Project: Data Science

Sean E. Curl & Steven Aguilar

April 25, 2018

About

- Global Terrorism Database (GTD)
- Study of Terrorism and Responses to Terrorism (START)
- University of Maryland

https://www.kaggle.com/START-UMD/gtd

Global Terrorism Database (GTD) is an open-source database with information about terrorist attacks worldwide from 1970 to 2016. The database is maintained by researchers at the National Consortium for the Study of Terrorism and Responses to Terrorism (START), headquartered at the University of Maryland.

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."

Observations

- The data-set included 170,350 observations.

Variables?

- Great than 100 variables on location, tactics, perpetrators, targets, and outcomes.

See the http://start.umd.edu/gtd/downloads/Codebook.pdf for important details on data collection methodology, definitions, and coding schema.

The Global Terrorism Database is funded through START, by the US Department of State and the US Department of Homeland Security Science and Technology Directorate's Office of University Programs. The coding decisions and classifications contained in the database are determined independently by START researchers and should not be interpreted as necessarily representing the official views or policies of the United States Government.

Load and Clean

```
## Observations: 170,350
## Variables: 18
## $ Year        <chr> "1970", "1970", "1970", "1970", "1970", "1970", "1...
## $ Month       <chr> "7", "0", "1", "1", "1", "1", "1", "1", "1", "1", ...
## $ Day         <chr> "2", "0", "0", "0", "0", "1", "2", "2", "2", "3", ...
## $ Country     <chr> "Dominican Republic", "Mexico", "Philippines", "Gr...
## $ Region      <chr> "Central America & Caribbean", "North America", "S...
## $ AttackType  <chr> "Assassination", "Hostage Taking (Kidnapping)", "A...
## $ Target      <chr> "Julio Guzman", "Nadine Chaval, daughter", "Employ...
## $ Killed      <dbl> 1, 0, 1, NA, NA, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, NA,...
## $ Wounded     <dbl> 0, 0, 0, NA, NA, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, NA,...
```

```
## $ Summary     <chr> NA, NA, NA, NA, NA, "1/1/1970: Unknown African Ame...
## $ Group       <chr> "MANO-D", "23rd of September Communist League", "U...
## $ Target_Type <chr> "Private Citizens & Property", "Government (Diplom...
## $ Weapon_type <chr> NA, NA, NA, "Unknown Explosive Type", NA, "Unknown...
## $ Motive      <chr> NA, NA, NA, NA, NA, "To protest the Cairo Illinois...
## $ City        <chr> "Santo Domingo", "Mexico city", "Unknown", "Athens...
## $ lat         <dbl> 18.45679, 19.43261, 15.47860, 37.98377, 33.58041, ...
## $ long        <dbl> -69.95116, -99.13321, 120.59974, 23.72816, 130.396...
## $ Casualties  <dbl> 1, 0, 1, NA, NA, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, NA,...
```

```
##        Year       Month         Day     Country      Region  AttackType
##           0           0           0           0           0           0
##      Target      Killed     Wounded     Summary       Group Target_Type
##         633        9682       15325       66138           0           0
## Weapon_type      Motive        City         lat        long  Casualties
##       19426      121764         446        4606        4606       15826
```

In the interested of time and sanity, we selected from the >100 variables, the best variables we thought would lead to a 'quick n' easy' interpretable user experience. As a result, we also created a variable of our own called 'causalities.' The variable is the summation of the Killed and Wounded variables. This variable allowed us a quick and easy way measure the impact a particular terrorist event(s) by using only a single variable.

As you can see, the terrorism database has a bunch of missing observations amongst Target, Killed, Wounded, Summary, Weapon_type, Motive, City, lat, long, and Casualties. Many of these variables are missing simply because the researchers lacked the fidelity for an exact location, weapon, motive, etc. In the case of Killed and Wounded, missing values represent that fact that no person was either killed or wounded in that particular attack. For example, an attack could be deemed a success if the intended target was destroyed (e.g. infrastructure), but there might not have been any casualties as a result of the attack.

High Impact Areas

```
## [1] "Year with Highest Terrorist Attacks: 2014"
```

```
## [1] "Month with Highest Terrorist Attacks: 5"
```

```
## [1] "Day with Highest Terrorist Attacks: 15"
```

```
## [1] "Country with Highest Terrorist Attacks: Iraq"
```

```
## [1] "Region with Highest Terrorist Attacks: Middle East & North Africa"
```

```
## [1] "AttackType with Highest Terrorist Attacks: Bombing/Explosion"
```

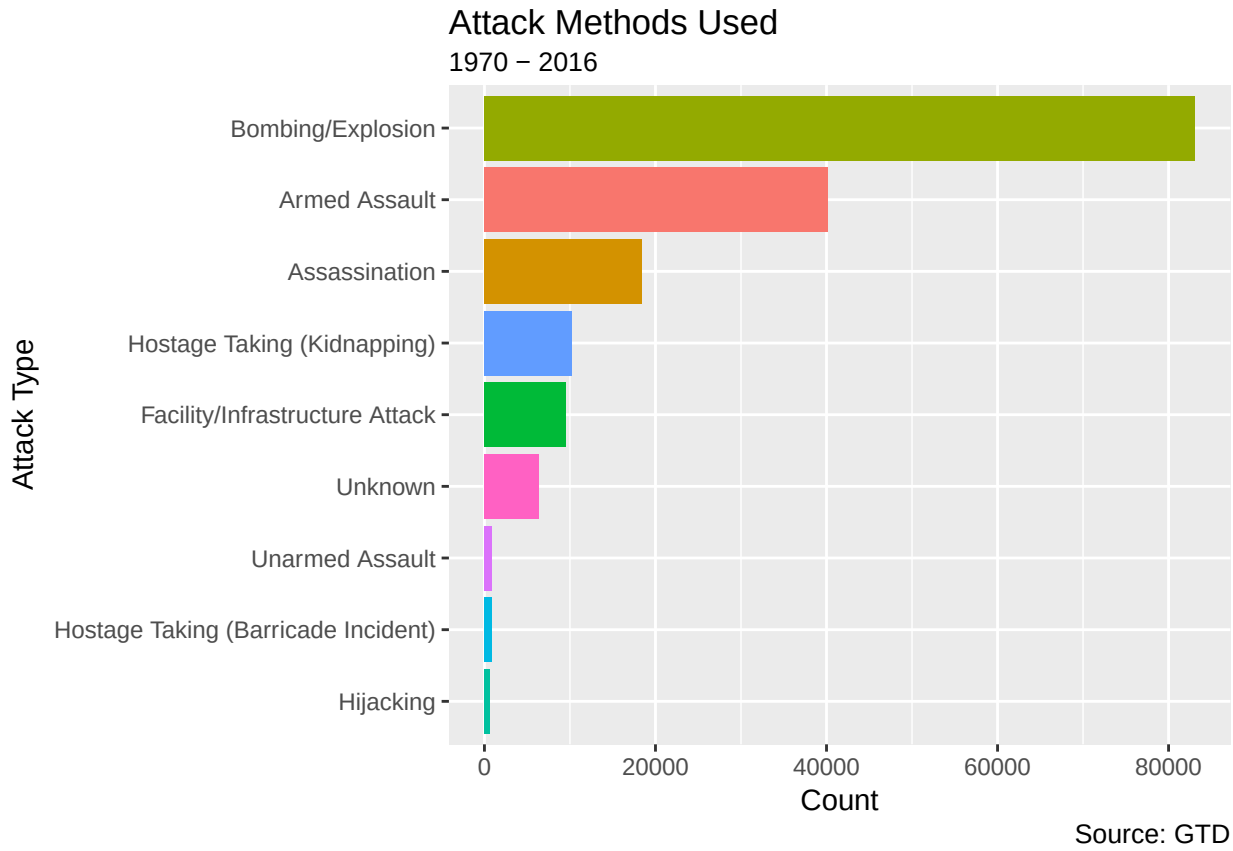
```
## [1] "Target with Highest Terrorist Attacks: Civilians"
```

```
## [1] "Maximum peopled killed in an attack are: In 2014 1500 peopled died in Iraq"
```

```
## [1] "Maximum peopled wounded in an attack are: In 2001 7366 peopled died in United States"
```

```
## [1] "Group with Highest Terrorist Attacks (not unk.): Taliban"
```

Popular Attack Types



Here's a table view. . .

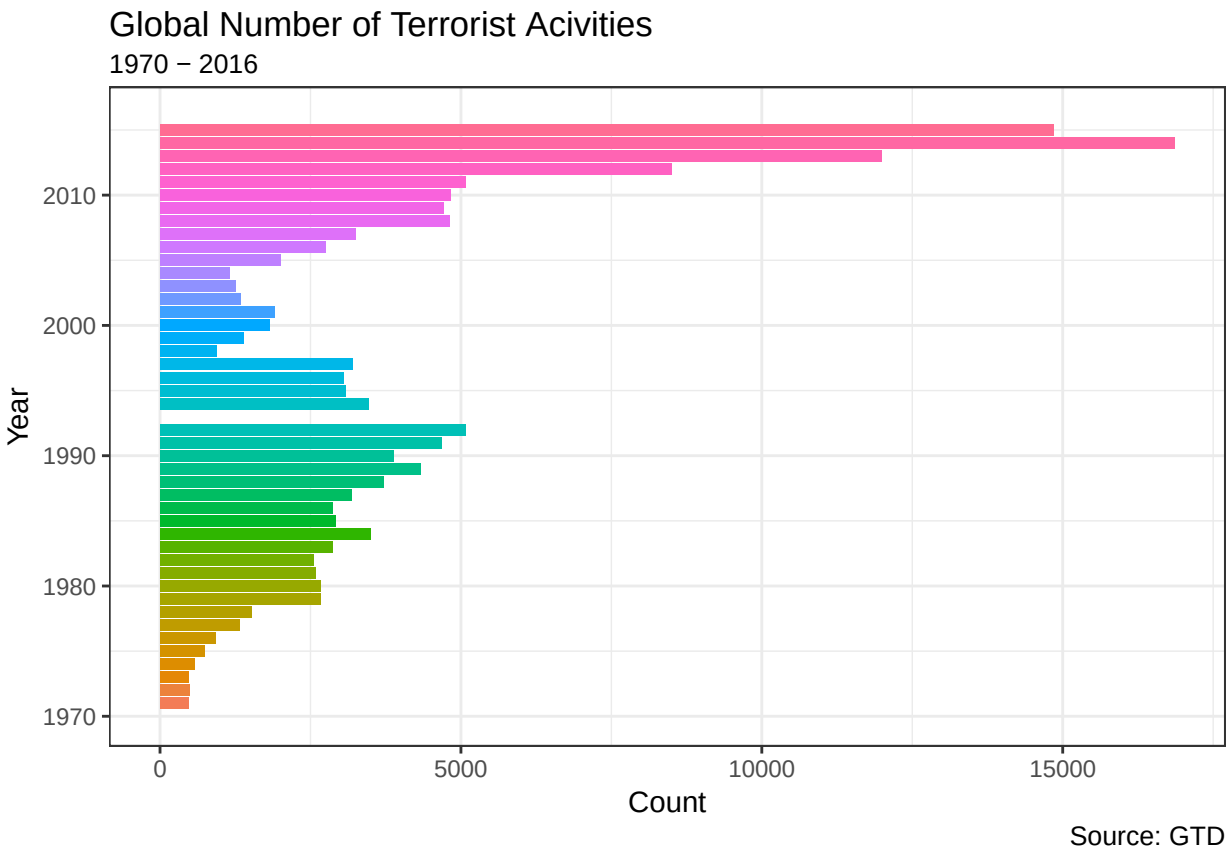
```
## # A tibble: 9 x 2
##   AttackType                          n
##   <chr>                           <int>
## 1 Bombing/Explosion               83073
## 2 Armed Assault                   40223
## 3 Assassination                   18402
## 4 Hostage Taking (Kidnapping)     10233
## 5 Facility/Infrastructure Attack   9581
## 6 Unknown                          6425
## 7 Unarmed Assault                   913
## 8 Hostage Taking (Barricade Incident)  902
## 9 Hijacking                         598
```

Let's take a look at the top three attack types. I'll include Year, Group, and Country too.

```
## # A tibble: 3 x 5
## # Groups:   AttackType, Year, Country [?]
##   AttackType        Year  Country  Group                                n
##   <chr>             <chr> <chr>    <chr>                            <int>
## 1 Bombing/Explosion 2016  Iraq     Islamic State of Iraq and the Le~   821
## 2 Assassination     1995  Pakistan Unknown                              242
```

```
## 3 Armed Assault     2014  Pakistan Unknown                          500
```

Number of Terrorist Acivities Year



We can see that the most recent years had a significant increase in violence.

```
## # A tibble: 6 x 2
##   Year      n
##   <chr> <int>
## 1 2014  16860
## 2 2015  14852
## 3 2016  13488
## 4 2013  11996
## 5 2012   8500
## 6 1992   5073
```

Let's take a look at the Year's > 2010.

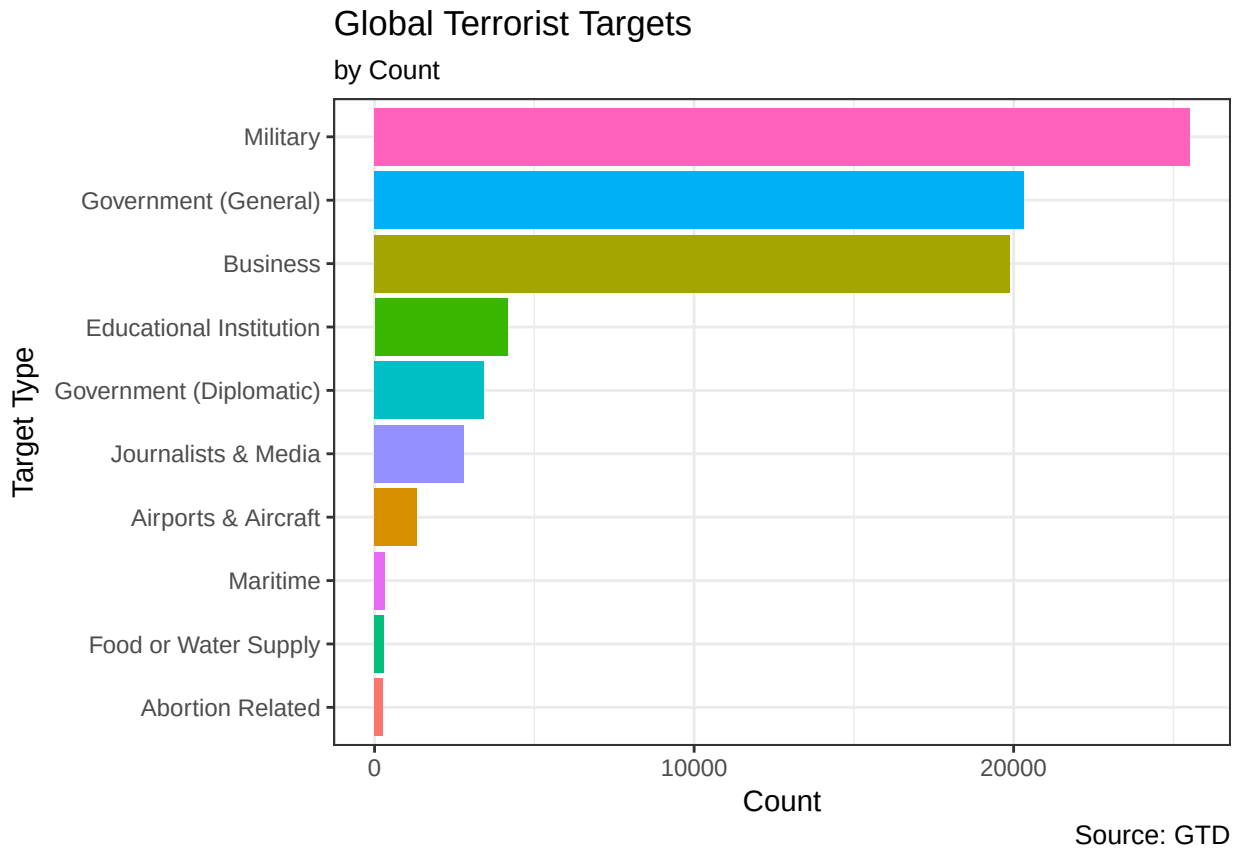
```
## # A tibble: 5 x 3
## # Groups:   Year [5]
##   Year  Group                                          n
##   <chr> <chr>                                      <int>
## 1 2016  Islamic State of Iraq and the Levant (ISIL)  1447
## 2 2015  Taliban                                     1249
## 3 2014  Islamic State of Iraq and the Levant (ISIL)  1247
```

```
## 4 2012  Taliban                                          800
## 5 2013  Taliban                                          773
```

How about looking at the Year's < 2010?

```
## # A tibble: 5 x 3
## # Groups:   Year [5]
##   Year  Group                                             n
##   <chr> <chr>                                         <int>
## 1 1989  Shining Path (SL)                                509
## 2 1984  Shining Path (SL)                                502
## 3 1983  Shining Path (SL)                                493
## 4 1991  Farabundo Marti National Liberation Front (FMLN)  492
## 5 1987  Shining Path (SL)                                464
```

Terrorist Targets by Count



What are the top 5 preferred targets for different groups?

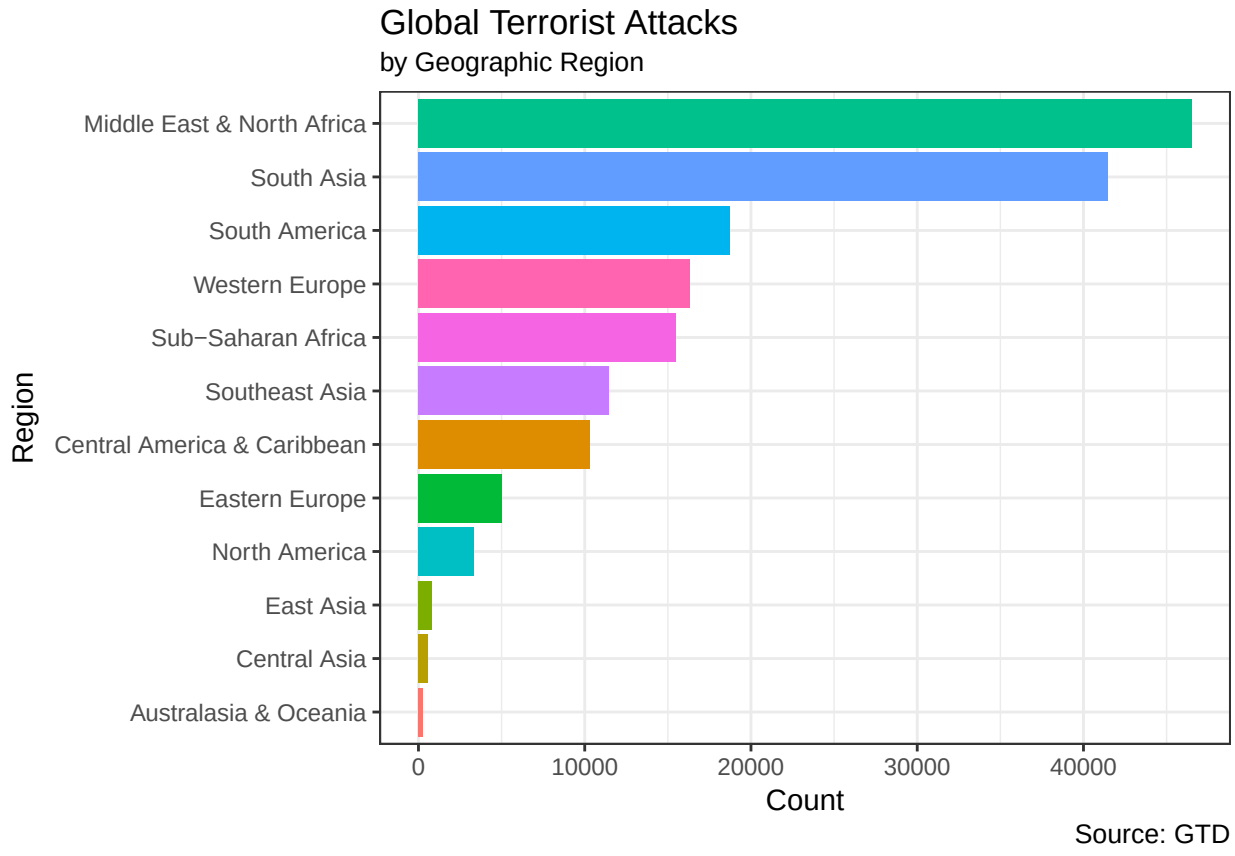
```
## # A tibble: 5 x 3
## # Groups:   Target_Type [5]
##   Target_Type                Group                          n
##   <chr>                      <chr>                      <int>
## 1 Police                     Taliban                     2201
```

```
## 2 Private Citizens & Property Islamic State of Iraq and the Levant ~  1724
## 3 Military                    Farabundo Marti National Liberation F~  1230
## 4 Utilities                   Farabundo Marti National Liberation F~   923
## 5 Government (General)        Taliban                                  851
```

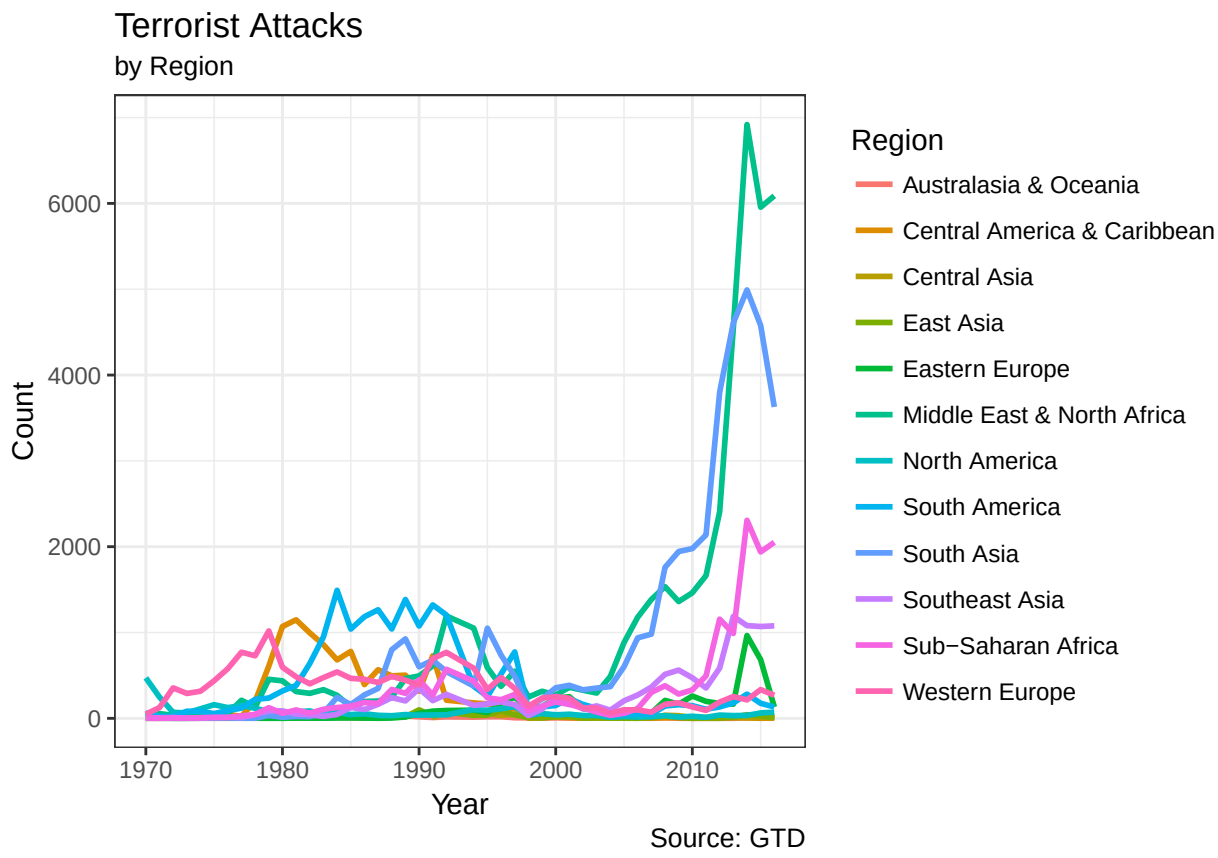
What are top 5 preferred targets within different countries?

```
## # A tibble: 5 x 3
## # Groups:   Target_Type [5]
##   Target_Type                 Country     n
##   <chr>                       <chr>   <int>
## 1 Private Citizens & Property Iraq     7794
## 2 Police                      Iraq     3502
## 3 Military                    Iraq     2768
## 4 Government (General)        Iraq     2142
## 5 Business                    Iraq     1897
```

Terrorist Attacks by Region



It's important to look across a timeframe. Let's compare Region across the years.



Which regions have the highest amount of terrorist events?

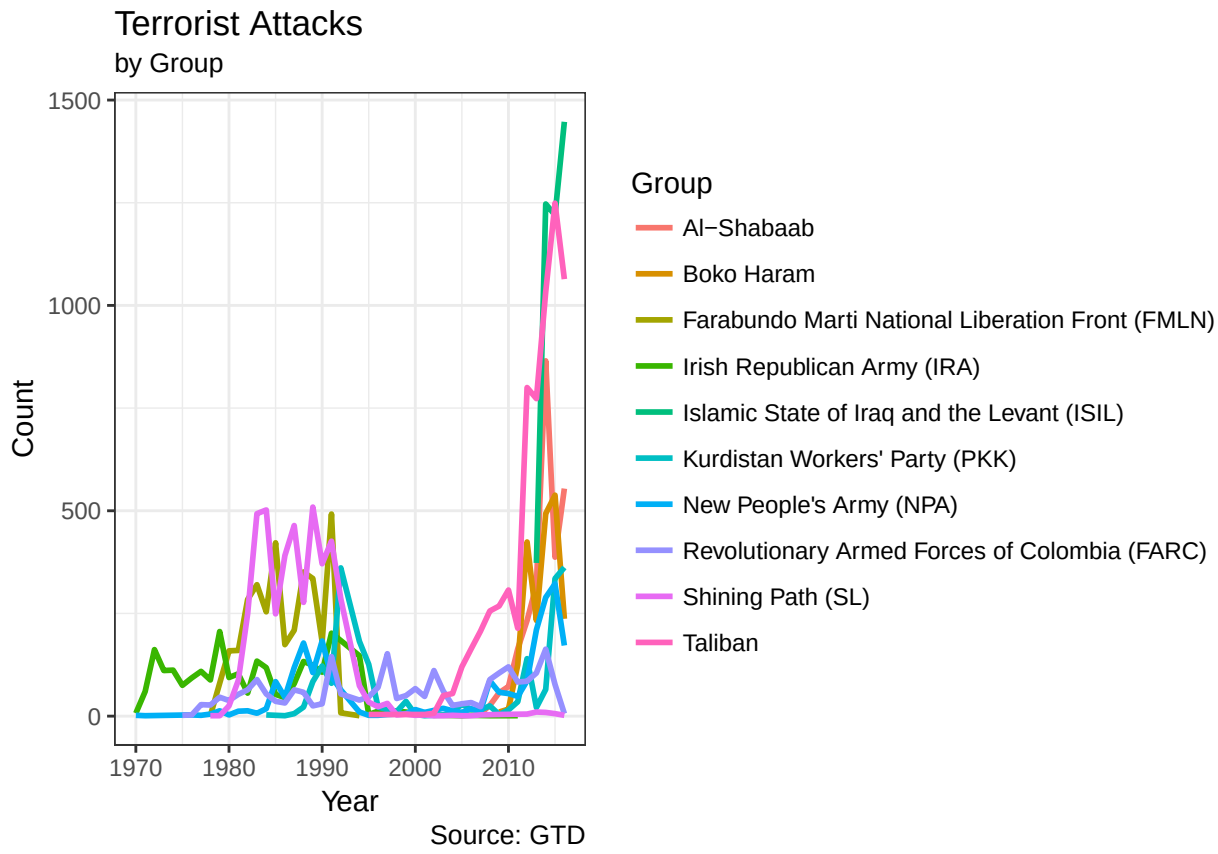
```
## # A tibble: 10 x 3
## # Groups:   Year [10]
##    Year  Region                        n
##    <chr> <chr>                     <int>
##  1 2014  Middle East & North Africa  6919
##  2 2016  Middle East & North Africa  6088
##  3 2015  Middle East & North Africa  5956
##  4 2013  South Asia                  4607
##  5 2012  South Asia                  3799
##  6 2011  South Asia                  2137
##  7 2010  South Asia                  1978
##  8 2009  South Asia                  1944
##  9 2008  South Asia                  1759
## 10 1984  South America               1492
```

. . . which regions have which groups? Here's the top 6.

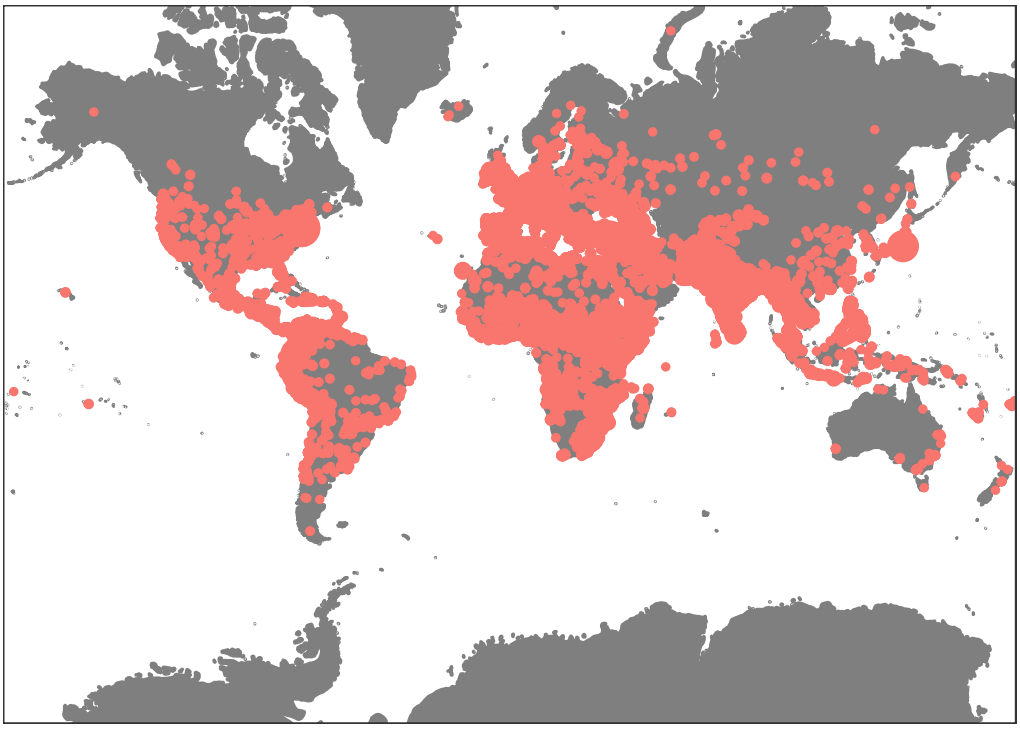
```
## # A tibble: 6 x 3
## # Groups:   Region [6]
##   Region                    Group                                 n
##   <chr>                     <chr>                             <int>
## 1 South Asia                Taliban                            6574
## 2 South America             Shining Path (SL)                  4550
```

```
## 3 Middle East & North Africa  Islamic State of Iraq and the Levant ~  4260
## 4 Central America & Caribbean Farabundo Marti National Liberation F~  3351
## 5 Sub-Saharan Africa          Al-Shabaab                               2683
## 6 Western Europe              Irish Republican Army (IRA)              2666
```

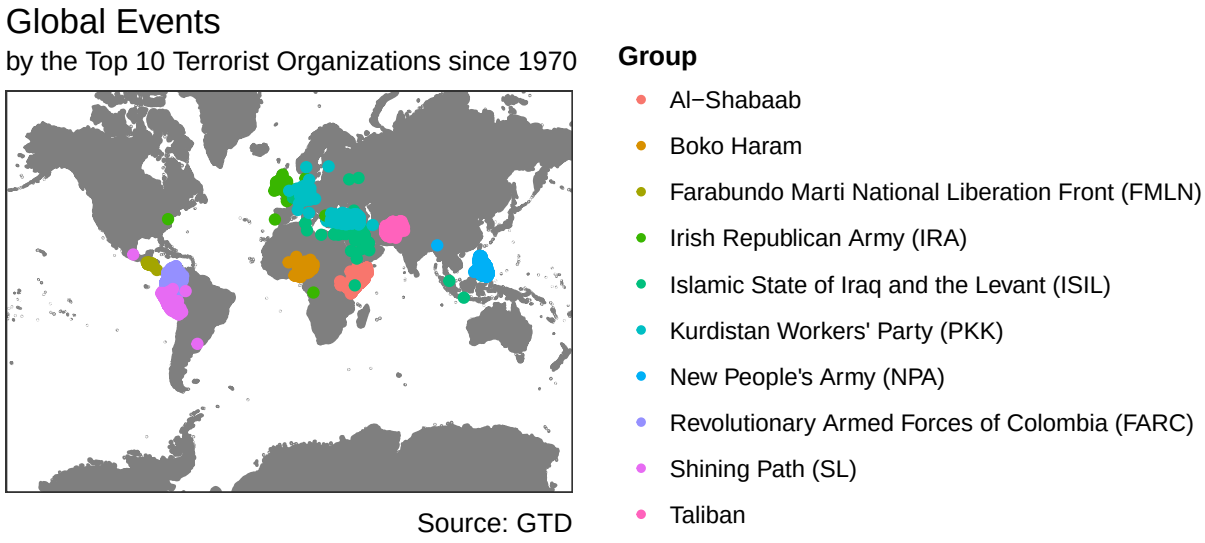
Terrorist attacks by Group. We'll show the top 10.



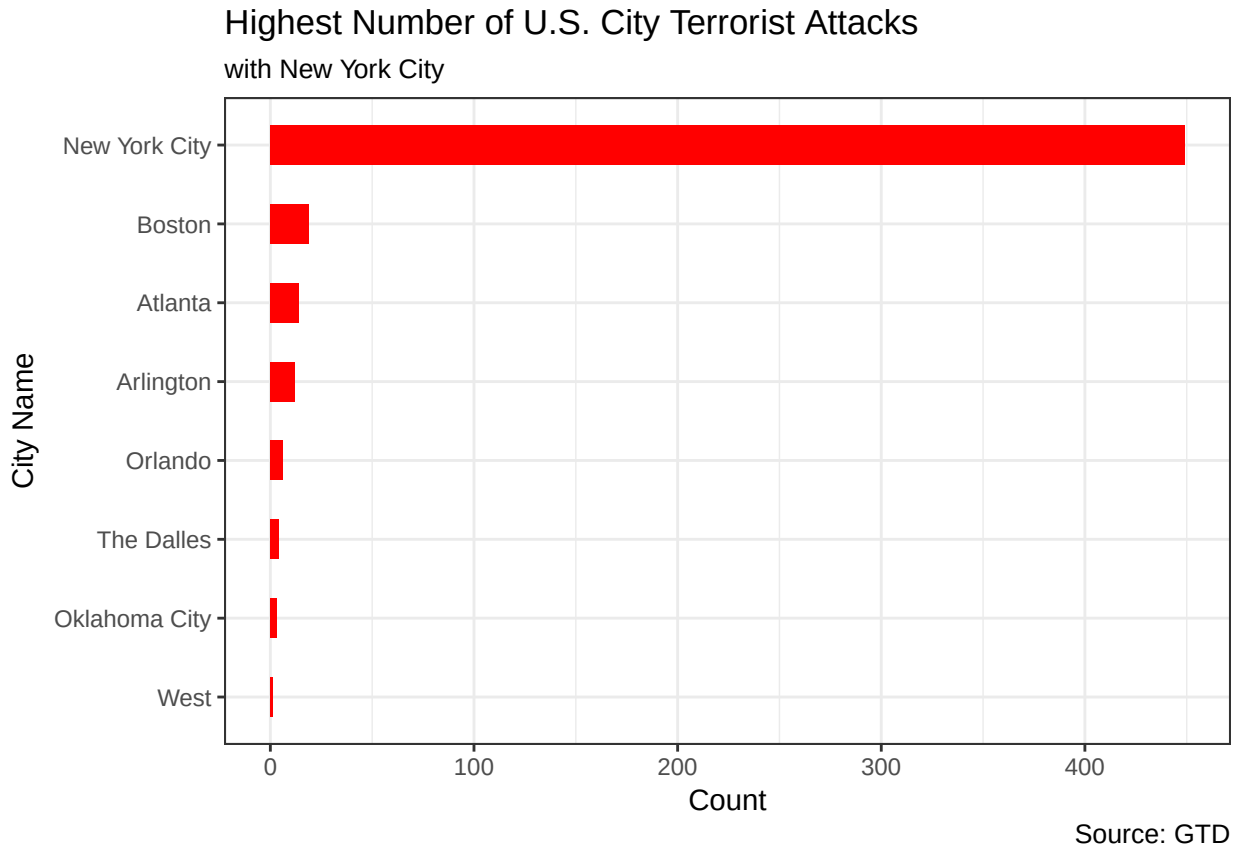
Global Maps Plots: Total Terrorism Attacks from 1970 - 2016



Global Maps Plots: Terrorism by Groups from 1970 - 2016



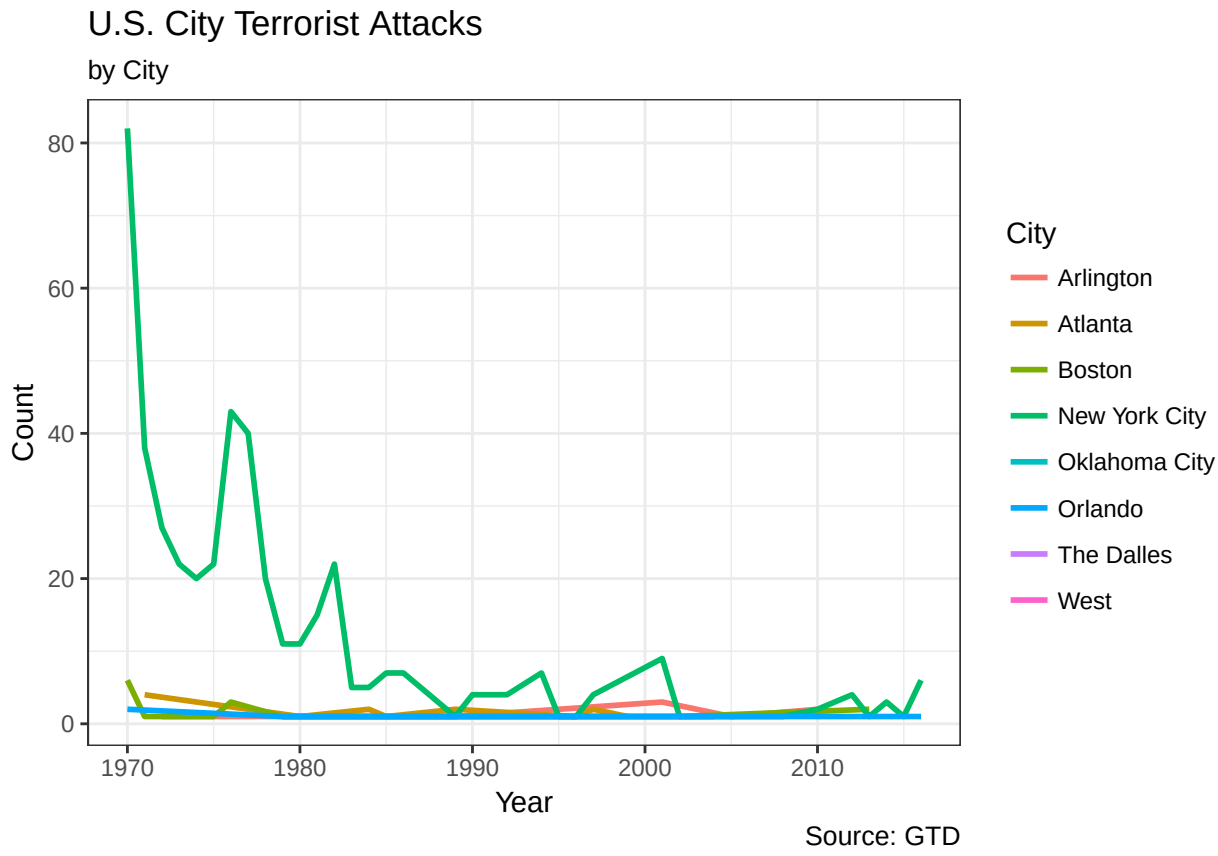
Terrorist Attacks by City within USA



New York City has the highest number of terrorist attacks since 1970.

```
## # A tibble: 10 x 3
## # Groups:   City [1]
##    City          Year  total
##    <chr>         <chr> <int>
##  1 New York City 1970     82
##  2 New York City 1976     43
##  3 New York City 1977     40
##  4 New York City 1971     38
##  5 New York City 1972     27
##  6 New York City 1973     22
##  7 New York City 1975     22
##  8 New York City 1982     22
##  9 New York City 1974     20
## 10 New York City 1978     20
```

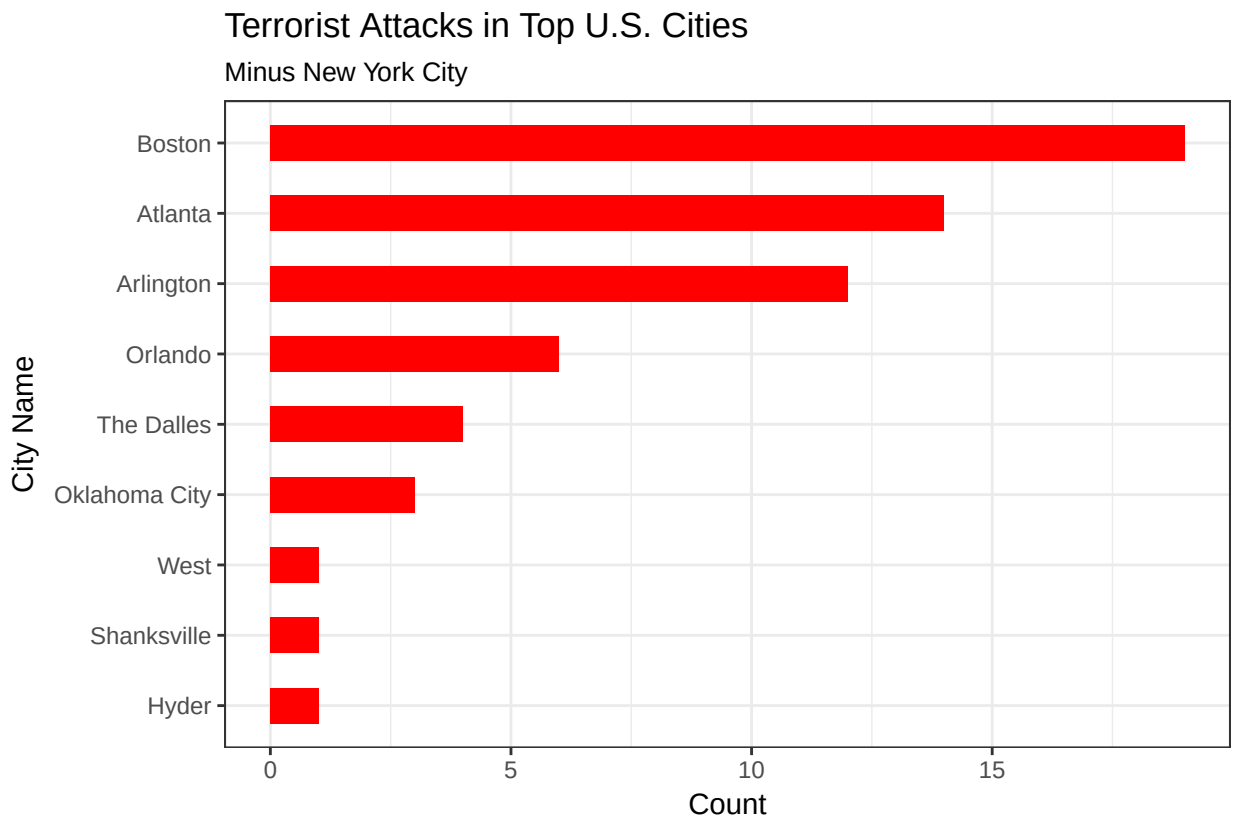
What do these attacks look like across time?



Across all U.S. cities, which had the deadliest attacks?

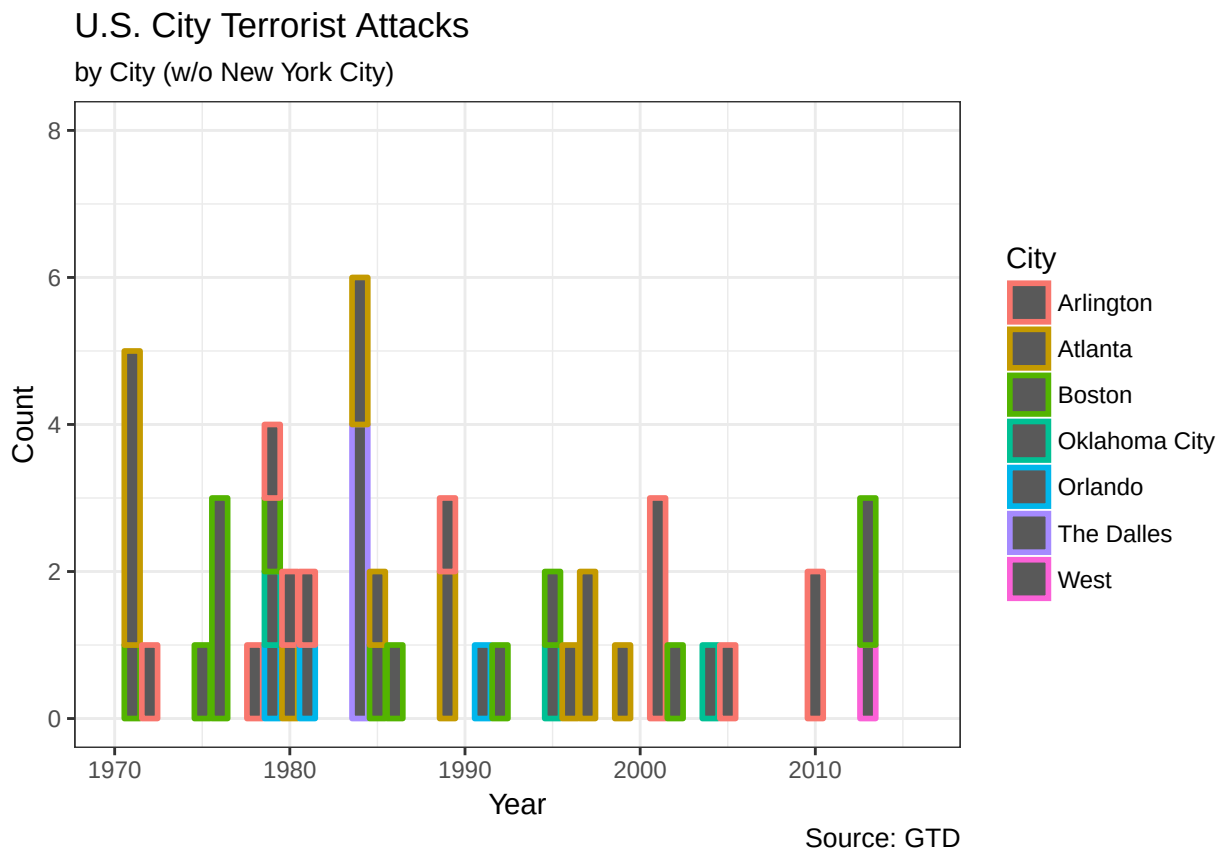
```
## # A tibble: 10 x 4
## # Groups:   Year, City [10]
##    Year  City          Casualties total
##    <chr> <chr>              <dbl> <dbl>
##  1 2001  New York City      8749. 8749.
##  2 1995  Oklahoma City       818.  818.
##  3 1984  The Dalles          751.  751.
##  4 2001  Arlington           295.  295.
##  5 2013  West                166.  166.
##  6 2013  Boston              134.  134.
##  7 1996  Atlanta             111.  111.
##  8 2016  Orlando             103.  103.
##  9 1975  New York City        85.   85.
## 10 1995  Hyder                79.   79.
```

Let's exclude New York City.




```
## # A tibble: 6 x 3
## # Groups:   City [6]
##   City          Year  total
##   <chr>         <chr> <int>
## 1 Boston        1970      6
## 2 Atlanta       1971      4
## 3 The Dalles    1984      4
## 4 Arlington     2001      3
## 5 Orlando       1970      2
## 6 Oklahoma City 1979      1
```

Again, what do these attacks look like across time?



Across these U.S. cities, which had the deadliest attacks?

```
## # A tibble: 6 x 4
## # Groups:   Year, City [6]
##   Year  City          Casualties total
##   <chr> <chr>              <dbl> <dbl>
## 1 1995  Oklahoma City       818.  818.
## 2 1984  The Dalles          751.  751.
## 3 2001  Arlington           295.  295.
## 4 2013  West                166.  166.
## 5 2013  Boston              134.  134.
## 6 1996  Atlanta             111.  111.
```

Shiny App Link:

https://saguilar.shinyapps.io/GTDB/

RPubs Presentation:

http://rpubs.com/securl/revealjspresent

Presentation created using REVEAL.JS

- https://revealjs.com/#/
- https://rmarkdown.rstudio.com/revealjs_presentation_format.html

Github Repo:

- https://github.com/dataSeanC

QUESTIONS?