Exploratory analysis of Global-Terrorism

Naveen

08/02/2021

## Including Plots

You can also embed plots, for example:



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.

#Loading the required packages for analysis  
library(readr)  
library(ggplot2)  
library(dplyr)

##   
## 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

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v tibble 3.0.4 v stringr 1.4.0  
## v tidyr 1.1.2 v forcats 0.5.0  
## v purrr 0.3.4

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(purrr)

#Reading the dataset using builtin function from readr library  
g\_terror <- read\_csv("Globalterrorism.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## .default = col\_double(),  
## approxdate = col\_character(),  
## resolution = col\_character(),  
## country\_txt = col\_character(),  
## region\_txt = col\_character(),  
## provstate = col\_character(),  
## city = col\_character(),  
## location = col\_character(),  
## summary = col\_character(),  
## alternative\_txt = col\_character(),  
## attacktype1\_txt = col\_character(),  
## attacktype2\_txt = col\_character(),  
## attacktype3 = col\_logical(),  
## attacktype3\_txt = col\_logical(),  
## targtype1\_txt = col\_character(),  
## targsubtype1\_txt = col\_character(),  
## corp1 = col\_character(),  
## target1 = col\_character(),  
## natlty1\_txt = col\_character(),  
## targtype2\_txt = col\_character(),  
## targsubtype2\_txt = col\_character()  
## # ... with 48 more columns  
## )  
## i Use `spec()` for the full column specifications.

## Warning: 7113 parsing failures.  
## row col expected actual file  
## 1687 ransomamtus 1/0/T/F/TRUE/FALSE 20000 'Globalterrorism.csv'  
## 1687 ransomnote 1/0/T/F/TRUE/FALSE A note, with a sticker of the Black Liberation Army, demanded $20,000 for the safe return of the bartender. However, the perpetrators left the bartender alone and he was able to escape. 'Globalterrorism.csv'  
## 1771 attacktype3 1/0/T/F/TRUE/FALSE 2 'Globalterrorism.csv'  
## 1771 attacktype3\_txt 1/0/T/F/TRUE/FALSE Armed Assault 'Globalterrorism.csv'  
## 3432 ransomnote 1/0/T/F/TRUE/FALSE Perpetrators demanded support by the Dutch in the UN for independence for the Moluccan Islands. 'Globalterrorism.csv'  
## .... ............... .................. .......................................................................................................................................................................................... .....................  
## See problems(...) for more details.

head(g\_terror)

## # A tibble: 6 x 135  
## eventid iyear imonth iday approxdate extended resolution country country\_txt  
## <dbl> <dbl> <dbl> <dbl> <chr> <dbl> <chr> <dbl> <chr>   
## 1 1.97e11 1970 7 2 <NA> 0 <NA> 58 Dominican ~  
## 2 1.97e11 1970 0 0 <NA> 0 <NA> 130 Mexico   
## 3 1.97e11 1970 1 0 <NA> 0 <NA> 160 Philippines  
## 4 1.97e11 1970 1 0 <NA> 0 <NA> 78 Greece   
## 5 1.97e11 1970 1 0 <NA> 0 <NA> 101 Japan   
## 6 1.97e11 1970 1 1 <NA> 0 <NA> 217 United Sta~  
## # ... with 126 more variables: region <dbl>, region\_txt <chr>, provstate <chr>,  
## # city <chr>, latitude <dbl>, longitude <dbl>, specificity <dbl>,  
## # vicinity <dbl>, location <chr>, summary <chr>, crit1 <dbl>, crit2 <dbl>,  
## # crit3 <dbl>, doubtterr <dbl>, alternative <dbl>, alternative\_txt <chr>,  
## # multiple <dbl>, success <dbl>, suicide <dbl>, attacktype1 <dbl>,  
## # attacktype1\_txt <chr>, attacktype2 <dbl>, attacktype2\_txt <chr>,  
## # attacktype3 <lgl>, attacktype3\_txt <lgl>, targtype1 <dbl>,  
## # targtype1\_txt <chr>, targsubtype1 <dbl>, targsubtype1\_txt <chr>,  
## # corp1 <chr>, target1 <chr>, natlty1 <dbl>, natlty1\_txt <chr>,  
## # targtype2 <dbl>, targtype2\_txt <chr>, targsubtype2 <dbl>,  
## # targsubtype2\_txt <chr>, corp2 <chr>, target2 <chr>, natlty2 <dbl>,  
## # natlty2\_txt <chr>, targtype3 <dbl>, targtype3\_txt <chr>,  
## # targsubtype3 <dbl>, targsubtype3\_txt <chr>, corp3 <chr>, target3 <chr>,  
## # natlty3 <dbl>, natlty3\_txt <chr>, gname <chr>, gsubname <chr>,  
## # gname2 <chr>, gsubname2 <lgl>, gname3 <lgl>, gsubname3 <lgl>, motive <chr>,  
## # guncertain1 <dbl>, guncertain2 <dbl>, guncertain3 <lgl>, individual <dbl>,  
## # nperps <dbl>, nperpcap <dbl>, claimed <dbl>, claimmode <dbl>,  
## # claimmode\_txt <chr>, claim2 <dbl>, claimmode2 <lgl>, claimmode2\_txt <lgl>,  
## # claim3 <lgl>, claimmode3 <lgl>, claimmode3\_txt <lgl>, compclaim <lgl>,  
## # weaptype1 <dbl>, weaptype1\_txt <chr>, weapsubtype1 <dbl>,  
## # weapsubtype1\_txt <chr>, weaptype2 <dbl>, weaptype2\_txt <chr>,  
## # weapsubtype2 <dbl>, weapsubtype2\_txt <chr>, weaptype3 <dbl>,  
## # weaptype3\_txt <chr>, weapsubtype3 <dbl>, weapsubtype3\_txt <chr>,  
## # weaptype4 <lgl>, weaptype4\_txt <lgl>, weapsubtype4 <lgl>,  
## # weapsubtype4\_txt <lgl>, weapdetail <chr>, nkill <dbl>, nkillus <dbl>,  
## # nkillter <dbl>, nwound <dbl>, nwoundus <dbl>, nwoundte <dbl>,  
## # property <dbl>, propextent <dbl>, propextent\_txt <chr>, propvalue <dbl>,  
## # propcomment <chr>, ...

#As you can see, there is a ton of information in the dataset and it's fairly ugly as well. We want to fix that. Right now, there are plenty of columns that are completely useless to us, such as the 'extended' and 'resolution' attributes, so we're going to get rid of them by using the select function to select the columns that I actually want from the original dataset.

#Considering the columns that are required for the analysis  
gta <- g\_terror %>% select(eventid, iyear, country, region\_txt, attacktype1\_txt, success, suicide, targtype1\_txt)  
  
#Changing the column names to a readable and understandable  
names(gta) <- c("Event\_ID", "Year", "Country", "Region", "Attack\_type", "Success", "Suicide", "Target\_type")  
gta

## # A tibble: 181,691 x 8  
## Event\_ID Year Country Region Attack\_type Success Suicide Target\_type   
## <dbl> <dbl> <dbl> <chr> <chr> <dbl> <dbl> <chr>   
## 1 1.97e11 1970 58 Central ~ Assassination 1 0 Private Citi~  
## 2 1.97e11 1970 130 North Am~ Hostage Taki~ 1 0 Government (~  
## 3 1.97e11 1970 160 Southeas~ Assassination 1 0 Journalists ~  
## 4 1.97e11 1970 78 Western ~ Bombing/Expl~ 1 0 Government (~  
## 5 1.97e11 1970 101 East Asia Facility/Inf~ 1 0 Government (~  
## 6 1.97e11 1970 217 North Am~ Armed Assault 1 0 Police   
## 7 1.97e11 1970 218 South Am~ Assassination 0 0 Police   
## 8 1.97e11 1970 217 North Am~ Bombing/Expl~ 1 0 Utilities   
## 9 1.97e11 1970 217 North Am~ Facility/Inf~ 1 0 Military   
## 10 1.97e11 1970 217 North Am~ Facility/Inf~ 1 0 Government (~  
## # ... with 181,681 more rows

#Checking if there are any NA values  
colSums(is.na(gta))

## Event\_ID Year Country Region Attack\_type Success   
## 0 0 0 0 0 0   
## Suicide Target\_type   
## 0 0

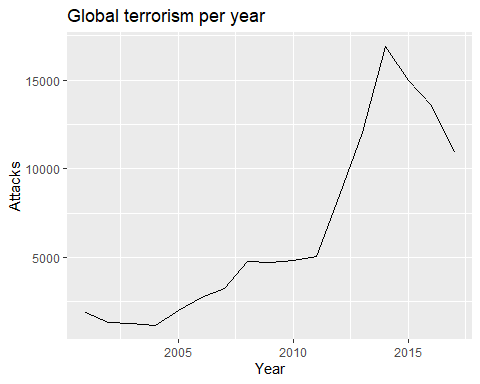
#Condensing the rows by filter the Year from 2000  
gta <- gta %>% filter(Year > 2000)  
gta

## # A tibble: 110,041 x 8  
## Event\_ID Year Country Region Attack\_type Success Suicide Target\_type   
## <dbl> <dbl> <dbl> <chr> <chr> <dbl> <dbl> <chr>   
## 1 2.00e11 2001 45 South Am~ Unknown 1 0 Private Citi~  
## 2 2.00e11 2001 228 Middle E~ Bombing/Expl~ 1 0 Religious Fi~  
## 3 2.00e11 2001 45 South Am~ Hostage Taki~ 1 0 Private Citi~  
## 4 2.00e11 2001 209 Middle E~ Bombing/Expl~ 1 0 Private Citi~  
## 5 2.00e11 2001 6 Middle E~ Unknown 1 0 Private Citi~  
## 6 2.00e11 2001 228 Middle E~ Bombing/Expl~ 1 0 Journalists ~  
## 7 2.00e11 2001 235 Eastern ~ Bombing/Expl~ 1 0 Police   
## 8 2.00e11 2001 217 North Am~ Facility/Inf~ 1 0 Business   
## 9 2.00e11 2001 209 Middle E~ Bombing/Expl~ 1 1 Police   
## 10 2.00e11 2001 199 Western ~ Bombing/Expl~ 1 0 Airports & A~  
## # ... with 110,031 more rows

#Terrorist attack over different year  
terror\_overtime <- gta %>% group\_by(Year) %>%   
 summarise(Attacks = n())

## `summarise()` ungrouping output (override with `.groups` argument)

ggplot(data = terror\_overtime,   
 aes(x = Year , y = Attacks)) +   
 geom\_line() +   
 labs(title="Global terrorism per year")

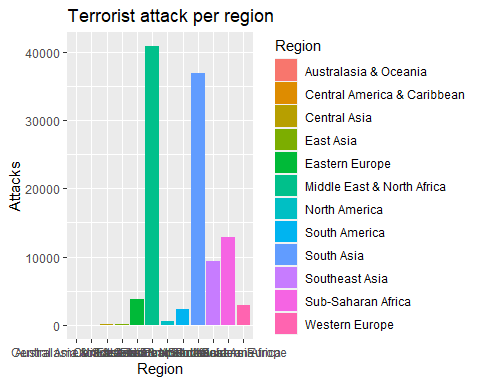


##Looking at this plot we can see that there's a huge uptick of terrorism in the most recent years, which, with the rise of ISIS and general instability in the world recently, is not a huge surprise.

#Terrorist attack on different regions  
Gta\_per\_region <- gta %>%   
 group\_by(Region) %>%   
 summarise(Attacks = n())

## `summarise()` ungrouping output (override with `.groups` argument)

ggplot(data = Gta\_per\_region,   
 aes(x = Region , y = Attacks, fill = Region)) +   
 geom\_bar(stat = "identity") +   
 labs(title="Terrorist attack per region", x = "Region", y = "Attacks")

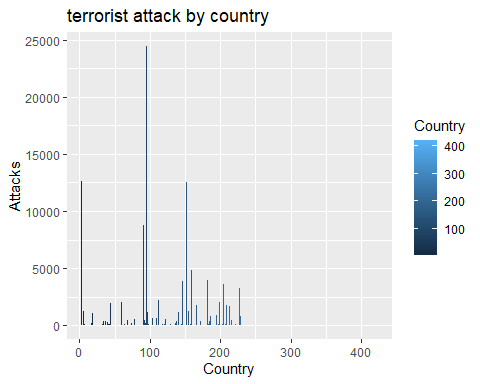


##With this plot we can see that the regions of the world with the highest amount of terrorism in the world are the Middle East + North Africa and South Asia, with Sub-Suharan Africa coming in a far third place. Generally speaking, these areas of the world are significantly worse off economically than the other regions of the world shown on this plot, like North America or Western Europe, which are represented far less in this visualization.

#Terrorist attack by country code   
t\_by\_cfilt <- gta %>%   
 filter(Country < 500)  
t\_by\_country <- t\_by\_cfilt %>%   
 group\_by(Country) %>%   
 summarise(Attacks = n())

## `summarise()` ungrouping output (override with `.groups` argument)

ggplot(data = t\_by\_country,   
 aes(x = Country, y = Attacks, fill = Country)) +   
 geom\_bar(stat = "identity") +   
 labs(title = "terrorist attack by country", x = "Country", y = "Attacks")

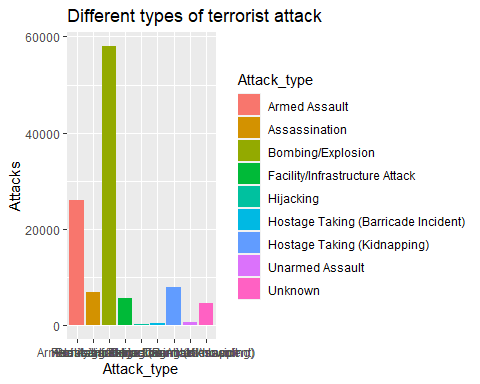


#We can see more number of attacks between the country code from 0-250. and also we can see the code 90-100 has max number of attacks.

#terrorist attack different attack types   
ter\_by\_attacktype <- gta %>%   
 group\_by(Attack\_type) %>%   
 summarise(Attacks = n())

## `summarise()` ungrouping output (override with `.groups` argument)

ggplot(data = ter\_by\_attacktype,   
 aes(x = Attack\_type, y = Attacks, fill = Attack\_type)) + geom\_bar(stat = "identity") +   
 labs(title = "Different types of terrorist attack", x = "Attack\_type", y = "Attacks")



#From the plot below we can see that the most of the attacks are done by Bomb/Explosion and Armed Assaults types.

##Conclusion: After exploring the data we can see that there’s a huge uptick of terrorism in the most recent years and also we can see that the regions of the world with the highest amount of terrorism in the world are the Middle East + North Africa and South Asia, with Sub-Suharan Africa coming in a far third place.most of the attacks are done by Bomb/Explosion and Armed Assaults types.