**Data Visualization**

**ISM 6419 Fall 2022**

Prof. Johannes Han Reichgelt, PhD

**Final Project Report**

**World Annual Deaths and their Causes**

By

Yogeshwar Reddy Kandati

U41071002

Information Systems and Management

**Introduction:**

The definitions of the terms "cause of death" and "risk factor for a premature death" should be understood. Each death has a unique cause according to the epidemiological framework used in the Global Burden of Disease study. Each death is linked to a single underlying cause, or, in their own words, "the reason that started the chain of events leading to death."

55% of the 55.4 million fatalities that occurred in the world in 2019 were due to the top 10 causes of death. The three major categories of cardiovascular (ischemic heart disease, stroke), respiratory (chronic obstructive pulmonary disease, lower respiratory infections), and neonatal conditions—which include birth asphyxia and birth trauma, neonatal sepsis and infections, and preterm birth complications—represent the top causes of death worldwide, in order of the total number of lives lost. The causes of death are classified into three types: communicable (infectious and parasitic disorders, as well as maternal, perinatal, and nutritional problems), non-communicable (chronic), and injuries.

Here are some of the queries this project will attempt to address:

1. Death rate due to various diseases in each country and comparing it to other countries with huge populations like China, India, and the USA to understand which is much more dangerous.
2. What is the percentage of GDP spent by each county on public health and the amount of external help received by each country?
3. Number of Physicians and average number of Physicians per 1000 of the population?

**Methodology:**

Three separate data sources, containing information from 1990 and 2019, were used to gather the data for this study.

<https://www.worldometers.info/>

<https://www.kaggle.com/>

<https://data.worldbank.org/>

for this project, I have not only collected the data from the above websites but also cleaned the data by using Excel and tableau prep builder since they have some missing values The real data values from the internet have been manually used to fill in any missing numbers in the GDP dataset.

The labels of each column in the Main data set are explained below. The data was normalized in Excel to enable more useful comparisons by transforming the raw data into death rates. If any country with more a death rate doesn’t mean it has more deaths because it should be compared to the country's population.

**Country name:** It contains the name of the country

**Country ID:** It contains the Code of the country

**Year:** It contains the range from 1990 – 2019

**WorldBankName:** The name of the country used by the World Bank

**HealthexppctGDP2016:** Level of current health expenditure expressed as a percentage of GDP.

**Healthexppublicpct2016:** Share of current health expenditures funded from domestic public sources for health.

**Physiciansper1000\_2009-18:** Physicians include generalists and specialist medical practitioners.

**Externalhealthexppct2016:** Share of current health expenditures funded from external sources.

The rest of the columns are represented with the name of each disease such as **Alzheimer’s disease, Parkinson’s disease, Interpersonal violence, Mental disorder**, and many more.

**Analysis:**

1. **Death rate due to various diseases in each country and comparing it to other countries with huge, populated countries like China, India, and the USA to understand which is much more dangerous.**

The below visualizations indicate the number of deaths caused by each disease all over the world over the years 2009-2019. Through this visualization, we can clearly define cardiovascular diseases as the main cause of more deaths to compare this with the Cause of death in the USA I have created a visualization by adding a parameter to know the top diseases that caused deaths in the USA. From the years 2009-2019.

Graphical user interface

Description automatically generated with low confidence

Now, I have created a visualization of the death rate in the USA by using a parameter for the cause of death so that here we can choose the top number of reasons for the deaths in the USA and compare them respectively.

Timeline

Description automatically generated

**Causes of death progression in the USA from 1990-2019.**

I took all the data on the cause of death in the USA from 1990 – 2019 to gradually see the change and how rapidly it has been increasing or decreasing for this here I used animation. I put the year in animation so that it shows the trials over the years.

Chart, line chart

Description automatically generated

**Comparison between China, India, and the USA**

Since China, India, and the USA are more populated countries comparing the Cause of death between them gives solutions to a few questions.

Table

Description automatically generated

1. What is the percentage of GDP spent by each county on public health and the amount of external help received by each country?

The below visualization shows the percentage of GDP spent by each country on public health here to get into deep research I used parameters for the top so that we can see the top number of countries and compare them.

Map

Description automatically generated

The below visualization represents the amount of help received from other countries, defined as external help. External sources compose of direct foreign transfers and foreign transfers distributed by the government encompassing all financial inflows into the national health system from outside the country. External sources either flow through the government scheme or are channeled through non-governmental organizations or other schemes.

Chart, bubble chart

Description automatically generated

1. Number of Physicians and average number of Physicians per 1000 of the population?

The visualizations below will address the above query. Here Physicians include generalist and specialist medical practitioners. Since there are many country regions and the visualization would get much more difficult I used parameters for the country region to grab the data of required countries.

Chart, bar chart

Description automatically generated

I have also created a visualization for countries with physician and nurse ratio from the year 2009 – 2018 here Nurses and midwives include professional nurses, professional midwives, auxiliary nurses, auxiliary midwives, enrolled nurses, enrolled midwives, and other associated personnel, such as dental nurses and primary care nurses.

Chart, scatter chart

Description automatically generated

**Dashboards**

Timeline

Description automatically generated

Chart

Description automatically generated

**Conclusion:**

In the Analysis section, we previously justified the study questions. Every nation is competing and making ongoing efforts to reduce the death rate. Many acts have been passed to control the death rates directly or indirectly. The world working and spreading awareness on obesity, healthy food, 5k runs, 10k runs, half marathons, and marathons these types of people gathering not only help us in fitness but also people do this to spread awareness on some sort of issues like water pollution, air pollution, soil pollution and many more which indirectly helps our lives to live long. There might be a few unexpected diseases like Covid-19 or unexpected reasons for death but the best way possible to overcome these is by keeping ourselves healthy.