

GTD Group Project

datING - BDA 503 - Fall 2017

07 December 2017

Table of Contents

Group Name	1
Group Members.....	1
Data	1
Description	2
Content and Data fields.....	2
Objectives.....	3
Abstract	3
Exploratory Analysis	3
Data Cleaning.....	4
Wordcloud Analysis	20
Analysis on Map.....	21
Interactive Data Map with Shiny module	22
References	23

Group Name

- **datING**

Group Members

1. Bilge Uzuner
2. Sefa Erbas
3. Türkan Sevgili

Data

- Global Terrorism Data

Description

The Global Terrorism Database ([GTD](#)) includes information on terrorist events around the World (205 countries, more than 33000 cities, from 1970 through 2016 on at least 45 variables for each case, with more recent incidents including information on more than 120 variables

Content and Data fields

Geography: Worldwide

Time period: 1970-2016, except 1993

Unit of analysis: Attack

Variables: >100 variables on location, tactics, perpetrators, targets, and outcomes

Sources: Unclassified media articles

Necessary information about mostly used fields can be found below. You can see the [GTD Codebook](#) for broadly explained definitions on fields, important details on data collection methodology, definitions, and coding schema.

- **eventid:** Incidents from the GTD follow a 12-digit Event ID system.
- **year:** This field contains the year in which the incident occurred.
- **month:** This field contains the number of the month in which the incident occurred.
- **day:** This field contains the numeric day of the month on which the incident occurred.
- **country:** This field identifies the country code country or location where the incident occurred.
- **region:** This field identifies the region code in which the incident occurred.
- **provstate:** This variable records the name (at the time of event) of the 1st order subnational administrative region in which the event occurs.
- **city:** This field contains the name of the city, village, or town in which the incident occurred.
- **latitude:** This field records the latitude (based on WGS1984 standards) of the city in which the event occurred.
- **longitude:** This field records the longitude (based on WGS1984 standards) of the city in which the event occurred.
- **location:** This field is used to specify additional information about the location of the incident.
- **success:** Success of a terrorist strike is defined according to the tangible effects of the attack. Success is not judged in terms of the larger goals of the perpetrators.

Objectives

- The purpose of the analysis is to understand the Turkey's Terrorism Statistics on the basis of the cities, districts, dates, seasons etc.
- Try to define the factors that influence the terror attacks in Turkey.
- A clear understanding of the data to see if we can locate useful insights about where and when terror attacks occur in Turkey.
- To create an opinion on the relatively risky and secure regions based on past experience, by in-depth analysis of the terrorist attacks in Turkey, also to see relatively safe and risky periods during the year.

Abstract

Definition of terrorism: "The threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation."

Terrorism is the largest human-oriented criminal organization that has become a common problem globally. Nowadays, it is known that the underlying reason of most of today's terrorist activities is strategic actions and international purposes. Terrorist attacks in a country can take place because of interior reasons (economic prosperity, socio-cultural, educational system) or other factors. The most important of factor is the geopolitical position of the country.

It is known that Turkey is the target of many terrorist organizations because it is a bridge in the point where the European, Asian and African continents are connected and it is close to the oil resources in the Middle East where there is a continuous and multifaceted conflict of interests and powers that can affect the world power balance. In this work, we will analyze the various details of terrorist acts in Turkey over the past 30 years.

Exploratory Analysis

- Load necessary frequently used R Packages and see whole GTD dataset.

```
# Necessary packages to load in
library(tidyverse)
library(ggplot2) # Data visualization
library(readr) # Read_csv function
library(dplyr)
options(dplyr.width=Inf)
#Sys.setlocale(locale = "Turkish_Turkey.1254")

# Get all data

gtd =
read.csv(file="https://bitbucket.org/sevgilit/bda503/raw/8ef59a06fd743de89955
610b2af9ca4e288bc1c1/globalterrorismdb_0617dist.csv", header=TRUE, sep=",")
```

```
# subset gtd data to attacks in Turkey
turkey.gtd <- gtd%>%
  filter(country_txt == "Turkey")
```

```
#if any problem
problems(turkey.gtd)
```

```
# Summary of turkey.gtd structure
glimpse(turkey.gtd)
```

Data Cleaning

Since there are more than 100 variables, first we selected which variables are meaningful for the analysis for Turkey. Other columns are the sub categories of these informations and not available for all events.

```
# Select columns that will use in EDA and renaming selected columns
gtd.turkey = select (turkey.gtd , eventid, year = iyear, month = imonth,
day=iday, country_code = country,
                    country_name=country_txt,region_code=region,
region_name=region_txt, provstate,city,
                    latitude,longitude,location,success,
attacktype1=attacktype1_txt,
attacktype2=attacktype2_txt,attacktype3=attacktype3_txt,targettype1=targettype1_t
xt,
                    targsubtype1=targsubtype1_txt,weaptype1=weaptype1_txt,
weapsubtype1=weapsubtype1_txt,
                    property,propextent txt,propvalue,gname,nkill,nwound)
```

```
# Summary of gtd.turkey structure
glimpse(gtd.turkey)
```

```
## Observations: 4,022
## Variables: 27
## $ eventid      <dbl> 197004250001, 197008310001, 197010020002, 19701...
## $ year         <int> 1970, 1970, 1970, 1970, 1970, 1970, 1970, 1970,...
## $ month        <int> 4, 8, 10, 10, 10, 10, 10, 11, 11, 11, 12, 12, 1...
## $ day          <int> 25, 31, 2, 3, 3, 6, 27, 10, 21, 23, 24, 29, 7, ...
## $ country_code <int> 209, 209, 209, 209, 209, 209, 209, 209, 209, 209, 20...
## $ country_name <fctr> Turkey, Turkey, Turkey, Turkey, Turkey, Turkey...
## $ region_code  <int> 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10,...
## $ region_name  <fctr> Middle East & North Africa, Middle East & Nort...
## $ provstate    <fctr> Istanbul, Ankara, Izmir, Ankara, Ankara, Ankar...
## $ city         <fctr> Istanbul, Ankara, Izmir, Ankara, Ankara, Ankar...
## $ latitude     <dbl> 41.01484, 39.91839, 38.42371, 39.91839, 39.9183...
## $ longitude    <dbl> 28.96141, 32.86560, 27.13421, 32.86560, 32.8656...
## $ location     <fctr> , , , , , , , , , , , , , , , , , , , , , , , ,
## $ success      <int> 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1,...
## $ attacktype1  <fctr> Bombing/Explosion, Bombing/Explosion, Bombing/...
```

[illegible]

The data set that will be used in this project and initial exploratory data analysis on it consists of 4,106 observations and 27 variables.

Secondly we checked the data for the missing/ NA or unmeaningful data;

- Event id is unique
- 30/31 February is not available
- No day 31 is available for the months 2,4,6,9,11

Country code and Region Code is unique for Turkey and Middle East & North Africa which are 209 and 10 approximately.

In the "City" column we see the district information for some of the data so we made the city analysis on "Provstate" column which we found more convenient.

- Initial exploratory analysis of GTD dataset and Turkey terrorism statistics can be found below.

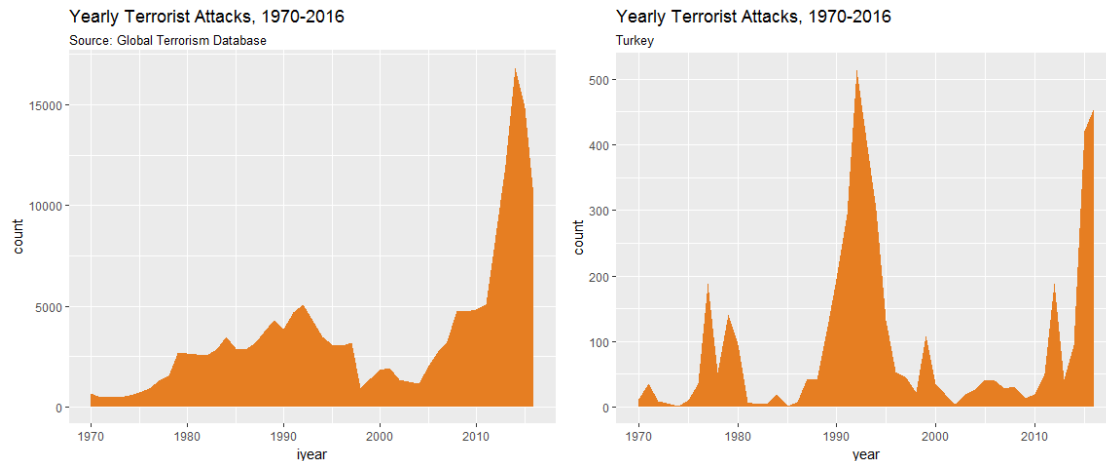
It is firstly examined how many terrorist attacks have been recorded in the years since the Global Terrorist Database. (1970-2016)

Terrorist Attacks, 1970-2016 (Global vs Turkey)

```
library(ggplot2)
library(gridExtra)
plot1<-ggplot(data=gtd) +
  geom_area(mapping=aes(x=iyear), stat="count", fill= ('#e67e22')) +
  ggtitle("Yearly Terrorist Attacks, 1970-2016", subtitle = "Source: Global
Terrorism Database")

plot2<-ggplot(data=gtd.turkey) +
  geom_area(mapping=aes(x=year), stat="count", fill= ('#e67e22')) +
  ggtitle("Yearly Terrorist Attacks, 1970-2016", subtitle = "Turkey")

grid.arrange(plot1,
              plot2,
              ncol=2)
```

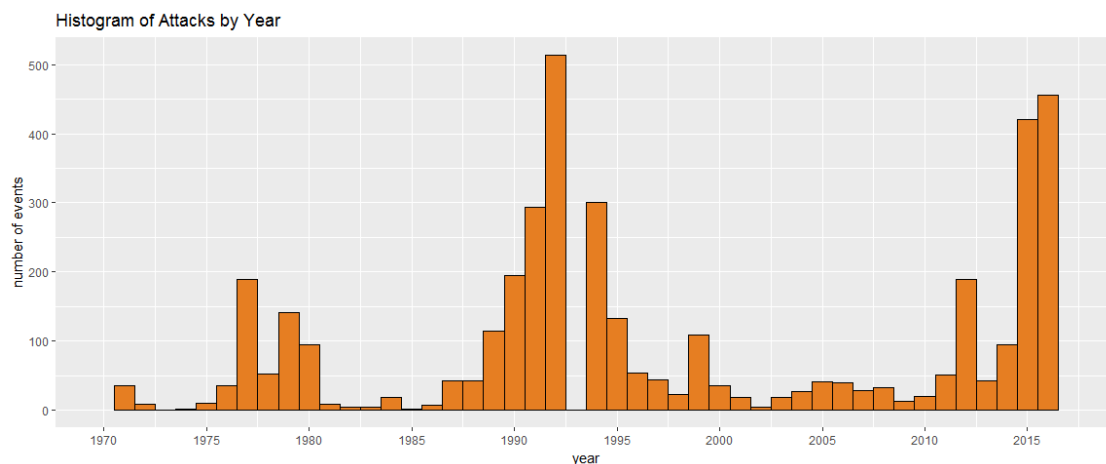


In global side, there is a steady increase over time until the early '90s, after which there is a steady decline until the mid 2000s. It is a dramatic increase (and record yearly high, in 2014) in reported attacks.

In Turkey, the upward trend observed at the beginning of the 90s declined along with the 2000s. It is observed that after 2010, it has entered a new upward trend.

Histogram of Attacks by Year

```
library(ggplot2)
ggplot(gtd.turkey, aes(gtd.turkey$year)) +
  geom_histogram(fill= ('#e67e22'), color='black', binwidth=1) +
  scale_x_continuous(limit=c(1970, 2017), breaks=seq(1970, 2017, by = 5)) +
  labs(x= 'year', y= 'number of events') +
  ggtitle('Histogram of Attacks by Year')
```



As shown above figure, the data for year 1993 is missing in the source data.

Histogram of Attacks by Month

```
# Data Preparation
gtd.turkey.month <- gtd.turkey %>%
  group_by(month) %>%
  summarise(numberOfEvents = n())
```

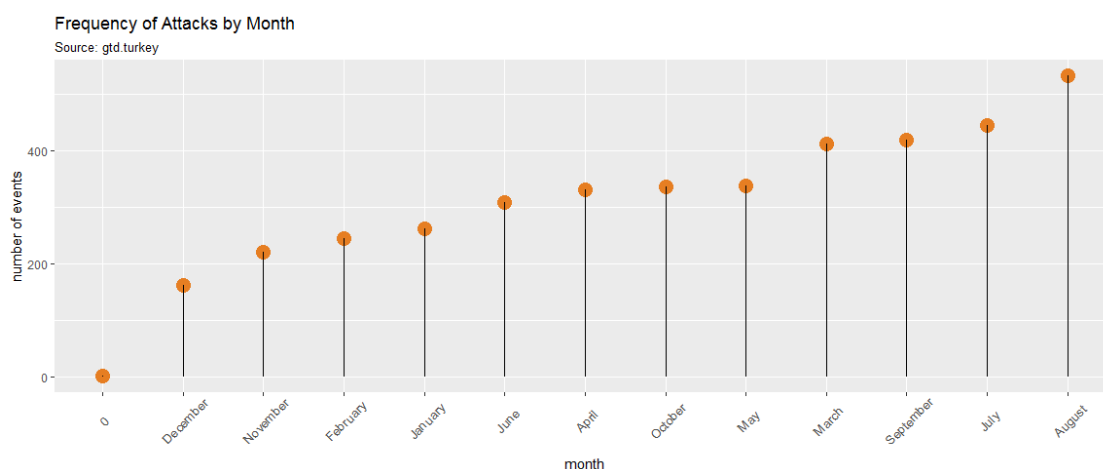
```

gtd.turkey.month$month[gtd.turkey.month$month == 0] <- 0
gtd.turkey.month$month[gtd.turkey.month$month == 1] <- "January"
gtd.turkey.month$month[gtd.turkey.month$month == 2] <- "February"
gtd.turkey.month$month[gtd.turkey.month$month == 3] <- "March"
gtd.turkey.month$month[gtd.turkey.month$month == 4] <- "April"
gtd.turkey.month$month[gtd.turkey.month$month == 5] <- "May"
gtd.turkey.month$month[gtd.turkey.month$month == 6] <- "June"
gtd.turkey.month$month[gtd.turkey.month$month == 7] <- "July"
gtd.turkey.month$month[gtd.turkey.month$month == 8] <- "August"
gtd.turkey.month$month[gtd.turkey.month$month == 9] <- "September"
gtd.turkey.month$month[gtd.turkey.month$month == 10] <- "October"
gtd.turkey.month$month[gtd.turkey.month$month == 11] <- "November"
gtd.turkey.month$month[gtd.turkey.month$month == 12] <- "December"

gtd.turkey.month <- gtd.turkey.month[order(gtd.turkey.month$numberOfEvents),
]
gtd.turkey.month$month <- factor(gtd.turkey.month$month, levels =
gtd.turkey.month$month)

# Plot
ggplot(gtd.turkey.month, aes(x=month, y=numberOfEvents)) +
  geom_point(size=5,color='#e67e22') +
  geom_segment(aes(x=month,
                  xend=month,
                  y=0,
                  yend=numberOfEvents)) +
  labs(x= 'month', y= 'number of events') +
  labs(title="Frequency of Attacks by Month",
        subtitle="Source: gtd.turkey") +
  theme(axis.text.x = element_text(angle=45, vjust=0.6, size=10))

```



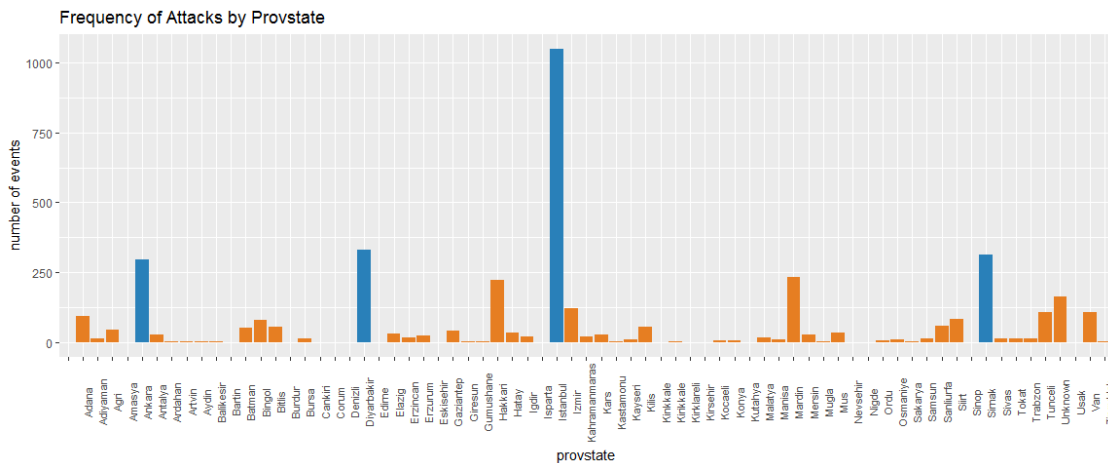
We checked which months are most dangerous. For some of the events the month information is "0". It is because the month information is not known during the creation of

the database. We do not need to manipulate the data and let it stay as zero. Figure shows the day information is missing and seen as zero in the data for some of the events.

When Turkey is busy with foreign visitors and the tourism season peaks, it is seen that the terrorist attacks also rise in summer months.

Frequency of Attacks by City

```
gtd.turkey %>%
  group_by(provstate)%>%
  count() %>% # count the number of times a provstate appear
  ggplot(aes(x=provstate,y=n))+
  geom_bar(stat = "identity",aes(fill=n>250)) +
  theme(axis.text.x = element_text(angle=90,size=8,vjust=1, hjust=0.5)) +
  labs(x= 'provstate', y= 'number of events') +
  ggtitle('Frequency of Attacks by Provstate')+
  scale_fill_manual(values = c('#e67e22', '#2980b9'),guide=FALSE)
```



When you look at the cities most subject to terrorist attacks; Istanbul, Ankara, Diyarbakir and Sirnak attract a great deal of attention.

Istanbul; which is a metropolis of Turkey and draws attention with its crowded population become the target of more than 1,000 terrorist attacks. While some cities face with less terrorist attacks and can be considered relatively safe (Amasya, Bartın, Burdur, Çankırı, Edirne, Isparta, Kırklareli, Kirsehir, etc.)at the same period.

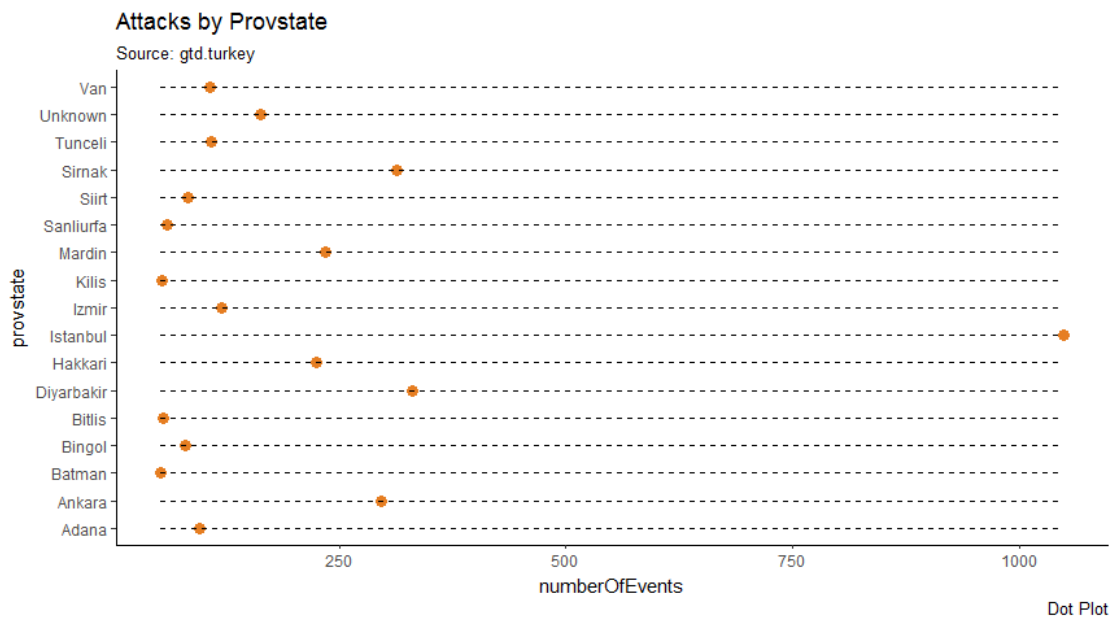
Where are the places that has been attacked > 50 in Turkey between 1970 - 2016?

```
library(scales)
theme_set(theme_classic())

gtd.turkey.provstate_50<-gtd.turkey%>%
  group_by(provstate)%>%
  summarise(numberOfEvents=n())%>%
  filter(numberOfEvents>50)
```

Plot


```
ggplot(gtd.turkey.provstate_50, aes(x=provstate, y=numberOfEvents)) +
  geom_point(col='#e67e22', size=3) + # Draw points
  geom_segment(aes(x=provstate,
                  xend=provstate,
                  y=min(numberOfEvents),
                  yend=max(numberOfEvents)),
              linetype="dashed",
              size=0.1) + # Draw dashed lines
  labs(title="Attacks by Provstate",
        subtitle="Source: gtd.turkey",
        caption="Dot Plot") +
  coord_flip()
```



Success/Failure Rates

```
library(plotrix)
library(ggplot2)
par(mfrow = c(1,2))
```

#Percentages Global&Turkey

```
a<-gtd.turkey %>%
  count(success)
```

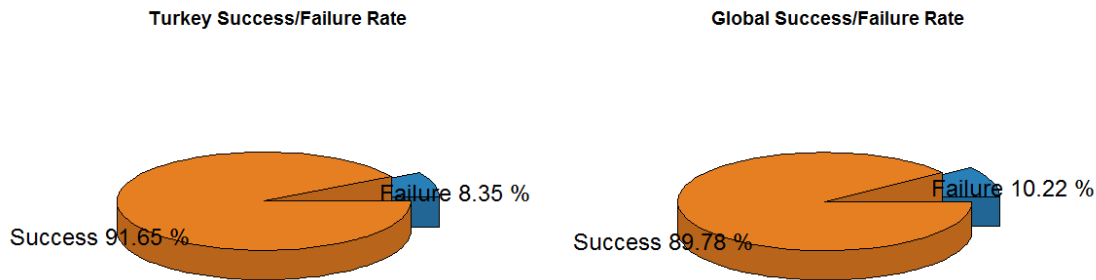
```
b<-gtd %>%
  count(success)
```

#Data Prep and Pie charts

```
data=c(a$n)
pct = (data/sum(data))*100
pct = round(pct,2)
labels = c("Failure", "Success")
labels = paste(labels,pct, "%")
```

```
col = c('#2980b9', '#e67e22')
pie1<-pie3D(pct,col = col, explode=0.1,
  labels =labels, main = "Turkey Success/Failure Rate")

data2=c(b$n)
pct2 = (data2/sum(data2))*100
pct2 = round(pct2,2)
labels2 = c("Failure", "Success")
labels2 = paste(labels2,pct2, "%")
pie2<-pie3D(pct2,col = col, explode=0.1,
  labels =labels2, main = "Global Success/Failure Rate")
```



Success failure rates deemed to be a good measure; about levels of operational risk regarding terrorism risk management, domestic intelligence and reliability of security countermeasures.

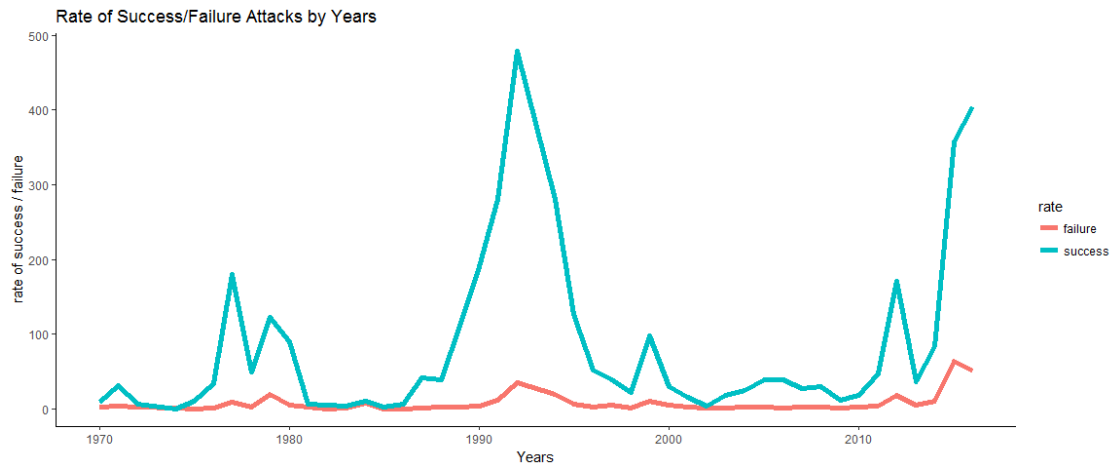
Between 1970 and 2016;the global failure rate is relatively higher than Turkey.

What is the failure rate of attacks over the years?

```
gtd.turkey %>%
  mutate(total = 1) %>% # total attacks that year (creates a new variable)
  count(year, wt=total-success) %>% # failed attempt
  cbind("failure") -> failure
  colnames(failure)[3] <- "rate"

gtd.turkey %>%
  count(year, wt = success) %>%
  cbind("success")-> success
  colnames(success)[3] <- "rate"

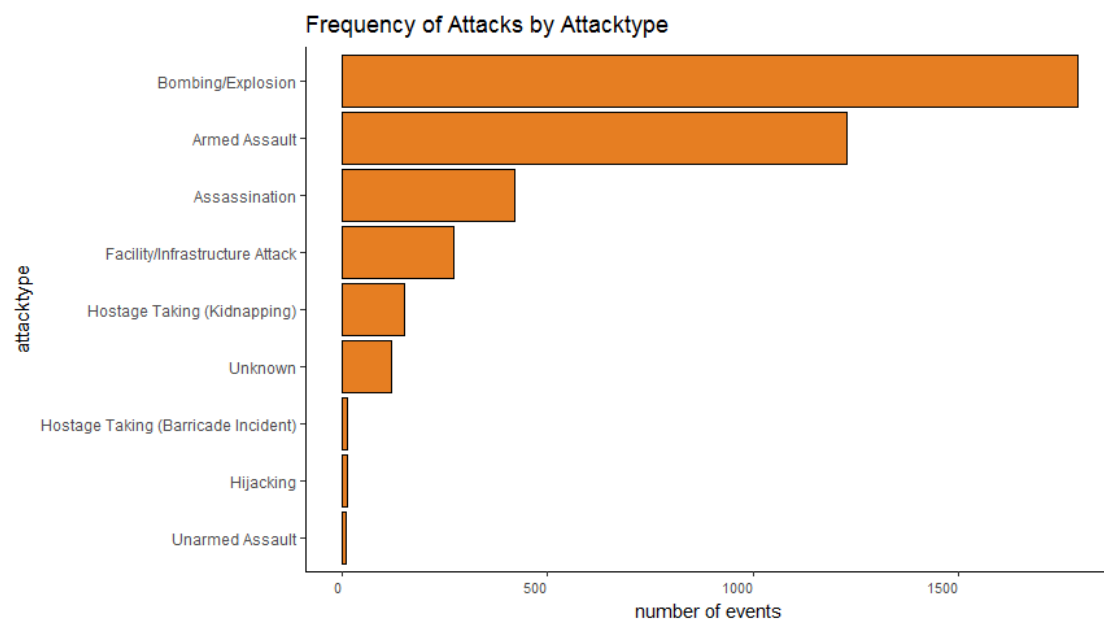
rbind(failure,success) %>%
  ggplot(aes(year,n)) +
  geom_line(aes(group=rate, colour=rate),size = 2)+
  ggtitle("Rate of Success/Failure Attacks by Years") +
  labs(x = "Years", y = "rate of success / failure")
```



It is seen that the ratio which is at a certain level in each period; shows an increasing tendency in recent years.

Frequency of Attacks by Attacktype

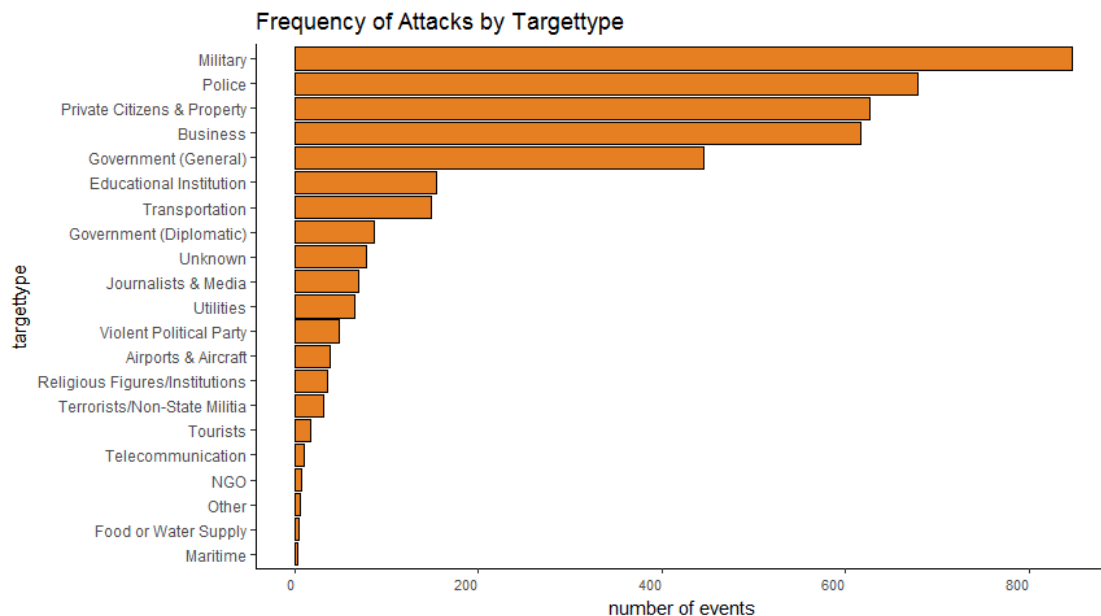
```
gtd.turkey %>%
  group_by(attacktype1)%>%
  count() %>%
  arrange(desc(n)) %>%
  ggplot(aes(x=reorder(attacktype1,n),y=n)) +
  geom_bar(stat = "identity",fill= ('#e67e22'), color='black') +
  theme(axis.text.x = element_text(size=8,vjust=1, hjust=1)) +
  labs(x= 'attacktype', y= 'number of events') +
  ggtitle('Frequency of Attacks by Attacktype') +
  coord_flip()
```



Above figure shows the number of terrorist attacks in the Turkey from 1970 to 2016, by attack type. The most common form of terrorist attack in the Turkey was a **Bombing/Explosion** type attack.

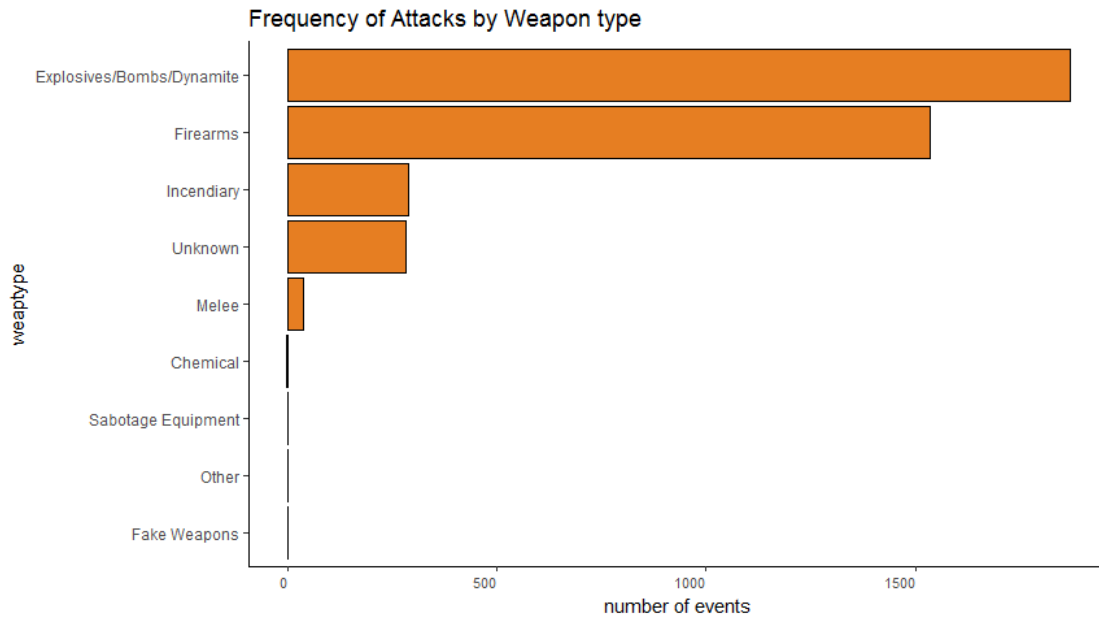
Frequency of Attacks by Target type

```
gtd.turkey %>%
  group_by(targettype1)%>%
  count() %>%
  arrange(desc(n)) %>%
  ggplot(aes(x=reorder(targettype1,n),y=n)) +
  geom_bar(stat = "identity",fill= ('#e67e22'), color='black') +
  theme(axis.text.x = element_text(size=8,vjust=1, hjust=1)) +
  labs(x= 'targettype', y= 'number of events') +
  ggtitle('Frequency of Attacks by Targettype') +
  coord_flip()
```



Frequency of Attacks by Weapon type

```
gtd.turkey %>%
  group_by(weaptype1)%>%
  count() %>%
  arrange(desc(n)) %>%
  ggplot(aes(x=reorder(weaptype1,n),y=n)) +
  geom_bar(stat = "identity",fill= ('#e67e22'), color='black') +
  theme(axis.text.x = element_text(size=8,vjust=1, hjust=1)) +
  labs(x= 'weaptype', y= 'number of events') +
  ggtitle('Frequency of Attacks by Weapon type') +
  coord_flip()
```



Frequency of Group Name

```
gtd.turkey %>%
```

```
  group_by(gname)%>% # group by gname
```

```
  count() %>% # count the number of times a gname appear
```

```
  arrange(desc(n)) %>% # subset by rows based on condition
```

```
  head(n=10) %>%
```

```
  ggplot(aes(x=reorder(gname,n),y=n))+
```

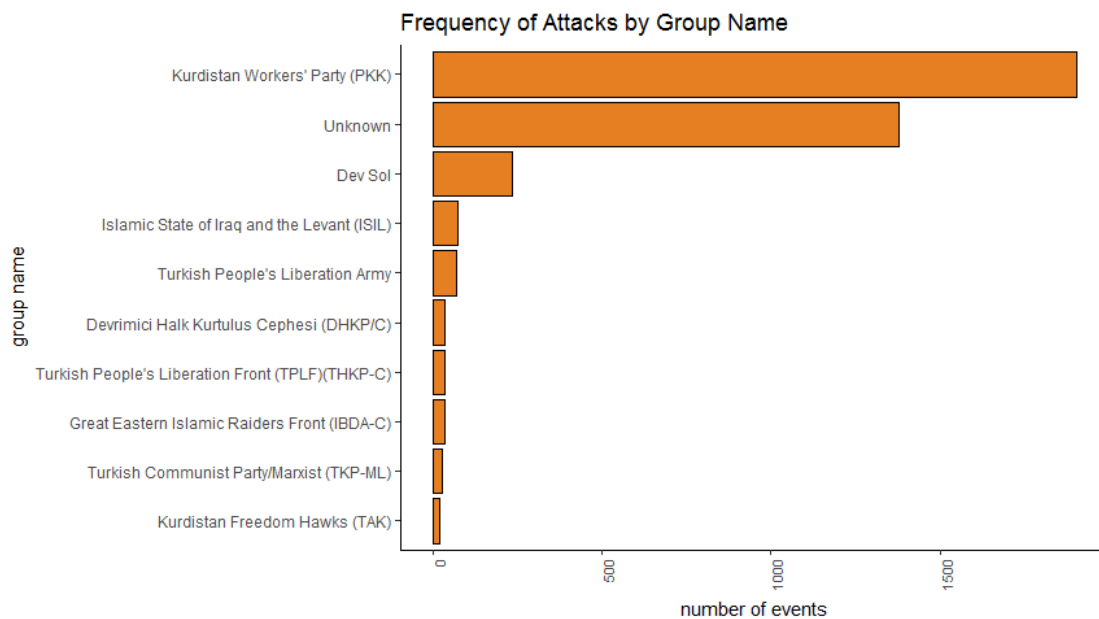
```
  geom_bar(stat = "identity",fill= ('#e67e22'), color='black') +
```

```
  theme(axis.text.x = element_text(angle=90,size=8,vjust=1, hjust=1)) +
```

```
  labs(x= 'group name', y= 'number of events') +
```

```
  ggtitle('Frequency of Attacks by Group Name') +
```

```
  coord_flip()
```



Change in Number of Attacks by Terrorist Groups in last two years

```
library(reshape)
target <- c("2015", "2016")
year_change<- gtd.turkey%>%
  filter(year %in% target) %>%
  group_by(gname, year)%>%
  summarise(number_of_attacks=n())
year_change <- year_change[-c(1, 4, 5, 12,13,14,15,16,17), ]
df<-cast(year_change, gname ~ year)

# prep data
colnames(year_change) <- c("gname", "2015", "2016")
left_label <- paste(df$continent, (df$gname),round(df$`2015`))
right_label <- paste(df$continent, (df$gname),round(df$`2016`))
df$class <- ifelse((df$`2016` - df$`2015`) < 0, "blue", "orange")

# Plot
plot <- ggplot(df) + geom_segment(aes(x=1, xend=2, y=`2015`, yend=`2016`,
col=class), size=1.5, show.legend=F) +
  geom_vline(xintercept=1, linetype="dashed", size=.1) +
  geom_vline(xintercept=2, linetype="dashed", size=.1) +
  scale_color_manual(labels = c("Up", "Down"),
    values = c("orange"="#f39c12", "blue"="#2980b9")) + #
  color of lines
  labs(x="", y="Attacks in 2015 & 2016 by Group") + # Axis Labels
  xlim(.5, 2.5) + ylim(0,(1.1*(max(df$`2015`, df$`2016`)))) # X and Y axis
  limits

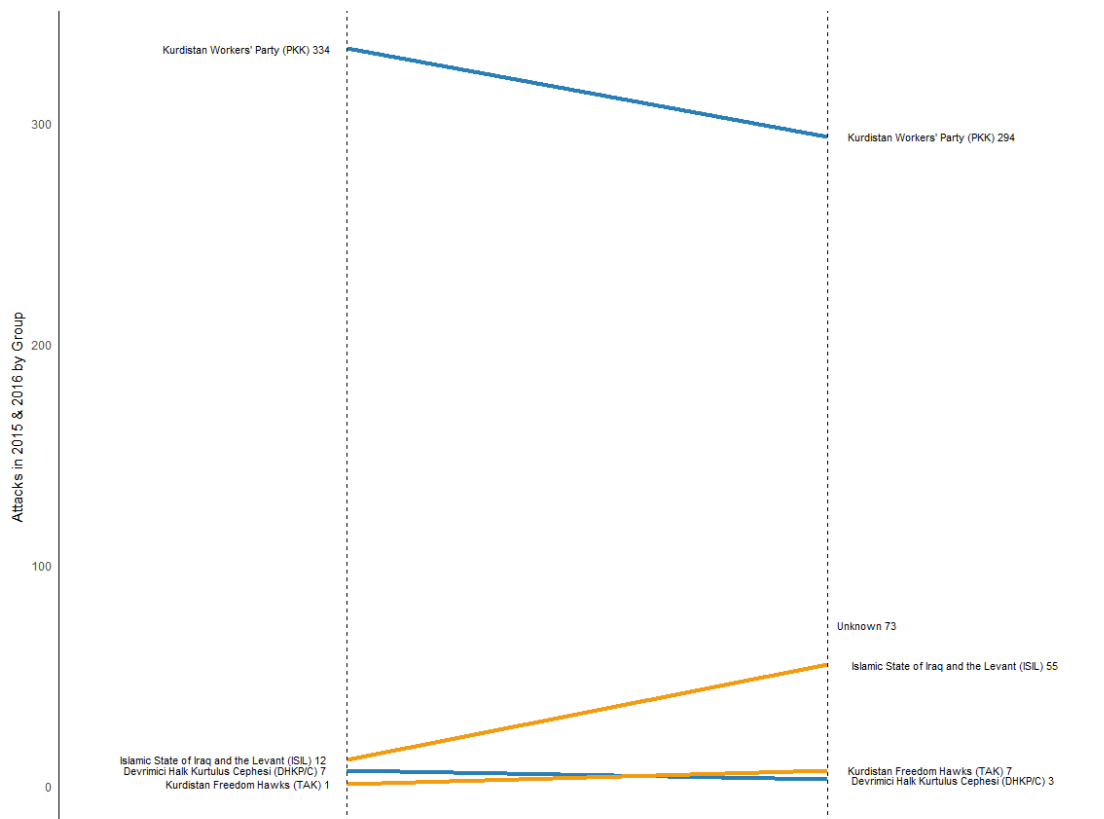
# Add texts
plot <- plot + geom_text(label=left_label, y=df$`2015`, x=rep(1, NROW(df)),
hjust=1.1, size=3)
plot <- plot + geom_text(label=right_label, y=df$`2016`, x=rep(2, NROW(df)),
hjust=-0.1, size=3)
plot <- plot + geom_text(label="2015", x=1, y=1.1*(max(df$`2015`,
df$`2016`)), hjust=1.2, size=5) # title
plot <- plot + geom_text(label="2016", x=2, y=1.1*(max(df$`2015`,
df$`2016`)), hjust=-0.1, size=5) # title

# Minify theme
plot + theme(panel.background = element_blank(),
  panel.grid = element_blank(),
  axis.ticks = element_blank(),
  axis.text.x = element_blank(),
  panel.border = element_blank(),
  plot.margin = unit(c(0.1,0.2,0.1,0.2), "cm"))

## Warning: Removed 1 rows containing missing values (geom_segment).
## Warning: Removed 1 rows containing missing values (geom_text).
```

```
## Warning: Removed 5 rows containing missing values (geom_text).
```

```
## Warning: Removed 5 rows containing missing values (geom_text).
```



This graphic shows the number of attacks of the terrorist groups which have attacked in last two years (2015 -2016) comparing the previous year. When we ignore the unknown attacks, it seems that 3 of 4 groups have increased the total number attacks.

Attacks vs Casualties

```
# Summary by year respect to event and casualties (Pivot)
```

```
gtd.turkey.year= group_by(gtd.turkey, year) %>%
```

```
  summarise(numberOfEvents = length(eventid), numberOfCasualties = sum(nkill,  
na.rm = TRUE))
```

```
ggplot(gtd.turkey.year, aes(x = year)) +
```

```
  geom_line(aes(y = numberOfEvents), size = 2, colour = '#e67e22') +
```

```
  geom_line(aes(y = numberOfCasualties), size = 2, ,colour = ('#2980b9'),
```

```
alpha=0.5) +
```

```
  scale_x_continuous(breaks=seq(1970,2017,1)) +
```

```
  annotate("text", x = c(2010,2010), y = c(1200,1100),
```

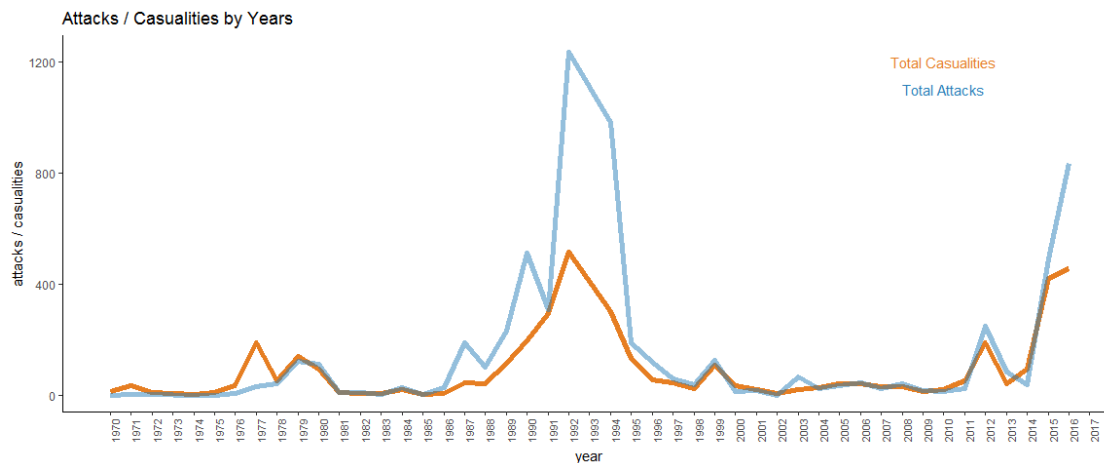
```
    label = c("Total Casualties", "Total Attacks"), colour =
```

```
c(('#e67e22'), ('#2980b9')), size = 4) +
```

```
  ggtitle("Attacks / Casualties by Years") +
```

```
  theme(axis.text.x = element_text(angle=90, size=8)) +
```

```
  labs(x = "year", y = "attacks / casualties")
```



This graph compares the attack number with total casualties per year. It is understood that up to 450 attacks, total casualties have occurred with same level with total attacks. However, even total attacks increase up to 800, total casualties are not influenced majorly from the attacks number.

Killed vs Wounded Rate

gtd.turkey %>%

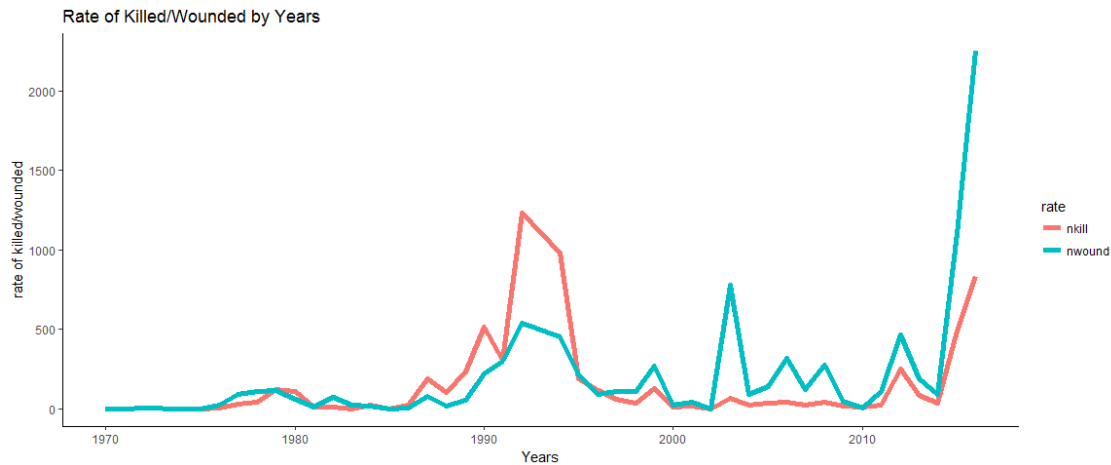
```
mutate(total = 1) %>% # total attacks that year (creates a new variable)
count(year, wt=nkill) %>% # failed attempt
cbind("nkill") ->killed
colnames(killed)[3] <- "rate"
```

gtd.turkey %>%

```
count(year, wt = nwound) %>%
cbind("nwound")-> wounded
colnames(wounded)[3] <- "rate"
```

rbind(killed,wounded) %>%

```
ggplot(aes(year,n)) +
geom_line(aes(group=rate, colour=rate),size = 2)+
ggtitle("Rate of Killed/Wounded by Years") +
labs(x = "Years", y = "rate of killed/wounded")
```

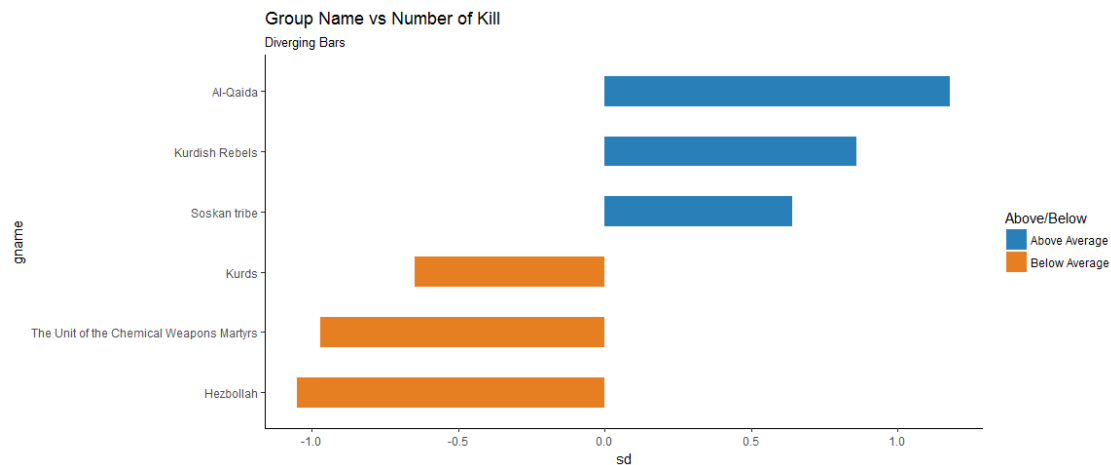
This graphic summarizes rate of death and wounded by years. It is obviously seen in 90's that death rate is much more than wounded rate for each attack. However, in recent years this situation has reversed and wounded rate by virtue of the attacks has dramatically started to increase. This case can be explained in two possibilities. One of them is, increasing the total attacks which may cause more people, the other is lower incidence of these attacks.

gname vs nkill

```
gtd.turkey$ncill[is.na(gtd.turkey$ncill)] <- 0
gnamevncill<-gtd.turkey%>%
  group_by(gname)%>%
  summarise(avgkillperattack= mean(ncill)) %>%
  filter(avgkillperattack > 5)

# Data Prep
gnamevncill$sd <- round((gnamevncill$avgkillperattack -
mean(gnamevncill$avgkillperattack))/sd(gnamevncill$avgkillperattack), 2) #
compute normalized mpg
gnamevncill$mpg_type <- ifelse(gnamevncill$sd < 0, "below", "above") # above
/ below avg flag
gnamevncill <- gnamevncill[order(gnamevncill$sd), ] # sort
gnamevncill$gname <- factor(gnamevncill$gname, levels = gnamevncill$gname) #
convert to factor to retain sorted order in plot.

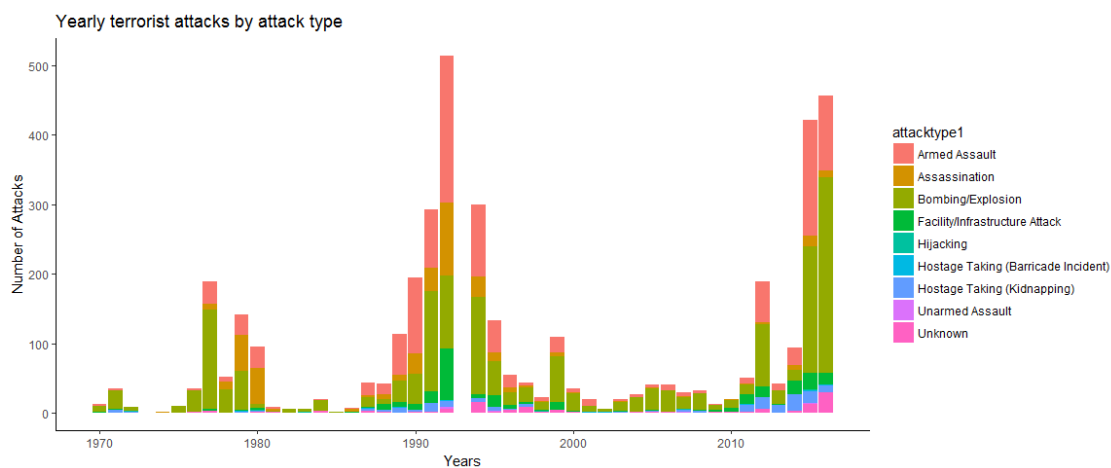
# Diverging Barcharts
ggplot(gnamevncill, aes(x=gname, y=sd, label=sd)) +
  geom_bar(stat='identity', aes(fill=mpg_type), width=.5) +
  scale_fill_manual(name="Above/Below", labels = c("Above Average", "Below
Average"),
                    values = c("above"="#2980b9", "below"="#e67e22")) +
  labs(subtitle="Diverging Bars",
        title= "Group Name vs Number of Kill") +
  coord_flip()
```



This graph compares each group which have caused more than 5 deaths in average due to their attacks with the average of the means the total group's. It is used standard normal distribution formula to see which terrorist group's death cause are in place above and below in terms of population mean.

Yearly terrorist attacks by Attack type

```
ggplot(data=gtd.turkey, aes(x=year, fill=attacktype1)) +
  geom_bar() + ggtitle("Yearly terrorist attacks by attack type") +
  labs(x = "Years", y = "Number of Attacks")
```

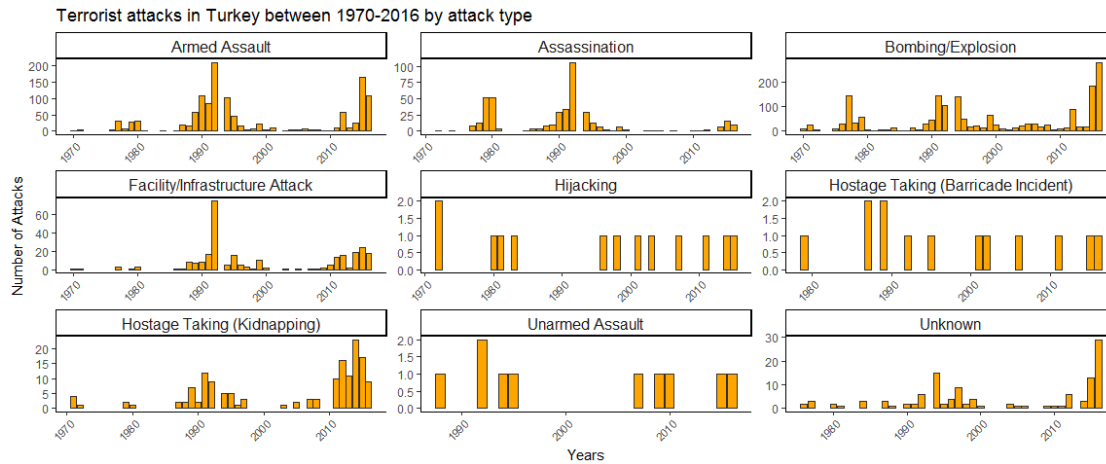


This graphics shows, which attack type is used by terrorist groups to organize an attack by years. It is seen that the most popular attack types are armed assault and bombing almost each year.

Terrorist attacks in Turkey between 1970-2016 by attack type

```
ggplot(gtd.turkey, aes(x = year)) + labs(title = "Terrorist attacks in Turkey
between 1970-2016 by attack type", x = "Years", y = "Number of Attacks", size
= 15) +
  geom_bar(colour = "grey19", fill = "orange") +
  facet_wrap(~attacktype1, scales = "free", ncol = 3) +
```

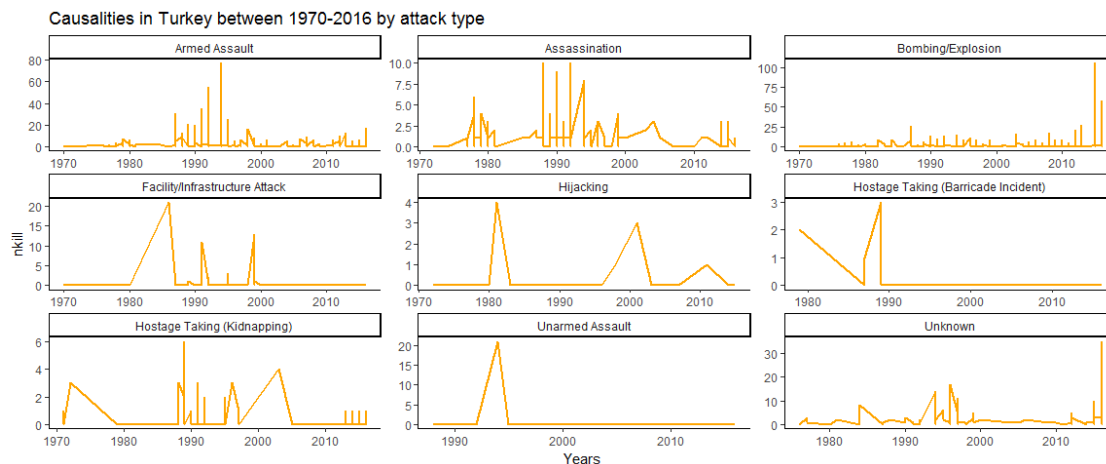
```
theme(axis.text.x = element_text(angle=45,hjust = 1, size = 8)) +
theme(strip.text = element_text(size = 12))
```



##Inspiration:"<https://www.kaggle.com/pranav84/visualizing-terrorist-attacks-on-india>"

Causalities in Turkey between 1970-2016 by attack type

```
ggplot(gtd.turkey, aes(x = year,y=nkill))+
  labs(title = "Terrorist attacks in Turkey by attack type", x = "Years", y =
"Number of Attacks", size = 15) +
  geom_line(colour = "orange",size = 1) +
  facet_wrap(~attacktype1,scales = "free", ncol = 3) +
  ggtitle("Causalities in Turkey between 1970-2016 by attack type")+
  labs(x = "Years", y = "nkill")
```

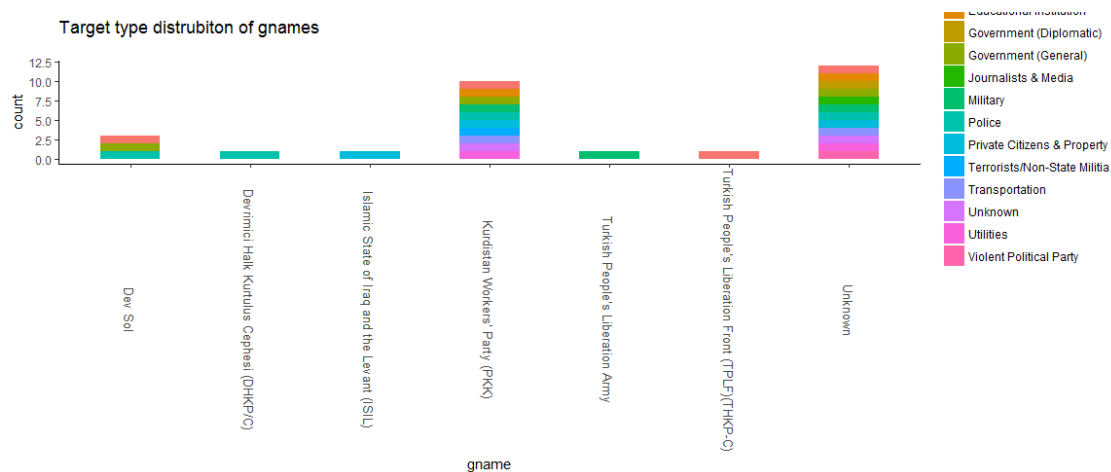


gname vs targtype1

```
gnamevstargtype1<-gtd.turkey%>%
  group_by(gname, targtype1)%>%
  summarise(numberOfEvents=n())%>%
  filter(numberOfEvents>15)
```

Histogram on a Categorical variable

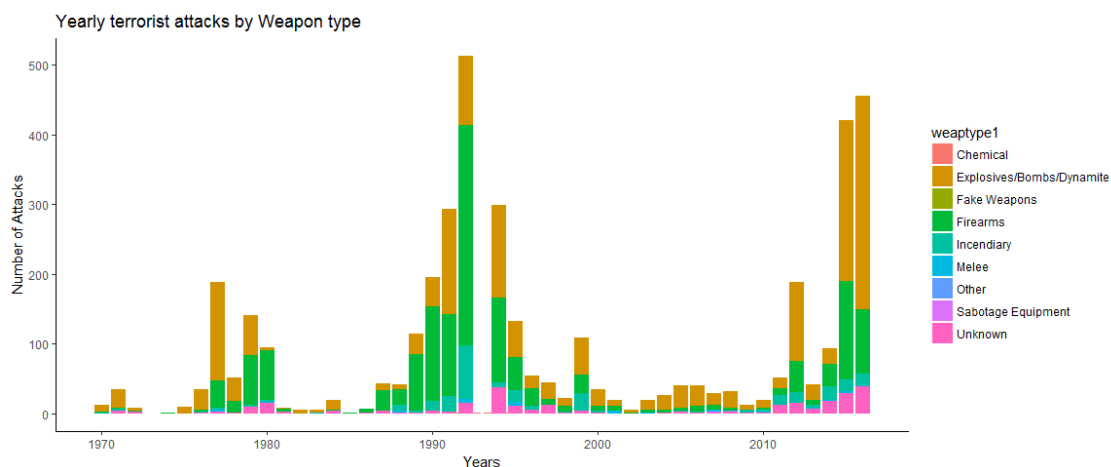
```
g <- ggplot(gnamevstargettype1, aes(gname))
g + geom_bar(aes(fill=targettype1), width = 0.5) +
  theme(axis.text.x = element_text(angle=-90, vjust=0.5, size=10)) +
  labs(title="Target type distrubiton of gnames",
        subtitle="")
```



Yearly terrorist attacks by Weapon type

```
ggplot(data=gtd.turkey, aes(x=year, fill=weaptype1)) +
  geom_bar() + ggtitle("Yearly terrorist attacks by Weapon type") +
  labs(x = "Years", y = "Number of Attacks")
```

Warning: position_stack requires non-overlapping x intervals



Wordcloud Analysis

```
# Install wordcloud if not already installed
if (!("wordcloud" %in% installed.packages())) {
  install.packages("wordcloud", repos = "https://cran.r-project.org")}

# Load wordcloud package
library(wordcloud)
par(mfrow = c(1,3))
```

```
wordcloud(gtd.turkey$targettype1,max.words = 100,random.order =
FALSE,colors=brewer.pal(8, "Dark2"))
wordcloud(gtd.turkey$attacktype1,max.words = 100,random.order =
FALSE,colors=brewer.pal(8, "Dark2"))
wordcloud(gtd.turkey$gname,max.words = 100,random.order =
FALSE,colors=brewer.pal(8, "Dark2"))
```

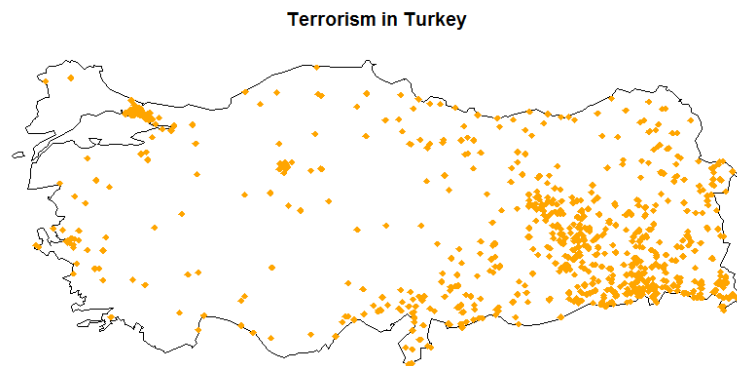


##Inspiration:"<https://www.kaggle.com/bharathsivaraman/an-untold-story-of-terrorism-in-india/notebook>"

Analysis on Map

```
library(sp) # spatial data classes and functions
library(maptools)
library(RColorBrewer)
library(maps) # powerful and flexible mapping package
library(mapdata)

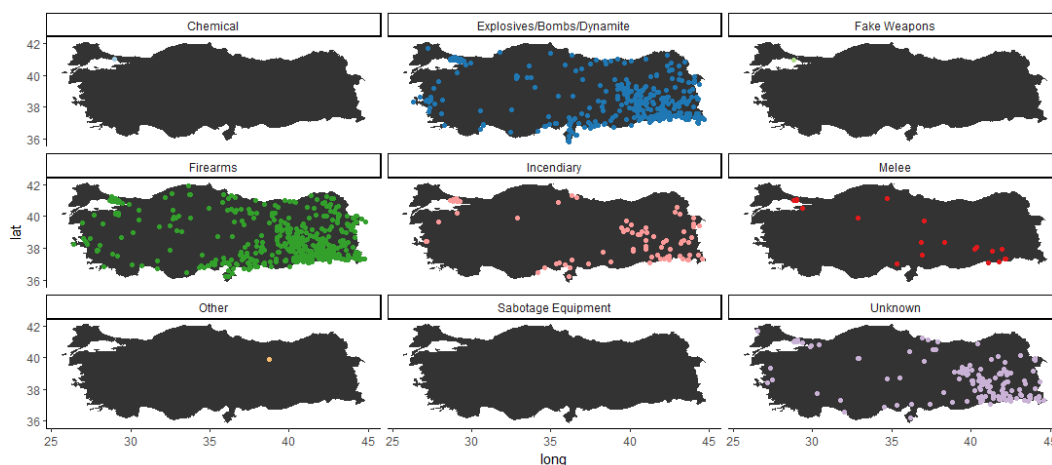
lon_lat <- data.frame(lon=gtd.turkey$longitude, lat=gtd.turkey$latitude)
map("world2Hires", "Turkey")
points(lon_lat$lon, lon_lat$lat, col="orange", pch=18)
title("Terrorism in Turkey")
```



#What are the most used weapons?

```
weapons<-unique(gtd.turkey$weaptype1)
library(ggmap)
map_tr<-map_data("world2Hires", "Turkey")
map_by_weapon<- ggplot() + geom_polygon(data = map_tr, aes(x=long, y = lat,
group = group)) +
  coord_fixed(1)
map_by_weapon<-
map_by_weapon+geom_point(data=gtd.turkey,aes(x=longitude,y=latitude,color=as.
factor(weaptype1)))
map_by_weapon<-
map_by_weapon+scale_color_brewer(palette="Paired")+facet_wrap(~weaptype1,ncol
=3)+guides(color=F)
map_by_weapon
```

Warning: Removed 164 rows containing missing values (geom_point).



##Inspiration:"<https://www.kaggle.com/aiswaryaramachandran/terror-attacks-against-india-since-2000>"

From the visualisation above, we can observe that Firearms and Bombing/Explosion is the most common type of attack and highly prevalent in especially southeast part of Turkey.

Interactive Data Map with Shiny module

Interactive Data Map

Terrorism in Turkey



References

Dataset source

1. [Global Terrorism Database](#)
2. <https://www.kaggle.com/dhrubajitdas/global-terrorism-full-analysis/data>
3. <http://start.umd.edu/gtd/downloads/Codebook.pdf>

Visualization and ggplot

1. <http://r-statistics.co/ggplot2-Tutorial-With-R.html>
2. <http://ggplot2.tidyverse.org/reference/>
3. <https://www.computerworld.com/article/2935394/business-intelligence/my-ggplot2-cheat-sheet-search-by-task.html>
4. <https://plot.ly/r/pie-charts/>
5. <https://www.statmethods.net/graphs/pie.html>
6. http://rstudio-pubs-static.s3.amazonaws.com/15358_25737b089f0042e086e68b1609e4f20b.html
7. <https://stackoverflow.com/questions/19926349/side-by-side-pie-charts-of-2-matrices-r>
8. <http://blocks.org/patilv/raw/7360425/>
9. <https://www.harding.edu/fmccown/r/>

10. <https://uc-r.github.io/lollipop>
11. <https://www.kaggle.com/iklotho/terror-attacks-in-turkey/notebook>

Shiny

1. <https://rstudio.github.io/leaflet/shiny.html>
2. <http://shiny.rstudio.com/tutorial/>

Other

1. <http://zevross.com/blog/2014/08/04/beautiful-plotting-in-r-a-ggplot2-cheatsheet-3/>
2. http://www.css.cornell.edu/faculty/dgr2/teach/R/R_corregr.pdf
3. <https://www.programiz.com/r-programming>
4. <http://lightonphiri.org/blog/ggplot2-multiple-plots-in-one-graph-using-gridextra>
5. [http://www.cookbook-r.com/Graphs/Multiple_graphs_on_one_page_\(ggplot2\)/](http://www.cookbook-r.com/Graphs/Multiple_graphs_on_one_page_(ggplot2)/)