

## **1.This Dataset contains Historical Data on Causes of Death around the World**

2. Reading the dataset
3. First 5 rows in the dataframe
4. Collecting Information about Dataframe
5. collecting the Statistical Information of Dataframe
6. Checking for null values in the Dataframe
7. Checking The unique Year data in the Dataframe
8. Checking the Correlation of various causes of death against year
9. Checking the Columns in the Dataframe
10. Checking the Total no. of Countries in the DataFrame

### **11. Feature Engineering**

#### **12. Top 10 Total\_no\_of\_Deaths**

	Total_no_of_Deaths	Country/Territory
1139	10442561	China
1138	10163943	China
1137	9978653	China
1119	9814213	China
1118	9591222	China
1117	9503904	China
1116	9411928	China
1114	9366974	China
1115	9364587	China
1113	9284664	China

**From the Above table, the "top 10 - Total\_no\_of\_Deaths" belongs to 'China' and the "top 60 - Total\_no\_of\_Deaths" belongs to 'China' and 'India' combined, followed by 'USA' and 'Russia'. This because of the fact "China" and "India" are the countries that stand in top 2 in terms of population.**

#### **13. Data Visualization-----**

- i) There is a clear rise in Total No. of Deaths recorded with each year for China in Jupyter notebook.
- ii) There is a rapid drop in Malaria Deaths recorded in China after 1999.
- iii) There is a drop in Nutritional Deficiencies Deaths recorded in China in 2007 and from 2008 the count of deaths again started to rise.
- iv) There is a continual drop in Interpersonal Violence Deaths recorded in China.
- v) There is a continual drop in Maternal Disorder Deaths recorded in China after year 1998.
- vi) There is a continual increase in HIV/AIDS Deaths recorded in China.
- vii) There is a continual decrease in Tuberculosis Deaths recorded in China.
- viii) There is a continual increase in Cardiovascular Diseases Deaths recorded in China.
- ix) There is a continual drop in Self-harm Deaths recorded in China after year 2004.
- x) There is a rapid rise in Exposure to Forces of Nature Deaths recorded in China in the year 2004 this is because of Sichuan earthquake of 2008.
- xi) The terrorist attack of 2014 is at Kunming Railway Station in Kunming, Yunnan, China on March 1st.
- xii) There is a continual rise in Poisonings Deaths recorded in China after year 1994.
- xiii) There is a continual rise in Protein-Energy Malnutrition Deaths recorded in China after year 2007.
- xiv) There is a continual drop in Road Injuries Deaths recorded in China after year 2009.
- xv) There is a continual drop in Chronic Respiratory Diseases Deaths recorded in China after year 2004.

*xvi) Overall there is a raise in Total No.of Deaths recorded with each year for India, even though there are fluctuations inbetween.*

*Xvii) There is a rapid drop in Malaria Deaths recorded in India from 1990, but the Deaths in 2018 and 2019 is greater than that of 2016 and 2017.*

*xviii) There is a continual drop in Nutritional Deficiencies Deaths recorded in India from 1990.*

*xix) There is a raise in Interpersonal Violence Deaths recorded in India from 2014.*

*xx) There is continual drop in Maternal Disorders Deaths recorded in India after 1999.*

*xxi) There is continual raise in HIV/AIDS Deaths recorded in India from 1992 and reached peak at 2006 from where there is continual drop.*

*xxii) Though there are deviations in Tuberculosis Deaths recorded in India, there is a drop after year 1998.*

*xxiii) There are continual raise in Cardiovascular Diseases Deaths recorded in India.*

*xxiv) There are continual raise in Self-harm Deaths recorded in India from 1990 to 2001, but then there are fluctuations.*

----\*The Peak incident: Bhachau Taluka of Kutch (Gujarat) experienced an earthquake of 7.6 to 7.9 on the Richter Scale and lasted for 120 seconds. This disaster killed around 20,000, injured 167,000 and nearly 400,000 people became homeless. The 2013 incident: The year 2013 witnessed two major events, the Uttarakhand Flash Floods, mainly affecting the state of Uttarakhand and the Cyclone "Phailin" affecting two coastal states of Odisha and Andhra Pradesh.\*

*xxv) Overall there is a gradual decrease in the Poisonings Deaths recorded in India.*

*xxvi) There is a continual drop in the Protein-Energy Malnutrition Deaths recorded in India.*

*xxvii) There is a continual raise in the Road Injuries Deaths recorded in India with fluctuations between 2013 to 2017.*

*xxviii) There is a continual raise in the Chronic Respiratory Diseases Deaths recorded in India with mere fluctuations inbetween.*

#### China - Top 10 Causes of Deaths

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

#### India - Top 10 Causes of Deaths

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

#### United States - Top 10 Causes of Deaths

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