

Data Visualization for Neoplasm-Related Death Numbers: The Past and The Future

Esra Tafralı

MSc. in Global Public Health, Deggendorf Institute of Technology

GPH-M-1: Digital Health (WS24/25)

Prof. Dr. Dominik Böhler

July 18, 2025

Background

Use Case: Due to the aging population, the increasing trend of neoplasm-related metrics such as incidence, prevalence, and mortality raises concerns globally. It is essential for researchers, policy-makers, institutional health organizations, and public health specialists to identify neoplasm-related mortality patterns in a dynamic, interpretable, and comparative format. Therefore, this application was developed to visualize trends and projections in neoplasm-related death numbers across countries over time (1990-2050). It can serve as a tool for all stakeholders to make informed decisions based on historical and projected data.

Data Source: The dataset belongs to the Global Burden of Disease (2019) study and it is available online. I compiled the neoplasm-related death numbers in the years of 1990, 2000, 2010, 2020, 2030, 2040, and 2050 into a single file.

The Solution

A Shiny application was built to enable all stakeholders to:

- Interactively explore mortality trends at 10-year intervals between 1990 and 2050 at country-specific levels,
- Compare multiple countries through line graph and heatmaps,
- Visualize percentage changes on a world map to highlight global disparities between 2020 and 2050 projections.

Implementation

1. Data Cleaning

The dataset was already clean because it belonged to an academic research. However, to practice data cleaning, I intentionally introduced some inconsistencies into the dataset. Specifically, I added leading and trailing spaces to some country names and inserted non-numeric text such as the word "deaths" into cells that were originally numeric. Afterwards, I followed the below steps on RStudio to ensure dataset is analysis-ready:

- Removed leading and trailing spaces from country names,
- Stripped out non-numeric characters (e.g., text like "deaths") from numerical columns,
- Converted all year columns from character to numeric format for proper data handling.

2. Line Graph Visualization

Users can select a single country or multiple countries to explore the neoplasm-related death numbers between 1990 and 2050. The x-axis of the graph includes the years at 10-year intervals while the y-axis shows the death numbers. Year 2020 is highlighted with a dash line for reference.

3. Heatmap Visualization

The heatmap allows users to compare death numbers over the years for a single country or multiple countries. The colors represent the intensity of death numbers, ranging from yellow (low intensity) to red (high intensity). This color structure provides a quick and intuitive interpretation of trends and disparities over the years.

4. World Map Visualization

This map is developed to display the percentage change in the death numbers between the reference year (2020) and the projected year (2050). The color gradient allows users to identify where the most significant increases or decreases are expected.

5. User Interface Features

- The application includes three tabs, showing a single visualization in each, and allows the user to navigate easily.
- Selection of countries creates an interactive and dynamic structure.
- Helper texts for Line Graph and Heatmap inform users about multiple selection options.