In [1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px
import scipy.stats

import warnings
warnings.filterwarnings("ignore")
```

In [2]:
```
#Setting the display for our graph display

sns.set_style('darkgrid')
plt.rcParams['font.size'] = 15
plt.rcParams ['figure.figsize'] = (10, 7)
plt.rcParams ['figure.facecolor'] = '#FFE5B4'
```

In [3]:
```
df=pd.read_csv("cause_of_deaths dataset.csv")
df
```

Out[3]:

	Country/Territory	Code	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Inter
0	Afghanistan	AFG	1990	2159	1116	371	2087	93	1370	
1	Afghanistan	AFG	1991	2218	1136	374	2153	189	1391	
2	Afghanistan	AFG	1992	2475	1162	378	2441	239	1514	
3	Afghanistan	AFG	1993	2812	1187	384	2837	108	1687	
4	Afghanistan	AFG	1994	3027	1211	391	3081	211	1809	
...	...	...	...	...	...	...	...	...	...	
6115	Zimbabwe	ZWE	2015	1439	754	215	3019	2518	770	
6116	Zimbabwe	ZWE	2016	1457	767	219	3056	2050	801	
6117	Zimbabwe	ZWE	2017	1460	781	223	2990	2116	818	
6118	Zimbabwe	ZWE	2018	1450	795	227	2918	2088	825	
6119	Zimbabwe	ZWE	2019	1450	812	232	2884	2068	827	

6120 rows × 34 columns

In [4]:
```
pd.set_option('display.max_columns',None) #  This will enable us to see truncated column
df.head()
```

Out[4]:

	Country/Territory	Code	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interper Vio
0	Afghanistan	AFG	1990	2159	1116	371	2087	93	1370	
1	Afghanistan	AFG	1991	2218	1136	374	2153	189	1391	

2	Afghanistan	AFG	1992	2475	1162	378	2441	239	1514
3	Afghanistan	AFG	1993	2812	1187	384	2837	108	1687
4	Afghanistan	AFG	1994	3027	1211	391	3081	211	1809

In [5]:
```
df.columns
```

Out[5]:
```
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')
```

In [6]:
```
# Data information
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6120 entries, 0 to 6119
Data columns (total 34 columns):
 #   Column                                        Non-Null Count  Dtype
---  ------                                        --------------  -----
 0   Country/Territory                             6120 non-null   object
 1   Code                                          6120 non-null   object
 2   Year                                          6120 non-null   int64
 3   Meningitis                                    6120 non-null   int64
 4   Alzheimer's Disease and Other Dementias       6120 non-null   int64
 5   Parkinson's Disease                           6120 non-null   int64
 6   Nutritional Deficiencies                      6120 non-null   int64
 7   Malaria                                       6120 non-null   int64
 8   Drowning                                      6120 non-null   int64
 9   Interpersonal Violence                        6120 non-null   int64
 10  Maternal Disorders                            6120 non-null   int64
 11  HIV/AIDS                                      6120 non-null   int64
 12  Drug Use Disorders                            6120 non-null   int64
 13  Tuberculosis                                  6120 non-null   int64
 14  Cardiovascular Diseases                       6120 non-null   int64
 15  Lower Respiratory Infections                  6120 non-null   int64
 16  Neonatal Disorders                            6120 non-null   int64
 17  Alcohol Use Disorders                         6120 non-null   int64
 18  Self-harm                                     6120 non-null   int64
 19  Exposure to Forces of Nature                  6120 non-null   int64
 20  Diarrheal Diseases                            6120 non-null   int64
 21  Environmental Heat and Cold Exposure          6120 non-null   int64
 22  Neoplasms                                     6120 non-null   int64
 23  Conflict and Terrorism                        6120 non-null   int64
 24  Diabetes Mellitus                             6120 non-null   int64
 25  Chronic Kidney Disease                        6120 non-null   int64
 26  Poisonings                                    6120 non-null   int64
 27  Protein-Energy Malnutrition                   6120 non-null   int64
 28  Road Injuries                                 6120 non-null   int64
 29  Chronic Respiratory Diseases                  6120 non-null   int64
 30  Cirrhosis and Other Chronic Liver Diseases    6120 non-null   int64
 31  Digestive Diseases                            6120 non-null   int64
 32  Fire, Heat, and Hot Substances                6120 non-null   int64
 33  Acute Hepatitis                               6120 non-null   int64
```

```
dtypes: int64(32), object(2)
memory usage: 1.6+ MB
```

In [7]:
```
df.duplicated().sum()
```

Out[7]:
```
0
```

No duplicates present

In [8]:
```
df.isin(['NAN','NA','N/A','-',' ','?',np.nan]).sum().any()
```

Out[8]:
```
False
```

No empty spaces present

In [9]:
```
df.isnull().sum().sum()
```

Out[9]:
```
0
```

No null values present

In [10]:
```
# Value counts for each feature data
for i in df.columns:
    print(df[i].value_counts())
    print('******************')
```

```
Afghanistan         30
Papua New Guinea    30
Niue                30
North Korea         30
North Macedonia     30
                    ..
Greenland           30
Grenada             30
Guam                30
Guatemala           30
Zimbabwe            30
Name: Country/Territory, Length: 204, dtype: int64
******************
AFG    30
PNG    30
NIU    30
PRK    30
MKD    30
       ..
GRL    30
GRD    30
GUM    30
GTM    30
ZWE    30
Name: Code, Length: 204, dtype: int64
******************
1990    204
1991    204
2018    204
2017    204
2016    204
2015    204
2014    204
2013    204
2012    204
2011    204
2010    204
```

```
2009    204
2008    204
2007    204
2006    204
2005    204
2004    204
2003    204
2002    204
2001    204
2000    204
1999    204
1998    204
1997    204
1996    204
1995    204
1994    204
1993    204
1992    204
2019    204
Name: Year, dtype: int64
*****************
0        316
1        261
2        221
3        160
6         80
        ...
3260       1
3412       1
3486       1
3562       1
55021      1
Name: Meningitis, Length: 2020, dtype: int64
*****************
0       78
2       66
8       55
1       55
16      54
        ..
6996     1
6766     1
6488     1
6220     1
795      1
Name: Alzheimer's Disease and Other Dementias, Length: 3037, dtype: int64
*****************
4       167
5       139
1       123
3       119
6       110
       ...
1093      1
995       1
1004      1
1416      1
256       1
Name: Parkinson's Disease, Length: 1817, dtype: int64
*****************
1       305
0       283
4       210
2       186
3       151
       ...
```

```
3320      1
3298      1
3271      1
3427      1
2990      1
Name: Nutritional Deficiencies, Length: 2147, dtype: int64
*****************
0       3431
1        169
2         52
4         51
5         44
        ...
1367       1
2020       1
1147       1
1292       1
2068       1
Name: Malaria, Length: 1723, dtype: int64
*****************
4       134
0       120
6       108
1        94
2        94
       ...
7884      1
7616      1
7379      1
6717      1
770       1
Name: Drowning, Length: 1875, dtype: int64
*****************
0       165
6       120
5       116
3       100
2        96
       ...
3080      1
3005      1
2878      1
2870      1
1434      1
Name: Interpersonal Violence, Length: 2142, dtype: int64
*****************
1       430
0       426
2       260
3       186
5       139
       ...
716       1
430       1
850       1
884       1
1312      1
Name: Maternal Disorders, Length: 1818, dtype: int64
*****************
1        306
2        249
0        235
3        146
4        117
        ...
1908       1
```

```
1976       1
1975       1
1933       1
20722      1
Name: HIV/AIDS, Length: 2412, dtype: int64
*****************
0       580
1       571
2       246
3       173
6       151
       ...
5692      1
5527      1
5434      1
5279      1
231       1
Name: Drug Use Disorders, Length: 876, dtype: int64
*****************
1        260
0        196
2        109
3         85
4         81
        ...
4722       1
4684       1
4621       1
4628       1
10465      1
Name: Tuberculosis, Length: 2843, dtype: int64
*****************
5        20
8        15
42       12
4        10
242       8
         ..
39650     1
38648     1
37991     1
36915     1
17810     1
Name: Cardiovascular Diseases, Length: 5225, dtype: int64
*****************
1        54
11       39
25       34
17       33
10       31
         ..
10368     1
10014     1
9852      1
9698      1
12897     1
Name: Lower Respiratory Infections, Length: 4106, dtype: int64
*****************
1        99
0        91
2        55
5        50
12       49
         ..
12576     1
12827     1
```

```
13000     1
13196     1
8609      1
Name: Neonatal Disorders, Length: 3553, dtype: int64
*****************
1       314
0       279
2       175
6       140
5       135
       ...
1628      1
1587      1
1570      1
1566      1
839       1
Name: Alcohol Use Disorders, Length: 1287, dtype: int64
*****************
3       145
2       126
8        76
5        66
4        65
       ...
2001      1
2122      1
2132      1
2517      1
2403      1
Name: Self-harm, Length: 2758, dtype: int64
*****************
0       3418
1        282
2        193
3        115
5        104
        ...
323        1
1497       1
694        1
2065       1
660        1
Name: Exposure to Forces of Nature, Length: 478, dtype: int64
*****************
1        261
0        159
2        149
3        146
4        138
        ...
12273      1
11149      1
10438      1
10115      1
4635       1
Name: Diarrheal Diseases, Length: 2874, dtype: int64
*****************
0       841
1       471
2       260
3       216
4       189
       ...
1593      1
1577      1
1523      1
```

```
1546      1
190       1
Name: Environmental Heat and Cold Exposure, Length: 714, dtype: int64
*****************
3        30
2        19
8        16
7        14
11       12
         ..
18509     1
23726     1
24536     1
25363     1
12353     1
Name: Neoplasms, Length: 4814, dtype: int64
*****************
0       3208
1        339
2        191
3        107
4        103
        ...
1712       1
2018       1
1790       1
2592       1
372        1
Name: Conflict and Terrorism, Length: 918, dtype: int64
*****************
3       44
5       41
4       39
6       32
1       30
        ..
215      1
220      1
231      1
255      1
3460     1
Name: Diabetes Mellitus, Length: 3366, dtype: int64
*****************
4       55
3       52
1       42
7       28
29      28
        ..
1717     1
1820     1
1848     1
1969     1
2160     1
Name: Chronic Kidney Disease, Length: 3246, dtype: int64
*****************
0       584
1       419
2       264
3       112
4        85
       ...
728        1
830        1
893        1
1179       1
```

```
304       1
Name: Poisonings, Length: 1087, dtype: int64
**********
0       413
1       353
2       265
3       243
4       196
       ...
3126      1
2941      1
2892      1
2854      1
2855      1
Name: Protein-Energy Malnutrition, Length: 2091, dtype: int64
**********
3       99
0       60
11      58
4       50
6       49
        ..
1013     1
967      1
995      1
1072     1
2554     1
Name: Road Injuries, Length: 3393, dtype: int64
**********
7       49
1       41
9       36
8       36
21      31
        ..
2387     1
2479     1
2576     1
2860     1
2891     1
Name: Chronic Respiratory Diseases, Length: 3803, dtype: int64
**********
2       92
7       61
8       56
0       51
9       49
        ..
382      1
387      1
427      1
435      1
2065     1
Name: Cirrhosis and Other Chronic Liver Diseases, Length: 3443, dtype: int64
**********
4       64
17      44
1       43
18      38
37      31
        ..
6075     1
6039     1
6352     1
6634     1
4377     1
```

```
Name: Digestive Diseases, Length: 4023, dtype: int64
*****************
0       334
2       276
1       166
3       151
4       105
       ...
1731      1
1708      1
1684      1
1574      1
590       1
Name: Fire, Heat, and Hot Substances, Length: 1406, dtype: int64
*****************
0       840
1       517
3       310
2       275
4       186
       ...
676       1
545       1
601       1
573       1
2985      1
Name: Acute Hepatitis, Length: 1059, dtype: int64
*****************
```

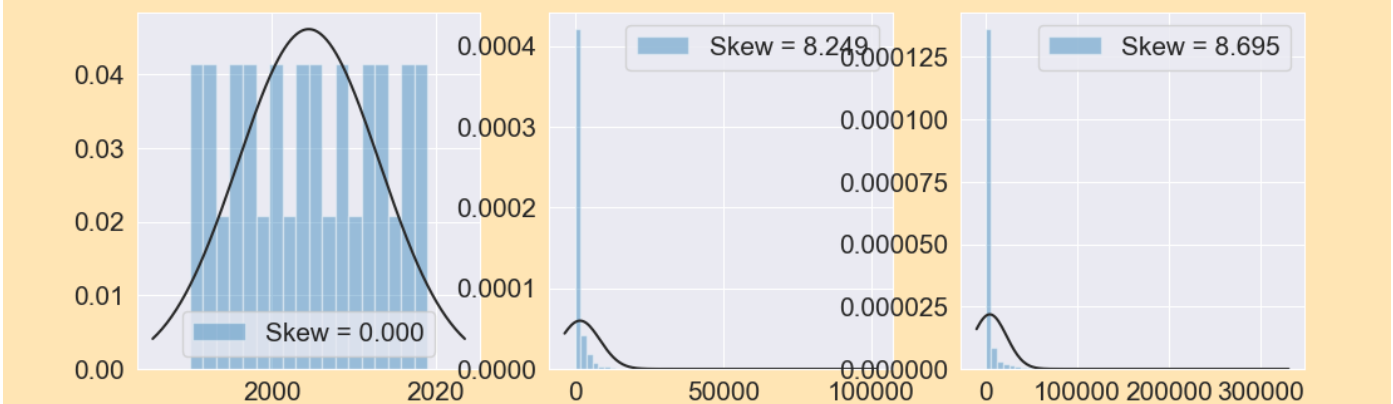
Dropping column "Code" as it is same as country.

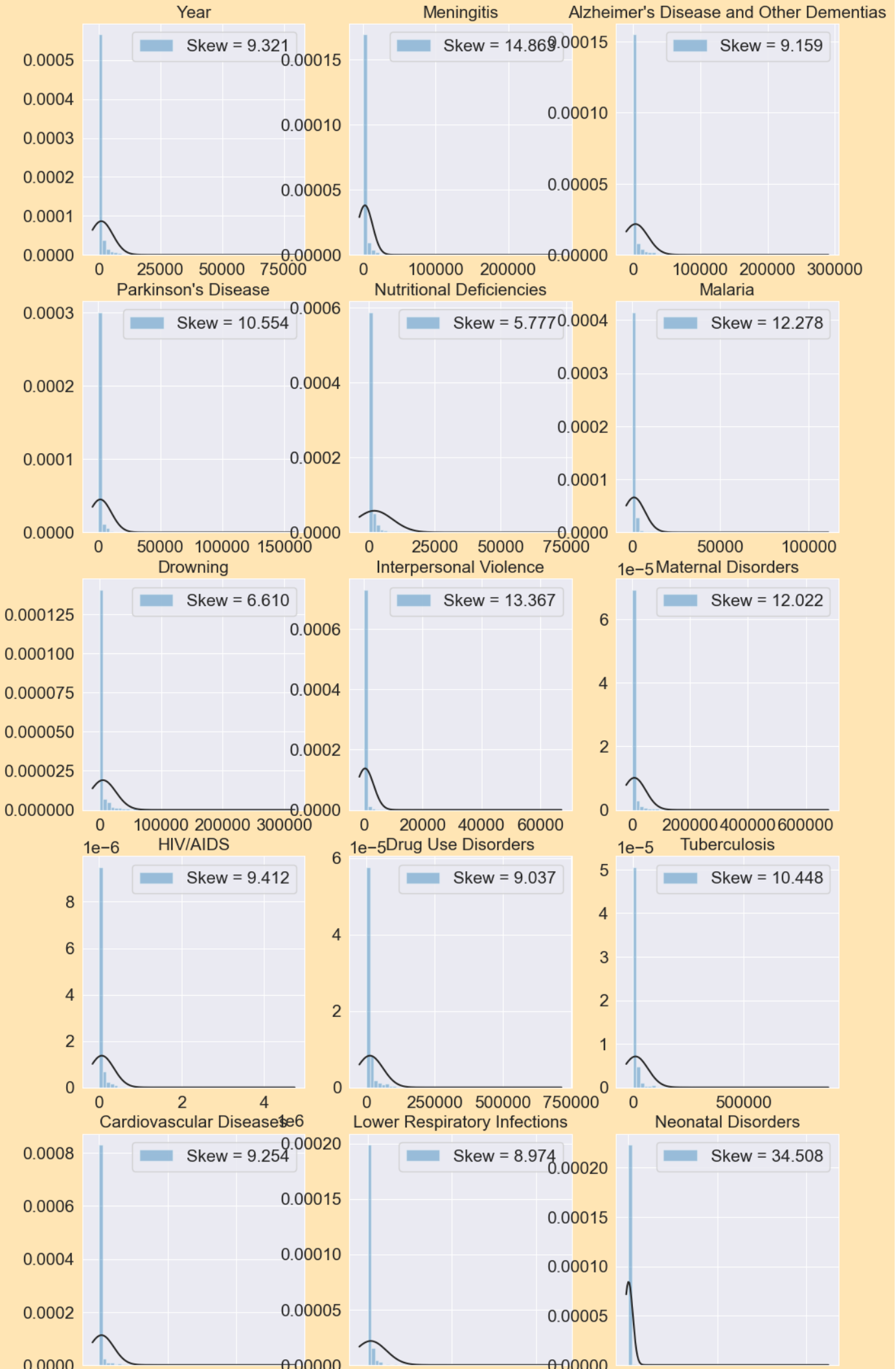
In [11]:
```
df.drop("Code",axis=1,inplace=True)
```

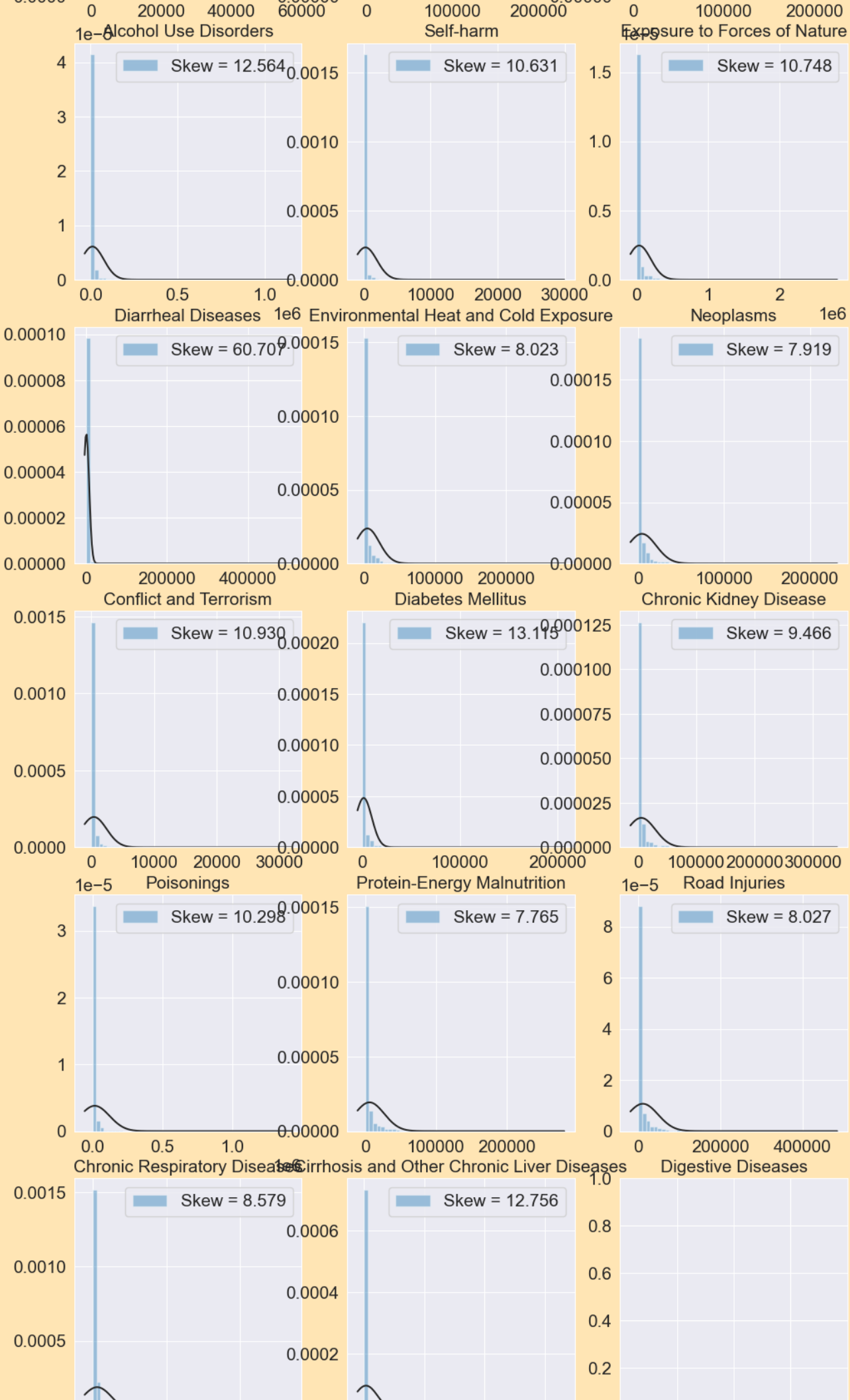
Exploratory Data Analysis

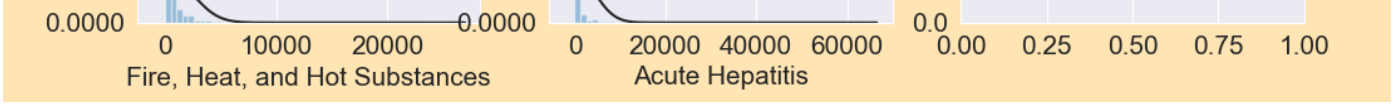
Univariate Analysis

In [12]:
```
#plotting a histogram of each independent feature for visualization
from scipy import stats
cols = df.columns[1:]
fig,ax = plt.subplots(11,3, figsize=(12, 48))
ax = ax.flatten()
i = 0
for col in cols:
    skew = df[col].skew()
    sns.distplot(df[col], ax = ax[i], fit= stats.norm, kde=False, label='Skew = %.3f' %(
    ax[i].legend(loc='best')
    i += 1
plt.show()
```




Year
Skew = 9.321
Meningitis
Skew = 14.869
Alzheimer's Disease and Other Dementias
Skew = 9.159
Parkinson's Disease
Skew = 10.554
Nutritional Deficiencies
Skew = 5.777
Malaria
Skew = 12.278
Drowning
Skew = 6.610
Interpersonal Violence
Skew = 13.367
1e−5 Maternal Disorders
Skew = 12.022
1e−6 HIV/AIDS
Skew = 9.412
1e−5 Drug Use Disorders
Skew = 9.037
1e−5 Tuberculosis
Skew = 10.448
Cardiovascular Diseases
1e6
Lower Respiratory Infections
Neonatal Disorders
Skew = 9.254
Skew = 8.974
Skew = 34.508


Alcohol Use Disorders
Skew = 12.564
Self-harm
Skew = 10.631
Exposure to Forces of Nature
Skew = 10.748
Diarrheal Diseases
Skew = 60.707
Environmental Heat and Cold Exposure
Skew = 8.023
Neoplasms
Skew = 7.919
Conflict and Terrorism
Skew = 10.930
Diabetes Mellitus
Skew = 13.115
Chronic Kidney Disease
Skew = 9.466
Poisonings
Skew = 10.298
Protein-Energy Malnutrition
Skew = 7.765
Road Injuries
Skew = 8.027
Chronic Respiratory Diseases
Skew = 8.579
Cirrhosis and Other Chronic Liver Diseases
Skew = 12.756
Digestive Diseases



In [13]:
```
df.skew()
```

Out[13]:
```
Year                                            0.000000
Meningitis                                      8.248599
Alzheimer's Disease and Other Dementias         8.695288
Parkinson's Disease                             9.321242
Nutritional Deficiencies                       14.863496
Malaria                                         9.159105
Drowning                                       10.553901
Interpersonal Violence                          5.777146
Maternal Disorders                             12.278358
HIV/AIDS                                        6.610169
Drug Use Disorders                             13.367100
Tuberculosis                                   12.022406
Cardiovascular Diseases                         9.411914
Lower Respiratory Infections                    9.036604
Neonatal Disorders                             10.447536
Alcohol Use Disorders                           9.253511
Self-harm                                       8.973653
Exposure to Forces of Nature                   34.507640
Diarrheal Diseases                             12.563758
Environmental Heat and Cold Exposure           10.631493
Neoplasms                                      10.748333
Conflict and Terrorism                         60.707004
Diabetes Mellitus                               8.022653
Chronic Kidney Disease                          7.919364
Poisonings                                     10.929760
Protein-Energy Malnutrition                    13.115196
Road Injuries                                   9.466209
Chronic Respiratory Diseases                   10.298131
Cirrhosis and Other Chronic Liver Diseases      7.764715
Digestive Diseases                              8.026536
Fire, Heat, and Hot Substances                  8.578848
Acute Hepatitis                                12.756098
dtype: float64
```

The skewness is very high for almost all the columns

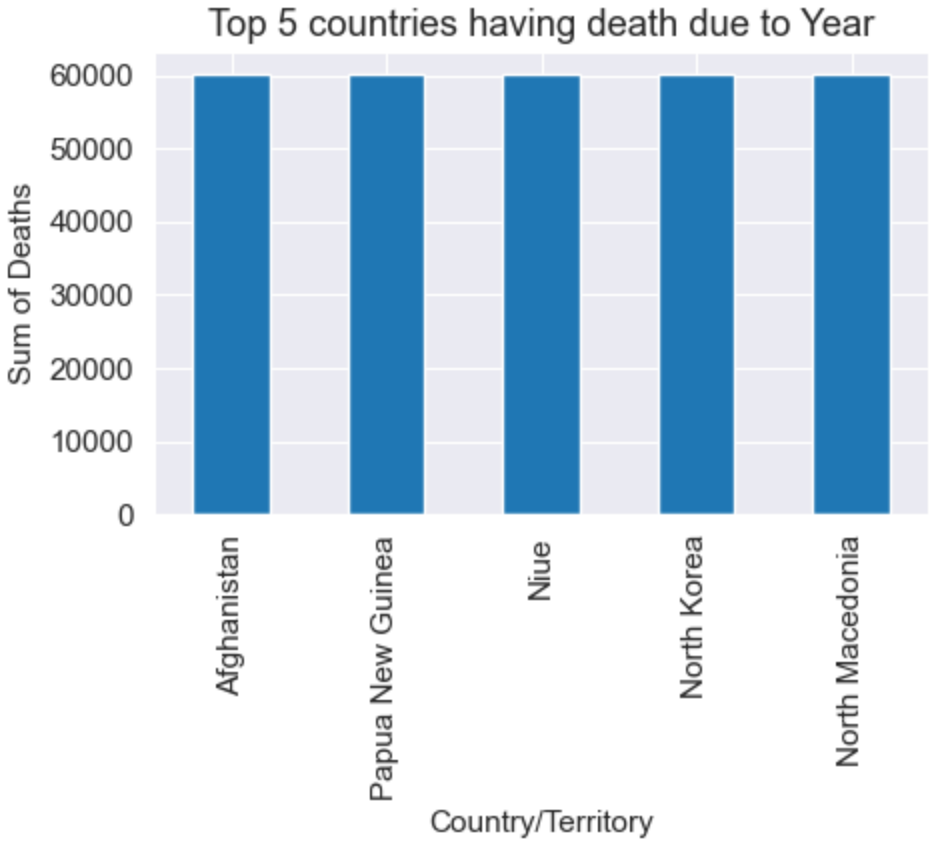
In [14]:
```
print(df['Year'].unique())
print(df['Year'].nunique())
```

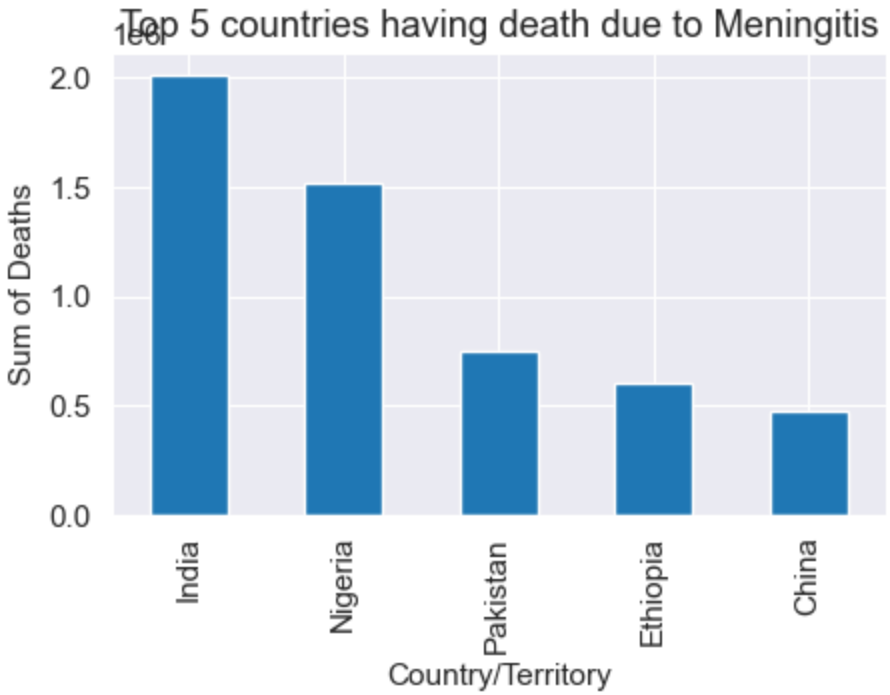
```
[1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
 2018 2019]
30
```

There are 30 years of statistics in this data set (1990-2019).

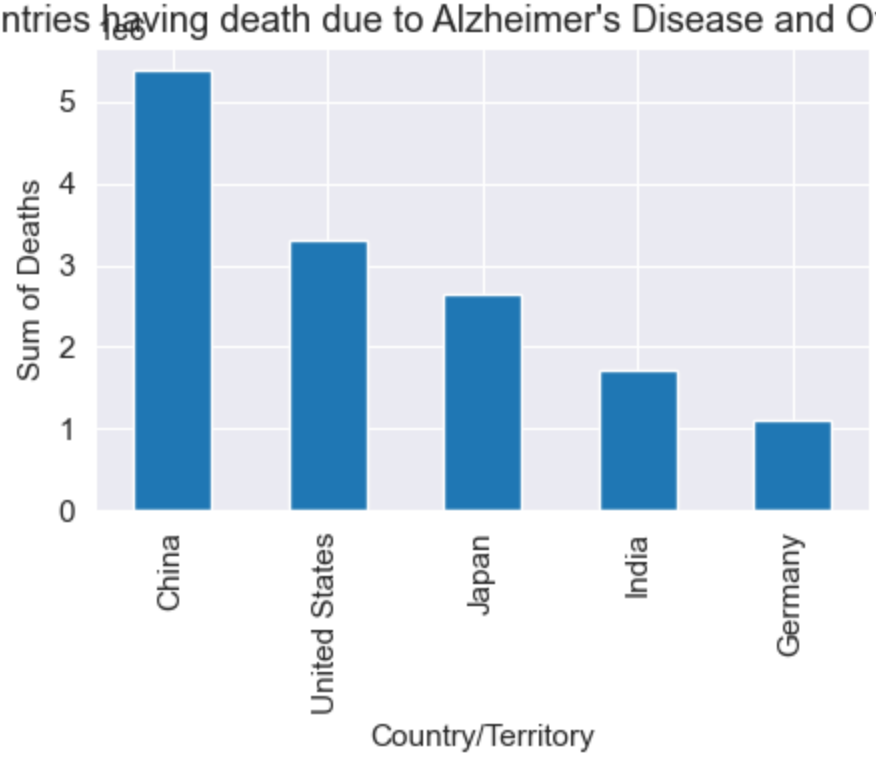
In [15]:
```
sns.set_style('darkgrid')
plt.rcParams['font.size'] = 11
plt.rcParams ['figure.figsize'] = (5, 3)

for i in df.columns[1:]:
    df.groupby('Country/Territory')[i].sum().sort_values(ascending=False).head().plot(ki
    plt.title('Top 5 countries having death due to '+i)
    plt.ylabel('Sum of Deaths')
    plt.show()
```

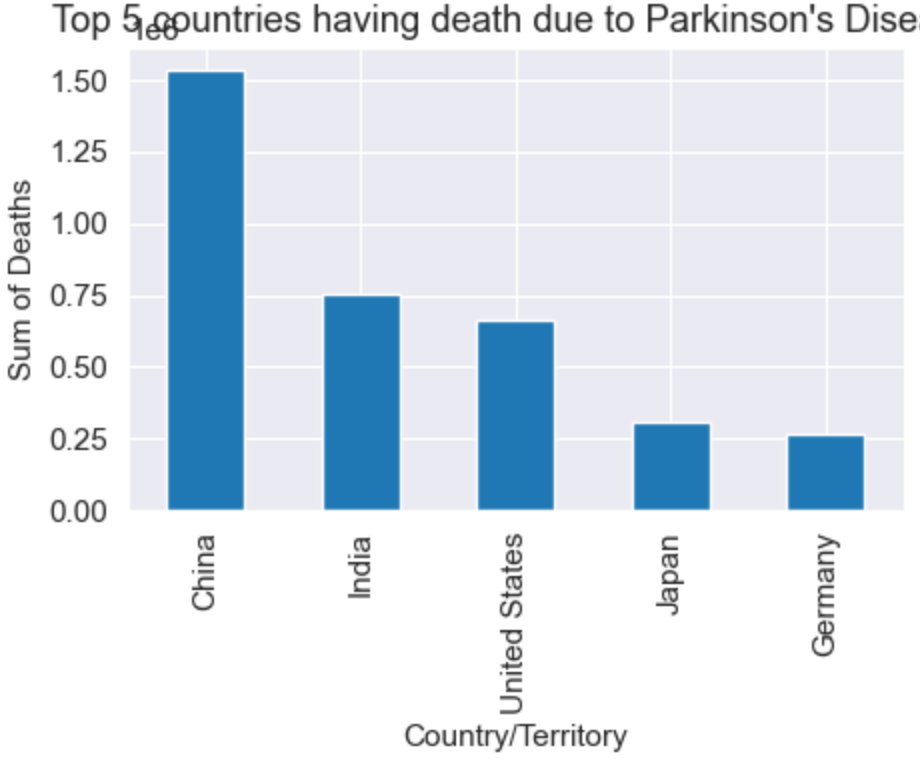

Top 5 countries having death due to Year
Sum of Deaths
60000
50000
40000
30000
20000
10000
0
Afghanistan
Papua New Guinea
Niue
North Korea
North Macedonia
Country/Territory

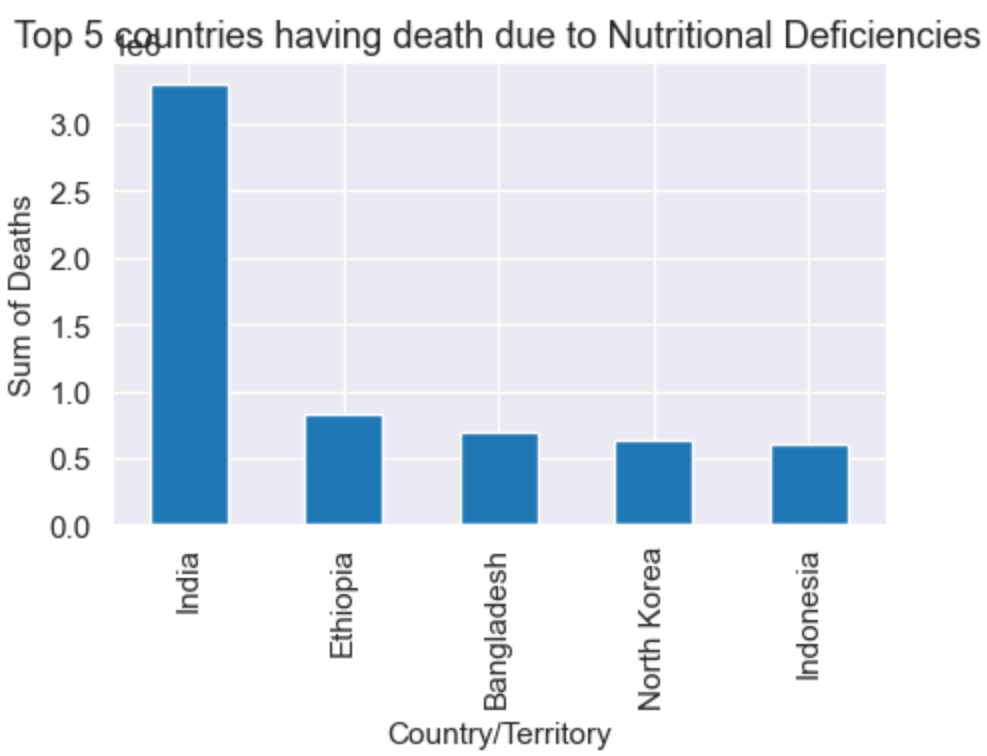

Top 5 countries having death due to Meningitis
1e6
Sum of Deaths
2.0
1.5
1.0
0.5
0.0
India
Nigeria
Pakistan
Ethiopia
China
Country/Territory

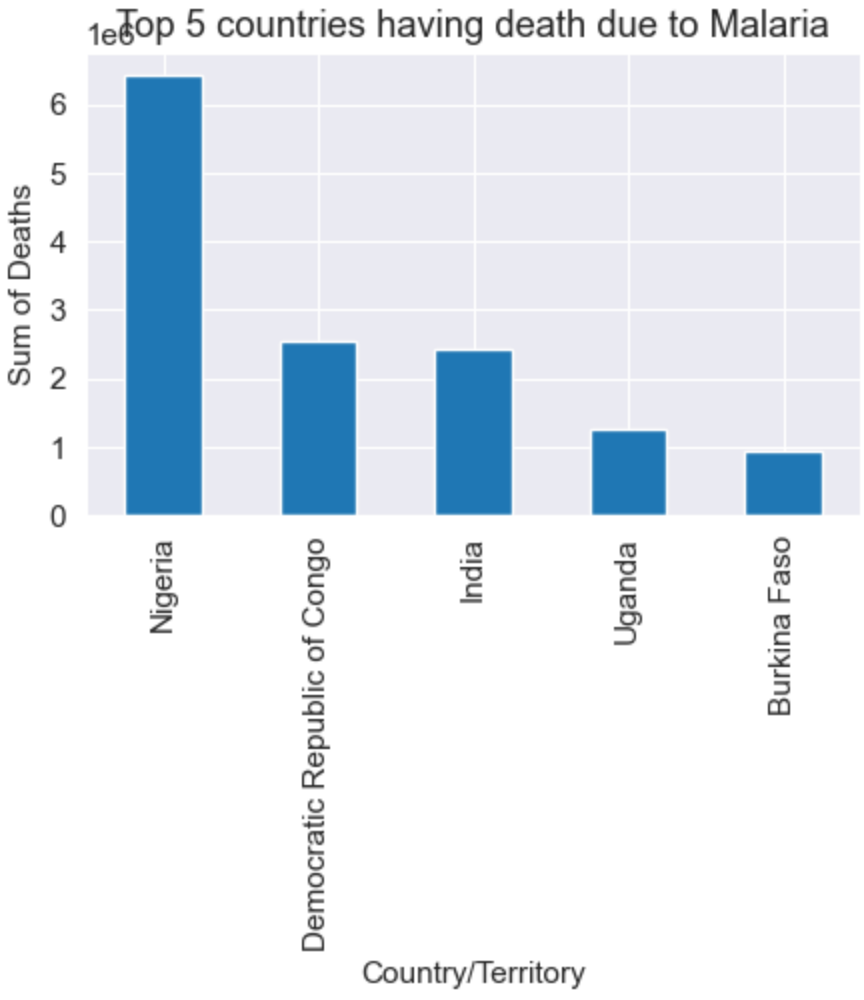
Top 5 countries having death due to Alzheimer's Disease and Other Dementias

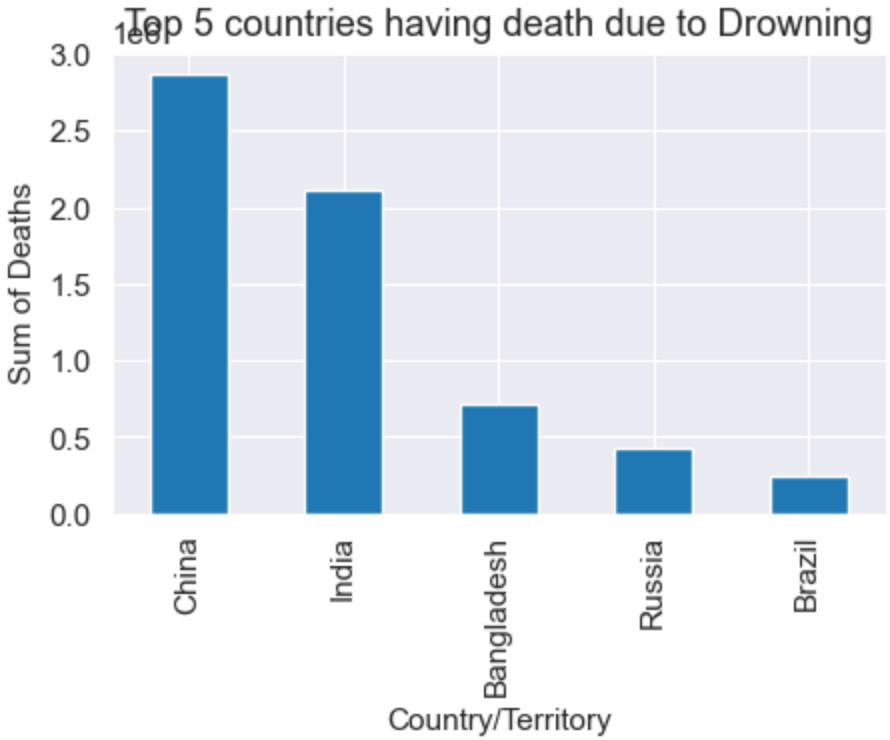


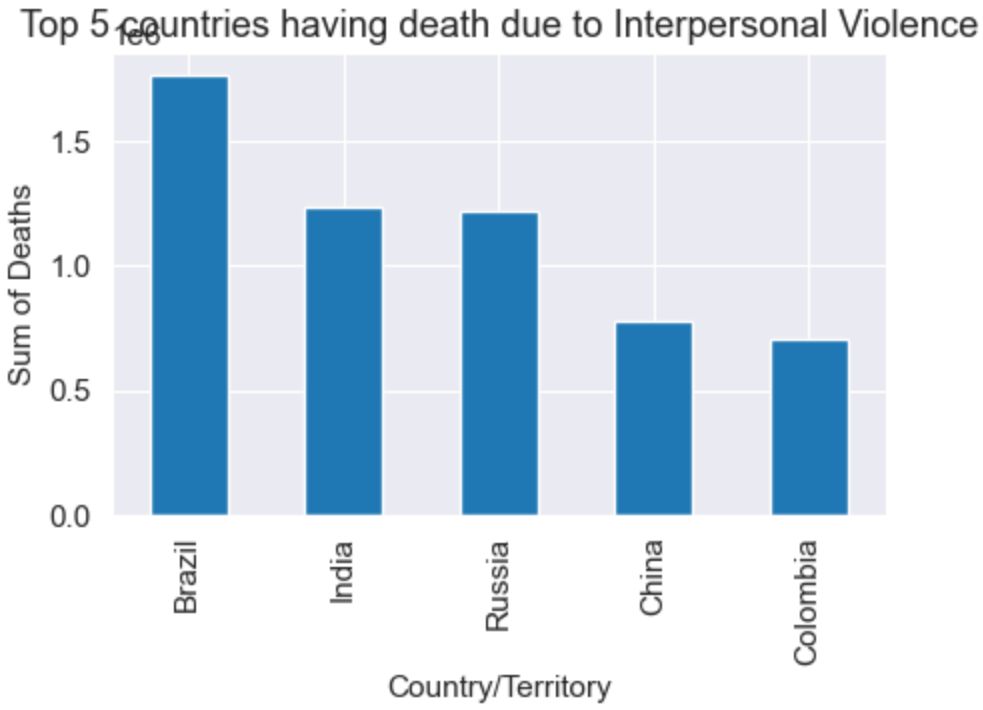
Top 5 countries having death due to Parkinson's Disease

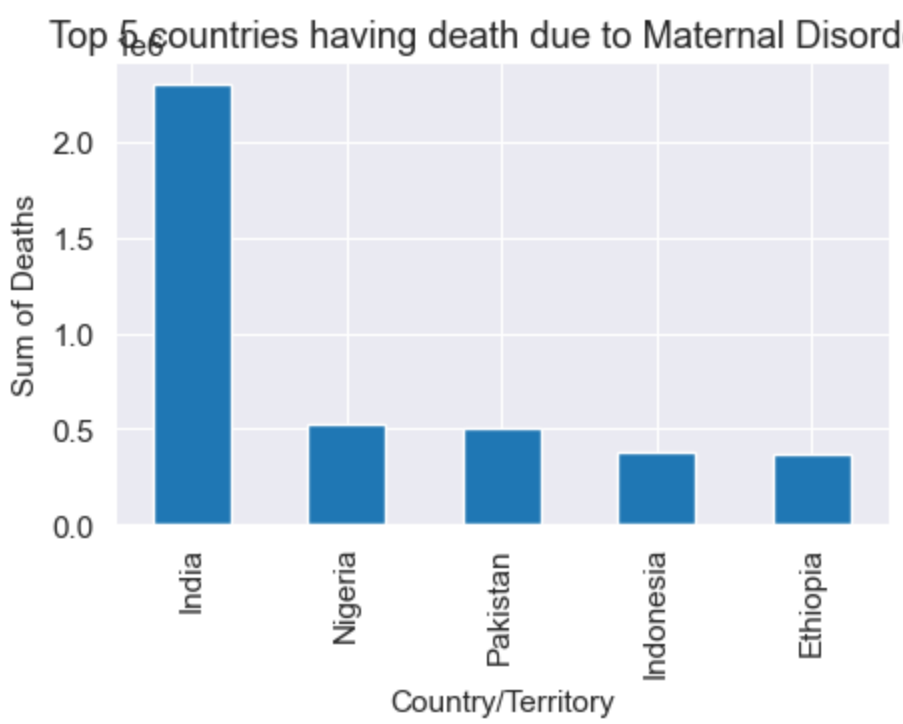


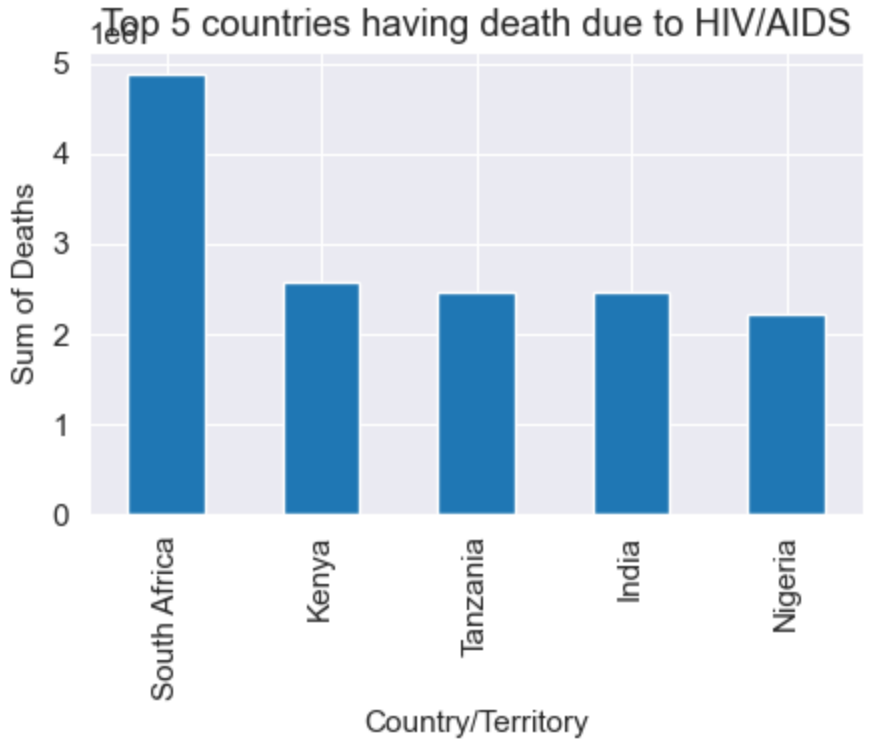

Top 5 countries having death due to Nutritional Deficiencies
1e8
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Sum of Deaths
India
Ethiopia
Bangladesh
North Korea
Indonesia
Country/Territory

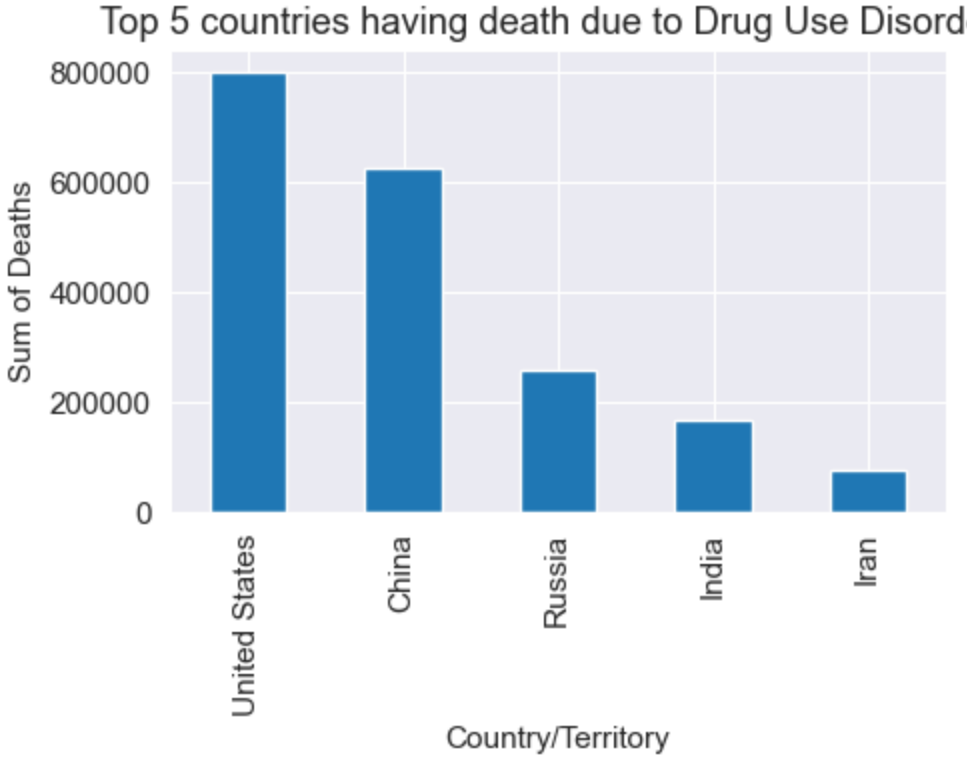

Top 5 countries having death due to Malaria
1e6
6
5
4
3
2
1
0
Sum of Deaths
Nigeria
Democratic Republic of Congo
India
Uganda
Burkina Faso
Country/Territory

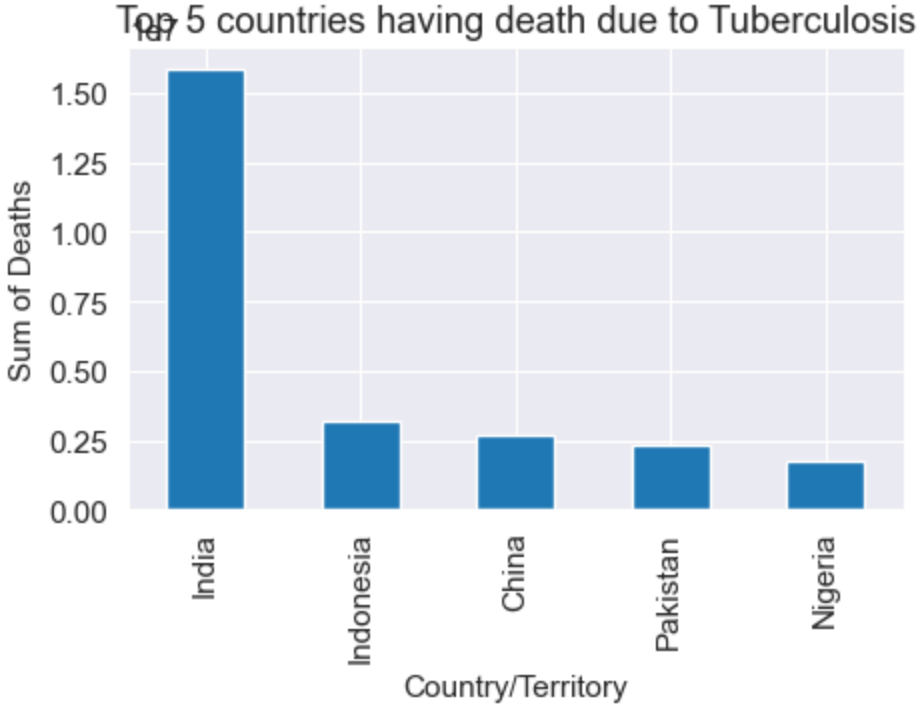

Top 5 countries having death due to Drowning
1e8
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Sum of Deaths
China
India
Bangladesh
Russia
Brazil
Country/Territory


Top 5 countries having death due to Interpersonal Violence
1e8
1.5
1.0
0.5
0.0
Sum of Deaths
Brazil
India
Russia
China
Colombia
Country/Territory

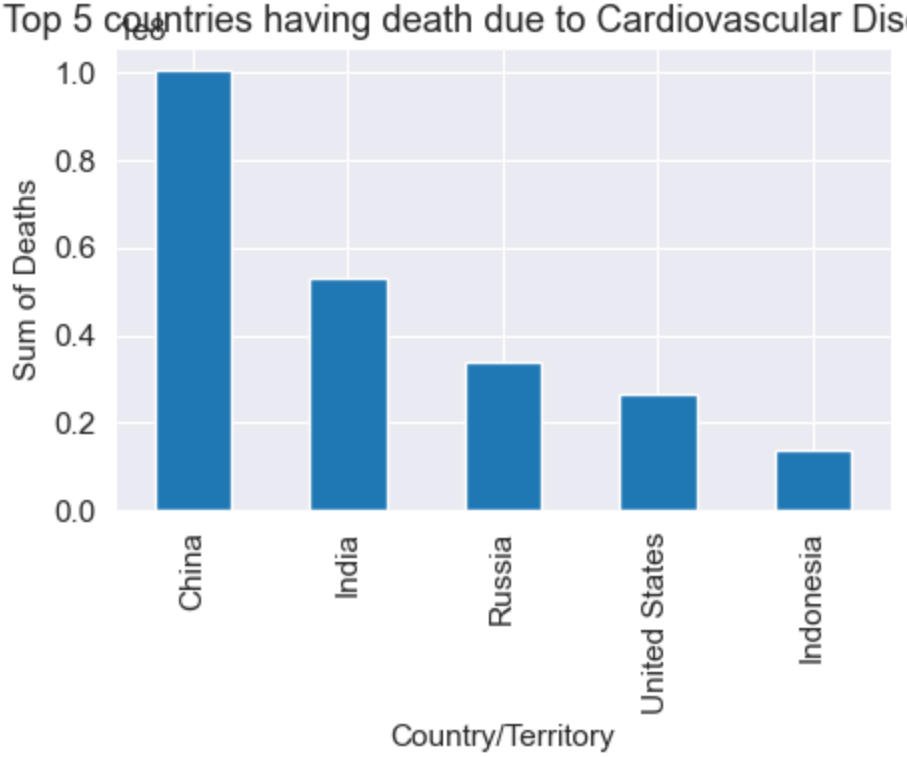

Top 5 countries having death due to Maternal Disorders
1e6
2.0
1.5
1.0
0.5
0.0
Sum of Deaths
India
Nigeria
Pakistan
Indonesia
Ethiopia
Country/Territory


Top 5 countries having death due to HIV/AIDS
1e6
5
4
3
2
1
0
Sum of Deaths
South Africa
Kenya
Tanzania
India
Nigeria
Country/Territory

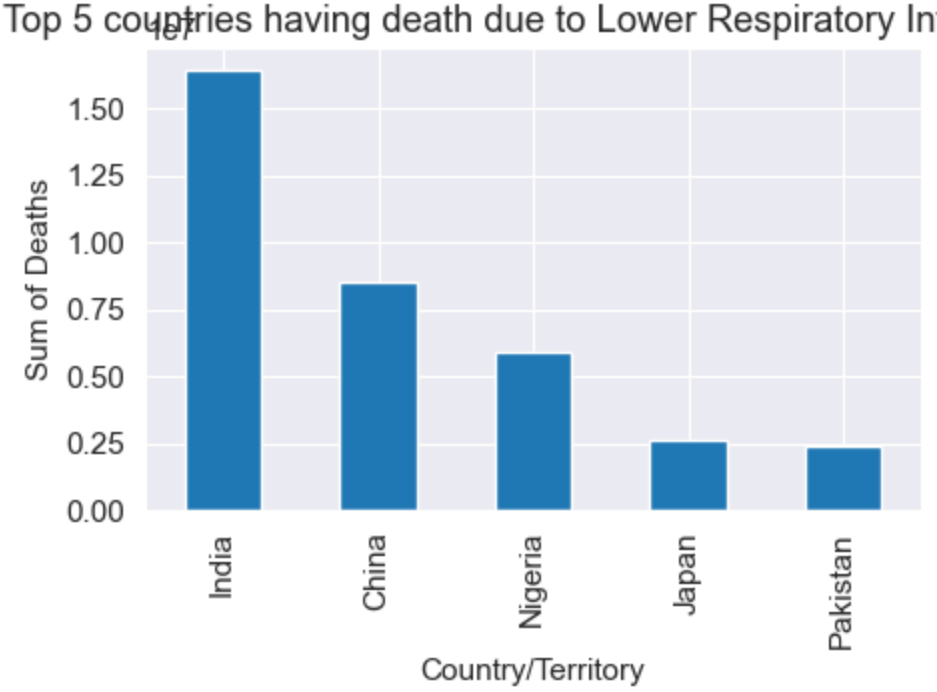

Top 5 countries having death due to Drug Use Disorders
800000
600000
400000
200000
0
Sum of Deaths
United States
China
Russia
India
Iran
Country/Territory

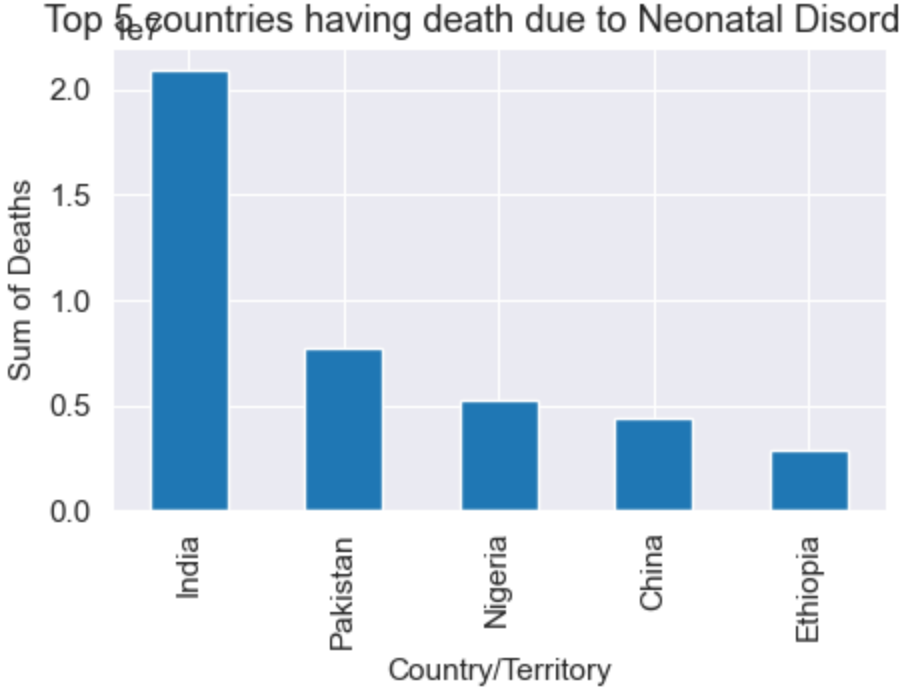

Top 5 countries having death due to Tuberculosis
1e7
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Sum of Deaths
India
Indonesia
China
Pakistan
Nigeria
Country/Territory

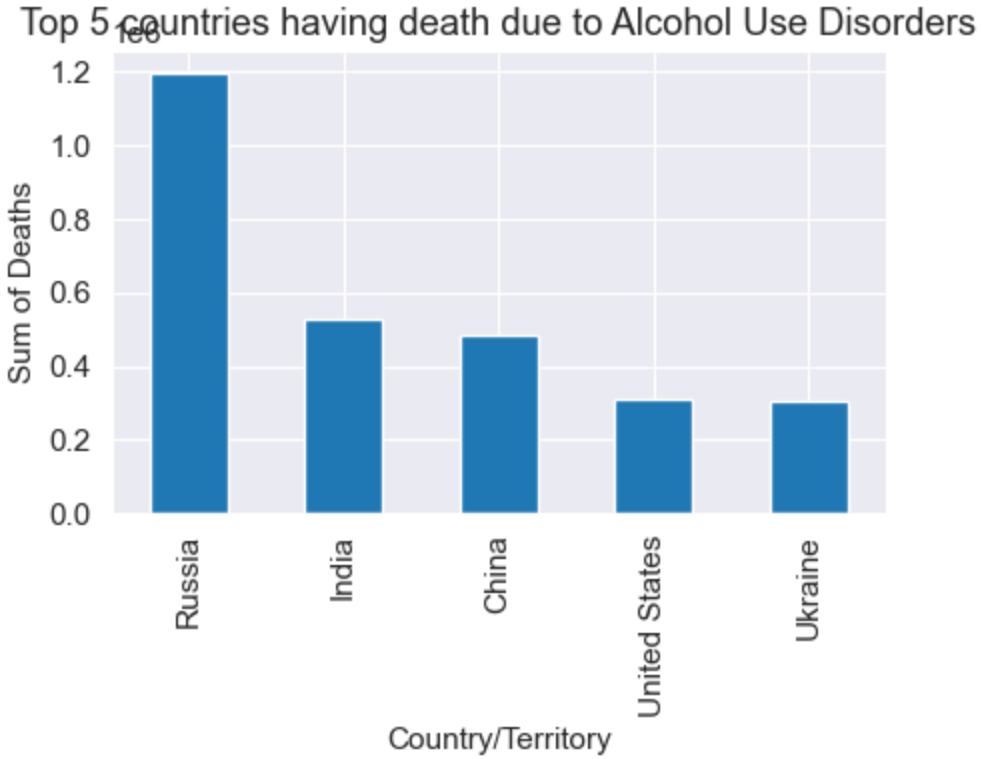
Top 5 countries having death due to Cardiovascular Diseases

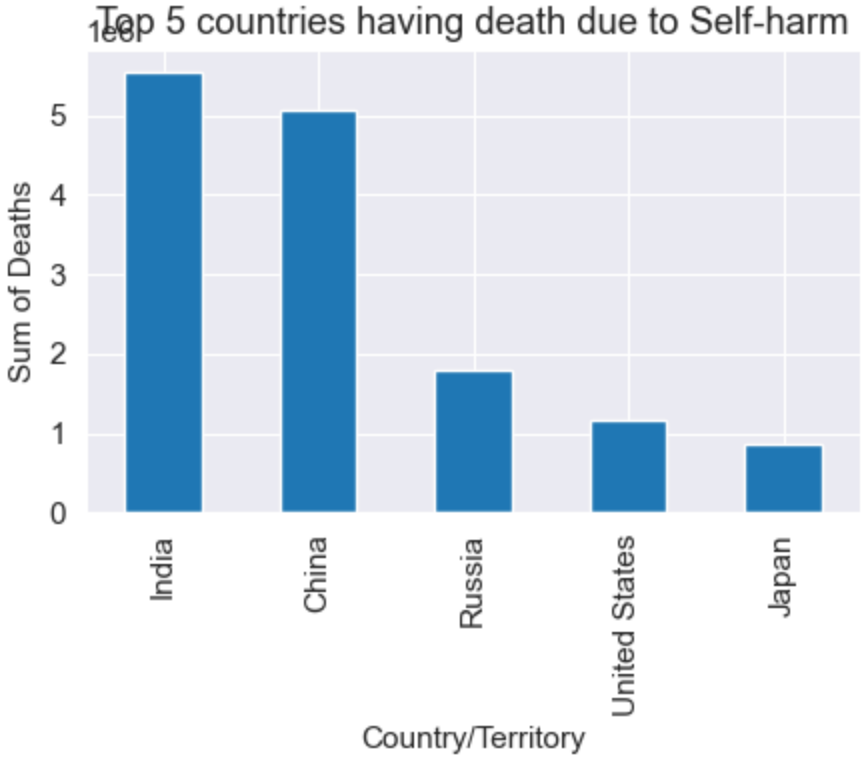


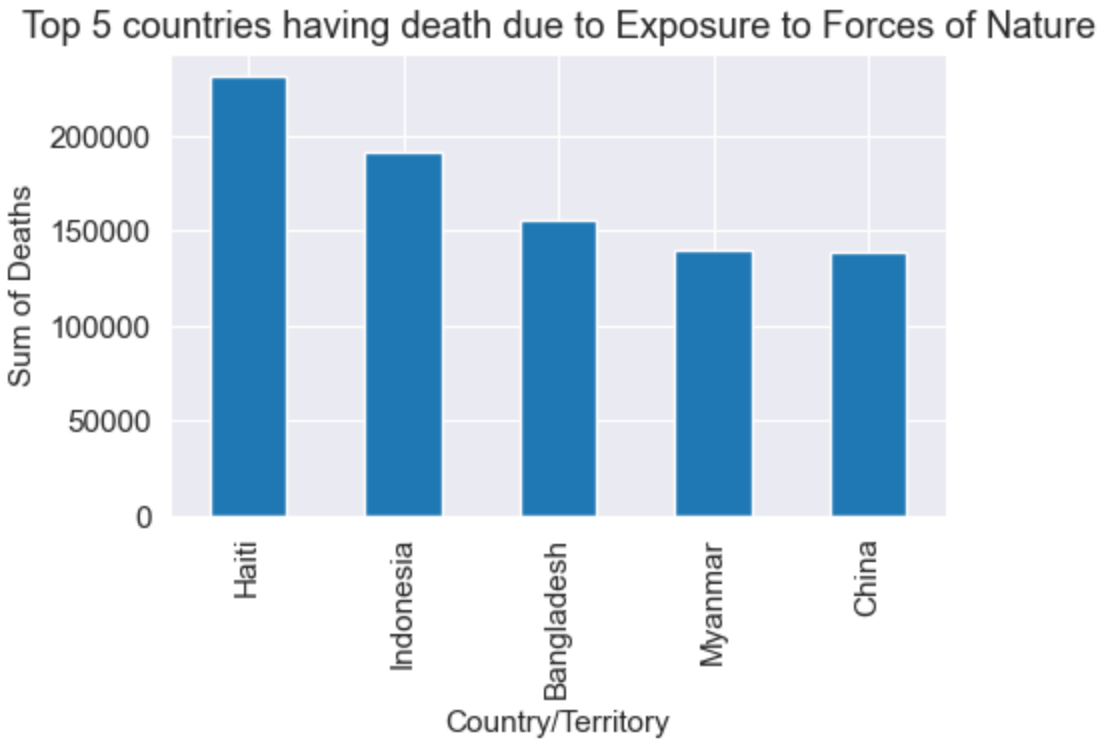
Top 5 countries having death due to Lower Respiratory Infections

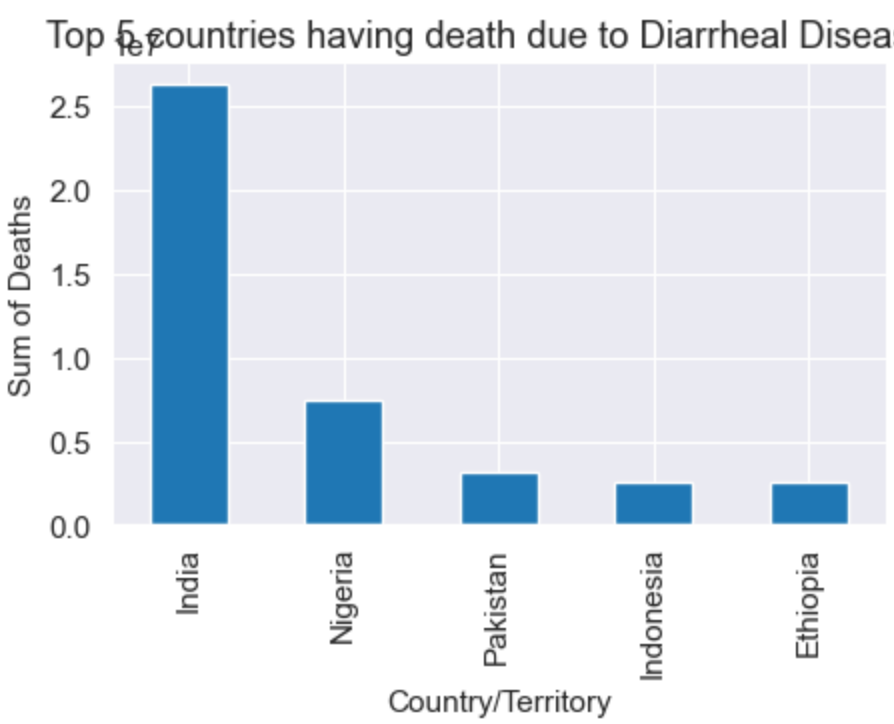


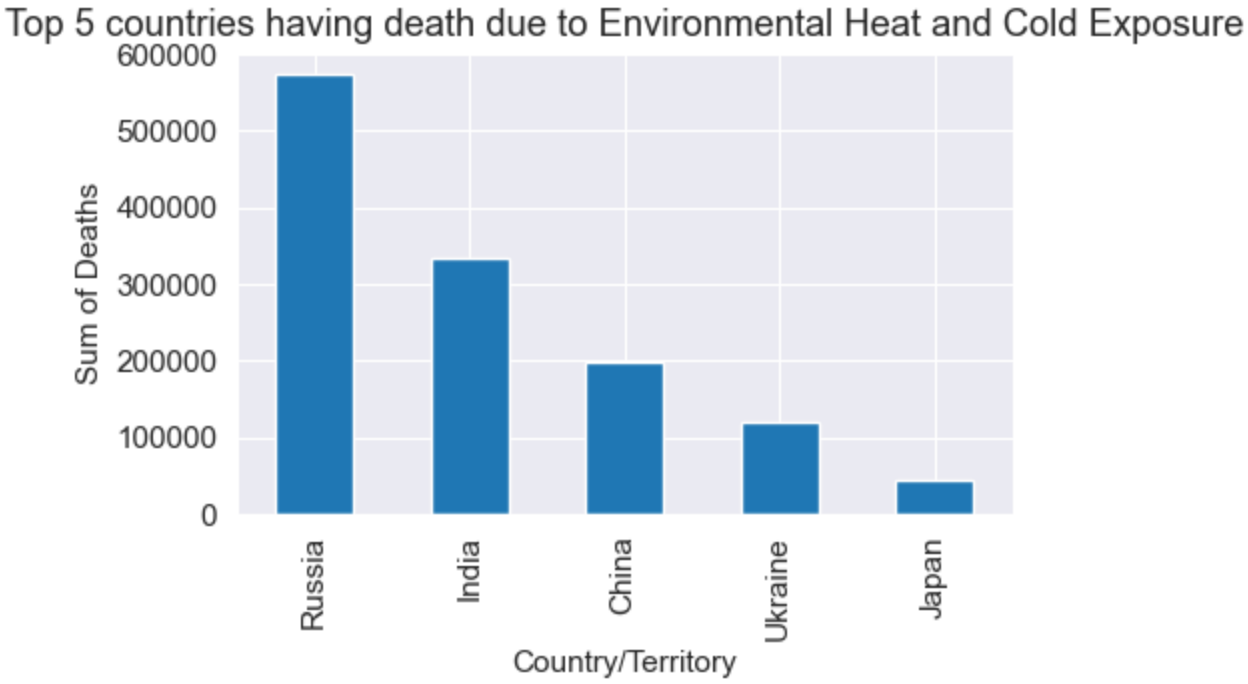

Top 5 countries having death due to Neonatal Disorders
1e7
2.0
1.5
1.0
0.5
0.0
Sum of Deaths
India
Pakistan
Nigeria
China
Ethiopia
Country/Territory

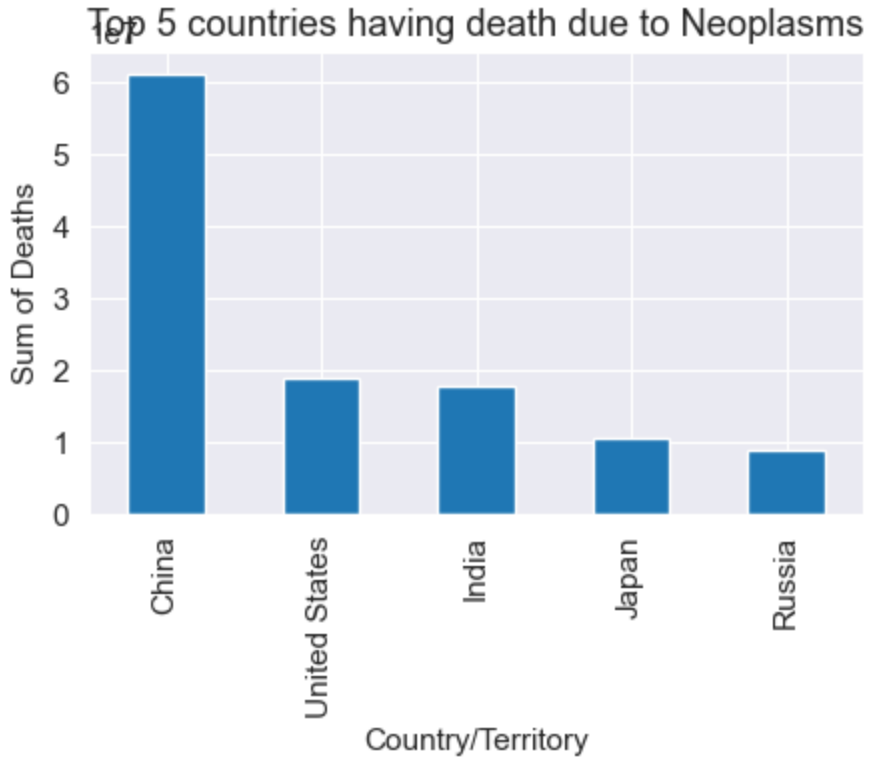

Top 5 countries having death due to Alcohol Use Disorders
1e8
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Sum of Deaths
Russia
India
China
United States
Ukraine
Country/Territory

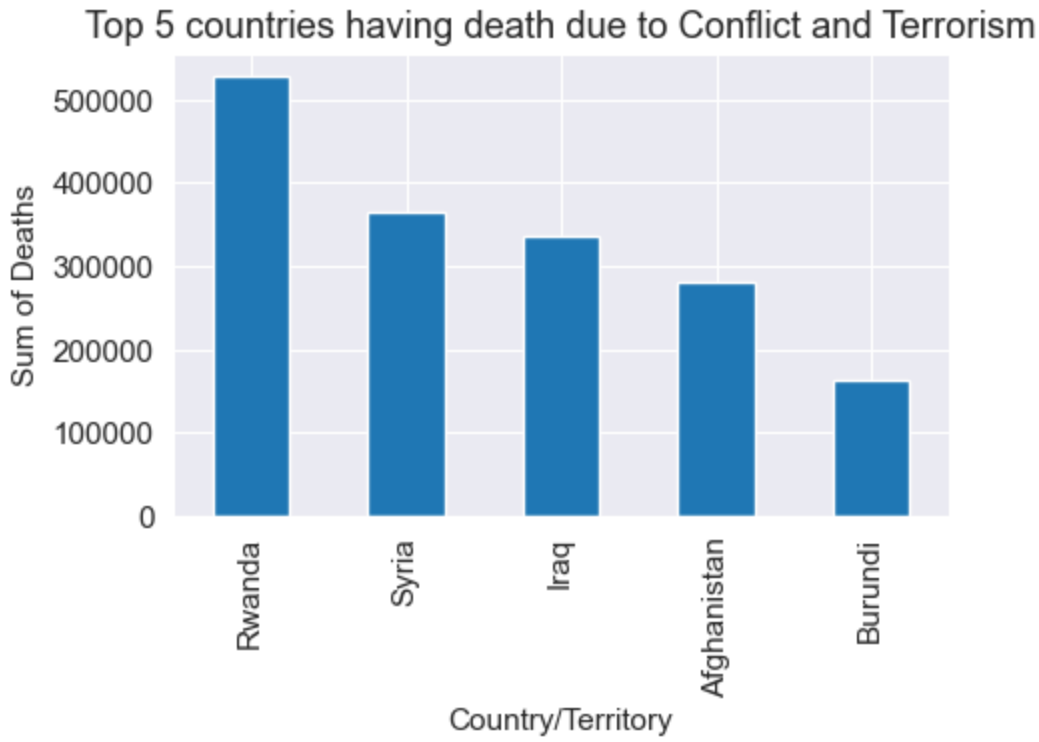

Top 5 countries having death due to Self-harm
1e6
5
4
3
2
1
0
Sum of Deaths
India
China
Russia
United States
Japan
Country/Territory


Top 5 countries having death due to Exposure to Forces of Nature
200000
150000
100000
50000
0
Sum of Deaths
Haiti
Indonesia
Bangladesh
Myanmar
China
Country/Territory

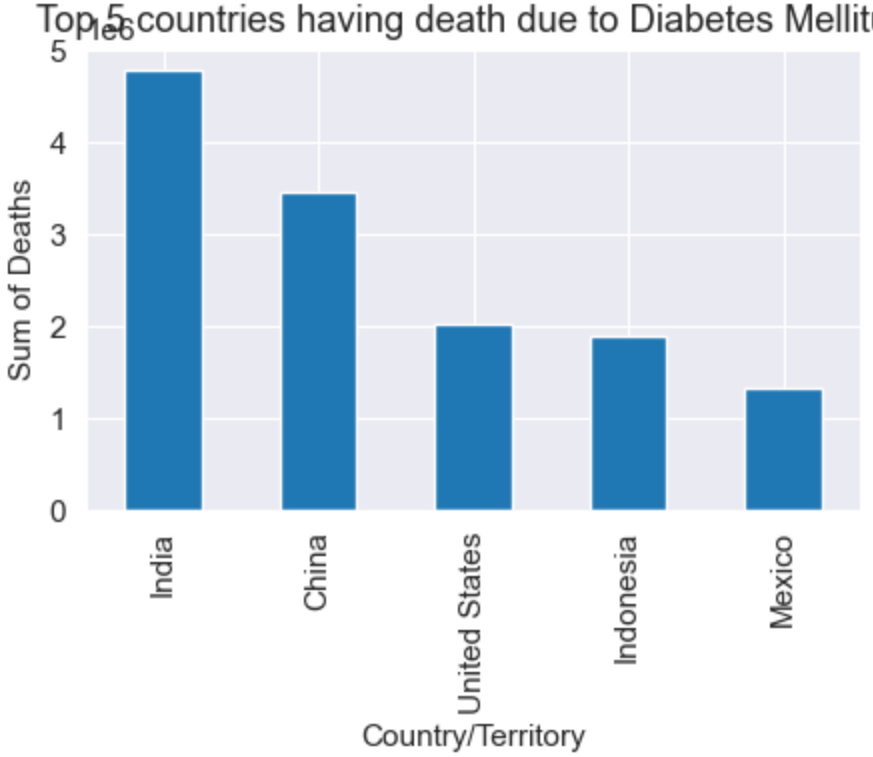

Top 5 countries having death due to Diarrheal Diseases
1e7
2.5
2.0
1.5
1.0
0.5
0.0
Sum of Deaths
India
Nigeria
Pakistan
Indonesia
Ethiopia
Country/Territory


Top 5 countries having death due to Environmental Heat and Cold Exposure
600000
500000
400000
300000
200000
100000
0
Sum of Deaths
Russia
India
China
Ukraine
Japan
Country/Territory

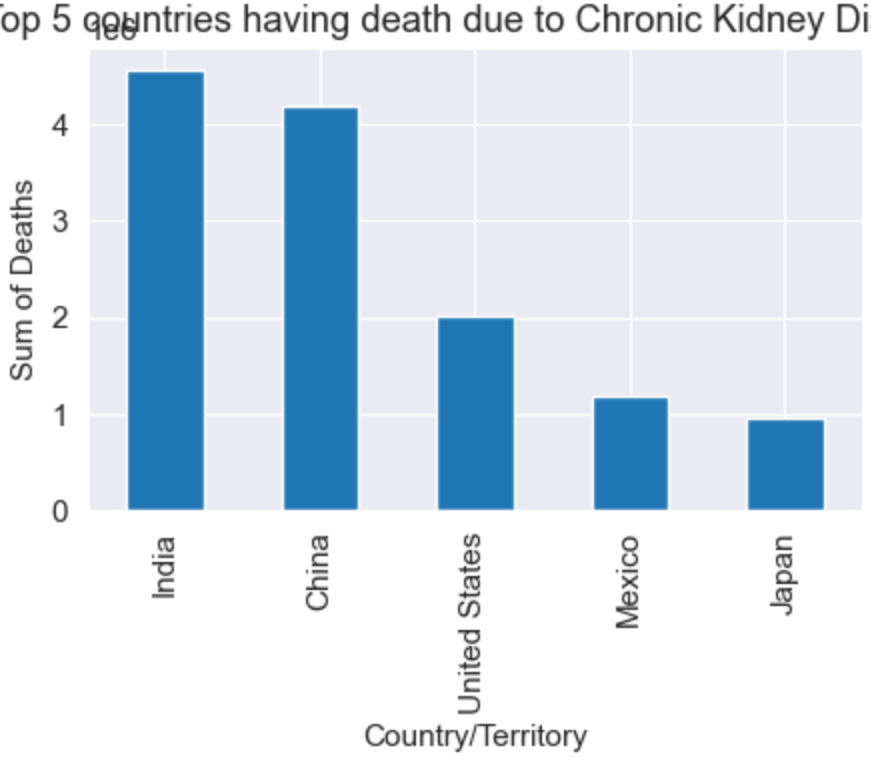

Top 5 countries having death due to Neoplasms
1e7
6
5
4
3
2
1
0
Sum of Deaths
China
United States
India
Japan
Russia
Country/Territory

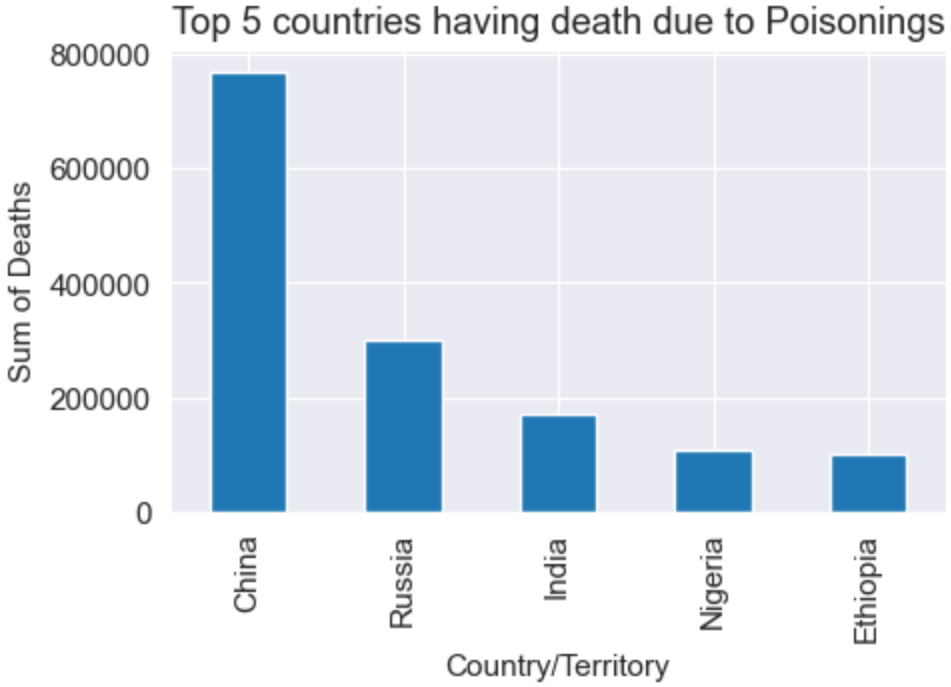

Top 5 countries having death due to Conflict and Terrorism
500000
400000
300000
200000
100000
0
Sum of Deaths
Rwanda
Syria
Iraq
Afghanistan
Burundi
Country/Territory

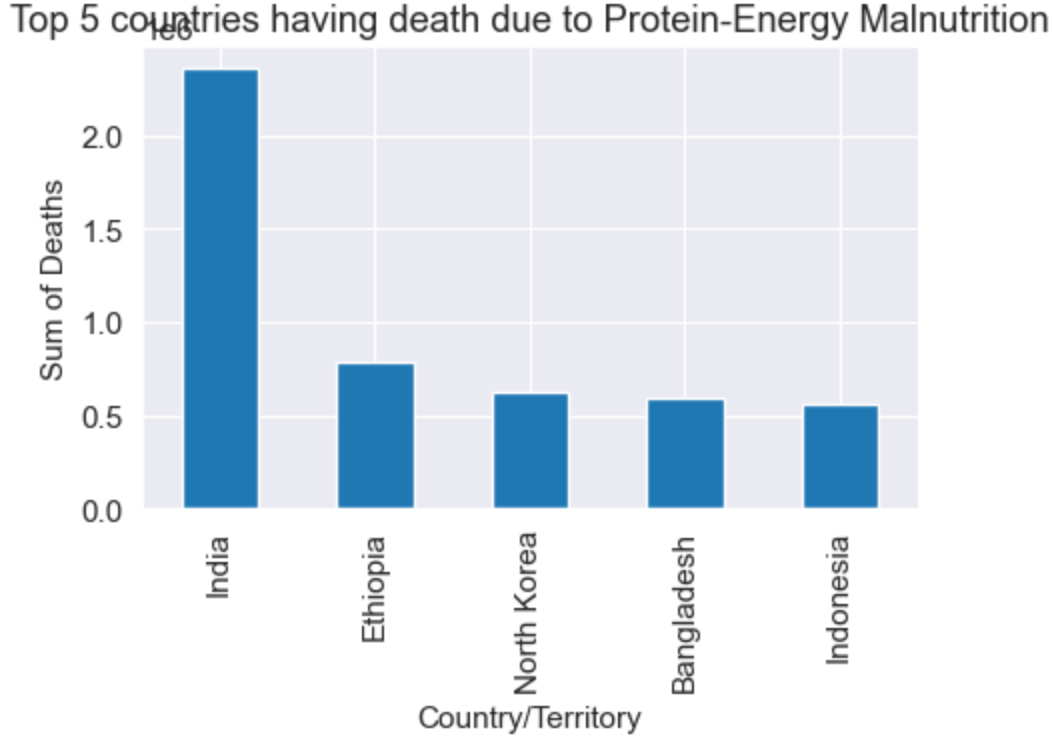
Top 5 countries having death due to Diabetes Mellitus

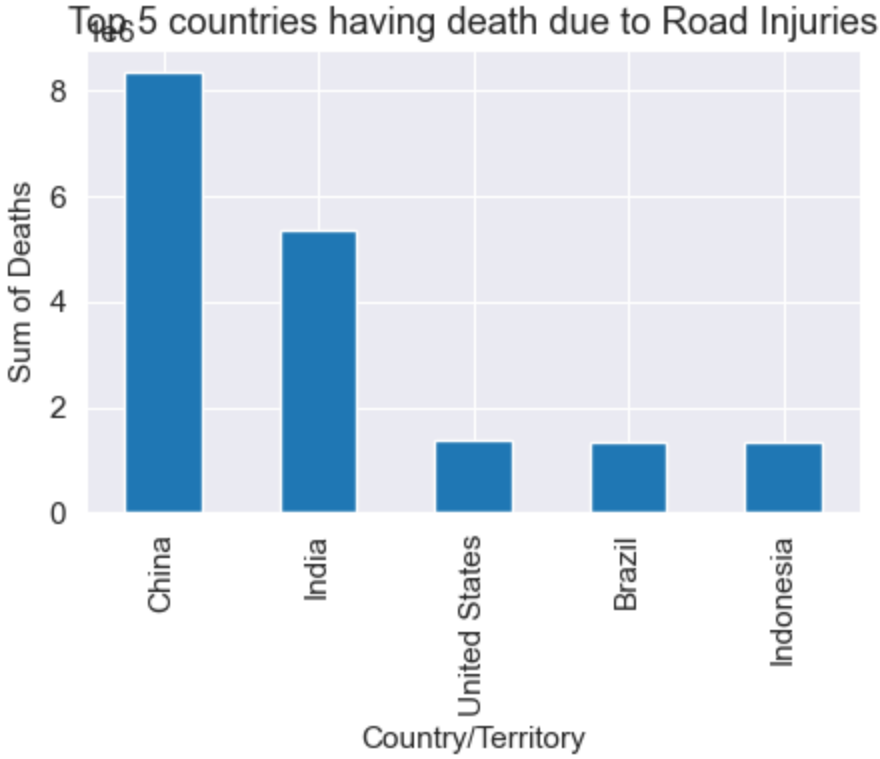


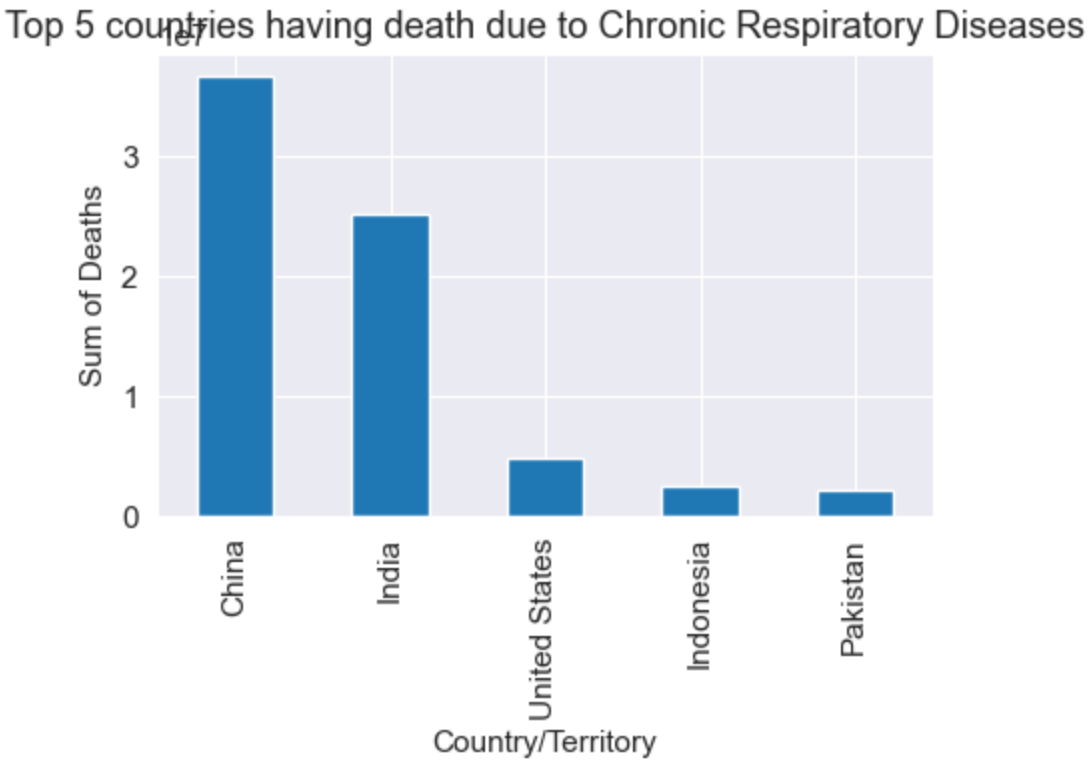
Top 5 countries having death due to Chronic Kidney Disease

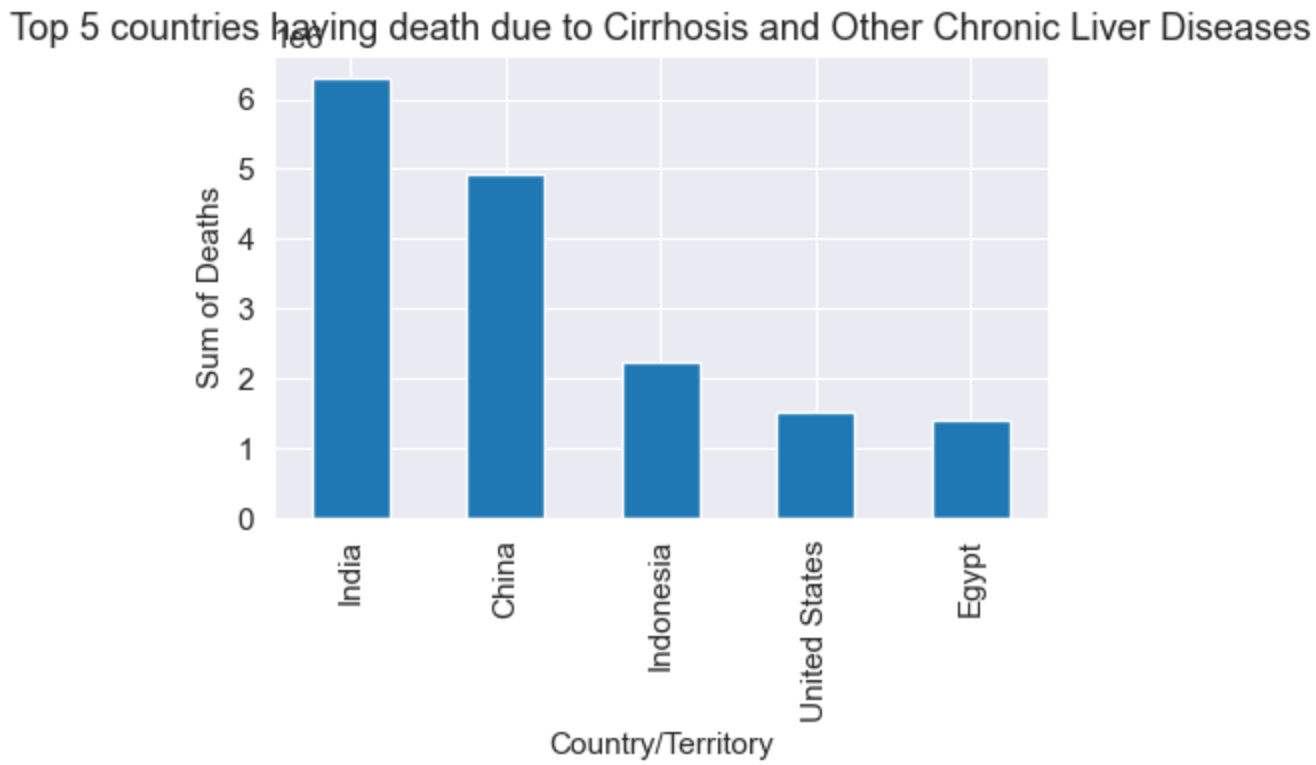


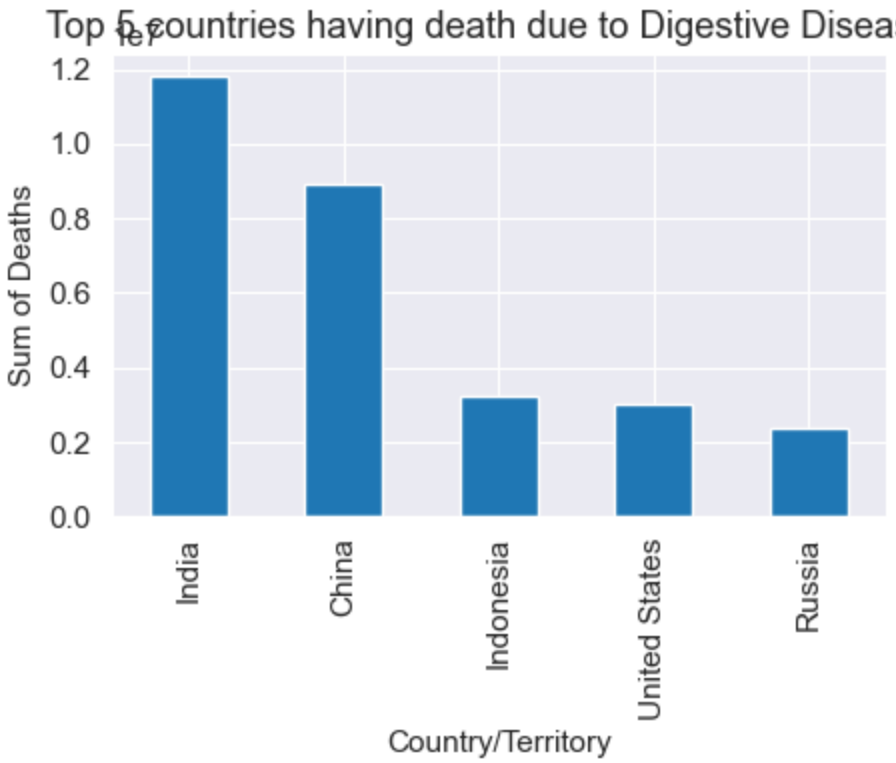

Top 5 countries having death due to Poisonings
800000
600000
400000
200000
0
Sum of Deaths
China
Russia
India
Nigeria
Ethiopia
Country/Territory

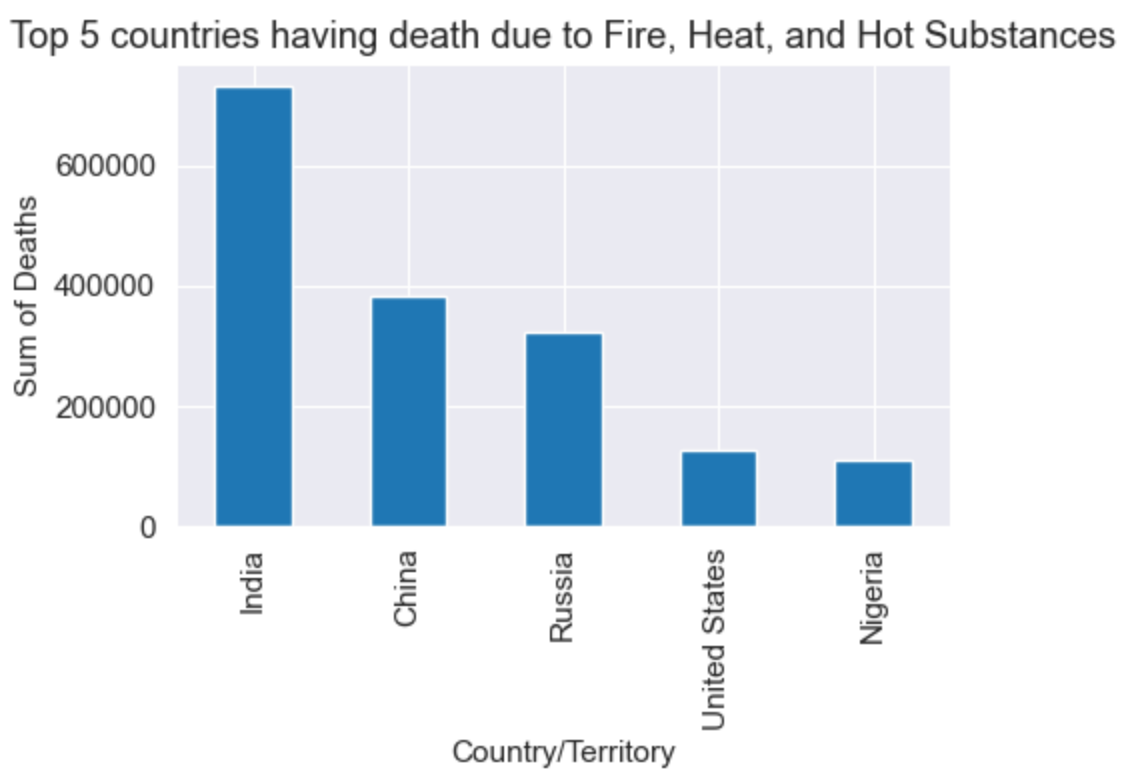

Top 5 countries having death due to Protein-Energy Malnutrition
1e6
2.0
1.5
1.0
0.5
0.0
Sum of Deaths
India
Ethiopia
North Korea
Bangladesh
Indonesia
Country/Territory

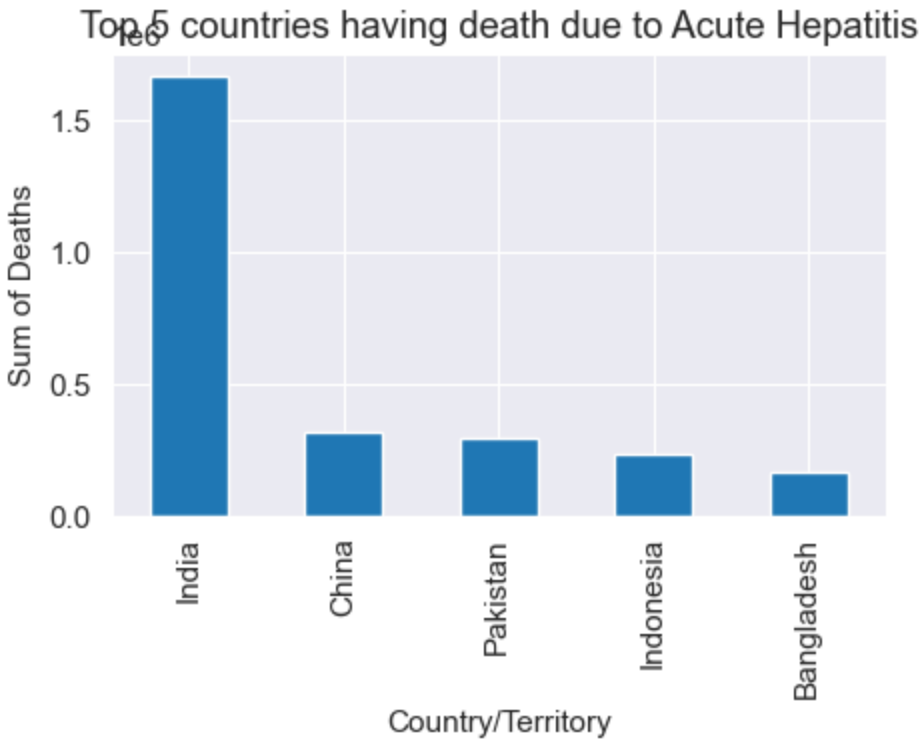

Top 5 countries having death due to Road Injuries
1e6
8
6
4
2
0
Sum of Deaths
China
India
United States
Brazil
Indonesia
Country/Territory


Top 5 countries having death due to Chronic Respiratory Diseases
1e7
3
2
1
0
Sum of Deaths
China
India
United States
Indonesia
Pakistan
Country/Territory


Top 5 countries having death due to Cirrhosis and Other Chronic Liver Diseases
1e6
6
5
4
3
2
1
0
Sum of Deaths
India
China
Indonesia
United States
Egypt
Country/Territory


Top 5 countries having death due to Digestive Diseases
1e7
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Sum of Deaths
India
China
Indonesia
United States
Russia
Country/Territory





Top countries in different death disease/enviroment from above graphs

India---- Meningits, Nutritional Deficiencies, Maternal Disorders, Tuberculosis, Lower Respiratory Infections, Self-harm

Neonatal Disorders,Diarrheal Diseases, Diabetes Mellitus,Choronic kidney,Protein-energy malnutrition,Digestive,Acute Hepatitis

Cirrhosis & other chronic liver, Fire/heat & hot substances

China---- Alzheimer, Parkinson, Drowning, Cardiovascular Diseases, Neoplasms, Poisonings, Road injuries, Chronic respiratory

Nigeria---- Malaria

Brazil---- Interpersonal Violence

South Africa---- HIV/AIDS

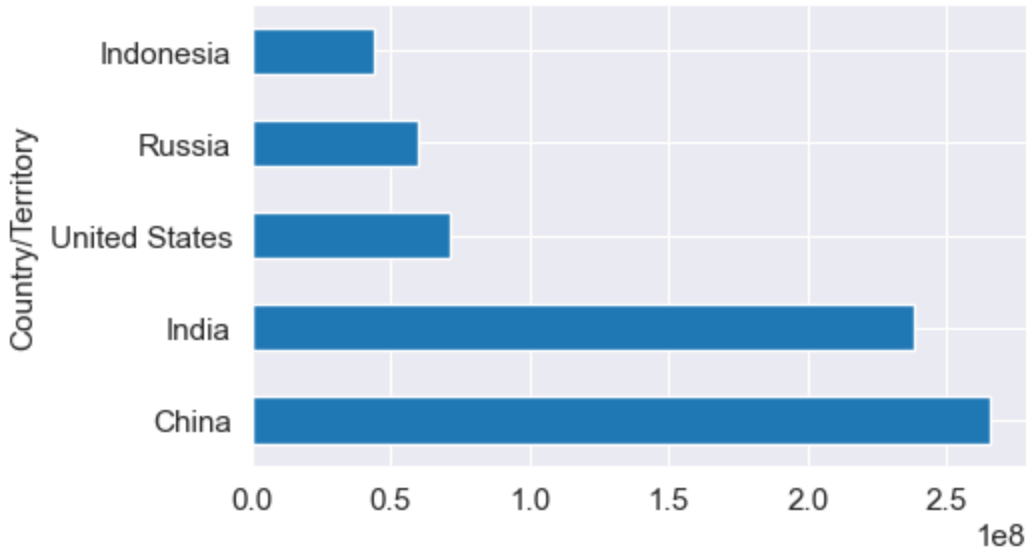
USA---- Drug use Disorders

Russia---- Alcohol use Disorders, Environmental Heat&cold exposure

Haiti---- Exposure to forces of nature

Rwanda---- Conflict & Terrorism

In [16]:
```
df["Total Deaths"]=df.sum(axis=1)
df.groupby('Country/Territory')["Total Deaths"].sum().sort_values(ascending=False).head(
plt.show()
```



In [17]:
```
summ=['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]))
```

```
df2=pd.DataFrame({
```

In [18]:
```
    'Name of diseases':summ,
    'Total diseases':sum_of
})
df2
```

Out[18]:

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

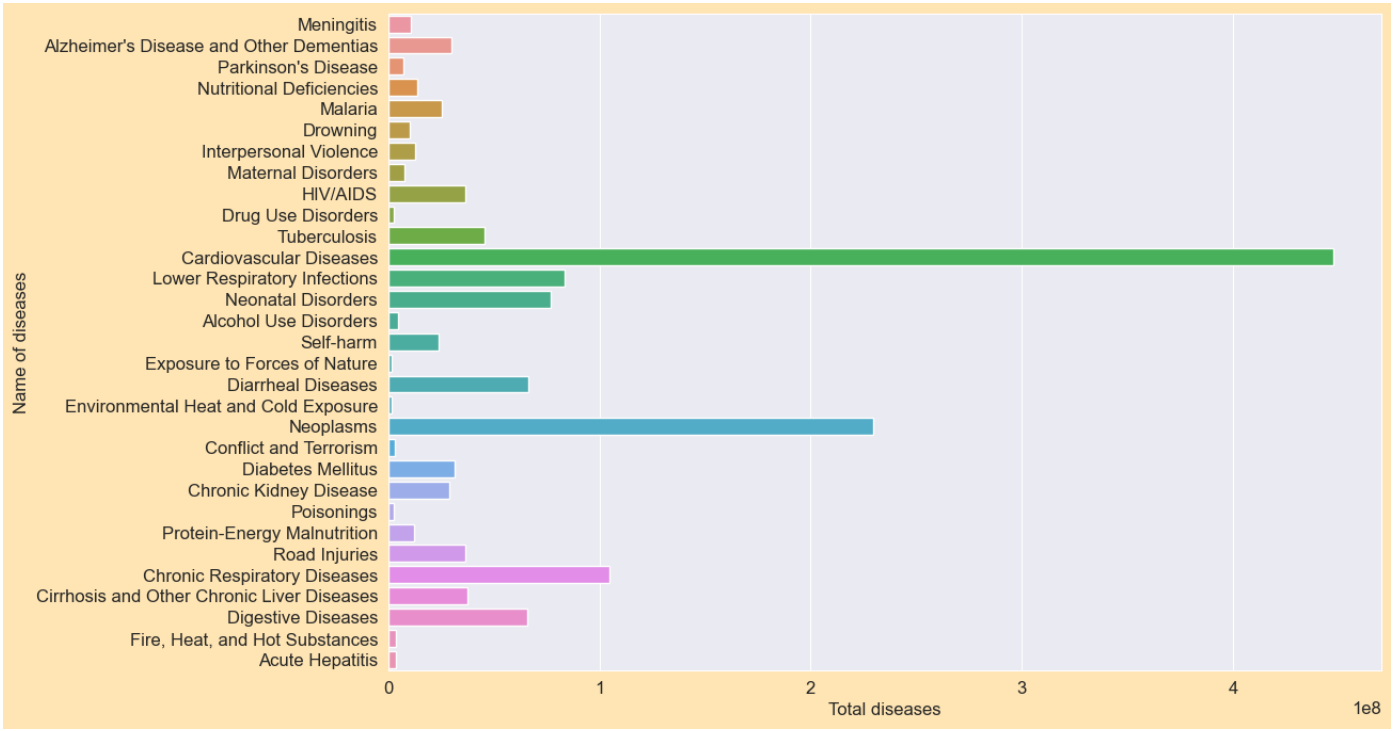
In [25]:
```
df2["Total diseases"]
```

Out[25]:
```
0       10524572
```

```
1       29768839
2        7179795
3       13792032
4       25342676
5       10301999
6       12752839
7        7727046
8       36364419
9        2656121
10      45850603
11     447741982
12      83770038
13      76860729
14       4819018
15      23713931
16       1490132
17      66235508
18       1788851
19     229758538
20       3294053
21      31448872
22      28911692
23       2601082
24      12031885
25      36296469
26     104605334
27      37479321
28      65638635
29       3602914
30       3784791
Name: Total diseases, dtype: int64
```

In [32]:
```
plt.rcParams['font.size'] = 12
plt.rcParams ['figure.figsize'] = (12, 8)
plt.rcParams ['figure.facecolor'] = '#FFE5B4'
sns.barplot(y=df2["Name of diseases"],x=df2["Total diseases"])
```

Out[32]:
```
<AxesSubplot:xlabel='Total diseases', ylabel='Name of diseases'>
```



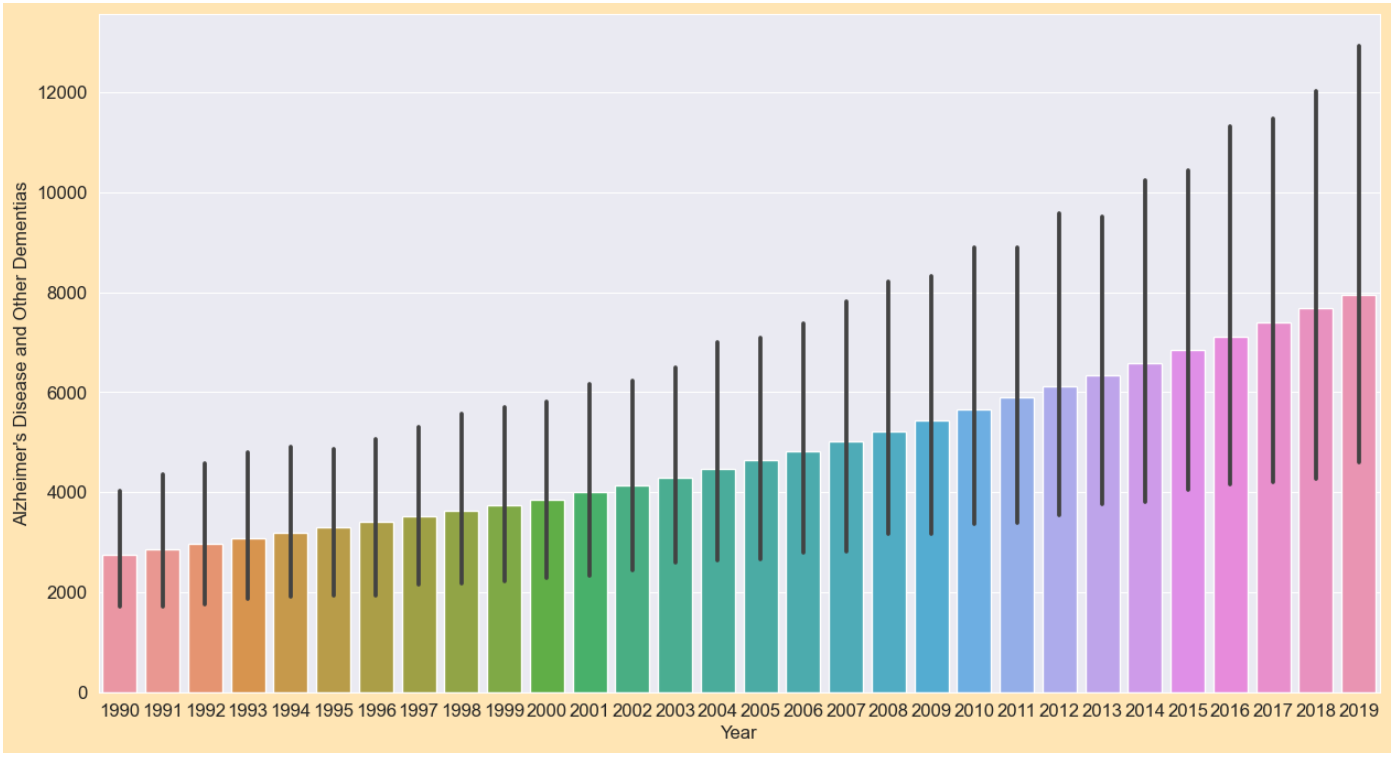
As we observe in the diseases of Cardiovascular there are maximum deth 447741982.

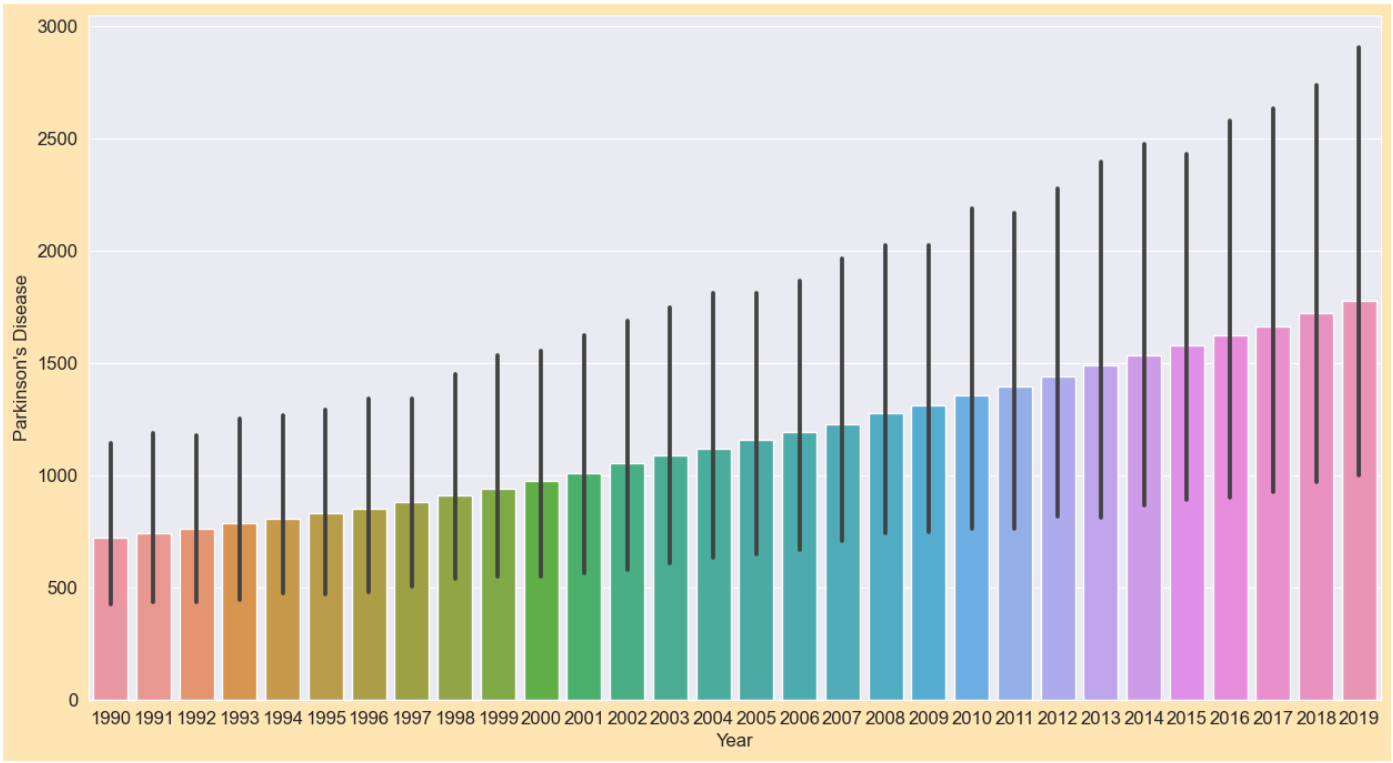
As we observe in the diseases of Exposure to Forces of Nature there are maximum deth 1490132.

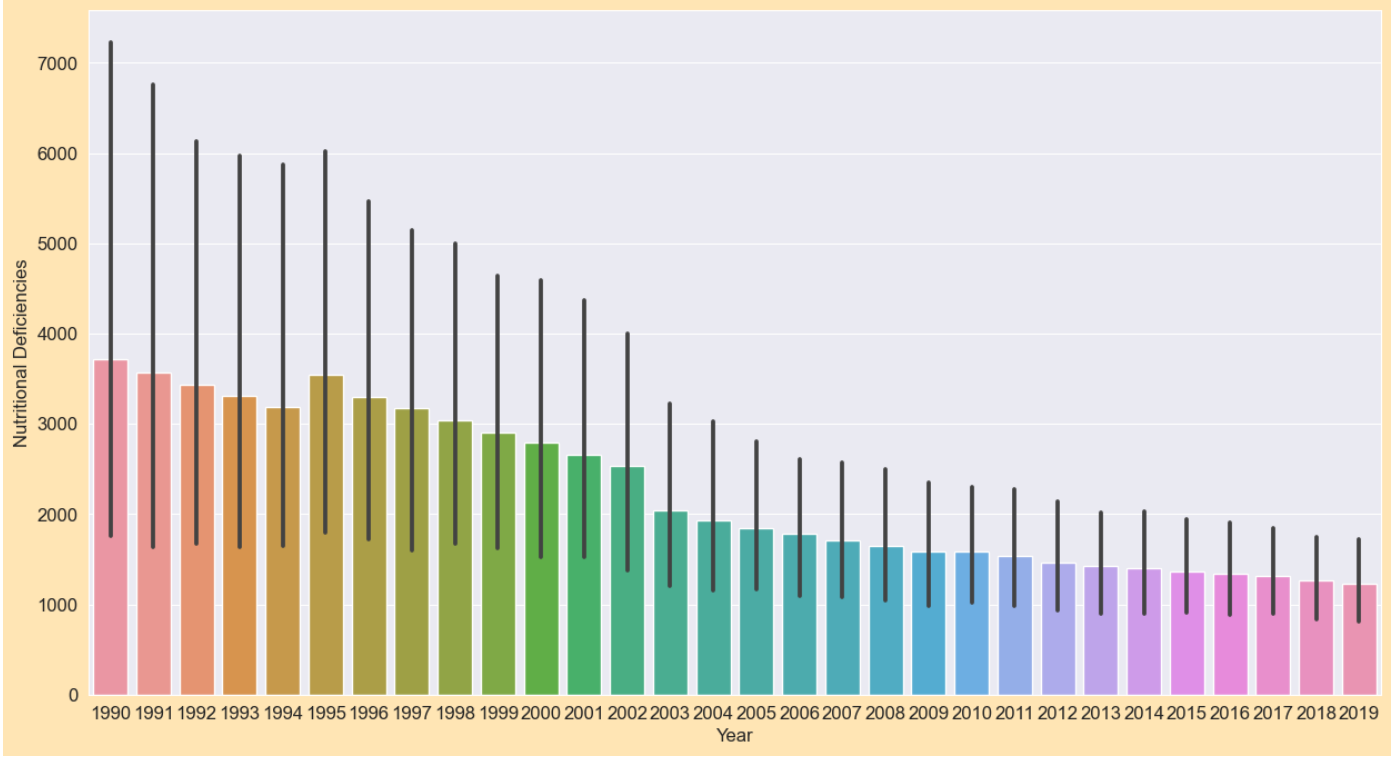
As we observe there are total 1480402256 persons are died.

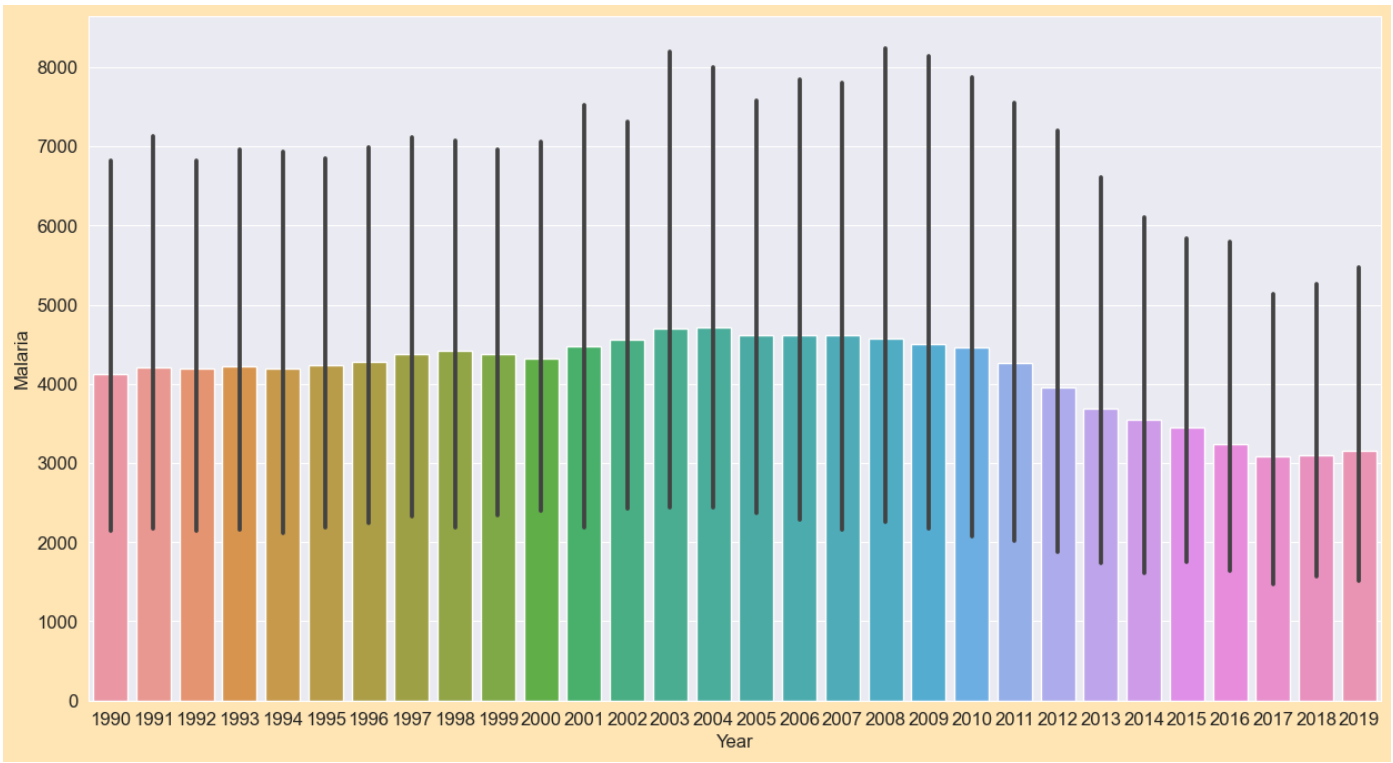
Bivariate Analysis

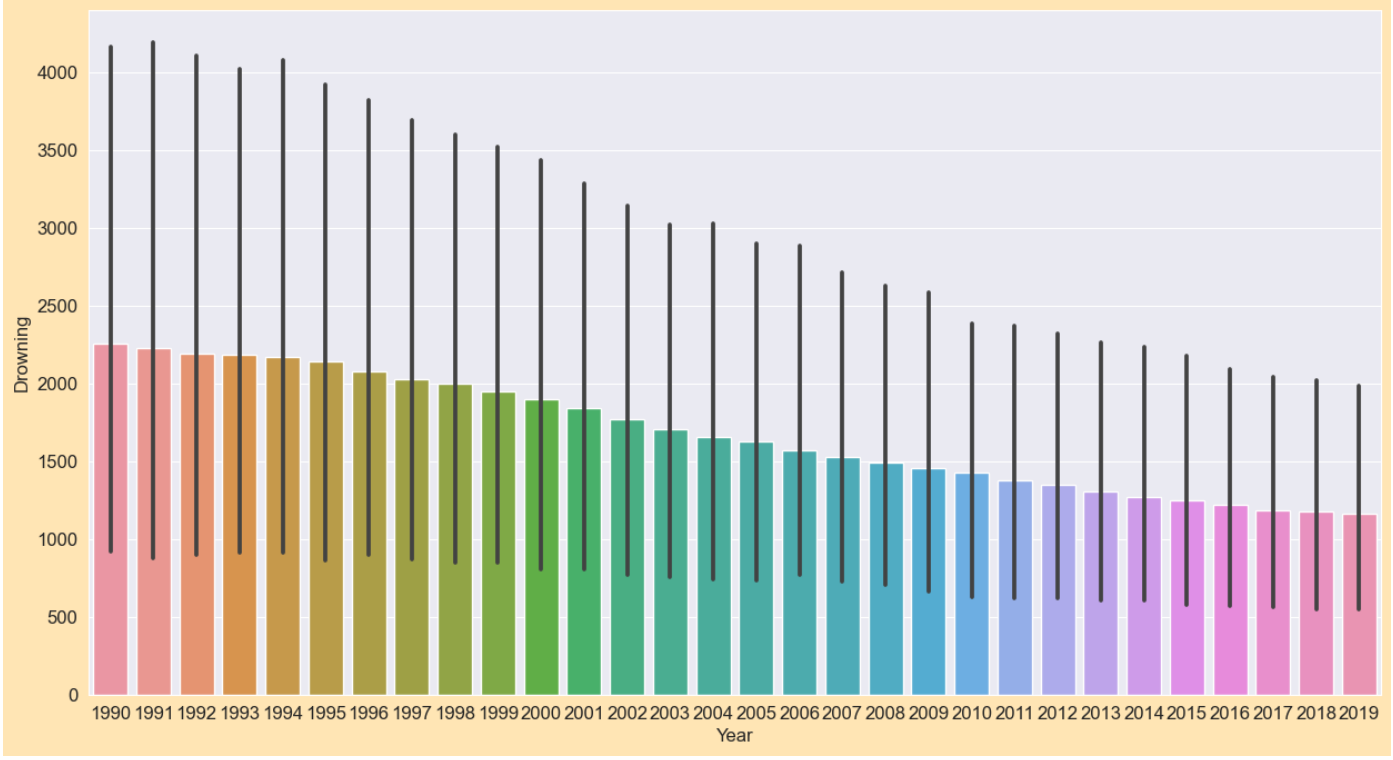
In [34]:
```
for i in summ[1:]:
    plt.figure(figsize=(15,8))
    sns.barplot(x='Year',y=i,data=df)
    plt.show()
```

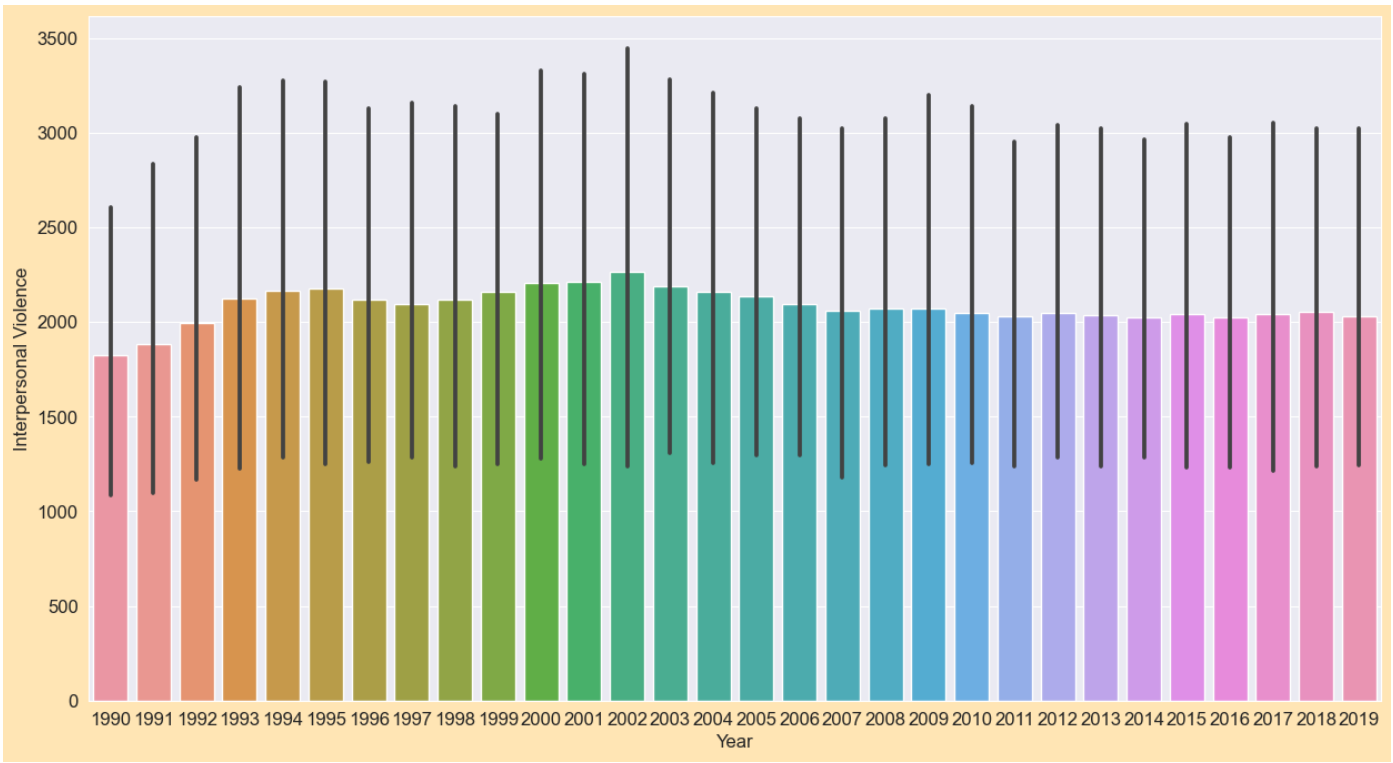


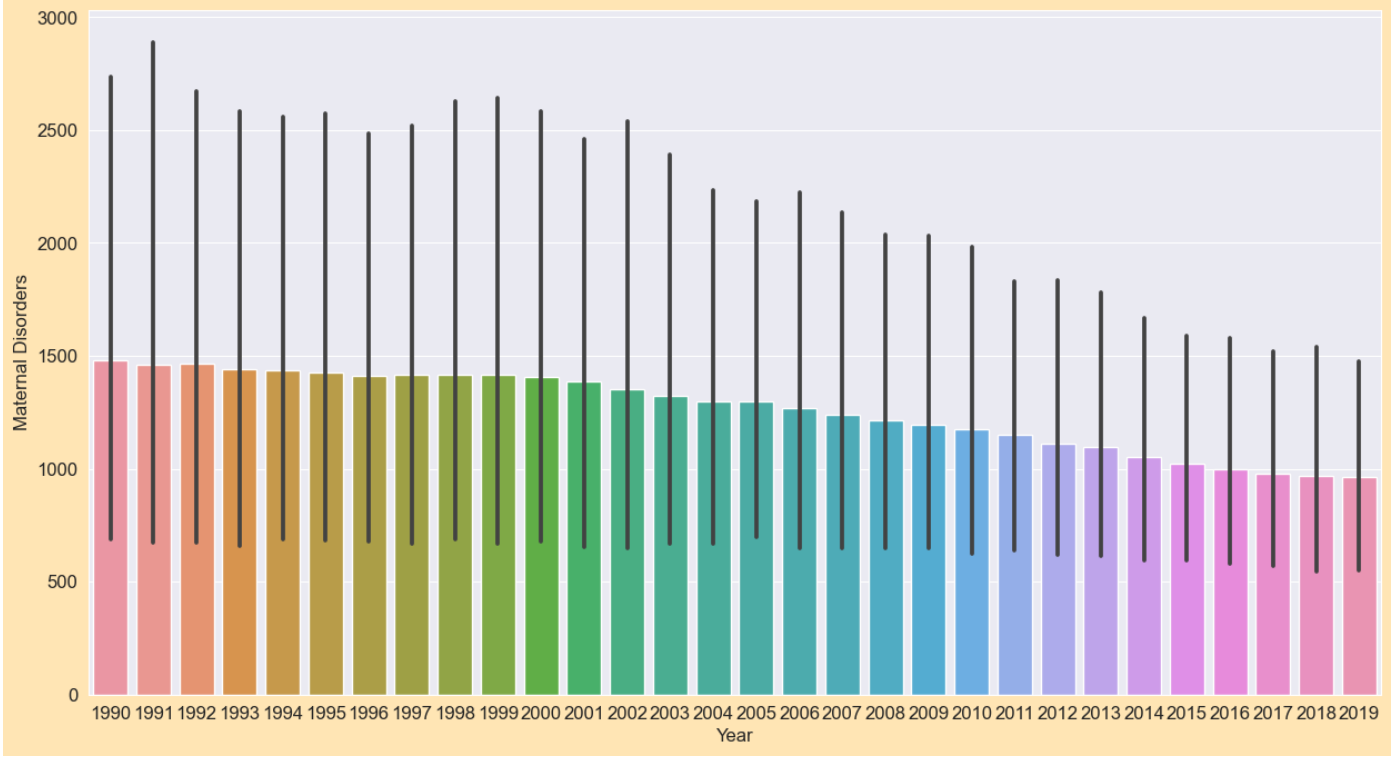


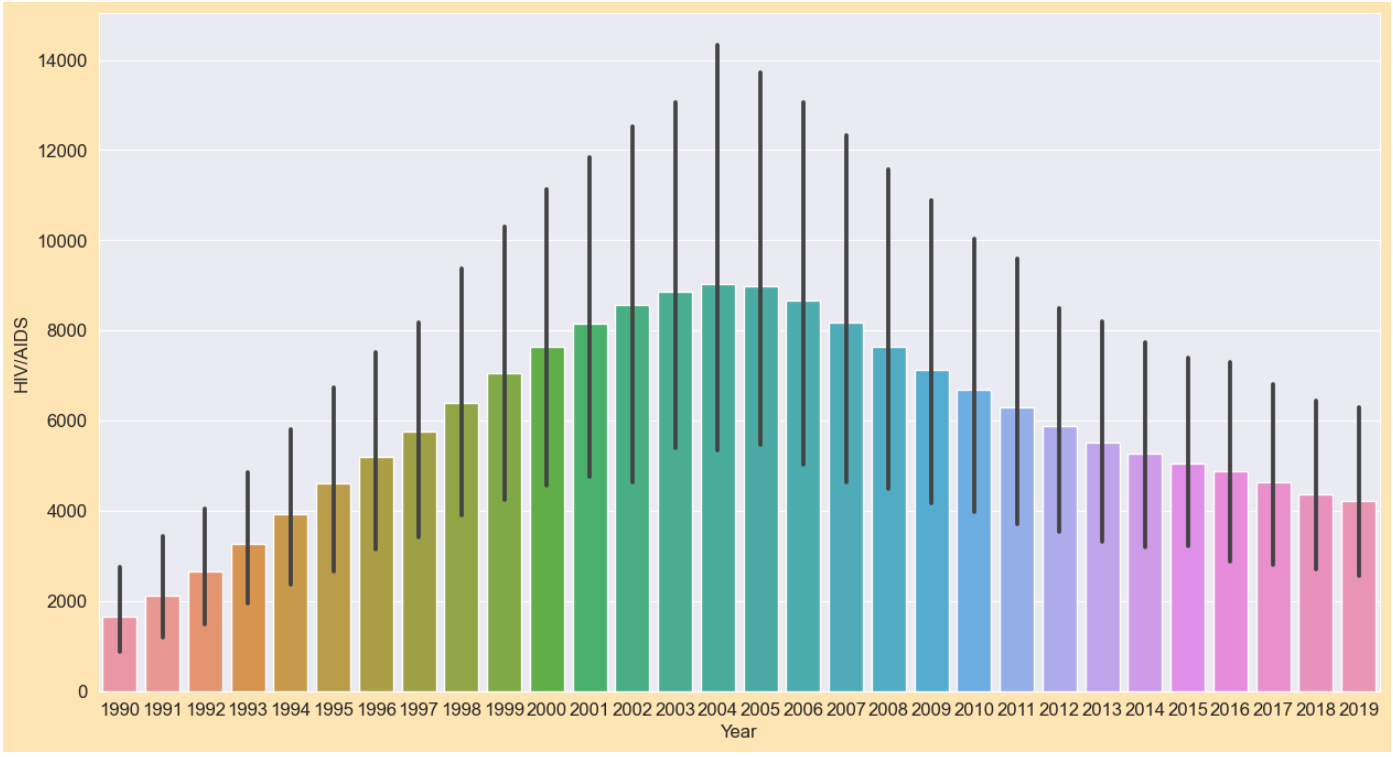

Nutritional Deficiencies
0
1000
2000
3000
4000
5000
6000
7000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

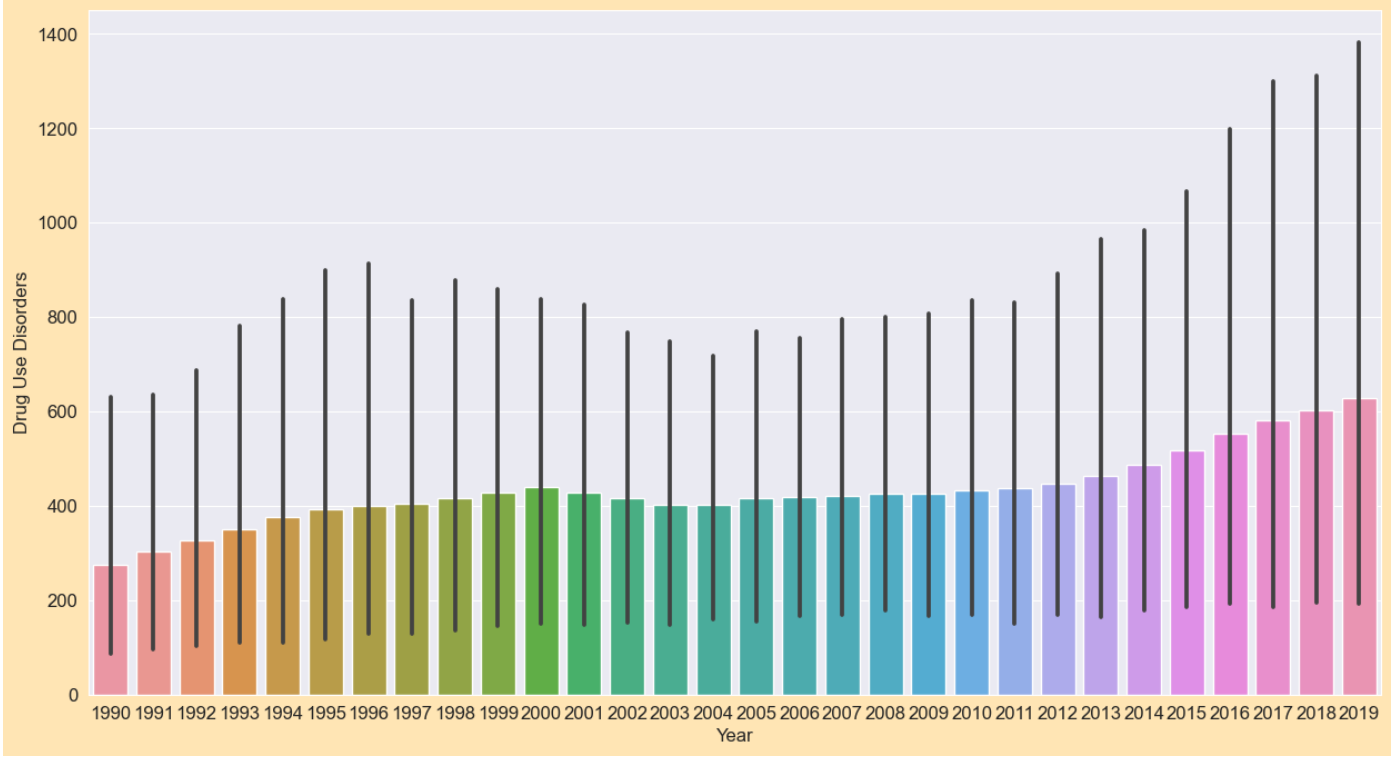

Malaria
0
1000
2000
3000
4000
5000
6000
7000
8000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

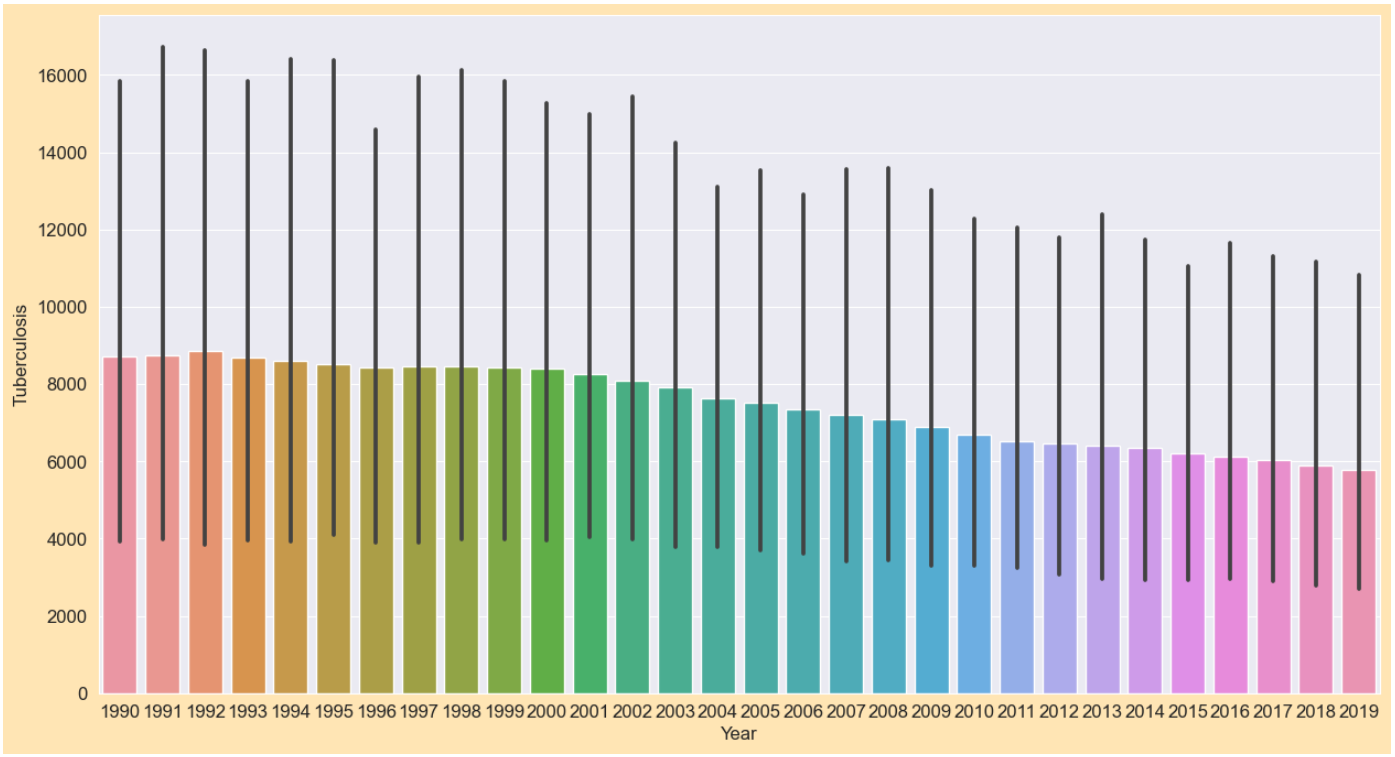

Drowning
0
500
1000
1500
2000
2500
3000
3500
4000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

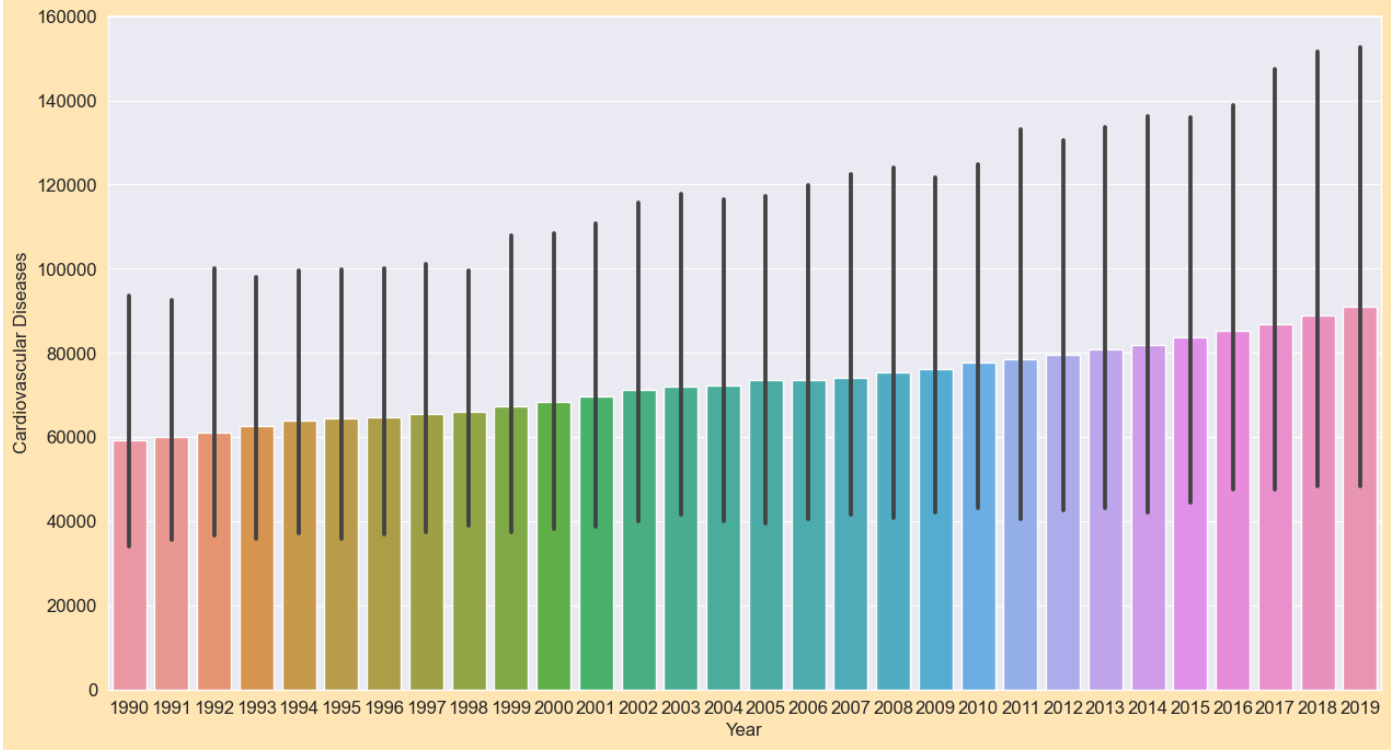

Interpersonal Violence
0
500
1000
1500
2000
2500
3000
3500
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

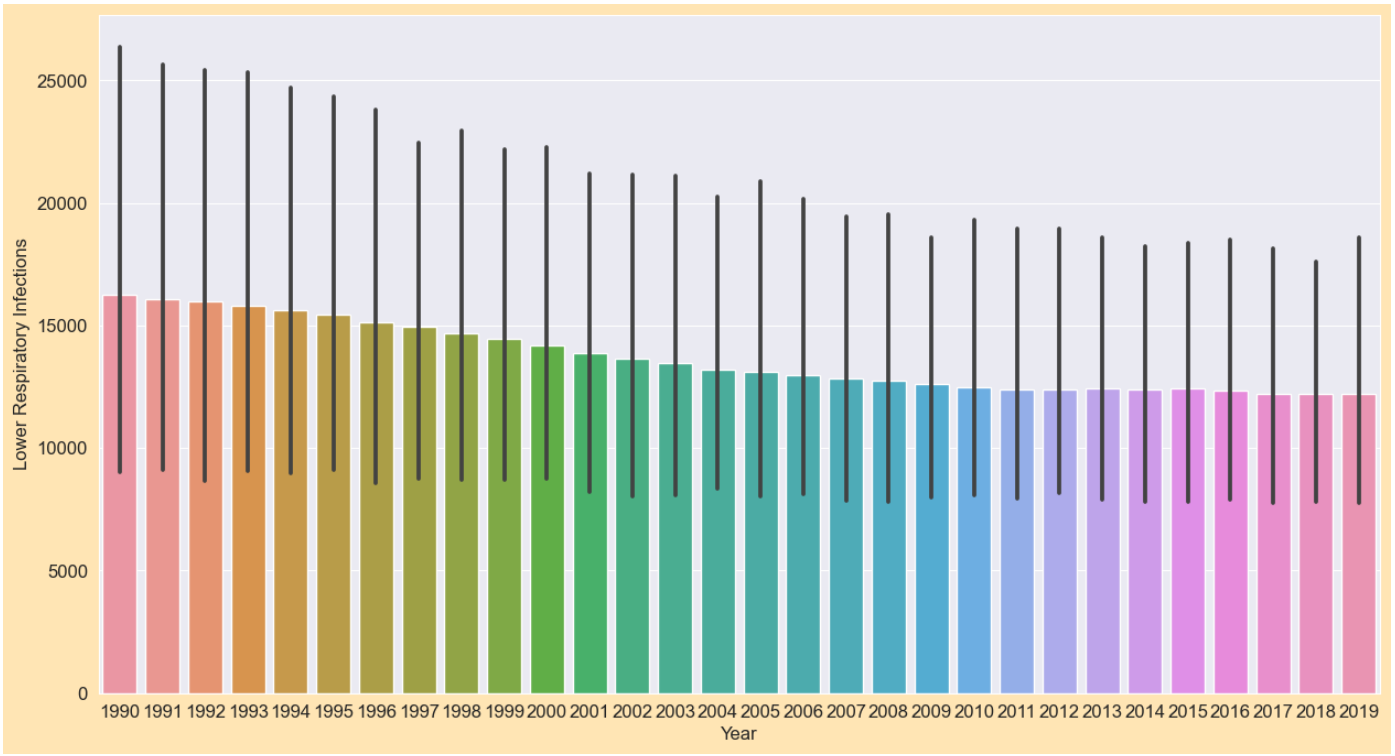

Maternal Disorders
3000
2500
2000
1500
1000
500
0
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

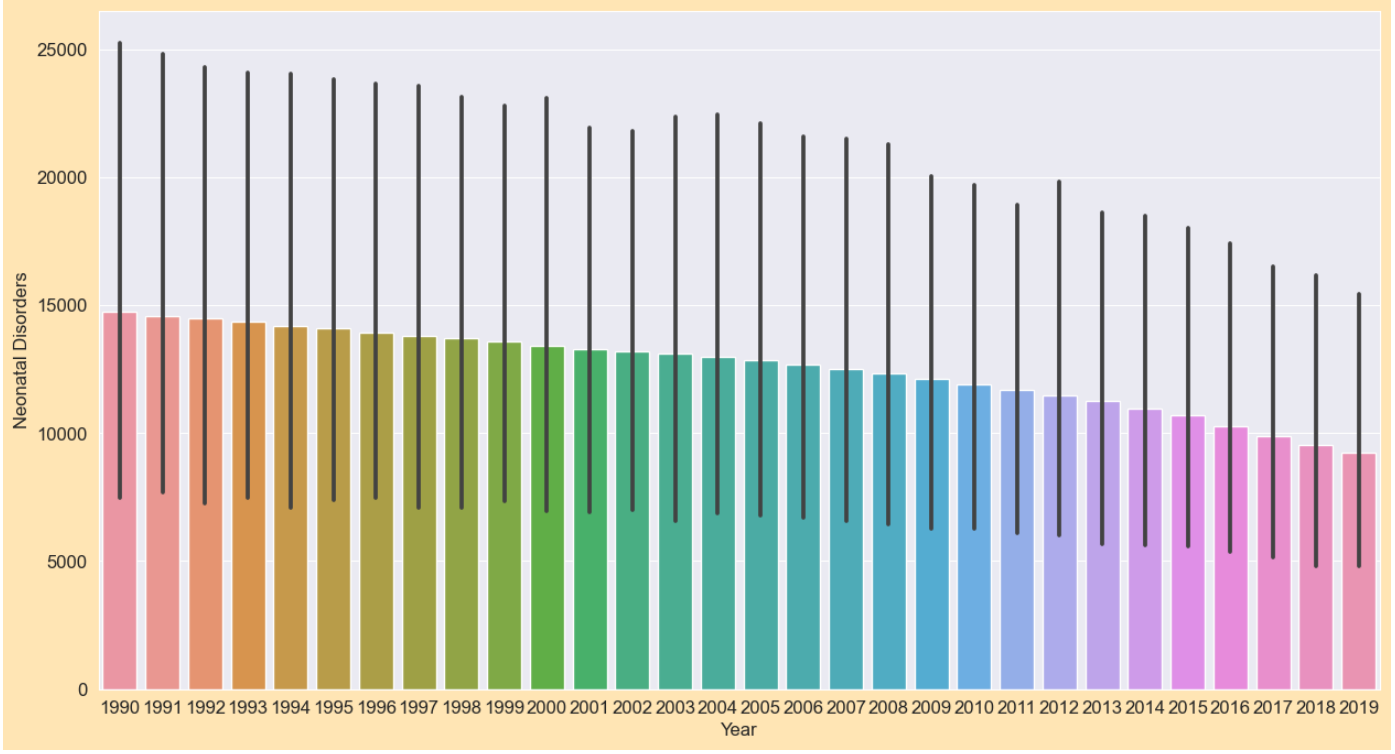

HIV/AIDS
14000
12000
10000
8000
6000
4000
2000
0
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

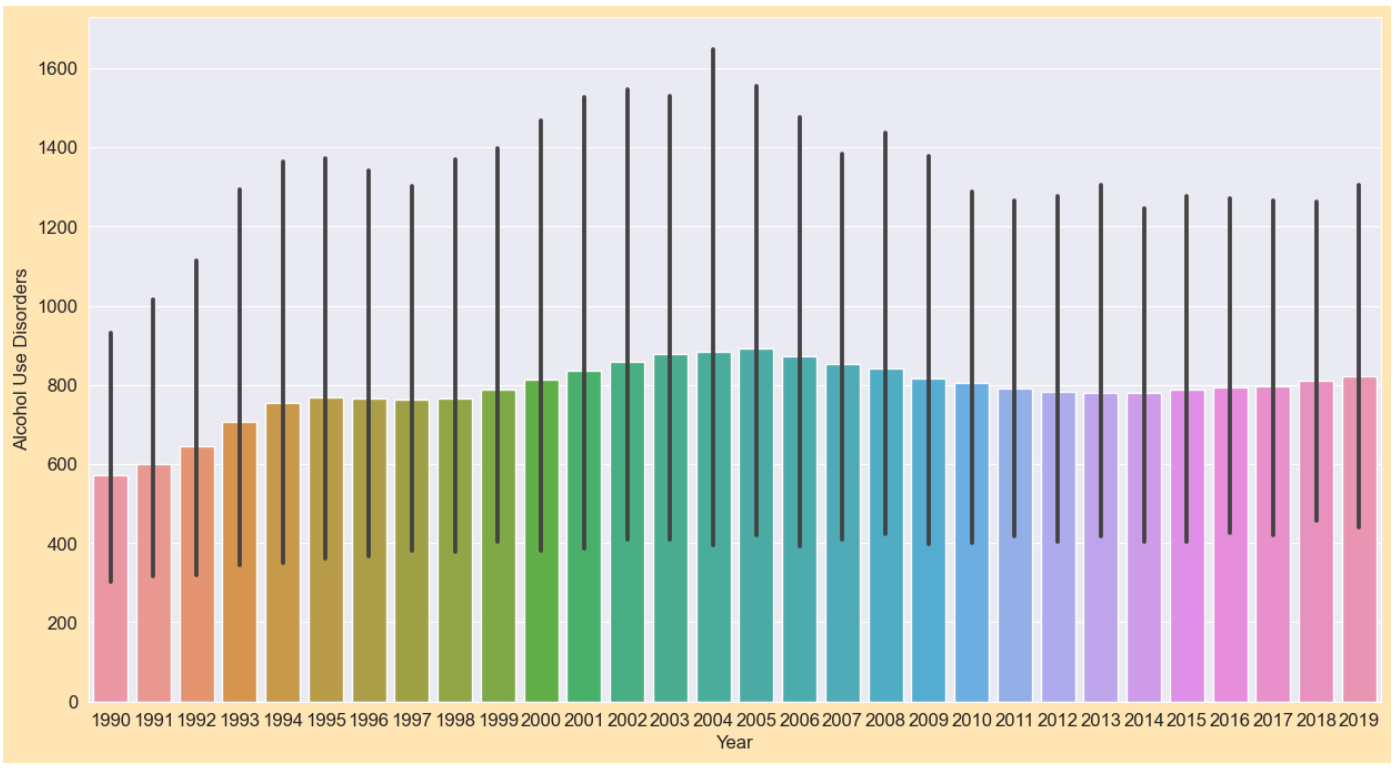

Drug Use Disorders
0
200
400
600
800
1000
1200
1400
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

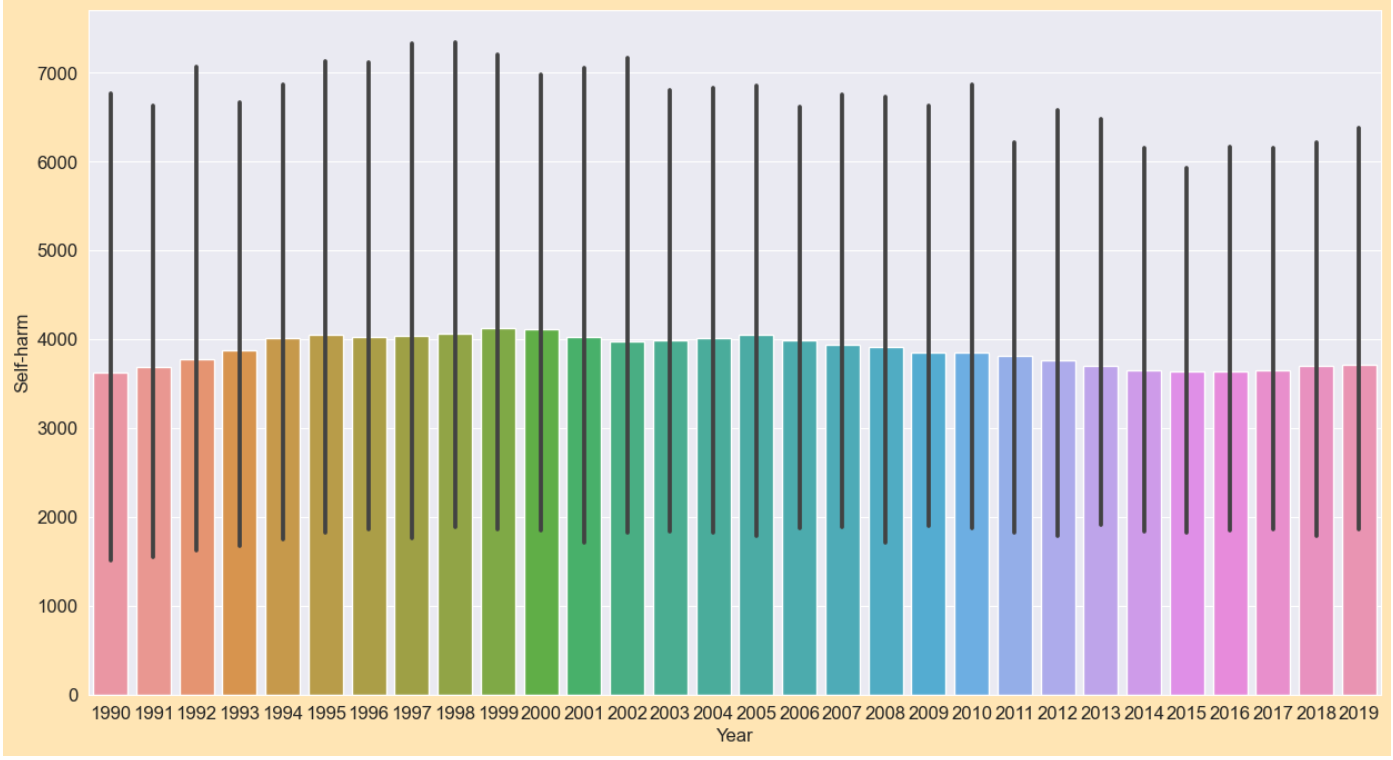

Tuberculosis
0
2000
4000
6000
8000
10000
12000
14000
16000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

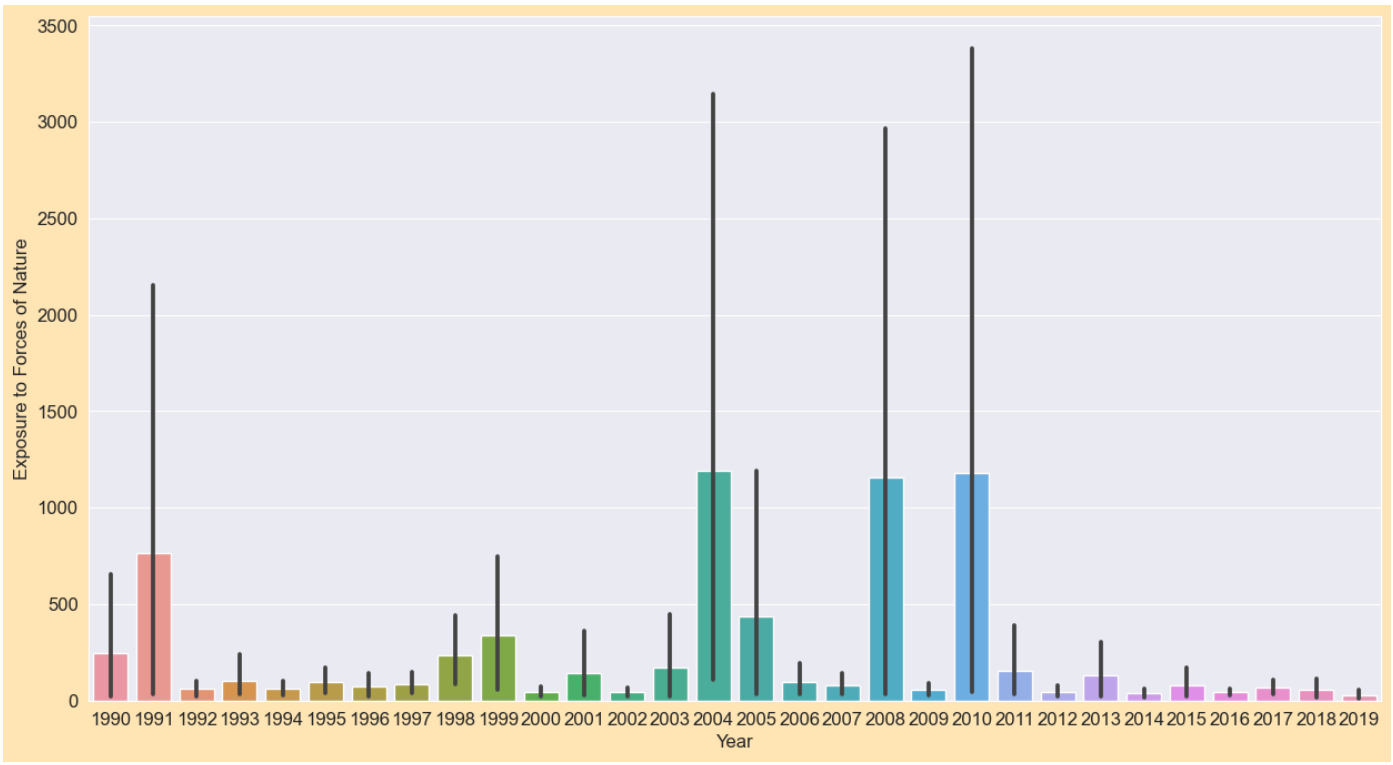

Cardiovascular Diseases
160000
140000
120000
100000
80000
60000
40000
20000
0
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

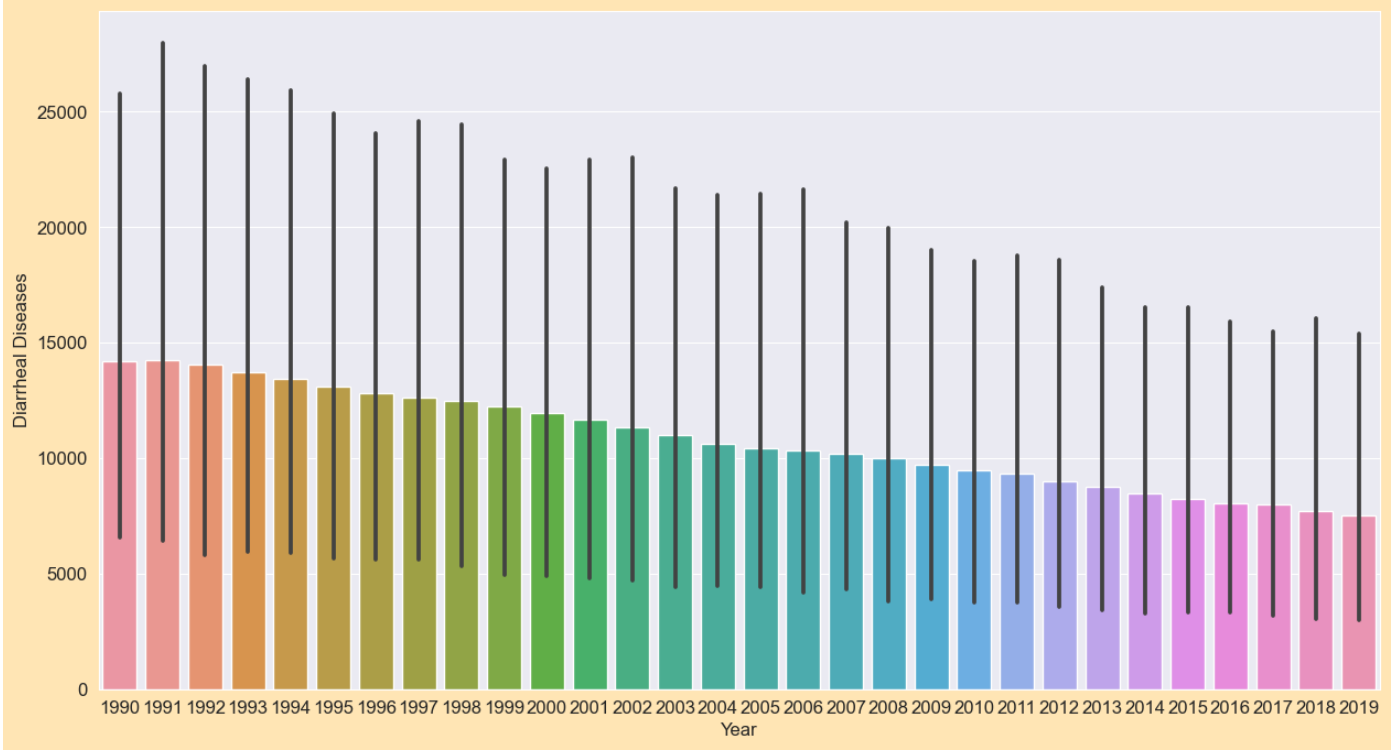

Lower Respiratory Infections
25000
20000
15000
10000
5000
0
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

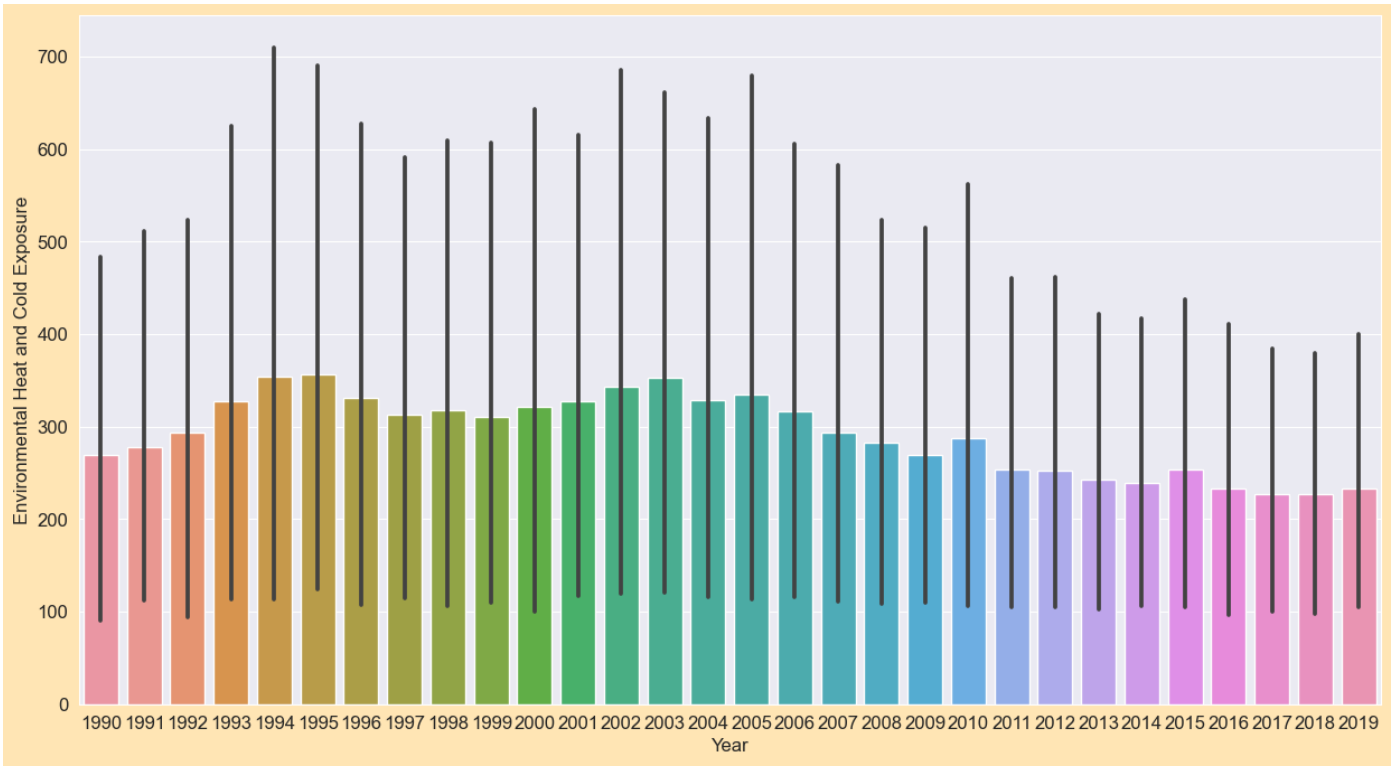

Neonatal Disorders
25000
20000
15000
10000
5000
0
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

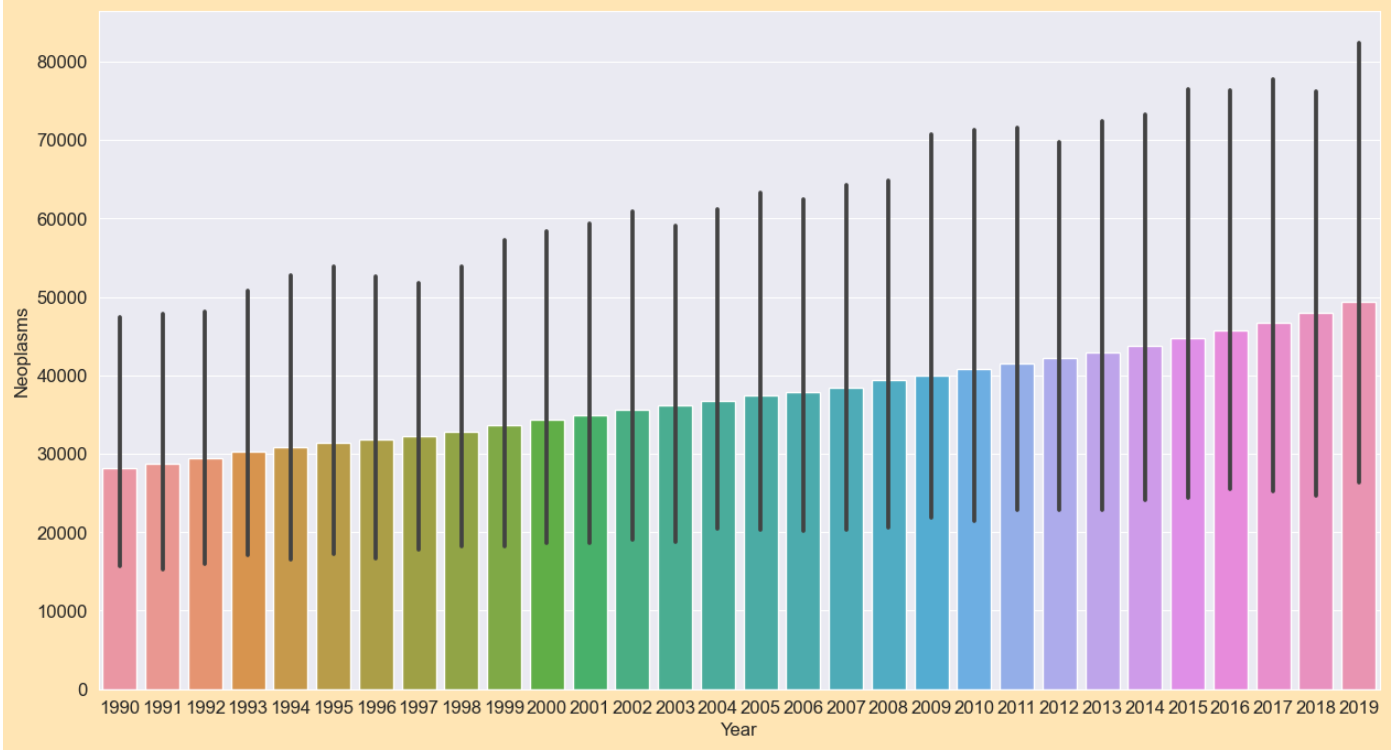

Alcohol Use Disorders
1600
1400
1200
1000
800
600
400
200
0
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

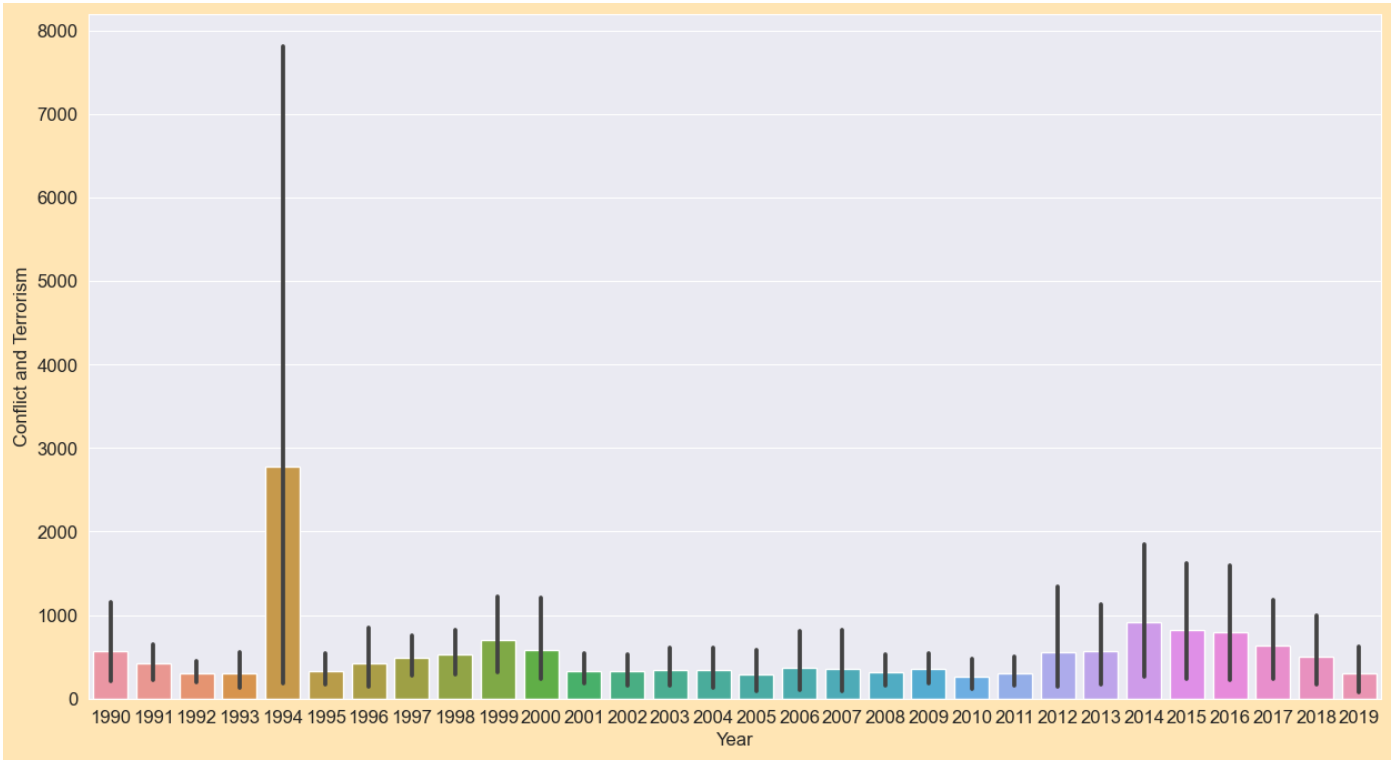

Self-harm
0
1000
2000
3000
4000
5000
6000
7000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

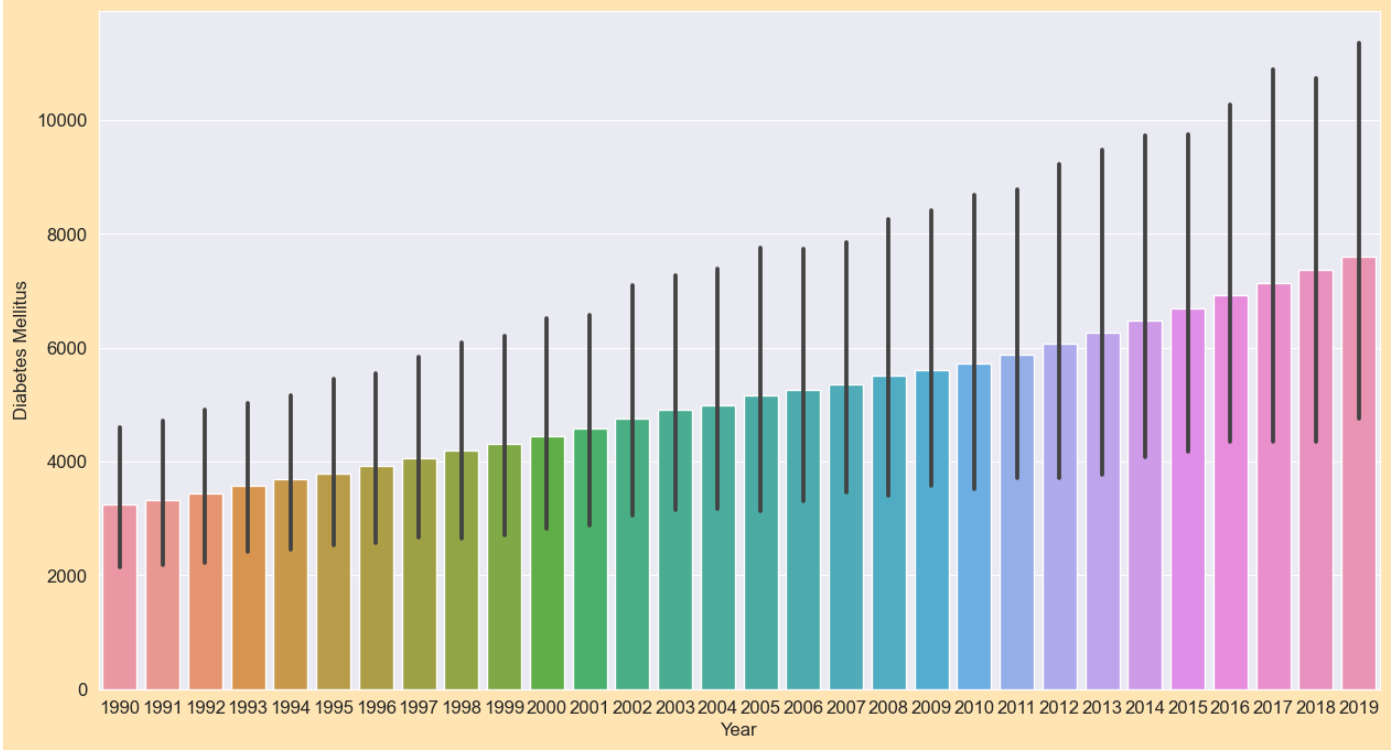

Exposure to Forces of Nature
0
500
1000
1500
2000
2500
3000
3500
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

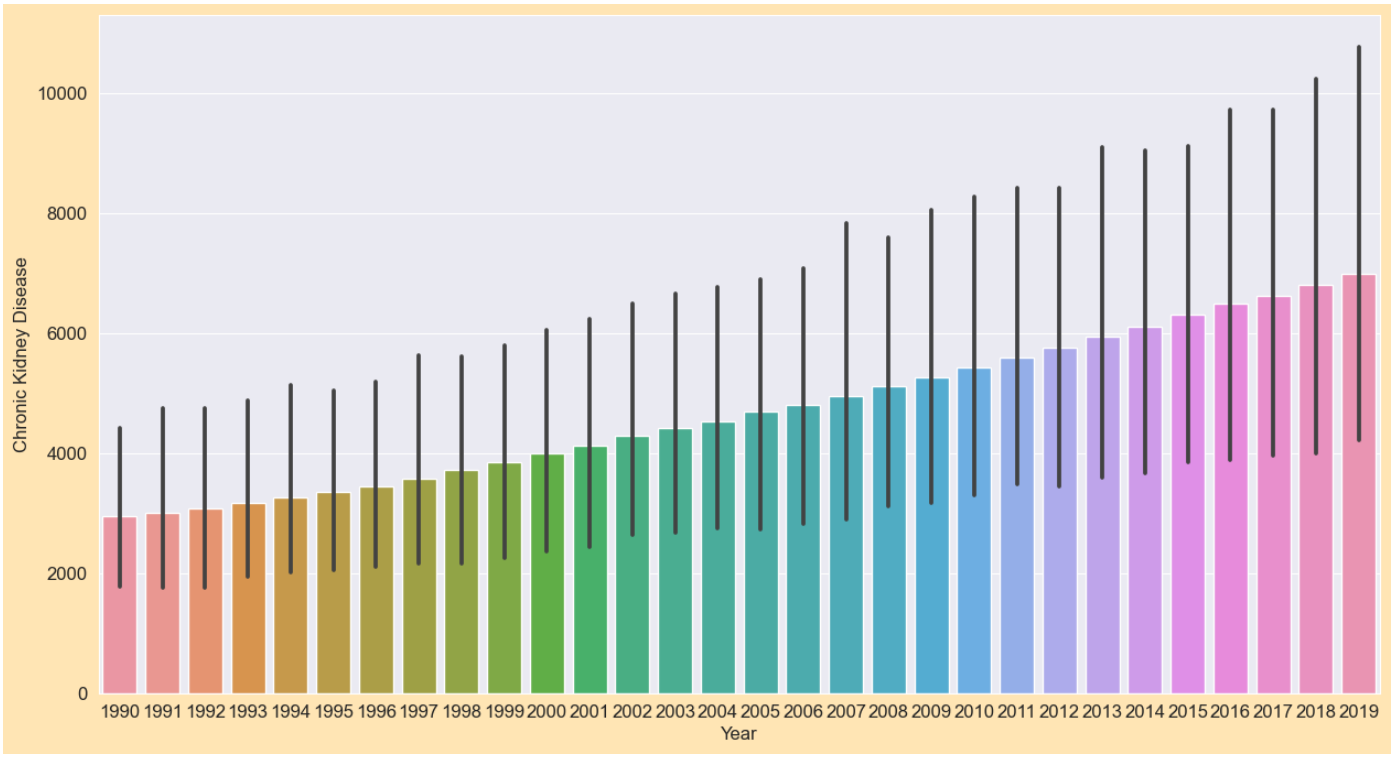

Diarrheal Diseases
0
5000
10000
15000
20000
25000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

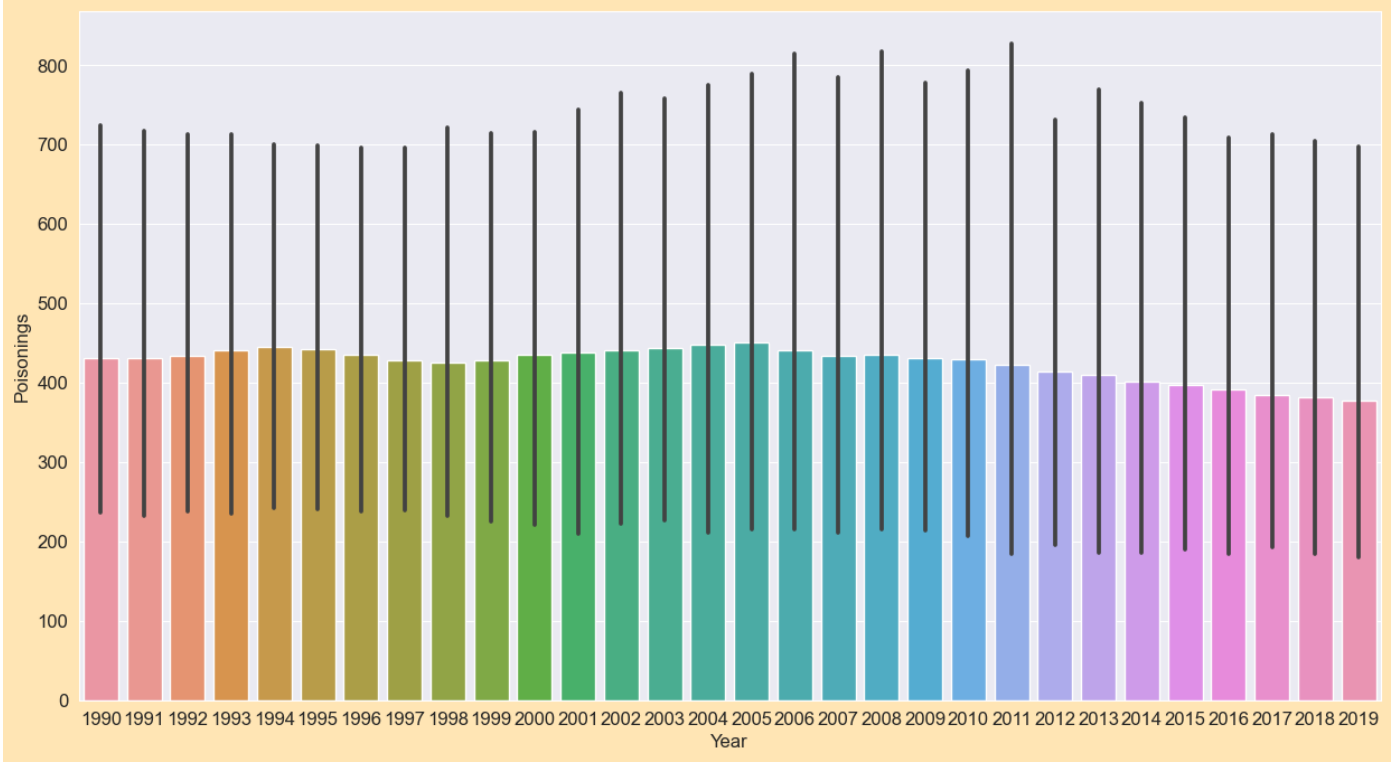

Environmental Heat and Cold Exposure
0
100
200
300
400
500
600
700
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

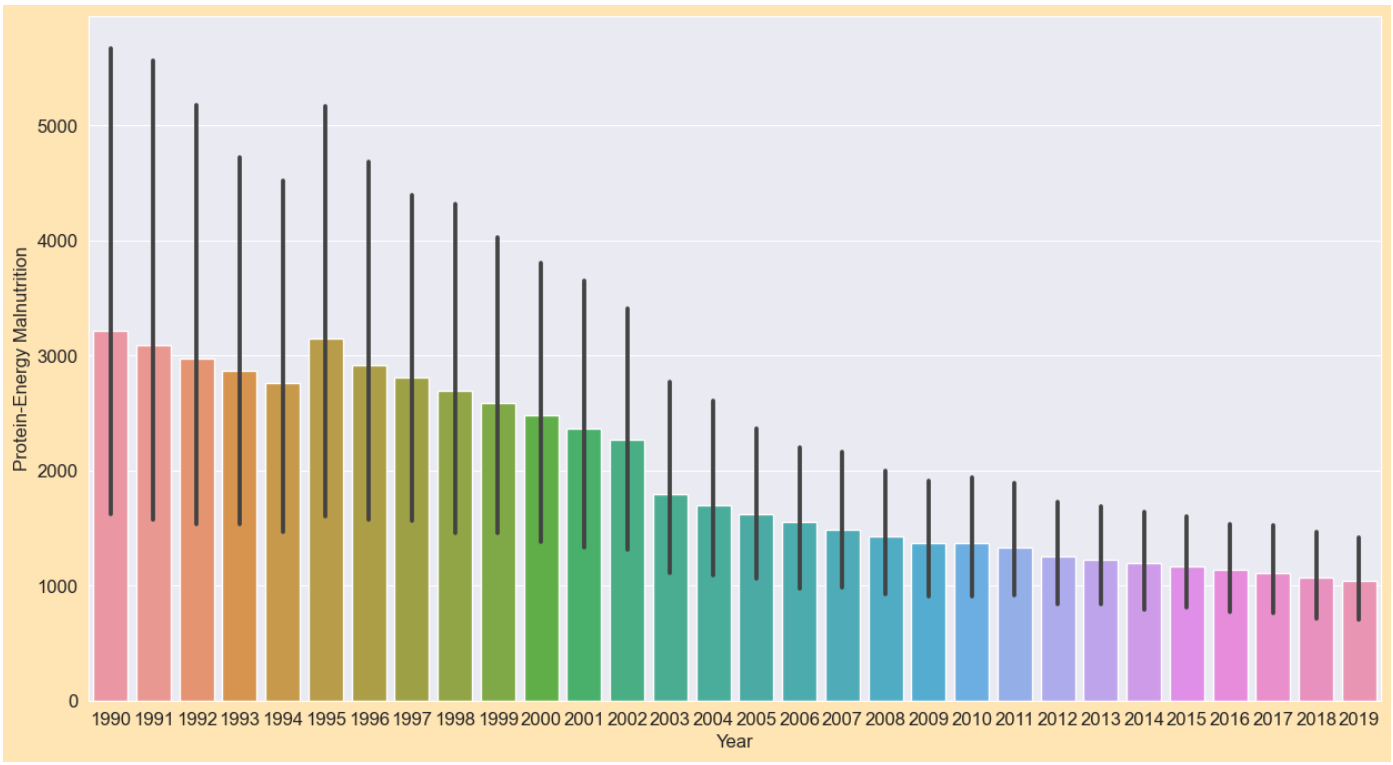

Neoplasms
0
10000
20000
30000
40000
50000
60000
70000
80000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

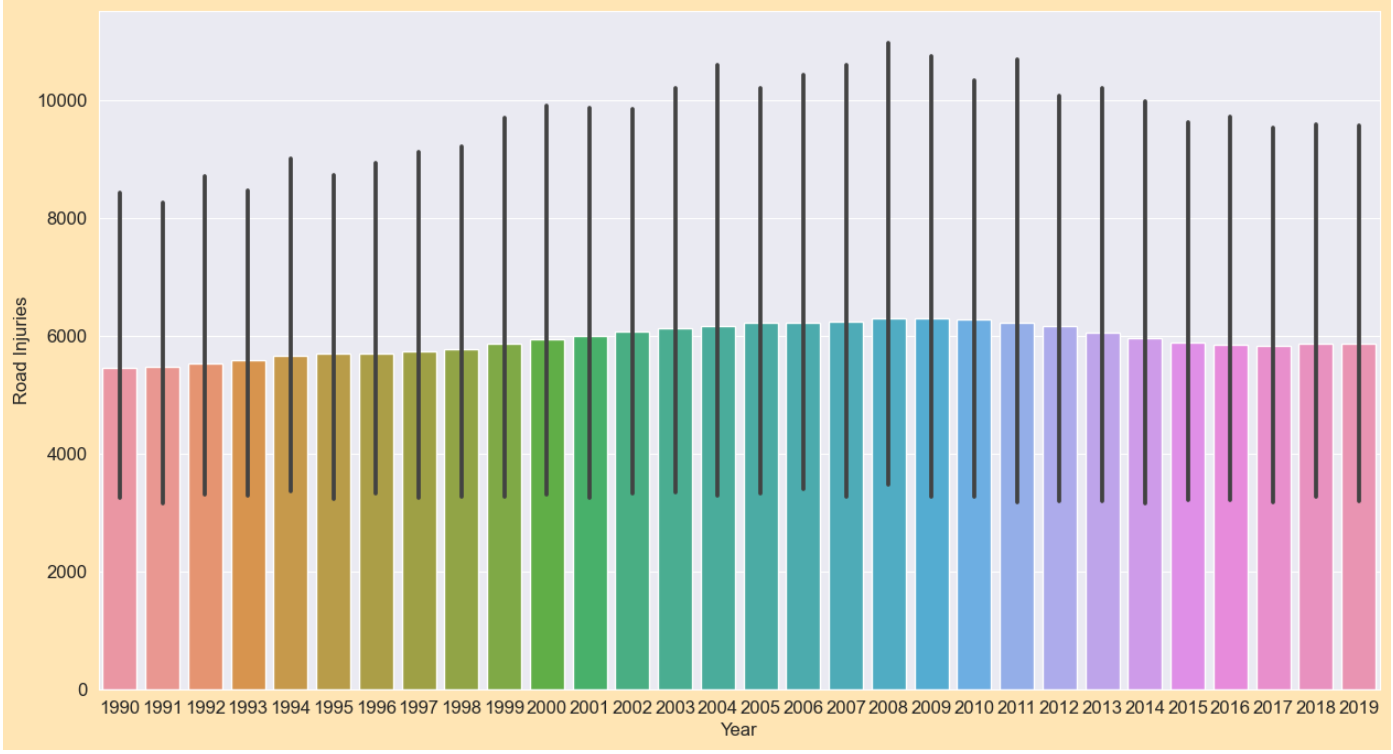

Conflict and Terrorism
0
1000
2000
3000
4000
5000
6000
7000
8000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

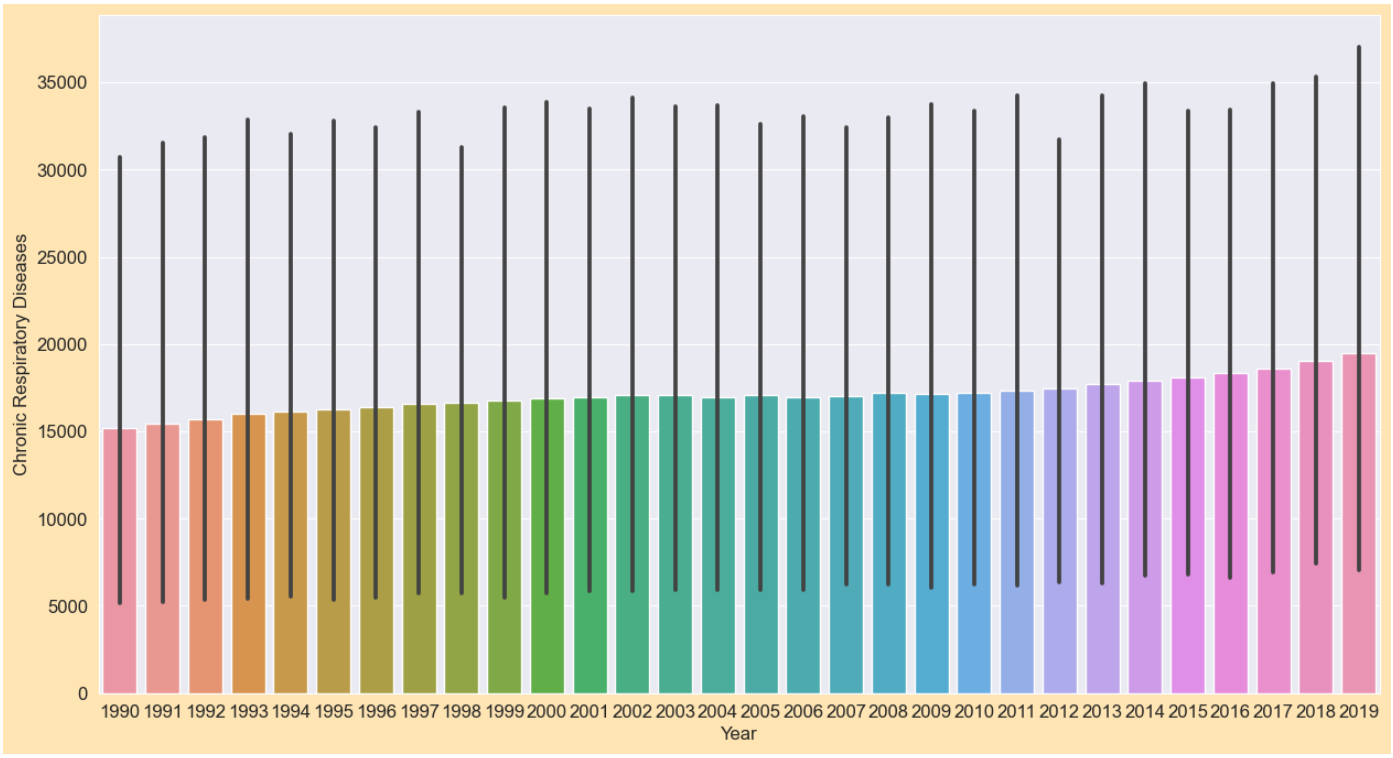

Diabetes Mellitus
0
2000
4000
6000
8000
10000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

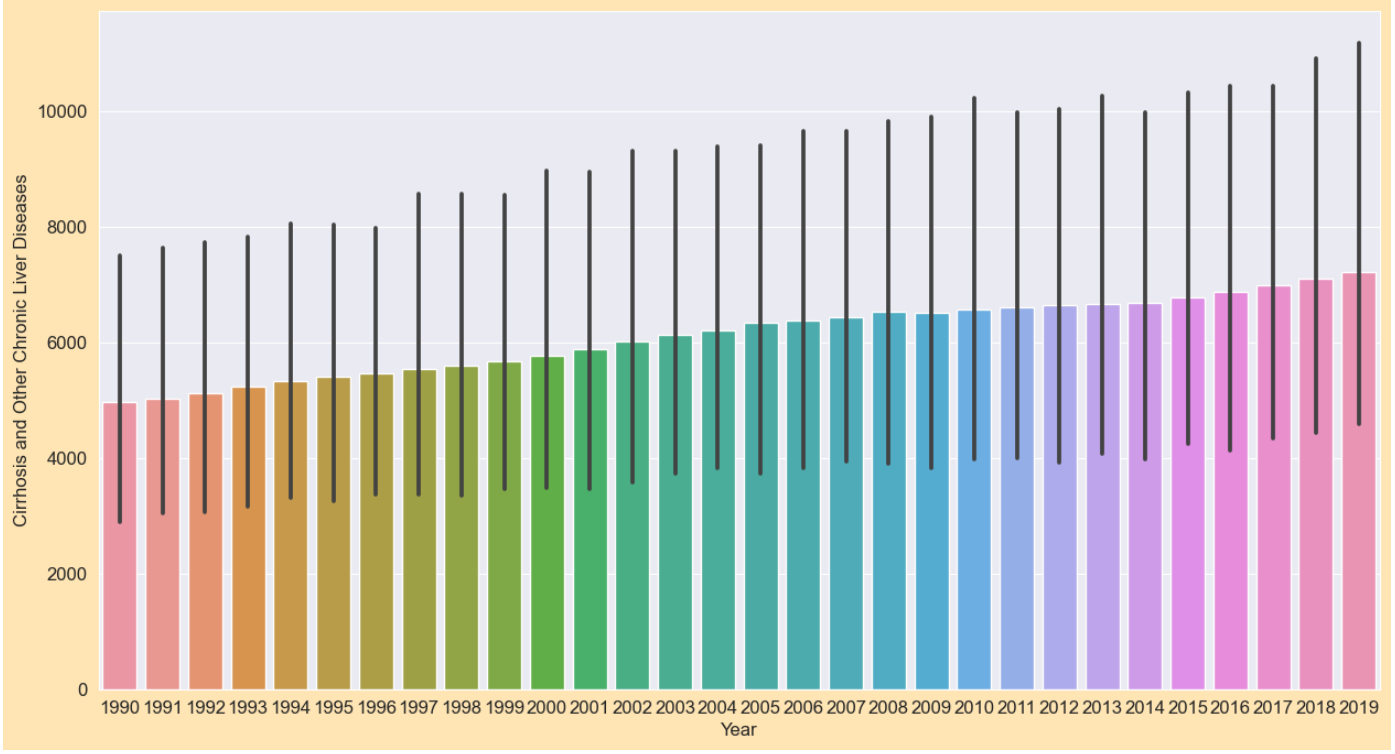

Chronic Kidney Disease
0
2000
4000
6000
8000
10000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

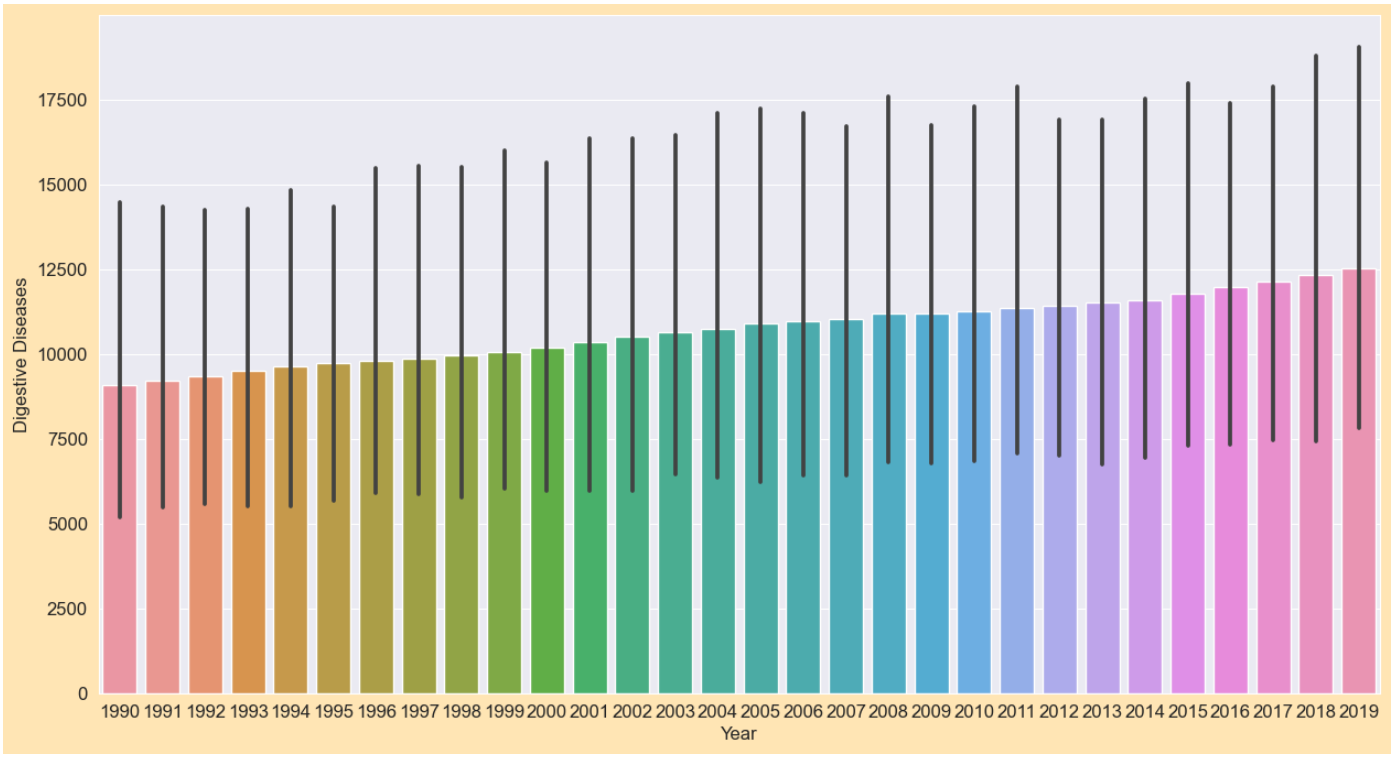

Poisonings
0
100
200
300
400
500
600
700
800
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

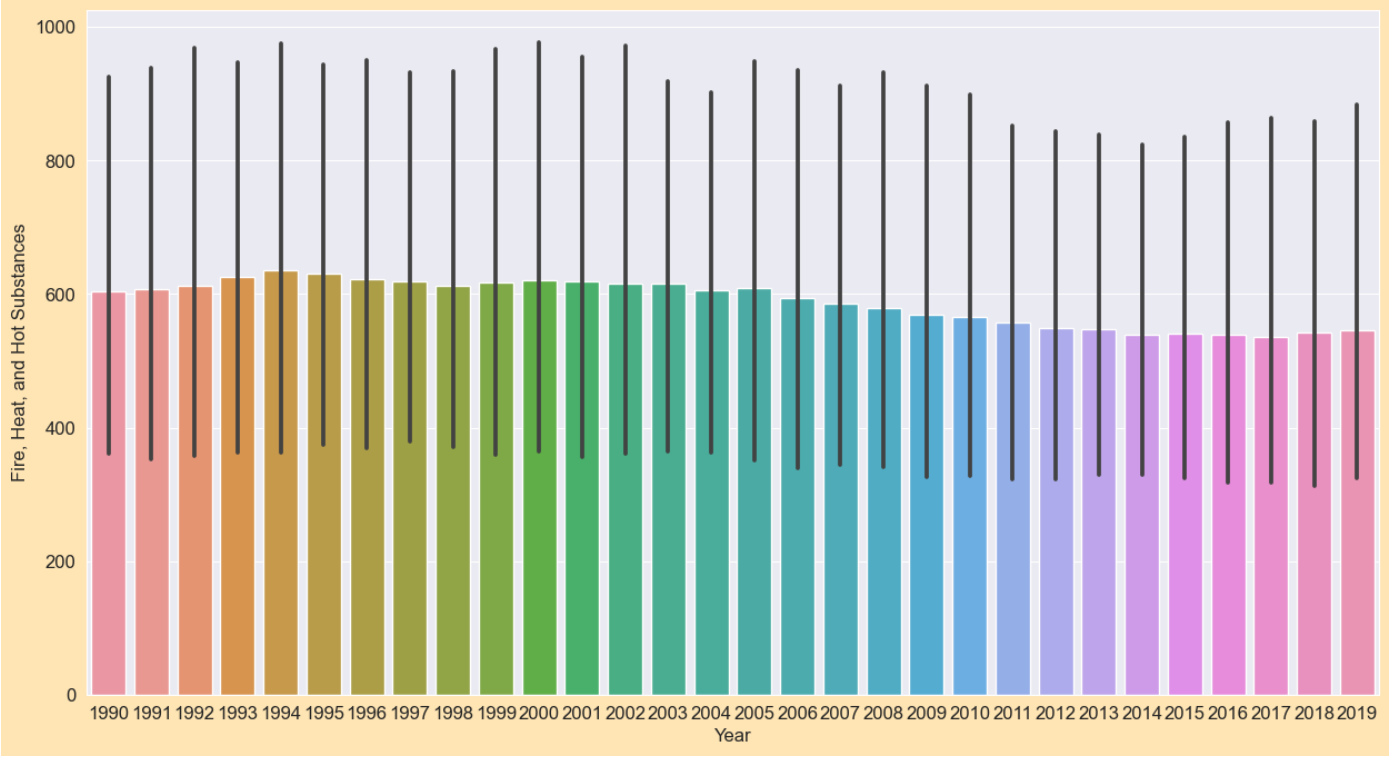

Protein-Energy Malnutrition
0
1000
2000
3000
4000
5000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year

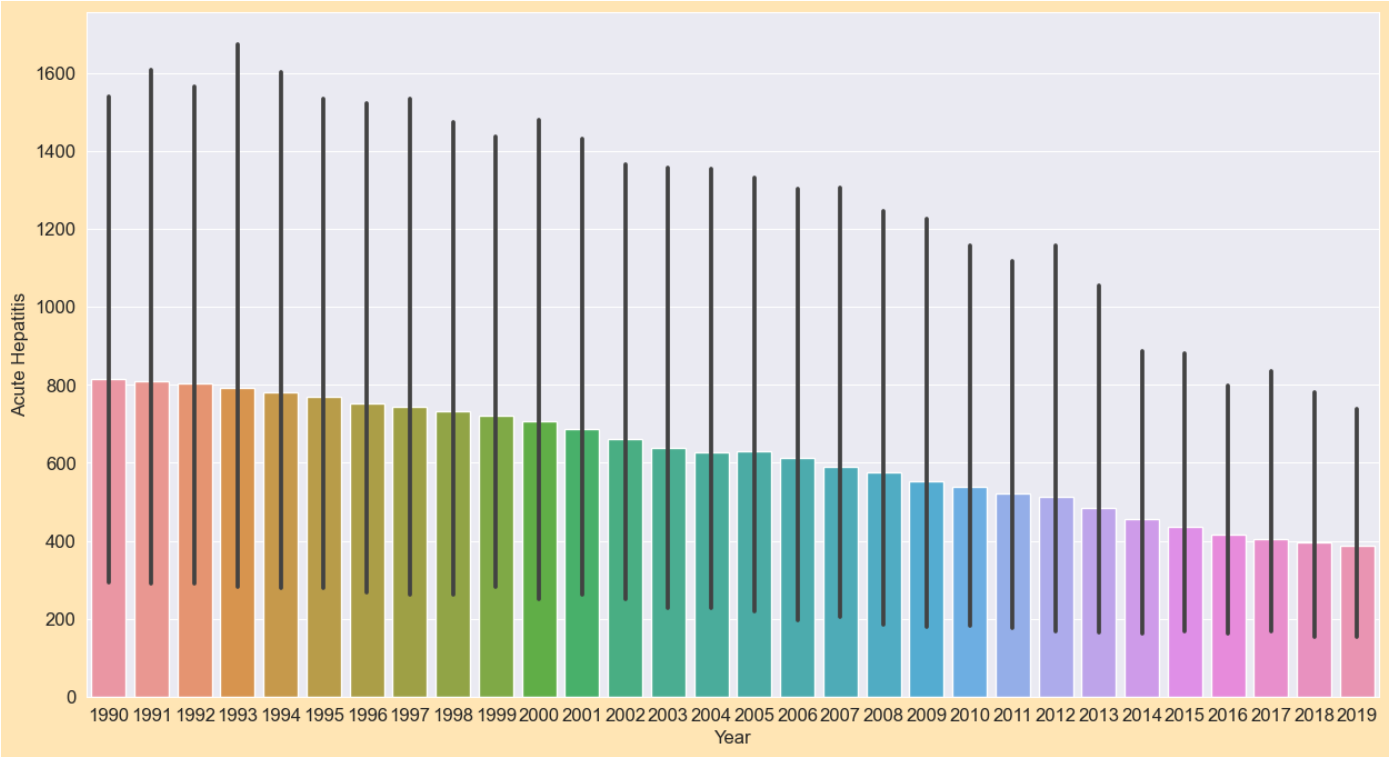

Road Injuries
0
2000
4000
6000
8000
10000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year


Chronic Respiratory Diseases
0
5000
10000
15000
20000
25000
30000
35000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year


Cirrhosis and Other Chronic Liver Diseases
0
2000
4000
6000
8000
10000
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year


Digestive Diseases
0
2500
5000
7500
10000
12500
15000
17500
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Year





Observation:- 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.

In [35]:
```
df.corr()
```

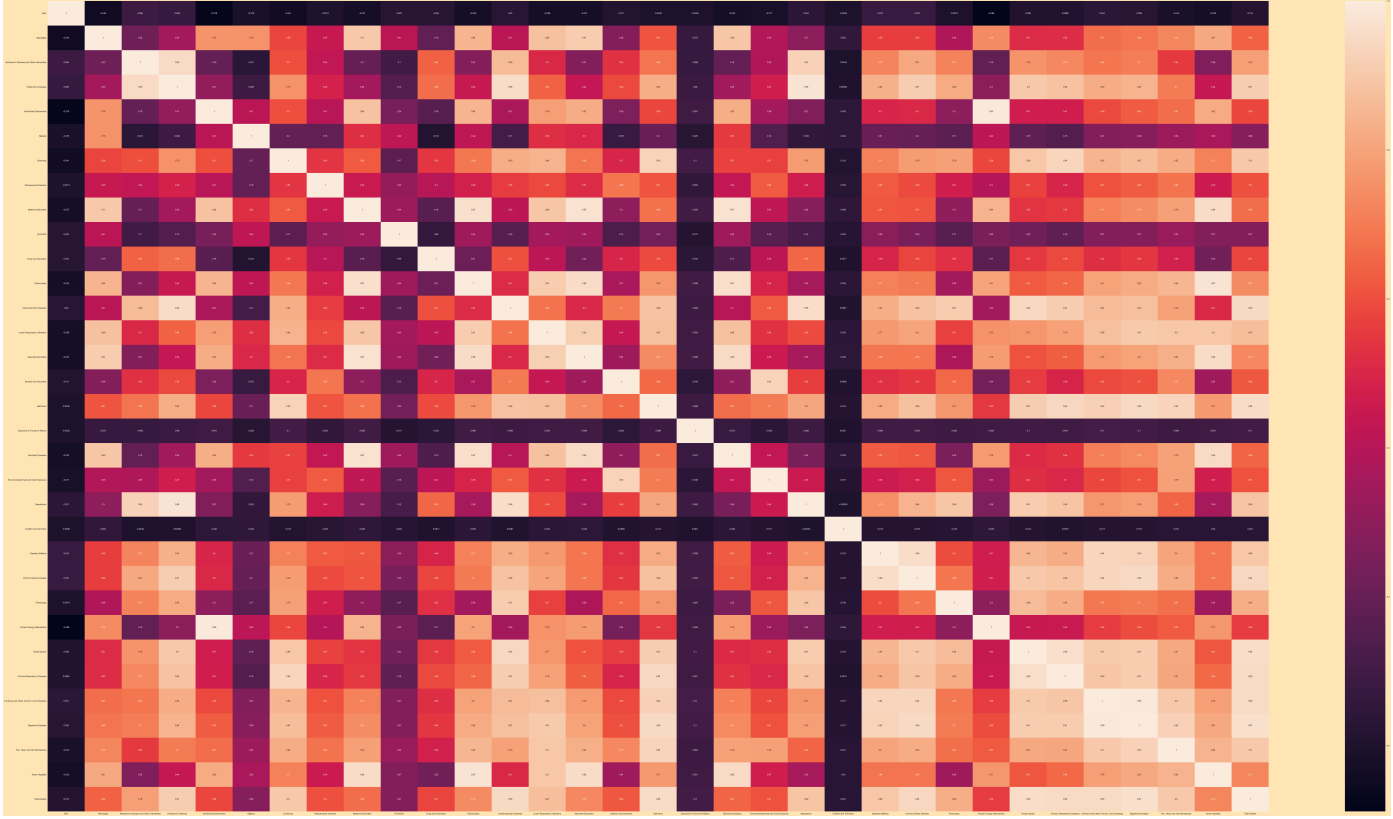
Out[35]:

	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence
Year	1.000000	-0.043288	0.083710	0.068756	-0.078266	-0.015964	-0.040910	-0.001122
Meningitis	-0.043288	1.000000	0.216713	0.351668	0.760851	0.755261	0.576347	0.447242
Alzheimer's Disease and Other Dementias	0.083710	0.216713	1.000000	0.950785	0.193209	0.031290	0.599403	0.429622
Parkinson's Disease	0.068756	0.351668	0.950785	1.000000	0.313033	0.084109	0.753663	0.485528
Nutritional Deficiencies	-0.078266	0.760851	0.193209	0.313033	1.000000	0.411149	0.596367	0.407065
Malaria	-0.015964	0.755261	0.031290	0.084109	0.411149	1.000000	0.195839	0.184469
Drowning	-0.040910	0.576347	0.599403	0.753663	0.596367	0.195839	1.000000	0.539339
Interpersonal Violence	-0.001122	0.447242	0.429622	0.485528	0.407065	0.184469	0.539339	1.000000
Maternal Disorders	-0.027460	0.899507	0.200315	0.356394	0.881740	0.523581	0.623558	0.449551
HIV/AIDS	0.022964	0.411881	0.101628	0.145521	0.241247	0.424471	0.171108	0.315642
Drug Use Disorders	0.023917	0.187050	0.641341	0.664385	0.163638	0.011560	0.545004	0.403071
Tuberculosis	-0.025297	0.844494	0.273336	0.445504	0.844321	0.423077	0.692165	0.478608
Cardiovascular Diseases	0.029813	0.411787	0.860759	0.956667	0.370829	0.107651	0.829424	0.554629
Lower Respiratory Infections	-0.027531	0.879827	0.503257	0.638771	0.783957	0.520901	0.840597	0.583966
Neonatal Disorders	-0.026949	0.908737	0.270157	0.435508	0.824924	0.501561	0.684802	0.512989
Alcohol Use Disorders	0.011315	0.275909	0.529750	0.584208	0.256448	0.073306	0.497126	0.687922
Self-harm	-0.004192	0.609952	0.682463	0.826083	0.581602	0.202015	0.929865	0.611616
Exposure to Forces of Nature	-0.005178	0.071674	0.066285	0.080246	0.091046	0.025273	0.104622	0.052323
Diarrheal Diseases	-0.031911	0.892564	0.189659	0.347421	0.829348	0.551661	0.573699	0.432915

Environmental Heat and Cold Exposure	-0.017286	0.390214	0.384578	0.469845	0.356665	0.144287	0.566503	0.623965
Neoplasms	0.036753	0.299265	0.925096	0.975203	0.266367	0.051927	0.782617	0.460441
Conflict and Terrorism	-0.005941	0.052123	-0.004190	0.000944	0.041965	0.043511	0.012436	0.023094
Diabetes Mellitus	0.074292	0.561177	0.723417	0.835941	0.495049	0.205493	0.714079	0.621728
Chronic Kidney Disease	0.074970	0.562013	0.814841	0.911376	0.504551	0.198686	0.781125	0.589217
Poisonings	-0.007414	0.381899	0.719373	0.818444	0.298918	0.169424	0.790391	0.471659
Protein-Energy Malnutrition	-0.088377	0.738272	0.185030	0.295749	0.991700	0.416539	0.579360	0.396216
Road Injuries	0.005982	0.511316	0.760421	0.897982	0.465285	0.178822	0.894074	0.574467
Chronic Respiratory Diseases	0.008645	0.516440	0.728337	0.882246	0.472233	0.146132	0.936633	0.494280
Cirrhosis and Other Chronic Liver Diseases	0.031126	0.665212	0.686722	0.825476	0.586889	0.271206	0.846058	0.606565
Digestive Diseases	0.025539	0.684047	0.703813	0.839807	0.626332	0.282995	0.869528	0.629069
Fire, Heat, and Hot Substances	-0.014134	0.724751	0.551121	0.699626	0.665677	0.335181	0.823063	0.679620
Acute Hepatitis	-0.032604	0.813554	0.263559	0.438842	0.820830	0.362693	0.711640	0.453349
Total Deaths	0.015173	0.641318	0.783616	0.910980	0.582967	0.277913	0.901813	0.604067

In [36]:
```
plt.figure(figsize=(150,80))
sns.heatmap(df.corr(),annot=True)
```

Out[36]:
```
<AxesSubplot:>
```



In [ ]: