HAMOYE PREMIERE PROJECT

PRESENTED BY

MODEL-MAKERS TEAM

TO

HAMOYE AI LABS

TOPIC: COMPARATIVE ANALYSIS OF DROUGHT IN AFRICA AND ITS EFFECTS SINCE THE 21ST CENTURY

ON

11TH NOVEMBER, 2023.

TEAM MEMBERS	ROLE
AMADI EMMANUEL	MEMBER

INTRODUCTION

Drought occurs as a result of little or no rainfall over a certain period of time in a particular area. It is a recurring and pervasive environmental challenge that affects numerous African countries. As the 21st century progresses, understanding the evolving patterns of drought and its consequences has become increasingly important. This comparative analysis seeks to provide insights into the distribution, severity, and impacts of drought in African nations during the 21st century. By examining available data, we aim to shed light on the changing dynamics of drought and the measures taken to address its effects.

PROBLEM STATEMENT

The problem at hand is the need to comprehensively understand these evolving patterns of drought and their multifaceted effects to inform targeted and effective mitigation strategies.

Data Source and Specification

For this analysis, we utilized a comprehensive dataset that includes information on drought occurrences in African countries since the year 2000. The dataset was taken from the EM-DAT page— https://public.emdat.be/graphql/files/public_emdat_custom_request_2023-11-08_c8191d8f-e11d-4a4e-86fe-6e1e591bea81.xlsx

The size of the dataset was 168 rows by 46 columns and it covered various parameters, including the duration, intensity and geographic location of drought events. It also provided socio-economic and environmental data to assess the consequences of these events.

AIM AND OBJECTIVES

The aim of this study is to comprehensively analyze the distribution, severity and impacts of drought in African countries since the beginning of the 21st century. By comparing and contrasting these aspects across regions and time, the research seeks to provide valuable insights into the evolving nature of drought in Africa.

SOLUTION PROCESS / METHODOLOGY

The following steps were undertaken in the execution of the project:

- Data Inspection and Cleaning
- Data Analysis
- Data Visualization

Data Inspection and Cleaning:

After a careful study of the dataset given, we came to a conclusion that the data on the dataset was not subsistent enough for model training and so, we decided to do an analysis of the dataset. The process of data inspection involved everyone trimming the dataset until an acceptable and cleaned data was realized. This process involved the following steps:

1. Handling Missing Data: The dataset contained missing data as well as null values which were inspected using the appropriate python code as shown below:

```
#check for null values
df.isnull().sum()
```

```
DisNo.                                    0
Historic                                  0
Classification Key                        0
Disaster Group                            0
Disaster Subgroup                         0
Disaster Type                             0
Disaster Subtype                          0
External IDs                            111
Event Name                              168
ISO                                       0
Country                                   0
Subregion                                 0
Region                                    0
Location                                 17
Origin                                  180
Associated Types                         47
OFDA Response                             0
Appeal                                    0
Declaration                               0
AID Contribution ('000 US$)             162
Magnitude                               164
Magnitude Scale                           0
Latitude                                168
Longitude                               168
River Basin                             168
Start Year                                0
Start Month                              35
Start Day                               167
End Year                                  0
End Month                                66
End Day                                 166
Total Deaths                            155
No. Injured                             168
No. Affected                             12
No. Homeless                            168
Total Affected                           12
Reconstruction Costs ('000 US$)         168
Reconstruction Costs, Adjusted ('000 US$)  168
Insured Damage ('000 US$)               168
Insured Damage, Adjusted ('000 US$)     168
Total Damage ('000 US$)                 157
Total Damage, Adjusted ('000 US$)       157
CPI                                       0
Admin Units                              24
Entry Date                                0
Last Update                               0
dtype: int64
```

2. The null columns were dropped as well as columns containing high amount of null values as seen below.

```
#removing other insignificant columns with large null values, columns with only one unique value and high missing values
cols_to_drop = ['DisNo.', 'Classification Key',  "Magnitude", "AID Contribution ('000 US$)","End Day", "Start Day","External IDs",
                "AID Contribution ('000 US$)",'Start Day','Historic','Disaster Group',
               'Disaster Subgroup', 'Disaster Type', 'Disaster Subtype', 'Region', 'Magnitude Scale', 'Admin Units', 'Entry Date', 'Last Update', 'No. Affected']
df.drop(columns=cols_to_drop, axis=1, inplace=True)
df.columns
```

```
Index(['ISO', 'Country', 'Subregion', 'Location', 'Origin', 'Associated Types',
       'OFDA Response', 'Appeal', 'Declaration', 'Start Year', 'Start Month',
       'End Year', 'End Month', 'Total Deaths', 'Total Affected',
       'Total Damage ('000 US$)', 'Total Damage, Adjusted ('000 US$)', 'CPI'],
      dtype='object')
```

3. Some numeric cells having null values were filled with zeros (e.g. the Total Affected column) to avoid errors in calculation and to improve accuracy of our analysis. However, some columns were filled with “Not Specified” for the same reason as stated before.

```
[ ] #Fill empty rows with 'not specified' for the columns with object data type
    columns_to_fill = ['Origin', 'Location','Associated Types']
    for column_name in columns_to_fill:
        df.loc[df[column_name].isna(), column_name] = 'Not specified'
    df.head()
```

	ISO	Country	Subregion	Location	Origin	Associated Types	OFDA Response	Appeal	Declaration	Start Year	Start Month	End Year	End Month	Total Deaths	Total Affected	Total Damage ('000 US$)
0	DJI	Djibouti	Sub-Saharan Africa	Ali Sabieh, Dikhil, Djibouti, Obock, Tadjourah...	Not specified	Not specified	Yes	No	No	2001	6.0	2001	0.0	NaN	100000.0	NaN
1	SDN	Sudan	Northern Africa	Northern Darfur, Northern Kordofan, Red Sea pr...	Not specified	Food shortage|Water shortage	No	No	No	2000	1.0	2001	0.0	NaN	2000000.0	NaN

Image 1: Python code assigning “Not Specified” to the null values of certain columns

```
#"Total Affected" : "int64"

df['Total Affected'] = pd.to_numeric(df['Total Affected'], errors='coerce').fillna(0).astype(int)
```

Image 2: Python code assigning “0” to the null values of a column

4. The datatypes of the columns to be used were changed due to the kind of values they contained e.g. Start Year was changed from datatype “object” to “category” and so on.

```
[13]
    # Change data types
    type_dict = { #"Total Affected" : "int64",
                  "Country": 'category',
                  "Subregion": "category",
                  "Origin": "category",
                  "OFDA Response": "category",
                  "Appeal": "category",
                  "Declaration": "category",
                  'Start Year': 'category',
                  'End Year': 'category'
    }

    for key, value in type_dict.items():
      df[key] = df[key].astype(value)
    df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 168 entries, 0 to 167
Data columns (total 46 columns):
 #   Column                                  Non-Null Count  Dtype
---  ------                                  --------------  -----
 0   DisNo.                                  168 non-null    object
 1   Historic                                168 non-null    object
 2   Classification Key                      168 non-null    object
 3   Disaster Group                          168 non-null    object
 4   Disaster Subgroup                       168 non-null    object
```

5. Finally, the dataset “JME_Regional-Classifications.xlsx” was imported from the internet and relevant columns from the dataset were merged to our own dataset as shown:

```
# Import the UNICEF regional data
regions_df = pd.read_excel("/content/drive/MyDrive/Colab Notebooks/JME_Regional-Classifications.xlsx")[['ISO Code', 'UN
regions_df.head()
```

	ISO Code	UN Sub Region	World Bank Income Groups Combined
0	AFG	Southern Asia	Low Income
1	ALB	Southern Europe	Middle Income
2	DZA	Northern Africa	Middle Income
3	AND	Southern Europe	High Income
4	AGO	Middle Africa	Middle Income

```
[ ] # merge the regions data with the original table
    df = pd.merge(left=df, right=regions_df, left_on = 'ISO', right_on = 'ISO Code', how='left')
    df = df.drop(columns=['ISO Code'])
    df
```

Data Analysis (Exploratory Data Analysis)

Research Questions

The dataset contains information on the occurrence of drought in Africa. This analysis aims to answer the following:

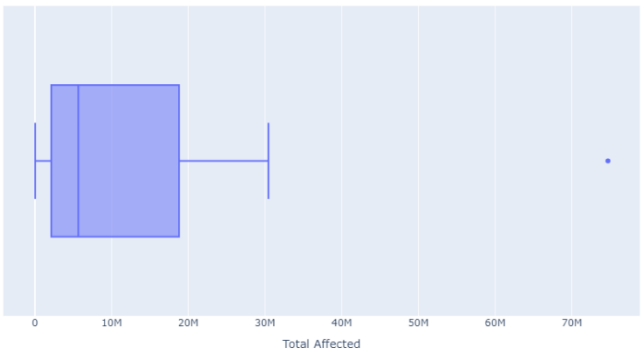
1. who were the worst hit? what is their distribution per:

- country
- region
- income group

After the data was cleaned, we went on to perform exploratory data analysis. Statistical computations were made and visualizations created with the goal of addressing the problem statements written in the earlier part of this documentation and some research questions as shown above. This exploratory data analysis was done in two parts: Uni-variate and Multi-variate analysis.

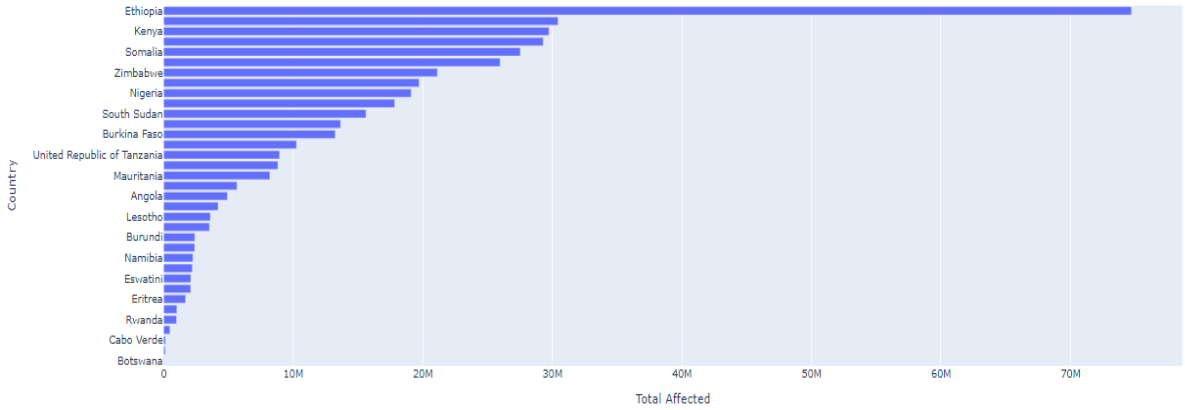
1. Uni-variate Analysis: Some of these are shown Below:

Box Plot of countries affected vs total affected



The box plot showed the quartile range, minimum, maximum points and the median. This information helped us know where the region of the most total affected was. Most values fell before **30 million,** while an outlier country had a value of **74.7million.**

Bar Plot



OBSERVATION

From the above analysis, the following countries were the most hit in terms of the number of persons affected by drought:

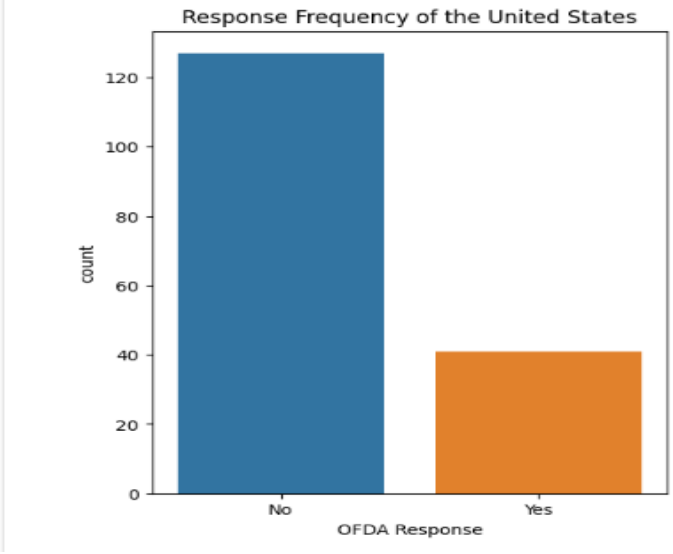
- Ethiopia
- Kenya
- Somalia
- Zimbabwe
- Nigeria.

Interestingly, top 3 of the top 5 countries most Affected were still eastern African countries. Why exactly is Eastern Africa most affected by drought? Could it be a geographical factor or as a result of poor emergency response on the part of the government?

This question calls for further investigations as our data cannot provide an answer to it!

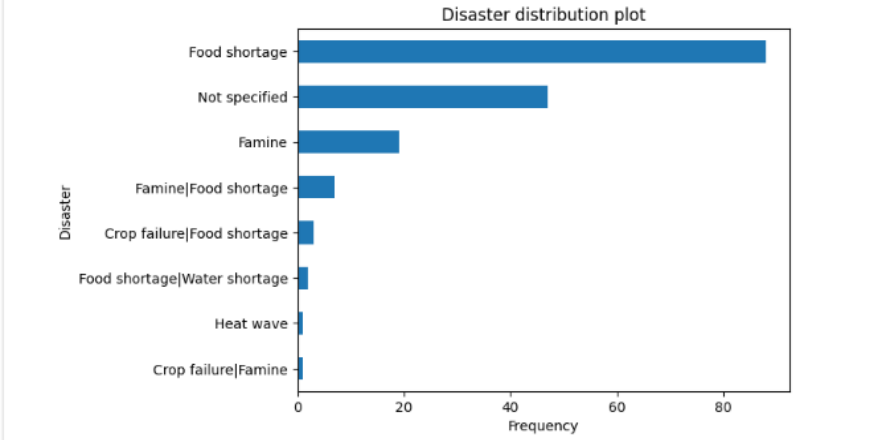
Other data visuals produced are as follows:

```
plt.figure(figsize=(5, 6))
sns.countplot(data=df, x="OFDA Response")
plt.title('Response Frequency of the United States')
plt.xlabel("OFDA Response")
plt.show()
```

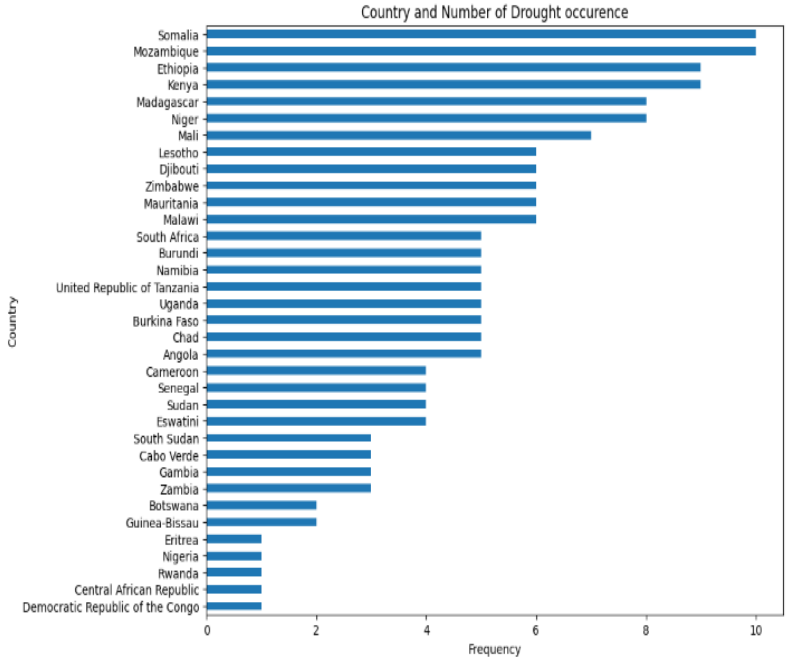


The bar chart above shows that the response of the united states department of foreign affairs to the plight of drought victims is very low. They respond about one in 3 times.

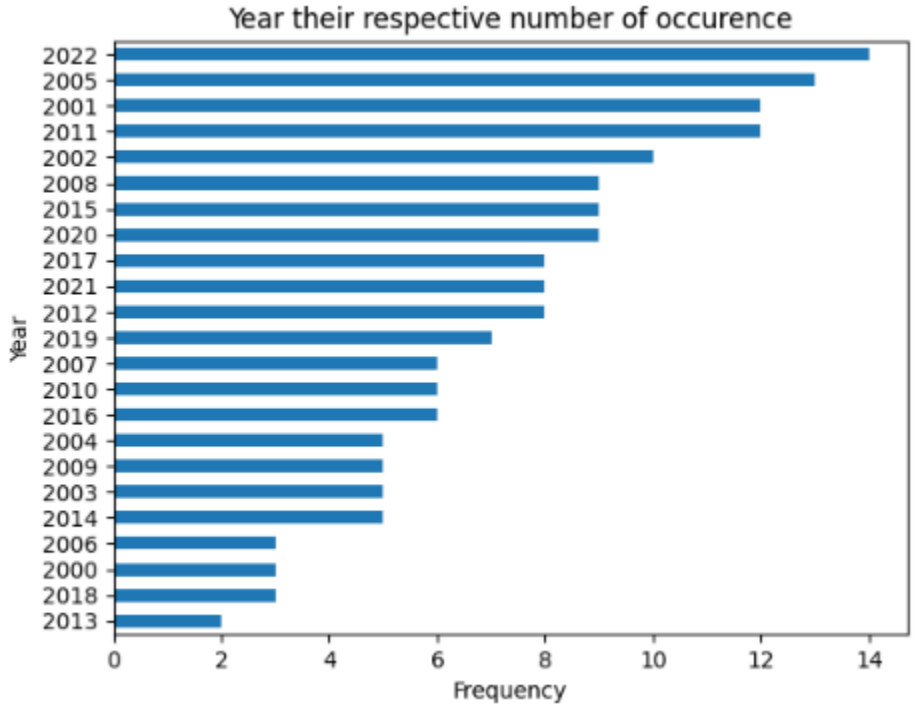
```
#countplot to show the Frequecy of disaster
df['Associated Types'].value_counts().sort_values().plot(kind='barh')
plt.title('Disaster distribution plot')
plt.xlabel('Frequency')
plt.ylabel('Disaster')
plt.show()
```



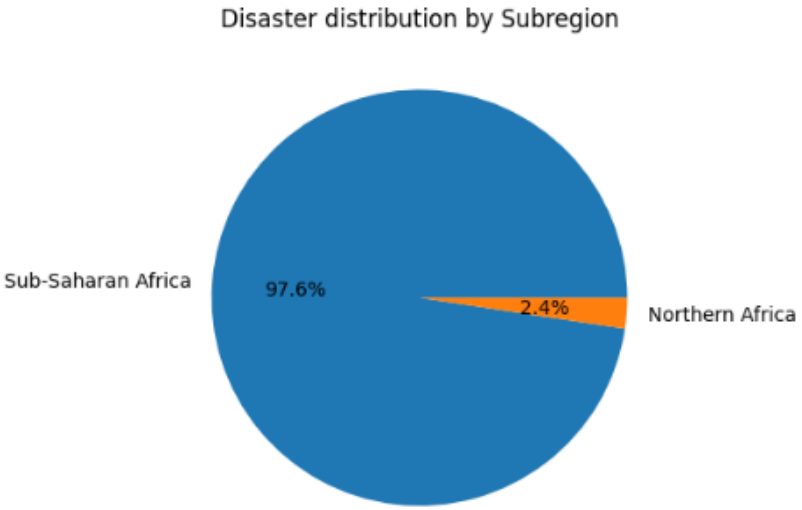
From the graph above and due to unspecified values in the dataset, it can be observed that the most occurring effect of drought is Food Shortage, followed by Famine. These are terrible living conditions for citizens and the government should look into this to avoid migration from areas prone to this disaster.



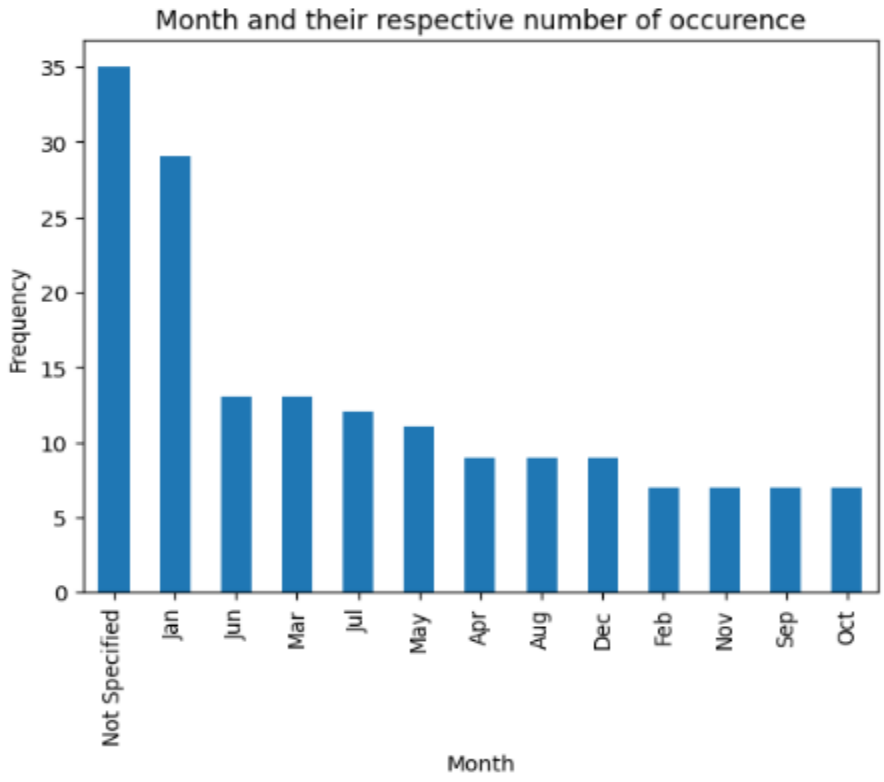
The chart above shows the frequency of occurrence of drought within the years 2000 – 2022 with Somalia and Mozambique toping the charts for the most times experienced (20 times) compared to other countries in Africa.



The bar chart above shows the distribution of drought according to year with the year 2022 recording the highest occurrence of drought, followed by 2005 and 2013 being the year with the least record of drought.



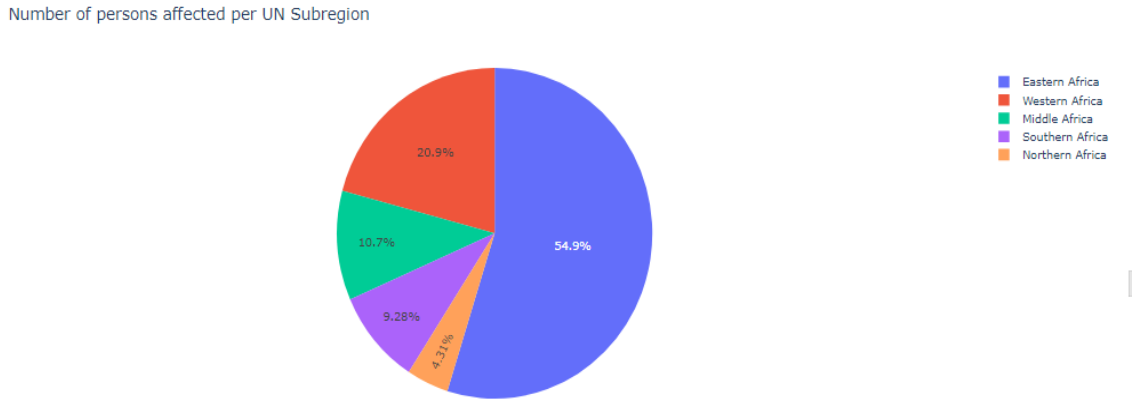
The pie chart above shows the distribution of drought within the regions in Africa. 97.6% of drought that happened in Africa has been from Sub-Saharan Africa and Northern Africa having 2.4%



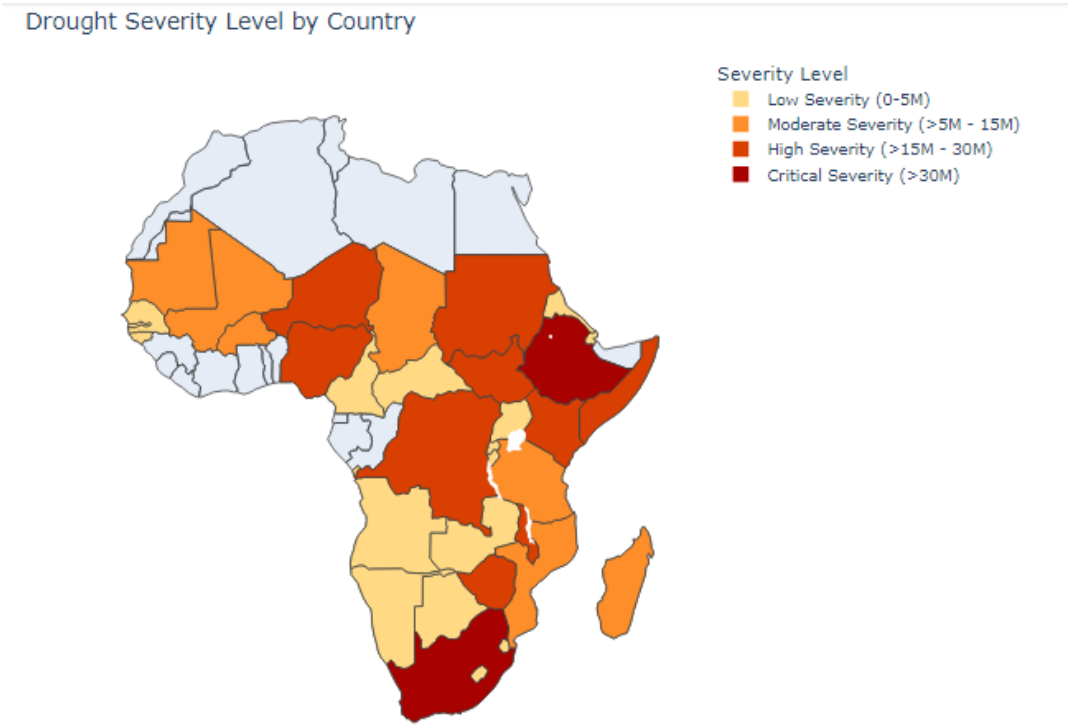
The bar chart showed the frequency of drought occurrence by month. From this, we can clearly state that drought occurs mostly in the month of January.

BIVARIATE ANALYSIS

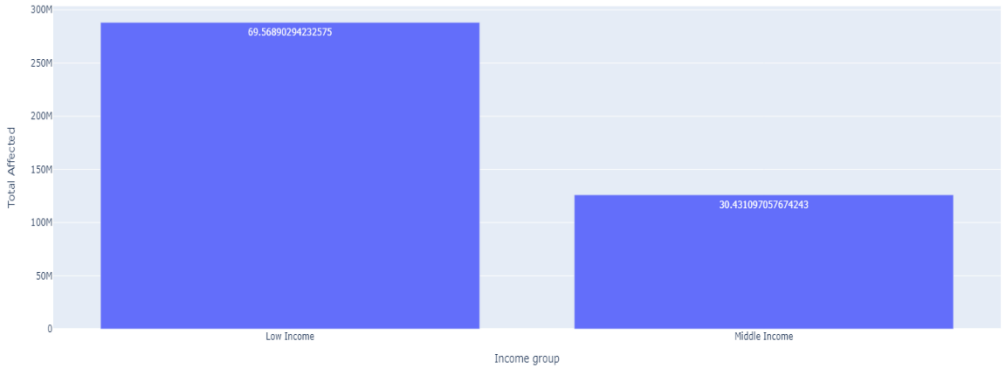
From the "JME_Regional-Classifications.xlsx" dataset we merged to ours, we further classified the distributions of drought according to sub regions from the United Nations. A pie chart was plotted to represent this:



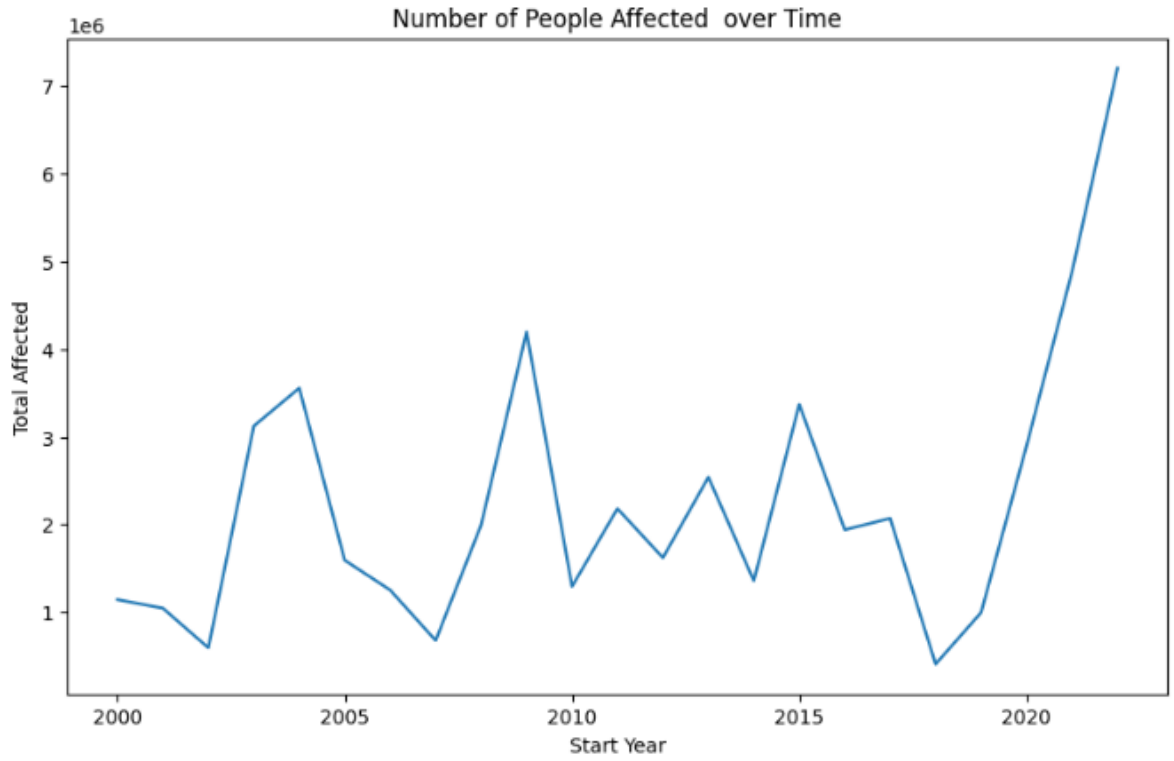
The pie chart clearly shows eastern Africa as the sub-region with the highest number of people affected by drought, the visual above shows that **54.9%** of persons affected come from eastern Africa. This means that one in every two persons affected by drought in Africa is from eastern Africa.



The map above gives a distribution of drought occurrence in African countries according to severity level.

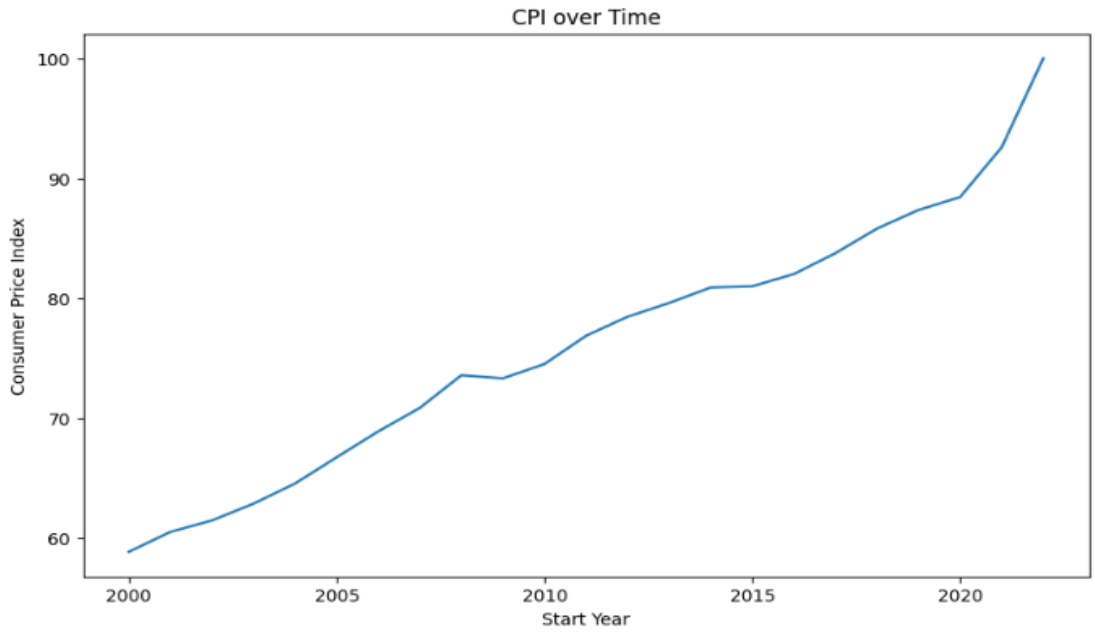


The bar chart above gives a relationship between middle-income and low-income countries and the number affected by drought each year. The bar chart shows that countries with low income were affected more than those in middle income by more than twice! It therefore gives a pointer that the income level of a country most definitely affects the number of persons involved in drought.

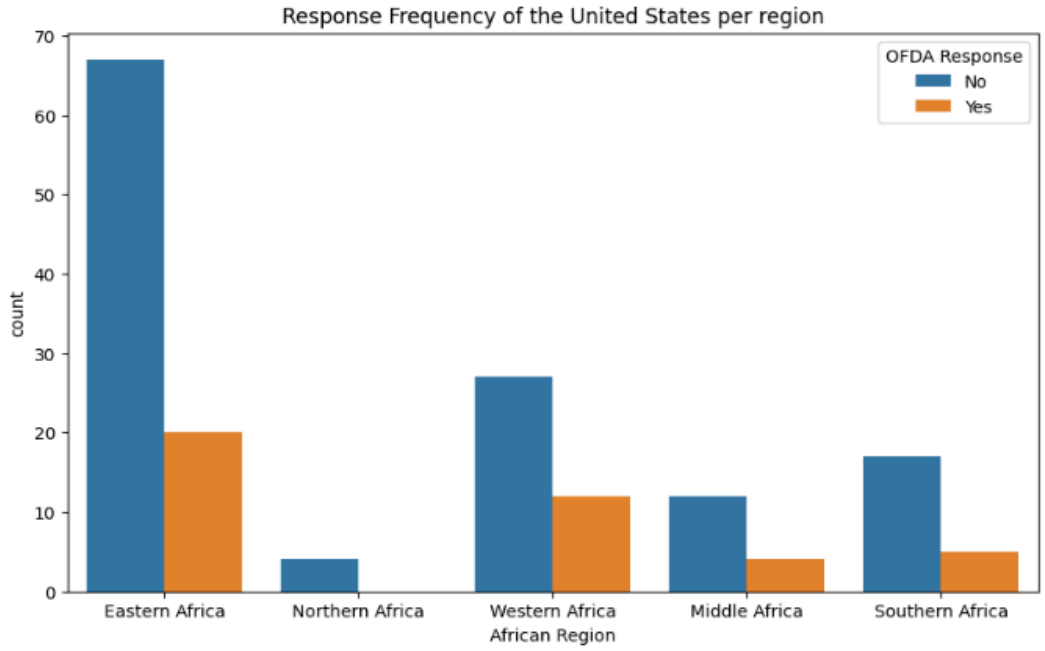


The line graph shows the correlation between year and the total affected. The line graph spikes up abruptly from the year 2020 and this could be due to 2 reasons:

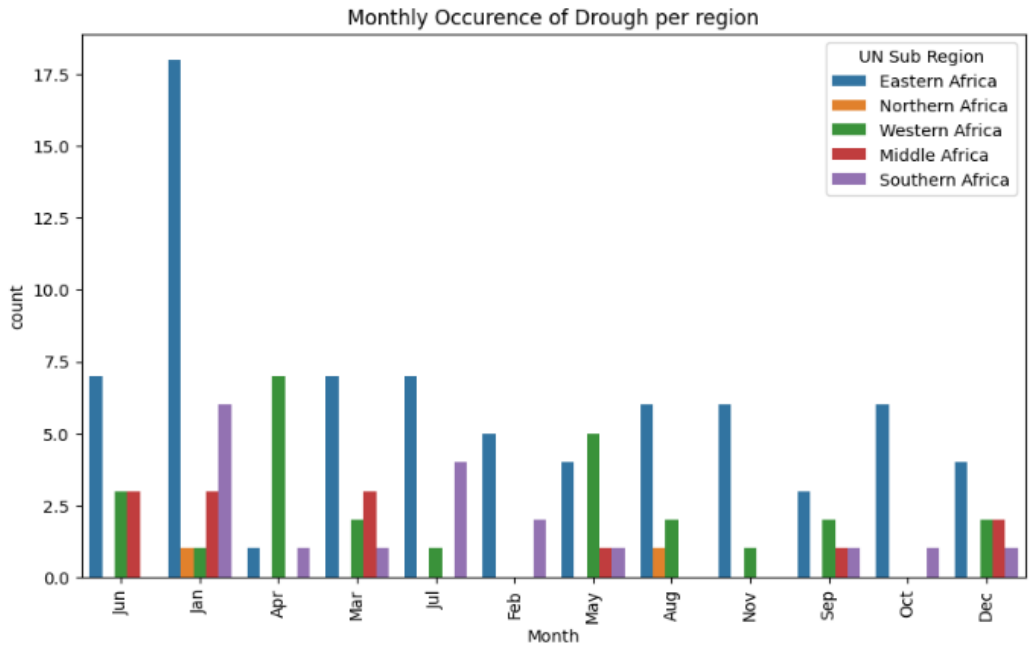
- More countries started to experience drought in Africa e.g. Nigeria who just had her first experience in 2022
- The effects of Covid-19 in 2020 on the economy of many African nations has not still worn off and as such still lead to high number of people affected.



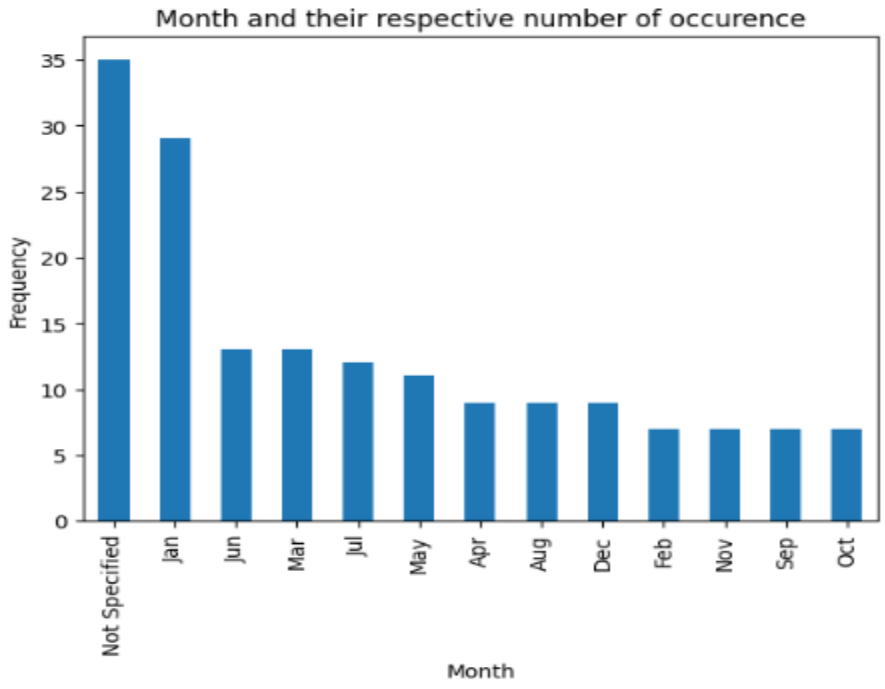
This gives a trend of the consumer price index of goods and services. The line graph shows an increasing trend and the graph tends to spike in 2020. This is most likely due to covid-19 problems and food shortage which is the main cause of rise in cost of goods and services.



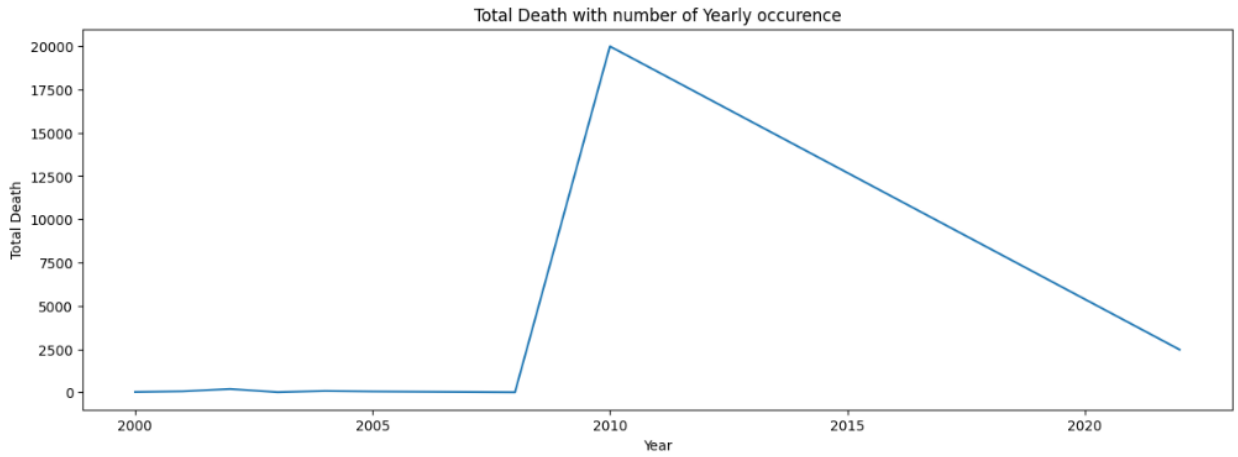
As shown earlier, this is a multiple bar chart showing response of the US dept. of foreign affairs.



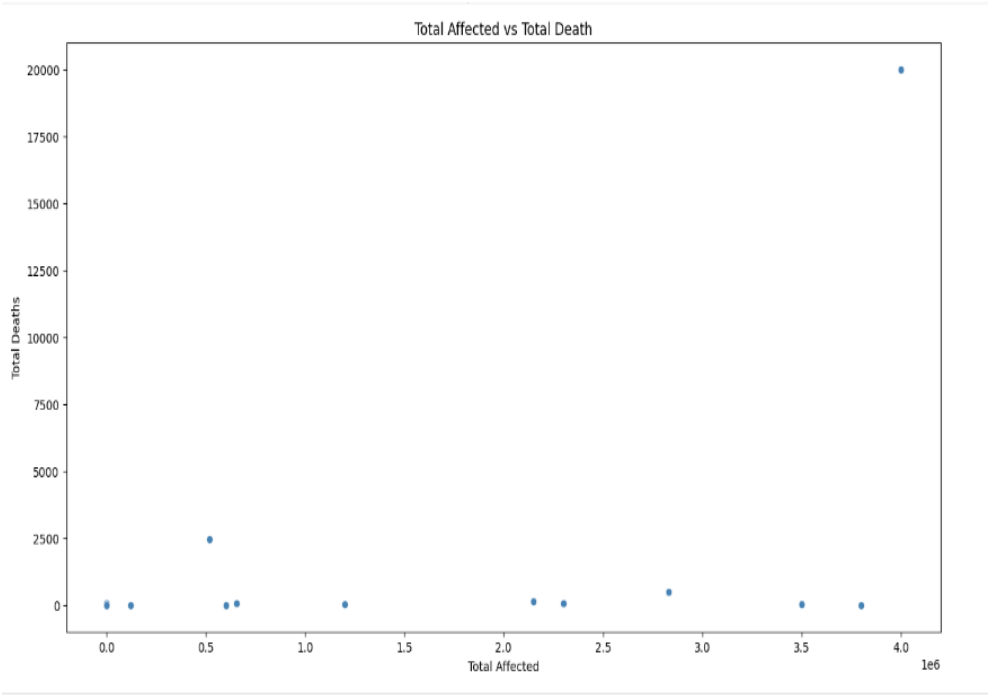
Multiple bar chart showing the distributions of sub- regions by month



From the multiple bar chart drawn, eastern Africa has its hit of drought 18 times in January, western Africa has experienced more drought in April (7 times) over the years. Northern Africa have experienced drought only in 2 months since the last 2 decades and southern Africa have not experienced drought in the months of June, August and November.

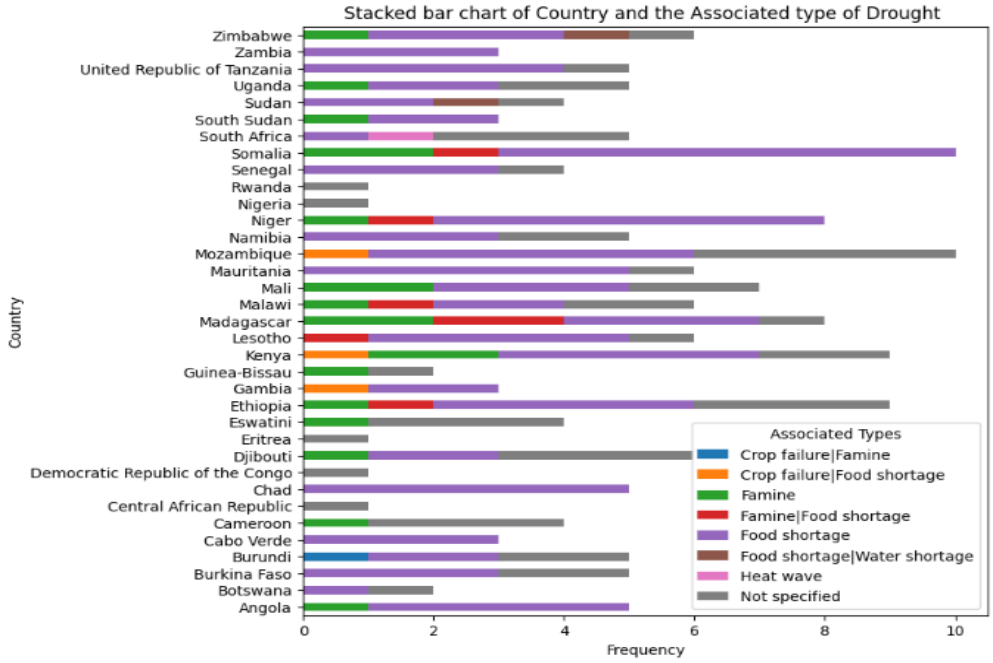


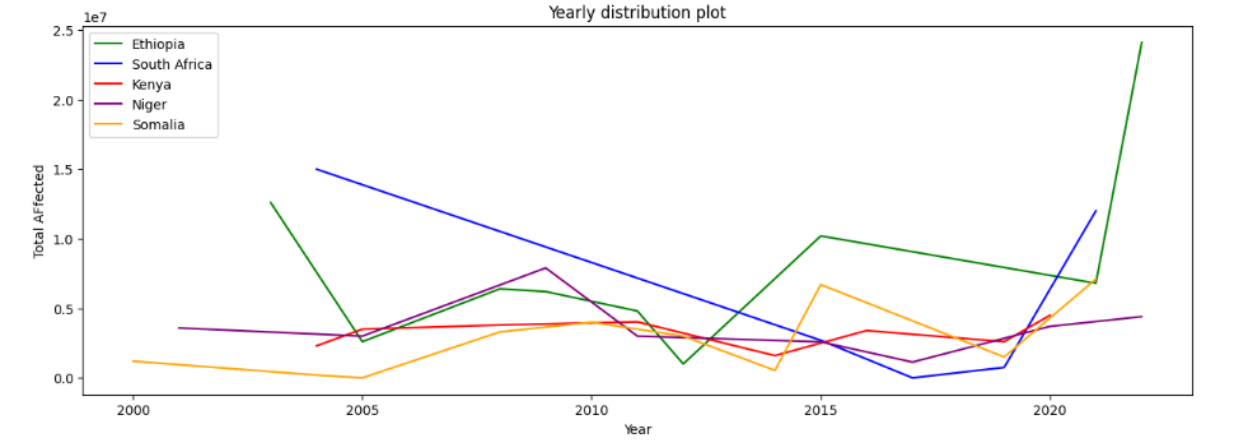
The line graph shows the relationship between total death and how it has been over the years. From the line graph, we see that the number of deaths had their highest value in 2010 and from the information given in the dataset, this number has been reducing ever since.



The scatter plot shows the distribution and relationship between the total number affected and the number of deaths caused by drought. From the plot, there is no relationship between number of death and total affected.

Other visualizations include:





This plot also confirms a spike from the year 2020 across all Sub-regions. This could be due to more countries affected or the after effects of the covid-19 fiasco.

RECOMMENDATIONS:

1. **Changing Trends in Drought Occurrence:** Our analysis revealed shifting trends in drought occurrence. Some regions have experienced an increase in drought events, while others have seen a decrease. Climate change indicators, such as rising temperatures and altered rainfall patterns, have played a significant role in these changes.

2. **Regional Disparities:** Drought severity varies widely across African regions. Countries in the sub-Saharan Africa have been particularly susceptible to severe and prolonged drought especially countries in east Africa, while other regions like Northern Africa have experienced milder events. This geographic variation necessitates region-specific drought management strategies.
3. **Socio-economic Impact:** Drought in the 21st century has had profound socio-economic implications. Many African nations have seen their economies and livelihoods negatively impacted, with reduced agricultural productivity, loss of livestock, and increased food insecurity being common consequences.
4. **Climate Change Link:** Our analysis points to a clear link between climate change and the increased vulnerability to drought. Rising temperatures and altered rainfall patterns exacerbate the likelihood and severity of drought, underscoring the need for climate adaptation measures.
5. **Mitigation Strategies:** African nations have undertaken various drought mitigation and adaptation strategies, including early warning systems, improved water resource management, and community-based resilience programs. Some of these measures have shown promise in reducing the impacts of drought.
6. **Food Security Concerns:** Drought has significantly impacted food security in African countries. The analysis revealed fluctuations in crop yields and a growing dependence on food aid during drought periods. This issue underscores the urgency of bolstering food security measures.

Conclusion:

The comparative analysis of drought in Africa since the 21st century highlights the multifaceted nature of this environmental challenge. As climate change continues to exert its influence, the need for effective drought management strategies becomes more critical. The findings suggest that while the effects of drought vary across the continent, common threads include the socio-economic vulnerabilities and the importance of proactive measures.

This analysis serves as a valuable resource for policy-makers, researchers, and organizations aiming to address the consequences of drought in Africa. It underscores the necessity of region-specific interventions, climate resilience, and collaborative efforts to mitigate the impacts of drought in the 21st century and beyond. As we move forward, a holistic and data-informed approach is essential to build a more resilient and adaptive Africa in the face of a changing climate.