Report

Team Names and ID:
Sai Rohith Pasala; 1001873156
Sri Naga Venkata Pavan Kalyan Sirigibattula; 1001886149
Rutvik Naga Sai Dondapati; 1001879177

Analysis of Disaster Management

Data set details:

Analysing the disasters over 63 years of Fedaral disasters dataset. The data set is obtained from database.csv file from kaggle repository(https://www.kaggle.com/fema/federal-disasters). database.csv contains 46,184 rows and 14 columns. This dataset begins with the year 1953, and runs up to the year 2017

Context:

The president can declare an emergency for any occasion or instance when the President determines federal assistance is needed. Emergency declarations supplement State and local or Indian tribal government efforts in providing emergency services, such as the protection of lives, property, public health, and safety, or to lessen or avert the threat of a catastrophe in any part of the United States. The total amount of assistance provided for in a single emergency may not exceed $5 million.

The president can declare a major disaster for any natural event, including any hurricane, tornado, storm, high water, wind-driven water, tidal wave, tsunami, earthquake, volcanic eruption, landslide, mudslide, snowstorm, or drought, or, regardless of cause, fire, flood, or explosion, that the President determines has caused damage of such severity that it is beyond the combined capabilities of state and local governments to respond. A major disaster declaration provides a wide range of federal assistance programs for individuals and public infrastructure, including funds for both emergency and permanent work.

Given below is the description of the columns of database.csv

Column Name	Description
Declaration Number	This is the unique Number assigned to the each of the situation emerged
Declaration Date	The first day the disaster has been recorded
State	State of occurence
County	County of the state
Disaster Type	What exactly is the type of disaster occured
Disaster Title	Title or the specific name given to the Disaster
Start Date	Starting date of the disaster occurred
End Date	Ending date of the disaster occurred
Close Date	Date of end Declaration
Individual Assistance Program	Whether Individual Assistance is provided or not
Individuals & Households Program	Whether Individual & Households is provided or not
Public Assistance Program	Whether Public Assistance is provided is provided or not
Hazard Mitigation Program	Whether Hazard Mitigation is provided or not

<u>Task 1</u> : Required to obtain some basic information of the database using pandas framework.

Printing the details of the df_data data frame (information such as number of rows,columns, name of columns, etc) using the function:
df_data.info( )

printing the number of rows and columns in the df_data data frame using the function:
num_rows = len(df_data)
num_cols = len(df_data.columns)

Printing the descriptive detail (count, unique, top, freq etc) for 'Start Date' column of the df_data using function:
df_data['Start Date'].describe( )

Printing all the unique values of Disaster Title and Counties:
list(df_data['Disaster Title'].unique( )
list(df_data['County'].unique( )

Output:

```
RangeIndex: 46185 entries, 0 to 46184
Data columns (total 14 columns):
 #    Column                           Non-Null Count    Dtype
---   ------                           --------------    -----
 0    Declaration Number               46185 non-null    object
 1    Declaration Type                 46185 non-null    object
 2    Declaration Date                 46185 non-null    object
 3    State                            46185 non-null    object
 4    County                           45988 non-null    object
 5    Disaster Type                    46185 non-null    object
 6    Disaster Title                   46185 non-null    object
 7    Start Date                       46185 non-null    object
 8    End Date                         45843 non-null    object
 9    Close Date                       35210 non-null    object
 10   Individual Assistance Program    46185 non-null    object
 11   Individuals & Households Program 46185 non-null    object
 12   Public Assistance Program        46185 non-null    object
 13   Hazard Mitigation Program        46185 non-null    object
dtypes: object(14)
memory usage: 4.9+ MB
```

```
>>Task 1-b: Number of rows:46185 and number of columns:14


>>Task 1-c: Descriptive details of 'Start Date' column are
 count         46185
unique         2324
top       8/29/2005
freq           2764
Name: Start Date, dtype: object


 >>Task 1-d:
['Tornado', 'Tornado and Heavy Rainfall', 'Flood', 'Floods', 'Forest Fire', 'Severe Hardship', 'Flood and Erosion',
 'Earthquake', 'Hurricanes', 'Hurricane', 'Volcano', 'Flood and Tornado', 'Hurricane and Flood', 'Floods and Rains',
 'Hurricane and Floods', 'Hurricane, Torrential Rain, and Floods', 'Tornadoes', 'Severe Storm', 'Wind Storm', 'Storm a
nd Flood', 'Storm', 'Tidal Wave', 'Hurricane, Rain, Wind, Hail, and Floods', 'Floods and Hurricane', 'Tornadoes and F
loods', 'Tornadoes, Rain, Hail, and Floods', 'Heavy Rainstorms and Flood', 'Heavy Rainstorms and Floods', 'Tornadoes,
Rain, Hail and Floods', 'Hurricane and Severe Storm', 'Hurricane Dot', 'Heavy Rains and Floods', 'Earthquakes and Vol
canic Disturbances', 'Severe Weather Conditions', 'Tidal Waves', 'Heavy Rains, Hail, Floods, and Tornadoes', 'Fires',
```

Task 2:

Finding out the declaration titles of type 'Emergency' which lasted greater than 120 days using function: pd.to_datetime

```
df_data['Start Date'] = pd.to_datetime(df_data['Start Date'])
df_data['End Date'] = pd.to_datetime(df_data['End Date'])
df_data['Declaration Date'] = pd.to_datetime(df_data['Declaration Date'])

Emergency_greater_120 = ", ".join(list(df_data[(df_data['End Date'] - df_data['Start
Date']).dt.days > 120][df_data['Declaration Type'] == 'Emergency']['Disaster
Title'].unique()))
print (">>Task 2-a: The declaration titles of type 'Emergency' which lasted more
than 120 days are %s"
    % (Emergency_greater_120))
```

Finding out the total number of declarations between 1998 and 1999 where Individual Assistance Program was provided for the state of Texas.

```
num_declarations_1998_1999 = len(df_data[(df_data['State'] == 'TX') &
(df_data['Declaration Date'].dt.year.isin( [1998,1999])) & (df_data['Individual
Assistance Program'] == 'Yes')])

print ("\n\n>>Task 2-b: The total number of declarations in 1998-1999 for the state
of Texas were %s"
    % (num_declarations_1998_1999))
```

Finding out the top 10 counties with the declaration type 'Disaster' and 'Emergency' for all years using functions:

```
top10_disasters="\n".join(list(df_data[df_data['Declaration
Type'].isin(['Disaster'])]['County'].value_counts().index[:n]))
```

top10_emergencies="\n".join(list(df_data[df_data['Declaration Type'].isin(['Emergency'])]['County'].value_counts().index[:n]))
Output:

```
>>Task 2-a: The declaration titles of type 'Emergency' which lasted more than 120 days are Extreme  Fire Hazard, Fire, Extreme Fire Hazards, West Nile Virus, Drought, Wildfires


>>Task 2-b: The total number of declarations in 1998-1999 for the state of Texas were 322


>>Task 2-c: top 10 counties with the most disasters for all years are:
Washington County
Jefferson County
Jackson County
Franklin County
Lincoln County
Clay County
Madison County
Montgomery County
Monroe County
Marion County


>>Task 2-c: top 10 counties with the most emergencies for all years are:
Washington County
Franklin County
Jackson County
Jefferson County
Montgomery County
Monroe County
Lincoln County
Madison County
Warren County
Marion County
```
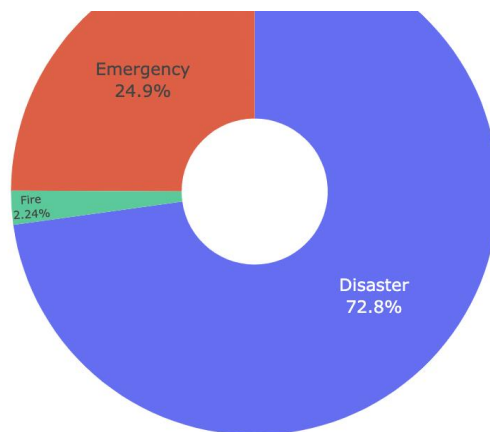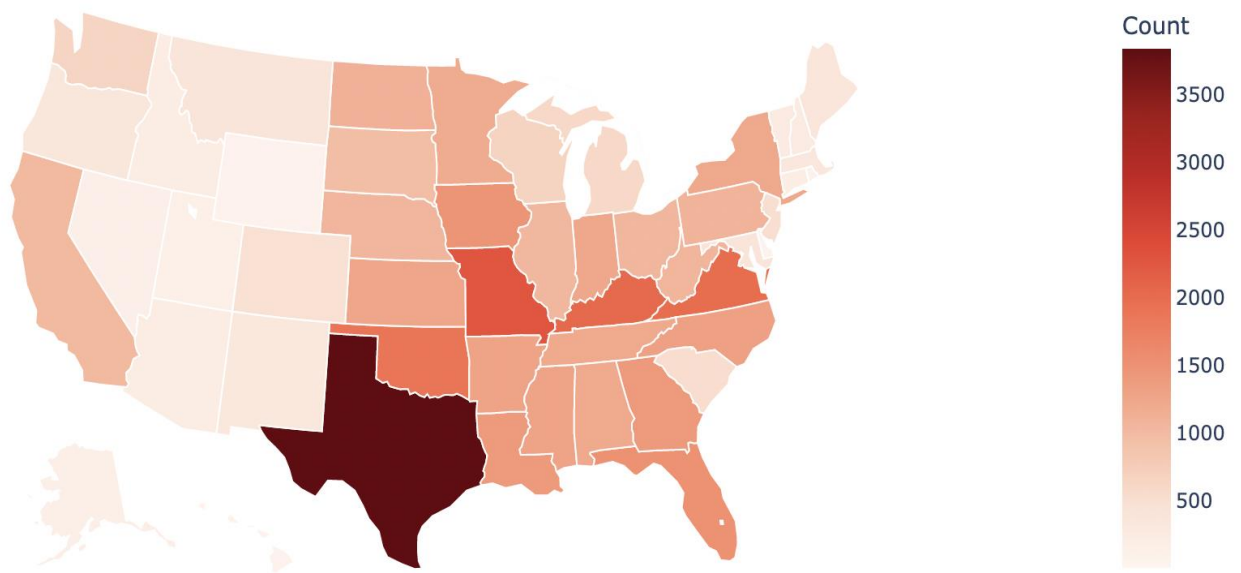
Task 3:

Plotting declaration count for all the declaration types with respect to all the states using the below function

Output:

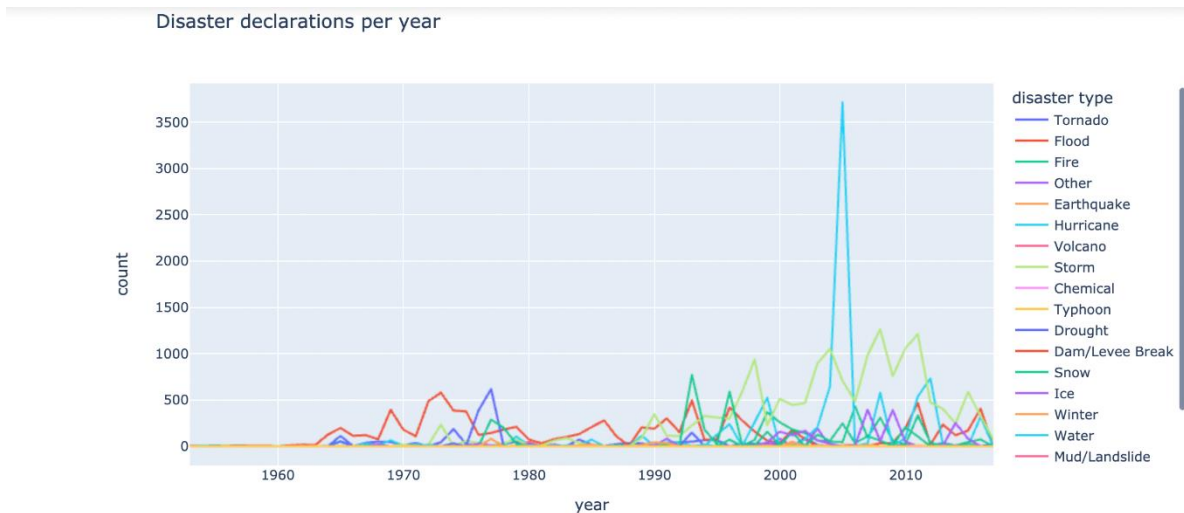United States Fedeal Disaster Declarations Count (1953-2017)

Task 4:

Finding out an 'interesting' information from the dataset and Creating a visualization for it.
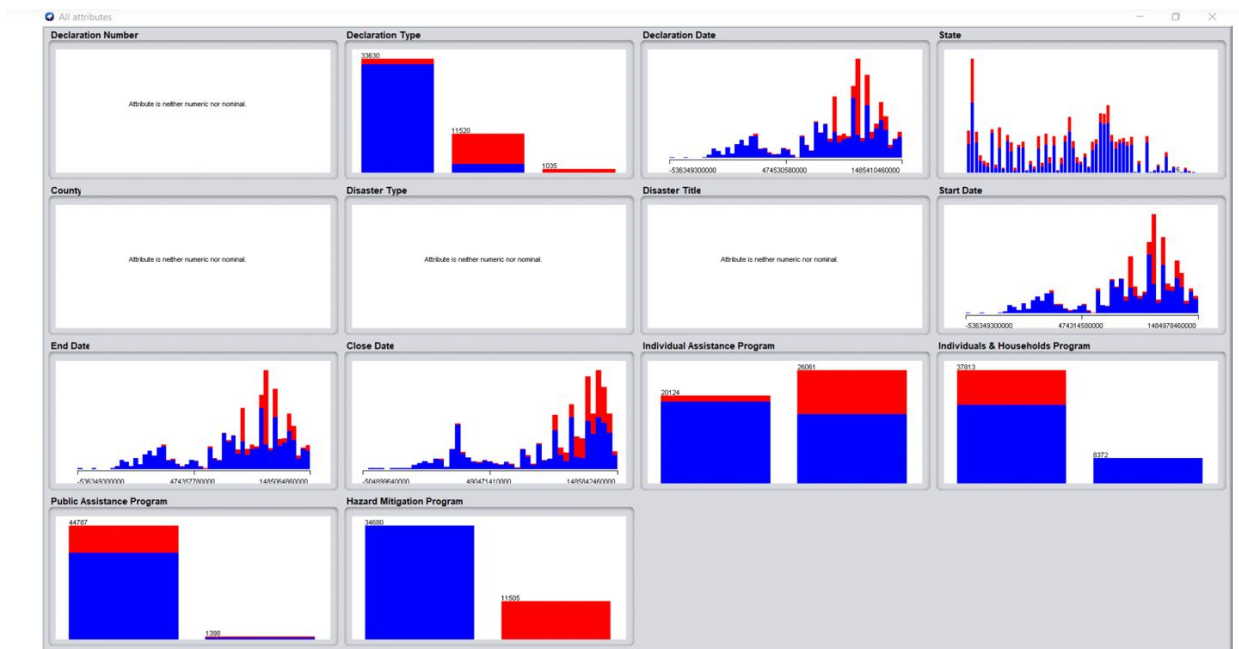
Output:

The above graph shows, Disaster types summed over year and averaged out to get line graph, hover mouse over to get more details.
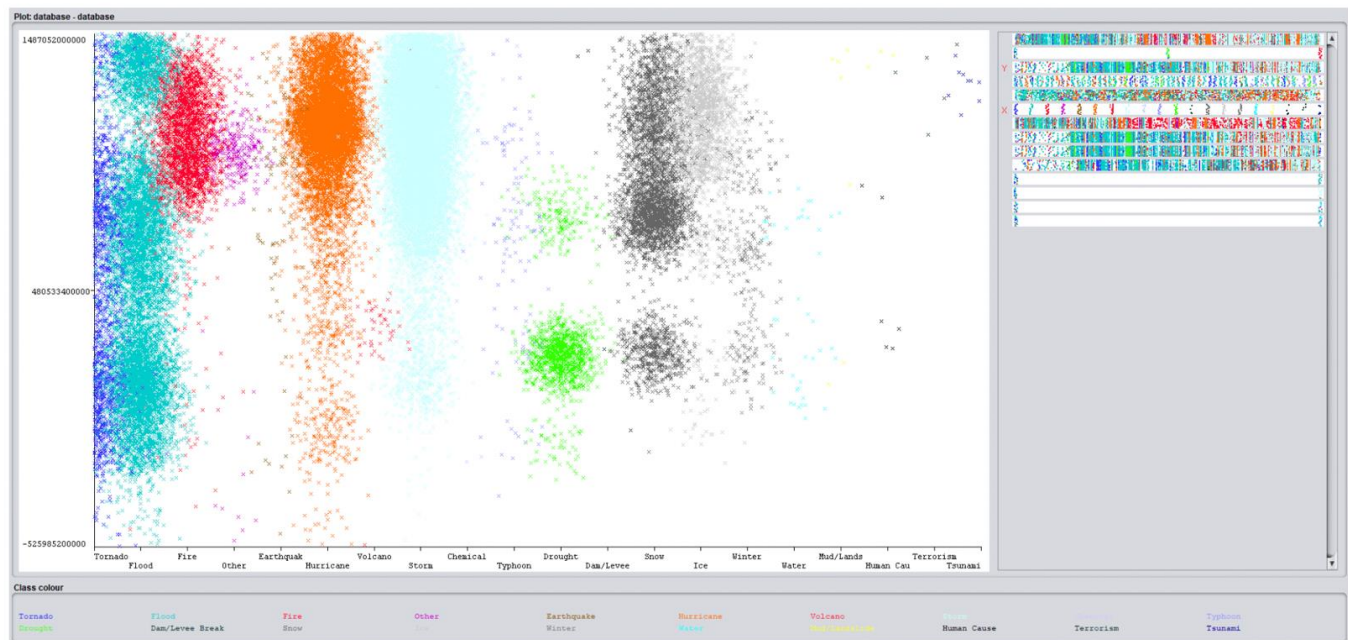
Using Weka

Below are some screenshots using the weka visualizer.

Disaster Types over time using weka

Special Pattern Observed is:
It is evident to see that we are observing more water, Storm, Flood, Hurricane related disasters recently and it has been a while, since mankind has seen any drought.