**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| NISHA AHIRE  EMAIL id [-nishaahire12345@gmail.com](mailto:-nishaahire12345@gmail.com)  Contribution –  The whole project was handled by me alone. I have done the first step i.e.,  DATA PREPROCESSING that had several steps mentioned below   * DATA CLEANING-   Removing missing values,and outliers   * DATA INTEGRATION-   Subcategories are put togeter to avoid the confusion.   * DATA TRANSFORMATION-   aggregation of subtypes to represent one value by summation of all similar subtypes.   * DIMENSION REDUCTION   removal of columns which carries less significance to our project.  After that using the given data base I have performed exploratory data analysis and founded some useful insights from the data base in the form of various visualizations. And later on concluded my insights in the conclusion. |
| **Please paste the GitHub Repo link.** |
| Github Link:- <https://github.com/Link/to/Repo> |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**  **Given dataset from the Global Terrorism Database had 135 columns,181691 rows and it is was a mix of categorical and numeric values. And given database had incomplete, inconsistent, contains many errors, missing attributes values, contains outliers, incorrect tags, and duplicate entries, name of the columns were not informatory enough to develop proper understanding. so, I had renamed the columns ,i.e. Data preprocessing and this has helped me to resolve these discrepancies, and formation of informatory data so ,that I can extract the required columns for exploring and analyzing the database, to discover the key findings pertaining to terrorist activities.**  **I have used all the rows of the given database, but I decreased the number of columns, because many of the column’s details were not that informatory for me.**  **\*\*\* INSIGHTS OF DATABASE \*\*\***  **\*\*Attacktype1\_txt\*\* = The type of attack happened. Attacktype1\_txt**  **consists of categories like explosion, armed assault,**  **assassination, kidnapping, unarmed assaults.**  **\*\*Target1\_txt\*\* = Type of target involved in the attack. Target1\_txt consists of categorical values like private citizens, military, police,**  **government officials, transportation, education, religious**  **institution, airports, etc.**  **\*\*Success\*\* = ‘1’ if attack was a success. ‘0’ if attack was a failure.**  **Multiple Value for the number of attacks conducted in a single**  **terrorist event.**  **\*\*Natlty1\*\* = Nationality of the attacker.**  **\*\*Weaptype1\*\* = Type of weapon used in the attack. Weaptype1 contains**  **values like firearms, explosives, melee, vehicles etc.**  **\*\*nkill\*\* = Number of people killed in any event**  **\*\*nwonded\*\* = Number of people wounded in any event**  **\*\*gname\*\* = Organization that claimed responsibility of an attack**  **\*\*Region\_txt\*\* = Name of the region where the attack happened. Region\_txt**  **consists values like East Asia, South Asia, Western Europe,**  **etc**  **\*\*Longitude\*\* = Longitude of the location**  **\*\*Latitude\*\* = Latitude of the location**  **\*\*Property\*\* = Total property damage happened in any event**  **\*\*Suicide\*\* = ‘1’ if attack was a suicide attempt. ‘0’ if attack was not a**  **suicide attempt.**  **\*\*Motive\*\* = Known motive of the attacker**  **\*\*Day\*\*, \*\*month\*\*, \*\*year\*\* = Calendar details of the event.**    **This were the columns used my me for the analysis. I have used matplotlib, seaborn for making the visualizations.**  **From the visualizations I founded some useful information mentioned below**  **1. TALIBAN is the most active organisation who is taking responsibility of many of the attacks that had happened in the world.**  **2. IRAQ is the most affected by the attacks, so security groups should be extra cautious in this country.**  **3 .IN 2014, maximum number of attacks has taken place in the world, so it was the dark year of terrorism.**  **4. IT is seen from the database that the explosives were majorly used by the**  **terrorist groups for the attacks.so security groups should try to locate their sources from where they acquire this explosive.**  **5. After doing the analysis we founded that that citizens are the major targets of the terrorist group, and the noticable part is that IRAQ is most affected by the terrorist groups, and maximum number of members of terrorist group belong to the nationality of IRAQ.** |