Visualizing Cancer Deaths by Type Worldwide (2019)

Team Olivedrab (CSC3007 P1)

SINGAPORE INSTITUTE OF TECHNOLOGY

Introduction

Cancer is a group of diseases characterized by the rapid proliferation of abnormal cells that can invade surrounding tissues and potentially spread to other parts of the body through the blood and lymphatic systems. It remains the second largest cause of death worldwide. The Global Burden of Disease study found that approximately 10 million deaths in 2019 were attributable to cancer, accounting for one in six deaths globally.

Due to the significance of cancers on public health on a global scale, it is crucial to understand the differences in cancer mortality based on gender. Analysing these differences can highlight disparities in cancer prevalence, which can aid in the diagnosis and treatments between men and women. This allows us to have informed targeted public health interventions and research priorities.

This project aims to improve upon a visualization of global cancer mortality data published by Our World in Data. The original visualization provides an detailed overview of the total number of deaths from different types of cancer each year, across all age groups and genders. Despite being a comprehensive visualisation of the data, we believe several aspects of the plot can be enhanced to provide clearer insights and more actionable information.

Previous Visualization

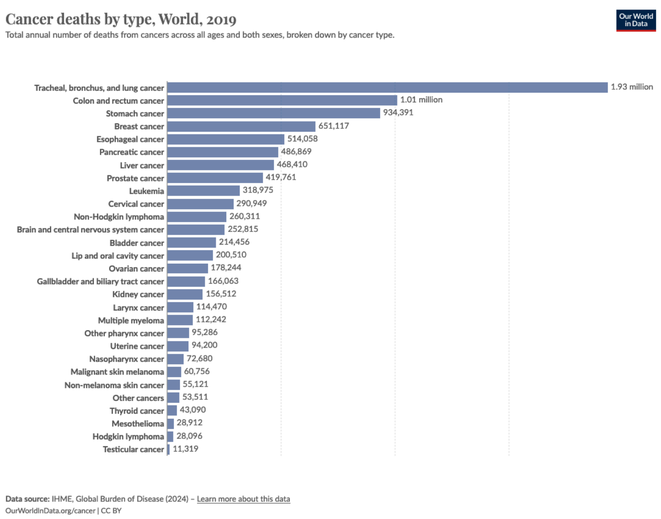

Figure 1: Cancer deaths by type, World, 2019, published by Our World in Data

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

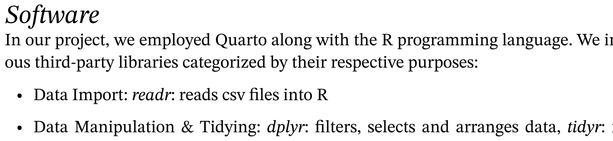

Figure 2: Strengths of Previous Visualization

Suggested Improvements

1. *Enhance informativeness* by adding demographic comparisons, such as cancer death rates by gender.
2. Group similar cancer types together to make the statistics more *accessible* and easier to understand for a general audience.
3. Add in labels to both the x and y-axis to *enhance the interpretability*.
4. Provide exact whole numbers instead of estimates to *enhance comprehension and clarity*.
5. Utilize distinct colors for each gender to *enhance clarity and visual differentiation*.
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 to over 1.5 million.

Implementation

Data

- The cancer deaths by type chart was sourced from Our World In Data[1]. 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)[2]

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

[1] https://ourworldindata.org/cancer
[2] https://vizhub.healthdata.org/gbd-results/

Improved Visualization

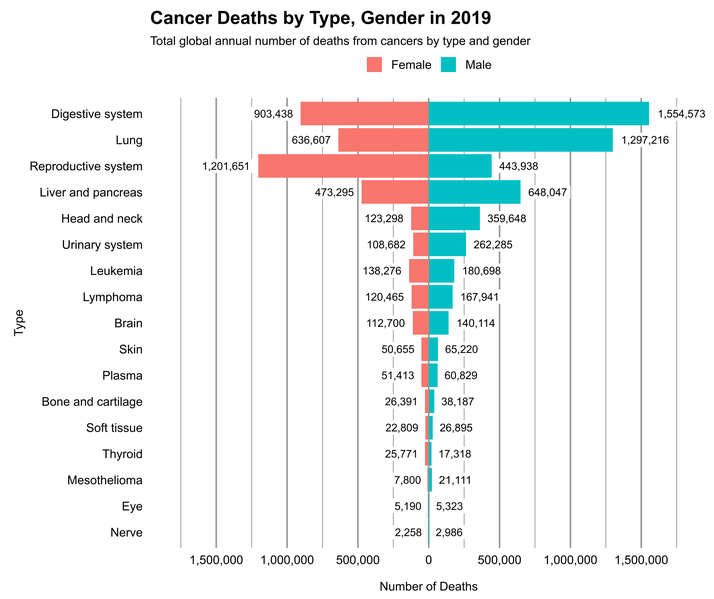

Figure 3: Revised visualization of Cancer Deaths by Gender in 2019.

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.