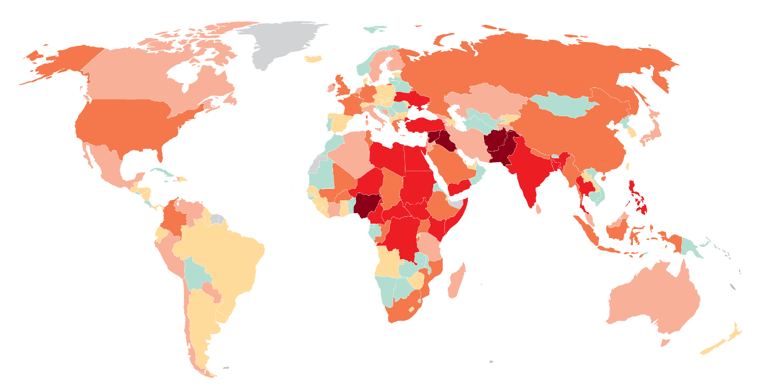
University of Houston-Clear Lake

Data Warehousing and Data Mining

ISAM 5332-Fall 2019



**Submitted By: - Greeshma Srinivasa Raju**

**Kavitha Sundar Ram**

**Raghavi Vijayakumar**

CONTENTS:

ABSTRACT-------------------------------------------------------------------------------------------------1

PROJECT OVERVIEW-----------------------------------------------------------------------------------1

INTROCUCTION TO DATAWAREHOUSING AND DATA MINING--------------------------1

NEED FOR DATAWAREHOUSING APPLICATION----------------------------------------------2

METHODOLOGY-----------------------------------------------------------------------------------------3

DATA CLEANSING--------------------------------------------------------------------------------------5

DATA TRANSFER TO MS ACCESS-----------------------------------------------------------------11

PROJECT DATASET------------------------------------------------------------------------------------32

DIMNENSIONAL MODELLING----------------------------------------------------------------------32

CUBE DEVELOPMENT AND DEPLOYMENT IN VISUAL STUDIO-------------------------41

SUCCESSFUL DATA TRANSFER TO SQL SERVER---------------------------------------------43

REPORTS---------------------------------------------------------------------------------------------------49

CONCLUSION---------------------------------------------------------------------------------------------65

REFERENCES----------------------------------------------------------------------------------------------65

**ABSTRACT:**

A **data warehouse** is constructed by integrating **data** from multiple heterogeneous sources that support analytical reporting, structured and/or ad hoc queries, and decision making. The Global Terrorism Database (GTD) is an open-source database including information on terrorist attacks around the world from 1970 through 2015.It includes systematic data on domestic as well as international terrorist incidents that have occurred during this time period and now includes more than 180,000 attacks. The system builds Global Terrorism data warehouse

using Microsoft SQL server database, Cube view and data source view.

**PROJECT OVERVIEW:**

Global Terrorism dataset had records of terrorist attacks that happened all over the world between the years 1970-2015. Given the data, we could calculate the tremendous impact it would leave the places it affected. Analysis on when the occurrences of these attacks were frequent were performed. Further analysis were concentrated on the type of attack on the type of targets that were successful. Countries such as Afghanistan, Indonesia, India, and Iraq that have been severely impacted by terrorism have been analyzed on what time of the year they were attacked the most.

**Introduction to Data Warehousing and Data mining**

**Data Warehousing:**

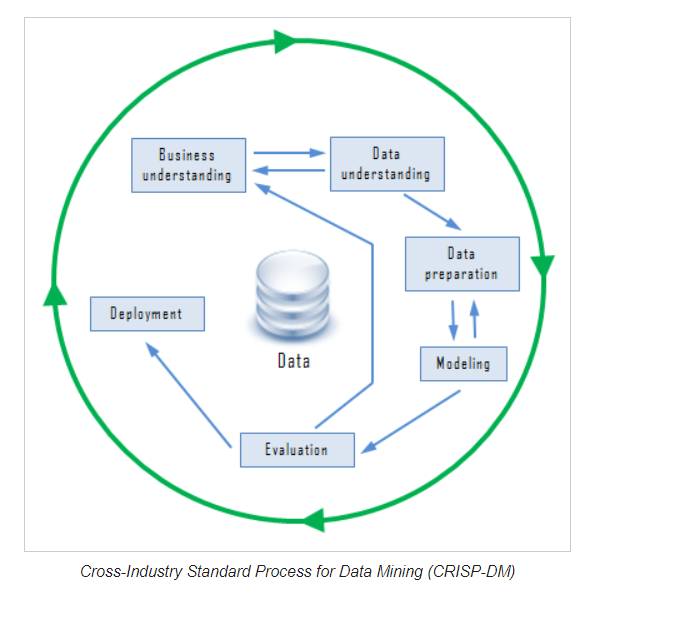
It is a technology that aggregates structured data from one or more sources so that it can be compared and analyzed rather than transaction processing. A **data warehouse** is designed to support management decision-making process by providing a platform for data cleaning, data integration and data consolidation. A data warehouse contains subject-oriented, integrated, time-variant and non-volatile data.

Data warehouse consolidates data from many sources while ensuring data quality, consistency and accuracy. Data warehouse improves system performance by separating analytics processing from transnational databases.

[**Data Mining**](https://www.geeksforgeeks.org/data-mining/)**:**  
Data mining is the process of finding anomalies, patterns and correlations within large data sets to predict outcomes. Using a broad range of techniques, you can use this information to increase revenues, cut costs, improve customer relationships, reduce risks and more. Data mining tools are used to build risk models and detect fraud. Data mining is used in market analysis and management, fraud detection, corporate analysis and risk management.

Some most important reasons for using Data mining are:

* Establish relevance and relationships amongst data. Use this information to generate profitable insights
* Business can make informed decisions quickly
* Helps to find out unusual shopping patterns in grocery stores.
* Optimize website business by providing customize offers to each visitor.
* Helps to measure customer's response rates in business marketing.
* Creating and maintaining new customer groups for marketing purposes.
* Predict customer defections, like which customers are more likely to switch to another supplier in the nearest future.
* Differentiate between profitable and unprofitable customers.
* Identify all kind of suspicious behavior, as part of a fraud detection process.



**The Need for Data Warehousing and Data Mining**

**Some most Important reasons for using Data warehouse are:**

* Integrates many sources of data and helps to decrease stress on a production system.
* Optimized Data for reading access and consecutive disk scans.
* Data Warehouse helps to protect Data from the source system upgrades.
* Allows users to perform master Data Management.
* Improve data quality in source systems.

**Some most important reasons for using Data mining are:**

* Establish relevance and relationships amongst data. Use this information to generate profitable insights
* Business can make informed decisions quickly
* Helps to find out unusual shopping patterns in grocery stores.
* Optimize website business by providing customize offers to each visitor.
* Helps to measure customer's response rates in business marketing.
* Creating and maintaining new customer groups for marketing purposes.
* Predict customer defections, like which customers are more likely to switch to another supplier in the nearest future.
* Differentiate between profitable and unprofitable customers.
* Identify all kind of suspicious behavior, as part of a fraud detection process.

**Functions of Data Warehouse Tools and Utilities**

The following are the functions of data warehouse tools and utilities:

* **Data Extraction** − Involves gathering data from multiple heterogeneous sources.
* **Data Cleaning** − Involves finding and correcting the errors in data.
* **Data Transformation** − Involves converting the data from legacy format to warehouse format.
* **Data Loading** − Involves sorting, summarizing, consolidating, checking integrity, and building indices and partitions.
* **Refreshing** − Involves updating from data sources to warehouse.

**Note** − Data cleaning and data transformation are important steps in improving the quality of data and data mining results.

1. **Methodology**

**3.1 Raw Data Definition and Format**

The Global Terrorism Database GTD is the most comprehensive unclassified database of terrorist attacks in the world. The GTD is produced by a dedicated team of researchers and technical staff, who are supported by an alumni of researchers.  The governance of the database is assured by an advisory board which includes some of the most respected academics in the field of terrorism research.

**Column name:** Year

**Definition:** This field shows the year in which the incident occurred.

**Data Type:** Numeric

**Value:** 1970 to 2015

**Column name:** Month

**Definition:** This field contains the details of the month in which the incident occurred.

**Data Type:** Text

**Column name:** Day

**Definition:** This field contains the numeric day of the month in which the incident occurred.

**Data Type:** Numeric

**Column name:** Country

**Definition:** This field identifies the country or location where the incident occurred.

**Data Type:** Text

**Column name:** State

**Definition:** Name (at the time of event) of the 1st order subnational administrative region.

**Data Type:** Text

**Column name:** City

**Definition:** Name of the city, village, or town in which the incident occurred.

**Data Type:** Text

**Column name:** Success

**Definition:** This provides information on the success of a terrorist strike.

**Data Type:** Numeric

**Column name:** Suicide

**Definition:** This field has two values 1 which means the incident was a suicide attack. 0 which means there is no indication that the incident was a suicide attack.

**Data Type:** Numeric

**Values:** 1,0

**Column name:** Attack Type

**Definition:** The general method of attack and broad class of tactics used during the occurrence of the incident.

**Data Type:** Text

**Values:** Armed Assault, Assassination, Hijacking, Bombing/Explosion, Hostage Taking (Kidnapping), Facility/Infrastructure Attack, Hostage Taking (Barricade Incident), Unarmed Assault.

**Column name:** Target Type

**Definition:** The general type of target/victim who were affected during the incident.

**Data Type:** Text

**Value:** Police, Airports & Aircraft, Government (Diplomatic), Business, Terrorists/Non-State Militia, Private Citizens & Property, Government (General), Telecommunication, Educational Institution, Military, Maritime, Tourists, Transportation, Religious Figures/Institutions, Violent Political Party, Journalists & Media, Utilities, Abortion Related, Food or Water Supply, NGO, Other.

**Column name:** Nationality Type

**Definition:** This field provides details of the nationality of the target that was attacked.

**Data Type:** Text

**Column name:** Weapon Type

**Definition:** This field has the details of General type of weapon used in the incident.

**Data Type:** Text

**Values:** Firearm, Explosives, Incendiary, Vehicle, Fake Weapons, Melee, Biological, Chemical, Radiological, Radiological, Sabotage Equipment, Other

**Column name:** Number Killed

**Definition:** The number of total confirmed fatalities for the incident

**Data Type:** Text

**Data Type:** Numeric

**3.2 Data Cleansing**

The raw data was collected from the Kaggle Website. Some data is irrelevant, repeated, missing, inconsistent and consists of null values. Data cleansing was necessary to achieve clean and accurate data to work with. This is where corrupt or inaccurate records are identified and replaced.

**Issues in the Data and Rectification of the Data:**

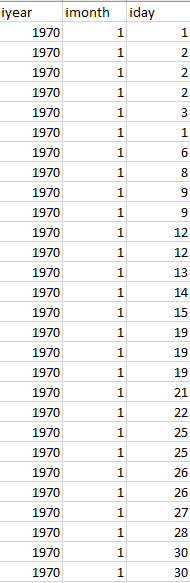
1. **Inconsistent data (Time details**)**:**

Data types and values not consistent in the CSV source file received.

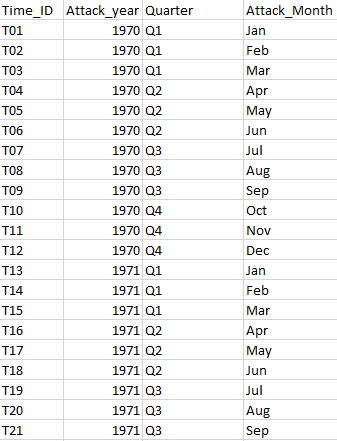
**Resolved:** Using consistent data types and meaningful values into a consolidated excel file.

**Example:**

**Uncleaned data:** We need to use a unique ID for Time details. Raw data had year, month and day with repeated values. We removed day since it was irrelevant. We combined year and month into a unique ID called Time\_ID.



**Cleaned data:**  Time\_ID is the primary key which is defined in Attack\_Time\_Details.



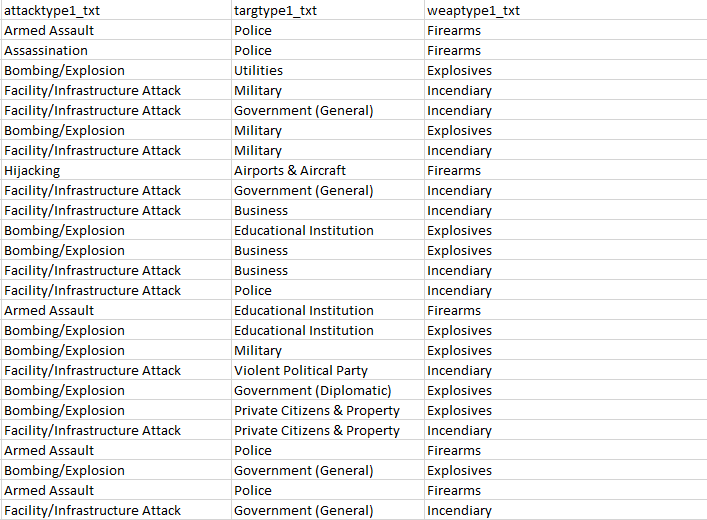
1. **Inconsistent data(Attack Details):**

Data types and values not consistent in the CSV source file received.

**Resolved:** Using consistent data types and meaningful values into a consolidated excel file.

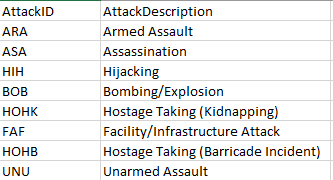
**Example:**

**Uncleaned data:** Attack\_type1, Target\_type and weapon\_type were repeated. we need to make it unique.

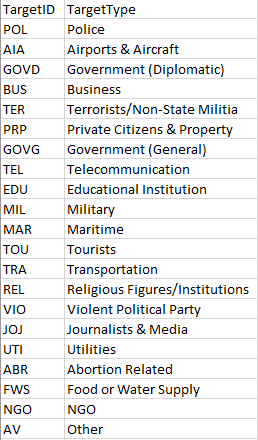


**Cleaned data:**  We created 3 dimension tables for Attack\_type1, Target\_type and weapon\_type with each table having a unique primary key for the values in the tables respectively.

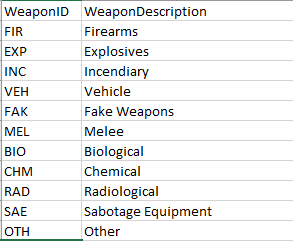
**Attack\_type1:**



**Targettype\_1:**



**Weapontype\_1:**



1. **Weakly related attributes:**

Non-key weakly related attributes found that do not generate any meaningful information.

**Resolved:** Weakly related attributes removed so that the data is better organized.

**Example:**

**Uncleaned data:** Weak attributes such as **region, region\_txt** and **provstate** carried no relevance.



**Cleaned data: region, region\_txt** and **provstate** were removed for consolidation.

1. **Missing data for certain non-key attributes:**

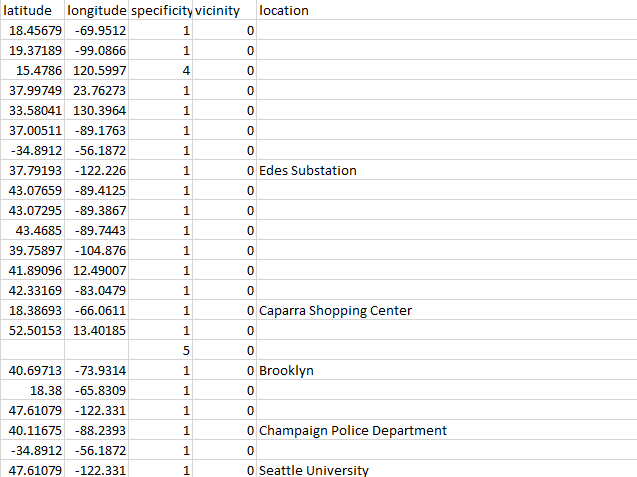
Data was missing in many records for the fields required.

**Resolved:** By including the correct and relevant data for the said fields or removing the said fields.

**Example:**

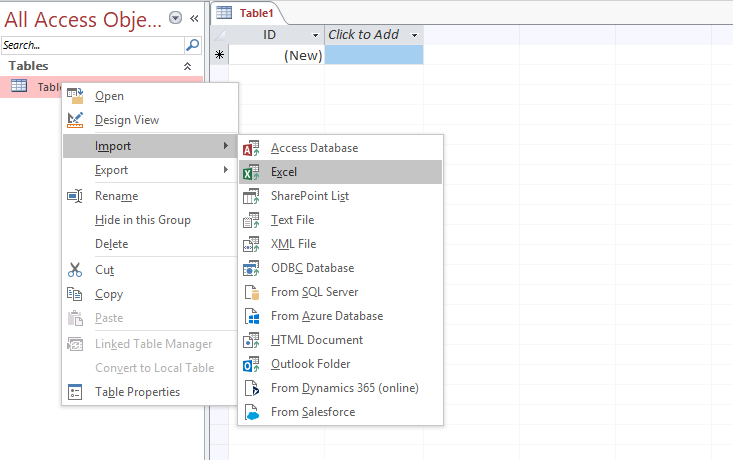
**Uncleaned date:** Location column has empty values.

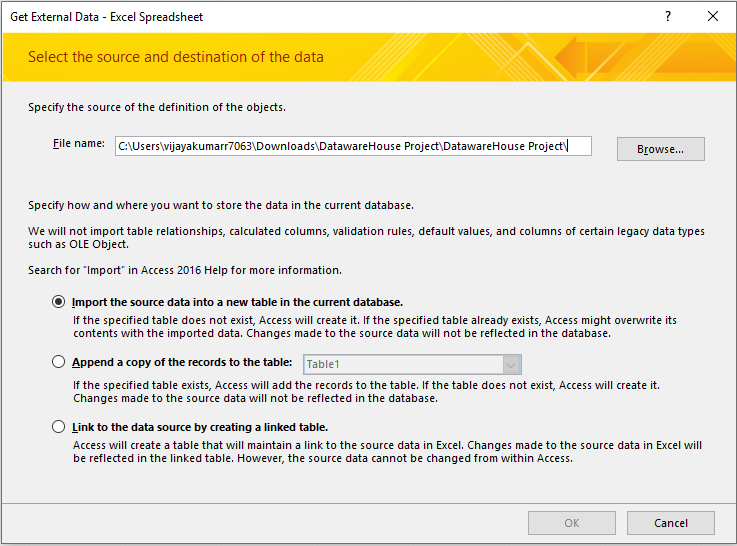
**Cleaned data:** Deleted records with empty values from the location column. Latitude, Longitude, Specificity, vicinity was removed since it’s not relevant in the context.

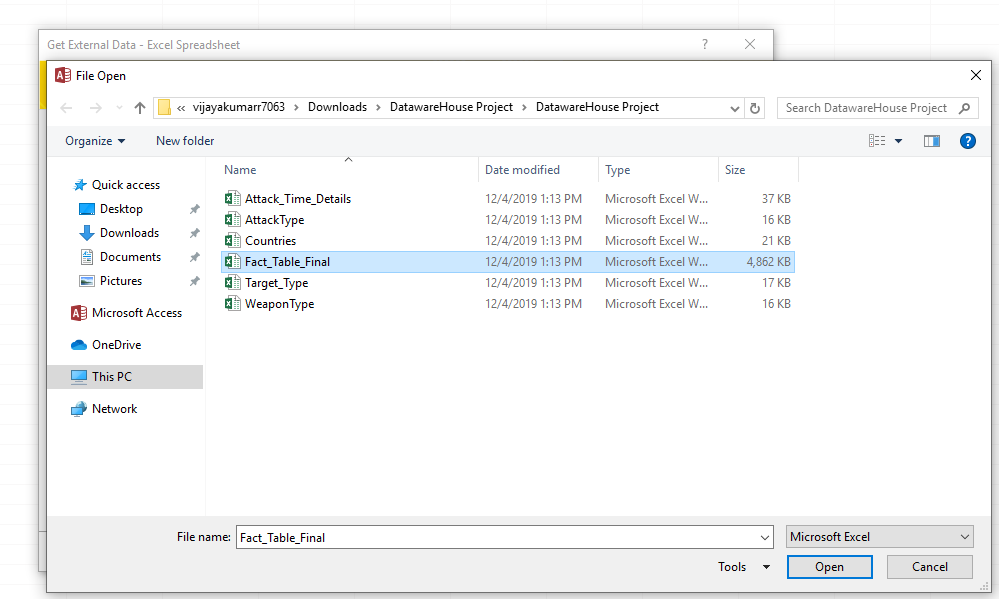


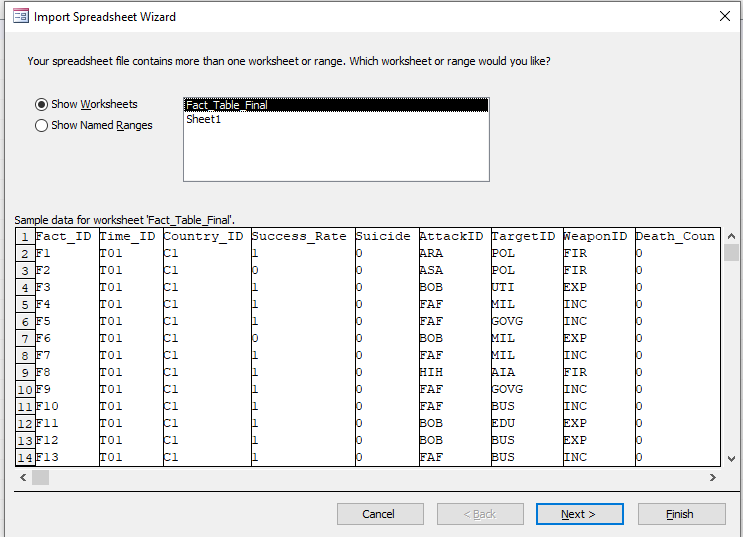
**Data Transfer to MS Access:**

1. Import Excel file **“Fact Table”** into Access

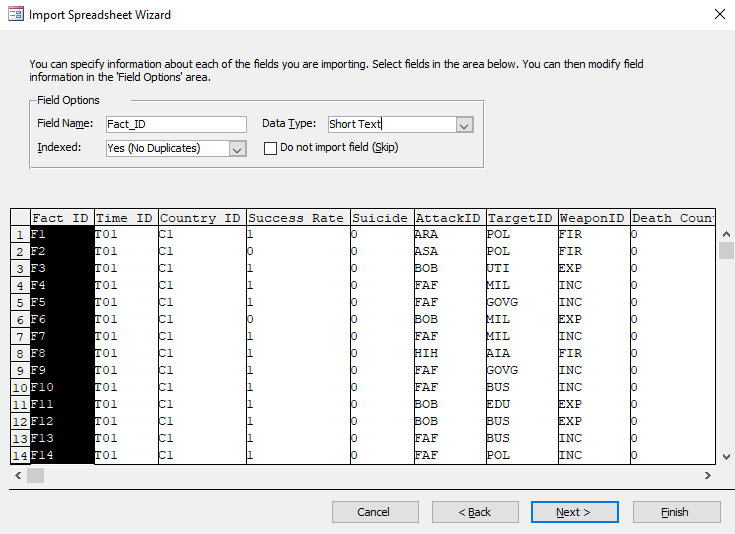




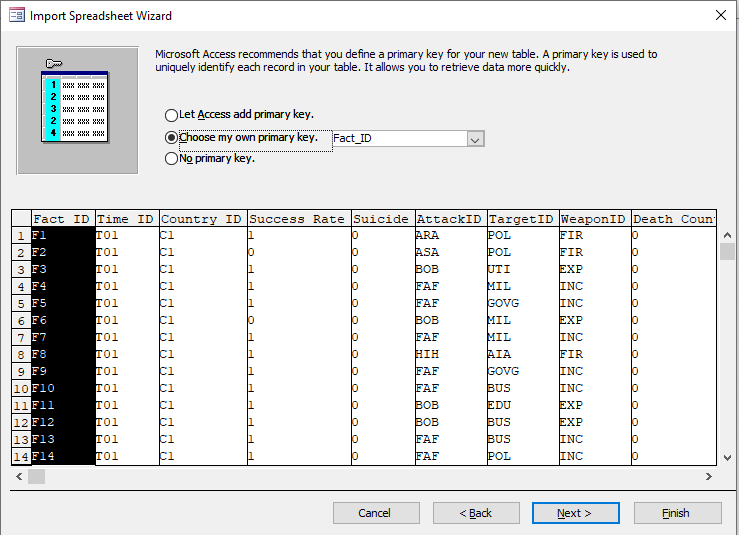




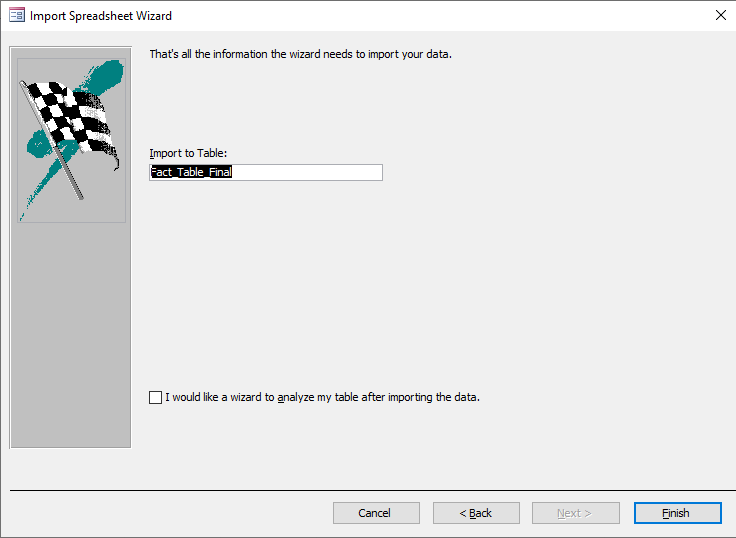
1. Choose **“Yes (No duplicates)”** for the chosen primary key of the fact table.



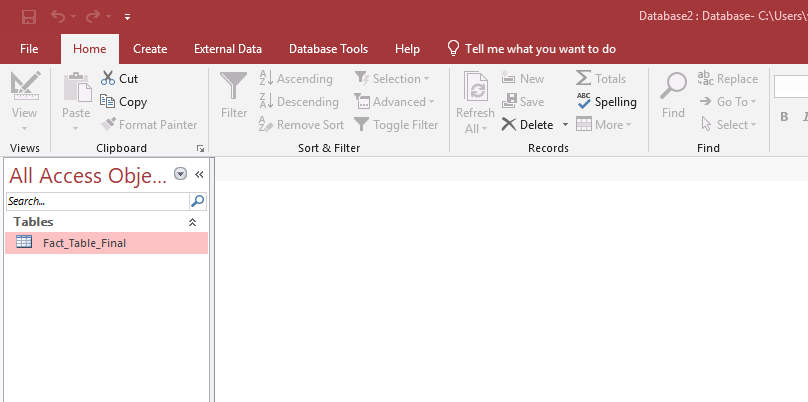
1. Choose Fact\_ID as a primary Key for the fact table.



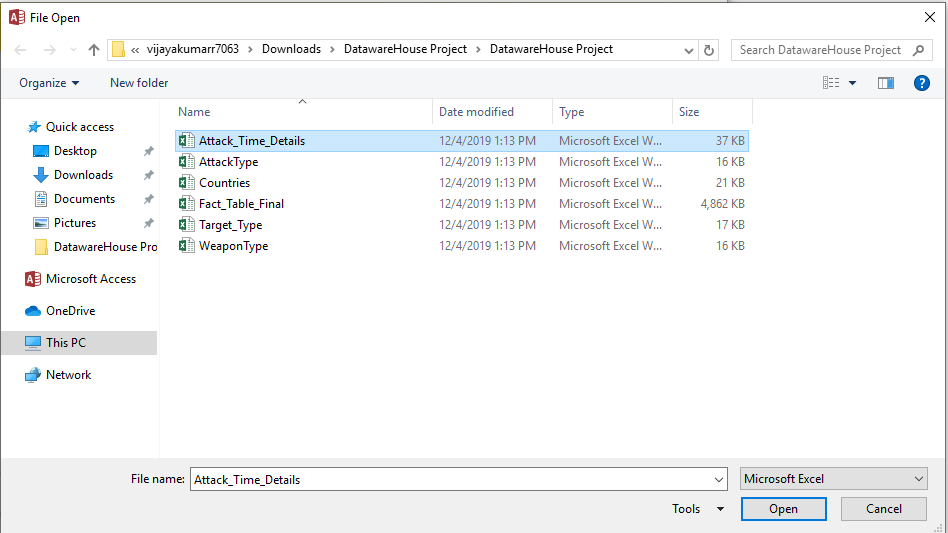
1. Naming the file being exported as Fact\_Table\_Final

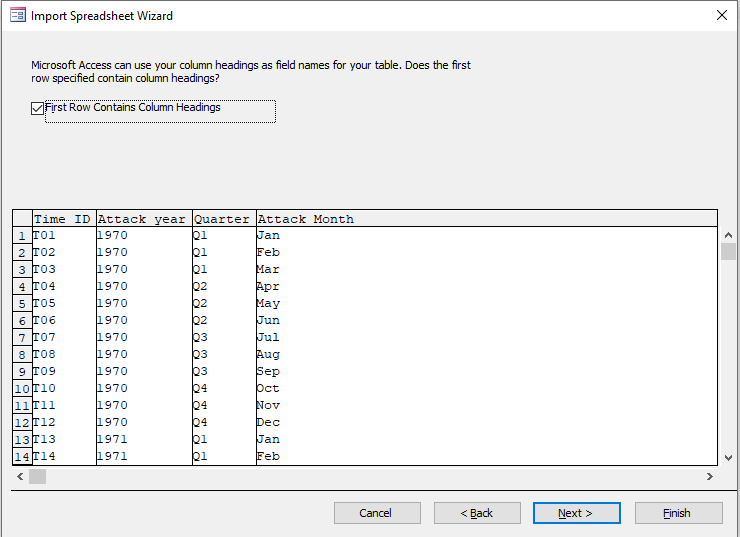


1. Fact\_Table\_Final is being exported successfully as seen in the screenshot below

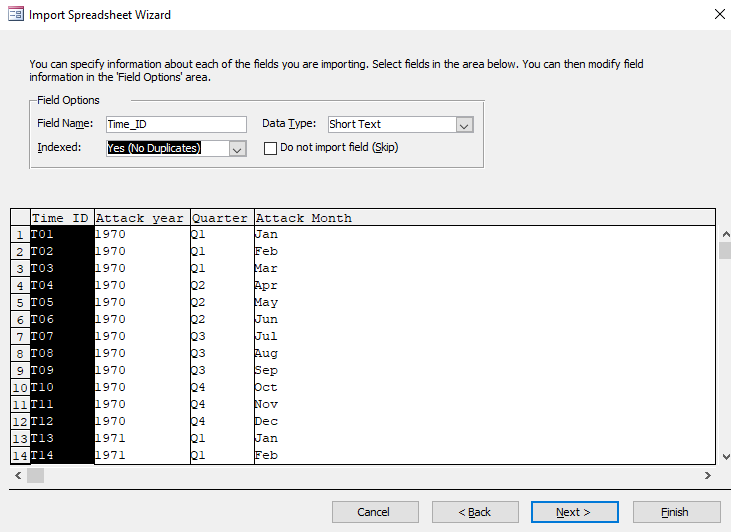


1. Import Excel file “Attack\_Time\_Details” into Access

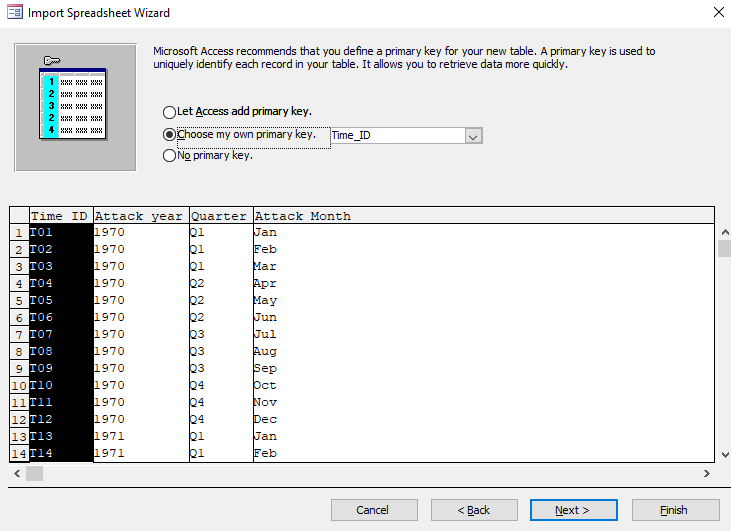




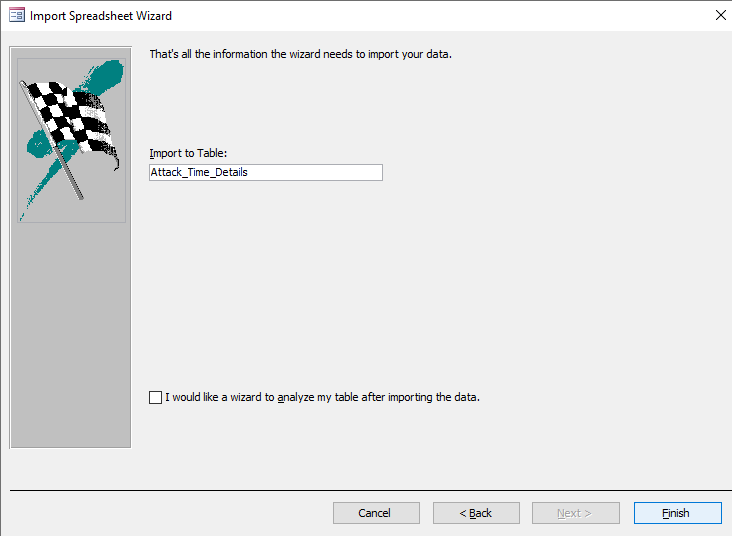
1. Choose “Yes (No duplicates)” for the chosen primary key of the fact table.

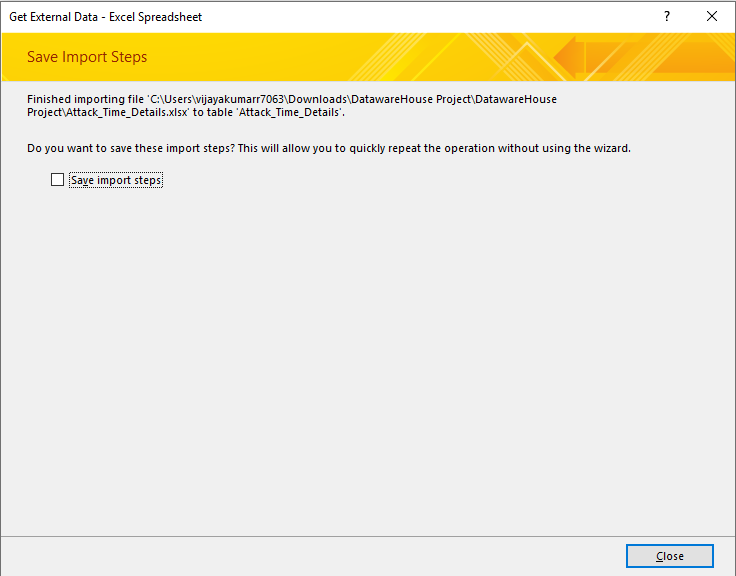


1. Choose Time\_ID as a primary Key for the fact table.

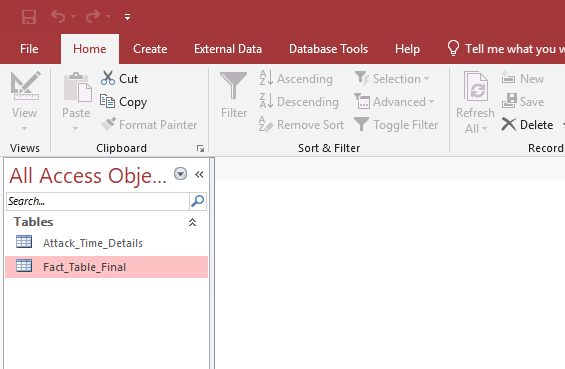


1. Naming the file being exported as Attack\_Time\_Details

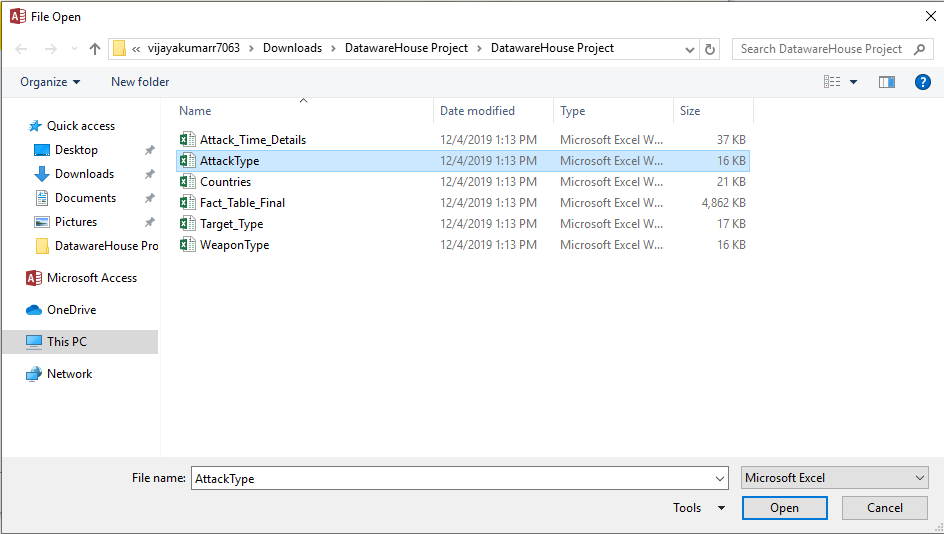


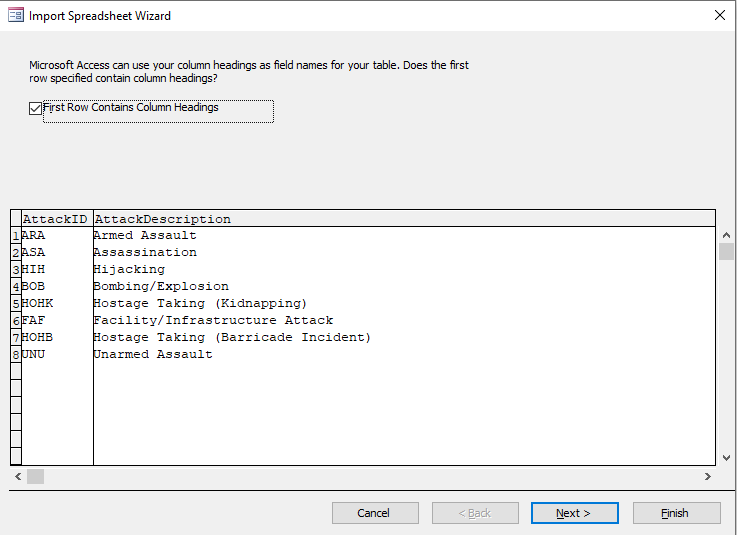


1. Attack\_Time\_Details is being exported successfully as seen in the screenshot below

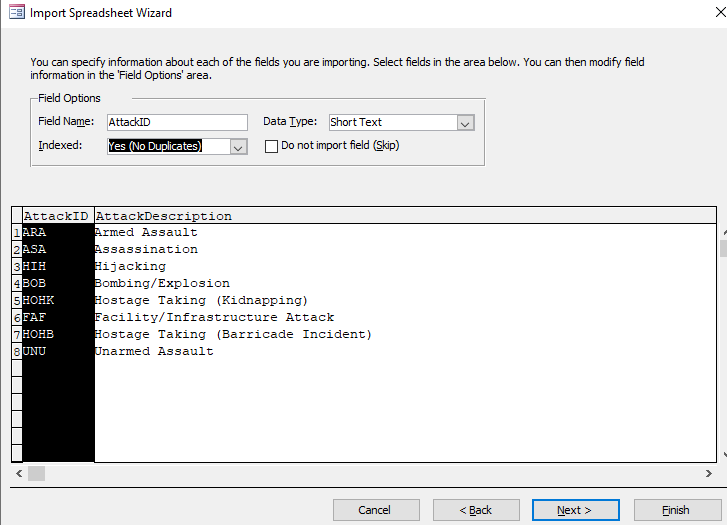


1. Import Excel file “AttackType” into Access

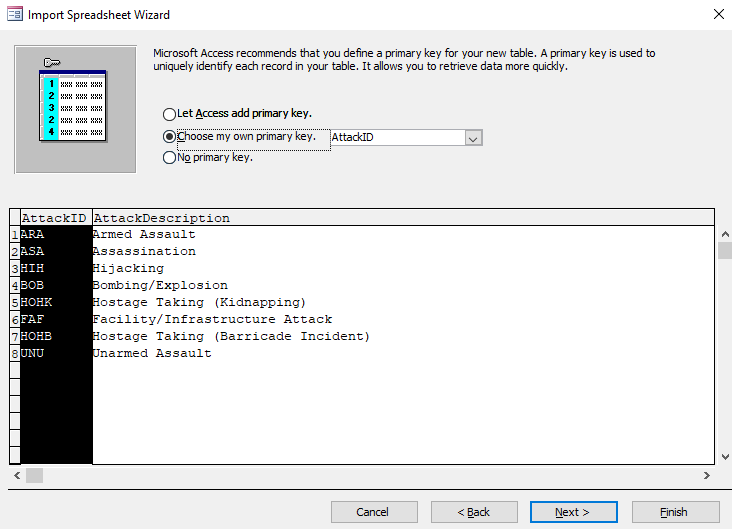


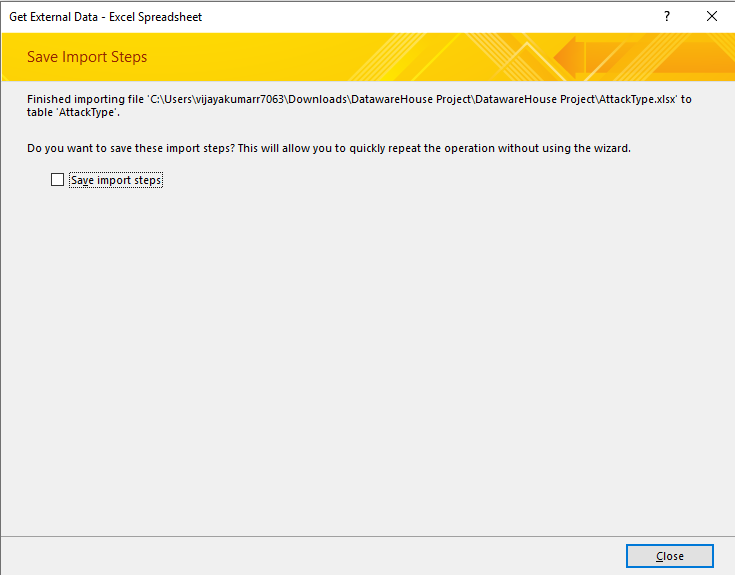


1. Choose “Yes (No duplicates)” for the chosen primary key of the fact table.

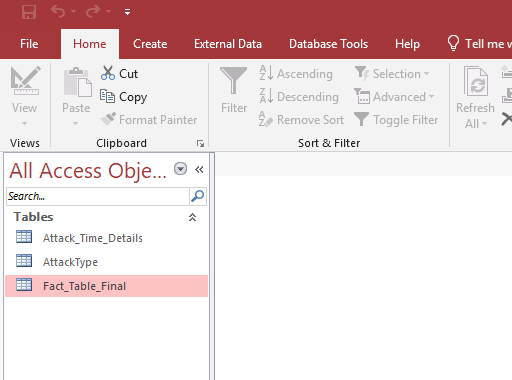


1. Choose AttackID as a primary Key for the fact table.

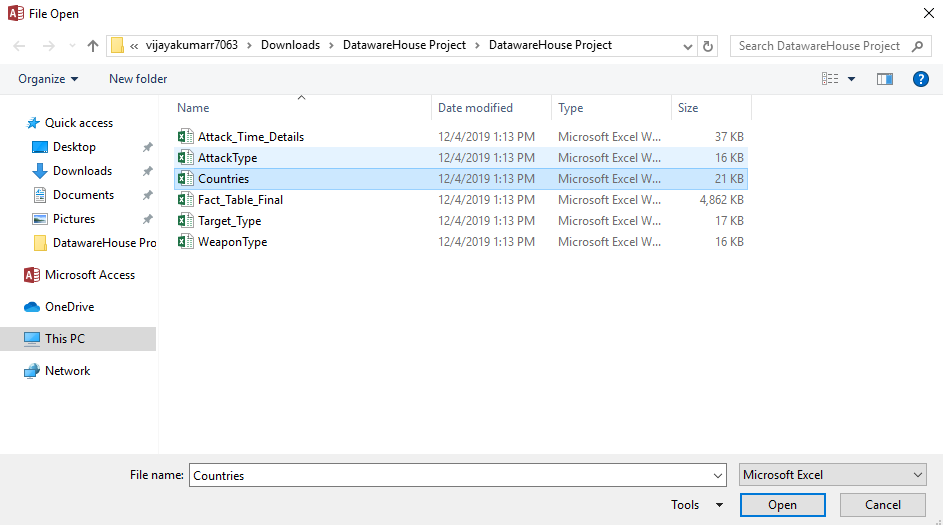


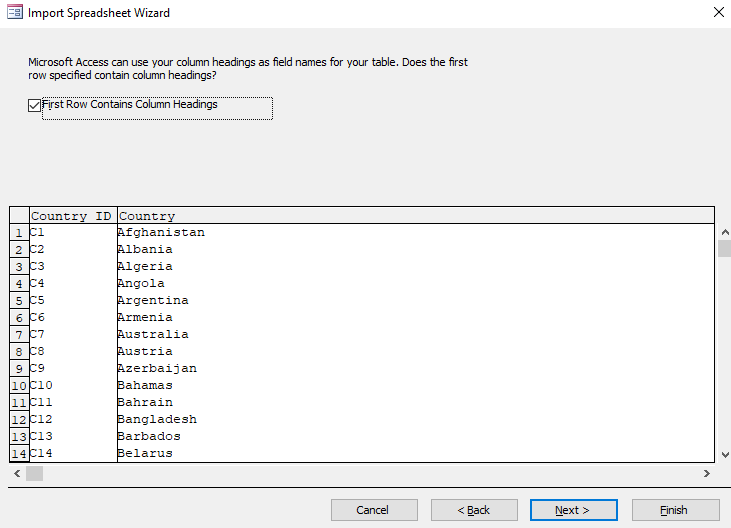


1. AttackType is being exported successfully as seen in the screenshot below

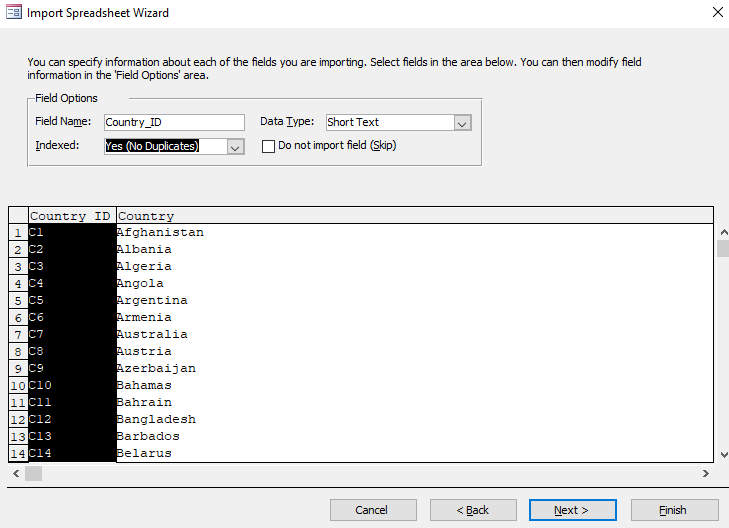


1. Import Excel file “Countries” into Access

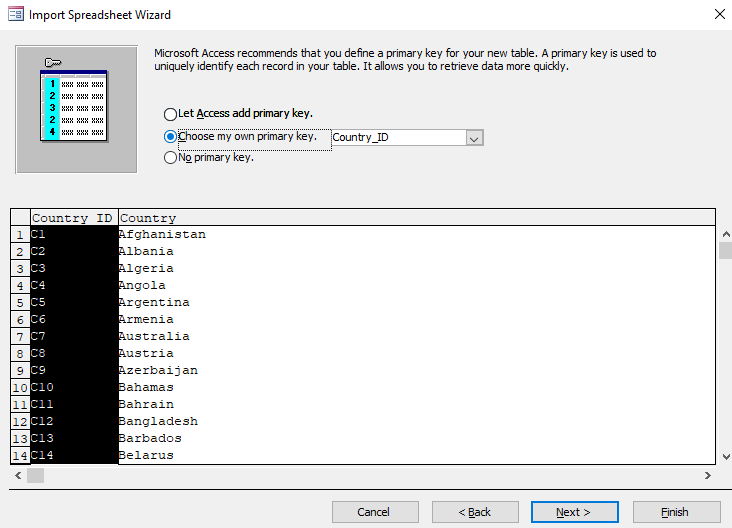




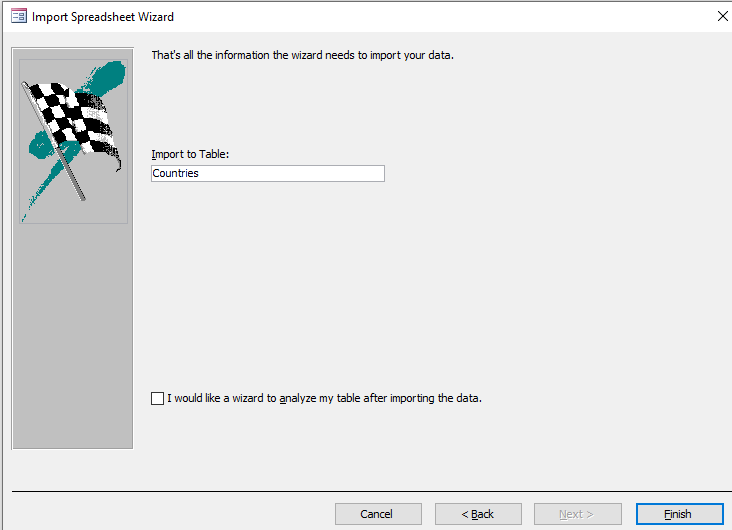
1. Choose “Yes (No duplicates)” for the chosen primary key of the fact table.

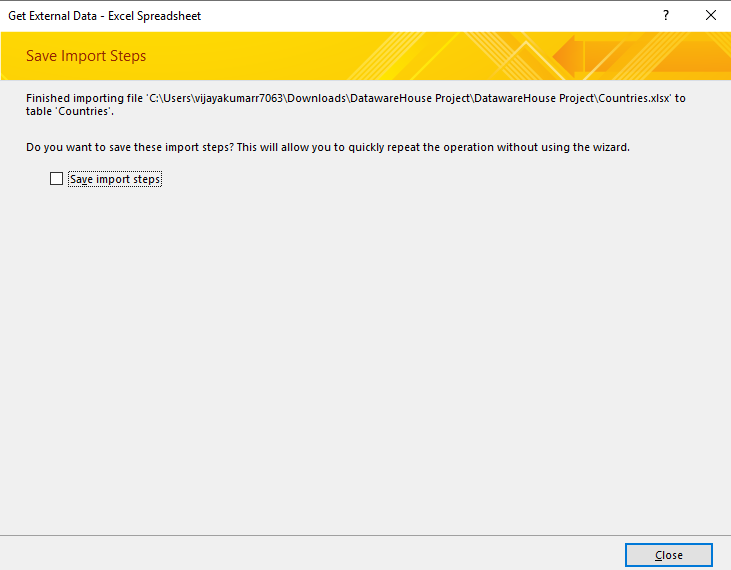


17.Choose Country\_ID as a primary Key for the fact table.

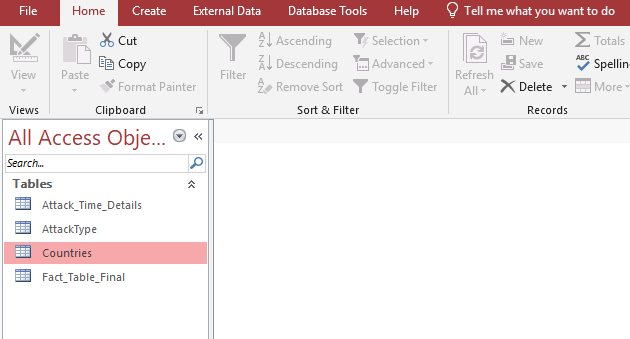


1. Naming the file being exported as Countries

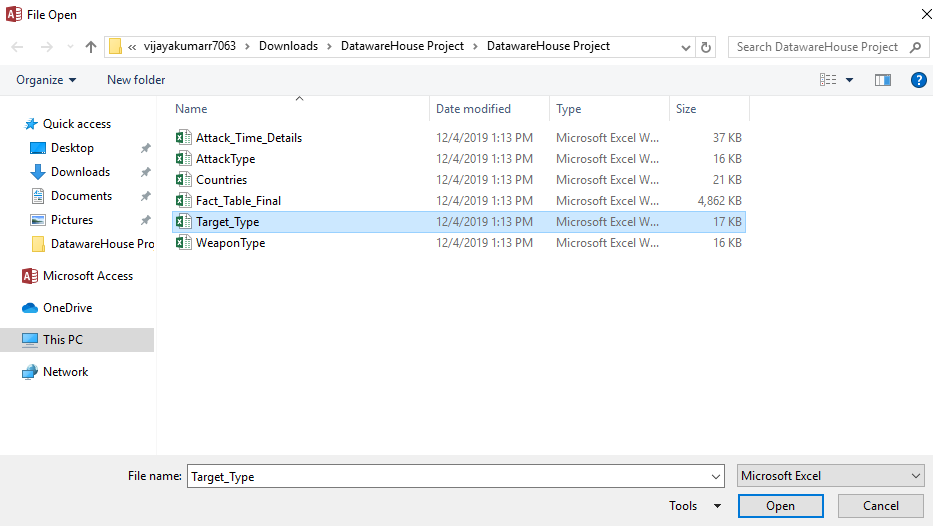


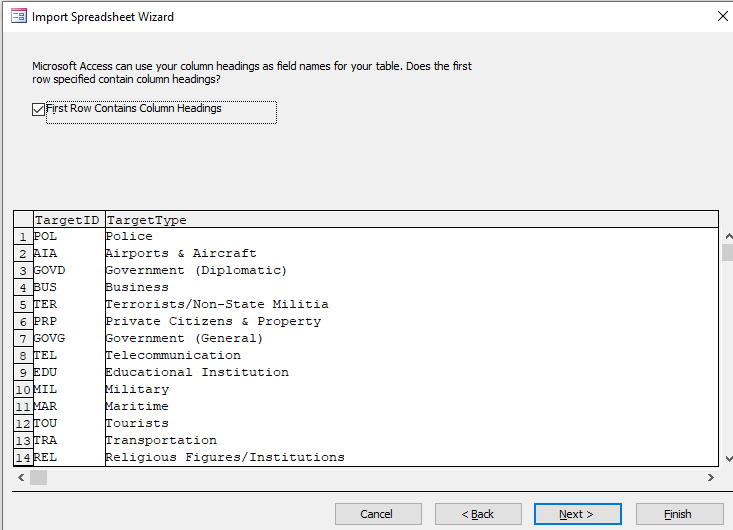


1. Countries is being exported successfully as seen in the screenshot below

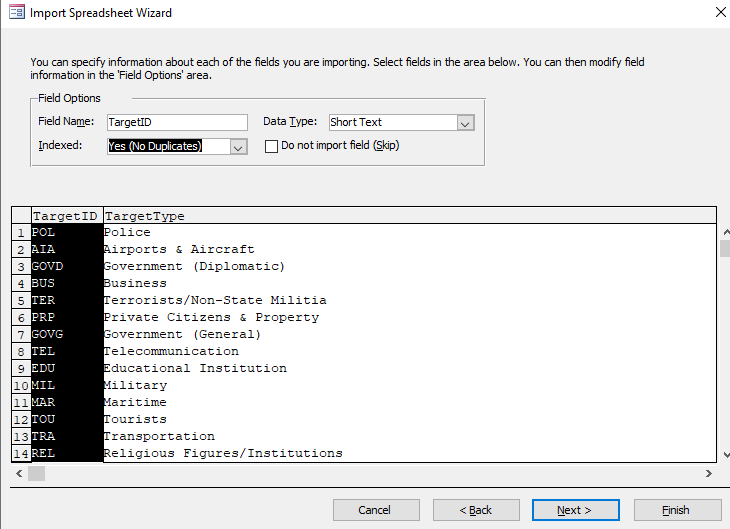


1. Import Excel file “Target\_Type” into Access

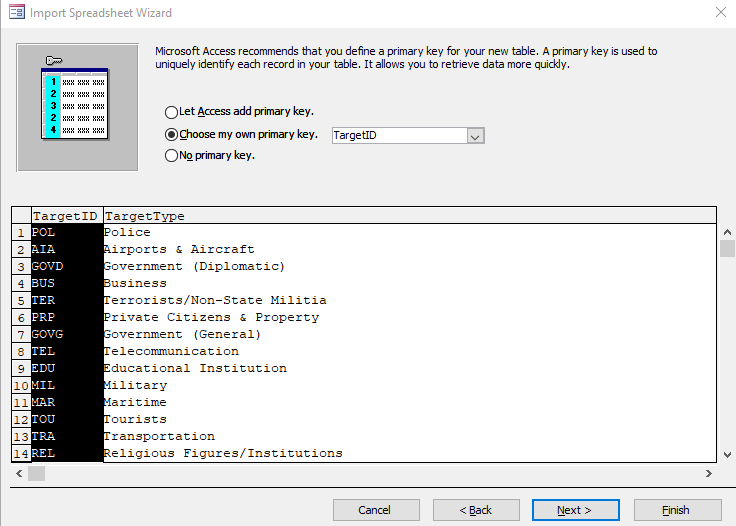




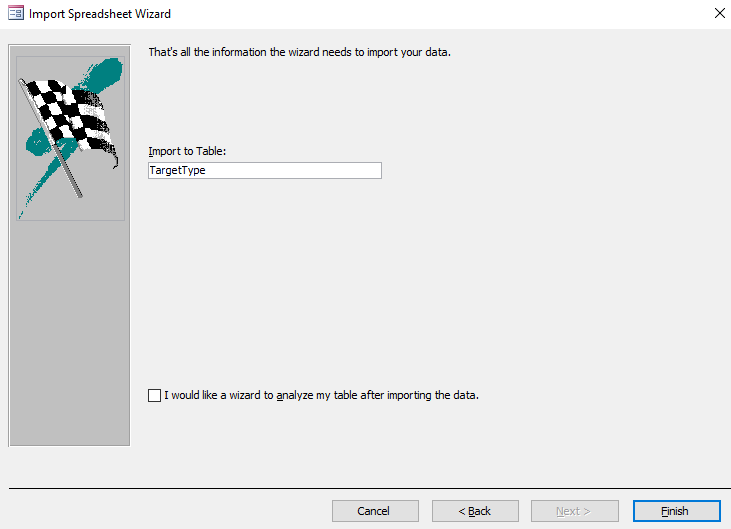
1. Choose “Yes (No duplicates)” for the chosen primary key of the fact table.

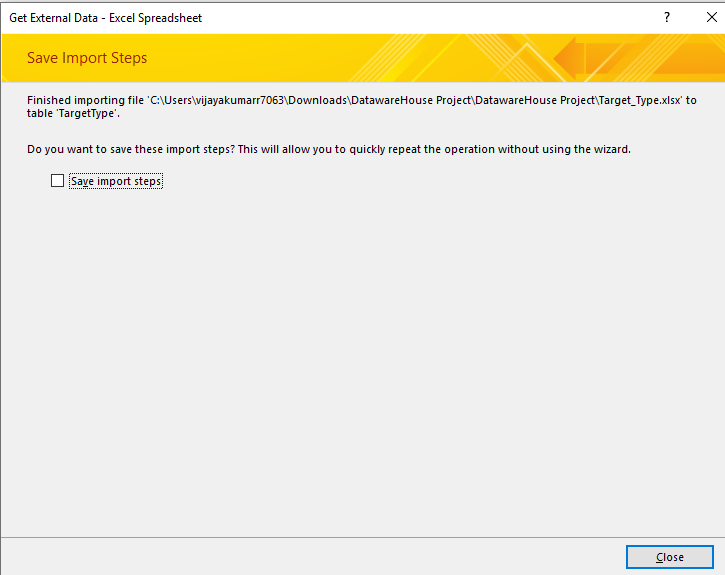


1. Choose TaregtID as a primary Key for the fact table.

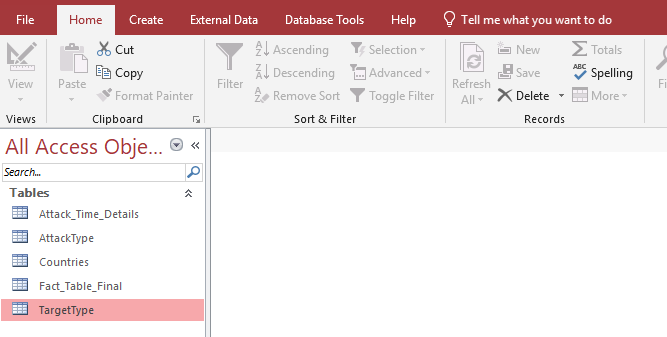


1. Naming the file being exported as TargetType

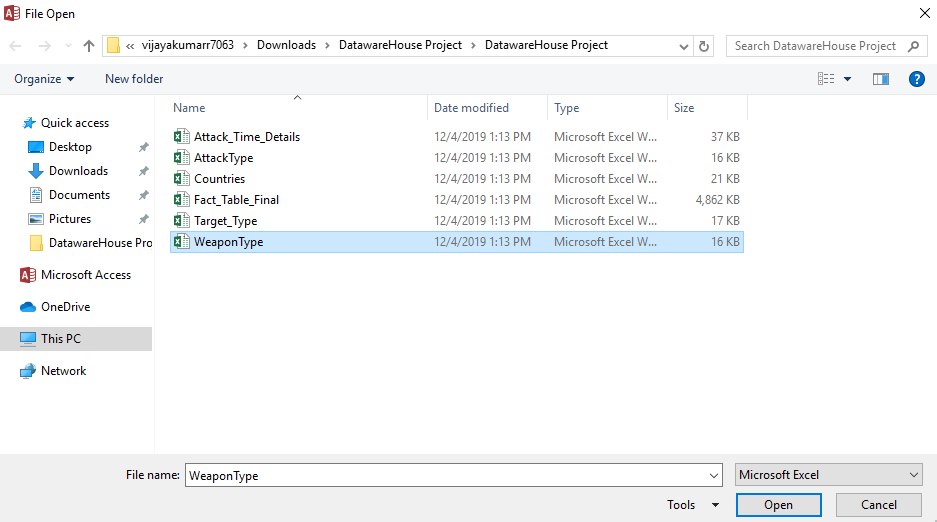


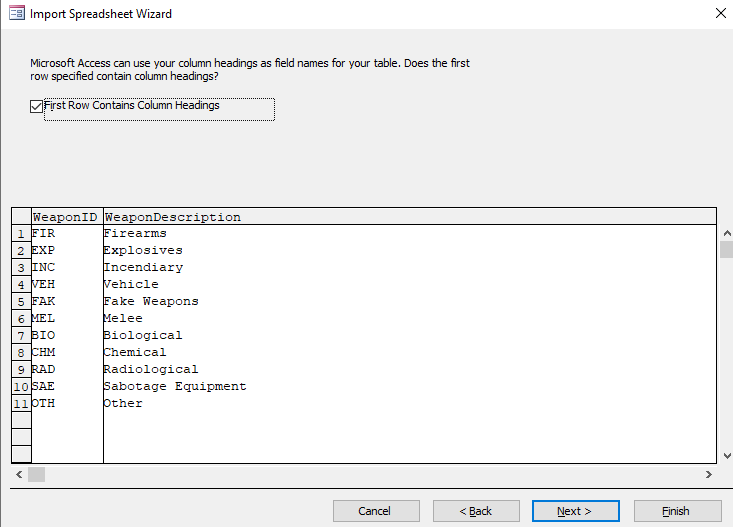


1. TargetType is being exported successfully as seen in the screenshot below

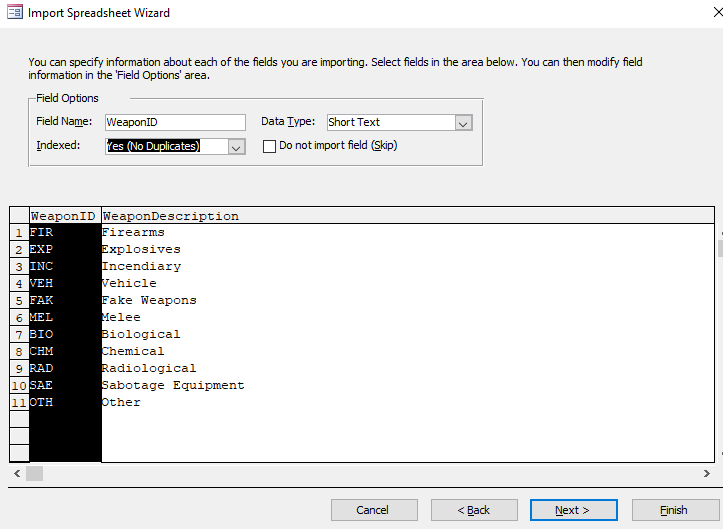


1. Import Excel file “WeaponType” into Access

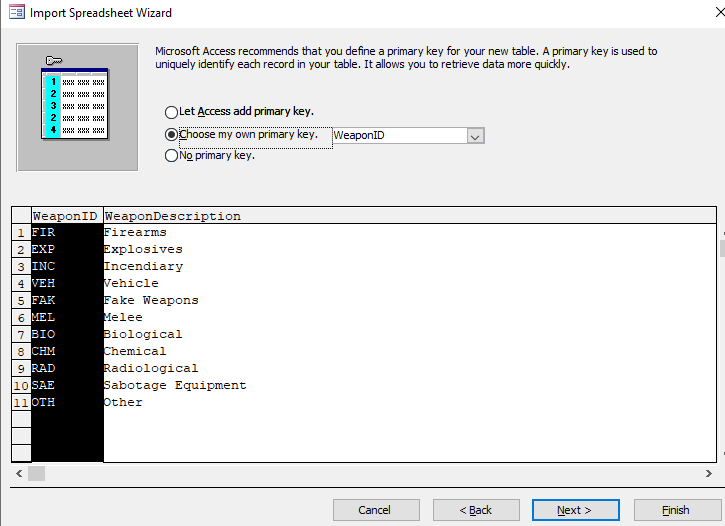




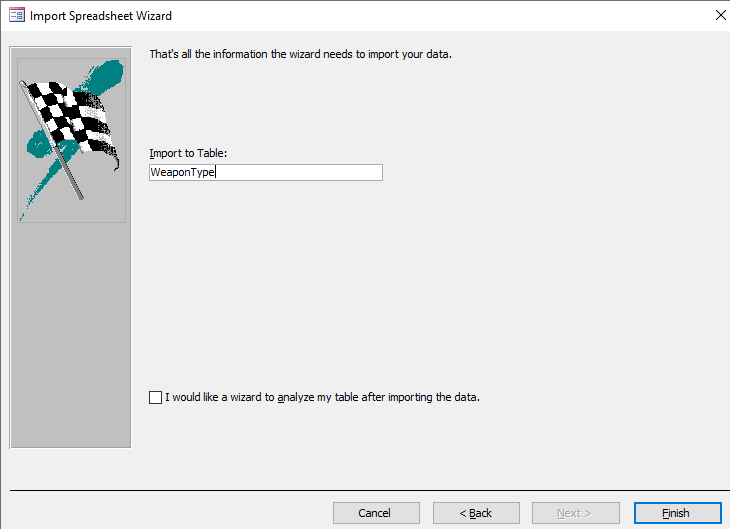
1. Choose “Yes (No duplicates)” for the chosen primary key of the fact table.

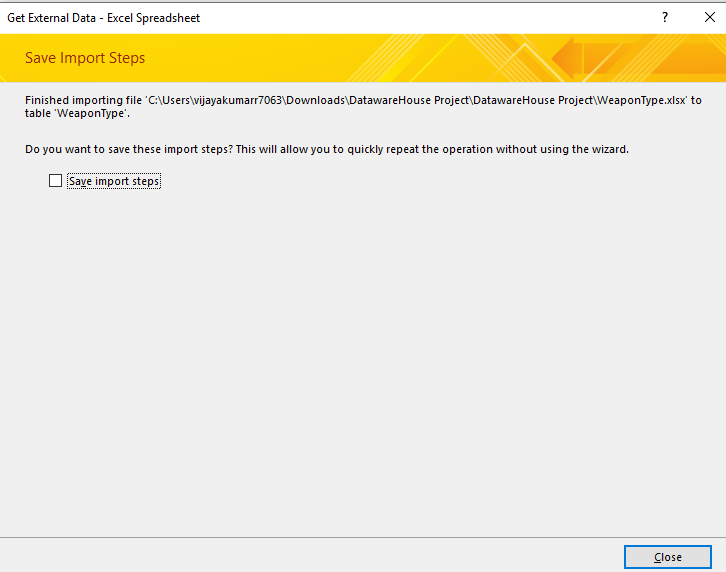


1. Choose WeaponID as a primary Key for the fact table.

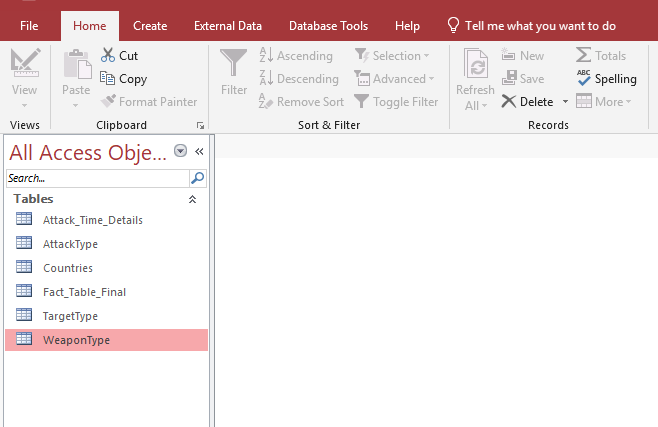


1. Naming the file being exported as WeoponType





1. WeoponType is being exported successfully as seen in the screenshot below



* 1. **PROJECT DATASET**

Steps to transfer Data

* Dataset: Retrieved Global Terrorism Dataset from

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

* Download all the dataset in Comma Separated value file which is in excel format.
* This data is then cleansed and combined and then imported to MS Access Database.

These Raw data in excel file are denormalization format. Some data is missing and repeated in the file. From Microsoft Access Environment, we also generated the Star schema diagram which we have displayed below. This MS Access Database was later used for developing the Cube in SQL Server Environment.

**3.4 DIMENSIONAL MODELING**

Data modeling is the process of creating a data model for the data to be stored in a Database. A Dimensional Model is a database structure that is optimized for queries and Data Warehousing tools .A Dimension model consists of a **Fact** and **Dimension** tables.

Data modeling is used to define and analyze data requirements needed to support the business processes within the scope of corresponding information systems in organizations. Therefore, the process of data modeling involves professional data modelers working closely with business stakeholders, as well as potential users of the information system.

Every attribute can have a unique level within a dimension. A Hierarchy is a specification of levels that represents relationship between different attributes within a dimension. In designing data models for data warehouses, the most commonly used schema types are [Star Schema](http://www.1keydata.com/datawarehousing/star-schema.html) and [Snowflake Schema](http://www.1keydata.com/datawarehousing/snowflake-schema.html). We are using the Star Schema.In real time OLAP Sytems we implement the data modelling with below objectives.

* **Identifying the business objective**
* **Identifying Granularity**
* **Identifying Dimensions and its Attributes**
* **Identifying the Fact**
* **Building of Schema**

In our Dataset the Fact table stores the measure for the following attributes in Global Terrorism are as follows:

1. Fact ID
2. Time ID
3. Country ID
4. Success Rate
5. Suicide
6. Attack Type
7. Target ID
8. Weapon ID
9. Death Count

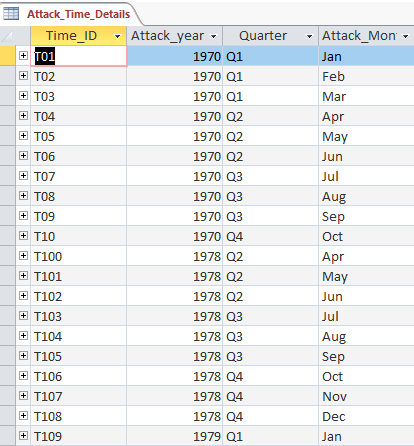
The Business Dimensions used in creating the Global Terrorism statistic Data warehouse are:

1. Attack Time Details
2. Attack Type
3. Countries
4. Target Type
5. Weapon Type

**Dimensions**:

Dimensional modeling are designed for reading, summarizing and analyzing numeric information. Relational Models are optimized for adding and maintaining data using real-time operational systems.

It is important to note that the dimensional modeling is not necessary depends on relational databases. The dimensional modeling approach, at the logical level, can be applied to any physical forms such as relational and multidimensional databases.

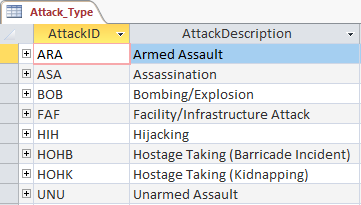
* **Attack Time Details** Attack time details dimensions has the hierarchies like Year, Month andQuarter. Time\_ID is the primary key.

**Attributes:**

1. Time\_ID – Unique ID for Attack details.
2. Attack Year – Year in which the incident occurred.
3. Attack Month – Number of Month in which the incident occurred.
4. Quarter – Quarter in which attack occurred.

**2.Attack Type**

This table is used to analyze the general method of attack. This dimension has a primary key Attack ID.

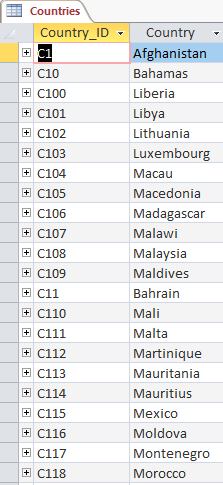


**Attributes:**

1. Attack ID – unique ID for the type of attack.
2. Attack Description – Type of the attack.

* **Countries**

The countries dimension is used to filter the records based on the country ID. This table provides the countries in which terrorist attack has occurred. This table has a primary key Country ID.

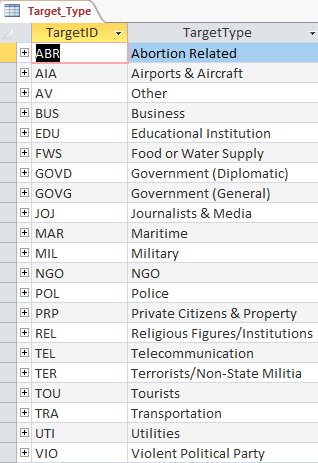


**Attributes:**

1. Country ID- Unique ID for identifying the countries.
2. Country – This field identifies the country where the incident occurred.

* **Target Type**

The Target type dimension is used to filter the records based on the target ID. This table provides the general type of target. This table has a primary key Target ID.

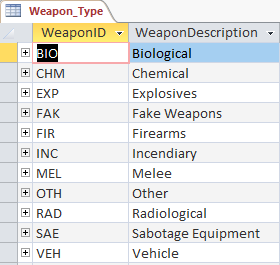


**Attributes:**

1. Target ID – Unique ID which is used to find the target type
2. Target Type – The general type of target/victim.

* **Weapon Type**

The weapon type dimension is used to filter the records based on the weapon ID. This table provides the general type of weapon. This table has a primary key weapon ID



**Attributes:**

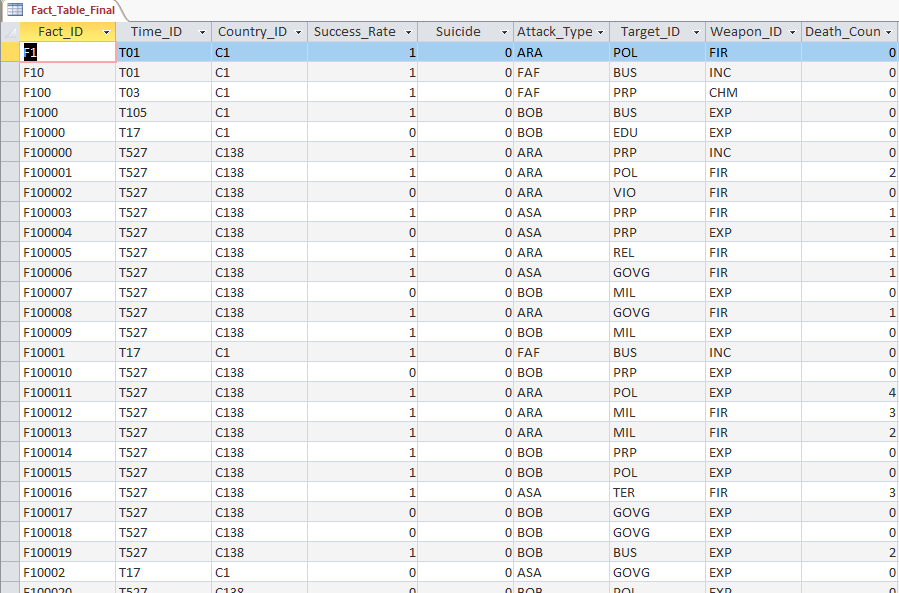
1. Weapon ID – Unique ID which is used to find the type of weapon used.
2. Weapon Description – General type of weapon used in the incident.

**FACTS:**

A fact table consists of the measurements, metrics or facts of a business process. It is located at the center of a star schema or a snowflake schema surrounded by dimension tables. Where multiple fact tables are used, these are arranged as a fact constellation schema.

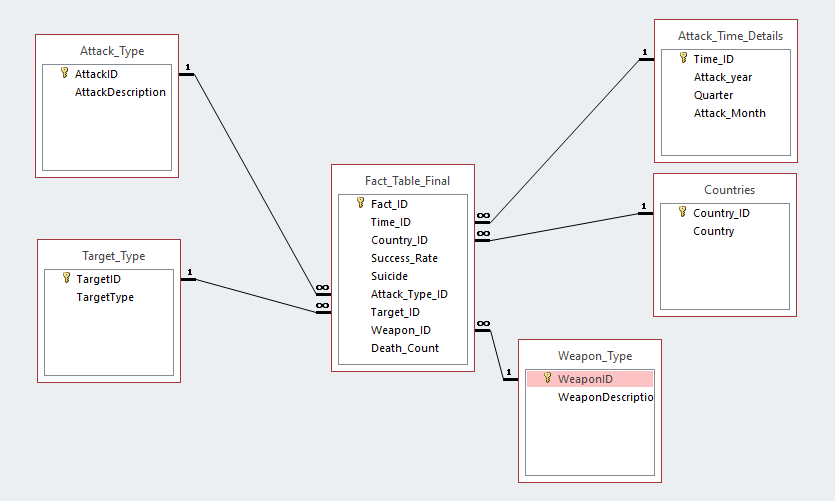
The facts table is all about:

1. Success Rate – It is nothing but success of a terrorist strike.
2. Suicide – If it is 1 = “yes”, the incident was a suicide attack. If it is 0 = “No”, there is no indication that the incident was a suicide attack.
3. Death Count – The number of total confirmed fatalities for the incident.



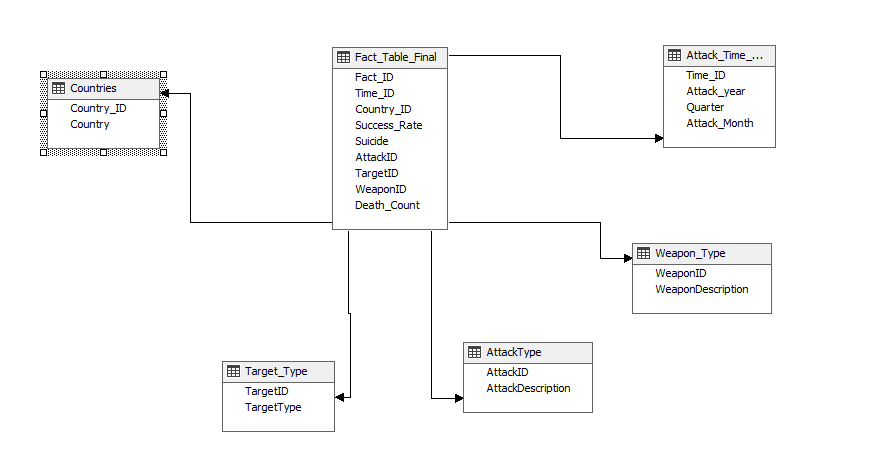
**STAR SCHEMA**

A star schema is composed of one or more central fact tables, a set of dimension tables, and the joins that relate the dimension tables to the fact tables. In STAR schema, the dimensional tables are arranged around a centralized fact table making it look like a star. A STAR schema is not normalized. Snow flaking is a method of normalizing the dimension tables in a STAR schema.



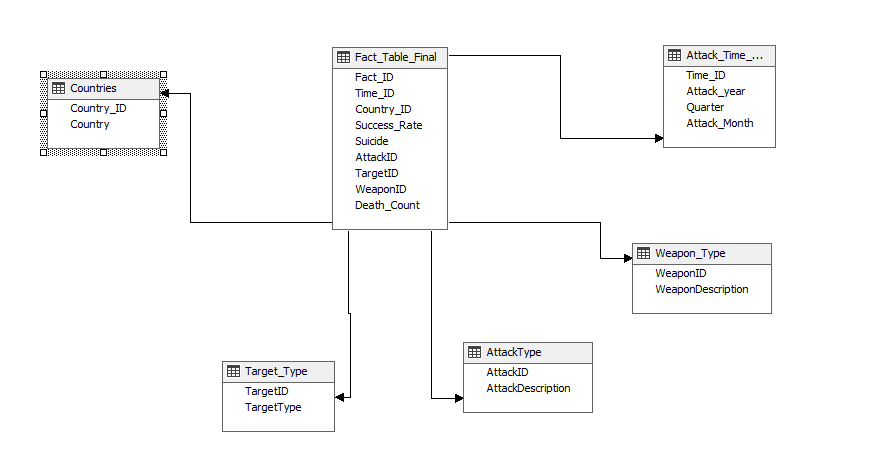
1. **DEVELOPING THE DATA SOURCE VIEW**

To create a Data Source View we need to create a SQL Server Database in SQL Server Management Studio, we first import the data from the already created MS Access Database and then transfer the files into the SQL Server Database and then use SQL Server Database for Data Source View generation and deployment.



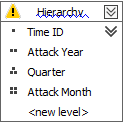
* **METHODS OF DEVELOPING THE DATA CUBE**

To create a Data Source Cube we need to create a SQL Server Database in SQL Server Management Studio, we first import the data from the already created MS Access Database and then transfer the files into the SQL Server Database and then use SQL Server Database for Data Source View generation and deployment.

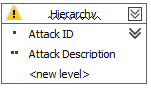


Below are the Hierarchies of the Relationships:

Attack Time Details:



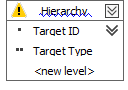
Attack Type:



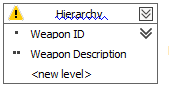
Countries:



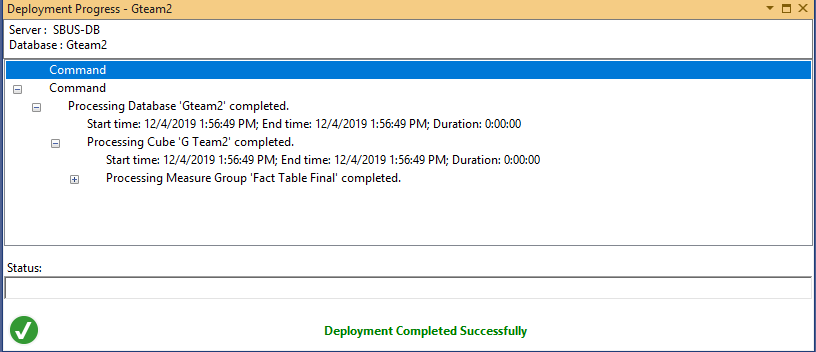
Target Type:



Weapon Type:



**DEPLOYMENT SUCCESS:**



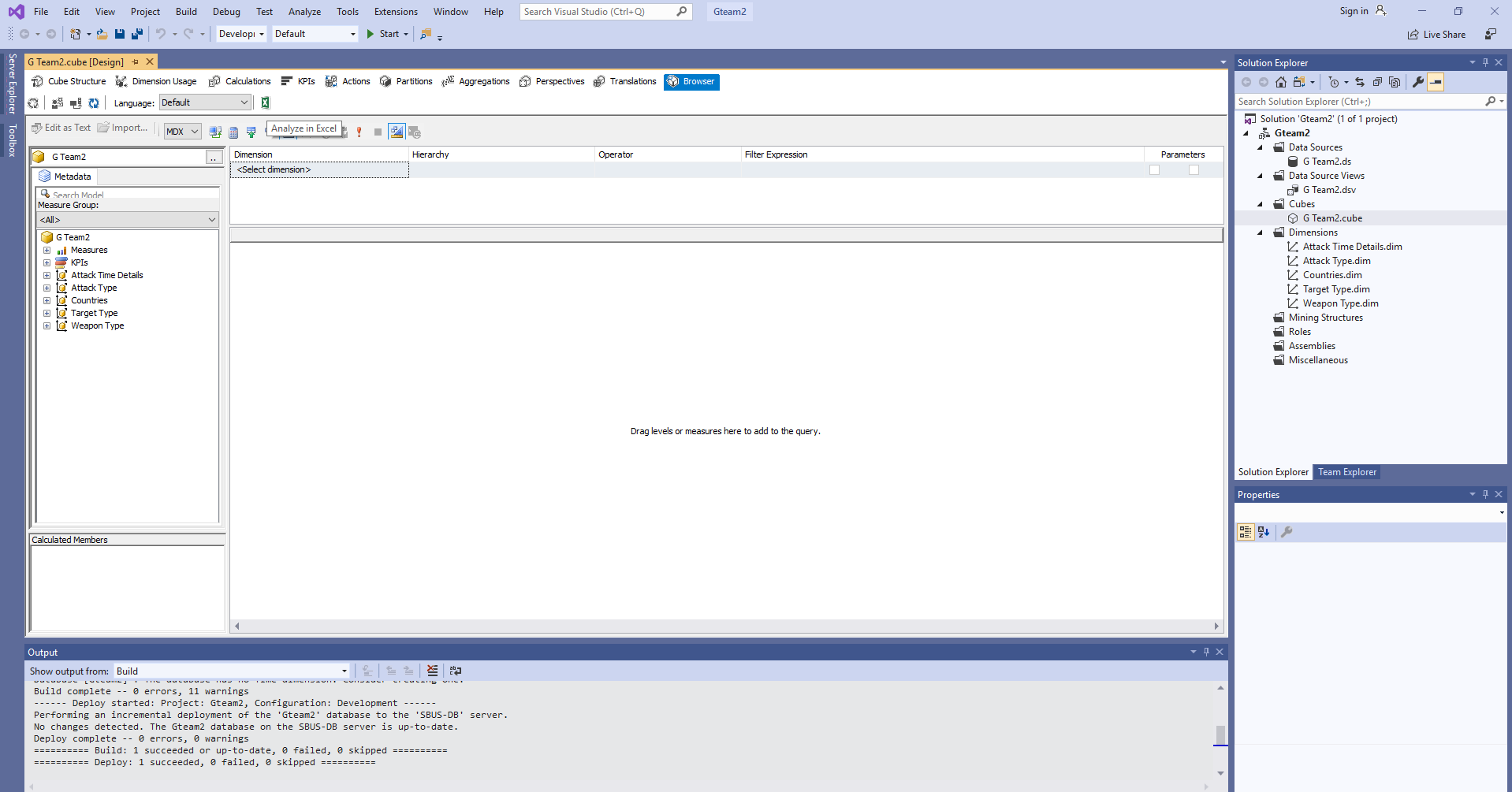
**REPORTS**:

1)Once the Data was successfully deployed we ran the reports. We pushed the data to Excel.

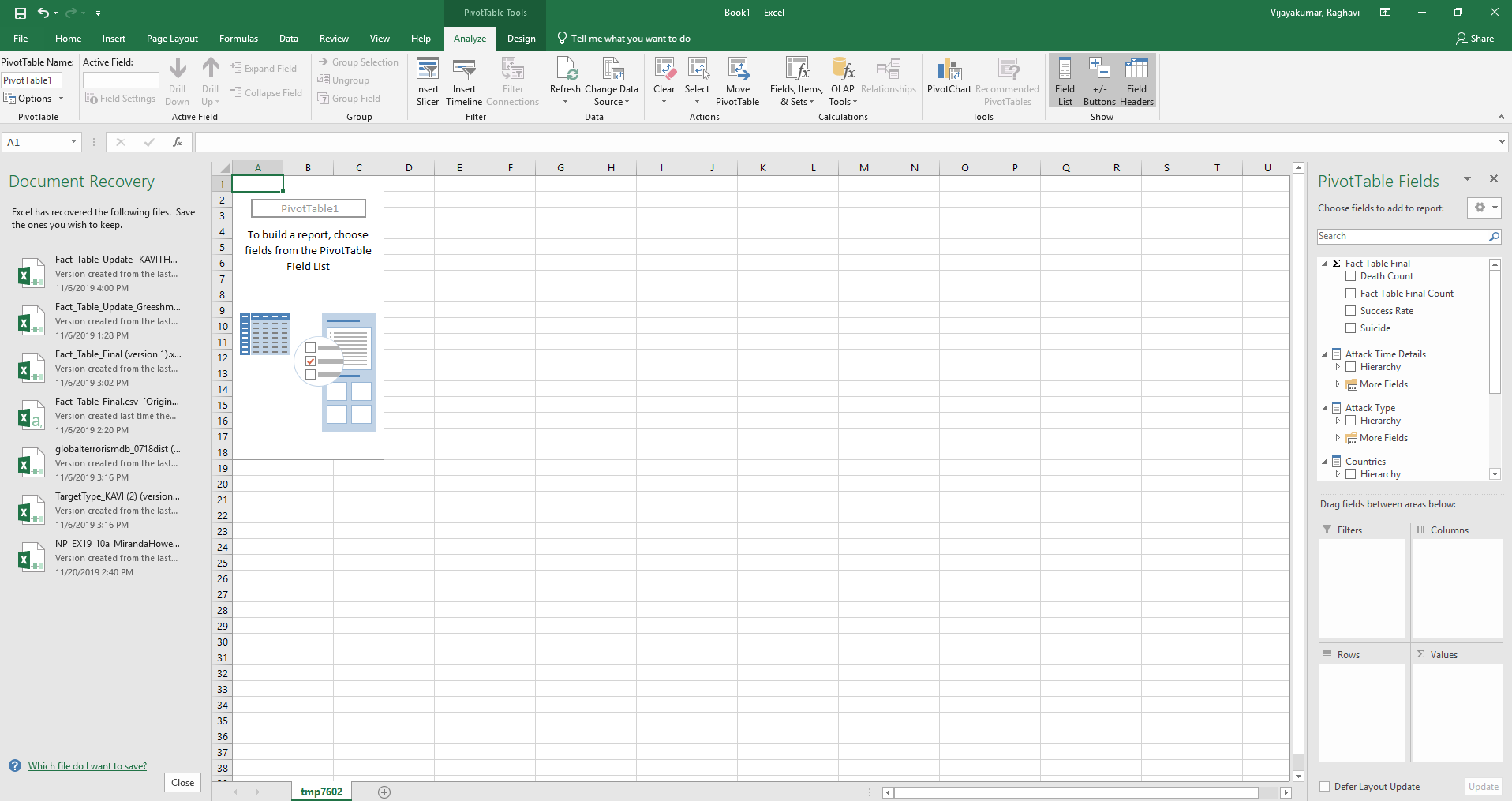
2)We then generated reports with Pivot Excel.

* The **FACT** values were put in the **Values** area.
* The **DIMENSIONS** were put in the **Rows** and **Columns** area.

The figures below show the visual studio and excel displays.

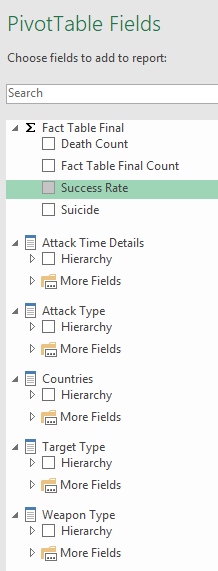


Visual Studio Display

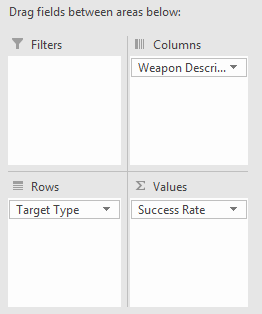


Excel Display

3)We can see that the Pivot is present on the left side of the Excel and the values we have used in our dataset have been preloaded already in the respective fields.

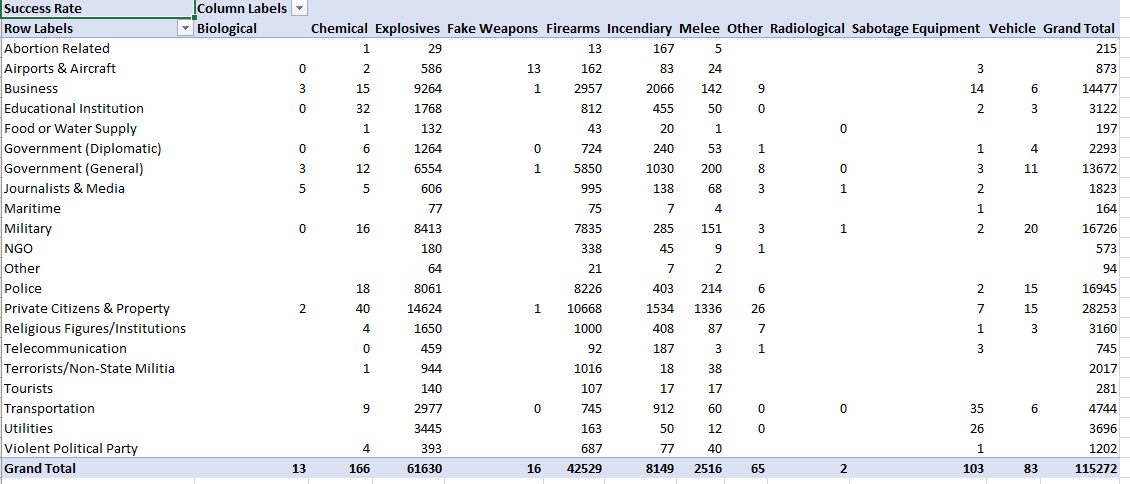


4)We have now chosen the **Fields** we want to do analysis on and placed them in the respectively within the **Values ,Columns and Rows**  area respectively.

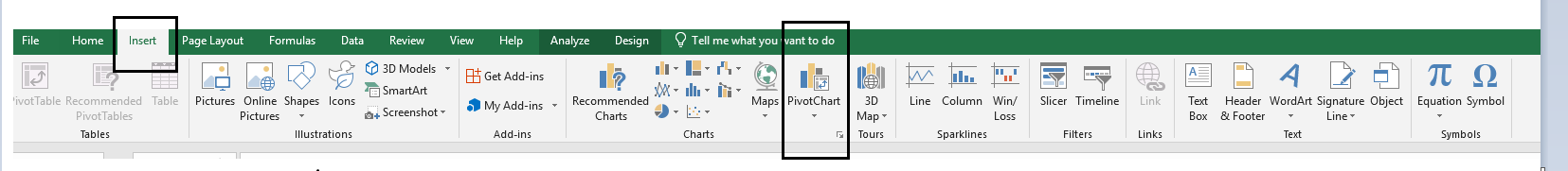


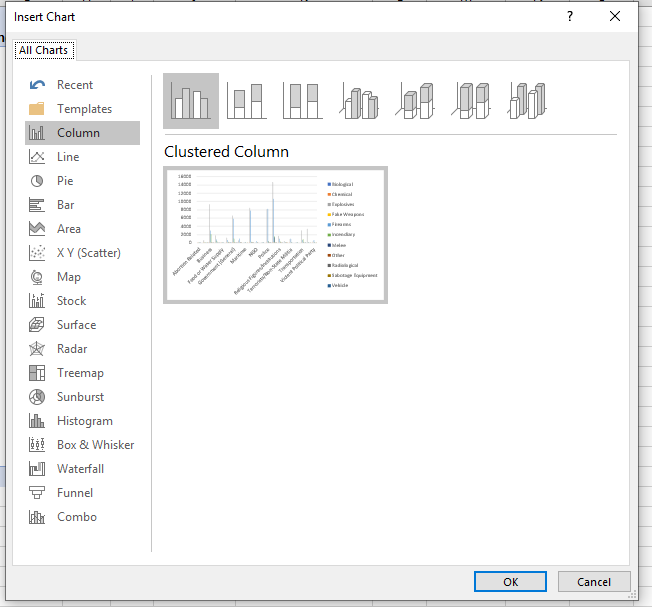
Choosing fields

5)This provided us with the below **Statistics.**

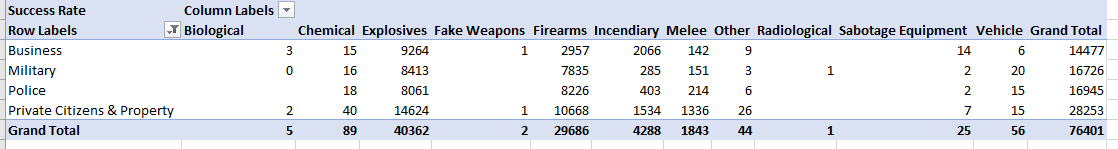


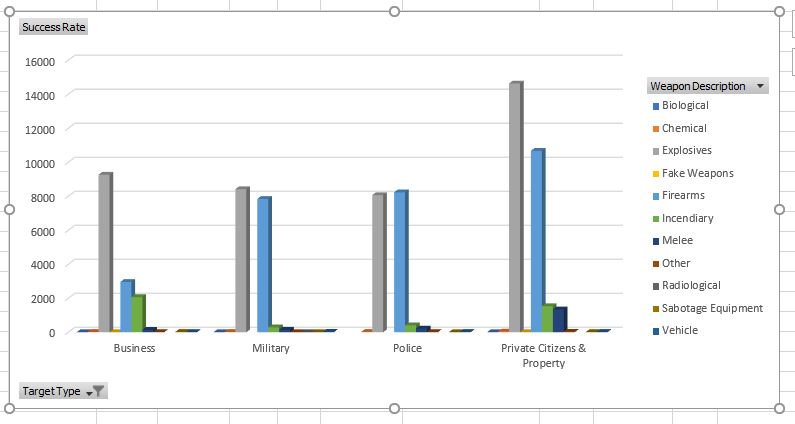
6)To generate a  **Chart** based on the above statistics,click on **Insert,**then click on the Pivot chart dialogue which then leads you to the pop up as shown below.





6) Now, choose the chart you want to display the analysis with. We in this project chose **3-D Clustered Chart** which generated the below chart.

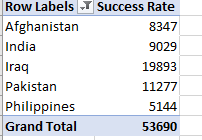


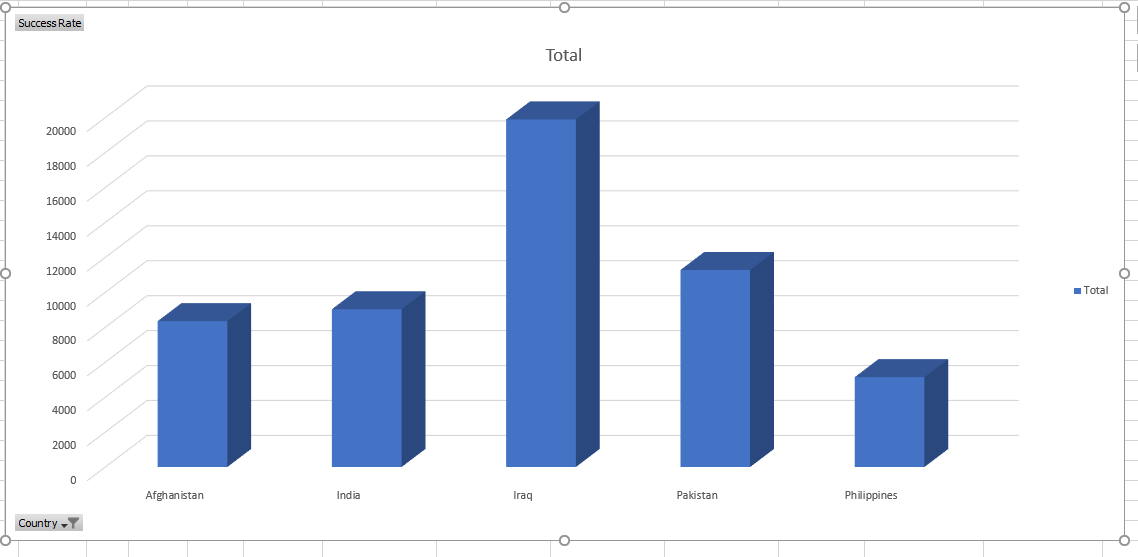


**In our project we have generated 6 Reports.**

1. **Top 6 Countries with Highest success Rate.**

We first analyzed the data with respect to all countries. Then, we chose the **Top 5 Countries** with highest attacks. We also did the analytics with only the **United States**.





**Observation:**

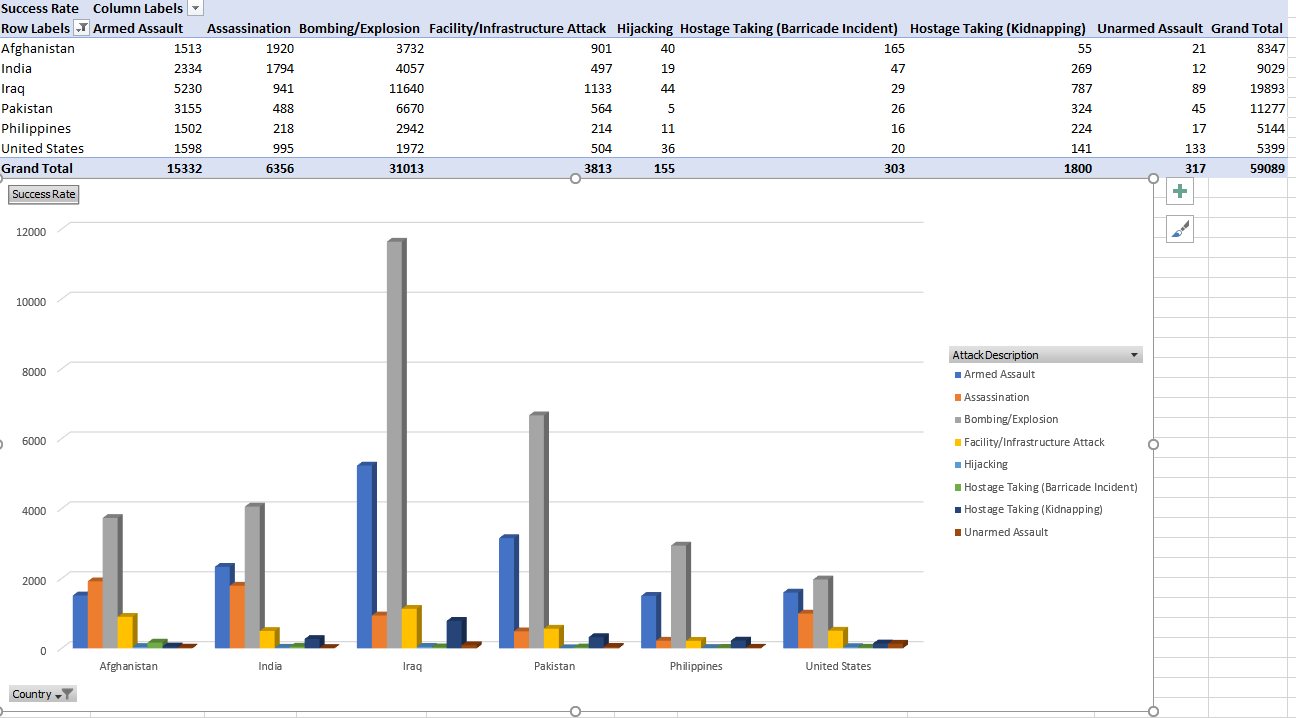
* From the figure above we see that the country with the highest success rate was **Iraq.**
* **Pakistan**  and **India** were among the second and third largest attacked countries.

**Pivot Details:**

* **Success Rate** in the Values area.
* **Country\_Description** in the Rows area.

1. **Countries with maximum Success Rate in particular Attack Type:**

Here we analyze the data with respect to the maximum success rate with the different types of attacks we have in our dataset. We were interested to know what type attack type were most successful in the top 6 countries we had shortlisted.



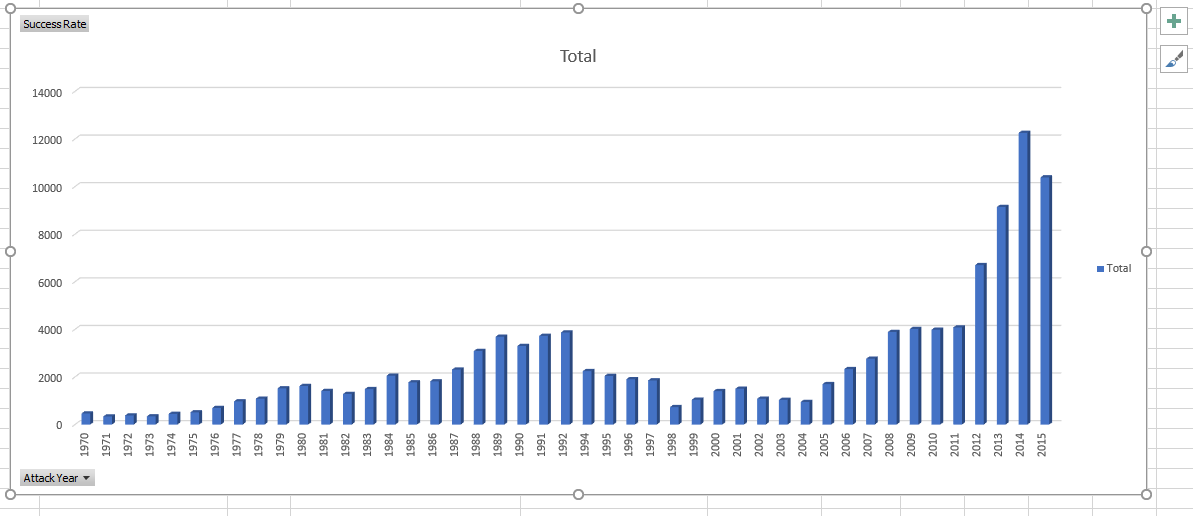
**Observation:**

* From the figure above we see that the country with the highest success rate was **Iraq** with Bombing has the highest successful attack Type**.**

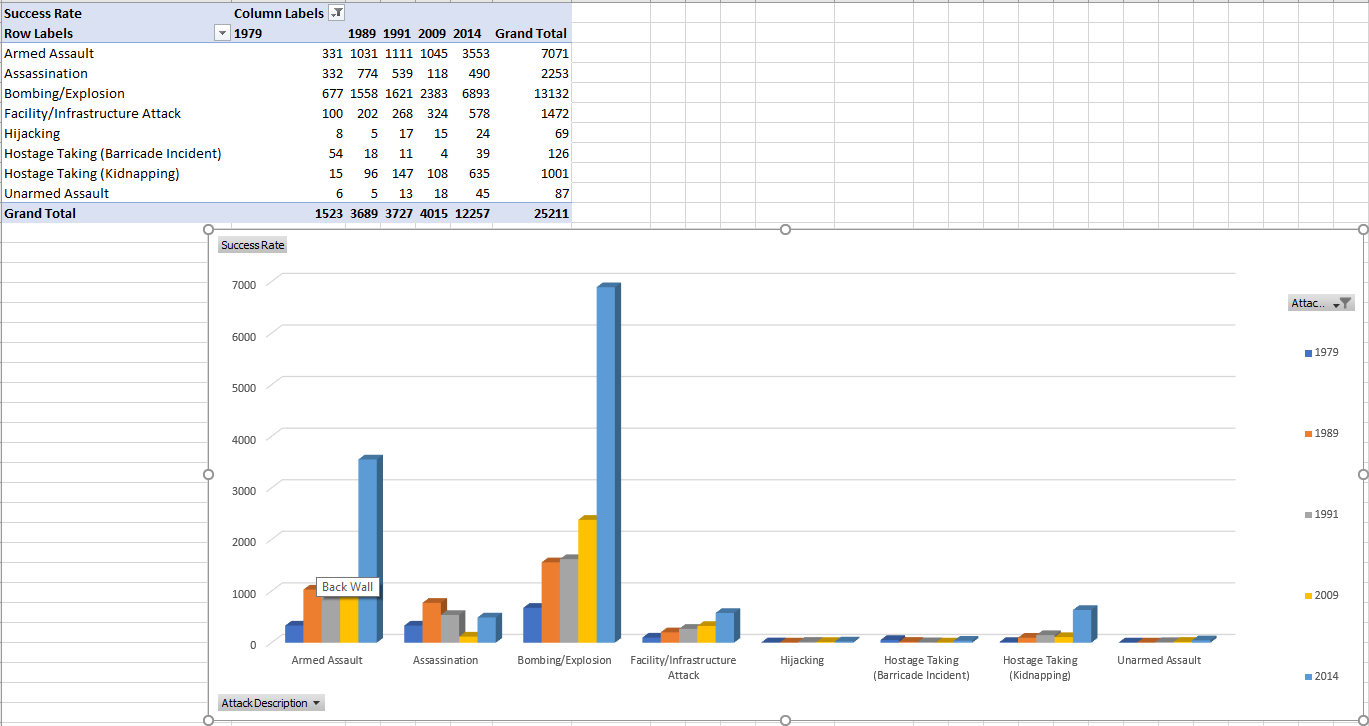
**Pivot Details:**

* **Success Rate** in the Values area.
* **CountryDescription** in the Rows area.
* **AttackDescription** in Columns area.

1. **Attacks that were Successful over different Years and Maximum attack type in a maximum attacked year.**
2. Here we picked data to find the different years in which the attacks were successful.This is displayed in the figure below.



2)We then picked the top 5 years and did the maximum successful attack type in those years.



**Observation:**

Concluding Bombing being the maximum attack.We compared the data:

1970-1979 .Maximum year with Bombing Attack-1979.

1980-1989.Maximum year with Bombing Attack-1989.

1990-1999.Maximum year with Bombing Attack-1991.

2000-2010.Maximum year with Bombing Attack-2009.

2010-2015.Maximum year with Bombing Attack-2014.

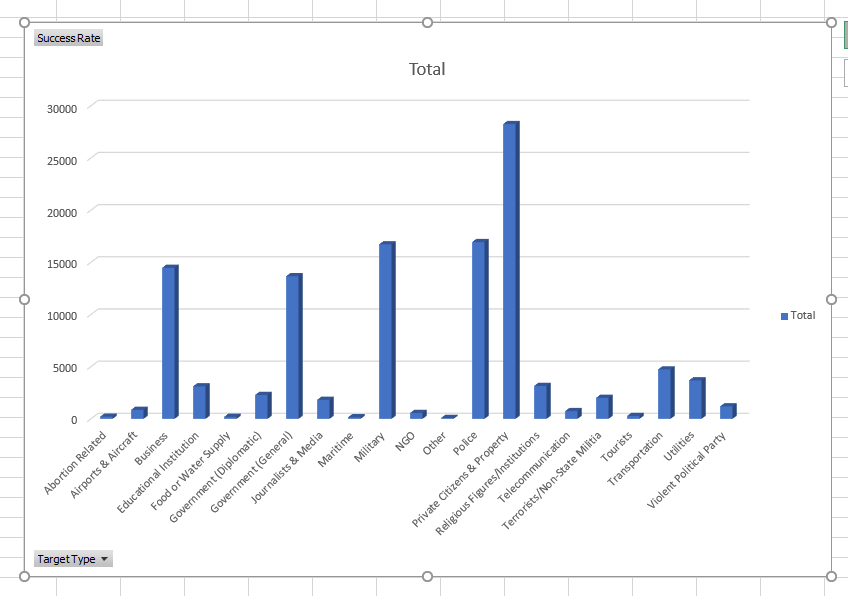
We also determined from our analysis that 2014 has the maximum attack.

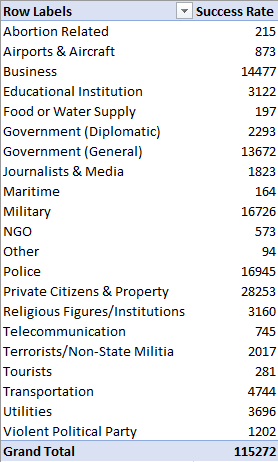
**Pivot Details:**

* **Success Rate** in the Values area.
* **AttackYear** in the Rows area.
* **AttackDescription** in Columns area.

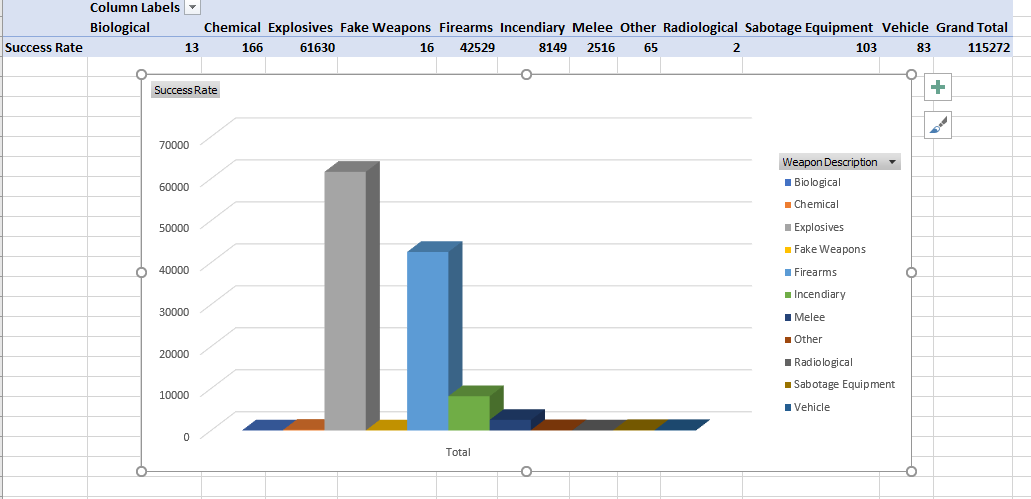
1. **Most used weapon on Major Targets:**

1)Analyzed the data with the Major Targets.

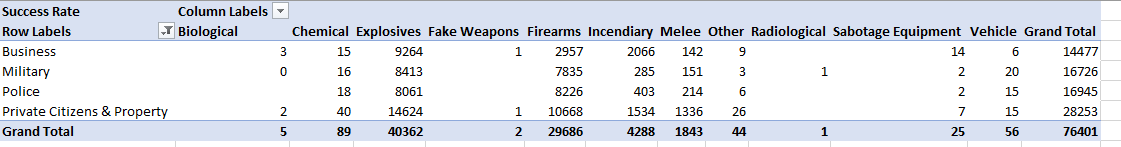


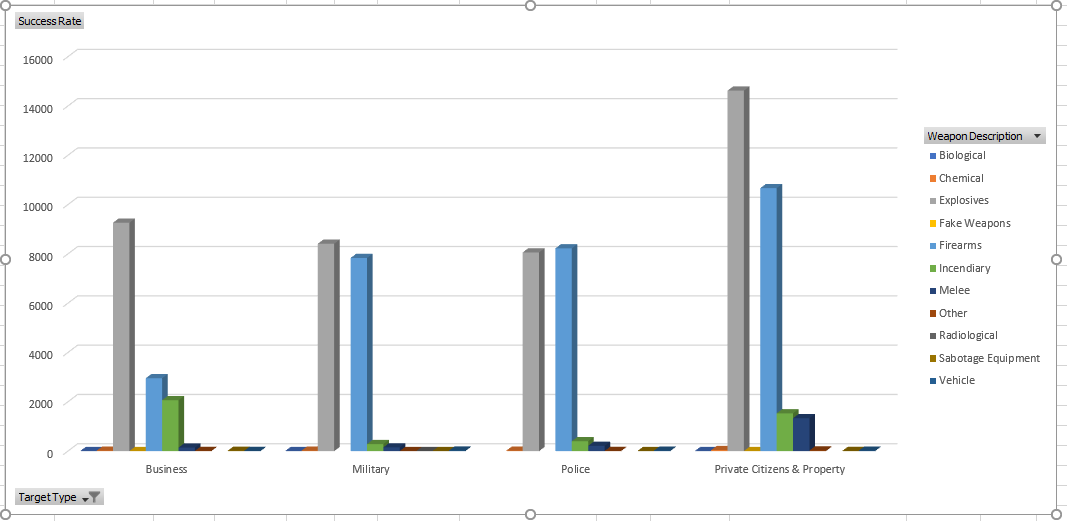


3)Major weapons used is displayed below:



2)Major Weapons on the Targets we defined above:





**Observation:**

From the above reports

* We have from the above reports the major targets and concluded the top 3 targets as below:
* **Private Properties and Citizens.**
* **Police and Military**
* **Business**
* From the reports we also found that targets were mostly attacked with the explosives.
* Most used top 3 weapons from the analysis is as below.
* **Bombing**
* **Firearms**
* **Incendiary**

We also determined from our analysis that 2014 has the maximum attack.

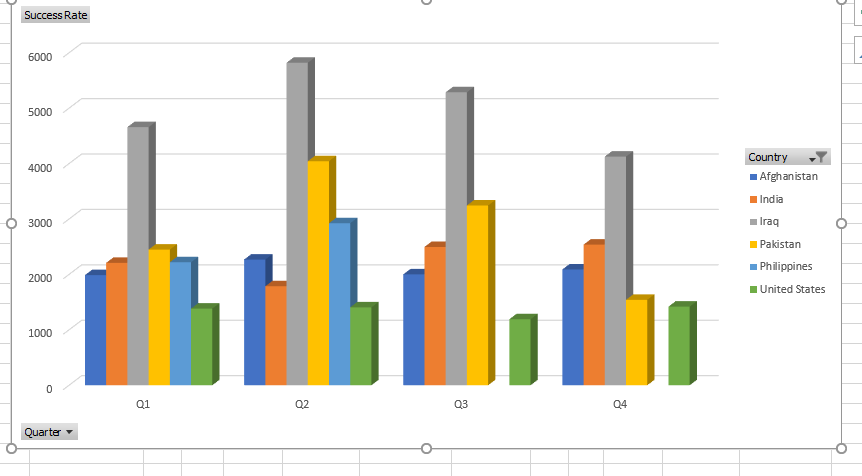
**Pivot Details:**

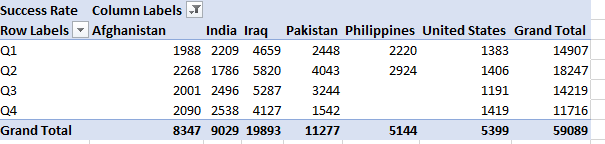
* **Success Rate** in the Values area.
* **TargetType** in the Rows area.
* **WeaponDescription** in Columns area.

1. **Attacks with time frame.**

Here we wanted to analyze the data because we wanted to see over what time period the attacks we maximum. Before doing the analysis we predicted that the Christmas/New years would be the highest time for a terror attack and began our analysis.

1)We picked our top 6 Countries for the analysis with success rate and Quarters of attack has the other two parameters.





**Observation:**

From the above reports:

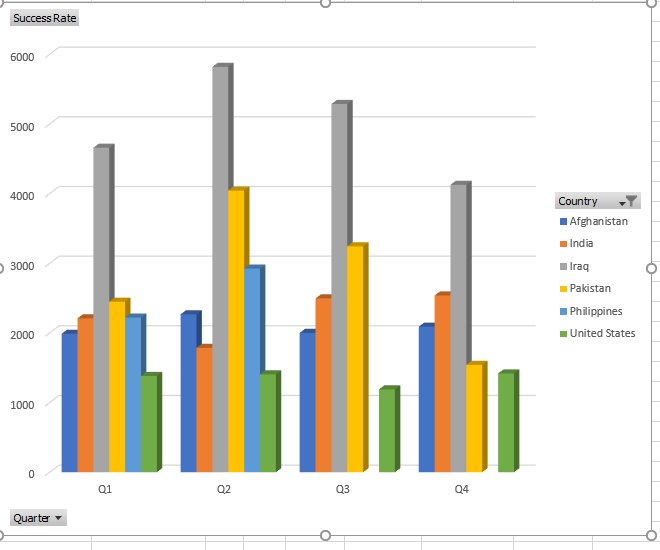
* Turned out that the prediction was wrong. The attacks were prominent around the months of April, May, June (Quarter 2) with no festive seasons or famous events even on Global platform.
* Iraq which the most attacked country had Q2 has the most attacked time period.

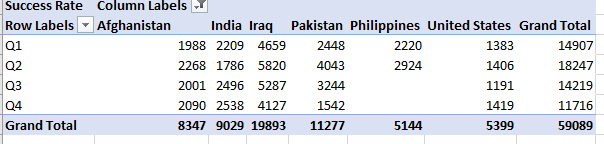
**Pivot Details:**

* **Success Rate** in the Values area.
* **Quarter** in the Rows area.
* **Country** in Columns area.
* **Success Rates and The death Counts.**

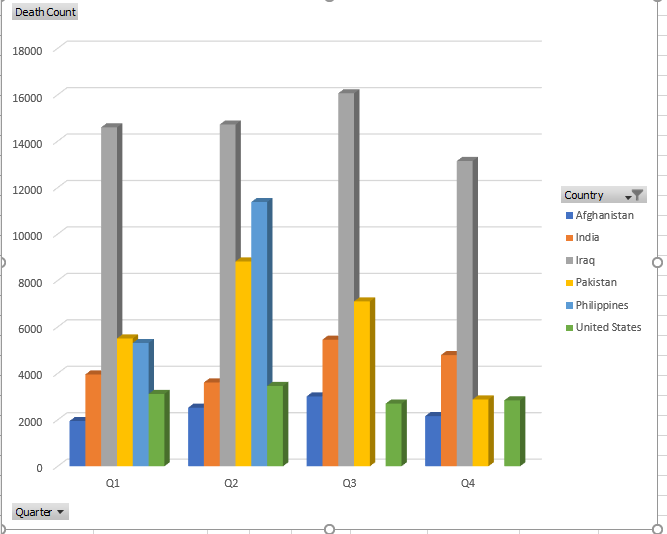
Here we analyze the data on the measure of its success rate and the death counts for each quarter. At the end of this step we also checked the country with maximum death count.

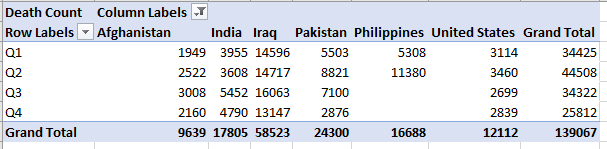
1)Below Chart displays the success rate in different quarters.



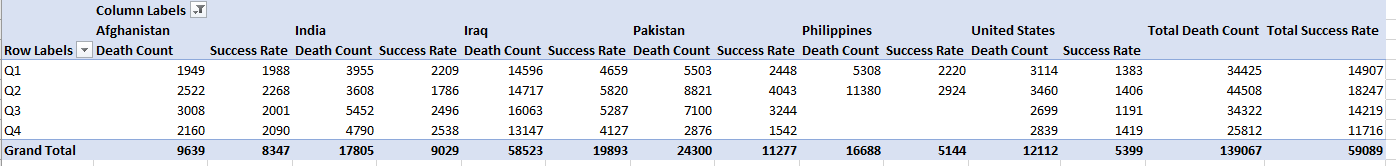


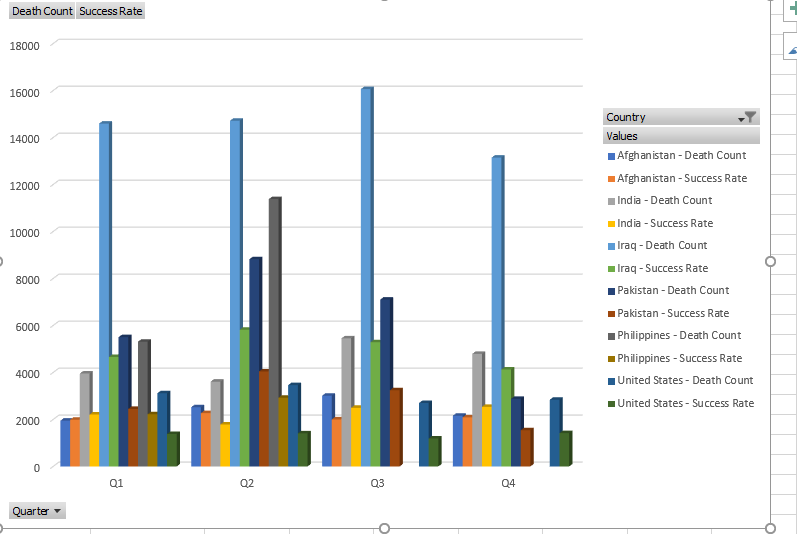
2)Below Charts display the Death Count in different quarters.





3)Cumulative chart of the death counts and success Rate.





**Observation:**

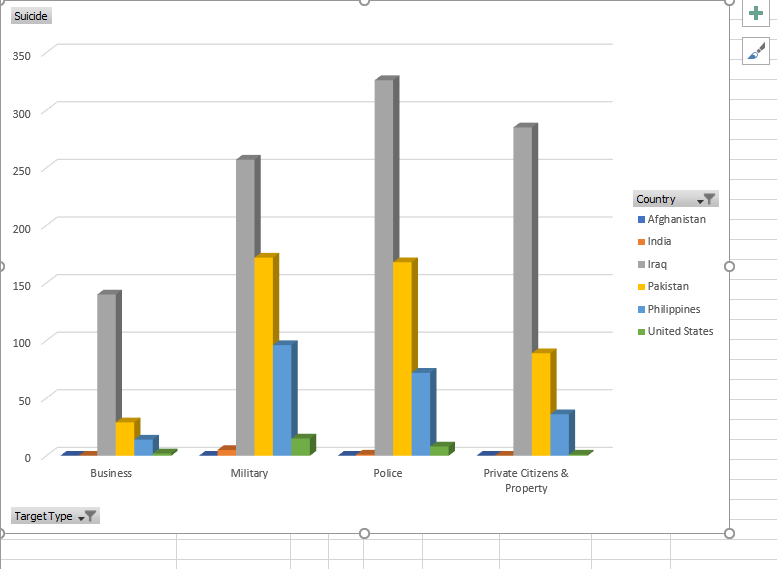
From the above reports:

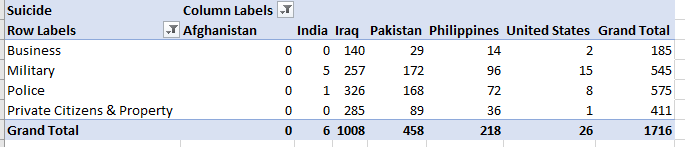
* Success rate was maximum in Quarter 2.
* Death Count was maximum in Quarter 3.
* Top 3 Countries with Maximum Success rates and death Counts.
* Iraq
* Pakistan
* India

**Pivot Details:**

* **Success Rate** in the Values area.
* **Quarter** in the Rows area.
* **Country** in Columns area.
* **Suicide Attacks on Top targets**

Here we are doing the analysis on suicide attacks on top targets.





**Observation:**

From the above reports:

* Suicide Attacks were Maximum on Police.

**Pivot Details:**

* **Success Rate** in the Values area.
* **Target Type** in the Rows area.
* **Country** in Columns area.

**CONCLUSION**

Data Warehousing has been important business tool to analysis data and to make decisions of past records for both successes and failures. The Global terrorism was implemented to recognize the terror attacks globally. It was developed by first cleaning the raw data with the required information for our analysis. The data was then divided to obtain the dimensions and facts which was then pushed into MS Access. Star Diagram was used to create the MS Access database. Cubes were generated using the MS Access database and SQL Server .Cubes played a primary role in creating and generating reports for the analysis.

**REFERENCES**

**1)Dataset: Retrieved Global Terrorism Dataset from**

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

**2) Google Images**