

ABDELGHAFOR'S HACKATHON

DATA ANALYSIS

Global Health Insights
Analysis



Introduction

This Global Health Data Analysis Project aims to uncover global mortality trends and identify regional patterns over time. By analyzing and visualizing data from various sources, the project seeks to provide insights that can inform public health decisions. Tools such as Python, Excel, and Power BI will be used to evaluate the dataset's structure, handle missing information, and generate meaningful insights. This foundational analysis will reveal mortality trends and enhance our understanding of global health statistics.

Purpose of the Analysis

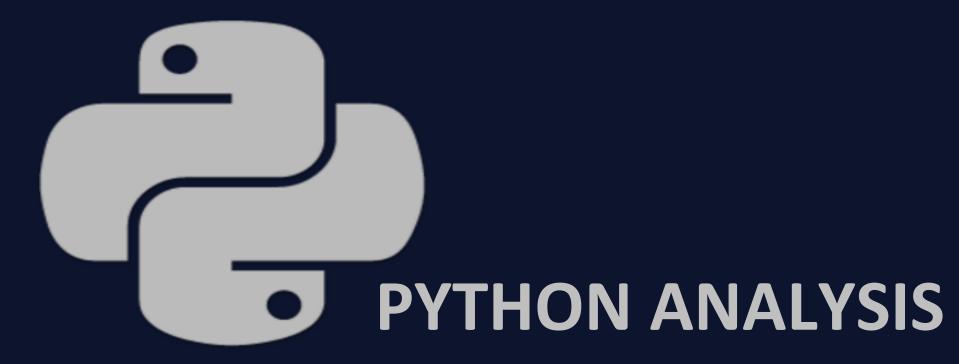
- To understand trends and patterns in causes of death over time.
- To compare statistics between different countries.
- To provide helpful insights and recommendation in making strategies regarding the public health.



CONTENTS

- Python Analysis
- Excel Workbook
- Power BI Dashboard
- Final Recommendations





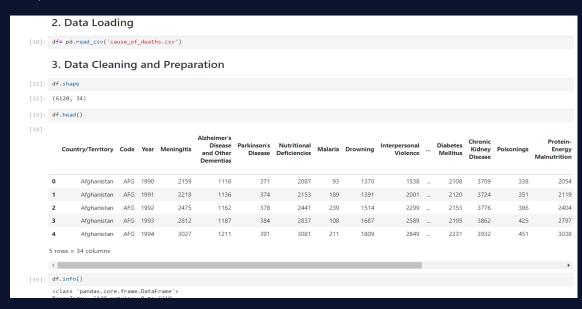


Data Overview

- Dataset Description
 - Source: cause of deaths.csv
 - Key columns: Year, Country/Territory, Various Causes of

Death

- Initial Observations
 - O Shape of the dataset
 - Data types





Data Cleaning

- Steps Taken:
 - Converted 'Year' to datetime
 - O Checked for null values and duplicates and none where found

```
[18]: df['Year'] = pd.to_datetime(df['Year'], format='%Y')

[35]: df.isnull().sum().sum()

[35]: 0

[22]: df.duplicated().sum()
[22]: 0
```



Descriptive Statistics of Causes of Death

Lower Respiratory Infections 6120.000000

```
#Statistical Analysis for only the disease column with highlighting the extreme values
disease columns=["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"]
disease df=df[disease columns]
disease df.describe().T.style.bar(subset=["mean","max"], color='#205ff2') \
                         .background gradient(subset=['std'], cmap='Reds') \
                         .background gradient(subset=['50%'], cmap='coolwarm')
                                                                                                         25%
                                                                                                                      50%
                                                                                                                                    75%
                                                                                   std
                                                                                            min
                                             count
                                                                  mean
                                                                                                                                                         max
                            Meningitis 6120.000000
                                                             1719.701307
                                                                           6672.006930 0.000000
                                                                                                    15.000000
                                                                                                                 109.000000
                                                                                                                              847.250000
                                                                                                                                                 98358.000000
Alzheimer's Disease and Other Dementias 6120.000000
                                                             4864.189379
                                                                                                    90.000000
                                                                                                                666.500000
                                                                                                                             2456.250000
                                                                           18220.659072
                                                                                        0.000000
                                                                                                                 164.000000
                                                                                                                              609.250000
                    Parkinson's Disease 6120.000000
                                                             1173,169118
                                                                           4616.156238
                                                                                        0.000000
                                                                                                    27.000000
                                                                                                                                                 76990,000000
                Nutritional Deficiencies 6120.000000
                                                             2253.600000
                                                                          10483.633601
                                                                                        0.000000
                                                                                                    9.000000
                                                                                                                119.000000
                                                                                                                             1167.250000
                                                                                                                                                268223.000000
                               Malaria 6120.000000
                                                             4140.960131
                                                                          18427.753137 0.000000
                                                                                                    0.000000
                                                                                                                  0.000000
                                                                                                                              393.000000
                                                                                                                                                280604.000000
                             Drowning 6120.000000
                                                             1683.333170
                                                                           8877.018366
                                                                                        0.000000
                                                                                                    34.000000
                                                                                                                177.000000
                                                                                                                              698,000000
                 Interpersonal Violence 6120.000000
                                                                                                    40.000000
                                                                                                                265.000000
                                                                                                                              877.000000
                                                             2083,797222
                                                                           6917.006075
                                                                                        0.000000
                                                                                                                                                 69640.000000
                    Maternal Disorders 6120.000000
                                                                                        0.000000
                                                                                                    5.000000
                                                                                                                 54.000000
                                                                                                                              734.000000
                                                             1262.589216
                                                                           6057.973183
                             HIV/AIDS 6120.000000
                                                             5941.898529
                                                                          21011.962487
                                                                                        0.000000
                                                                                                    11.000000
                                                                                                                 136.000000
                                                                                                                             1879.000000
                                                                                                                                                305491.000000
                    Drug Use Disorders 6120.000000
                                                              434.006699
                                                                                        0.000000
                                                                                                     3.000000
                                                                                                                 20.000000
                                                                                                                              129.000000
                                                                           2898.761628
                                                                                                                                                 65717.000000
                          Tuberculosis 6120.000000
                                                             7491.928595
                                                                          39549.977578 0.000000
                                                                                                    35.000000
                                                                                                                             2924.250000
                                                                                                                                                657515.000000
                Cardiovascular Diseases 6120.000000
                                                                         291577.537794
                                                                                                 2028.000000
                                                                                                               11742.000000
                                                                                                                            42546.500000
                                                                                        4.000000
```

13687.914706

48031.720009 0.000000

345.000000

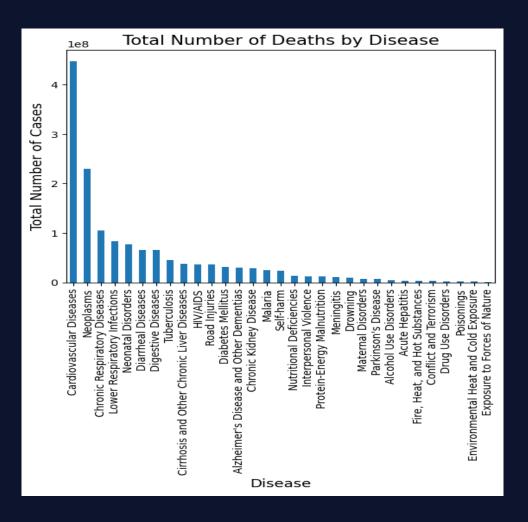
10161.250000

690913.000000



Visual Representations

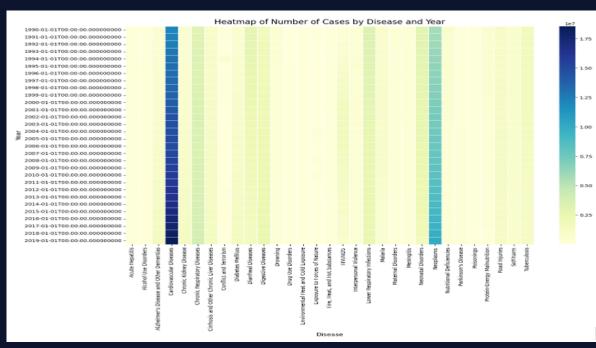
- Bar chart showing total causes of death by each disease.
- Both the descriptive statics and bar chart show that the most prevalent cause of death is **Cardiovascular Diseases** followed by **Neoplasms** then **Chronic Respiratory Diseases**.





Visual Representations

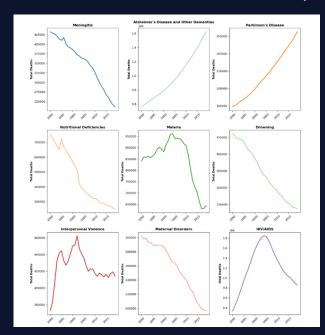
- Heatmap summarizing patterns, correlations and trends over time in the causes of death data.

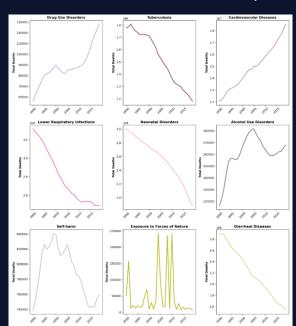


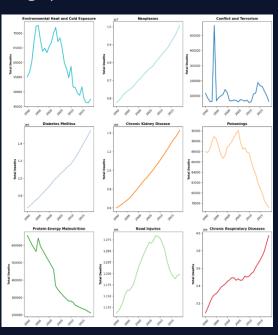


Trends Over Time

Subplots showing progression of disease deaths overtime. Some disease show constant increase like **Cardiovascular Diseases** and **Diabetes** while others decline like **Meningitis**, and causes like **Exposure to Forces of Nature** show relatively low count of deaths but with some spikes of large portions of death at once.



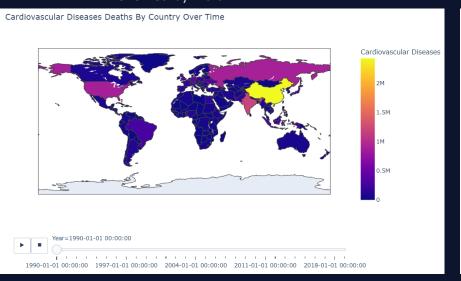


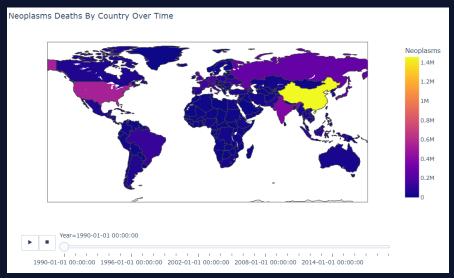




Geographic Distribution of Most Common Causes of Deaths

- Choropleth Maps:
 - O Cardiovascular Diseases, Neoplasms, Chronic Respiratory Diseases as they're most common causes of death.
- Animation Over Time:
 - O Show how deaths are distributed globally across the years with China showing the most number of deaths for all disease followed by India.

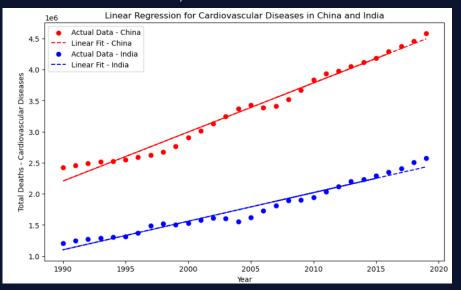


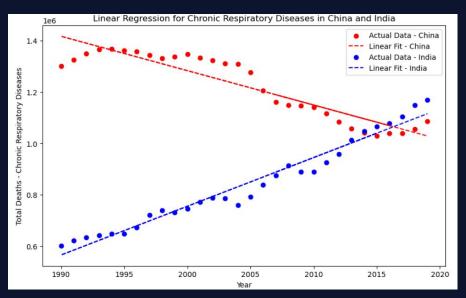




Predictive Analysis of Countries with Most Deaths

- Linear Regression Models
- Selected Diseases:
 - O Cardiovascular Diseases, Neoplasms, Chronic Respiratory Diseases
- Selected Countries:
 - o China, India



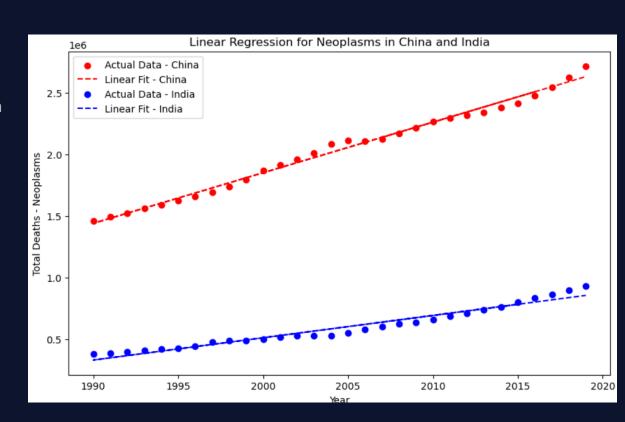




Predictive Analysis of Countries with Most Deaths

Shows that both Cardiovascular

Diseases and Neoplasms deaths are more likely to increase over time in both China and India while Chronic Respiratory Diseases will increase in China and decline in India.





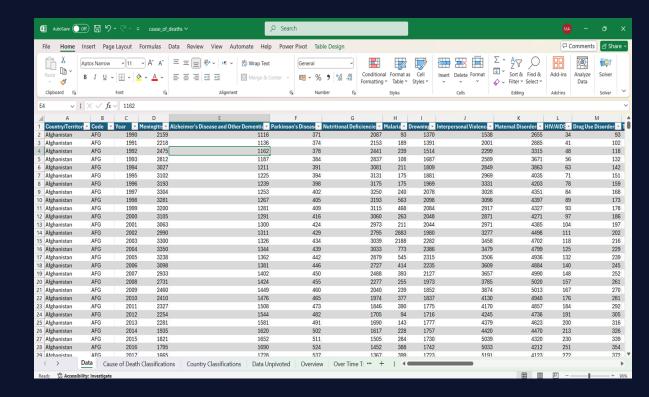


EXCEL Workbook



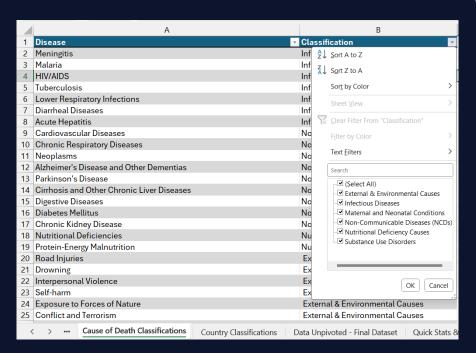
DATA OUTLINE

- 6121 rows
- 34 columns
 - Country
 - Country Code
 - Year
 - 31 different causes of death & diseases
- No duplicates or missing values found





ADDED CLASSIFICATIONS



In order to make the analysis process more meaningful and insightful, official disease and country classifications were added to the dataset.

- WHO disease classification helps categorize diseases based on their causes and impacts, guiding public health strategies and interventions:
 - **Infectious Diseases**: Caused by pathogens, spread through infection.
 - Non-Communicable Diseases (NCDs): Chronic diseases not caused by infections, often linked to lifestyle factors.
 - Nutritional Deficiency Causes: Result from inadequate intake of essential nutrients.
 - External & Environmental Causes: Related to injuries or conditions caused by environmental or external factors.
 - **Substance Use Disorders**: Disorders due to misuse or dependency on substances.
 - **Maternal and Neonatal Conditions**: Health issues affecting mothers during pregnancy and newborns.



World Bank Income Level Classification: Classifies countries into categories like Low-

income, Lower-middle-income, Upper-middle-income, and High-income based on their Gross National Income (GNI).

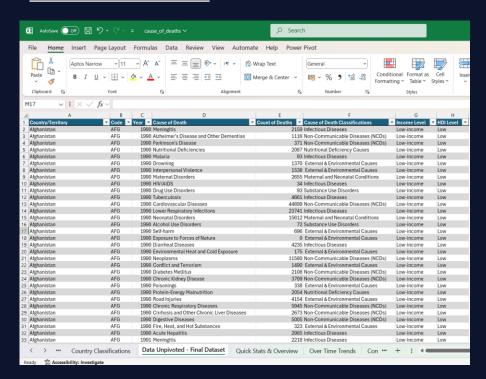
• UNDP HDI Level Classification:

Ranks countries based on the Human Development Index, which assesses overall human development and highlights the importance of improved healthcare systems to enhance life expectancy and quality of life. Levels include Low, Medium, High, and Very High.

	А	В		С						
1	Country	Income Level	¥	HDI Level						
2	Afghanistan	Low-income		Low						
3	Albania	Upper-middle-in	come	High						
4	Algeria	Upper-middle-ir	come	High						
5	American Samoa	High-income		High						
6	Andorra	High-income		Very High						
7	Angola	Lower-middle-in	come	Low						
8	Antigua and Barbuda	High-income		High						
9	Argentina	Upper-middle-in	come	High						
10	Armenia	Lower-middle-in	come	High						
11	Australia	High-income		Very High						
12	Austria	High-income		Very High						
13	Azerbaijan	Upper-middle-in	come	High						
14	Bahamas	High-income		High						
15	Bahrain	High-income		Very High						
16	Bangladesh	Lower-middle-in	come	Medium						
17	Barbados	High-income		Very High						
18	Belarus	Upper-middle-in	come	High						
19	Belgium	High-income		Very High						
20	Belize	Upper-middle-in	come	Medium						
21	Benin	Low-income		Low						
22	Bermuda	High-income		High						
23	Bhutan	Lower-middle-in	come	Medium						
24	Bolivia	Lower-middle-in	come	Medium						
25	Bosnia and			Litela						
Country Classifications Data Unpivoted - Final Dataset										



FINAL DATASET



After unpivoting the disease columns using Power Query and using Xlookup function to add the classifications to each corresponding row, the resulting dataset is in a tidy, consistent format and is ready for analysis and visualizations using pivot tables.



QUICK STATS & OVERVIEW

Cause of Death	¥	Total Number of Deaths	Average Number of Deaths	Maximum Number of Deaths
Acute Hepatitis		3,784,791		64,305
Alcohol Use Disorders		4,819,018		55,200
Alzheimer's Disease and Other Dementias		29,768,839	4,864	320,715
Cardiovascular Diseases		447,741,982	73,160	4,584,273
Chronic Kidney Disease	ı	28,911,692	4,724	222,922
Chronic Respiratory Diseases		104,605,334	17,092	1,366,039
Cirrhosis and Other Chronic Liver Diseases	ı	37,479,321	6,124	270,037
Conflict and Terrorism		3,294,053	538	503,532
Diabetes Mellitus	ı	31,448,872	5,139	273,089
Diarrheal Diseases		66,235,508	10,823	1,119,477
Digestive Diseases		65,638,635	10,725	464,914
Drowning		10,301,999	1,683	153,773
Drug Use Disorders		2,656,121	434	65,717
Environmental Heat and Cold Exposure		1,788,851	292	29,048
Exposure to Forces of Nature		1,490,132	243	222,641
Fire, Heat, and Hot Substances		3,602,914	589	25,876
HIV/AIDS		36,364,419	5,942	305,491
Interpersonal Violence		12,752,839	2,084	69,640
Lower Respiratory Infections		83,770,038	13,688	690,913
Malaria		25,342,676		
Maternal Disorders		7,727,046		
Meningitis		10,524,572	1,720	98,358
Neonatal Disorders		76,860,729	12,559	852,761
Neoplasms		229,758,538	37,542	2,716,551
Nutritional Deficiencies		13,792,032		
Parkinson's Disease		7,179,795	1,173	76,990
Poisonings		2,601,082	425	30,883
Protein-Energy Malnutrition		12,031,885		
Road Injuries		36,296,469	5,931	
Self-harm		23,713,931		220,357
Tuberculosis		45,850,603	7,492	657,515

Using Excel pivot table and conditional formatting, the following insights can be derived:

- Leading Causes of Death:

- Cardiovascular Diseases have the highest total number of deaths (447.7 million), followed by Neoplasms (cancers) (229.8 million) and Chronic Respiratory Diseases (104.6 million). These are significant global health concerns.
- Other high-mortality causes include Lower Respiratory Infections (83.8 million) and Diarrheal Diseases (66.2 million).

- High-Impact Events:

• Causes like Conflict and Terrorism and Exposure to Forces of Nature show relatively low total deaths (3.2 million and 1.5 million respectively), but their maximum death figures (503,532 and 222,641 respectively) indicate that extreme events or disasters account for a large portion of deaths at once.



QUICK STATS & OVERVIEW

Cause of Death	₩	Total Number of Deaths	Average Number of Deaths	Maximum Number of Deaths
Acute Hepatitis		3,784,791	618	64,305
Alcohol Use Disorders		4,819,018	787	55,200
Alzheimer's Disease and Other Dementias		29,768,839	4,864	320,715
Cardiovascular Diseases		447,741,982	73,160	4,584,273
Chronic Kidney Disease		28,911,692	4,724	222,922
Chronic Respiratory Diseases		104,605,334	17,092	1,366,039
Cirrhosis and Other Chronic Liver Diseases		37,479,321		270,037
Conflict and Terrorism		3,294,053	538	503,532
Diabetes Mellitus		31,448,872	5,139	273,089
Diarrheal Diseases		66,235,508	10,823	1,119,477
Digestive Diseases		65,638,635	10,725	464,914
Drowning		10,301,999	1,683	153,773
Drug Use Disorders		2,656,121	434	65,717
Environmental Heat and Cold Exposure		1,788,851	292	29,048
Exposure to Forces of Nature		1,490,132	243	222,641
Fire, Heat, and Hot Substances		3,602,914	589	25,876
HIV/AIDS		36,364,419	5,942	305,491
Interpersonal Violence		12,752,839	2,084	69,640
Lower Respiratory Infections		83,770,038	13,688	690,913
Malaria		25,342,676	4,141	280,604
Maternal Disorders		7,727,046	1,263	107,929
Meningitis		10,524,572	1,720	98,358
Neonatal Disorders		76,860,729	12,559	852,761
Neoplasms		229,758,538	37,542	2,716,551
Nutritional Deficiencies		13,792,032	2,254	268,223
Parkinson's Disease		7,179,795	1,173	76,990
Poisonings		2,601,082	425	30,883
Protein-Energy Malnutrition	ı	12,031,885	1,966	202,241
Road Injuries		36,296,469	5,931	329,237
Self-harm		23,713,931	3,875	220,357
Tuberculosis		45,850,603	7,492	657,515

- Low-Impact Events

 Poisonings, Drug Use Disorders, and Environmental Heat and Cold Exposure show relatively smaller maximum deaths, which suggests that their impact is more spread out over time rather than being tied to singular, catastrophic events

- Consistent Killers:

 Non-communicable diseases (Cardiovascular Diseases and Neoplasms) have very high average deaths per event (73,160 and 37,542 respectively), showing that they are widespread and consistently deadly. This reflects the global shift towards lifestylerelated conditions as the primary health burden.



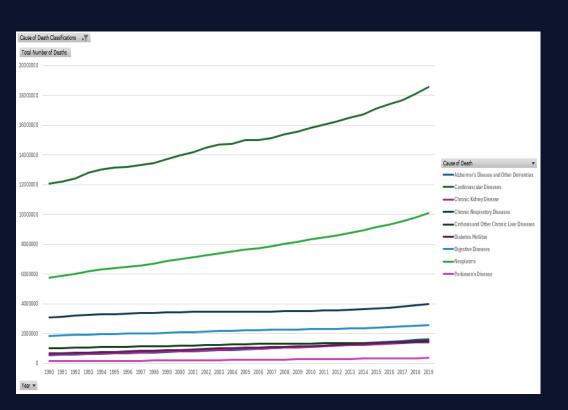
QUICK STATS & OVERVIEW

Using XLookup and Concatenation to highlight the years and countries at which the leading causes of death and impactful events caused the most number of deaths.

=XLOOKUP(A5&D5,	'Data U	npivote	d - Fir	al Data	set'!D	:D&'Dat	a Unpivo	ted - F	inal Da	taset'!	E:E,'Da	nta Unpi	ivoted -	- Final	Dataset	'!c:c)
G	Н	- 1	J	K	L	М	N	0	Р	Q	R	S	Т	U	V	W
The most prevalent cause of death appears to be Cardiovascular diseases followed by Neoplasms. At what year and which country they caused the most deaths?							At what year and countries high impactful events like Conflict and Terrorism and Exposure to Forces of Nature caused the most number of deaths?									
Cardiovascular Diseases Year Country	2019 China						Conflict an Year Countrty	d Terrorism	1994 Rwanda							
Neoplasms Year Country	2019 China						Exposure to Year Country	o Forces of I	Nature 2010 Haiti							



TRENDS OVER TIME

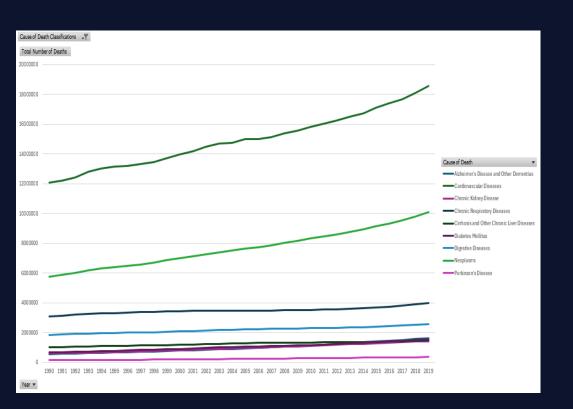


Overtime trend analysis focused on **NCDS** shows that:

- The global burden of **NCDs** has substantially increased over the past three decades, with **Cardiovascular Diseases** remaining the leading cause of death and consistently accounting for the highest number of deaths throughout the period, more than all other **NCDs** combined.
- Neoplasms were the second-leading cause of death after Cardiovascular Diseases, with deaths rising from 5.75 million in 1990 to 10.07 million in 2019.



TRENDS OVER TIME



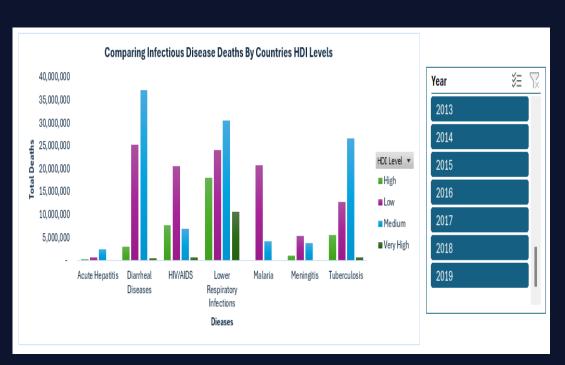
The rise in **Alzheimer's and related dementias** deaths suggests growing public health challenges related to aging populations and lifestyle changes.

Diabetes-related deaths rose from 661,085 in 1990 to 1,549,593 in 2019, showing a steady but moderate increase, likely reflecting global lifestyle changes.

Deaths from cirrhosis and other **chronic liver diseases** also increased, but at a relatively slower pace compared to other diseases.



COMPARATIVE ANALYSIS



Comparing death counts of **Infectious Diseases** by HDI levels over time shows that:

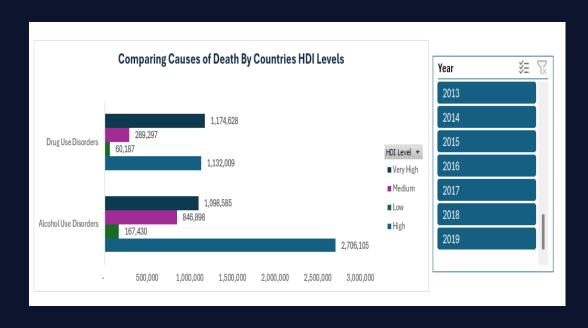
- Diseases like HIV/AIDS, Diarrheal Diseases, and Malaria have significantly higher death counts in low and medium HDI regions.
- Lower Respiratory Infections affect both low and high HDI countries, though the burden is higher in less developed regions.
- Tuberculosis and Meningitis have a higher mortality rate in less developed countries.
- Malaria is nearly eradicated in very high HDI regions but remains a major cause of death in low HDI countries.

Which all indicate the weaker healthcare infrastructure in countries with lower HDI levels.



COMPARATIVE ANALYSIS

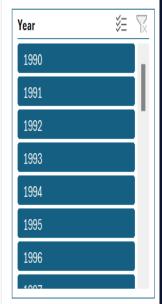
Comparing deaths by **Alcohol** and **Drug Use Disorders** by countries' HDI levels overtime shows that they result in more deaths in high and very high HDI countries, reflecting lifestyle-related health issues.



The comparative analysis highlights the differing health challenges across development levels, with infectious diseases dominating in lower HDI countries and lifestyle-related issues prevalent in higher HDI countries.



Correlation Between Countries Income Levels and Nutritional Causes of Death 6.000 5.000 Nutritional Deficiencies 1,000 Protein-Energy Malnutrition



Correlation

Is there a correlation between countries income levels and the prevalence of nutritional deficiency causes of death?

Low and lower-middle income countries consistently had higher deaths due to nutritional deficiencies across each and all years.







Power BI Dashboard





Power BI Dashboard

- The visuals included in the dashboard are filtered by the country, year, and cause of death using slicers.
- A bar chart was created to visualise the top 10 causes of death.
- The map chart demonstrates the top 5 countries based on the number of deaths.
- The line chart shows the number of deaths through time and prediction for the following 5 years.



Final Recommendations

 Governments should prioritize addressing the top 10 causes of death, with health ministries developing programs to raise public awareness about early disease detection and promoting safety measures to reduce mortality from these causes.

 Governmental authorities must inform citizens if a pandemic exists or is suspected, or if there are any spreading diseases or fatal risks in the countries they are visiting or residing in. This can help limit the spread of disease and protect individuals from harm.



Final Recommendations

 Continuous analysis of health data and causes of death can provide valuable, data-driven insights into mortality trends. This information can be used to predict the number of deaths and their causes, enabling authorities to take preventive actions, improve healthcare services, and respond more rapidly to save lives.

 Establishing a global network to track common causes of death and share best practices could lead to better strategies for saving lives worldwide.



Thank You!