Causes of Death Worldwide: 1990-2019

March 26, 2024

Jan 2024 Cohort

G. Gorham, T. Ngo, H. Kang, J. Creanza,J.Bein, KK.Kothandath.

Overview

Causes of Death Worldwide: 1990-2019

Why investigate this?

We seek to test the following hypotheses, the results of which would be meaningful to several stakeholders:

1. Lifestyle risk factors rather than environmental risk factors are the leading cause of death worldwide, irrespective of a country's income classification
2. Deaths stemming from environmental risk factors are higher in low income and lower middle income countries than in higher income countries
3. There exists a strong, positive correlation between the two leading causes of death in each country

Causes of Death Worldwide: 1990-2019

Who would benefit from this analysis?

Stakeholders at many levels could use the results to effectuate lasting, positive change.

1. Policymakers: the results could assist in determining where, and to what extent resources are deployed
2. Healthcare professionals: could evaluate overarching treatment options and, perhaps education curricula, e.g. methods of sanitation
3. Biologists: in the case of environmental risk factors, biologists could use the results to determine whether the efficacy of previous remediation methods have deteriorated, e.g., if organisms are developing resistance to pesticides

Causes of Death Worldwide: 1990-2019

Tools used in this analysis

1. Jupyter notebooks: A “main” notebook (Common Causes of Death) was created and team members created their own notebooks to focus on specific questions.
2. Python libraries:
    a. pandas
    b. numpy
    c. mathplotlib
    d. scipy
    e. seaborn
    f. hvplot
    g. geoviews
    h. geopandas
3. GitHub: https://github.com/JoeyForgetabbait/Common-Causes-of-Death-

Sources of Data

Source Data: Overview

For our analysis we obtained four datasets:

Cause of Deaths: Obtained from Kaggle, this dataset was comprised of 6,121 rows and 34 columns.
[country/territory; country code; year; causes of death]

U.S. Foreign Aid: Obtained from the U.S. Agency for International Development (USAID), this dataset was comprised of 21,675 rows and 11 columns.
[country code and name; region ID and name; income group; transaction type; aid amounts]

Population Data: Obtained from The World Bank Group, this dataset was comprised of this dataset contains 5,760 rows and 3 columns
[country, years]

Geographic Data: Obtained from Google's public data, this dataset was comprised of 246 rows and 4 columns.
[country abbreviation, latitude, longitude, country name]

Locations of source data and acknowledgements are included in the Appendix.

Source Data: Examples

Causes of Death:

Country/Territory	Code	Year	Meningitis	Malaria
Afghanistan	AFG	1990	2159	93
Afghanistan	AFG	1991	2218	189
Afghanistan	AFG	1992	2475	239
Bahrain	BHR	1992	5	0

Source Data: Examples

U.S. Foreign Aid:

Country Code	Country Name	Income Group Name	Transaction Type Name	Fiscal Year	current_amount	constant_amount
ABW	Aruba	High Income Country	Disbursements	2004	1000	1487
ABW	Aruba	High Income Country	Disbursements	2005	29270	42255
ABW	Aruba	High Income Country	Disbursements	2006	1000	1398
ABW	Aruba	High Income Country	Disbursements	2007	61577	83794

Source Data: Examples

Population Data:

Code	Year	Population
AFG	1990	10694796
AFG	1991	10745167
AFG	1992	12057433
AFG	1993	14003760
AFG	1994	15455555
AFG	1995	16418912

Source Data: Examples

Geographic Data:

country	latitude	longitude	name
AD	42.54625	1.601554	Andorra
AE	23.42408	53.84782	United Arab Emirates
AF	33.93911	67.70995	Afghanistan
AG	17.06082	-61.7964	Antigua and Barbuda
AI	18.22055	-63.0686	Anguilla
AL	41.15333	20.16833	Albania
AM	40.0691	45.03819	Armenia

Data Processing

Data Processing

- After importing the source data, we performed data cleaning and data preparation tasks, e.g.:
  - filtering the data to capture the columns of interest
  - renaming columns for ease of viewing
  - changing data types to facilitate file merging
- Next, we merged the causes of death dataset with the population dataset.
- Then, we merged the USA foreign aid dataset, Country API Google dataset, and the World Bank data set.
- Lastly, we created several dataframes specific to the questions for which we are interested. For example:
  - cause_aid_pop.df
  - per_capita_death.df
  - top_death_country_year.df
  - bottom_country_year.df:
  - afg_death_top.df & afg_death_bottom.df
  - usa_death.df, usa_death_top.df, usa_death_bottom.df

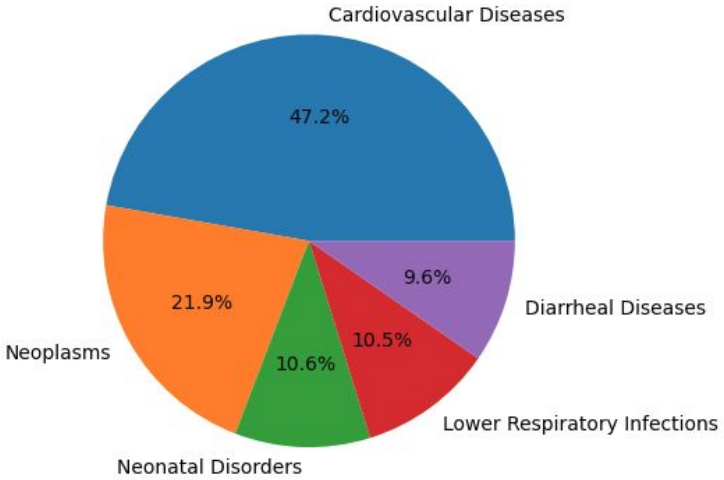
Data Analysis

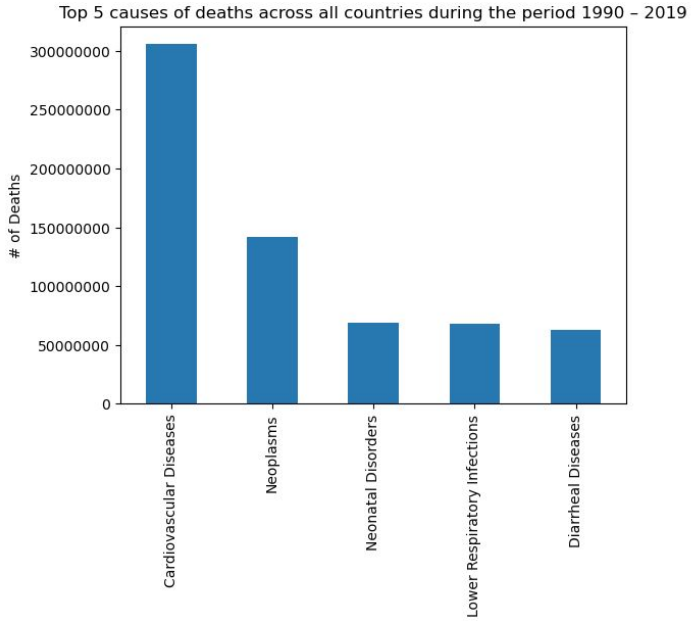
Data Analysis: Overview

To evaluate our hypotheses, analyses pertaining to the following questions were conducted simultaneously:

- What were the top 5 and bottom 5 causes of death worldwide?
- Did U.S. aid to other countries impact causes of death, and if so, which causes?
- What were top 5 and bottom 5 causes of death in high income and low income countries?
- What was the correlation, if any, between the top 2 causes of death within specific countries?

What were the top 5 causes of deaths across all countries 1990 – 2019?

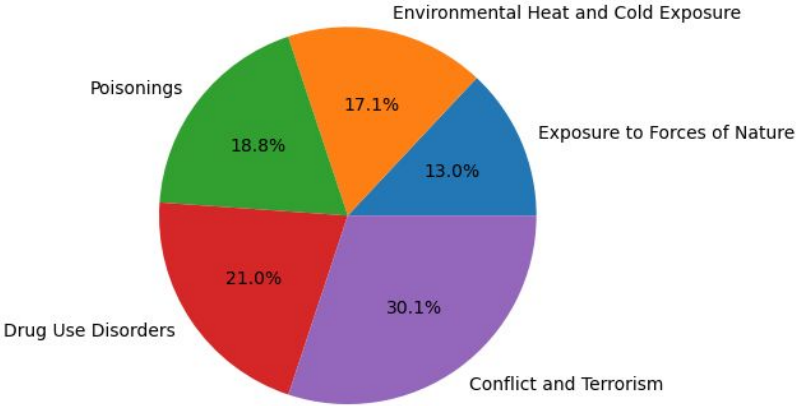


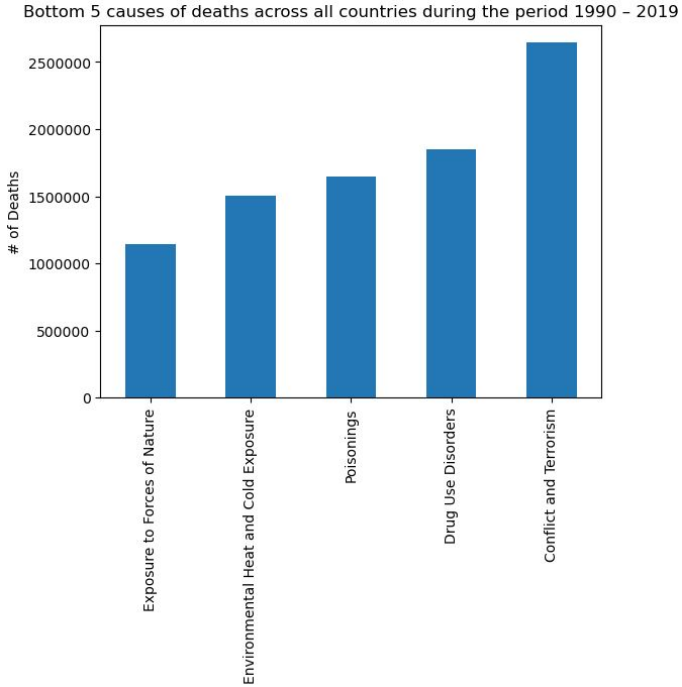


Observations and Conclusions

- Cardiovascular Diseases account for *nearly half* of the deaths out of the top 5.
- Diarrheal Diseases - which are preventable and treatable - accounted for the least percentage of deaths.

What were the bottom 5 causes of death across all countries 1990 – 2019?

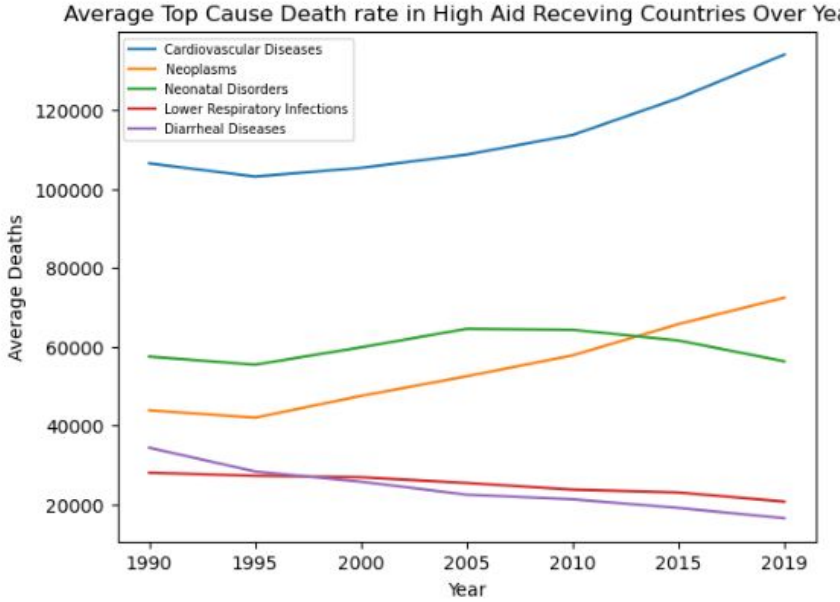


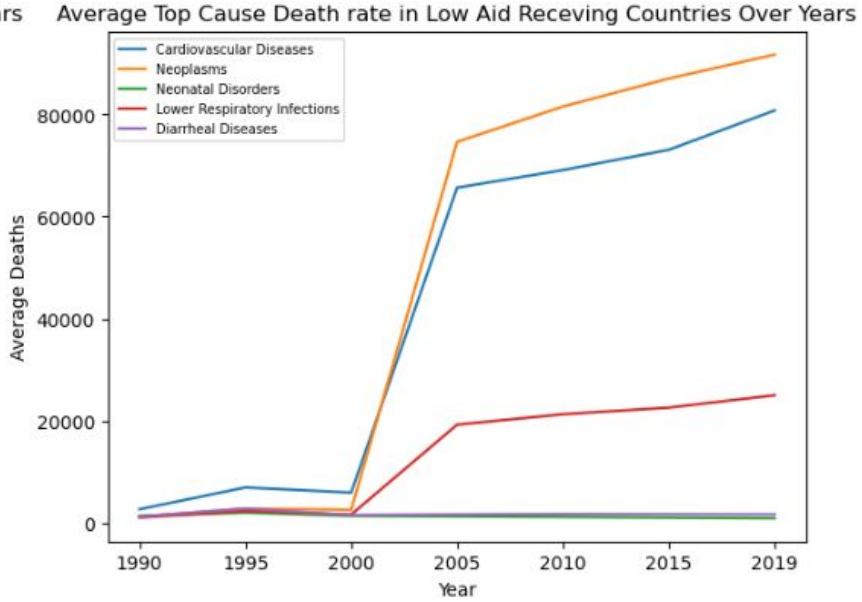


Observations and Conclusions

- Of the bottom 5 causes of death across all countries, nearly 20% were the result of accidental poisonings. This suggests adverse environmental conditions and lack of healthcare resources, e.g. poison control centers.
- The number of deaths arising from conflict and terrorism were 50% greater than those arising from drug use disorders.

Did US foreign aid affect the average top causes of deaths in the countries receiving the most and least aid?

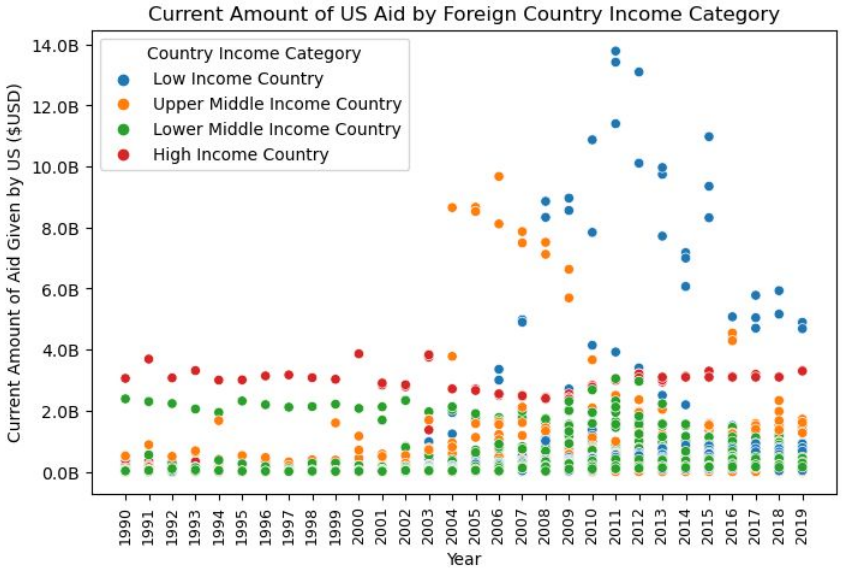


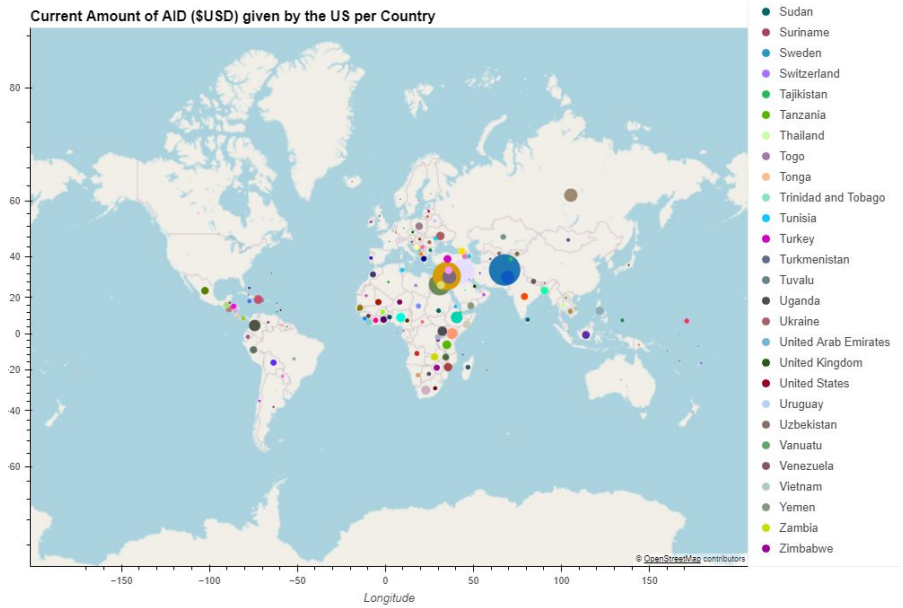


Observations and Conclusions

- High Aid Receiving Countries: the data suggests that US aid efforts had a non-trivial, positive impact on causes of death arising from public health-related factors (respiratory infections and diarrheal disorders).
- Low Aid Receiving Countries: the data suggests that the reduction/elimination of US aid had a tremendous impact from 2000 - 2019.

US Foreign Aid by Income Category

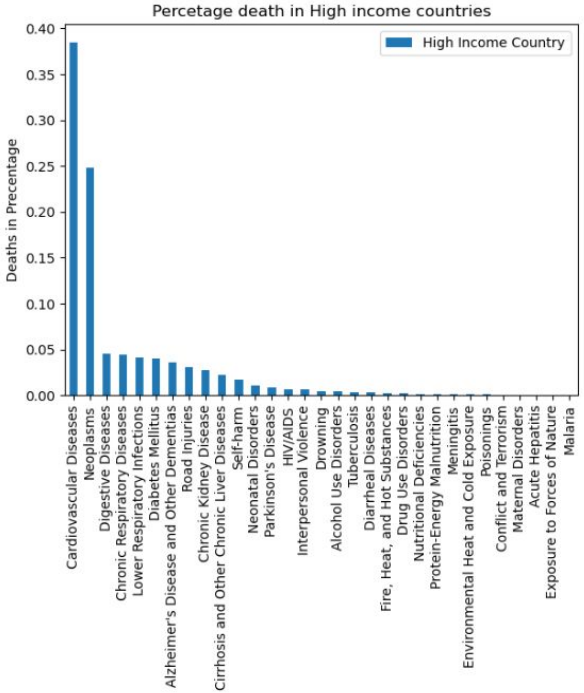


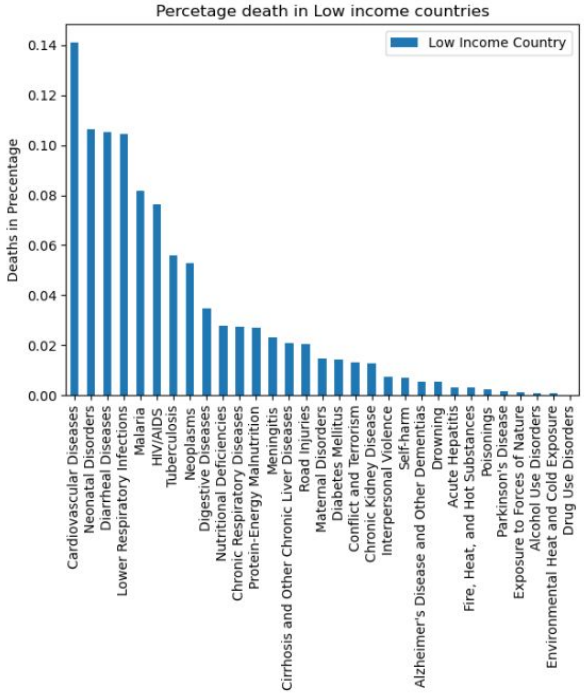


Observations and Conclusions

- Hvplot demonstrates heatmap of all the aid given over the 30 years observed. You can see the big clusters of aid given to the Middle East during this time period.
- Low Income countries, e.g. Afghanistan and Nigeria, didn't begin receiving material aid until the early 21st century.

Causes of Deaths: High Income Countries vs. Low Income Countries

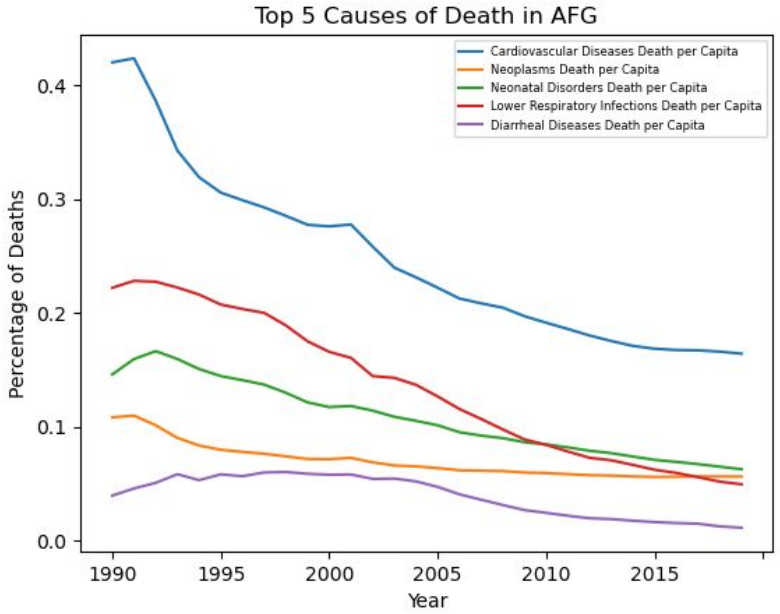




Observations and Conclusions

- Cardiovascular diseases and neoplasms are a worldwide problem.
- Deaths that can be attributable in large part to public health conditions (diarrheal, respiratory) are twice as high in low income countries than they are in high income countries.

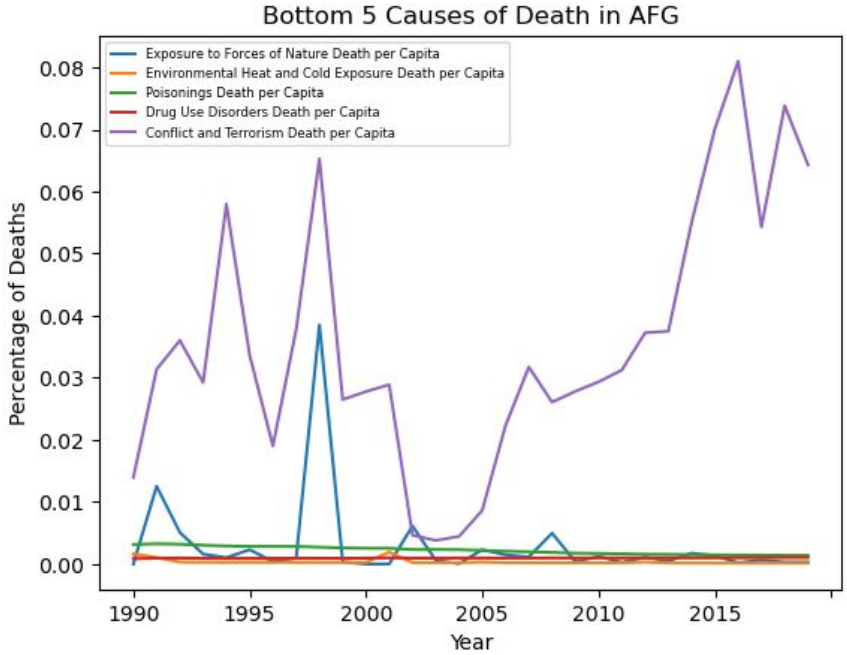
What were the top 5 causes of death in Afghanistan 1990 – 2019?



Observations and Conclusions

- The general trend: a decrease in the deaths per capita for each of the top 5 causes.
- Deaths per capita associated with cardiovascular diseases and respiratory infections decreased by 50%.

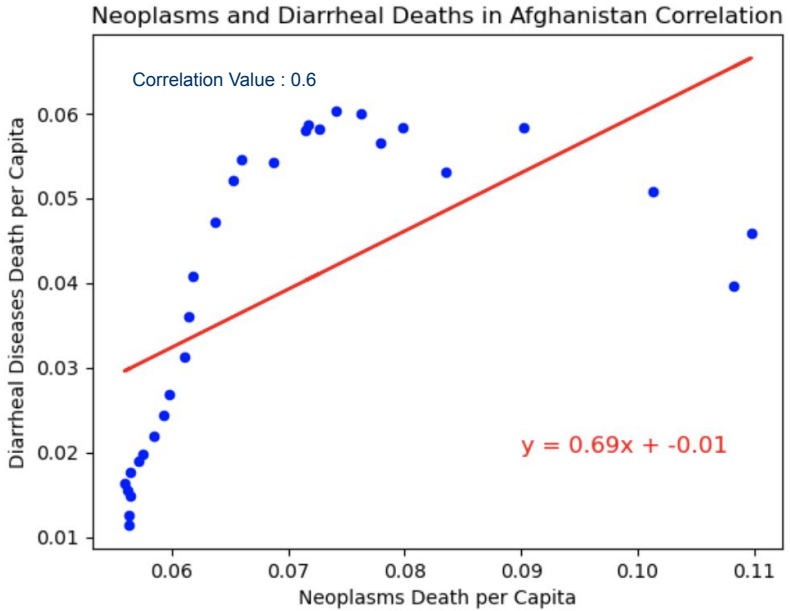
What were the bottom 5 causes of death in Afghanistan 1990 – 2019?



Observations and Conclusions

- Deaths arising from conflict and terrorism skyrocketed from the mid-90s to the late 90s. The graph suggests that during this period there was a high correlation between deaths caused by conflict and terrorism and deaths caused by exposure to forces of nature.
- Deaths arising from conflict and terrorism increased meteorically every year since the early 2000s.

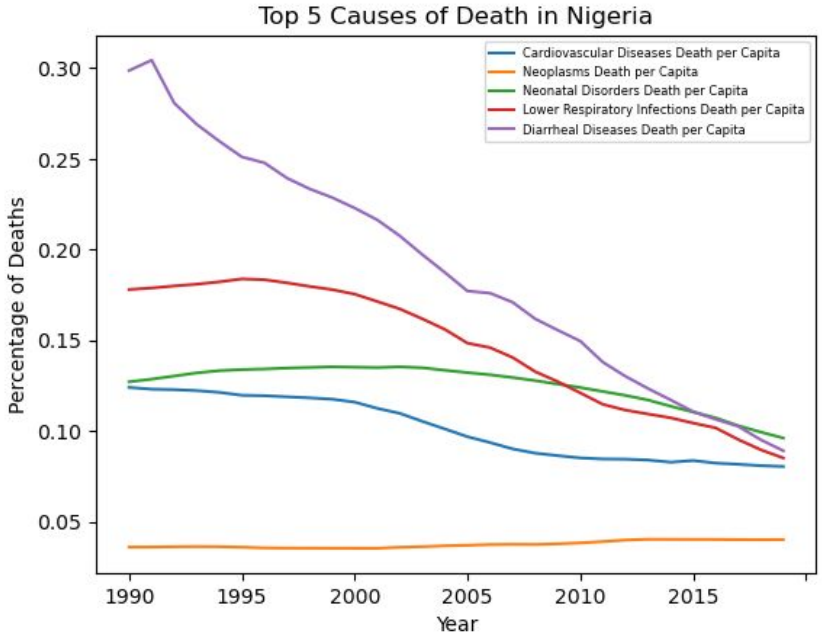
What is the correlation between the top two causes of death in Afghanistan (1990 - 2019)?



Observations and Conclusions

- There is mild correlation between the top two causes of death in Afghanistan, diarrheal diseases and neoplasms
- These two conditions are particularly virulent in countries with with a booming population, very young children, and inadequate/insufficient infrastructure

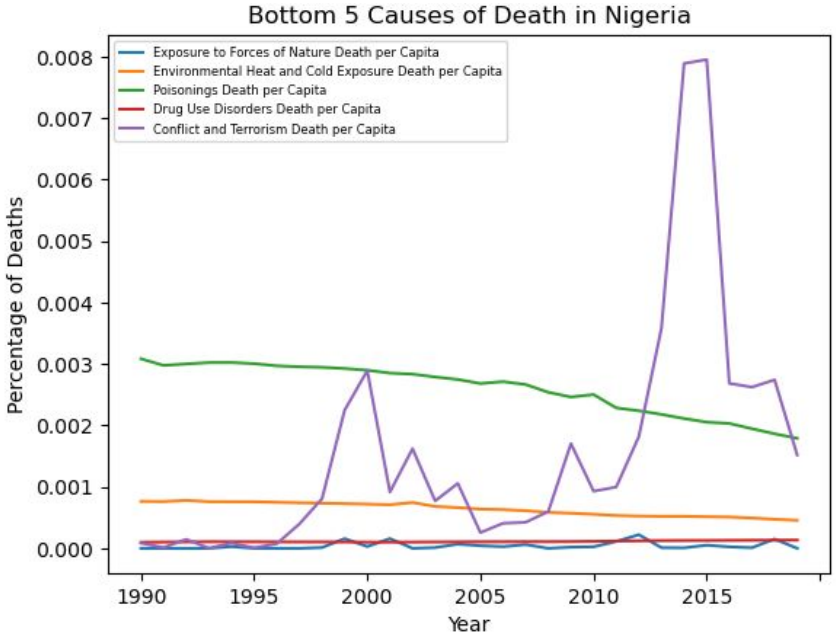
What were the top 5 causes of death in Nigeria 1990 – 2019?



Observations and Conclusions

- In Nigeria, the highest number of deaths resulted from diarrheal diseases.

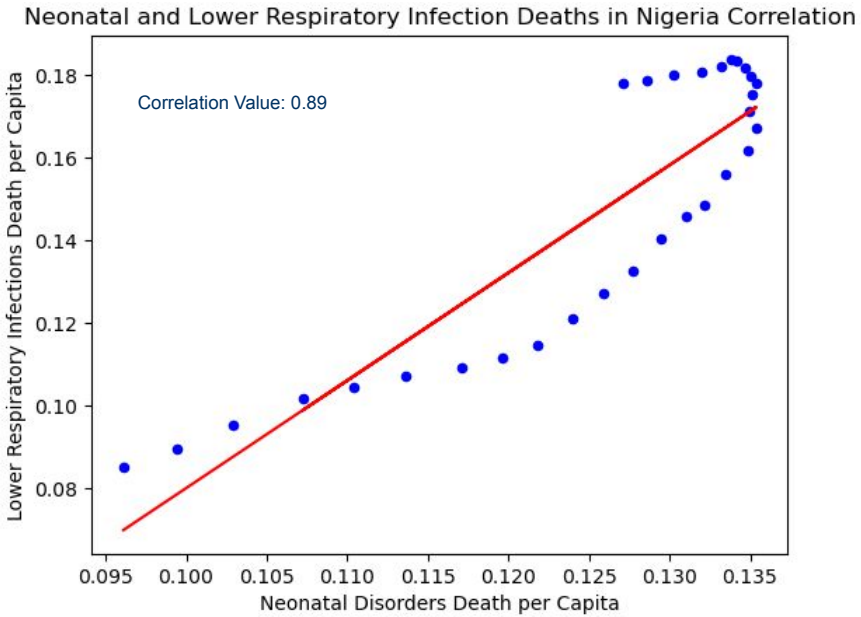
What were the bottom 5 causes of death in Nigeria 1990 – 2019?



Observations and Conclusions

- Conflict and terrorism have taken a substantial toll on the inhabitants of Nigeria.
- In relative terms, the number of deaths due to poisonings was 3x higher than deaths due to drug use disorders

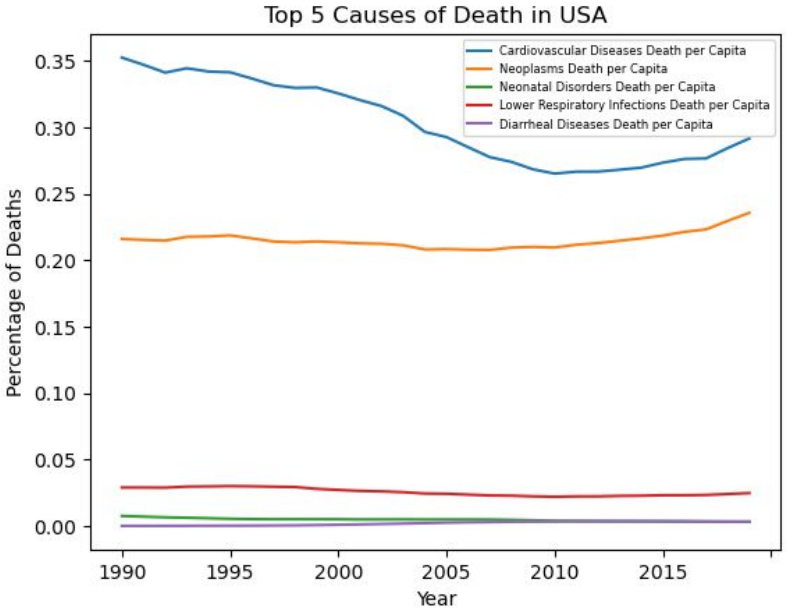
What is the correlation between two of the five top causes of death in Nigeria (1990 - 2019)?



Observations and Conclusions

- The top two causes of death in Nigeria, lower respiratory infections and neonatal deaths, were highly correlated
- The nature of these conditions and their apparent correlation suggests substantial inadequacy of public health resources, including sanitation, prenatal care, and malaria abatement.

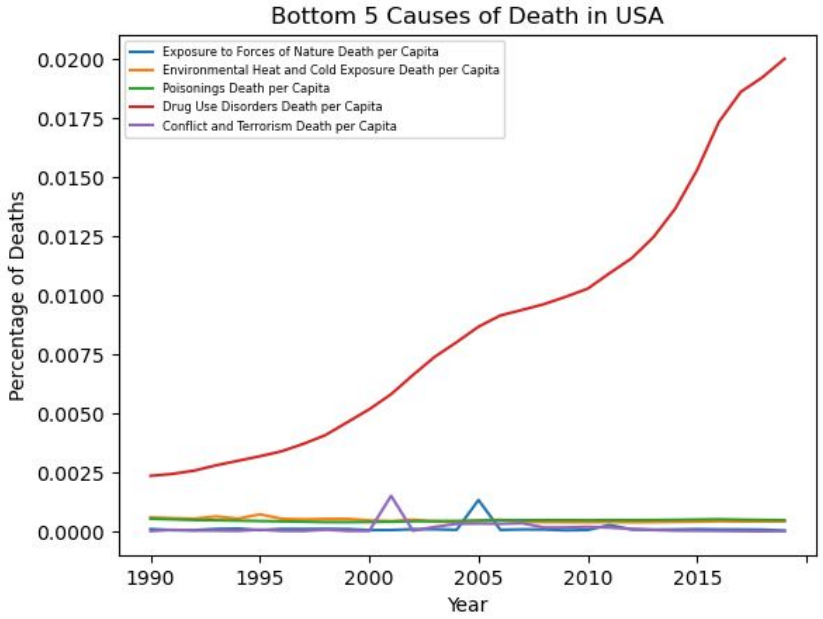
What were the top 5 causes of death in the U.S.A. 1990 – 2019?



Observations and Conclusions

- The US has very little Diarrheal Deaths this is likely to how accessible clean drinking water is to citizens

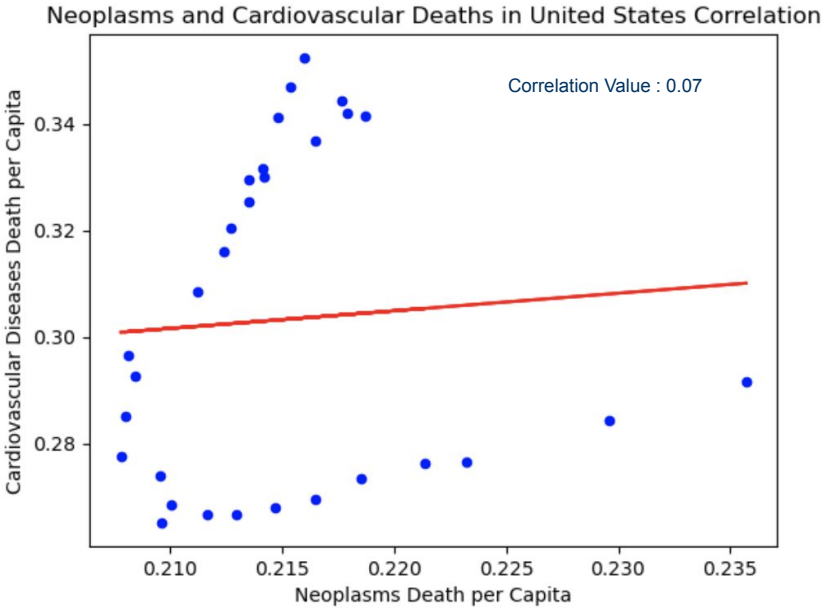
What were the bottom 5 causes of death in USA 1990 – 2019?



Observations and Conclusions

- Deaths resulting from drug overdose have increased alarmingly since 1990. Per the Centers for Disease Control, synthetic opioids (primarily fentanyl) is the leading cause.
- The plot shows two distinct spikes: the terrorist attacks on 9/11/2021, and Hurricane Katrina, 8/23/2005 - 08/31/2005.

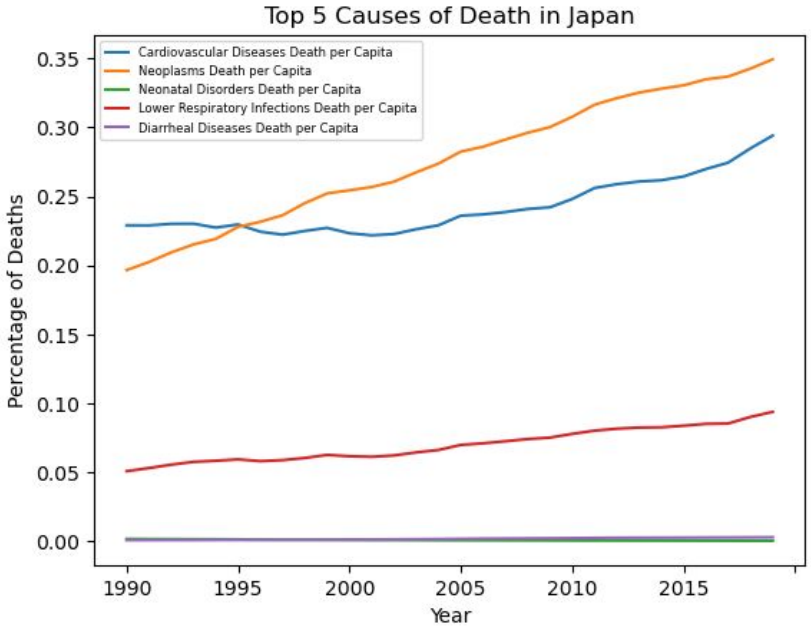
What is the correlation between the top two causes of death in the U.S.A. (1990 - 2019)?



Observations and Conclusions

- The top causes of death in the US were not correlated, suggesting they both may have different precursors.

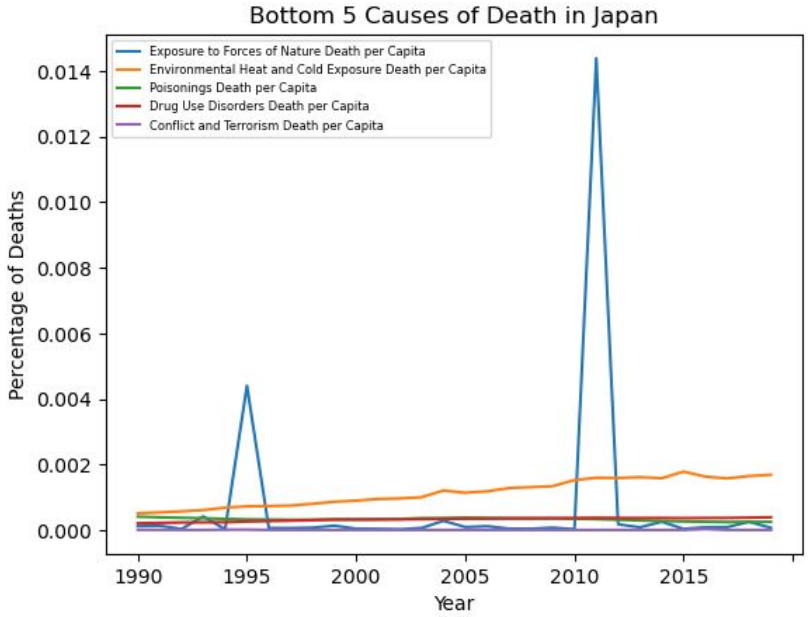
What were the top 5 causes of death in Japan 1990 – 2019?



Observations and Conclusions

- The top cause of death in Japan was Cardiovascular disease, seemingly just like the rest of other high income nations. Maybe high smokers? Societal norms?

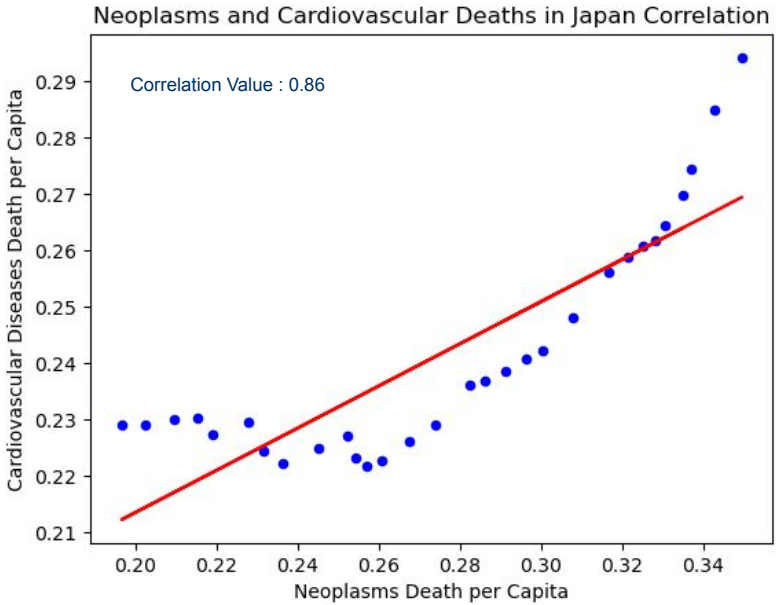
What were the bottom 5 causes of death in Japan 1990 – 2019?



Observations and Conclusions

- The most impactful of the bottom five was exposure to forces of nature, seeing a huge spike in 2011 when Japan experienced one of the worlds strongest earthquakes.

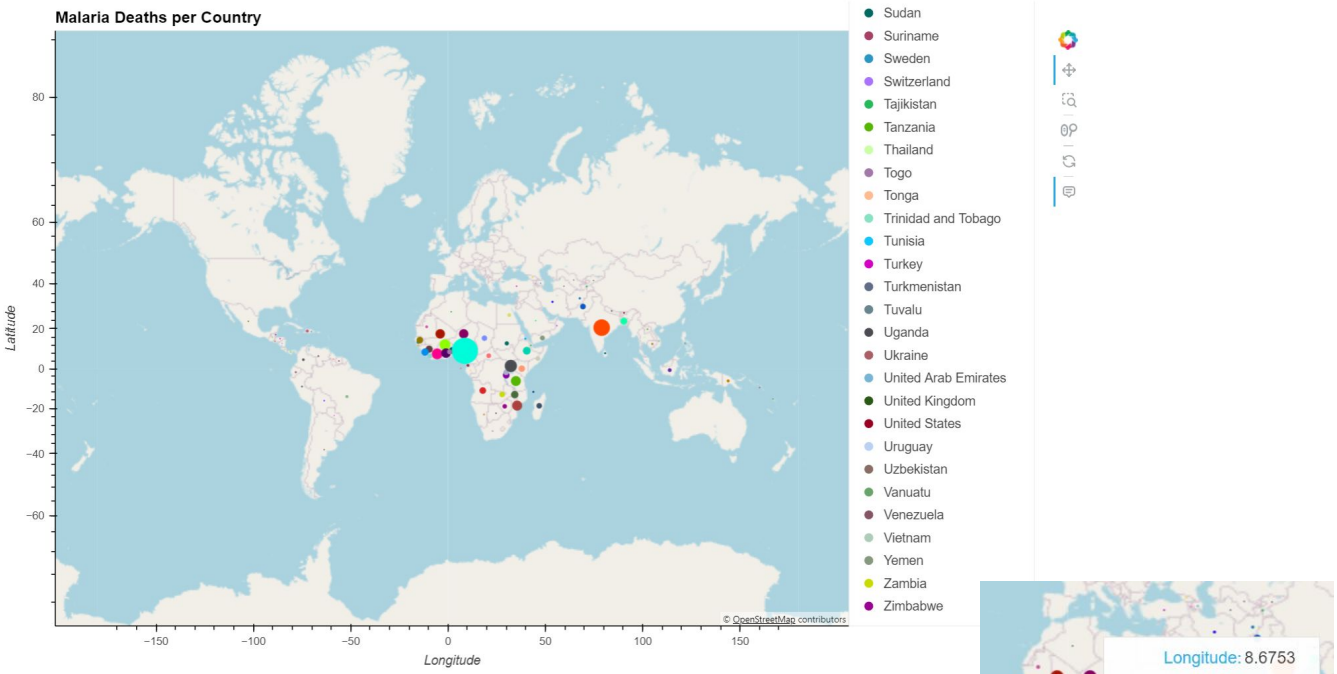
What is the correlation between the top two causes of death in Japan (1990 - 2019)

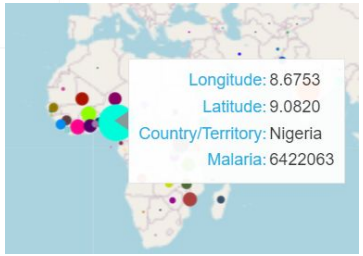


Observations and Conclusions

- Neoplasms and Cardiovascular deaths were the top two causes and they were highly correlated. They also had the same top two highest causes of death as the USA, but were much more highly correlated. Does this suggest things about health care in each country? Also, maybe we should entertain that both C.O.D.s have the same starting cause, unlike the US.

Which countries have the highest rates of death caused by malaria?





Observations and Conclusions

- Malaria is transmitted via mosquitos and thus malaria has higher prevalence in areas with mosquitos
- Lower respiratory infection deaths were one of the top causes of death in Nigeria (related to malaria)
- Countries bordering the Gulf of Guinea are categorized as “Low Income” or “Lower Middle Income” may have a lack of resources

Conclusions

Conclusions

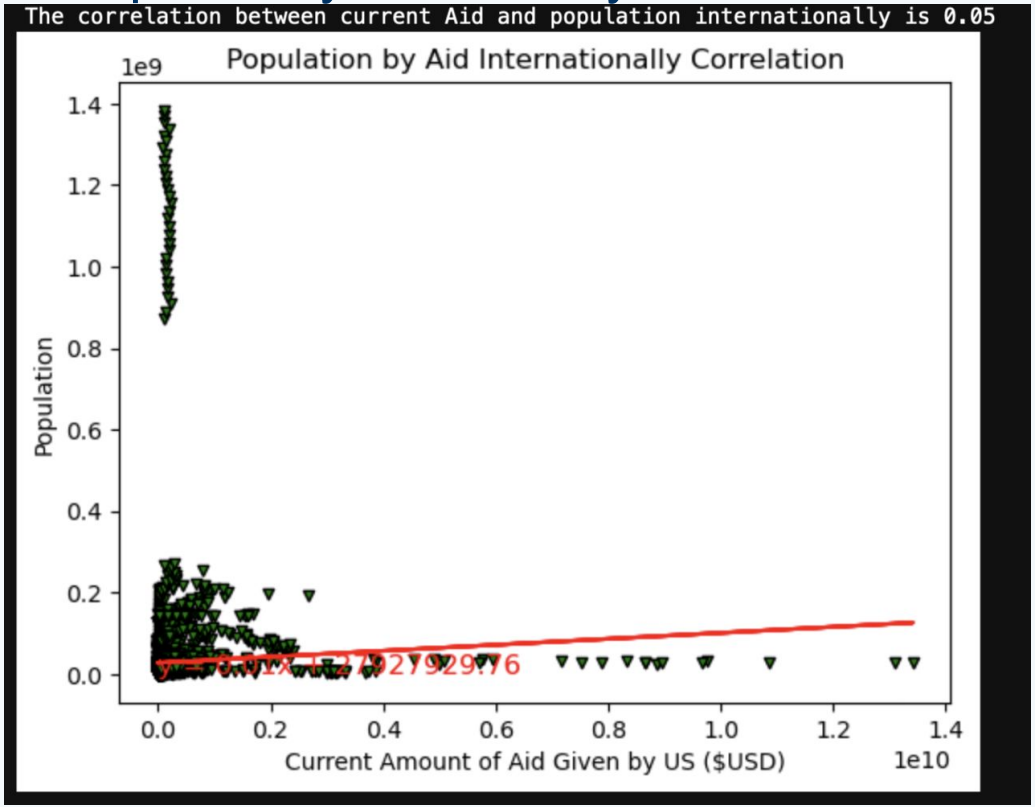
- Death arising from lifestyle risk factors rather than environmental factors are the leading causes of death irrespective of a country's income classification.
- Deaths arising from environmental risk factors (e.g., inadequate sanitation, safe drinking water, vector-borne disease transmission) are far greater in lower income countries than they are in high income countries.
- Lower/non-aid receiving countries the death rates are due to the lifestyle diseases like cardiovascular diseases, and neoplasm. But in low-income countries, even though foreign aid has a positive impact on the death rate public health-related deaths are pretty high compared to the high-income countries.

Challenges and Next Steps

- As pertaining to neonatal deaths in low income countries, further inquiry is warranted to determine the availability and efficacy of both prenatal care and antenatal care.
- It’s essential to include other countries that provide foreign aid and merge that data into our dataframe to analyze its effects on causes of death per country. We can generate different visualizations that show the countries offering foreign aid and how that affects the causes of death per country.
- It would be important to look at the rate a singular country is receiving aid across the years, while simultaneously the highest and lowest causes of death there, to make a tool used for predicting based on rates of Common Causes of Deaths, how much aid should be granted as a sufficient source to mitigate that C.O.D.

Appendix

Exploratory Data Analysis



The correlation between Current Aid and Population is 0.05

Y = 0.01x + 279279

Endnotes

1. For a discussion of the The World Bank's country income group determination:
   [https://datahelpdesk.worldbank.org/knowledgebase/articles/378833-how-are-the-income-group-thresholds-determined]
2. Diarrheal disease is the third leading cause of death in children under 5 years old. It is both preventable (potable water, sanitation, hygiene) and treatable. [https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease]
3. Poisonings (accidental): As of January 1, 2023, only 47% of World Health Organization Member States had a poisons center. The most notable gaps being in the African, Eastern Mediterranean and Western Pacific regions.
   [https://www.who.int/data/gho/data/themes/topics/indicator-groups/poison-control-and-unintentional-poisoning]
4. Cardiovascular diseases[1]: Known as "silent killers", these are a group of disorders of the heart and blood vessels. Behavioral risk factors are responsible for 80% of coronary heart disease and strokes. Poverty, lack of education, and unplanned urbanization can increase exposure to cardiovascular risk factors. [https://www.afro.who.int/node/5537]
5. Malaria: Caused by a parasite spread by mosquitoes. Children younger than 5 accounted for approximately 80% of all malaria deaths in the World Health Organization's African Region. [https://www.afro.who.int/node/5537]
6. Events in Afghanistan during the period 1988 - 1999
   a. 1988: Al Qaeda is formed
   b. 1989: Peace accords signed, Soviet military withdraws from Afghanistan
   c. 1992: The Mujahadeen storm the capital and form a largely Islamic state
   d. 1995: Newly formed militia, the Taliban, rises to power. Islamic law is enforced. The USA refuses to recognize their authority
   e. 1995-1999: Continuing drought devastates farmers, renders many rural areas uninhabitable. +1M Afghans flee and languish in refugee camps.
   [A historical timeline of Afghanistan, https://www.pbs.org/newshour/politics/asia-jan-june11-timeline-afghanistan]
7. In 2015 Nigeria saw suicide bombings and attacks by Boko Haram militants, in which thousands were killed.
8. Japan: On March 11, 2011, a magnitude 9.1 earthquake struck off the northeast coast of Honshu on the Japan Trench, resulting in over 18,000 dead. This earthquake was the largest in the world since 1900 [https://www.ncei.noaa.gov/news/day-2011-japan-earthquake-and-tsunam]

Source Data

- [cause_of_deaths.csv, Sourav Banerjee] https://www.kaggle.com/datasets/iamsouravbanerjee/cause-of-deaths-around-the-world, accessed March 13, 2024

- [us_foreign_aid_country.csv, U.S. Agency for International Development (USAID) and U.S. Department of State on behalf of United States Government agencies reporting foreign assistance.] https://www.foreignassistance.gov/data, accessed March 13, 2024

- [P_Data_Extract_From_World_Development_Indicators.zip, The World Bank Group] https://databank.worldbank.org/reports.aspx?source=2&series=SP.POP.TOTL&country= , accessed March 13, 2024

- [https://developers.google.com/public-data/docs/canonical/countries_csv, Google] accessed March 18, 2024