Cause of Death

Submitted by:

Saurav Kumar

ACKNOWLEDGMENT

I would like to thank Flip Robo Technologies for providing me with the opportunity to work on this project from which I have learned a lot. I am also grateful to Shwetank Mishra for his constant guidance and support.

CONTENTS:

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.

INTRODUCTION

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.

Dataset Glossary (Column-wise)

- 01. Country/Territory - Name of the Country/Territory
- 02. Code - Country/Territory Code
- 03. Year - Year of the Incident
- 04. Meningitis - No. of People died from Meningitis
- 05. Alzheimer's Disease and Other Dementias - No. of People died from Alzheimer's Disease and Other Dementias
- 06. Parkinson's Disease - No. of People died from Parkinson's Disease
- 07. Nutritional Deficiencies - No. of People died from Nutritional Deficiencies
- 08. Malaria - No. of People died from Malaria
- 09. Drowning - No. of People died from Drowning
- 10. Interpersonal Violence - No. of People died from Interpersonal Violence
- 11. Maternal Disorders - No. of People died from Maternal Disorders
- 12. Drug Use Disorders - No. of People died from Drug Use Disorders
- 13. Tuberculosis - No. of People died from Tuberculosis
- 14. Cardiovascular Diseases - No. of People died from Cardiovascular Diseases
- 15. Lower Respiratory Infections - No. of People died from Lower Respiratory Infections
- 16. Neonatal Disorders - No. of People died from Neonatal Disorders
- 17. Alcohol Use Disorders - No. of People died from Alcohol Use Disorders
- 18. Self-harm - No. of People died from Self-harm
- 19. Exposure to Forces of Nature - No. of People died from Exposure to Forces of Nature
- 20. Diarrheal Diseases - No. of People died from Diarrheal Diseases
- 21. Environmental Heat and Cold Exposure - No. of People died from Environmental Heat and Cold Exposure
- 22. Neoplasms - No. of People died from Neoplasms
- 23. Conflict and Terrorism - No. of People died from Conflict and Terrorism
- 24. Diabetes Mellitus - No. of People died from Diabetes Mellitus
- 25. Chronic Kidney Disease - No. of People died from Chronic Kidney Disease
- 26. Poisonings - No. of People died from Poisoning
- 27. Protein-Energy Malnutrition - No. of People died from Protein-Energy Malnutrition
- 28. Chronic Respiratory Diseases - No. of People died from Chronic Respiratory Diseases
- 29. Cirrhosis and Other Chronic Liver Diseases - No. of People died from Cirrhosis and Other Chronic Liver Diseases
- 30. Digestive Diseases - No. of People died from Digestive Diseases
- 31. Fire, Heat, and Hot Substances - No. of People died from Fire or Heat or any Hot Substances

32. Acute Hepatitis - No. of People died from Acute Hepatitis

Analytical Problem Framing

- In this dataset we have to find the mortality as per country and the chances of causes of death with disease . We have list of disease name columns and in column we have the total number of death by disease . 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). And we have to predict the Cause of death Datset . In this dataset we have 2 Object column and left all are integer columns.

```
In [2]: #Lets Load The DataSet
df=pd.read_csv(r"C:\Users\saura\OneDrive\Desktop\cause_of_deaths dataset.csv")
df.head()
```

Out[2]:

	Country/Territory	Code	Year	Meningitis	Alzheimer's Disease and Other Dementias	Parkinson's Disease	Nutritional Deficiencies	Malaria	Drowning	Interpersonal Violence	...	Diabetes Mellitus	Chronic Kidney Disease	Poisonings	Pr E Malnu
0	Afghanistan	AFG	1990	2159	1116	371	2087	93	1370	1538	...	2108	3709	338	
1	Afghanistan	AFG	1991	2218	1136	374	2153	189	1391	2001	...	2120	3724	351	
2	Afghanistan	AFG	1992	2475	1162	378	2441	239	1514	2299	...	2153	3776	386	
3	Afghanistan	AFG	1993	2812	1187	384	2837	108	1687	2589	...	2195	3862	425	
4	Afghanistan	AFG	1994	3027	1211	391	3081	211	1809	2849	...	2231	3932	451	

5 rows × 34 columns

Data:

In this project we are given excel file containing dataset of Cause of death. There are 34 columns by which we can analyse how much chances to live with any disease .In this dataset we have 2 Object column and left are integer datatype. The dataset contains 3120 rows and 34 columns. The data contains no null value however there is a lot of unwanted characters in the columns.

Hardware and Software Requirements and Tools Used

• A mid level computer that runs on Intel i3- i5 8th generation, 4gb ram or A10/A11or any other equivalent chipset and a suitable processor.

• Juypter Notebook/Google chrome

• Libraries and packeges used:

import pandas as pd

import numpy as np import matplotlib. pyplot as plt

import seaborn as sns

import warnings warnings.filterwarnings("ignore")

We will use Python through Jupyter notebook for data processing. Also we will use liabraries here is numpy, matplotlib, pandas and seaborn. The matpotplotlib and seaborn library has been used to make charts to visualize and understand the problem, correlation, outliers and many other things, the pandas and numpy library issued to handle dataset and perform various tasks.

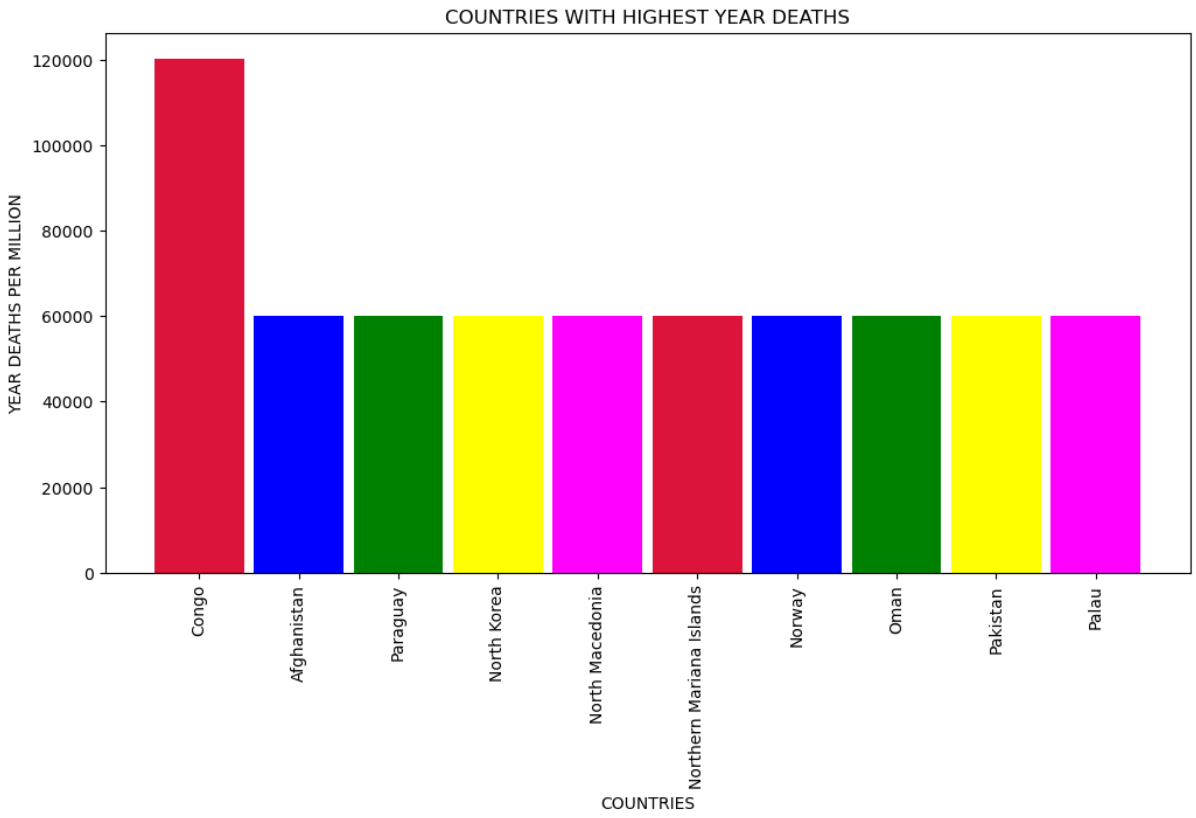
Exploratory Data Analysis:

All Columns Name ->

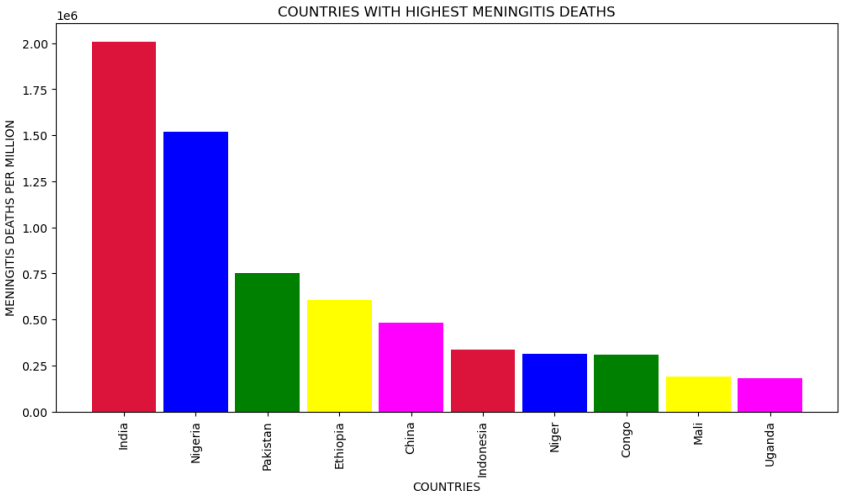
```
In [9]: #Lets plot all unique values.
        df.columns

Out[9]: 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')
```

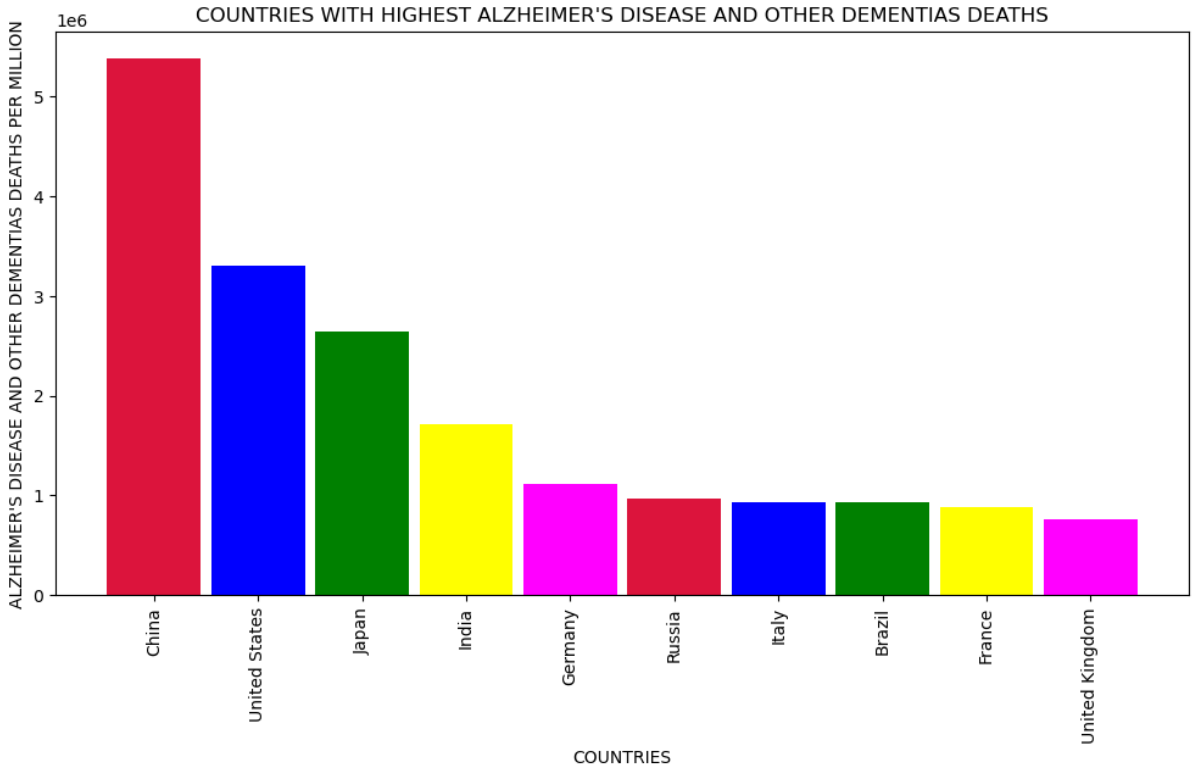
- **Countries with Highest year Death .**



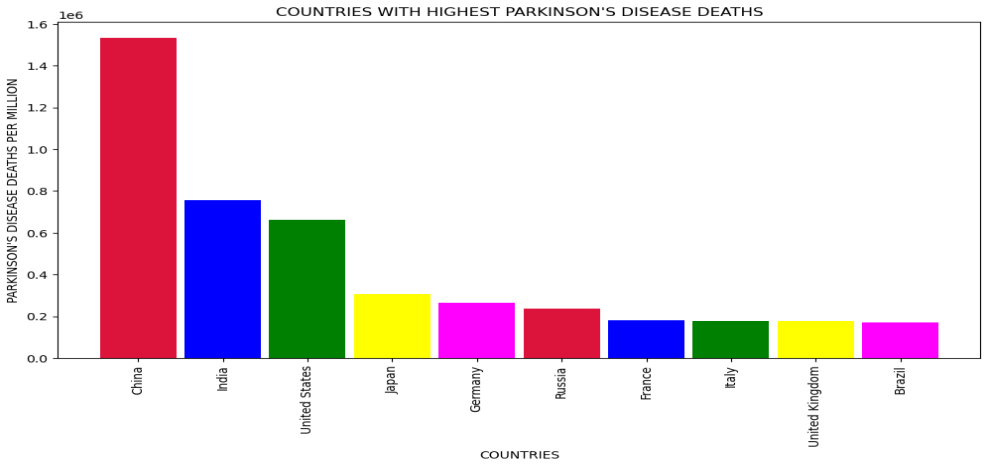
Countries with Highest Meningitis Death



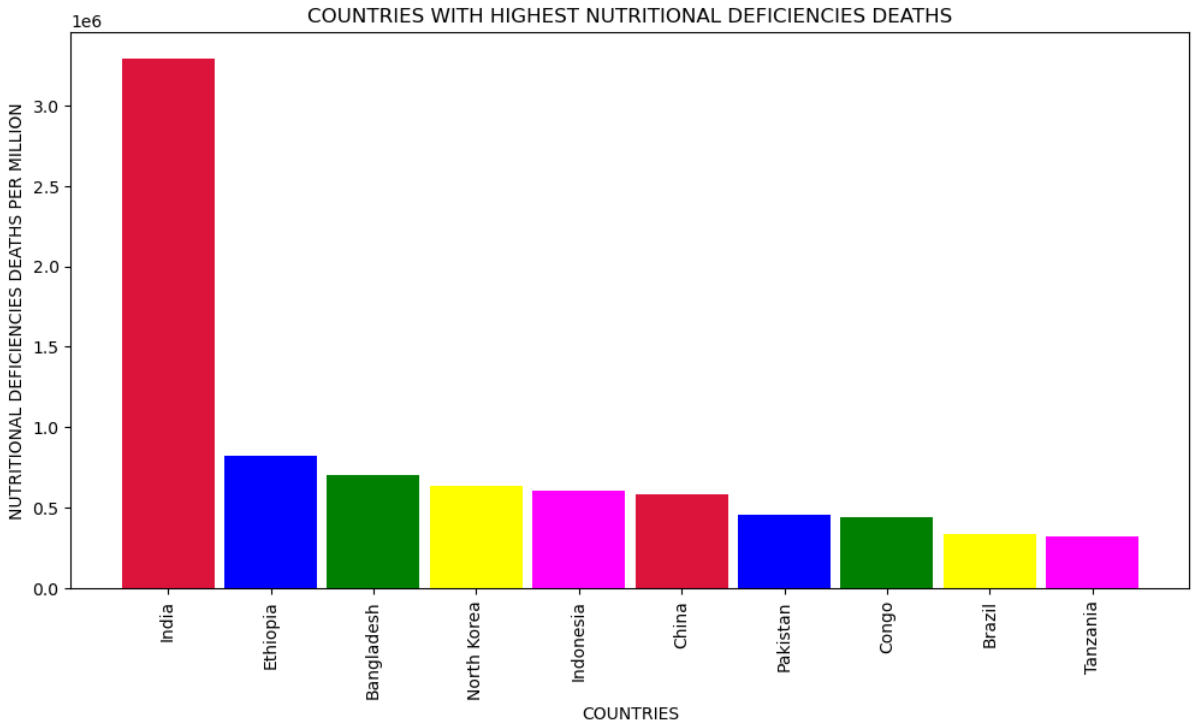
Countries with Highest Alzheimer's Disease and Other Dementias Death



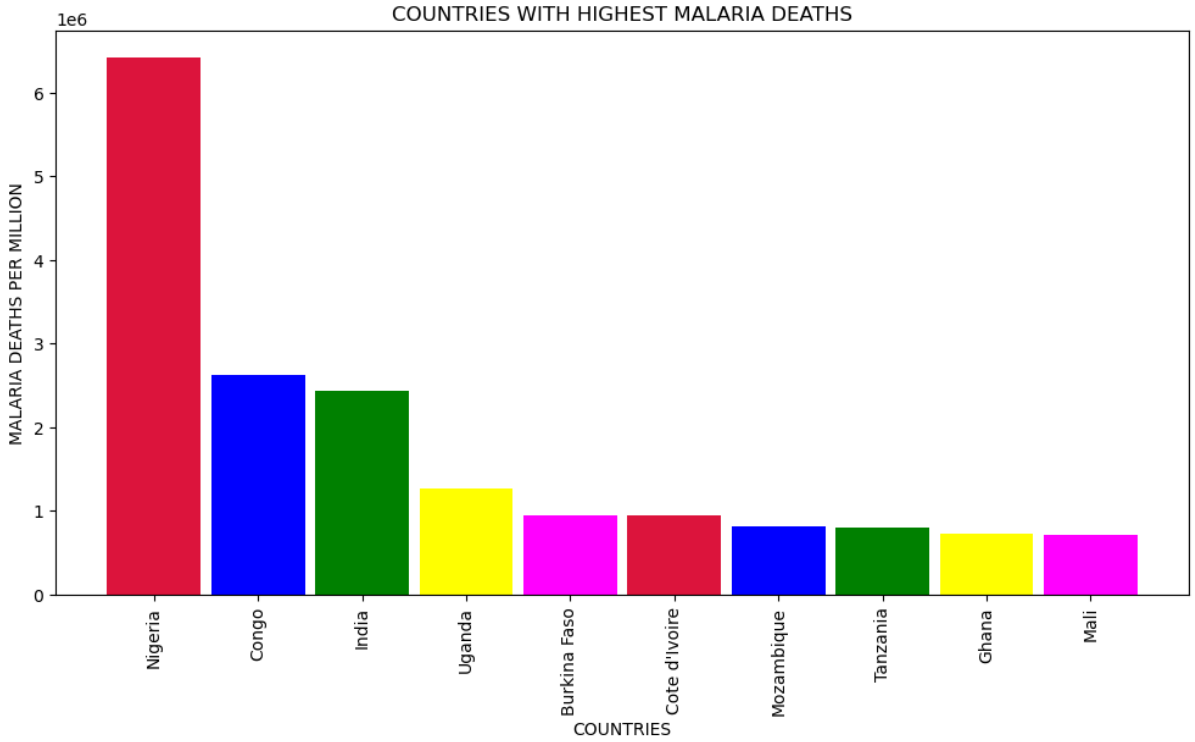
Countries with Highest Parkinson's Disease Death



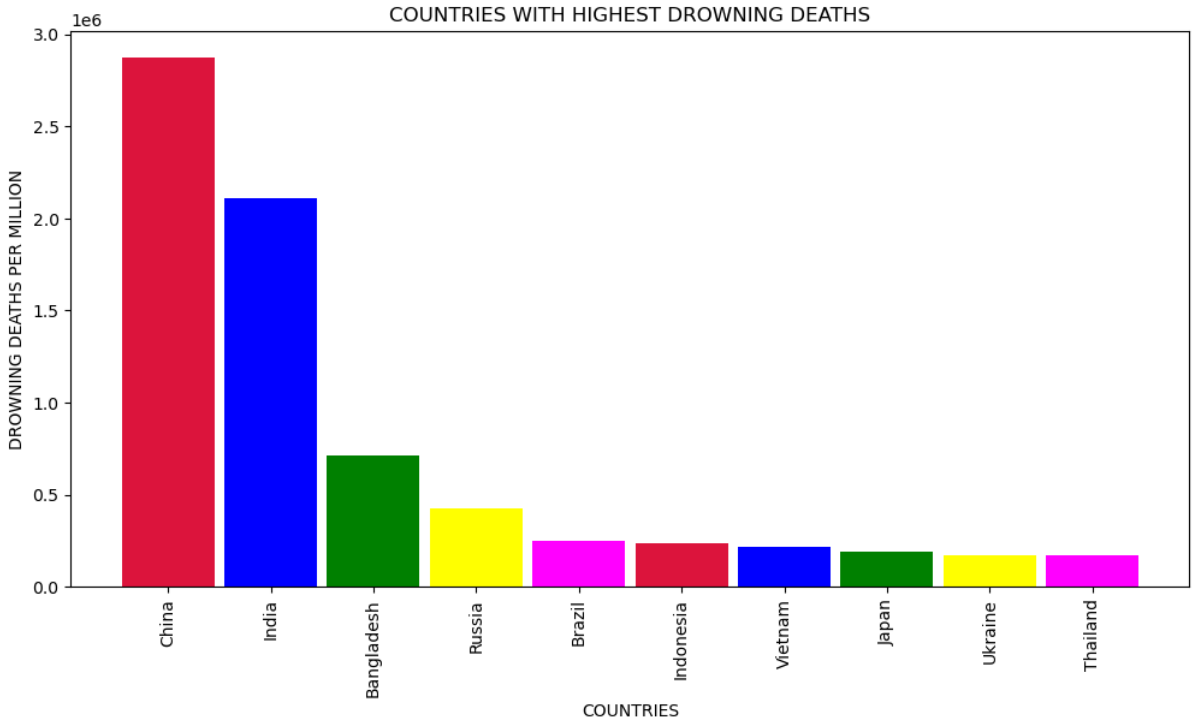
Countries with Highest Nutritional Deficiencies Death



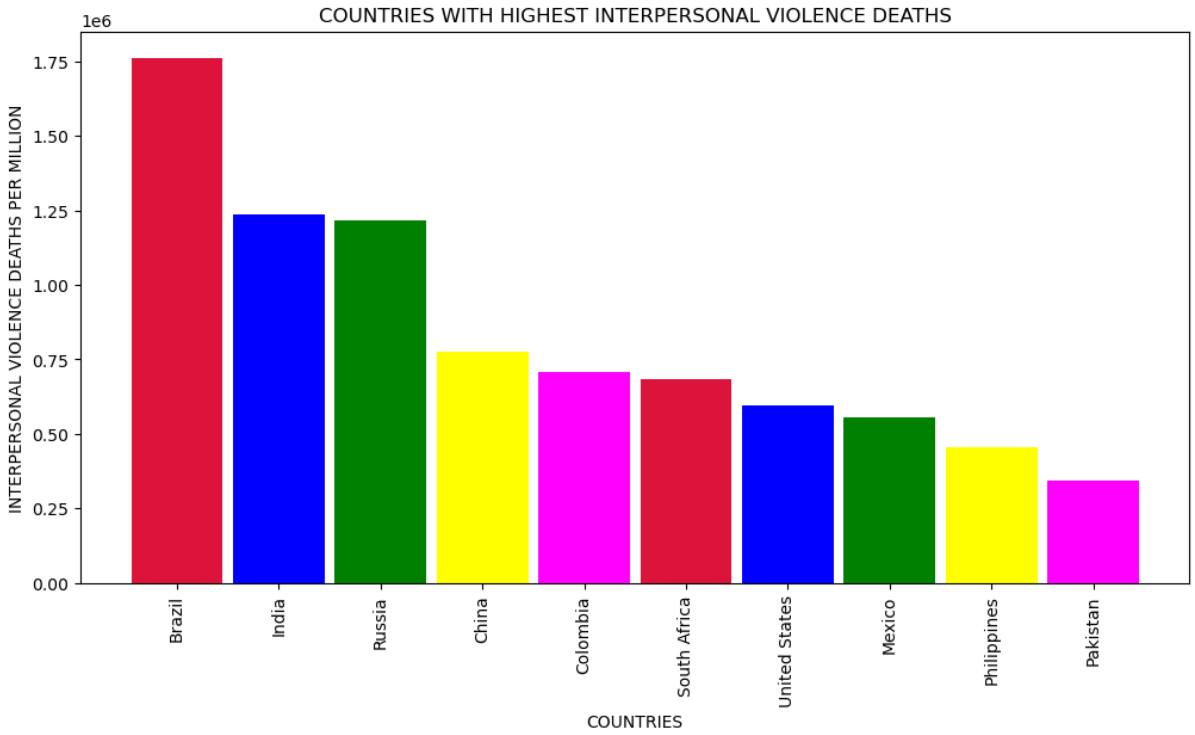
Countries with Highest Malaria Death



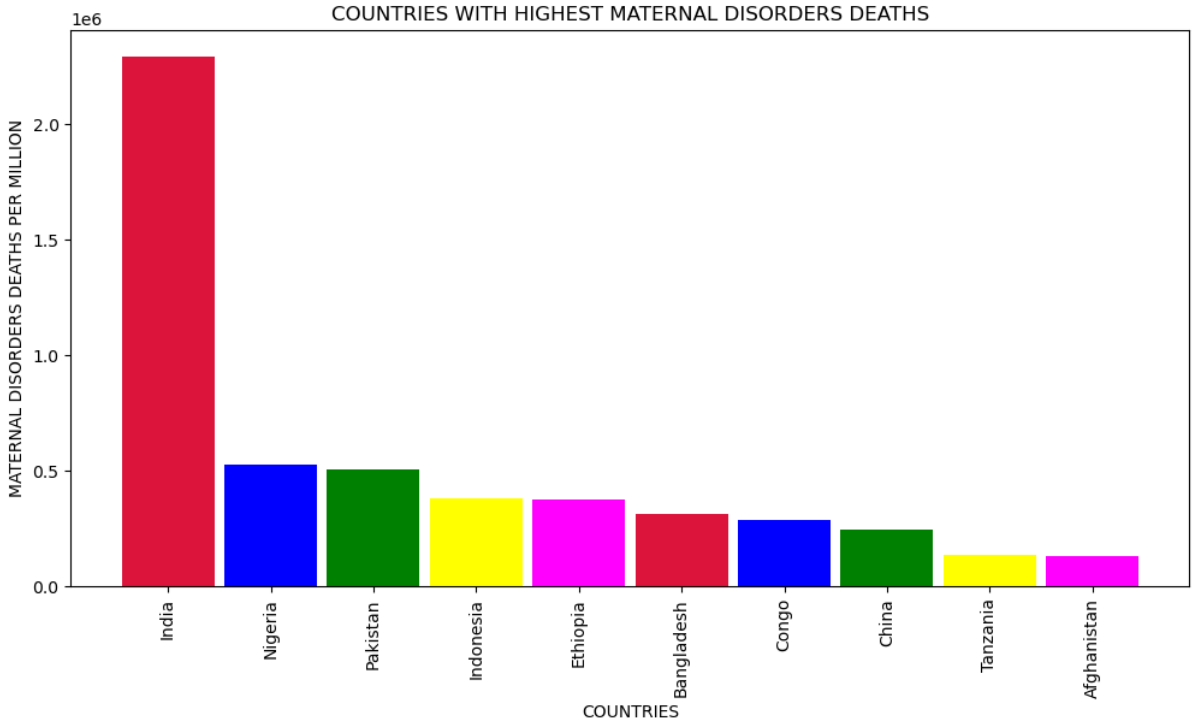
Countries with Highest Drowning Death



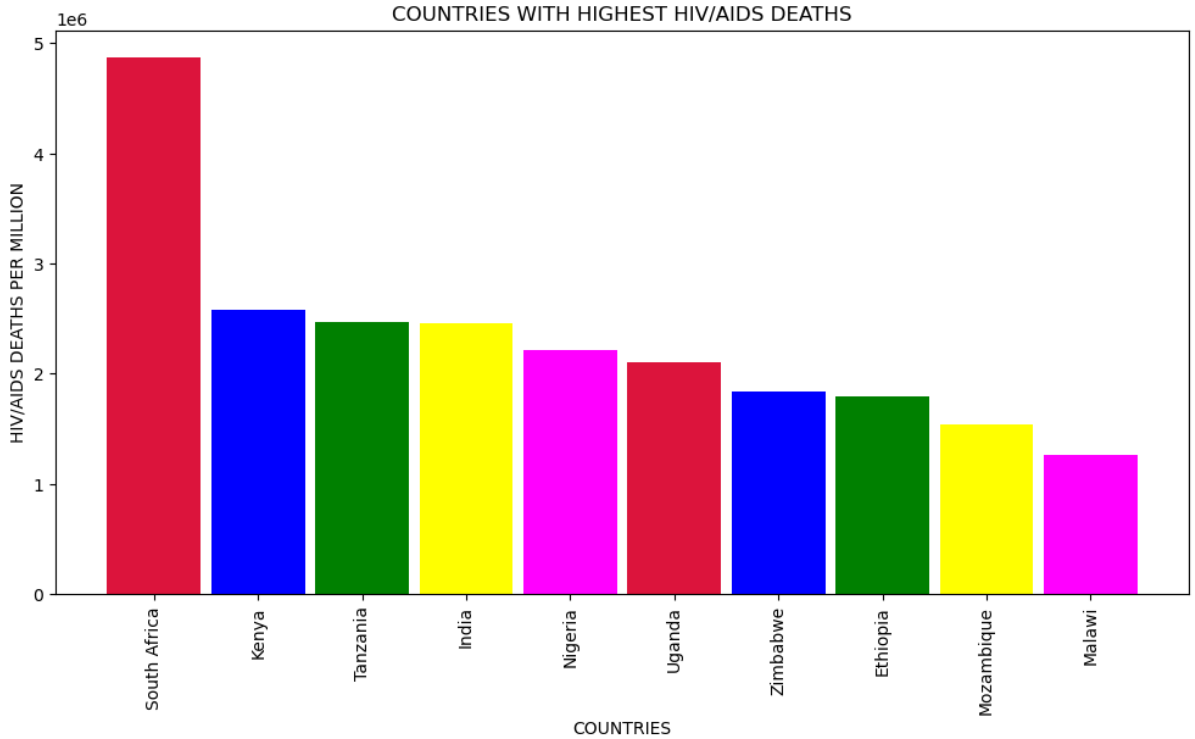
Countries with Highest Interpersonal Violence Death



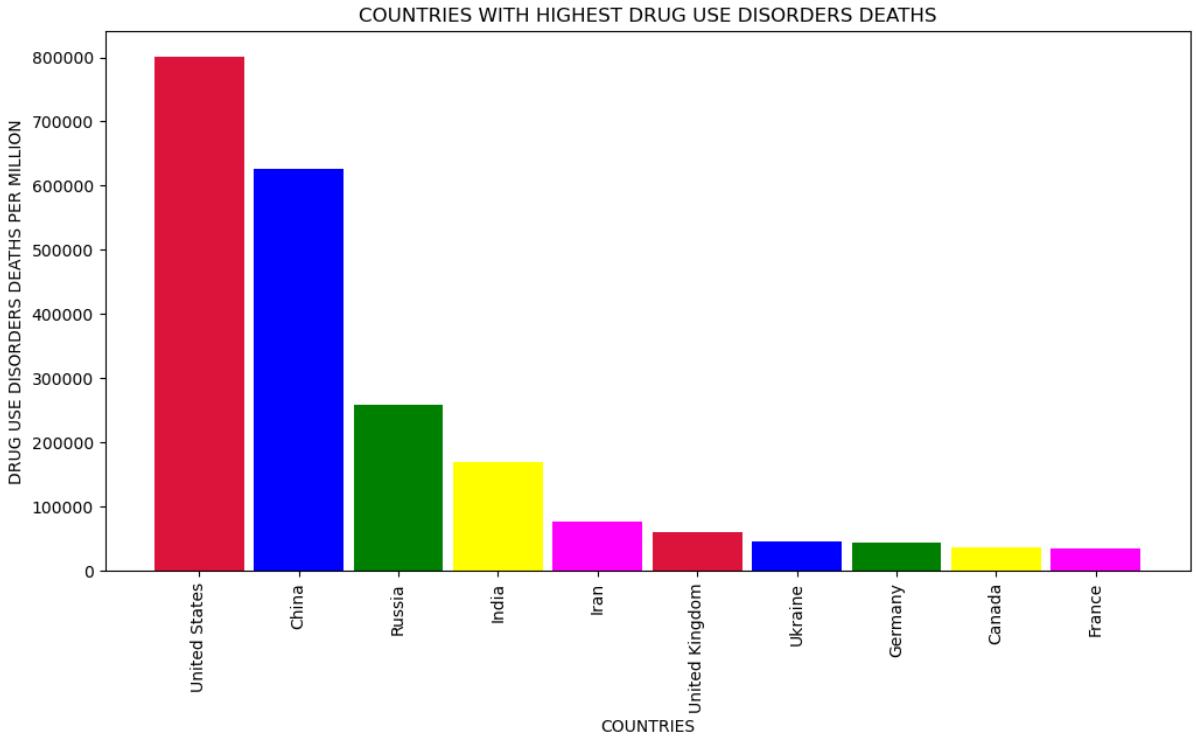
Countries with Highest Maternal Disorders Death



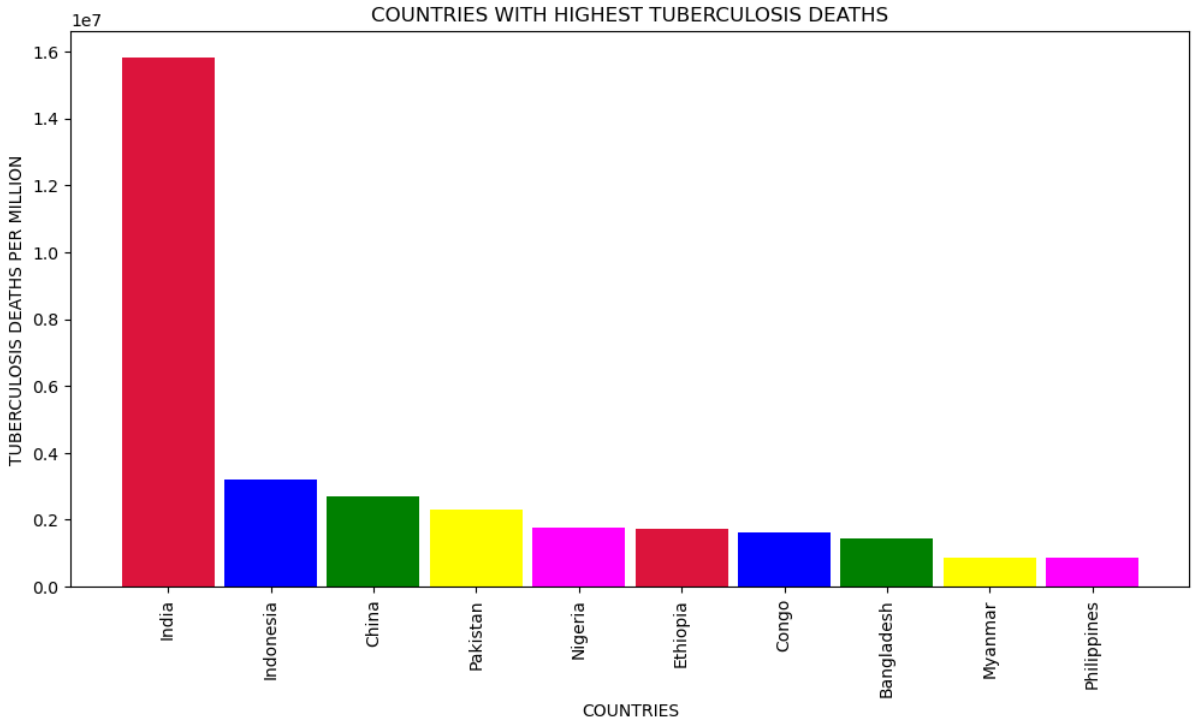
Countries with Highest HIV/AIDS Death



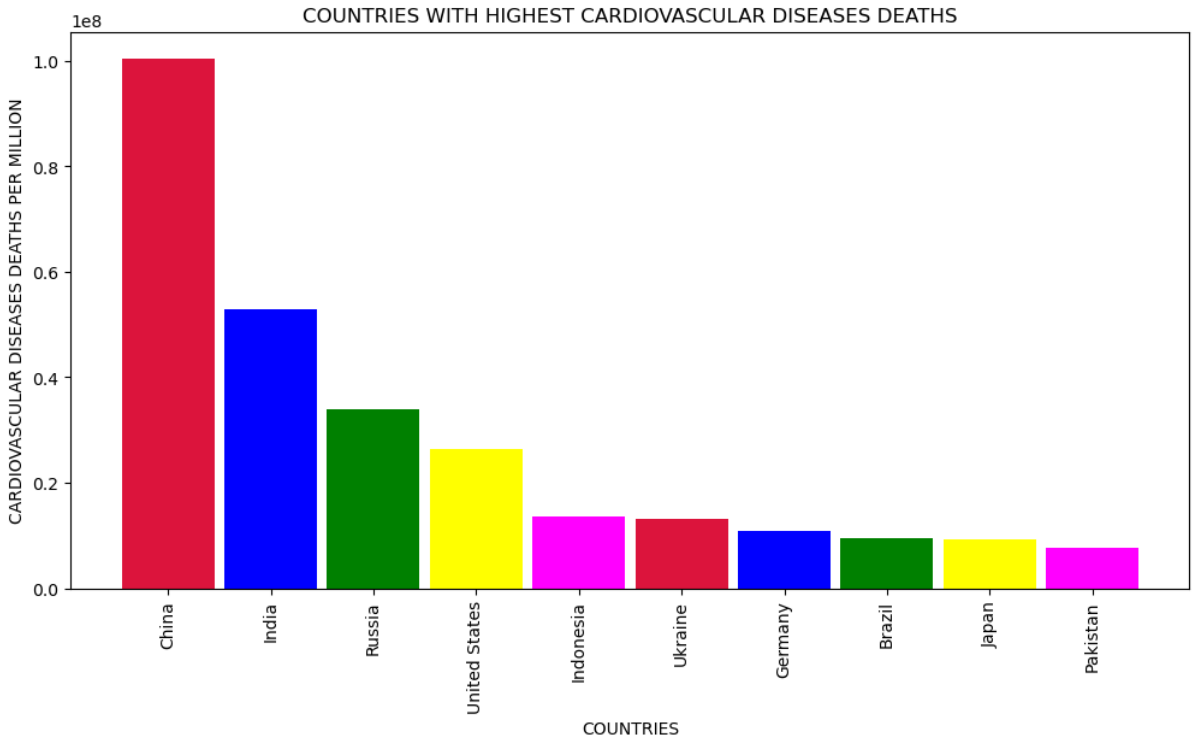
Countries with Highest Drug Use Disorders Death



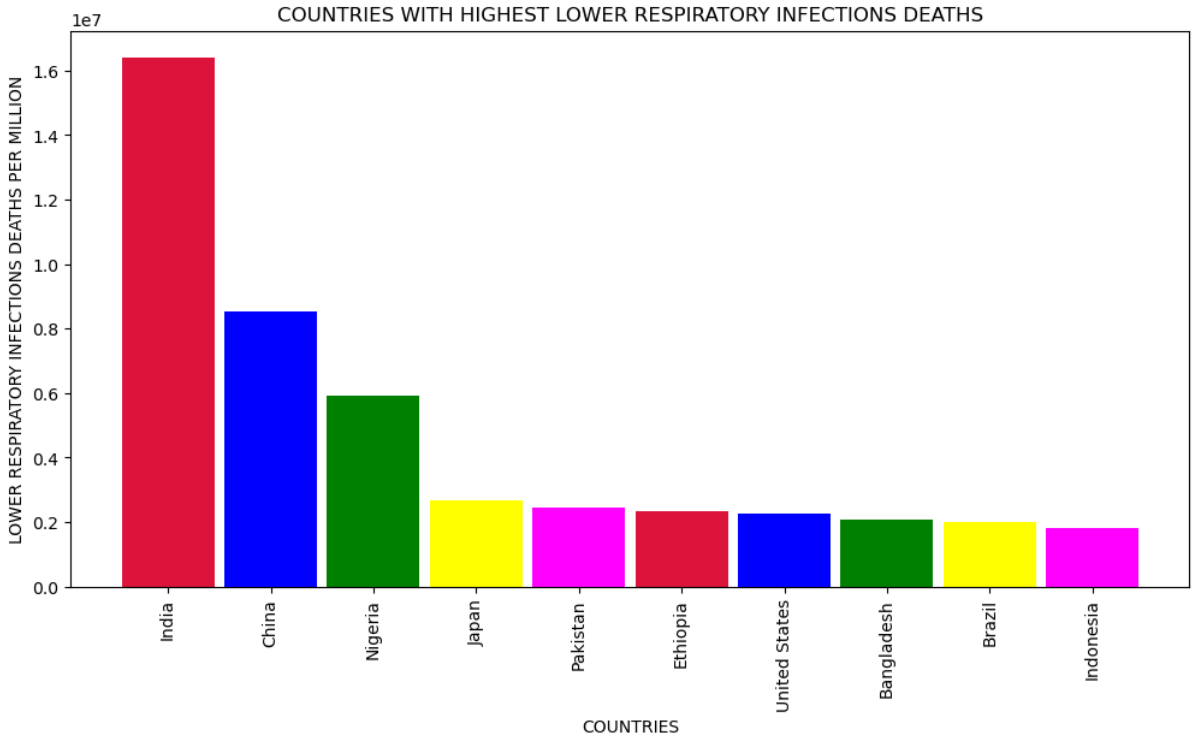
Countries with Highest Tuberculosis Death



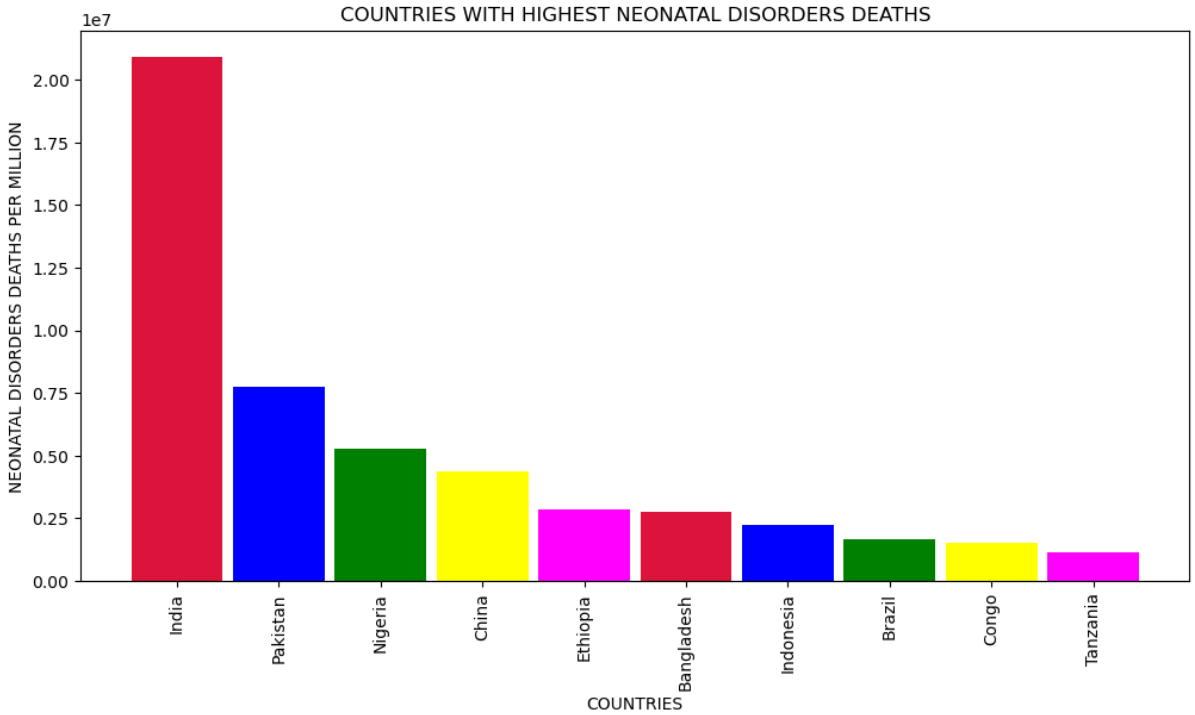
Countries with Highest Cardiovascular Diseases Death



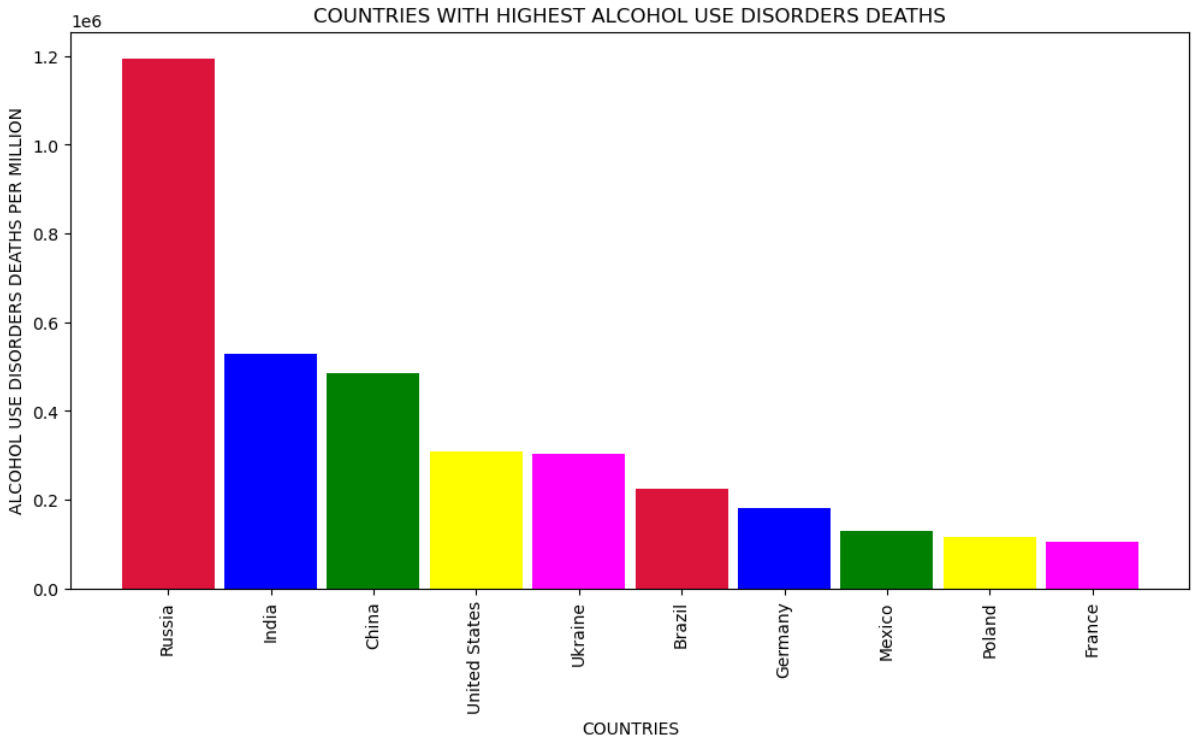
Countries with Highest Lower Respiratory Infections Death



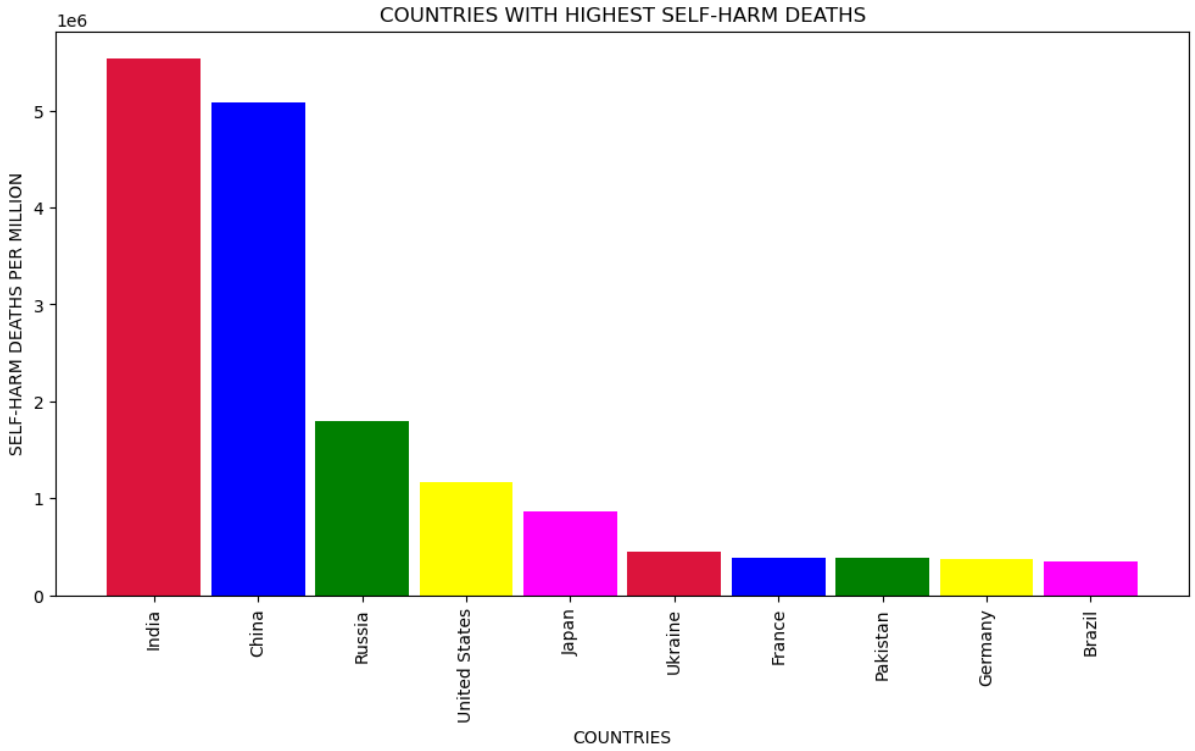
Countries with Highest Neonatal Disorders Death



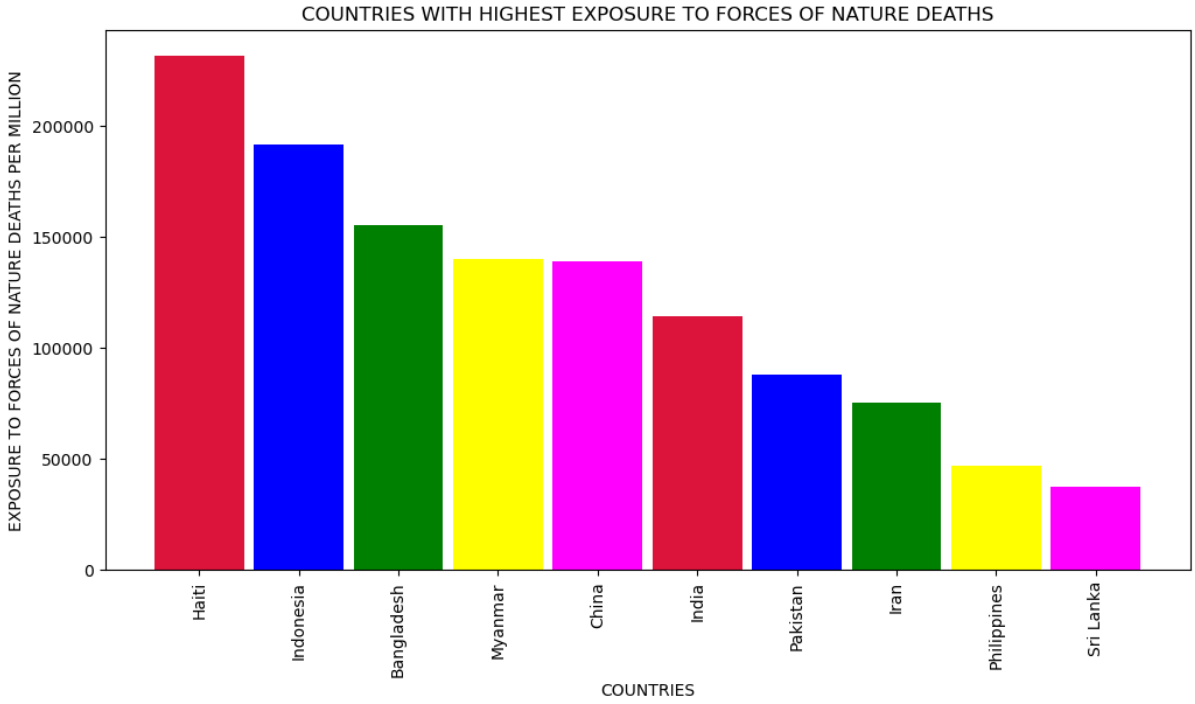
Countries with Highest Alcohol Use Disorders Death



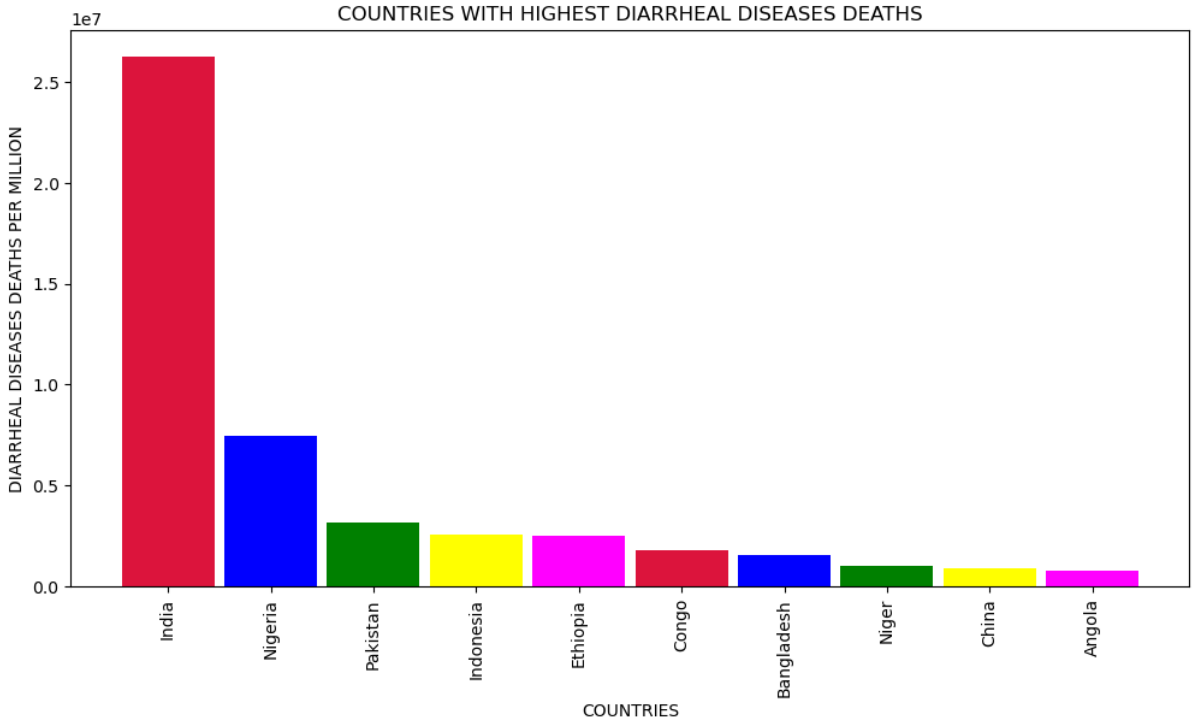
Countries with Highest Self-harm Death



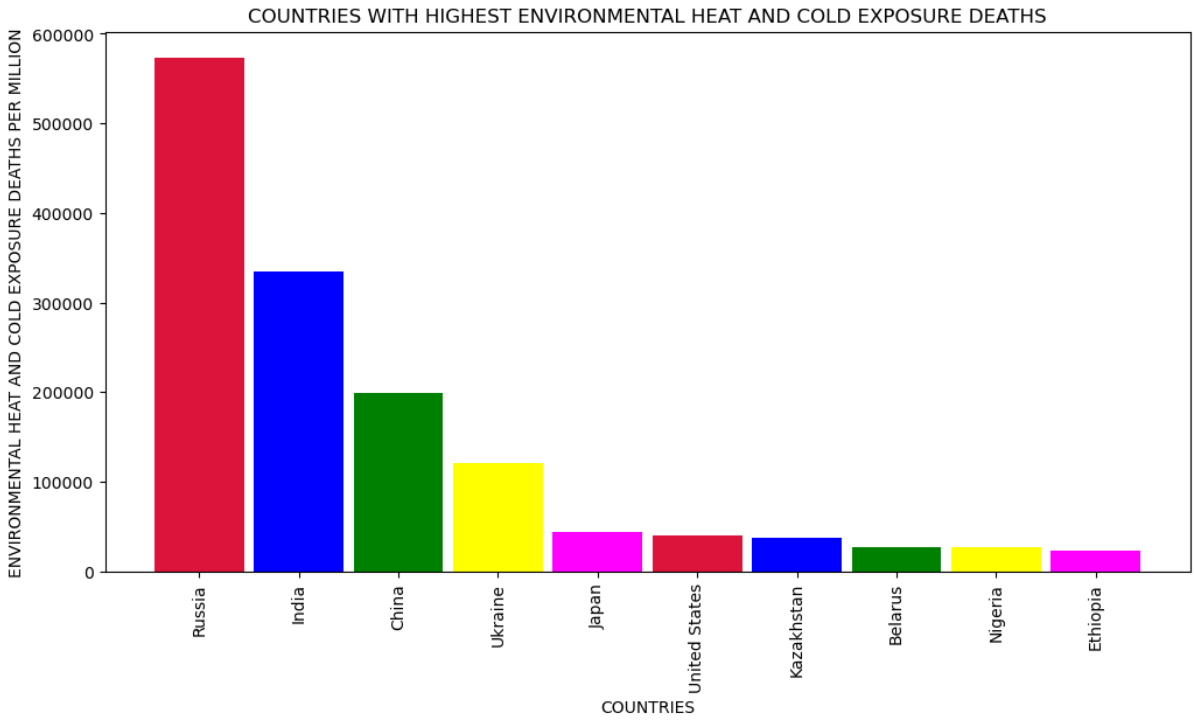
Countries with Highest Exposure to Forces of Nature Death



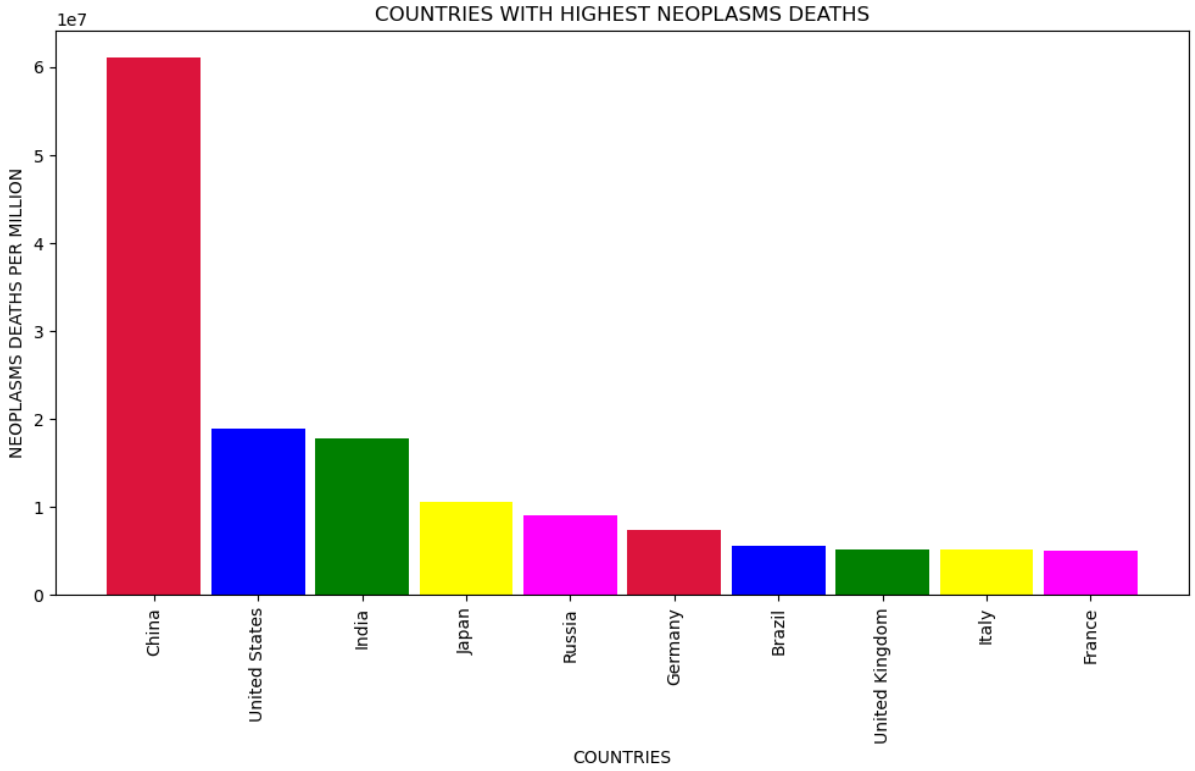
Countries with Highest Diarrheal Diseases Death



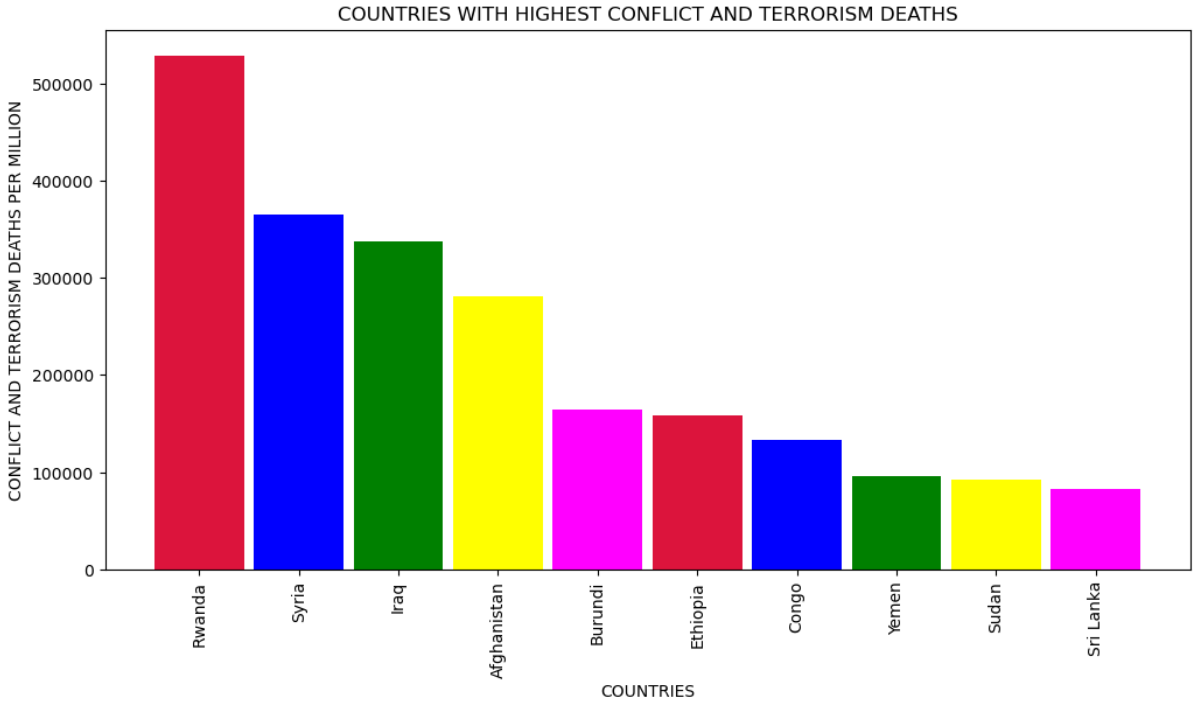
Countries with Highest Environmental Heat and Cold Exposure Death



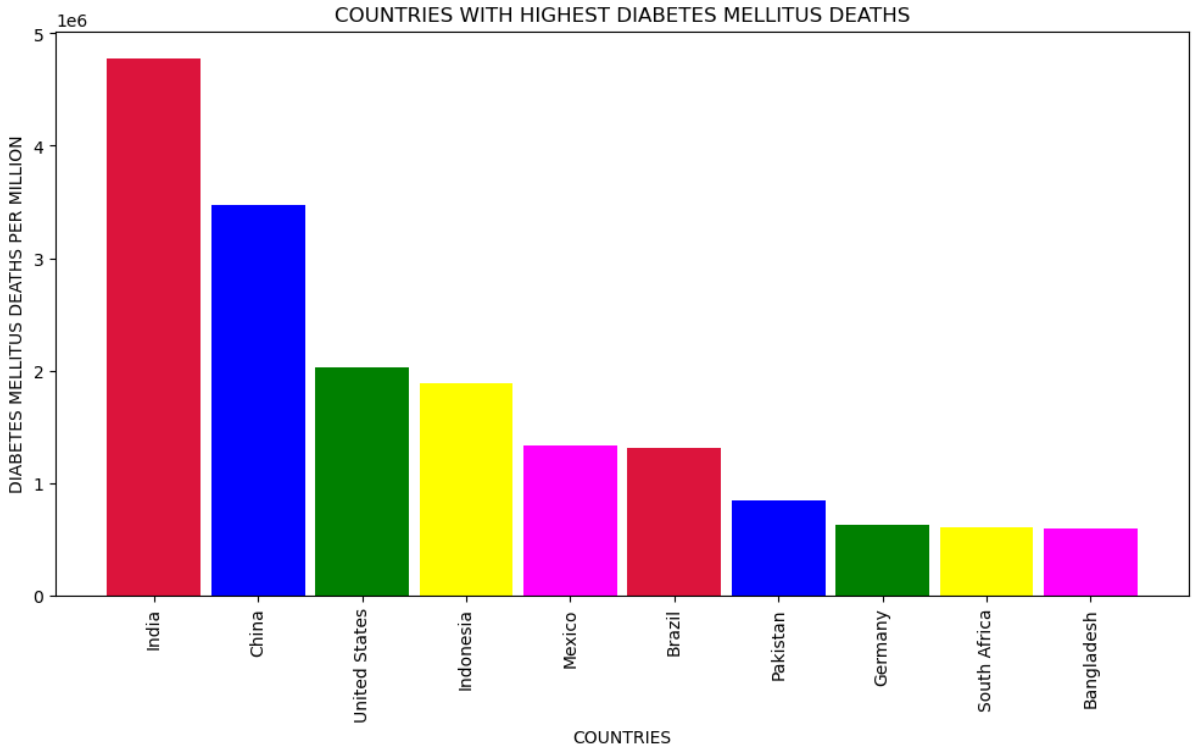
Countries with Highest Neoplasms Death



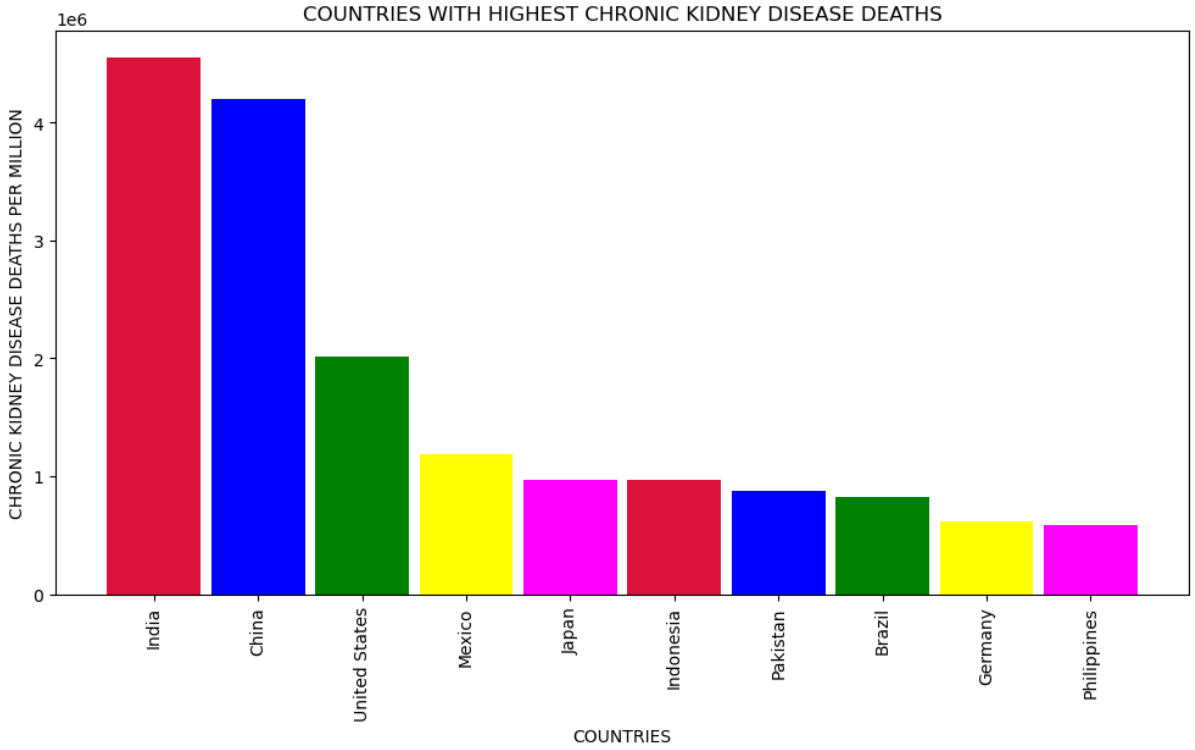
Countries with Highest Conflict and Terrorism Death



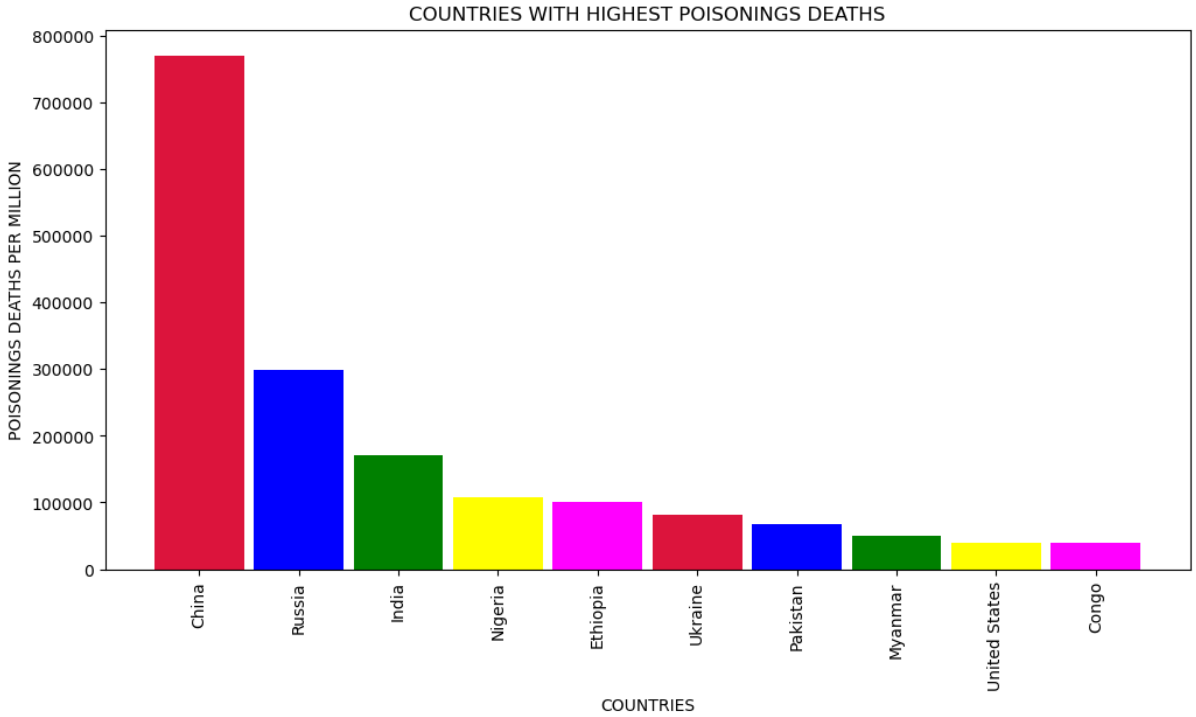
Countries with Highest Diabetes Mellitus Death



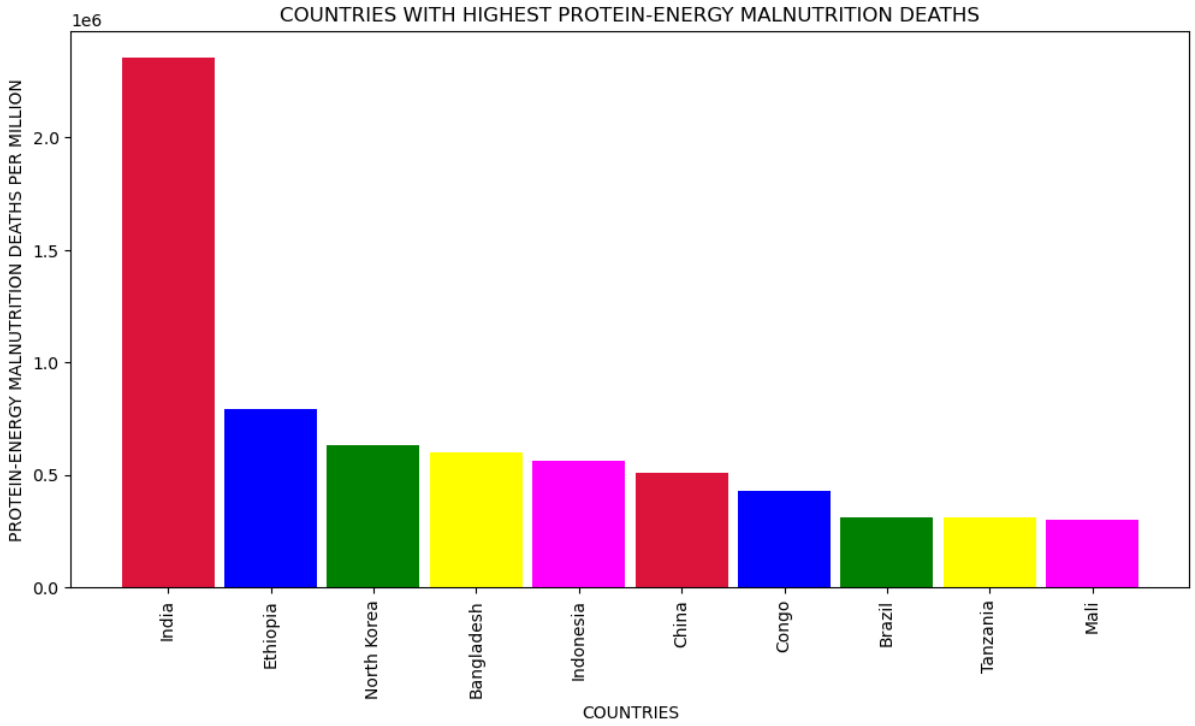
Countries with Highest Chronic Kidney Disease Death



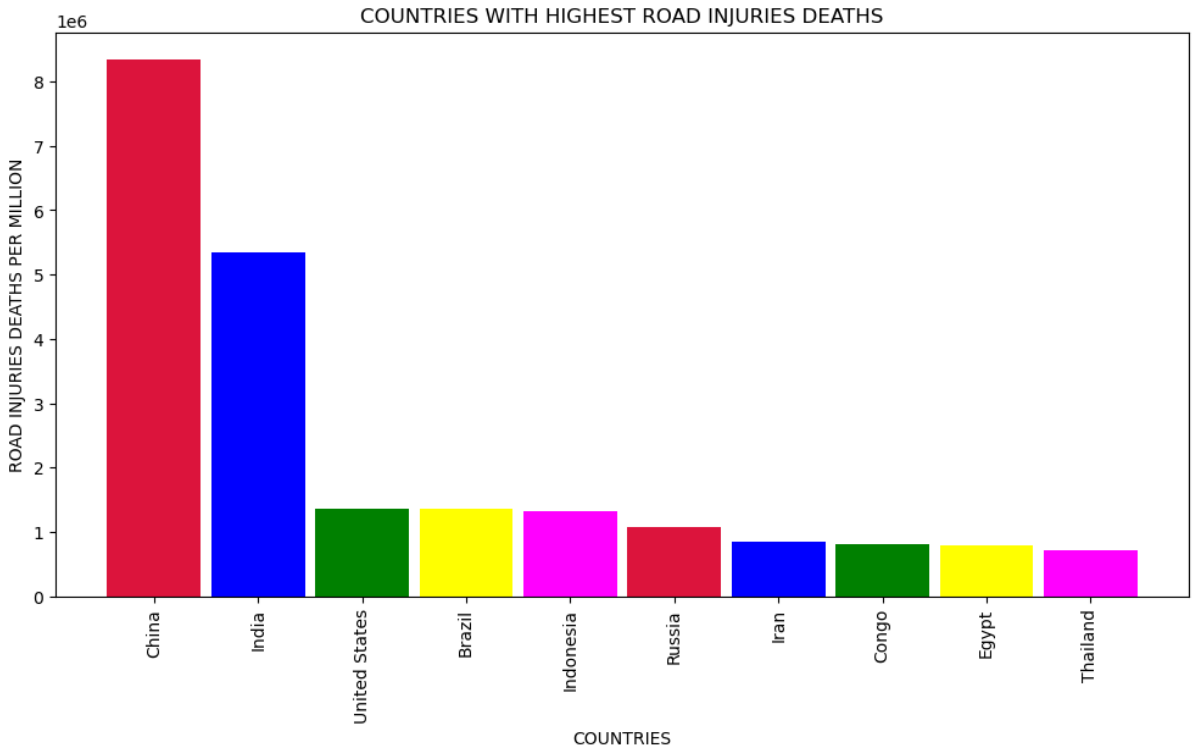
Countries with Highest Poisonings Disease Death



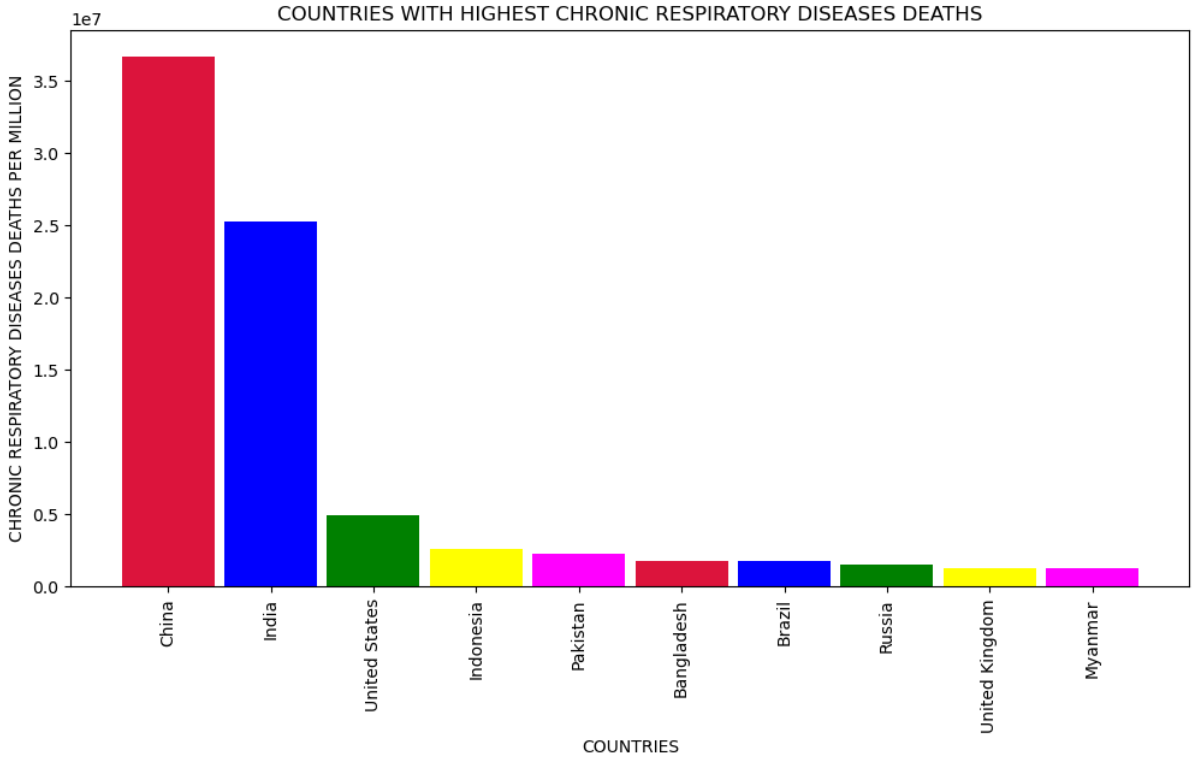
Countries with Highest Protein-Energy Malnutrition Death



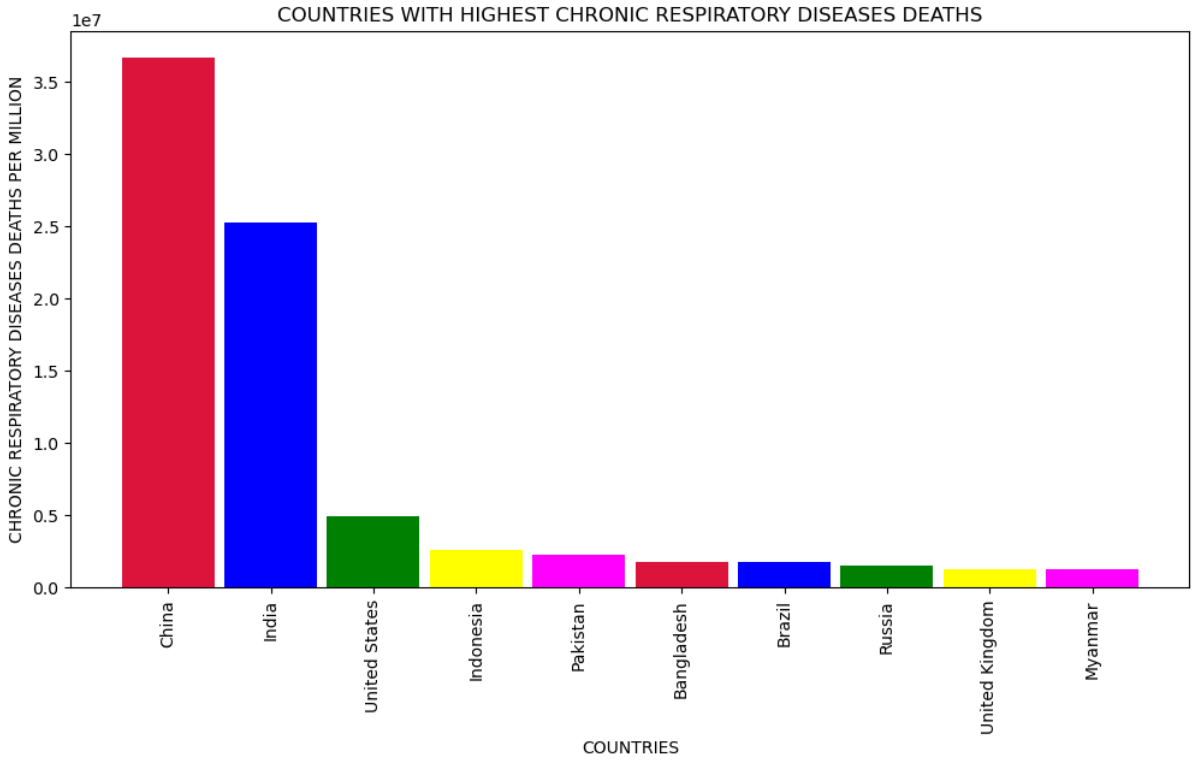
Countries with Highest Road Injuries Death



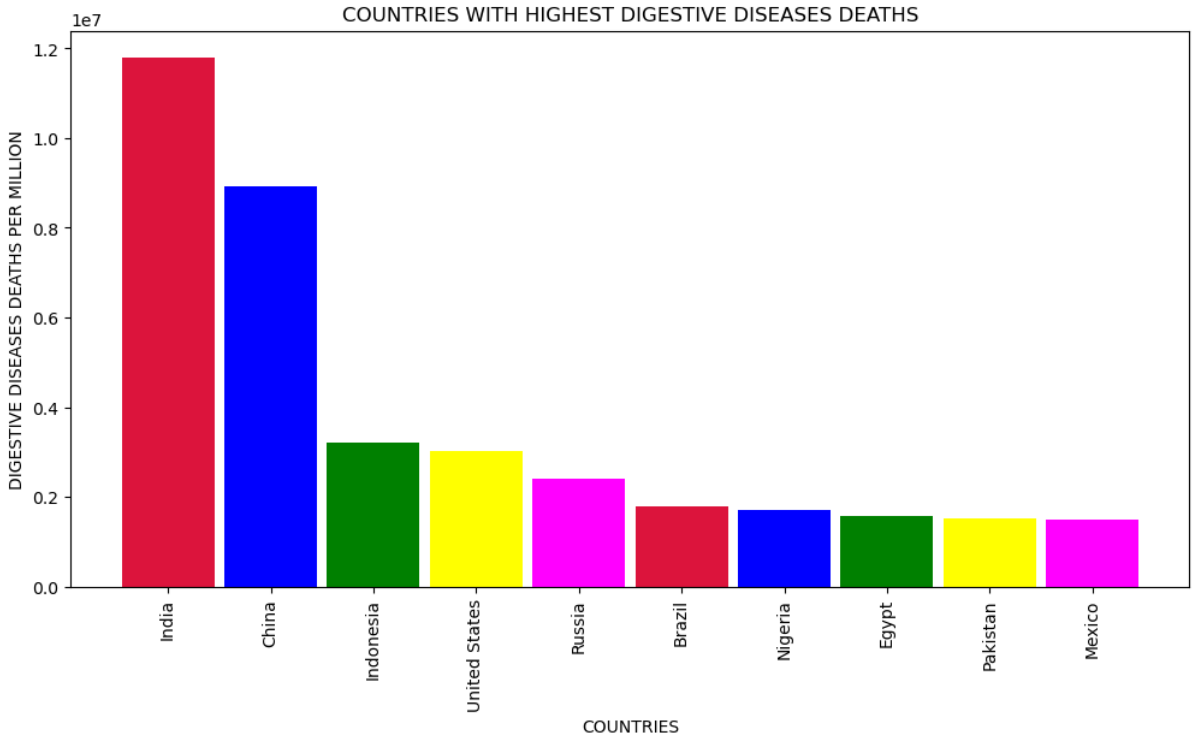
Countries with Chronic Respiratory Diseases Death



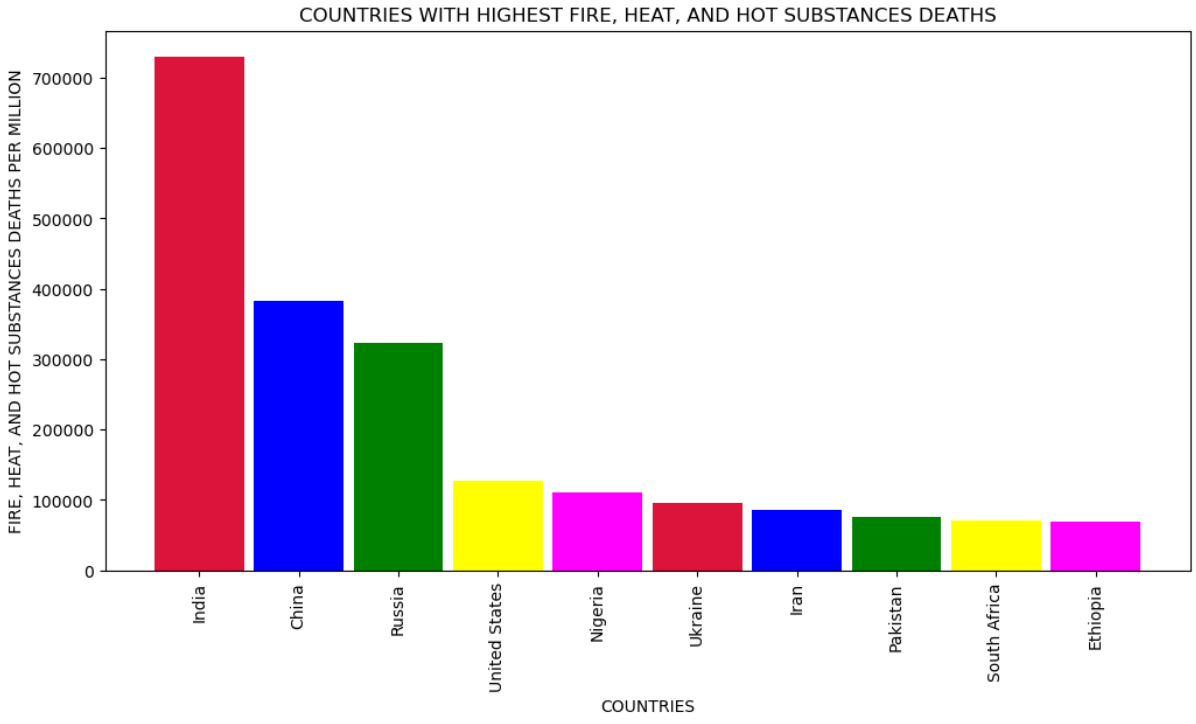
Countries with Cirrhosis and Other Chronic Liver Diseases Death



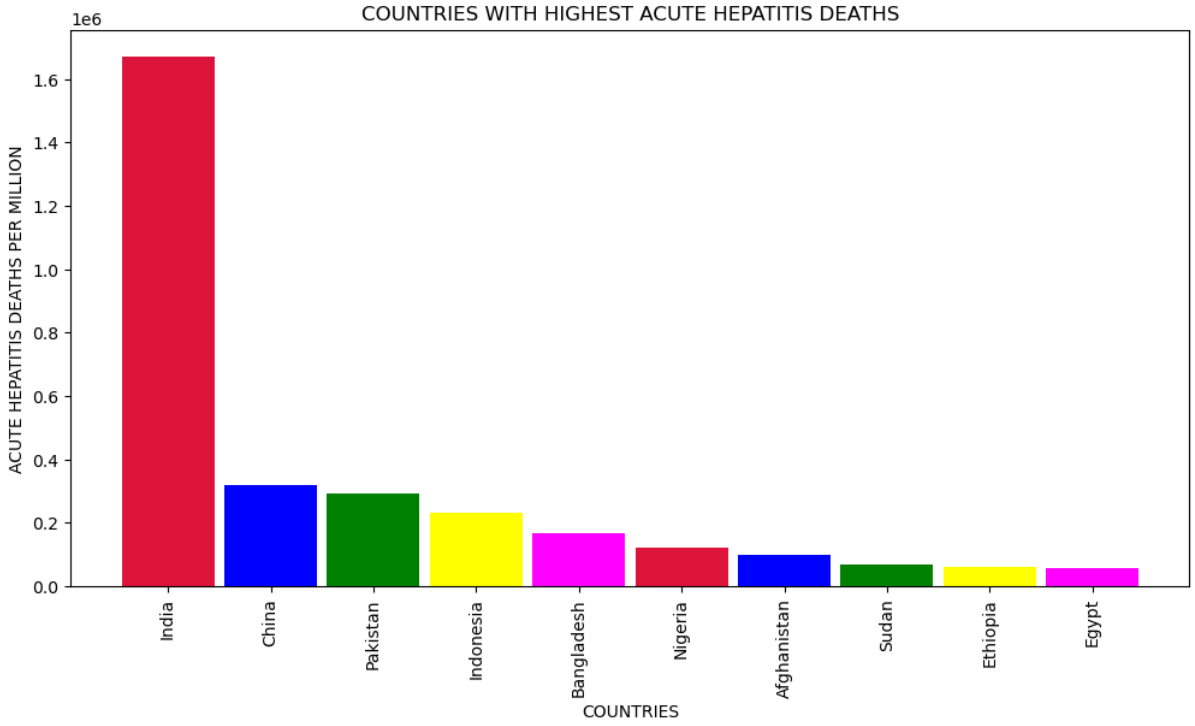
Countries with Digestive Diseases Death



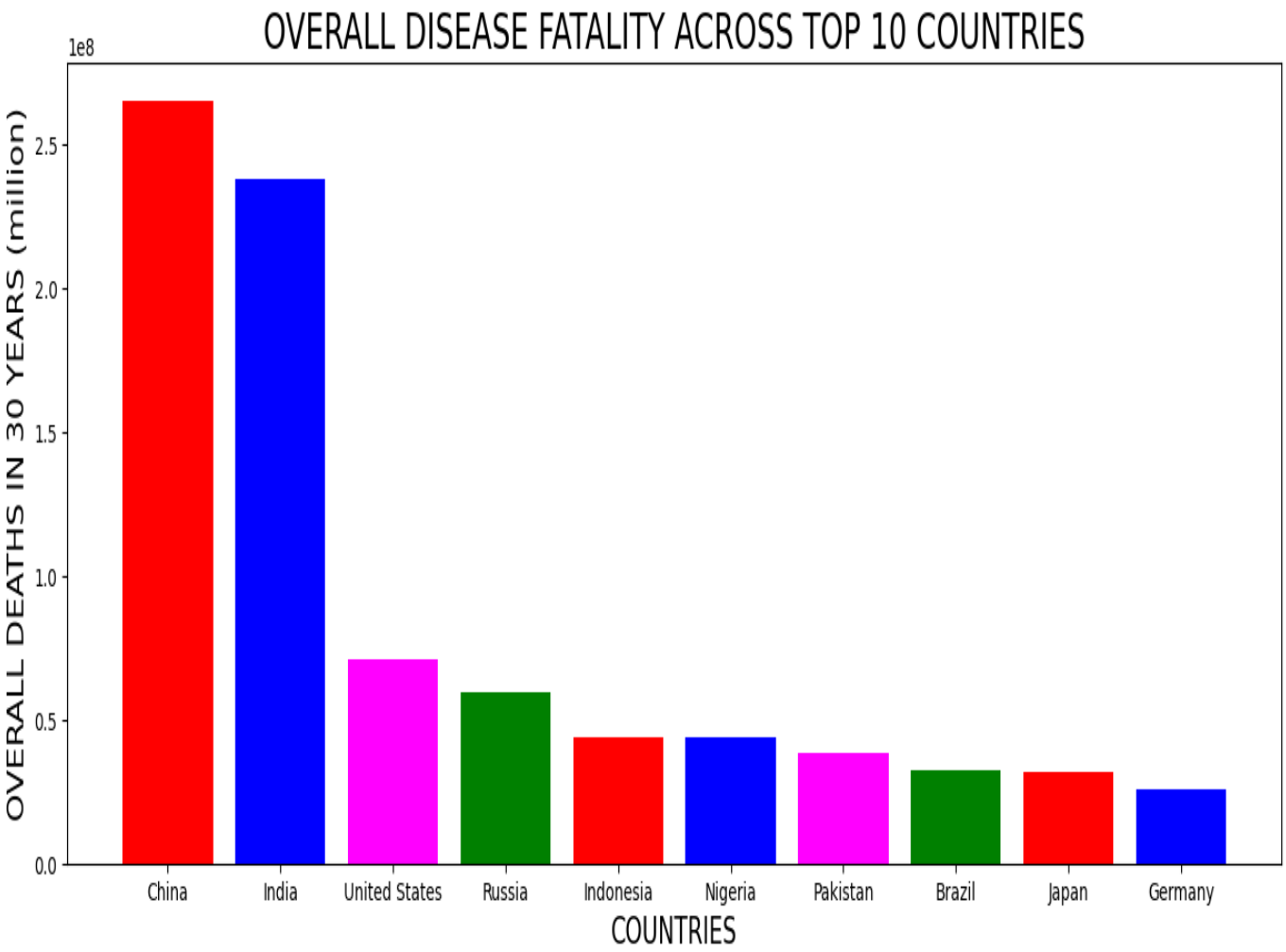
Countries with Fire, Heat, and Hot Substances Death



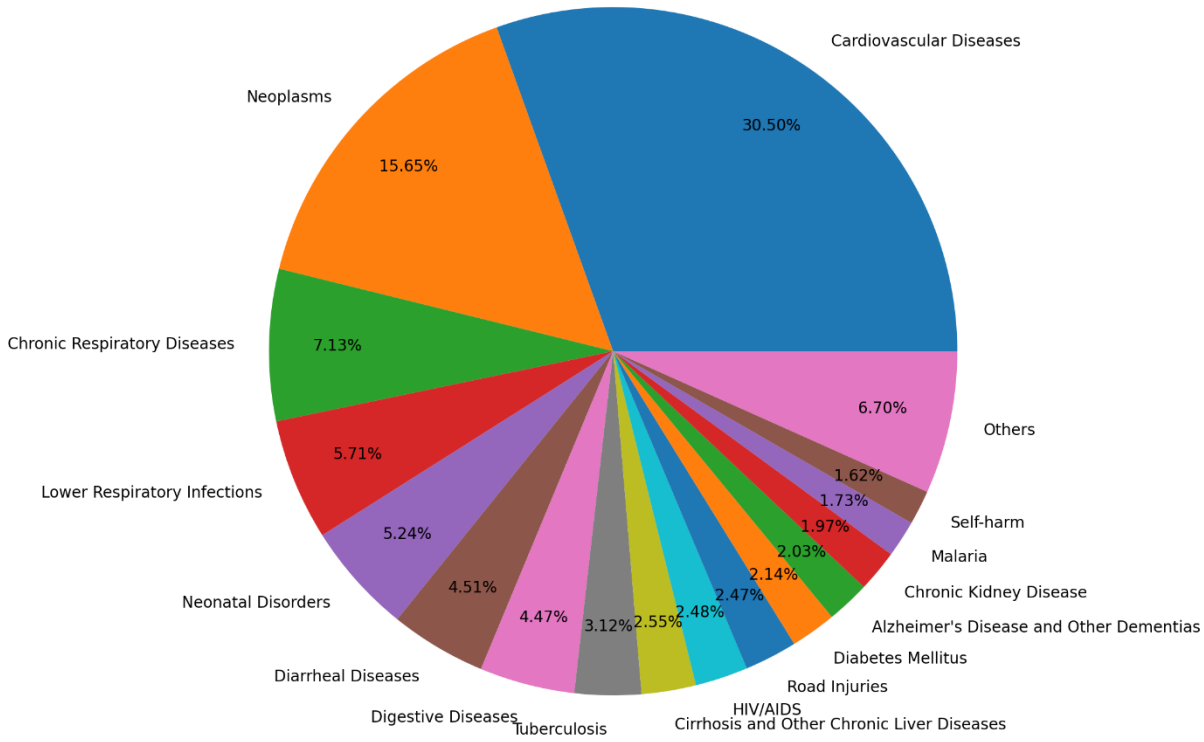
Countries with Acute Hepatitis Death



WHICH ARE THE COUNTRIES FACING OVERALL HIGHEST DEATHS DUE TO DISEASES –



WHICH ARE THE TOP KILLER DISEASES IN THE WORLD



CONCLUSION :-

Check Death Percent by Disease in World Wide :-

- **Cardiovascular Diseases 30.50% Cause of Death.**
- **Neoplasms Disease 15.67% Cause is Death.**
- **Chronic Kidney Disease 6.70% Cause of Death**
- **Others (Like Fire , Road Accident , Self Damage) 6.70% Cause of Death.**
- **Lower Respiratory Infections 5.71% Cause of Death .**
- **Neonatal Disorders 5.24% is Cause of Death**
- **Diarrheal Diseases 4.51% is Cause of Death**
- **Digestive Disease 4.51% Cause of Death.**
- **Tuberculosis Disease 3.12% is Cause of Death**
- **Cirrhosis and Other Chronic Liver Diseases 2.5 % Casue of death .**

In World Wide.

Overall Deaths in 30 Years (In Million)

- China -> More then 2.5 Million
- India -> More then 2.4 Million
- US -> Almost 1 Million
- Russia -> Almost 0.5 Million
- Indonesia -> Almost 0.5 Milliom
- Nigeria -> Almost 0.5 Million
- Pakistan -> Almost 0.4 Million
- Brazil -> Almost 0.3 Million

- Japan -> Almost 0.3 Million
- Germany -> Almost 0.3 Million

As according to dataset CHINA , INDIA AND USA face the largest brunt of deaths due to diseases in the world Cardiovascular diseases , Neoplasms (Malignancy/Cancer) and Lower Respiratory Tract Infections (for example : Pneumonia) are the top 3 killer diseases in the world. Overall Deaths in 30 Years. And left disease are also cause of death but not more the the Cardiovascular diseases, Neoplasms (Malignancy/Cancer) and Lower Respiratory Tract Infections (for example : Pneumonia.

This is observation of me from Given Datset.