# Visualizing Cancer Deaths by Type Worldwide (2019)

*Team Olivedrab* (CSC3007 P1)

SINGAPORE INSTITUTE OF TECHNOLOGY

## Introduction

Cancer is a group of diseases defined by the rapid proliferation of abnormal cells. It is the second largest cause of death worldwide. These cells have the potential to severely compromise one's health by invading surrounding tissues and, in some instances, spreading through other parts of the human body through the blood and lymphatic systems. In 2019, the Global Burden of Disease study found that approximately 10 million deaths worldwide, or one in six deaths were attributable to cancer.[1]

Measles is a highly contagious viral infection that can lead to severe complications, including pneumonia, encephalitis, and death. The average number of secondary cases resulting from a single infected individual is estimated to be between 10 and 16,[2] but vaccination is reported to be 97% effective at preventing measles[3]. Consequently, the introduction of the vaccine in 1963 was followed by a significant decline in measles incidence in the USA.

Given persistent skepticism in some communities about the efficacy of vaccines, it is important to communicate the impact of vaccination on public health. In this project, we built on a visualization of the measles incidence in the USA published by the Wall Street Journal[4] (Figure 1). This visualization has garnered praise as "One of the more compelling data visualizations produced in recent years"[5]. However, we contend that, despite its effectiveness in summarizing data, several aspects of the plot can be improved.

## Previous Visualization

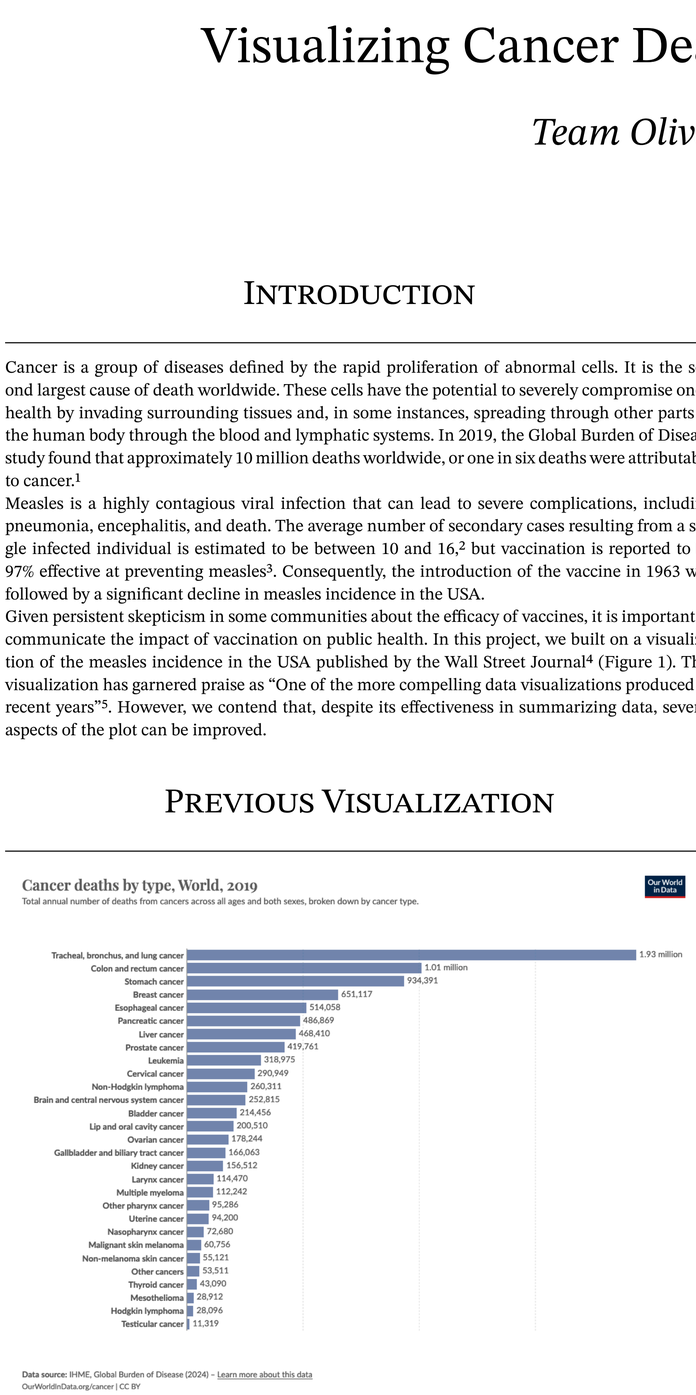


Figure 1: Measles incidence by state of the USA and year, published by the Wall Street Journal.

[1]National Cancer Institute. Cancer Statistics [Internet]. Bethesda, MD: National Cancer Institute; [cited 2024 Jun 28]

[2]National Cancer Institute. Cancer Statistics [Internet]. Bethesda, MD: National Cancer Institute; [cited 2024 Jun 28]

[3]L. Franconeri et al., "Two-dose measles vaccine effectiveness remains high over time: A French observational study, 2017–2019," Vaccine, vol. 41, no. 39, pp. 5797–5804, 2023]

[4]https://graphics.wsj.com/infectious-diseases-and-vaccines/

[5]https://www.mikelee.co/posts/2017-06-28-wsj-measles-vaccination-chart

## Strengths

- Straightforward and clear representation of data, providing a clear visual indication of the most deadly cancers
- Horizontal bar chart format makes it easy to compare the number of deaths caused by different types of cancer, viewers quickly identify which cancers have the highest and lowest mortality rates.
- Exact figures are alongside the bars providing precise data which is useful for detailed analysis and comparison

## Suggested Improvements

1. *Enhance informativeness* by adding demographic comparisons, such as cancer death rates by gender.
2. *Reduce cluttering and improve understanding* on cancer death statistics by grouping similar cancer types together.
3. Add in labels to both the x and y-axis for *improved clarity*.
4. Provide exact whole numbers instead of estimates to *enhance comprehension and clarity*.
5. Incorporate a straightforward and *visually appealing* colour scheme to *improve the chart's appearance*.
6. Include a colour legend to explain the meaning of the colours used.
7. Sort the values in descending order according to the total deaths to *highlight the most critical cancer types*.
8. Include grid lines at *meaningful intervals*, such as every 500,000, considering the value range extends up to a million.

## Implementation

### *Data*

- The cancer deaths by type chart was sourced from Our World In Data[6]. During data preparation, unnecessary columns were removed and cancer names were abbreviated for better visualization. We rounded death values for consistency and grouped cancer types by gender, allowing us to create a bi-directional bar chart for comparing different types of cancer deaths across genders.
- Data source was gathered from IHME, Global Burden of Disease. (2024)[7]

### *Software*

In our project, we employed Quarto along with the R programming language. We integrated various third-party libraries categorized by their respective purposes:

Data Import: *readr*: reads csv files into R

Data Manipulation & Tidying: *dplyr*: filters, selects and arranges data, *tidyr*: reshapes data frames to a tidy format

Data Visualization: *knitr*: integrates R code and output, affects how figures are displayed in the final document, *ggplot2*: creates graphics based on Grammer of Graphics, *ggnewscale*: applies multiple colours and fill scales within plots, *RColorBrewer*: provides addition colour palettes for graphics, *scales*: handles various aspects of scaling data

[6]https://ourworldindata.org/cancer

[7]https://vizhub.healthdata.org/gbd-results/

## Improved Visualization

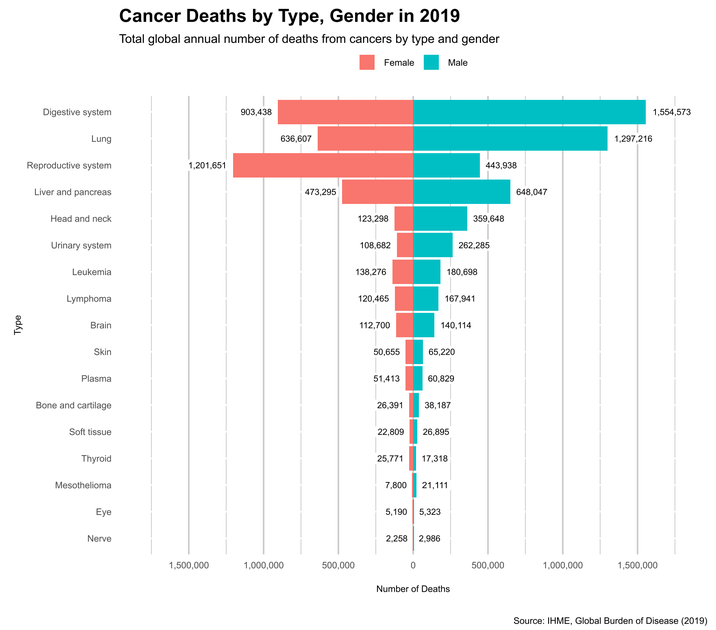


Figure 2: Revised visualization of measles incidence by state of the USA and year.

## Further Suggestions for Interactivity

As our visualization is designed for static clarity on a poster, it might be lacking some interactivity. Moving to a digital format like HTML with tools such as Plotly in R could enhance user engagement. Features like hover-over tooltips (values, etc), detailed categorical breakdowns (specific cancer organ), and yearly trend analysis would provide further dynamic insights and enrich user interaction.

## Conclusion

We successfully addressed our initial criticisms of the old visualization and implemented our improvements in our new visualization. By ensuring consistent cancer type categorization and including data split by gender in a bi-directional graph, the updated chart now offers a clearer and

more detailed view. We also improved accessibility with a colorblind-safe palette and added a detailed legend, enhancing overall visual appeal. These enhancements aim to make the visualization more user-friendly and inclusive, enabling deeper exploration and understanding of disease trends across genders for a diverse audience.