

Analysis of Death Risk Factors

Iqra Bismi
Dept of Applied Data
Science, San Jose State
University, USA

Megha Gupta
Dept of Applied Data
Science, San Jose State
University, USA

Saniya Lande
Dept of Applied Data
Science, San Jose State
University, USA

Madhura Pandit
Dept of Applied Data
Science, San Jose State
University, USA

Abstract— Death is a part of the life cycle but throughout years many man-made factors have been observed causing millions of deaths. Statistics show that on an average there are over thousands of deaths caused due to poor lifestyle choices, environmental factors and other health issues. Although the maximum number of deaths have been observed due to high blood pressure, there are still many other factors widely responsible for mortality such as poor diet content, lack of physical activity, smoking etc. Further correlation between these factors can also be analyzed as each of these factors affects others.

The death rates dataset taken from Kaggle shows death rates by various factors like air pollution, child mortality, dietary and health issues, lifestyle choices. In-order to get some useful insights from this dataset,

I. INTRODUCTION

Pandemic has made us realize the importance of life and what steps should be taken to improve health and overall lifestyle. Covid-19 made us realize that there can be other crucial factors leading to higher death rates. It is important to know which are the leading causes for increased death rate as this will help us to understand what preventive measures can be taken such as improving the lifestyle and dietary habits thereby leading to decrease in death rate

These factors can be grouped into 4 main categories which are as follows:

- Pollution
- Dietary habits
- Lifestyle choices
- Child Mortality

By measuring the death rates of each cause, the concerned authorities can improve resources in that area. Mortality data can help improve essential resources among sectors such as food & agriculture, environment as well as health.

Also, it is important to analyze which continents and countries show the highest overall death rates to analyze how a country's industrial and economic growth is affecting the overall death rate.

we joined the development indicators dataset as well as the population dataset to correlate these death factors with GDP per capita and population of the respective countries. The analysis of this dataset is done through Tableau visualizations by creating various types of graphs, charts and maps to find insight trends and correlations between different parameters which are responsible for death. In addition to this, insightful results about the dataset have been provided for further prediction purposes.

Keywords — pollution, child mortality, dietary and health issues, lifestyle choices, visualization, Number of deaths, Tableau

II. OBJECTIVE

1. To determine which is the leading factors of death country wise
2. To compare different leading causes of death and to analyze which parameter leads to highest number of deaths throughout years
3. To study the effect on GDP, population and economic development of countries on mortality rate

Forecasting of the major parameters of death to study the trend over the years.

III. RELATED WORK

We referred to some research papers to understand the death rate statistics for various countries and the reasons leading to these factors

Pollution

Air pollution is one of the largest environmental problems. It affects the health adversely causing various diseases, specifically respiratory diseases. The effect of growing pollution directly impacts Life expectancy. The Environmental Protection Agency closely monitors the trend and gathers the data to know what are the factors and how it can be improvised. The purpose of this data visualization is to estimate the global death from air pollution [1]. We are considering 4 important types of air pollution for this review:

- *Outdoor Pollution:* It is a pollution mainly caused by combustion processes from motor vehicles, solid fuel burning and industry. The death rate has increased significantly over the years [4]. The main cause of this increase is total population.
- *Ozone Pollution:* It is the result of hydrocarbon and nitrogen oxide emissions. It has adverse impacts on human health like Lung function impairment, eye and nose irritation.
- *Matter Pollution:* Particulate matter is mixture of solid particles and liquid droplets present in the air. They can cause respiratory diseases and can enter bloodstream as well. Examples of particles which cause this pollution are sulfate, nitrates, ammonia, etc. Sources of particulate matter are construction sites, unpaved roads, fires and smokestacks
- *Household pollution:* It is caused by burning solid fuel sources – such as firewood, crop waste, and dung – for cooking and heating. Around 2.6 million people died due to the practices like cooking in open fires and inefficient stoves. It affects low and middle income countries [3]. These cooking practices are inefficient and affect mostly women and young children who spend most of their time at home.

Child Mortality

Research paper titled ‘Income and child mortality in developing countries: a systematic review and meta-analysis’ studied the relationship between national income and child and under-five mortality in a few developing countries. It was concluded that as the GDP of the country increases by 10%, the child mortality will decrease from 50 per 1000 births to 45 per 1000 births. Hence it can be stated that income is an important factor in determining the survival of a child.

Another research paper titled ‘Global, regional, and national causes of child mortality in 2008: a systematic analysis’ proposed that the country-wise estimate of the leading factors of child deaths should be used to reduce child-mortality by addressing the maternal, newborn and child health interventions.

Various factors leading to child mortality are as described below:

- *Child Stunting:* It means Low height for the child’s weight. It is caused due to poor nutrition, Lack of proper sanitation facilities, etc.
- *Child wasting:* It means low weight for the child’s height. It is observed in children that experience short periods of undernutrition that leads to the reduction in muscle and fat tissue.
- *Low birth weight:* New-born babies that weigh less than 2500 g are considered to be

underweight. It is caused due to the lack of oxygen and nutrients supplied to the baby when in the mother’s womb. Mothers that are malnourished, suffer from chronic kidney diseases, respiratory problems, addicted to drugs, alcohol, smoking give birth to low-weight babies.

- *Discontinued breastfeeding:* It means stopping breastfeeding to infants before the age of six months.
- *Non-exclusive breastfeeding:* Feeding babies less than 6 months of age with fluids and food other than breast milk, prescribed vitamins and minerals is non-exclusive breastfeeding.
- Discontinued and non-exclusive breastfeeding usually is observed in mothers that are malnourished and are less than 17 years or greater than 35 years of age. Infants not been fed exclusively face health problems such as acute respiratory diseases, diarrhea, etc.

Dietary Habits & Health Issues

❖ *Under Dietary Habits*

Research Paper titled “*Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study*” studied the causes of death due to dietary habits.

Following are the factors responsible for deaths as the optimal level of intake is not adequate:

- Diet-High in Sodium
- Diet Low in Fruits and Nut & Seeds
- Diet Low in Vegetables
- Diet Low in Wholegrain
- Vitamin- A Deficiency
- Iron Deficiency

The optimal level of intake for each factor is the level of risk exposure that minimizes the risk from all causes of death.

Table 1

Factors	Optimal Level of Intake
Diet-High in Sodium	3 g (1–5) per day
Diet Low in Fruits and Nut & Seeds	250 g (200–300) per day & 125 g (100–150) per day
Diet Low in Vegetables	360 g (290–430) per day
Diet Low in Wholegrain	125 g (100–150) per day

- *Under Health Issues*

Research paper titled “*Combined Effect of Blood Pressure and Total Cholesterol Levels on Long-Term Risks of Subtypes of Cardiovascular Death*” studied the causes of death due to Health issues.

Following are the factors responsible for death due to health issues:

Table 2

Factors	Leading Causes
High Blood Pressure & High Cholesterol	Due to diet high in salt, fat, and/or cholesterol; chronic conditions etc
High Plasma Glucose	Due to stress, lack of exercise, eating junk food
High BMI & Low Physical Activity	Due to diet, lack of exercise, stress
Low Bone Mineral Density	Due to smoking, alcohol, age, poor diet, medications

- High Plasma Glucose
- High BMI & Low Physical Activity
- Low Bone Mineral Density
- High Blood Pressure
- High Cholesterol

Lifestyle Choices:

We will be mainly considering deaths due to two of the following categories of Lifestyle choices.

- *Unsafe Sanitation*

Unsafe Sanitation and poor water-source can have far-reaching and ill-fated effects for those who are required to use unsanitary toilet facilities and drink, bathe, cook with and use contaminated water. In first world countries, we take clean water for granted. We can simply open the tap and have unlimited access to clean water. But that is not the case with rest of the world! Water-borne pathogens and bacterias are spread through contaminated water, causing diseases like diarrhoea, cholera, typhoid etc. Factors contributing Water pollution are - Industrial waste, Untreated Sewage and wastewater, Mining & marine dumping, Accidental oil and chemical leakages, Mixing of pesticides and fertilizer in groundwater due to rain, acid rainfall etc. Causes for water pollution and contamination are discussed in [1]. Benefits of improving sanitation and challenges faced in world are discussed along with key facts & Distribution of global deaths and its variations in low-income countries and middle, high-income countries are discussed in [3]. Categories, common types of sanitation and water contamination, the preventive measures to be taken are discussed further.

- *Addictions: Alcohol, Drugs and Smoking*

Addiction, also called substance use disorder, is a disease that affects a person's brain and behavior and leads to an inability to control the use of a legal or illegal drug or medication. Substances such as alcohol, marijuana and nicotine also are considered drugs. When you're addicted, you may continue using the drug despite the harm it causes. Alcohol and tobacco are among the top causes of preventable deaths in the United States. Moreover, these substances often are used together: Studies have found that people who smoke are much more likely to drink, and people who drink are much more likely to smoke. The effects of these addictions on our body is discussed in [6]. Almost every part of the human body is affected by these addictions. Diseases like Cardiovascular, Heart Diseases like stroke, various types of cancers, HIV/AIDS, Lung and Kidney diseases as well as Brain damage are some major ones. Almost six million people die from tobacco use and 2.5 million from harmful use of alcohol each year worldwide, the World Health Organization (WHO) reports. Overall statistics of impact of consumption of alcohol, drugs and smoking in the world are further discussed.

TABLEAU VISUALISATIONS

Research paper titled ‘Data visualization view with Tableau’ talks about various visualizations that can be used to create simple views on a dataset. Various visualizations like donut charts, geospatial maps, bar charts, heat-maps, treemaps, bubble charts, line charts, as well as some custom visualizations can be drawn as per the requirements. Interactive visualizations can also be built using filters, pages, drop-down lists, calculated field, sets, groups, etc.

After referring to these research papers, we have used appropriate data visualization techniques and tools to analyze leading parameters of death and also to identify any correlation and patterns between different leading parameters. Also, our analysis showcases the death trends with respect to geographical attributes like continents, sub-regions and countries.

III. DATASET ON DEATH RATES

Dataset plays an essential role in Visualization as the data will determine how the visualization will look like. For a small dataset, the results may not be able to define all the variables. So, it’s important to choose the dataset which has a large number of values to get meaningful results. We selected our dataset from kaggle, where five data files from 3 different data sources were used.

Files used:

1. Number of Deaths by risk factors - mentions about total death cases under various parameters
2. Population - mentions about population wrt countries and continents
3. Development Indicators - mentions about Years, Geographical information from Continents, Sub-region and countries attributes, GDP per capita
4. Death rates from air pollution - mentions about death rates caused to various categories under pollution
5. Death rate total air pollution- mentions about overall deaths due to pollution

IV. METHODS

• Data Wrangling

The dataset is a publicly accessible dataset available on Kaggle at

<https://www.kaggle.com/datasets/pavan9065/air-pollution> and is called “Air Pollution”. It consists of data in 4 different sheets with deaths due to air pollution and deaths due to other risk factors ‘Development Indicators’ was taken from the tableau official website whereas the data file of Population was taken from Kaggle.. After joining, our data consists of about 4.5k records covering 5 continents from years 1990 to 2014.

ER diagram of parameters from each files are shown below:

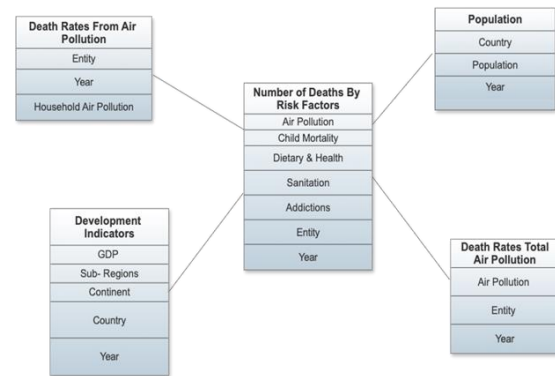


Fig. 1. ER Diagram

number-of-deaths-by-risk-factor.csv is made of 6 tables. ①



Fig. 2. Column Names in Dataset

Five datasets were combined using tableau join. Two of them were joined using left join whereas other three were joined by inner join.

V. VISUALIZATION DESIGN

Analysis of the number of deaths by risk factors is complicated as well as sensitive. To address this issue with all possible angles, we decided to divide this project into several parts: from time, geographical distribution, and environmental perspectives.

On one particular time scale. For example, we wanted to look at the number of Deaths distributed in different continents or countries and compare it with GDP per capita of respective countries. We selected parameters causing the highest number of deaths in all categories like Child mortality, Pollution etc. We put a filter to select each parameter as needed and year to be selected from 1990 to 2014. GDP per capita is grouped to get a better distribution of low-income, middle and higher-income countries. Size of dots in countries represents the number of deaths. Bigger the dot, higher is the number of deaths caused by selected parameters in that country. For this dual map one graph showed the GDP per capita with varying colors whereas another map showed the death cases with varying size. Finally both the maps were overlapped using the dual axis.

We decided to embed some graphs in the tooltip of other graphs for better visualization with more information. For example, the graph of Child mortality rate per continent is showing distribution within each continent per subregion. We created a hierarchy of sub-region and country and respective child mortality rate per 1000 births as years increased. By embedding this table into a graph of Child mortality for continents, we get the whole picture of distribution of child mortality rate per country in selected continents and subregions under it.

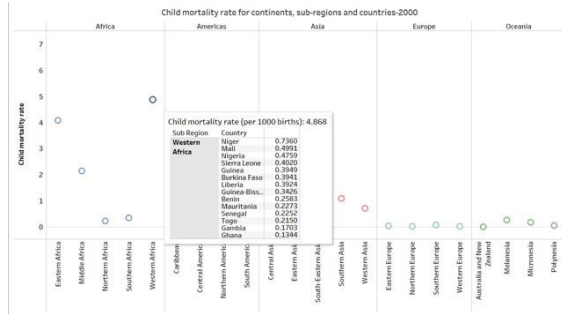


Fig. 3. Child Mortality rate for continents

VI. TABLEAU OPERATION

The main objective for this data visualization project was to show interesting insights obtained from the death rate dataset with the help of interactive dashboards. We made four dashboards, one dashboard summarizing the statistics for each of the death parameters.

- Death Analysis wrt Pollution

Bar graph was used to measure the categorical values of various factors leading to total air pollution. We used year in pages for making the dashboard interactive.

We created a heat map to show the total deaths caused by Air pollution. This chart was preferred as the size of the continents with highest deaths can be easily visualized.

Dual combination graph was used to show and compare the trends of household pollution with outdoor pollution. In the graph, bars are representing the deaths by household pollution, whereas on the other hand lines are representing the deaths caused by outdoor pollution. We created the line graph comparing the death caused by air pollution specifically in 3 countries i.e. India, China and Finland to analyse and compare the trends of the deaths with respect to time. We have used filters for country attribute for user interactivity.

- Death Analysis wrt to Nutritional Deficiency and Health Conditions:

To analyze death cases due to dietary and health issues, a geospatial map was created which is showing death cases in two major parameters along with GDP per capita for all countries. We created a parameter consisting of total death cases due to dietary habits and health issues which the user can modify as per the preference and the values were represented by circles. In which higher values had higher diameter. Also, the GDP was shown with varying colors, and the countries with high GDP and death cases were highlighted using annotation. In addition to this, factors under two major categories were embedded in the tooltip. This map helped us to gain insights about the death cases wrt to developed and developing countries.

Fig. 4.Parameter

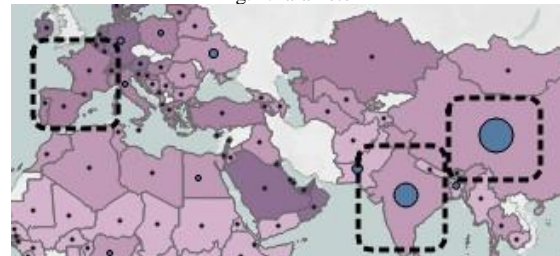


Fig. 5. map

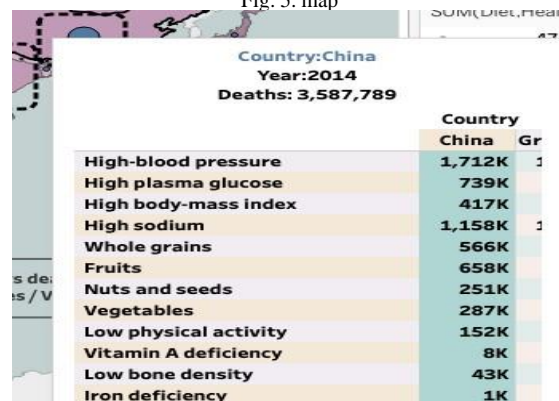


Fig. 6. Parameter

Next to show visualization wrt to continents, two parameters were created consisting of factors under two major categories such as dietary habits

parameter consisting of diet high in sodium, diet low in fruits and vegetables etc. Same pattern was followed for health condition parameter as well. Furthermore, for better comparison and analysis, bar plot was plotted along with two line graphs. Bar plot showed the cumulative death cases under two major categories which users can change. Likewise for line graphs, factors under two categories were plotted which users can modify as per requirement. In addition to this, a pie chart was embedded in the tooltip to drill down further in the sub-regions wrt. to continents.

Fig. 7. Parameter (Diet Content)

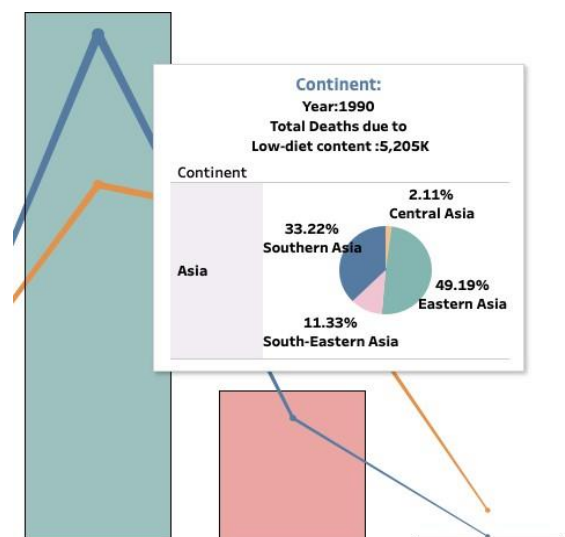


Fig. 8. Graph

To do in depth analysis, between various factors, again a bar plot was plotted showing comparison between top factors under health issues such as high B.P, High cholesterol and High BMI. Also, an embedded mountain graph showing death cases

along with population was plotted wrt to sub-regions. This graph helped us to gain insights about death cases wrt to population of the continents/sub-regions.

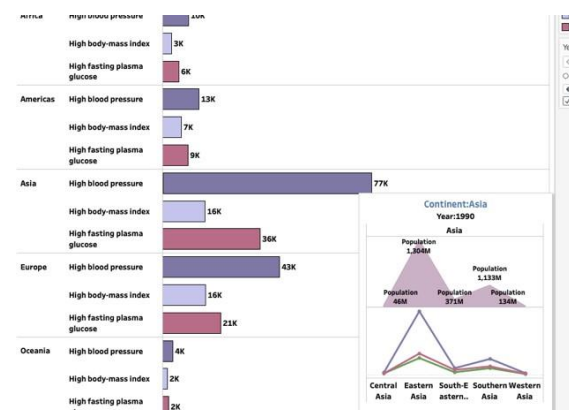


Fig. 9. Graph

Lastly, to see county-wise death cases due to diet high in sodium a pie chart was created showing total death cases. A table was embedded in the tooltip showing total death cases due to diet high in sodium with respect to countries. With this visualization we were able to find out the countries with highest death cases due to diet high in sodium.

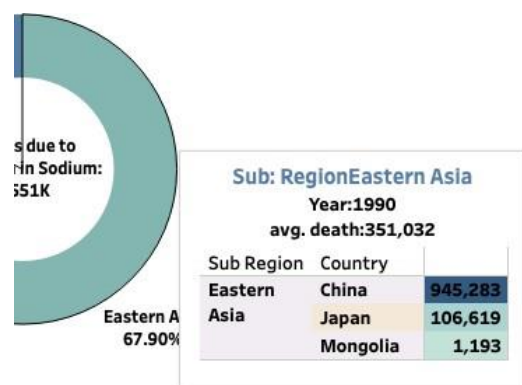


Fig. 10. Graph

Overall, in all the graphs, year was embedded as pages so that the user can change the year as per the preference/ requirement. Also, the overall trend can be seen by using the play button in the year mark card.

Fig. 11. Graph

- Analysis of Child mortality:

We used a bar chart for displaying various factors leading to child mortality as the bar chart explains categorical data clearly. We filtered the null values present in the continent by adding it into the filters mark card. We also displayed the continents as well as countries using drop down options for user interaction. We added years in the pages to make it more interactive as the years increased.

We created a calculated field to show the total child mortality by adding the deaths due to various factors like child stunting, child wasting, etc . We created a set of countries to show Top N countries based on the total child mortality and used a bubble chart for representing the same as this chart will show the sizes of the bubbles according to the level of total child mortality. Easier to understand. Figure below shows a set created for Top N countries showing child mortality. We also embedded the heat map into the bubble chart in the tooltip to show the population for each country.

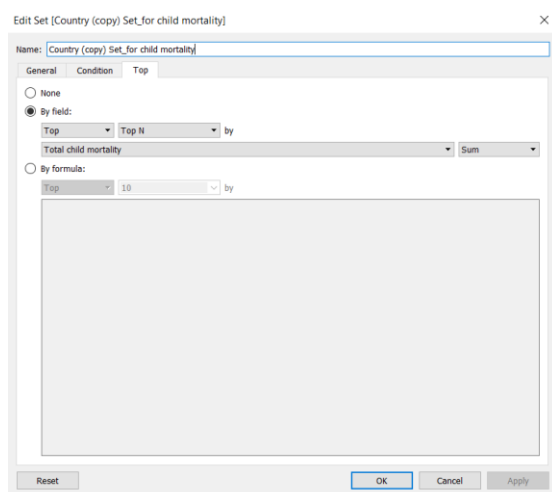


Fig. 12 Set (Child Mortality)

We created a calculated field to show the child mortality rate (per 1000 births) for each country by dividing the total child mortality with the population of the respective countries and dividing by 1000. We then plotted a circle view showing child mortality rate wrt continents and subregions. Circle views was used as it summarizes huge amounts of data into one plot easily. We embedded the hierarchy chart into the tool tip of circle views to show countries for a particular continent and sub-region. Parameters were created for allowing the users to select the continent and sub-region of their interest.

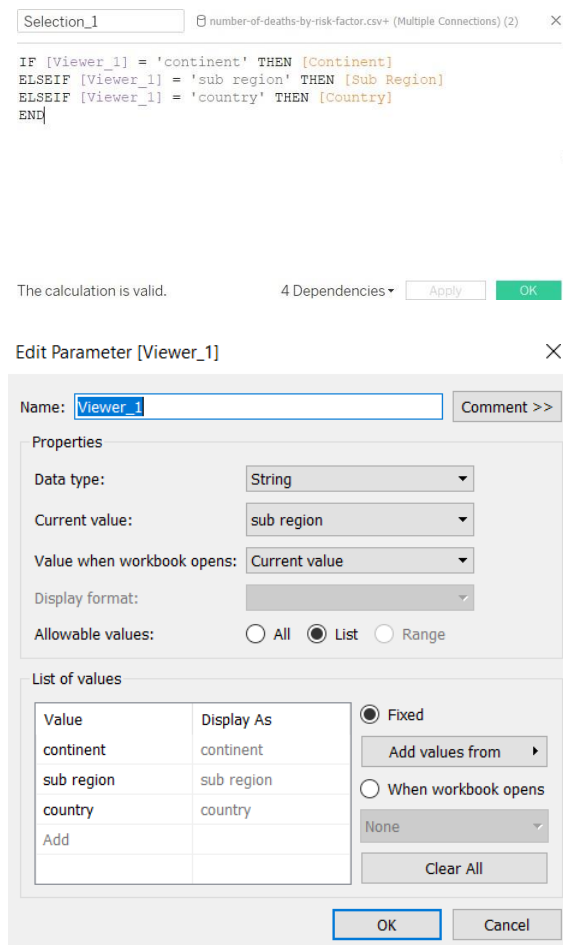


Fig. 13. Parameter

Donut chart was used to show the distribution of child mortality rate wrt continents as it is useful to show proportion of the categorical data with respect to the whole. We embedded a heat map showing GDP per capita wrt continents into the tool tip of the donut chart.

Line chart was used to show the trend for GDP per capita and child mortality rate as this visualization is best used for showing time series data. Annotations and labels were used to label the maximum and minimum values for GDP and child mortality rate.

Analysis of Deaths due to Lifestyle choices:

The dashboard for Lifestyle choices, mainly deaths due to Addiction issues and Unsafe sanitation is making use of some new types of charts like Butterfly chart, Word cloud and Sunburst chart. For example, In butterfly chart, we created a calculated field for all addictions (addition of alcohol, drugs & smoking) and compared it with deaths due to low physical activity, to understand the correlation with these two parameters. By selecting top 10 countries with highest deaths in these parameters, this graph gives a good visual idea about the pattern. Like

butterfly wings, we can observe that the pattern or ratio with the number of deaths in top 10 countries is almost similar due to two parameters. Similarly, word cloud is a visualization type where words are given weightage as per parameter value. For example, word clouds for total deaths due to sanitation issues have names of subregions. And higher the number of deaths bigger is the size of that subregion name. Also we embedded the bubble chart in this word cloud, i.e. as we hover over any subregion, the respective countries as shown in the bubble chart showing total deaths & countries with maximum number of deaths in it.



Fig. 14. Regions with most Death due to Sanitation issues

Another new type of chart used here is Sunburst chart. It is a type of pie or donut chart which shows distribution of deaths per continent in inner disc, and distribution of deaths per subregions in outer disc. Calculated field dummy-zero is created for this which has basically value 0 and it is used to get empty circles or rings between discs. Value of parameter to be shown, total addition deaths in this case is used in angle to create angles based of values per continent and subregions in it.

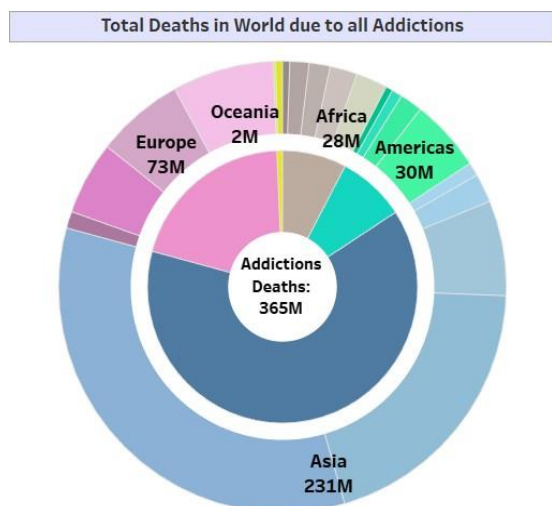


Fig. 15. Total Deaths in world due to all Addictions

VII. DATA ANALYSIS

Death by Pollution Parameters: In the stacked bar plot shown in fig 16, we are comparing deaths caused by Household Pollution, Matter Pollution and Ozone Pollution across the continents and over the years. As we can see clearly from the Graph, Deaths by Household pollution is more comparatively to other two pollution parameters i.e. Ozone and Matter Pollution.

Household pollution: From 1990 to 2014, the deaths have gone down in each continent. Deaths have gone down by 40%, i.e. 1.6% drop annually. Maximum decrease in number of deaths is for Europe Continent which is 69% and least decrease in deaths is for Oceania continent i.e. for 23%.

Matter Pollution: From 1990 to 2014, deaths have dropped by 33.3% for the European continent. No change was observed in the number of deaths in the African continent.

Ozone pollution: Deaths in ozone pollution have also decreased from 14 to 10 which is a 28% drop in the number of deaths.

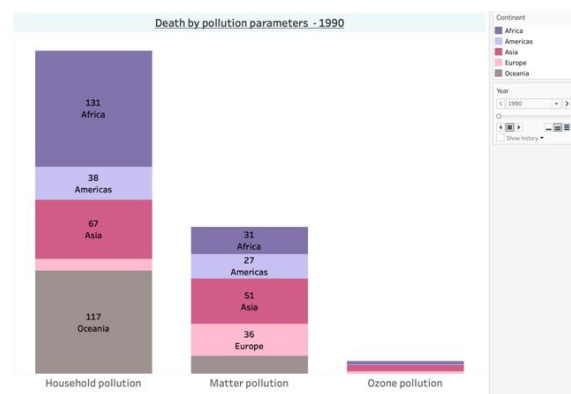


Fig. 16. Deaths by Pollution Parameters

Heat map of Total Death by Air Pollution: We are measuring total deaths across continents and it is clearly visible that from 1990 to 2014 deaths due to outdoor air pollution has continuously gone up. Asia continent remained the leading continent with the maximum number of deaths due to outdoor air pollution. Death have increased by 71.75% in Asia continent from 1990 to 2014 i.e. 2.8 % increase annually.

Deaths in Oceania continent had remained 7K which also shows the continent with least increase in number of deaths.

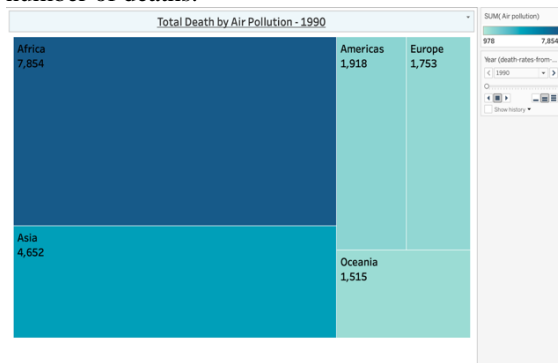


Fig. 17. Total Death by Air Pollution

In fig 18, we are comparing the household pollution with Outdoor pollution. Primitive techniques used in household cooking and handling have changed significantly over the years with new innovations. Deaths caused by these pollution have gone down by 49%. On the other hand, outdoor pollution increased significantly (58.3%) as a result of industrialization and transition of countries from low income to middle income, more use of resources.

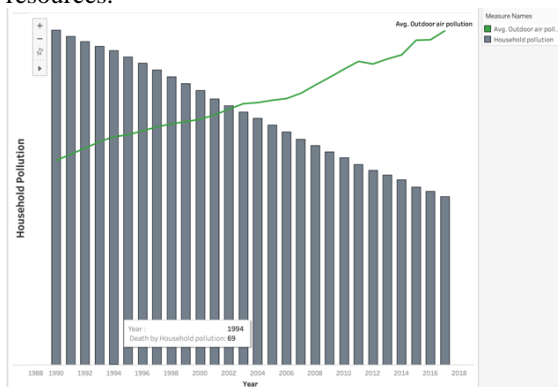


Fig. 18. Death by household vs Outdoor Pollution

In Fig 19, we are comparing the death rates in 3 countries i.e India, China and Finland. In countries like China and India, the population is high as well as the death rates. From 1990 to 2017, deaths in China from Air Pollution have gone up from 618K to 1030K . It's a 66% increase in the number of deaths or 2.38% increase annually. Deaths in India from Air Pollution have gone up as well from 368K to 819K i.e. 4.47% increase annually or 122% increase in deaths due to Air Pollution from 1990 to 2017. The reason for this much increase is excessive use of coal and vehicles. Deaths in Finland from Air Pollution have gone down as well from 2K to 1K. Countries like Finland which Identified the potential issues with Air pollution early on added measures to reduce the effects of Air pollution.

It has greatly benefited the country since the deaths due to air pollution has gone down. It's a 50%

reduction in deaths due to air pollution or 1.78% reduction in deaths annually as a successful result of the government. strong environmental regulations.

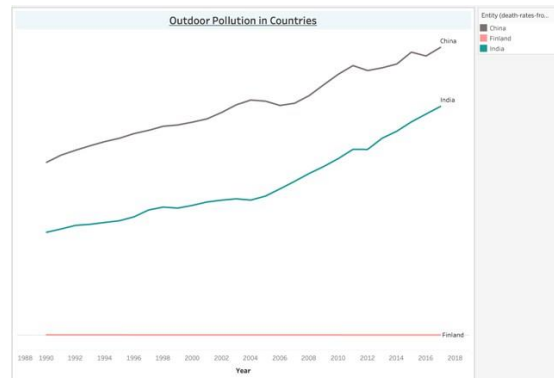


Fig. 19. Death by Outdoor Pollution in countries

Child Mortality:

It can be seen from the fig 19 that deaths due to child wasting was the highest as compared to other factors for the year 2014.

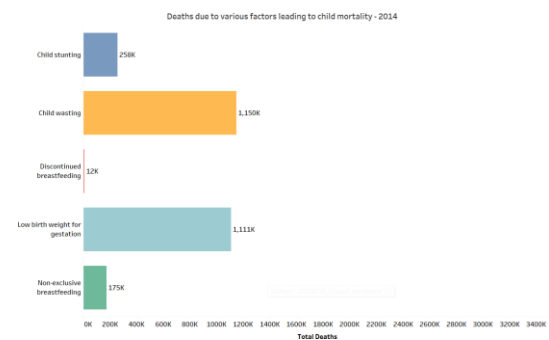


Fig. 19. Deaths due to various factors

Fig. 20 shows the Top N countries with highest child mortality for the year 2014. It can be seen that India had the highest child mortality of 699K. Such high value may be due to the higher population of India which is around 1296M.

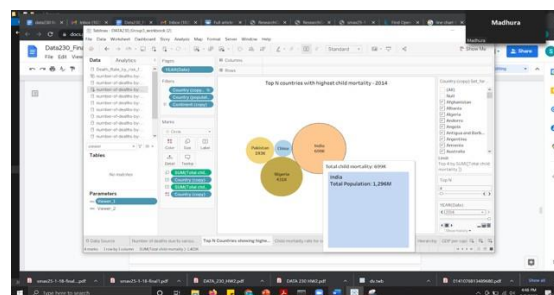


Fig. 20. Top N countries with highest Child Mortality

It can be seen from the Fig 21 that Western Africa sub-region of the African continent had the highest child mortality rate (per 1000 births) which is around 2.079. Further, we can also see that Niger from Western African sub-region contributed to the

highest child mortality rate whereas it was least for Ghana in Western African sub-region.

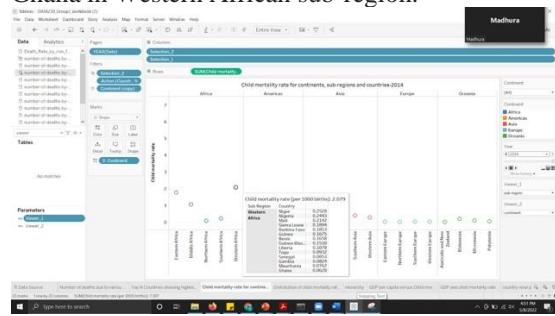


Fig. 21. Child mortality rate

Donut chart below shows that the highest child mortality rate was in the African continent (around 73%) whereas the least child mortality rate was observed in the European continent.

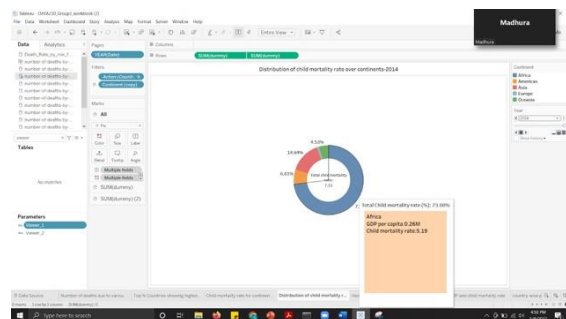


Fig. 22. Distribution of child Mortality for continents

We can see a decreasing trend in the child mortality rate (from 24.452 in 1990 to 7.107 in 2014) as the GDP per capita increases (from around 2M in 1990 to around 3 M in 2014)

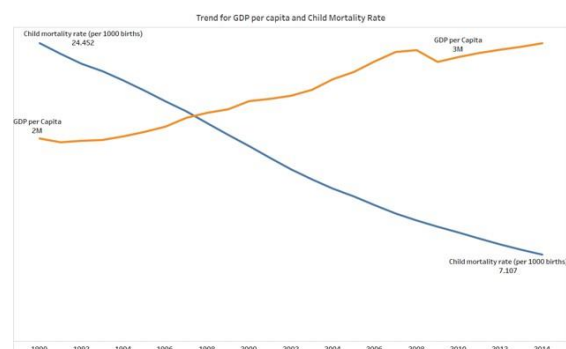


Fig. 23. Trends for GDP per capita and child Mortality

Death Analysis wrt to Nutritional Deficiency and Health Conditions:

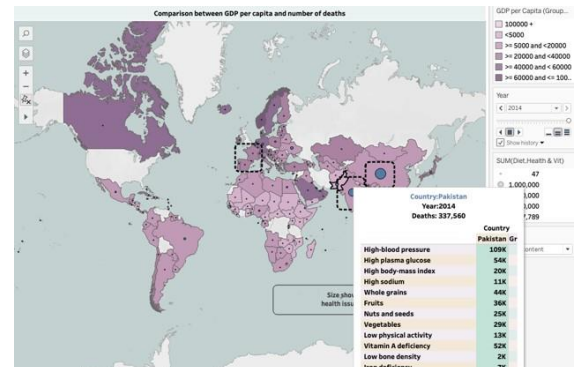


Fig. 24. Comparison between GDP per capita and number of deaths

From the geospatial map above (fig 24), we can see that the circle size is larger for India and China for year- 2014 ie. The death cases due to diet issues were high in India and China. Also, in India, the highest number of death cases were due to high Blood Pressure. Same trend was observed for health issues as well.

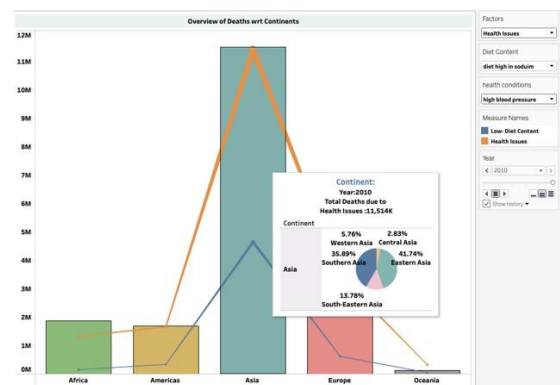


Fig. 25. Overview of Deaths wrt continents

Bar and line graph above (fig 25) shows that Asia had the highest number of death cases for health issues as well as for low diet content. Also, In Asia more deaths were due to high blood pressure as compared to death due to the diet high in sodium. Furthermore, the embedded pie chart shows that eastern Asia had the highest number of death cases (around 41%) for 2010.

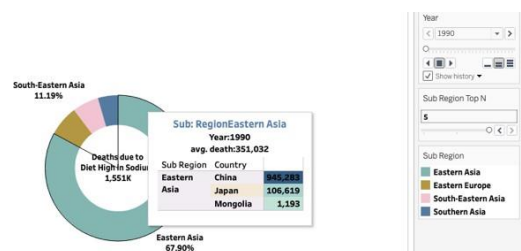


Fig. 26. Sub Region Eastern Asia

The pie chart drawn above shows that Eastern Asia had the highest number of death cases i.e. around 67% due to a diet high in sodium. Under this, China had the highest cases, approximately 945k cases for 1990. This trend was the same for all years as well.

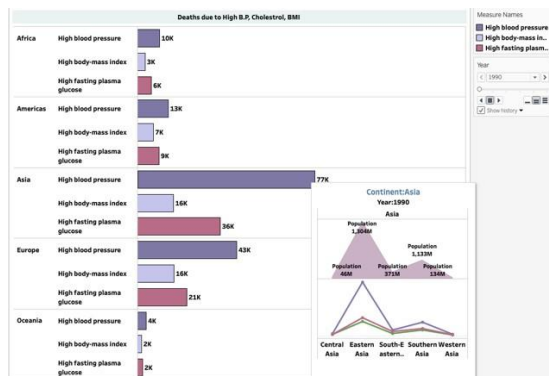


Fig. 27. Deaths due to high np

From the bar plot in fig 27, we can see that for year 1990, Europe and Asia had almost the same number of death cases due to High BMI (i.e. around 16k). However, Asia had the highest number of death cases due to High BP. Under Asia, Eastern Asia had the most cases as its population was also high. Likewise, the same behavior was observed for all years.

Analysis of deaths due to Lifestyle Choices:

We have compared number of deaths due to all addiction parameters i.e. Alcohol, Drugs and Smoking. By creating a filter, we can select top N countries with highest number of deaths in respective parameters. As we can see, China is country with maximum deaths due to all three parameter followed by India. We also showed forecast for number of deaths for next 10 years which has clear upward trend.

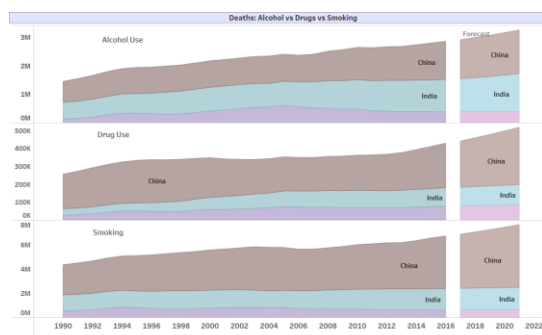


Fig. 28. Deaths Alcohol vs Drugs vs smoking

Then we compared deaths due to smoking and secondhand smoking. As we can see for the top 3 countries, though the deaths due to secondhand smoking is not as much as deaths due to smoking, it is still significant in number, infact almost one third of deaths due to smoking. China, India & Russia

being leading countries in deaths due to smoking and secondhand smoking.

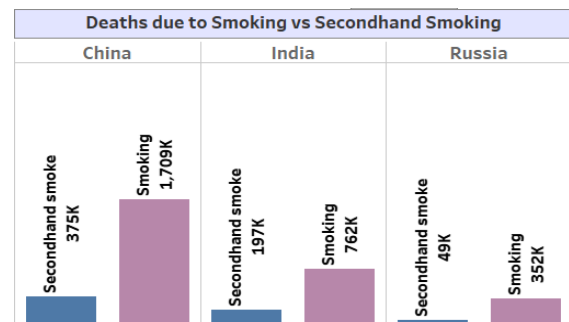


Fig. 29. Deaths due to Smoking vs Secondhand Smoking

Butterfly chart below compares total deaths due to all addictions (alcohol+Drugs+Smoking) and comparing it with deaths due to low physical activity. As we can relate that people who are addicted are less likely to indulge into physical activity hence causing diseases due to low physical activity resulting in death. By looking at top 10 countries, we can see that though numbers vary in scale, the ratio of deaths due to left parameter is same as right parameter.

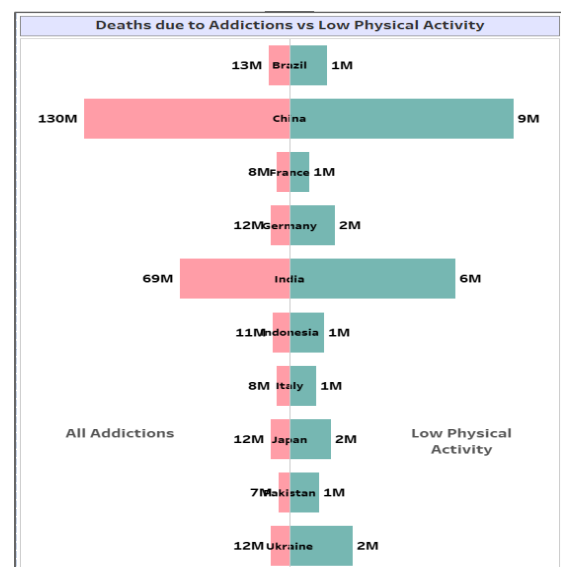


Fig. 30. Deaths due to Addiction vs Low physical Activity

Following graph shows correlation between two parameters namely, addictions and health issues. In particular, comparing the number of deaths due to high cholesterol, high blood pressure, and high BMI with the number of deaths due to Alcohol use, drug use and smoking. As we know one of the major factors for causing the above health issues is addictions. This positive strong correlation is clear from the comparison graph as they go hand-in-hand.

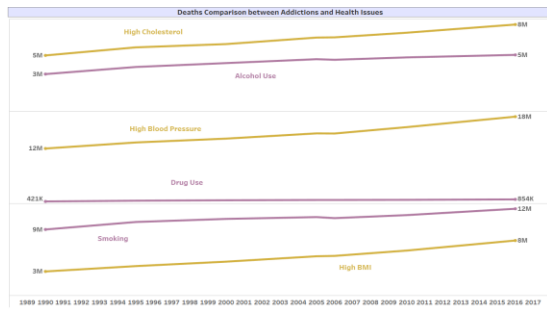


Fig. 31 Deaths comparison between Addiction and Health issues

From the research, we know that a diet low in nutrients leads to child stunting and child wasting. To show a correlation between them, we plotted a graph for deaths due to child stunting and deaths due to vitamin and nutrient deficiency. It was observed that both the graphs show a similar decreasing trend over the years.

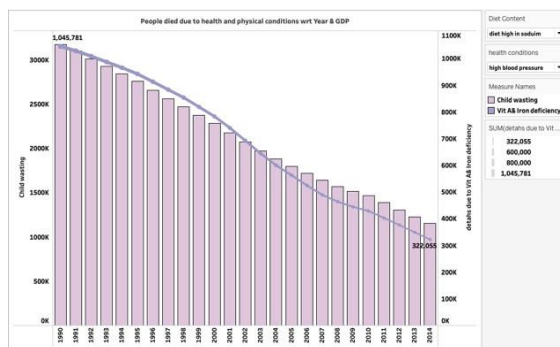


Fig. 32. Deaths due to health and physical conditions

VIII. RESULTS

Results for each of the mentioned death parameters have been summarized into four dashboards. Two of dashboards are as shown in fig 33 and fig 34 respectively.

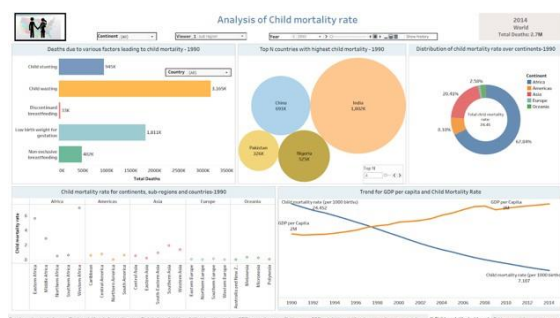


Fig. 33. Analysis of Child mortality rate

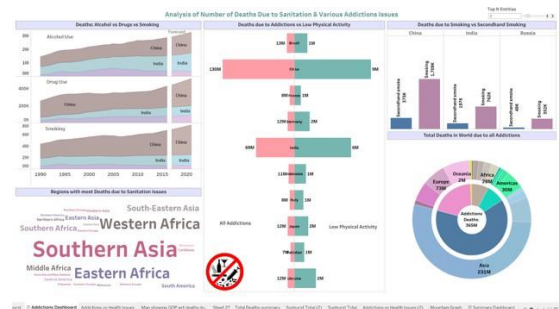


Fig. 34. Deaths due to sanitation and various addictions

IX. CONCLUSION

- **Air Pollution** - From 1990 to 2014 Deaths due to air pollution has continuously gone up i.e 58.3%. Asia continent remained the leading continent with the maximum number of deaths due to outdoor air pollution. However, the death by household pollution is reduced by 49%.
- **Child Mortality** – Around 20% of global deaths accounts for child mortality. 70% decrease in child mortality rate from 1990 to 2014.Expected to decrease further by 67% till 2030.
- **Diet low in Nutrition**- Approximately 40% of global deaths accounted due to diet low nutrition
- **Health Conditions**- Around 62% of global deaths were due to health issues specifically, deaths due to high BMI was highest overall i.e. around 19%
- **Unsafe Sanitation** – 1.4% of Global deaths are caused by poor sanitation and unsafe water source. In low Income countries poor sanitary conditions accounts for 5% Deaths.
- **Alcohol, Drugs, Smoking and Secondhand Smoking** – 74% of Global Deaths are due to all addiction related issues. Drug overdose deaths have more than tripled since 1990.

X. KEY LEARNINGS

- Adopt preventive measures for reducing air pollution
- Improve dietary content to reduce deaths and reduce child mortality
- Implement strict regulations about alcohol smoking & drugs availability and consumption
- Improve sanitary conditions worldwide and avoid water contamination to make clean water source

XI. FUTURE WORKS

- Recent few years had us all talking about number of deaths due to Covid-19 pandemic. We would like to further expand our analysis by involving deaths due such parameters
- We can involve other main parameters like accidental deaths or factors causing mass deaths i.e. natural calamities like earthquake, tsunami etc
- We would like to dig deeper into our analysis by normalizing our data by population of each country. As countries have different population density, it is only fair to put data in context for better and meaningful comparison world-wide

XI. REFERENCES

[1] United States Environmental Protection Agency, Air Data: Air Quality Data Collected at Outdoor Monitors Across the US, URL: <https://www.epa.gov/outdoor-airquality-data>.

[2] Visualization and Analysis of Air Pollution in US East Coast Cities, Diya Peng, Zhaopeng Xu, Jani Pallis, Xingguo Xiong, URL: [Visualization and Analysis of Air Pollution in US East Coast Cities](#)

[3] Household air pollution and health, URL: [Household air pollution and health](#)

[4] Outdoor Air Pollution, URL: [Outdoor Air Pollution - Our World in Data](#)

[5] Johnson, H. L., Fischer Walker, C., Liu, L., Black, R. E., Cousens, S., Johnson, H. L., Lawn, J. E., Rudan, I., Bassani, D. G., Jha, P., Campbell, H., Fischer Walker, C., Cibulskis, R., Eisele, T., & Liu, L. (2010).

[6] Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet*, 375, 1969–1987. <https://doi.org/10.1016/S0140>

[7] Dietary Habits, “Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017”, URL: [https://www.thelancet.com/article/S0140-6736\(19\)30041-8/fulltext](https://www.thelancet.com/article/S0140-6736(19)30041-8/fulltext)

[8] Deaths due to Health Conditions, “Combined Effect of Blood Pressure and Total Cholesterol Levels on Long-Term Risks of Subtypes of Cardiovascular Death”,

URL:

<https://www.ahajournals.org/doi/10.1161/HYPERTENSIONAHA.114.04639>

[9] Sellis, Timos., ACM Digital Library., & Association for Computing Machinery. Special Interest Group on Management of Data. (2011). *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*. ACM.

[10] Kosara, R., & MacKinlay, J. (2013). Storytelling: The next step for visualization. *Computer*, 46(5), 44–50. <https://doi.org/10.1109/MC.2013.36>

[11] Vasundhara, S. (n.d.). *DATA VISUALIZATION VIEW WITH TABLEAU*. <https://www.researchgate.net/publication/353847059>