

# **ANALYSIS OF WORLD'S LEADING RISK FACTORS FOR DEATHS**

## **STAT 515 - PROJECT REPORT**

### **GROUP-8**

#### **INTRODUCTION:**

The world is changing remarkably over the past few years. The one persistent thing is at the end we die. The health care over the last 20 years has changed drastically. Today we live longer than our ancestors have lived due to the advancements in the health care facilities. It also led to the increase in life expectancy, however the modern lifestyle, change in daily habits, improved infrastructure is also making us lethargic. As per the World Health Organization, approximately 60 – 85 percent of people around the globe lead inactive lifestyles that contribute to their health issues. The leading causes of death across the world still vary significantly, it is crucial to understand the causes that lead to death and the risk factors related to a precipitous death. The aim of this project is to analyze the data and find patterns in the deaths caused due to various risk factors. We believe this project would help the governments and organizations to focus on better policy making and targeted spending on health care.

#### **DATA SOURCE:** (Roser, 2019)

The dataset consists of all the countries in the world and the corresponding number of deaths caused by the selected 10 risk factors between the years 2000 and 2017.

Column [Population] is separately taken from the UN official website and integrated into the dataset.

#### **ABOUT THE DATASOURCE:**

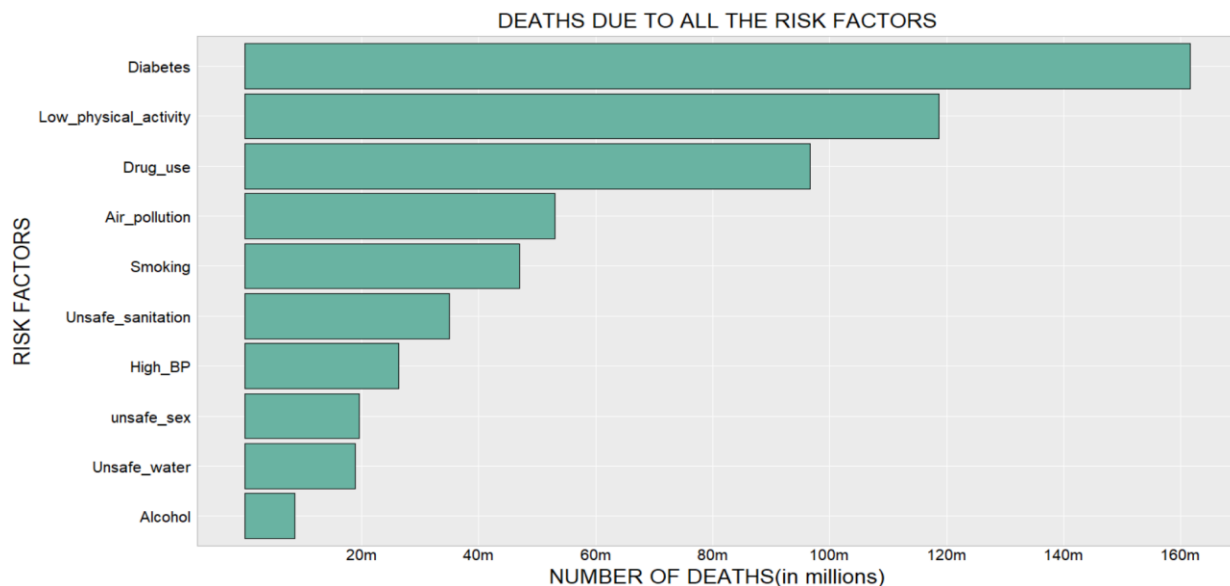
The website “Our World in Data” has the datasets on various topics that concern the world. The dataset “Number of deaths by risk factors” has the data regarding various factors that contribute to the deaths around the world over the past few years. Analyzing the data and observing the patterns can help us get deeper insights into the issues. The team at “Our World in Data” refers to the work of others which is available online as the team believes that there is enough data available which is important and useful that goes inaccessible. Our World in Data and the SDG-Tracker are outcomes of the collaborative efforts between researchers at the University of Oxford. (measuring progress towards SDG, 2018)

## DATA CLEANING:

- **Removed Unwanted columns:** The original data set had the data of deaths caused by 35 risk factors out of which we have picked 10 risk factors that are most common concerns in day-to-day life.
- **Renamed the columns:** The data set had long names with extra spaces. The spaces have been eliminated and the columns have been renamed.
- **Round off decimal values:** The number of deaths were in decimal. The values have been rounded off to the integer value.
- **Added extra population column:** The column population was not in the original data set. An extra column of the population has been added.
- **Derived data subsets:** Additional two data subsets have been derived from the original dataset.
  1. The data subset contains column: Year, Risk Factors, Dead\_B, values in column Dead\_B has been rounded off.
  2. Filtered the world data (from country column) and created a separate data set.

## VISUALISATIONS AND ANALYSIS:

To start with, we wanted to have an overall picture of the number of deaths caused in the world due to various risk factors. We wanted to analyze the most concerning factors and least concerning factors in the world. To facilitate that we have used Bar plot as we believe that this is the best fit to visualize the values.



**Figure1: Total number of deaths by risk factors**

Plot used: Bar Plot

The data taken is the sum of total number of deaths over the last 17 years. Diabetes is becoming one of the most crucial public health concerns of our time. The above plot shows that the number of affected people has increased significantly in the last two decades. Diabetes traditionally affecting adults and people associated with obesity and an inactive lifestyle, is now leading cause of deaths worldwide. The x-axis has number of deaths in millions, y axis has the risk factors. The bar plot shows that diabetes is the reason for maximum number of deaths (more than 160m) and alcohol for least number of deaths (8.5m). The bar plot was used as each bar represents value of the corresponding risk factor. When the bars are placed next to one another, it is feasible and easy to compare deaths due to one risk factor to another.

Moreover, after visualizing the overall deaths due to various risk factors, we wanted to observe the yearly trends for 17 years as well. To observe the trends, we have used interactive scatter plot as it is the best method to show trends and the range of data flow. With this plot, the maximum and minimum number of deaths can be determined for each risk factor over the period of 17 years.

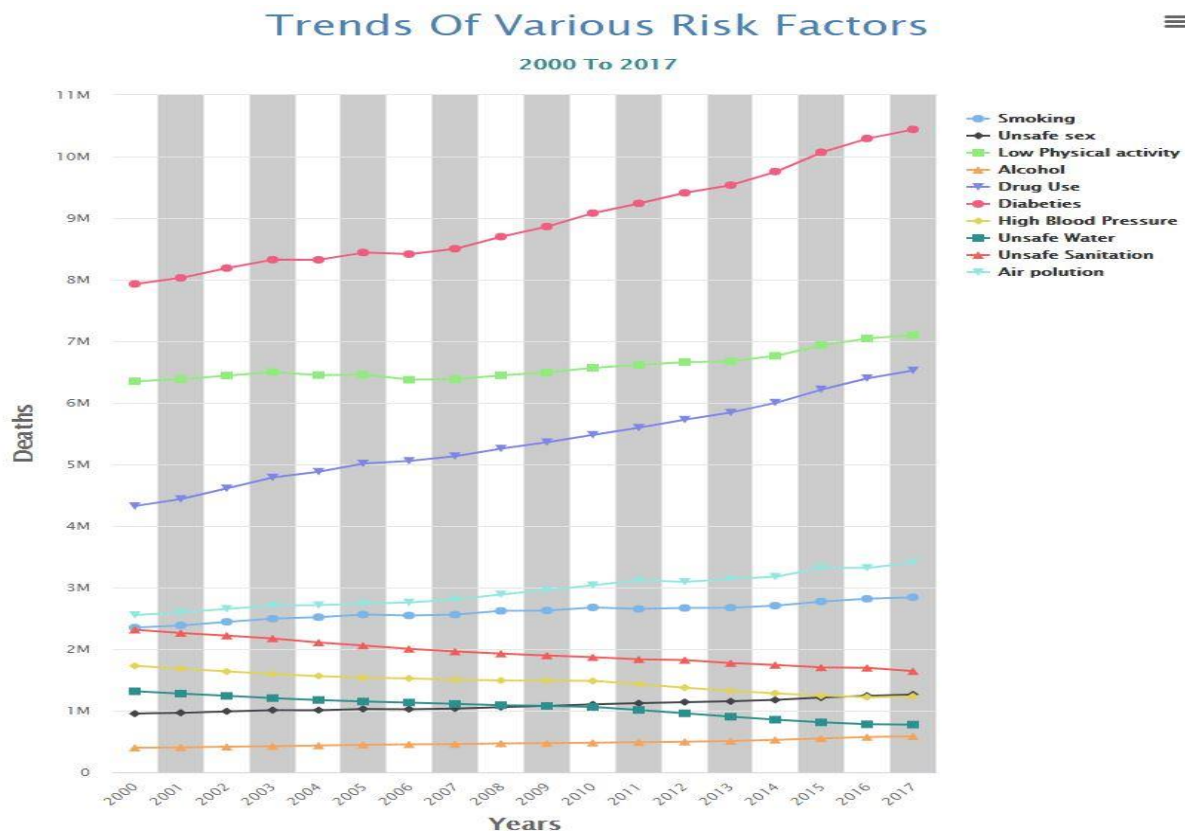


Figure2: Yearly death caused by different risk factors

Plot used: Scatter plot. (Deaths in millions)

The x-axis has year data, y-axis has population data, the plot shows evident positive trend in the number of deaths caused due to Diabetes, low physical activity, Drug use, Smoking, Unsafe sex, and Air pollution. This trend is concerning to observe as the deaths due to these factors are increasing constantly on yearly basis. Moreover, it can be observed that there is no significant increase in the death rate due to Alcohol over the period of 17 years. Meanwhile, negative trend can be seen in the number of deaths caused due to High blood pressure, unsafe sanitation, and unsafe water. Though there is negative trend in these factors, we wanted to analyze in depth if this trend applies to all the countries. To facilitate that we considered the categorization of countries based on Sustainable Development Index (SDI).

The Sustainable Development Index (SDI) measures the ecological capability of human development, realizing all the development must be achieved within global boundaries.

SDI is calculated based on two parameters: (United Nations, 2018)

- Human Development Index which is calculated by the combining life expectancy, education, and income index of every country.
- Ecological impact index which is calculated by measuring the consumption of CO2 emission/ resources and generation of waste.

The reason to consider below mentioned factors is that basic facilities such as good sanitation, clean air, safe water, and health factors (i.e., Diabetes and Physical activity as they are playing a major role in deaths) are essential for sustainability.

H-M = High-Medium, L-M = Low-Medium

### Trend of deaths in countries with different SDI

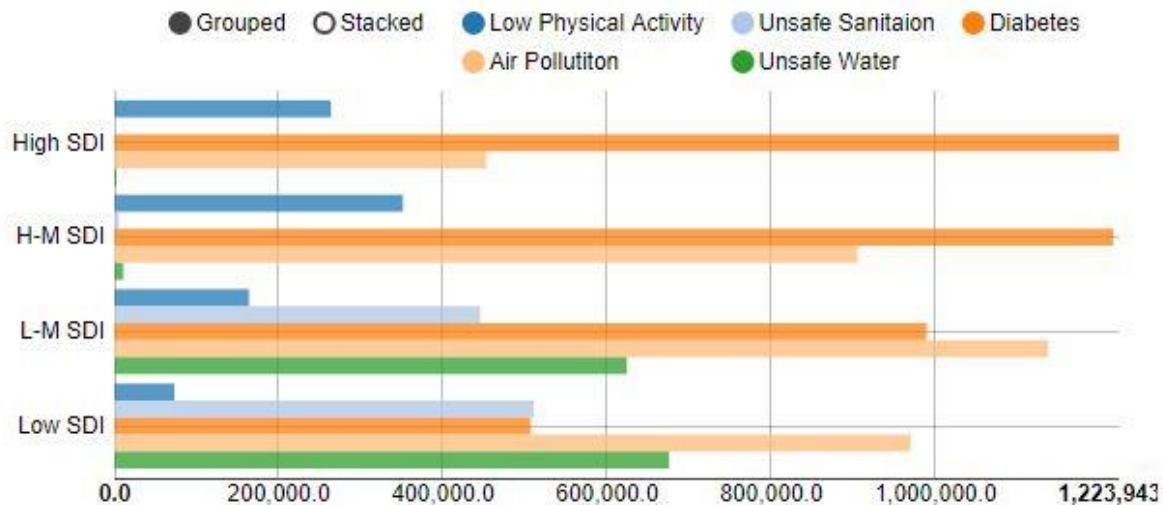


Figure 3: Trend of deaths in countries with different SDI

Plot used: Dodged Bar Plot

To check trends according to the Country's SDI we have taken a Dodged Bar Plot because it was difficult to compare different parameters for each country as there are 192 countries. We have categorized the countries based on their SDI as the countries with similar development will have similar SDI. It can be observed that in High SDI and High-Middle SDI countries, the deaths caused due to Diabetes, Air pollution and low physical activity are more. The deaths due to unsafe sanitation and unsafe water are negligible when compared to other 3 factors. It is good to observe this trend as the High SDI and High-Middle SDI countries are focusing more on hygiene, sanitation, and safe water. Meanwhile, an interesting observation can be made from Low-Middle SDI and Low SDI countries. Along with Diabetes and Air pollution, the factors such as unsafe sanitation and unsafe water are major reasons for deaths in Low-Middle SDI countries. It is concerning to notice that deaths due to unsafe water, unsafe sanitation and Air pollution are high when compared to other categories. This is an alarming situation because even in 21<sup>st</sup> century deaths are more due to lack of basic facilities.

In this alarming situation, we want to consider the death rates. We are visualizing all the factors with the help of a Box plot.



Figure 4: Perspective of the Risk Factors. (Deaths in millions)

By looking at this Boxplot, there is a slow positive trend in few factors such as Diabetes, Low physical activity, unsafe sanitation. There is high positive trend in Smoking, unsafe water, and high blood pressure. Rest of the other factors are growing at a normal rate. It is good to see that the rate of Diabetes, Low physical activity and unsafe sanitation is on a lower side. It is also concerning to see that factors such as Smoking, unsafe water, and high blood pressure are on a higher side. It says that the situation can be brought under control if the governments and health organizations work together in making better policies and increase spending on health and basic facilities. (We have made an interactive Boxplot, which explains the above-mentioned phenomenon. The Boxplot can be found in rmd file which is attached with this document)

## Conclusion:

1. It can be observed from Figure 1 that that the deaths due to Alcohol and Smoking are considerably low when compared to Diabetes, Drug use and Low physical activity.
2. Looking at the trends of death cases in Figure 2 for each of the factors, we can conclude that the deaths due to Diabetes, low physical activity, and Drug use are increasing at an alarming rate and the deaths due to unsafe sanitation, unsafe water, high blood pressure are decreasing gradually all over the world except in Low and Low-Medium SDI countries.

3. We can observe a trend in Figure 3 that the countries with high SDI suffer the most from diabetes and low physical activity. Deaths due to unsafe sanitation and unsafe water are negligible. People need to change the lethargic way of living and start living a healthy lifestyle. The governments of these countries can work on better policies to implement healthy way of living to decrease the death rates.
4. We can also see that, though the overall death cases are decreasing because of unsafe water and unsafe sanitation, the countries with low SDI and Low-Medium SDI are still suffering in this aspect. The death rates in these countries due to these factors are much higher when compared to countries with High SDI and High-Medium SDI. The governments of these countries should focus on improving the policies and provide good sanitation and water to decrease the deaths due to these factors.
5. It can be suggested that focused spending on Health and sanitation along with better policies by the respective governments can drastically bring down the death rates due to all the factors discussed above.

## REFERENCES

- measuring progress towards SDG*. (2018, 06 28). Retrieved from SDG Tracker: <https://sdg-tracker.org/>
- Novia, D. (2021, 03 09). *Data Novia*. Retrieved from Highcharter R Package essentials for easy interactive graphs: <https://www.datanovia.com/en/lessons/highchart-interactive-boxplot-in-r/>
- Roser, H. R. (2019, 12 01). *Causes of Death*. Retrieved from Our world in Data: <https://ourworldindata.org/causes-of-death>
- United Nations*. (2018, 06 01). Retrieved from United Nations: <https://www.un.org/en/>