

Data Warehouses – Project

Wroclaw University of Science and Technology, Date: June 15, 2022

Student:	-----	Grade
Identifier	<u>245784</u>	
First name	<u>Rahul</u>	
Last name	<u>Vijayvargiya</u>	

This mini project assignment consists of 1 task.

1. Title of the project

Global Terrorism Database

More than 180,000 terrorist attacks worldwide, 1970-2017

2. General description of the domain

The Global Terrorism Database (GTD) is the most comprehensive unclassified database of terrorist attacks in the world. The National Consortium for the Study of Terrorism and Responses to Terrorism (START) makes the GTD available via this site to improve understanding of terrorist violence, so that it can be more readily studied and defeated.

The GTD is produced by a dedicated team of researchers and technical staff.

The GTD is an open-source database, which provides information on domestic and international terrorist attacks around the world since 1970, and now includes more than 200,000 events. For each event, a wide range of information is available, including the date and location of the incident, the weapons used, nature of the target, the number of casualties, and – when identifiable – the group or individual responsible

Data analysis is to enable, among others, to establish regions of the city and country in which the highest number of Terrorist attacks happened, they also become more and more brutal., What kind of method or weapon they used, and Which Org has done more damage to the world and its people, we can gain insight how many people has been killed till now in those attacks.

3. Description of the analysis area (selected fragment of the domain, intended for detailed analysis and development of the data warehouse)

The Goal of the project is to create a Data Warehouse for Business Intelligence so we can do a detailed analysis of the GTD Dataset b/w 1970-2017. It will be interesting to analyze the cities where the arrests took place, the time, which org has planned these attacks, what kind of weapon they used, how many people killed until 2017 in those attacks, we will complete exploratory data analysis of Global Terrorism Dataset.

3.1. Problems

The significant problems include: 1. Attacks Increased over the years 2. More and More People have been killed 3. Type of weapon or Type of attack 4. Finding most destructive terror org which planned alot of attacks 5. Finding the trend of the root cause

3.2. Project goal

3.2.1. Expectations and needs for decision-making support (research questions)

The analytical database should enable an in-depth analysis of the facts of GTD, Type of attack, Weapon, Time, Terror Org,

1. Are there more suicide bombing, killing with guns, or hijacking airplane
2. At what place highest number of attacks took place
3. Which terror org has taken place most attacks
4. Which attack killed most people
5. Which year has the most attacks took place

A proper analysis of historical data should provide factual answers to the above questions and provide information from which it will be possible to draw conclusions that constitute the basis for making correct decisions

3.2.2. Scope of analysis - Aspects examined

The Scope will be analyzed under fixed circumstances such as Year, Org Type, Killed, Type of attacks, place and so on

Which places and for what reasons does a large number of offenses occur?

What time of day is conducive to the highest number of offenses?

What are the most common causes and effects of detentions?

We should analyze furthermore the ways to stop such attacks, Predictive Analysis could help us her, In depth analysis can be done in Tableau, Python

3.2.3. Potential data warehouse users

This analytical database will support the decision-making processes and understanding the root cause for such attacks will be used by a government entities and anti-terrorism squad of respective countries

furthermore, more the information will be provided to journalists interested in the subject matter

4. Source of data

4.1. Data Sources

The source data needed for the creation of a data warehouse are following, it's an open-source database, free to use and available in csv file format

No	File	Type	Rec.	Size [MB]	Description
1.	Global Terrorism Database	CSV	181619	155 MB	This Dataset is about Global Terrorist Attacks Between 1970 – 2017 Details

4.2. Location, availability of source data

The data comes from the websites:

We Use Kaggle to Download the Dataset, it's an Open-Source dataset, I attached following link to the dataset, attached below

file available in csv format

[Global Terrorism Database](#)

[Global Terrorism Database | Kaggle](#)

Publications: The GTD has been leveraged extensively in scholarly publications, reports, and media articles. Putting Terrorism in Context: Lessons from the Global Terrorism Database, by GTD principal investigators LaFree, Dugan, and Miller investigates patterns of terrorism and provides perspective on the challenges of data collection and analysis. The GTD's data collection manager, Michael Jensen, discusses important Benefits and Drawbacks of Methodological Advancements in Data Collection and Coding.

4.3. Data Dictionary – Interpretation, data semantics

I renamed the columns so they can be more readable to the end user and can understand what sort data it consists in more meaningful words, I Use Python for renaming the Columns and find out which column has the most null values and cannot use

Code below for renaming columns:

```
terror_df.rename(columns={'iyear':'Year','imonth':'Month','iday':'Day',  
'country_txt':'Country','provstate':'state',  
'region_txt':'Region','attacktype1_txt':'AttackType','target1':'Target','nkill':'Killed',  
'nwound':'Wounded','summary':'Summary','gname':'Group','targtype1_txt':'Target_type',  
'weaptype1_txt':'Weapon_type','motive':'Motive'}, inplace=True)
```

Column approxdate has been dropped since nulls percentage is 95 %

Column resolution has been dropped since nulls percentage is 99 %

Column location has been dropped since nulls percentage is 69 %

Column summary has been dropped since nulls percentage is 36 %

Column alternative has been dropped since nulls percentage is 84 %

Column alternative_txt has been dropped since nulls percentage is 84 %

Column attacktype2 has been dropped since nulls percentage is 97 %

Column attacktype2_txt has been dropped since nulls percentage is 97 %

Column attacktype3 has been dropped since nulls percentage is 100 %

Column attacktype3_txt has been dropped since nulls percentage is 100 %

Column targtype2 has been dropped since nulls percentage is 94 %

Column targtype2_txt has been dropped since nulls percentage is 94 %

Column targsubtype2 has been dropped since nulls percentage is 94 %

Column targsubtype2_txt has been dropped since nulls percentage is 94 %

Column corp2 has been dropped since nulls percentage is 94 %

Column target2 has been dropped since nulls percentage is 94 %

Column natlty2 has been dropped since nulls percentage is 94 %

Column natlty2_txt has been dropped since nulls percentage is 94 %

Column targtype3 has been dropped since nulls percentage is 99 %

Column targtype3_txt has been dropped since nulls percentage is 99 %

Column targsubtype3 has been dropped since nulls percentage is 99 %

Column targsubtype3_txt has been dropped since nulls percentage is 99 %

Column corp3 has been dropped since nulls percentage is 99 %

Column target3 has been dropped since nulls percentage is 99 %

Column natlty3 has been dropped since nulls percentage is 99 %

Column natlty3_txt has been dropped since nulls percentage is 99 %

Column gsubname has been dropped since nulls percentage is 97 %

Column gname2 has been dropped since nulls percentage is 99 %

Column gsubname2 has been dropped since nulls percentage is 100 %

Column gname3 has been dropped since nulls percentage is 100 %

Column gsubname3 has been dropped since nulls percentage is 100 %

Column motive has been dropped since nulls percentage is 72 %

Column guncertain2 has been dropped since nulls percentage is 99 %

Column guncertain3 has been dropped since nulls percentage is 100 %

Column nperps has been dropped since nulls percentage is 39 %

Column nperpcap has been dropped since nulls percentage is 38 %

Column claimed has been dropped since nulls percentage is 36 %

Column claimmode has been dropped since nulls percentage is 89 %

Column claimmode_txt has been dropped since nulls percentage is 89 %

Column claim2 has been dropped since nulls percentage is 99 %

Column claimmode2 has been dropped since nulls percentage is 100 %

Column claimmode2_txt has been dropped since nulls percentage is 100 %

Column claim3 has been dropped since nulls percentage is 100 %

Column claimmode3 has been dropped since nulls percentage is 100 %

Column claimmode3_txt has been dropped since nulls percentage is 100 %

Column complaim has been dropped since nulls percentage is 97 %

Column weaptype2 has been dropped since nulls percentage is 93 %

Column weaptype2_txt has been dropped since nulls percentage is 93 %

Column weapsubtype2 has been dropped since nulls percentage is 94 %

Column weapsubtype2_txt has been dropped since nulls percentage is 94 %

Column weaptype3 has been dropped since nulls percentage is 99 %

Column weaptype3_txt has been dropped since nulls percentage is 99 %

Column weapsubtype3 has been dropped since nulls percentage is 99 %

Column weapsubtype3_txt has been dropped since nulls percentage is 99 %

Column weaptype4 has been dropped since nulls percentage is 100 %

Column weaptype4_txt has been dropped since nulls percentage is 100 %

Column weapsubtype4 has been dropped since nulls percentage is 100 %

Column weapsubtype4_txt has been dropped since nulls percentage is 100 %

Column weapdetail has been dropped since nulls percentage is 37 %

Column nkillus has been dropped since nulls percentage is 35 %

Column nkillter has been dropped since nulls percentage is 37 %

Column nwoundus has been dropped since nulls percentage is 36 %

Column nwoundte has been dropped since nulls percentage is 38 %

Column propextent has been dropped since nulls percentage is 65 %

Column propextent_txt has been dropped since nulls percentage is 65 %

Column propvalue has been dropped since nulls percentage is 79 %

Column propcomment has been dropped since nulls percentage is 68 %

Column nhostkid has been dropped since nulls percentage is 93 %

Column nhostkidus has been dropped since nulls percentage is 93 %

Column nhours has been dropped since nulls percentage is 98 %

Column ndays has been dropped since nulls percentage is 96 %

Column divert has been dropped since nulls percentage is 100 %

Column kidhijcountry has been dropped since nulls percentage is 98 %

Column ransom has been dropped since nulls percentage is 57 %

Column ransomamt has been dropped since nulls percentage is 99 %

Column ransomamtus has been dropped since nulls percentage is 100 %

Column ransompaid has been dropped since nulls percentage is 100 %

Column ransompaidus has been dropped since nulls percentage is 100 %

Column ransomnote has been dropped since nulls percentage is 100 %

Column hostkidoutcome has been dropped since nulls percentage is 94 %

Column hostkidoutcome_txt has been dropped since nulls percentage is 94 %

Column nreleased has been dropped since nulls percentage is 94 %

Column addnotes has been dropped since nulls percentage is 84 %

Column scite1 has been dropped since nulls percentage is 36 %

Column scite2 has been dropped since nulls percentage is 58 %

Column scite3 has been dropped since nulls percentage is 76 %

Column related has been dropped since nulls percentage is 86 %

4.4. Qualitative assessment of data

Data Source: globalterrorism.csv	Shape Of records (181691, 18)
----------------------------------	-------------------------------

Column	Dtype	Meaning	Remark
Year	int64	Year of attack	Integer Number of year
Month	int64	Month of attack	Contain Month in Integer
Day	int64	Day of attack	Contain Day in Int
Country	object	Country of attack	Country as String Object
State	object	State where it occurs	Store State as String Object
Region	object	Region of attack	Store Region as String
City	object	City Where it took place	Store City as String
Latitude	float64	Latitude of the attack	Store Float Value of Latitude
Longitude	float64	Longitude of the attack	Store Float Value of Longitude
AttackType	object	Attack type eg: bomb or gun attack	Attack type as String Object
Killed	float64	People killed in attack	Killed People store as a Float Data Type
Wounded	float64	People who are injured in attack	Wounded People Store as a Float data type
Target	object	Target of attack	String Object of taget
Summary	object	Summary or attack	String Object of Summary
Group	object	Terror org name	String object of Group
Target_type	object	Target type	String object of Target Type
Weapon_type	object	Weapon used	Weapon Type as a string object
Motive	Object	Motive of attack	Motive for the cause of attack is a string object

Columns	Null Values
Year	0
Month	0
Day	0
Country	0
state	421
Region	0
city	434
latitude	4556
longitude	4557
AttackType	0
Killed	10313
Wounded	16311
Target	636
Summary	66129
Group	0
Target_type	0
Weapon_type	0
Motive	131130

5. Multidimensional analytical models

5.1. Fact-analysis context

Table 4. Facts subject to analysis

Lp.	Fakt	Measure(s)
1.	Fact_Events	EventID, Group, Killed, Suicide, Wounded, Success, TimeID, attack_id, weapon_id, target_id, country_id
2.		
...	...	

5.2. Facts analysis context

Table 5. Facts Analysis Dimensions

No.	Dimension	Properties, characteristics
1.	Time	Contains Year, Month and Day
2.	Attack	Contains Attack Type and Attack Type ID
3.	Country	Contains Country Name, Region, State, City, Longitude and Latitude of attacked place
4.	Target	Contains Target ID, Target Type and Target Type ID
5.	Weapon	Contains Weapon Type and Weapon ID

5.3. Multidimensional Models (UML)



6. Design of the ETL process

6.1. HD database schema (SQL script)

BEGIN

```
IF NOT EXISTS (SELECT * FROM sys.databases WHERE name = 'GlobalTerrorDW2022')
CREATE DATABASE GlobalTerrorDW2022;
END;
GO
```

```
CREATE TABLE [dbo].[attack](
    [attack_id] [bigint] NOT NULL,
    [attacktype1] [bigint] NULL,
    [AttackType] [nvarchar](4000) NULL,
    CONSTRAINT [PK_attack] PRIMARY KEY CLUSTERED
(
    [attack_id] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO
```

```
CREATE TABLE [dbo].[country](
    [country_id] [bigint] NOT NULL,
    [country_code] [bigint] NULL,
    [Country] [nvarchar](4000) NULL,
    [region] [bigint] NULL,
    [state] [nvarchar](4000) NULL,
    [city] [nvarchar](4000) NULL,
    [longitude] [float] NULL,
    [latitude] [float] NULL,
    CONSTRAINT [PK_country] PRIMARY KEY CLUSTERED
(
    [country_id] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO
```

```
CREATE TABLE [dbo].[factEvents](
    [eventid] [bigint] NOT NULL,
```

```

[Group] [nvarchar](4000) NULL,
[attack_id] [bigint] NOT NULL,
[weapon_id] [bigint] NOT NULL,
[TimeID] [bigint] NOT NULL,
[Wounded] [bigint] NULL,
[Killed] [bigint] NULL,
[suicide] [bigint] NULL,
[target_id] [bigint] NOT NULL,
[success] [bigint] NULL,
[country_id] [bigint] NOT NULL,
CONSTRAINT [PK_fact_events] PRIMARY KEY CLUSTERED
(
[eventid] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

```

```

ALTER TABLE [dbo].[factEvents] WITH CHECK ADD CONSTRAINT [FK_fact_events_attack]
FOREIGN KEY([attack_id])
REFERENCES [dbo].[attack] ([attack_id])
GO

```

```

ALTER TABLE [dbo].[factEvents] CHECK CONSTRAINT [FK_fact_events_attack]
GO

```

```

ALTER TABLE [dbo].[factEvents] WITH CHECK ADD CONSTRAINT
[FK_fact_events_country] FOREIGN KEY([country_id])
REFERENCES [dbo].[country] ([country_id])
GO

```

```

ALTER TABLE [dbo].[factEvents] CHECK CONSTRAINT [FK_fact_events_country]
GO

```

```

ALTER TABLE [dbo].[factEvents] WITH CHECK ADD CONSTRAINT [FK_fact_events_Target]
FOREIGN KEY([target_id])
REFERENCES [dbo].[Target] ([target_id])
GO

```

```

ALTER TABLE [dbo].[factEvents] CHECK CONSTRAINT [FK_fact_events_Target]
GO

```

```

ALTER TABLE [dbo].[factEvents] WITH CHECK ADD CONSTRAINT [FK_fact_events_time]
FOREIGN KEY([TimeID])
REFERENCES [dbo].[time] ([TimeID])
GO

```

```

ALTER TABLE [dbo].[factEvents] CHECK CONSTRAINT [FK_fact_events_time]
GO

```

```
ALTER TABLE [dbo].[factEvents] WITH CHECK ADD CONSTRAINT  
[FK_fact_events_Weapon] FOREIGN KEY([weapon_id])  
REFERENCES [dbo].[Weapon] ([weapon_id])  
GO
```

```
ALTER TABLE [dbo].[factEvents] CHECK CONSTRAINT [FK_fact_events_Weapon]  
GO
```

```
CREATE TABLE [dbo].[Target](  
    [target_id] [bigint] NOT NULL,  
    [target_category_id] [bigint] NULL,  
    [Target_category] [nvarchar](4000) NULL,  
    [Target] [nvarchar](4000) NULL,  
    CONSTRAINT [PK_Target] PRIMARY KEY CLUSTERED  
(  
        [target_id] ASC  
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =  
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON,  
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]  
) ON [PRIMARY]  
GO
```

```
CREATE TABLE [dbo].[time](  
    [TimeID] [bigint] NOT NULL,  
    [year] [bigint] NULL,  
    [month] [bigint] NULL,  
    [day] [bigint] NULL,  
    CONSTRAINT [PK_time] PRIMARY KEY CLUSTERED  
(  
        [TimeID] ASC  
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =  
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON,  
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]  
) ON [PRIMARY]  
GO
```

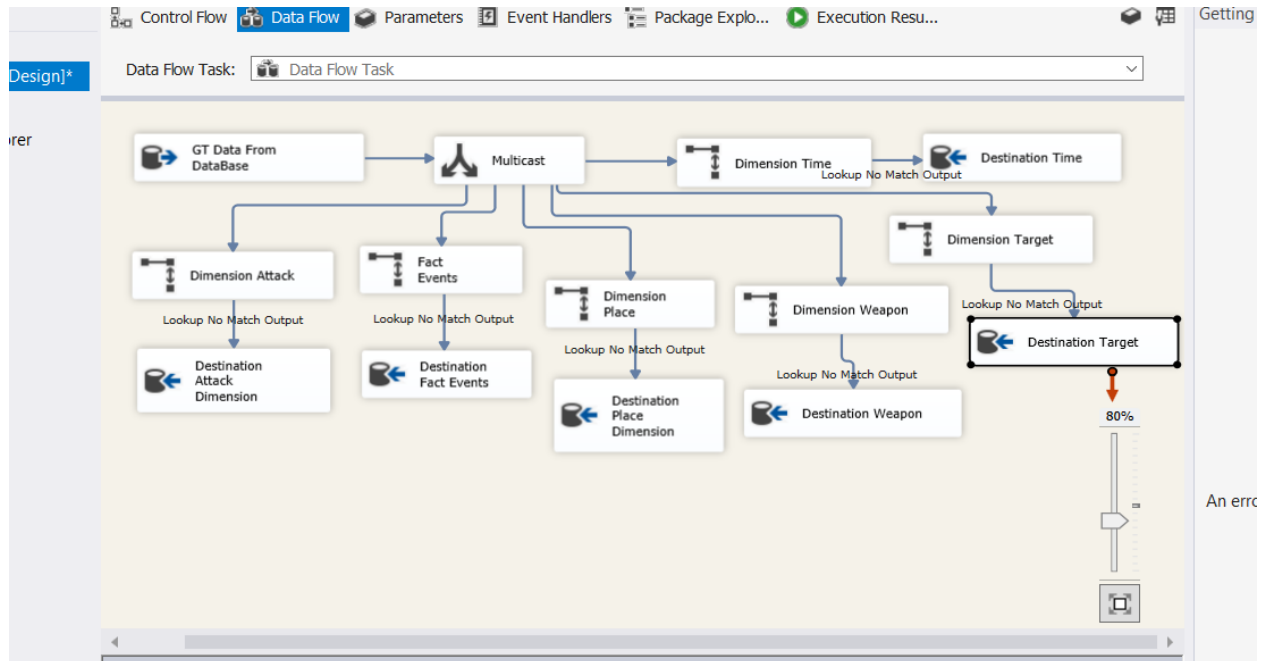
```
CREATE TABLE [dbo].[Weapon](  
    [weapon_id] [bigint] NOT NULL,  
    [weaptype_category_id] [bigint] NULL,  
    [Weapon] [nvarchar](4000) NULL,  
    CONSTRAINT [PK_Weapon] PRIMARY KEY CLUSTERED  
(  
        [weapon_id] ASC
```

```

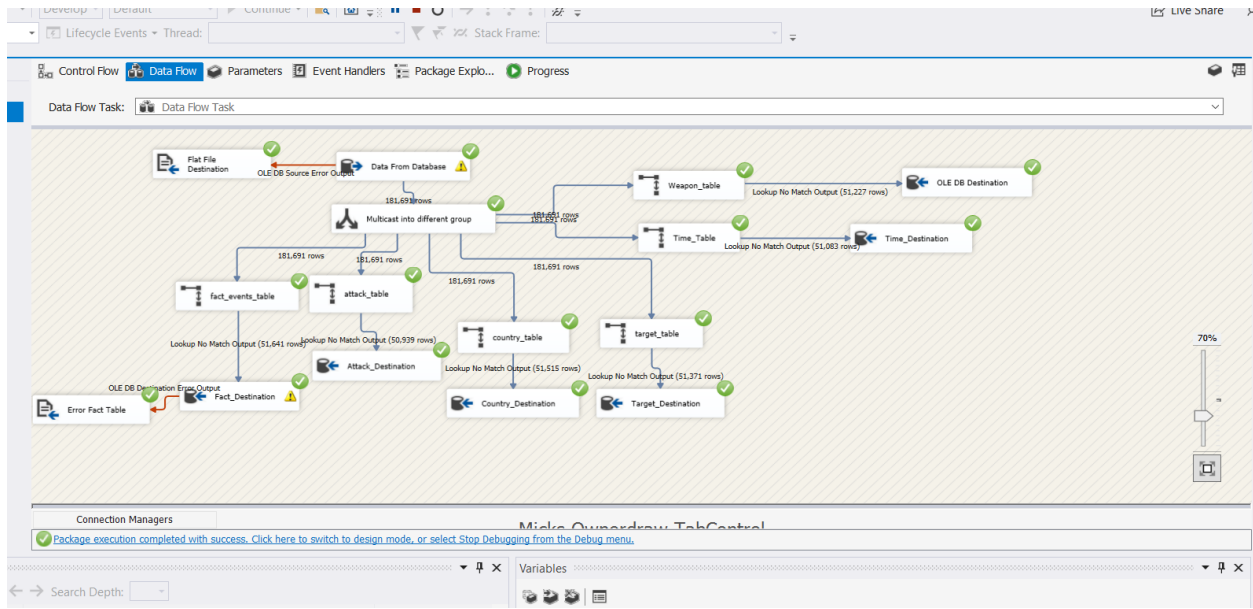
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

```

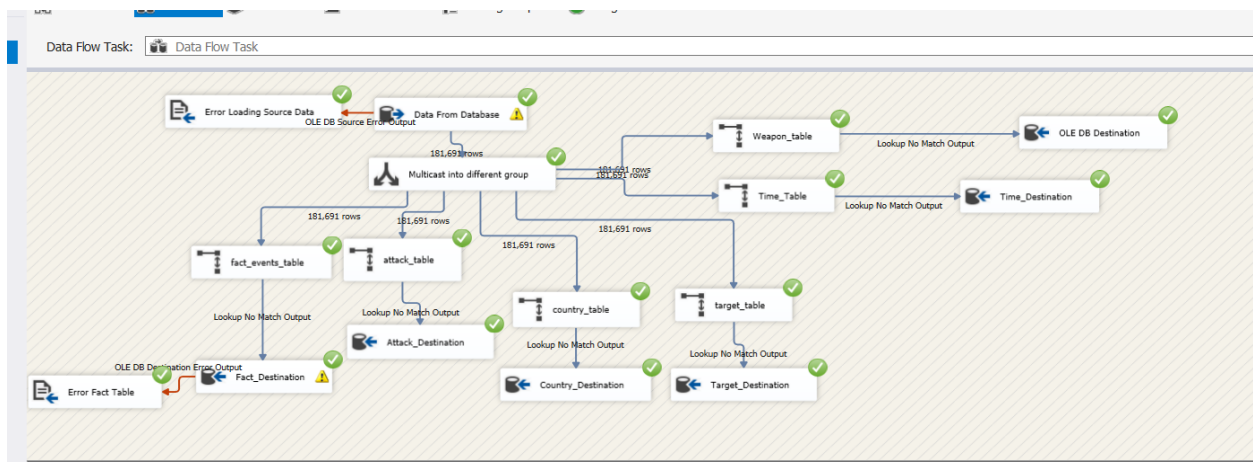
6.2. Specification of ETL (Control Flow + Data Flow) processes



Picture 1: ETL Implementation



Picture 2: Running ETL Flow

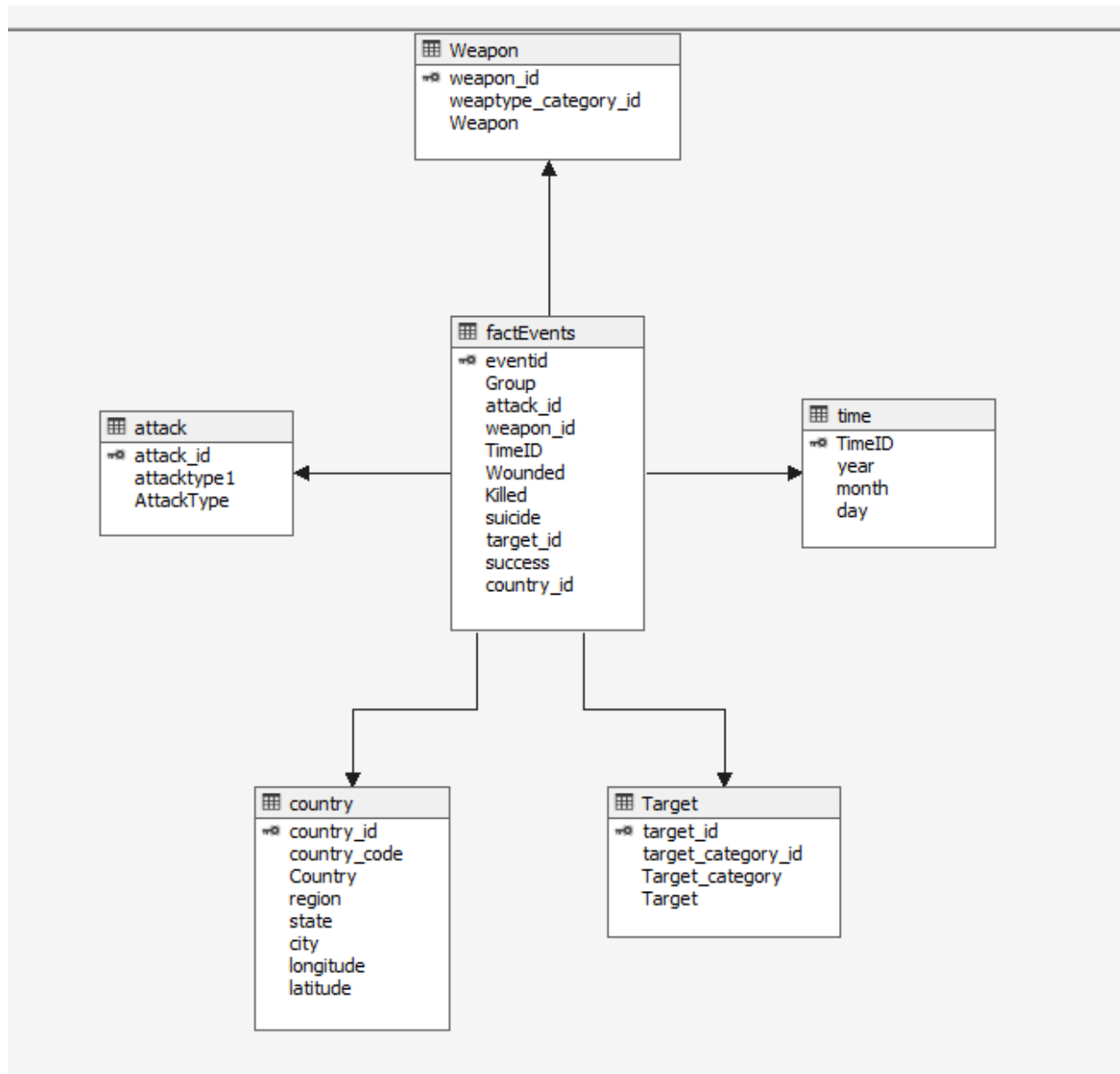


Picture 3: Prevention of Duplication of records

In the Above pictures as you can see, there are multiple flows of ETL, in first picture we implemented ETL and in 2nd picture we run our solution, tested okay, we successfully populate our analytical database, in the 3rd picture it's a checking prevention of duplicate records in DW if we run our flow again and we tested it successfully, turned out we cannot put duplicate records even if we run the whole ETL Flow

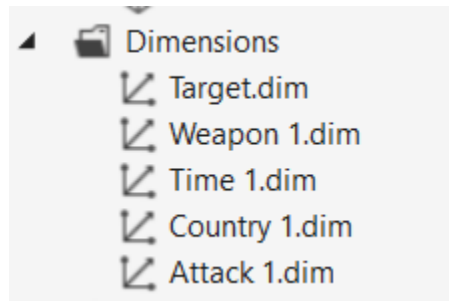
7. Implementation of multidimensional models

7.1. Data View

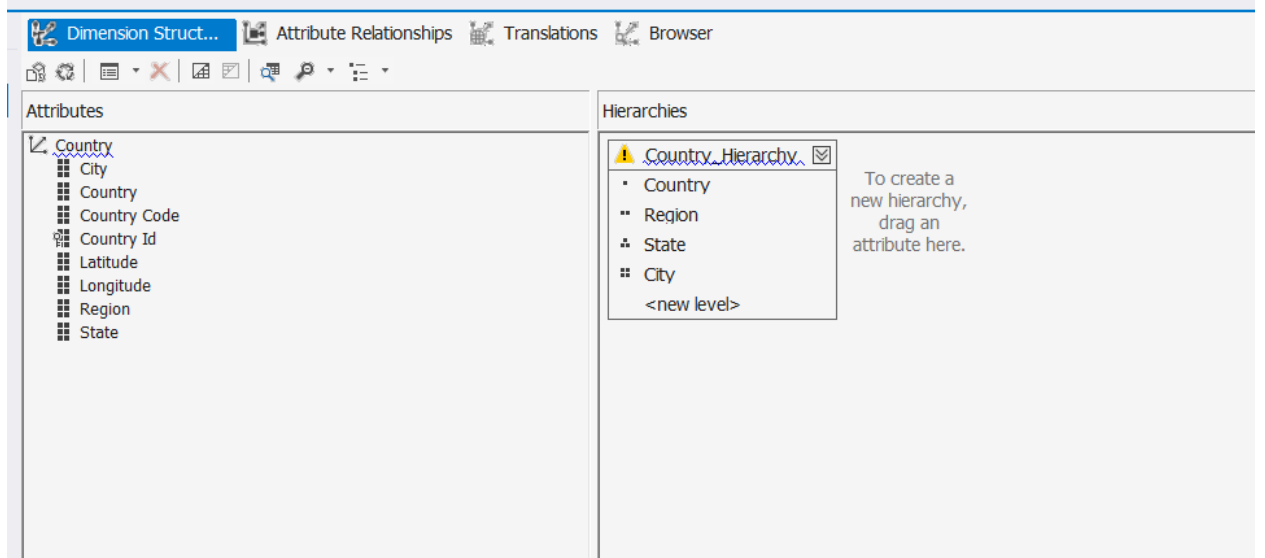
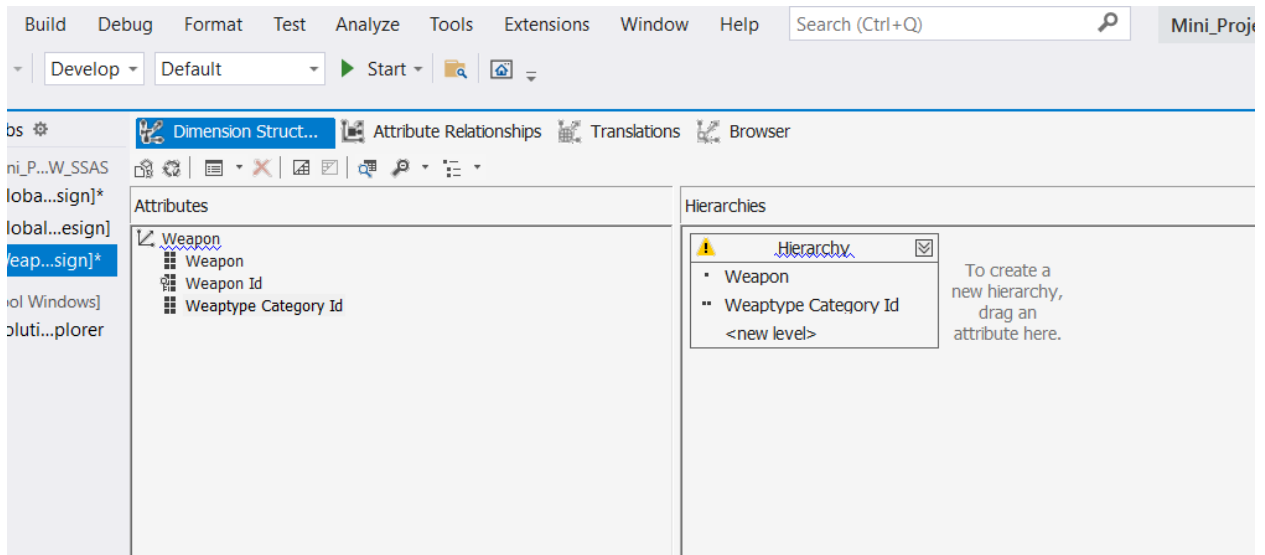


Picture 4: Data Source View

7.2. Dimensions

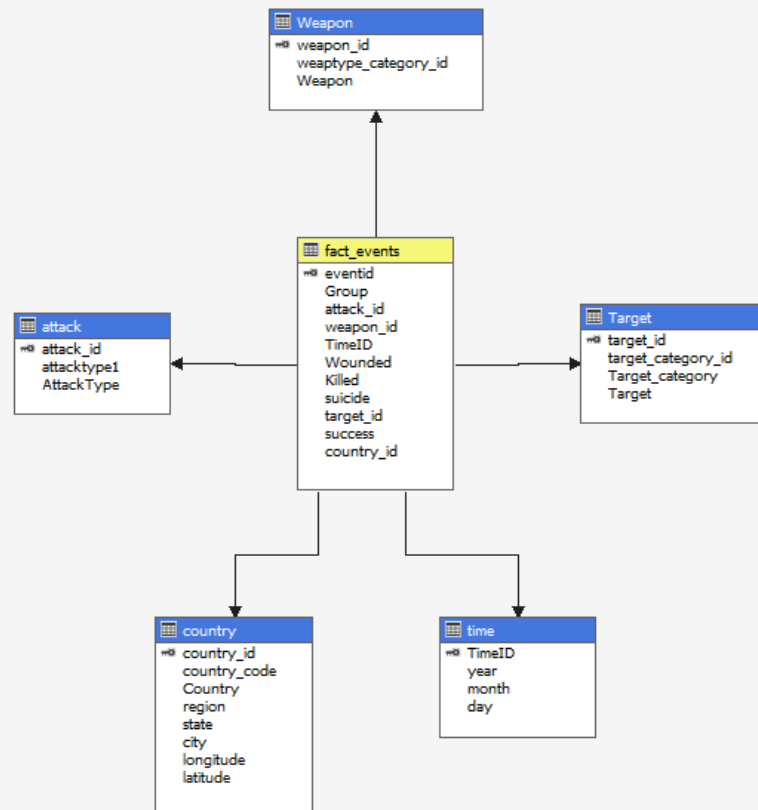


Picture 5 Dimension: Contains list of Dimension in Cubes

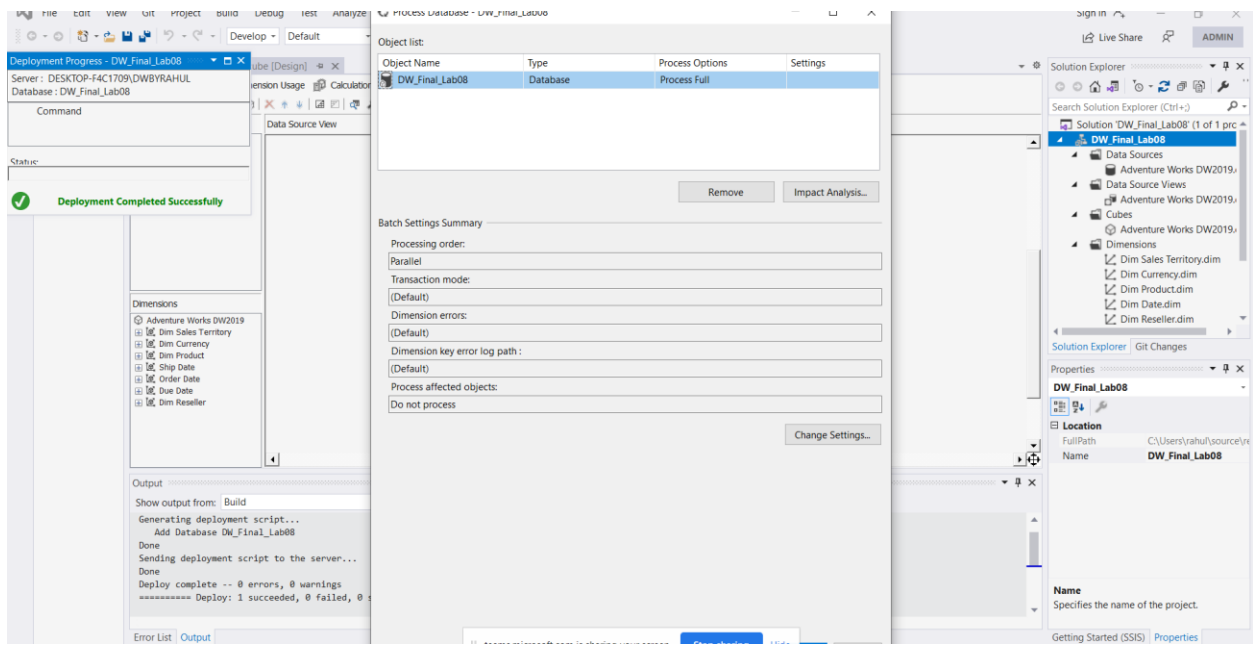


Picture 6.1 and 6.2: Hierarchy in Dimensions

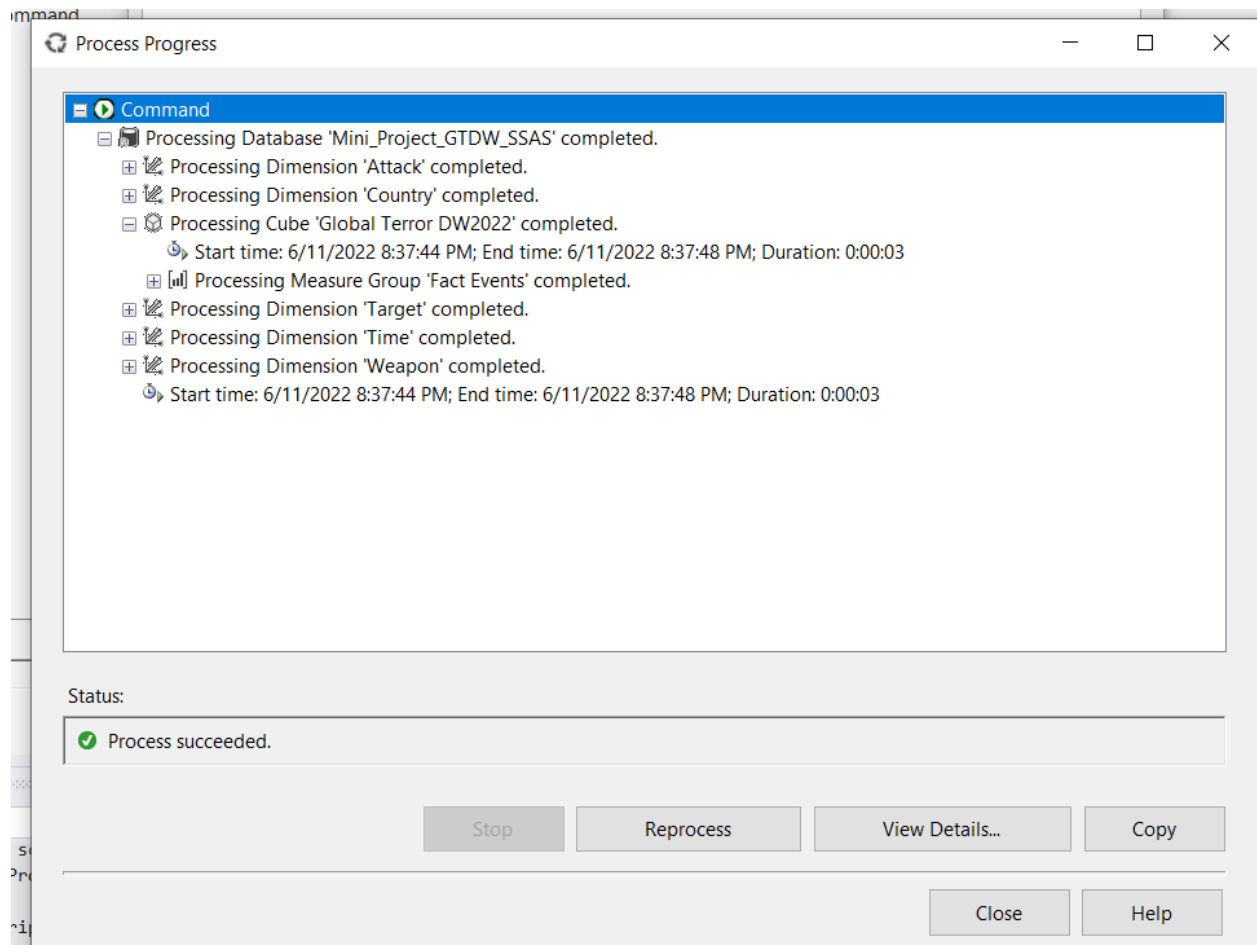
7.3. Multidimensional Models – Cubes



Picture7: Implementation of Cubes for our Project



Picture8: Process Cubes



Picture 9: Successfully Processed Cube

Develop Default Start Live Share ADMIN

Mini_Project_GTDW_SSAS

Attackdim [Design]

Country.dim [Design]

Global Terror DW2022.cube [Design]

Role.role [Design]

Targetdim [Design]

Time.dim [Design]

Weapon.dim [Design]

[Tool Windows]

Deployment Progress..._Project_GTDW_SSAS

Solution Explorer

Cube Structure

Measures

Global Terror DW2022

Fact Events

Dimensions

Global Terror DW2022

Time

Country

Target

Weapon

Attack

Data Source View

target_id
target_category_id
Target_category
Target

factevents
eventid
Group
attack_id
Weapon_id
TimeID
Wounded
Killed
suicide
target_id
success
country_id

attack
attack_id
attacktype1
Attack Type

Weapon
weapon_id
weapon_type_category_id
Weapon

country
country_id
country_code
Country
region
state
city
longitude
latitude

time
year
month
day

Output

Show output from: Build

Process Database Mini_Project_GTDW_SSAS

Done

Sending deployment script to the server...

Done

Deploy complete -- 0 errors, 0 warnings

===== Build: 1 succeeded on up-to-date, 0 failed, 0 skipped =====

===== Deploy: 1 succeeded, 0 failed, 0 skipped =====

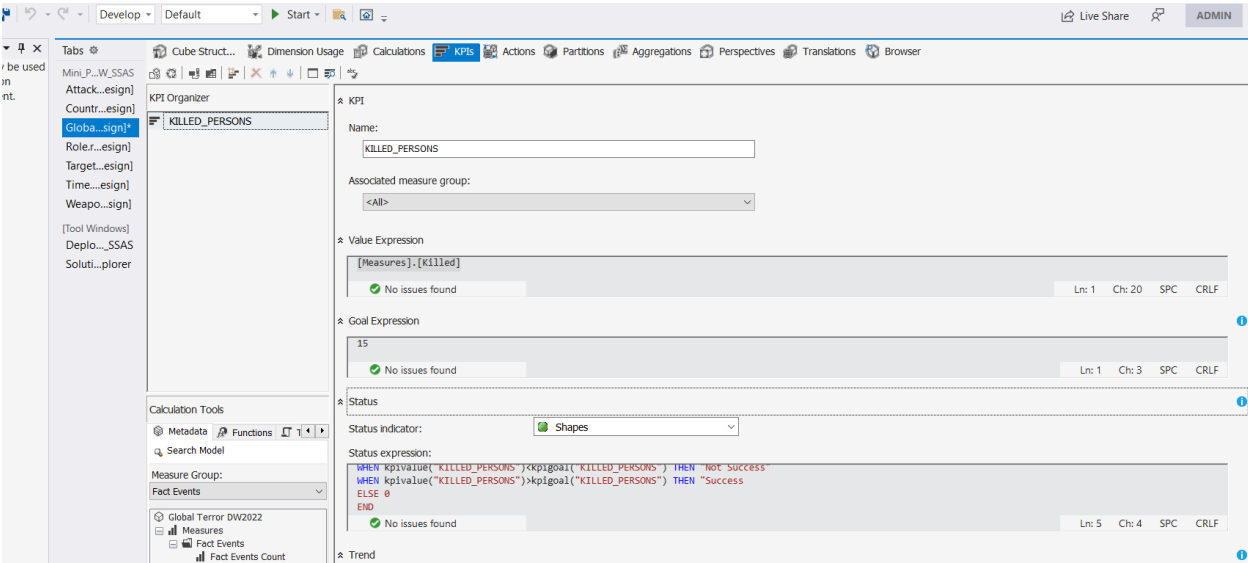
Error List Output

Picture 10: Successfully Deployed Cube

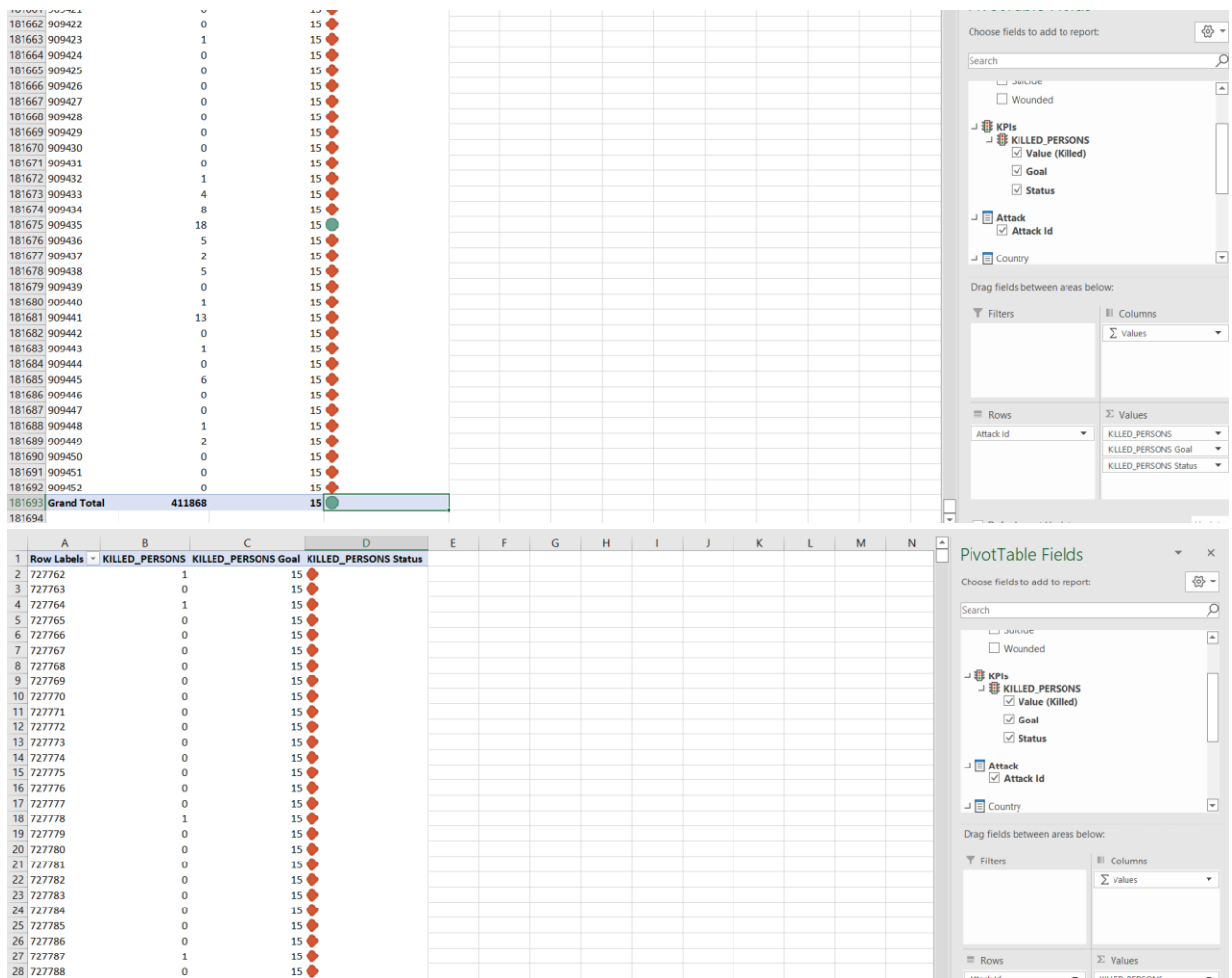
----- The end of the obligatory part of the report documentation -----

8. Data analysis

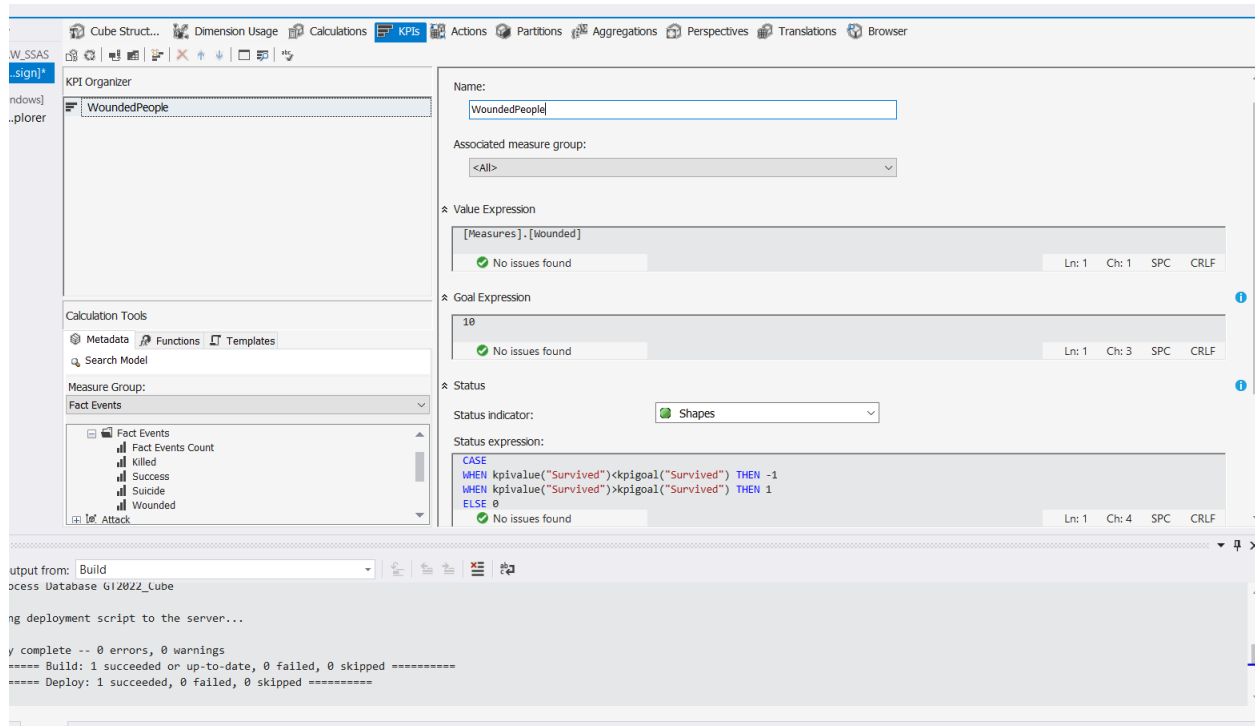
8.1. Implementation of analytical processes



Picture 11: Defining KPI for Killed Persons in Attacks



Pictures 11-12: KPI Results in Excel, Red and Green Shape Indicates us the person killed in attacks are above 15 or not in one attack, as you can see in most attacks people got killed less than 15



Picture 13: Defining KPI For Wounded People, I took goal Expression as 10 when in Fact Table wounded people are more than 10 we will have a green shape and otherwise red shape in our analysis

Explorer

```

SELECT
(
{[Target].[Target Category].MEMBERS}
) ON ROWS
,
{
[Measures].[Killed]
}
ON COLUMNS
FROM [Global Terror DW2022]

```

100 %

Messages Results

	Killed
All	411868
Abortion Related	10
Airports & Aircraft	3767
Business	23487
Educational Institution	3745
Food or Water Supply	313
Government (Diplomatic)	3039
Government (General)	26071
Journalists & Media	1501
Maritime	1191
Military	106047
NGO	1057
Other	255
Police	53704
Private Citizens & Property	140504
Religious Figures/Institutions	13413
Telecommunication	182

DESKTOP-F4C1709\DWBYRAHUL | DESKTOP-F4C1709\rahul | GT2022_Cube | 00:00:01

Picture 14: MDX Query for Killed people in different target attacks over the years

SQL Explorer

```

SELECT
{
[Attack].[Attack Type].MEMBERS
}
ON COLUMNS,
MEASUREGROUPMEASURES( "Fact Events" )
ON ROWS
FROM [Global Terror DW2022]

```

100 %

Messages Results

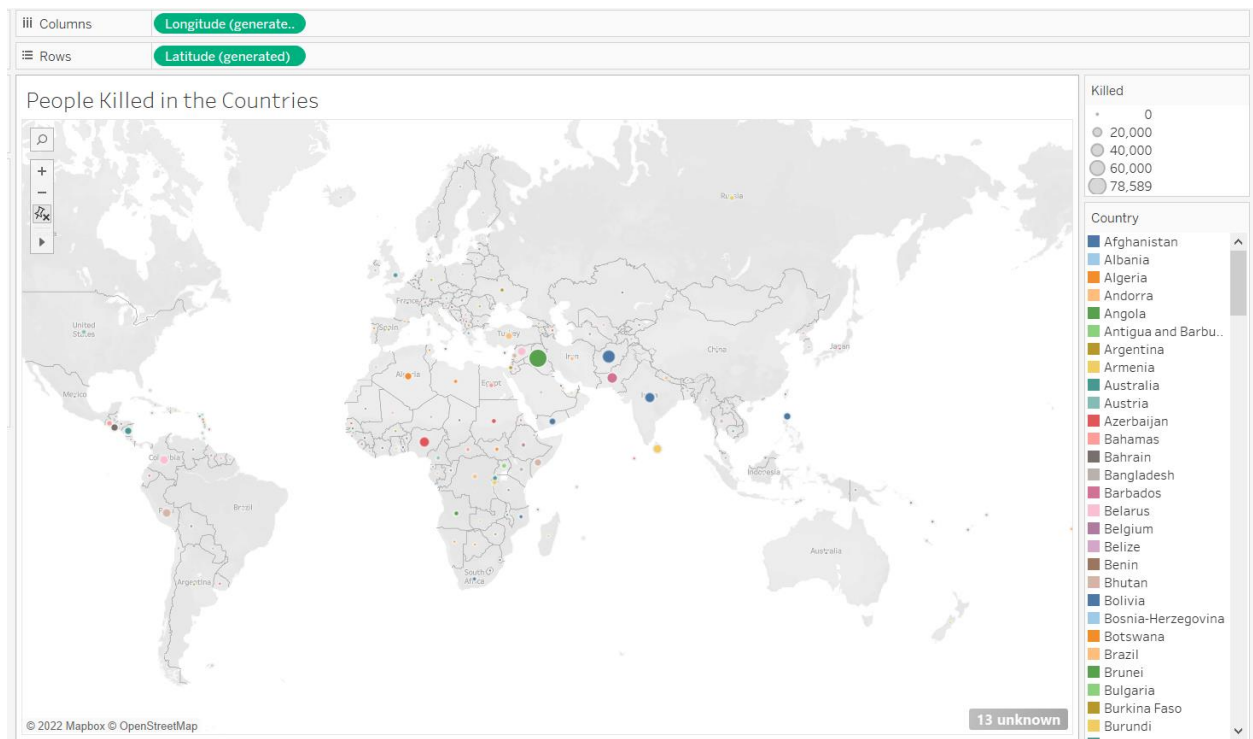
	All	Armed Assault	Assassination	Bombing/Explosion	Facility/Infrastructure Attack	Hijacking	Hostage Taking (Barricade Incident)	Hostage Taking (Barricade Incident)
Wounded	523852	77349	13887	372686	3765	17001	3966	
Killed	411868	160297	24920	157321	3642	3718	4478	
Suicide	6633	49	262	6214	1	11	54	
Success	161632	40345	14615	77530	9779	579	983	
Fact Events Count	181691	42669	19312	88255	10356	659	991	

DESKTOP-F4C1709\DWBYRAHUL | DESKTOP-F4C1709\rahul | GT2022_Cube | 00:00:01

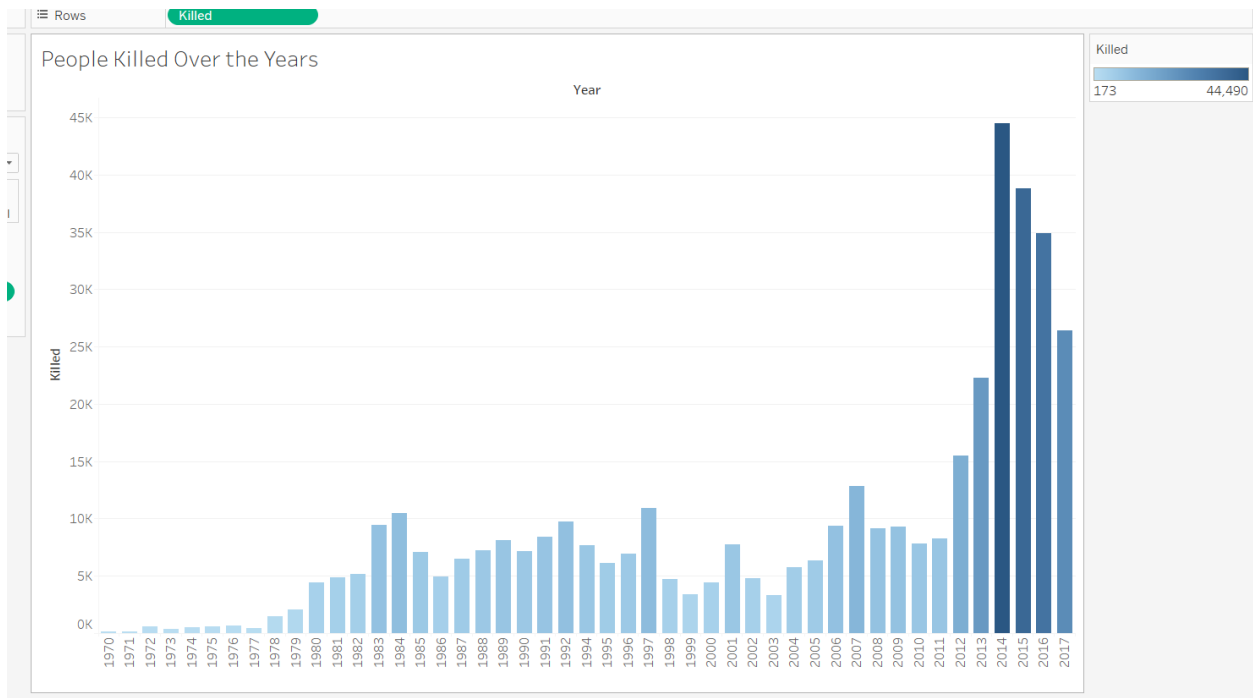
Picture 15: MDX Query for Group of Measures in people affected in types of attacks



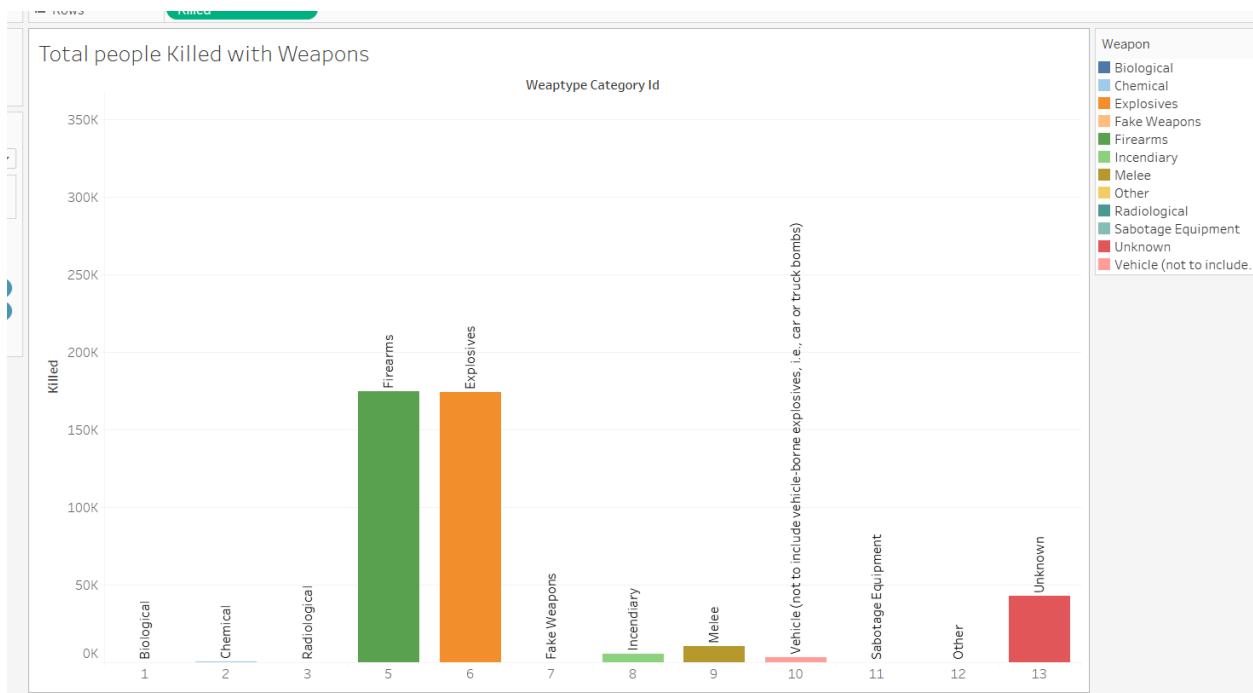
Picture 16: Countries with the most attacks plot, as you can see list of countries affected with the terror attacks



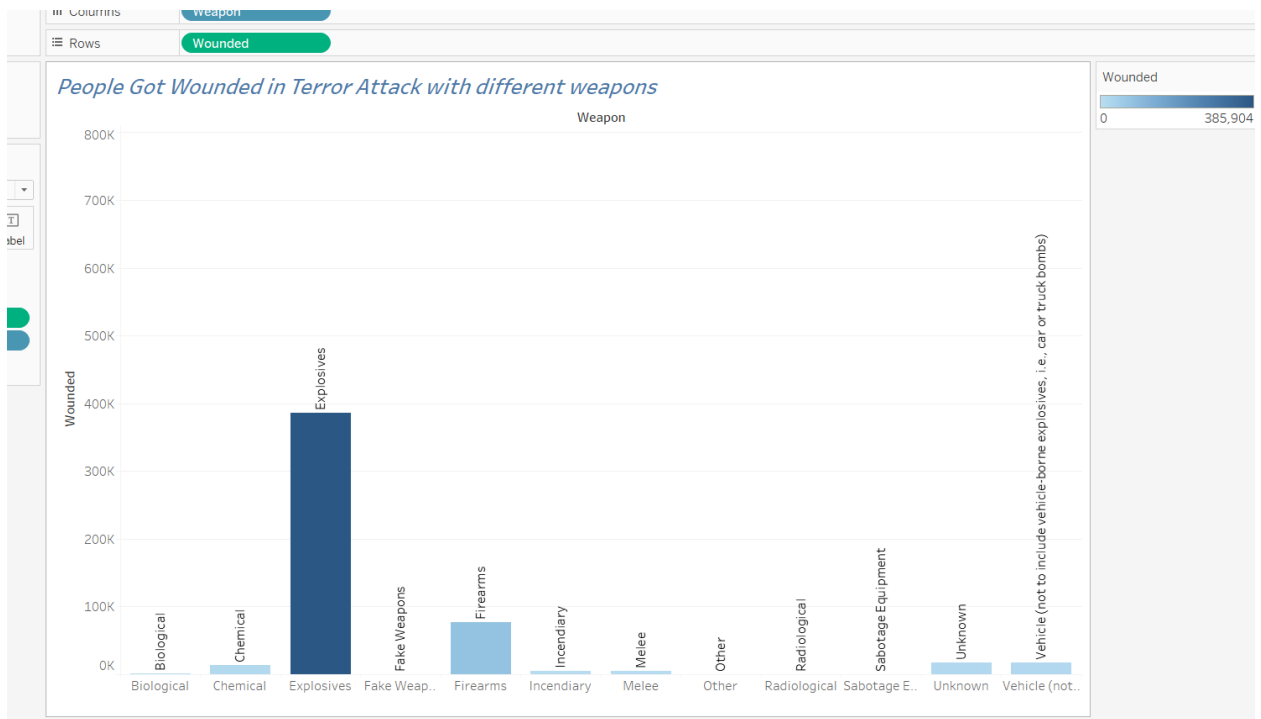
Picture 17: Countries with Most People Killed, as you can here countries who lost people in terror attacks



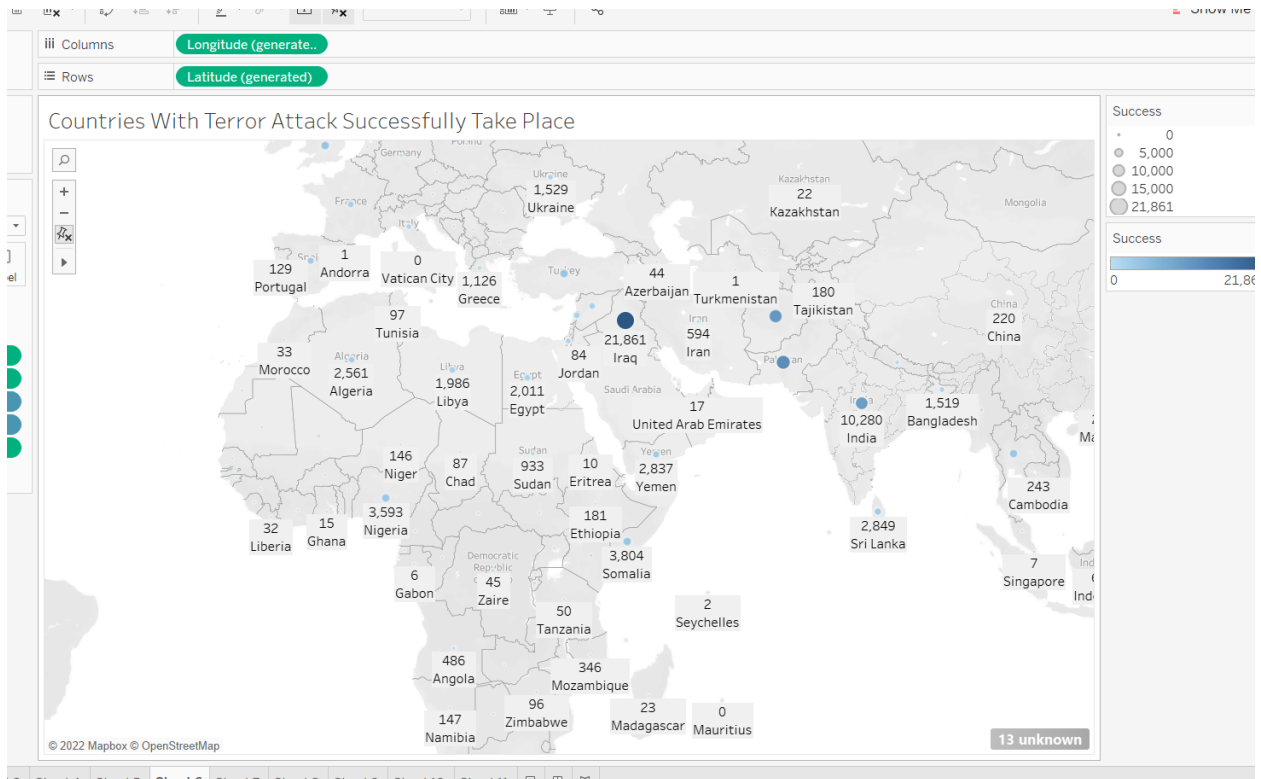
Picture 18: Total People Killed in Terror Attacks Over the years, as you can see a bar plot for total people killed in these attacks over the period of time



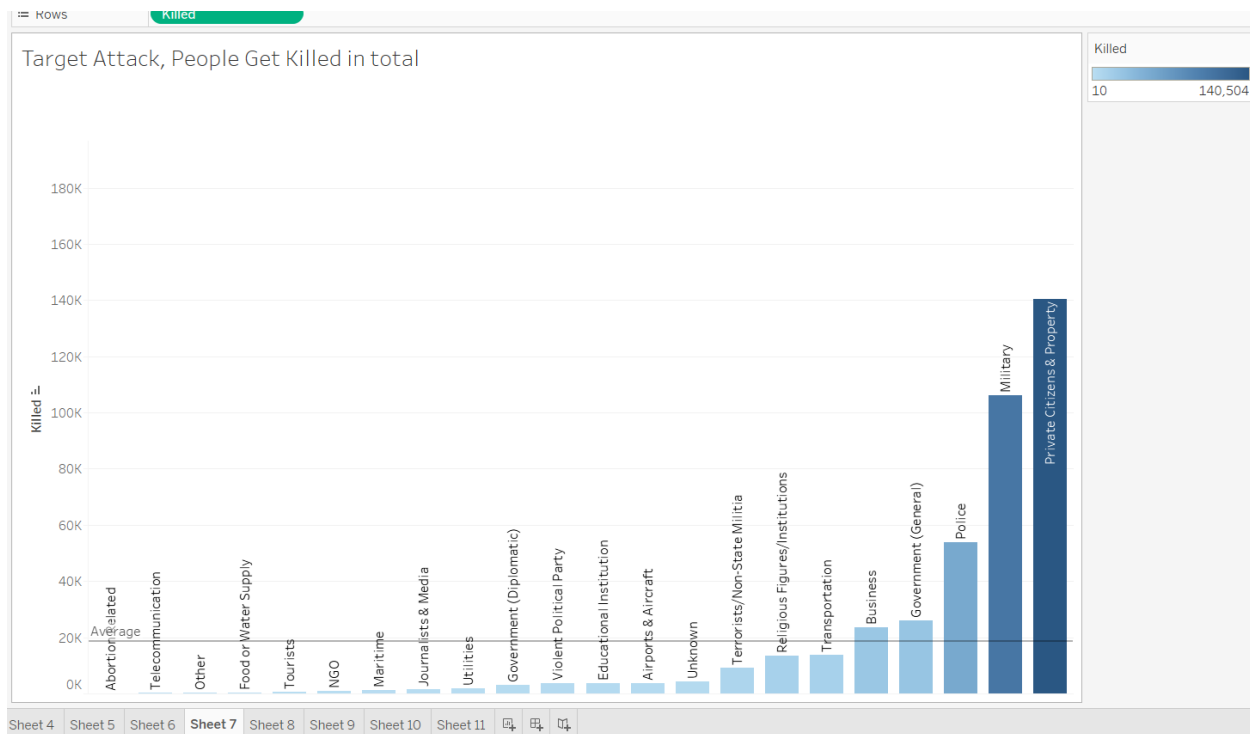
Picture 19: People Killed with different Weapons, bar plot for weapon use in terror attacks



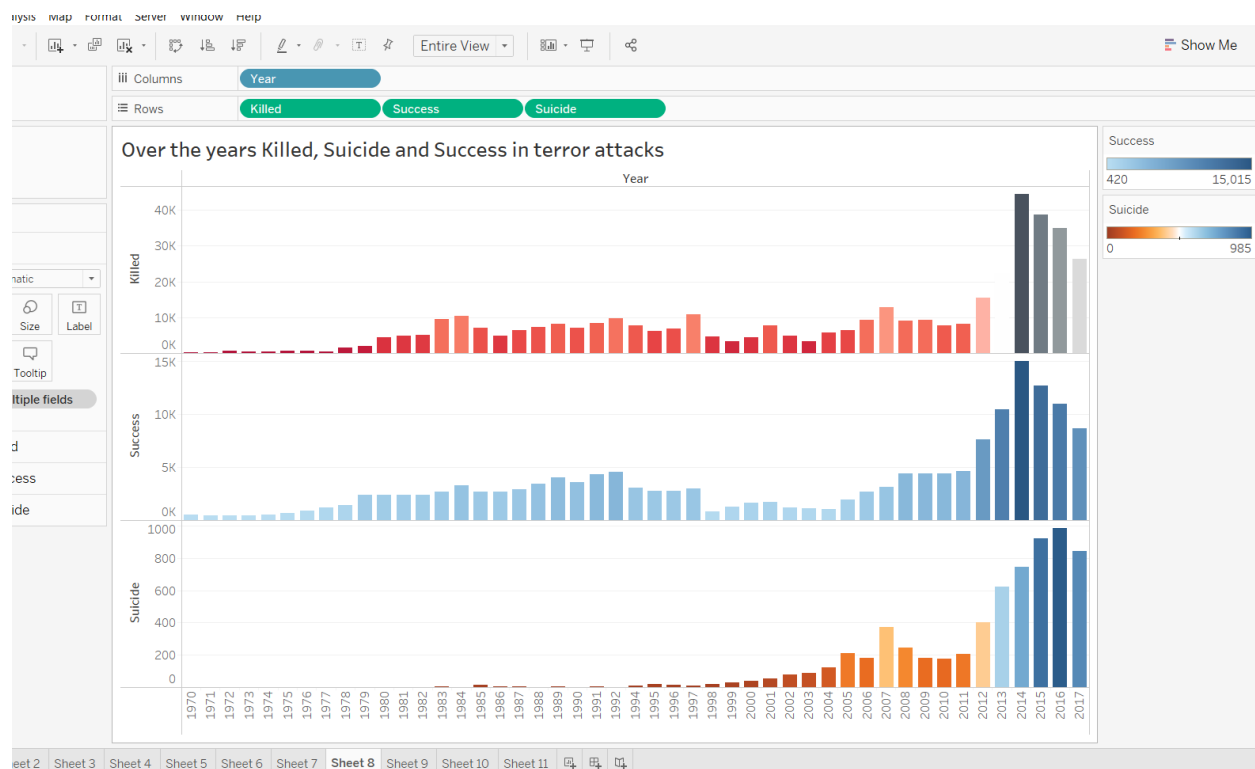
Picture 20: People Got wounded with different Weapon, Bar plot for people got wounded with different weapons



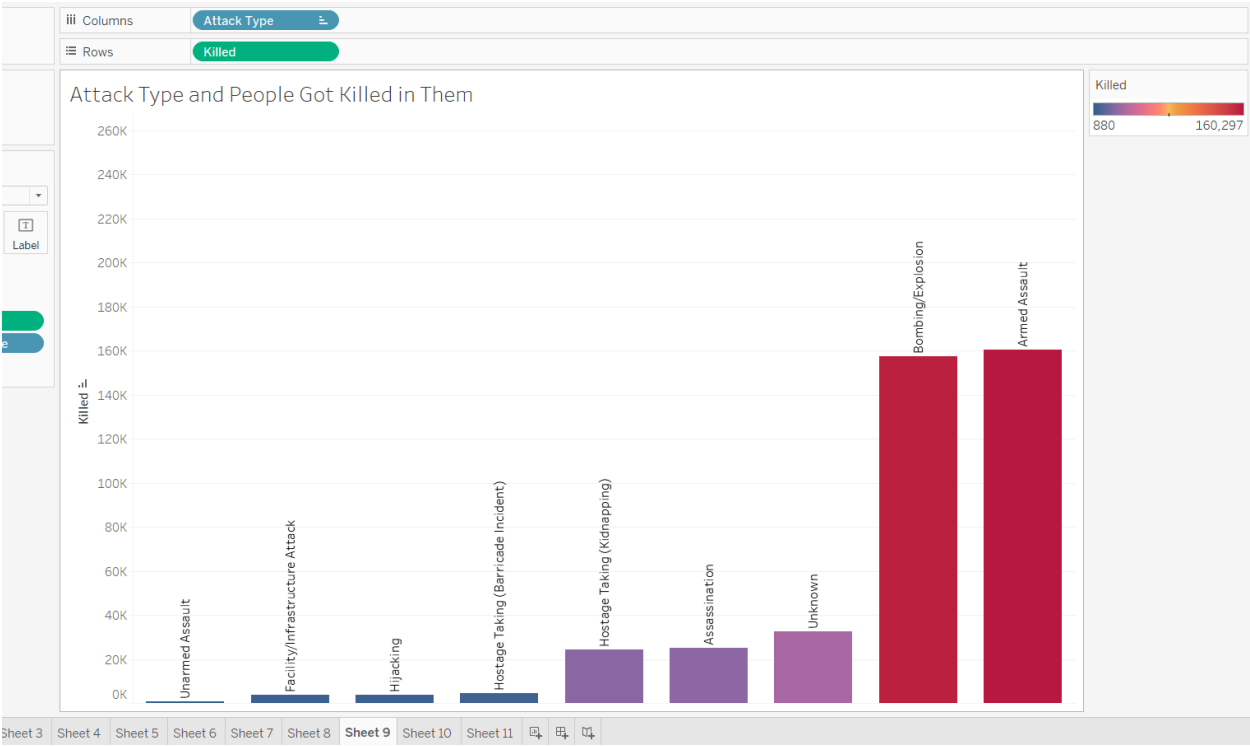
Picture 21: Countries With Terror Attacks take place successfully



Picture 22: Target Attack and Total People got killed in it, there is a average line in the plot which indicate us 20k average die in target attacks

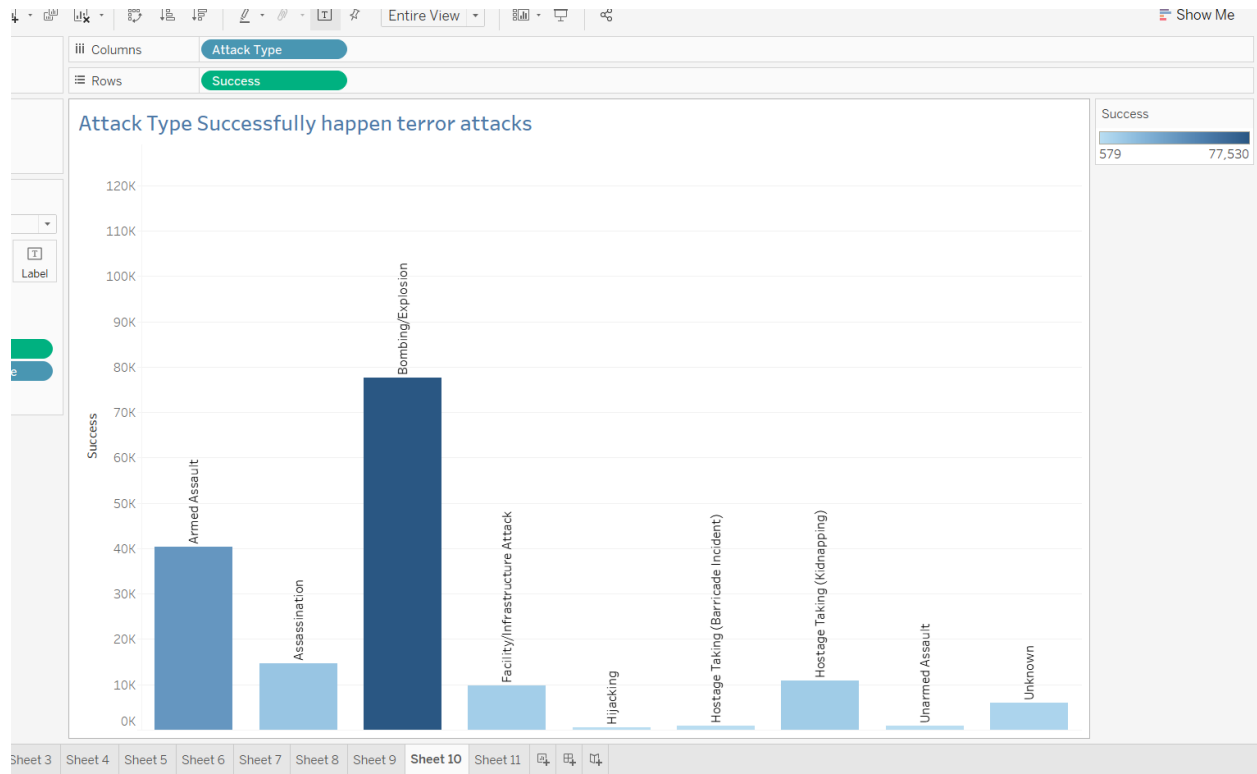


Picture 23: Comparison of Killed, Suicide, Success of attacks over the years, bar plot comparison

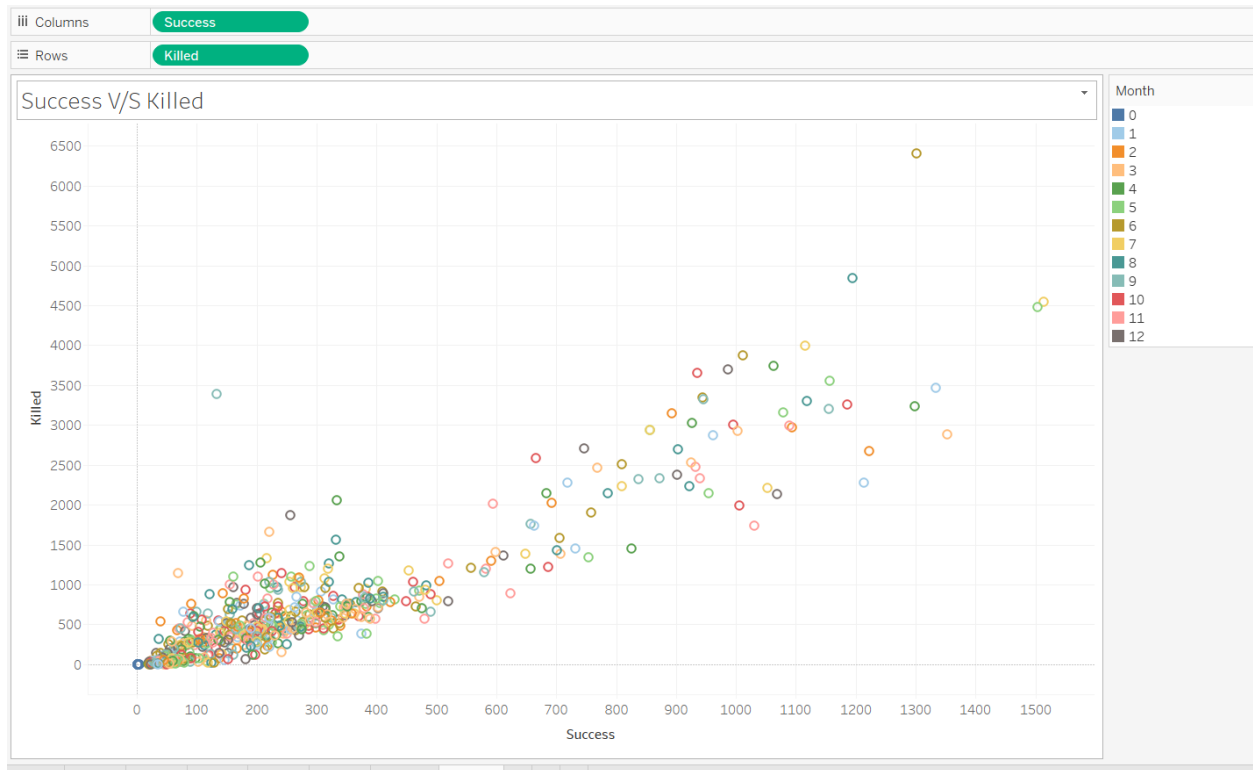


Picture 24: Attack Type and People got killed in them, bar plot

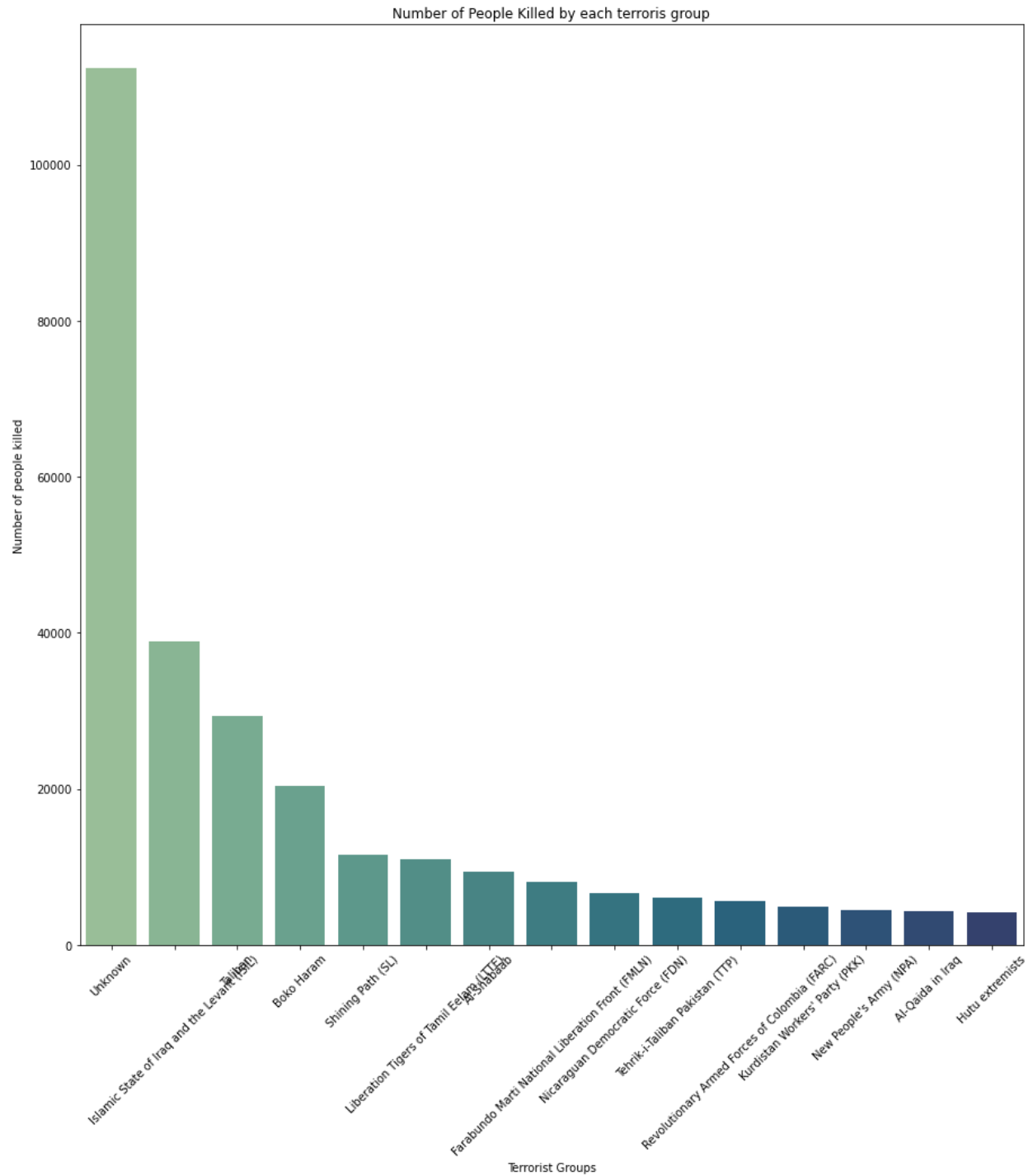
As you can see bombing and armed assault killed the most people



Picture 25: Attack Type Successfully take place, Explosive Attacks Took place successfully



Picture 26: Scatter plot of killed v/s Success over the years



Picture 27: No. of people Killed by Different Terror Groups

Source: Python

8.2. Summary - conclusions from the analysis

As you can see, we have completed our data analysis and successfully deployed ETL and Cubes for Business Intelligence, we can see the after analysis I have created following tableau below in order to explain the following outcome of our project:

Country with most attacks:	Iraq
City with most attacks:	Baghdad
Region with the most attacks:	Middle East & North Africa
Year with the most attacks:	2014
Month with the most attacks:	5
Group with the most attacks:	Taliban
Most Attack Types:	Bombing/Explosion

Terror Groups Killed Total People	
Group	Killed
Unknown	112367
Islamic State of Iraq and the Levant (ISIL)	38923
Taliban	29410
Boko Haram	20328
Shining Path (SL)	11601
Liberation Tigers of Tamil Eelam (LTTE)	10989
Al-Shabaab	9330
Farabundo Marti National Liberation Front (FMLN)	8065
Nicaraguan Democratic Force (FDN)	6662
Tehrik-i-Taliban Pakistan (TTP)	6042

Revolutionary Armed Forces of Colombia (FARC)	5661
Kurdistan Workers' Party (PKK)	4955
New People's Army (NPA)	4386
Al-Qaida in Iraq	4381
Hutu extremists	4102

Total Attacks by Terror Groups	
Taliban	7478
Islamic State of Iraq and the Levant (ISIL)	5613
Shining Path (SL)	4555
Farabundo Marti National Liberation Front (FMLN)	3351
Al-Shabaab	3288
New People's Army (NPA)	2772
Irish Republican Army (IRA)	2671
Revolutionary Armed Forces of Colombia (FARC)	2487
Boko Haram	2418
Kurdistan Workers' Party (PKK)	2310
Basque Fatherland and Freedom (ETA)	2024
Communist Party of India - Maoist (CPI-Maoist)	1878
Maoists	1630
Liberation Tigers of Tamil Eelam (LTTE)	1606

Total type of attacks	
Bombing/Explosion	88255
Armed Assault	42669
Assassination	19312
Hostage Taking (Kidnapping)	11158

Facility/Infrastructure Attack	10356
Unknown	7276
Unarmed Assault	1015
Hostage Taking (Barricade Incident)	991
Hijacking	659

Cities with most attacks	
Unknown	9775
Baghdad	7589
Karachi	2652
Lima	2359
Mosul	2265
Belfast	2171
Santiago	1621
Mogadishu	1581
San Salvador	1558
Istanbul	1048
Athens	1019
Bogota	984
Kirkuk	925
Beirut	918
Medellin	848

Countries with most Attacks	
Iraq	24636
Pakistan	14368
Afghanistan	12731
India	11960
Colombia	8306
Philippines	6908
Peru	6096
El Salvador	5320
United Kingdom	5235
Turkey	4292
Somalia	4142
Nigeria	3907
Thailand	3849
Yemen	3347
Spain	3249

9. Final conclusions from project implementation

9.1. The Problems

During the implementation of the project, I encountered two major problems and, of course, a whole bunch of smaller ones. The bigger problems include the conceptual model and the ETL process. The conceptual model, as I later found out, was one of the most important stages of the project, but it took me quite a long time to separate the attributes into appropriate tables so that the future warehouse it made sense. However, I cannot compare the difficulty level of the conceptual model with the difficulty level of the ETL process. The concept itself took me 2 days, and the implementation another 2. Even though we had classes in the ETL process, quite a large abstraction was the combination of subsequent processes into a whole. I know we don't have much time in the semester, but I think one more hour of ETL classes could speed up my work.

9.2. Acquired knowledge and experience

I must admit that I really learned a lot while implementing the project! Even though ETL took me 4 days, thanks to the fact that I solved each problem myself, I understand this process better. Besides, ... Plus, I really enjoyed working with Tableau.

Data Analysis is fun, I learned a lot about Multidimensional Models, MDX Query, ETL Workflow and Data Flow Constraints, Implementation of Data Warehouse, I use Python, SQL, SSIS and SSAS, Tableau, Tableau Prep, I learned all of them in this laboratory.

10. Information sources used in the data analysis stage

During the analysis of the data, to better understand the analyzed charts, as well as to draw appropriate conclusions, I used the following sources:

[Global Terrorism Database](#)

[Global Terrorism Database | Kaggle](#)

[Tools for Data Analysis and Staging:](#)

Python, SQL, SSIS and SSAS, Tableau, Tableau Prep, Visual Studio Code, Visual Studio, YouTube, Kaggle