```
# -*- coding: utf-8 -*-
"""EDA on  golbal terrorism.ipynb

Automatically
generated by Colaboratory.

Original file is located at

https://colab.research.google.com/drive/1JfrlD4v61CIor6QX4enUDLaWmac1bV_4

## EDA Analysis on
Terrorism
##Import libraries
"""

import math
import numpy as np
import pandas
as pd
import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in
filenames:
        print(os.path.join(dirname, filename))

"""##Load and explore
the dataset

"""



import plotly.offline as py
import plotly.graph_objs as
go
import matplotlib.pyplot as plt
import seaborn as sns
import
warnings
warnings.filterwarnings('ignore')

df =
pd.read_csv('/content/globalterrorismdb_0718dist.csv',encoding = "ISO-8859-1")

#
Display the last few rows of the dataset
df.tail()

df.columns

#Renaming the
columns

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)

#As there are too
many columns in dataset, we are taking only important columns from the dataset for data
preprocessing

df =
df[['Year','Month','Day','Country','state','Region','city','latitude','longitude','AttackType',
'Killed',

'Wounded','Target','Summary','Group','Target_type','Weapon_type','Motive']]

#Cheching the null
```

```
values

df.isnull().sum()

df.shape

df.info()

df.describe

df['Year'].value_counts(dropna =
False).sort_index()

from wordcloud import WordCloud
from scipy import signal
cities =
df.state.dropna()
plt.subplots(figsize=(10,10))
wordcloud = WordCloud(background_color =
'white',
                      width = 550,
                      height = 390).generate('
'.join(cities))
plt.axis('off')
plt.imshow(wordcloud)
plt.show()

"""##Data
visualization
people killed and wounded in each
year
"""

b=df[["Year","Wounded"]].groupby("Year")
.sum()
b.head()

b.plot(kind="bar",color="purple",figsize=(15,6))
plt.title
("Year wise
Wounded",fontsize=13)
plt.xlabel("Years",fontsize=13)
plt.xticks(fontsize=12)
pl
t.ylabel("Number of
Wounded",fontsize=13)
plt.show()

k=df[["Year","Killed"]].groupby(&quo
t;Year").sum()
k.head()

w=df[["Year","Wounded"]].groupby("Year&q
uot;).sum()
w.head()

fig=plt.figure()
ax0=fig.add_subplot(2,1,1)
ax1=fig.add_subplot(2,1,2)

#
Killed
k.plot(kind="bar",color="green",figsize=(15,15),ax=ax0)
ax0.set_titl
e("People Killed in each
Year")
ax0.set_xlabel("Years")
ax0.set_ylabel("Number of People
Killed")

#Wounded
w.plot(kind="bar",color="blue",figsize=(15,15),ax=a
x1)
ax1.set_title("People Wounded in each
Year")
```

```
ax1.set_xlabel("Years")
ax1.set_ylabel("Number of People
Wounded")

"""##Number of Terrorist Activities each
Year"""

x_year = df['Year'].unique()
y_count_years =
df['Year'].value_counts(dropna = False).sort_index()
plt.figure(figsize =
(18,10))
sns.barplot(x = x_year,
            y = y_count_years,
            palette =
'rocket')
plt.xticks(rotation = 45)
plt.xlabel('Attack Year')
plt.ylabel('Number of Attacks
each year')
plt.title('Attack Of Years')
plt.show()

import seaborn as sns
import
matplotlib.pyplot as plt

plt.subplots(figsize=(15, 6))
sns.countplot(x='Year', data=df,
palette='RdYlGn_r', edgecolor=sns.color_palette("YlOrBr",
10))
plt.xticks(rotation=45)
plt.title('Number Of Terrorist Activities Each
Year')

"""##Terrorist Activities by Region in each Year through Area
Plot"""

pd.crosstab(df.Year,
df.Region).plot(kind='area',figsize=(15,6))
plt.title('Terrorist Activities by Region in each
Year')
plt.ylabel('Number of Attacks')
plt.show()

df['Wounded'] =
df['Wounded'].fillna(0).astype(int)
df['Killed'] =
df['Killed'].fillna(0).astype(int)
df['casualities'] = df['Killed'] +
df['Wounded']

"""##Heatmap"""

t1 =
df.sort_values(by='casualities',ascending=False)[:40]
heat=t1.pivot_table(index='Country',colum
ns='Year',values='casualities')
heat.fillna(0,inplace=True)
heat.head()

import plotly.offline
as py
py.init_notebook_mode(connected=True)
import plotly.graph_objs as go
colorscale = [[0,
'#edf8fb'], [.3, '#00BFFF'],  [.6, '#8856a7'],  [1, '#810f7c']]
heatmap =
go.Heatmap(z=heat.values, x=heat.columns, y=heat.index, colorscale=colorscale)
data =
[heatmap]
layout = go.Layout(
    title='Top 40 Worst Terror Attacks in History from 1982 to
2016',
    xaxis = dict(ticks='', nticks=20),
```

```
    yaxis = dict(ticks='')
)
fig =
go.Figure(data=data, layout=layout)
py.iplot(fig,
filename='heatmap',show_link=False)

reg=pd.crosstab(df.Year,df.Region)
reg.head()

regt=reg.tr
anspose()
regt["Total"]=regt.sum(axis=1)
a=regt["Total"].sort_values(ascend
ing=False)
a

a.plot(kind="bar",figsize=(15,6))
plt.title("Total Number of
Attacks in each Region from
1970-2017")
plt.xlabel("Region")
plt.ylabel("Number of
Attacks")
plt.show()

"""##Top Countries affected by Terror
Attacks"""

df.Country.value_counts()[:15]

import seaborn as sns
import
matplotlib.pyplot as plt

plt.subplots(figsize=(15, 6))
top_countries =
df['Country'].value_counts()[:15]
sns.barplot(x=top_countries.index, y=top_countries.values,
palette='rainbow')
plt.title('Top Countries
Affected')
plt.xlabel('Countries')
plt.ylabel('Count')
plt.xticks(rotation=
90)
plt.show()

"""##Customized Data Analysis


Terrorist Attacks of a
Particular year and their Locations
"""

import folium
from folium.plugins
import MarkerCluster
filterYear = df['Year'] == 1974

filterData = df[filterYear] # filter
data
# filterData.info()
reqFilterData = filterData.loc[:,'city':'longitude'] #We are getting
the required fields
reqFilterData = reqFilterData.dropna() # drop NaN values in latitude and
longitude
reqFilterDataList = reqFilterData.values.tolist()
# reqFilterDataList

map =
folium.Map(location = [0, 30], tiles='CartoDB positron', zoom_start=2)
# clustered
marker
markerCluster = folium.plugins.MarkerCluster().add_to(map)
```

```
for point in range(0,
len(reqFilterDataList)):

folium.Marker(location=[reqFilterDataList[point][1],reqFilterDataList[point][2]],

    popup = reqFilterDataList[point][0]).add_to(markerCluster)
map

"""In 1970,
the American continent accounted for 84% of all terrorist attacks, while the Middle East and
North Africa, which are now known as hotspots for conflicts and terrorism, experienced just one
terrorist incident during that year.

##Types of terrorist attacks that cause
deaths
"""

# Total Number of people killed in terror attack
killData =
df.loc[:,'Killed']
print('Number of people killed by terror attack:',
int(sum(killData.dropna())))# drop the NaN values

# Let's look at what types of attacks these
deaths were made of.
attackData = df.loc[:,'AttackType']
# attackData
typeKillData =
pd.concat([attackData, killData], axis=1)
typeKillData.head()

typeKillFormatData =
typeKillData.pivot_table(columns='AttackType', values='Killed',
aggfunc='sum')
typeKillFormatData

typeKillFormatData.info()

import numpy as np

labels =
typeKillFormatData.columns.tolist() # convert line to list
transpoze = typeKillFormatData.T #
transpoze

# Assuming values is a 2D array
values = transpoze.values.tolist()
values =
np.array(values).flatten()  # Flatten the 2D array to make it 1D

fig, ax =
plt.subplots(figsize=(20, 20), subplot_kw=dict(aspect="equal"))
plt.pie(values,
startangle=90, autopct='%.2f%%')
plt.title('Types of terrorist attacks that cause
deaths')
plt.legend(labels, loc='upper right', bbox_to_anchor=(1.3, 0.9), fontsize=15)  #
location legend
plt.show()

"""The combination of armed assaults and
bombings/explosions accounts for a significant portion, specifically 77%, of the fatalities
resulting from these attacks. This high fatality rate underscores the recurring use of these
tactics in terrorist activities, highlighting the profound threat posed by weapons and
explosives on a global scale.

##Number of People Killed in Terrorist Attacks (By
Countries)
"""

#Number of Killed in Terrorist Attacks by Countries
countryData
= df.loc[:,'Country']
# countyData
countryKillData = pd.concat([countryData, killData],
```

```
axis=1)

countryKillFormatData = countryKillData.pivot_table(columns='Country',
values='Killed', aggfunc='sum')
countryKillFormatData

fig_size =
plt.rcParams["figure.figsize"]
fig_size[0]=25
fig_size[1]=25
plt.rcParams["figur
e.figsize"] = fig_size

labels = countryKillFormatData.columns.tolist()
labels =
labels[152:206]
index = np.arange(len(labels))
transpoze = countryKillFormatData.T
values =
transpoze.values.tolist()
values = values[152:206]
values = [int(i[0]) for i in values]
colors
= ['red', 'black', 'blue', 'purple', 'yellow', 'brown', 'black', 'teal', 'magenta',
'orange']
fig, ax = plt.subplots(1, 1)
ax.yaxis.grid(True)
fig_size =
plt.rcParams["figure.figsize"]
fig_size[0]=25
fig_size[1]=25
plt.rcParams["figur
e.figsize"] = fig_size
plt.bar(index, values, color = colors, width =
0.9)
plt.ylabel('Killed People', fontsize=20)
plt.xlabel('Countries', fontsize =
20)
plt.xticks(index, labels, fontsize=18, rotation=90)
plt.title('Number of people killed by
countries', fontsize = 20)
plt.show()

"""##Region Wise
Casuality"""

c=df[["Region","Wounded"]].groupby("Regio
n").sum().sort_values(by="Wounded",ascending=False)
c

c.plot(kind="bar&quo
t;,color="khaki",figsize=(15,6))
plt.title("Region wise
Casualty",fontsize=13)
plt.xlabel("Regions",fontsize=13)
plt.xticks(fontsize=12)

plt.ylabel("Number of Casualty",fontsize=13)
plt.show()

"""##People
Killed and Wounded In Each
Region"""

k=df[["Region","Killed"]].groupby("Region&qu
ot;).sum().sort_values(by="Killed",ascending=False)
k

w=df[["Region",&quot
;Wounded"]].groupby("Region").sum().sort_values(by="Wounded",ascending
=False)
w

fig=plt.figure()
ax0=fig.add_subplot(1,2,1)
```

```
ax1=fig.add_subplot(1,2,2)

#People
Killed
k.plot(kind="bar",color="indigo",figsize=(15,6),ax=ax0)
ax0.set_titl
e("People Killed in each
Region")
ax0.set_xlabel("Regions")
ax0.set_ylabel("Number of People
Killed")

#People
Wounded
w.plot(kind="bar",color="olive",figsize=(15,6),ax=ax1)
ax1.set_titl
e("People Wounded in each
Region")
ax1.set_xlabel("Regions")
ax1.set_ylabel("Number of People
Wounded")

plt.show

"""##Conclusion
Terrorist incidents in the Middle East
and Northern Africa region have been associated with severe and deadly outcomes. This region
has been identified as a hotspot for significant terrorist attacks. Contrary to a common
perception linking Muslims with terrorism, it is worth noting that Muslims are often the
primary victims of these terrorist acts. A closer examination of the data visualizations
reveals that Iraq, Afghanistan, and Pakistan are among the most adversely affected nations, and
it's noteworthy that these countries are predominantly Muslim.
"""
```