Yukon}; criterion = 1, statistic = 37.154  
## 95)\* weights = 290   
## 94) Juris\_Long == {Alberta, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island}  
## 96) Juris\_Long == {Newfoundland and Labrador}; criterion = 1, statistic = 37.031  
## 97)\* weights = 55   
## 96) Juris\_Long == {Alberta, New Brunswick, Nova Scotia, Prince Edward Island}  
## 98)\* weights = 130   
## 75) Fire\_Cause\_Human == {Human, Unspecified}  
## 99) Juris\_Long == {Saskatchewan}; criterion = 1, statistic = 413.701  
## 100) Response.category == {Modified}; criterion = 1, statistic = 40.914  
## 101) Year <= 2014; criterion = 1, statistic = 16.136  
## 102) Year <= 1997; criterion = 0.958, statistic = 6.924  
## 103)\* weights = 30   
## 102) Year > 1997  
## 104)\* weights = 88   
## 101) Year > 2014  
## 105)\* weights = 19   
## 100) Response.category == {None}  
## 106) Fire\_Cause\_Human == {Unspecified}; criterion = 0.955, statistic = 6.8  
## 107)\* weights = 20   
## 106) Fire\_Cause\_Human == {Human}  
## 108)\* weights = 136   
## 99) Juris\_Long == {Alberta, British Columbia, Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 109) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 630.182  
## 110) Juris\_Long == {British Columbia, Manitoba, Northwest Territories, Ontario}; criterion = 1, statistic = 484.52  
## 111) Response.category == {Modified}; criterion = 1, statistic = 248.662  
## 112) Year <= 2012; criterion = 1, statistic = 130.797  
## 113) Juris\_Long == {Manitoba, Northwest Territories, Ontario}; criterion = 1, statistic = 29.379  
## 114)\* weights = 475   
## 113) Juris\_Long == {British Columbia}  
## 115)\* weights = 152   
## 112) Year > 2012  
## 116) Juris\_Long == {British Columbia}; criterion = 1, statistic = 45.072  
## 117)\* weights = 17   
## 116) Juris\_Long == {Manitoba, Northwest Territories, Ontario}  
## 118)\* weights = 98   
## 111) Response.category == {None}  
## 119) Juris\_Long == {Manitoba, Ontario}; criterion = 1, statistic = 86.878  
## 120) Fire\_Cause\_Human == {Human}; criterion = 0.989, statistic = 9.336  
## 121)\* weights = 369   
## 120) Fire\_Cause\_Human == {Unspecified}  
## 122)\* weights = 51   
## 119) Juris\_Long == {British Columbia, Northwest Territories}  
## 123)\* weights = 399   
## 110) Juris\_Long == {Alberta, National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Quebec, Yukon}  
## 124) Juris\_Long == {National parks, Quebec}; criterion = 1, statistic = 625.901  
## 125)\* weights = 818   
## 124) Juris\_Long == {Alberta, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Yukon}  
## 126)\* weights = 1805   
## 109) Protection.zone == {Limited}  
## 127) Response.category == {None}; criterion = 1, statistic = 39.767  
## 128) Juris\_Long == {Manitoba, Ontario, Quebec, Yukon}; criterion = 0.998, statistic = 22.397  
## 129)\* weights = 728   
## 128) Juris\_Long == {Newfoundland and Labrador, Northwest Territories}  
## 130)\* weights = 225   
## 127) Response.category == {Modified}  
## 131) Juris\_Long == {Manitoba, Ontario}; criterion = 1, statistic = 52.421  
## 132)\* weights = 363   
## 131) Juris\_Long == {Newfoundland and Labrador, Northwest Territories, Quebec, Yukon}  
## 133)\* weights = 577

plot(fire\_ctree, type="simple")



train\_index <- sample(1:nrow(data), 0.7 \* nrow(data))  
train.set <- data[train\_index,]  
test.set <- data[-train\_index,]  
  
fire\_ctree\_model <- ctree(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data=train.set)  
fire\_ctree\_model

##   
## Conditional inference tree with 54 terminal nodes  
##   
## Response: Number   
## Inputs: Juris\_Long, Year, Fire\_Cause\_Human, Protection.zone, Response.category   
## Number of observations: 8063   
##   
## 1) Response.category == {Full, Unspecified}; criterion = 1, statistic = 451.074  
## 2) Juris\_Long == {Alberta, British Columbia}; criterion = 1, statistic = 411.854  
## 3) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 208.128  
## 4) Juris\_Long == {British Columbia}; criterion = 1, statistic = 15.803  
## 5)\* weights = 18   
## 4) Juris\_Long == {Alberta}  
## 6)\* weights = 22   
## 3) Fire\_Cause\_Human == {Human, Unspecified}  
## 7) Fire\_Cause\_Human == {Human}; criterion = 0.997, statistic = 19.172  
## 8)\* weights = 288   
## 7) Fire\_Cause\_Human == {Unspecified}  
## 9)\* weights = 40   
## 2) Juris\_Long == {Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Ontario, Prince Edward Island, Quebec, Saskatchewan, Yukon}  
## 10) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 256.901  
## 11) Juris\_Long == {Ontario, Saskatchewan}; criterion = 1, statistic = 76.895  
## 12) Protection.zone == {Intensive, Unspecified}; criterion = 0.996, statistic = 14.517  
## 13)\* weights = 37   
## 12) Protection.zone == {Limited}  
## 14)\* weights = 22   
## 11) Juris\_Long == {Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Prince Edward Island, Quebec, Yukon}  
## 15) Juris\_Long == {Manitoba, Northwest Territories, Quebec}; criterion = 1, statistic = 60.286  
## 16) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 27.638  
## 17) Year <= 2007; criterion = 0.966, statistic = 7.324  
## 18)\* weights = 42   
## 17) Year > 2007  
## 19)\* weights = 20   
## 16) Protection.zone == {Limited}  
## 20)\* weights = 33   
## 15) Juris\_Long == {National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Yukon}  
## 21) Juris\_Long == {National parks, New Brunswick, Yukon}; criterion = 1, statistic = 35.664  
## 22) Protection.zone == {Intensive, Unspecified}; criterion = 0.998, statistic = 15.386  
## 23)\* weights = 63   
## 22) Protection.zone == {Limited}  
## 24)\* weights = 9   
## 21) Juris\_Long == {Newfoundland and Labrador, Nova Scotia, Prince Edward Island}  
## 25) Juris\_Long == {Newfoundland and Labrador, Nova Scotia}; criterion = 0.952, statistic = 9.269  
## 26)\* weights = 57   
## 25) Juris\_Long == {Prince Edward Island}  
## 27)\* weights = 17   
## 10) Fire\_Cause\_Human == {Human, Unspecified}  
## 28) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 295.317  
## 29) Juris\_Long == {Manitoba, New Brunswick, Nova Scotia, Ontario, Quebec, Saskatchewan}; criterion = 1, statistic = 339.289  
## 30) Juris\_Long == {Ontario, Quebec}; criterion = 1, statistic = 73.347  
## 31) Fire\_Cause\_Human == {Human}; criterion = 1, statistic = 35.505  
## 32) Year <= 1991; criterion = 0.998, statistic = 31.799  
## 33)\* weights = 19   
## 32) Year > 1991  
## 34)\* weights = 255   
## 31) Fire\_Cause\_Human == {Unspecified}  
## 35) Juris\_Long == {Ontario}; criterion = 1, statistic = 28.354  
## 36)\* weights = 21   
## 35) Juris\_Long == {Quebec}  
## 37)\* weights = 20   
## 30) Juris\_Long == {Manitoba, New Brunswick, Nova Scotia, Saskatchewan}  
## 38) Year <= 2003; criterion = 0.985, statistic = 30.36  
## 39)\* weights = 311   
## 38) Year > 2003  
## 40)\* weights = 299   
## 29) Juris\_Long == {National parks, Newfoundland and Labrador, Northwest Territories, Prince Edward Island, Yukon}  
## 41) Juris\_Long == {Newfoundland and Labrador}; criterion = 1, statistic = 68.339  
## 42) Fire\_Cause\_Human == {Human}; criterion = 0.985, statistic = 8.844  
## 43)\* weights = 127   
## 42) Fire\_Cause\_Human == {Unspecified}  
## 44)\* weights = 19   
## 41) Juris\_Long == {National parks, Northwest Territories, Prince Edward Island, Yukon}  
## 45) Year <= 1998; criterion = 0.999, statistic = 13.649  
## 46) Juris\_Long == {Yukon}; criterion = 0.968, statistic = 12.284  
## 47)\* weights = 53   
## 46) Juris\_Long == {National parks, Northwest Territories, Prince Edward Island}  
## 48)\* weights = 136   
## 45) Year > 1998  
## 49) Juris\_Long == {National parks, Yukon}; criterion = 0.997, statistic = 17.057  
## 50)\* weights = 210   
## 49) Juris\_Long == {Northwest Territories, Prince Edward Island}  
## 51)\* weights = 183   
## 28) Protection.zone == {Limited}  
## 52) Juris\_Long == {Ontario}; criterion = 1, statistic = 123.063  
## 53)\* weights = 148   
## 52) Juris\_Long == {Manitoba, Newfoundland and Labrador, Northwest Territories, Quebec, Saskatchewan, Yukon}  
## 54) Juris\_Long == {Manitoba, Quebec}; criterion = 1, statistic = 115.834  
## 55)\* weights = 268   
## 54) Juris\_Long == {Newfoundland and Labrador, Northwest Territories, Saskatchewan, Yukon}  
## 56)\* weights = 250   
## 1) Response.category == {Modified, None}  
## 57) Fire\_Cause\_Human == {Lightning}; criterion = 1, statistic = 714.698  
## 58) Juris\_Long == {Northwest Territories, Saskatchewan}; criterion = 1, statistic = 117.365  
## 59) Response.category == {Modified}; criterion = 1, statistic = 45.529  
## 60)\* weights = 34   
## 59) Response.category == {None}  
## 61)\* weights = 38   
## 58) Juris\_Long == {Alberta, British Columbia, Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 62) Protection.zone == {Limited, Unspecified}; criterion = 1, statistic = 66.858  
## 63) Response.category == {None}; criterion = 1, statistic = 47.692  
## 64) Juris\_Long == {Ontario}; criterion = 1, statistic = 24.764  
## 65)\* weights = 16   
## 64) Juris\_Long == {Manitoba, Newfoundland and Labrador, Quebec, Yukon}  
## 66) Juris\_Long == {Manitoba, Quebec, Yukon}; criterion = 0.992, statistic = 15.182  
## 67)\* weights = 53   
## 66) Juris\_Long == {Newfoundland and Labrador}  
## 68)\* weights = 19   
## 63) Response.category == {Modified}  
## 69) Juris\_Long == {Manitoba, National parks}; criterion = 1, statistic = 39.625  
## 70)\* weights = 19   
## 69) Juris\_Long == {Alberta, New Brunswick, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 71)\* weights = 77   
## 62) Protection.zone == {Intensive}  
## 72) Juris\_Long == {Manitoba, National parks}; criterion = 1, statistic = 38.993  
## 73) Year <= 2011; criterion = 0.992, statistic = 9.863  
## 74)\* weights = 49   
## 73) Year > 2011  
## 75)\* weights = 15   
## 72) Juris\_Long == {Alberta, British Columbia, New Brunswick, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 76)\* weights = 275   
## 57) Fire\_Cause\_Human == {Human, Unspecified}  
## 77) Juris\_Long == {Saskatchewan}; criterion = 1, statistic = 322.656  
## 78) Response.category == {Modified}; criterion = 1, statistic = 28.707  
## 79) Year <= 2014; criterion = 0.998, statistic = 12.235  
## 80)\* weights = 86   
## 79) Year > 2014  
## 81)\* weights = 13   
## 78) Response.category == {None}  
## 82)\* weights = 110   
## 77) Juris\_Long == {Alberta, British Columbia, Manitoba, National parks, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Ontario, Prince Edward Island, Quebec, Yukon}  
## 83) Protection.zone == {Intensive, Unspecified}; criterion = 1, statistic = 655.357  
## 84) Juris\_Long == {British Columbia, Manitoba, Northwest Territories, Ontario}; criterion = 1, statistic = 518.066  
## 85) Response.category == {Modified}; criterion = 0.999, statistic = 323.689  
## 86) Year <= 2017; criterion = 0.999, statistic = 179.502  
## 87) Year <= 2010; criterion = 0.982, statistic = 8.436  
## 88) Juris\_Long == {Manitoba, Northwest Territories, Ontario}; criterion = 1, statistic = 27.792  
## 89)\* weights = 313   
## 88) Juris\_Long == {British Columbia}  
## 90)\* weights = 106   
## 87) Year > 2010  
## 91) Juris\_Long == {British Columbia}; criterion = 1, statistic = 41.561  
## 92)\* weights = 7   
## 91) Juris\_Long == {Manitoba, Northwest Territories, Ontario}  
## 93)\* weights = 85   
## 86) Year > 2017  
## 94)\* weights = 9   
## 85) Response.category == {None}  
## 95) Juris\_Long == {Manitoba, Ontario}; criterion = 1, statistic = 58.929  
## 96)\* weights = 302   
## 95) Juris\_Long == {British Columbia, Northwest Territories}  
## 97)\* weights = 275   
## 84) Juris\_Long == {Alberta, National parks, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Quebec, Yukon}  
## 98) Juris\_Long == {National parks, Quebec}; criterion = 1, statistic = 382.065  
## 99)\* weights = 594   
## 98) Juris\_Long == {Alberta, New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island, Yukon}  
## 100)\* weights = 1257   
## 83) Protection.zone == {Limited}  
## 101) Response.category == {None}; criterion = 1, statistic = 33.5  
## 102) Juris\_Long == {Ontario, Quebec, Yukon}; criterion = 0.997, statistic = 21.867  
## 103)\* weights = 386   
## 102) Juris\_Long == {Manitoba, Newfoundland and Labrador, Northwest Territories}  
## 104)\* weights = 274   
## 101) Response.category == {Modified}  
## 105) Juris\_Long == {Manitoba, Ontario}; criterion = 1, statistic = 30.4  
## 106)\* weights = 268   
## 105) Juris\_Long == {Newfoundland and Labrador, Northwest Territories, Quebec, Yukon}  
## 107)\* weights = 396

fire\_ctree\_prediction <- predict(fire\_ctree\_model, test.set)   
# gives the probability for each class  
head(fire\_ctree\_prediction)

## Number  
## [1,] 98.29514  
## [2,] 98.29514  
## [3,] 98.29514  
## [4,] 98.29514  
## [5,] 98.29514  
## [6,] 98.29514

table(fire\_ctree\_prediction, test.set$Number)

##   
## fire\_ctree\_prediction 0 1 2 3 4 5 6 7 8 9 10 11 12  
## 0 206 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 179 2 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 545 2 0 0 0 0 0 1 0 0 0 0 0  
## 0.04 111 6 1 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 8 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 119 5 0 1 0 1 0 0 0 1 0 0 0  
## 0.118181818181818 46 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 81 10 4 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 198 17 7 1 0 0 1 0 0 0 0 0 0  
## 0.380829015544041 140 18 3 5 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 90 14 8 6 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 82 19 8 4 1 4 1 0 1 0 0 0 0  
## 0.623529411764706 27 8 4 1 2 1 0 0 0 0 1 0 0  
## 1.30223880597015 73 23 7 8 5 4 2 2 1 3 0 1 0  
## 2.21857923497268 49 15 4 3 2 4 3 0 3 2 2 0 0  
## 2.35636363636364 76 9 4 2 1 0 1 1 1 0 1 1 0  
## 3.30232558139535 18 2 1 1 1 1 1 1 3 1 1 0 0  
## 3.50649350649351 17 4 2 2 2 0 0 0 0 0 0 0 0  
## 4.45714285714286 32 9 5 6 4 2 1 5 3 2 2 0 2  
## 4.74264705882353 33 3 4 2 4 5 0 2 1 2 0 2 0  
## 4.88888888888889 1 0 0 0 0 0 0 0 1 0 0 0 0  
## 6.91228070175439 3 1 6 7 2 3 1 1 2 0 0 0 0  
## 7.50675675675676 13 11 2 1 6 1 3 1 1 2 0 5 2  
## 7.53061224489796 8 2 0 0 1 2 1 1 0 2 0 0 1  
## 8 2 0 1 0 0 0 1 2 0 0 1 0 0  
## 9.73584905660377 5 0 0 0 0 2 0 1 0 0 1 0 0  
## 10 0 0 1 1 1 0 1 1 0 0 1 0 0  
## 11.2307692307692 4 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 1 1 2 4 0 0 1 1 2 0 0 0 0  
## 14.8421052631579 1 0 0 1 0 0 0 0 0 0 1 0 0  
## 15.0708661417323 13 4 5 4 4 4 2 2 0 1 0 2 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 1 1 0  
## 23.0666666666667 3 1 0 0 1 0 0 0 0 0 0 0 0  
## 29.0969899665552 23 5 2 7 8 3 7 3 4 5 2 2 1  
## 33.3962264150943 0 0 0 0 0 1 1 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 1 2 1 0 0 0 0  
## 40.7395498392283 12 3 3 3 3 3 2 7 4 4 4 3 0  
## 41.775 1 0 0 0 0 0 0 0 0 0 1 0 0  
## 57.7272727272727 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 1 2 1 1 3 1 1 2 0 2 4 2 1  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 0 0 0 1 2 1 0 2 2 1 2 2 0  
## 100.294117647059 2 0 0 0 0 0 0 0 0 0 1 0 0  
## 140.052631578947 0 0 1 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 374.459459459459 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 13 14 15 16 17 18 19 20 21 22 23 24 25  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 2 0 0 2 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 1 1 0 0 0 1 1 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 3.30232558139535 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 1 1 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 2 1 0 0 1 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 1 0 0 1 0 1 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 1 0 1 0 1 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 1 5 0 2 1 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 1 0 1 0 0  
## 9.73584905660377 2 1 0 0 1 0 0 0 0 0 2 0 0  
## 10 0 0 1 1 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 1 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 2 0 2 0 0 0 0 0 0 1 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 1 0 0 0 0 0  
## 15.0708661417323 0 2 1 0 4 1 0 0 1 0 0 3 3  
## 19.3333333333333 0 0 0 0 0 1 0 1 0 0 0 0 1  
## 23.0666666666667 0 0 0 1 0 0 0 1 1 0 0 0 0  
## 29.0969899665552 2 5 7 3 0 1 4 1 5 0 1 1 0  
## 33.3962264150943 1 2 0 1 0 0 0 0 0 0 1 0 0  
## 39.4285714285714 1 1 1 0 0 0 0 1 0 0 0 1 2  
## 40.7395498392283 1 0 0 1 2 4 5 1 0 1 2 1 1  
## 41.775 0 0 1 1 0 0 0 0 0 0 1 0 0  
## 57.7272727272727 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 1 0 0 0 0 0 1 0 0 0 0 0 0  
## 65.7686274509804 4 4 3 2 0 2 1 0 2 2 0 5 2  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 1 2 2 3 0 0 0 3 2 1 0 0 1  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 1 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 374.459459459459 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 26 27 28 29 30 31 32 33 34 35 36 37 38  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 1 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 1 0 0 0 0 1 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 1 0 1 1 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 1 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 1 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 1 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 1 1 1 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 1 0 0 0 0 0 0 0 0 1  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 1 0 2 0 1 1 1 1 3 2 0 0 0  
## 33.3962264150943 0 0 1 0 1 0 1 0 1 0 0 1 0  
## 39.4285714285714 0 0 0 2 0 1 1 1 0 0 1 0 0  
## 40.7395498392283 4 1 0 2 0 0 0 1 0 3 1 1 0  
## 41.775 0 0 0 0 0 0 1 0 0 0 1 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 1 0 0 0 0 0 0 0 0 0 2 0 0  
## 65.7686274509804 0 0 1 2 1 1 2 1 1 1 0 0 0  
## 81.1 0 0 0 0 0 0 0 1 0 0 0 0 0  
## 98.2951388888889 1 1 0 1 1 1 0 0 0 1 0 1 0  
## 100.294117647059 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 140.052631578947 0 0 0 0 1 0 0 1 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 374.459459459459 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 39 40 41 42 43 44 45 46 47 48 49 50 51  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 1 0 0 1 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 1 1 0 1 0 1 2 0 2 0 1 0  
## 33.3962264150943 0 0 1 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 1 0  
## 40.7395498392283 3 2 0 0 0 3 2 0 0 1 1 1 0  
## 41.775 1 0 0 1 1 0 0 1 0 0 1 0 0  
## 57.7272727272727 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 0 0 0 0 3 0 1 1 1 3 0 0 1  
## 81.1 0 0 0 1 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 1 0 1 0 0 2 2 0 1 1 0 0 0  
## 100.294117647059 0 0 0 0 0 0 0 0 1 0 0 0 0  
## 140.052631578947 0 1 0 0 0 0 0 0 0 0 1 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0 0 1  
## 374.459459459459 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 52 53 54 55 56 58 59 60 61 62 63 64 65  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 1 0 0 0 0 0 1 1 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 1 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 1 0 0 0 0 1 0 0 0 0 0  
## 33.3962264150943 0 0 0 0 1 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 1 3 0 3 0 0 0 0 0 0 1 0  
## 41.775 0 0 0 0 0 1 1 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 1 0 0 0 0 0 0 2 1 2 0 0 0  
## 81.1 0 0 0 1 0 0 0 0 0 1 0 1 0  
## 98.2951388888889 1 1 0 0 0 0 2 0 0 1 0 1 1  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 1 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 374.459459459459 0 0 0 0 0 0 0 0 0 0 1 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 66 67 69 70 71 72 73 75 76 77 78 79 80  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 1 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 1 1 1 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 1 1 0 1 1 0 0 0 0 0 0 0 0  
## 33.3962264150943 0 0 0 0 2 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 1 0 0 0 0  
## 40.7395498392283 2 0 0 1 2 0 1 0 0 0 1 0 1  
## 41.775 1 0 0 0 1 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 1 0 1 0 0 0 0 1 0  
## 63.5625 0 0 0 0 0 0 1 0 0 1 0 0 0  
## 65.7686274509804 0 0 0 0 1 0 1 0 0 0 1 0 0  
## 81.1 0 0 0 0 0 0 1 0 0 1 0 0 0  
## 98.2951388888889 0 0 0 0 0 1 0 0 3 0 0 0 0  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 1 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 374.459459459459 0 1 0 1 0 0 0 1 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 82 83 84 85 86 87 88 90 91 93 94 96 98  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 1 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 1 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 2 0 0 0 0 0 0 0 1 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 1 0  
## 39.4285714285714 0 0 0 0 1 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 1 0 0 1 1 1 2 1 0 0 0 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 1 0 0 0 0 0 1 0 0 0 0  
## 65.7686274509804 0 0 0 0 1 0 0 1 1 0 0 0 0  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 0 0 0 0 0 1 1 0 0 1 1 0 0  
## 100.294117647059 0 0 0 0 0 0 0 1 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 1 0 0 0 0 0 0 0  
## 374.459459459459 1 0 0 0 0 0 0 0 0 0 0 0 1  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 99 100 101 102 103 104 105 107 109 110 112 113 115  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 1 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 1 0 2  
## 33.3962264150943 0 0 1 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 1 0 0 1 1 0 1 0 0 0 1 0 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 65.7686274509804 0 1 0 0 1 1 2 2 1 0 0 1 0  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 0 1 0 1 1 0 0 1 0 0 1 3 0  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 374.459459459459 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 116 117 118 120 121 123 124 125 126 127 128 129 130  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 1 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 1 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 0 0 1 0 0 0 0 0 0 0 1 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 1 0  
## 63.5625 0 0 0 0 0 0 1 0 0 0 0 0 1  
## 65.7686274509804 0 1 1 1 0 0 0 1 0 2 0 0 0  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 1 0  
## 98.2951388888889 1 0 1 0 0 0 2 1 0 0 0 0 1  
## 100.294117647059 0 0 0 1 0 1 0 0 1 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0 0 1  
## 374.459459459459 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 134 135 136 137 138 141 143 144 145 146 147 148 154  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 1 0 0 0 0 1 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 0 1 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 1 0 0 0 1 1 0 0 0 1 0 0 2  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 0 0 0 0 0 0 0 1 0 0 0 0 0  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 1 0 2 0 0 1 0 0 0 0 1 0 0  
## 100.294117647059 0 0 0 0 0 0 1 0 0 0 0 1 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 374.459459459459 0 1 0 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 155 156 158 160 161 162 163 165 167 170 172 174 176  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 1 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 1 1 0 0 1 0 0 0 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 0 0 0 0 0 1 0 0 0 1 1 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 0 0 2 1 0 0 0 0 0 2 0 0 0  
## 81.1 0 0 0 0 1 0 0 0 0 0 0 0 1  
## 98.2951388888889 1 0 1 0 0 1 0 0 0 0 0 1 0  
## 100.294117647059 0 0 0 0 0 0 0 1 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 1 0 0 0 0  
## 374.459459459459 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 178 179 180 181 185 186 187 188 190 194 195 196 197  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 1 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 1 0 0 1 0 0 1 0 0 0 1 0 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 0 1 0 0 0 1 0 0 0 0 0 0 0  
## 81.1 0 0 0 0 0 0 0 0 1 0 0 0 1  
## 98.2951388888889 0 0 1 0 1 0 0 0 1 1 0 0 0  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 374.459459459459 0 0 1 0 0 0 0 1 0 0 0 1 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 198 209 211 213 215 216 217 218 225 228 232 241 244  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 0 1 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 1 1 0 0 0 0 0 0 0 0 0 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 0 0 0 0 0 0 0 0 0 1 1 0 1  
## 81.1 0 0 0 0 0 1 0 0 0 0 0 0 0  
## 98.2951388888889 0 0 0 0 0 0 0 1 1 1 0 0 0  
## 100.294117647059 0 0 0 0 1 0 0 0 0 0 0 0 0  
## 140.052631578947 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 1 0 0 1 0 0 0 0 0 0  
## 374.459459459459 0 0 0 0 0 0 0 0 1 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 246 247 262 263 266 270 282 288 289 290 292 296 300  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 1 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 1 0 1 0 0 1 0 0 0 0 1 1 0  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 0 0 0 1 1 0 0 0 0 1 0 0 1  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 1 1 0 0 0 0 0  
## 374.459459459459 0 0 1 0 0 0 1 0 1 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## fire\_ctree\_prediction 313 315 317 321 322 323 327 340 342 352 377 381 384  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 1 1 0 2 1 0 0 1 1 1 1 1 0  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 1 0 0 0 0 0 0 0  
## 374.459459459459 0 0 1 0 0 0 0 0 0 0 0 0 0  
## 556.772727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 0 1 0 0 0 1  
##   
## fire\_ctree\_prediction 396 426 436 451 458 466 467 479 486 490 517 533 585  
## 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 81.1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 0 1 0 0 1 0 1 0 0 0 0 0 0  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 374.459459459459 0 0 0 1 0 0 1 0 0 0 0 1 0  
## 556.772727272727 0 0 1 0 0 1 0 0 0 1 1 0 0  
## 1356.11111111111 0 0 0 0 0 0 0 1 1 0 0 0 1  
##   
## fire\_ctree\_prediction 675 777 782 890 912 944 988 992 1141 1175 1192  
## 0 0 0 0 0 0 0 0 0 0 0 0  
## 0.00757575757575758 0 0 0 0 0 0 0 0 0 0 0  
## 0.0119331742243437 0 0 0 0 0 0 0 0 0 0 0  
## 0.04 0 0 0 0 0 0 0 0 0 0 0  
## 0.0588235294117647 0 0 0 0 0 0 0 0 0 0 0  
## 0.0912408759124088 0 0 0 0 0 0 0 0 0 0 0  
## 0.118181818181818 0 0 0 0 0 0 0 0 0 0 0  
## 0.130597014925373 0 0 0 0 0 0 0 0 0 0 0  
## 0.203703703703704 0 0 0 0 0 0 0 0 0 0 0  
## 0.380829015544041 0 0 0 0 0 0 0 0 0 0 0  
## 0.536423841059603 0 0 0 0 0 0 0 0 0 0 0  
## 0.584664536741214 0 0 0 0 0 0 0 0 0 0 0  
## 0.623529411764706 0 0 0 0 0 0 0 0 0 0 0  
## 1.30223880597015 0 0 0 0 0 0 0 0 0 0 0  
## 2.21857923497268 0 0 0 0 0 0 0 0 0 0 0  
## 2.35636363636364 0 0 0 0 0 0 0 0 0 0 0  
## 3.30232558139535 0 0 0 0 0 0 0 0 0 0 0  
## 3.50649350649351 0 0 0 0 0 0 0 0 0 0 0  
## 4.45714285714286 0 0 0 0 0 0 0 0 0 0 0  
## 4.74264705882353 0 0 0 0 0 0 0 0 0 0 0  
## 4.88888888888889 0 0 0 0 0 0 0 0 0 0 0  
## 6.91228070175439 0 0 0 0 0 0 0 0 0 0 0  
## 7.50675675675676 0 0 0 0 0 0 0 0 0 0 0  
## 7.53061224489796 0 0 0 0 0 0 0 0 0 0 0  
## 8 0 0 0 0 0 0 0 0 0 0 0  
## 9.73584905660377 0 0 0 0 0 0 0 0 0 0 0  
## 10 0 0 0 0 0 0 0 0 0 0 0  
## 11.2307692307692 0 0 0 0 0 0 0 0 0 0 0  
## 11.3333333333333 0 0 0 0 0 0 0 0 0 0 0  
## 14.8421052631579 0 0 0 0 0 0 0 0 0 0 0  
## 15.0708661417323 0 0 0 0 0 0 0 0 0 0 0  
## 19.3333333333333 0 0 0 0 0 0 0 0 0 0 0  
## 23.0666666666667 0 0 0 0 0 0 0 0 0 0 0  
## 29.0969899665552 0 0 0 0 0 0 0 0 0 0 0  
## 33.3962264150943 0 0 0 0 0 0 0 0 0 0 0  
## 39.4285714285714 0 0 0 0 0 0 0 0 0 0 0  
## 40.7395498392283 0 0 0 0 0 0 0 0 0 0 0  
## 41.775 0 0 0 0 0 0 0 0 0 0 0  
## 57.7272727272727 0 0 0 0 0 0 0 0 0 0 0  
## 63.5625 0 0 0 0 0 0 0 0 0 0 0  
## 65.7686274509804 0 0 0 0 0 0 0 0 0 0 0  
## 81.1 0 0 0 0 0 0 0 0 0 0 0  
## 98.2951388888889 0 1 0 0 0 0 0 0 0 0 0  
## 100.294117647059 0 0 0 0 0 0 0 0 0 0 0  
## 140.052631578947 0 0 0 0 0 0 0 0 0 0 0  
## 198.095238095238 0 0 0 0 0 0 0 0 0 0 0  
## 374.459459459459 0 0 0 0 0 0 1 0 1 0 0  
## 556.772727272727 0 0 0 1 0 0 0 0 0 0 1  
## 1356.11111111111 1 0 1 0 1 1 0 1 0 1 0

## 32.

## Model below is separate from models above

## Additional for testing: Random Forest

# Check data before Random Forest  
summary(data)

## Cause Jurisdiction Number   
## Lightning :1325 QC :1516 Min. : 0.00   
## Unspecified :1290 NL :1486 1st Qu.: 0.00   
## Miscellaneous known causes:1279 ON :1368 Median : 0.00   
## Recreation :1274 MB :1304 Mean : 18.64   
## Incendiary :1271 YT :1059 3rd Qu.: 4.00   
## Residents :1268 NT : 773 Max. :2913.00   
## (Other) :3812 (Other):4013   
## Protection.zone Response.category Year   
## Intensive :8190 Full :4204 Min. :1990   
## Limited :3281 Modified :3521 1st Qu.:1997   
## Unspecified: 48 None :3681 Median :2004   
## Unspecified: 113 Mean :2004   
## 3rd Qu.:2011   
## Max. :2018   
##   
## Juris\_Long Cause\_Grouped Time1   
## Quebec :1516 Length:11519 Length:11519   
## Newfoundland and Labrador:1486 Class :character Class :character   
## Ontario :1368 Mode :character Mode :character   
## Manitoba :1304   
## Yukon :1059   
## Northwest Territories : 773   
## (Other) :4013   
## Time2 Region Fire\_Cause\_Human  
## Length:11519 Length:11519 Human :8904   
## Class :character Class :character Lightning :1325   
## Mode :character Mode :character Unspecified:1290   
##   
##   
##   
##

class(data$Fire\_Cause\_Human) # [1] "factor"

## [1] "factor"

unique(data$Fire\_Cause\_Human)

## [1] Human Lightning Unspecified  
## Levels: Human Lightning Unspecified

data$Fire\_Cause\_Human <- as.factor(data$Fire\_Cause\_Human) #Change to factor  
class(data$Fire\_Cause\_Human)

## [1] "factor"

unique(data$Fire\_Cause\_Human)

## [1] Human Lightning Unspecified  
## Levels: Human Lightning Unspecified

names(data)

## [1] "Cause" "Jurisdiction" "Number"   
## [4] "Protection.zone" "Response.category" "Year"   
## [7] "Juris\_Long" "Cause\_Grouped" "Time1"   
## [10] "Time2" "Region" "Fire\_Cause\_Human"

str(data)

## Classes 'data.table' and 'data.frame': 11519 obs. of 12 variables:  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "AB","BC","MB",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Number : int 22 14 12 11 13 14 8 29 10 20 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Cause\_Grouped : chr "People" "People" "People" "People" ...  
## $ Time1 : chr "Early 90s" "Early 90s" "Early 90s" "Early 90s" ...  
## $ Time2 : chr "1990s" "1990s" "1990s" "1990s" ...  
## $ Region : chr "Prairie Region" "Prairie Region" "Prairie Region" "Prairie Region" ...  
## $ Fire\_Cause\_Human : Factor w/ 3 levels "Human","Lightning",..: 1 1 1 1 1 1 1 1 1 1 ...  
## - attr(\*, ".internal.selfref")=<externalptr>

train\_index <- sample(1:nrow(data), 0.7 \* nrow(data))  
train.set <- data[train\_index,]  
test.set <- data[-train\_index,]  
  
rf1 <- randomForest(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set)  
rf1

##   
## Call:  
## randomForest(formula = Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set)   
## Type of random forest: regression  
## Number of trees: 500  
## No. of variables tried at each split: 1  
##   
## Mean of squared residuals: 4550.046  
## % Var explained: 35.75

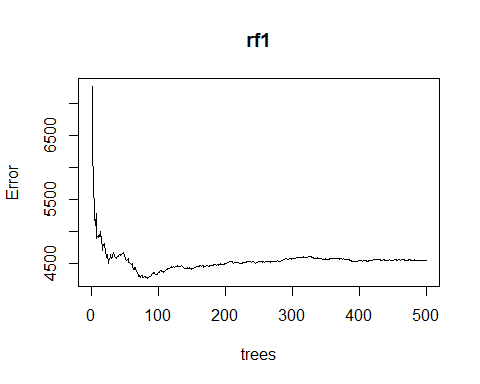
print(rf1)

##   
## Call:  
## randomForest(formula = Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set)   
## Type of random forest: regression  
## Number of trees: 500  
## No. of variables tried at each split: 1  
##   
## Mean of squared residuals: 4550.046  
## % Var explained: 35.75

randomForest::importance(rf1) # Higher the value, the greater the importance.

## IncNodePurity  
## Juris\_Long 4776969.8  
## Year 942840.9  
## Fire\_Cause\_Human 4964138.4  
## Protection.zone 985171.6  
## Response.category 4325837.3

plot(rf1) # Error levels of output



table(train.set$Number)/nrow(train.set)

##   
## 0 1 2 3 4   
## 0.6435569887 0.0613915416 0.0256728265 0.0190995907 0.0178593576   
## 5 6 7 8 9   
## 0.0110380752 0.0099218653 0.0095497954 0.0074413990 0.0079374922   
## 10 11 12 13 14   
## 0.0076894456 0.0071933524 0.0040927694 0.0055810492 0.0047128860   
## 15 16 17 18 19   
## 0.0043408161 0.0052089793 0.0027285130 0.0040927694 0.0035966762   
## 20 21 22 23 24   
## 0.0026044896 0.0031005829 0.0026044896 0.0021083964 0.0027285130   
## 25 26 27 28 29   
## 0.0026044896 0.0023564430 0.0021083964 0.0026044896 0.0022324197   
## 30 31 32 33 34   
## 0.0017363264 0.0017363264 0.0027285130 0.0016123031 0.0012402332   
## 35 36 37 38 39   
## 0.0007441399 0.0022324197 0.0008681632 0.0013642565 0.0018603497   
## 40 41 42 43 44   
## 0.0016123031 0.0017363264 0.0011162098 0.0013642565 0.0016123031   
## 45 46 47 48 49   
## 0.0014882798 0.0016123031 0.0011162098 0.0018603497 0.0006201166   
## 50 51 52 53 54   
## 0.0011162098 0.0007441399 0.0008681632 0.0009921865 0.0012402332   
## 55 56 57 58 59   
## 0.0011162098 0.0012402332 0.0002480466 0.0008681632 0.0003720699   
## 60 61 62 63 64   
## 0.0009921865 0.0008681632 0.0011162098 0.0003720699 0.0009921865   
## 65 66 67 68 69   
## 0.0008681632 0.0009921865 0.0004960933 0.0004960933 0.0011162098   
## 70 71 72 73 74   
## 0.0012402332 0.0011162098 0.0002480466 0.0008681632 0.0003720699   
## 75 76 77 78 79   
## 0.0003720699 0.0008681632 0.0006201166 0.0006201166 0.0009921865   
## 80 81 82 83 84   
## 0.0004960933 0.0009921865 0.0006201166 0.0006201166 0.0004960933   
## 85 86 87 88 89   
## 0.0006201166 0.0003720699 0.0006201166 0.0007441399 0.0006201166   
## 90 91 92 93 94   
## 0.0006201166 0.0006201166 0.0007441399 0.0004960933 0.0006201166   
## 95 96 97 98 99   
## 0.0008681632 0.0003720699 0.0008681632 0.0002480466 0.0006201166   
## 100 101 102 103 104   
## 0.0004960933 0.0003720699 0.0007441399 0.0004960933 0.0003720699   
## 105 106 107 108 109   
## 0.0003720699 0.0004960933 0.0004960933 0.0008681632 0.0002480466   
## 110 111 112 113 114   
## 0.0003720699 0.0002480466 0.0002480466 0.0008681632 0.0002480466   
## 115 116 117 119 120   
## 0.0001240233 0.0004960933 0.0002480466 0.0003720699 0.0003720699   
## 121 123 124 125 126   
## 0.0004960933 0.0006201166 0.0003720699 0.0004960933 0.0002480466   
## 127 128 129 130 131   
## 0.0003720699 0.0004960933 0.0002480466 0.0003720699 0.0003720699   
## 132 133 134 135 136   
## 0.0002480466 0.0004960933 0.0002480466 0.0003720699 0.0003720699   
## 137 138 139 140 141   
## 0.0006201166 0.0002480466 0.0001240233 0.0002480466 0.0004960933   
## 143 145 146 147 148   
## 0.0004960933 0.0003720699 0.0003720699 0.0006201166 0.0002480466   
## 149 150 151 152 153   
## 0.0002480466 0.0001240233 0.0001240233 0.0002480466 0.0001240233   
## 154 155 156 158 159   
## 0.0001240233 0.0002480466 0.0001240233 0.0003720699 0.0002480466   
## 160 161 162 163 164   
## 0.0001240233 0.0006201166 0.0002480466 0.0001240233 0.0003720699   
## 165 166 167 170 172   
## 0.0003720699 0.0001240233 0.0003720699 0.0002480466 0.0003720699   
## 173 174 178 179 180   
## 0.0001240233 0.0001240233 0.0003720699 0.0002480466 0.0003720699   
## 181 183 184 185 186   
## 0.0001240233 0.0001240233 0.0001240233 0.0002480466 0.0001240233   
## 187 188 190 191 194   
## 0.0002480466 0.0001240233 0.0002480466 0.0001240233 0.0003720699   
## 195 196 198 200 201   
## 0.0002480466 0.0001240233 0.0001240233 0.0002480466 0.0001240233   
## 202 203 204 207 209   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0002480466   
## 211 213 215 216 217   
## 0.0002480466 0.0002480466 0.0001240233 0.0002480466 0.0001240233   
## 218 219 220 221 222   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 225 226 228 232 233   
## 0.0003720699 0.0001240233 0.0002480466 0.0003720699 0.0001240233   
## 235 238 240 241 242   
## 0.0001240233 0.0001240233 0.0001240233 0.0002480466 0.0001240233   
## 244 246 247 249 250   
## 0.0001240233 0.0001240233 0.0001240233 0.0002480466 0.0001240233   
## 253 256 260 262 266   
## 0.0001240233 0.0002480466 0.0001240233 0.0003720699 0.0003720699   
## 269 270 272 275 276   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 280 282 285 287 288   
## 0.0001240233 0.0002480466 0.0001240233 0.0001240233 0.0001240233   
## 289 290 296 299 300   
## 0.0002480466 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 305 306 313 316 321   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0002480466   
## 322 323 327 337 338   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 342 343 344 352 355   
## 0.0001240233 0.0001240233 0.0001240233 0.0002480466 0.0001240233   
## 357 361 377 384 396   
## 0.0001240233 0.0001240233 0.0001240233 0.0002480466 0.0001240233   
## 398 403 407 417 421   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 426 432 436 440 451   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 460 466 467 478 490   
## 0.0001240233 0.0001240233 0.0002480466 0.0001240233 0.0001240233   
## 505 511 514 515 530   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 545 563 566 585 609   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 626 723 726 741 759   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 772 777 779 780 782   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 787 828 831 842 886   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 912 914 988 1141 1151   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 1192 1237 1249 1433 1513   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233 0.0001240233   
## 1536 1716 1773 2913   
## 0.0001240233 0.0001240233 0.0001240233 0.0001240233

## 33.Model below is separate from models above

## Additional for testing: Linear Regression for Predicting Forest Fires.

# Create total forest fires per year variable.   
require(dplyr)  
  
print(unique(data[, 1])) #Print column 1, "Cause"

## Cause  
## 1: Forest industry  
## 2: Incendiary  
## 3: Lightning  
## 4: Miscellaneous known causes  
## 5: Other industry  
## 6: Railways  
## 7: Recreation  
## 8: Residents  
## 9: Unspecified  
## 10: Unspecified human activities

# Cause  
 # <fctr>  
 # Forest industry   
 # Incendiary   
 # Lightning   
 # Miscellaneous known causes   
 # Other industry   
 # Railways   
 # Recreation   
 # Residents   
 # Unspecified   
 # Unspecified human activities  
  
data$Fire\_Cause\_Human = as.character(data$Cause)  
#character is easier to use than factor  
#difference between factor and character  
  
data$Fire\_Cause\_Human[ data$Fire\_Cause\_Human != "Lightning" & data$Fire\_Cause\_Human != "Unspecified" ] = "Human"  
# There are three types of fire causes, Lightning, Human and Unspecified.   
# We don't know what unspecified is. Unspecified could be Lightning or Human cause.  
  
unique(data$Fire\_Cause\_Human)

## [1] "Human" "Lightning" "Unspecified"

#[1] "Human" "Lightning" "Unspecified"  
  
names(data)

## [1] "Cause" "Jurisdiction" "Number"   
## [4] "Protection.zone" "Response.category" "Year"   
## [7] "Juris\_Long" "Cause\_Grouped" "Time1"   
## [10] "Time2" "Region" "Fire\_Cause\_Human"

# [1] "Cause" "Jurisdiction" "Number" "Protection.zone" "Response.category" "Year" "Juris\_Long" "Cause\_Grouped" "Time1" "Time2"   
# [11] "Region" "Fire\_Cause\_Human"   
  
unique(data$Cause\_Grouped)

## [1] "People" "Lightning"

#[1] "People" "Lightning"  
  
unique(data$Year)

## [1] 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003  
## [15] 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017  
## [29] 2018

#[1] 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018  
############################  
fm1 = Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Response.category  
#class(fm1) #[1] "formula"  
#str(fm1)  
 # Class 'formula' language Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Response.category  
 # ..- attr(\*, ".Environment")=<environment: R\_GlobalEnv>   
  
model1 = lm(fm1, data = data)  
#class(model1) #[1] "lm"  
#str(model1) #Output is too long  
summary(model1)

##   
## Call:  
## lm(formula = fm1, data = data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -129.23 -27.39 0.16 16.86 2740.08   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 226.05347 190.23602 1.188 0.23475  
## Juris\_LongBritish Columbia 36.31642 4.32305 8.401 < 2e-16  
## Juris\_LongManitoba -39.30700 3.68226 -10.675 < 2e-16  
## Juris\_LongNational parks -45.09698 4.31909 -10.441 < 2e-16  
## Juris\_LongNew Brunswick -38.33090 5.78814 -6.622 3.69e-11  
## Juris\_LongNewfoundland and Labrador -46.29485 3.58016 -12.931 < 2e-16  
## Juris\_LongNorthwest Territories -40.44122 4.10710 -9.847 < 2e-16  
## Juris\_LongNova Scotia -37.21587 4.16929 -8.926 < 2e-16  
## Juris\_LongOntario -25.30748 3.63396 -6.964 3.48e-12  
## Juris\_LongPrince Edward Island -56.22543 4.97999 -11.290 < 2e-16  
## Juris\_LongQuebec -34.83754 3.56814 -9.764 < 2e-16  
## Juris\_LongSaskatchewan -27.01072 4.43673 -6.088 1.18e-09  
## Juris\_LongYukon -47.24102 3.82322 -12.356 < 2e-16  
## Year -0.07714 0.09492 -0.813 0.41644  
## Fire\_Cause\_HumanLightning 64.36514 2.36423 27.225 < 2e-16  
## Fire\_Cause\_HumanUnspecified -7.04797 2.39209 -2.946 0.00322  
## Response.categoryModified -44.00517 1.86045 -23.653 < 2e-16  
## Response.categoryNone -44.07735 1.84119 -23.940 < 2e-16  
## Response.categoryUnspecified 20.61424 7.91526 2.604 0.00922  
##   
## (Intercept)   
## Juris\_LongBritish Columbia \*\*\*  
## Juris\_LongManitoba \*\*\*  
## Juris\_LongNational parks \*\*\*  
## Juris\_LongNew Brunswick \*\*\*  
## Juris\_LongNewfoundland and Labrador \*\*\*  
## Juris\_LongNorthwest Territories \*\*\*  
## Juris\_LongNova Scotia \*\*\*  
## Juris\_LongOntario \*\*\*  
## Juris\_LongPrince Edward Island \*\*\*  
## Juris\_LongQuebec \*\*\*  
## Juris\_LongSaskatchewan \*\*\*  
## Juris\_LongYukon \*\*\*  
## Year   
## Fire\_Cause\_HumanLightning \*\*\*  
## Fire\_Cause\_HumanUnspecified \*\*   
## Response.categoryModified \*\*\*  
## Response.categoryNone \*\*\*  
## Response.categoryUnspecified \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 80.28 on 11500 degrees of freedom  
## Multiple R-squared: 0.1709, Adjusted R-squared: 0.1696   
## F-statistic: 131.7 on 18 and 11500 DF, p-value: < 2.2e-16

############################  
  
model1 = lm(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = data)  
#class(model1) #[1] "lm"  
#all years  
  
summary(model1)

##   
## Call:  
## lm(formula = Number ~ Juris\_Long + Year + Fire\_Cause\_Human +   
## Protection.zone + Response.category, data = data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -185.37 -27.16 0.44 14.72 2741.73   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 331.0932 190.7816 1.735 0.08269  
## Juris\_LongBritish Columbia 36.1393 4.3004 8.404 < 2e-16  
## Juris\_LongManitoba -32.4533 3.7679 -8.613 < 2e-16  
## Juris\_LongNational parks -45.1380 4.2963 -10.506 < 2e-16  
## Juris\_LongNew Brunswick -39.0378 5.7585 -6.779 1.27e-11  
## Juris\_LongNewfoundland and Labrador -39.0149 3.6881 -10.579 < 2e-16  
## Juris\_LongNorthwest Territories -39.3198 4.0867 -9.621 < 2e-16  
## Juris\_LongNova Scotia -37.2304 4.1473 -8.977 < 2e-16  
## Juris\_LongOntario -18.2214 3.7371 -4.876 1.10e-06  
## Juris\_LongPrince Edward Island -56.7625 4.9545 -11.457 < 2e-16  
## Juris\_LongQuebec -27.5689 3.6763 -7.499 6.89e-14  
## Juris\_LongSaskatchewan -25.7759 4.4169 -5.836 5.50e-09  
## Juris\_LongYukon -41.0647 3.8860 -10.567 < 2e-16  
## Year -0.1298 0.0952 -1.364 0.17263  
## Fire\_Cause\_HumanLightning 62.9286 2.3584 26.683 < 2e-16  
## Fire\_Cause\_HumanUnspecified -6.7846 2.3797 -2.851 0.00437  
## Protection.zoneLimited -14.0857 1.9217 -7.330 2.46e-13  
## Protection.zoneUnspecified 94.5828 11.6949 8.088 6.70e-16  
## Response.categoryModified -43.9075 1.8512 -23.718 < 2e-16  
## Response.categoryNone -43.6226 1.8321 -23.810 < 2e-16  
## Response.categoryUnspecified 17.1594 7.8833 2.177 0.02952  
##   
## (Intercept) .   
## Juris\_LongBritish Columbia \*\*\*  
## Juris\_LongManitoba \*\*\*  
## Juris\_LongNational parks \*\*\*  
## Juris\_LongNew Brunswick \*\*\*  
## Juris\_LongNewfoundland and Labrador \*\*\*  
## Juris\_LongNorthwest Territories \*\*\*  
## Juris\_LongNova Scotia \*\*\*  
## Juris\_LongOntario \*\*\*  
## Juris\_LongPrince Edward Island \*\*\*  
## Juris\_LongQuebec \*\*\*  
## Juris\_LongSaskatchewan \*\*\*  
## Juris\_LongYukon \*\*\*  
## Year   
## Fire\_Cause\_HumanLightning \*\*\*  
## Fire\_Cause\_HumanUnspecified \*\*   
## Protection.zoneLimited \*\*\*  
## Protection.zoneUnspecified \*\*\*  
## Response.categoryModified \*\*\*  
## Response.categoryNone \*\*\*  
## Response.categoryUnspecified \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 79.86 on 11498 degrees of freedom  
## Multiple R-squared: 0.1798, Adjusted R-squared: 0.1784   
## F-statistic: 126 on 20 and 11498 DF, p-value: < 2.2e-16

# avg of the reference category for  
# (Alberta)  
# (cause is human)  
# (response category is Full)  
# (protection.zone is Intensive)  
#Start with 331.0932 and add or subtract items to calculate specific query.  
  
#Three \*\*\*, evidence there is an effect on y variable  
# Unit of measure is number of fires  
# predict y is number of fires  
# Year variable has zero linear effect, no asterix \*  
# asterix \*, Evaluating P value if there is a linear effect or not  
#Estimate column is important  
#Multiple R-squared: 0.1798  
#R-squared: percentage of variation explained by the model. 18%  
#Adjusted R-squared: takes into account how many variables in model total.  
#p value for entire model: p-value: < 2.2e-16  
#p value: we reject the NULL. All the coefficients are zero.  
#p value comes from F-statistic  
  
unique(data$Response.category)

## [1] Full Modified None Unspecified  
## Levels: Full Modified None Unspecified

# [1] Full Modified None Unspecified  
# Levels: Full Modified None Unspecified  
  
unique(data$Protection.zone)

## [1] Intensive Limited Unspecified  
## Levels: Intensive Limited Unspecified

# [1] Intensive Limited Unspecified  
# Levels: Intensive Limited Unspecified

## 34.Model below is separate from models above

## Additional for testing: prediction

#Prediction for year 2018  
  
#Line below is for 2018  
#Train set  
data2018 = filter(data, Year == 2018)  
  
#Line below is everything except 2018  
#Test set  
datarm2018 = filter(data, Year != 2018)  
  
#Linear model for all years except 2018?  
modelrm2018 = lm(fm1, data = datarm2018)  
#Summary on Test set  
summary(modelrm2018)

##   
## Call:  
## lm(formula = fm1, data = datarm2018)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -127.08 -26.70 0.73 16.04 2743.53   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 413.99423 188.82655 2.192 0.02837  
## Juris\_LongBritish Columbia 34.90878 4.26057 8.193 2.81e-16  
## Juris\_LongManitoba -38.07578 3.62666 -10.499 < 2e-16  
## Juris\_LongNational parks -43.72007 4.25649 -10.271 < 2e-16  
## Juris\_LongNew Brunswick -36.87495 5.72484 -6.441 1.23e-10  
## Juris\_LongNewfoundland and Labrador -44.99968 3.52573 -12.763 < 2e-16  
## Juris\_LongNorthwest Territories -38.90683 4.07636 -9.545 < 2e-16  
## Juris\_LongNova Scotia -35.93674 4.10797 -8.748 < 2e-16  
## Juris\_LongOntario -24.91407 3.57872 -6.962 3.55e-12  
## Juris\_LongPrince Edward Island -54.81433 4.91482 -11.153 < 2e-16  
## Juris\_LongQuebec -33.83135 3.51387 -9.628 < 2e-16  
## Juris\_LongSaskatchewan -25.91140 4.37326 -5.925 3.21e-09  
## Juris\_LongYukon -45.65426 3.76580 -12.123 < 2e-16  
## Year -0.17184 0.09422 -1.824 0.06821  
## Fire\_Cause\_HumanLightning 63.20913 2.34469 26.958 < 2e-16  
## Fire\_Cause\_HumanUnspecified -6.79209 2.35367 -2.886 0.00391  
## Response.categoryModified -43.08178 1.83678 -23.455 < 2e-16  
## Response.categoryNone -43.16625 1.81365 -23.801 < 2e-16  
## Response.categoryUnspecified 20.27430 7.77966 2.606 0.00917  
##   
## (Intercept) \*   
## Juris\_LongBritish Columbia \*\*\*  
## Juris\_LongManitoba \*\*\*  
## Juris\_LongNational parks \*\*\*  
## Juris\_LongNew Brunswick \*\*\*  
## Juris\_LongNewfoundland and Labrador \*\*\*  
## Juris\_LongNorthwest Territories \*\*\*  
## Juris\_LongNova Scotia \*\*\*  
## Juris\_LongOntario \*\*\*  
## Juris\_LongPrince Edward Island \*\*\*  
## Juris\_LongQuebec \*\*\*  
## Juris\_LongSaskatchewan \*\*\*  
## Juris\_LongYukon \*\*\*  
## Year .   
## Fire\_Cause\_HumanLightning \*\*\*  
## Fire\_Cause\_HumanUnspecified \*\*   
## Response.categoryModified \*\*\*  
## Response.categoryNone \*\*\*  
## Response.categoryUnspecified \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 78.89 on 11425 degrees of freedom  
## Multiple R-squared: 0.1687, Adjusted R-squared: 0.1674   
## F-statistic: 128.8 on 18 and 11425 DF, p-value: < 2.2e-16

#Predicting on (Train Set)  
number\_predict\_2018 = predict(modelrm2018, newdata = dplyr::select(data2018, Juris\_Long, Year, Fire\_Cause\_Human, Response.category), type = "response" )  
  
print(number\_predict\_2018) #estimating y values, how many data points in new data.

## 1 2 3 4 5 6   
## 28.3230758 28.3230758 91.5322075 28.3230758 28.3230758 28.3230758   
## 7 8 9 10 11 12   
## 28.3230758 28.3230758 21.5309841 -14.7587091 -14.7587091 48.4504225   
## 13 14 15 16 17 18   
## -14.7587091 -14.7587091 -14.7587091 -14.7587091 -14.7587091 -21.5508009   
## 19 20 21 22 23 24   
## -14.8431751 -14.8431751 48.3659565 -14.8431751 -14.8431751 -14.8431751   
## 25 26 27 28 29 30   
## -14.8431751 -14.8431751 -21.6352669 130.4390409 165.3478193 92.3632654   
## 31 32 33 34 35 36   
## 86.7189661 93.5640908 85.4393577 94.5023057 105.5249741 75.6247112   
## 37 38 39 40 41 42   
## 96.6076917 104.5276416 84.7847832 67.2299093 102.1386877 29.1541337   
## 43 44 45 46 47 48   
## 23.5098345 30.3549592 22.2302261 31.2931740 42.3158425 12.4155795   
## 49 50 51 52 53 54   
## 33.3985601 41.3185099 21.5756515 87.3572560 122.2660343 49.2814804   
## 55 56 57 58 59 60   
## 43.6371812 50.4823059 42.3575728 51.4205207 62.4431892 32.5429262   
## 61 62 63 64 65 66   
## 53.5259068 61.4458566 41.7029982 24.1481243 59.0569027 -13.9276512   
## 67 68 69 70 71 72   
## -19.5719504 -12.7268258 -20.8515589 -11.7886109 -0.7659424 -30.6662054   
## 73 74 75   
## -9.6832249 -1.7632750 -21.5061334

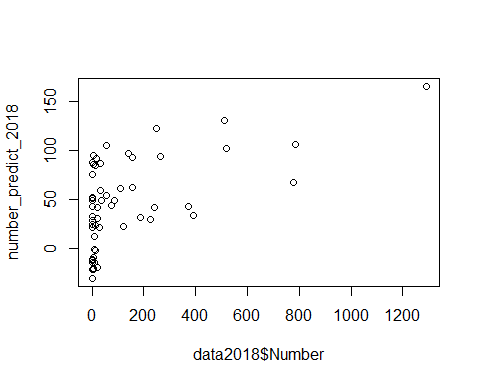
length(number\_predict\_2018) #75 data points for 2018

## [1] 75

length(data2018$Number) #75 data points in 2018

## [1] 75

plot(data2018$Number, number\_predict\_2018) #comparing prediction with actual results



#plotting (actual y values of 2018) vs the (estimated y values of 2018)  
#We are predicting less number of forest fires.  
#Model only captures 18% of the variation.  
#We are over estimating.  
  
# The MSE of the model.   
mse\_18 = mean((number\_predict\_2018 - data2018$Number)^2)  
mse\_18 #[1] 40489.22

## [1] 40489.22

# The RMSE of the model. ORIGINAL UNITS, NUMBER OF FIRES  
# How many forest fires we are off by.  
# We most likely overpredicted by 200 fires  
rmse\_18 = sqrt(mean((number\_predict\_2018 - data2018$Number)^2))  
rmse\_18 #[1] 201.2193

## [1] 201.2193

# The RMSE of the model. ORIGINAL UNITS, NUMBER OF FIRES  
# How many forest fires we are off by.  
# We most likely overpredicted by 200 fires  
rmse\_18\_v2 = (mean((number\_predict\_2018 - data2018$Number)^2))^.5  
rmse\_18\_v2 #[1] 201.2193

## [1] 201.2193

## 35.RMSE Prediction for Years

# Function that calculates the rmse for the year specified. The rmse is calculated as the number of predicted fires minus the actual number of fires in that year and taking the squared mean of the differnce.   
rmse\_year\_function = function(data\_set, year\_predict, model\_formula = Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Response.category ) {  
   
 data\_test = filter(data\_set, Year == year\_predict)  
 data\_train = filter(data\_set, Year < year\_predict)  
   
 model = lm( model\_formula, data = data\_train )  
   
 RMSE = sqrt(mean((predict(model, newdata = dplyr::select(data\_test, Juris\_Long, Year, Fire\_Cause\_Human, Response.category)) - data\_test$Number)^2) )  
   
 return(data.frame( Year = year\_predict, RMSE = RMSE))  
   
}  
   
result18 = rmse\_year\_function(data\_set = data, year\_predict = 2018)   
result18

## Year RMSE  
## 1 2018 201.2193

result17 = rmse\_year\_function(data\_set = data, year\_predict = 2017)   
result17

## Year RMSE  
## 1 2017 46.09691

result16 = rmse\_year\_function(data\_set = data, year\_predict = 2016)   
result16

## Year RMSE  
## 1 2016 44.43549

final\_result = rbind(result18, rmse\_year\_function(data\_set = data, year\_predict = 2017) )  
final\_result

## Year RMSE  
## 1 2018 201.21932  
## 2 2017 46.09691

class(rmse\_year\_function(data\_set = data, year\_predict = 2018))

## [1] "data.frame"

rmse\_year\_function(data\_set = data, year\_predict = 2017)

## Year RMSE  
## 1 2017 46.09691

rmse\_year\_function(data\_set = data, year\_predict = 2016)

## Year RMSE  
## 1 2016 44.43549

final\_result = NULL  
class(final\_result)

## [1] "NULL"

for( i in 1991:2018){  
   
 rmse\_i = rmse\_year\_function(data\_set = data, year\_predict = i)   
   
 final\_result = rbind(final\_result, rmse\_i)  
   
}

## Warning in predict.lm(model, newdata = dplyr::select(data\_test,  
## Juris\_Long, : prediction from a rank-deficient fit may be misleading

dim(final\_result)

## [1] 28 2

class(final\_result)

## [1] "data.frame"

final\_result

## Year RMSE  
## 1 1991 76.09586  
## 2 1992 120.59318  
## 3 1993 52.57477  
## 4 1994 147.28320  
## 5 1995 64.81554  
## 6 1996 53.45090  
## 7 1997 50.71129  
## 8 1998 123.27760  
## 9 1999 63.30460  
## 10 2000 54.85529  
## 11 2001 62.75889  
## 12 2002 70.21400  
## 13 2003 81.67327  
## 14 2004 88.99145  
## 15 2005 76.61250  
## 16 2006 104.14469  
## 17 2007 59.30258  
## 18 2008 68.97695  
## 19 2009 107.81176  
## 20 2010 64.48858  
## 21 2011 43.75223  
## 22 2012 66.30571  
## 23 2013 62.14653  
## 24 2014 52.52352  
## 25 2015 76.63914  
## 26 2016 44.43549  
## 27 2017 46.09691  
## 28 2018 201.21932

## 36.Prepare Time Series Analysis.

# install.packages("tseries")  
require(forecast) #for forecast function

## Loading required package: forecast

## Warning: package 'forecast' was built under R version 3.5.3

##   
## Attaching package: 'forecast'

## The following object is masked from 'package:caretEnsemble':  
##   
## autoplot

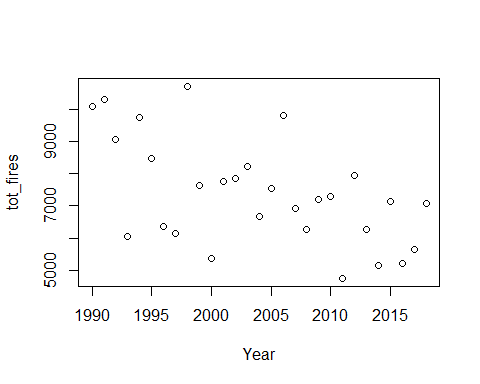
## The following object is masked from 'package:ggpubr':  
##   
## gghistogram

require(tseries)

## Loading required package: tseries

## Warning: package 'tseries' was built under R version 3.5.3

time\_s = data %>% group\_by(Year) %>% summarise(tot\_fires = sum(Number))  
plot(time\_s) # Plot ALL, group by years



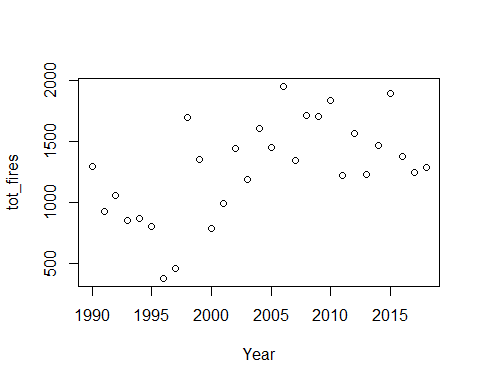
length(unique(data$Year)) #[1] 29

## [1] 29

###################################################################################################################################  
# Subset data for Alberta  
time\_s\_alberta = subset(data, Juris\_Long == "Alberta")  
head(time\_s\_alberta)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry AB 22 Intensive Full  
## 2: Forest industry AB 14 Intensive Full  
## 3: Forest industry AB 12 Intensive Full  
## 4: Forest industry AB 11 Intensive Full  
## 5: Forest industry AB 13 Intensive Full  
## 6: Forest industry AB 14 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 1990 Alberta People Early 90s 1990s Prairie Region  
## 2: 1991 Alberta People Early 90s 1990s Prairie Region  
## 3: 1992 Alberta People Early 90s 1990s Prairie Region  
## 4: 1993 Alberta People Early 90s 1990s Prairie Region  
## 5: 1994 Alberta People Early 90s 1990s Prairie Region  
## 6: 1995 Alberta People Early 90s 1990s Prairie Region  
## Fire\_Cause\_Human  
## 1: Human  
## 2: Human  
## 3: Human  
## 4: Human  
## 5: Human  
## 6: Human

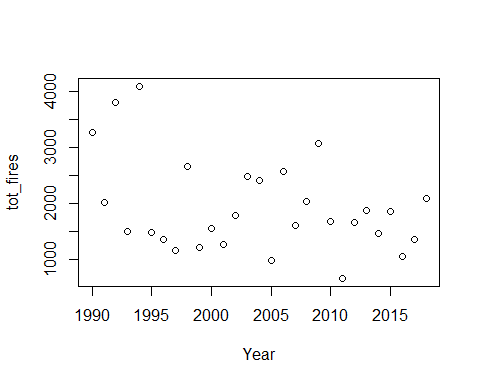
# Group subsetted data for Alberta  
plot\_time\_s\_alberta = time\_s\_alberta %>% group\_by(Year) %>% summarise(tot\_fires = sum(Number))  
plot(plot\_time\_s\_alberta) # Plot Alberta



###################################################################################################################################  
# Subset data for BC  
time\_s\_BC = subset(data, Juris\_Long == "British Columbia")  
head(time\_s\_BC)

## Cause Jurisdiction Number Protection.zone Response.category  
## 1: Forest industry BC 185 Intensive Full  
## 2: Forest industry BC 143 Intensive Full  
## 3: Forest industry BC 152 Intensive Full  
## 4: Forest industry BC 149 Intensive Full  
## 5: Forest industry BC 113 Intensive Full  
## 6: Forest industry BC 134 Intensive Full  
## Year Juris\_Long Cause\_Grouped Time1 Time2 Region  
## 1: 1990 British Columbia People Early 90s 1990s Pacific Region  
## 2: 1991 British Columbia People Early 90s 1990s Pacific Region  
## 3: 1992 British Columbia People Early 90s 1990s Pacific Region  
## 4: 1993 British Columbia People Early 90s 1990s Pacific Region  
## 5: 1994 British Columbia People Early 90s 1990s Pacific Region  
## 6: 1995 British Columbia People Early 90s 1990s Pacific Region  
## Fire\_Cause\_Human  
## 1: Human  
## 2: Human  
## 3: Human  
## 4: Human  
## 5: Human  
## 6: Human

# Group subsetted data for Alberta  
plot\_time\_s\_BC = time\_s\_BC %>% group\_by(Year) %>% summarise(tot\_fires = sum(Number))  
plot(plot\_time\_s\_BC) # Plot Alberta



###################################################################################################################################

## 37.Model: Time Series Analysis.

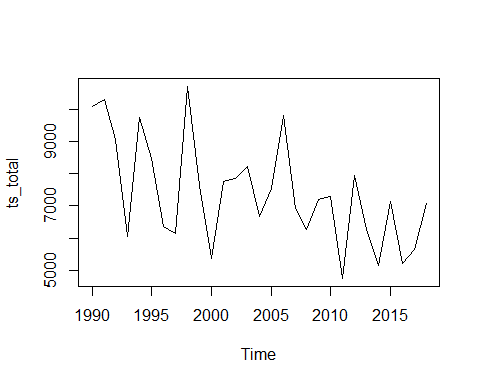
# One liner below replaces the four lines below.  
time\_s$y = time\_s$tot\_fires  
time\_series\_data = time\_s  
y = "Number of Total Fires"  
y\_time\_series = time\_series\_data$y  
  
#########################################################  
  
y\_time\_series = time\_s$tot\_fires  
class(y\_time\_series) # [1] "integer"

## [1] "integer"

#ts\_total = ts(y\_time\_series, start = 1990, end = 2018, frequency = 1)  
ts\_total = ts(y\_time\_series, start = 1990, frequency = 1)  
str(ts\_total)

## Time-Series [1:29] from 1990 to 2018: 10111 10327 9068 6043 9763 8486 6349 6148 10723 7633 ...

plot(ts\_total)



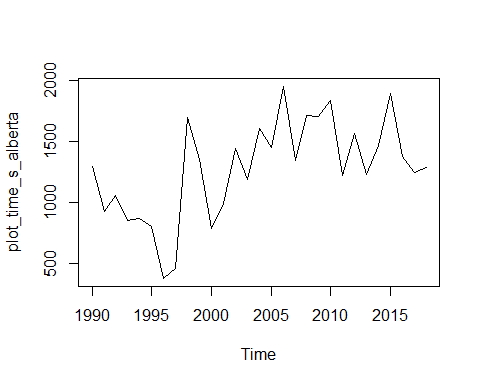
###################################################################################################################################  
# Alberta  
y\_time\_series <- plot\_time\_s\_alberta$tot\_fires  
class(y\_time\_series) # [1] "integer"

## [1] "integer"

#ts\_total = ts(y\_time\_series, start = 1990, end = 2018, frequency = 1)  
plot\_time\_s\_alberta <- ts(y\_time\_series, start = 1990, frequency = 1)  
str(plot\_time\_s\_alberta)

## Time-Series [1:29] from 1990 to 2018: 1296 923 1055 848 872 803 376 456 1698 1355 ...

plot(plot\_time\_s\_alberta)



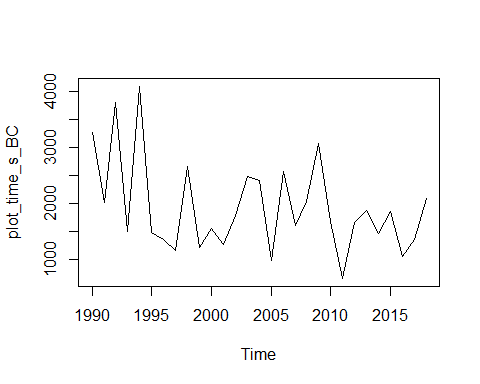
###################################################################################################################################  
# BC  
y\_time\_series <- plot\_time\_s\_BC$tot\_fires  
class(y\_time\_series) # [1] "integer"

## [1] "integer"

#ts\_total = ts(y\_time\_series, start = 1990, end = 2018, frequency = 1)  
plot\_time\_s\_BC <- ts(y\_time\_series, start = 1990, frequency = 1)  
str(plot\_time\_s\_BC)

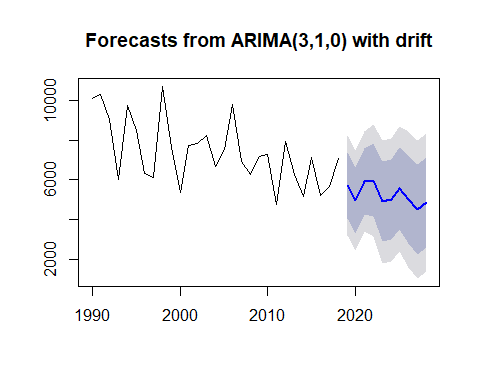
## Time-Series [1:29] from 1990 to 2018: 3255 2013 3805 1497 4088 1474 1346 1161 2662 1214 ...

plot(plot\_time\_s\_BC)

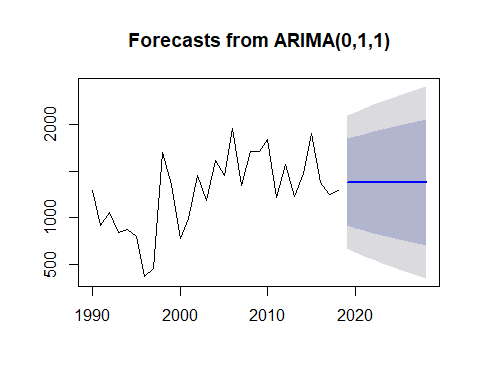


## 38.Model: Forecasting Time Series ARIMA Analysis.

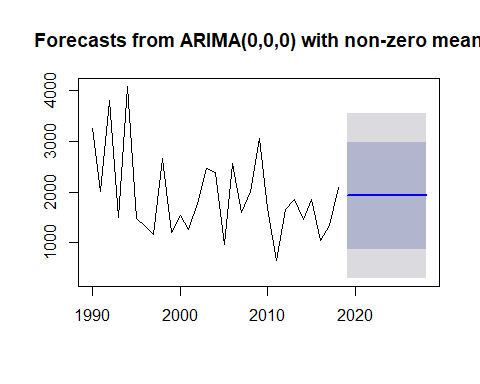
# Directly plotting a forecast of a model  
plot(forecast(auto.arima(ts\_total)))



# Forecast for Alberta  
plot(forecast(auto.arima(plot\_time\_s\_alberta)))



# Forecast for BC  
plot(forecast(auto.arima(plot\_time\_s\_BC)))



In time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. Both of these models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). ARIMA models are applied in some cases where data show evidence of non-stationarity, where an initial differencing step (corresponding to the “integrated” part of the model) can be applied one or more times to eliminate the non-stationarity.[1]

The AR part of ARIMA indicates that the evolving variable of interest is regressed on its own lagged (i.e., prior) values. The MA part indicates that the regression error is actually a linear combination of error terms whose values occurred contemporaneously and at various times in the past. The I (for “integrated”) indicates that the data values have been replaced with the difference between their values and the previous values (and this differencing process may have been performed more than once). The purpose of each of these features is to make the model fit the data as well as possible.

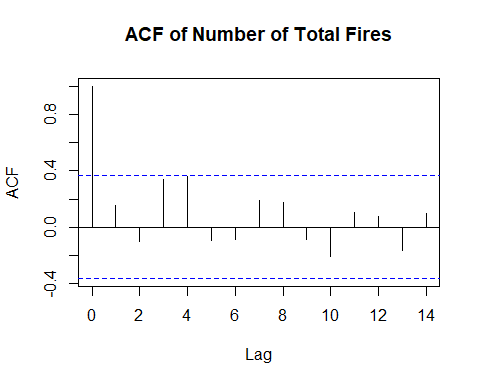
Non-seasonal ARIMA models are generally denoted ARIMA(p,d,q) where parameters p, d, and q are non-negative integers, p is the order (number of time lags) of the autoregressive model, d is the degree of differencing (the number of times the data have had past values subtracted), and q is the order of the moving-average model. Seasonal ARIMA models are usually denoted ARIMA(p,d,q)(P,D,Q)m, where m refers to the number of periods in each season, and the uppercase P,D,Q refer to the autoregressive, differencing, and moving average terms for the seasonal part of the ARIMA model.[2][3]

When two out of the three terms are zeros, the model may be referred to based on the non-zero parameter, dropping “AR”, “I” or “MA” from the acronym describing the model. For example, ARIMA (1,0,0) is AR(1), ARIMA(0,1,0) is I(1), and ARIMA(0,0,1) is MA(1).

## 39.Analysis of Stationarity

## Autocorrelation function (ACF) of the time series

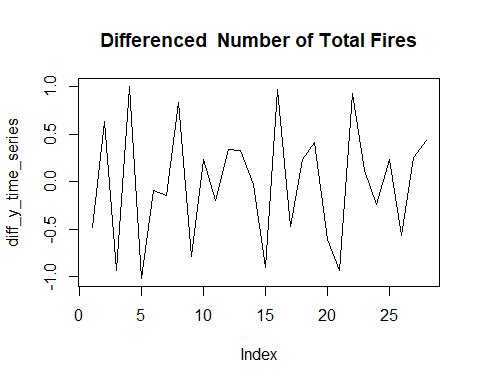
#checking for asumptions of model.  
acf(time\_series\_data$y, main = paste("ACF of", y)) #Autocorrelation (lags outside 95% confidence bands) implies stochastic or deterministic trend



## 40.Plot of the differenced time series

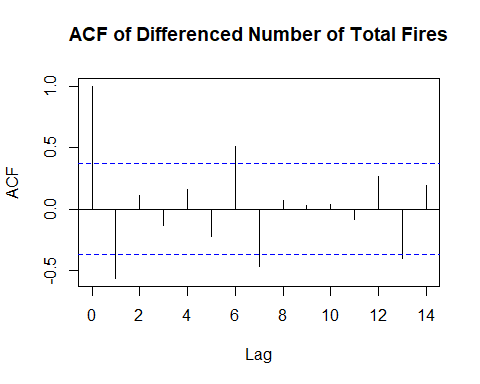
diff\_y\_time\_series = diff(log(y\_time\_series))  
plot(diff\_y\_time\_series, type = "line", main = paste('Differenced ', y)) #Visual proof of weakly stationary mean and variance

## Warning in plot.xy(xy, type, ...): plot type 'line' will be truncated to  
## first character



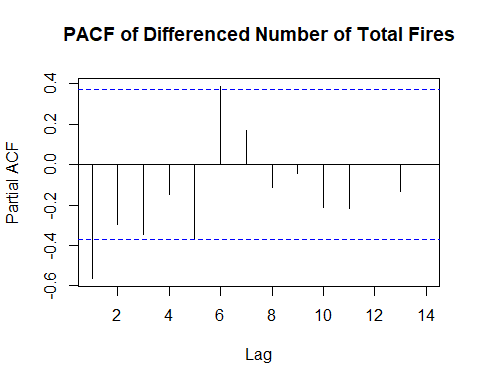
## 41.Autocorrelation function (ACF) of the differenced time series

acf(na.omit(diff\_y\_time\_series), main = paste('ACF of Differenced', y)) #Lags within the 95% confidence bands are considered noise and are statistically insignificant (trend removed). Lags outside 95% confidence bands determine the order of the MA(q) process -- an MA process of the order q has an ACF that cuts off after q lags.



## 42.Partial autocorrelation function (PACF) of the differenced time series

pacf(na.omit(diff\_y\_time\_series), main = paste('PACF of Differenced', y)) #Lags within the 95% confidence bands are considered noise and are statistically insignificant (trend removed). Lags outside 95% confidence bands determine the order of the AR(p) process -- an AR process of the order p has a PACF that cuts off after p lags.



## 43.Ljung-Box test of autocorrelation

Box.test(diff\_y\_time\_series, lag = log(length(diff\_y\_time\_series)), type = "Ljung-Box") #Formal test of autocorrelation and/or associated trend. The null hypothesis is that there is no autocorrelation (confirmed by p-value > 5%)

##   
## Box-Ljung test  
##   
## data: diff\_y\_time\_series  
## X-squared = 10.856, df = 3.3322, p-value = 0.01676

## 44.

## Augmented Dickey-Fuller test of stationarity

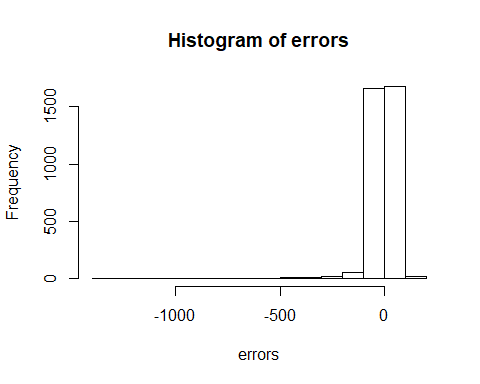
library(tseries)  
adf.test(na.omit(diff\_y\_time\_series)) #Formal test of stationarity. The null hypothesis is that the process is NOT stationary. We look for a p-value < 5% to reject the null hypothesis in favor of the alternative hypothesis (in which case the process is stationary)

##   
## Augmented Dickey-Fuller Test  
##   
## data: na.omit(diff\_y\_time\_series)  
## Dickey-Fuller = -3.9787, Lag order = 3, p-value = 0.02356  
## alternative hypothesis: stationary

## 40.Additional testing

## Variable Selection for Multiple Linear Regression in R

# Split the dataset to 70% of training and 30% of test sets.   
# We want to make sure that the training set and the test set do not have any common data points.  
set.seed(1235)  
  
train\_index <- sample(1:nrow(data), 0.7 \* nrow(data))  
train.set <- data[train\_index,]  
test.set <- data[-train\_index,]  
  
# Train our model on the training set  
model2 = lm(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set)  
#summary(model2)  
  
# Prediction on the test set  
prediction <- predict(model2, interval="prediction", newdata =test.set)  
#model2  
  
# Calculate error (prediction Number - test Number) in predictions and show the histogram of error  
errors <- prediction[,"fit"] - test.set$Number  
hist(errors)



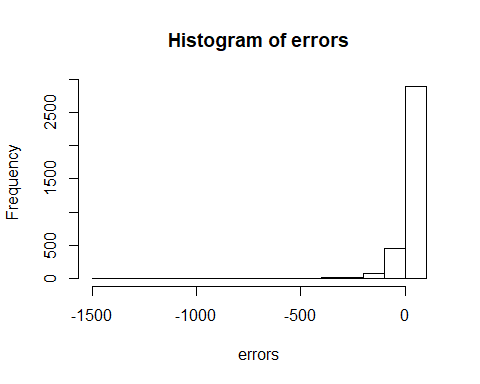
# Compute the root mean square error and find the percentage of cases with less than 25% error.  
# Calculate mean square error (mse) and find the percentage of cases with less than 25% error.  
  
rmse <- sqrt(sum((prediction[,"fit"] - test.set$Number)^2)/nrow(test.set))  
rel\_change <- 1 - ((test.set$Number - abs(errors)) / test.set$Number)  
  
pred25 <- table(rel\_change<0.25)["TRUE"] / nrow(test.set)  
paste("RMSE:", rmse)

## [1] "RMSE: 71.9922996373549"

paste("PRED(25):", pred25)

## [1] "PRED(25): 0.0306712962962963"

# Use simple linear regression model by using 'Juris\_Long' as an independent variable. Compare the results with the multiple linear regression.  
  
rn\_train <- sample(nrow(data), floor(nrow(data)\*0.7))  
train <- data[rn\_train,c("Number","Juris\_Long")]  
test <- data[-rn\_train,c("Number","Juris\_Long")]  
model\_ulm <- lm(Number~Juris\_Long, data=train)  
prediction <- predict(model\_ulm, interval="prediction", newdata =test)  
errors <- prediction[,"fit"] - test$Number  
hist(errors)



rmse <- sqrt(sum((prediction[,"fit"] - test$Number)^2)/nrow(test))  
rel\_change <- 1 - ((test$Number - abs(errors)) / test$Number)  
pred25 <- table(rel\_change<0.25)["TRUE"] / nrow(test)  
paste("RMSE:", rmse)

## [1] "RMSE: 72.5042950450685"

paste("PRED(25):", pred25)

## [1] "PRED(25): 0.0434027777777778"

# Both Pred(25) and RMSE values are better for multiple linear regression.

## 41.Additional testing

## Forward and Backward selection algorithm

library(MASS) # stepwise regression

## Warning: package 'MASS' was built under R version 3.5.3

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

#install.packages('leaps')  
library(leaps) # all subsets regression

## Warning: package 'leaps' was built under R version 3.5.3

#install.packages('FNN')  
library(FNN)

## Warning: package 'FNN' was built under R version 3.5.3

# Forward selection   
full <- lm(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = data)  
null <- lm(Number~1,data=data)  
stepF <- stepAIC(null, scope=list(lower=null, upper=full), direction= "forward", trace=TRUE)

## Start: AIC=103176.1  
## Number ~ 1  
##   
## Df Sum of Sq RSS AIC  
## + Response.category 3 5342703 84055596 102472  
## + Fire\_Cause\_Human 2 5023926 84374373 102514  
## + Juris\_Long 12 5043747 84354553 102531  
## + Protection.zone 2 1813104 87585195 102944  
## + Year 1 41537 89356763 103173  
## <none> 89398300 103176  
##   
## Step: AIC=102472.2  
## Number ~ Response.category  
##   
## Df Sum of Sq RSS AIC  
## + Fire\_Cause\_Human 2 5041208 79014389 101764  
## + Juris\_Long 12 4889175 79166421 101806  
## + Protection.zone 2 1511519 82544077 102267  
## + Year 1 36667 84018930 102469  
## <none> 84055596 102472  
##   
## Step: AIC=101763.8  
## Number ~ Response.category + Fire\_Cause\_Human  
##   
## Df Sum of Sq RSS AIC  
## + Juris\_Long 12 4892653 74121736 101051  
## + Protection.zone 2 1260564 77753824 101583  
## + Year 1 49320 78965069 101759  
## <none> 79014389 101764  
##   
## Step: AIC=101051.5  
## Number ~ Response.category + Fire\_Cause\_Human + Juris\_Long  
##   
## Df Sum of Sq RSS AIC  
## + Protection.zone 2 784684 73337051 100933  
## <none> 74121736 101051  
## + Year 1 4256 74117479 101053  
##   
## Step: AIC=100932.9  
## Number ~ Response.category + Fire\_Cause\_Human + Juris\_Long +   
## Protection.zone  
##   
## Df Sum of Sq RSS AIC  
## <none> 73337051 100933  
## + Year 1 11863 73325189 100933

summary(stepF)

##   
## Call:  
## lm(formula = Number ~ Response.category + Fire\_Cause\_Human +   
## Juris\_Long + Protection.zone, data = data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -185.32 -27.04 0.46 14.45 2742.77   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 70.923 3.108 22.820 < 2e-16  
## Response.categoryModified -43.883 1.851 -23.705 < 2e-16  
## Response.categoryNone -43.520 1.831 -23.773 < 2e-16  
## Response.categoryUnspecified 18.787 7.793 2.411 0.01593  
## Fire\_Cause\_HumanLightning 62.918 2.358 26.678 < 2e-16  
## Fire\_Cause\_HumanUnspecified -6.763 2.380 -2.842 0.00449  
## Juris\_LongBritish Columbia 36.385 4.297 8.468 < 2e-16  
## Juris\_LongManitoba -32.682 3.764 -8.682 < 2e-16  
## Juris\_LongNational parks -45.256 4.296 -10.535 < 2e-16  
## Juris\_LongNew Brunswick -39.076 5.759 -6.786 1.21e-11  
## Juris\_LongNewfoundland and Labrador -38.951 3.688 -10.562 < 2e-16  
## Juris\_LongNorthwest Territories -39.494 4.085 -9.669 < 2e-16  
## Juris\_LongNova Scotia -37.242 4.147 -8.980 < 2e-16  
## Juris\_LongOntario -18.003 3.734 -4.822 1.44e-06  
## Juris\_LongPrince Edward Island -56.388 4.947 -11.398 < 2e-16  
## Juris\_LongQuebec -27.496 3.676 -7.480 7.98e-14  
## Juris\_LongSaskatchewan -25.936 4.415 -5.874 4.38e-09  
## Juris\_LongYukon -41.284 3.883 -10.632 < 2e-16  
## Protection.zoneLimited -14.232 1.919 -7.417 1.28e-13  
## Protection.zoneUnspecified 92.720 11.615 7.983 1.57e-15  
##   
## (Intercept) \*\*\*  
## Response.categoryModified \*\*\*  
## Response.categoryNone \*\*\*  
## Response.categoryUnspecified \*   
## Fire\_Cause\_HumanLightning \*\*\*  
## Fire\_Cause\_HumanUnspecified \*\*   
## Juris\_LongBritish Columbia \*\*\*  
## Juris\_LongManitoba \*\*\*  
## Juris\_LongNational parks \*\*\*  
## Juris\_LongNew Brunswick \*\*\*  
## Juris\_LongNewfoundland and Labrador \*\*\*  
## Juris\_LongNorthwest Territories \*\*\*  
## Juris\_LongNova Scotia \*\*\*  
## Juris\_LongOntario \*\*\*  
## Juris\_LongPrince Edward Island \*\*\*  
## Juris\_LongQuebec \*\*\*  
## Juris\_LongSaskatchewan \*\*\*  
## Juris\_LongYukon \*\*\*  
## Protection.zoneLimited \*\*\*  
## Protection.zoneUnspecified \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 79.86 on 11499 degrees of freedom  
## Multiple R-squared: 0.1797, Adjusted R-squared: 0.1783   
## F-statistic: 132.5 on 19 and 11499 DF, p-value: < 2.2e-16

# We end up using all the variables. We set 'trace=TRUE' to see all the steps.  
  
# We can also use 'backward' elimination, which will start with 'full'.  
full <- lm(Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = data)  
stepB <- stepAIC(full, direction= "backward", trace=TRUE)

## Start: AIC=100933  
## Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone +   
## Response.category  
##   
## Df Sum of Sq RSS AIC  
## - Year 1 11863 73337051 100933  
## <none> 73325189 100933  
## - Protection.zone 2 792291 74117479 101053  
## - Juris\_Long 12 4369049 77694238 101576  
## - Fire\_Cause\_Human 2 4792069 78117257 101658  
## - Response.category 3 5105989 78431178 101702  
##   
## Step: AIC=100932.9  
## Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone + Response.category  
##   
## Df Sum of Sq RSS AIC  
## <none> 73337051 100933  
## - Protection.zone 2 784684 74121736 101051  
## - Juris\_Long 12 4416773 77753824 101583  
## - Fire\_Cause\_Human 2 4789846 78126898 101658  
## - Response.category 3 5107444 78444495 101702

summary(stepB)

##   
## Call:  
## lm(formula = Number ~ Juris\_Long + Fire\_Cause\_Human + Protection.zone +   
## Response.category, data = data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -185.32 -27.04 0.46 14.45 2742.77   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 70.923 3.108 22.820 < 2e-16  
## Juris\_LongBritish Columbia 36.385 4.297 8.468 < 2e-16  
## Juris\_LongManitoba -32.682 3.764 -8.682 < 2e-16  
## Juris\_LongNational parks -45.256 4.296 -10.535 < 2e-16  
## Juris\_LongNew Brunswick -39.076 5.759 -6.786 1.21e-11  
## Juris\_LongNewfoundland and Labrador -38.951 3.688 -10.562 < 2e-16  
## Juris\_LongNorthwest Territories -39.494 4.085 -9.669 < 2e-16  
## Juris\_LongNova Scotia -37.242 4.147 -8.980 < 2e-16  
## Juris\_LongOntario -18.003 3.734 -4.822 1.44e-06  
## Juris\_LongPrince Edward Island -56.388 4.947 -11.398 < 2e-16  
## Juris\_LongQuebec -27.496 3.676 -7.480 7.98e-14  
## Juris\_LongSaskatchewan -25.936 4.415 -5.874 4.38e-09  
## Juris\_LongYukon -41.284 3.883 -10.632 < 2e-16  
## Fire\_Cause\_HumanLightning 62.918 2.358 26.678 < 2e-16  
## Fire\_Cause\_HumanUnspecified -6.763 2.380 -2.842 0.00449  
## Protection.zoneLimited -14.232 1.919 -7.417 1.28e-13  
## Protection.zoneUnspecified 92.720 11.615 7.983 1.57e-15  
## Response.categoryModified -43.883 1.851 -23.705 < 2e-16  
## Response.categoryNone -43.520 1.831 -23.773 < 2e-16  
## Response.categoryUnspecified 18.787 7.793 2.411 0.01593  
##   
## (Intercept) \*\*\*  
## Juris\_LongBritish Columbia \*\*\*  
## Juris\_LongManitoba \*\*\*  
## Juris\_LongNational parks \*\*\*  
## Juris\_LongNew Brunswick \*\*\*  
## Juris\_LongNewfoundland and Labrador \*\*\*  
## Juris\_LongNorthwest Territories \*\*\*  
## Juris\_LongNova Scotia \*\*\*  
## Juris\_LongOntario \*\*\*  
## Juris\_LongPrince Edward Island \*\*\*  
## Juris\_LongQuebec \*\*\*  
## Juris\_LongSaskatchewan \*\*\*  
## Juris\_LongYukon \*\*\*  
## Fire\_Cause\_HumanLightning \*\*\*  
## Fire\_Cause\_HumanUnspecified \*\*   
## Protection.zoneLimited \*\*\*  
## Protection.zoneUnspecified \*\*\*  
## Response.categoryModified \*\*\*  
## Response.categoryNone \*\*\*  
## Response.categoryUnspecified \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 79.86 on 11499 degrees of freedom  
## Multiple R-squared: 0.1797, Adjusted R-squared: 0.1783   
## F-statistic: 132.5 on 19 and 11499 DF, p-value: < 2.2e-16

# We end up using all the variables.

## 42.Additional testing

## Variable selection using automatic methods

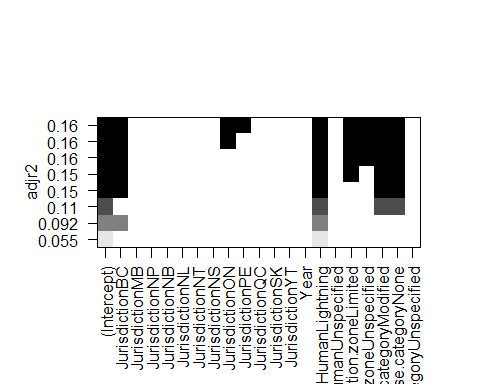
# Best combination of the 6 attributes.  
subsets<-regsubsets(Number ~ Jurisdiction + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = data, nbest=1)  
sub.sum <- summary(subsets)  
as.data.frame(sub.sum$outmat)

## JurisdictionBC JurisdictionMB JurisdictionNP JurisdictionNB  
## 1 ( 1 )   
## 2 ( 1 ) \*   
## 3 ( 1 )   
## 4 ( 1 ) \*   
## 5 ( 1 ) \*   
## 6 ( 1 ) \*   
## 7 ( 1 ) \*   
## 8 ( 1 ) \*   
## JurisdictionNL JurisdictionNT JurisdictionNS JurisdictionON  
## 1 ( 1 )   
## 2 ( 1 )   
## 3 ( 1 )   
## 4 ( 1 )   
## 5 ( 1 )   
## 6 ( 1 )   
## 7 ( 1 ) \*  
## 8 ( 1 ) \*  
## JurisdictionPE JurisdictionQC JurisdictionSK JurisdictionYT Year  
## 1 ( 1 )   
## 2 ( 1 )   
## 3 ( 1 )   
## 4 ( 1 )   
## 5 ( 1 )   
## 6 ( 1 )   
## 7 ( 1 )   
## 8 ( 1 ) \*   
## Fire\_Cause\_HumanLightning Fire\_Cause\_HumanUnspecified  
## 1 ( 1 ) \*   
## 2 ( 1 ) \*   
## 3 ( 1 ) \*   
## 4 ( 1 ) \*   
## 5 ( 1 ) \*   
## 6 ( 1 ) \*   
## 7 ( 1 ) \*   
## 8 ( 1 ) \*   
## Protection.zoneLimited Protection.zoneUnspecified  
## 1 ( 1 )   
## 2 ( 1 )   
## 3 ( 1 )   
## 4 ( 1 )   
## 5 ( 1 ) \*   
## 6 ( 1 ) \* \*  
## 7 ( 1 ) \* \*  
## 8 ( 1 ) \* \*  
## Response.categoryModified Response.categoryNone  
## 1 ( 1 )   
## 2 ( 1 )   
## 3 ( 1 ) \* \*  
## 4 ( 1 ) \* \*  
## 5 ( 1 ) \* \*  
## 6 ( 1 ) \* \*  
## 7 ( 1 ) \* \*  
## 8 ( 1 ) \* \*  
## Response.categoryUnspecified  
## 1 ( 1 )   
## 2 ( 1 )   
## 3 ( 1 )   
## 4 ( 1 )   
## 5 ( 1 )   
## 6 ( 1 )   
## 7 ( 1 )   
## 8 ( 1 )

summary(subsets)

## Subset selection object  
## Call: regsubsets.formula(Number ~ Jurisdiction + Year + Fire\_Cause\_Human +   
## Protection.zone + Response.category, data = data, nbest = 1)  
## 20 Variables (and intercept)  
## Forced in Forced out  
## JurisdictionBC FALSE FALSE  
## JurisdictionMB FALSE FALSE  
## JurisdictionNP FALSE FALSE  
## JurisdictionNB FALSE FALSE  
## JurisdictionNL FALSE FALSE  
## JurisdictionNT FALSE FALSE  
## JurisdictionNS FALSE FALSE  
## JurisdictionON FALSE FALSE  
## JurisdictionPE FALSE FALSE  
## JurisdictionQC FALSE FALSE  
## JurisdictionSK FALSE FALSE  
## JurisdictionYT FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 8  
## Selection Algorithm: exhaustive  
## JurisdictionBC JurisdictionMB JurisdictionNP JurisdictionNB  
## 1 ( 1 ) " " " " " " " "   
## 2 ( 1 ) "\*" " " " " " "   
## 3 ( 1 ) " " " " " " " "   
## 4 ( 1 ) "\*" " " " " " "   
## 5 ( 1 ) "\*" " " " " " "   
## 6 ( 1 ) "\*" " " " " " "   
## 7 ( 1 ) "\*" " " " " " "   
## 8 ( 1 ) "\*" " " " " " "   
## JurisdictionNL JurisdictionNT JurisdictionNS JurisdictionON  
## 1 ( 1 ) " " " " " " " "   
## 2 ( 1 ) " " " " " " " "   
## 3 ( 1 ) " " " " " " " "   
## 4 ( 1 ) " " " " " " " "   
## 5 ( 1 ) " " " " " " " "   
## 6 ( 1 ) " " " " " " " "   
## 7 ( 1 ) " " " " " " "\*"   
## 8 ( 1 ) " " " " " " "\*"   
## JurisdictionPE JurisdictionQC JurisdictionSK JurisdictionYT Year  
## 1 ( 1 ) " " " " " " " " " "   
## 2 ( 1 ) " " " " " " " " " "   
## 3 ( 1 ) " " " " " " " " " "   
## 4 ( 1 ) " " " " " " " " " "   
## 5 ( 1 ) " " " " " " " " " "   
## 6 ( 1 ) " " " " " " " " " "   
## 7 ( 1 ) " " " " " " " " " "   
## 8 ( 1 ) "\*" " " " " " " " "   
## Fire\_Cause\_HumanLightning Fire\_Cause\_HumanUnspecified  
## 1 ( 1 ) "\*" " "   
## 2 ( 1 ) "\*" " "   
## 3 ( 1 ) "\*" " "   
## 4 ( 1 ) "\*" " "   
## 5 ( 1 ) "\*" " "   
## 6 ( 1 ) "\*" " "   
## 7 ( 1 ) "\*" " "   
## 8 ( 1 ) "\*" " "   
## Protection.zoneLimited Protection.zoneUnspecified  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## 5 ( 1 ) "\*" " "   
## 6 ( 1 ) "\*" "\*"   
## 7 ( 1 ) "\*" "\*"   
## 8 ( 1 ) "\*" "\*"   
## Response.categoryModified Response.categoryNone  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) "\*" "\*"   
## 4 ( 1 ) "\*" "\*"   
## 5 ( 1 ) "\*" "\*"   
## 6 ( 1 ) "\*" "\*"   
## 7 ( 1 ) "\*" "\*"   
## 8 ( 1 ) "\*" "\*"   
## Response.categoryUnspecified  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "   
## 5 ( 1 ) " "   
## 6 ( 1 ) " "   
## 7 ( 1 ) " "   
## 8 ( 1 ) " "

plot(subsets, scale="adjr2")



# In the output \* denotes the included variables.   
# The best combination of 4 attributes is: 'Fire\_Cause\_HumanLightning', 'JurisdictionBC', 'Response.categoryModified' and 'Response.categoryNone'.   
# The best combination of 5 attributes is: 'Fire\_Cause\_HumanLightning', 'JurisdictionBC', 'Response.categoryModified', 'Response.categoryNone' and Protection.zoneLimited

## 43.

## Additional testing

# Prediction using k Nearest Neighbor Regression

set.seed(1235)  
unique(data$Fire\_Cause\_Human) # [1] Human Lightning Unspecified

## [1] "Human" "Lightning" "Unspecified"

unique(data$Juris\_Long) # [1] Alberta British Columbia ...

## [1] Alberta British Columbia   
## [3] Manitoba National parks   
## [5] New Brunswick Newfoundland and Labrador  
## [7] Nova Scotia Ontario   
## [9] Prince Edward Island Quebec   
## [11] Saskatchewan Yukon   
## [13] Northwest Territories   
## 13 Levels: Alberta British Columbia Manitoba ... Yukon

unique(data$Protection.zone) # [1] Intensive Limited Unspecified

## [1] Intensive Limited Unspecified  
## Levels: Intensive Limited Unspecified

unique(data$Response.category) # [1] Full Modified None Unspecified

## [1] Full Modified None Unspecified  
## Levels: Full Modified None Unspecified

unique(data$Year) # 1990 1991 1992 1993 1994 1995 1996

## [1] 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003  
## [15] 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017  
## [29] 2018

# rm(data1)  
# rm(dataset.numeric)  
  
data1 <- data.frame(data[, c("Number","Fire\_Cause\_Human", "Juris\_Long", "Protection.zone", "Response.category", "Year")])  
  
#data1 <- data[, c("Number","Fire\_Cause\_Human", "Juris\_Long", "Protection.zone", "Response.category", "Year")]  
  
head(data1)

## Number Fire\_Cause\_Human Juris\_Long Protection.zone Response.category  
## 1 22 Human Alberta Intensive Full  
## 2 14 Human Alberta Intensive Full  
## 3 12 Human Alberta Intensive Full  
## 4 11 Human Alberta Intensive Full  
## 5 13 Human Alberta Intensive Full  
## 6 14 Human Alberta Intensive Full  
## Year  
## 1 1990  
## 2 1991  
## 3 1992  
## 4 1993  
## 5 1994  
## 6 1995

tail(data1)

## Number Fire\_Cause\_Human Juris\_Long Protection.zone  
## 11514 0 Human Nova Scotia Unspecified  
## 11515 11 Human Ontario Unspecified  
## 11516 0 Human Prince Edward Island Unspecified  
## 11517 6 Human Quebec Unspecified  
## 11518 13 Human Saskatchewan Unspecified  
## 11519 6 Human Yukon Unspecified  
## Response.category Year  
## 11514 Modified 2018  
## 11515 Modified 2018  
## 11516 Modified 2018  
## 11517 Modified 2018  
## 11518 Modified 2018  
## 11519 Modified 2018

str(data1)

## 'data.frame': 11519 obs. of 6 variables:  
## $ Number : int 22 14 12 11 13 14 8 29 10 20 ...  
## $ Fire\_Cause\_Human : chr "Human" "Human" "Human" "Human" ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...

summary(data1)

## Number Fire\_Cause\_Human Juris\_Long   
## Min. : 0.00 Length:11519 Quebec :1516   
## 1st Qu.: 0.00 Class :character Newfoundland and Labrador:1486   
## Median : 0.00 Mode :character Ontario :1368   
## Mean : 18.64 Manitoba :1304   
## 3rd Qu.: 4.00 Yukon :1059   
## Max. :2913.00 Northwest Territories : 773   
## (Other) :4013   
## Protection.zone Response.category Year   
## Intensive :8190 Full :4204 Min. :1990   
## Limited :3281 Modified :3521 1st Qu.:1997   
## Unspecified: 48 None :3681 Median :2004   
## Unspecified: 113 Mean :2004   
## 3rd Qu.:2011   
## Max. :2018   
##

# new\_fire <- fires[, c("Cause", "Jurisdiction", "Number", "Protection.zone", "Response.category", "Year")]  
  
# sum(is.na(data1$Number))  
# sum(is.na(data1$Fire\_Cause\_Human))  
# sum(is.na(data1$Juris\_Long))  
# sum(is.na(data1$Protection.zone))  
# sum(is.na(data1$Response.category))  
# sum(is.na(data1$Year))  
  
# addlist1 <- c(0,'Lightning','Alberta','Intensive','Full','1996')  
# dataset <- rbind(data1, addlist1)  
  
dataset <- rbind(data1, c(0,'Lightning','Alberta','Intensive','Full','2020'))  
  
str(dataset)

## 'data.frame': 11520 obs. of 6 variables:  
## $ Number : chr "22" "14" "12" "11" ...  
## $ Fire\_Cause\_Human : chr "Human" "Human" "Human" "Human" ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : chr "1990" "1991" "1992" "1993" ...

unique(dataset$Number) # [1] Human Lightning Unspecified

## [1] "22" "14" "12" "11" "13" "8" "29" "10" "20" "18"   
## [11] "26" "19" "41" "48" "30" "15" "23" "32" "185" "143"   
## [21] "152" "149" "113" "134" "130" "62" "3" "7" "4" "5"   
## [31] "6" "21" "1" "0" "2" "24" "27" "17" "37" "16"   
## [41] "9" "38" "25" "53" "47" "33" "58" "49" "36" "59"   
## [51] "71" "66" "52" "81" "60" "34" "55" "40" "45" "50"   
## [61] "28" "73" "56" "42" "31" "46" "69" "39" "123" "121"   
## [71] "104" "146" "102" "94" "256" "145" "181" "93" "147" "133"   
## [81] "141" "125" "79" "257" "116" "266" "262" "300" "352" "202"   
## [91] "173" "302" "235" "250" "200" "107" "168" "76" "131" "87"   
## [101] "97" "99" "105" "64" "61" "327" "35" "67" "43" "83"   
## [111] "80" "91" "106" "63" "120" "84" "100" "88" "86" "96"   
## [121] "77" "115" "151" "44" "68" "914" "466" "626" "517" "499"   
## [131] "357" "217" "242" "1192" "890" "429" "490" "868" "527" "732"   
## [141] "432" "746" "513" "779" "566" "741" "215" "436" "287" "563"   
## [151] "772" "514" "416" "2015" "759" "2344" "609" "2913" "342" "723"   
## [161] "675" "1773" "585" "842" "479" "870" "1513" "1716" "384" "1536"  
## [171] "912" "1175" "2184" "992" "209" "944" "1151" "726" "1237" "486"   
## [181] "782" "213" "178" "361" "101" "323" "110" "176" "118" "161"   
## [191] "129" "197" "190" "451" "988" "191" "188" "831" "344" "545"   
## [201] "1141" "282" "167" "886" "533" "417" "98" "1249" "1433" "467"   
## [211] "70" "305" "580" "780" "103" "82" "135" "249" "211" "421"   
## [221] "89" "92" "446" "493" "299" "157" "220" "238" "119" "683"   
## [231] "241" "216" "244" "338" "201" "828" "495" "180" "317" "440"   
## [241] "289" "196" "225" "207" "136" "126" "75" "239" "74" "65"   
## [251] "276" "165" "226" "51" "78" "218" "306" "260" "460" "171"   
## [261] "156" "184" "195" "163" "272" "194" "273" "179" "72" "54"   
## [271] "95" "90" "108" "155" "142" "148" "224" "240" "243" "111"   
## [281] "124" "117" "219" "263" "158" "162" "138" "109" "232" "313"   
## [291] "275" "280" "355" "321" "290" "381" "403" "343" "515" "458"   
## [301] "392" "340" "315" "322" "407" "153" "85" "137" "114" "174"   
## [311] "159" "127" "285" "144" "292" "296" "198" "140" "252" "164"   
## [321] "222" "337" "221" "258" "170" "246" "347" "139" "166" "183"   
## [331] "204" "154" "426" "478" "505" "362" "398" "388" "150" "228"   
## [341] "377" "112" "187" "128" "172" "57" "233" "270" "160" "186"   
## [351] "132" "203" "208" "253" "288" "316" "396" "530" "269" "511"   
## [361] "1290" "787" "777" "518" "373" "247"

unique(dataset$Fire\_Cause\_Human) # [1] Human Lightning Unspecified

## [1] "Human" "Lightning" "Unspecified"

unique(dataset$Juris\_Long) # [1] Alberta British Columbia ...

## [1] Alberta British Columbia   
## [3] Manitoba National parks   
## [5] New Brunswick Newfoundland and Labrador  
## [7] Nova Scotia Ontario   
## [9] Prince Edward Island Quebec   
## [11] Saskatchewan Yukon   
## [13] Northwest Territories   
## 13 Levels: Alberta British Columbia Manitoba ... Yukon

unique(dataset$Protection.zone) # [1] Intensive Limited Unspecified

## [1] Intensive Limited Unspecified  
## Levels: Intensive Limited Unspecified

unique(dataset$Response.category) # [1] Full Modified None Unspecified

## [1] Full Modified None Unspecified  
## Levels: Full Modified None Unspecified

unique(dataset$Year) # 1990 1991 1992 1993 1994 1995 1996

## [1] "1990" "1991" "1992" "1993" "1994" "1995" "1996" "1997" "1998" "1999"  
## [11] "2000" "2001" "2002" "2003" "2004" "2005" "2006" "2007" "2008" "2009"  
## [21] "2010" "2011" "2012" "2013" "2014" "2015" "2016" "2017" "2018"

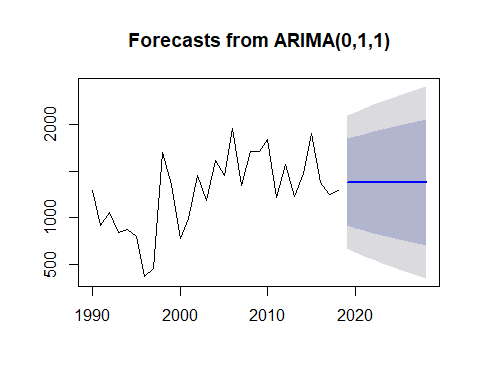
dataset[is.na(dataset)] <- 0  
dataset.numeric <- sapply( dataset[,1:6], as.numeric )

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

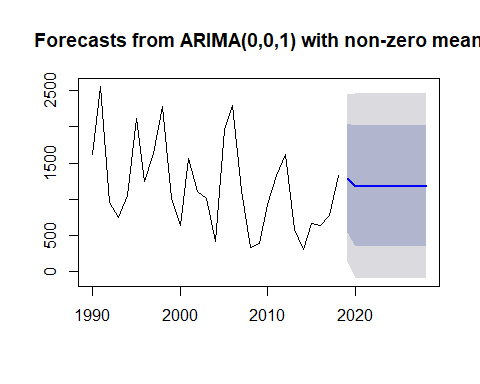
#Should convert data to numeric to use knn.reg  
  
#Remove NAs from matrix, replace with zero  
dataset.numeric[is.na(dataset.numeric)] <- 0  
#Matrix now has NAs removed, now replaced with zero  
  
#unique(dataset.numeric$Fire\_Cause\_Human)   
  
dataset.numeric <- as.data.frame(dataset.numeric)  
prediction <- knn.reg(dataset.numeric[1:nrow(data1),-1],  
test = dataset.numeric[nrow(data1)+1,-1],  
 dataset.numeric[1:nrow(data1),]$Number, k = 7 , algorithm="kd\_tree")  
  
prediction$pred ## [1] 268.7143

## Additional time series data

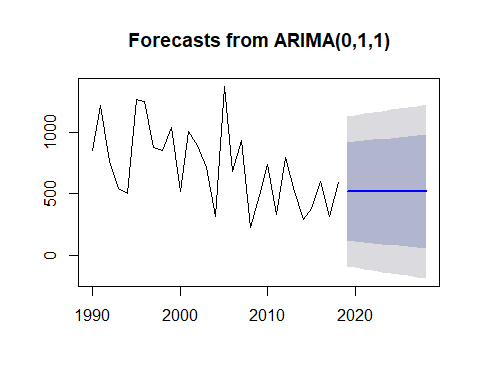
ts\_forecast\_cov = function(data, covariate , subset = "Province" ) {  
 # Subset = "Cause" or "Province"  
 # Covariate = "Human", "Lightning","Unspecified" or one of the Provinces.  
  
 if (subset == "Cause"){  
 time\_s = subset(data,Fire\_Cause\_Human == covariate) %>% group\_by(Year) %>% summarise(tot\_fires = sum(Number))  
 } else {  
 time\_s = subset(data, Juris\_Long == covariate) %>% group\_by(Year) %>% summarise(tot\_fires = sum(Number))  
 }  
 y\_time\_series = time\_s$tot\_fires  
 ts\_total = ts(y\_time\_series, start = 1990, frequency = 1)  
 plot(forecast(auto.arima(ts\_total)))  
}  
#y = "Number of Total Fires"  
# Different subsets and covariates of the data.  
ts\_forecast\_cov(data, covariate = "Alberta", subset = "Province" )



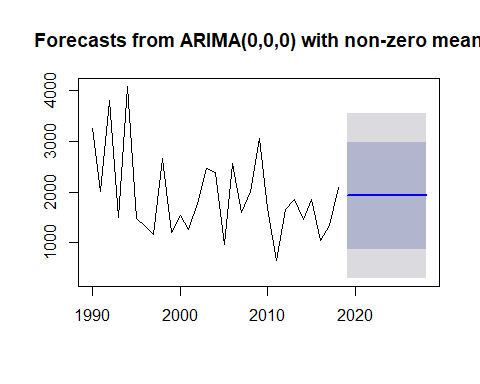
ts\_forecast\_cov(data, covariate = "Ontario", subset = "Province" )



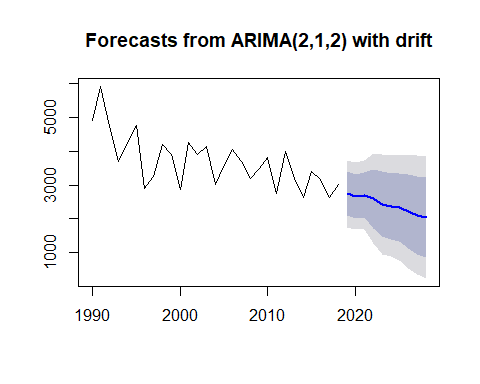
ts\_forecast\_cov(data, covariate = "Quebec", subset = "Province" )



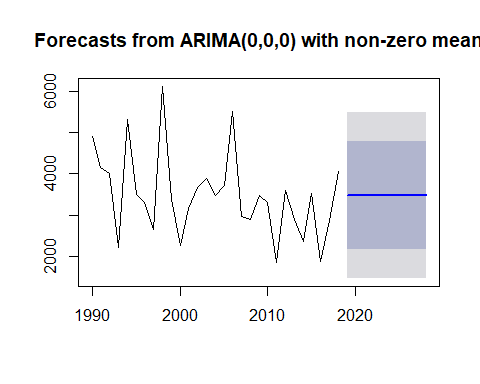
ts\_forecast\_cov(data, covariate = "British Columbia", subset = "Province" )



ts\_forecast\_cov(data, covariate = "Human", subset = "Cause" )



ts\_forecast\_cov(data, covariate = "Lightning", subset = "Cause" )



Notes for myself

RMSE

The RMSE is the square root of the variance of the residuals. It indicates the absolute fit of the model to the data-how close the observed data points are to the model’s predicted values. Whereas R-squared is a relative measure of fit, RMSE is an absolute measure of fit. As the square root of a variance, RMSE can be interpreted as the standard deviation of the unexplained variance, and has the useful property of being in the same units as the response variable. Lower values of RMSE indicate better fit. RMSE is a good measure of how accurately the model predicts the response, and it is the most important criterion for fit if the main purpose of the model is prediction.

The F-test

The F-test evaluates the null hypothesis that all regression coefficients are equal to zero versus the alternative that at least one is not. An equivalent null hypothesis is that R-squared equals zero. A significant F-test indicates that the observed R-squared is reliable and is not a spurious result of oddities in the data set. Thus the F-test determines whether the proposed relationship between the response variable and the set of predictors is statistically reliable and can be useful when the research objective is either prediction or explanation.

R-squared and Adjusted R-squared

The difference between SST and SSE is the improvement in prediction from the regression model, compared to the mean model. Dividing that difference by SST gives R-squared. It is the proportional improvement in prediction from the regression model, compared to the mean model. It indicates the goodness of fit of the model.

R-squared has the useful property that its scale is intuitive: it ranges from zero to one, with zero indicating that the proposed model does not improve prediction over the mean model, and one indicating perfect prediction. Improvement in the regression model results in proportional increases in R-squared.

One pitfall of R-squared is that it can only increase as predictors are added to the regression model. This increase is artificial when predictors are not actually improving the model’s fit. To remedy this, a related statistic, Adjusted R-squared, incorporates the model’s degrees of freedom. Adjusted R-squared will decrease as predictors are added if the increase in model fit does not make up for the loss of degrees of freedom. Likewise, it will increase as predictors are added if the increase in model fit is worthwhile. Adjusted R-squared should always be used with models with more than one predictor variable. It is interpreted as the proportion of total variance that is explained by the model.

There are situations in which a high R-squared is not necessary or relevant. When the interest is in the relationship between variables, not in prediction, the R-square is less important. An example is a study on how religiosity affects health outcomes. A good result is a reliable relationship between religiosity and health. No one would expect that religion explains a high percentage of the variation in health, as health is affected by many other factors. Even if the model accounts for other variables known to affect health, such as income and age, an R-squared in the range of 0.10 to 0.15 is reasonable.

## OUTLIERS CAPPED

head(outlier\_data\_frame3)

## Cause Jurisdiction capped\_Number Protection.zone  
## 1 Forest industry AB 96 Intensive  
## 2 Forest industry AB 96 Intensive  
## 3 Forest industry AB 96 Intensive  
## 4 Forest industry AB 96 Intensive  
## 5 Forest industry AB 96 Intensive  
## 6 Forest industry AB 96 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 1 Full 1990 Alberta People Early 90s 1990s  
## 2 Full 1991 Alberta People Early 90s 1990s  
## 3 Full 1992 Alberta People Early 90s 1990s  
## 4 Full 1993 Alberta People Early 90s 1990s  
## 5 Full 1994 Alberta People Early 90s 1990s  
## 6 Full 1995 Alberta People Early 90s 1990s  
## Region Fire\_Cause\_Human  
## 1 Prairie Region Human  
## 2 Prairie Region Human  
## 3 Prairie Region Human  
## 4 Prairie Region Human  
## 5 Prairie Region Human  
## 6 Prairie Region Human

tail(outlier\_data\_frame3)

## Cause Jurisdiction capped\_Number  
## 11514 Unspecified human activities NS 0  
## 11515 Unspecified human activities ON 96  
## 11516 Unspecified human activities PE 0  
## 11517 Unspecified human activities QC 6  
## 11518 Unspecified human activities SK 96  
## 11519 Unspecified human activities YT 6  
## Protection.zone Response.category Year Juris\_Long  
## 11514 Unspecified Modified 2018 Nova Scotia  
## 11515 Unspecified Modified 2018 Ontario  
## 11516 Unspecified Modified 2018 Prince Edward Island  
## 11517 Unspecified Modified 2018 Quebec  
## 11518 Unspecified Modified 2018 Saskatchewan  
## 11519 Unspecified Modified 2018 Yukon  
## Cause\_Grouped Time1 Time2 Region Fire\_Cause\_Human  
## 11514 People Late 20s 2010s Atlantic Region Human  
## 11515 People Late 20s 2010s Central Region Human  
## 11516 People Late 20s 2010s Atlantic Region Human  
## 11517 People Late 20s 2010s Central Region Human  
## 11518 People Late 20s 2010s Prairie Region Human  
## 11519 People Late 20s 2010s North Region Human

str(outlier\_data\_frame3)

## 'data.frame': 11519 obs. of 12 variables:  
## $ Cause : Factor w/ 10 levels "Forest industry",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Jurisdiction : Factor w/ 13 levels "AB","BC","MB",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ capped\_Number : num 96 96 96 96 96 96 8 96 10 96 ...  
## $ Protection.zone : Factor w/ 3 levels "Intensive","Limited",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Response.category: Factor w/ 4 levels "Full","Modified",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Year : int 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 ...  
## $ Juris\_Long : Factor w/ 13 levels "Alberta","British Columbia",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Cause\_Grouped : chr "People" "People" "People" "People" ...  
## $ Time1 : chr "Early 90s" "Early 90s" "Early 90s" "Early 90s" ...  
## $ Time2 : chr "1990s" "1990s" "1990s" "1990s" ...  
## $ Region : chr "Prairie Region" "Prairie Region" "Prairie Region" "Prairie Region" ...  
## $ Fire\_Cause\_Human : chr "Human" "Human" "Human" "Human" ...

dim(outlier\_data\_frame3) # 11519 rows, 6 columns

## [1] 11519 12

## OUTLIERS CAPPED

## Prepare packages for models with outliers removed

library(caret)  
library(lars)  
library(elasticnet)

## OUTLIERS CAPPED

## Split the data into training and test set

## Set the seed to make your partition reproducible

## We want to make sure that the training set and the test set do not have any common data points.

## R built in function “Sample”" randomly selects samples

set.seed(1235)   
train\_index <- sample(1:nrow(outlier\_data\_frame3), 0.80 \* nrow(outlier\_data\_frame3))  
train.set <- outlier\_data\_frame3[train\_index,]  
test.set <- outlier\_data\_frame3[-train\_index,]

## OUTLIERS CAPPED

## Set up model

## Set up K-fold cross-validation

## Defining the training controls for multiple models

train.control <- trainControl(method = "repeatedcv",   
 number = 3, repeats = 5)

## OUTLIERS CAPPED

## Build the models with start time and end time for each model

#############################################################################################################################################  
# Model 1: lm model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
lm\_model <- train(capped\_Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lm",  
 trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
lm\_model\_time.taken <- end.time - start.time  
#lm\_model\_time.taken  
  
#############################################################################################################################################  
# Model 2: glm model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
glm\_model <- train(capped\_Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "glm",  
 trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
glm\_model\_time.taken <- end.time - start.time  
#glm\_model\_time.taken  
  
#############################################################################################################################################  
# Model 3: lasso model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
lasso\_Mod <- train(capped\_Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lasso",  
 trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
lasso\_Mod\_time.taken <- end.time - start.time  
#lasso\_Mod\_time.taken  
#############################################################################################################################################  
# Model 4: knn model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
knn\_model <- train(capped\_Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "knn",  
 trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
knn\_model\_time.taken <- end.time - start.time  
#knn\_model\_time.taken  
#############################################################################################################################################  
# Model 5: leapForward model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
LF\_model <- train(capped\_Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "leapForward",  
 trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
LF\_model\_time.taken <- end.time - start.time  
#LF\_model\_time.taken  
#############################################################################################################################################  
# Model 6: leapBackward model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
LB\_model <- train(capped\_Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "leapBackward",  
 trControl = train.control)  
  
# Stop the clock  
end.time <- Sys.time()  
  
LB\_model\_time.taken <- end.time - start.time  
#LB\_model\_time.taken  
#############################################################################################################################################  
# Model 7: lmStepAIC model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)  
lmStepAIC\_Mod <- train(capped\_Number ~ Juris\_Long + Year + Fire\_Cause\_Human + Protection.zone + Response.category, data = train.set, method = "lmStepAIC",  
 trControl = train.control)

## Start: AIC=41243.38  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongOntario 1 1126 5027022 41243  
## <none> 5025897 41243  
## - `Juris\_LongBritish Columbia` 1 2006 5027902 41244  
## - Year 1 2584 5028481 41245  
## - Juris\_LongSaskatchewan 1 2667 5028563 41245  
## - Juris\_LongManitoba 1 5686 5031582 41248  
## - Juris\_LongQuebec 1 8664 5034561 41252  
## - Response.categoryUnspecified 1 10550 5036447 41254  
## - `Juris\_LongNova Scotia` 1 11282 5037179 41255  
## - Fire\_Cause\_HumanUnspecified 1 15303 5041200 41260  
## - Protection.zoneUnspecified 1 20353 5046250 41266  
## - `Juris\_LongNew Brunswick` 1 21849 5047745 41268  
## - `Juris\_LongNewfoundland and Labrador` 1 48941 5074838 41301  
## - `Juris\_LongNational parks` 1 49476 5075372 41302  
## - Juris\_LongYukon 1 49861 5075758 41302  
## - `Juris\_LongNorthwest Territories` 1 62607 5088503 41317  
## - `Juris\_LongPrince Edward Island` 1 115876 5141773 41381  
## - Protection.zoneLimited 1 145190 5171087 41416  
## - Fire\_Cause\_HumanLightning 1 414923 5440820 41729  
## - Response.categoryModified 1 1312441 6338338 42667  
## - Response.categoryNone 1 1339019 6364916 42692  
##   
## Step: AIC=41242.76  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - `Juris\_LongBritish Columbia` 1 1013 5028035 41242  
## <none> 5027022 41243  
## - Year 1 2756 5029778 41244  
## - Juris\_LongSaskatchewan 1 6699 5033721 41249  
## - Response.categoryUnspecified 1 10590 5037612 41254  
## - Fire\_Cause\_HumanUnspecified 1 15319 5042341 41259  
## - Juris\_LongManitoba 1 16679 5043701 41261  
## - Protection.zoneUnspecified 1 20417 5047439 41266  
## - `Juris\_LongNew Brunswick` 1 21351 5048373 41267  
## - `Juris\_LongNova Scotia` 1 21801 5048823 41267  
## - Juris\_LongQuebec 1 24751 5051773 41271  
## - `Juris\_LongNational parks` 1 79667 5106689 41337  
## - Juris\_LongYukon 1 98384 5125406 41360  
## - `Juris\_LongNorthwest Territories` 1 106338 5133360 41369  
## - `Juris\_LongNewfoundland and Labrador` 1 109898 5136920 41374  
## - Protection.zoneLimited 1 148593 5175616 41420  
## - `Juris\_LongPrince Edward Island` 1 162415 5189437 41436  
## - Fire\_Cause\_HumanLightning 1 415301 5442324 41728  
## - Response.categoryModified 1 1314554 6341576 42668  
## - Response.categoryNone 1 1339402 6366424 42692  
##   
## Step: AIC=41242  
## .outcome ~ Juris\_LongManitoba + `Juris\_LongNational parks` +   
## `Juris\_LongNew Brunswick` + `Juris\_LongNewfoundland and Labrador` +   
## `Juris\_LongNorthwest Territories` + `Juris\_LongNova Scotia` +   
## `Juris\_LongPrince Edward Island` + Juris\_LongQuebec + Juris\_LongSaskatchewan +   
## Juris\_LongYukon + Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 5028035 41242  
## - Year 1 2822 5030858 41243  
## - Juris\_LongSaskatchewan 1 8508 5036543 41250  
## - Response.categoryUnspecified 1 10507 5038542 41253  
## - Fire\_Cause\_HumanUnspecified 1 15396 5043432 41259  
## - `Juris\_LongNew Brunswick` 1 20387 5048423 41265  
## - Juris\_LongManitoba 1 20514 5048550 41265  
## - Protection.zoneUnspecified 1 20523 5048559 41265  
## - `Juris\_LongNova Scotia` 1 26145 5054181 41272  
## - Juris\_LongQuebec 1 30117 5058153 41277  
## - `Juris\_LongNational parks` 1 90101 5118137 41349  
## - Juris\_LongYukon 1 110885 5138920 41374  
## - `Juris\_LongNorthwest Territories` 1 121154 5149189 41386  
## - `Juris\_LongNewfoundland and Labrador` 1 126289 5154324 41392  
## - Protection.zoneLimited 1 156871 5184907 41429  
## - `Juris\_LongPrince Edward Island` 1 176391 5204426 41452  
## - Fire\_Cause\_HumanLightning 1 415466 5443502 41728  
## - Response.categoryModified 1 1315129 6343164 42667  
## - Response.categoryNone 1 1339767 6367803 42691  
## Start: AIC=41141.63  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongSaskatchewan 1 655 4939409 41140  
## <none> 4938754 41142  
## - Juris\_LongOntario 1 2005 4940758 41142  
## - Year 1 4709 4943463 41145  
## - `Juris\_LongNew Brunswick` 1 5907 4944661 41147  
## - Juris\_LongManitoba 1 6321 4945075 41147  
## - `Juris\_LongBritish Columbia` 1 7170 4945924 41149  
## - Juris\_LongQuebec 1 9868 4948622 41152  
## - Response.categoryUnspecified 1 11722 4950476 41154  
## - `Juris\_LongNova Scotia` 1 12417 4951171 41155  
## - Protection.zoneUnspecified 1 22986 4961739 41168  
## - Fire\_Cause\_HumanUnspecified 1 23996 4962749 41169  
## - Juris\_LongYukon 1 48669 4987422 41200  
## - `Juris\_LongNewfoundland and Labrador` 1 49244 4987997 41201  
## - `Juris\_LongNational parks` 1 57311 4996065 41211  
## - `Juris\_LongNorthwest Territories` 1 63854 5002608 41219  
## - `Juris\_LongPrince Edward Island` 1 131724 5070477 41301  
## - Protection.zoneLimited 1 176799 5115553 41356  
## - Fire\_Cause\_HumanLightning 1 361805 5300559 41574  
## - Response.categoryModified 1 1323502 6262256 42598  
## - Response.categoryNone 1 1395478 6334232 42669  
##   
## Step: AIC=41140.44  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongYukon + Year + Fire\_Cause\_HumanLightning +   
## Fire\_Cause\_HumanUnspecified + Protection.zoneLimited + Protection.zoneUnspecified +   
## Response.categoryModified + Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4939409 41140  
## - Juris\_LongOntario 1 4673 4944082 41144  
## - Year 1 4782 4944191 41144  
## - Juris\_LongManitoba 1 6074 4945483 41146  
## - `Juris\_LongNew Brunswick` 1 8282 4947691 41149  
## - Juris\_LongQuebec 1 10290 4949699 41151  
## - Response.categoryUnspecified 1 11204 4950612 41152  
## - `Juris\_LongBritish Columbia` 1 11611 4951020 41153  
## - `Juris\_LongNova Scotia` 1 12670 4952079 41154  
## - Protection.zoneUnspecified 1 22852 4962261 41167  
## - Fire\_Cause\_HumanUnspecified 1 24061 4963470 41168  
## - Juris\_LongYukon 1 57854 4997263 41210  
## - `Juris\_LongNewfoundland and Labrador` 1 60685 5000094 41213  
## - `Juris\_LongNational parks` 1 64970 5004379 41219  
## - `Juris\_LongNorthwest Territories` 1 74979 5014388 41231  
## - `Juris\_LongPrince Edward Island` 1 148222 5087631 41320  
## - Protection.zoneLimited 1 177722 5117131 41356  
## - Fire\_Cause\_HumanLightning 1 361914 5301323 41573  
## - Response.categoryModified 1 1322921 6262330 42596  
## - Response.categoryNone 1 1395180 6334589 42667  
## Start: AIC=41197.94  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongOntario 1 899 4989751 41197  
## <none> 4988852 41198  
## - Year 1 3097 4991949 41200  
## - `Juris\_LongBritish Columbia` 1 3332 4992184 41200  
## - Response.categoryUnspecified 1 4217 4993069 41201  
## - Juris\_LongSaskatchewan 1 5296 4994148 41202  
## - `Juris\_LongNew Brunswick` 1 7117 4995969 41205  
## - Fire\_Cause\_HumanUnspecified 1 7963 4996815 41206  
## - Juris\_LongManitoba 1 12464 5001316 41211  
## - `Juris\_LongNova Scotia` 1 13072 5001924 41212  
## - Protection.zoneUnspecified 1 14655 5003507 41214  
## - Juris\_LongQuebec 1 20588 5009440 41221  
## - `Juris\_LongNational parks` 1 60703 5049556 41270  
## - Juris\_LongYukon 1 62934 5051786 41273  
## - `Juris\_LongNewfoundland and Labrador` 1 63517 5052369 41274  
## - `Juris\_LongNorthwest Territories` 1 68887 5057739 41280  
## - `Juris\_LongPrince Edward Island` 1 116070 5104922 41337  
## - Protection.zoneLimited 1 136530 5125383 41362  
## - Fire\_Cause\_HumanLightning 1 390587 5379439 41659  
## - Response.categoryModified 1 1337803 6326655 42655  
## - Response.categoryNone 1 1369143 6357995 42686  
##   
## Step: AIC=41197.04  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4989751 41197  
## - `Juris\_LongBritish Columbia` 1 2435 4992186 41198  
## - Year 1 3184 4992935 41199  
## - Response.categoryUnspecified 1 4251 4994001 41200  
## - `Juris\_LongNew Brunswick` 1 6253 4996004 41203  
## - Fire\_Cause\_HumanUnspecified 1 8045 4997795 41205  
## - Juris\_LongSaskatchewan 1 11133 5000883 41209  
## - Protection.zoneUnspecified 1 14727 5004478 41213  
## - `Juris\_LongNova Scotia` 1 24678 5014429 41225  
## - Juris\_LongManitoba 1 30701 5020451 41233  
## - Juris\_LongQuebec 1 50999 5040749 41258  
## - `Juris\_LongNational parks` 1 95701 5085451 41312  
## - `Juris\_LongNorthwest Territories` 1 116952 5106703 41337  
## - Juris\_LongYukon 1 124581 5114332 41347  
## - Protection.zoneLimited 1 138753 5128504 41364  
## - `Juris\_LongNewfoundland and Labrador` 1 141418 5131169 41367  
## - `Juris\_LongPrince Edward Island` 1 160782 5150533 41390  
## - Fire\_Cause\_HumanLightning 1 390430 5380181 41658  
## - Response.categoryModified 1 1339961 6329711 42656  
## - Response.categoryNone 1 1368734 6358484 42684  
## Start: AIC=41070.25  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongOntario 1 16 4881729 41068  
## <none> 4881712 41070  
## - Juris\_LongSaskatchewan 1 2375 4884088 41071  
## - `Juris\_LongBritish Columbia` 1 3105 4884818 41072  
## - Year 1 3771 4885483 41073  
## - `Juris\_LongNew Brunswick` 1 8279 4889991 41079  
## - Fire\_Cause\_HumanUnspecified 1 9118 4890830 41080  
## - Response.categoryUnspecified 1 13246 4894958 41085  
## - Juris\_LongManitoba 1 14787 4896499 41087  
## - `Juris\_LongNova Scotia` 1 16603 4898316 41089  
## - Juris\_LongQuebec 1 16754 4898466 41089  
## - Protection.zoneUnspecified 1 19053 4900766 41092  
## - `Juris\_LongNational parks` 1 54986 4936698 41137  
## - `Juris\_LongNewfoundland and Labrador` 1 64555 4946267 41149  
## - Juris\_LongYukon 1 69624 4951336 41155  
## - `Juris\_LongNorthwest Territories` 1 72039 4953751 41158  
## - `Juris\_LongPrince Edward Island` 1 119371 5001084 41217  
## - Protection.zoneLimited 1 137117 5018829 41238  
## - Fire\_Cause\_HumanLightning 1 384050 5265762 41534  
## - Response.categoryModified 1 1271861 6153573 42491  
## - Response.categoryNone 1 1313660 6195372 42532  
##   
## Step: AIC=41068.27  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4881729 41068  
## - Juris\_LongSaskatchewan 1 3646 4885374 41071  
## - Year 1 3793 4885522 41071  
## - `Juris\_LongBritish Columbia` 1 4122 4885850 41071  
## - Fire\_Cause\_HumanUnspecified 1 9127 4890856 41078  
## - `Juris\_LongNew Brunswick` 1 9768 4891497 41079  
## - Response.categoryUnspecified 1 13254 4894983 41083  
## - Protection.zoneUnspecified 1 19062 4900790 41090  
## - `Juris\_LongNova Scotia` 1 25023 4906752 41098  
## - Juris\_LongManitoba 1 28353 4910081 41102  
## - Juris\_LongQuebec 1 33858 4915587 41109  
## - `Juris\_LongNational parks` 1 78670 4960398 41164  
## - `Juris\_LongNorthwest Territories` 1 110109 4991838 41203  
## - Juris\_LongYukon 1 122525 5004254 41219  
## - `Juris\_LongNewfoundland and Labrador` 1 127347 5009075 41224  
## - Protection.zoneLimited 1 144688 5026417 41246  
## - `Juris\_LongPrince Edward Island` 1 158426 5040155 41262  
## - Fire\_Cause\_HumanLightning 1 384499 5266227 41532  
## - Response.categoryModified 1 1272314 6154043 42489  
## - Response.categoryNone 1 1313666 6195395 42530  
## Start: AIC=41297.95  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 5070743 41298  
## - Juris\_LongOntario 1 2509 5073252 41299  
## - Juris\_LongSaskatchewan 1 2757 5073500 41299  
## - `Juris\_LongBritish Columbia` 1 4298 5075042 41301  
## - Year 1 4453 5075196 41301  
## - Juris\_LongManitoba 1 6602 5077345 41304  
## - Response.categoryUnspecified 1 6697 5077440 41304  
## - `Juris\_LongNew Brunswick` 1 6735 5077478 41304  
## - `Juris\_LongNova Scotia` 1 10306 5081049 41308  
## - Juris\_LongQuebec 1 11190 5081933 41309  
## - Fire\_Cause\_HumanUnspecified 1 19877 5090620 41320  
## - Protection.zoneUnspecified 1 23590 5094333 41324  
## - Juris\_LongYukon 1 43844 5114587 41349  
## - `Juris\_LongNewfoundland and Labrador` 1 54001 5124744 41361  
## - `Juris\_LongNational parks` 1 60281 5131024 41369  
## - `Juris\_LongNorthwest Territories` 1 60283 5131026 41369  
## - `Juris\_LongPrince Edward Island` 1 111166 5181909 41429  
## - Protection.zoneLimited 1 178303 5249046 41508  
## - Fire\_Cause\_HumanLightning 1 366869 5437613 41725  
## - Response.categoryModified 1 1351221 6421964 42747  
## - Response.categoryNone 1 1443171 6513914 42834  
## Start: AIC=41215.65  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 5003258 41216  
## - Juris\_LongSaskatchewan 1 2075 5005333 41216  
## - Year 1 2282 5005541 41216  
## - Juris\_LongOntario 1 2931 5006189 41217  
## - Juris\_LongManitoba 1 4071 5007329 41219  
## - `Juris\_LongBritish Columbia` 1 4391 5007649 41219  
## - Response.categoryUnspecified 1 6131 5009390 41221  
## - `Juris\_LongNova Scotia` 1 10144 5013402 41226  
## - Juris\_LongQuebec 1 10181 5013439 41226  
## - Protection.zoneUnspecified 1 14656 5017914 41232  
## - Fire\_Cause\_HumanUnspecified 1 17234 5020493 41235  
## - `Juris\_LongNew Brunswick` 1 17976 5021235 41236  
## - `Juris\_LongNewfoundland and Labrador` 1 43435 5046694 41267  
## - Juris\_LongYukon 1 48206 5051465 41273  
## - `Juris\_LongNational parks` 1 51880 5055138 41277  
## - `Juris\_LongNorthwest Territories` 1 62352 5065611 41290  
## - `Juris\_LongPrince Edward Island` 1 131472 5134731 41373  
## - Protection.zoneLimited 1 144031 5147290 41388  
## - Fire\_Cause\_HumanLightning 1 418412 5421670 41707  
## - Response.categoryModified 1 1356007 6359265 42687  
## - Response.categoryNone 1 1356035 6359293 42687  
## Start: AIC=41210.78  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongOntario 1 22 4999316 41209  
## <none> 4999294 41211  
## - Juris\_LongSaskatchewan 1 1644 5000937 41211  
## - `Juris\_LongBritish Columbia` 1 2542 5001836 41212  
## - Year 1 4440 5003734 41214  
## - Response.categoryUnspecified 1 13313 5012607 41225  
## - Fire\_Cause\_HumanUnspecified 1 13693 5012987 41226  
## - Juris\_LongManitoba 1 13824 5013118 41226  
## - Protection.zoneUnspecified 1 15182 5014476 41227  
## - `Juris\_LongNew Brunswick` 1 17605 5016898 41230  
## - Juris\_LongQuebec 1 18678 5017972 41232  
## - `Juris\_LongNova Scotia` 1 22503 5021797 41236  
## - `Juris\_LongNational parks` 1 58175 5057469 41280  
## - `Juris\_LongNewfoundland and Labrador` 1 63054 5062348 41286  
## - `Juris\_LongNorthwest Territories` 1 67088 5066382 41291  
## - Juris\_LongYukon 1 68528 5067822 41292  
## - `Juris\_LongPrince Edward Island` 1 143844 5143138 41383  
## - Protection.zoneLimited 1 144930 5144224 41384  
## - Fire\_Cause\_HumanLightning 1 386349 5385642 41666  
## - Response.categoryModified 1 1326707 6326001 42655  
## - Response.categoryNone 1 1334456 6333749 42662  
##   
## Step: AIC=41208.81  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4999316 41209  
## - Juris\_LongSaskatchewan 1 2561 5001877 41210  
## - `Juris\_LongBritish Columbia` 1 3222 5002538 41211  
## - Year 1 4468 5003784 41212  
## - Response.categoryUnspecified 1 13324 5012640 41223  
## - Fire\_Cause\_HumanUnspecified 1 13703 5013019 41224  
## - Protection.zoneUnspecified 1 15202 5014518 41225  
## - `Juris\_LongNew Brunswick` 1 20401 5019717 41232  
## - Juris\_LongManitoba 1 25744 5025060 41238  
## - `Juris\_LongNova Scotia` 1 33112 5032428 41247  
## - Juris\_LongQuebec 1 36576 5035892 41252  
## - `Juris\_LongNational parks` 1 80551 5079867 41305  
## - `Juris\_LongNorthwest Territories` 1 100578 5099894 41329  
## - Juris\_LongYukon 1 114741 5114057 41346  
## - `Juris\_LongNewfoundland and Labrador` 1 122300 5121616 41355  
## - Protection.zoneLimited 1 153889 5153205 41393  
## - `Juris\_LongPrince Edward Island` 1 187482 5186798 41433  
## - Fire\_Cause\_HumanLightning 1 386381 5385697 41664  
## - Response.categoryModified 1 1327424 6326740 42653  
## - Response.categoryNone 1 1334678 6333994 42660  
## Start: AIC=41120.98  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongOntario 1 1362 4928109 41121  
## <none> 4926747 41121  
## - Year 1 2350 4929097 41122  
## - `Juris\_LongBritish Columbia` 1 3834 4930581 41124  
## - Juris\_LongSaskatchewan 1 5811 4932558 41126  
## - `Juris\_LongNew Brunswick` 1 7407 4934154 41128  
## - Juris\_LongManitoba 1 9287 4936034 41131  
## - `Juris\_LongNova Scotia` 1 9776 4936523 41131  
## - Juris\_LongQuebec 1 12185 4938933 41134  
## - Response.categoryUnspecified 1 14437 4941184 41137  
## - Fire\_Cause\_HumanUnspecified 1 18900 4945647 41143  
## - Protection.zoneUnspecified 1 35269 4962016 41163  
## - `Juris\_LongNewfoundland and Labrador` 1 58612 4985359 41192  
## - Juris\_LongYukon 1 59490 4986237 41193  
## - `Juris\_LongNational parks` 1 61394 4988141 41195  
## - `Juris\_LongNorthwest Territories` 1 68751 4995499 41204  
## - `Juris\_LongPrince Edward Island` 1 114054 5040801 41260  
## - Protection.zoneLimited 1 141462 5068209 41293  
## - Fire\_Cause\_HumanLightning 1 395657 5322404 41594  
## - Response.categoryModified 1 1334761 6261508 42592  
## - Response.categoryNone 1 1338140 6264887 42595  
##   
## Step: AIC=41120.68  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4928109 41121  
## - `Juris\_LongBritish Columbia` 1 2496 4930605 41122  
## - Year 1 2512 4930621 41122  
## - `Juris\_LongNew Brunswick` 1 6052 4934161 41126  
## - Juris\_LongSaskatchewan 1 12879 4940989 41135  
## - Response.categoryUnspecified 1 14500 4942610 41137  
## - Fire\_Cause\_HumanUnspecified 1 18914 4947024 41142  
## - `Juris\_LongNova Scotia` 1 20664 4948774 41144  
## - Juris\_LongManitoba 1 26140 4954249 41151  
## - Juris\_LongQuebec 1 34503 4962613 41162  
## - Protection.zoneUnspecified 1 35431 4963540 41163  
## - `Juris\_LongNational parks` 1 99490 5027599 41241  
## - `Juris\_LongNorthwest Territories` 1 118774 5046884 41265  
## - Juris\_LongYukon 1 121598 5049707 41268  
## - `Juris\_LongNewfoundland and Labrador` 1 133984 5062094 41283  
## - Protection.zoneLimited 1 143457 5071566 41295  
## - `Juris\_LongPrince Edward Island` 1 161780 5089889 41317  
## - Fire\_Cause\_HumanLightning 1 396489 5324599 41594  
## - Response.categoryModified 1 1336547 6264656 42593  
## - Response.categoryNone 1 1338171 6266281 42594  
## Start: AIC=41243.06  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongSaskatchewan 1 923 5021893 41242  
## <none> 5020970 41243  
## - Response.categoryUnspecified 1 1735 5022705 41243  
## - Juris\_LongManitoba 1 2874 5023844 41245  
## - Year 1 3738 5024708 41246  
## - Juris\_LongOntario 1 4266 5025236 41246  
## - `Juris\_LongBritish Columbia` 1 5732 5026702 41248  
## - `Juris\_LongNova Scotia` 1 6909 5027878 41250  
## - `Juris\_LongNew Brunswick` 1 7154 5028124 41250  
## - Juris\_LongQuebec 1 7920 5028890 41251  
## - Protection.zoneUnspecified 1 11693 5032663 41255  
## - Fire\_Cause\_HumanUnspecified 1 13366 5034335 41257  
## - Juris\_LongYukon 1 35770 5056740 41285  
## - `Juris\_LongNewfoundland and Labrador` 1 41223 5062193 41291  
## - `Juris\_LongNational parks` 1 47735 5068704 41299  
## - `Juris\_LongNorthwest Territories` 1 58853 5079823 41313  
## - `Juris\_LongPrince Edward Island` 1 105901 5126870 41369  
## - Protection.zoneLimited 1 172400 5193370 41448  
## - Fire\_Cause\_HumanLightning 1 383197 5404167 41693  
## - Response.categoryModified 1 1319654 6340624 42675  
## - Response.categoryNone 1 1433453 6454422 42784  
##   
## Step: AIC=41242.19  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongYukon + Year + Fire\_Cause\_HumanLightning +   
## Fire\_Cause\_HumanUnspecified + Protection.zoneLimited + Protection.zoneUnspecified +   
## Response.categoryModified + Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Response.categoryUnspecified 1 1437 5023330 41242  
## <none> 5021893 41242  
## - Juris\_LongManitoba 1 1955 5023848 41243  
## - Year 1 3858 5025751 41245  
## - `Juris\_LongNova Scotia` 1 6065 5027958 41248  
## - Juris\_LongQuebec 1 7440 5029333 41249  
## - Juris\_LongOntario 1 9270 5031163 41252  
## - `Juris\_LongBritish Columbia` 1 10108 5032001 41253  
## - `Juris\_LongNew Brunswick` 1 10187 5032080 41253  
## - Protection.zoneUnspecified 1 11676 5033569 41254  
## - Fire\_Cause\_HumanUnspecified 1 13383 5035275 41257  
## - Juris\_LongYukon 1 40834 5062727 41290  
## - `Juris\_LongNewfoundland and Labrador` 1 48981 5070874 41300  
## - `Juris\_LongNational parks` 1 52961 5074854 41305  
## - `Juris\_LongNorthwest Territories` 1 67992 5089885 41323  
## - `Juris\_LongPrince Edward Island` 1 116251 5138143 41381  
## - Protection.zoneLimited 1 173601 5195493 41449  
## - Fire\_Cause\_HumanLightning 1 383608 5405500 41692  
## - Response.categoryModified 1 1318958 6340851 42673  
## - Response.categoryNone 1 1432959 6454852 42782  
##   
## Step: AIC=41241.95  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongYukon + Year + Fire\_Cause\_HumanLightning +   
## Fire\_Cause\_HumanUnspecified + Protection.zoneLimited + Protection.zoneUnspecified +   
## Response.categoryModified + Response.categoryNone  
##   
## Df Sum of Sq RSS AIC  
## <none> 5023330 41242  
## - Juris\_LongManitoba 1 1891 5025221 41242  
## - Year 1 4731 5028061 41246  
## - `Juris\_LongNova Scotia` 1 6527 5029857 41248  
## - Juris\_LongQuebec 1 7970 5031300 41250  
## - Juris\_LongOntario 1 8743 5032073 41251  
## - `Juris\_LongBritish Columbia` 1 9566 5032896 41252  
## - `Juris\_LongNew Brunswick` 1 9705 5033035 41252  
## - Protection.zoneUnspecified 1 11787 5035116 41254  
## - Fire\_Cause\_HumanUnspecified 1 13503 5036833 41256  
## - Juris\_LongYukon 1 41936 5065266 41291  
## - `Juris\_LongNewfoundland and Labrador` 1 50522 5073852 41301  
## - `Juris\_LongNational parks` 1 54438 5077768 41306  
## - `Juris\_LongNorthwest Territories` 1 68027 5091357 41323  
## - `Juris\_LongPrince Edward Island` 1 118913 5142243 41384  
## - Protection.zoneLimited 1 174995 5198325 41450  
## - Fire\_Cause\_HumanLightning 1 382998 5406328 41691  
## - Response.categoryModified 1 1346958 6370288 42699  
## - Response.categoryNone 1 1469377 6492707 42816  
## Start: AIC=41215.87  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 5003440 41216  
## - Juris\_LongOntario 1 1771 5005211 41216  
## - Juris\_LongSaskatchewan 1 2212 5005653 41217  
## - Year 1 5420 5008860 41221  
## - `Juris\_LongBritish Columbia` 1 6093 5009533 41221  
## - Response.categoryUnspecified 1 6856 5010296 41222  
## - Juris\_LongManitoba 1 7005 5010446 41222  
## - Juris\_LongQuebec 1 9942 5013382 41226  
## - Fire\_Cause\_HumanUnspecified 1 10315 5013755 41227  
## - `Juris\_LongNova Scotia` 1 10890 5014330 41227  
## - Protection.zoneUnspecified 1 11759 5015200 41228  
## - `Juris\_LongNew Brunswick` 1 12769 5016209 41230  
## - `Juris\_LongNorthwest Territories` 1 45540 5048981 41270  
## - `Juris\_LongNewfoundland and Labrador` 1 47728 5051169 41272  
## - Juris\_LongYukon 1 48120 5051560 41273  
## - `Juris\_LongNational parks` 1 54210 5057650 41280  
## - `Juris\_LongPrince Edward Island` 1 111842 5115282 41350  
## - Protection.zoneLimited 1 136053 5139493 41379  
## - Fire\_Cause\_HumanLightning 1 513361 5516801 41814  
## - Response.categoryModified 1 1270775 6274215 42604  
## - Response.categoryNone 1 1335796 6339237 42668  
## Start: AIC=41199.67  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongSaskatchewan 1 1012 4986648 41199  
## <none> 4985636 41200  
## - `Juris\_LongBritish Columbia` 1 1646 4987282 41200  
## - Juris\_LongOntario 1 2033 4987669 41200  
## - Year 1 3827 4989463 41202  
## - Juris\_LongManitoba 1 5183 4990819 41204  
## - `Juris\_LongNew Brunswick` 1 9477 4995112 41209  
## - Response.categoryUnspecified 1 11505 4997141 41212  
## - `Juris\_LongNova Scotia` 1 11742 4997378 41212  
## - Juris\_LongQuebec 1 13117 4998753 41214  
## - Fire\_Cause\_HumanUnspecified 1 19304 5004939 41221  
## - Protection.zoneUnspecified 1 27342 5012978 41231  
## - `Juris\_LongNational parks` 1 49646 5035282 41259  
## - `Juris\_LongNewfoundland and Labrador` 1 51758 5037394 41261  
## - Juris\_LongYukon 1 52623 5038259 41262  
## - `Juris\_LongNorthwest Territories` 1 72513 5058149 41286  
## - `Juris\_LongPrince Edward Island` 1 123930 5109566 41349  
## - Protection.zoneLimited 1 165482 5151118 41398  
## - Fire\_Cause\_HumanLightning 1 317089 5302725 41577  
## - Response.categoryModified 1 1366233 6351869 42686  
## - Response.categoryNone 1 1404781 6390416 42723  
##   
## Step: AIC=41198.92  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongYukon + Year + Fire\_Cause\_HumanLightning +   
## Fire\_Cause\_HumanUnspecified + Protection.zoneLimited + Protection.zoneUnspecified +   
## Response.categoryModified + Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4986648 41199  
## - `Juris\_LongBritish Columbia` 1 3832 4990480 41202  
## - Year 1 3952 4990600 41202  
## - Juris\_LongManitoba 1 4241 4990889 41202  
## - Juris\_LongOntario 1 5330 4991977 41203  
## - Response.categoryUnspecified 1 10805 4997452 41210  
## - `Juris\_LongNova Scotia` 1 11191 4997839 41211  
## - `Juris\_LongNew Brunswick` 1 13313 4999961 41213  
## - Juris\_LongQuebec 1 13378 5000026 41213  
## - Fire\_Cause\_HumanUnspecified 1 19446 5006093 41221  
## - Protection.zoneUnspecified 1 27236 5013884 41230  
## - `Juris\_LongNational parks` 1 54447 5041094 41264  
## - Juris\_LongYukon 1 61469 5048117 41272  
## - `Juris\_LongNewfoundland and Labrador` 1 62799 5049447 41274  
## - `Juris\_LongNorthwest Territories` 1 84867 5071514 41301  
## - `Juris\_LongPrince Edward Island` 1 137312 5123959 41364  
## - Protection.zoneLimited 1 166922 5153569 41399  
## - Fire\_Cause\_HumanLightning 1 316851 5303499 41575  
## - Response.categoryModified 1 1365225 6351873 42684  
## - Response.categoryNone 1 1405396 6392044 42722  
## Start: AIC=41158.6  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongOntario 1 362 4957373 41157  
## <none> 4957011 41159  
## - Year 1 1636 4958647 41159  
## - `Juris\_LongBritish Columbia` 1 4687 4961698 41162  
## - Juris\_LongSaskatchewan 1 4699 4961710 41162  
## - Response.categoryUnspecified 1 7018 4964029 41165  
## - `Juris\_LongNew Brunswick` 1 9554 4966565 41168  
## - Juris\_LongManitoba 1 12771 4969782 41172  
## - `Juris\_LongNova Scotia` 1 14027 4971038 41174  
## - Juris\_LongQuebec 1 15295 4972305 41176  
## - Fire\_Cause\_HumanUnspecified 1 16148 4973159 41177  
## - Protection.zoneUnspecified 1 18914 4975925 41180  
## - Juris\_LongYukon 1 60404 5017415 41231  
## - `Juris\_LongNewfoundland and Labrador` 1 61538 5018549 41232  
## - `Juris\_LongNational parks` 1 63340 5020351 41235  
## - `Juris\_LongNorthwest Territories` 1 79029 5036040 41254  
## - `Juris\_LongPrince Edward Island` 1 127186 5084197 41312  
## - Protection.zoneLimited 1 156784 5113795 41348  
## - Fire\_Cause\_HumanLightning 1 348603 5305613 41574  
## - Response.categoryModified 1 1340111 6297122 42627  
## - Response.categoryNone 1 1355451 6312462 42642  
##   
## Step: AIC=41157.05  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4957373 41157  
## - Year 1 1700 4959073 41157  
## - `Juris\_LongBritish Columbia` 1 4799 4962172 41161  
## - Response.categoryUnspecified 1 7044 4964417 41164  
## - Juris\_LongSaskatchewan 1 8906 4966279 41166  
## - `Juris\_LongNew Brunswick` 1 9699 4967072 41167  
## - Fire\_Cause\_HumanUnspecified 1 16094 4973467 41175  
## - Protection.zoneUnspecified 1 19039 4976412 41179  
## - `Juris\_LongNova Scotia` 1 23895 4981268 41185  
## - Juris\_LongManitoba 1 29143 4986516 41191  
## - Juris\_LongQuebec 1 36544 4993917 41200  
## - `Juris\_LongNational parks` 1 96480 5053853 41273  
## - Juris\_LongYukon 1 114332 5071705 41295  
## - `Juris\_LongNorthwest Territories` 1 128816 5086189 41313  
## - `Juris\_LongNewfoundland and Labrador` 1 130171 5087543 41314  
## - Protection.zoneLimited 1 163397 5120770 41354  
## - `Juris\_LongPrince Edward Island` 1 173053 5130426 41366  
## - Fire\_Cause\_HumanLightning 1 348974 5306347 41573  
## - Response.categoryModified 1 1341715 6299088 42626  
## - Response.categoryNone 1 1355284 6312657 42640  
## Start: AIC=41247  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongOntario 1 1560 5030416 41247  
## - Year 1 1601 5030457 41247  
## <none> 5028856 41247  
## - `Juris\_LongBritish Columbia` 1 4205 5033061 41250  
## - `Juris\_LongNew Brunswick` 1 6530 5035386 41253  
## - Juris\_LongSaskatchewan 1 7461 5036317 41254  
## - Protection.zoneUnspecified 1 8034 5036890 41255  
## - Juris\_LongManitoba 1 8198 5037054 41255  
## - Juris\_LongQuebec 1 11352 5040208 41259  
## - `Juris\_LongNova Scotia` 1 11813 5040669 41259  
## - Response.categoryUnspecified 1 13675 5042531 41262  
## - Fire\_Cause\_HumanUnspecified 1 16817 5045673 41266  
## - `Juris\_LongNational parks` 1 61271 5090127 41319  
## - `Juris\_LongNewfoundland and Labrador` 1 61324 5090180 41319  
## - Juris\_LongYukon 1 64324 5093179 41323  
## - `Juris\_LongNorthwest Territories` 1 70601 5099457 41331  
## - `Juris\_LongPrince Edward Island` 1 112861 5141716 41381  
## - Protection.zoneLimited 1 143817 5172672 41418  
## - Fire\_Cause\_HumanLightning 1 407895 5436750 41724  
## - Response.categoryNone 1 1381531 6410387 42736  
## - Response.categoryModified 1 1400495 6429351 42754  
##   
## Step: AIC=41246.9  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 5030416 41247  
## - Year 1 1738 5032154 41247  
## - `Juris\_LongBritish Columbia` 1 2697 5033113 41248  
## - `Juris\_LongNew Brunswick` 1 5036 5035452 41251  
## - Protection.zoneUnspecified 1 8107 5038523 41255  
## - Response.categoryUnspecified 1 13732 5044148 41262  
## - Juris\_LongSaskatchewan 1 15512 5045928 41264  
## - Fire\_Cause\_HumanUnspecified 1 16741 5047157 41265  
## - Juris\_LongManitoba 1 23237 5053653 41273  
## - `Juris\_LongNova Scotia` 1 23779 5054195 41274  
## - Juris\_LongQuebec 1 32189 5062605 41284  
## - `Juris\_LongNational parks` 1 97918 5128334 41363  
## - `Juris\_LongNorthwest Territories` 1 118920 5149336 41388  
## - Juris\_LongYukon 1 126958 5157374 41398  
## - `Juris\_LongNewfoundland and Labrador` 1 136940 5167356 41410  
## - Protection.zoneLimited 1 145411 5175827 41420  
## - `Juris\_LongPrince Edward Island` 1 158836 5189252 41436  
## - Fire\_Cause\_HumanLightning 1 407701 5438117 41724  
## - Response.categoryNone 1 1382031 6412447 42736  
## - Response.categoryModified 1 1403383 6433799 42756  
## Start: AIC=41147.95  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongOntario 1 329 4948748 41146  
## <none> 4948419 41148  
## - `Juris\_LongBritish Columbia` 1 2434 4950853 41149  
## - Juris\_LongSaskatchewan 1 3606 4952025 41150  
## - Response.categoryUnspecified 1 5368 4953786 41153  
## - Year 1 5427 4953845 41153  
## - Juris\_LongManitoba 1 8434 4956852 41156  
## - Juris\_LongQuebec 1 13826 4962245 41163  
## - `Juris\_LongNova Scotia` 1 14979 4963398 41165  
## - `Juris\_LongNew Brunswick` 1 15310 4963729 41165  
## - Fire\_Cause\_HumanUnspecified 1 15802 4964220 41166  
## - Protection.zoneUnspecified 1 26514 4974933 41179  
## - Juris\_LongYukon 1 56956 5005374 41216  
## - `Juris\_LongNewfoundland and Labrador` 1 57193 5005612 41217  
## - `Juris\_LongNational parks` 1 60274 5008693 41220  
## - `Juris\_LongNorthwest Territories` 1 70871 5019290 41233  
## - `Juris\_LongPrince Edward Island` 1 129625 5078043 41305  
## - Protection.zoneLimited 1 149499 5097917 41329  
## - Fire\_Cause\_HumanLightning 1 396751 5345170 41620  
## - Response.categoryModified 1 1284218 6232637 42563  
## - Response.categoryNone 1 1349572 6297990 42627  
##   
## Step: AIC=41146.36  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4948748 41146  
## - `Juris\_LongBritish Columbia` 1 2222 4950971 41147  
## - Response.categoryUnspecified 1 5390 4954138 41151  
## - Year 1 5528 4954276 41151  
## - Juris\_LongSaskatchewan 1 7034 4955782 41153  
## - Fire\_Cause\_HumanUnspecified 1 15823 4964571 41164  
## - `Juris\_LongNew Brunswick` 1 16220 4964968 41164  
## - Juris\_LongManitoba 1 19843 4968591 41169  
## - `Juris\_LongNova Scotia` 1 25776 4974524 41176  
## - Protection.zoneUnspecified 1 26583 4975331 41177  
## - Juris\_LongQuebec 1 32358 4981107 41184  
## - `Juris\_LongNational parks` 1 92936 5041684 41259  
## - Juris\_LongYukon 1 107823 5056571 41277  
## - `Juris\_LongNorthwest Territories` 1 115375 5064123 41286  
## - `Juris\_LongNewfoundland and Labrador` 1 121603 5070351 41293  
## - Protection.zoneLimited 1 155788 5104536 41335  
## - `Juris\_LongPrince Edward Island` 1 177740 5126488 41361  
## - Fire\_Cause\_HumanLightning 1 397173 5345921 41619  
## - Response.categoryModified 1 1285540 6234288 42563  
## - Response.categoryNone 1 1349250 6297998 42625  
## Start: AIC=41188.61  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## - Juris\_LongSaskatchewan 1 10 4976678 41187  
## <none> 4976668 41189  
## - Juris\_LongOntario 1 2512 4979180 41190  
## - Year 1 3529 4980196 41191  
## - `Juris\_LongBritish Columbia` 1 5514 4982182 41193  
## - Response.categoryUnspecified 1 6714 4983382 41195  
## - Juris\_LongManitoba 1 7185 4983853 41195  
## - `Juris\_LongNova Scotia` 1 10272 4986940 41199  
## - `Juris\_LongNew Brunswick` 1 10304 4986971 41199  
## - Juris\_LongQuebec 1 12471 4989139 41202  
## - Fire\_Cause\_HumanUnspecified 1 12906 4989573 41203  
## - Protection.zoneUnspecified 1 24032 5000700 41216  
## - Juris\_LongYukon 1 40983 5017651 41237  
## - `Juris\_LongNewfoundland and Labrador` 1 43653 5020321 41240  
## - `Juris\_LongNational parks` 1 46284 5022952 41243  
## - `Juris\_LongNorthwest Territories` 1 54078 5030746 41253  
## - `Juris\_LongPrince Edward Island` 1 121479 5098147 41335  
## - Protection.zoneLimited 1 164759 5141427 41387  
## - Fire\_Cause\_HumanLightning 1 366460 5343128 41623  
## - Response.categoryModified 1 1299051 6275719 42612  
## - Response.categoryNone 1 1371515 6348183 42682  
##   
## Step: AIC=41186.62  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongYukon + Year + Fire\_Cause\_HumanLightning +   
## Fire\_Cause\_HumanUnspecified + Protection.zoneLimited + Protection.zoneUnspecified +   
## Response.categoryModified + Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 4976678 41187  
## - Year 1 3529 4980206 41189  
## - Juris\_LongOntario 1 3719 4980397 41189  
## - Response.categoryUnspecified 1 6763 4983441 41193  
## - `Juris\_LongBritish Columbia` 1 7211 4983889 41194  
## - Juris\_LongManitoba 1 9631 4986308 41197  
## - `Juris\_LongNew Brunswick` 1 11863 4988541 41199  
## - `Juris\_LongNova Scotia` 1 12595 4989273 41200  
## - Fire\_Cause\_HumanUnspecified 1 12910 4989588 41201  
## - Juris\_LongQuebec 1 17066 4993743 41206  
## - Protection.zoneUnspecified 1 24024 5000701 41214  
## - Juris\_LongYukon 1 54014 5030691 41251  
## - `Juris\_LongNational parks` 1 56491 5033169 41254  
## - `Juris\_LongNewfoundland and Labrador` 1 60346 5037024 41259  
## - `Juris\_LongNorthwest Territories` 1 69442 5046120 41270  
## - `Juris\_LongPrince Edward Island` 1 142483 5119161 41358  
## - Protection.zoneLimited 1 165264 5141942 41385  
## - Fire\_Cause\_HumanLightning 1 366460 5343138 41621  
## - Response.categoryModified 1 1300203 6276881 42611  
## - Response.categoryNone 1 1371900 6348578 42680  
## Start: AIC=61781.96  
## .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified  
##   
## Df Sum of Sq RSS AIC  
## <none> 7485882 61782  
## - Juris\_LongOntario 1 1917 7487799 61782  
## - Juris\_LongSaskatchewan 1 3653 7489535 61784  
## - Year 1 5183 7491065 61786  
## - `Juris\_LongBritish Columbia` 1 5849 7491731 61787  
## - Juris\_LongManitoba 1 11873 7497755 61795  
## - Response.categoryUnspecified 1 12440 7498322 61795  
## - `Juris\_LongNew Brunswick` 1 15758 7501640 61799  
## - `Juris\_LongNova Scotia` 1 18390 7504272 61803  
## - Juris\_LongQuebec 1 18869 7504751 61803  
## - Fire\_Cause\_HumanUnspecified 1 22570 7508452 61808  
## - Protection.zoneUnspecified 1 28438 7514320 61815  
## - Juris\_LongYukon 1 80356 7566238 61878  
## - `Juris\_LongNewfoundland and Labrador` 1 80651 7566533 61879  
## - `Juris\_LongNational parks` 1 83543 7569425 61882  
## - `Juris\_LongNorthwest Territories` 1 97365 7583248 61899  
## - `Juris\_LongPrince Edward Island` 1 182106 7667988 62001  
## - Protection.zoneLimited 1 229058 7714940 62058  
## - Fire\_Cause\_HumanLightning 1 584052 8069934 62472  
## - Response.categoryModified 1 1990915 9476797 63953  
## - Response.categoryNone 1 2054568 9540450 64015

# Stop the clock  
end.time <- Sys.time()  
  
lmStepAIC\_Mod\_time.taken <- end.time - start.time  
#lmStepAIC\_Mod\_time.taken  
#############################################################################################################################################

## OUTLIERS CAPPED

## View the total amount of time taken to run each model.

Model\_Time <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 Time\_train=c(lm\_model\_time.taken, glm\_model\_time.taken, lasso\_Mod\_time.taken, knn\_model\_time.taken, LF\_model\_time.taken, LB\_model\_time.taken, lmStepAIC\_Mod\_time.taken))  
Model\_Time

## Name Time\_train  
## 1 lm\_model 1.643645 secs  
## 2 glm\_model 1.738344 secs  
## 3 lasso\_Mod 2.068453 secs  
## 4 knn\_model 25.646455 secs  
## 5 LF\_model 1.012250 secs  
## 6 LB\_model 1.079079 secs  
## 7 lmStepAIC\_Mod 4.663523 secs

## OUTLIERS CAPPED

## View summaries of the models.

############################################  
# Model 1  
# Summarize the results for lm\_model  
print(lm\_model)

## Linear Regression   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 28.59885 0.392571 21.82526  
##   
## Tuning parameter 'intercept' was held constant at a value of TRUE

summary(lm\_model)

##   
## Call:  
## lm(formula = .outcome ~ ., data = dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -89.597 -19.241 -3.102 13.433 115.532   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 242.30849 76.15445 3.182  
## `Juris\_LongBritish Columbia` 4.63208 1.72819 2.680  
## Juris\_LongManitoba -5.80193 1.51934 -3.819  
## `Juris\_LongNational parks` -17.65427 1.74287 -10.129  
## `Juris\_LongNew Brunswick` 10.17541 2.31298 4.399  
## `Juris\_LongNewfoundland and Labrador` -14.81272 1.48833 -9.953  
## `Juris\_LongNorthwest Territories` -17.90664 1.63750 -10.935  
## `Juris\_LongNova Scotia` -7.99015 1.68126 -4.752  
## Juris\_LongOntario 2.31954 1.51156 1.535  
## `Juris\_LongPrince Edward Island` -29.50534 1.97291 -14.955  
## Juris\_LongQuebec -7.14402 1.48402 -4.814  
## Juris\_LongSaskatchewan -3.73044 1.76118 -2.118  
## Juris\_LongYukon -15.63102 1.57343 -9.934  
## Year -0.09588 0.03800 -2.523  
## Fire\_Cause\_HumanLightning 25.35580 0.94672 26.783  
## Fire\_Cause\_HumanUnspecified -5.01223 0.95200 -5.265  
## Protection.zoneLimited -12.86508 0.76702 -16.773  
## Protection.zoneUnspecified 27.41256 4.63841 5.910  
## Response.categoryModified -36.60685 0.74030 -49.449  
## Response.categoryNone -36.76037 0.73179 -50.233  
## Response.categoryUnspecified 12.20025 3.12123 3.909  
## Pr(>|t|)   
## (Intercept) 0.001468 \*\*   
## `Juris\_LongBritish Columbia` 0.007369 \*\*   
## Juris\_LongManitoba 0.000135 \*\*\*  
## `Juris\_LongNational parks` < 2e-16 \*\*\*  
## `Juris\_LongNew Brunswick` 1.10e-05 \*\*\*  
## `Juris\_LongNewfoundland and Labrador` < 2e-16 \*\*\*  
## `Juris\_LongNorthwest Territories` < 2e-16 \*\*\*  
## `Juris\_LongNova Scotia` 2.04e-06 \*\*\*  
## Juris\_LongOntario 0.124932   
## `Juris\_LongPrince Edward Island` < 2e-16 \*\*\*  
## Juris\_LongQuebec 1.50e-06 \*\*\*  
## Juris\_LongSaskatchewan 0.034190 \*   
## Juris\_LongYukon < 2e-16 \*\*\*  
## Year 0.011649 \*   
## Fire\_Cause\_HumanLightning < 2e-16 \*\*\*  
## Fire\_Cause\_HumanUnspecified 1.43e-07 \*\*\*  
## Protection.zoneLimited < 2e-16 \*\*\*  
## Protection.zoneUnspecified 3.54e-09 \*\*\*  
## Response.categoryModified < 2e-16 \*\*\*  
## Response.categoryNone < 2e-16 \*\*\*  
## Response.categoryUnspecified 9.34e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 28.53 on 9194 degrees of freedom  
## Multiple R-squared: 0.3964, Adjusted R-squared: 0.3951   
## F-statistic: 301.9 on 20 and 9194 DF, p-value: < 2.2e-16

lm\_model$finalModel

##   
## Call:  
## lm(formula = .outcome ~ ., data = dat)  
##   
## Coefficients:  
## (Intercept)   
## 242.30849   
## `Juris\_LongBritish Columbia`   
## 4.63208   
## Juris\_LongManitoba   
## -5.80193   
## `Juris\_LongNational parks`   
## -17.65427   
## `Juris\_LongNew Brunswick`   
## 10.17541   
## `Juris\_LongNewfoundland and Labrador`   
## -14.81272   
## `Juris\_LongNorthwest Territories`   
## -17.90664   
## `Juris\_LongNova Scotia`   
## -7.99015   
## Juris\_LongOntario   
## 2.31954   
## `Juris\_LongPrince Edward Island`   
## -29.50534   
## Juris\_LongQuebec   
## -7.14402   
## Juris\_LongSaskatchewan   
## -3.73044   
## Juris\_LongYukon   
## -15.63102   
## Year   
## -0.09588   
## Fire\_Cause\_HumanLightning   
## 25.35580   
## Fire\_Cause\_HumanUnspecified   
## -5.01223   
## Protection.zoneLimited   
## -12.86508   
## Protection.zoneUnspecified   
## 27.41256   
## Response.categoryModified   
## -36.60685   
## Response.categoryNone   
## -36.76037   
## Response.categoryUnspecified   
## 12.20025

lm\_model$modelType

## [1] "Regression"

############################################  
# Model 2  
# Summarize the results for glm model  
print(glm\_model)

## Generalized Linear Model   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 28.59885 0.392571 21.82526

summary(glm\_model)

##   
## Call:  
## NULL  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -89.597 -19.241 -3.102 13.433 115.532   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 242.30849 76.15445 3.182  
## `Juris\_LongBritish Columbia` 4.63208 1.72819 2.680  
## Juris\_LongManitoba -5.80193 1.51934 -3.819  
## `Juris\_LongNational parks` -17.65427 1.74287 -10.129  
## `Juris\_LongNew Brunswick` 10.17541 2.31298 4.399  
## `Juris\_LongNewfoundland and Labrador` -14.81272 1.48833 -9.953  
## `Juris\_LongNorthwest Territories` -17.90664 1.63750 -10.935  
## `Juris\_LongNova Scotia` -7.99015 1.68126 -4.752  
## Juris\_LongOntario 2.31954 1.51156 1.535  
## `Juris\_LongPrince Edward Island` -29.50534 1.97291 -14.955  
## Juris\_LongQuebec -7.14402 1.48402 -4.814  
## Juris\_LongSaskatchewan -3.73044 1.76118 -2.118  
## Juris\_LongYukon -15.63102 1.57343 -9.934  
## Year -0.09588 0.03800 -2.523  
## Fire\_Cause\_HumanLightning 25.35580 0.94672 26.783  
## Fire\_Cause\_HumanUnspecified -5.01223 0.95200 -5.265  
## Protection.zoneLimited -12.86508 0.76702 -16.773  
## Protection.zoneUnspecified 27.41256 4.63841 5.910  
## Response.categoryModified -36.60685 0.74030 -49.449  
## Response.categoryNone -36.76037 0.73179 -50.233  
## Response.categoryUnspecified 12.20025 3.12123 3.909  
## Pr(>|t|)   
## (Intercept) 0.001468 \*\*   
## `Juris\_LongBritish Columbia` 0.007369 \*\*   
## Juris\_LongManitoba 0.000135 \*\*\*  
## `Juris\_LongNational parks` < 2e-16 \*\*\*  
## `Juris\_LongNew Brunswick` 1.10e-05 \*\*\*  
## `Juris\_LongNewfoundland and Labrador` < 2e-16 \*\*\*  
## `Juris\_LongNorthwest Territories` < 2e-16 \*\*\*  
## `Juris\_LongNova Scotia` 2.04e-06 \*\*\*  
## Juris\_LongOntario 0.124932   
## `Juris\_LongPrince Edward Island` < 2e-16 \*\*\*  
## Juris\_LongQuebec 1.50e-06 \*\*\*  
## Juris\_LongSaskatchewan 0.034190 \*   
## Juris\_LongYukon < 2e-16 \*\*\*  
## Year 0.011649 \*   
## Fire\_Cause\_HumanLightning < 2e-16 \*\*\*  
## Fire\_Cause\_HumanUnspecified 1.43e-07 \*\*\*  
## Protection.zoneLimited < 2e-16 \*\*\*  
## Protection.zoneUnspecified 3.54e-09 \*\*\*  
## Response.categoryModified < 2e-16 \*\*\*  
## Response.categoryNone < 2e-16 \*\*\*  
## Response.categoryUnspecified 9.34e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 814.2138)  
##   
## Null deviance: 12402753 on 9214 degrees of freedom  
## Residual deviance: 7485882 on 9194 degrees of freedom  
## AIC: 87935  
##   
## Number of Fisher Scoring iterations: 2

glm\_model$finalModel

##   
## Call: NULL  
##   
## Coefficients:  
## (Intercept)   
## 242.30849   
## `Juris\_LongBritish Columbia`   
## 4.63208   
## Juris\_LongManitoba   
## -5.80193   
## `Juris\_LongNational parks`   
## -17.65427   
## `Juris\_LongNew Brunswick`   
## 10.17541   
## `Juris\_LongNewfoundland and Labrador`   
## -14.81272   
## `Juris\_LongNorthwest Territories`   
## -17.90664   
## `Juris\_LongNova Scotia`   
## -7.99015   
## Juris\_LongOntario   
## 2.31954   
## `Juris\_LongPrince Edward Island`   
## -29.50534   
## Juris\_LongQuebec   
## -7.14402   
## Juris\_LongSaskatchewan   
## -3.73044   
## Juris\_LongYukon   
## -15.63102   
## Year   
## -0.09588   
## Fire\_Cause\_HumanLightning   
## 25.35580   
## Fire\_Cause\_HumanUnspecified   
## -5.01223   
## Protection.zoneLimited   
## -12.86508   
## Protection.zoneUnspecified   
## 27.41256   
## Response.categoryModified   
## -36.60685   
## Response.categoryNone   
## -36.76037   
## Response.categoryUnspecified   
## 12.20025   
##   
## Degrees of Freedom: 9214 Total (i.e. Null); 9194 Residual  
## Null Deviance: 12400000   
## Residual Deviance: 7486000 AIC: 87930

glm\_model$modelType

## [1] "Regression"

############################################  
# Model 3  
# Summarize the results  
print(lasso\_Mod)

## The lasso   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results across tuning parameters:  
##   
## fraction RMSE Rsquared MAE   
## 0.1 34.61358 0.2854614 26.50920  
## 0.5 29.82419 0.3633080 21.42606  
## 0.9 28.61676 0.3919238 21.67310  
##   
## RMSE was used to select the optimal model using the smallest value.  
## The final value used for the model was fraction = 0.9.

summary(lasso\_Mod)

## Length Class Mode   
## call 4 -none- call   
## actions 23 -none- list   
## allset 20 -none- numeric   
## beta.pure 460 -none- numeric   
## vn 20 -none- character  
## mu 1 -none- numeric   
## normx 20 -none- numeric   
## meanx 20 -none- numeric   
## lambda 1 -none- numeric   
## L1norm 23 -none- numeric   
## penalty 23 -none- numeric   
## df 23 -none- numeric   
## Cp 23 -none- numeric   
## sigma2 1 -none- numeric   
## xNames 20 -none- character  
## problemType 1 -none- character  
## tuneValue 1 data.frame list   
## obsLevels 1 -none- logical   
## param 0 -none- list

lasso\_Mod$finalModel

##   
## Call:  
## elasticnet::enet(x = as.matrix(x), y = y, lambda = 0)  
## Cp statistics of the Lasso fit   
## Cp: 6019.796 5958.733 4987.575 3565.541 2349.214 1759.613 1554.527 1548.863 1520.272 1078.429 900.780 761.628 759.114 715.145 623.335 193.778 181.134 82.816 84.098 50.352 43.852 28.793 21.000   
## DF: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 20 20 21   
## Sequence of moves:  
## Response.categoryNone Response.categoryModified  
## Var 19 18  
## Step 1 2  
## Fire\_Cause\_HumanLightning Protection.zoneLimited  
## Var 14 16  
## Step 3 4  
## Juris\_LongNew Brunswick Juris\_LongBritish Columbia  
## Var 4 1  
## Step 5 6  
## Juris\_LongPrince Edward Island Juris\_LongOntario  
## Var 9 8  
## Step 7 8  
## Juris\_LongNewfoundland and Labrador Juris\_LongYukon  
## Var 5 12  
## Step 9 10  
## Juris\_LongNorthwest Territories Juris\_LongNational parks  
## Var 6 3  
## Step 11 12  
## Response.categoryUnspecified Protection.zoneUnspecified  
## Var 20 17  
## Step 13 14  
## Fire\_Cause\_HumanUnspecified Juris\_LongSaskatchewan Year  
## Var 15 11 13  
## Step 15 16 17  
## Juris\_LongNova Scotia Juris\_LongQuebec Juris\_LongManitoba  
## Var 7 10 2  
## Step 18 19 20  
## Juris\_LongSaskatchewan Juris\_LongSaskatchewan   
## Var -11 11 23  
## Step 21 22 23

lasso\_Mod$modelType

## [1] "Regression"

############################################  
# Model 4  
# Summarize the results for knn model  
print(knn\_model)

## k-Nearest Neighbors   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results across tuning parameters:  
##   
## k RMSE Rsquared MAE   
## 5 26.90877 0.4714403 17.17772  
## 7 27.27735 0.4669475 18.28519  
## 9 27.49980 0.4608009 18.64220  
##   
## RMSE was used to select the optimal model using the smallest value.  
## The final value used for the model was k = 5.

summary(knn\_model)

## Length Class Mode   
## learn 2 -none- list   
## k 1 -none- numeric   
## theDots 0 -none- list   
## xNames 20 -none- character  
## problemType 1 -none- character  
## tuneValue 1 data.frame list   
## obsLevels 1 -none- logical   
## param 0 -none- list

knn\_model$finalModel

## 5-nearest neighbor regression model

knn\_model$modelType

## [1] "Regression"

############################################  
# Model 5  
# Summarize the results for LF model  
print(LF\_model)

## Linear Regression with Forward Selection   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results across tuning parameters:  
##   
## nvmax RMSE Rsquared MAE   
## 2 31.62485 0.2571868 21.23867  
## 3 30.50836 0.3087205 19.88286  
## 4 29.93894 0.3342066 21.93797  
##   
## RMSE was used to select the optimal model using the smallest value.  
## The final value used for the model was nvmax = 4.

summary(LF\_model)

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: forward  
## Juris\_LongBritish Columbia Juris\_LongManitoba  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongNational parks Juris\_LongNew Brunswick  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongNewfoundland and Labrador  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "   
## Juris\_LongNorthwest Territories Juris\_LongNova Scotia  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongOntario Juris\_LongPrince Edward Island Juris\_LongQuebec  
## 1 ( 1 ) " " " " " "   
## 2 ( 1 ) " " " " " "   
## 3 ( 1 ) " " " " " "   
## 4 ( 1 ) " " " " " "   
## Juris\_LongSaskatchewan Juris\_LongYukon Year  
## 1 ( 1 ) " " " " " "   
## 2 ( 1 ) " " " " " "   
## 3 ( 1 ) " " " " " "   
## 4 ( 1 ) " " " " " "   
## Fire\_Cause\_HumanLightning Fire\_Cause\_HumanUnspecified  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) "\*" " "   
## 4 ( 1 ) "\*" " "   
## Protection.zoneLimited Protection.zoneUnspecified  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) "\*" " "   
## Response.categoryModified Response.categoryNone  
## 1 ( 1 ) " " "\*"   
## 2 ( 1 ) "\*" "\*"   
## 3 ( 1 ) "\*" "\*"   
## 4 ( 1 ) "\*" "\*"   
## Response.categoryUnspecified  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "

LF\_model$finalModel

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: forward

LF\_model$modelType

## [1] "Regression"

############################################  
# Model 6  
# Summarize the results for LB Modle  
print(LB\_model)

## Linear Regression with Backwards Selection   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results across tuning parameters:  
##   
## nvmax RMSE Rsquared MAE   
## 2 31.62485 0.2571868 21.23867  
## 3 30.50836 0.3087205 19.88286  
## 4 29.93894 0.3342066 21.93797  
##   
## RMSE was used to select the optimal model using the smallest value.  
## The final value used for the model was nvmax = 4.

summary(LB\_model)

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: backward  
## Juris\_LongBritish Columbia Juris\_LongManitoba  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongNational parks Juris\_LongNew Brunswick  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongNewfoundland and Labrador  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "   
## Juris\_LongNorthwest Territories Juris\_LongNova Scotia  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) " " " "   
## Juris\_LongOntario Juris\_LongPrince Edward Island Juris\_LongQuebec  
## 1 ( 1 ) " " " " " "   
## 2 ( 1 ) " " " " " "   
## 3 ( 1 ) " " " " " "   
## 4 ( 1 ) " " " " " "   
## Juris\_LongSaskatchewan Juris\_LongYukon Year  
## 1 ( 1 ) " " " " " "   
## 2 ( 1 ) " " " " " "   
## 3 ( 1 ) " " " " " "   
## 4 ( 1 ) " " " " " "   
## Fire\_Cause\_HumanLightning Fire\_Cause\_HumanUnspecified  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) "\*" " "   
## 4 ( 1 ) "\*" " "   
## Protection.zoneLimited Protection.zoneUnspecified  
## 1 ( 1 ) " " " "   
## 2 ( 1 ) " " " "   
## 3 ( 1 ) " " " "   
## 4 ( 1 ) "\*" " "   
## Response.categoryModified Response.categoryNone  
## 1 ( 1 ) " " "\*"   
## 2 ( 1 ) "\*" "\*"   
## 3 ( 1 ) "\*" "\*"   
## 4 ( 1 ) "\*" "\*"   
## Response.categoryUnspecified  
## 1 ( 1 ) " "   
## 2 ( 1 ) " "   
## 3 ( 1 ) " "   
## 4 ( 1 ) " "

LB\_model$finalModel

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: backward

LB\_model$modelType

## [1] "Regression"

############################################  
# Model 7  
# Summarize the results for lmStepAIC  
print(lmStepAIC\_Mod)

## Linear Regression with Stepwise Selection   
##   
## 9215 samples  
## 5 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 5 times)   
## Summary of sample sizes: 6143, 6144, 6143, 6144, 6143, 6143, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 28.60671 0.3922352 21.82608

summary(lmStepAIC\_Mod)

##   
## Call:  
## lm(formula = .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified, data = dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -89.597 -19.241 -3.102 13.433 115.532   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 242.30849 76.15445 3.182  
## `Juris\_LongBritish Columbia` 4.63208 1.72819 2.680  
## Juris\_LongManitoba -5.80193 1.51934 -3.819  
## `Juris\_LongNational parks` -17.65427 1.74287 -10.129  
## `Juris\_LongNew Brunswick` 10.17541 2.31298 4.399  
## `Juris\_LongNewfoundland and Labrador` -14.81272 1.48833 -9.953  
## `Juris\_LongNorthwest Territories` -17.90664 1.63750 -10.935  
## `Juris\_LongNova Scotia` -7.99015 1.68126 -4.752  
## Juris\_LongOntario 2.31954 1.51156 1.535  
## `Juris\_LongPrince Edward Island` -29.50534 1.97291 -14.955  
## Juris\_LongQuebec -7.14402 1.48402 -4.814  
## Juris\_LongSaskatchewan -3.73044 1.76118 -2.118  
## Juris\_LongYukon -15.63102 1.57343 -9.934  
## Year -0.09588 0.03800 -2.523  
## Fire\_Cause\_HumanLightning 25.35580 0.94672 26.783  
## Fire\_Cause\_HumanUnspecified -5.01223 0.95200 -5.265  
## Protection.zoneLimited -12.86508 0.76702 -16.773  
## Protection.zoneUnspecified 27.41256 4.63841 5.910  
## Response.categoryModified -36.60685 0.74030 -49.449  
## Response.categoryNone -36.76037 0.73179 -50.233  
## Response.categoryUnspecified 12.20025 3.12123 3.909  
## Pr(>|t|)   
## (Intercept) 0.001468 \*\*   
## `Juris\_LongBritish Columbia` 0.007369 \*\*   
## Juris\_LongManitoba 0.000135 \*\*\*  
## `Juris\_LongNational parks` < 2e-16 \*\*\*  
## `Juris\_LongNew Brunswick` 1.10e-05 \*\*\*  
## `Juris\_LongNewfoundland and Labrador` < 2e-16 \*\*\*  
## `Juris\_LongNorthwest Territories` < 2e-16 \*\*\*  
## `Juris\_LongNova Scotia` 2.04e-06 \*\*\*  
## Juris\_LongOntario 0.124932   
## `Juris\_LongPrince Edward Island` < 2e-16 \*\*\*  
## Juris\_LongQuebec 1.50e-06 \*\*\*  
## Juris\_LongSaskatchewan 0.034190 \*   
## Juris\_LongYukon < 2e-16 \*\*\*  
## Year 0.011649 \*   
## Fire\_Cause\_HumanLightning < 2e-16 \*\*\*  
## Fire\_Cause\_HumanUnspecified 1.43e-07 \*\*\*  
## Protection.zoneLimited < 2e-16 \*\*\*  
## Protection.zoneUnspecified 3.54e-09 \*\*\*  
## Response.categoryModified < 2e-16 \*\*\*  
## Response.categoryNone < 2e-16 \*\*\*  
## Response.categoryUnspecified 9.34e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 28.53 on 9194 degrees of freedom  
## Multiple R-squared: 0.3964, Adjusted R-squared: 0.3951   
## F-statistic: 301.9 on 20 and 9194 DF, p-value: < 2.2e-16

lmStepAIC\_Mod$finalModel

##   
## Call:  
## lm(formula = .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified, data = dat)  
##   
## Coefficients:  
## (Intercept)   
## 242.30849   
## `Juris\_LongBritish Columbia`   
## 4.63208   
## Juris\_LongManitoba   
## -5.80193   
## `Juris\_LongNational parks`   
## -17.65427   
## `Juris\_LongNew Brunswick`   
## 10.17541   
## `Juris\_LongNewfoundland and Labrador`   
## -14.81272   
## `Juris\_LongNorthwest Territories`   
## -17.90664   
## `Juris\_LongNova Scotia`   
## -7.99015   
## Juris\_LongOntario   
## 2.31954   
## `Juris\_LongPrince Edward Island`   
## -29.50534   
## Juris\_LongQuebec   
## -7.14402   
## Juris\_LongSaskatchewan   
## -3.73044   
## Juris\_LongYukon   
## -15.63102   
## Year   
## -0.09588   
## Fire\_Cause\_HumanLightning   
## 25.35580   
## Fire\_Cause\_HumanUnspecified   
## -5.01223   
## Protection.zoneLimited   
## -12.86508   
## Protection.zoneUnspecified   
## 27.41256   
## Response.categoryModified   
## -36.60685   
## Response.categoryNone   
## -36.76037   
## Response.categoryUnspecified   
## 12.20025

lmStepAIC\_Mod$modelType

## [1] "Regression"

############################################

## OUTLIERS CAPPED

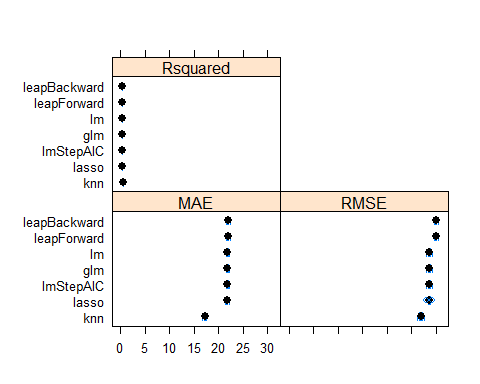
## Evaluation of techniques

## View the results of the models

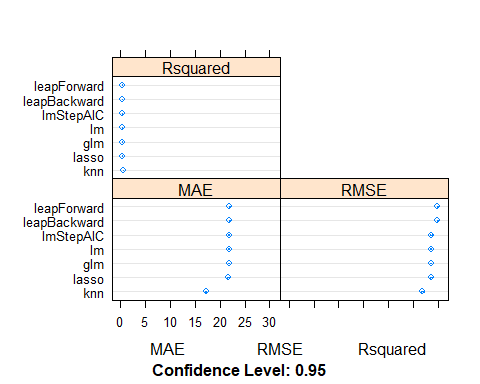
# Evaluation of techniques  
results <- resamples(list(lm=lm\_model, glm=glm\_model,lasso=lasso\_Mod, knn=knn\_model, leapForward=LF\_model,leapBackward=LB\_model, lmStepAIC=lmStepAIC\_Mod))  
summary(results)

##   
## Call:  
## summary.resamples(object = results)  
##   
## Models: lm, glm, lasso, knn, leapForward, leapBackward, lmStepAIC   
## Number of resamples: 15   
##   
## MAE   
## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's  
## lm 21.53339 21.68737 21.82130 21.82526 21.93564 22.24242 0  
## glm 21.53339 21.68737 21.82130 21.82526 21.93564 22.24242 0  
## lasso 21.35685 21.54987 21.68147 21.67310 21.77985 22.09815 0  
## knn 16.71977 17.01028 17.17379 17.17772 17.36791 17.56501 0  
## leapForward 21.64467 21.73973 21.97749 21.93797 22.05939 22.30420 0  
## leapBackward 21.64467 21.73973 21.97749 21.93797 22.05939 22.30420 0  
## lmStepAIC 21.54732 21.70436 21.82054 21.82608 21.93276 22.22935 0  
##   
## RMSE   
## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's  
## lm 28.08313 28.43720 28.55949 28.59885 28.74751 29.16927 0  
## glm 28.08313 28.43720 28.55949 28.59885 28.74751 29.16927 0  
## lasso 28.10365 28.48871 28.59152 28.61676 28.74342 29.17116 0  
## knn 26.25559 26.67743 26.89454 26.90877 27.17055 27.65418 0  
## leapForward 29.59733 29.76946 29.97219 29.93894 30.06610 30.46950 0  
## leapBackward 29.59733 29.76946 29.97219 29.93894 30.06610 30.46950 0  
## lmStepAIC 28.08313 28.46010 28.56481 28.60671 28.75393 29.17147 0  
##   
## Rsquared   
## Min. 1st Qu. Median Mean 3rd Qu. Max.  
## lm 0.3775842 0.3885947 0.3898913 0.3925710 0.3982131 0.4089038  
## glm 0.3775842 0.3885947 0.3898913 0.3925710 0.3982131 0.4089038  
## lasso 0.3776691 0.3889653 0.3898007 0.3919238 0.3965534 0.4062071  
## knn 0.4441612 0.4587133 0.4694606 0.4714403 0.4857192 0.5024296  
## leapForward 0.3251086 0.3288203 0.3351751 0.3342066 0.3381589 0.3452045  
## leapBackward 0.3251086 0.3288203 0.3351751 0.3342066 0.3381589 0.3452045  
## lmStepAIC 0.3774531 0.3883748 0.3896676 0.3922352 0.3981590 0.4069640  
## NA's  
## lm 0  
## glm 0  
## lasso 0  
## knn 0  
## leapForward 0  
## leapBackward 0  
## lmStepAIC 0

# Compare results with boxplots   
bwplot(results)



# Compare results with dot plots   
dotplot(results)



## OUTLIERS CAPPED

## Find the best results for each model.

get\_best\_result = function(caret\_fit) {  
 best = which(rownames(caret\_fit$results) == rownames(caret\_fit$bestTune))  
 best\_result = caret\_fit$results[best, ]  
 rownames(best\_result) = NULL  
 best\_result  
}  
  
###################################################################################  
lm\_best <- data.frame(get\_best\_result(lm\_model) %>%  
 dplyr::select(2:4))  
lm\_model$finalModel

##   
## Call:  
## lm(formula = .outcome ~ ., data = dat)  
##   
## Coefficients:  
## (Intercept)   
## 242.30849   
## `Juris\_LongBritish Columbia`   
## 4.63208   
## Juris\_LongManitoba   
## -5.80193   
## `Juris\_LongNational parks`   
## -17.65427   
## `Juris\_LongNew Brunswick`   
## 10.17541   
## `Juris\_LongNewfoundland and Labrador`   
## -14.81272   
## `Juris\_LongNorthwest Territories`   
## -17.90664   
## `Juris\_LongNova Scotia`   
## -7.99015   
## Juris\_LongOntario   
## 2.31954   
## `Juris\_LongPrince Edward Island`   
## -29.50534   
## Juris\_LongQuebec   
## -7.14402   
## Juris\_LongSaskatchewan   
## -3.73044   
## Juris\_LongYukon   
## -15.63102   
## Year   
## -0.09588   
## Fire\_Cause\_HumanLightning   
## 25.35580   
## Fire\_Cause\_HumanUnspecified   
## -5.01223   
## Protection.zoneLimited   
## -12.86508   
## Protection.zoneUnspecified   
## 27.41256   
## Response.categoryModified   
## -36.60685   
## Response.categoryNone   
## -36.76037   
## Response.categoryUnspecified   
## 12.20025

###################################################################################  
glm\_best <- data.frame(get\_best\_result(glm\_model) %>%   
 dplyr::select(2:4))  
glm\_model$finalModel

##   
## Call: NULL  
##   
## Coefficients:  
## (Intercept)   
## 242.30849   
## `Juris\_LongBritish Columbia`   
## 4.63208   
## Juris\_LongManitoba   
## -5.80193   
## `Juris\_LongNational parks`   
## -17.65427   
## `Juris\_LongNew Brunswick`   
## 10.17541   
## `Juris\_LongNewfoundland and Labrador`   
## -14.81272   
## `Juris\_LongNorthwest Territories`   
## -17.90664   
## `Juris\_LongNova Scotia`   
## -7.99015   
## Juris\_LongOntario   
## 2.31954   
## `Juris\_LongPrince Edward Island`   
## -29.50534   
## Juris\_LongQuebec   
## -7.14402   
## Juris\_LongSaskatchewan   
## -3.73044   
## Juris\_LongYukon   
## -15.63102   
## Year   
## -0.09588   
## Fire\_Cause\_HumanLightning   
## 25.35580   
## Fire\_Cause\_HumanUnspecified   
## -5.01223   
## Protection.zoneLimited   
## -12.86508   
## Protection.zoneUnspecified   
## 27.41256   
## Response.categoryModified   
## -36.60685   
## Response.categoryNone   
## -36.76037   
## Response.categoryUnspecified   
## 12.20025   
##   
## Degrees of Freedom: 9214 Total (i.e. Null); 9194 Residual  
## Null Deviance: 12400000   
## Residual Deviance: 7486000 AIC: 87930

###################################################################################  
lasso\_best <- data.frame(get\_best\_result(lasso\_Mod) %>%   
 dplyr::select(2:4))  
lasso\_Mod$finalModel

##   
## Call:  
## elasticnet::enet(x = as.matrix(x), y = y, lambda = 0)  
## Cp statistics of the Lasso fit   
## Cp: 6019.796 5958.733 4987.575 3565.541 2349.214 1759.613 1554.527 1548.863 1520.272 1078.429 900.780 761.628 759.114 715.145 623.335 193.778 181.134 82.816 84.098 50.352 43.852 28.793 21.000   
## DF: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 20 20 21   
## Sequence of moves:  
## Response.categoryNone Response.categoryModified  
## Var 19 18  
## Step 1 2  
## Fire\_Cause\_HumanLightning Protection.zoneLimited  
## Var 14 16  
## Step 3 4  
## Juris\_LongNew Brunswick Juris\_LongBritish Columbia  
## Var 4 1  
## Step 5 6  
## Juris\_LongPrince Edward Island Juris\_LongOntario  
## Var 9 8  
## Step 7 8  
## Juris\_LongNewfoundland and Labrador Juris\_LongYukon  
## Var 5 12  
## Step 9 10  
## Juris\_LongNorthwest Territories Juris\_LongNational parks  
## Var 6 3  
## Step 11 12  
## Response.categoryUnspecified Protection.zoneUnspecified  
## Var 20 17  
## Step 13 14  
## Fire\_Cause\_HumanUnspecified Juris\_LongSaskatchewan Year  
## Var 15 11 13  
## Step 15 16 17  
## Juris\_LongNova Scotia Juris\_LongQuebec Juris\_LongManitoba  
## Var 7 10 2  
## Step 18 19 20  
## Juris\_LongSaskatchewan Juris\_LongSaskatchewan   
## Var -11 11 23  
## Step 21 22 23

###################################################################################  
knn\_best <- data.frame(get\_best\_result(knn\_model) %>%   
 dplyr::select(2:4))  
knn\_model$finalModel

## 5-nearest neighbor regression model

###################################################################################  
LF\_best <- data.frame(get\_best\_result(LF\_model) %>%   
 dplyr::select(2:4))  
LF\_model$finalModel

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: forward

###################################################################################  
LB\_best <- data.frame(get\_best\_result(LB\_model) %>%   
 dplyr::select(2:4))  
LB\_model$finalModel

## Subset selection object  
## 20 Variables (and intercept)  
## Forced in Forced out  
## Juris\_LongBritish Columbia FALSE FALSE  
## Juris\_LongManitoba FALSE FALSE  
## Juris\_LongNational parks FALSE FALSE  
## Juris\_LongNew Brunswick FALSE FALSE  
## Juris\_LongNewfoundland and Labrador FALSE FALSE  
## Juris\_LongNorthwest Territories FALSE FALSE  
## Juris\_LongNova Scotia FALSE FALSE  
## Juris\_LongOntario FALSE FALSE  
## Juris\_LongPrince Edward Island FALSE FALSE  
## Juris\_LongQuebec FALSE FALSE  
## Juris\_LongSaskatchewan FALSE FALSE  
## Juris\_LongYukon FALSE FALSE  
## Year FALSE FALSE  
## Fire\_Cause\_HumanLightning FALSE FALSE  
## Fire\_Cause\_HumanUnspecified FALSE FALSE  
## Protection.zoneLimited FALSE FALSE  
## Protection.zoneUnspecified FALSE FALSE  
## Response.categoryModified FALSE FALSE  
## Response.categoryNone FALSE FALSE  
## Response.categoryUnspecified FALSE FALSE  
## 1 subsets of each size up to 4  
## Selection Algorithm: backward

###################################################################################  
lmStep\_best <- data.frame(get\_best\_result(lmStepAIC\_Mod) %>%   
 dplyr::select(2:4))  
lmStepAIC\_Mod$finalModel

##   
## Call:  
## lm(formula = .outcome ~ `Juris\_LongBritish Columbia` + Juris\_LongManitoba +   
## `Juris\_LongNational parks` + `Juris\_LongNew Brunswick` +   
## `Juris\_LongNewfoundland and Labrador` + `Juris\_LongNorthwest Territories` +   
## `Juris\_LongNova Scotia` + Juris\_LongOntario + `Juris\_LongPrince Edward Island` +   
## Juris\_LongQuebec + Juris\_LongSaskatchewan + Juris\_LongYukon +   
## Year + Fire\_Cause\_HumanLightning + Fire\_Cause\_HumanUnspecified +   
## Protection.zoneLimited + Protection.zoneUnspecified + Response.categoryModified +   
## Response.categoryNone + Response.categoryUnspecified, data = dat)  
##   
## Coefficients:  
## (Intercept)   
## 242.30849   
## `Juris\_LongBritish Columbia`   
## 4.63208   
## Juris\_LongManitoba   
## -5.80193   
## `Juris\_LongNational parks`   
## -17.65427   
## `Juris\_LongNew Brunswick`   
## 10.17541   
## `Juris\_LongNewfoundland and Labrador`   
## -14.81272   
## `Juris\_LongNorthwest Territories`   
## -17.90664   
## `Juris\_LongNova Scotia`   
## -7.99015   
## Juris\_LongOntario   
## 2.31954   
## `Juris\_LongPrince Edward Island`   
## -29.50534   
## Juris\_LongQuebec   
## -7.14402   
## Juris\_LongSaskatchewan   
## -3.73044   
## Juris\_LongYukon   
## -15.63102   
## Year   
## -0.09588   
## Fire\_Cause\_HumanLightning   
## 25.35580   
## Fire\_Cause\_HumanUnspecified   
## -5.01223   
## Protection.zoneLimited   
## -12.86508   
## Protection.zoneUnspecified   
## 27.41256   
## Response.categoryModified   
## -36.60685   
## Response.categoryNone   
## -36.76037   
## Response.categoryUnspecified   
## 12.20025

###################################################################################  
total <- rbind(lm\_best, glm\_best, lasso\_best, knn\_best, LF\_best, LB\_best, lmStep\_best)  
total\_best\_train <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 total,  
 Model\_Time)%>%   
 dplyr::select(-5)  
total\_best\_train

## Name RMSE Rsquared MAE Time\_train  
## 1 lm\_model 28.59885 0.3925710 21.82526 1.643645 secs  
## 2 glm\_model 28.59885 0.3925710 21.82526 1.738344 secs  
## 3 lasso\_Mod 28.61676 0.3919238 21.67310 2.068453 secs  
## 4 knn\_model 26.90877 0.4714403 17.17772 25.646455 secs  
## 5 LF\_model 29.93894 0.3342066 21.93797 1.012250 secs  
## 6 LB\_model 29.93894 0.3342066 21.93797 1.079079 secs  
## 7 lmStepAIC\_Mod 28.60671 0.3922352 21.82608 4.663523 secs

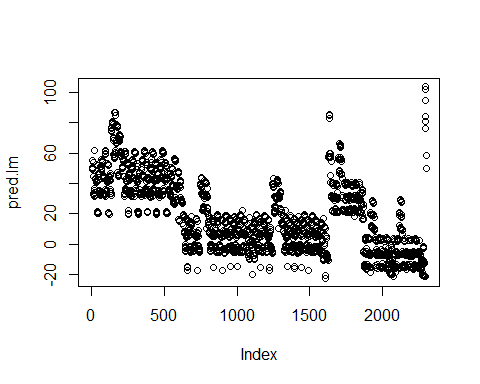
## OUTLIERS CAPPED

## Predict on test set

#############################################################################################################################################  
# Model 1: lm model  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.lm = predict(lm\_model, newdata = test.set)  
output <- cbind(test.set, pred.lm)  
head(output)

## Cause Jurisdiction capped\_Number Protection.zone  
## 13 Forest industry AB 96 Intensive  
## 17 Forest industry AB 96 Intensive  
## 18 Forest industry AB 96 Intensive  
## 19 Forest industry AB 96 Intensive  
## 26 Forest industry AB 96 Intensive  
## 37 Forest industry BC 3 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 13 Full 2002 Alberta People Early 10s 2000s  
## 17 Full 2006 Alberta People Late 10s 2000s  
## 18 Full 2007 Alberta People Late 10s 2000s  
## 19 Full 2008 Alberta People Late 10s 2000s  
## 26 Full 2015 Alberta People Early 20s 2010s  
## 37 Full 1998 British Columbia People Late 90s 1990s  
## Region Fire\_Cause\_Human pred.lm  
## 13 Prairie Region Human 50.35293  
## 17 Prairie Region Human 49.96940  
## 18 Prairie Region Human 49.87352  
## 19 Prairie Region Human 49.77764  
## 26 Prairie Region Human 49.10647  
## 37 Pacific Region Human 55.36853

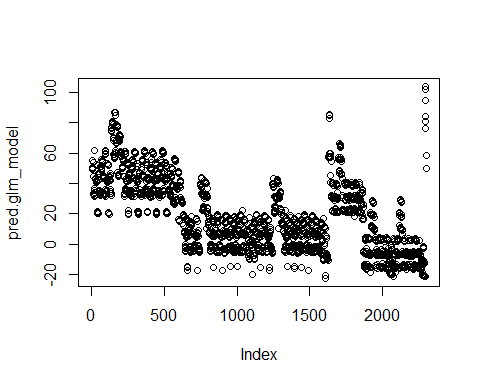
plot(pred.lm)



# Stop the clock  
end.time <- Sys.time()  
  
pre.lm\_model\_time.taken <- end.time - start.time  
#############################################################################################################################################  
# Model 2: glm model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.glm\_model = predict(glm\_model, newdata = test.set)  
output <- cbind(test.set, pred.glm\_model)  
head(output)

## Cause Jurisdiction capped\_Number Protection.zone  
## 13 Forest industry AB 96 Intensive  
## 17 Forest industry AB 96 Intensive  
## 18 Forest industry AB 96 Intensive  
## 19 Forest industry AB 96 Intensive  
## 26 Forest industry AB 96 Intensive  
## 37 Forest industry BC 3 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 13 Full 2002 Alberta People Early 10s 2000s  
## 17 Full 2006 Alberta People Late 10s 2000s  
## 18 Full 2007 Alberta People Late 10s 2000s  
## 19 Full 2008 Alberta People Late 10s 2000s  
## 26 Full 2015 Alberta People Early 20s 2010s  
## 37 Full 1998 British Columbia People Late 90s 1990s  
## Region Fire\_Cause\_Human pred.glm\_model  
## 13 Prairie Region Human 50.35293  
## 17 Prairie Region Human 49.96940  
## 18 Prairie Region Human 49.87352  
## 19 Prairie Region Human 49.77764  
## 26 Prairie Region Human 49.10647  
## 37 Pacific Region Human 55.36853

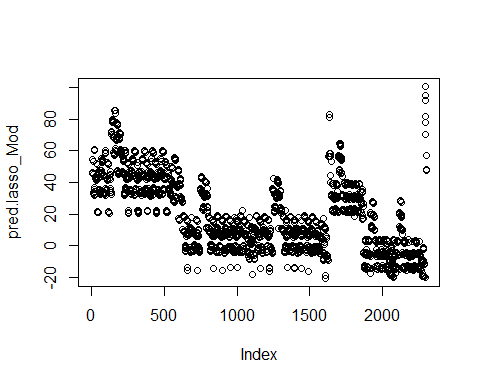
plot(pred.glm\_model)



# Stop the clock  
end.time <- Sys.time()  
  
pre.glm\_model\_time.taken <- end.time - start.time  
#glm\_model\_time.taken  
  
#############################################################################################################################################  
# Model 3: lasso model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.lasso\_Mod = predict(lasso\_Mod, newdata = test.set)  
output <- cbind(test.set, pred.lasso\_Mod)  
head(output)

## Cause Jurisdiction capped\_Number Protection.zone  
## 13 Forest industry AB 96 Intensive  
## 17 Forest industry AB 96 Intensive  
## 18 Forest industry AB 96 Intensive  
## 19 Forest industry AB 96 Intensive  
## 26 Forest industry AB 96 Intensive  
## 37 Forest industry BC 3 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 13 Full 2002 Alberta People Early 10s 2000s  
## 17 Full 2006 Alberta People Late 10s 2000s  
## 18 Full 2007 Alberta People Late 10s 2000s  
## 19 Full 2008 Alberta People Late 10s 2000s  
## 26 Full 2015 Alberta People Early 20s 2010s  
## 37 Full 1998 British Columbia People Late 90s 1990s  
## Region Fire\_Cause\_Human pred.lasso\_Mod  
## 13 Prairie Region Human 46.45870  
## 17 Prairie Region Human 46.13247  
## 18 Prairie Region Human 46.05091  
## 19 Prairie Region Human 45.96936  
## 26 Prairie Region Human 45.39846  
## 37 Pacific Region Human 54.38707

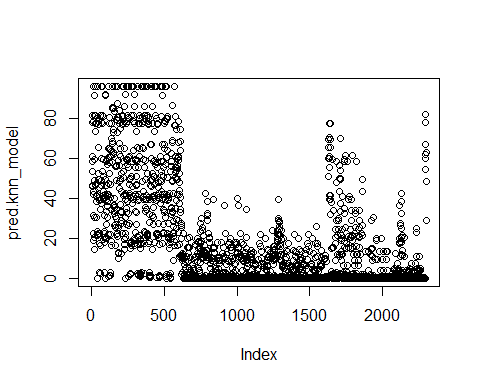
plot(pred.lasso\_Mod)



# Stop the clock  
end.time <- Sys.time()  
  
pre.lasso\_Mod\_time.taken <- end.time - start.time  
#lasso\_Mod\_time.taken  
#############################################################################################################################################  
# Model 4: knn model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.knn\_model = predict(knn\_model, newdata = test.set)  
output <- cbind(test.set, pred.knn\_model)  
head(output)

## Cause Jurisdiction capped\_Number Protection.zone  
## 13 Forest industry AB 96 Intensive  
## 17 Forest industry AB 96 Intensive  
## 18 Forest industry AB 96 Intensive  
## 19 Forest industry AB 96 Intensive  
## 26 Forest industry AB 96 Intensive  
## 37 Forest industry BC 3 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 13 Full 2002 Alberta People Early 10s 2000s  
## 17 Full 2006 Alberta People Late 10s 2000s  
## 18 Full 2007 Alberta People Late 10s 2000s  
## 19 Full 2008 Alberta People Late 10s 2000s  
## 26 Full 2015 Alberta People Early 20s 2010s  
## 37 Full 1998 British Columbia People Late 90s 1990s  
## Region Fire\_Cause\_Human pred.knn\_model  
## 13 Prairie Region Human 81.33333  
## 17 Prairie Region Human 46.30682  
## 18 Prairie Region Human 45.73626  
## 19 Prairie Region Human 78.00000  
## 26 Prairie Region Human 61.00000  
## 37 Pacific Region Human 58.20000

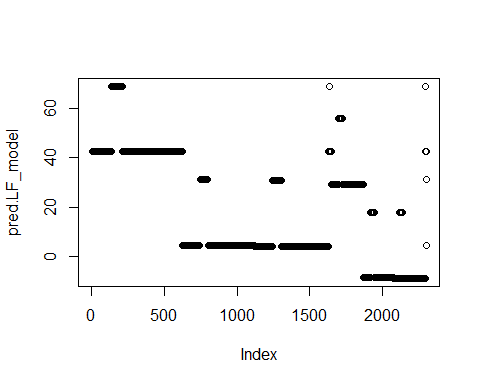
plot(pred.knn\_model)



# Stop the clock  
end.time <- Sys.time()  
  
pre.knn\_model\_time.taken <- end.time - start.time  
#knn\_model\_time.taken  
#############################################################################################################################################  
# Model 5: leapForward model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.LF\_model = predict(LF\_model, newdata = test.set)  
output <- cbind(test.set, pred.LF\_model)  
head(output)

## Cause Jurisdiction capped\_Number Protection.zone  
## 13 Forest industry AB 96 Intensive  
## 17 Forest industry AB 96 Intensive  
## 18 Forest industry AB 96 Intensive  
## 19 Forest industry AB 96 Intensive  
## 26 Forest industry AB 96 Intensive  
## 37 Forest industry BC 3 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 13 Full 2002 Alberta People Early 10s 2000s  
## 17 Full 2006 Alberta People Late 10s 2000s  
## 18 Full 2007 Alberta People Late 10s 2000s  
## 19 Full 2008 Alberta People Late 10s 2000s  
## 26 Full 2015 Alberta People Early 20s 2010s  
## 37 Full 1998 British Columbia People Late 90s 1990s  
## Region Fire\_Cause\_Human pred.LF\_model  
## 13 Prairie Region Human 42.42734  
## 17 Prairie Region Human 42.42734  
## 18 Prairie Region Human 42.42734  
## 19 Prairie Region Human 42.42734  
## 26 Prairie Region Human 42.42734  
## 37 Pacific Region Human 42.42734

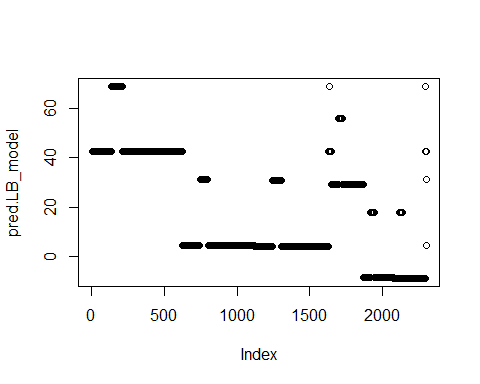
plot(pred.LF\_model)



# Stop the clock  
end.time <- Sys.time()  
  
pre.LF\_model\_time.taken <- end.time - start.time  
#LF\_model\_time.taken  
#############################################################################################################################################  
# Model 6: leapBackward model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.LB\_model = predict(LB\_model, newdata = test.set)  
output <- cbind(test.set, pred.LB\_model)  
head(output)

## Cause Jurisdiction capped\_Number Protection.zone  
## 13 Forest industry AB 96 Intensive  
## 17 Forest industry AB 96 Intensive  
## 18 Forest industry AB 96 Intensive  
## 19 Forest industry AB 96 Intensive  
## 26 Forest industry AB 96 Intensive  
## 37 Forest industry BC 3 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 13 Full 2002 Alberta People Early 10s 2000s  
## 17 Full 2006 Alberta People Late 10s 2000s  
## 18 Full 2007 Alberta People Late 10s 2000s  
## 19 Full 2008 Alberta People Late 10s 2000s  
## 26 Full 2015 Alberta People Early 20s 2010s  
## 37 Full 1998 British Columbia People Late 90s 1990s  
## Region Fire\_Cause\_Human pred.LB\_model  
## 13 Prairie Region Human 42.42734  
## 17 Prairie Region Human 42.42734  
## 18 Prairie Region Human 42.42734  
## 19 Prairie Region Human 42.42734  
## 26 Prairie Region Human 42.42734  
## 37 Pacific Region Human 42.42734

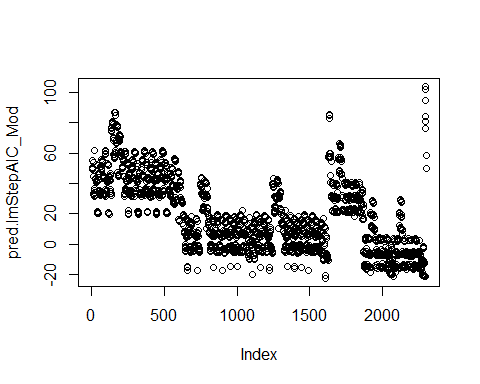
plot(pred.LB\_model)



# Stop the clock  
end.time <- Sys.time()  
  
pre.LB\_model\_time.taken <- end.time - start.time  
#LB\_model\_time.taken  
#############################################################################################################################################  
# Model 7: lmStepAIC model  
# Start the clock  
start.time <- Sys.time()  
  
set.seed(1235)   
pred.lmStepAIC\_Mod = predict(lmStepAIC\_Mod, newdata = test.set)  
output <- cbind(test.set, pred.lmStepAIC\_Mod)  
head(output)

## Cause Jurisdiction capped\_Number Protection.zone  
## 13 Forest industry AB 96 Intensive  
## 17 Forest industry AB 96 Intensive  
## 18 Forest industry AB 96 Intensive  
## 19 Forest industry AB 96 Intensive  
## 26 Forest industry AB 96 Intensive  
## 37 Forest industry BC 3 Intensive  
## Response.category Year Juris\_Long Cause\_Grouped Time1 Time2  
## 13 Full 2002 Alberta People Early 10s 2000s  
## 17 Full 2006 Alberta People Late 10s 2000s  
## 18 Full 2007 Alberta People Late 10s 2000s  
## 19 Full 2008 Alberta People Late 10s 2000s  
## 26 Full 2015 Alberta People Early 20s 2010s  
## 37 Full 1998 British Columbia People Late 90s 1990s  
## Region Fire\_Cause\_Human pred.lmStepAIC\_Mod  
## 13 Prairie Region Human 50.35293  
## 17 Prairie Region Human 49.96940  
## 18 Prairie Region Human 49.87352  
## 19 Prairie Region Human 49.77764  
## 26 Prairie Region Human 49.10647  
## 37 Pacific Region Human 55.36853

plot(pred.lmStepAIC\_Mod)



# Stop the clock  
end.time <- Sys.time()  
  
pre.lmStepAIC\_Mod\_time.taken <- end.time - start.time  
#lmStepAIC\_Mod\_time.taken  
#############################################################################################################################################

## OUTLIERS CAPPED

## Time taken for prediction

Model\_TimePred <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 Predict\_Time=c(pre.lm\_model\_time.taken, pre.glm\_model\_time.taken, pre.lasso\_Mod\_time.taken, pre.knn\_model\_time.taken, pre.LF\_model\_time.taken, pre.LB\_model\_time.taken, pre.lmStepAIC\_Mod\_time.taken))  
Model\_TimePred

## Name Predict\_Time  
## 1 lm\_model 0.10870695 secs  
## 2 glm\_model 0.12167406 secs  
## 3 lasso\_Mod 0.09374905 secs  
## 4 knn\_model 0.60438299 secs  
## 5 LF\_model 0.09973288 secs  
## 6 LB\_model 0.10173392 secs  
## 7 lmStepAIC\_Mod 0.11070514 secs

## OUTLIERS CAPPED

## Compare correlation between actual and predicted

#############################################################################################################################################  
# Formula to calculate correlation   
corr\_lm\_model <- round(cor(test.set$capped\_Number, predict(lm\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_glm\_model <- round(cor(test.set$capped\_Number, predict(glm\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_lasso\_Mod <- round(cor(test.set$capped\_Number, predict(lasso\_Mod, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_knn\_model <- round(cor(test.set$capped\_Number, predict(knn\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_LF\_model <- round(cor(test.set$capped\_Number, predict(LF\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_LB\_model <- round(cor(test.set$capped\_Number, predict(LB\_model, test.set), method = c("pearson", "kendall", "spearman")),2)  
corr\_lmStepAIC\_Mod <- round(cor(test.set$capped\_Number, predict(lmStepAIC\_Mod, test.set), method = c("pearson", "kendall", "spearman")),2)  
  
total\_corr <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 corr\_pred=c(corr\_lm\_model, corr\_glm\_model, corr\_lasso\_Mod, corr\_knn\_model, corr\_LF\_model, corr\_LB\_model, corr\_lmStepAIC\_Mod))  
total\_corr

## Name corr\_pred  
## 1 lm\_model 0.64  
## 2 glm\_model 0.64  
## 3 lasso\_Mod 0.64  
## 4 knn\_model 0.75  
## 5 LF\_model 0.58  
## 6 LB\_model 0.58  
## 7 lmStepAIC\_Mod 0.64

## OUTLIERS CAPPED

## RMSE between actual and predicted

#############################################################################################################################################  
# Formula to calculate RMSE on test set  
calc\_rmse = function(actual, predicted) {  
 sqrt(mean((actual - predicted) ^ 2))  
}  
  
# RMSE value on test set  
rmse\_lm\_model <- calc\_rmse(actual = test.set$capped\_Number,  
 predicted = predict(lm\_model, test.set))  
  
# RMSE value on test set  
rmse\_glm\_model <- calc\_rmse(actual = test.set$capped\_Number,  
 predicted = predict(glm\_model, test.set))  
  
# RMSE value on test set  
rmse\_lasso\_Mod <- calc\_rmse(actual = test.set$capped\_Number,  
 predicted = predict(lasso\_Mod, test.set))  
  
# RMSE value on test set  
rmse\_knn\_model <- calc\_rmse(actual = test.set$capped\_Number,  
 predicted = predict(knn\_model, test.set))  
  
# RMSE value on test set  
rmse\_LF\_model <- calc\_rmse(actual = test.set$capped\_Number,  
 predicted = predict(LF\_model, test.set))  
  
# RMSE value on test set  
rmse\_LB\_model <- calc\_rmse(actual = test.set$capped\_Number,  
 predicted = predict(LB\_model, test.set))  
  
# RMSE value on test set  
rmse\_lmStepAIC\_Mod <- calc\_rmse(actual = test.set$capped\_Number,  
 predicted = predict(lmStepAIC\_Mod, test.set))  
  
total\_RMSE <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 RMSE\_Pred=c(rmse\_lm\_model, rmse\_glm\_model, rmse\_lasso\_Mod, rmse\_knn\_model, rmse\_LF\_model, rmse\_LB\_model, rmse\_lmStepAIC\_Mod))  
total\_RMSE

## Name RMSE\_Pred  
## 1 lm\_model 28.21732  
## 2 glm\_model 28.21732  
## 3 lasso\_Mod 28.24713  
## 4 knn\_model 24.58882  
## 5 LF\_model 29.94359  
## 6 LB\_model 29.94359  
## 7 lmStepAIC\_Mod 28.21732

## OUTLIERS CAPPED

## MAE between actual and predicted

#############################################################################################################################################  
  
MAE\_test\_lm\_model <- MAE(test.set$capped\_Number, predict(lm\_model, test.set))  
MAE\_test\_glm\_model <- MAE(test.set$capped\_Number, predict(glm\_model, test.set))  
MAE\_test\_lasso\_Mod <- MAE(test.set$capped\_Number, predict(lasso\_Mod, test.set))  
MAE\_test\_knn\_model <- MAE(test.set$capped\_Number, predict(knn\_model, test.set))  
MAE\_test\_LF\_model <- MAE(test.set$capped\_Number, predict(LF\_model, test.set))  
MAE\_test\_LB\_model <- MAE(test.set$capped\_Number, predict(LB\_model, test.set))  
MAE\_test\_lmStepAIC\_Mod <- MAE(test.set$capped\_Number, predict(lmStepAIC\_Mod, test.set))  
  
total\_MAE <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 MAE\_Pred=c(MAE\_test\_lm\_model, MAE\_test\_glm\_model, MAE\_test\_lasso\_Mod, MAE\_test\_knn\_model, MAE\_test\_LF\_model, MAE\_test\_LB\_model, MAE\_test\_lmStepAIC\_Mod))  
total\_MAE

## Name MAE\_Pred  
## 1 lm\_model 21.65778  
## 2 glm\_model 21.65778  
## 3 lasso\_Mod 21.49014  
## 4 knn\_model 13.63390  
## 5 LF\_model 22.06808  
## 6 LB\_model 22.06808  
## 7 lmStepAIC\_Mod 21.65778

## OUTLIERS CAPPED

## r squared between actual and predicted

calc\_rss = function(actual, predicted) {  
 sum((predicted - actual) ^ 2) ## residual sum of squares  
}  
  
calc\_tss = function(actual, predicted) {  
 sum((actual - mean(actual)) ^ 2) ## total sum of squares  
}  
  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$capped\_Number,  
 predicted = predict(lm\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$capped\_Number,  
 predicted = predict(lm\_model, test.set))  
  
rsq\_lm\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$capped\_Number,  
 predicted = predict(glm\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$capped\_Number,  
 predicted = predict(glm\_model, test.set))  
  
rsq\_glm\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$capped\_Number,  
 predicted = predict(lasso\_Mod, test.set))  
  
caltss <- calc\_tss(actual = test.set$capped\_Number,  
 predicted = predict(lasso\_Mod, test.set))  
  
rsq\_lasso\_Mod <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$capped\_Number,  
 predicted = predict(knn\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$capped\_Number,  
 predicted = predict(knn\_model, test.set))  
  
rsq\_knn\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$capped\_Number,  
 predicted = predict(LF\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$capped\_Number,  
 predicted = predict(LF\_model, test.set))  
  
rsq\_LF\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$capped\_Number,  
 predicted = predict(LB\_model, test.set))  
  
caltss <- calc\_tss(actual = test.set$capped\_Number,  
 predicted = predict(LB\_model, test.set))  
  
rsq\_LB\_model <- 1 - calrss/caltss  
#######################################################  
# rss calculation  
calrss <- calc\_rss(actual = test.set$capped\_Number,  
 predicted = predict(lmStepAIC\_Mod, test.set))  
  
caltss <- calc\_tss(actual = test.set$capped\_Number,  
 predicted = predict(lmStepAIC\_Mod, test.set))  
  
rsq\_lmStepAIC\_Mod <- 1 - calrss/caltss  
#######################################################  
  
total\_rsq <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 R\_squared\_Pred=c(rsq\_lm\_model, rsq\_glm\_model, rsq\_lasso\_Mod, rsq\_knn\_model, rsq\_LF\_model, rsq\_LB\_model, rsq\_lmStepAIC\_Mod))  
total\_rsq

## Name R\_squared\_Pred  
## 1 lm\_model 0.4139288  
## 2 glm\_model 0.4139288  
## 3 lasso\_Mod 0.4126898  
## 4 knn\_model 0.5549650  
## 5 LF\_model 0.3400263  
## 6 LB\_model 0.3400263  
## 7 lmStepAIC\_Mod 0.4139288

## OUTLIERS CAPPED

## Combine predicted RMSE, MAE R squared, time

total\_combALLpred <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 Model\_TimePred,  
 total\_RMSE,  
 total\_MAE,  
 total\_rsq) %>%   
 dplyr::select(-2,-4,-6,-8)  
total\_combALLpred

## Name Predict\_Time RMSE\_Pred MAE\_Pred R\_squared\_Pred  
## 1 lm\_model 0.10870695 secs 28.21732 21.65778 0.4139288  
## 2 glm\_model 0.12167406 secs 28.21732 21.65778 0.4139288  
## 3 lasso\_Mod 0.09374905 secs 28.24713 21.49014 0.4126898  
## 4 knn\_model 0.60438299 secs 24.58882 13.63390 0.5549650  
## 5 LF\_model 0.09973288 secs 29.94359 22.06808 0.3400263  
## 6 LB\_model 0.10173392 secs 29.94359 22.06808 0.3400263  
## 7 lmStepAIC\_Mod 0.11070514 secs 28.21732 21.65778 0.4139288

## OUTLIERS CAPPED

## Compare the two sets of RMSE, MAE, r squared, time

total\_setscomp <- data.frame(Name=c("lm\_model","glm\_model","lasso\_Mod", "knn\_model","LF\_model","LB\_model","lmStepAIC\_Mod"),  
 total\_combALLpred,  
 total\_best\_train) %>%   
 dplyr::select(-2,-7)  
# Reorder columns  
total\_OUTLIERS\_CAPPED <- total\_setscomp[,c(1,6,3,8,4,7,5,9,2)]  
  
# Change name of column in data.frame  
total\_OUTLIERS\_CAPPED %>%   
 rename(  
 RMSE\_Train = RMSE,  
 MAE\_Train = MAE,  
 Rsquared\_Train = Rsquared  
 )

## Name RMSE\_Train RMSE\_Pred MAE\_Train MAE\_Pred Rsquared\_Train  
## 1 lm\_model 28.59885 28.21732 21.82526 21.65778 0.3925710  
## 2 glm\_model 28.59885 28.21732 21.82526 21.65778 0.3925710  
## 3 lasso\_Mod 28.61676 28.24713 21.67310 21.49014 0.3919238  
## 4 knn\_model 26.90877 24.58882 17.17772 13.63390 0.4714403  
## 5 LF\_model 29.93894 29.94359 21.93797 22.06808 0.3342066  
## 6 LB\_model 29.93894 29.94359 21.93797 22.06808 0.3342066  
## 7 lmStepAIC\_Mod 28.60671 28.21732 21.82608 21.65778 0.3922352  
## R\_squared\_Pred Time\_train Predict\_Time  
## 1 0.4139288 1.643645 secs 0.10870695 secs  
## 2 0.4139288 1.738344 secs 0.12167406 secs  
## 3 0.4126898 2.068453 secs 0.09374905 secs  
## 4 0.5549650 25.646455 secs 0.60438299 secs  
## 5 0.3400263 1.012250 secs 0.09973288 secs  
## 6 0.3400263 1.079079 secs 0.10173392 secs  
## 7 0.4139288 4.663523 secs 0.11070514 secs

## OUTLIERS CAPPED

## RMSE, Rsquared MAE on training set

# RMSE, Rsquared MAE on training set  
lm\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 28.6 0.39 21.83

# RMSE, Rsquared MAE on training set  
glm\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 28.6 0.39 21.83

# RMSE, Rsquared MAE on training set  
lasso\_Mod$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 34.61 0.29 26.51  
## 2 29.82 0.36 21.43  
## 3 28.62 0.39 21.67

# RMSE, Rsquared MAE on training set  
knn\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 26.91 0.47 17.18  
## 2 27.28 0.47 18.29  
## 3 27.50 0.46 18.64

# RMSE, Rsquared MAE on training set  
LF\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 31.62 0.26 21.24  
## 2 30.51 0.31 19.88  
## 3 29.94 0.33 21.94

# RMSE, Rsquared MAE on training set  
LB\_model$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 31.62 0.26 21.24  
## 2 30.51 0.31 19.88  
## 3 29.94 0.33 21.94

# RMSE, Rsquared MAE on training set  
lmStepAIC\_Mod$results[c("RMSE","Rsquared","MAE")] %>%  
 round(2)

## RMSE Rsquared MAE  
## 1 28.61 0.39 21.83