Capstone Project Submission

Team Member's Name, Email and Contribution:

Team Member's Name	Email
Ajinkya Shingote	shingoteajinkya65@gmail.com
Syed Aquib	syedaquib153@gmail.com
Ganesh Subramanian	gsubramanian5@gmail.com
Mayur Chulbhare	chulbharemayur@gmail.com
Abhishek Kirar	abhishekkirar27@gmail.com

Contribution:

Ajinkya Shingote:
- Data Wrangling
    1. Loading and Preprocessing
    2. Structuring data
    3. Enriching data
- Data Mining
- Data Analysis
- Visualizations
    - Bar graphs and Pie charts

    - Observations
    - Summarization
    - Conclusions
    - Technical Document
    - Power Point Presentation

Ganesh Subramanian:
- Data Wrangling
    1. Loading and Preprocessing
    2. Structuring data
    3. Enriching data
- Data Mining
- Data Analysis
- Visualizations
    - Bar graphs and Pie charts

    - Observations
    - Summarization
    - Conclusions
    - Technical Document
    - Power Point Presentation

Mayur Chulbhare:

- Data Wrangling
    1. Loading and Preprocessing
    2. Structuring data
    3. Enriching data

- Data Mining
- Data Analysis
- Visualizations
  - ➢ Bar graphs and Pie charts

  - Observations
  - Summarization
  - Conclusions
  - Technical Document
  - Power Point Presentation

Abhishek Kirar:
- Data Wrangling
  - 1.Loading and Preprocessing
  - 2.Structuring data
  - 3.Enriching data
- Data Mining
- Data Analysis
- Visualizations
  - ➢ Bar graphs and Pie charts

  - Observations
  - Summarization
  - Conclusions
  - Technical Document
  - Power Point Presentation

Sayed Aquib:
- Data Wrangling
  - 1.Loading and Preprocessing
  - 2.Structuring data
  - 3.Enriching data
- Data Mining
- Data Analysis
- Visualizations
  - ➢ Bar graphs and Pie charts

  - Observations
  - Summarization
  - Conclusions
  - Technical Document
  - Power Point Presentation

Please paste the GitHub Repo link.

Github Link: https://github.com/Ajinkya6597

Project Name: The Global Terrorism Analysis (GTA).

Problem Statement:

Data provided by the Global Terrorism Database.csv file is in unformatted manner, uneven data, and duplicate data and also some data columns it is irrelevant, because it's a piled-up data coming from various different countries. For doing the analysis on the data the data needs to be in correct format and well organized formed. The main objective of the analysis is to obtain the meaning full insights and facts from the given huge datasets.

Approach:
As a team we read the data present in the file and gone through the data in each column detailed a s individual. The data set was huge in which some of the dataset was not required for the analysis. The data was cleaned and created new data frame with the columns we required for the analysis. The first problem we faced was the name of the columns and nan values present in the data.We re named   the columns by using a dictionary format, we replaced all nan values to Zero in int dtype and unknown in object dtype by using .replace syntax.Each and every column were compared to gain the knowledge about the data.
Worked   individually gaining some insights doing some EDA. By cleaning the datasets, doing a proper analysis and visualization and plotting the useful information into different graph and charts so that the trend and relationship between the various indicators on which the analysis is done can be understand easily.

Conclusion:
* The year in which maximum attacks took place was during the year 2014.
*   The attacks were seen gradually increasing from 2005 to 2014
* The country which was attacked the most are Iraq, Pakistan, Afghanistan   and India
* Different kind of attacks were done.The major was through Bombing/Explosion as it gives high er damage and Armed Assault.
* The Middle East and North Africa was the most affected region
* The most affected city in the world is Baghdad and Indian City is 'Srinagar'
* We observed that attack mostly takes place on weekdays(Mon,Tue,Wed,Thur) compared to wee kends(Fri, Sat, Sun)
* When we analysed the attack monthwise we observed that every month has equal contribution and date wise most the attack took place on 1st and 15th of the month and minimum attack on the end of the month
* Overall 171378 people were killed during the attack and Iraq country has the highest count of p eople
 killed.
*   The motive behind the attacks was some unknown means only reason behind attack is to sprea d terrorism worldwide and some of them known such as "Protest the Vietnam War and the R.O.T. C." & "Protest the War in Vietnam and sabotage the draft".
* The most common Targets are Private Citizens & Property followed by Military sectors and pol ice
*   Taliban is the most notorious group