Cause Of Death Analisys

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.

Dataset Glossary (Column-wise)

1. Country/Territory - Name of the Country/Territory
2. Code - Country/Territory Code
3. Year - Year of the Incident
4. Meningitis - No. of People died from Meningitis
5. Alzheimer's Disease and Other Dementias - No. of People died from Alzheimer's Disease and Other Dementias
6. Parkinson's Disease - No. of People died from Parkinson's Disease
7. Nutritional Deficiencies - No. of People died from Nutritional Deficiencies
8. Malaria - No. of People died from Malaria
9. 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

Importing Necessary Libraries:

```
Start coding or generate with AI.
```

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib as mpl
import matplotlib.pyplot as plt
%matplotlib inline

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

Importing dataset from csv_file to pandas_dataframe:

```
df = pd.read_csv('cause_of_deaths dataset1.csv')
#pd.set_option('display.max_columns', None) # This will enable us to see truncated columns
#pd.set_option('display.max_rows', None) # This will enable us to see truncated rows
df.head()
```

	Country/Territory	Code	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	...	Diabetes Mellitus
0	Afghanistan	AFG	1990	2159	1116	371	2087	93	1370	1538	...	2108
1	Afghanistan	AFG	1991	2218	1136	374	2153	189	1391	2001	...	2120
2	Afghanistan	AFG	1992	2475	1162	378	2441	239	1514	2299	...	2153
3	Afghanistan	AFG	1993	2812	1187	384	2837	108	1687	2589	...	2195
4	Afghanistan	AFG	1994	3027	1211	391	3081	211	1809	2849	...	2231

5 rows × 34 columns

Data Profiling:

```
print('No. of Rows :',df.shape[0])
print('No. of Columns :',df.shape[1])
```

```
No. of Rows : 6120
No. of Columns : 34
```

```
df.columns.to_series().groupby(df.dtypes).groups
```

```
{int64: ['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'], object:
['Country/Territory', 'Code']}
```

```
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
```

```
# Check unique values in each column
df.nunique()
```

```
Country/Territory                            204
Code                                         204
Year                                          30
Meningitis                                  2020
Alzheimer's Disease and Other Dementias     3037
Parkinson's Disease                         1817
Nutritional Deficiencies                    2147
Malaria                                     1723
Drowning                                    1875
Interpersonal Violence                      2142
Maternal Disorders                          1818
HIV/AIDS                                    2412
Drug Use Disorders                           876
Tuberculosis                                2843
Cardiovascular Diseases                     5225
Lower Respiratory Infections                4106
Neonatal Disorders                          3553
Alcohol Use Disorders                       1287
Self-harm                                   2758
Exposure to Forces of Nature                 478
Diarrheal Diseases                          2874
Environmental Heat and Cold Exposure         714
Neoplasms                                   4814
Conflict and Terrorism                       918
Diabetes Mellitus                           3366
Chronic Kidney Disease                      3246
Poisonings                                  1087
Protein-Energy Malnutrition                 2091
Road Injuries                               3393
Chronic Respiratory Diseases                3803
Cirrhosis and Other Chronic Liver Diseases  3443
Digestive Diseases                          4023
Fire, Heat, and Hot Substances              1406
Acute Hepatitis                             1059
dtype: int64
```

Observations:

- We have **6120 Rows and 34 Columns** in this Dataset.
- We have **32 Numerical dtype Columns and 02 Object dtype columns**.
- So, far **we don't have any null values** in any of the columns of the dataset.

- We may **drop** either 'Country/Territory' or 'Code' Column, as they boths gives the same info.
- Object data type columns need to be converted to numerical data.

Data Integrity Check

- Dataset may contain whitespaces, missing value etc. let's investigate the integrity of data before proceeding for further analysis.

```
# Checking for the presense of whitespaces, '-', 'null', 'NA':

df.isin(['None','-999','0',' ','?','-',' ?','null','NA','N/A','nan','Nan','NAN','NaN']).sum().any()
```

```
False
```

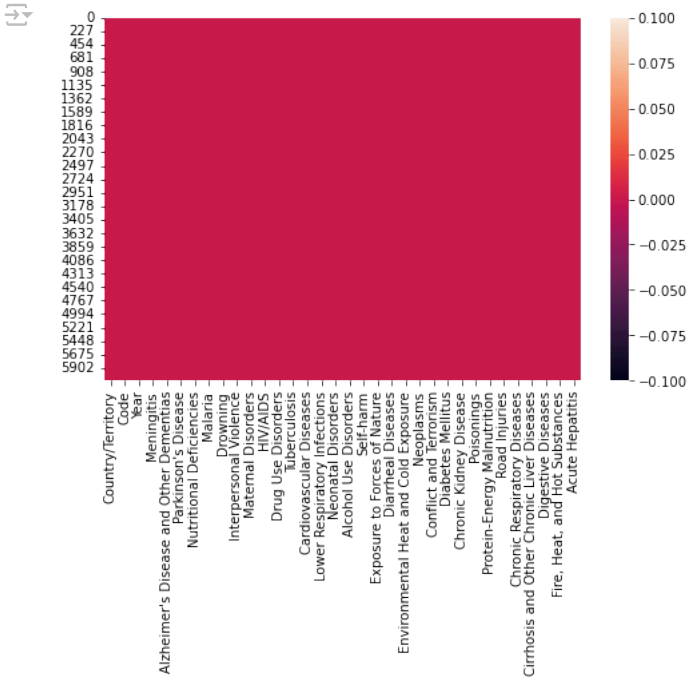
```
df.isnull().sum().any()
```

```
False
```

```
df.duplicated().sum() # Checking for duplicacy.
```

```
0
```

```
# Missing value check

sns.set_palette('coolwarm_r')
plt.figure(figsize = (8,5))
sns.heatmap(df.isnull())
plt.show()
```




```
# getting NaN Columns data:
pd.set_option('display.max_rows',None) # to display all rows.
miss_val = df.isnull().sum().sort_values(ascending = False)
percent_miss_val =(miss_val/len(df))*100
print(pd.concat([miss_val, percent_miss_val], axis =1, keys =["No. of NaN's", "% NaN data"]))
```

```
                                    No. of NaN's  % NaN data
Country/Territory                              0         0.0
Chronic Kidney Disease                         0         0.0
Exposure to Forces of Nature                   0         0.0
Diarrheal Diseases                             0         0.0
Environmental Heat and Cold Exposure           0         0.0
Neoplasms                                      0         0.0
Conflict and Terrorism                         0         0.0
Diabetes Mellitus                              0         0.0
Poisonings                                     0         0.0
Code                                           0         0.0
Protein-Energy Malnutrition                    0         0.0
Road Injuries                                  0         0.0
Chronic Respiratory Diseases                   0         0.0
Cirrhosis and Other Chronic Liver Diseases     0         0.0
Digestive Diseases                             0         0.0
Fire, Heat, and Hot Substances                 0         0.0
```

```
Self-harm                                   0         0.0
Alcohol Use Disorders                       0         0.0
Neonatal Disorders                          0         0.0
Lower Respiratory Infections                0         0.0
Cardiovascular Diseases                     0         0.0
Tuberculosis                                0         0.0
Drug Use Disorders                          0         0.0
HIV/AIDS                                    0         0.0
Maternal Disorders                          0         0.0
Interpersonal Violence                      0         0.0
Drowning                                    0         0.0
Malaria                                     0         0.0
Nutritional Deficiencies                    0         0.0
Parkinson's Disease                         0         0.0
Alzheimer's Disease and Other Dementias     0         0.0
Meningitis                                  0         0.0
Year                                        0         0.0
Acute Hepatitis                             0         0.0
```

Observation:

- Dataset do not contain whitespaces, missing value, 'None','-999','0',' ','?','-',' ?','null','NA','N/A'.
- Dataset don't have any duplicate values in it.

```
Start coding or generate with AI.
```

Statistical Analysis:

```
# Segregating Columns Based on data Types:

Num_data = ['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']

Cat_data = ['Country/Territory','Code']


df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
Year	6120.0	2004.500000	8.656149	1990.0	1997.00	2004.5	2012.00	2019.0
Meningitis	6120.0	1719.701307	6672.006930	0.0	15.00	109.0	847.25	98358.0
Alzheimer's Disease and Other Dementias	6120.0	4864.189379	18220.659072	0.0	90.00	666.5	2456.25	320715.0
Parkinson's Disease	6120.0	1173.169118	4616.156238	0.0	27.00	164.0	609.25	76990.0
Nutritional Deficiencies	6120.0	2253.600000	10483.633601	0.0	9.00	119.0	1167.25	268223.0
Malaria	6120.0	4140.960131	18427.753137	0.0	0.00	0.0	393.00	280604.0
Drowning	6120.0	1683.333170	8877.018366	0.0	34.00	177.0	698.00	153773.0
Interpersonal Violence	6120.0	2083.797222	6917.006075	0.0	40.00	265.0	877.00	69640.0
Maternal Disorders	6120.0	1262.589216	6057.973183	0.0	5.00	54.0	734.00	107929.0
HIV/AIDS	6120.0	5941.898529	21011.962487	0.0	11.00	136.0	1879.00	305491.0
Drug Use Disorders	6120.0	434.006699	2898.761628	0.0	3.00	20.0	129.00	65717.0
Tuberculosis	6120.0	7491.928595	39549.977578	0.0	35.00	417.0	2924.25	657515.0
Cardiovascular Diseases	6120.0	73160.454575	291577.537794	4.0	2028.00	11742.0	42546.50	4584273.0
Lower Respiratory Infections	6120.0	13687.914706	48031.720009	0.0	345.00	2126.5	10161.25	690913.0
Neonatal Disorders	6120.0	12558.942647	56058.366412	0.0	131.00	916.0	7419.75	852761.0
Alcohol Use Disorders	6120.0	787.421242	3545.823616	0.0	9.00	80.0	316.00	55200.0
Self-harm	6120.0	3874.825327	18425.616418	0.0	94.00	533.0	1882.25	220357.0
Exposure to Forces of Nature	6120.0	243.485621	4717.104377	0.0	0.00	0.0	12.00	222641.0
Diarrheal Diseases	6120.0	10822.795425	65416.174485	0.0	20.00	296.5	3946.75	1119477.0
Environmental Heat and Cold Exposure	6120.0	292.295915	1704.466356	0.0	2.00	21.0	109.00	29048.0
Neoplasms	6120.0	37542.244771	161558.365445	1.0	809.75	5629.5	20147.75	2716551.0
Conflict and Terrorism	6120.0	538.243954	7033.308187	0.0	0.00	0.0	23.00	503532.0
Diabetes Mellitus	6120.0	5138.704575	16773.081040	1.0	236.00	1087.0	2954.00	273089.0
Chronic Kidney Disease	6120.0	4724.132680	16470.429969	0.0	145.75	822.0	2922.50	222922.0
Poisonings	6120.0	425.013399	2022.640521	0.0	6.00	52.5	254.00	30883.0
Protein-Energy Malnutrition	6120.0	1965.994281	8255.999063	0.0	5.00	92.0	1042.50	202241.0
Road Injuries	6120.0	5930.795588	24097.784291	0.0	174.75	966.5	3435.25	329237.0
Chronic Respiratory Diseases	6120.0	17092.374837	105157.179839	1.0	289.00	1689.0	5249.75	1366039.0
Cirrhosis and Other Chronic Liver Diseases	6120.0	6124.072059	20688.118580	0.0	154.00	1210.0	3547.25	270037.0
Digestive Diseases	6120.0	10725.267157	37228.051096	0.0	284.00	2185.0	6080.00	464914.0
Fire, Heat, and Hot Substances	6120.0	588.711438	2128.595120	0.0	17.00	126.0	450.00	25876.0
Acute Hepatitis	6120.0	618.429902	4186.023497	0.0	2.00	15.0	160.00	64305.0

```
df[Cat_data].describe()
```

	Country/Territory	Code
count	6120	6120
unique	204	204
top	Afghanistan	AFG
freq	30	30

Feature Engineering:

```
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')
```

```
death_dcs = ['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']
```

```
# dropping 'Code' column, as 'Country/Territory' & 'Code' Column both gives the same info.
df = df.drop("Code", axis='columns')
```

```
# Creating a new column for 'Total_Deaths' for individual Country and Year
df['Total_Deaths'] = df[death_dcs].sum(axis=1)
```

This dataset is interesting as it contains a lot of important information, but there is not one clearly defined dependent, or target, variable. Hence, full discretion is given to the user of this dataset while exploring the limitations of the given data and performing analysis.

Segregating the Cause of deaths:

```
df.columns
```

```
Index(['Country/Territory', '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', 'Total_Deaths'],
      dtype='object')
```

```
Total_brain_Deaths = ['Meningitis',"Alzheimer's Disease and Other Dementias", "Parkinson's Disease"]

Total_gastrointestinal_Deaths = ['Nutritional Deficiencies','Digestive Diseases','Diarrheal Diseases',
                                 'Protein-Energy Malnutrition']

Total_respiratory_Deaths = ['Tuberculosis','Lower Respiratory Infections','Chronic Respiratory Diseases']

heart_blood_cancer_Deaths_respectively = ['Cardiovascular Diseases','HIV/AIDS','Neoplasms']

Total_Pregnancy_Deaths = ['Neonatal Disorders','Maternal Disorders']

Total_kidney_Deaths= ['Chronic Kidney Disease','Diabetes Mellitus']

Total_liver_Deaths = ['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases','Malaria','Acute Hepatitis']

Nat_or_man_made_Deaths = ['Interpersonal Violence','Drug Use Disorders','Self-harm', 'Exposure to Forces of Nature',
         "Fire, Heat, and Hot Substances",'Poisonings','Conflict and Terrorism','Environmental Heat and Cold Exposure',
        'Road Injuries','Drowning']
```

```
Start coding or generate with AI.
```

Statistical & Exploratory Data Analysis (EDA):

```
# listing top 'Total_Deaths' countries.

Country_list = df.groupby("Country/Territory").agg({"Total_Deaths":"sum"})
Country_list.sort_values(by='Total_Deaths',ascending=False)[0:25]
```

Country/Territory	Total_Deaths
China	265408106
India	238158165
United States	71197802
Russia	59591155
Indonesia	44046941
Nigeria	43670014
Pakistan	38151878
Brazil	32674112
Japan	31922807
Germany	25559667
Bangladesh	24803502
Ukraine	21245451
Ethiopia	20880668
Democratic Republic of Congo	17446538
United Kingdom	17281600
Italy	16779302
South Africa	15807129
Mexico	15720801
France	15093782
Egypt	14878359
Vietnam	14444650
Philippines	13680040
Myanmar	13486764
Thailand	11358484
Tanzania	11195695

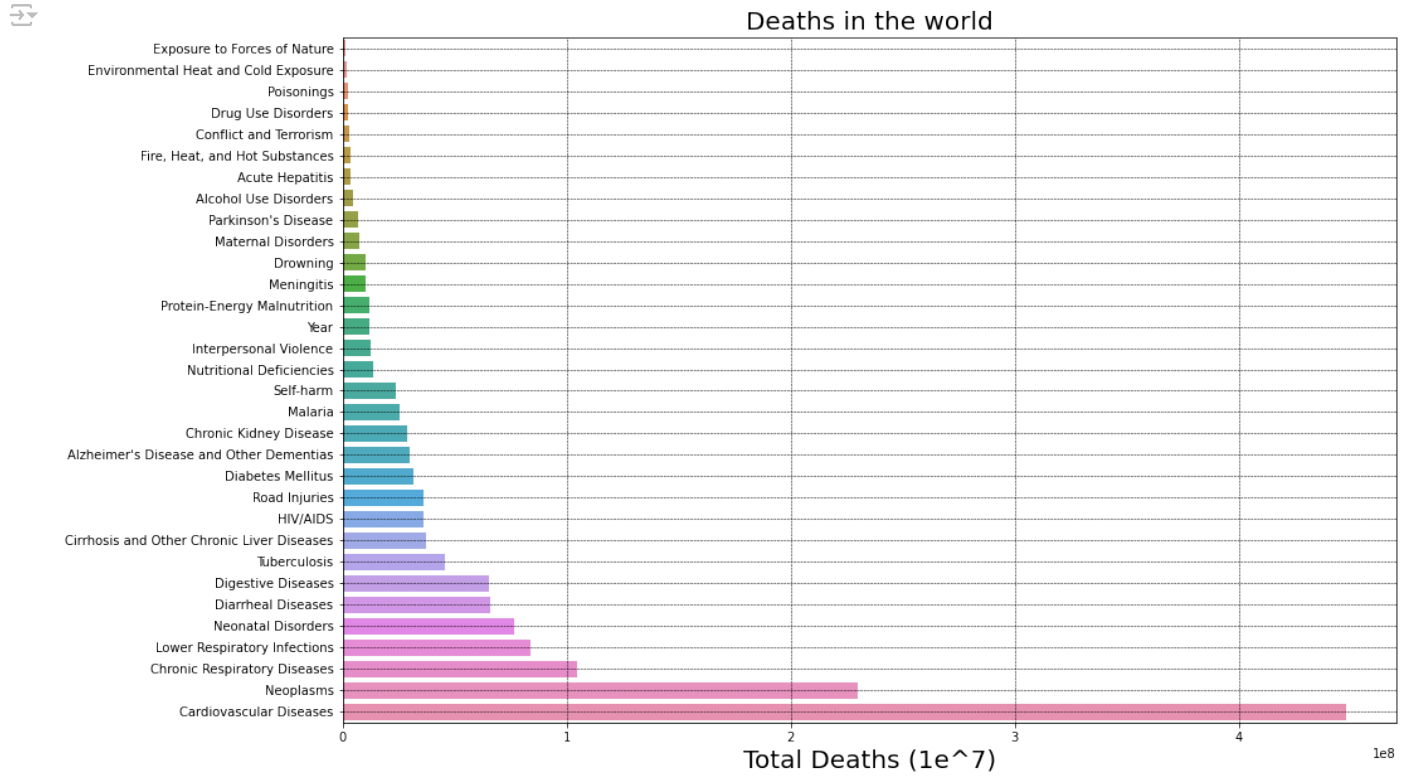
- From the Above table, it clear that, the "top Total_Deaths" belongs to 'China' and 'India' , followed by 'USA', 'Russia' & 'Indonesia'. This because of the fact "China" and "India" are the countries that stand in top 2 in terms of population.

```
# "Number of Deaths" due to Individual Diseases/Causes:

countries= df.sum(axis=0)[1:33].sort_values(ascending=True).index
total_death = df.sum(axis=0)[1:33].sort_values(ascending=True).values

print("\033[1m"+"Causes of death (sorted ascending) "+"\033[0m")
for i in range(0,len(countries)):
    print(countries[i],"=",f'{total_death[i]:,}')

print("\033[1m"+"\nTotal death ="+"\033[0m",
     f'{ np.sum(df.sum(axis=0)[1:].sort_values(ascending=False).values) :,}')
```

```
Causes of death (sorted ascending)
Exposure to Forces of Nature = 1,490,132
Environmental Heat and Cold Exposure = 1,788,851
Poisonings = 2,601,082
Drug Use Disorders = 2,656,121
Conflict and Terrorism = 3,294,053
Fire, Heat, and Hot Substances = 3,602,914
Acute Hepatitis = 3,784,791
Alcohol Use Disorders = 4,819,018
Parkinson's Disease = 7,179,795
Maternal Disorders = 7,727,046
Drowning = 10,301,999
Meningitis = 10,524,572
Protein-Energy Malnutrition = 12,031,885
Year = 12,267,540
Interpersonal Violence = 12,752,839
```

```
Nutritional Deficiencies = 13,792,032
Self-harm = 23,713,931
Malaria = 25,342,676
Chronic Kidney Disease = 28,911,692
Alzheimer's Disease and Other Dementias = 29,768,839
Diabetes Mellitus = 31,448,872
Road Injuries = 36,296,469
HIV/AIDS = 36,364,419
Cirrhosis and Other Chronic Liver Diseases = 37,479,321
Tuberculosis = 45,850,603
Digestive Diseases = 65,638,635
Diarrheal Diseases = 66,235,508
Neonatal Disorders = 76,860,729
Lower Respiratory Infections = 83,770,038
Chronic Respiratory Diseases = 104,605,334
Neoplasms = 229,758,538
Cardiovascular Diseases = 447,741,982

Total death = 2,948,536,972
```

```
fig, ax = plt.subplots(figsize =(15, 10))
sns.barplot(total_death,countries)
plt.grid(color='black', linestyle='--', linewidth=.5)
plt.title("Deaths in the world",fontsize=20)
plt.xlabel("Total Deaths (1e^7)",fontsize=20)
plt.show()
```




```
# Which country suffers from what disease the most?

df["Country/Territory"] = df["Country/Territory"].dropna().apply(str.lower)
cause_by_country = df.groupby("Country/Territory").sum().drop(["Year","Total_Deaths"],axis=1).idxmax(axis=1)
cause_by_country.value_counts()
```

```
Cardiovascular Diseases         156
HIV/AIDS                         17
Neoplasms                        10
Diarrheal Diseases                7
Malaria                           5
Neonatal Disorders                4
Lower Respiratory Infections      4
Conflict and Terrorism            1
dtype: int64
```

- **It is obvious that Cardiovascular Disease dominate death toll all over the world.**

```
# Top Countries in Deaths across 20 years:

Countries_Total_no_of_Deaths_20year_data = df.groupby('Country/Territory').sum()
Countries_Total_no_of_Deaths_20year_data.drop('Year',axis=1,inplace=True)
Countries_Total_no_of_Deaths_20year_data.sort_values(by='Total_Deaths',ascending =False)[:5]
```

Country/Territory	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	Maternal Disorders	HIV/AIDS	Drug Disord
China	480899	5381846	1533092	584236	13418	2873619	776275	243257	433709	626
India	2008944	1707561	756832	3290569	2439244	2110438	1237163	2292449	2454374	168
United States	40032	3302609	661288	133044	0	114752	596818	25206	528417	800
Russia	60519	972305	236367	15906	0	423044	1215179	15028	350679	259
Indonesia	337724	487566	145752	604467	74664	237902	81342	376966	74981	12

5 rows × 32 columns

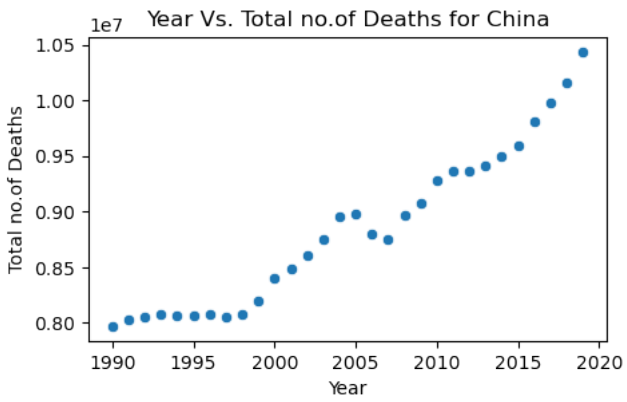
- From the above table, we can see a mixture of developed and developing countries with a sizeable population. So, by doing ample EDA & Statistical analysis of these top 5 countries we can get the affect of each disease with respect to the rest of countries.

1. For China:

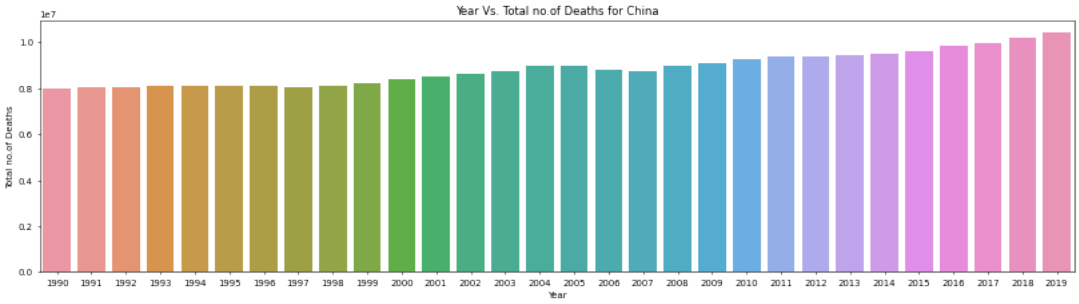
```
# China: "Total_Deaths" against "Year"

China_Total_Deaths = df[df['Country/Territory']=='China'].sort_values(by='Total_Deaths',ascending=False)
China_Total_Deaths[["Country/Territory",'Year','Total_Deaths']]
```

	Country/Territory	Year	Total_Deaths
1139	China	2019	10442561
1138	China	2018	10163943
1137	China	2017	9978653
1119	China	2016	9814213
1118	China	2015	9591222
1117	China	2014	9503904
1116	China	2013	9411928
1114	China	2011	9366974
1115	China	2012	9364587
1113	China	2010	9284664
1112	China	2009	9074833
1135	China	2005	8982702
1111	China	2008	8972670
1134	China	2004	8960684
1136	China	2006	8794396
1110	China	2007	8755201
1133	China	2003	8750361
1132	China	2002	8610956
1131	China	2001	8490928
1130	China	2000	8404159
1129	China	1999	8203474
1128	China	1998	8082343
1126	China	1996	8079610
1123	China	1993	8077792
1124	China	1994	8070458
1125	China	1995	8067172
1122	China	1992	8059068
1127	China	1997	8051031
1121	China	1991	8033487
1120	China	1990	7964132

```
# China: "Total_Deaths" v/s "Year"

plt.figure(figsize=(5,3),dpi=100)
sns.scatterplot(data=China_Total_Deaths, x='Year', y='Total_Deaths')
plt.xlabel("Year")
plt.ylabel("Total no.of Deaths")
plt.title("Year Vs. Total no.of Deaths for China")
plt.show();
```



```
plt.figure(figsize=(20,5),dpi=50)
sns.barplot(data=China_Total_Deaths, x='Year', y='Total_Deaths')
plt.xlabel("Year")
plt.ylabel("Total no.of Deaths")
plt.title("Year Vs. Total no.of Deaths for China")
plt.show();
```



- **There is a clear raise in Total No.of Deaths recorded with each incrementing year for China.**

Now, let's Investigate the Major cause of Deaths in China:

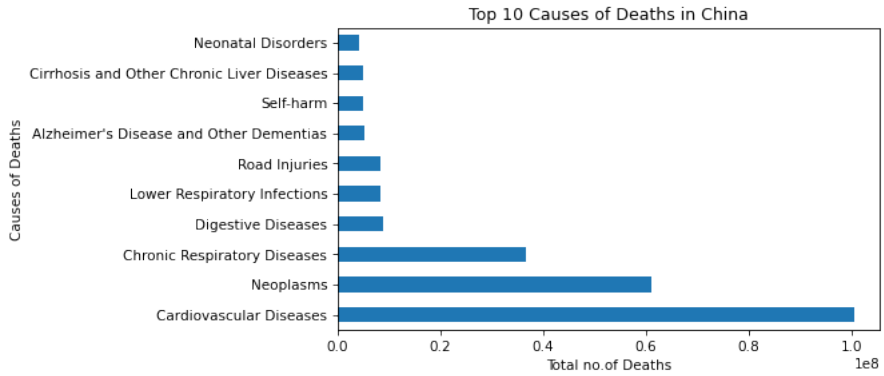
```
# Get maximum deaths causing Disease wrt Year:

China_Total_Deaths.set_index(["Country/Territory", "Total_Deaths"], inplace = True) # making County & Total_Deaths as index.
China_Total_Deaths['Max'] = China_Total_Deaths.idxmax(axis=1) # getting desease for max cause of deaths in China
China_Total_Deaths[['Year','Max']] # displying reqd data only.
```

		Year	Max
Country/Territory	Total_Deaths		
China	**10442561**	2019	Cardiovascular Diseases
	10163943	2018	Cardiovascular Diseases
	9978653	2017	Cardiovascular Diseases
	9814213	2016	Cardiovascular Diseases
	9591222	2015	Cardiovascular Diseases
	9503904	2014	Cardiovascular Diseases
	9411928	2013	Cardiovascular Diseases
	9366974	2011	Cardiovascular Diseases
	9364587	2012	Cardiovascular Diseases
	9284664	2010	Cardiovascular Diseases
	9074833	2009	Cardiovascular Diseases
	8982702	2005	Cardiovascular Diseases
	8972670	2008	Cardiovascular Diseases
	8960684	2004	Cardiovascular Diseases
	8794396	2006	Cardiovascular Diseases
	8755201	2007	Cardiovascular Diseases
	8750361	2003	Cardiovascular Diseases
	8610956	2002	Cardiovascular Diseases
	8490928	2001	Cardiovascular Diseases
	8404159	2000	Cardiovascular Diseases
	8203474	1999	Cardiovascular Diseases
	8082343	1998	Cardiovascular Diseases
	8079610	1996	Cardiovascular Diseases
	8077792	1993	Cardiovascular Diseases
	8070458	1994	Cardiovascular Diseases
	8067172	1995	Cardiovascular Diseases
	8059068	1992	Cardiovascular Diseases
	8051031	1997	Cardiovascular Diseases
	8033487	1991	Cardiovascular Diseases
	7964132	1990	Cardiovascular Diseases

```
# China - Top 10 Causes of Deaths:

china_10 = Countries_Total_no_of_Deaths_20year_data.sort_values(by='Total_Deaths',ascending =False)[:1]
china_10.iloc[0].sort_values(ascending=False)[1:11]
```

```
Cardiovascular Diseases                      100505973
Neoplasms                                     61060527
Chronic Respiratory Diseases                  36676826
Digestive Diseases                             8924906
Lower Respiratory Infections                   8525819
Road Injuries                                  8350399
Alzheimer's Disease and Other Dementias        5381846
Self-harm                                      5078550
Cirrhosis and Other Chronic Liver Diseases     4918899
Neonatal Disorders                             4353666
Name: China, dtype: int64
```

```
plt.figure(figsize=(7,4),dpi=75)
china_10.iloc[0].sort_values(ascending=False)[1:11].plot(kind='barh')
plt.xlabel("Total no.of Deaths")
plt.ylabel("Causes of Deaths")
plt.title("Top 10 Causes of Deaths in China")
plt.show();
```



- **Major Causes of Deaths for China are:**

1. Cardiovascular Diseases
2. Neoplasms
3. Chronic Respiratory Diseases
4. Digestive Diseases
5. Lower Respiratory Infections
6. Road Injuries

```
China_Total_Deaths
```

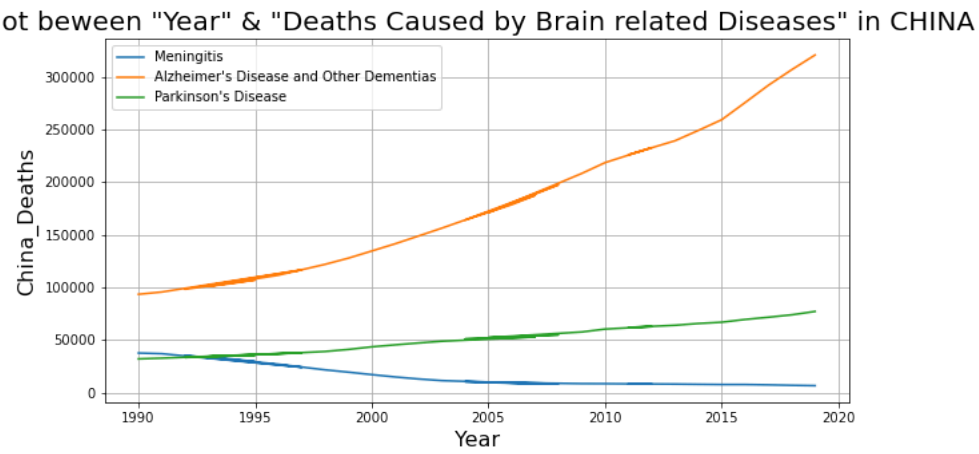
Country/Territory	Total_Deaths	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	Matern Disorde
China	**10442561**	2019	6465	320715	76990	16863	0	56524	11970	15
	10163943	2018	6798	306747	73789	16630	0	57898	12197	17
	9978653	2017	7228	291962	71490	16572	0	59354	12523	20
	9814213	2016	7550	275481	69364	15827	0	61049	13076	25
	9591222	2015	7553	259217	66761	14487	0	61489	13512	23
	9503904	2014	7740	249056	65555	13528	0	62733	14194	23
	9411928	2013	7971	239240	63853	12748	0	63152	15134	27
	9366974	2011	8193	225852	61781	11739	11	66696	17477	33
	9364587	2012	8057	232656	62640	12120	0	65110	16177	30
	9284664	2010	8341	218436	60241	11293	12	70955	18926	37
	9074833	2009	8394	208139	57485	10874	9	72888	20659	42
	8982702	2005	9502	170967	52138	11457	44	85898	26579	65
	8972670	2008	8384	197380	54828	10623	15	73624	21775	45
	8960684	2004	10297	163879	50802	12030	51	90532	27869	72
	8794396	2006	8824	178510	51900	11083	36	79135	25011	57
	8755201	2007	8407	187358	52766	10611	21	76277	23339	50
	8750361	2003	11265	156132	48660	12952	46	94274	27976	78
	8610956	2002	12828	148675	47020	14403	60	100416	27998	92
	8490928	2001	14740	141288	45240	16332	35	105437	29311	111
	8404159	2000	16947	134438	43357	18083	120	111301	31227	123
	8203474	1999	19295	127660	40894	19966	1142	115349	33074	133
	8082343	1998	21485	121766	38955	22088	558	121543	34164	138
	8079610	1996	26788	111481	36703	27043	930	130813	36192	143
	8077792	1993	33396	100822	34129	35265	1327	142119	37525	128
	8070458	1994	31752	103621	34680	32825	1146	138993	38146	141
	8067172	1995	29319	106958	35450	29567	1068	136083	37763	148
	8059068	1992	35086	98135	33287	37980	2061	144949	38179	138
	8051031	1997	23987	116489	37807	24625	946	124856	34972	137
	8033487	1991	36834	95470	32607	42446	2245	150399	39187	155

7964132	1990	37473	93316	31920	42176	1535	153773	40143	174

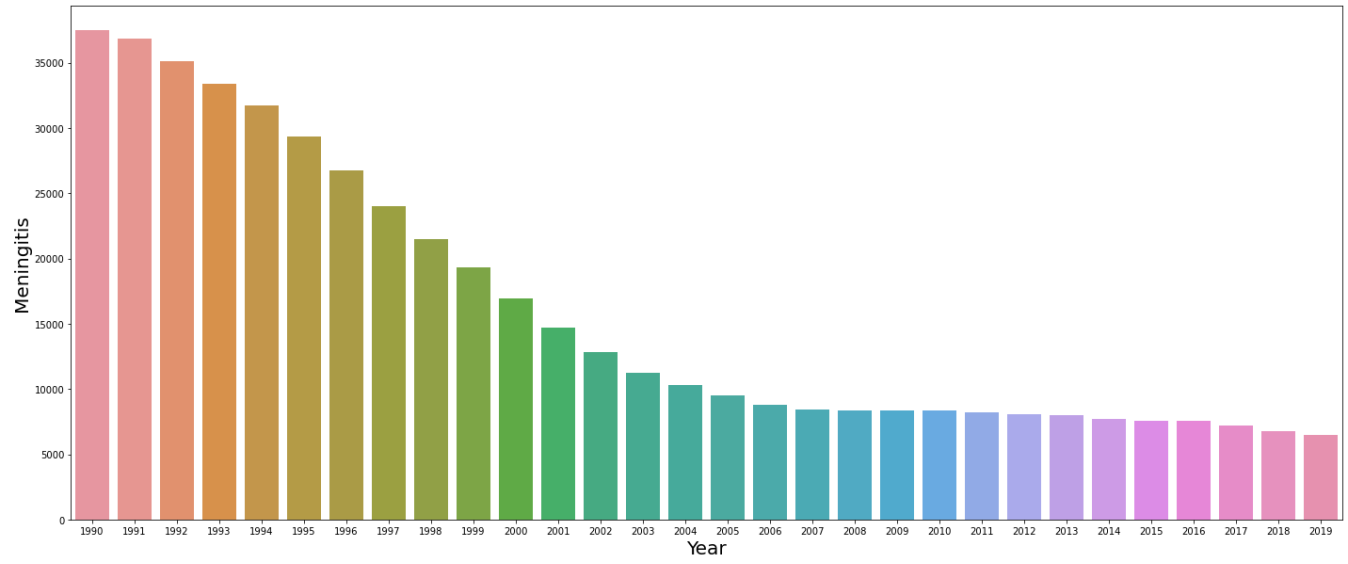
30 rows × 33 columns

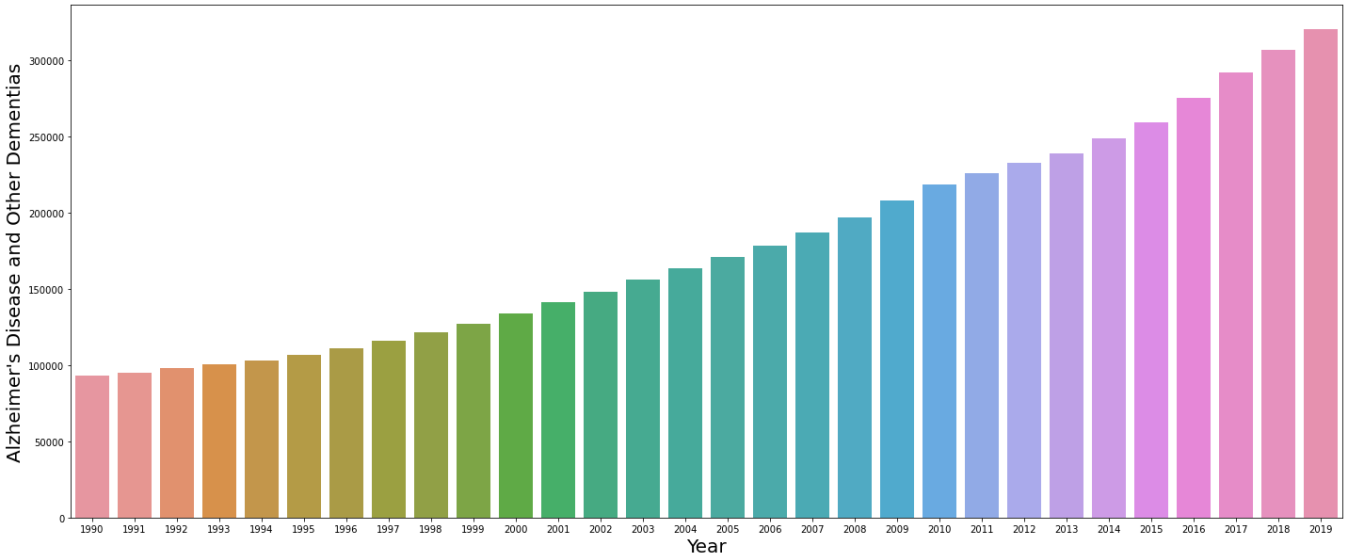
Segregating on the basis of Cause of deaths:

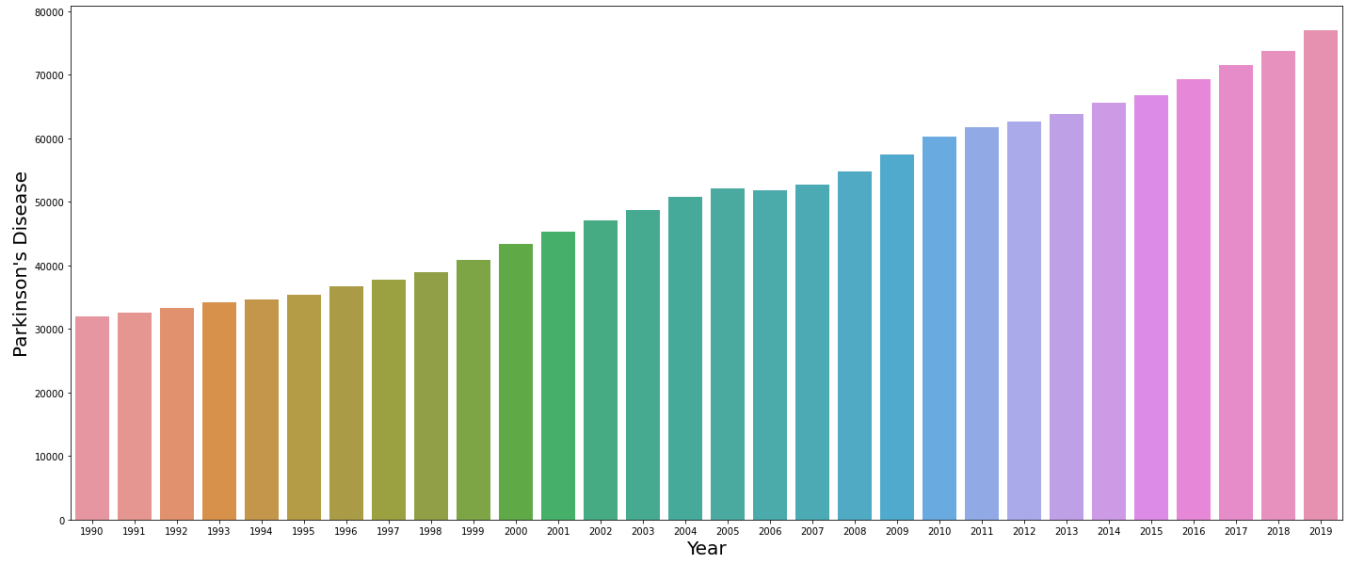
```
# Plot beween 'Year' & "Deaths Caused by Brain related Diseases":
China_Total_Deaths.plot(x='Year', y=['Meningitis',"Alzheimer's Disease and Other Dementias","Parkinson's Disease"],
                        figsize=(10,5), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Brain related Diseases" in CHINA',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('China_Deaths',fontsize =16)
plt.show()
```




```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = China_Total_Deaths[['Meningitis',"Alzheimer's Disease and Other Dementias","Parkinson's Disease"]]
X = China_Total_Deaths['Year']
for col in y:
    if plotnumber<=8:
        plt.subplot(3,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

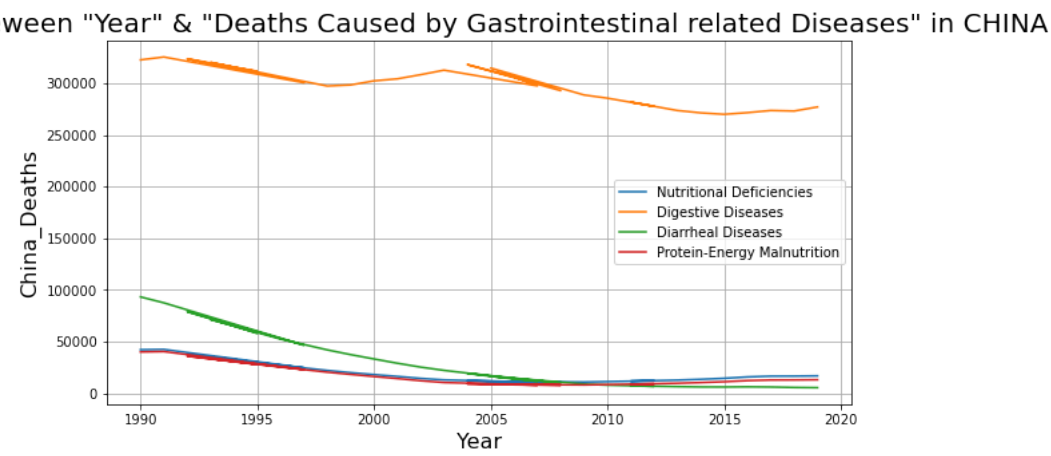

Meningitis
Year


Alzheimer's Disease and Other Dementias
Year

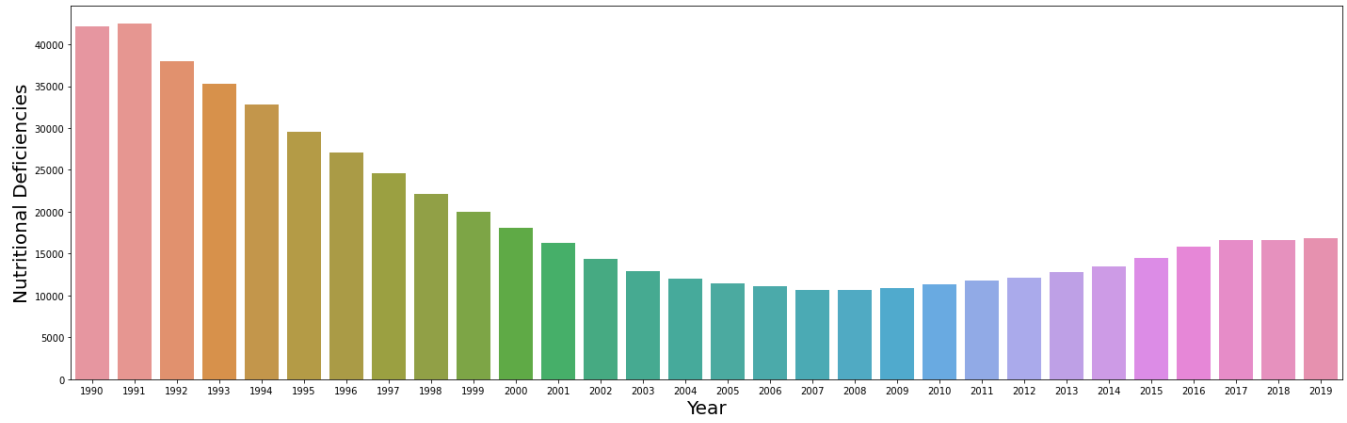

Parkinson's Disease
Year

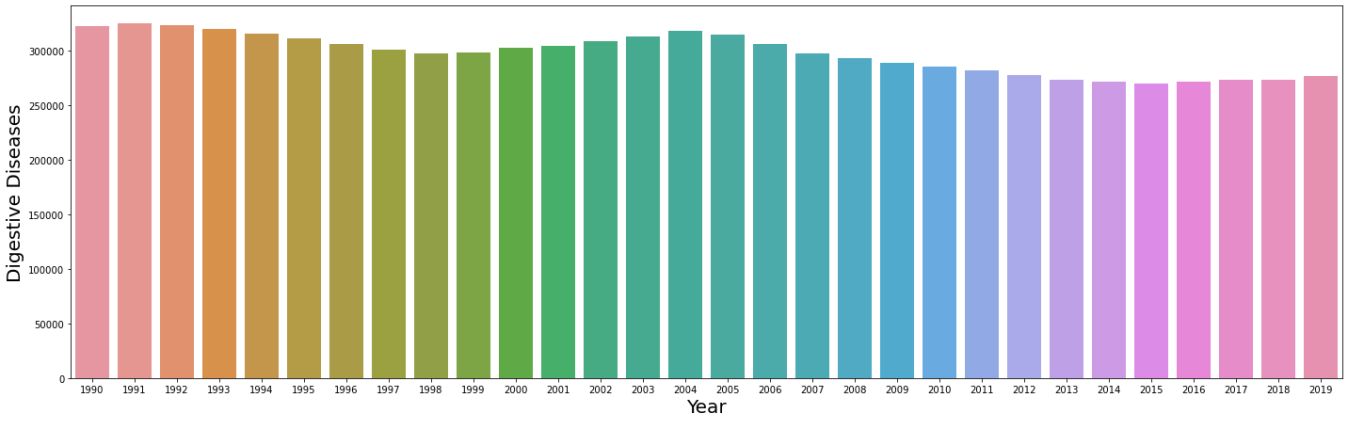
- Deaths in China due to **Alzheimer's Disease and Other Dementias** is becoming more dominant among all Deaths Caused by Brain related Diseases with every incrementing year.

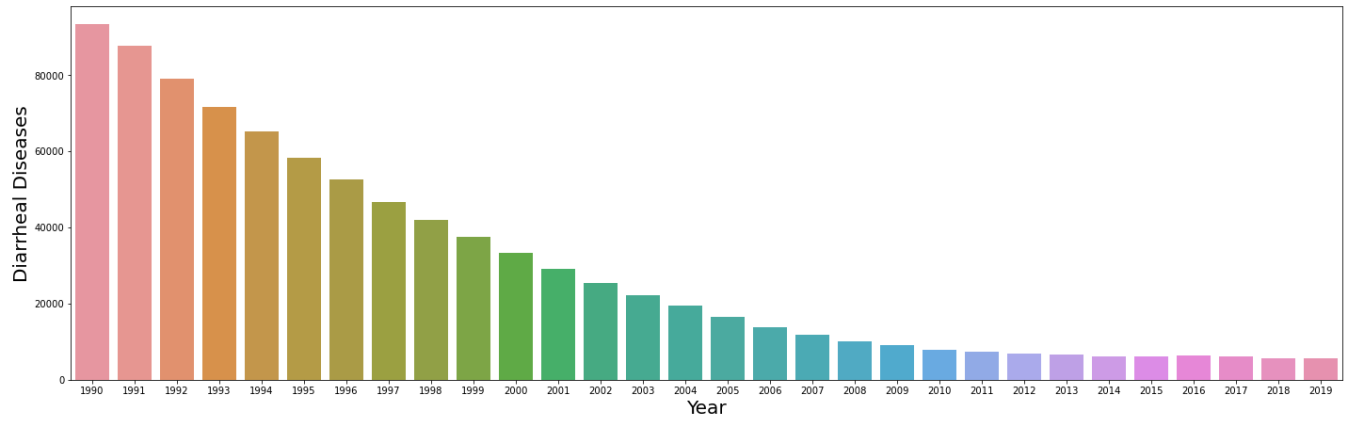
```
# Plot beween 'Year' & "Deaths Caused by Gastrointestinal related Diseases":
China_Total_Deaths.plot(x='Year', y=['Nutritional Deficiencies','Digestive Diseases',
        'Diarrheal Diseases','Protein-Energy Malnutrition'],figsize=(10,5), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Gastrointestinal related Diseases" in CHINA',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('China_Deaths',fontsize =16)
plt.show()
```

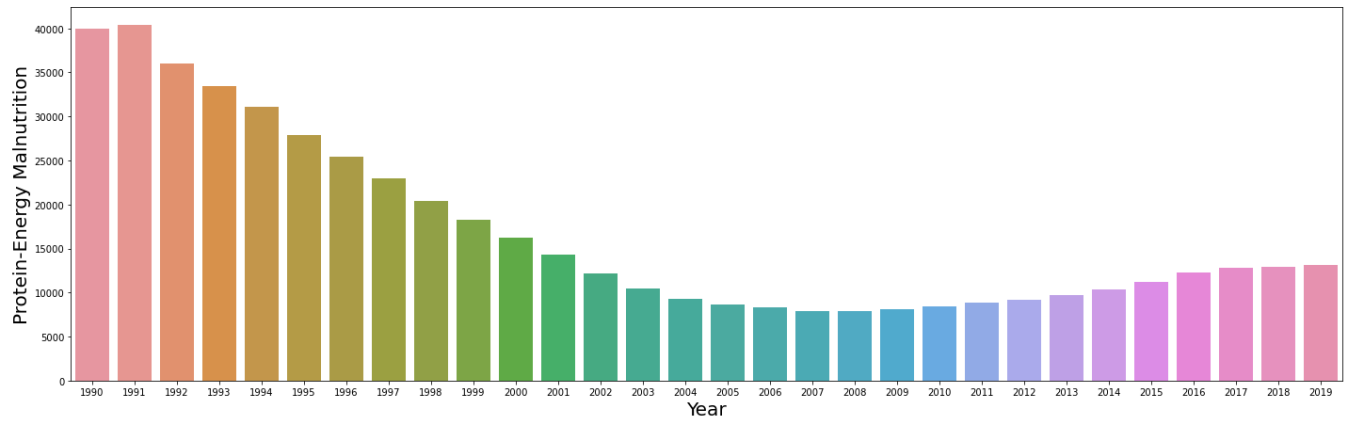



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = China_Total_Deaths[['Nutritional Deficiencies','Digestive Diseases',
        'Diarrheal Diseases','Protein-Energy Malnutrition']]
X = China_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```


Nutritional Deficiencies
Year

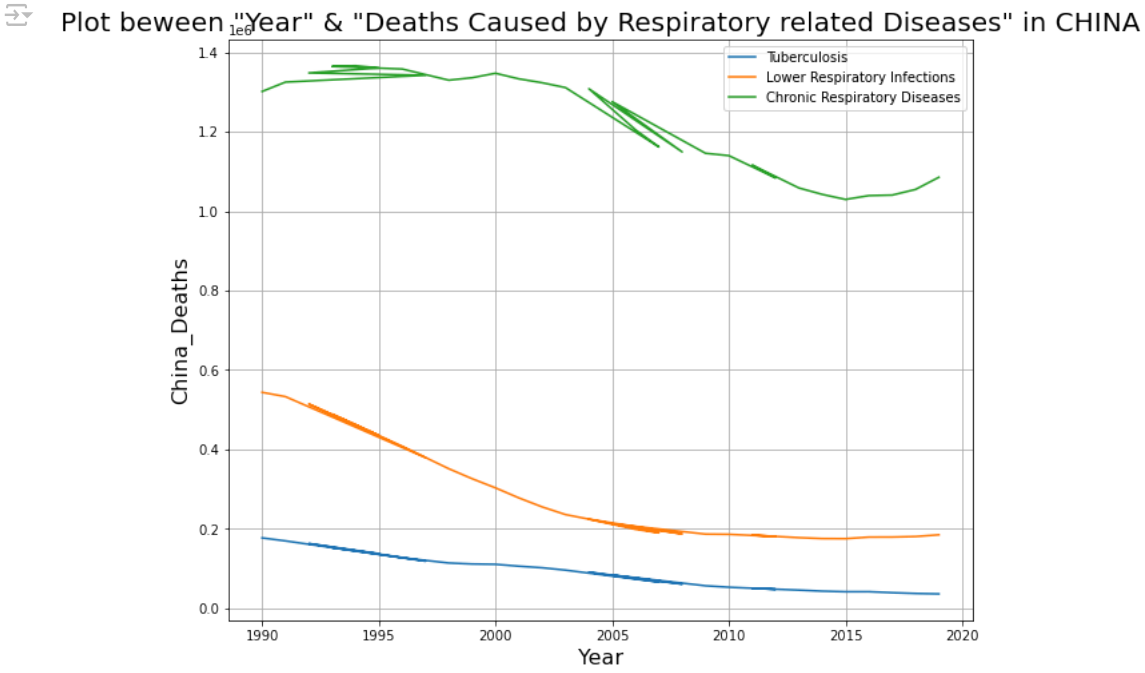

Digestive Diseases
Year


Diarrheal Diseases
Year

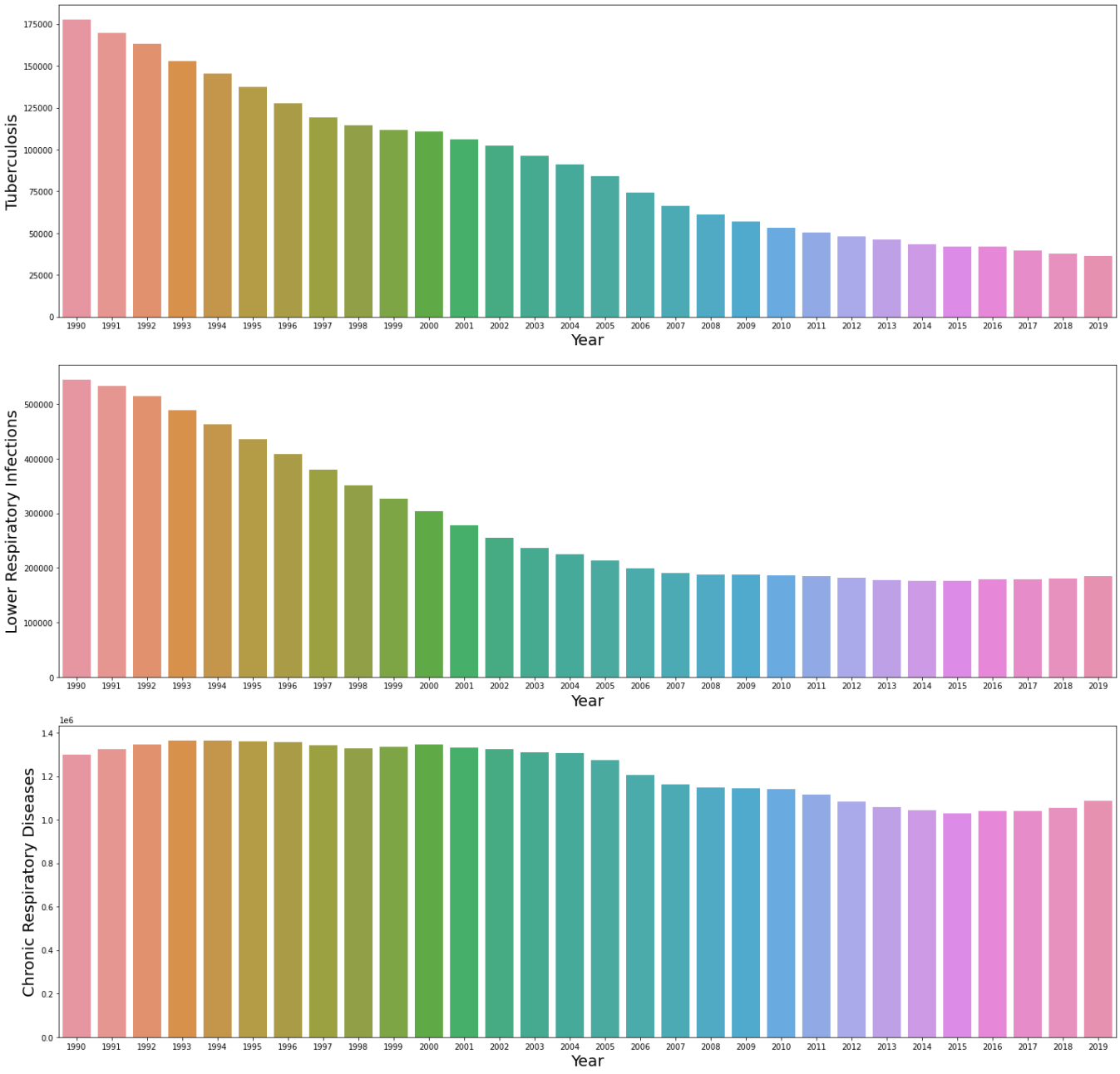

Protein-Energy Malnutrition
Year

- Deaths in China due to **Digestive Diseases** is almost constant & Dominant among all Deaths Caused by Gastrointestinal related Diseases with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by Respiratory related Diseases":
China_Total_Deaths.plot(x='Year', y=['Tuberculosis','Lower Respiratory Infections',
                                     'Chronic Respiratory Diseases'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in CHINA',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('China_Deaths',fontsize =16)
plt.show()
```



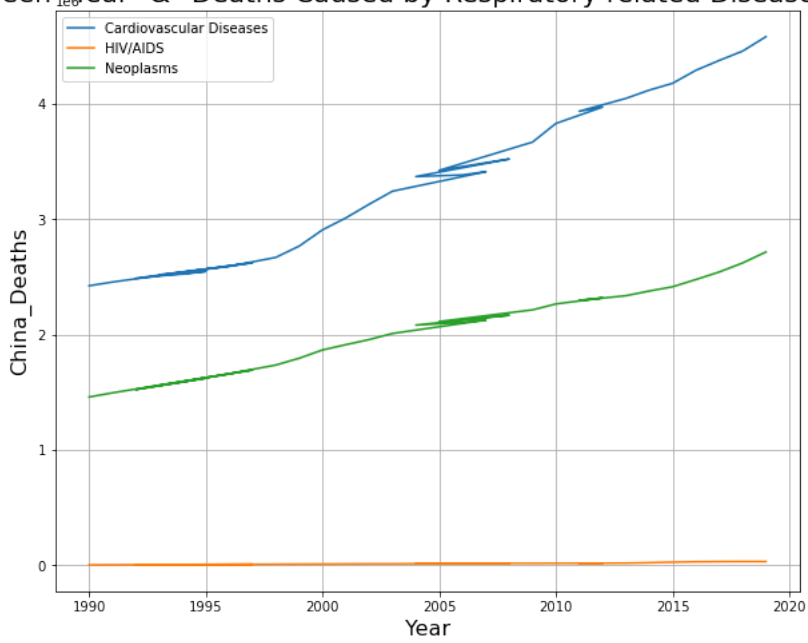

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = China_Total_Deaths[['Tuberculosis','Lower Respiratory Infections','Chronic Respiratory Diseases']]
X = China_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



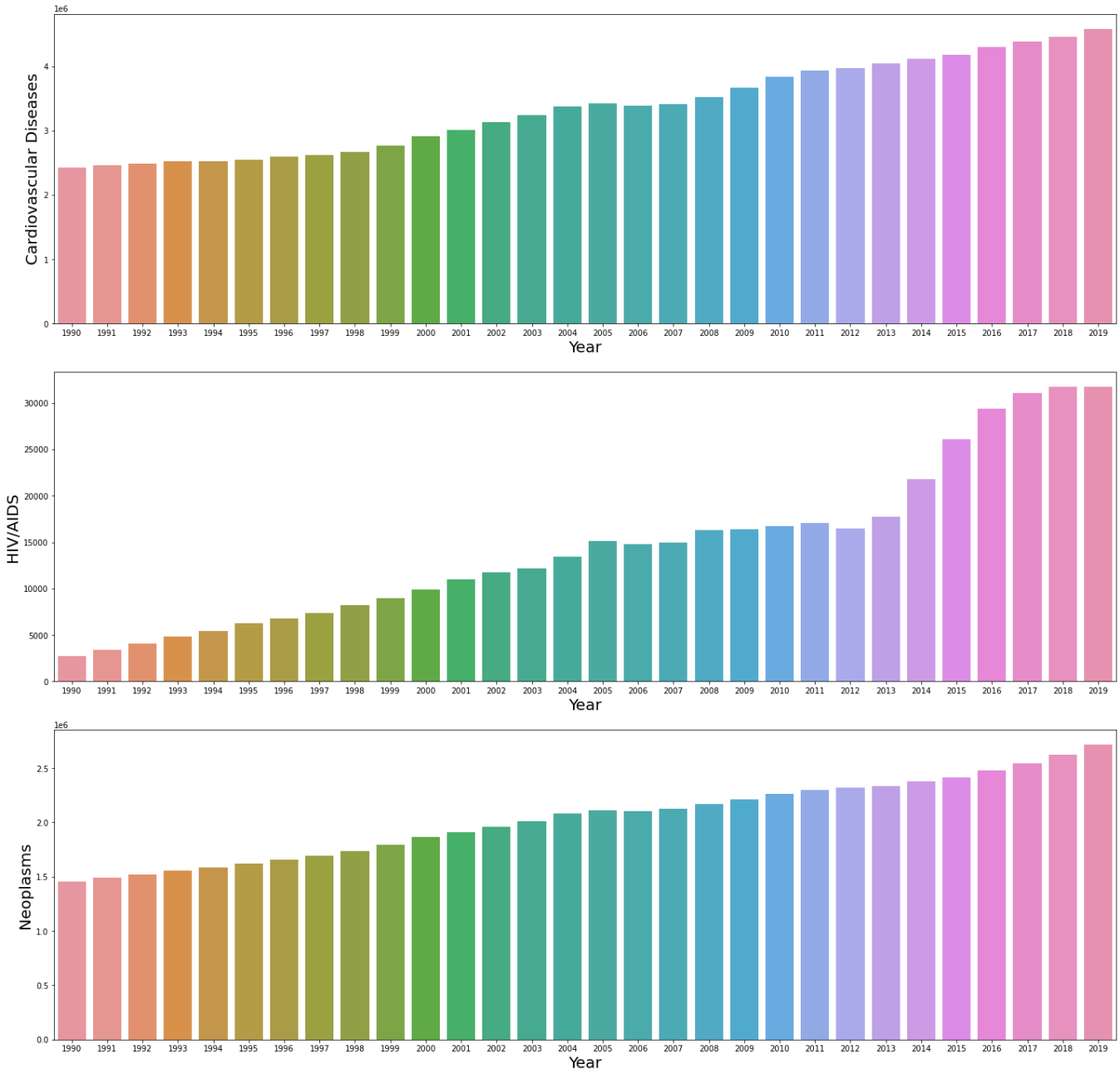
- Deaths in China due to **Chronic Respiratory Diseases** is almost constant(between 10.3lakh/year & 13.5lakh/year) & Dominant among all Deaths Caused by Respiratory related Diseases with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by 'Cardiovascular Diseases','HIV/AIDS' & 'Neoplasms' Diseases":
China_Total_Deaths.plot(x='Year', y=['Cardiovascular Diseases','HIV/AIDS','Neoplasms'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in CHINA',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('China_Deaths',fontsize =16)
plt.show()
```

Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in CHINA



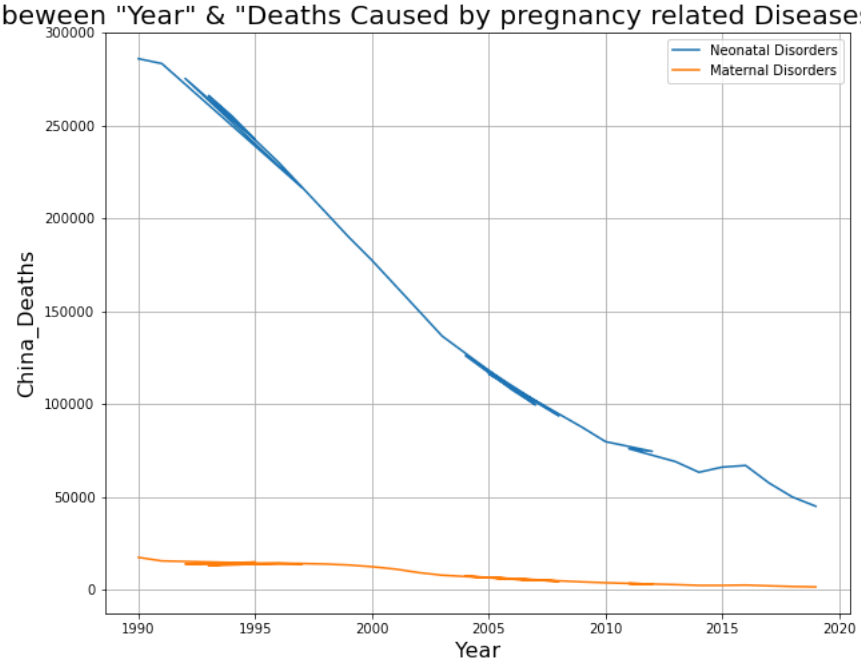

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = China_Total_Deaths[['Cardiovascular Diseases','HIV/AIDS','Neoplasms']]
X = China_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



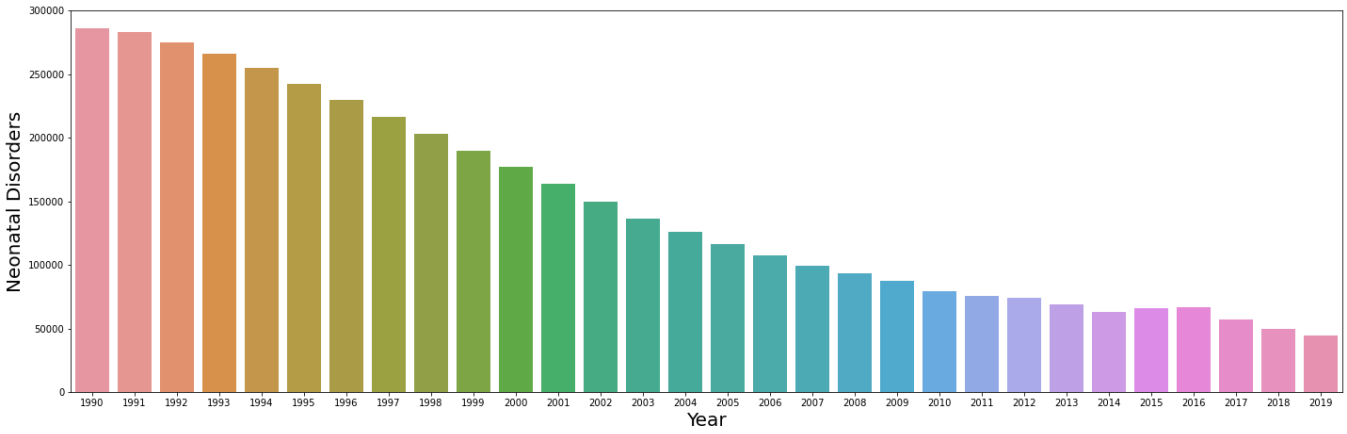
- Deaths in China due to **Cardiovascular Diseases** is most Dominant among all Deaths Caused with every incrementing year.

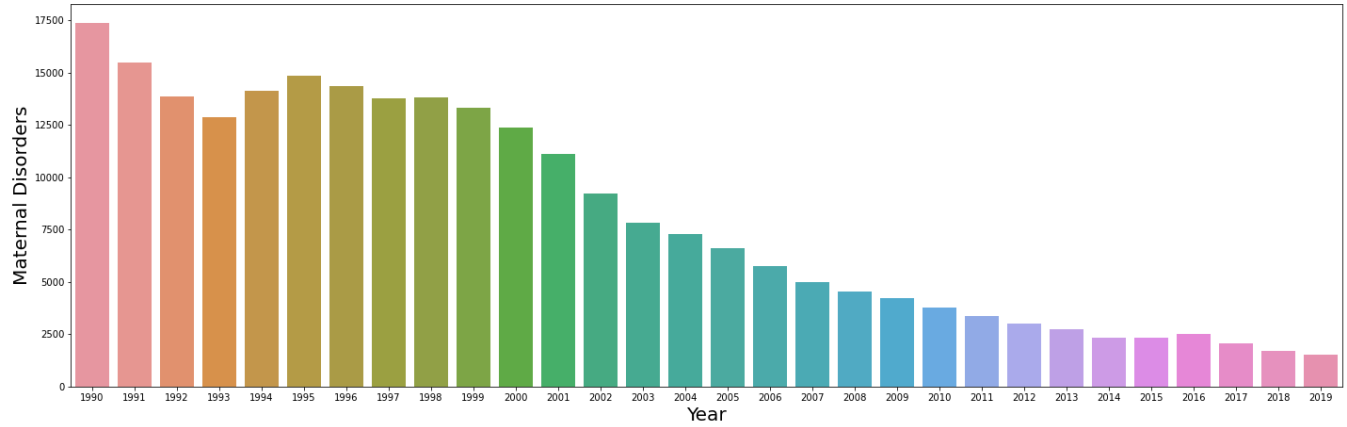
```
# Plot beween 'Year' & "Deaths Caused by pregnancy related Diseases":
China_Total_Deaths.plot(x='Year', y=['Neonatal Disorders','Maternal Disorders'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by pregnancy related Diseases" in CHINA',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('China_Deaths',fontsize =16)
plt.show()
```

Plot beween "Year" & "Deaths Caused by pregnancy related Diseases" in CHINA



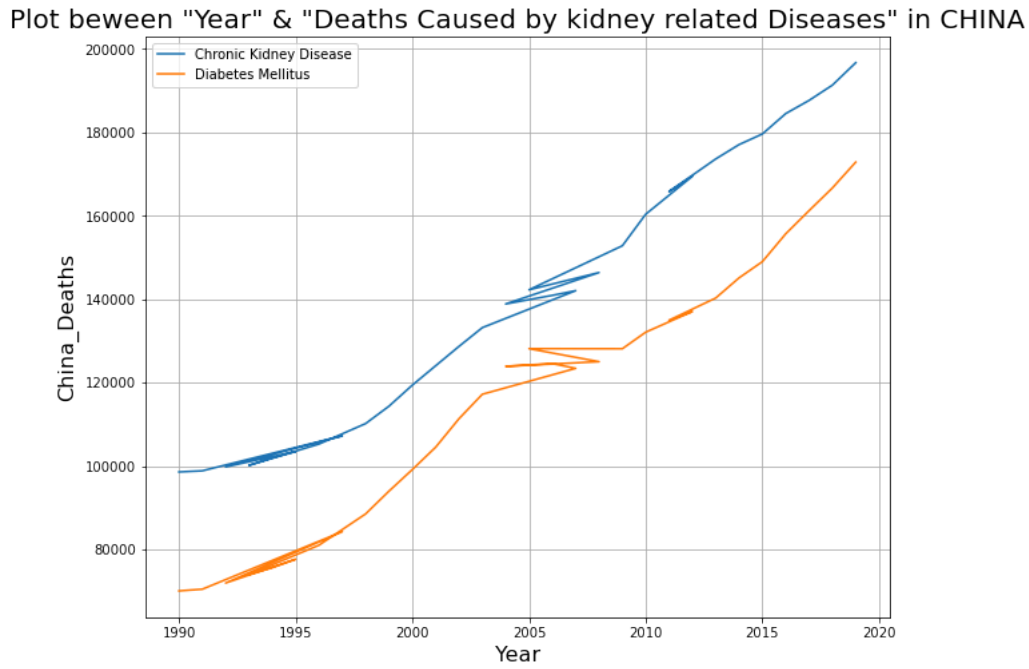

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = China_Total_Deaths[['Neonatal Disorders','Maternal Disorders']]
X = China_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



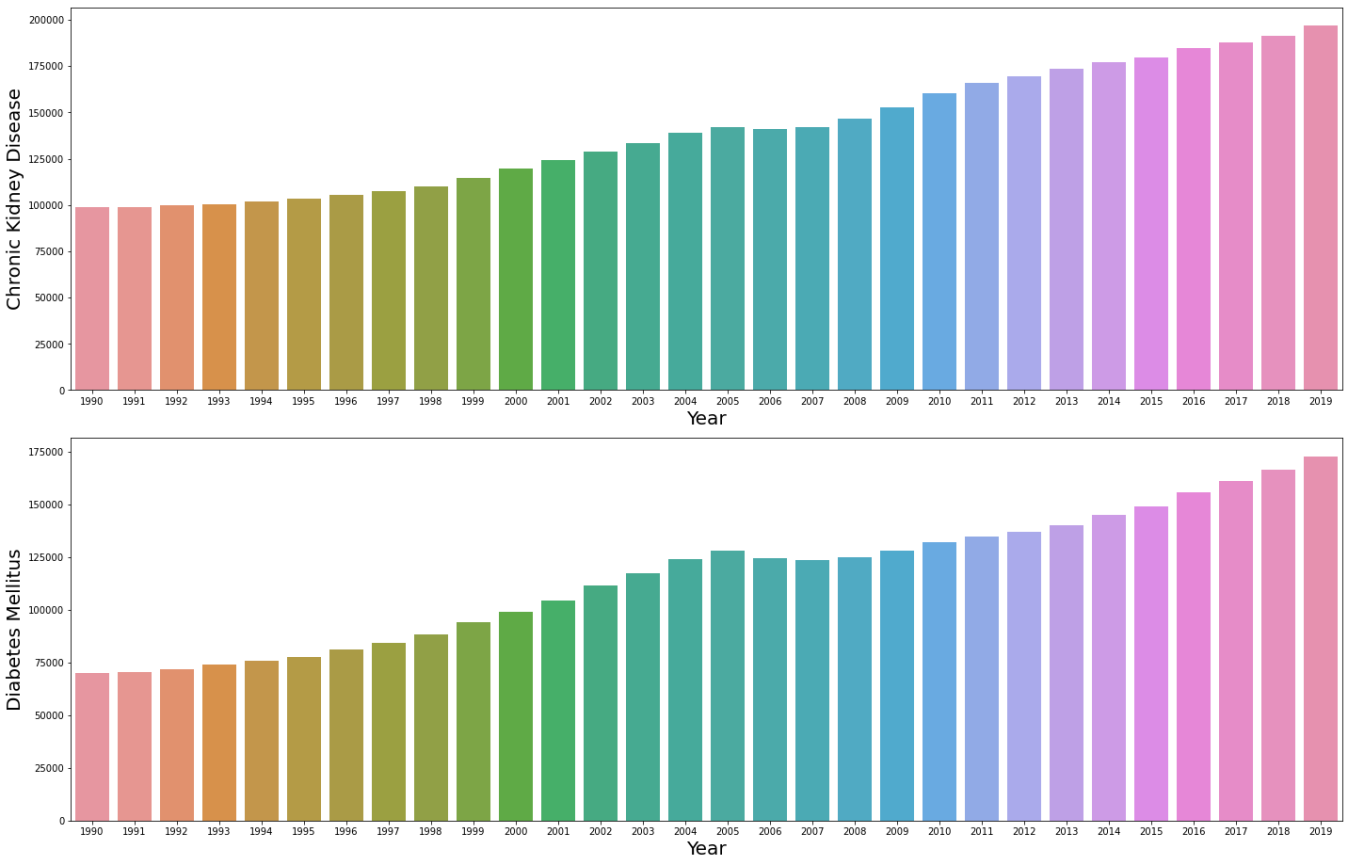


- Deaths in China due to **Pregnancy related** issues is **on decline** with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by kidney related Diseases":
China_Total_Deaths.plot(x='Year', y=['Chronic Kidney Disease','Diabetes Mellitus'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by kidney related Diseases" in CHINA',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('China_Deaths',fontsize =16)
plt.show()
```

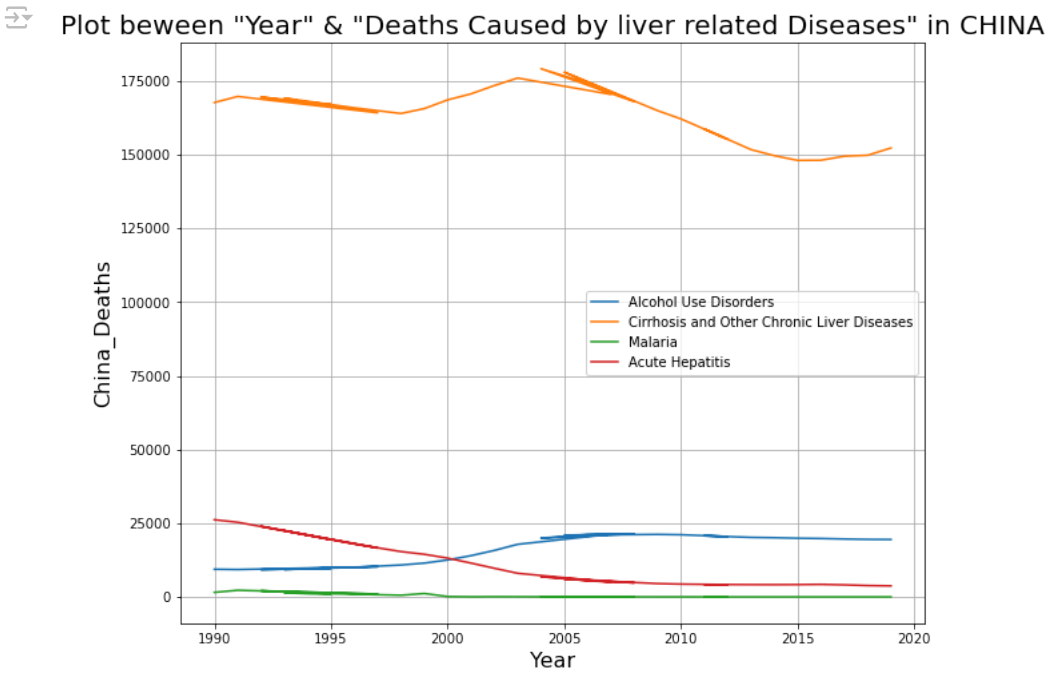


```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = China_Total_Deaths[['Chronic Kidney Disease','Diabetes Mellitus']]
X = China_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

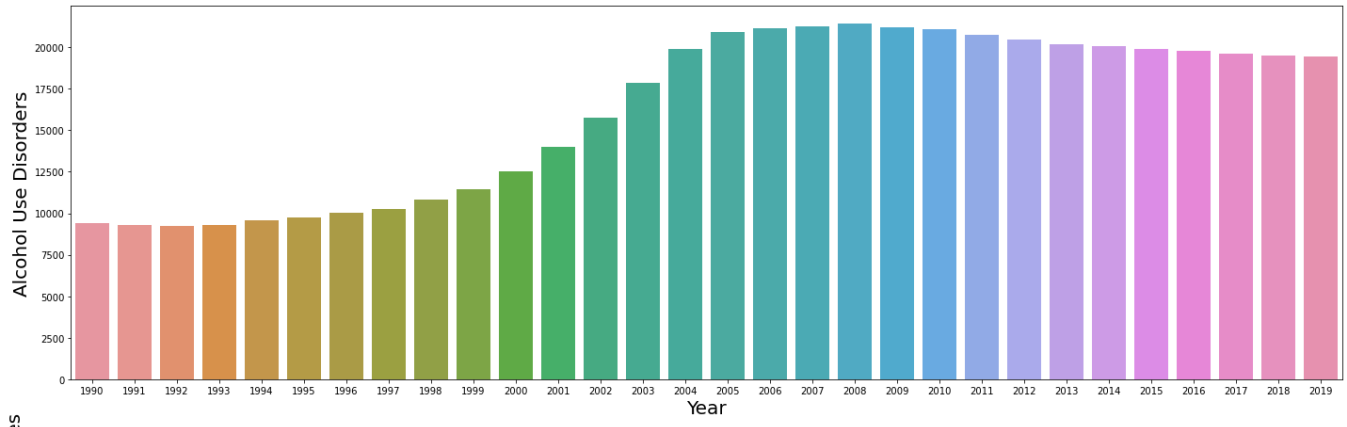


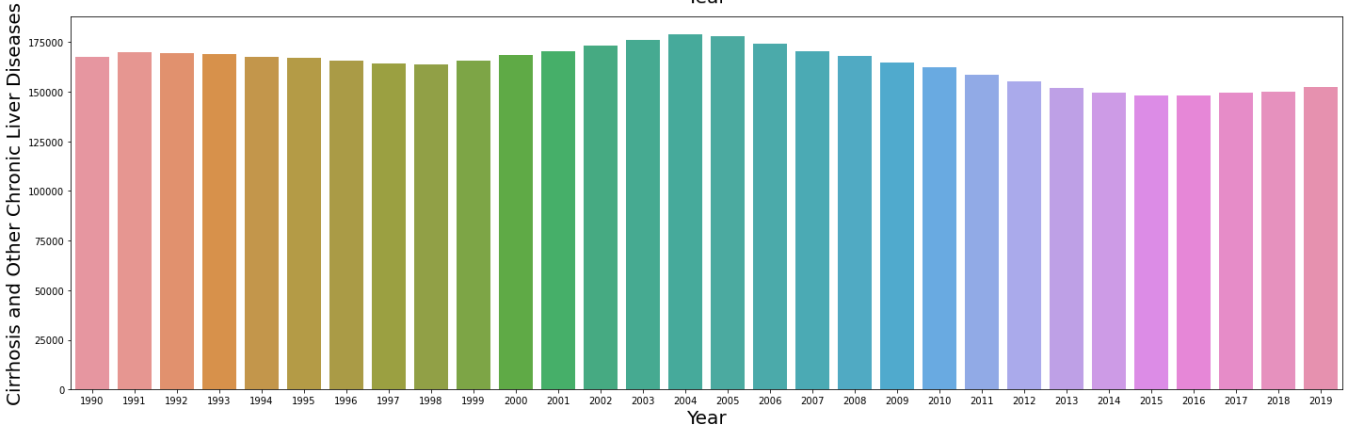
- Deaths in China due to **all kidney related Diseases** is Dominant with every **incrementing** year.

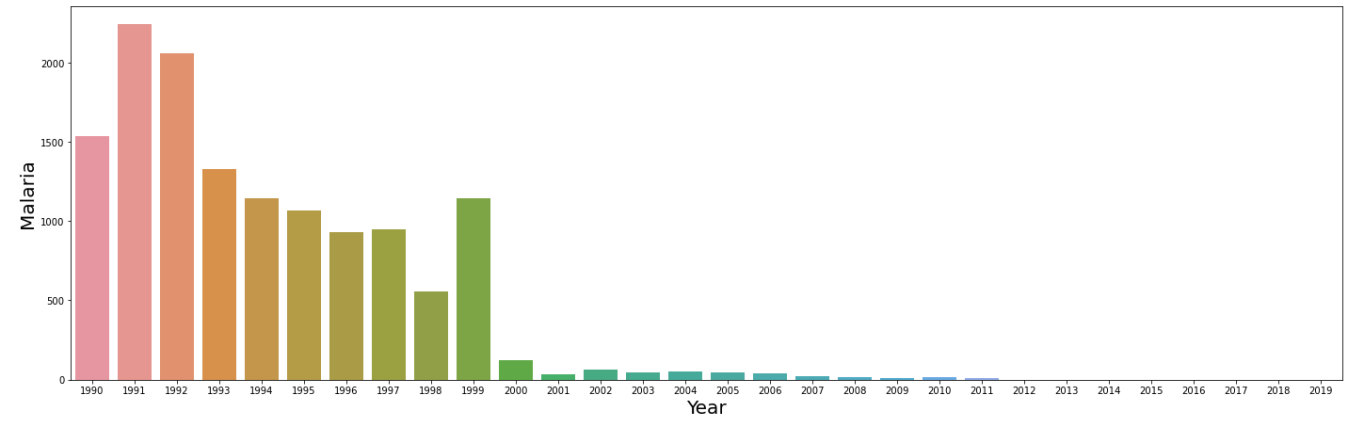
```
# Plot beween 'Year' & "Deaths Caused by liver related Diseases":
China_Total_Deaths.plot(x='Year', y=['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases',
                                     'Malaria','Acute Hepatitis'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by liver related Diseases" in CHINA',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('China_Deaths',fontsize =16)
plt.show()
```

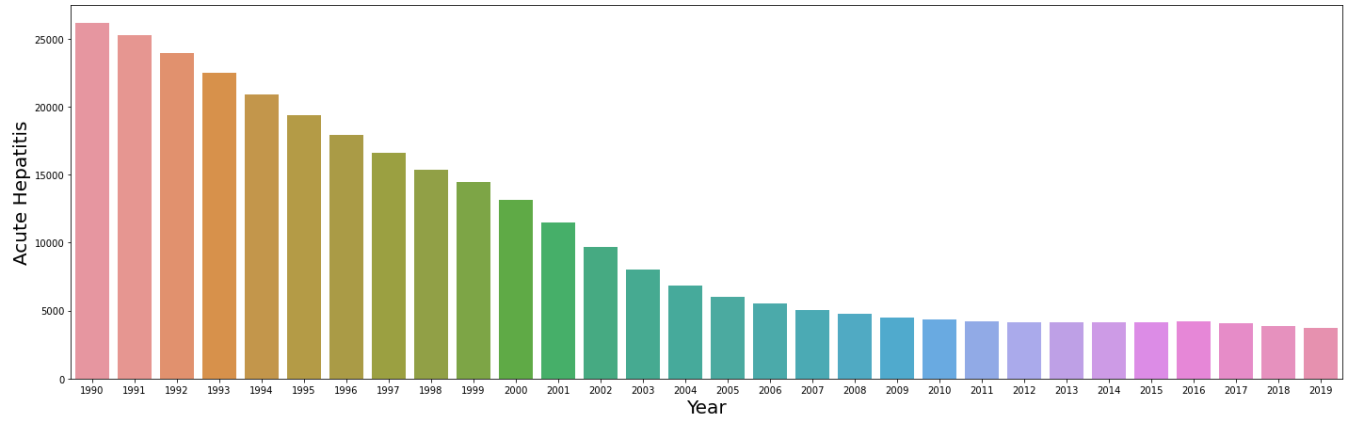



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = China_Total_Deaths[['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases',
                        'Malaria','Acute Hepatitis']]
X = China_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```


Alcohol Use Disorders
Year

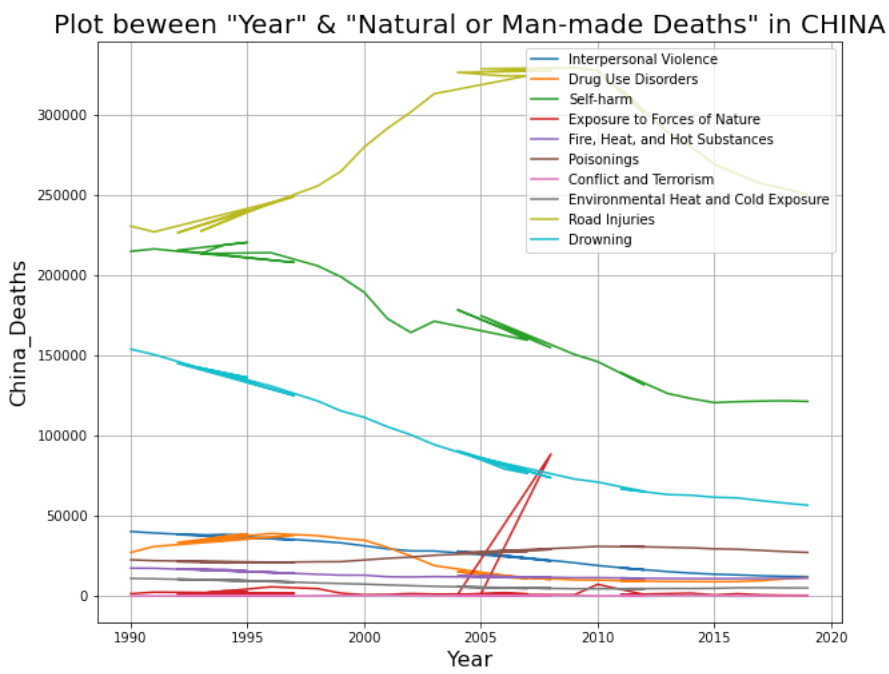

Cirrhosis and Other Chronic Liver Diseases
Year


Malaria
Year

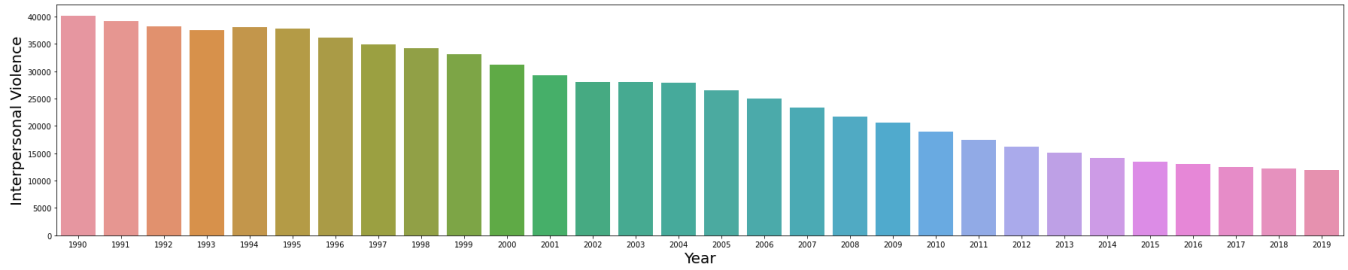

Acute Hepatitis
Year

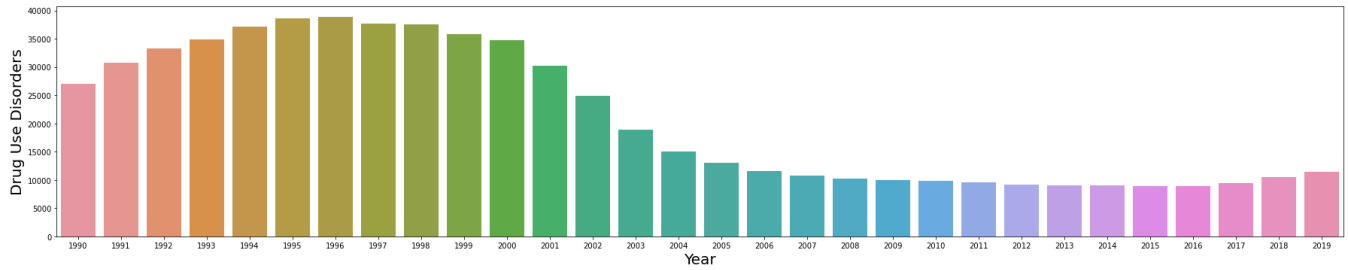
- Deaths in China due to **Cirrhosis and Other Chronic Liver Diseases** is Dominant among all Deaths Caused by **liver** related Diseases with every incrementing year.

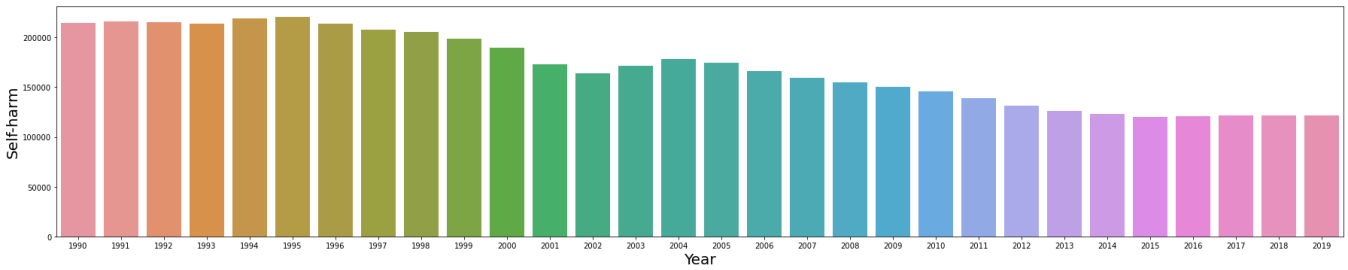
```
# Plot beween 'Year' & "Natural or Man-made Deaths":
China_Total_Deaths.plot(x='Year', y=['Interpersonal Violence','Drug Use Disorders',
                                     'Self-harm', 'Exposure to Forces of Nature',
                                     "Fire, Heat, and Hot Substances",
                                     'Poisonings','Conflict and Terrorism',
                                     'Environmental Heat and Cold Exposure',
                                     'Road Injuries','Drowning'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Natural or Man-made Deaths" in CHINA',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('China_Deaths',fontsize =16)
plt.show()
```

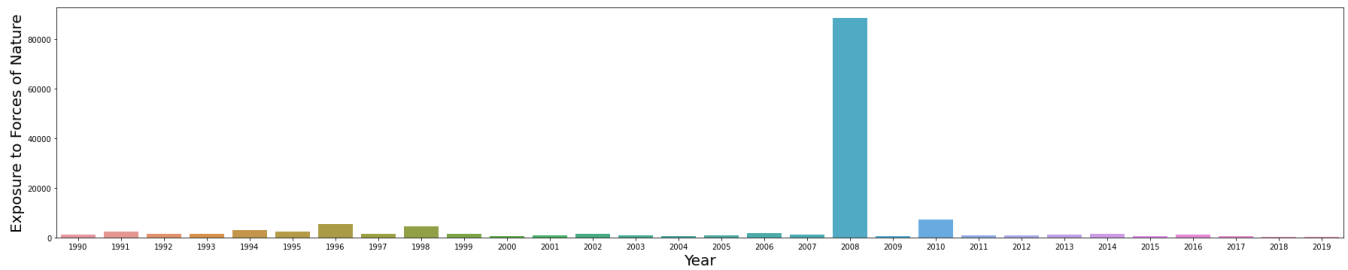


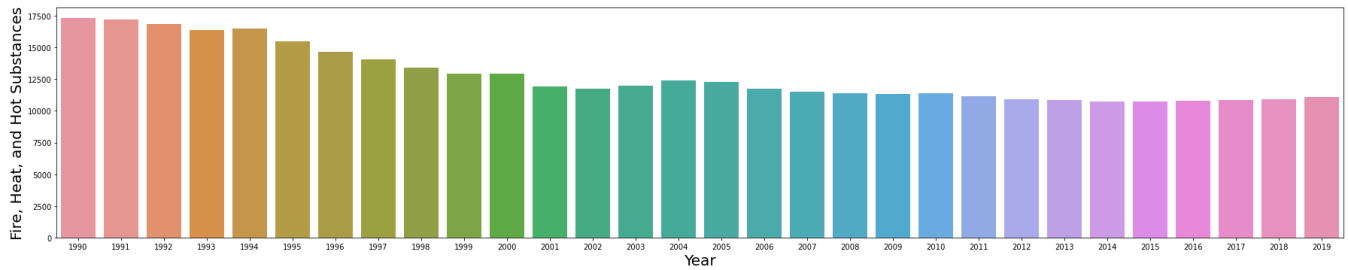

```
plt.figure(figsize=(25,50),facecolor='white')
plotnumber=1
y = China_Total_Deaths[['Interpersonal Violence','Drug Use Disorders',
                                     'Self-harm', 'Exposure to Forces of Nature',
                                     "Fire, Heat, and Hot Substances",
                                     'Poisonings','Conflict and Terrorism',
                                     'Environmental Heat and Cold Exposure',
                                     'Road Injuries','Drowning']]
X = China_Total_Deaths['Year']
for col in y:
    if plotnumber<=10:
        plt.subplot(10,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

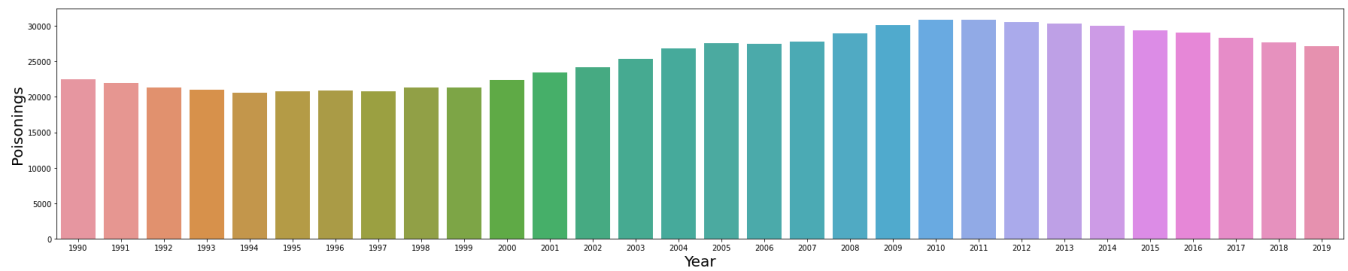

Interpersonal Violence
Year

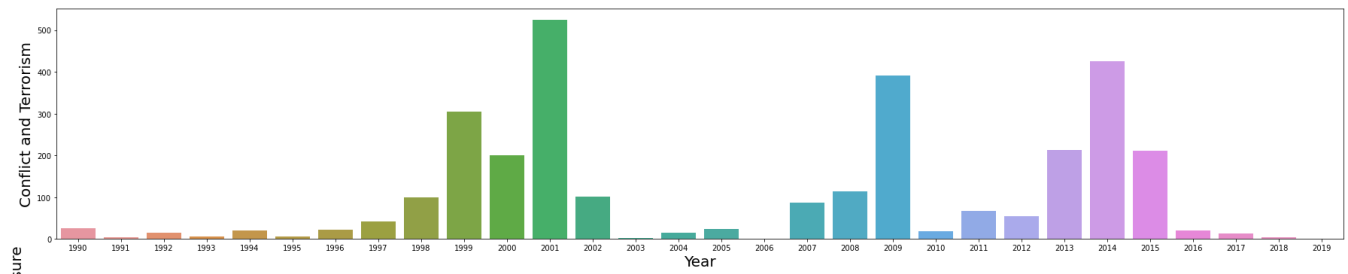

Drug Use Disorders
Year

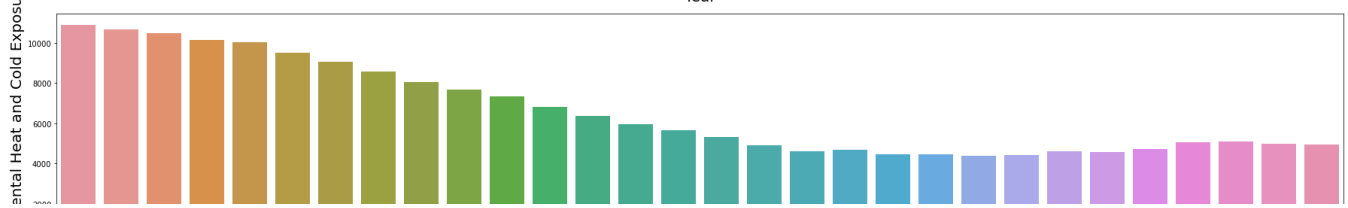

Self-harm
Year

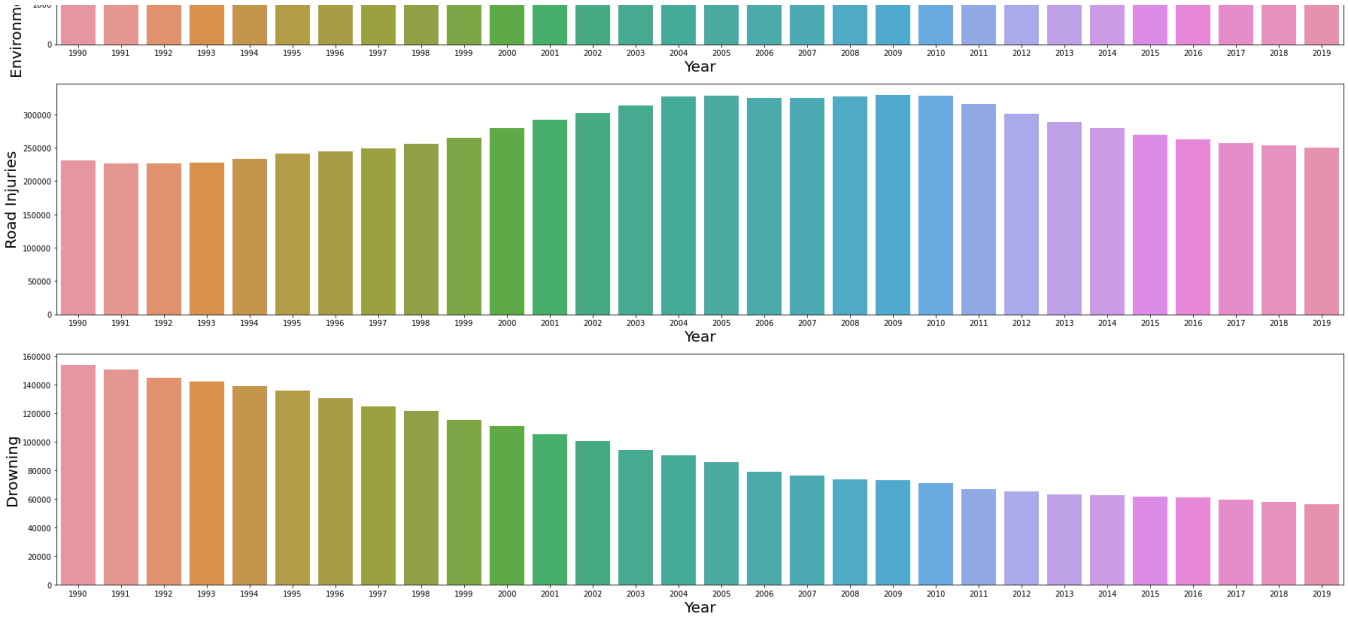

Exposure to Forces of Nature
Year


Fire, Heat, and Hot Substances
Year


Poisonings
Year


Conflict and Terrorism
Year


ental Heat and Cold Exposure



- Deaths in China due to **Natural or Man-made Deaths** are summarised below:

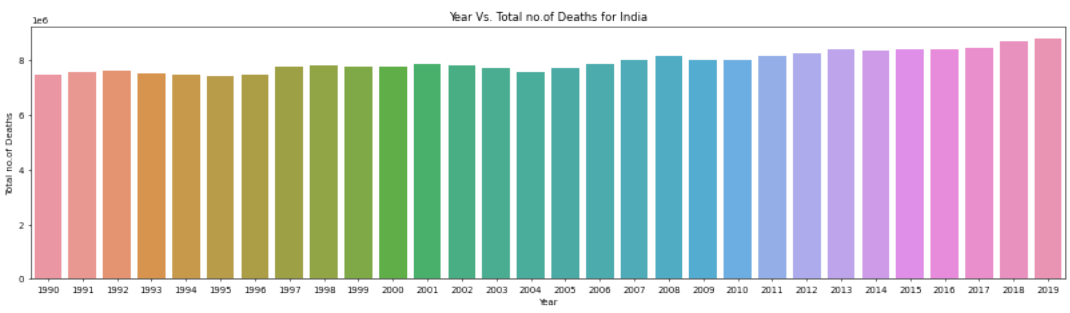
  - Deaths due to 'Interpersonal Violence' is declining with every incrementing year.
  - Deaths due to 'Drug Use Disorders' is declining with every incrementing year.
  - Deaths due to 'Self-harm' is declining with every incrementing year.
  - Deaths due to 'Exposure to Forces of Nature' is hisghest in 2008 due to the **cold spell** mortality in subtropical China.The 2008 cold spell increased mortality by 43.8% compared to non-cold spell days with the highest effects in southern and central China.
  - Deaths due to "Fire, Heat, and Hot Substances" is declining with every incrementing year.
  - Deaths due to 'Poisonings',is increasing with every incrementing year.
  - Deaths due to 'Conflict and Terrorism' is negligible when compared to other causes of deaths.
  - Deaths due to 'Environmental Heat and Cold Exposure' is declining with every incrementing year.
  - Deaths due to 'Road Injuries' was increased between 1990-2010 & declining thereafter.
  - Deaths due to 'Drowning' is declining with every incrementing year.

2. For India:

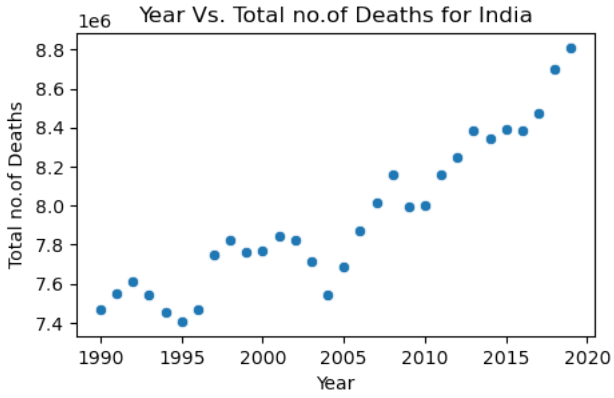
```
# India: "Total_Deaths" against "Year"

India_Total_Deaths = df[df['Country/Territory']=='India'].sort_values(by='Total_Deaths',ascending=False)
India_Total_Deaths[["Country/Territory",'Year','Total_Deaths']]
```

	Country/Territory	Year	Total_Deaths
2459	India	2019	8812747
2458	India	2018	8698039
2457	India	2017	8471931
2435	India	2015	8394566
2456	India	2016	8383986
2434	India	2013	8382673
2455	India	2014	8343254
2433	India	2012	8245684
2454	India	2011	8160776
2431	India	2008	8159092
2430	India	2007	8014215
2453	India	2010	8001850
2432	India	2009	7997048
2452	India	2006	7874588
2447	India	2001	7843434
2444	India	1998	7827043
2448	India	2002	7823400
2446	India	2000	7770778
2445	India	1999	7761464
2443	India	1997	7747248
2449	India	2003	7711081
2451	India	2005	7690096
2438	India	1992	7613105
2437	India	1991	7552853
2450	India	2004	7544432
2439	India	1993	7540003
2442	India	1996	7467868
2436	India	1990	7466270
2440	India	1994	7455322
2441	India	1995	7403319

```
plt.figure(figsize=(20,5),dpi=50)
sns.barplot(data=India_Total_Deaths, x='Year', y='Total_Deaths')
plt.xlabel("Year")
plt.ylabel("Total no.of Deaths")
plt.title("Year Vs. Total no.of Deaths for India")
plt.show();
```




```
# India: "Total_Deaths" v/s "Year"

plt.figure(figsize=(5,3),dpi=100)
sns.scatterplot(data=India_Total_Deaths, x='Year', y='Total_Deaths')
plt.xlabel("Year")
plt.ylabel("Total no.of Deaths")
plt.title("Year Vs. Total no.of Deaths for India")
plt.show();
```



- **We can see slight dips in the total death counts, in India.**

Now, let's Investigate the Major cause of Deaths in India:

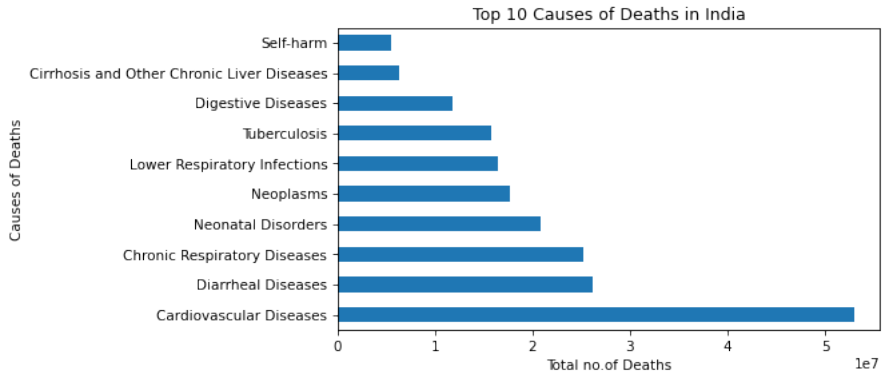
```
# Get maximum deaths causing Disease wrt Year:

India_Total_Deaths.set_index(["Country/Territory", "Total_Deaths"], inplace = True) # making County & Total_Deaths as index.
India_Total_Deaths['Max'] = India_Total_Deaths.idxmax(axis=1) # getting desease for max cause of deaths in China
India_Total_Deaths[['Year','Max']] # displying reqd data only.
```

		Year	Max
Country/Territory	Total_Deaths		
India	**8812747**	2019	Cardiovascular Diseases
	8698039	2018	Cardiovascular Diseases
	8471931	2017	Cardiovascular Diseases
	8394566	2015	Cardiovascular Diseases
	8383986	2016	Cardiovascular Diseases
	8382673	2013	Cardiovascular Diseases
	8343254	2014	Cardiovascular Diseases
	8245684	2012	Cardiovascular Diseases
	8160776	2011	Cardiovascular Diseases
	8159092	2008	Cardiovascular Diseases
	8014215	2007	Cardiovascular Diseases
	8001850	2010	Cardiovascular Diseases
	7997048	2009	Cardiovascular Diseases
	7874588	2006	Cardiovascular Diseases
	7843434	2001	Cardiovascular Diseases
	7827043	1998	Cardiovascular Diseases
	7823400	2002	Cardiovascular Diseases
	7770778	2000	Cardiovascular Diseases
	7761464	1999	Cardiovascular Diseases
	7747248	1997	Cardiovascular Diseases
	7711081	2003	Cardiovascular Diseases
	7690096	2005	Cardiovascular Diseases
	7613105	1992	Cardiovascular Diseases
	7552853	1991	Cardiovascular Diseases
	7544432	2004	Cardiovascular Diseases
	7540003	1993	Cardiovascular Diseases
	7467868	1996	Cardiovascular Diseases
	7466270	1990	Cardiovascular Diseases
	7455322	1994	Cardiovascular Diseases
	7403319	1995	Cardiovascular Diseases

```
# India - Top 10 Causes of Deaths:

India_10 = Countries_Total_no_of_Deaths_20year_data.sort_values(by='Total_Deaths',ascending =False)
India_10.iloc[1].sort_values(ascending=False)[1:11]
```

```
Cardiovascular Diseases                      52994710
Diarrheal Diseases                           26243547
Chronic Respiratory Diseases                 25232974
Neonatal Disorders                           20911570
Neoplasms                                    17762703
Lower Respiratory Infections                 16419404
Tuberculosis                                 15820922
Digestive Diseases                           11804380
Cirrhosis and Other Chronic Liver Diseases    6294910
Self-harm                                     5543395
Name: India, dtype: int64
```

```
plt.figure(figsize=(7,4),dpi=75)
India_10.iloc[1].sort_values(ascending=False)[1:11].plot(kind='barh')
plt.xlabel("Total no.of Deaths")
plt.ylabel("Causes of Deaths")
plt.title("Top 10 Causes of Deaths in India")
plt.show();
```



- **Major Causes of Deaths for India are:**

1. Cardiovascular Diseases
2. Diarrheal Diseases
3. Chronic Respiratory Diseases
4. Neonatal Disorders
5. Neoplasms
6. Lower Respiratory Infections
7. Tuberculosis
8. Digestive Diseases
9. Cirrhosis and Other Chronic Liver Diseases
10. Self-harm

```
India_Total_Deaths
```

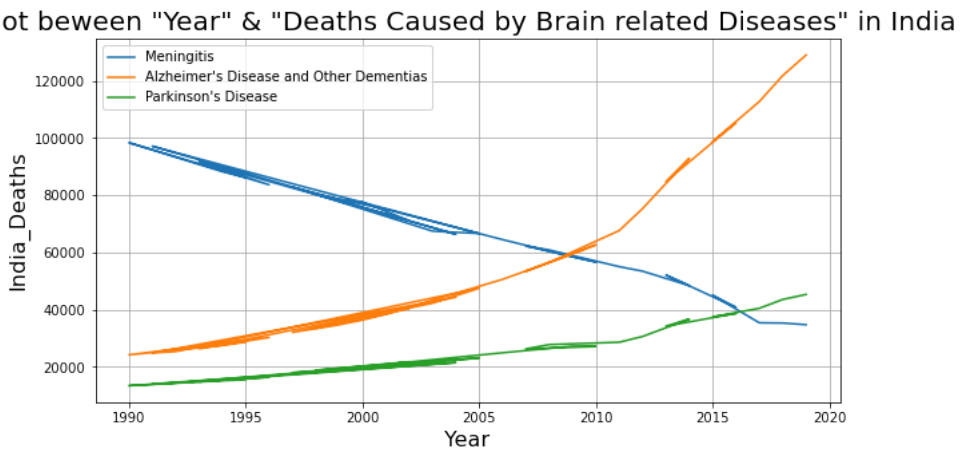
Country/Territory	Total_Deaths	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	Matern Disorde
India	**8812747**	2019	34736	129011	45305	26868	33372	54046	40954	418
	8698039	2018	35306	121785	43550	28440	34030	53720	40402	417
	8471931	2017	35394	112741	40466	29994	29355	51599	39452	403
	8394566	2015	45025	98764	37450	37961	50605	56352	38837	441
	8383986	2016	40773	105315	38673	33198	34592	53423	39190	418
	8382673	2013	52065	84796	34226	47070	32921	63122	39234	530
	8343254	2014	48398	92895	36704	42065	45334	57974	38180	472
	8245684	2012	53433	75330	30619	51001	42686	66282	39520	550
	8160776	2011	55021	67634	28597	52412	63142	66452	42195	598
	8159092	2008	60843	56488	27779	63747	62028	67494	44464	666
	8014215	2007	62224	53407	26269	67671	68053	67595	44087	693
	8001850	2010	56658	62664	27110	55050	60236	66156	42607	625
	7997048	2009	58449	59418	27243	58734	58617	65811	43077	637
	7874588	2006	64479	50576	24911	71995	71129	69017	44173	740
	7843434	2001	75064	38442	20880	115737	73843	74611	43580	945
	7827043	1998	80366	33498	18778	154443	108147	79681	41945	985
	7823400	2002	71130	40483	21757	103615	80138	70212	44600	885
	7770778	2000	77606	36351	19608	126904	70076	75883	43405	978
	7761464	1999	78524	34704	18778	138994	84608	77952	42901	987
	7747248	1997	82891	32050	17949	169334	118550	79082	40703	986
	7711081	2003	67440	42257	21911	91452	78316	67185	43147	815
	7690096	2005	66618	47478	23093	74847	71265	68974	44211	771
	7613105	1992	94891	25353	14394	237280	151759	86073	38859	1079
	7552853	1991	97134	24862	14011	253431	159473	88014	37806	1067
	7544432	2004	66330	44336	21480	81994	77383	67265	43323	778
	7540003	1993	91783	26306	14824	222390	142187	85025	39355	1036
	7467868	1996	83642	30275	16385	181431	120014	77890	39721	957
	7466270	1990	98358	24195	13419	268223	162369	88688	38720	1061
	7455322	1994	88221	27469	15174	208659	131169	83638	39070	1004

7403319	1995	86142	28678	15489	195629	123847	81222	39445	969

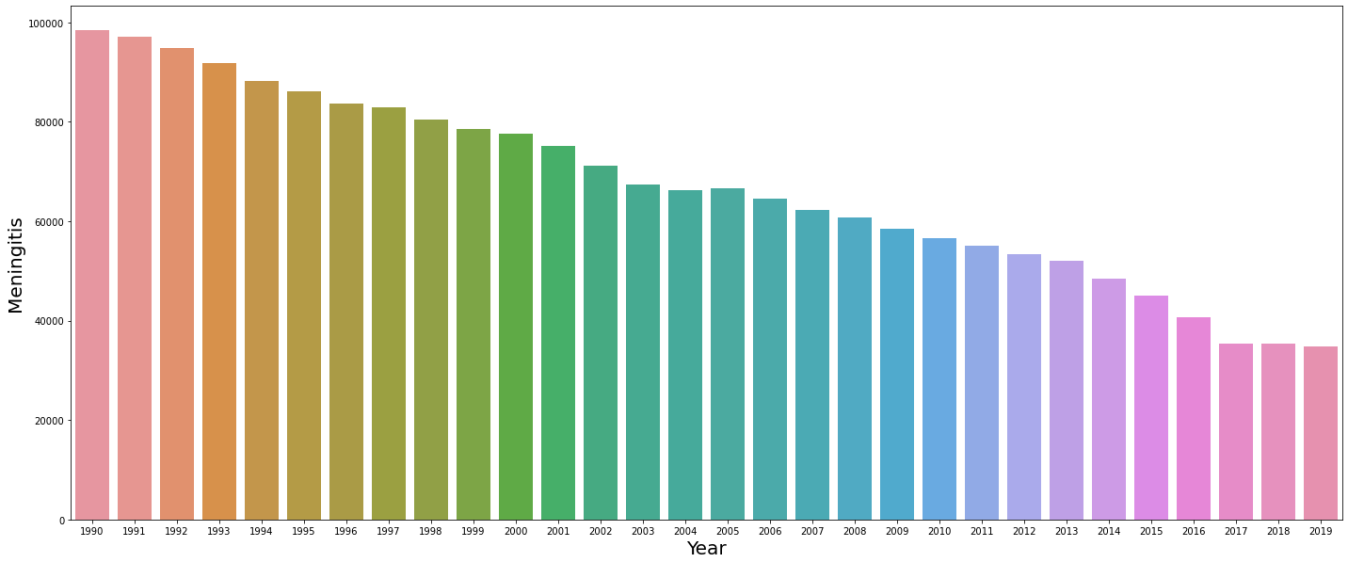
30 rows × 33 columns

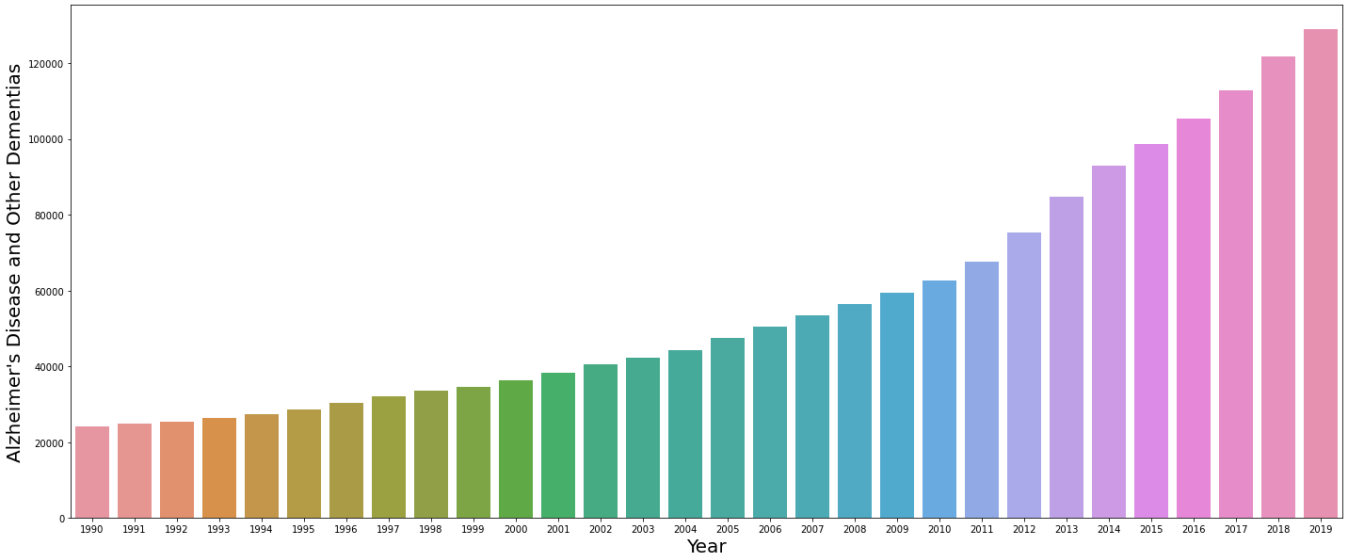
Segregating on the basis of Cause of deaths:

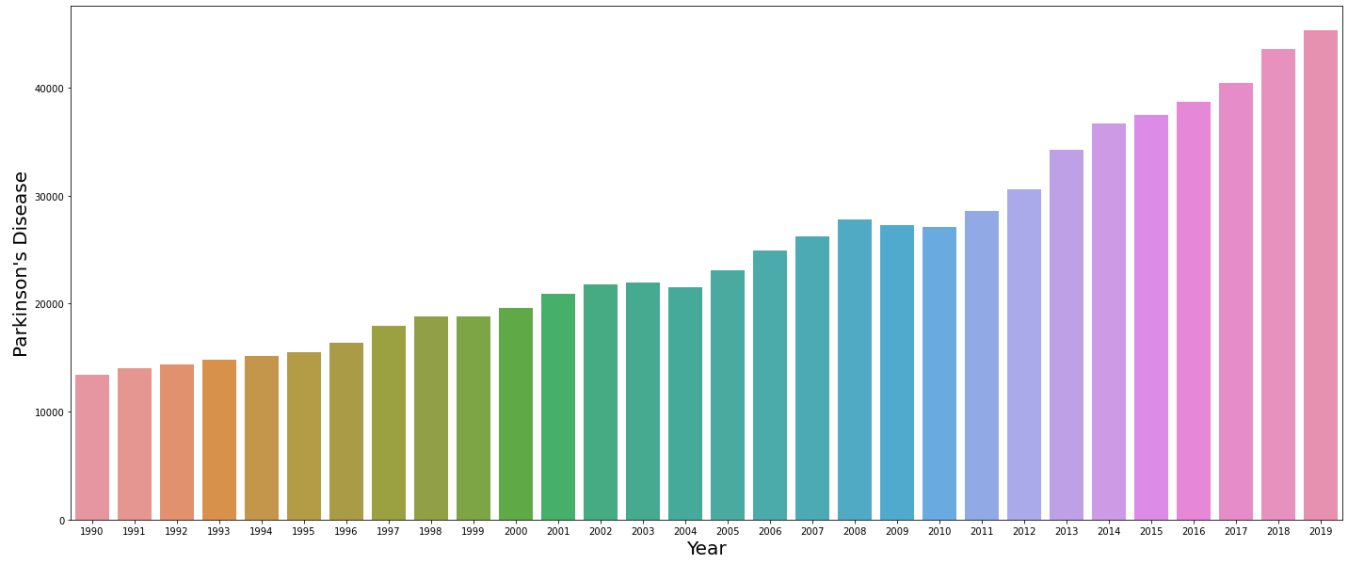
```
# Plot beween 'Year' & "Deaths Caused by Brain related Diseases":
India_Total_Deaths.plot(x='Year', y=['Meningitis',"Alzheimer's Disease and Other Dementias","Parkinson's Disease"],
                        figsize=(10,5), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Brain related Diseases" in India',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('India_Deaths',fontsize =16)
plt.show()
```




```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = India_Total_Deaths[['Meningitis',"Alzheimer's Disease and Other Dementias","Parkinson's Disease"]]
X = India_Total_Deaths['Year']
for col in y:
    if plotnumber<=8:
        plt.subplot(3,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

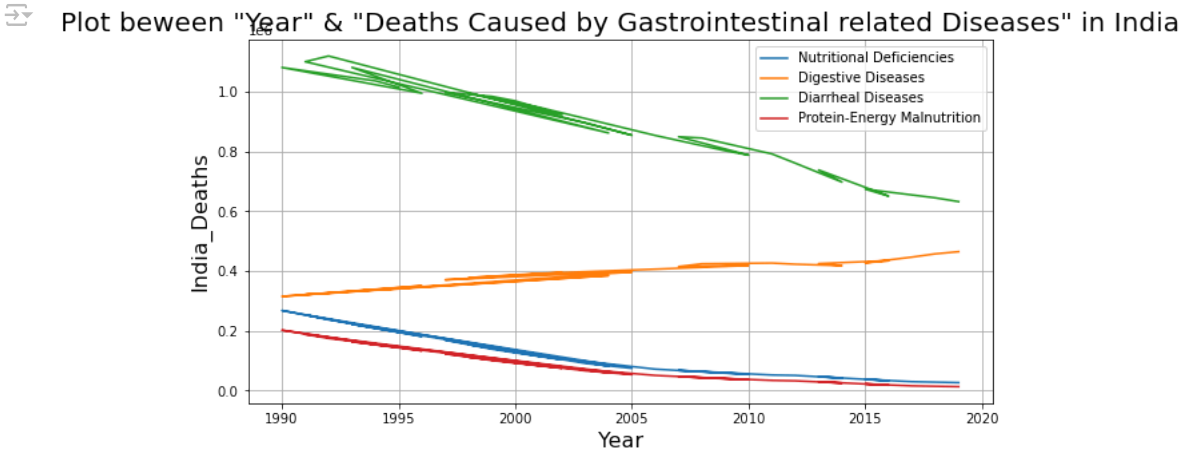

Meningitis
Year


Alzheimer's Disease and Other Dementias
Year

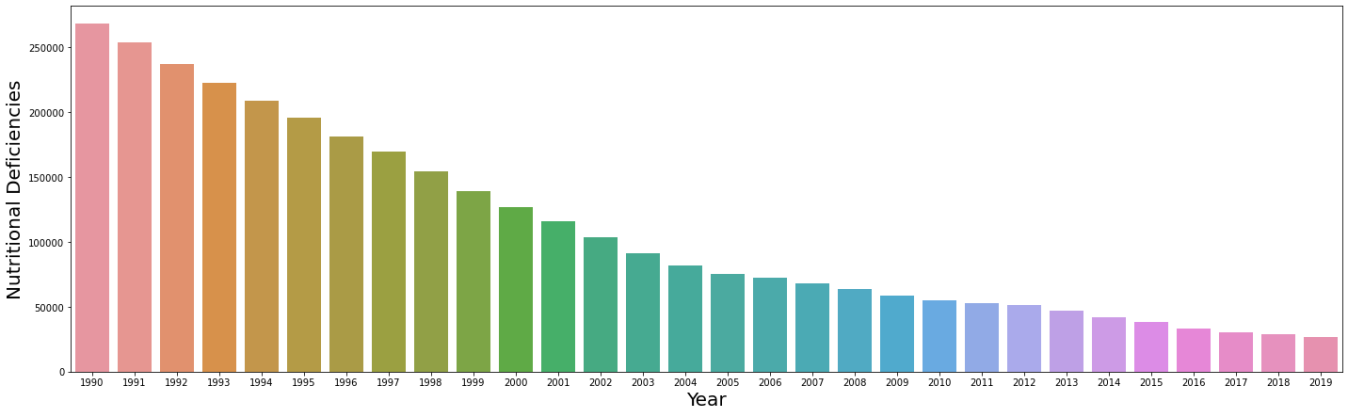

Parkinson's Disease
Year

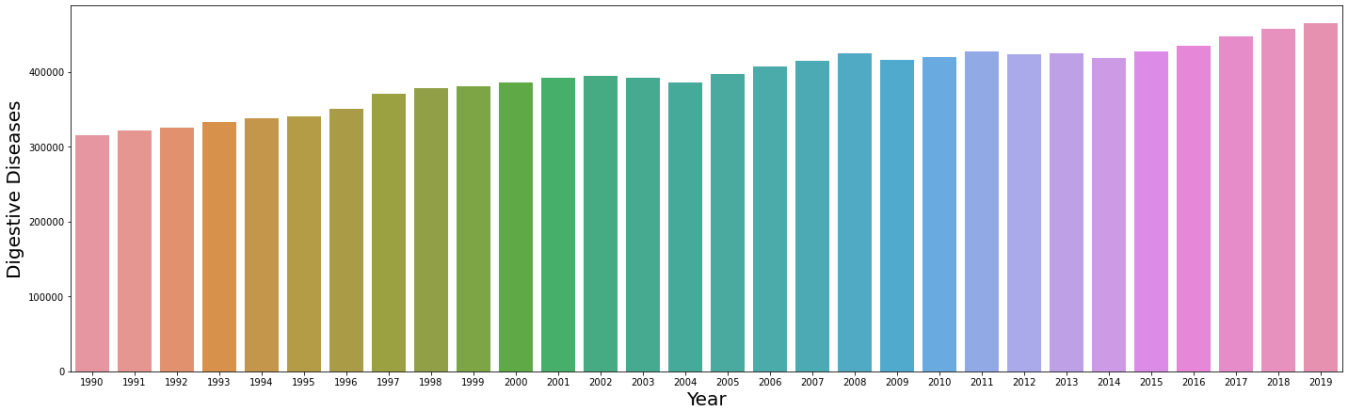
- Deaths in India due to **Alzheimer's Disease and Other Dementias** is becoming more dominant among all Deaths Caused by Brain related Diseases with every incrementing year.

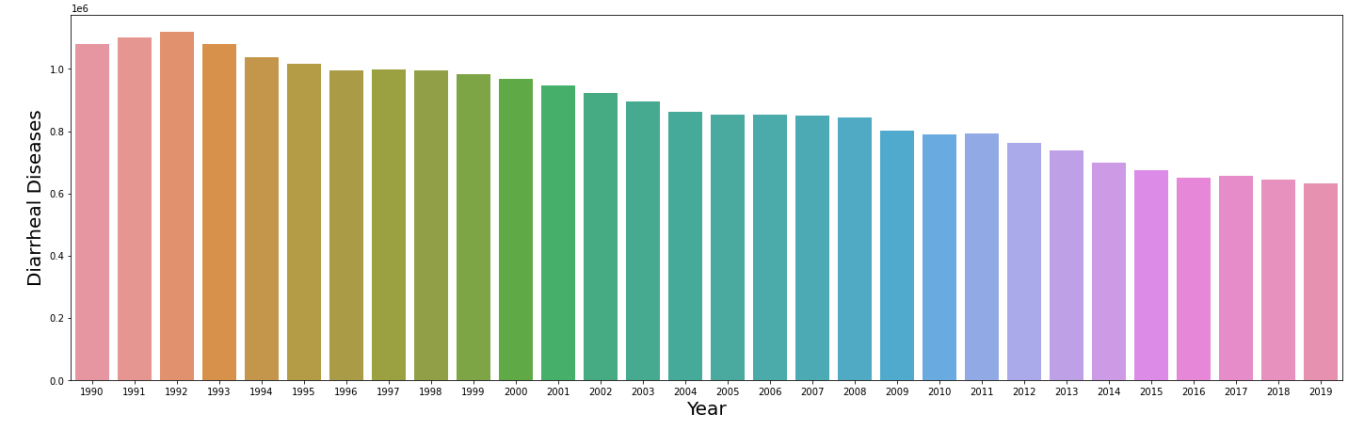
```
# Plot beween 'Year' & "Deaths Caused by Gastrointestinal related Diseases":
India_Total_Deaths.plot(x='Year', y=['Nutritional Deficiencies','Digestive Diseases',
        'Diarrheal Diseases','Protein-Energy Malnutrition'],figsize=(10,5), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Gastrointestinal related Diseases" in India',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('India_Deaths',fontsize =16)
plt.show()
```

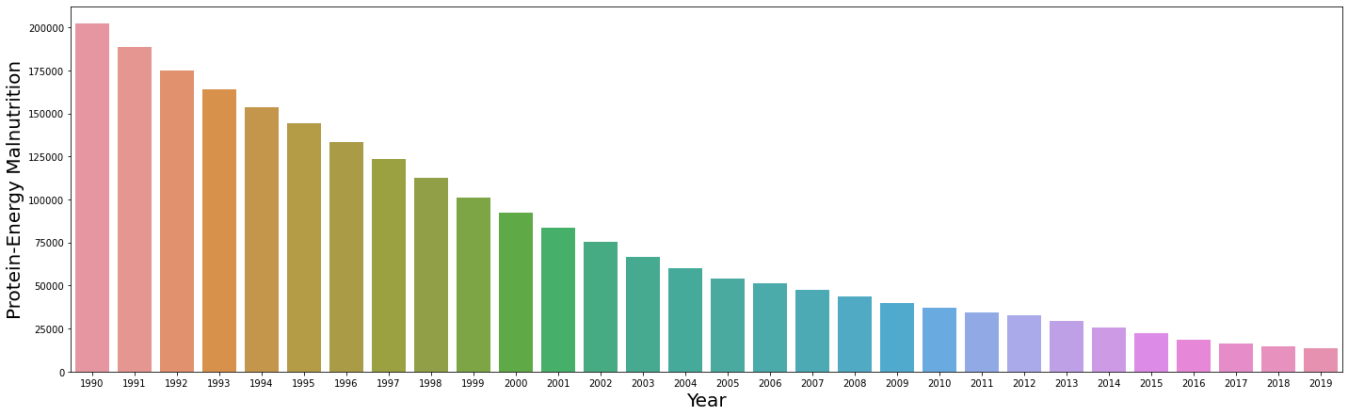



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = India_Total_Deaths[['Nutritional Deficiencies','Digestive Diseases',
        'Diarrheal Diseases','Protein-Energy Malnutrition']]
X = India_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```


Nutritional Deficiencies
Year

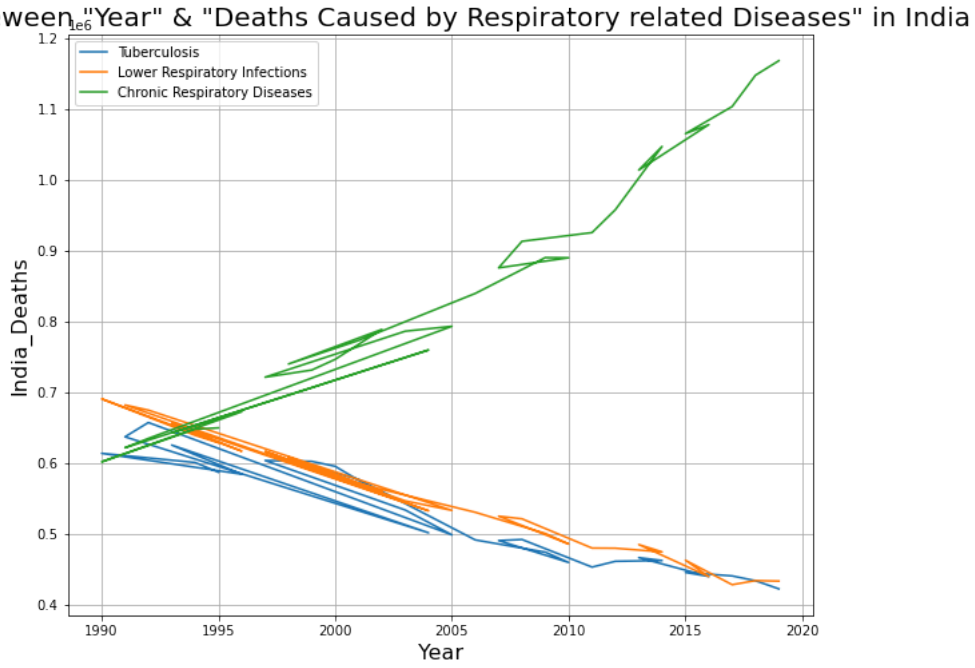

Digestive Diseases
Year


1e6
Diarrheal Diseases
Year

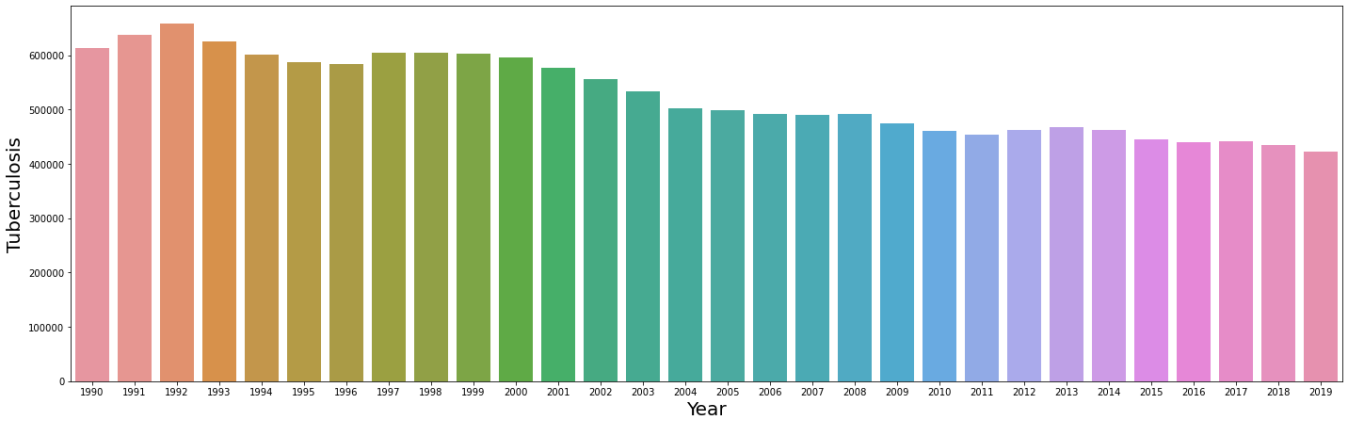

Protein-Energy Malnutrition
Year

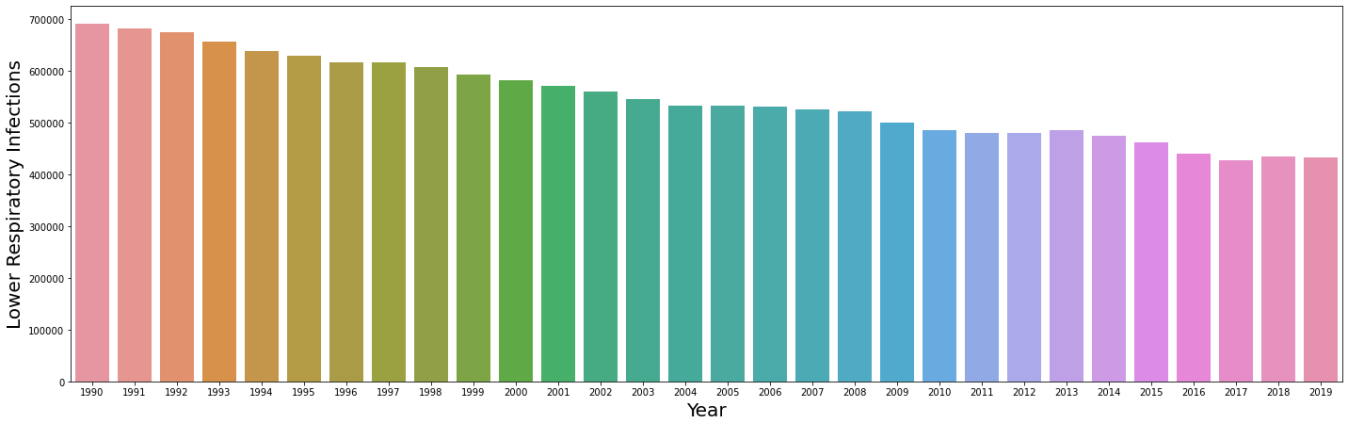
- Deaths in India due to Diarrheal Diseases is dominant among all Deaths Caused by Gastrointestinal related Diseases with every incrementing year.

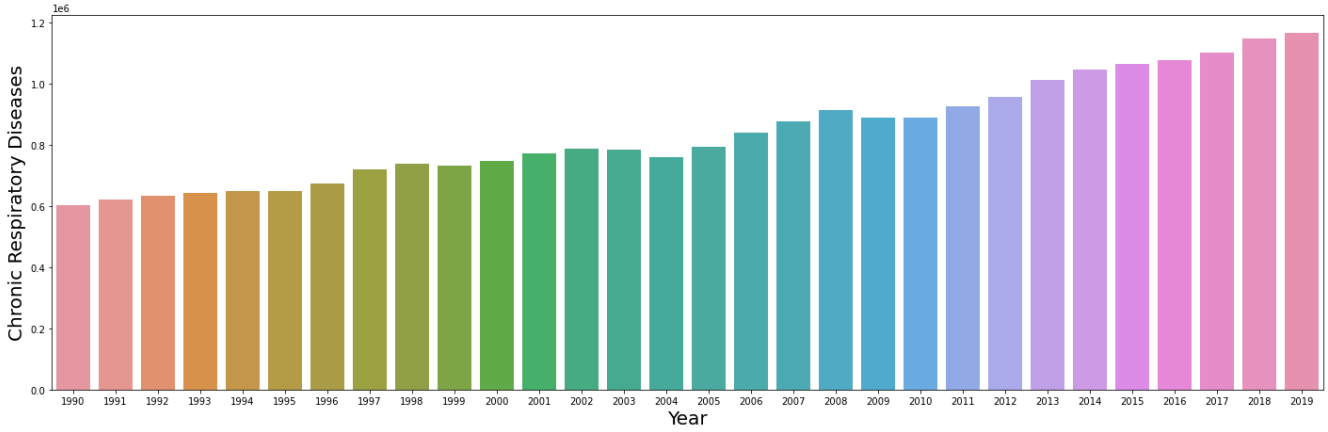
```
# Plot beween 'Year' & "Deaths Caused by Respiratory related Diseases":
India_Total_Deaths.plot(x='Year', y=['Tuberculosis','Lower Respiratory Infections',
                                     'Chronic Respiratory Diseases'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in India',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('India_Deaths',fontsize =16)
plt.show()
```




```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = India_Total_Deaths[['Tuberculosis','Lower Respiratory Infections','Chronic Respiratory Diseases']]
X = India_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



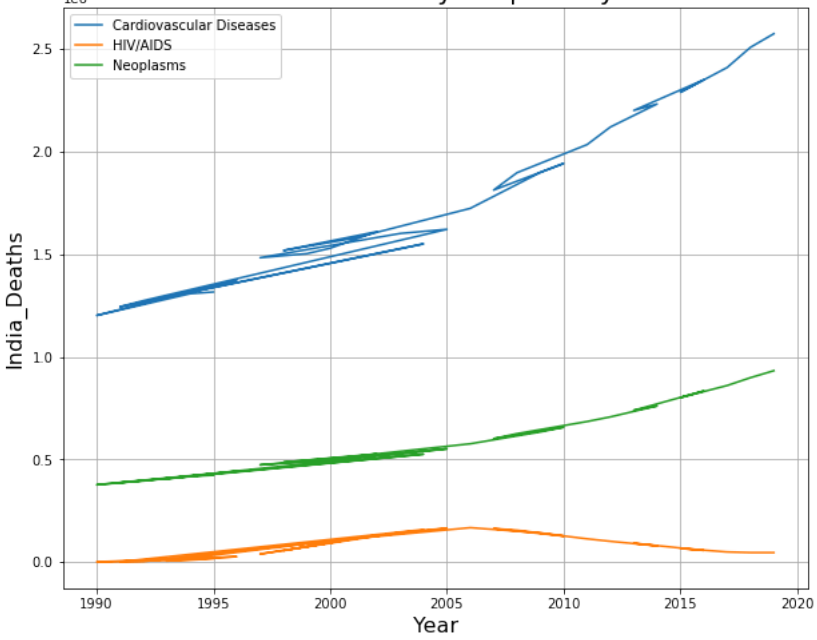




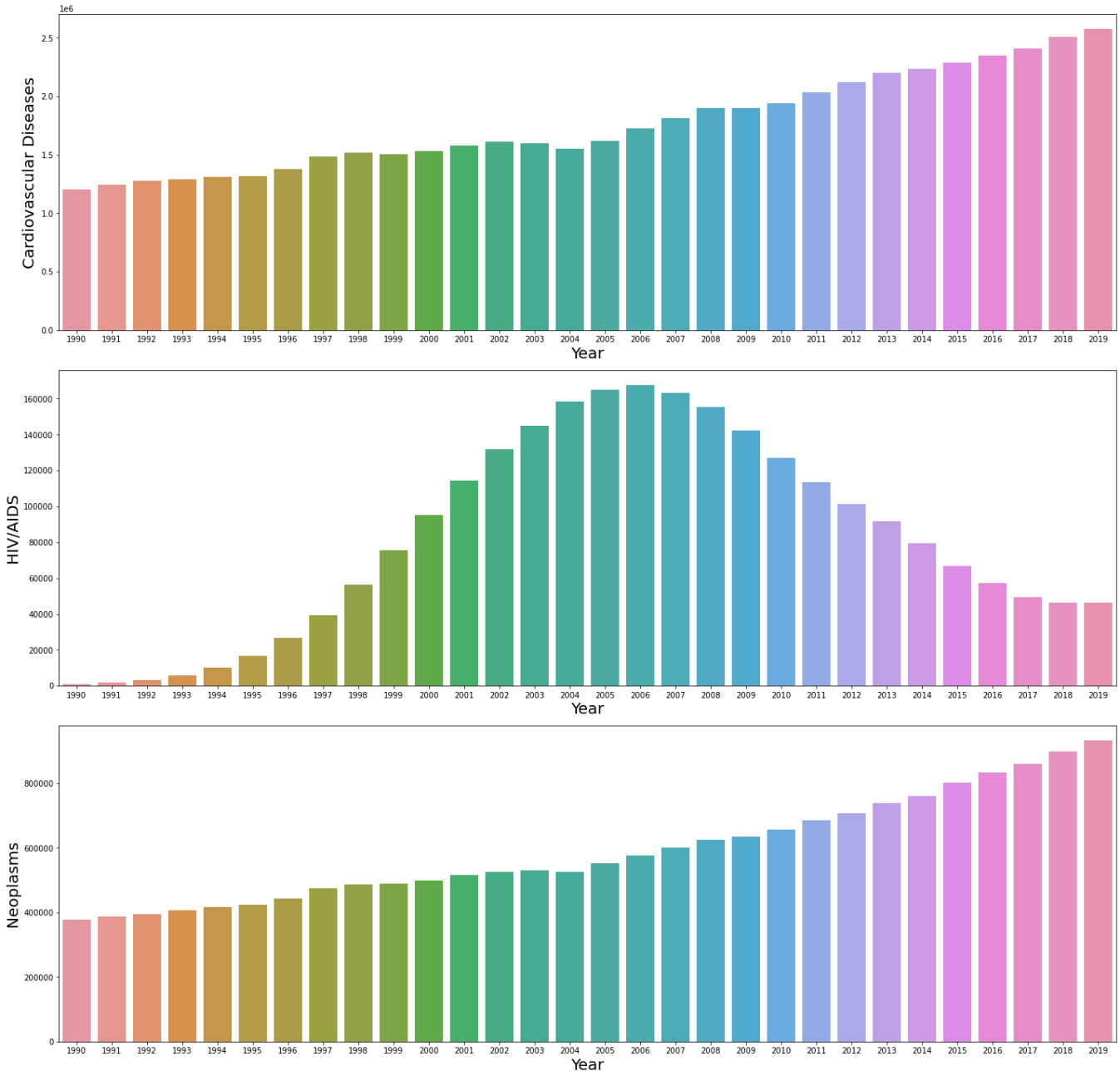
• Deaths in India due to Chronic Respiratory Diseases is dominant among all Deaths Caused by Respiratory related Diseases with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by 'Cardiovascular Diseases','HIV/AIDS' & 'Neoplasms' Diseases":
India_Total_Deaths.plot(x='Year', y=['Cardiovascular Diseases','HIV/AIDS','Neoplasms'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in India',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('India_Deaths',fontsize =16)
plt.show()
```

Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in India



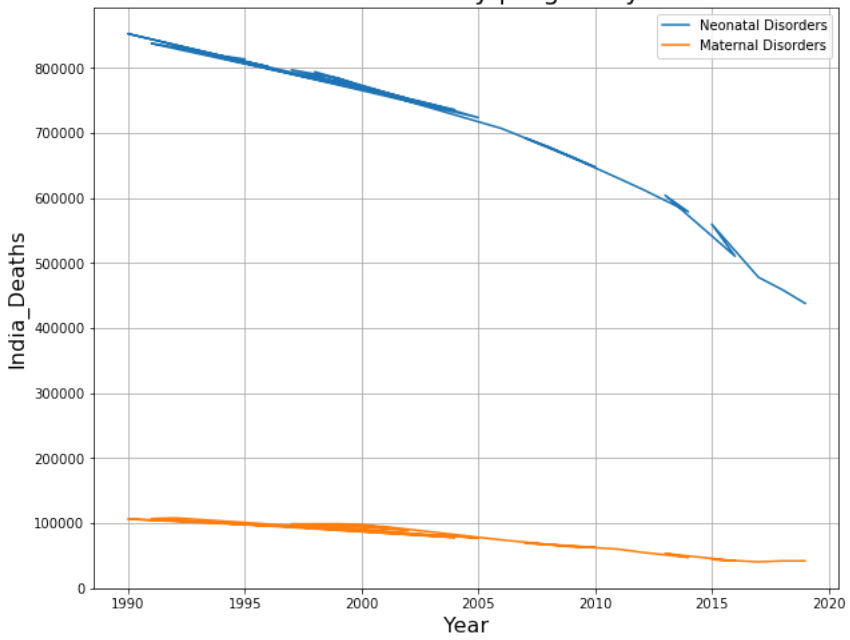

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = India_Total_Deaths[['Cardiovascular Diseases','HIV/AIDS','Neoplasms']]
X = India_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



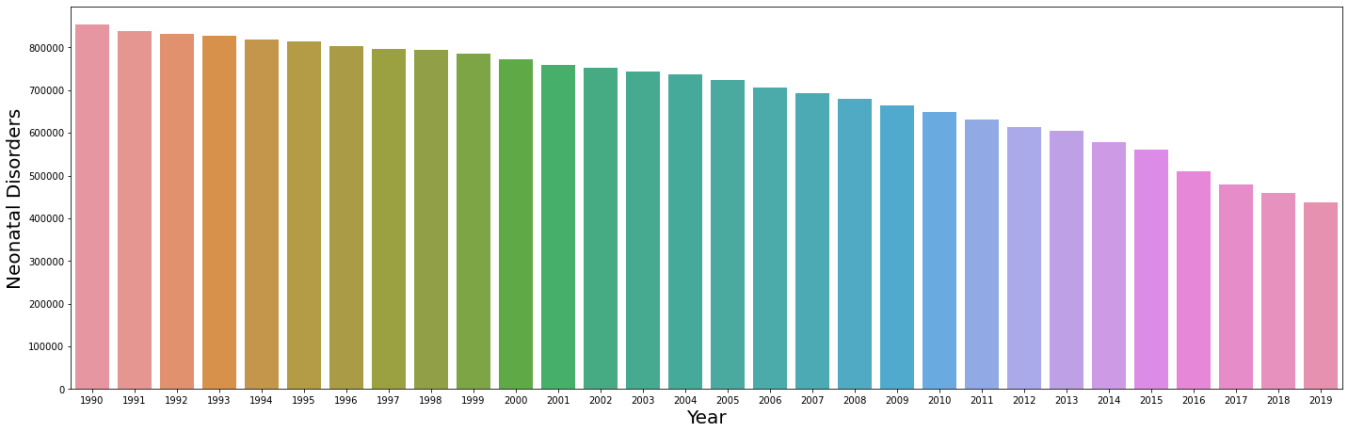
• Deaths in India due to cardiovascular diseases is most Dominant among all Deaths Caused with every incrementing year.

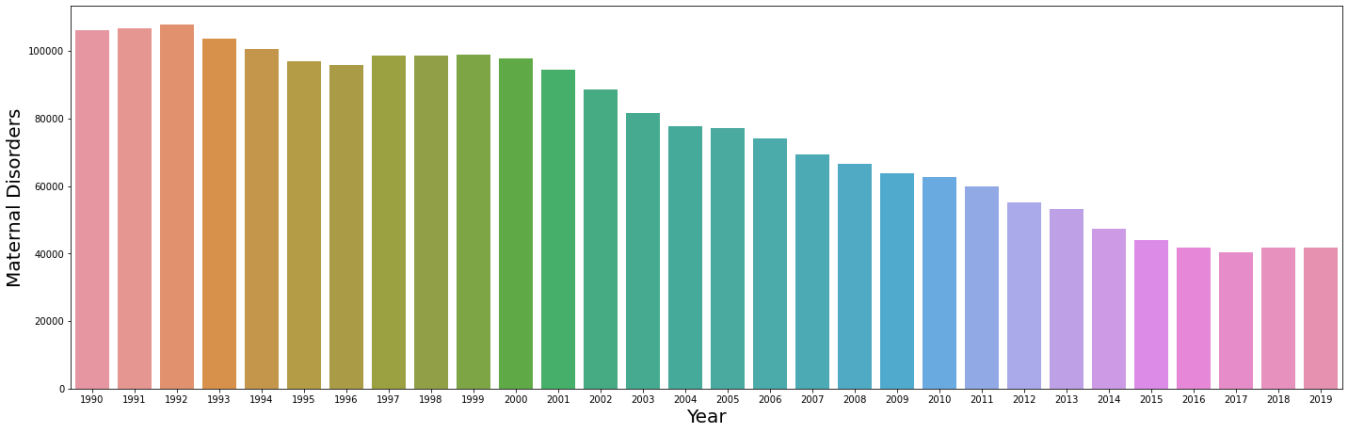
```
# Plot beween 'Year' & "Deaths Caused by pregnancy related Diseases":
India_Total_Deaths.plot(x='Year', y=['Neonatal Disorders','Maternal Disorders'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by pregnancy related Diseases" in India',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('India_Deaths',fontsize =16)
plt.show()
```

Plot beween "Year" & "Deaths Caused by pregnancy related Diseases" in India




```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = India_Total_Deaths[['Neonatal Disorders','Maternal Disorders']]
X = India_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

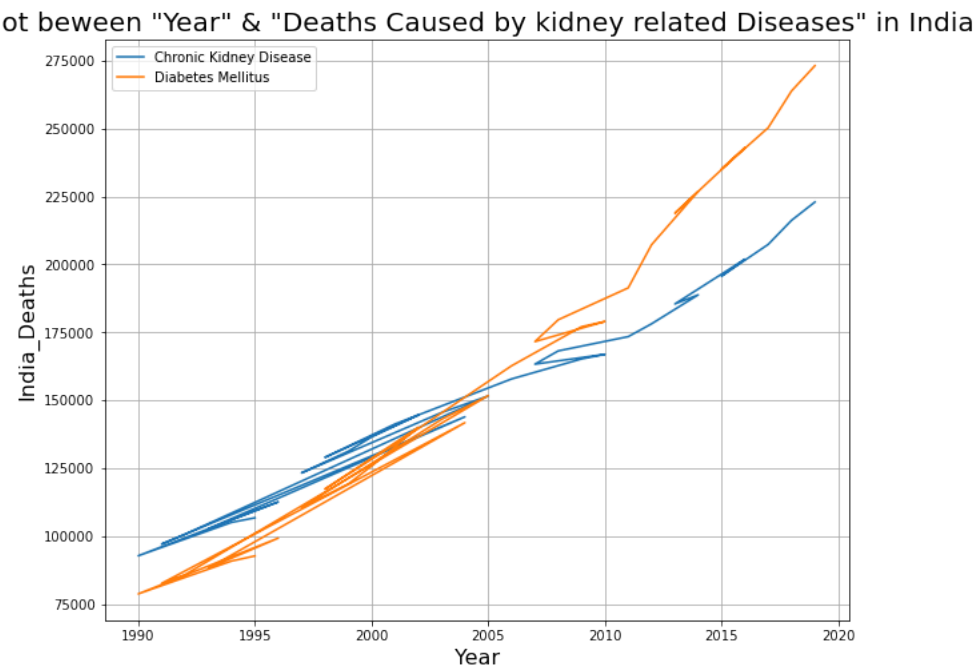




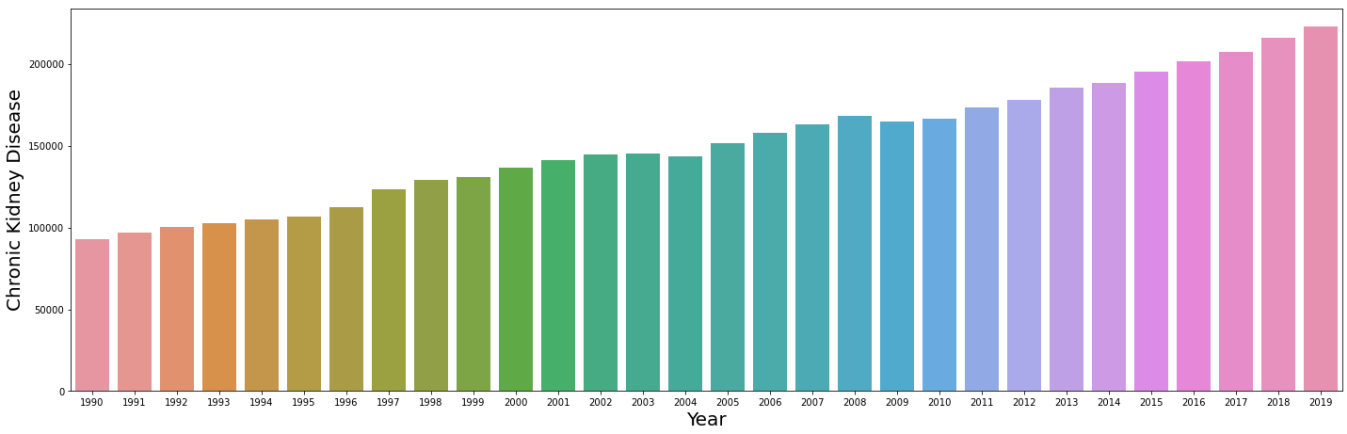
• Deaths in India due to Pregnancy related issues is on decline with every incrementing year.

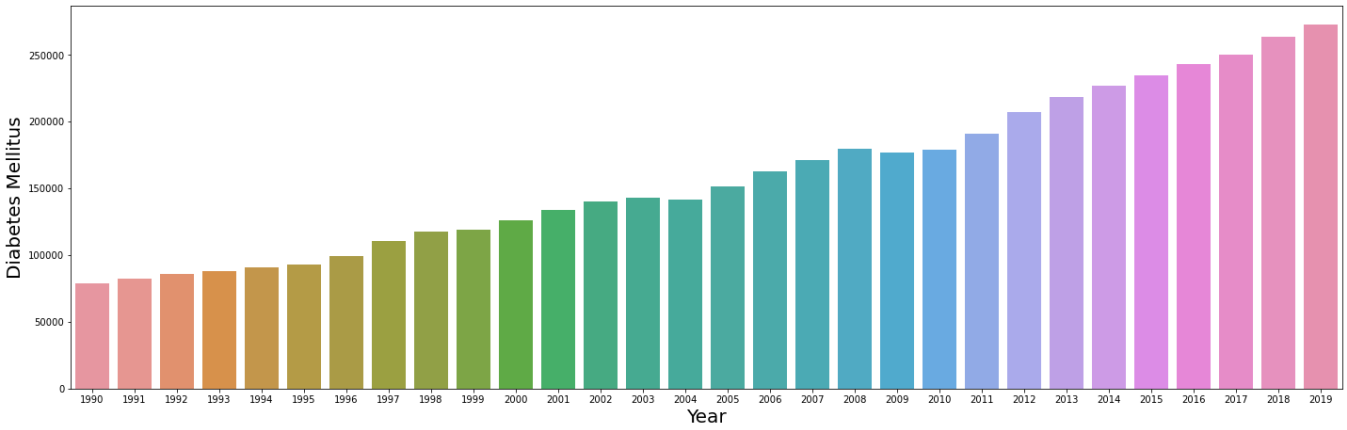
Double-click (or enter) to edit

```
# Plot beween 'Year' & "Deaths Caused by kidney related Diseases":
India_Total_Deaths.plot(x='Year', y=['Chronic Kidney Disease','Diabetes Mellitus'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by kidney related Diseases" in India',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('India_Deaths',fontsize =16)
plt.show()
```



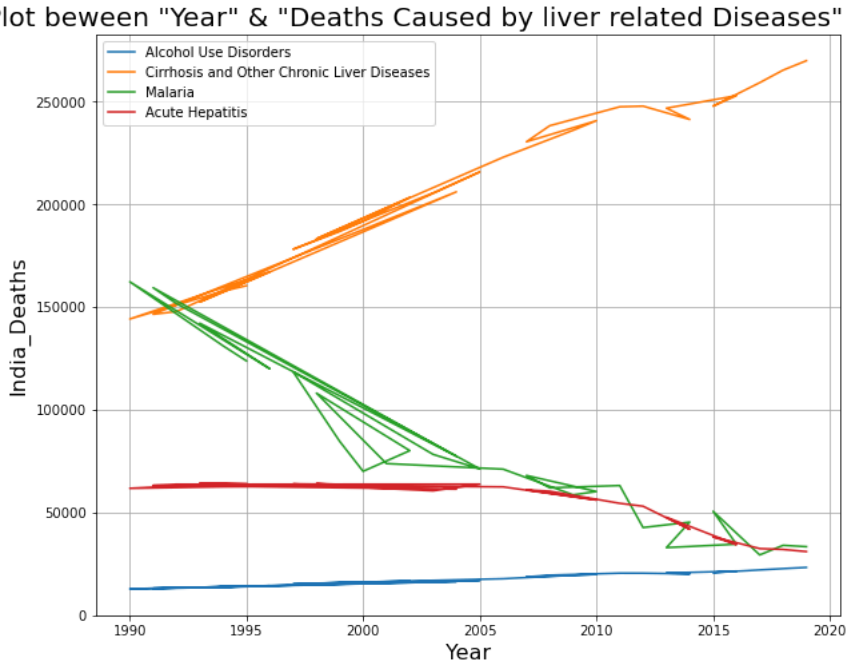

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = India_Total_Deaths[['Chronic Kidney Disease','Diabetes Mellitus']]
X = India_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



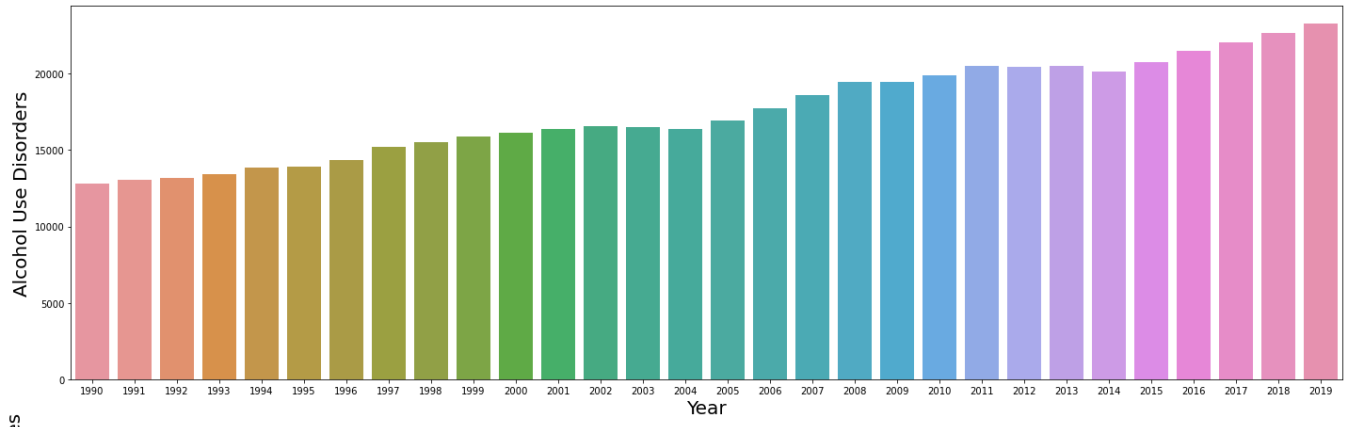


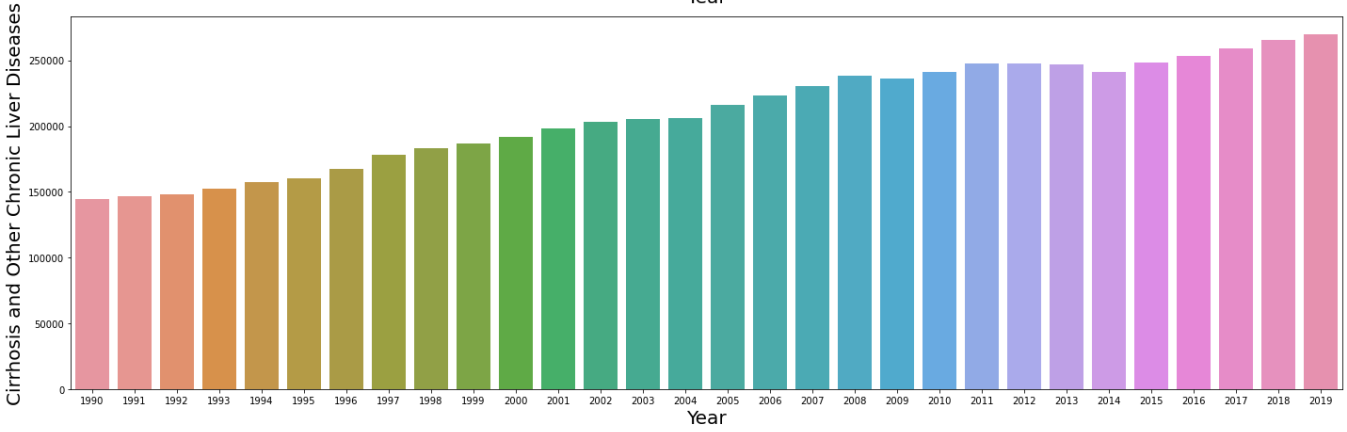
• Deaths in India due to all kidney related Diseases is Dominant with every incrementing year.

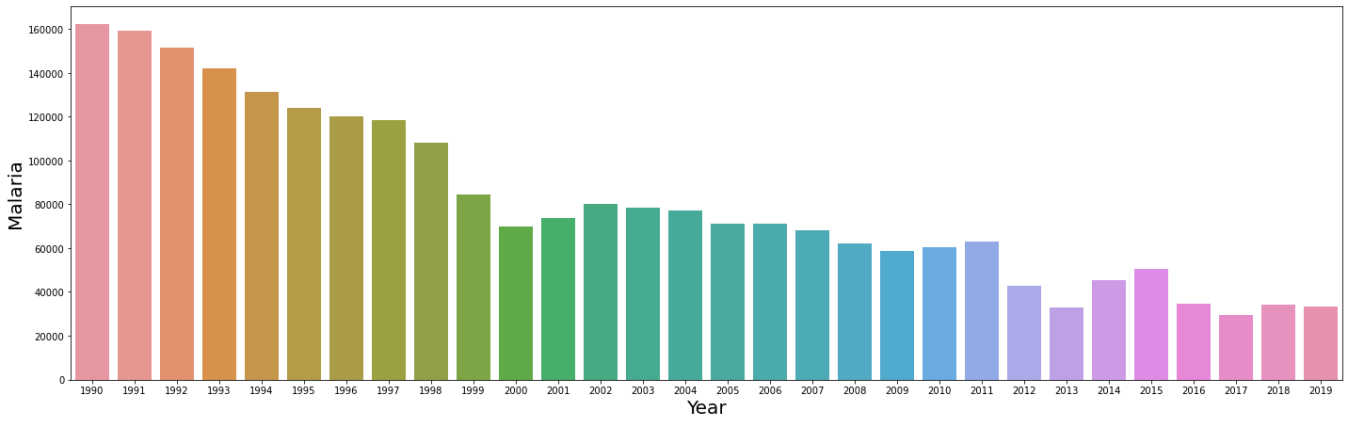
```
# Plot beween 'Year' & "Deaths Caused by liver related Diseases":
India_Total_Deaths.plot(x='Year', y=['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases',
                                    'Malaria','Acute Hepatitis'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by liver related Diseases" in India',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('India_Deaths',fontsize =16)
plt.show()
```

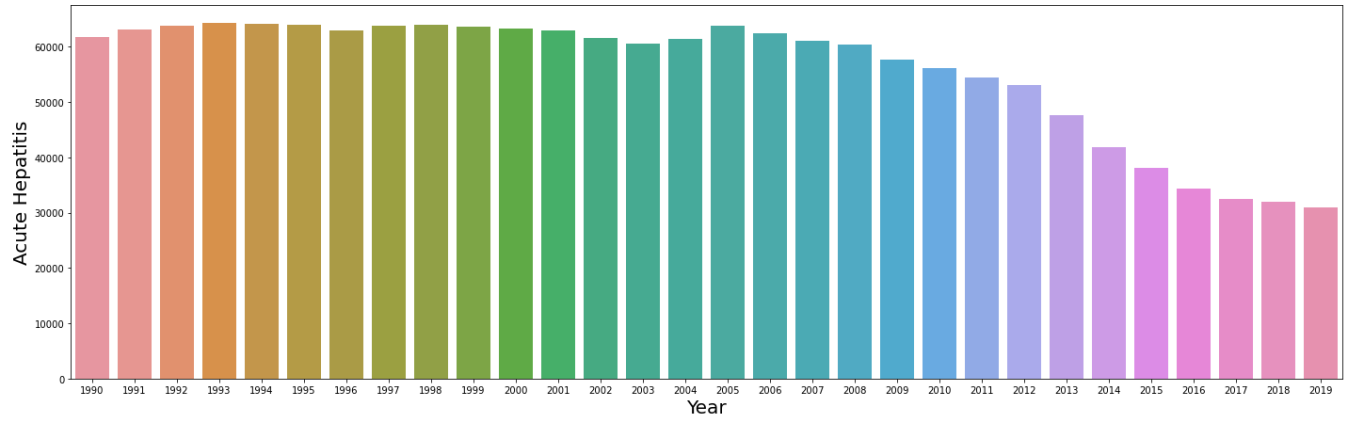


```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = India_Total_Deaths[['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases',
                        'Malaria','Acute Hepatitis']]
X = India_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```


Alcohol Use Disorders
Year

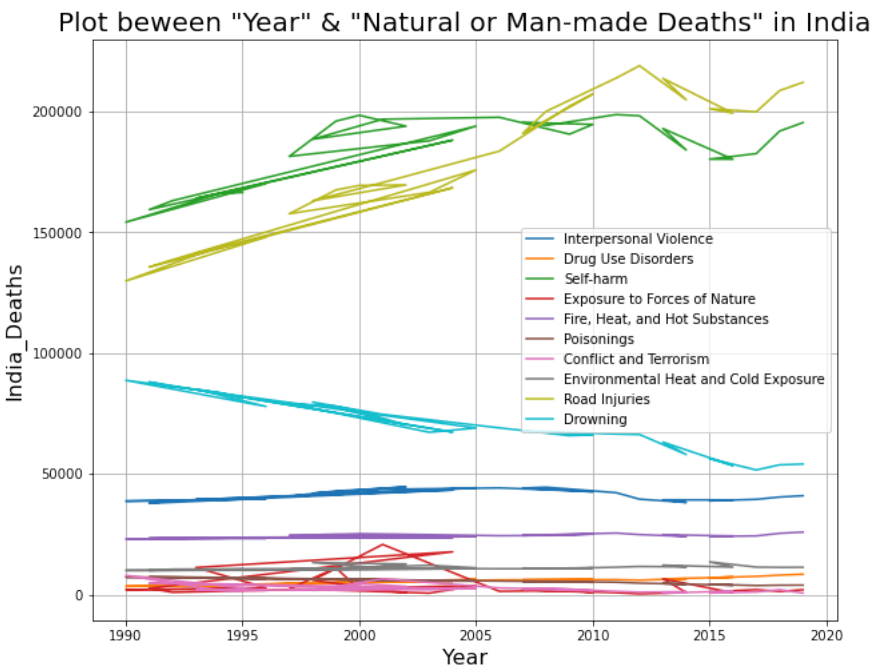

Cirrhosis and Other Chronic Liver Diseases
Year


Malaria
Year

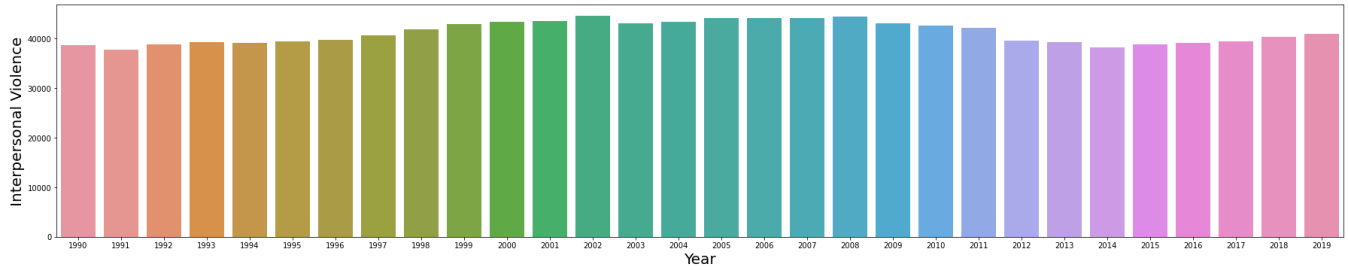

Acute Hepatitis
Year

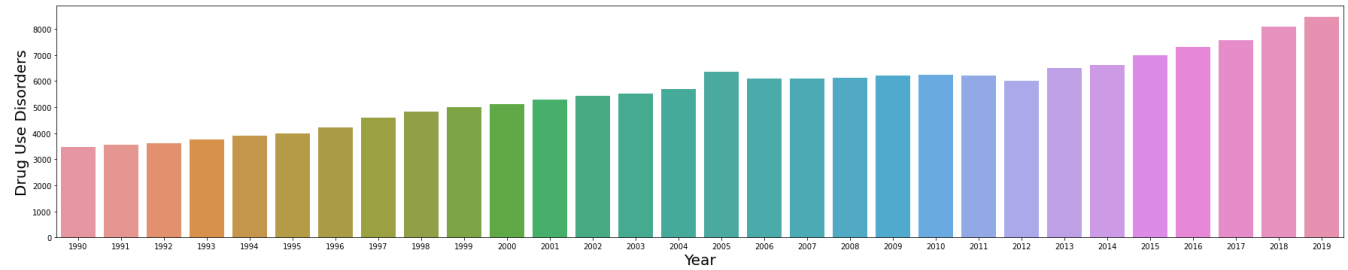
• Deaths in India due to Cirrhosis and Other Chronic Liver Diseases is Dominant among all Deaths Caused by liver related Diseases with every incrementing year.

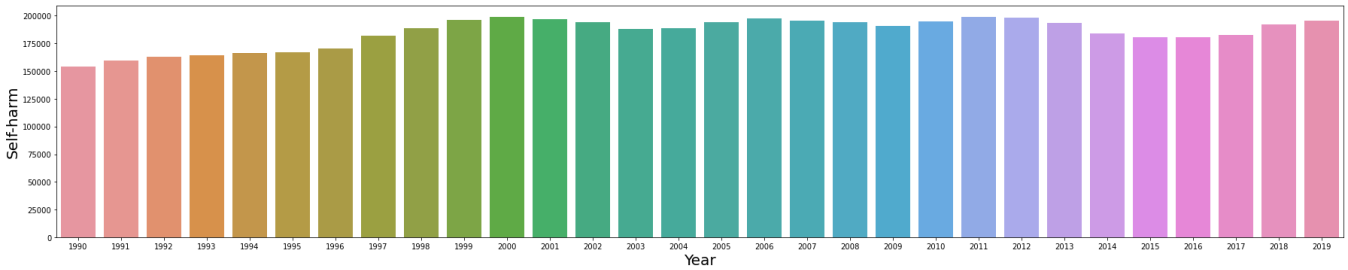
```
# Plot beween 'Year' & "Natural or Man-made Deaths":
India_Total_Deaths.plot(x='Year', y=['Interpersonal Violence','Drug Use Disorders',
                                     'Self-harm', 'Exposure to Forces of Nature',
                                     "Fire, Heat, and Hot Substances",
                                     'Poisonings','Conflict and Terrorism',
                                     'Environmental Heat and Cold Exposure',
                                     'Road Injuries','Drowning'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Natural or Man-made Deaths" in India',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('India_Deaths',fontsize =16)
plt.show()
```

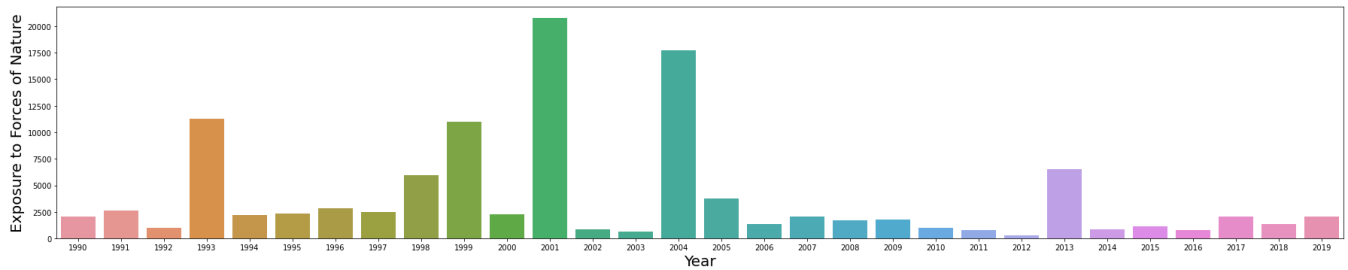


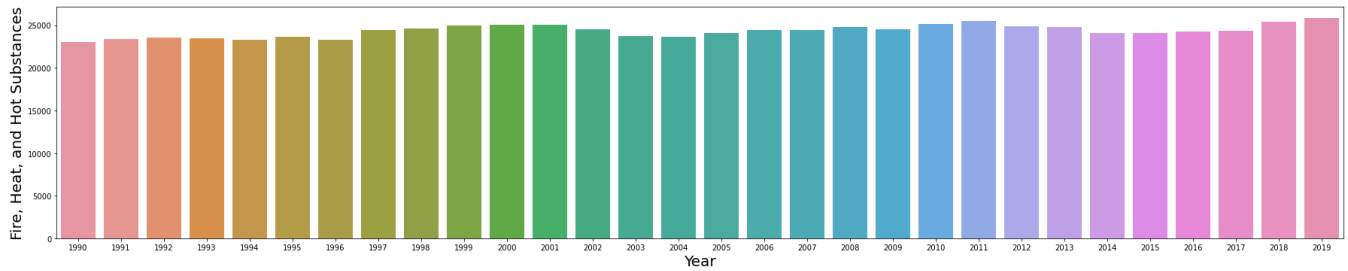

```
plt.figure(figsize=(25,50),facecolor='white')
plotnumber=1
y = India_Total_Deaths[['Interpersonal Violence','Drug Use Disorders',
                                     'Self-harm', 'Exposure to Forces of Nature',
                                     "Fire, Heat, and Hot Substances",
                                     'Poisonings','Conflict and Terrorism',
                                     'Environmental Heat and Cold Exposure',
                                     'Road Injuries','Drowning']]
X = India_Total_Deaths['Year']
for col in y:
    if plotnumber<=10:
        plt.subplot(10,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

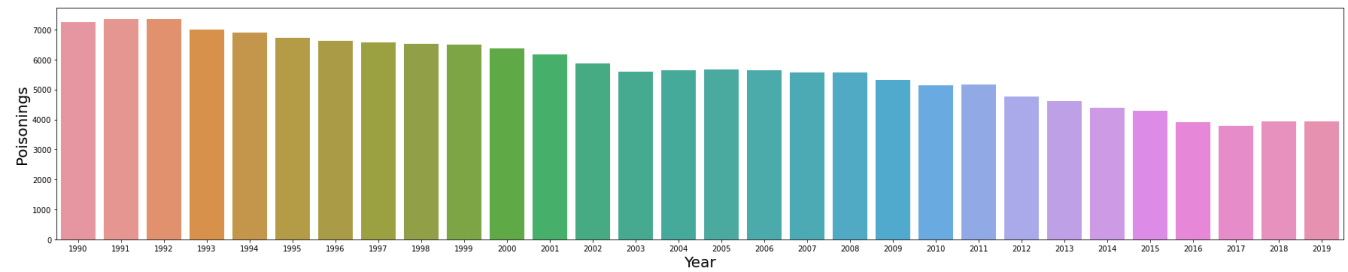

Interpersonal Violence
Year

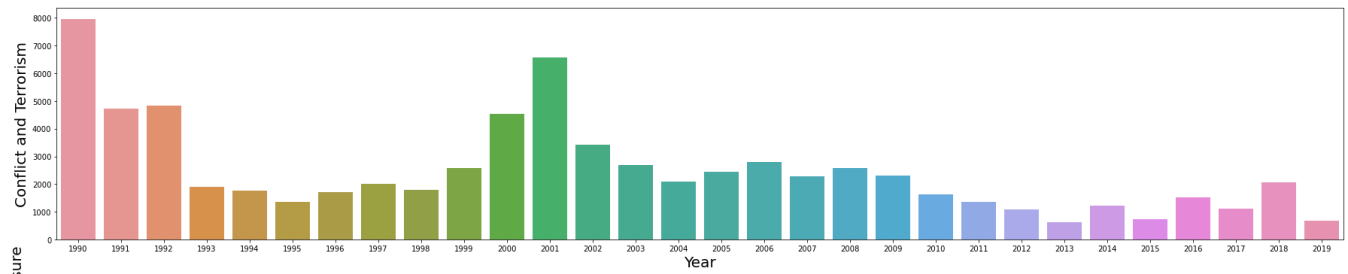

Drug Use Disorders
Year

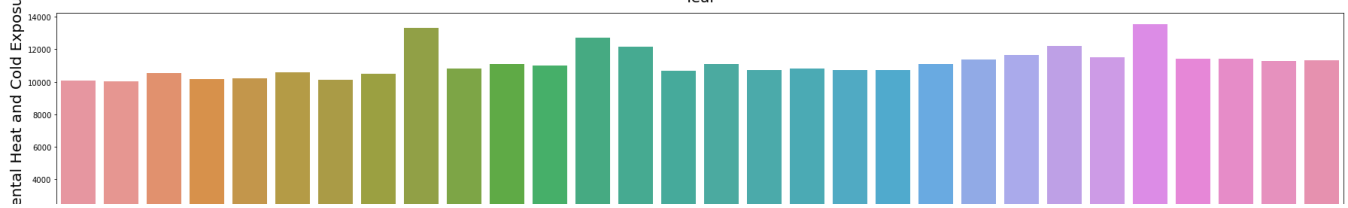

Self-harm
Year

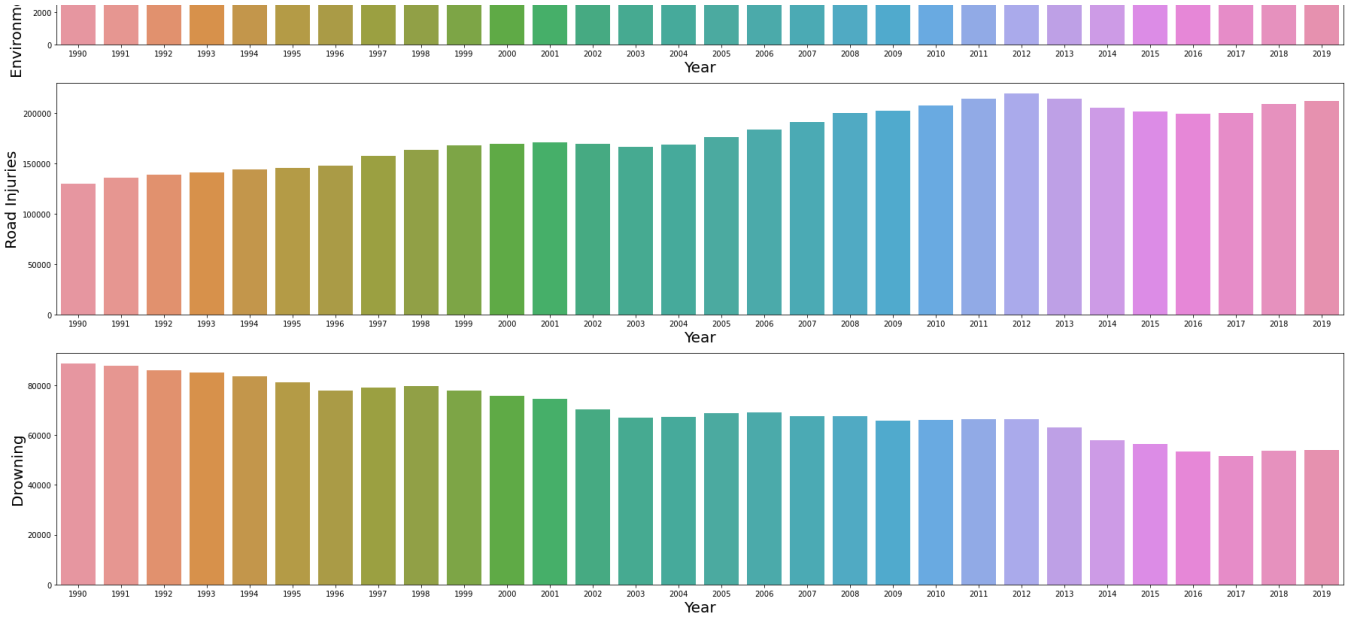

Exposure to Forces of Nature
Year


Fire, Heat, and Hot Substances
Year


Poisonings
Year


Conflict and Terrorism
Year


ental Heat and Cold Exposure



Deaths in India due to Natural or Man-made Deaths are summarised below:

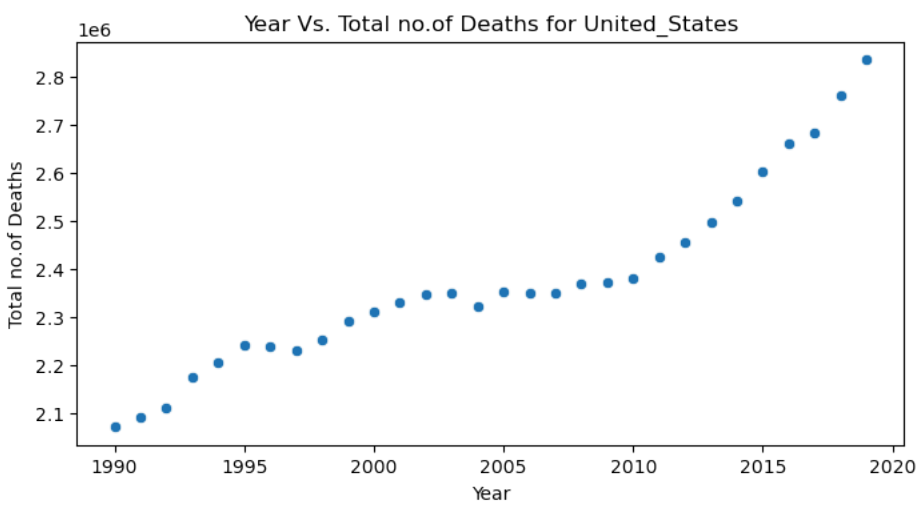
- Deaths due to 'Interpersonal Violence' is constant with every incrementing year.
- Deaths due to 'Drug Use Disorders' is incrementing with every incrementing year.
- Deaths due to 'Self-harm' is constant with every incrementing year.
- Deaths due to 'Exposure to Forces of Nature' is highest in 1993, 2001, 2004 & 2013.
- Deaths due to "Fire, Heat, and Hot Substances" is constant with every incrementing year.
- Deaths due to 'Poisonings', is declining with every incrementing year.
- Deaths due to 'Conflict and Terrorism' is declining with every incrementing year, but need to be taken care of.
- Deaths due to 'Environmental Heat and Cold Exposure' is constant with every incrementing year.
- Deaths due to 'Road Injuries' is incrementing with every incrementing year.
- Deaths due to 'Drowning' is declining with every incrementing year.

3. For United States of America:

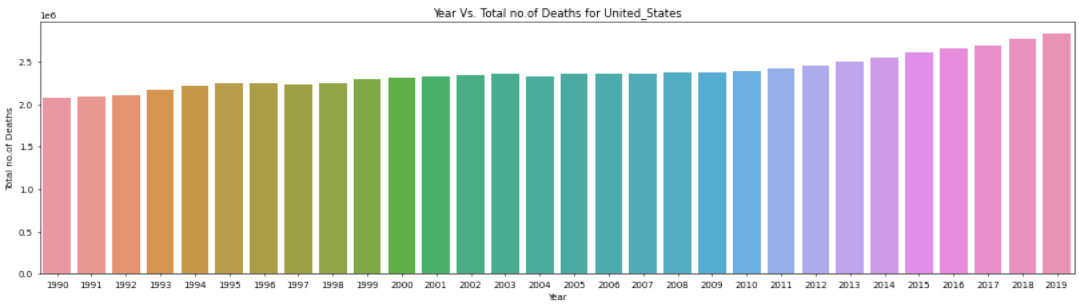
```
# USA: "Total_Deaths" against "Year"

United_States_Total_Deaths = df[df['Country/Territory']=='United States'].sort_values(by='Total_Deaths',ascending=False)
United_States_Total_Deaths[["Country/Territory",'Year','Total_Deaths']]
```

	Country/Territory	Year	Total_Deaths
5849	United States	2019	2834964
5848	United States	2018	2761982
5847	United States	2017	2683821
5829	United States	2016	2659829
5828	United States	2015	2602225
5827	United States	2014	2540672
5826	United States	2013	2496904
5825	United States	2012	2455675
5824	United States	2011	2424621
5823	United States	2010	2380923
5822	United States	2009	2372134
5821	United States	2008	2369139
5845	United States	2005	2352254
5846	United States	2006	2349964
5843	United States	2003	2349782
5820	United States	2007	2348450
5842	United States	2002	2346876
5841	United States	2001	2330838
5844	United States	2004	2322785
5840	United States	2000	2309875
5839	United States	1999	2291373
5838	United States	1998	2252332
5835	United States	1995	2240998
5836	United States	1996	2237660
5837	United States	1997	2230711
5834	United States	1994	2205930
5833	United States	1993	2174147
5832	United States	1992	2108829
5831	United States	1991	2091119
5830	United States	1990	2070990

```
# United States: "Total_Deaths" v/s "Year"

plt.figure(figsize=(8,4),dpi=100)
sns.scatterplot(data=United_States_Total_Deaths, x='Year', y='Total_Deaths')
plt.xlabel("Year")
plt.ylabel("Total no.of Deaths")
plt.title("Year Vs. Total no.of Deaths for United_States")
plt.show();
```




```
plt.figure(figsize=(20,5),dpi=50)
sns.barplot(data=United_States_Total_Deaths, x='Year', y='Total_Deaths')
plt.xlabel("Year")
plt.ylabel("Total no.of Deaths")
plt.title("Year Vs. Total no.of Deaths for United_States")
plt.show();
```



- **There is a clear raise in Total No.of Deaths recorded with each incrementing year for United_States.**

Now, let's Investigate the Major cause of Deaths in United_States:

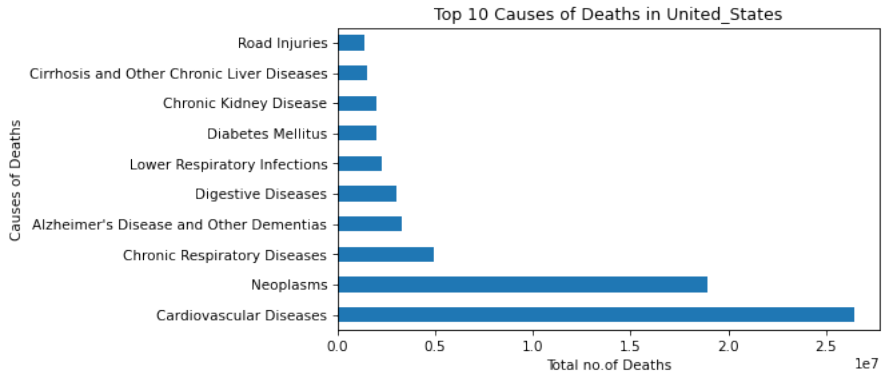
```
# Get maximum deaths causing Disease wrt Year:

United_States_Total_Deaths.set_index(["Country/Territory", "Total_Deaths"], inplace = True) # making County & Total_Deaths as index.
United_States_Total_Deaths['Max'] = United_States_Total_Deaths.idxmax(axis=1) # getting desease for max cause of deaths in China
United_States_Total_Deaths[['Year','Max']] # displying reqd data only.
```

		Year	Max
Country/Territory	Total_Deaths		
United States	**2834964**	2019	Cardiovascular Diseases
	2761982	2018	Cardiovascular Diseases
	2683821	2017	Cardiovascular Diseases
	2659829	2016	Cardiovascular Diseases
	2602225	2015	Cardiovascular Diseases
	2540672	2014	Cardiovascular Diseases
	2496904	2013	Cardiovascular Diseases
	2455675	2012	Cardiovascular Diseases
	2424621	2011	Cardiovascular Diseases
	2380923	2010	Cardiovascular Diseases
	2372134	2009	Cardiovascular Diseases
	2369139	2008	Cardiovascular Diseases
	2352254	2005	Cardiovascular Diseases
	2349964	2006	Cardiovascular Diseases
	2349782	2003	Cardiovascular Diseases
	2348450	2007	Cardiovascular Diseases
	2346876	2002	Cardiovascular Diseases
	2330838	2001	Cardiovascular Diseases
	2322785	2004	Cardiovascular Diseases
	2309875	2000	Cardiovascular Diseases
	2291373	1999	Cardiovascular Diseases
	2252332	1998	Cardiovascular Diseases
	2240998	1995	Cardiovascular Diseases
	2237660	1996	Cardiovascular Diseases
	2230711	1997	Cardiovascular Diseases
	2205930	1994	Cardiovascular Diseases
	2174147	1993	Cardiovascular Diseases
	2108829	1992	Cardiovascular Diseases
	2091119	1991	Cardiovascular Diseases
	2070990	1990	Cardiovascular Diseases

```
# United_States - Top 10 Causes of Deaths:

United_States_10 = Countries_Total_no_of_Deaths_20year_data.sort_values(by='Total_Deaths',ascending =False)
United_States_10.iloc[2].sort_values(ascending=False)[1:11]
```

```
Cardiovascular Diseases                      26438346
Neoplasms                                    18905315
Chronic Respiratory Diseases                  4949052
Alzheimer's Disease and Other Dementias       3302609
Digestive Diseases                            3026943
Lower Respiratory Infections                  2248625
Diabetes Mellitus                             2030631
Chronic Kidney Disease                        2018497
Cirrhosis and Other Chronic Liver Diseases    1514325
Road Injuries                                 1359744
Name: United States, dtype: int64
```

```
plt.figure(figsize=(7,4),dpi=75)
United_States_10.iloc[2].sort_values(ascending=False)[1:11].plot(kind='barh')
plt.xlabel("Total no.of Deaths")
plt.ylabel("Causes of Deaths")
plt.title("Top 10 Causes of Deaths in United_States")
plt.show();
```



- **Major Causes of Deaths for China are:**

1. Cardiovascular Diseases
2. Neoplasms
3. Chronic Respiratory Diseases
4. Alzheimer's Disease and Other Dementias
5. Digestive Diseases
6. Lower Respiratory Infections
7. Diabetes Mellitus
8. Chronic Kidney Disease
9. Cirrhosis and Other Chronic Liver Diseases
10. Road Injuries

```
United_States_Total_Deaths
```

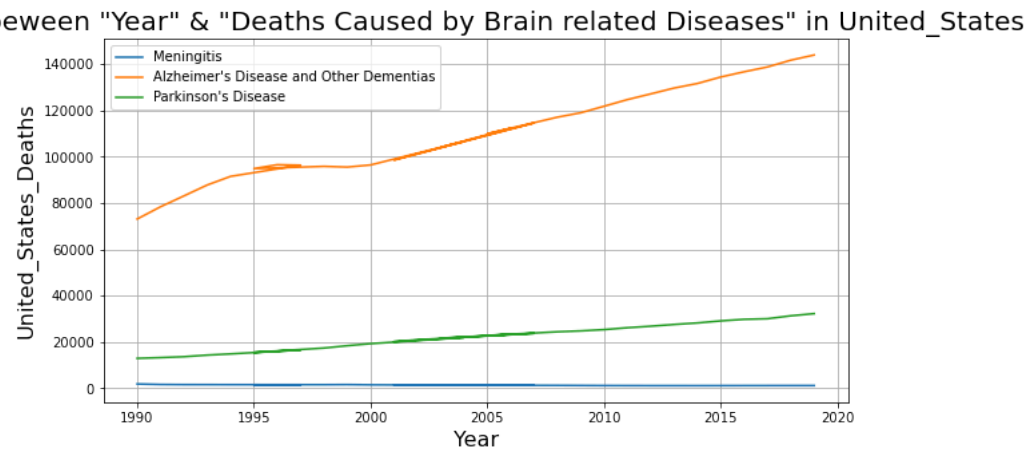
Country/Territory	Total_Deaths	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	Matern Disorde
United States	2834964	2019	1146	143919	32211	6090	0	3615	17709	9
	2761982	2018	1142	141681	31277	5934	0	3687	18086	10
	2683821	2017	1137	138750	30014	5814	0	3769	18831	11
	2659829	2016	1131	136673	29748	5952	0	3846	19254	12
	2602225	2015	1115	134411	29081	5661	0	3727	18164	12
	2540672	2014	1102	131616	28193	5268	0	3627	16618	11
	2496904	2013	1106	129640	27531	4943	0	3606	16767	11
	2455675	2012	1108	127120	26781	4707	0	3647	17577	10
	2424621	2011	1126	124656	26123	4538	0	3669	17224	10
	2380923	2010	1141	121781	25309	4367	0	3684	17329	9
	2372134	2009	1186	118965	24760	4277	0	3753	17944	10
	2369139	2008	1223	117064	24369	4280	0	3819	18716	11
	2352254	2005	1288	109900	22755	4372	0	3850	19380	10
	2349964	2006	1273	112291	23243	4257	0	3888	19765	10
	2349782	2003	1344	104024	21562	4570	0	3761	18911	7
	2348450	2007	1249	114585	23705	4238	0	3864	19370	10
	2346876	2002	1367	101170	20907	4704	0	3772	18979	6
	2330838	2001	1399	98558	20082	4736	0	3720	18723	6
	2322785	2004	1293	106572	21922	4392	0	3747	18767	9
	2309875	2000	1458	96426	19242	4760	0	3728	18204	6
	2291373	1999	1566	95512	18354	4617	0	3749	18421	5
	2252332	1998	1517	95816	17349	4263	0	3831	19286	5
	2240998	1995	1514	94855	15434	3530	0	3967	23477	4
	2237660	1996	1504	96526	15962	3741	0	3857	21706	5
	2230711	1997	1514	96342	16545	3978	0	3810	20385	5
	2205930	1994	1530	91491	14816	3312	0	3999	24898	5
	2174147	1993	1550	87768	14301	3148	0	4075	25598	5
	2108829	1992	1558	83053	13600	2945	0	4069	25307	5
	2091119	1991	1615	78365	13217	2855	0	4246	26081	5

2070990	1990	1830	73079	12895	2795	0	4370	25341	5

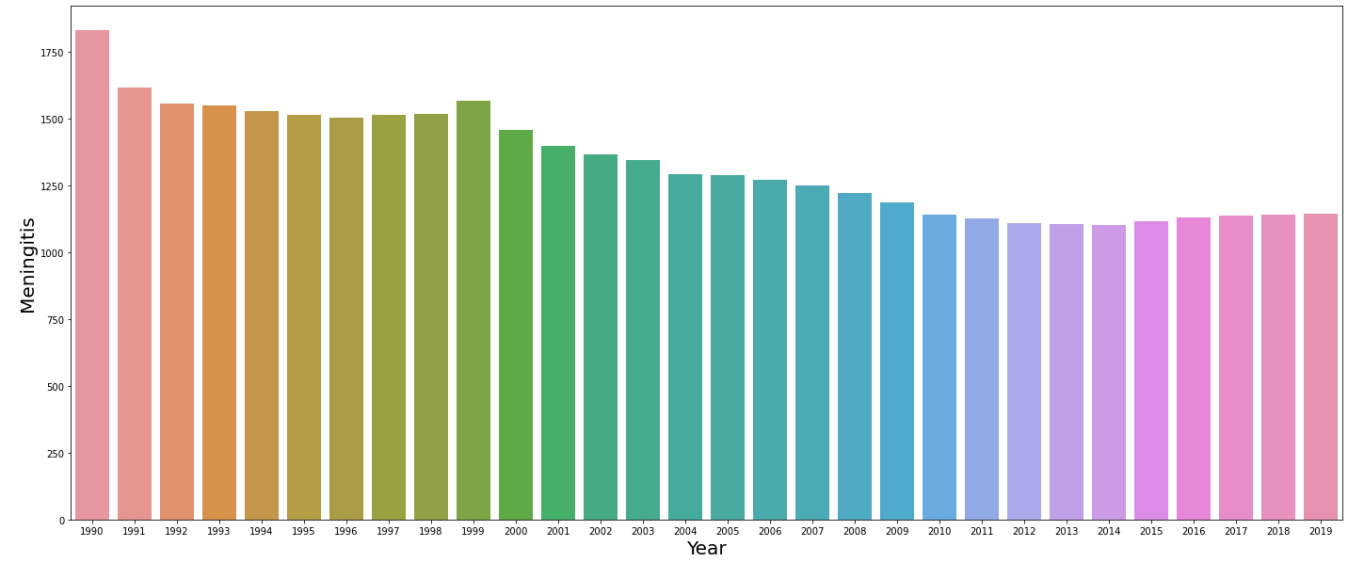
30 rows × 33 columns

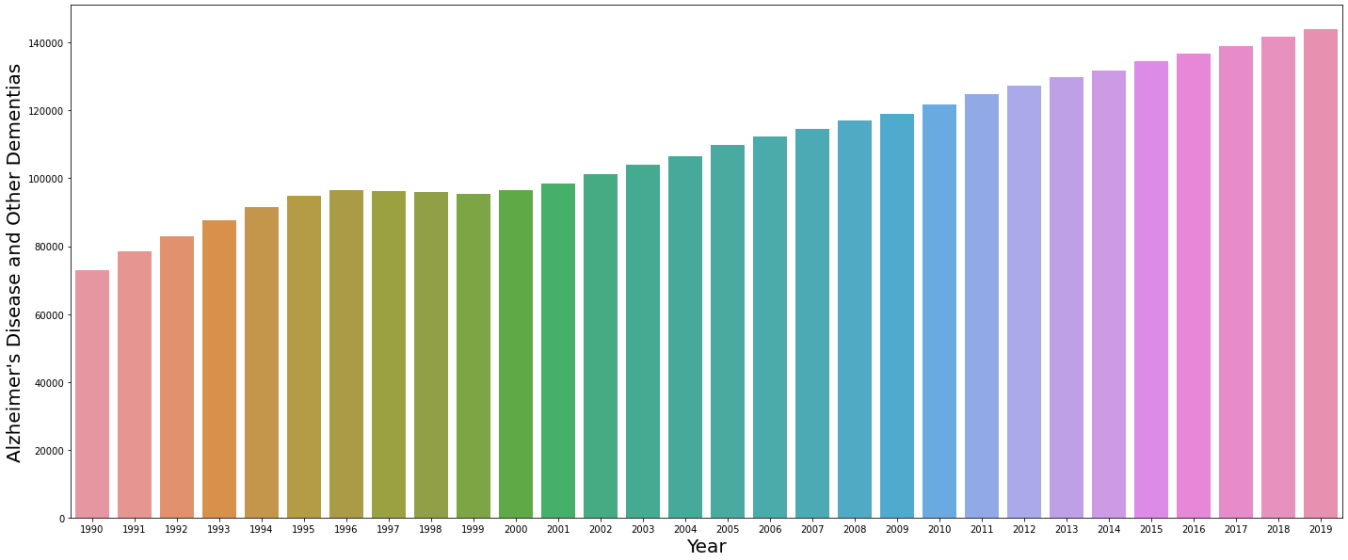
Segregating on the basis of Cause of deaths:

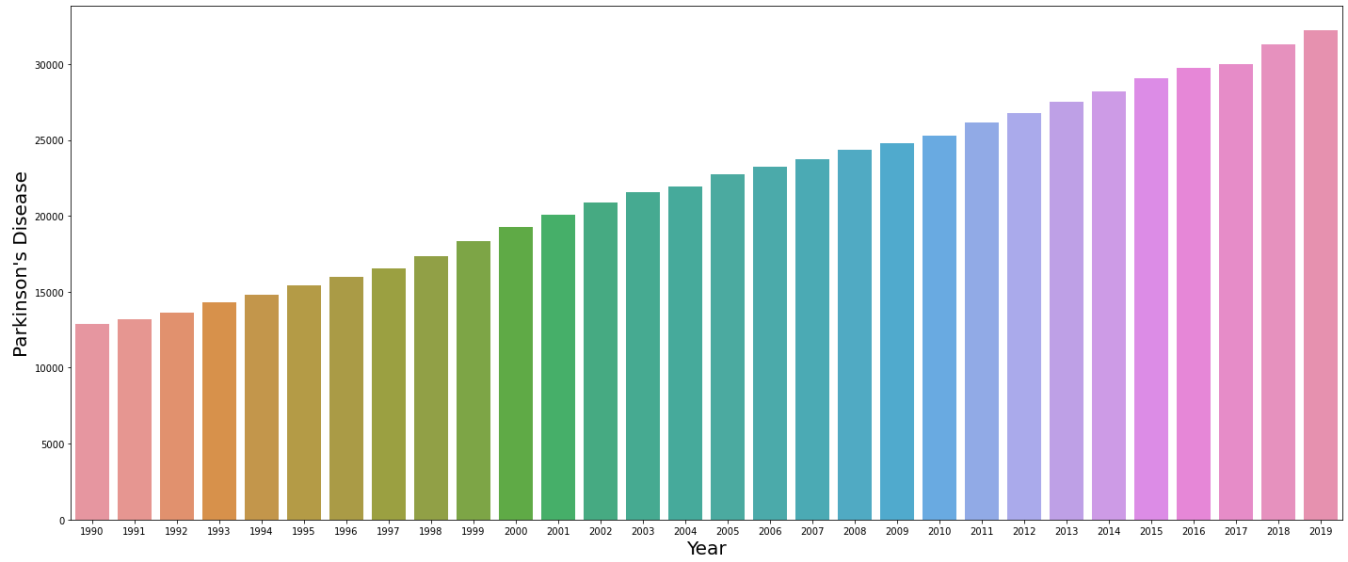
```
# Plot beween 'Year' & "Deaths Caused by Brain related Diseases":
United_States_Total_Deaths.plot(x='Year', y=['Meningitis',"Alzheimer's Disease and Other Dementias","Parkinson's Disease"],
                        figsize=(10,5), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Brain related Diseases" in United_States',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('United_States_Deaths',fontsize =16)
plt.show()
```




```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = United_States_Total_Deaths[['Meningitis',"Alzheimer's Disease and Other Dementias","Parkinson's Disease"]]
X = United_States_Total_Deaths['Year']
for col in y:
    if plotnumber<=8:
        plt.subplot(3,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

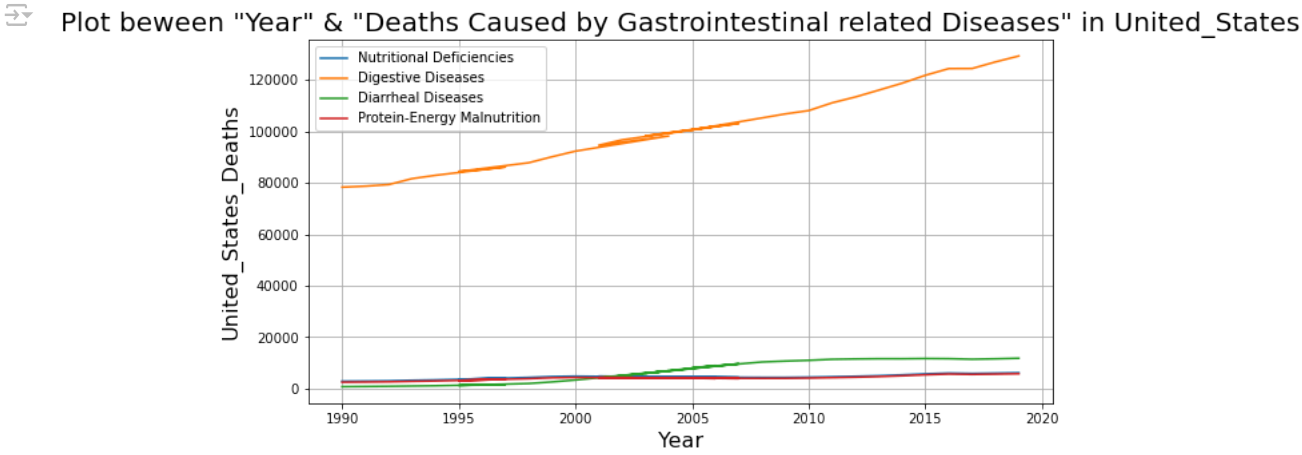

Meningitis
Year


Alzheimer's Disease and Other Dementias
Year

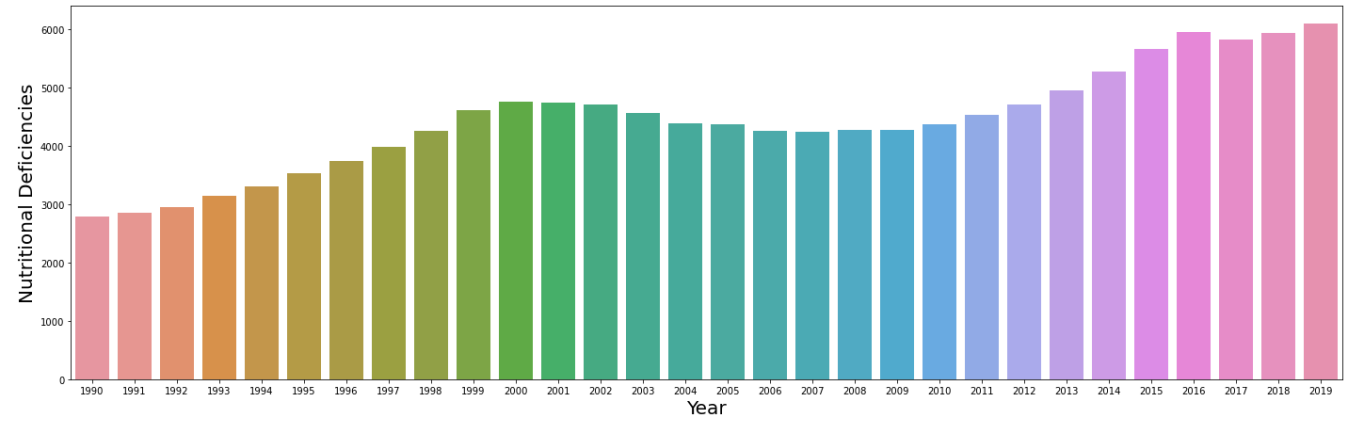

Parkinson's Disease
Year

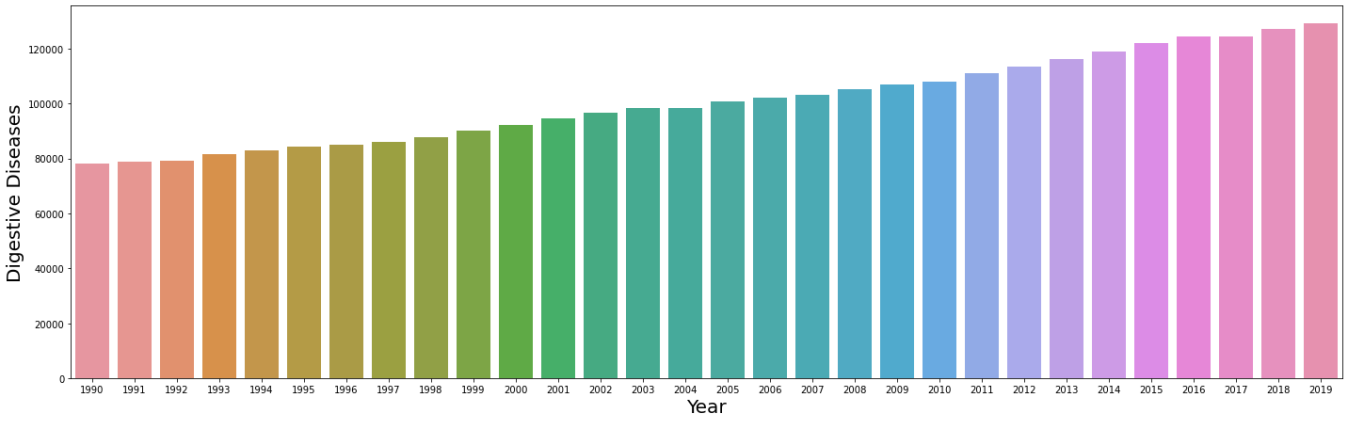
- Deaths in United_States due to **Alzheimer's Disease and Other Dementias** is becoming more dominant among all Deaths Caused by Brain related Diseases with every incrementing year.

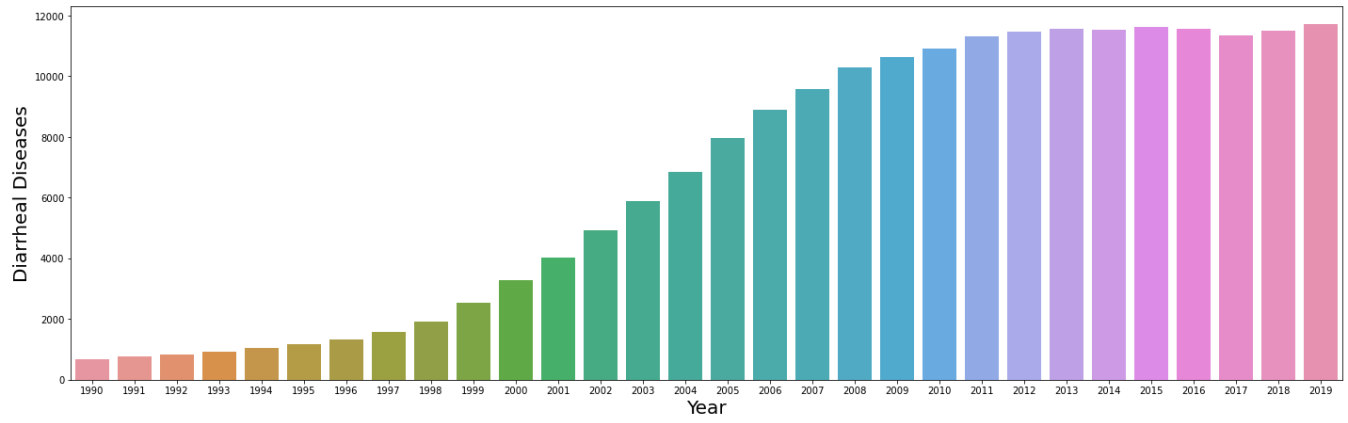
```
# Plot beween 'Year' & "Deaths Caused by Gastrointestinal related Diseases":
United_States_Total_Deaths.plot(x='Year', y=['Nutritional Deficiencies','Digestive Diseases',
        'Diarrheal Diseases','Protein-Energy Malnutrition'],figsize=(10,5), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Gastrointestinal related Diseases" in United_States',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('United_States_Deaths',fontsize =16)
plt.show()
```

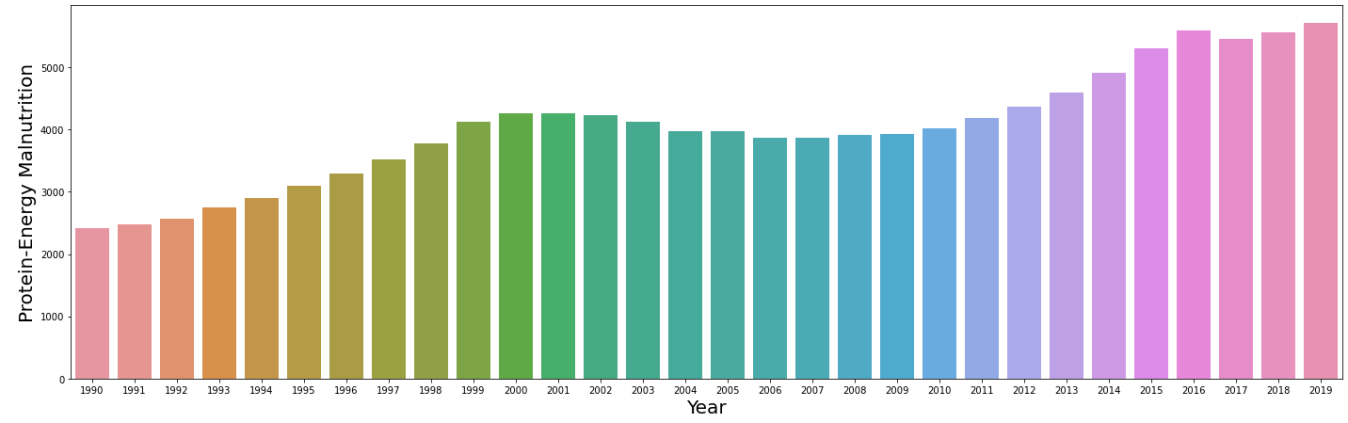



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = United_States_Total_Deaths[['Nutritional Deficiencies','Digestive Diseases',
        'Diarrheal Diseases','Protein-Energy Malnutrition']]
X = United_States_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```


Nutritional Deficiencies
Year

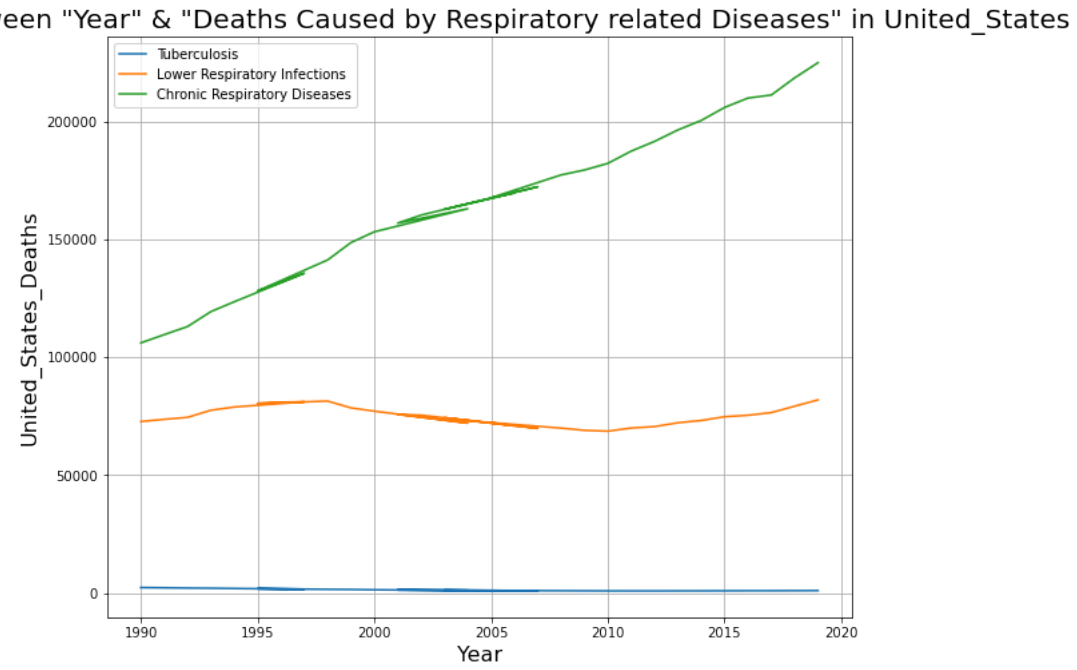

Digestive Diseases
Year


Diarrheal Diseases
Year

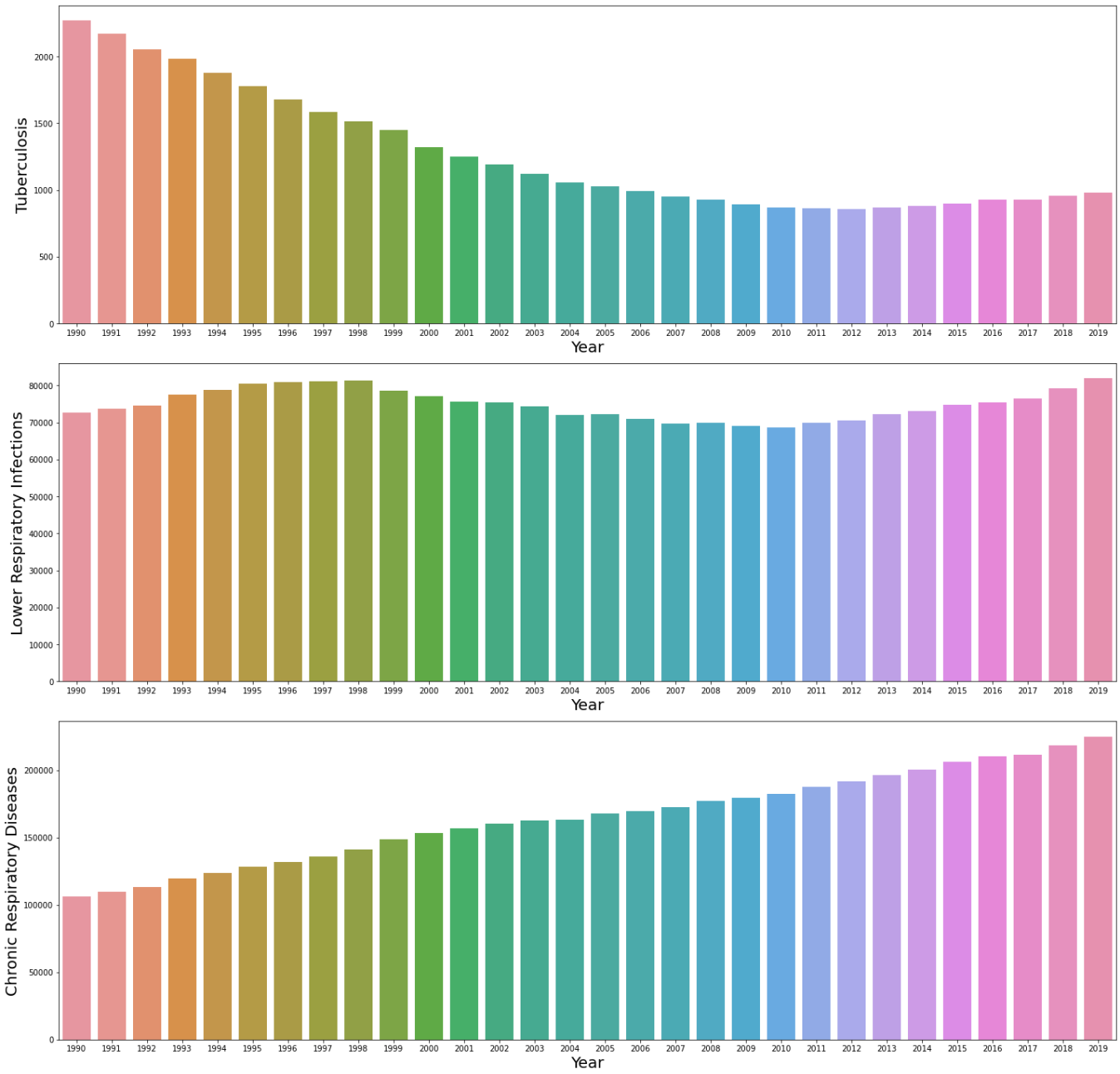

Protein-Energy Malnutrition
Year

- Deaths in United_States due to **Digestive Diseases** is increasing almost exponentially & is Dominant among all Deaths Caused by Gastrointestinal related Diseases with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by Respiratory related Diseases":
United_States_Total_Deaths.plot(x='Year', y=['Tuberculosis','Lower Respiratory Infections',
                                  'Chronic Respiratory Diseases'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in United_States',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('United_States_Deaths',fontsize =16)
plt.show()
```



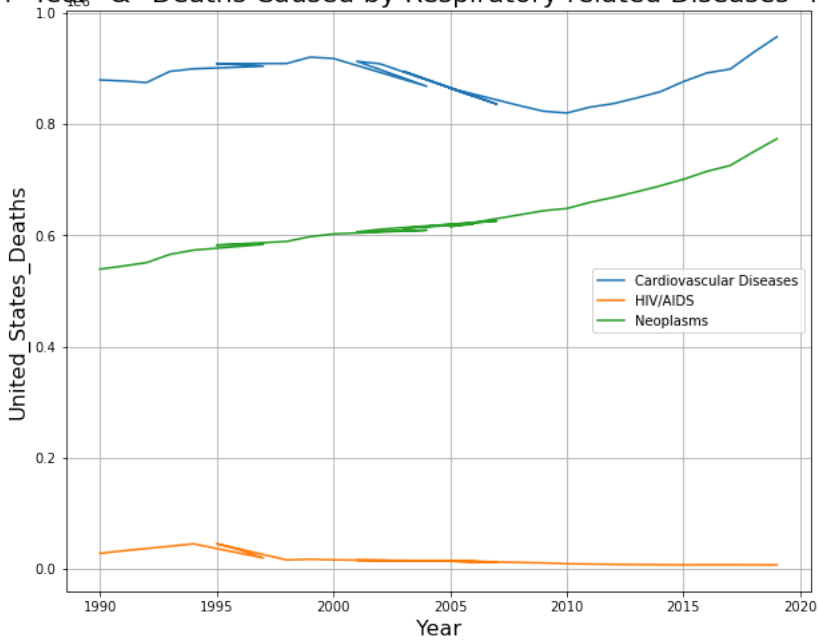

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = United_States_Total_Deaths[['Tuberculosis','Lower Respiratory Infections','Chronic Respiratory Diseases']]
X = United_States_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



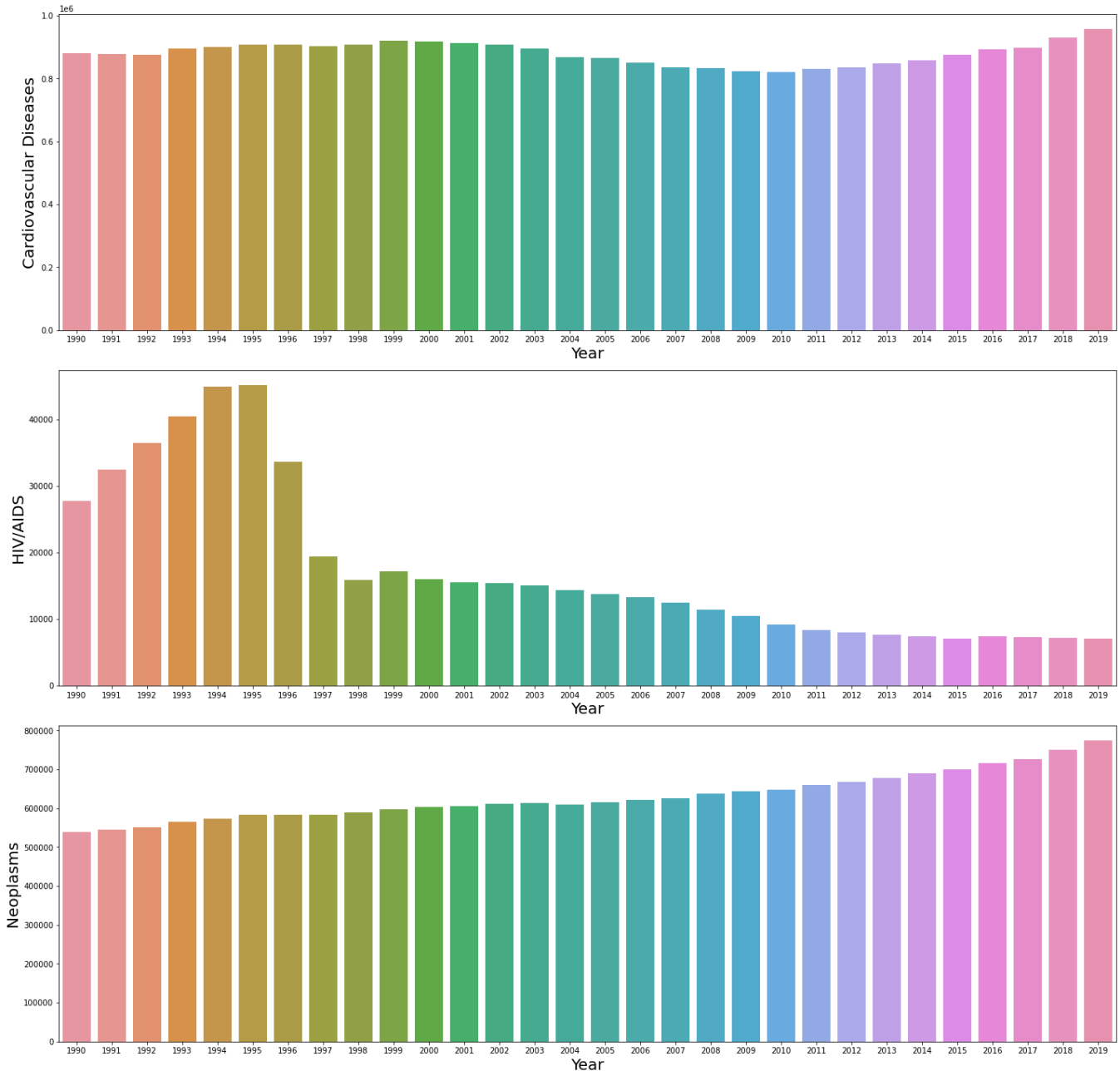
- Deaths in United_States due to **Chronic Respiratory Diseases** is increasing & Dominant among all Deaths Caused by Respiratory related Diseases with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by 'Cardiovascular Diseases','HIV/AIDS' & 'Neoplasms' Diseases":
United_States_Total_Deaths.plot(x='Year', y=['Cardiovascular Diseases','HIV/AIDS','Neoplasms'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in United_States',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('United_States_Deaths',fontsize =16)
plt.show()
```

Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in United_States

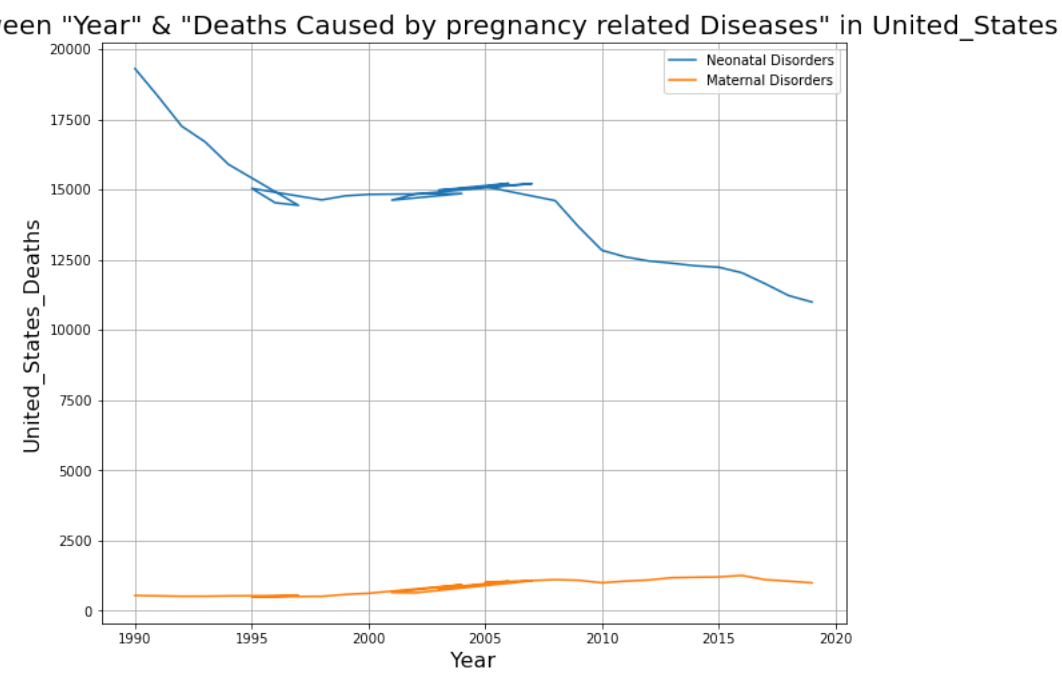



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = United_States_Total_Deaths[['Cardiovascular Diseases','HIV/AIDS','Neoplasms']]
X = United_States_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

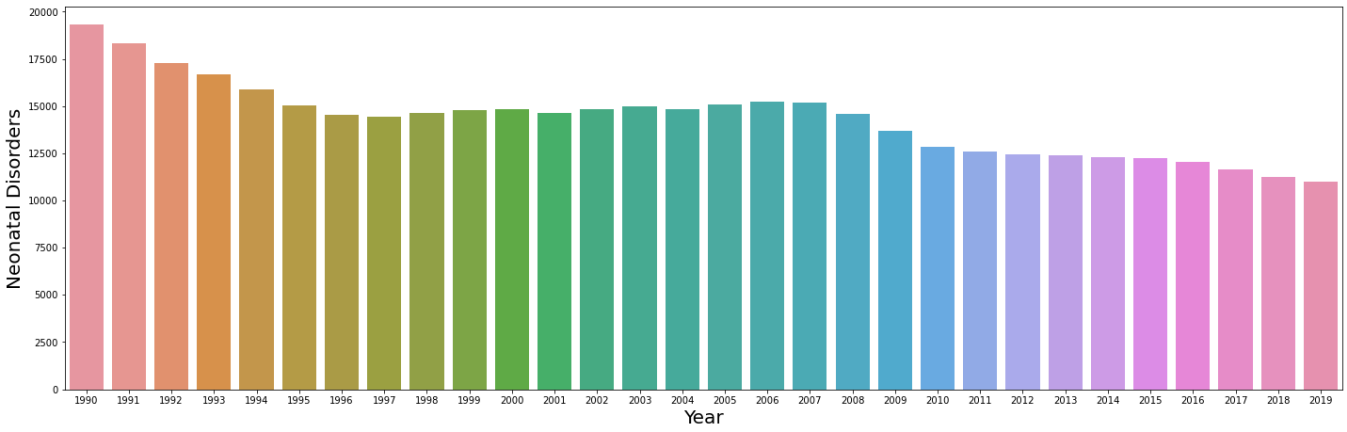


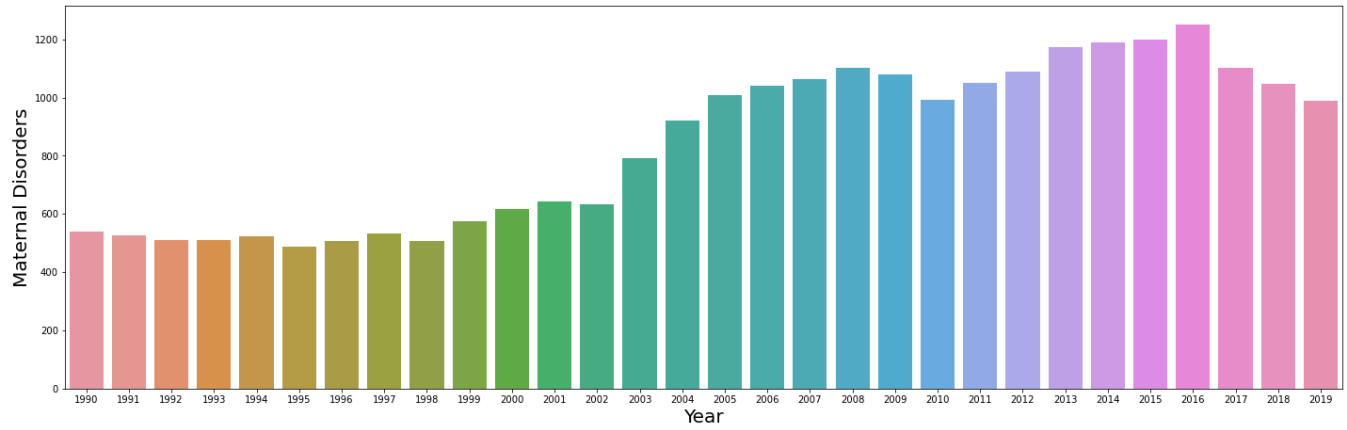
- Deaths in United_States due to **Cardiovascular Diseases** is most Dominant among all Deaths Caused with every incrementing year. Also deaths due to **Neoplasms(Cancer)** is increasing.

```
# Plot beween 'Year' & "Deaths Caused by pregnancy related Diseases":
United_States_Total_Deaths.plot(x='Year', y=['Neonatal Disorders','Maternal Disorders'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by pregnancy related Diseases" in United_States',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('United_States_Deaths',fontsize =16)
plt.show()
```



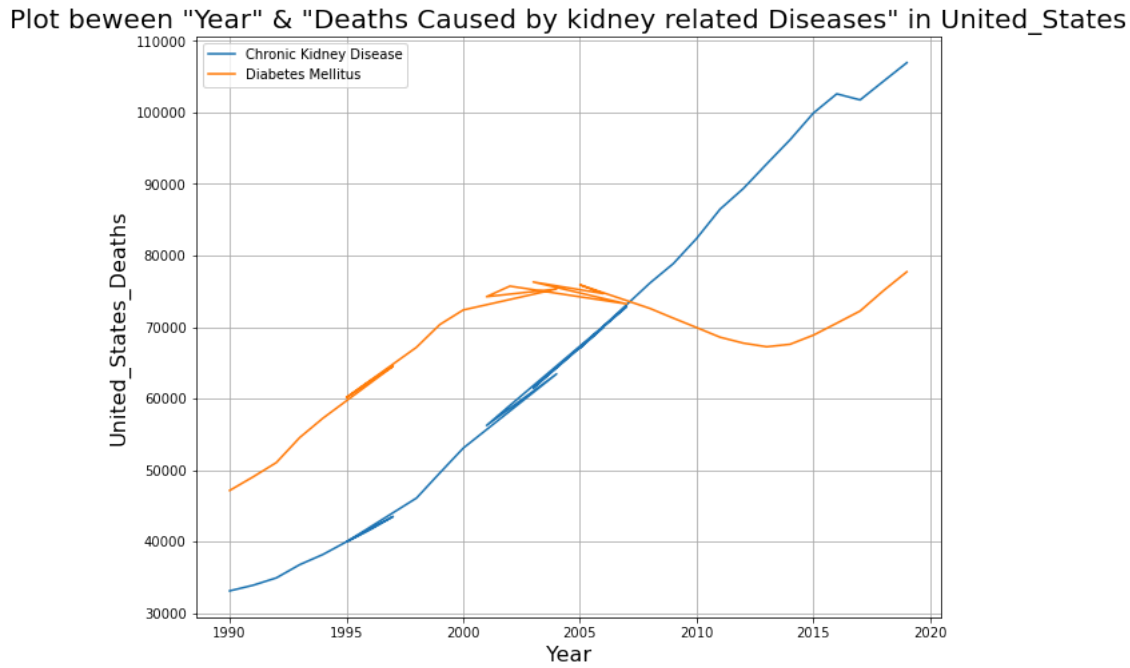

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = United_States_Total_Deaths[['Neonatal Disorders','Maternal Disorders']]
X = United_States_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



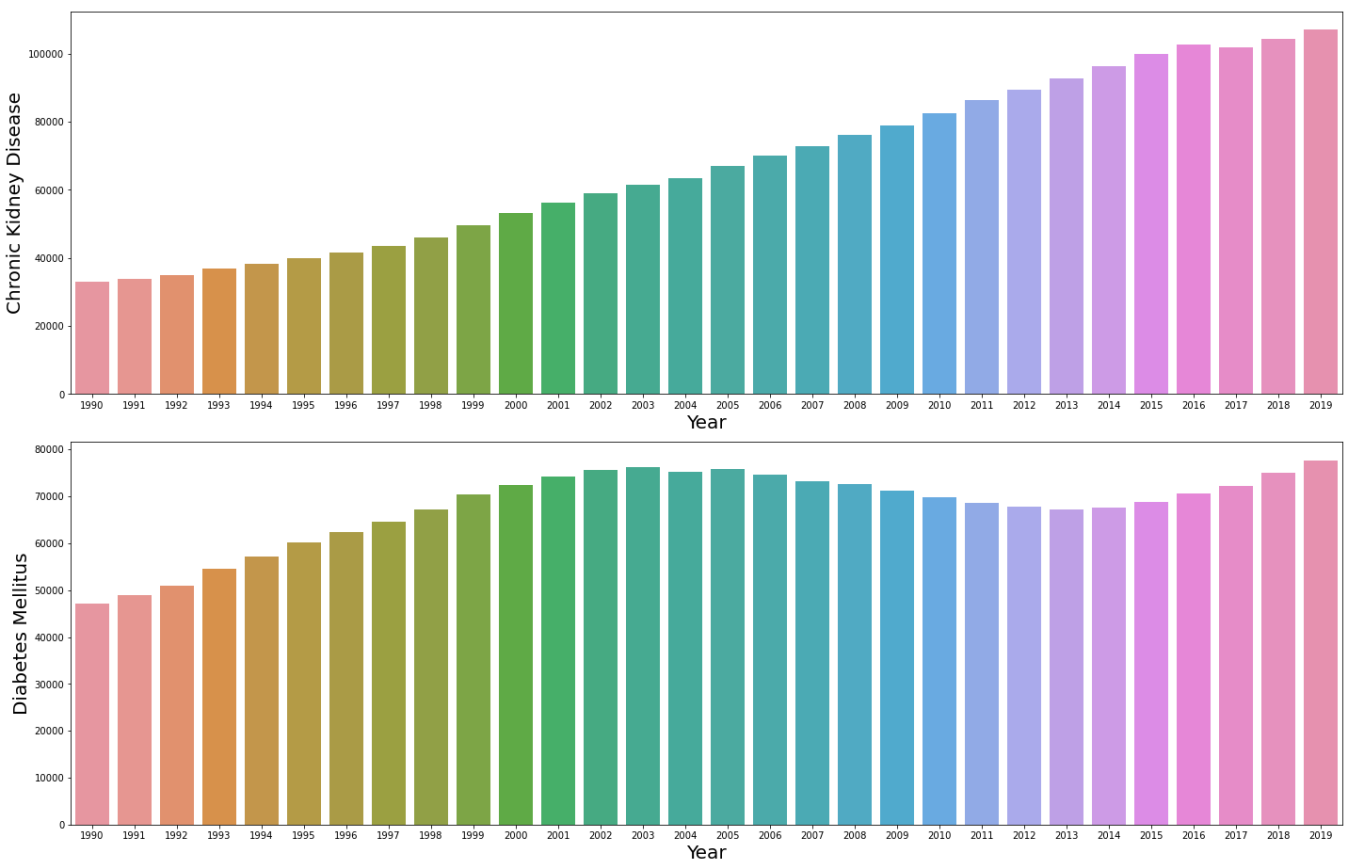


- Deaths in United_States due to **Pregnancy related** issues is **on decline** with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by kidney related Diseases":
United_States_Total_Deaths.plot(x='Year', y=['Chronic Kidney Disease','Diabetes Mellitus'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by kidney related Diseases" in United_States',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('United_States_Deaths',fontsize =16)
plt.show()
```

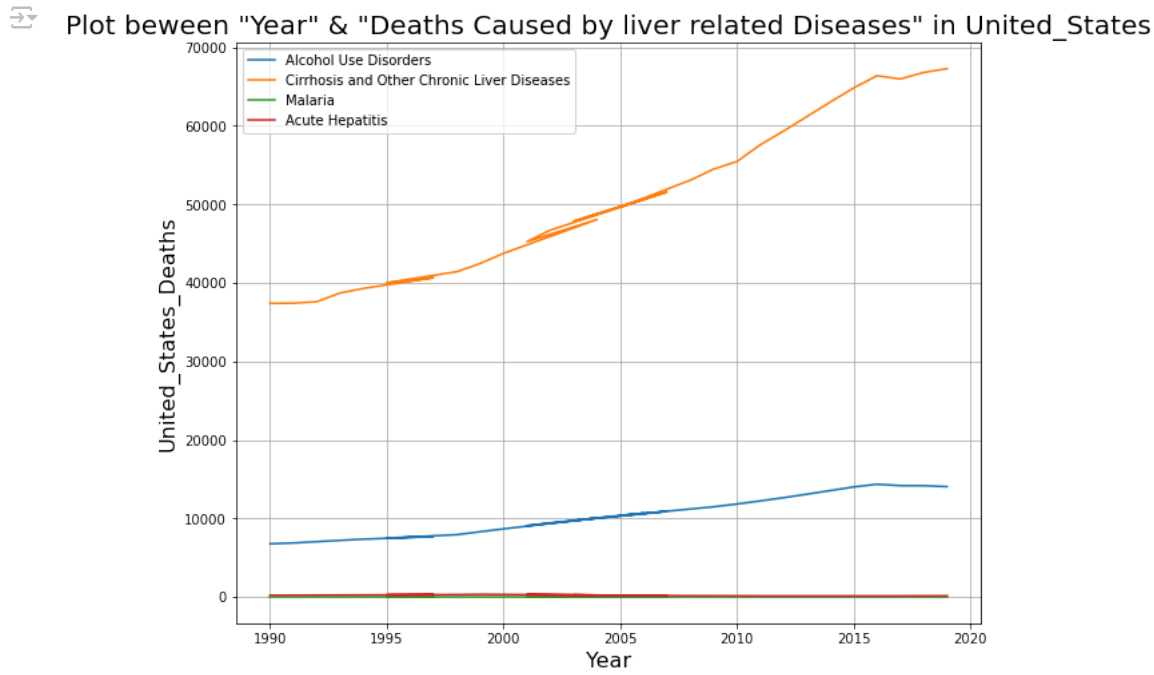


```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = United_States_Total_Deaths[['Chronic Kidney Disease','Diabetes Mellitus']]
X = United_States_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

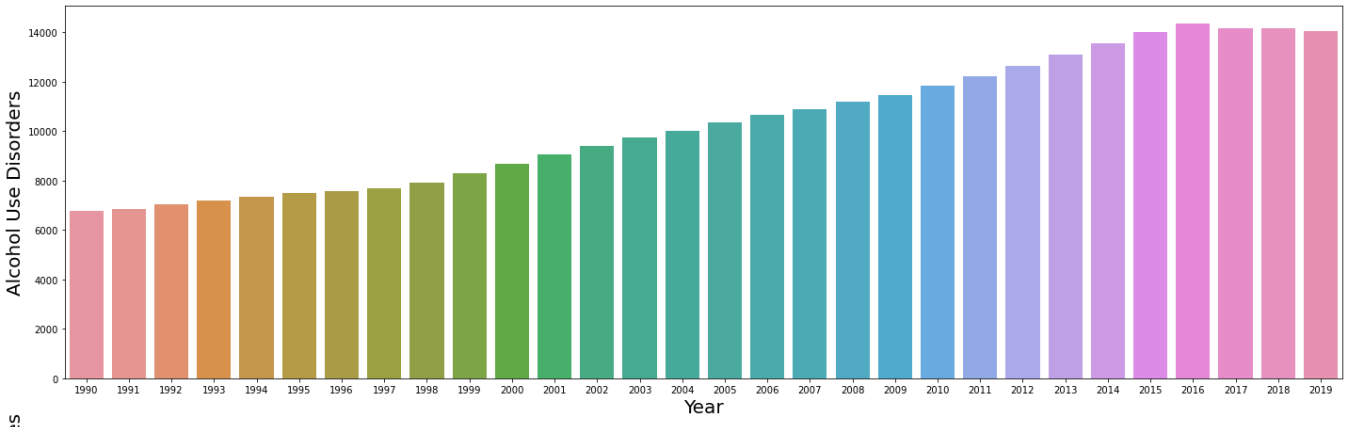


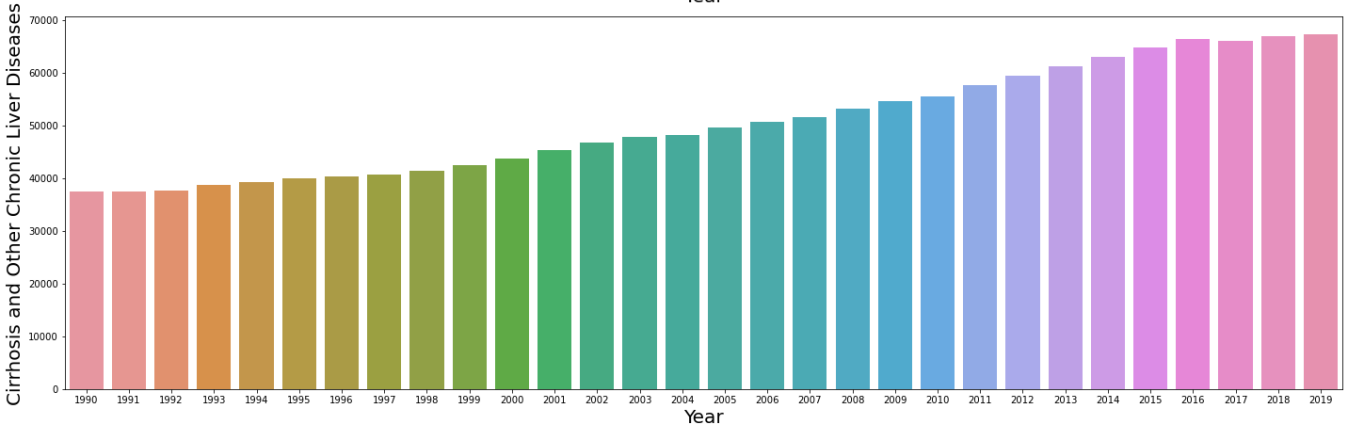
- Deaths in United_States due to **all kidney related Diseases** is Dominant with every **incrementing** year.

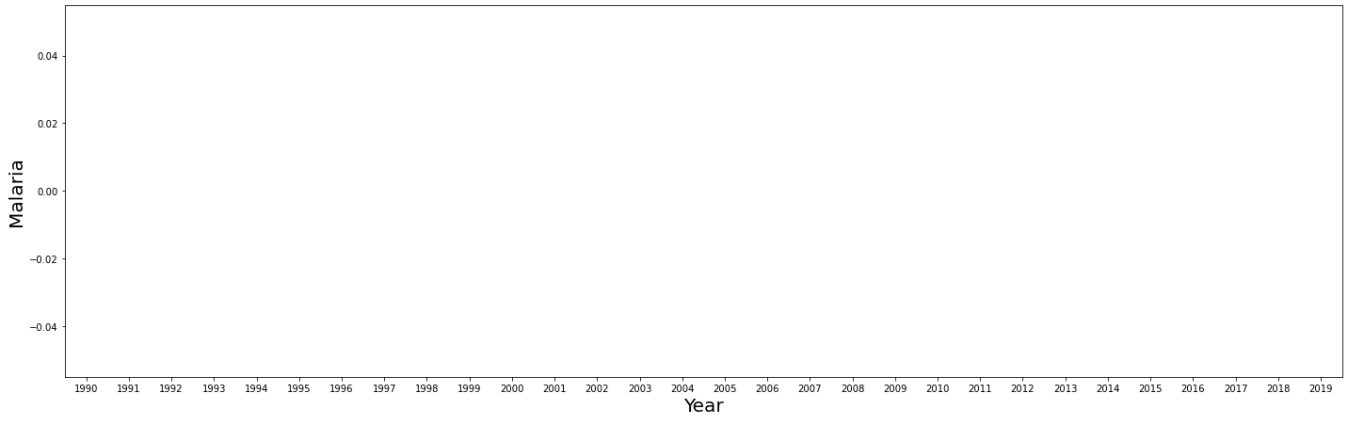
```
# Plot beween 'Year' & "Deaths Caused by liver related Diseases":
United_States_Total_Deaths.plot(x='Year', y=['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases',
                                'Malaria','Acute Hepatitis'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by liver related Diseases" in United_States',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('United_States_Deaths',fontsize =16)
plt.show()
```

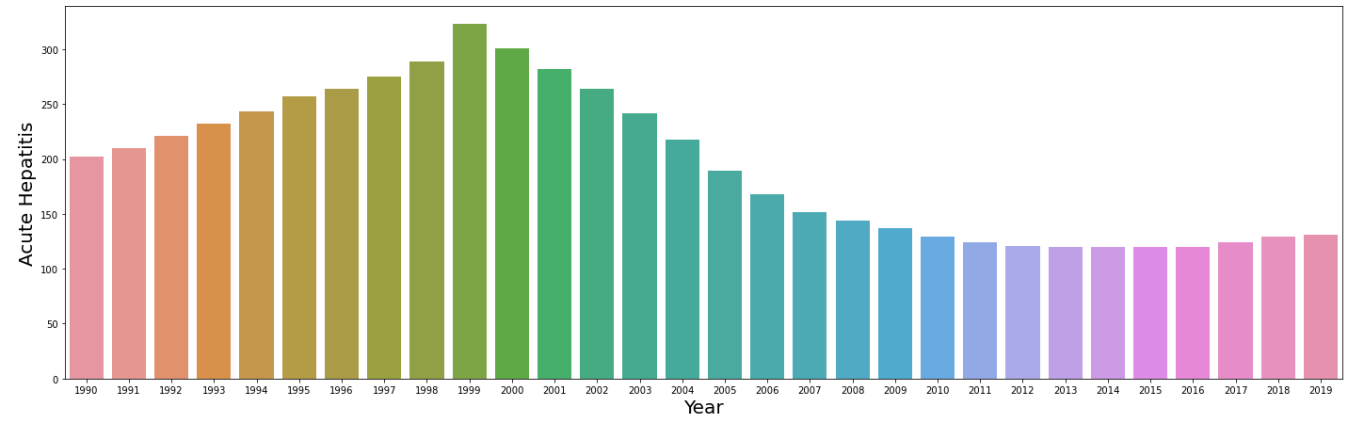



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = United_States_Total_Deaths[['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases',
                        'Malaria','Acute Hepatitis']]
X = United_States_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```


Alcohol Use Disorders
Year

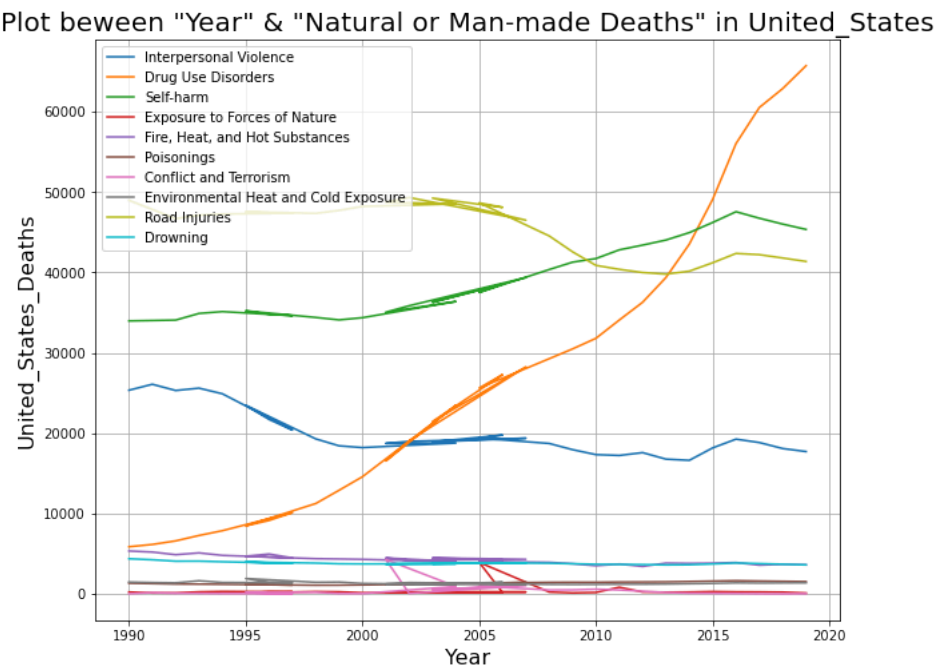

Cirrhosis and Other Chronic Liver Diseases
Year


Malaria
Year

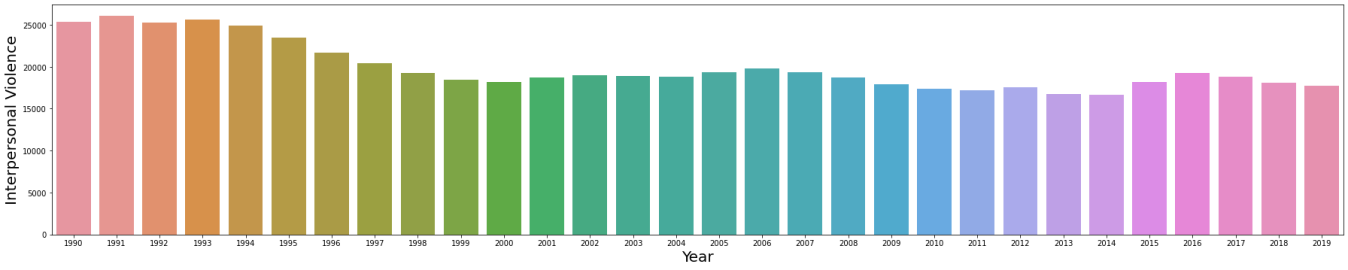

Acute Hepatitis
Year

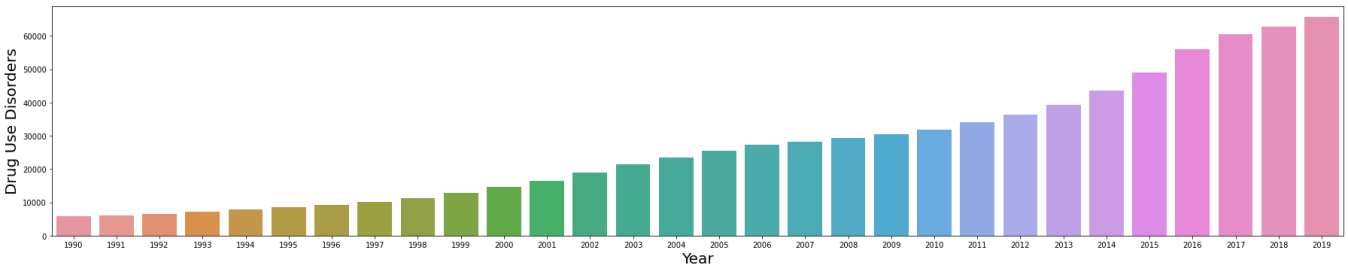
- Deaths in United_States due to **Cirrhosis and Other Chronic Liver Diseases** is Dominant among all Deaths Caused by **liver** related Diseases with every incrementing year.

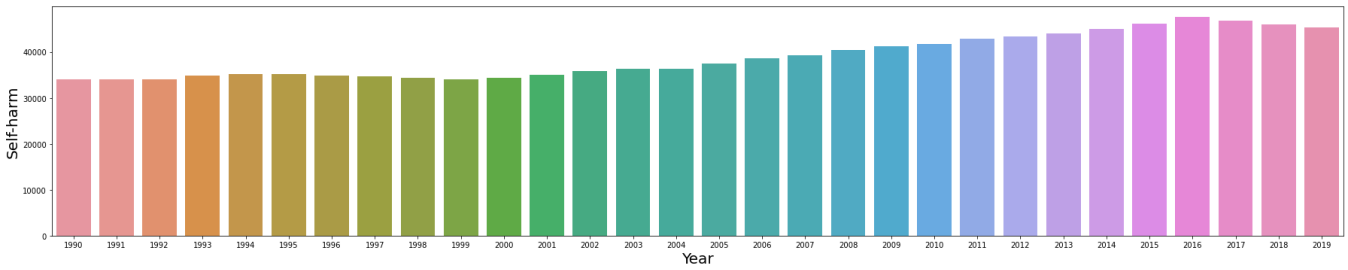
```
# Plot beween 'Year' & "Natural or Man-made Deaths":
United_States_Total_Deaths.plot(x='Year', y=['Interpersonal Violence','Drug Use Disorders',
                                    'Self-harm', 'Exposure to Forces of Nature',
                                    "Fire, Heat, and Hot Substances",
                                    'Poisonings','Conflict and Terrorism',
                                    'Environmental Heat and Cold Exposure',
                                    'Road Injuries','Drowning'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Natural or Man-made Deaths" in United_States',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('United_States_Deaths',fontsize =16)
plt.show()
```

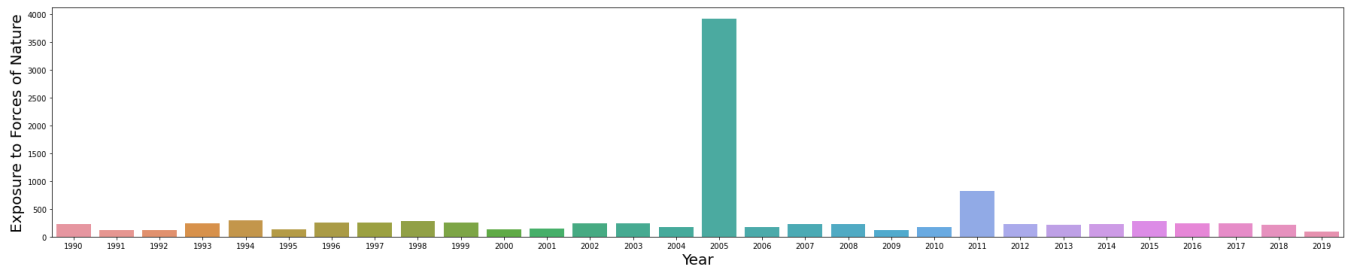


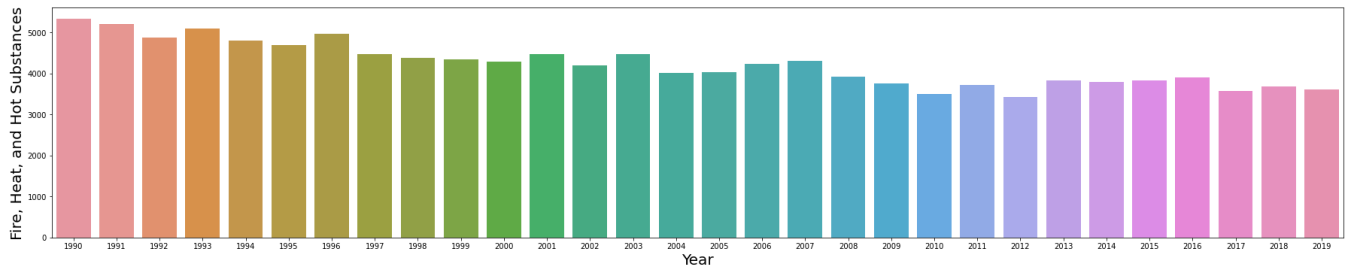

```
plt.figure(figsize=(25,50),facecolor='white')
plotnumber=1
y = United_States_Total_Deaths[['Interpersonal Violence','Drug Use Disorders',
                                    'Self-harm', 'Exposure to Forces of Nature',
                                    "Fire, Heat, and Hot Substances",
                                    'Poisonings','Conflict and Terrorism',
                                    'Environmental Heat and Cold Exposure',
                                    'Road Injuries','Drowning']]
X = United_States_Total_Deaths['Year']
for col in y:
    if plotnumber<=10:
        plt.subplot(10,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

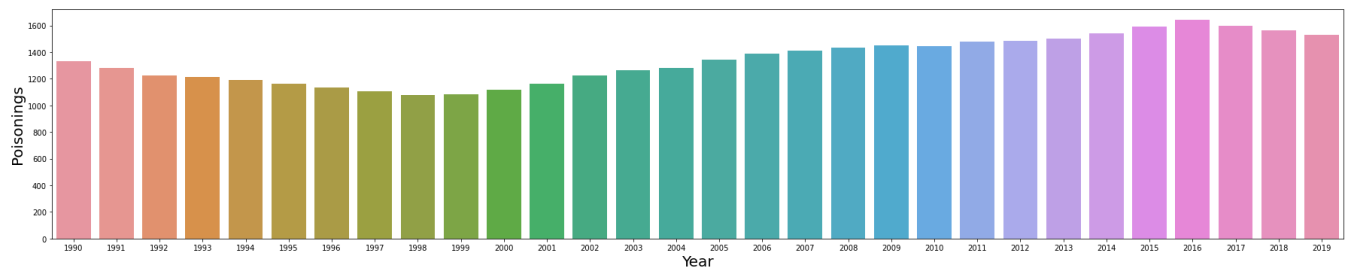

Interpersonal Violence
Year

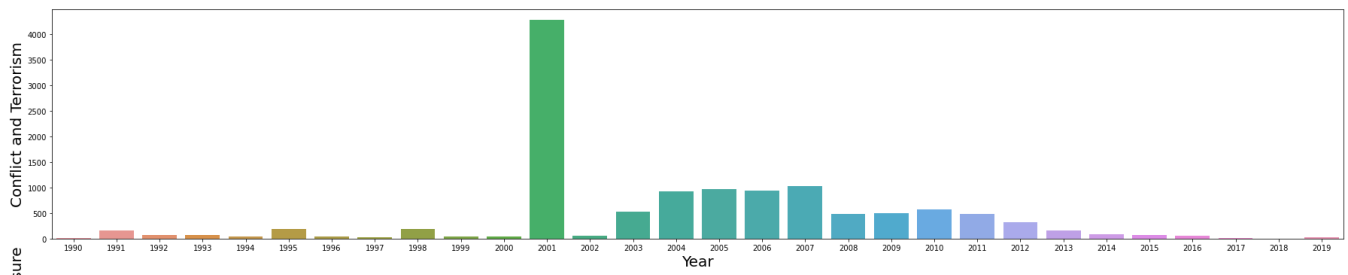

Drug Use Disorders
Year

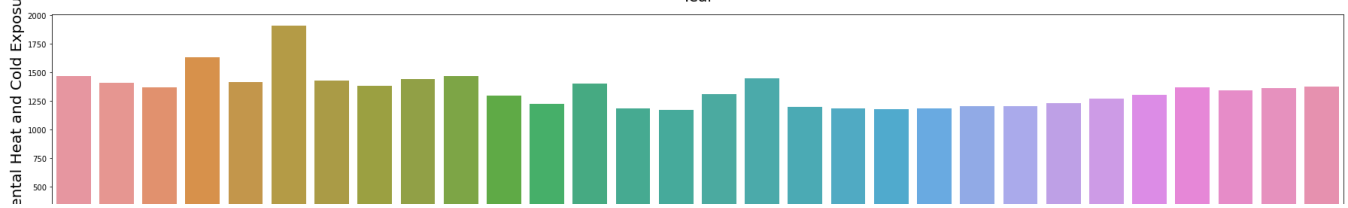

Self-harm
Year

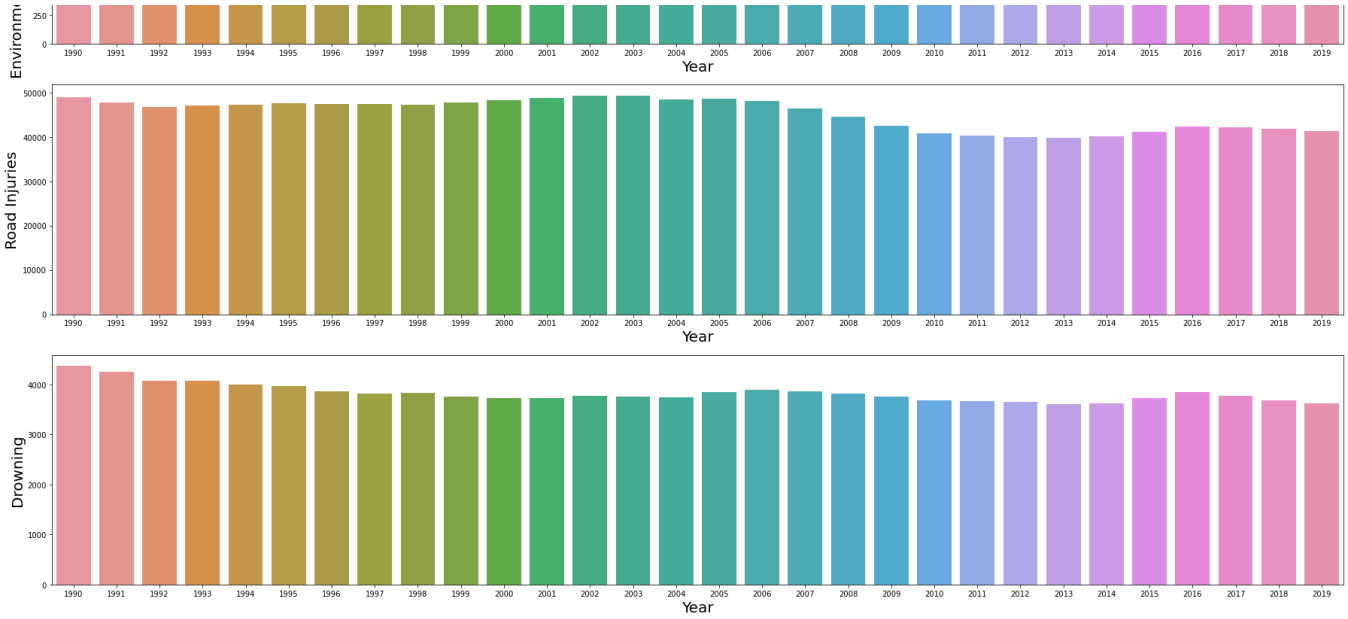

Exposure to Forces of Nature
Year


Fire, Heat, and Hot Substances
Year


Poisonings
Year


Conflict and Terrorism
Year


ental Heat and Cold Exposure



• Deaths in United States due to Natural or Man-made Deaths are summarised below:

 Deaths due to 'Interpersonal Violence' is decreasing with every incrementing year.

 Deaths due to 'Drug Use Disorders' is incrementing with every incrementing year.

 Deaths due to 'Self-harm' is almost constant with every incrementing year.

 Deaths due to 'Exposure to Forces of Nature' is highest in 2005.

 Deaths due to "Fire, Heat, and Hot Substances" is constant with every incrementing year.

 Deaths due to 'Poisonings', is slightly increasing with every incrementing year.

 Deaths due to 'Conflict and Terrorism' is declining with every incrementing year, except in 2001 due to WTC Attack.

 Deaths due to 'Environmental Heat and Cold Exposure' is constant with every incrementing year.

 Deaths due to 'Road Injuries' is slightly decreasing with every incrementing year.

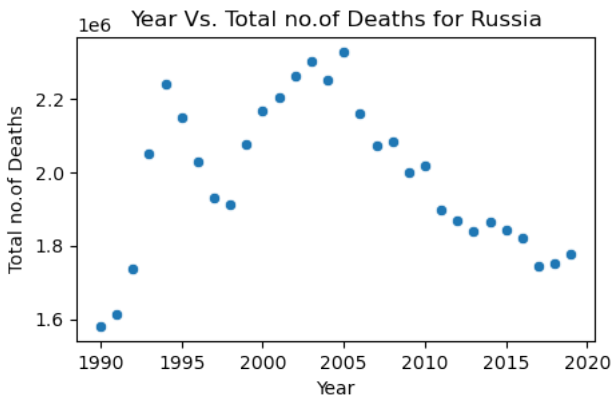
 Deaths due to 'Drowning' is slightly declining with every incrementing year.

4. For Russia:

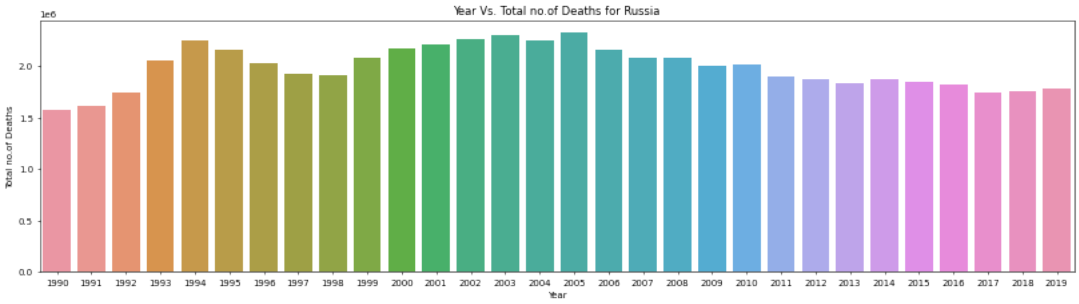
```
# Russia: "Total_Deaths" against "Year"

Russia_Total_Deaths = df[df['Country/Territory']=='Russia'].sort_values(by='Total_Deaths',ascending=False)
Russia_Total_Deaths[["Country/Territory",'Year','Total_Deaths']]
```

	Country/Territory	Year	Total_Deaths
4485	Russia	2005	2330502
4483	Russia	2003	2301507
4482	Russia	2002	2264207
4484	Russia	2004	2250901
4474	Russia	1994	2242147
4481	Russia	2001	2204306
4480	Russia	2000	2169598
4486	Russia	2006	2161926
4475	Russia	1995	2151195
4488	Russia	2008	2083629
4479	Russia	1999	2078659
4487	Russia	2007	2074246
4473	Russia	1993	2051078
4476	Russia	1996	2028083
4490	Russia	2010	2019565
4489	Russia	2009	2000625
4477	Russia	1997	1929473
4478	Russia	1998	1913563
4491	Russia	2011	1898941
4492	Russia	2012	1867010
4494	Russia	2014	1865234
4495	Russia	2015	1842512
4493	Russia	2013	1838368
4496	Russia	2016	1820655
4499	Russia	2019	1777223
4498	Russia	2018	1752517
4497	Russia	2017	1744289
4472	Russia	1992	1737492
4471	Russia	1991	1612808
4470	Russia	1990	1578896

```
# Russia: "Total_Deaths" v/s "Year"

plt.figure(figsize=(5,3),dpi=100)
sns.scatterplot(data=Russia_Total_Deaths, x='Year', y='Total_Deaths')
plt.xlabel("Year")
plt.ylabel("Total no.of Deaths")
plt.title("Year Vs. Total no.of Deaths for Russia")
plt.show();
```



```
plt.figure(figsize=(20,5),dpi=50)
sns.barplot(data=Russia_Total_Deaths, x='Year', y='Total_Deaths')
plt.xlabel("Year")
plt.ylabel("Total no.of Deaths")
plt.title("Year Vs. Total no.of Deaths for Russia")
plt.show();
```



- There is a clear decrease in Total No. of Deaths after 2005 with each incrementing year in Russia.

Now, let's Investigate the Major cause of Deaths in Russia:

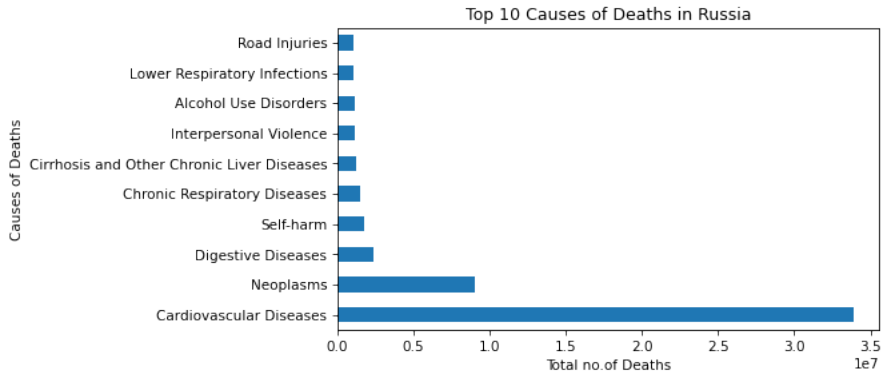
```
# Get maximum deaths causing Disease wrt Year:

Russia_Total_Deaths.set_index(["Country/Territory", "Total_Deaths"], inplace = True) # making County & Total_Deaths as index.
Russia_Total_Deaths['Max'] = Russia_Total_Deaths.idxmax(axis=1) # getting desease for max cause of deaths in Russia
Russia_Total_Deaths[['Year','Max']] # displaying reqd data only.
```

		Year	Max
Country/Territory	Total_Deaths		
Russia	**2330502**	2005	Cardiovascular Diseases
	2301507	2003	Cardiovascular Diseases
	2264207	2002	Cardiovascular Diseases
	2250901	2004	Cardiovascular Diseases
	2242147	1994	Cardiovascular Diseases
	2204306	2001	Cardiovascular Diseases
	2169598	2000	Cardiovascular Diseases
	2161926	2006	Cardiovascular Diseases
	2151195	1995	Cardiovascular Diseases
	2083629	2008	Cardiovascular Diseases
	2078659	1999	Cardiovascular Diseases
	2074246	2007	Cardiovascular Diseases
	2051078	1993	Cardiovascular Diseases
	2028083	1996	Cardiovascular Diseases
	2019565	2010	Cardiovascular Diseases
	2000625	2009	Cardiovascular Diseases
	1929473	1997	Cardiovascular Diseases
	1913563	1998	Cardiovascular Diseases
	1898941	2011	Cardiovascular Diseases
	1867010	2012	Cardiovascular Diseases
	1865234	2014	Cardiovascular Diseases
	1842512	2015	Cardiovascular Diseases
	1838368	2013	Cardiovascular Diseases
	1820655	2016	Cardiovascular Diseases
	1777223	2019	Cardiovascular Diseases
	1752517	2018	Cardiovascular Diseases
	1744289	2017	Cardiovascular Diseases
	1737492	1992	Cardiovascular Diseases
	1612808	1991	Cardiovascular Diseases
	1578896	1990	Cardiovascular Diseases

```
# Russia - Top 10 Causes of Deaths:

Russia_10 = Countries_Total_no_of_Deaths_20year_data.sort_values(by='Total_Deaths',ascending =False)
Russia_10.iloc[3].sort_values(ascending=False)[1:11]
```

```
Cardiovascular Diseases                      33903781
Neoplasms                                     9079313
Digestive Diseases                            2398456
Self-harm                                     1793855
Chronic Respiratory Diseases                  1518195
Cirrhosis and Other Chronic Liver Diseases    1233608
Interpersonal Violence                        1215179
Alcohol Use Disorders                         1194672
Lower Respiratory Infections                  1113475
Road Injuries                                 1067225
Name: Russia, dtype: int64
```

```
plt.figure(figsize=(7,4),dpi=75)
Russia_10.iloc[3].sort_values(ascending=False)[1:11].plot(kind='barh')
plt.xlabel("Total no.of Deaths")
plt.ylabel("Causes of Deaths")
plt.title("Top 10 Causes of Deaths in Russia")
plt.show();
```



- **Major Causes of Deaths for Russia are:**

1. Cardiovascular Diseases
2. Neoplasms
3. Digestive Diseases
4. Self-harm
5. Chronic Respiratory Diseases
6. Cirrhosis and Other Chronic Liver Diseases
7. Interpersonal Violence
8. Alcohol Use Disorders
9. Lower Respiratory Infections
10. Road Injuries

```
Russia_Total_Deaths
```

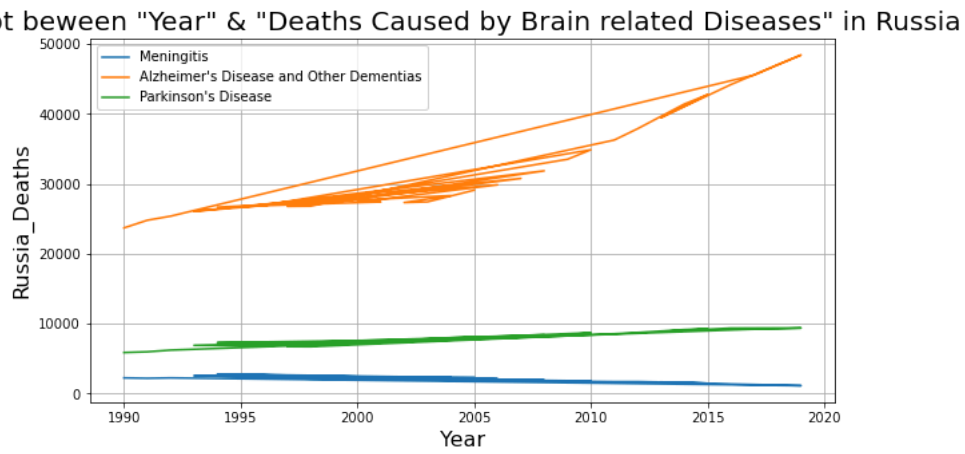
Country/Territory	Total_Deaths	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	Matern Disorde
Russia	**2330502**	2005	2337	29104	8145	622	0	17587	52105	4
	2301507	2003	2402	27435	7832	642	0	17987	55536	5
	2264207	2002	2380	27337	7642	635	0	18248	56916	5
	2250901	2004	2310	28328	7796	607	0	17231	53512	4
	2242147	1994	2803	26677	7286	691	0	21430	63629	8
	2204306	2001	2362	27427	7410	621	0	18579	55879	5
	2169598	2000	2315	27403	7313	624	0	18632	54883	5
	2161926	2006	2121	29869	7992	551	0	15267	43282	4
	2151195	1995	2689	26650	7038	659	0	20808	61876	8
	2083629	2008	1929	31868	8413	490	0	13235	37643	4
	2078659	1999	2230	27413	7157	602	0	18008	52417	5
	2074246	2007	1982	30782	8075	502	0	13946	39587	4
	2051078	1993	2564	26077	6912	623	0	19098	56102	8
	2028083	1996	2387	26918	6846	605	0	18380	52944	6
	2019565	2010	1803	34885	8709	442	0	11690	31777	3
	2000625	2009	1820	33526	8416	453	0	11896	33142	4
	1929473	1997	2203	26786	6729	558	0	16558	47709	6
	1913563	1998	2112	26820	6718	547	0	16448	46589	5
	1898941	2011	1711	36275	8559	402	0	10260	28830	3
	1867010	2012	1714	37893	8735	396	0	9649	27111	2
	1865234	2014	1578	41422	9157	432	0	8945	25463	2
	1842512	2015	1466	42768	9276	456	0	8172	24223	2
	1838368	2013	1661	39475	8865	399	0	9098	26158	2
	1820655	2016	1375	44217	9371	495	0	7619	22308	2
	1777223	2019	1143	48437	9390	456	0	7086	22285	2
	1752517	2018	1180	47040	9260	459	0	7038	21896	2
	1744289	2017	1235	45575	9228	483	0	6884	20474	1
	1737492	1992	2253	25388	6228	516	0	15559	41586	8
	1612808	1991	2199	24804	5987	473	0	14178	30982	9

1578896	1990	2255	23706	5882	465	0	13528	28335	9

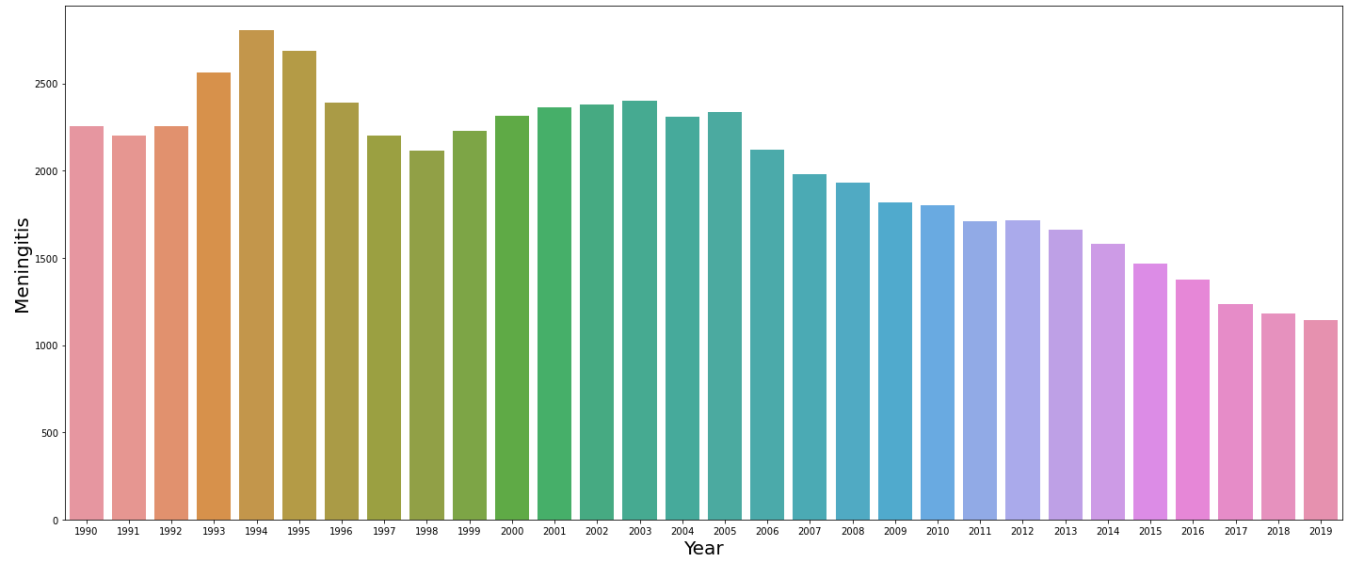
30 rows × 33 columns

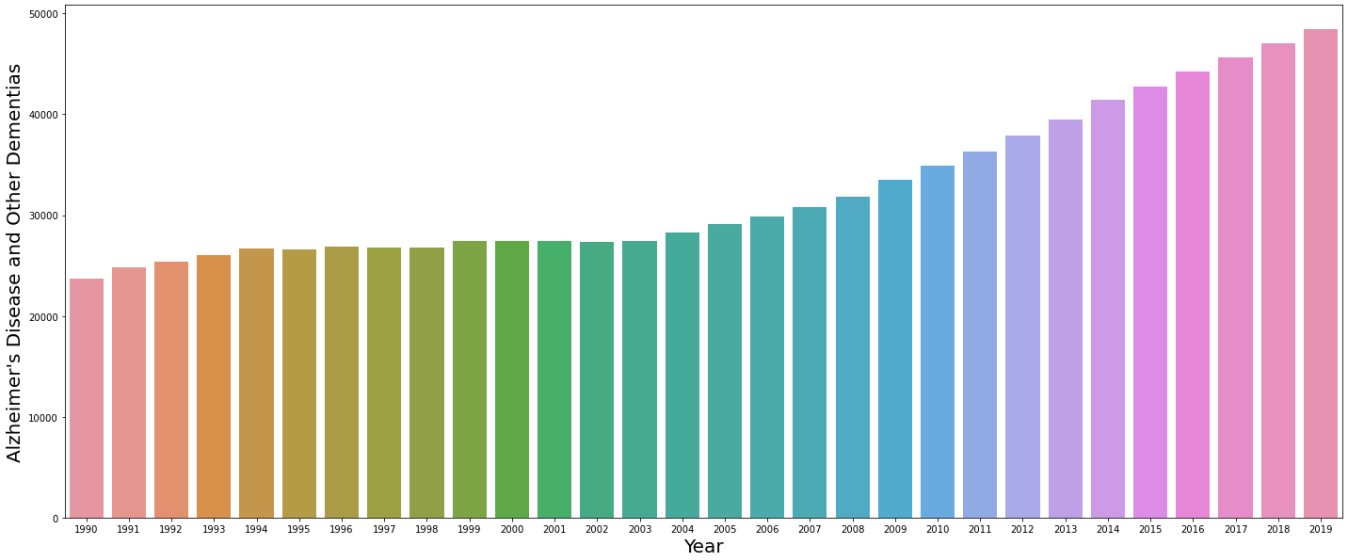
Segregating on the basis of Cause of deaths:

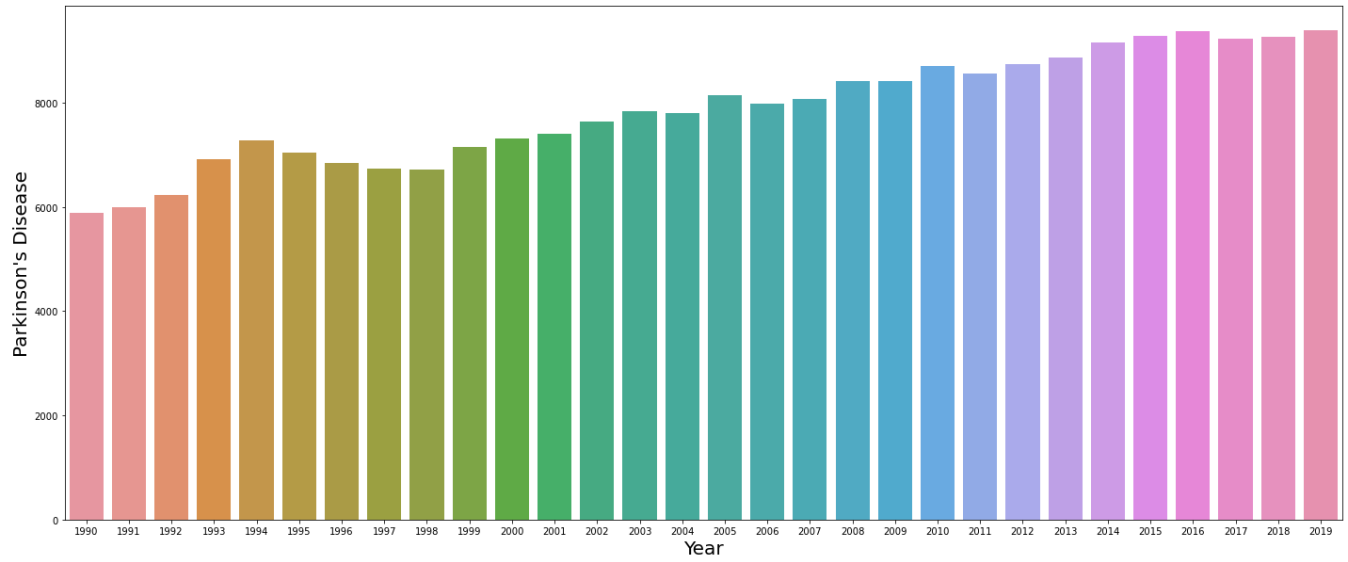
```
# Plot beween 'Year' & "Deaths Caused by Brain related Diseases":
Russia_Total_Deaths.plot(x='Year', y=['Meningitis',"Alzheimer's Disease and Other Dementias","Parkinson's Disease"],
                         figsize=(10,5), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Brain related Diseases" in Russia',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('Russia_Deaths',fontsize =16)
plt.show()
```




```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = Russia_Total_Deaths[['Meningitis',"Alzheimer's Disease and Other Dementias","Parkinson's Disease"]]
X = Russia_Total_Deaths['Year']
for col in y:
    if plotnumber<=8:
        plt.subplot(3,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

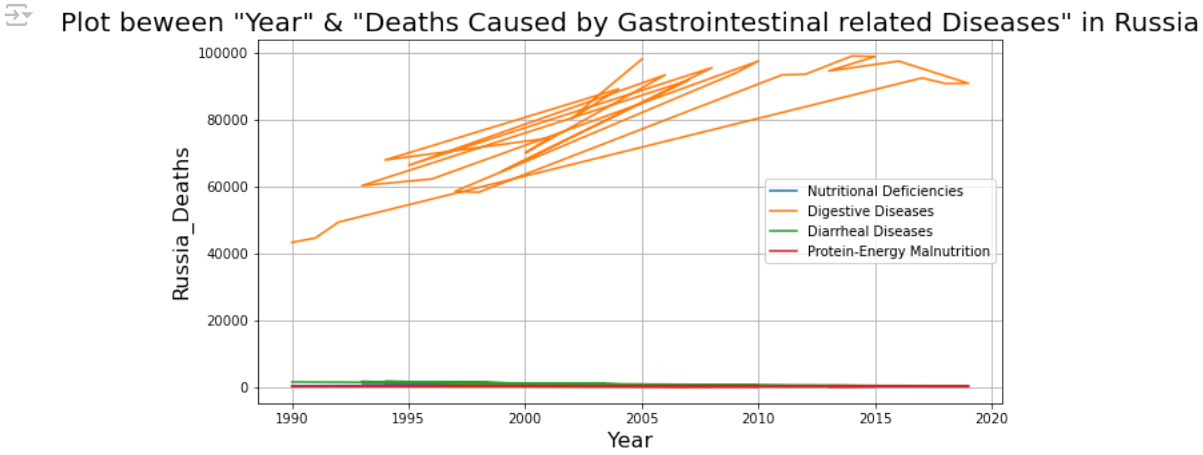

Meningitis
Year


Alzheimer's Disease and Other Dementias
Year

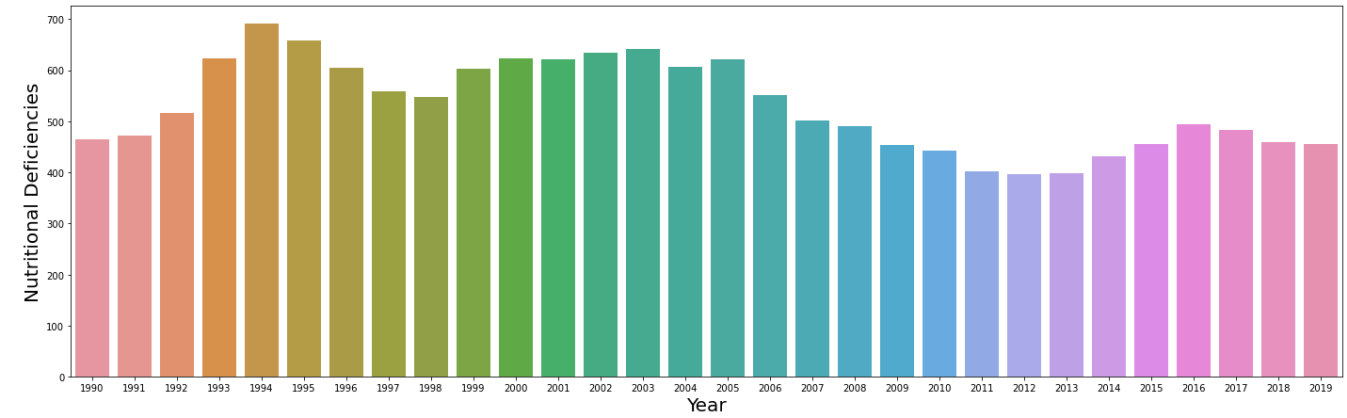

Parkinson's Disease
Year

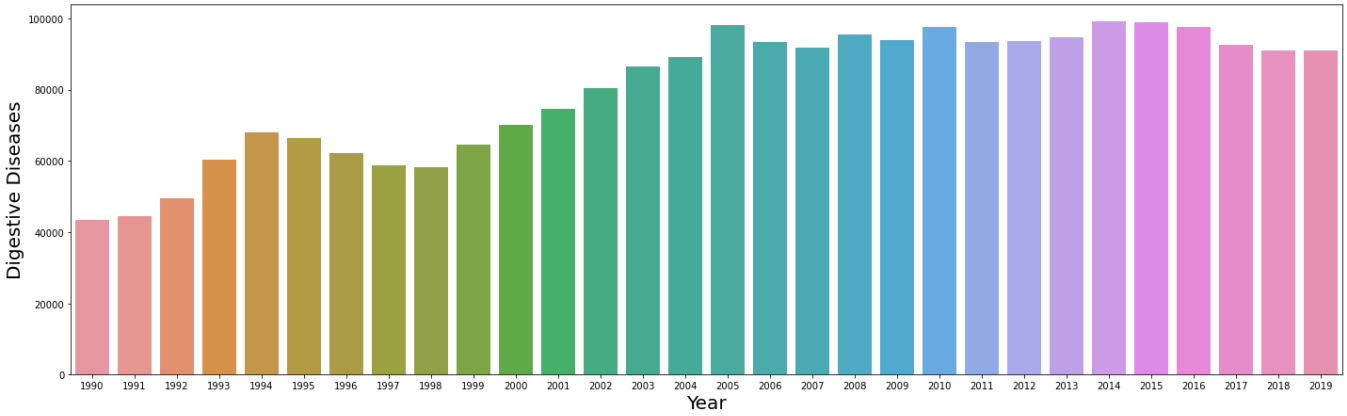
• Deaths in Russia due to Alzheimer's Disease and Other Dementias is becoming more dominant among all Deaths Caused by Brain related Diseases with every incrementing year.

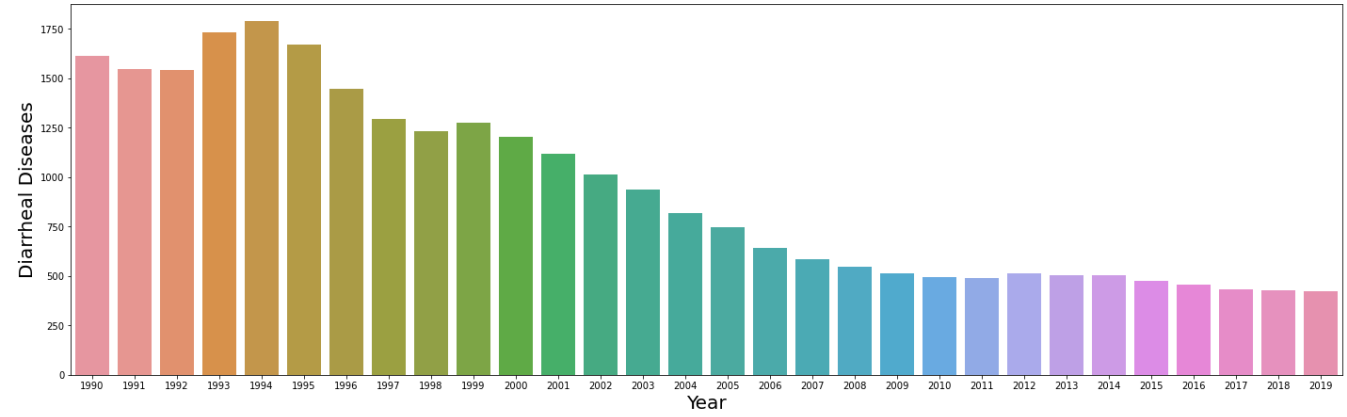
```
# Plot beween 'Year' & "Deaths Caused by Gastrointestinal related Diseases":
Russia_Total_Deaths.plot(x='Year', y=['Nutritional Deficiencies','Digestive Diseases',
        'Diarrheal Diseases','Protein-Energy Malnutrition'],figsize=(10,5), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Gastrointestinal related Diseases" in Russia',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('Russia_Deaths',fontsize =16)
plt.show()
```

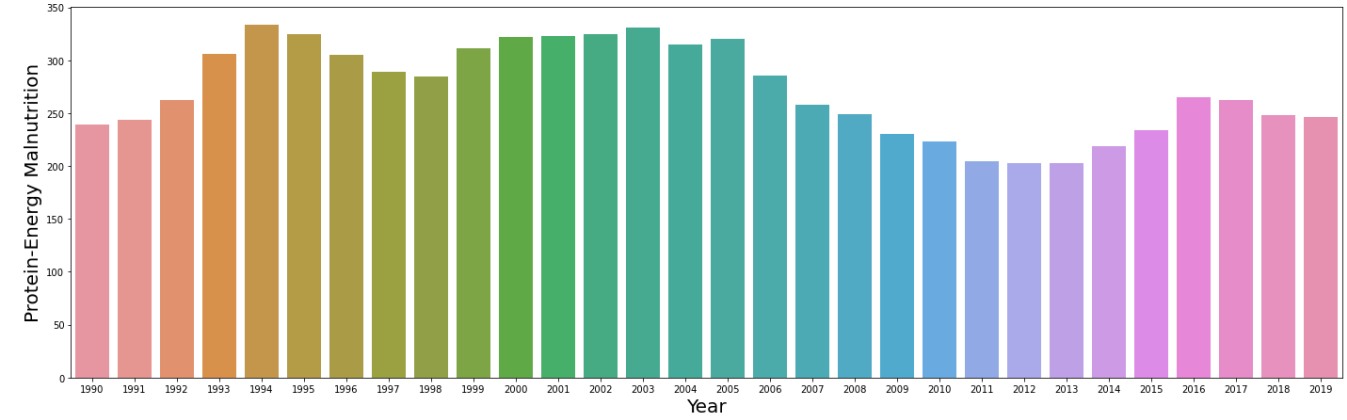



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = Russia_Total_Deaths[['Nutritional Deficiencies','Digestive Diseases',
        'Diarrheal Diseases','Protein-Energy Malnutrition']]
X = Russia_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```


Nutritional Deficiencies
Year

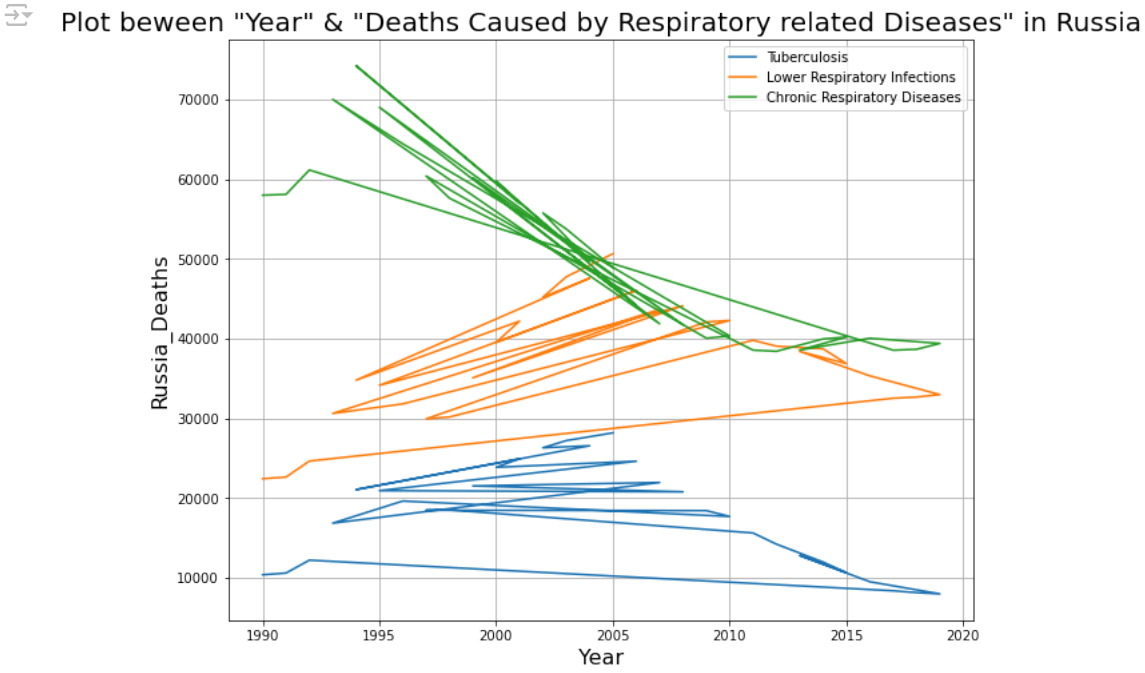

Digestive Diseases
Year


Diarrheal Diseases
Year

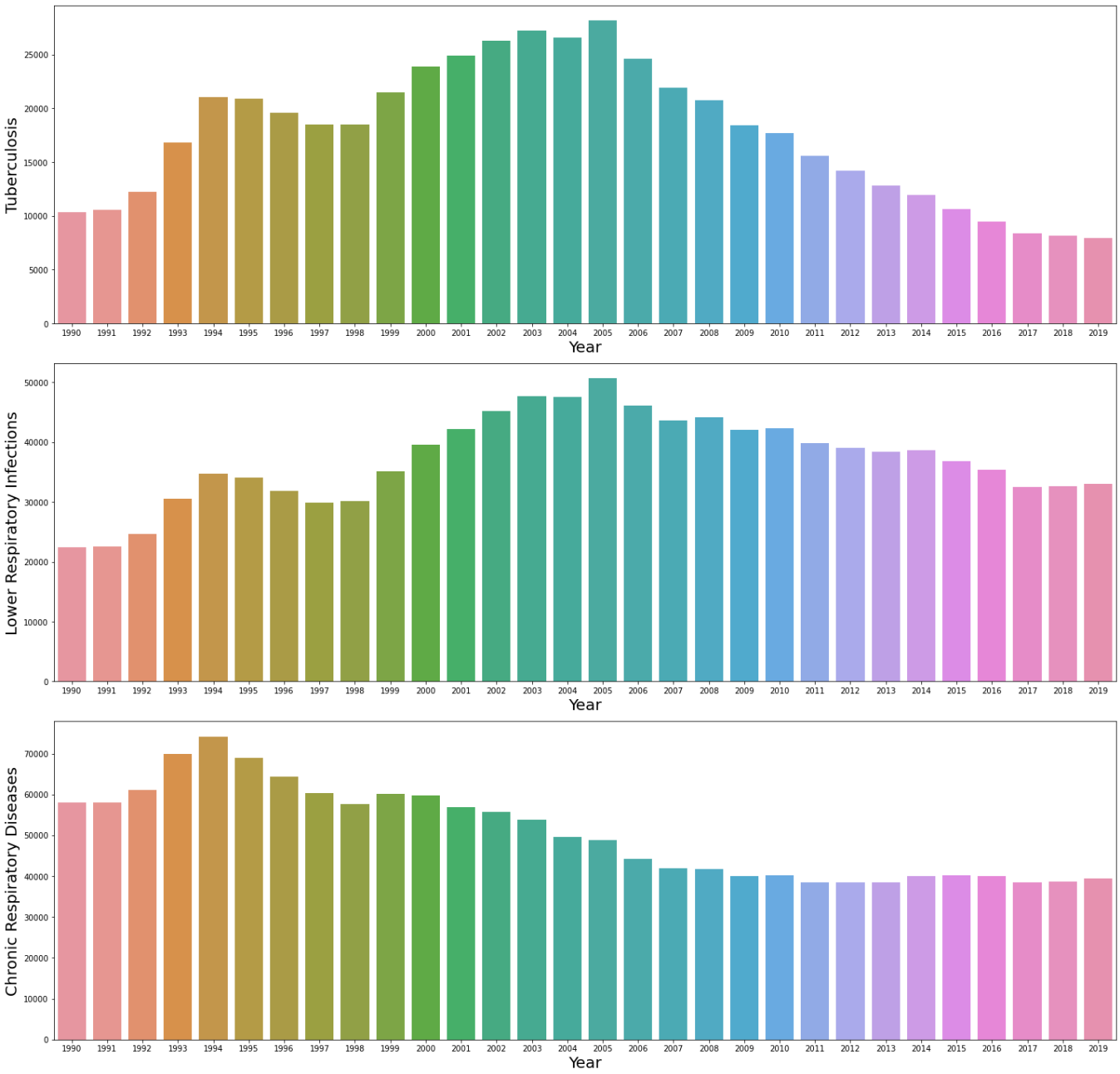

Protein-Energy Malnutrition
Year

• Deaths in Russia due to Digestive Diseases is increasing almost exponentially & is Dominant among all Deaths Caused by Gastrointestinal related Diseases with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by Respiratory related Diseases":
Russia_Total_Deaths.plot(x='Year', y=['Tuberculosis','Lower Respiratory Infections',
                                      'Chronic Respiratory Diseases'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in Russia',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('Russia_Deaths',fontsize =16)
plt.show()
```

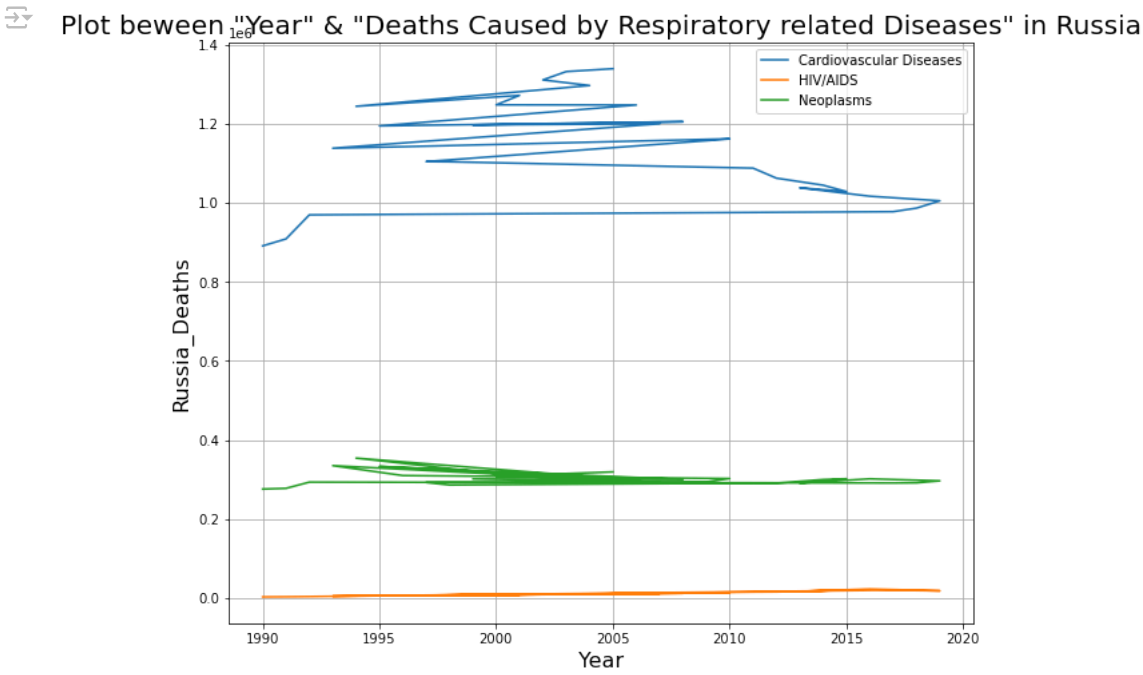



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = Russia_Total_Deaths[['Tuberculosis','Lower Respiratory Infections','Chronic Respiratory Diseases']]
X = Russia_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

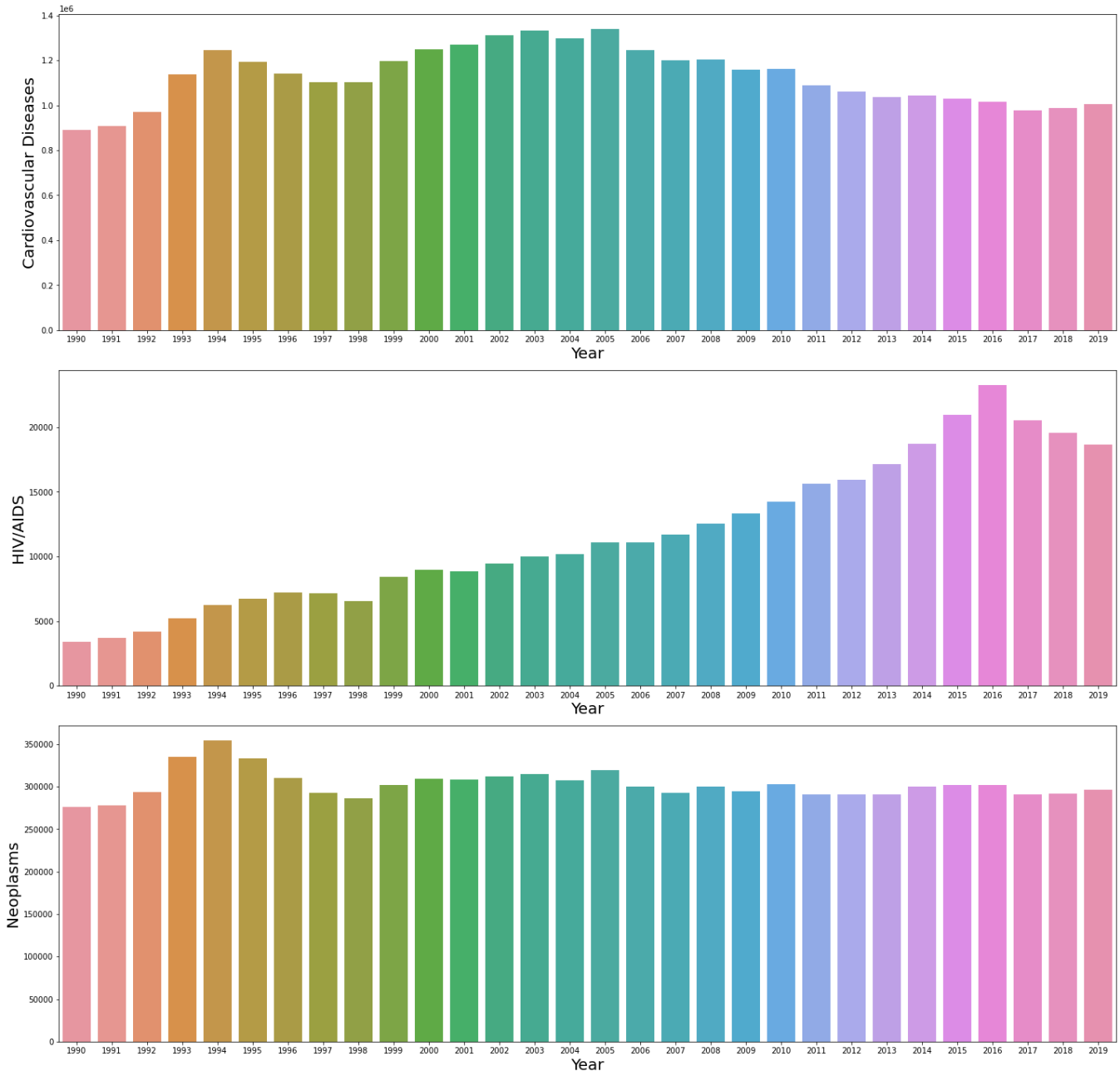


• Deaths in Russia due to Chronic Respiratory Diseases is decreasing but dominant among all Deaths Caused by Respiratory related Diseases with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by 'Cardiovascular Diseases','HIV/AIDS' & 'Neoplasms' Diseases":
Russia_Total_Deaths.plot(x='Year', y=['Cardiovascular Diseases','HIV/AIDS','Neoplasms'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by Respiratory related Diseases" in Russia',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('Russia_Deaths',fontsize =16)
plt.show()
```



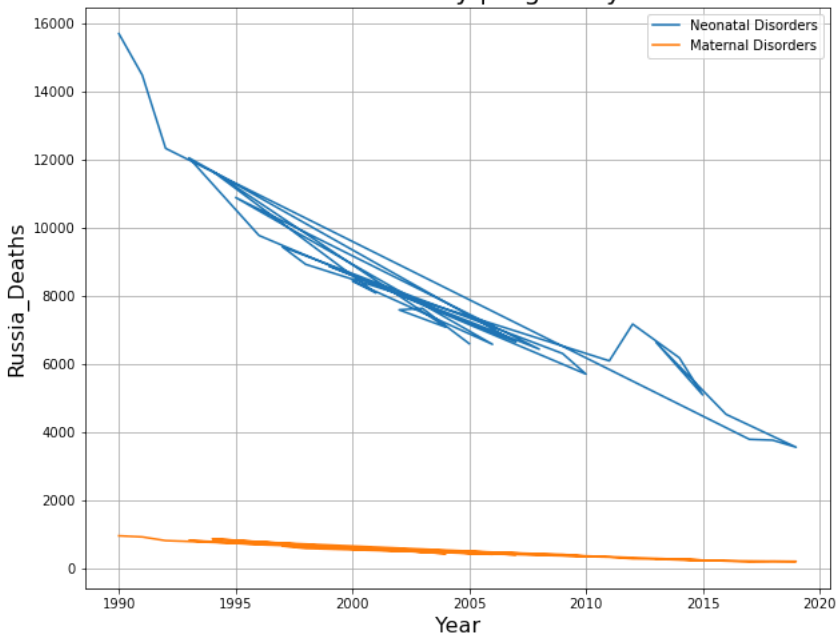

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = Russia_Total_Deaths[['Cardiovascular Diseases','HIV/AIDS','Neoplasms']]
X = Russia_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



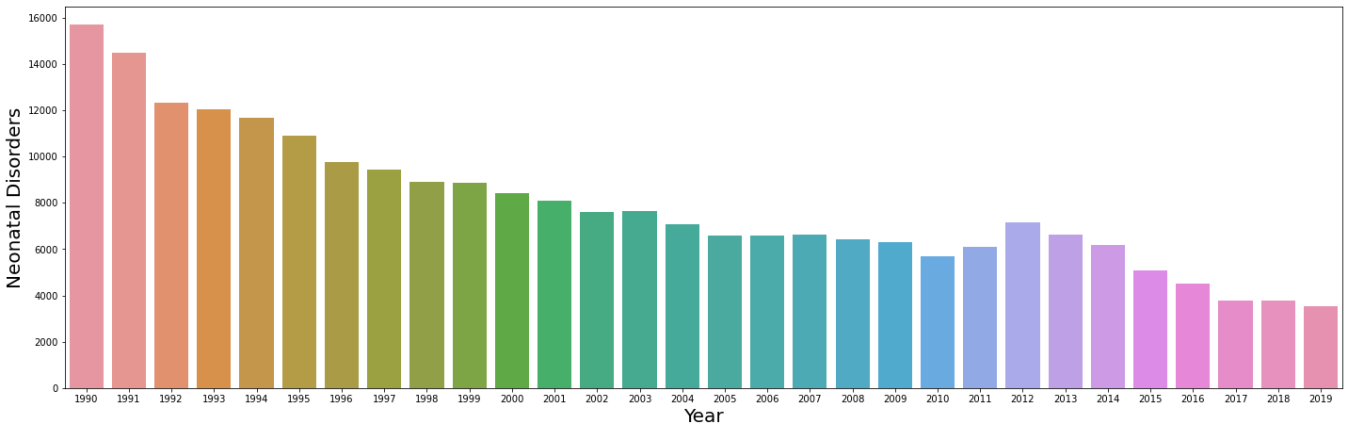
• Deaths in Russia due to cardiovascular diseases is most Dominant among all Deaths Caused with every incrementing year.

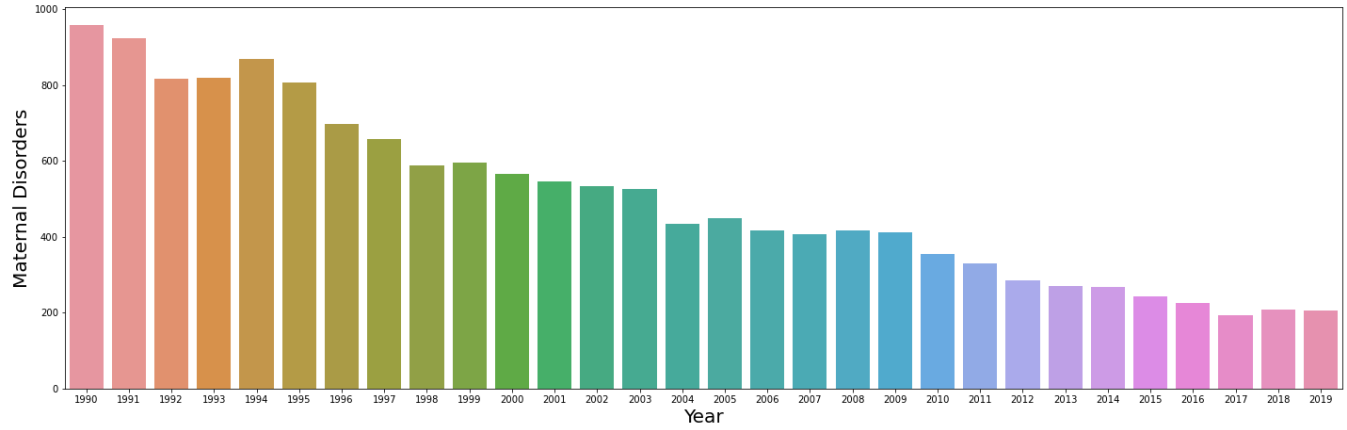
```
# Plot beween 'Year' & "Deaths Caused by pregnancy related Diseases":
Russia_Total_Deaths.plot(x='Year', y=['Neonatal Disorders','Maternal Disorders'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by pregnancy related Diseases" in Russia',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('Russia_Deaths',fontsize =16)
plt.show()
```

Plot beween "Year" & "Deaths Caused by pregnancy related Diseases" in Russia



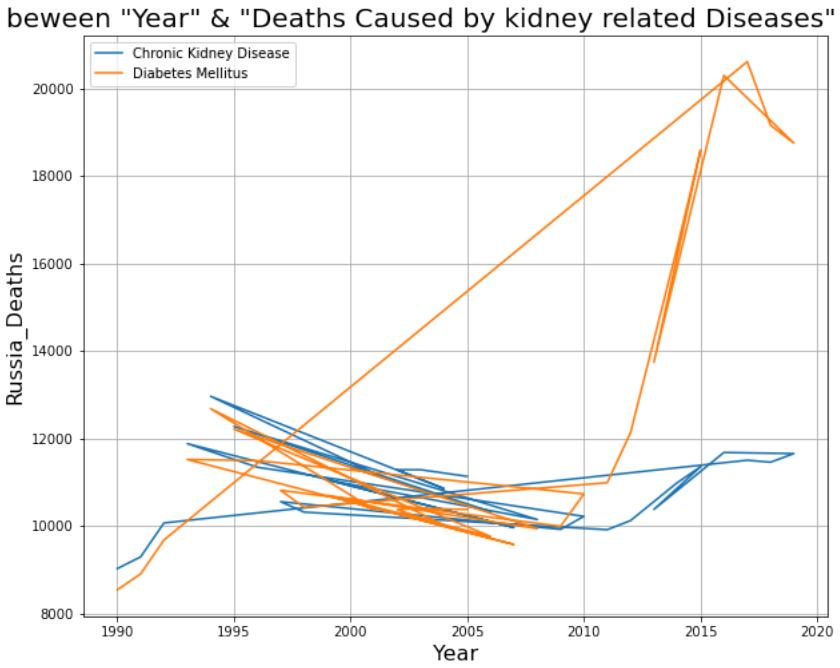

```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = Russia_Total_Deaths[['Neonatal Disorders','Maternal Disorders']]
X = Russia_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```



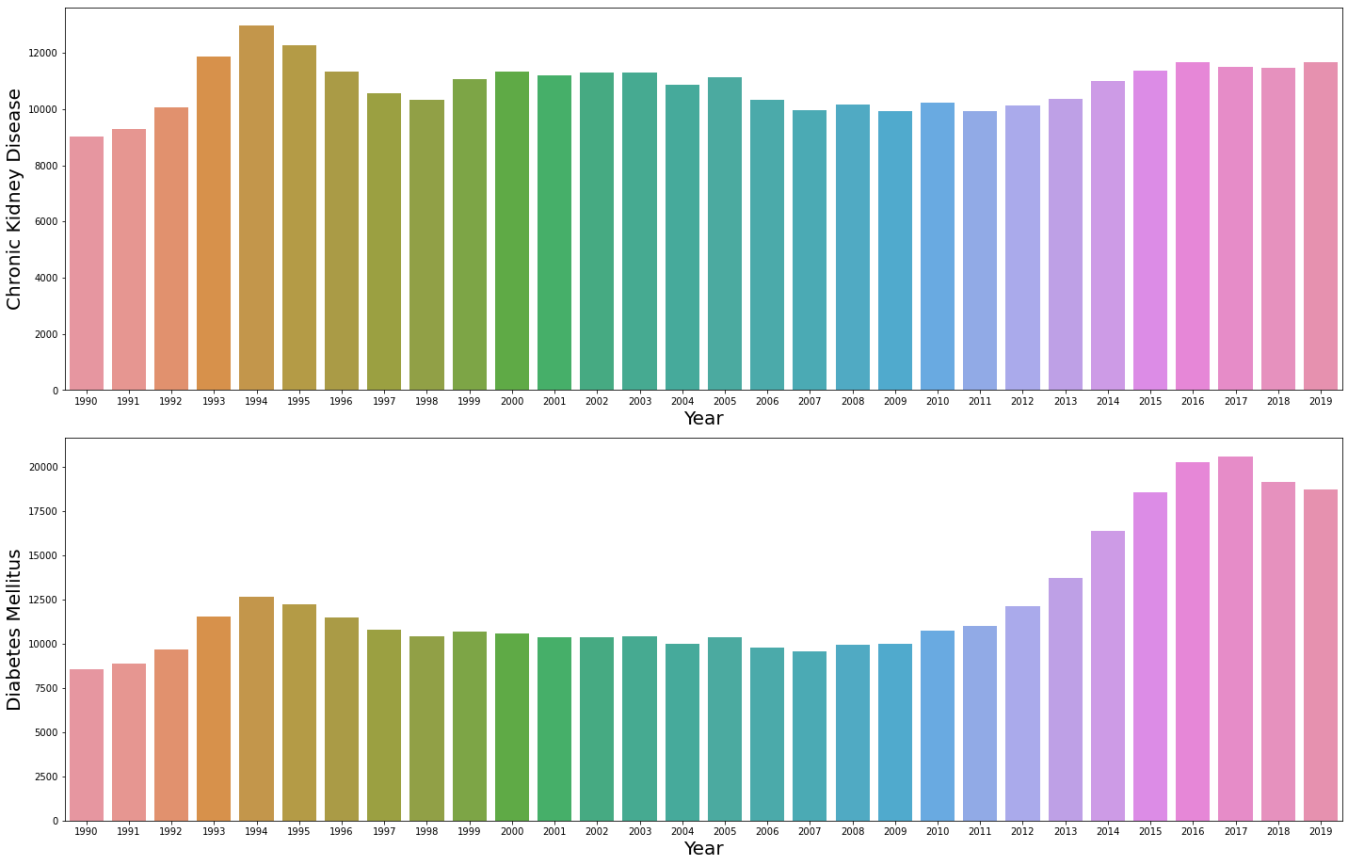


• Deaths in Russia due to Pregnancy related issues is on decline with every incrementing year.

```
# Plot beween 'Year' & "Deaths Caused by kidney related Diseases":
Russia_Total_Deaths.plot(x='Year', y=['Chronic Kidney Disease','Diabetes Mellitus'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by kidney related Diseases" in Russia',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('Russia_Deaths',fontsize =16)
plt.show()
```

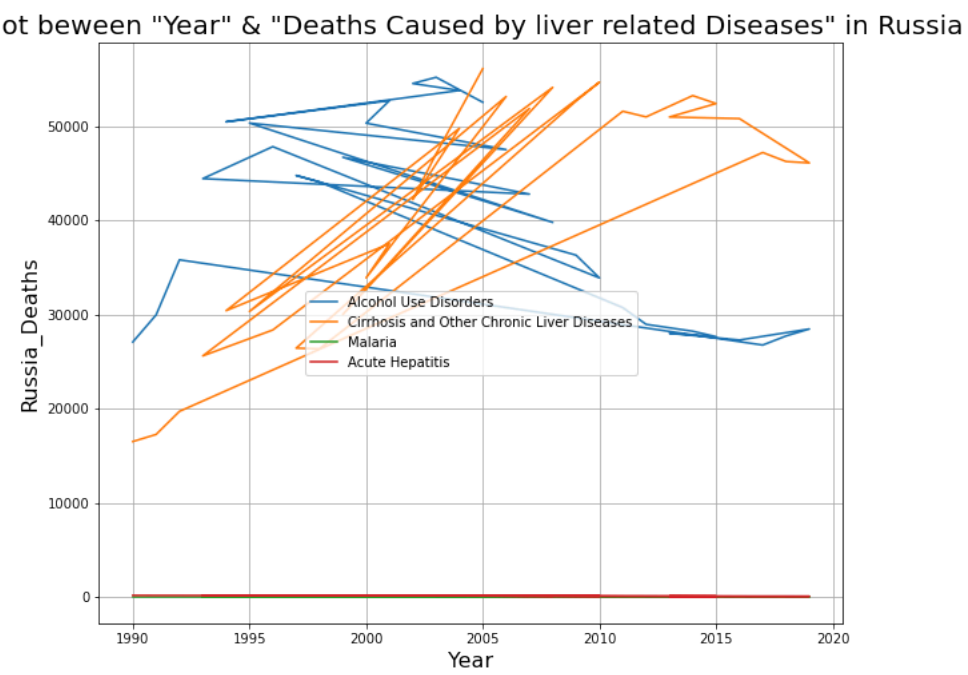


```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = Russia_Total_Deaths[['Chronic Kidney Disease','Diabetes Mellitus']]
X = Russia_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```

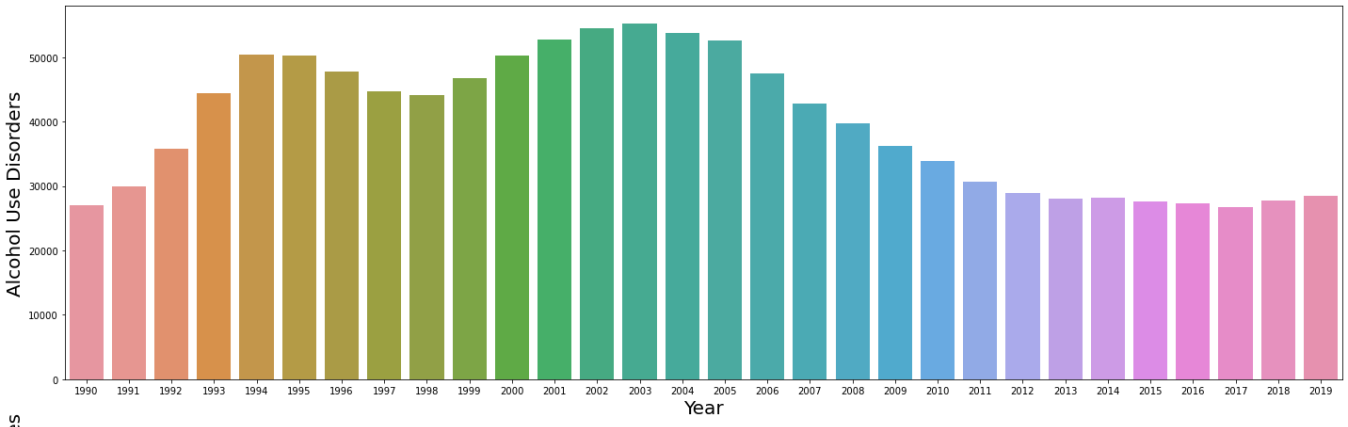


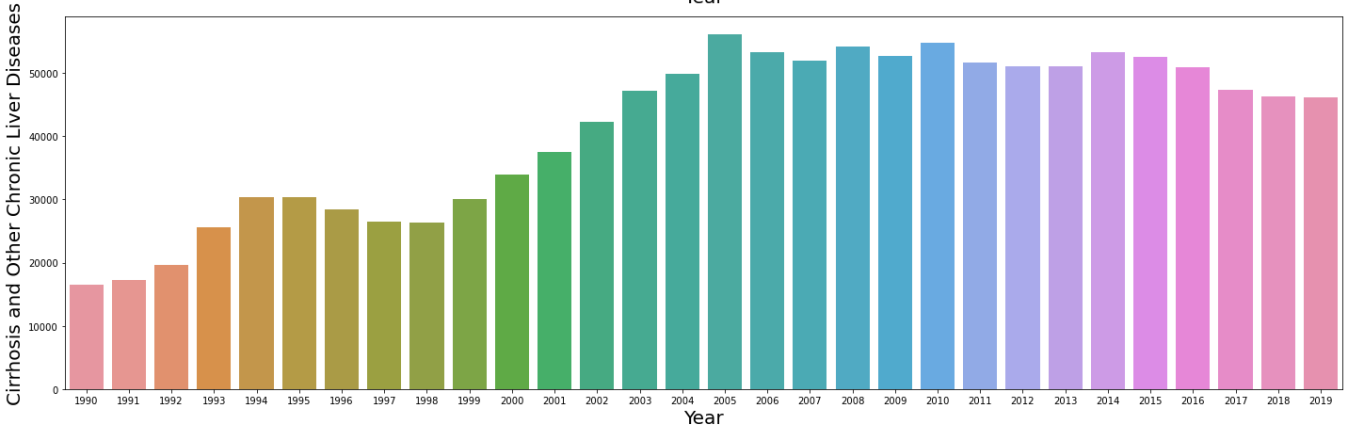
• Deaths in Russia due to all kidney related Diseases is Dominant with every incrementing year.

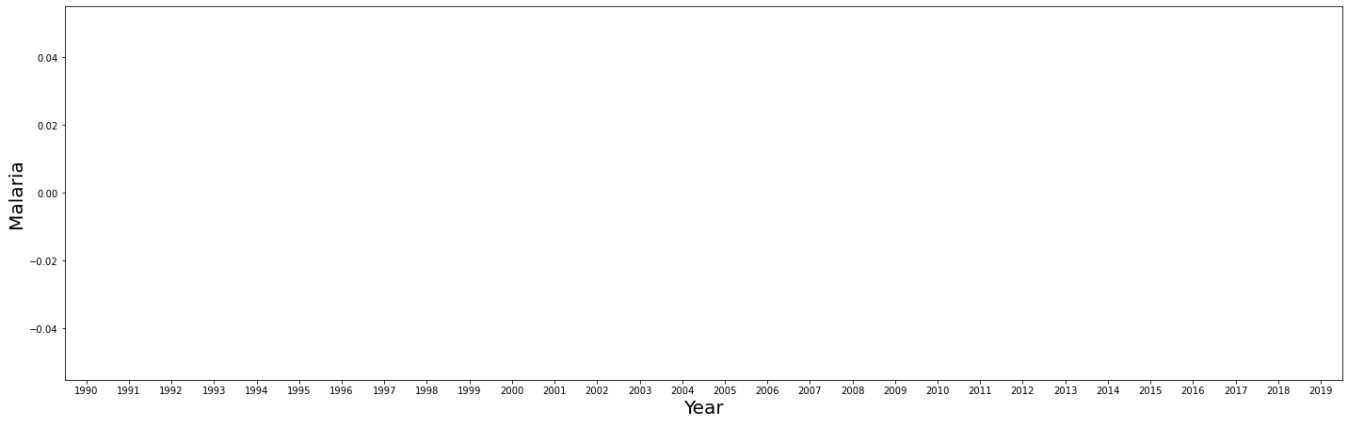
```
# Plot beween 'Year' & "Deaths Caused by liver related Diseases":
Russia_Total_Deaths.plot(x='Year', y=['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases',
                                      'Malaria','Acute Hepatitis'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Deaths Caused by liver related Diseases" in Russia',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('Russia_Deaths',fontsize =16)
plt.show()
```

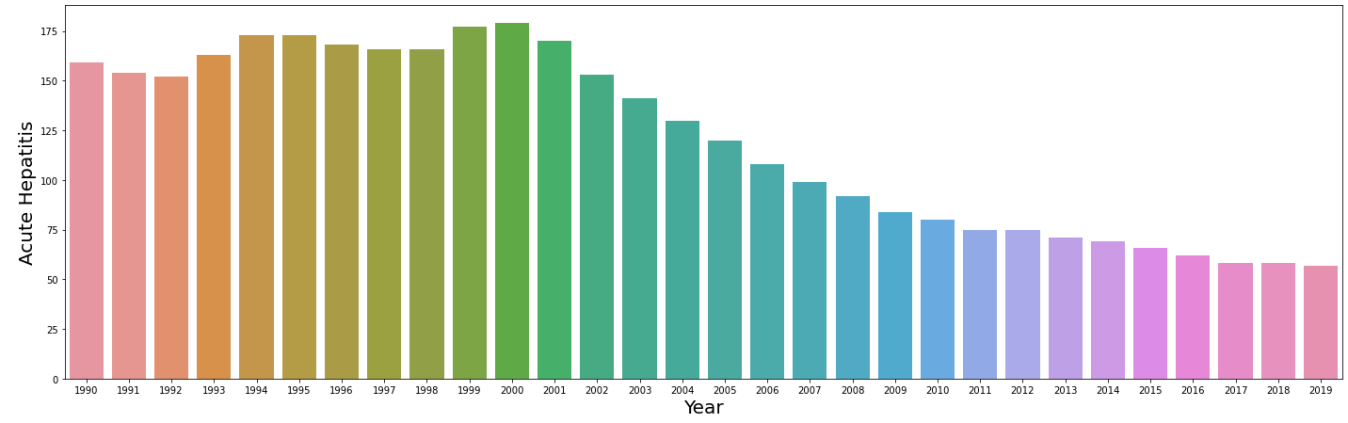



```
plt.figure(figsize=(20,25),facecolor='white')
plotnumber=1
y = Russia_Total_Deaths[['Alcohol Use Disorders','Cirrhosis and Other Chronic Liver Diseases',
                         'Malaria','Acute Hepatitis']]
X = Russia_Total_Deaths['Year']
for col in y:
    if plotnumber<=4:
        plt.subplot(4,1,plotnumber)
        sns.barplot(X,y[col])
        plt.xlabel('Year',fontsize=20)
        plt.ylabel(col,fontsize=20)
    plotnumber+=1


plt.tight_layout()
```


Alcohol Use Disorders
Year


Cirrhosis and Other Chronic Liver Diseases
Year


Malaria
Year


Acute Hepatitis
Year

• Deaths in Russia due to Cirrhosis and Other Chronic Liver Diseases is Dominant among all Deaths Caused by liver related Diseases with every incrementing year.

```
# Plot beween 'Year' & "Natural or Man-made Deaths":
Russia_Total_Deaths.plot(x='Year', y=['Interpersonal Violence','Drug Use Disorders',
                                      'Self-harm', 'Exposure to Forces of Nature',
                                     "Fire, Heat, and Hot Substances",
                                      'Poisonings','Conflict and Terrorism',
                                      'Environmental Heat and Cold Exposure',
                                      'Road Injuries','Drowning'],figsize=(10,8), grid=True)
plt.title('Plot beween "Year" & "Natural or Man-made Deaths" in Russia',fontsize =20)
plt.xlabel('Year',fontsize =16)
plt.ylabel('Russia_Deaths',fontsize =16)
plt.show()
```