PROJECT: DATA ANALYTICS ON RELIGIOUS TERRORIST ATTACKS

Use data from The Religion of Peace to predict attacks and save lives

1. Data Cleaning and Exploration

Reading the dataset

```
In [6]: import pandas as pd

df=pd.read_csv(r'C:\Users\sourabh singh\Downloads\Datasets\attacks_data_UTF8.csv')
df
```

Out[6]:

	Unnamed: 0	Date	Country	City	Killed	Injured	Description
0	1	2002-01-01	Indonesia	Palu	1	0	Four bombs explode at Christian churches and o...
1	2	2002-01-01	India	Baramulla	1	0	Terrorists enter the home of a civilian and ki...
2	3	2002-01-01	India	Poshkar	2	0	Two civilians are abducted from their homes th...
3	4	2002-01-02	India	Rajouri	6	9	Three separate terror attacks in the district ...
4	5	2002-01-02	India	Jehangir Chowk	2	25	A Muslim militant kills two and injures twenty...
...	...	...	...	...	...	...	...
29359	29360	2016-07-16	Syria	Aleppo	5	9	Two brothers are two women are among five civi...
29360	29361	2016-07-18	Syria	Aleppo	3	15	Sunnis send rockets into a residential neighbo...
29361	29362	2016-07-18	Yemen	Mukalla	10	18	Ten people are blown to bits by a suicide car ...
29362	29363	2016-07-18	Kazakhstan	Almaty	4	0	An extremist shouting praises to Allah murders...
29363	29364	2016-07-18	Germany	Wuerzburg	0	5	A 'refugee' with an axe hacks at people on a t...

29364 rows × 7 columns

Removing the duplicate values

```
In [12]: df=df.drop_duplicates()
df.shape
```
Out[12]: (29363, 7)

Dropping the null values

```
In [15]: df.isnull().sum()
```
Out[15]:
```
Unnamed: 0     0
Date           0
Country        0
City           0
Killed         0
Injured        0
Description    0
dtype: int64
```

There is no null values.

Printing a concise summary of attacks dataframe

```
In [71]:  df.info()

          <class 'pandas.core.frame.DataFrame'>
          RangeIndex: 29364 entries, 0 to 29363
          Data columns (total 6 columns):
           #   Column       Non-Null Count  Dtype
          ---  ------       --------------  -----
           0   Date         29364 non-null  object
           1   Country      29364 non-null  object
           2   City         29364 non-null  object
           3   Killed       29364 non-null  int64
           4   Injured      29364 non-null  int64
           5   Description  29363 non-null  object
          dtypes: int64(2), object(4)
          memory usage: 1.3+ MB
```

Columns avaliable in the data set

```
In [18]:  df.columns

Out[18]:  Index(['Unnamed: 0', 'Date', 'Country', 'City', 'Killed', 'Injured',
                 'Description'],
                dtype='object')
```

Have a look at a random sample of n rows

```
In [19]:  df.head(10)
```

Out[19]:

	Unnamed: 0	Date	Country	City	Killed	Injured	Description
0	1	2002-01-01	Indonesia	Palu	1	0	Four bombs explode at Christian churches and o...
1	2	2002-01-01	India	Baramulla	1	0	Terrorists enter the home of a civilian and ki...
2	3	2002-01-01	India	Poshkar	2	0	Two civilians are abducted from their homes th...
3	4	2002-01-02	India	Rajouri	6	9	Three separate terror attacks in the district ...
4	5	2002-01-02	India	Jehangir Chowk	2	25	A Muslim militant kills two and injures twenty...
5	6	2002-01-03	Pakistan	Karachi	1	0	A cleric is gunned down in a sectarian attack.
6	7	2002-01-04	India	Kashmir	2	0	Two civilians are abducted and killed by the M...
7	8	2002-01-05	India	Mahore	2	3	An ambush on a security patrol leaves two memb...
8	9	2002-01-06	India	Ramsu	3	12	Three brothers are killed in their home by Las...
9	10	2002-01-06	India	Luddu	4	4	Four civilians, three of whom were in the same...

Basic Statistics functions in following columns

```
In [202]:  df[['Killed','Injured','Total_Attacks']].describe()
```

Out[202]:

	Killed	Injured	Total_Attacks
count	29364.000000	29364.000000	29364.000000
mean	6.659822	9.399537	16.059358
std	24.164587	31.126073	46.156065
min	0.000000	0.000000	0.000000
25%	1.000000	0.000000	2.000000
50%	2.000000	1.000000	5.000000
75%	6.000000	8.000000	15.000000
max	2000.000000	1841.000000	2500.000000

2. Data Analysis Using Different Functions and Visualisations

Total number of Terrorist Attacks

```
In [20]:  Total = df.apply(lambda x: x['Killed'] + x['Injured'], axis=1)

          print('Total number of Terrorist Attacks=',Total.sum())

          Total number of Terrorist Attacks= 471553
```

Attack in top 10 countries

```
In [22]: import pandas as pd
         import matplotlib.pyplot as plt
         df=pd.read_csv(r'C:\Users\sourabh singh\Downloads\Datasets\attacks_data_UTF8.csv')
         df['Total_Attacks'] = Total

         %matplotlib inline
         df.groupby('Country').Total_Attacks.sum().sort_values(ascending=False).head(10).plot(kind='bar',color='Maroon')
         plt.style.use('_classic_test_patch')
         plt.title('Attack in top 10 countries',color='maroon')
         label=plt.xlabel('Country')
         label.set_color('maroon')
         plt.show()
```
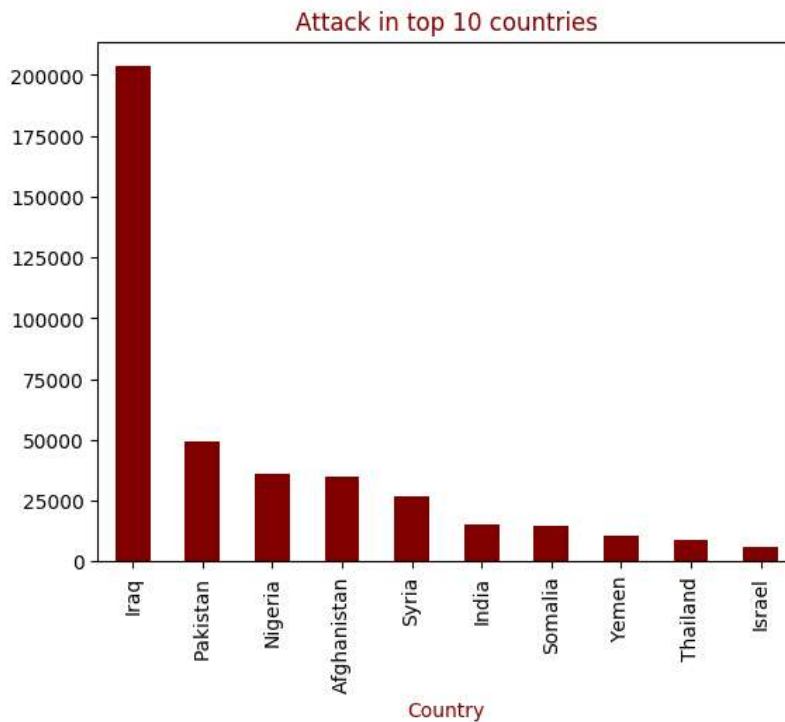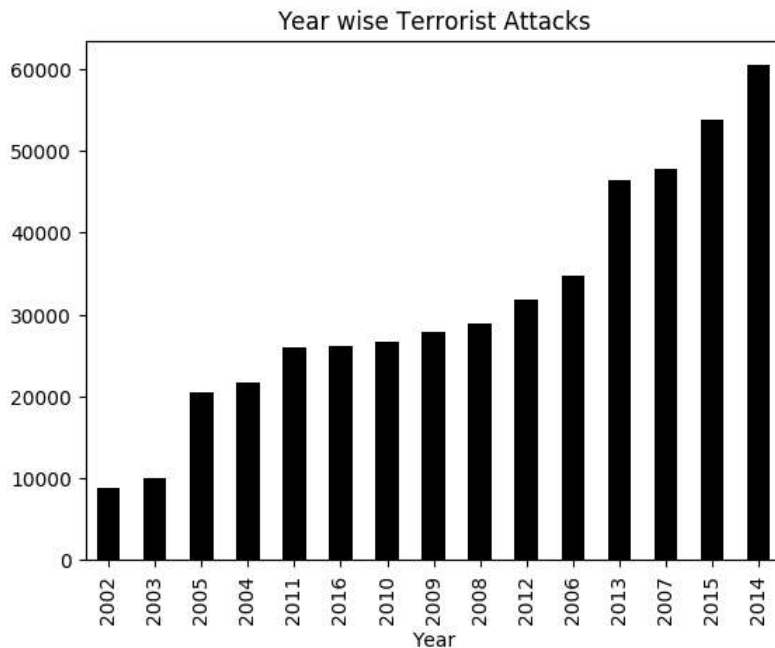

Year wise Terrorist Attacks

```
In [42]:  df['Date']=pd.to_datetime(df.Date)
          df['Year']=df['Date'].dt.year

          %matplotlib inline
          df.groupby('Year').Total_Attacks.sum().sort_values(ascending=True).plot(kind='bar',color='black')
          plt.style.use('_classic_test_patch')
          plt.title('Year wise Terrorist Attacks',color='Black')

Out[42]:  Text(0.5, 1.0, 'Year wise Terrorist Attacks')
```
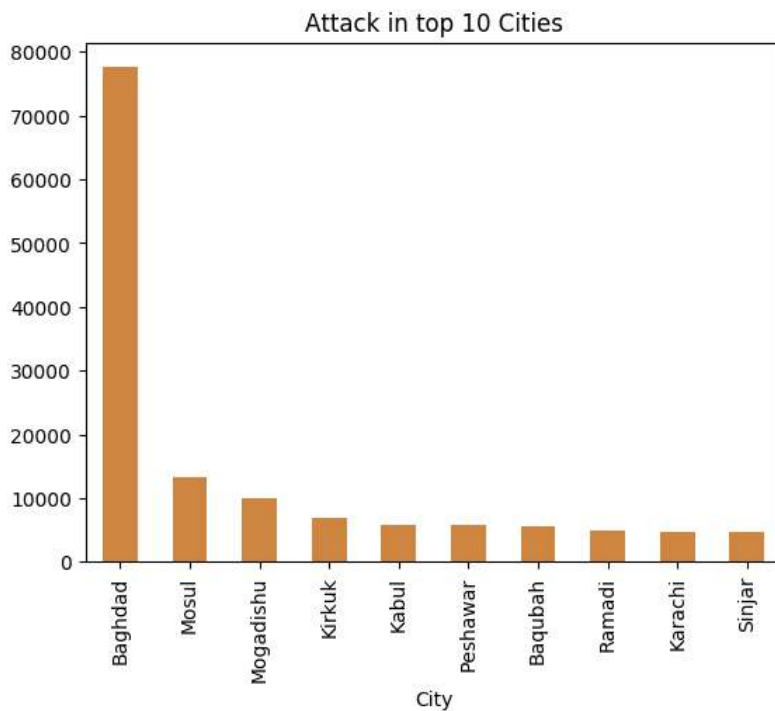

Attack in Top 10 Cities

```
In [29]:  %matplotlib inline
          top10City=df.groupby('City').Total_Attacks.sum().sort_values(ascending=False).head(10).plot(kind='bar',color='peru
          plt.title('Attack in top 10 Cities',color='Black')
          plt.style.use('_classic_test_patch')
          label=plt.xlabel('City')
          label.set_color('black')
```


KILLED Vs INJURED

```
In [41]: Total_killed=df['Killed'].sum()
         Total_Injured=df['Injured'].sum()
         Total_Injured,Total_killed,df['Total_Attacks'].sum()
         value=[276008, 195559]
         label=['Killed','Injured']
         colors=['black','red']
         plt.pie(value,labels=label,colors=colors,autopct = '%1.1f%%',wedgeprops = {"linewidth": 1, "edgecolor": "white"})
         plt.title('KILLED Vs INJURED')
```

Out[41]: Text(0.5, 1.0, 'KILLED Vs INJURED')

KILLED Vs INJURED