Cause of death

Report

Prepared by KISHAN SAPARIYA

Data Science Intern at Flip Robo Technologies

BATCH 34

FLIP ROBO

SME Name:

Khushboo Garg

Acknowledgement

- Here we discuss all the Acknowledgement which is help me to complete this project for that first I thank to fliprobo technologies for help complete this project when I stacked in the one of function then I discuss with the team of fliprobo technologies and they are help me a lot. Also, one of my best friends is google when I forget any command or any name of function or library then I will use google by use of I complete the steps.

About Dataset

Context

A straightforward way to assess the health status of a population is to focus on mortality – or concepts like child mortality or life expectancy, which are based on mortality estimates. A focus on mortality, however, does not take into account that the burden of diseases is not only that they kill people, but that they cause suffering to people who live with them. Assessing health outcomes by both mortality and morbidity (the prevalent diseases) provides a more encompassing view on health outcomes. This is the topic of this entry. The sum of mortality and morbidity is referred to as the 'burden of disease' and can be measured by a metric called 'Disability Adjusted Life Years' (DALYs). DALYs are measuring lost health and are a standardized metric that allow for direct comparisons of disease burdens of different diseases across countries, between different populations, and over time. Conceptually, one DALY is the equivalent of losing one year in good health because of either premature death or disease or disability. One DALY represents one lost year of healthy life. The first 'Global Burden of Disease' (GBD) was GBD 1990 and the DALY metric was prominently featured in the World Bank's 1993 World Development Report. Today it is published by both the researchers at the Institute of Health Metrics and Evaluation (IHME) and the 'Disease Burden Unit' at the World Health Organization (WHO), which was created in 1998. The IHME continues the work that was started in the early 1990s and publishes the Global Burden of Disease study.

Content

In this Dataset, we have Historical Data of different cause of deaths for all ages around the World. The key features of this Dataset are: Meningitis, Alzheimer's Disease and Other Dementias, Parkinson's Disease, Nutritional Deficiencies, Malaria, Drowning, Interpersonal Violence, Maternal Disorders, HIV/AIDS, Drug Use Disorders, Tuberculosis, Cardiovascular Diseases, Lower Respiratory Infections, Neonatal Disorders, Alcohol Use Disorders, Self-harm, Exposure to Forces of Nature, Diarrheal Diseases, Environmental Heat and Cold Exposure, Neoplasms, Conflict and Terrorism, Diabetes Mellitus, Chronic Kidney Disease, Poisonings, Protein-Energy Malnutrition, Road Injuries, Chronic Respiratory Diseases, Cirrhosis and Other Chronic Liver Diseases, Digestive Diseases, Fire, Heat, and Hot Substances, Acute Hepatitis.

Data Set columns name

- 01. Country/Territory - Name of the Country/Territory
- 02. Code - Country/Territory Code
- 03. Year - Year of the Incident
- 04. Meningitis - No. of People died from Meningitis
- 05. Alzheimer's Disease and Other Dementias - No. of People died from Alzheimer's Disease and Other Dementias
- 06. Parkinson's Disease - No. of People died from Parkinson's Disease
- 07. Nutritional Deficiencies - No. of People died from Nutritional Deficiencies
- 08. Malaria - No. of People died from Malaria
- 09. Drowning - No. of People died from Drowning
- 10. Interpersonal Violence - No. of People died from Interpersonal Violence
- 11. Maternal Disorders - No. of People died from Maternal Disorders

- 12. Drug Use Disorders - No. of People died from Drug Use Disorders
- 13. Tuberculosis - No. of People died from Tuberculosis
- 14. Cardiovascular Diseases - No. of People died from Cardiovascular Diseases
- 15. Lower Respiratory Infections - No. of People died from Lower Respiratory Infections
- 16. Neonatal Disorders - No. of People died from Neonatal Disorders
- 17. Alcohol Use Disorders - No. of People died from Alcohol Use Disorders
- 18. Self-harm - No. of People died from Self-harm
- 19. Exposure to Forces of Nature - No. of People died from Exposure to Forces of Nature
- 20. Diarrheal Diseases - No. of People died from Diarrheal Diseases
- 21. Environmental Heat and Cold Exposure - No. of People died from Environmental Heat and Cold Exposure
- 22. Neoplasms - No. of People died from Neoplasms
- 23. Conflict and Terrorism - No. of People died from Conflict and Terrorism
- 24. Diabetes Mellitus - No. of People died from Diabetes Mellitus
- 25. Chronic Kidney Disease - No. of People died from Chronic Kidney Disease
- 26. Poisonings - No. of People died from Poisoning
- 27. Protein-Energy Malnutrition - No. of People died from Protein-Energy Malnutrition
- 28. Chronic Respiratory Diseases - No. of People died from Chronic Respiratory Diseases
- 29. Cirrhosis and Other Chronic Liver Diseases - No. of People died from Cirrhosis and Other Chronic Liver Diseases
- 30. Digestive Diseases - No. of People died from Digestive Diseases
- 31. Fire, Heat, and Hot Substances - No. of People died from Fire or Heat or any Hot Substances
- 32. Acute Hepatitis - No. of People died from Acute Hepatitis

Analytical Problem Framing

- For starting data analysis, we should first import require basic Libraries for that we import pandas, NumPy, matplotlib and seaborn for import data and visualization it. If it is not installed in our pc first, we have to install it and then import it otherwise we will get error such as module not found.

```
import pandas as pd
import numpy as np
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
import matplotlib.pyplot as plt
```

- Now our next step is to import data by use of pandas there are many functions in pandas for import libraries such as csv, xlsx and many more. But we observe that we have CSV file (Comma Separated value) so we use csv function. If our data is in other format, we will use any other method to import it.

```
df=pd.read_csv('cause_of_deaths dataset.csv')
df
```

- Here we import our data which is in CSV data for use it to analysis. Now our data is imported. So now we take a look of it.

	Country/Territory	Code	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	...	Diabetes Mellitus	Chronic Kidney Disease	Poisonings
0	Afghanistan	AFG	1990	2159	1116	371	2087	93	1370	1538	...	2108	3709	338
1	Afghanistan	AFG	1991	2218	1136	374	2153	189	1391	2001	...	2120	3724	351
2	Afghanistan	AFG	1992	2475	1162	378	2441	239	1514	2299	...	2153	3776	386
3	Afghanistan	AFG	1993	2812	1187	384	2837	108	1687	2589	...	2195	3862	425
4	Afghanistan	AFG	1994	3027	1211	391	3081	211	1809	2849	...	2231	3932	451
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
6115	Zimbabwe	ZWE	2015	1439	754	215	3019	2518	770	1302	...	3176	2108	381
6116	Zimbabwe	ZWE	2016	1457	767	219	3056	2050	801	1342	...	3259	2160	393
6117	Zimbabwe	ZWE	2017	1460	781	223	2990	2116	818	1363	...	3313	2196	398
6118	Zimbabwe	ZWE	2018	1450	795	227	2918	2088	825	1396	...	3381	2240	400
6119	Zimbabwe	ZWE	2019	1450	812	232	2884	2068	827	1434	...	3460	2292	405

6120 rows × 34 columns

- As we observe there are many columns and row so now for check how many rows and columns, we use pandas shape function.

```
df.shape
```

```
(6120, 34)
```

- We observe that there are total 6120 rows and 34 columns in this data there are two columns are same Country and Code.

```
df.columns
```

```
Index(['Country/Territory', 'Code', 'Year', 'Meningitis',
       'Alzheimer's Disease and Other Dementias', 'Parkinson's Disease',
       'Nutritional Deficiencies', 'Malaria', 'Drowning',
       'Interpersonal Violence', 'Maternal Disorders', 'HIV/AIDS',
       'Drug Use Disorders', 'Tuberculosis', 'Cardiovascular Diseases',
       'Lower Respiratory Infections', 'Neonatal Disorders',
       'Alcohol Use Disorders', 'Self-harm', 'Exposure to Forces of Nature',
       'Diarrheal Diseases', 'Environmental Heat and Cold Exposure',
       'Neoplasms', 'Conflict and Terrorism', 'Diabetes Mellitus',
       'Chronic Kidney Disease', 'Poisonings', 'Protein-Energy Malnutrition',
       'Road Injuries', 'Chronic Respiratory Diseases',
       'Cirrhosis and Other Chronic Liver Diseases', 'Digestive Diseases',
       'Fire, Heat, and Hot Substances', 'Acute Hepatitis'],
      dtype='object')
```

- By observe this above function we observe the name of the columns.
- Now we check for null values for that we use is null function with the help of sum function to check it and also we use heat map to plot the null values if it has.

```
df.isnull()
```

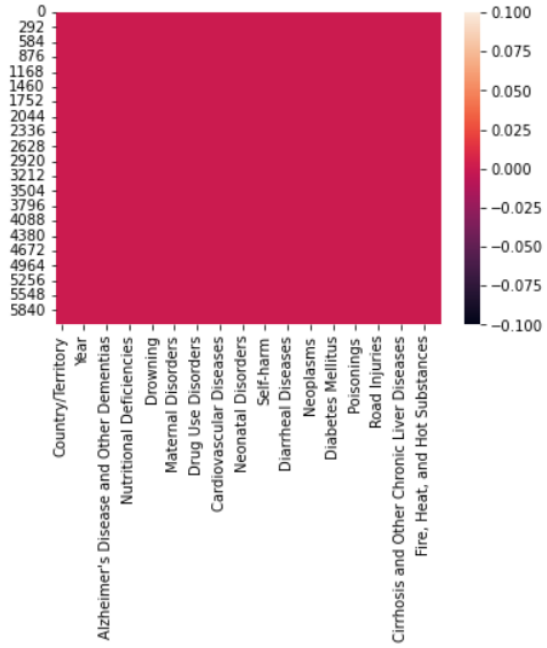
	Country/Territory	Code	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	...	Diabetes Mellitus	Chronic Kidney Disease	Poisonings
0	False	False	False	False	False	False	False	False	False	False	...	False	False	False
1	False	False	False	False	False	False	False	False	False	False	...	False	False	False
2	False	False	False	False	False	False	False	False	False	False	...	False	False	False
3	False	False	False	False	False	False	False	False	False	False	...	False	False	False
4	False	False	False	False	False	False	False	False	False	False	...	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...	...	...	..
6115	False	False	False	False	False	False	False	False	False	False	...	False	False	False
6116	False	False	False	False	False	False	False	False	False	False	...	False	False	False
6117	False	False	False	False	False	False	False	False	False	False	...	False	False	False
6118	False	False	False	False	False	False	False	False	False	False	...	False	False	False
6119	False	False	False	False	False	False	False	False	False	False	...	False	False	False

```
df.isnull().sum()
```

```
Country/Territory                           0
Code                                        0
Year                                        0
Meningitis                                  0
Alzheimer's Disease and Other Dementias     0
Parkinson's Disease                         0
Nutritional Deficiencies                    0
Malaria                                     0
Drowning                                    0
Interpersonal Violence                      0
Maternal Disorders                          0
HIV/AIDS                                    0
Drug Use Disorders                          0
Tuberculosis                                0
Cardiovascular Diseases                     0
Lower Respiratory Infections                0
Neonatal Disorders                          0
Alcohol Use Disorders                       0
Self-harm                                   0
Exposure to Forces of Nature                0
Diarrheal Diseases                          0
Environmental Heat and Cold Exposure        0
Neoplasms                                   0
Conflict and Terrorism                      0
Diabetes Mellitus                           0
Chronic Kidney Disease                      0
Poisonings                                  0
Protein-Energy Malnutrition                 0
Road Injuries                               0
Chronic Respiratory Diseases                0
Cirrhosis and Other Chronic Liver Diseases  0
Digestive Diseases                          0
Fire, Heat, and Hot Substances              0
Acute Hepatitis                             0
dtype: int64
```

```
sns.heatmap(df.isnull())
```

```
<AxesSubplot:>
```



- Now we observe that there are no null values in our data set so we no need to remove it.
- Now we observe for data types for that we use pandas function dtypes.

```
df.dtypes
```

```
Country/Territory                           object
Code                                        object
Year                                         int64
Meningitis                                   int64
Alzheimer's Disease and Other Dementias      int64
Parkinson's Disease                          int64
Nutritional Deficiencies                     int64
Malaria                                      int64
Drowning                                     int64
Interpersonal Violence                       int64
Maternal Disorders                           int64
HIV/AIDS                                     int64
Drug Use Disorders                           int64
Tuberculosis                                 int64
Cardiovascular Diseases                      int64
Lower Respiratory Infections                 int64
Neonatal Disorders                           int64
Alcohol Use Disorders                        int64
Self-harm                                    int64
Exposure to Forces of Nature                 int64
Diarrheal Diseases                           int64
Environmental Heat and Cold Exposure         int64
Neoplasms                                    int64
Conflict and Terrorism                       int64
Diabetes Mellitus                            int64
Chronic Kidney Disease                       int64
Poisonings                                   int64
Protein-Energy Malnutrition                  int64
Road Injuries                                int64
Chronic Respiratory Diseases                 int64
Cirrhosis and Other Chronic Liver Diseases   int64
Digestive Diseases                           int64
Fire, Heat, and Hot Substances               int64
Acute Hepatitis                              int64
dtype: object
```

- As we observe there are total two columns Country/Territory and Code are in object type data rest of are in int64.
- Now we observe unique values in the columns of which contains object type data. For that we use pandas unique and nunique functions.

```
df['Country/Territory'].unique()
```

```
array(['Afghanistan', 'Albania', 'Algeria', 'American Samoa', 'Andorra',
       'Angola', 'Antigua and Barbuda', 'Argentina', 'Armenia',
       'Australia', 'Austria', 'Azerbaijan', 'Bahamas', 'Bahrain',
       'Bangladesh', 'Barbados', 'Belarus', 'Belgium', 'Belize', 'Benin',
       'Bermuda', 'Bhutan', 'Bolivia', 'Bosnia and Herzegovina',
       'Botswana', 'Brazil', 'Brunei', 'Bulgaria', 'Burkina Faso',
       'Burundi', 'Cambodia', 'Cameroon', 'Canada', 'Cape Verde',
       'Central African Republic', 'Chad', 'Chile', 'China', 'Colombia',
       'Comoros', 'Congo', 'Cook Islands', 'Costa Rica', "Cote d'Ivoire",
       'Croatia', 'Cuba', 'Cyprus', 'Czechia',
       'Democratic Republic of Congo', 'Denmark', 'Djibouti', 'Dominica',
       'Dominican Republic', 'Ecuador', 'Egypt', 'El Salvador',
       'Equatorial Guinea', 'Eritrea', 'Estonia', 'Eswatini', 'Ethiopia',
       'Fiji', 'Finland', 'France', 'Gabon', 'Gambia', 'Georgia',
       'Germany', 'Ghana', 'Greece', 'Greenland', 'Grenada', 'Guam',
       'Guatemala', 'Guinea', 'Guinea-Bissau', 'Guyana', 'Haiti',
       'Honduras', 'Hungary', 'Iceland', 'India', 'Indonesia', 'Iran',
       'Iraq', 'Ireland', 'Israel', 'Italy', 'Jamaica', 'Japan', 'Jordan',
       'Kazakhstan', 'Kenya', 'Kiribati', 'Kuwait', 'Kyrgyzstan', 'Laos',
       'Latvia', 'Lebanon', 'Lesotho', 'Liberia', 'Libya', 'Lithuania',
       'Luxembourg', 'Madagascar', 'Malawi', 'Malaysia', 'Maldives',
       'Mali', 'Malta', 'Marshall Islands', 'Mauritania', 'Mauritius',
       'Mexico', 'Micronesia', 'Moldova', 'Monaco', 'Mongolia',
       'Montenegro', 'Morocco', 'Mozambique', 'Myanmar', 'Namibia',
       'Nauru', 'Nepal', 'Netherlands', 'New Zealand', 'Nicaragua',
       'Niger', 'Nigeria', 'Niue', 'North Korea', 'North Macedonia',
       'Northern Mariana Islands', 'Norway', 'Oman', 'Pakistan', 'Palau',
       'Palestine', 'Panama', 'Papua New Guinea', 'Paraguay', 'Peru',
       'Philippines', 'Poland', 'Portugal', 'Puerto Rico', 'Qatar',
       'Romania', 'Russia', 'Rwanda', 'Saint Kitts and Nevis',
       'Saint Lucia', 'Saint Vincent and the Grenadines', 'Samoa',
       'San Marino', 'Sao Tome and Principe', 'Saudi Arabia', 'Senegal',
       'Serbia', 'Seychelles', 'Sierra Leone', 'Singapore', 'Slovakia',
       'Slovenia', 'Solomon Islands', 'Somalia', 'South Africa',
       'South Korea', 'South Sudan', 'Spain', 'Sri Lanka', 'Sudan',
       'Suriname', 'Sweden', 'Switzerland', 'Syria', 'Taiwan',
       'Tajikistan', 'Tanzania', 'Thailand', 'Timor', 'Togo', 'Tokelau',
       'Tonga', 'Trinidad and Tobago', 'Tunisia', 'Turkey',
       'Turkmenistan', 'Tuvalu', 'Uganda', 'Ukraine',
       'United Arab Emirates', 'United Kingdom', 'United States',
       'United States Virgin Islands', 'Uruguay', 'Uzbekistan', 'Vanuatu',
       'Venezuela', 'Vietnam', 'Yemen', 'Zambia', 'Zimbabwe'],
      dtype=object)
```

- Check for number of unique values in column.

```
df['Country/Territory'].nunique()
```

```
204
```

- There is total 204 unique values in this column.

```
df['Code'].unique()
```

```
array(['AFG', 'ALB', 'DZA', 'ASM', 'AND', 'AGO', 'ATG', 'ARG', 'ARM',
       'AUS', 'AUT', 'AZE', 'BHS', 'BHR', 'BGD', 'BRB', 'BLR', 'BEL',
       'BLZ', 'BEN', 'BMU', 'BTN', 'BOL', 'BIH', 'BWA', 'BRA', 'BRN',
       'BGR', 'BFA', 'BDI', 'KHM', 'CMR', 'CAN', 'CPV', 'CAF', 'TCD',
       'CHL', 'CHN', 'COL', 'COM', 'COG', 'COK', 'CRI', 'CIV', 'HRV',
       'CUB', 'CYP', 'CZE', 'COD', 'DNK', 'DJI', 'DMA', 'DOM', 'ECU',
       'EGY', 'SLV', 'GNQ', 'ERI', 'EST', 'SWZ', 'ETH', 'FJI', 'FIN',
       'FRA', 'GAB', 'GMB', 'GEO', 'DEU', 'GHA', 'GRC', 'GRL', 'GRD',
       'GUM', 'GTM', 'GIN', 'GNB', 'GUY', 'HTI', 'HND', 'HUN', 'ISL',
       'IND', 'IDN', 'IRN', 'IRQ', 'IRL', 'ISR', 'ITA', 'JAM', 'JPN',
       'JOR', 'KAZ', 'KEN', 'KIR', 'KWT', 'KGZ', 'LAO', 'LVA', 'LBN',
       'LSO', 'LBR', 'LBY', 'LTU', 'LUX', 'MDG', 'MWI', 'MYS', 'MDV',
       'MLI', 'MLT', 'MHL', 'MRT', 'MUS', 'MEX', 'FSM', 'MDA', 'MCO',
       'MNG', 'MNE', 'MAR', 'MOZ', 'MMR', 'NAM', 'NRU', 'NPL', 'NLD',
       'NZL', 'NIC', 'NER', 'NGA', 'NIU', 'PRK', 'MKD', 'MNP', 'NOR',
       'OMN', 'PAK', 'PLW', 'PSE', 'PAN', 'PNG', 'PRY', 'PER', 'PHL',
       'POL', 'PRT', 'PRI', 'QAT', 'ROU', 'RUS', 'RWA', 'KNA', 'LCA',
       'VCT', 'WSM', 'SMR', 'STP', 'SAU', 'SEN', 'SRB', 'SYC', 'SLE',
       'SGP', 'SVK', 'SVN', 'SLB', 'SOM', 'ZAF', 'KOR', 'SSD', 'ESP',
       'LKA', 'SDN', 'SUR', 'SWE', 'CHE', 'SYR', 'TWN', 'TJK', 'TZA',
       'THA', 'TLS', 'TGO', 'TKL', 'TON', 'TTO', 'TUN', 'TUR', 'TKM',
       'TUV', 'UGA', 'UKR', 'ARE', 'GBR', 'USA', 'VIR', 'URY', 'UZB',
       'VUT', 'VEN', 'VNM', 'YEM', 'ZMB', 'ZWE'], dtype=object)
```

- Check for number of unique values in this column.

```
df['Code'].nunique()
```

```
204
```

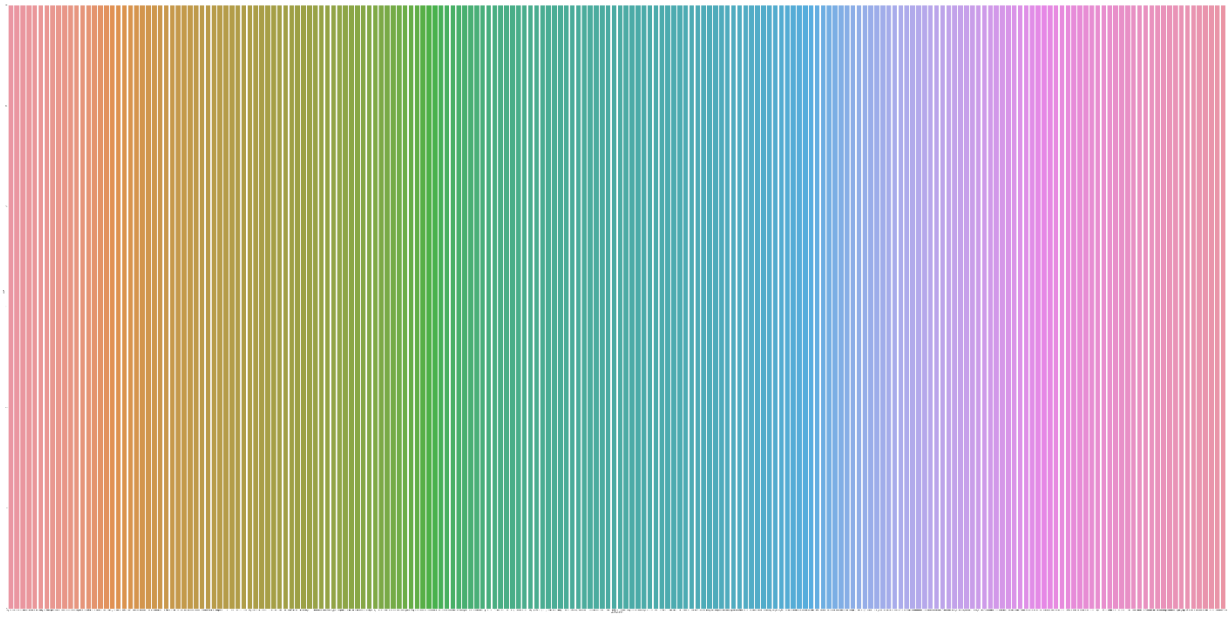
- As we observe there are total 204 unique values.
- In the column of Code there are total 204 unique values. This columns is shows that the country code.

Exploratory Data Analysis

1. Univariate Analysis

```
plt.figure(figsize=(150,80))
sns.countplot(df['Country/Territory'])
```

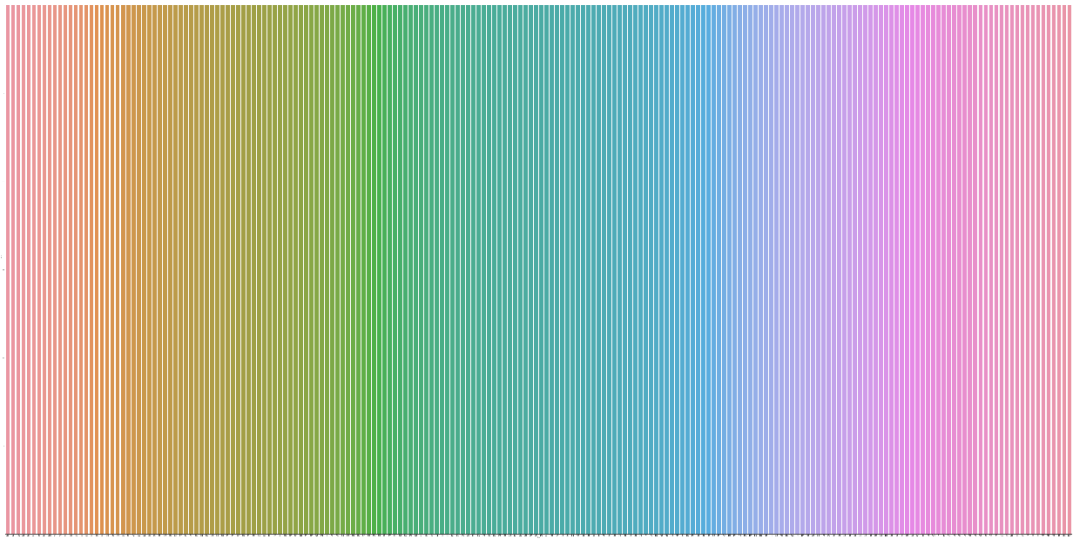
- Now check the output of above graph.



- As we observe this graph by zoom it we observe that there are total count of each value is 30 and it shows that there are total 204 unique value and 204*30=6120. our row size is 6120.

```
plt.figure(figsize=(150,80))
sns.countplot(df['Code'])
```

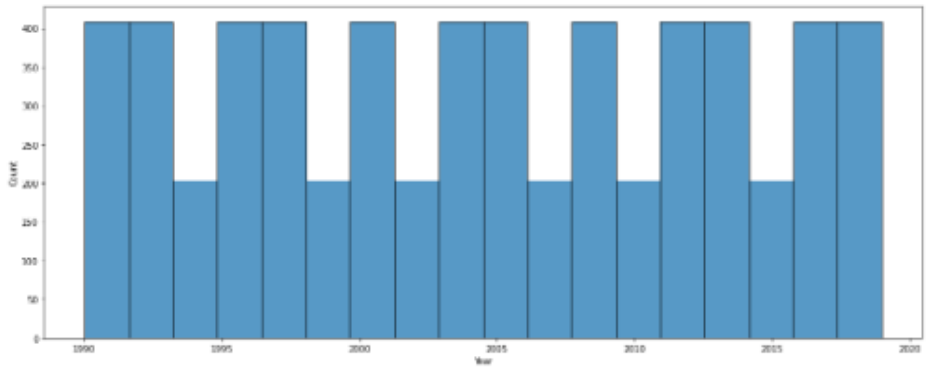
- Now check output of above graph.

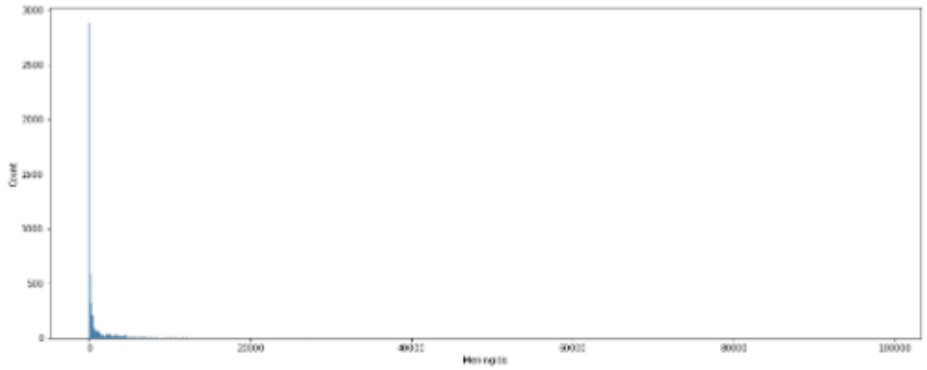


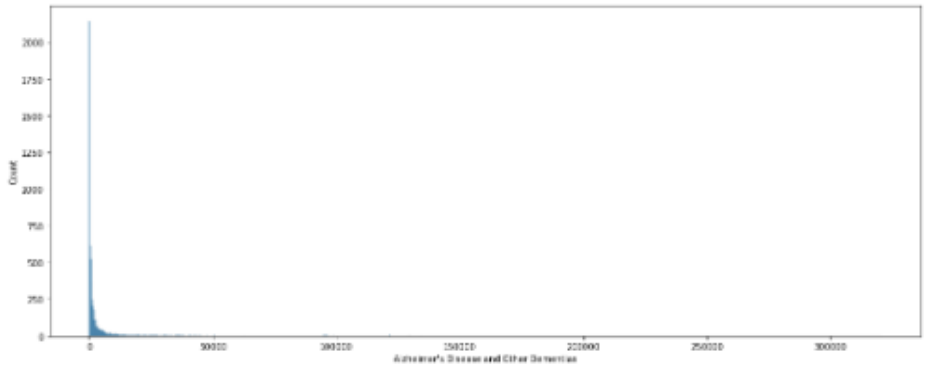
- Now we plot hist graph for our numerical values of data. By use of for loop.

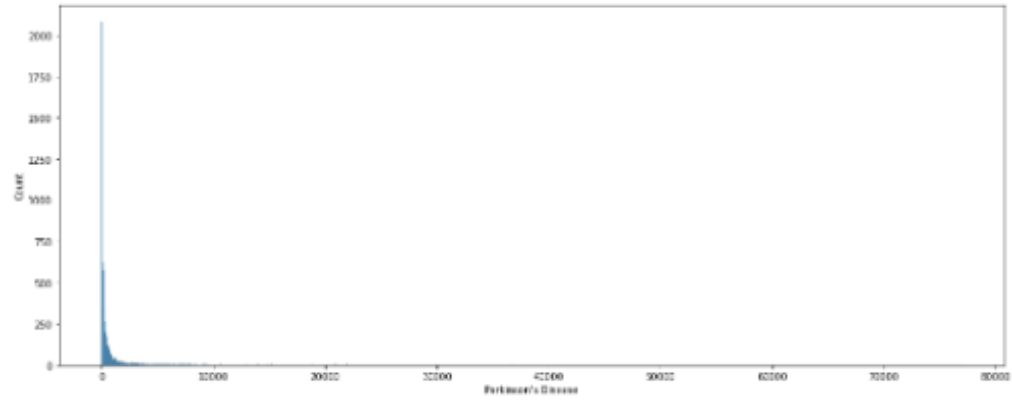
```
columns=['Year', 'Meningitis',
       "Alzheimer's Disease and Other Dementias", "Parkinson's Disease",
       'Nutritional Deficiencies', 'Malaria', 'Drowning',
       'Interpersonal Violence', 'Maternal Disorders', 'HIV/AIDS',
       'Drug Use Disorders', 'Tuberculosis', 'Cardiovascular Diseases',
       'Lower Respiratory Infections', 'Neonatal Disorders',
       'Alcohol Use Disorders', 'Self-harm']

for i in columns:
    plt.figure(figsize=(18,7))
    sns.histplot(df[i])
    plt.show()
```

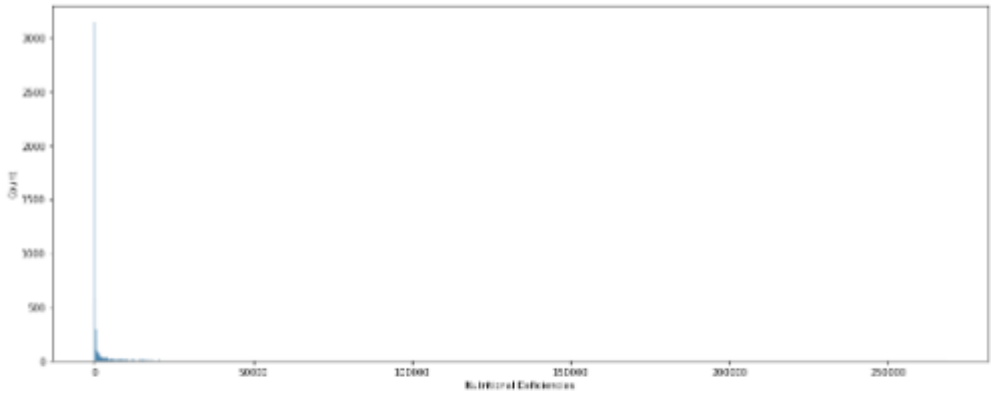
- Plot output of graph and analysis it.

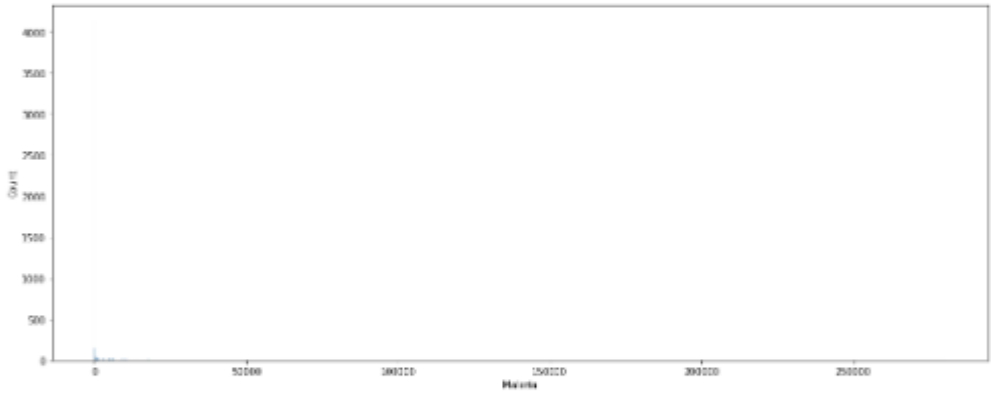


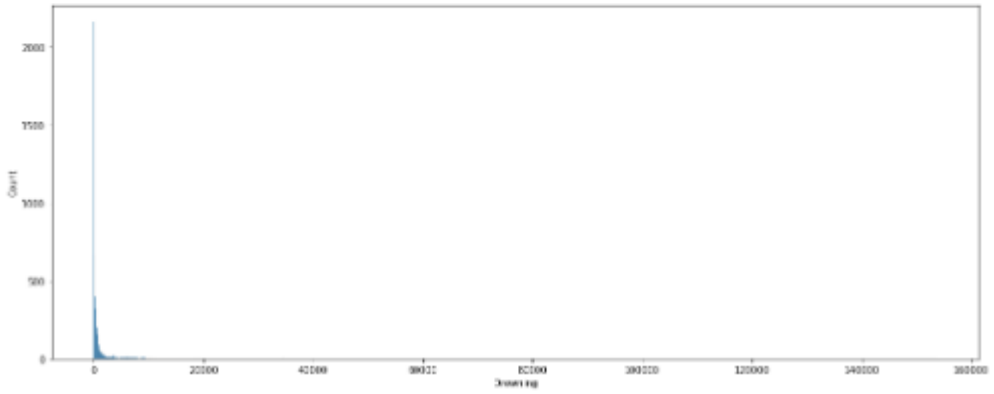


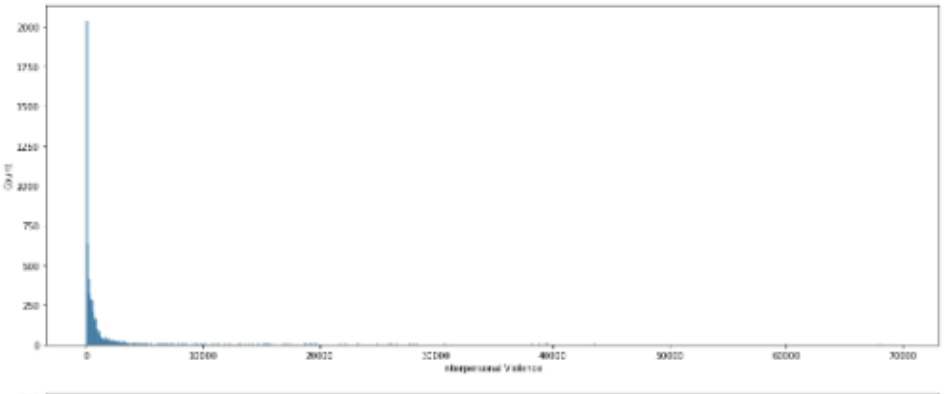


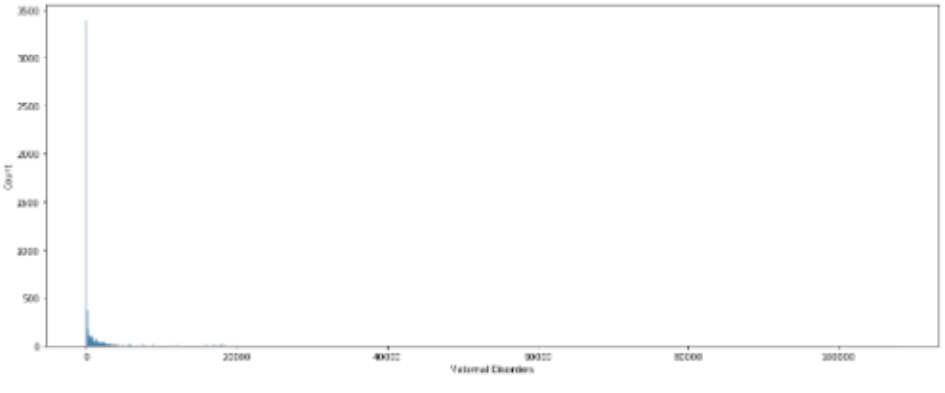


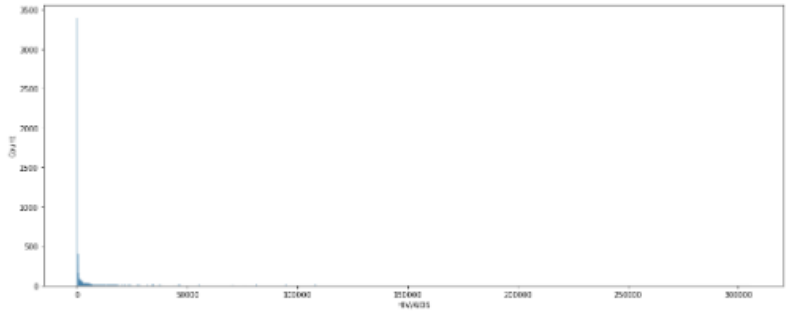

Count
0
500
1000
1500
2000
2500
3000
0
50000
100000
150000
200000
250000

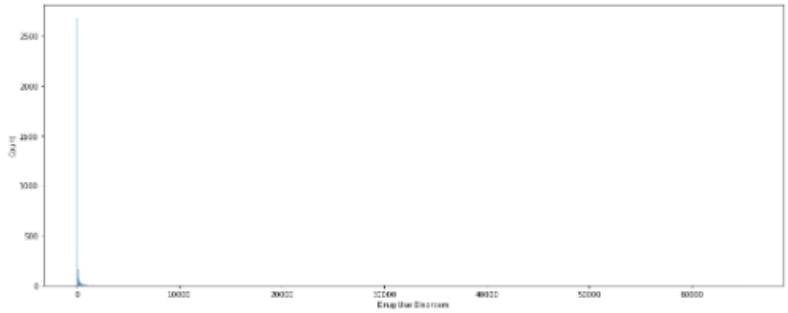

Count
0
500
1000
1500
2000
2500
3000
3500
4000
0
50000
100000
150000
200000
250000

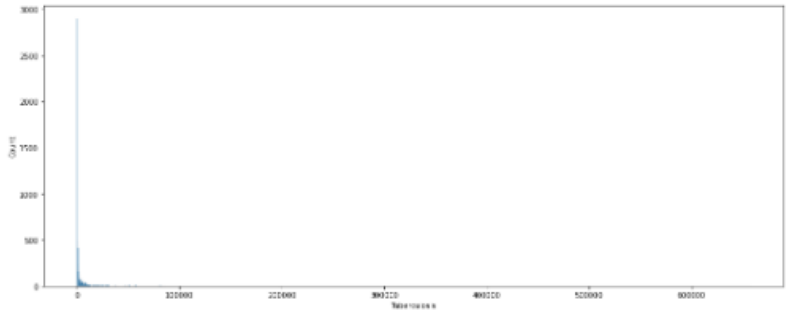

Count
0
500
1000
1500
2000
0
20000
40000
60000
80000
100000
120000
140000
160000

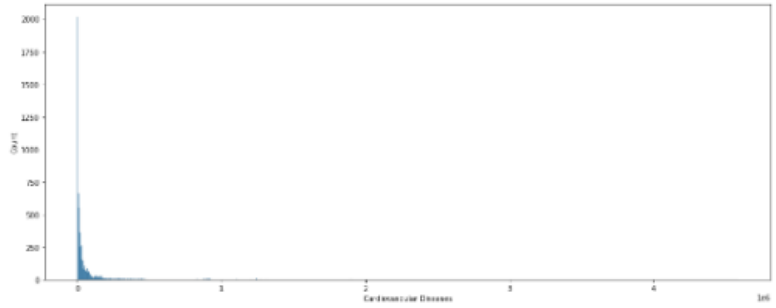

Count
0
250
500
750
1000
1250
1500
1750
2000
0
10000
20000
30000
40000
50000
60000
70000

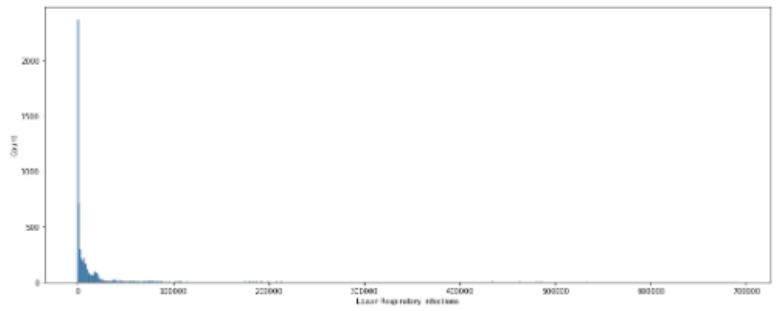

Count
0
500
1000
1500
2000
2500
3000
3500
0
20000
40000
60000
80000
100000

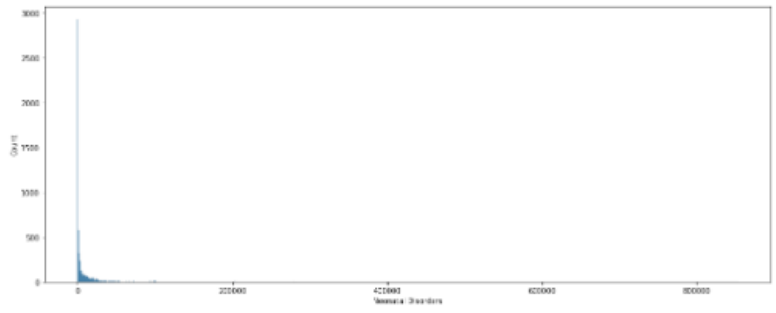

Count

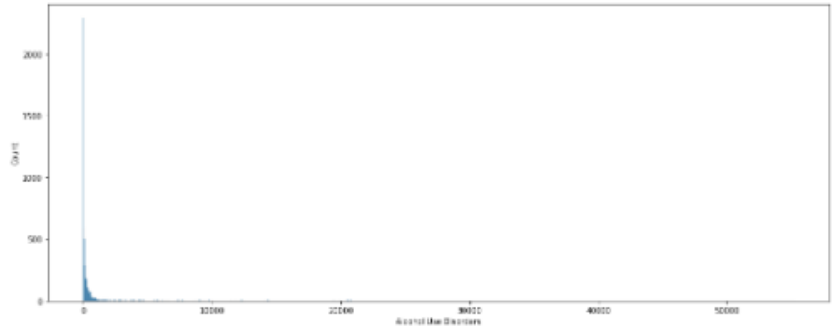

Count

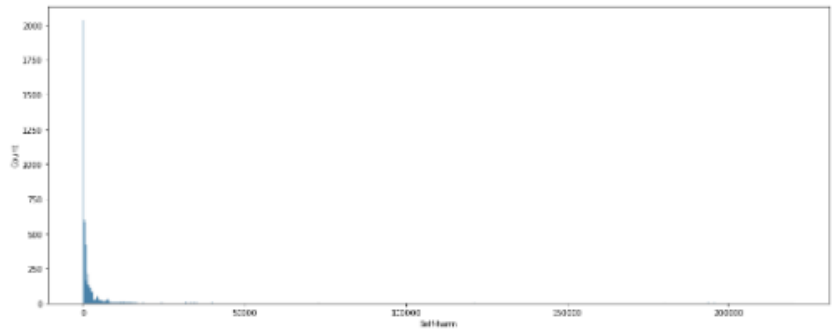

Count

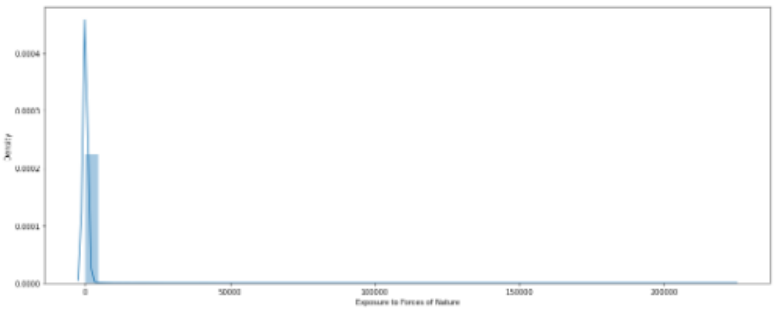

Count

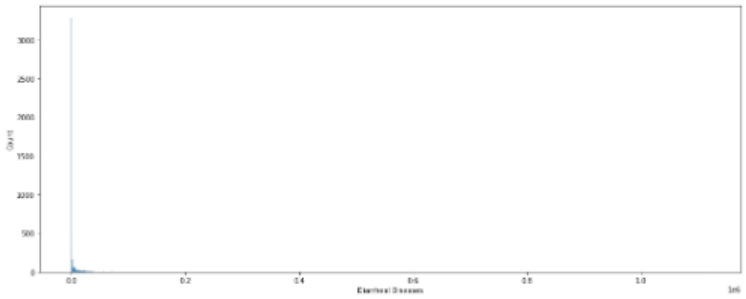

Count

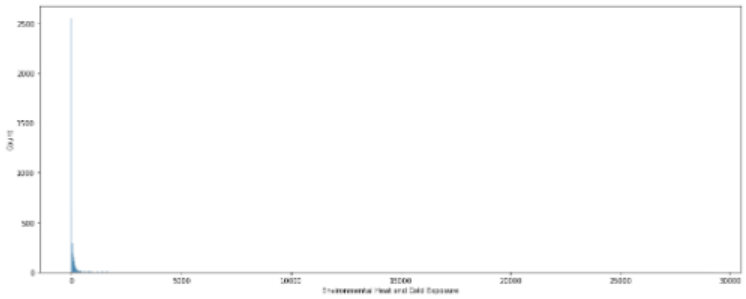

Count

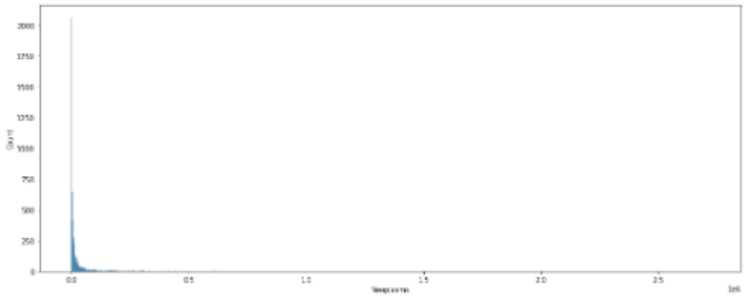

Count
0
500
1000
1500
2000
0
10000
20000
30000
40000
50000

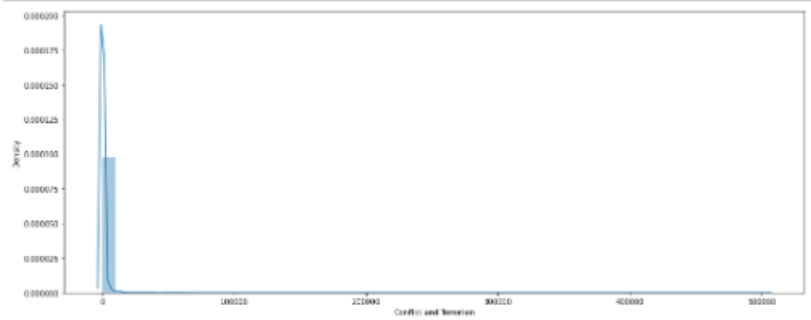

Count
0
250
500
750
1000
1250
1500
1750
2000
0
50000
100000
150000
200000

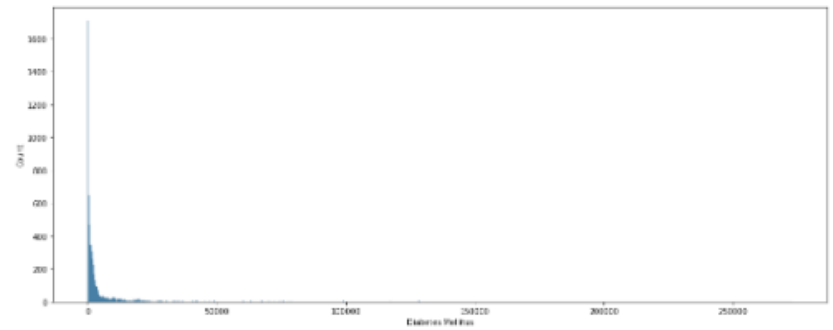

Density
0.0000
0.0001
0.0002
0.0003
0.0004
0
50000
100000
150000
200000
Exposure to Forces of Nature

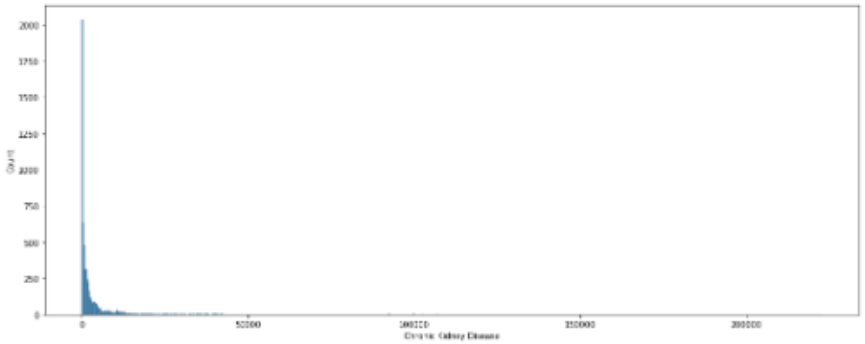

Count
0
500
1000
1500
2000
2500
3000
0.0
0.2
0.4
0.6
0.8
1.0
1e6

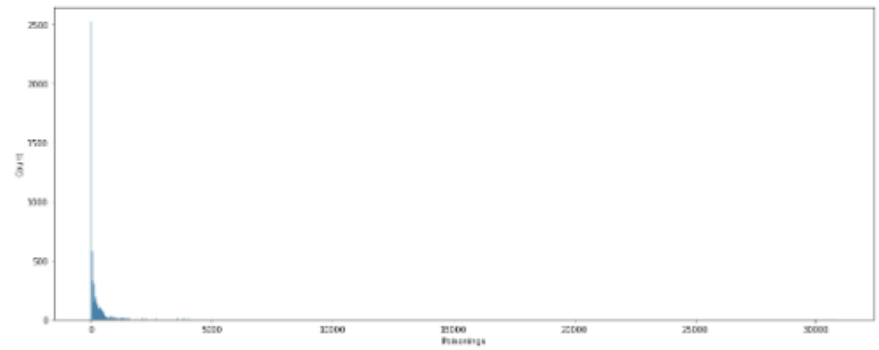

Count
0
500
1000
1500
2000
2500
0
5000
10000
15000
20000
25000
30000

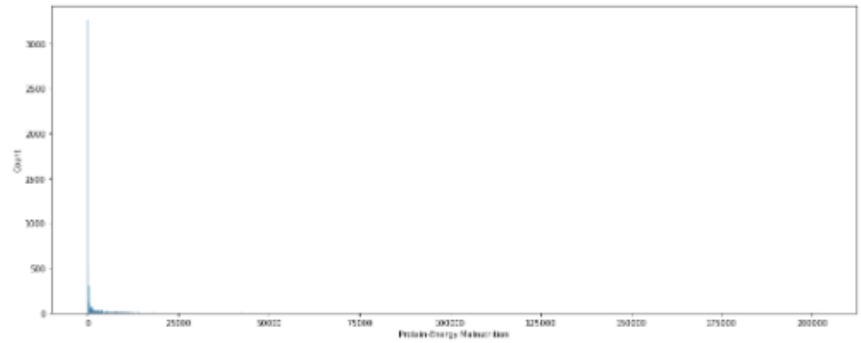

Count
0
250
500
750
1000
1250
1500
1750
2000
0.0
0.5
1.0
1.5
2.0
2.5
1e6

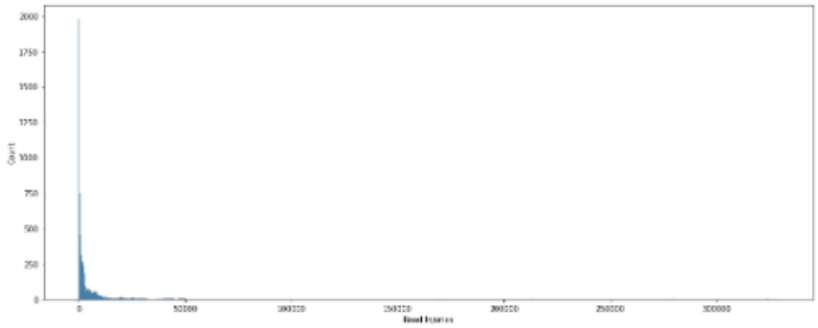


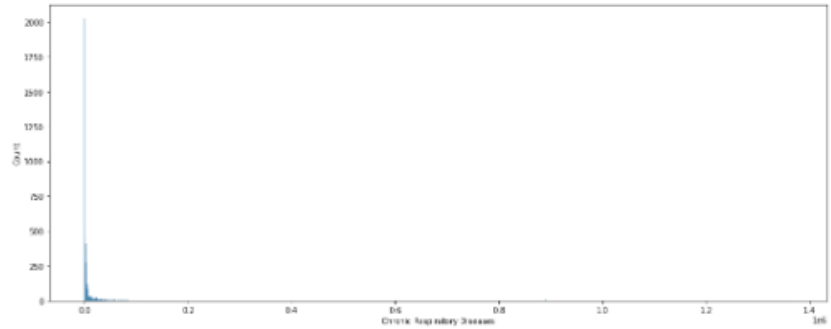


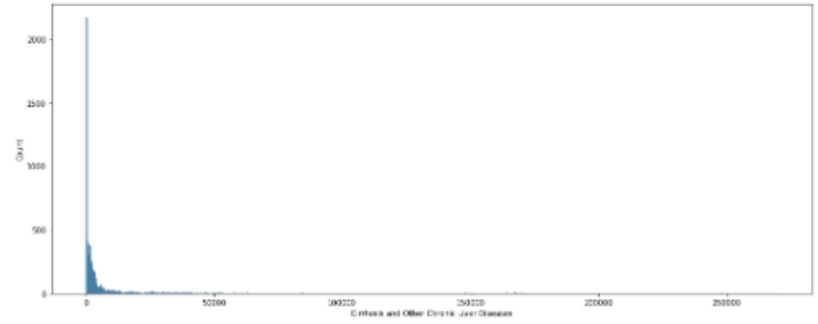


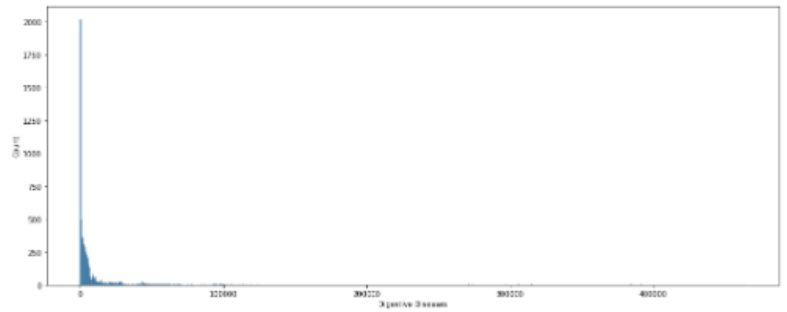


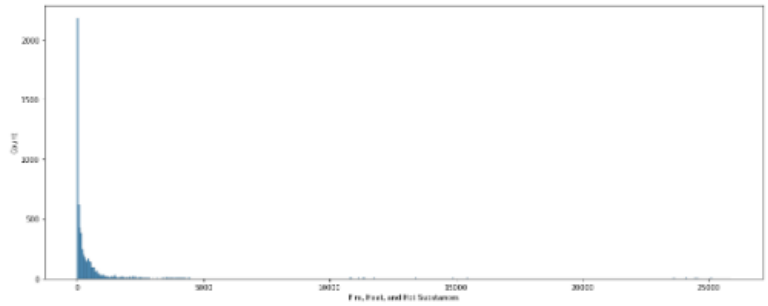


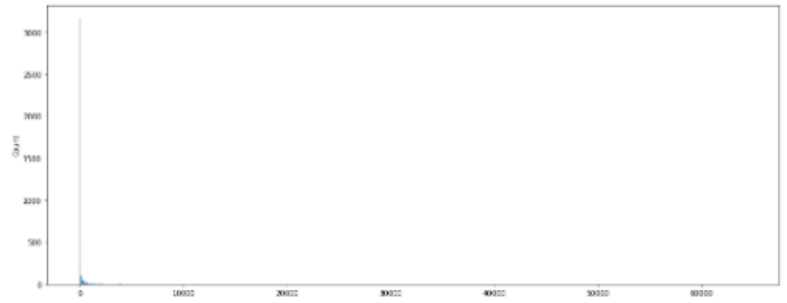

Count
Head Injuries


Count
Chronic Respiratory Diseases
1e6


Count
Cirrhosis and Other Chronic Liver Diseases


Count
Digestive Diseases


Count
Fire, Heat, and Hot Substances


Count

- By observing above graphs, we write some observations.

  Observations: -

  1. Year in this column we observe that majority of data are equal with the year.
  2. Menignits in this column we observe that majority of data are says that around 5000 persons died.
  3. Alzheimer's Disease and Other Dementias in this column we observe that majority of data are says that around 12500 persons died.
  4. Parkinson's Disease in this column we observe that majority of data are says that around 3000 persons died.
  5. Nutritional Deficiencies in this column we observe that majority of data are says that around 10000 persons died.
  6. Malaria in this column we observe that majority of data are says that very less persons died.
  7. Drowning in this column we observe that majority of data are says that around 5000 persons died.
  8. Interpersonal Violence in this column we observe that majority of data are says that around 5000 persons died.
  9. Maternal Disorders in this column we observe that majority of data are says that around 4000 persons died.
  10. HIV/AIDS in this column we observe that majority of data are says that around 10000 persons died.
  11. Drug Use Disorders in this column we observe that majority of data are says that very less persons are died.
  12. Tuberculosis in this column we observe that majority of data are says that around 20000 persons died.
  13. cardiovascular diseases in this column we observe that majority of data are says that around 3000 persons died.
  14. Lower Respiratory Infections in this column we observe that majority of data are says that around 25000 persons died.
  15. Neonatal Disorders in this column we observe that majority of data are says that around 10000 persons died.
  16. Alcohol Use Disorders in this column we observe that majority of data are says that around 5000 persons died.
  17. Self-harm in this column we observe that majority of data are says that around 2500 persons died.
  18. Exposure to Forces of Nature in this column we observe that majority of data are says that around 2000 persons died.
  19. Diarrheal Diseases in this column we observe that majority of data are says that around 2000 persons died.
  20. Environmental Heat and Cold Exposure in this column we observe that majority of data are says that around 3000 persons died.
  21. Neoplasms in this column we observe that majority of data are says that around 2500 persons died.
  22. Conflict and Terrorism in this column we observe that majority of data are says that around 1000 persons died.
  23. Diabetes Mellitus in this column we observe that majority of data are says that around 25000 persons died.
  24. chronic kidney disease in this column we observe that majority of data are says that around 25000 persons died.
  25. Poisonings in this column we observe that majority of data are says that around 2500 persons died.
  26. Protein-Energy Malnutrition in this column we observe that majority of data are says that very less persons died.

27. Road Injuries in this column we observe that majority of data are says that around 12500 persons died.
28. Chronic Respiratory Diseases in this column we observe that majority of data are says that around 2500 persons died.
29. Cirrhosis and Other Chronic Liver Diseases in this column we observe that majority of data are says that around 10000 persons died.
30. Digestive Diseases in this column we observe that majority of data are says that around 3000 persons died.
31. Fire, Heat, and Hot Substances in this column we observe that majority of data are says that around 2500 persons died.
32. Acute Hepatitis in this column we observe that majority of data are says that very less persons died.

- Now we observe for check the how much death in each disease.
- For that we make one data frame by use of for loop first we make one list for total disease in the use of sum.

```
summ=['Year', 'Meningitis',
       "Alzheimer's Disease and Other Dementias", "Parkinson's Disease",
       'Nutritional Deficiencies', 'Malaria', 'Drowning',
       'Interpersonal Violence', 'Maternal Disorders', 'HIV/AIDS',
       'Drug Use Disorders', 'Tuberculosis', 'Cardiovascular Diseases',
       'Lower Respiratory Infections', 'Neonatal Disorders',
       'Alcohol Use Disorders', 'Self-harm', 'Exposure to Forces of Nature',
       'Diarrheal Diseases', 'Environmental Heat and Cold Exposure',
       'Neoplasms', 'Conflict and Terrorism', 'Diabetes Mellitus',
       'Chronic Kidney Disease', 'Poisonings', 'Protein-Energy Malnutrition',
       'Road Injuries', 'Chronic Respiratory Diseases',
       'Cirrhosis and Other Chronic Liver Diseases', 'Digestive Diseases',
       'Fire, Heat, and Hot Substances', 'Acute Hepatitis']

sum_of=[]
for i in summ:
    sum_of.append(sum(df[i]))
```

```
df1=pd.DataFrame({
    'Name of diseases':summ,
    'Total diseases':sum_of
})
df1
```

- Now we take a look of it our Data Frame.

	Name of diseases	Total diseases
0	Year	12267540
1	Meningitis	10524572
2	Alzheimer's Disease and Other Dementias	29768839
3	Parkinson's Disease	7179795
4	Nutritional Deficiencies	13792032
5	Malaria	25342676
6	Drowning	10301999
7	Interpersonal Violence	12752839
8	Maternal Disorders	7727046
9	HIV/AIDS	36364419
10	Drug Use Disorders	2656121
11	Tuberculosis	45850603
12	Cardiovascular Diseases	447741982
13	Lower Respiratory Infections	83770038
14	Neonatal Disorders	76860729
15	Alcohol Use Disorders	4819018
16	Self-harm	23713931
17	Exposure to Forces of Nature	1490132
18	Diarrheal Diseases	66235508
19	Environmental Heat and Cold Exposure	1788851
20	Neoplasms	229758538
21	Conflict and Terrorism	3294053
22	Diabetes Mellitus	31448872
23	Chronic Kidney Disease	28911692
24	Poisonings	2601082
25	Protein-Energy Malnutrition	12031885
26	Road Injuries	36296469
27	Chronic Respiratory Diseases	104605334
28	Cirrhosis and Other Chronic Liver Diseases	37479321
29	Digestive Diseases	65638635
30	Fire, Heat, and Hot Substances	3602914
31	Acute Hepatitis	3784791

- Now take a look of our Data Frame.

  observation: -

  1. Year in these diseases there are total 12267540 persons are died.
  2. Meningitis in these diseases there are total 10524572 persons are died.
  3. Alzheimer's Disease and Other Dementias in this diseases there are total 29768839 persons are died.
  4. Parkinson's Disease in this diseases there are total 7179795 persons are died.
  5. Nutritional Deficiencies in this diseases there are total 13792032 persons are died.
  6. Malaria in this diseases there are total 25342676 persons are died.
  7. Drowning in this diseases there are total 10301999 persons are died.
  8. Interpersonal Violence in this diseases there are total 12752839 persons are died.
  9. Maternal Disorders in this diseases there are total 7727046 persons are died.

10. HIV/AIDS in this diseases there are total 36364419 persons are died.
11. Drug Use Disorders in this diseases there are total 2656121 persons are died.
12. Tuberculosis in this diseases there are total 45850603 persons are died.
13. Cardiovascular Diseases in this diseases there are total 447741982 persons are died.
14. Lower Respiratory Infections in this diseases there are total 83770038 persons are died.
15. Neonatal Disorders in this diseases there are total 76860729 persons are died.
16. Alcohol Use Disorders in this diseases there are total 4819018 persons are died.
17. Self-harm in this diseases there are total 23713931 persons are died.
18. Exposure to Forces of Nature in this diseases there are total 1490132 persons are died.
19. Diarrheal Diseases in this diseases there are total 66235508 persons are died.
20. Environmental Heat and Cold Exposure in this diseases there are total 1788851 persons are died.
21. Neoplasms in this diseases there are total 229758538 persons are died.
22. Conflict and Terrorism in this diseases there are total 3294053 persons are died.
23. Diabetes Mellitus in this diseases there are total 31448872 persons are died.
24. Chronic Kidney Disease in this diseases there are total 28911692 persons are died.
25. Poisonings in this diseases there are total 2601082 persons are died.
26. Protein-Energy Malnutrition in this diseases there are total 12031885 persons are died.
27. Road Injuries in this diseases there are total 36296469 persons are died.
28. Chronic Respiratory Diseases in this diseases there are total 104605334 persons are died.
29. Cirrhosis and Other Chronic Liver Diseases in this diseases there are total 37479321 persons are died.
30. Digestive Diseases in this diseases there are total 65638635 persons are died.
31. Fire, Heat, and Hot Substances in this diseases there are total 3602914 persons are died.
32. Acute Hepatitis in this diseases there are total 3784791 persons are died.

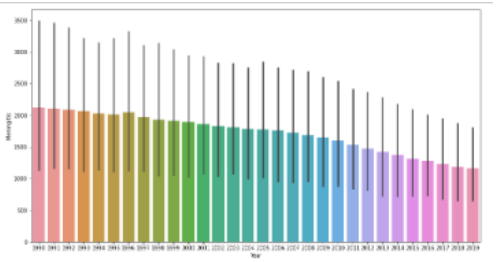
- **As we observe in the diseases of Cardiovascular there are maximum death 447741982.**
- **As we observe in the diseases of Exposure to Forces of Nature there are maximum death 1490132.**
- **As we observe there are total 1480402256 persons are died.**

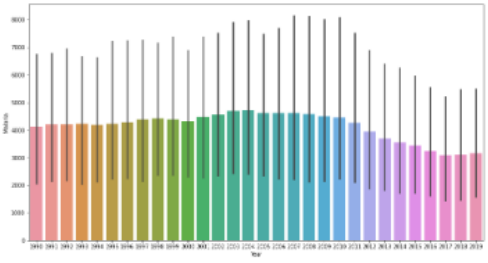
Now we start for Bivariate Analysis.

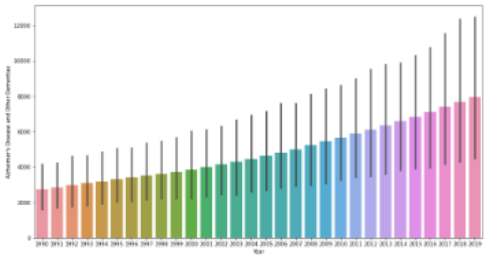
- For plot graph we use for loop for that.

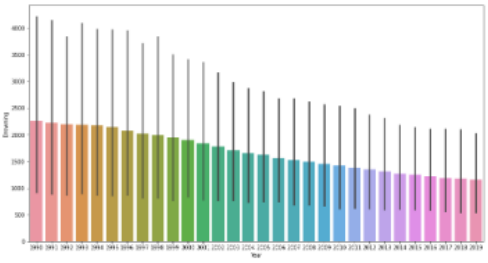
```
for i in summ[1:]:
    plt.figure(figsize=(15,8))
    sns.barplot(x='Year',y=i,data=df)
    plt.show()
```

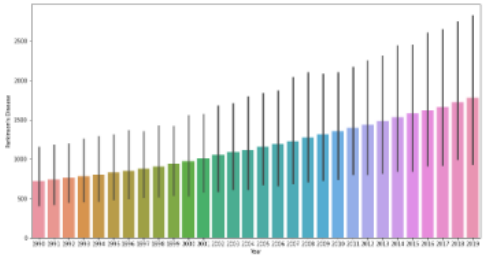
- Now we plot all the graph and then we analysis each by each.

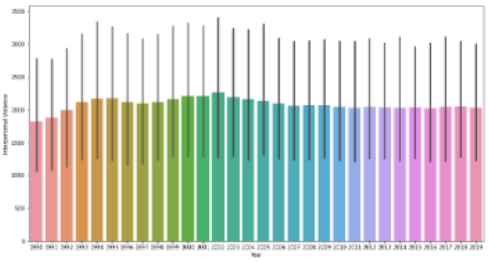

Measles
Year

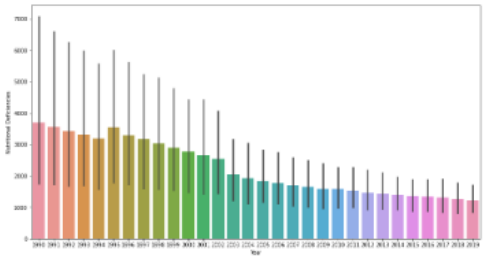

Malaria
Year

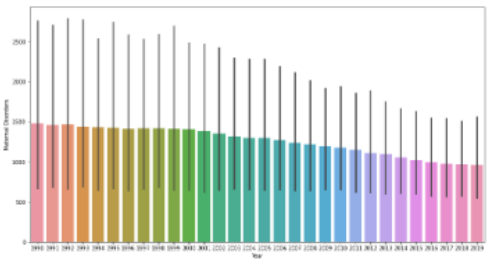

Alzheimer's Disease and Other Dementias
Year

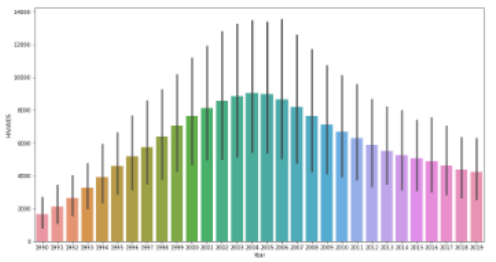

Ethiopia
Year

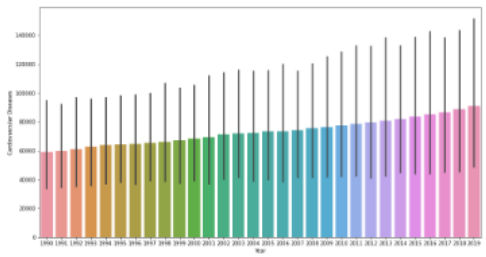

Parkinson's Disease
Year

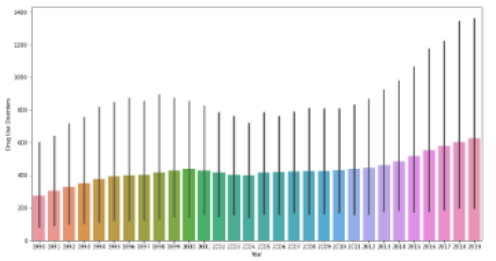

Interpersonal Violence
Year

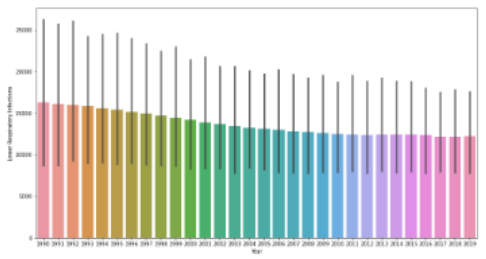

Nutritional Deficiencies
Year

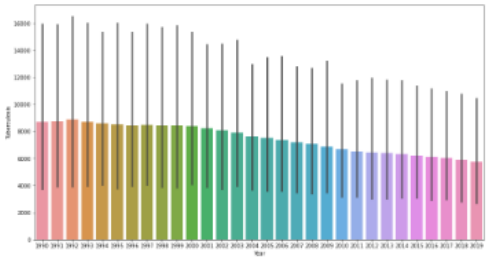

Malaria Disorders
Year

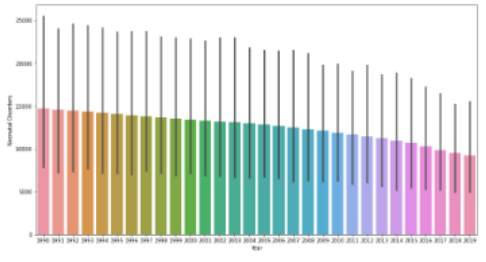

HIV/AIDS
Year

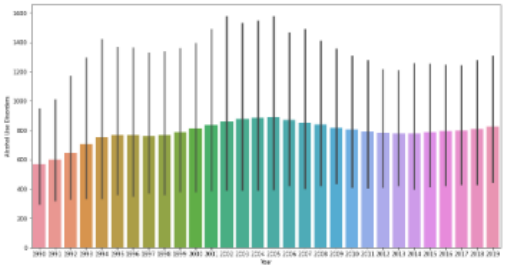

Cardiovascular Diseases
Year

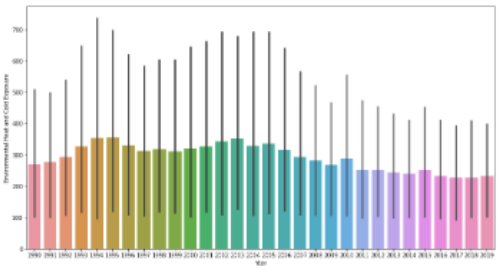

Drug Use Disorders
Year

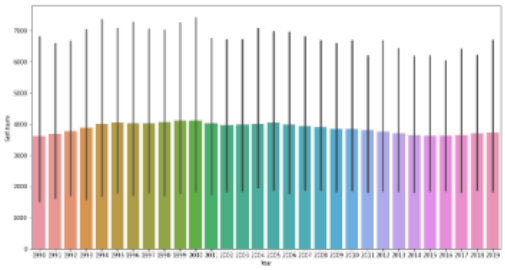

Lower Respiratory Infections
Year

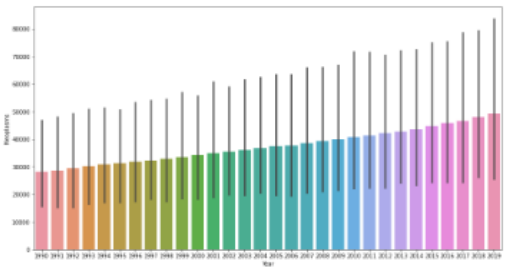

Tuberculosis
Year

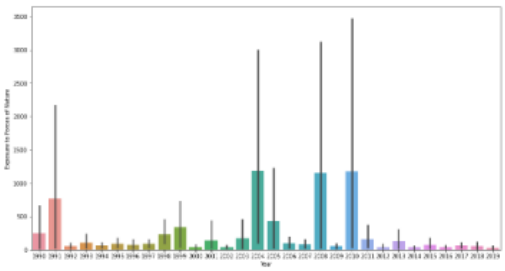

Neonatal Disorders
Year

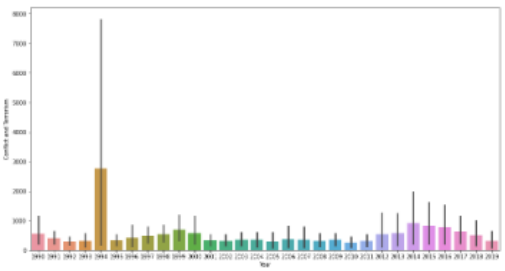

Alcohol Use Disorders
Year

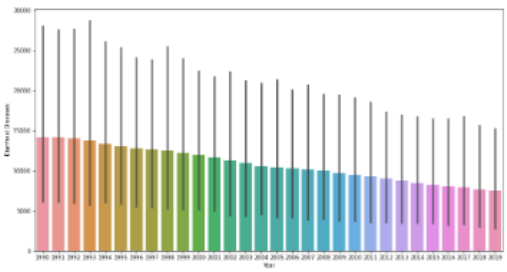

Environmental Heat and Cold Exposure
Year

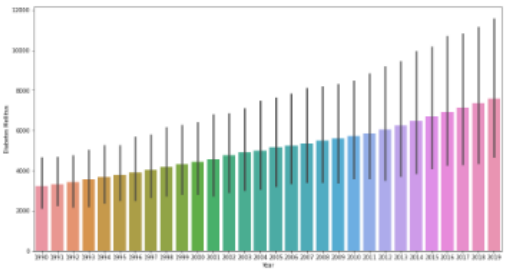

Self-harm
Year

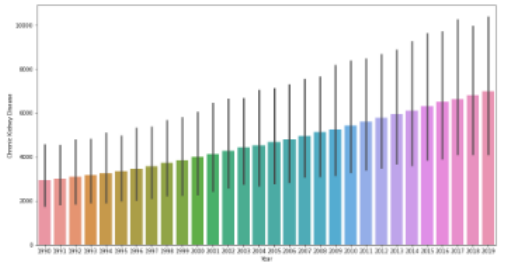

Respiratory
Year

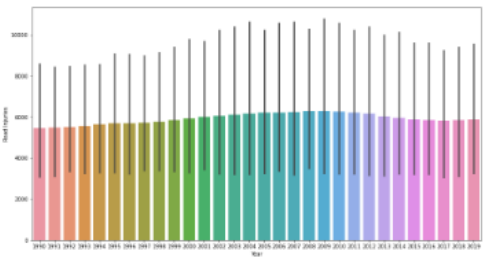

Exposure to Forces of Nature
Year

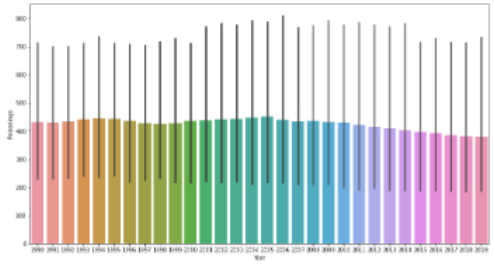

Conflict and Terrorism
Year

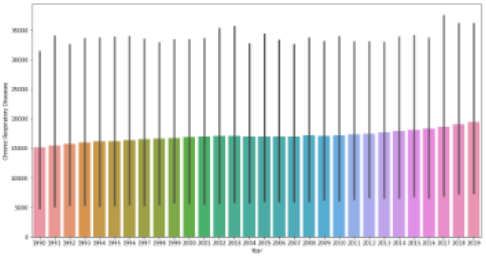

Diarrheal Diseases
Year

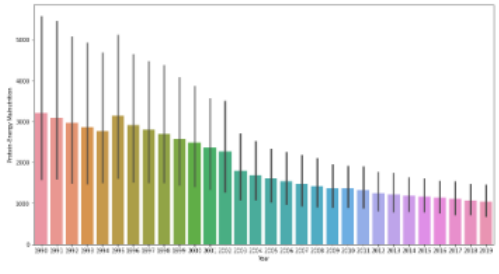

Diabetes Mellitus
Year

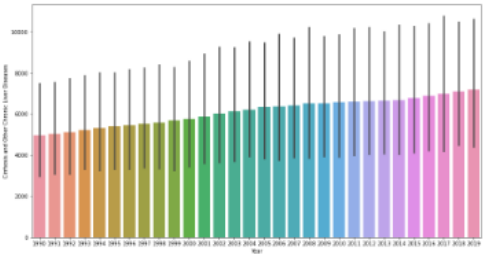

Chronic Kidney Disease
Year

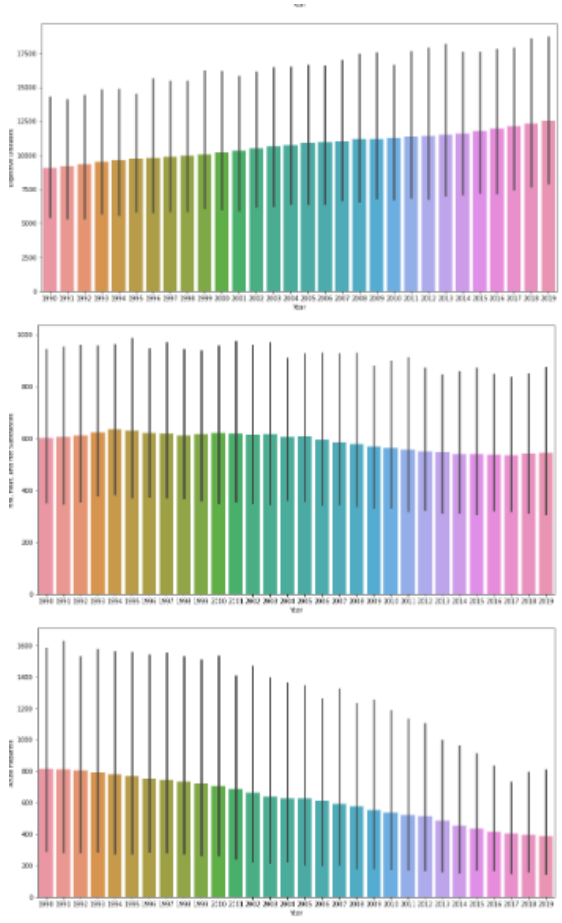

Road Injuries
Year


Poisonings
Year


Chronic Respiratory Diseases
Year


Protein-Energy Malnutrition
Year


Cirrhosis and Other Chronic Liver Diseases
Year



- By observing this graphs we write some observations.

  Observation:-

  1. As we observe in this figure we observe that by increase year the death is decrease in the Meningitis diseases.

  2. As we observe in this figure we observe that by increase year the death is also increase in the Alzheimer's Disease and Other Dementias diseases.

  3. As we observe in this figure we observe that by increase year the death is also increase in the Parkinson's Disease diseases.

  4. As we observe in this figure we observe that by increase year the death is decrease in the Nutritional Deficiencies diseases.

  5. As we observe in this figure we observe that by year this is saturated but in last 4 to 5 years it will decrease the death of Malaria diseases.

  6. As we observe in this figure we observe that by increase year the death is decrease in the Drowning diseases.

  7. As we observe in this figure we observe that this is saturated and also stable in all the years. for Interpersonal Violence diseases death.

  8. As we observe in this figure we observe that by increase year the death is decrease in the Maternal Disorders diseases.

  9. As we observe in this figure we observe that in the year of 2001 to 2008 the death is higher now a days it will decrease by year in the HIV//AIDS diseases.

10. As we observe in this figure we observe that by increase year the death is also increase in the Drug Use Disorders diseases.

11. As we observe in this figure we observe that by increase year the death is decrease in the Tuberculosis diseases.

12. As we observe in this figure we observe that by increase year the death is also increase in the Cardiovascular diseases.

13. As we observe in this figure we observe that by increase year the death is decrease in the Lower Respiratory Infections diseases.

14. As we observe in this figure we observe that by increase year the death is decrease in the Neonatal Disorders diseases.

15. As we observe in this figure we observe that by increase year it will increase but in the year of 2004 and 2005 Alcohol Use Disorders diseases is high.

16. As we observe in this figure we observe that this is saturated and also stable in all the years. for Self-harm diseases death.

17. As we observe for Exposure to Forces of Nature death there are only the year of 1990,1991,2004,2008,2010 it was higher.

18. As we observe in this figure we observe that by increase year the death is decrease in the Diarrheal diseases.

19. As we observe in this figure we observe that this is saturated and also stable in all the years. for Environmental Heat and Cold Exposure diseases death.

20. As we observe in this figure we observe that by increase year the death is also increase in the Neoplasms diseases.

21. As we observe for Conflict and Terrorism there are only the year of 1994,2014 it was higher.

22. As we observe in this figure we observe that by increase year the death is also increase in the Diabetes Mellitus diseases.

23. As we observe in this figure we observe that by increase year the death is also increase in the Chronic Kidney Disease diseases.

24. As we observe in this figure we observe that this is saturated and also stable in all the years. for Poisonings diseases death.

25. As we observe in this figure we observe that by increase year the death is decrease in the Protein-Energy Malnutrition diseases.

26. As we observe in this figure we observe that this is saturated and also stable in all the years. for Road Injuries diseases death.

27. As we observe in this figure we observe that by increase year the death is also increase in the Chronic Respiratory diseases.

28. As we observe in this figure we observe that by increase year the death is also increase in the Cirrhosis and Other Chronic Liver diseases.

29. As we observe in this figure we observe that by increase year the death is also increase in the Digestive diseases.

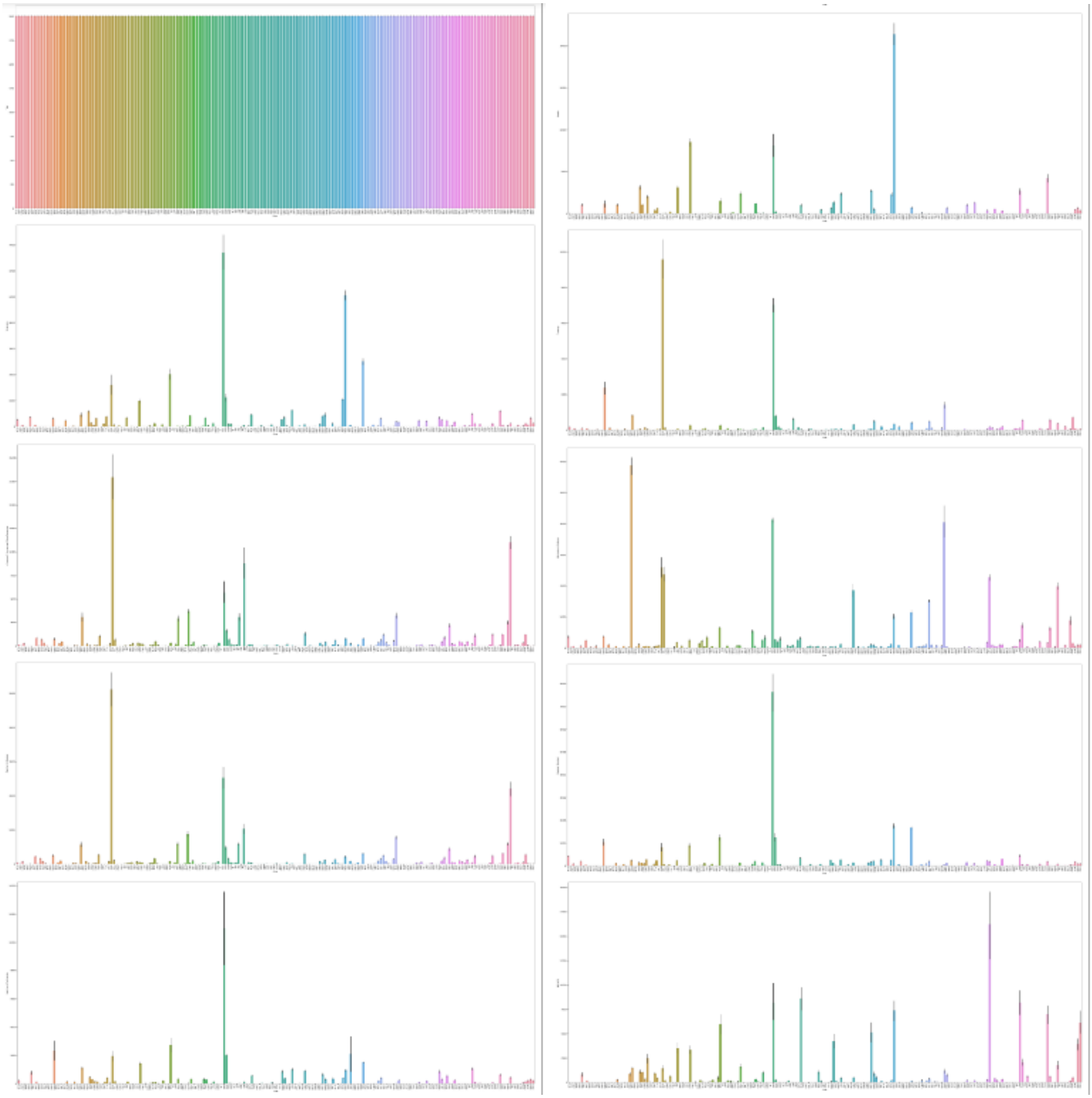
30. As we observe in this figure we observe that this is saturated and also stable in all the years. for Fire, Heat, and Hot Substances diseases death.

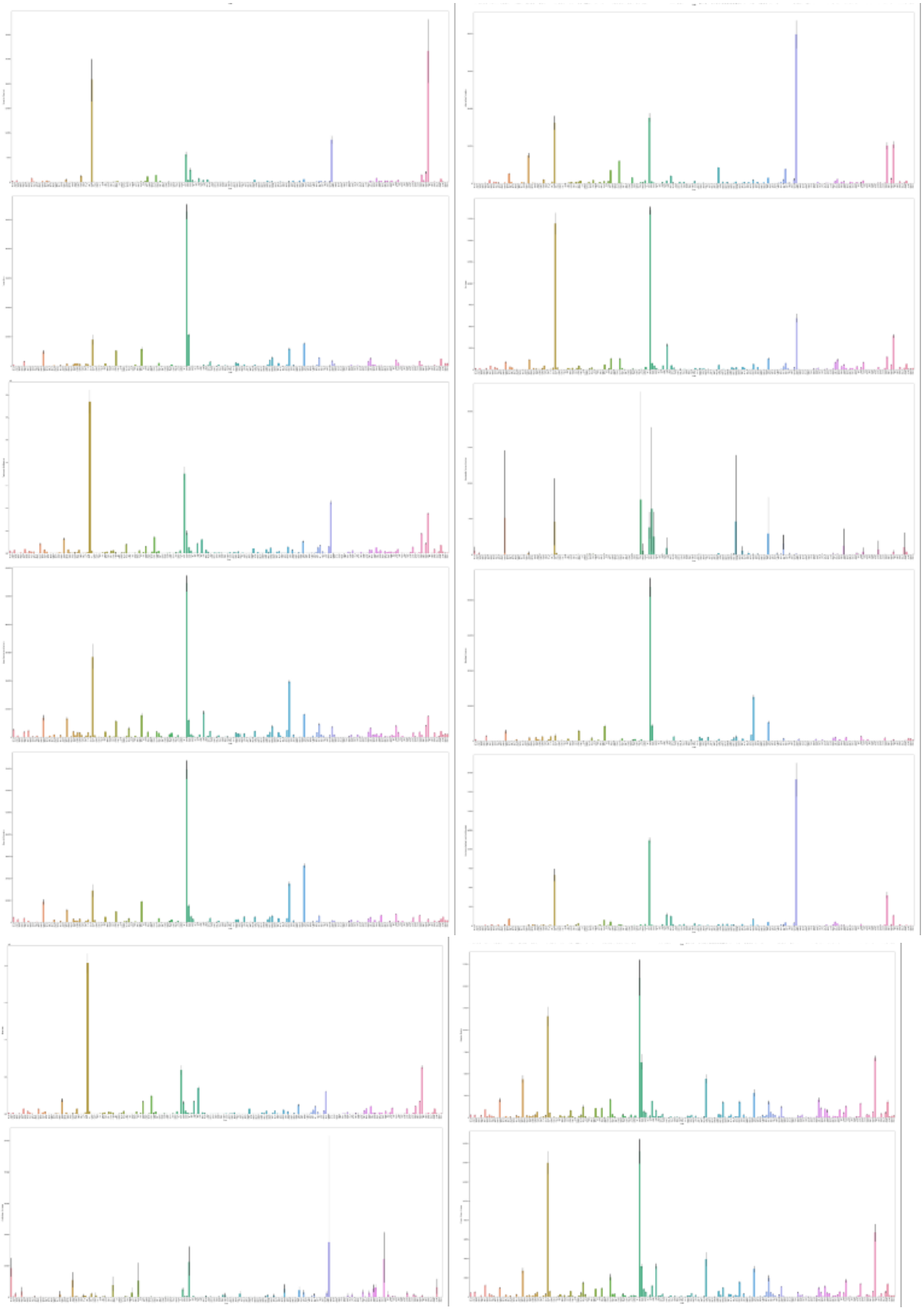
31. As we observe in this figure we observe that by increase year the death is decrease in the Acute Hepatitis diseases.

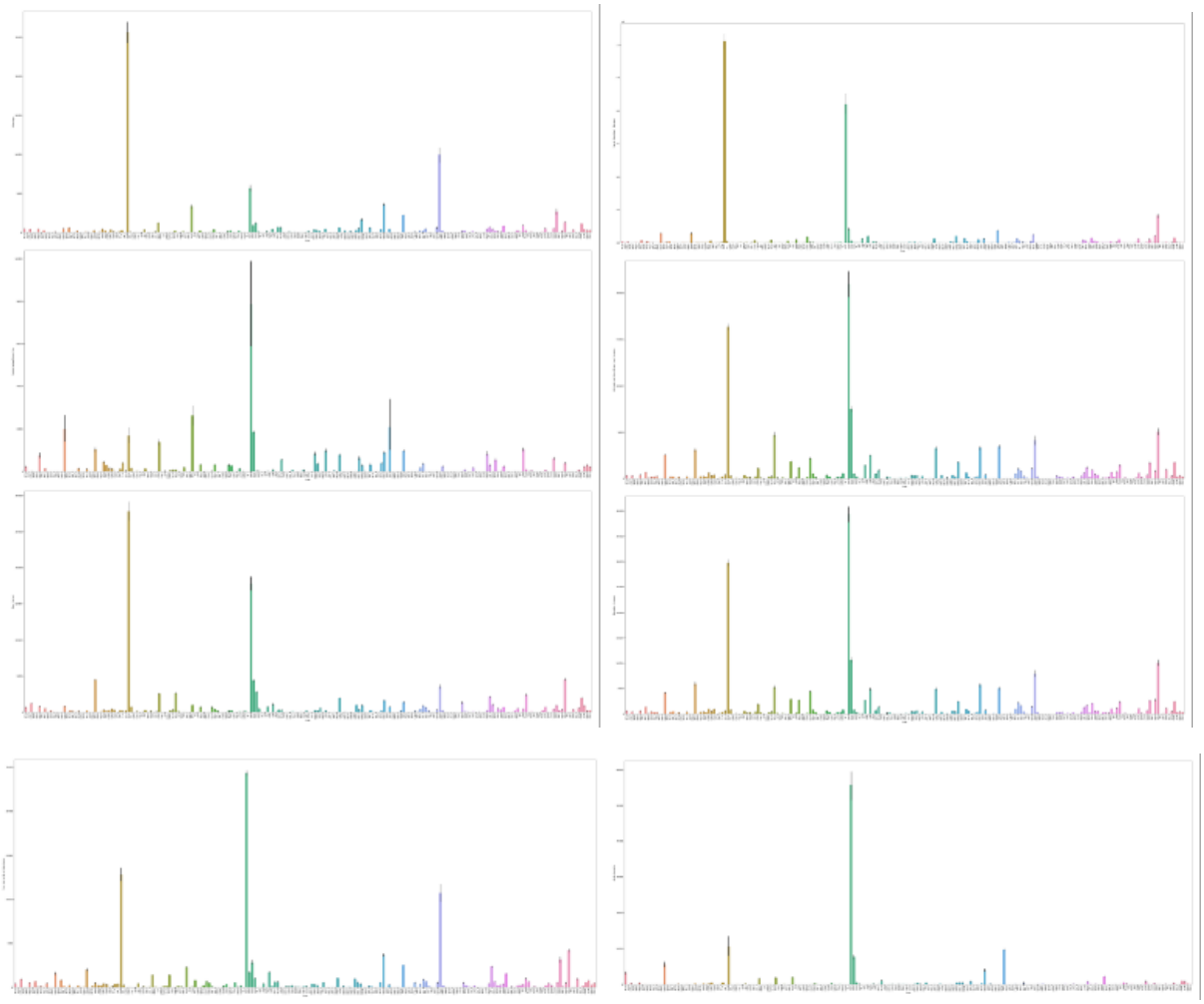
- Above we observe our data by use of year now we observe for code meand contry code.
- For that first we plot all the graph by use of for loop and then analysis it.

```
for i in summ:
    plt.figure(figsize=(50,20))
    sns.barplot(x='Code',y=i,data=df)
    plt.xticks(rotation=90)
    plt.show()
```

- Now check output of it.







- Now observe data and write observation of it.

observation:-

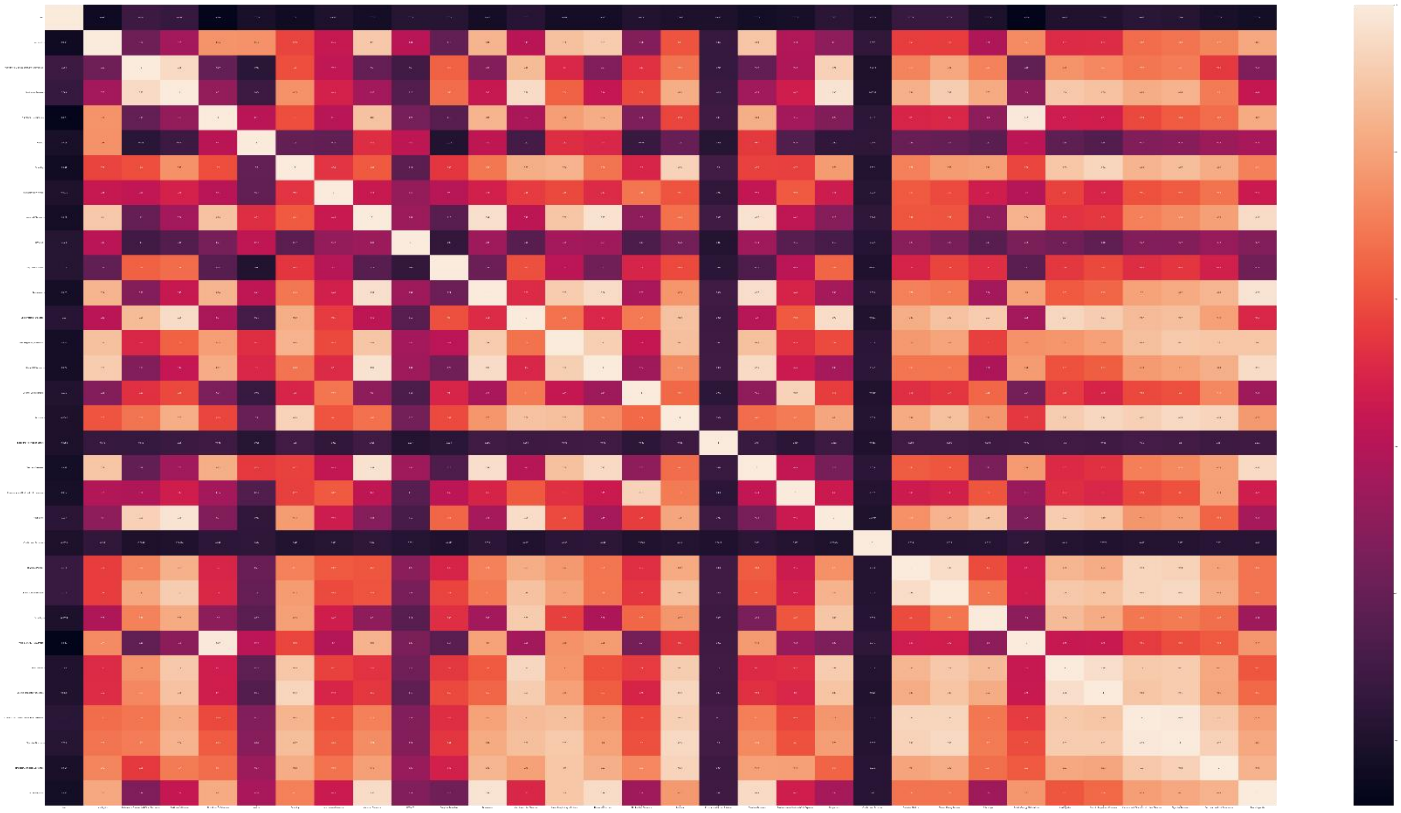
1. As we observe in the graph we observe that all the country are same death rate by year.
2. As we observe in the graph we observe that IND,NGK and PAK has highest death in the diseases of Meningitis.
3. As we observe in the graph we observe that CHN,USA,JPN and IND has highest death in the diseases of Alzheimer's Disease and Other Dementias.
4. As we observe in the graph we observe that CHN,USA and IND has highest death in the diseases of Parkinson's Disease.
5. As we observe in the graph we observe that ETH and IND has highest death in the diseases of Nutritional Deficiencies.
6. As we observe in the graph we observe that NGA,COD and IND has highest death in the diseases of Malaria.
7. As we observe in the graph we observe that CHN and IND has highest death in the diseases of Drowning.
8. As we observe in the graph we observe that BRA,CHN,COL,IND and RUS has highest death in the diseases of Interpersonal Violence.
9. As we observe in the graph we observe that IND has highest death in the diseases of Maternal Disorders.
10. As we observe in the graph we observe that ZAF,KEN and IND has highest death in the diseases of HIV/AIDS.
11. As we observe in the graph we observe that USA,CHN and RUS has highest death in the diseases of Drug Use Disorders.
12. As we observe in the graph we observe that IDN and IND has highest death in the diseases of Tuberculosis.

13. As we observe in the graph we observe that CHN and IND has highest death in the diseases of Cardiovascular Diseases.
14. As we observe in the graph we observe that NGA,CHN and IND has highest death in the diseases of Lower Respiratory Infections.
15. As we observe in the graph we observe that PAK and IND has highest death in the diseases of Neonatal Disorders.
16. As we observe in the graph we observe that RUS,CHN and IND has highest death in the diseases of Alcohol Use Disorders.
17. As we observe in the graph we observe that CHN and IND has highest death in the diseases of Self-harm.
18. As we observe in the graph we observe that IDN,HTI,BGD,CHN and MMR has highest death in the diseases of Exposure to Forces of Nature.
19. As we observe in the graph we observe that NGA and IND has highest death in the diseases of Diarrheal Diseases.
20. As we observe in the graph we observe that RUS,CHA and IND has highest death in the diseases of Environmental Heat and Cold Exposure.
21. As we observe in the graph we observe that CHA and USA has highest death in the diseases of Neoplasms.
22. As we observe in the graph we observe that RWA and SYR has highest death in the diseases of Conflict and Terrorism.
23. As we observe in the graph we observe that IDN,CHA and IND has highest death in the diseases of Diabetes Mellitus.
24. As we observe in the graph we observe that CHA and IND has highest death in the diseases of Chronic Kidney Disease.
25. As we observe in the graph we observe that CHA and RUS has highest death in the diseases of Poisonings.
26. As we observe in the graph we observe that ETH and IND has highest death in the diseases of Protein-Energy Malnutrition.
27. As we observe in the graph we observe that CHA and IND has highest death in the diseases of Road Injuries.
28. As we observe in the graph we observe that CHA and IND has highest death in the diseases of Chronic Respiratory Diseases.
29. As we observe in the graph we observe that IDN,CHA and IND has highest death in the diseases of Cirrhosis and Other Chronic Liver Diseases.
30. As we observe in the graph we observe that IDN,CHA and IND has highest death in the diseases of Digestive Diseases.
31. As we observe in the graph we observe that CHA,RUS and IND has highest death in the diseases of Fire, Heat, and Hot Substances.
32. As we observe in the graph we observe that IND has highest death in the diseases of Acute Hepatitis.

- Now we plot Multivariate data by use of df.corr() function.

```
plt.figure(figsize=(150,80))
sns.heatmap(df.corr(),annot=True)
```

- By use of above code we plot graph.



- As we observe in above graph there are majority of the columns are in positively correlated

Power BI dashboard

- We made one simple PowerBI Dashboard for data analysis here the outlook of it.