Global Terrorism Analysis Project

“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 or intimidation.”

United Nations

1. **Largest human-oriented criminal activity**

2. **UN's main peace and security agendas**

3. **Evaluate and analyze past terrorist attacks**

4. **Identify risky regions:** Develop State's capacity to defeat terrorism

5. **Identify target terrorist organization:**
   1. Deny terrorists the means to carry out an attack
   2. Dissuade people and nations from supporting terrorism

Original Dataset - 181691 obs. of 135 columns

```
terror.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 181691 entries, 0 to 181690
Columns: 135 entries, eventid to related
dtypes: float64(55), int64(22), object(58)
memory usage: 187.1+ MB
```

```
terror.head()
```

	eventid	iyear	imonth	iday	approxdate	extended	resolution	country	country_txt	region	...	addnotes	scite1	scite2	scite3	dbsource	INT_LOG
0	197000000001	1970	7	2	NaN	0	NaN	58	Dominican Republic	2	...	NaN	NaN	NaN	NaN	PGIS	0
1	197000000002	1970	0	0	NaN	0	NaN	130	Mexico	1	...	NaN	NaN	NaN	NaN	PGIS	0
2	197001000001	1970	1	0	NaN	0	NaN	160	Philippines	5	...	NaN	NaN	NaN	NaN	PGIS	-9
3	197001000002	1970	1	0	NaN	0	NaN	78	Greece	8	...	NaN	NaN	NaN	NaN	PGIS	-9
4	197001000003	1970	1	0	NaN	0	NaN	101	Japan	4	...	NaN	NaN	NaN	NaN	PGIS	-9

5 rows × 135 columns

Normalization plan

Data transformation：

Concatenated year, month, day variables to Date variable

1NF:

Remove duplicate columns: Weapon_type1, Weapon_type2

2NF & 3NF:

Entity tables(Primary Key(ID), Entity Name): gname_id, gname_name

Relation tables: Eventid, gname_id

Extract, Transform, Load (ETL) process - 15 tables

01 Step

02 Step

03 Step

Extract

- Remove duplicate columns
- Design appropriate names (rename columns)
- Make a draft for constraints
- Extract useful columns, drop columns with too many null values, and drop duplicated data

Transform

- Check for duplicated data
- Data munging (transforming and mapping data)
- Push data properly based on our database schema

Load

- Load data
- Check for data loading accuracy by running SQL queries in PostgreSQL database

Process of interacting with data

For analysts

Jupyter notebook scripts, **direct SQL queries** and **ER Diagrams**

Analysts are already proficient in **programming languages** so that they can understand our rationale by reading our scripts

For c-level officers

Metabase - create **visualizations** to demonstrate our findings/insights

Especially for people who **don't have the technique background**

Results For Analysts

1. The **Middle East & North Africa** area suffers from terrorist attacks the most.
2. The most dangerous country is **Iraq** with the top 5 cities: **Baghdad, Mosul, Kirkuk, Baqubah and Fallujah**.
3. **Taliban** and **ISIL** are two most active terrorist groups in the **recent 20 years**.
4. The **bombing/explosion** is the most used attack method by terrorists
5. The most common target in a terrorist attack is mainly **private citizens & property (civilians)**.
6. The **year 2014** has the most terrorist attacks. An interesting fact is that **the top 8 years with most attacks are all post-2010.** In other words, since 2010, the number of terrorists grew exponentially until 2014.

Dashboard

http://f19server.apan5310.com:3208/public/dashboard/320483fc-0745-4881-a528-b47ccfa40dab

http://f19server.apan5310.com:3208/public/dashboard/96e285f1-7659-47af-88c0-12fab0bc8cc7