

1. Data Sources and Integration

The Centers for Disease Control and Prevention (CDC)'s "Suicide Data and Statistics" dataset, available on the official U.S. Centers for Disease Control and Prevention website (<https://www.cdc.gov>), provides timely and authoritative data on suicide, one of the leading causes of death in the United States. This dataset breaks down suicide statistics by age, sex, race/ethnicity, and method, offering valuable insights for researchers, policymakers, and public health professionals. Aligned with the FAIR principles, the data is *findable* through a well-maintained government domain, *accessible* to all users without restriction, *interoperable* thanks to standardized formats that allow integration with broader health datasets, and *reusable* due to the CDC's clear documentation and metadata. Furthermore, the dataset reflects the CARE principles by serving the *collective benefit* of suicide prevention, ensuring *authority to control* through CDC oversight, upholding *responsibility* by sensitively presenting potentially distressing information, and maintaining *ethical* standards in public health reporting. Overall, this resource supports data-driven approaches to mental health, enabling evidence-based action to reduce suicide rates and improve outcomes across different demographic groups.

By integrating the **OECD Causes of Mortality dataset** with the **WHO Suicide Data (1950-2021)**, we can gain a deeper and more comprehensive understanding of the global patterns and trends in mortality, particularly focusing on suicide and its relationship with other causes of death. While the OECD dataset provides detailed information on the leading causes of death across OECD countries, categorized by demographic factors such as age, sex, and health conditions, the WHO dataset offers a long-term, global view of suicide data, spanning over seven decades. Combining these datasets enables us to explore not only the prevalence and demographics of suicide but also how suicide trends correlate with shifts in other causes of death, such as chronic diseases, mental health conditions, and accidents. Additionally, it allows for cross-country comparisons, highlighting how different nations address suicide prevention and related public health issues.

CDC dataset:

<https://www.cdc.gov/suicide/facts/data.html>

OECD Causes of Mortality dataset:

<https://www.kaggle.com/code/raselmeya/country-based-suicides-analysis>

WHO Suicide Data (1950-2021):

https://www.kaggle.com/datasets/kumaranand05/who-suicide-data-1950-2021?select=combined_processed_data.csv

2. Research Question

How can age-specific mortality patterns inform targeted healthcare resource allocation for ageing populations worldwide?

Sub-question:

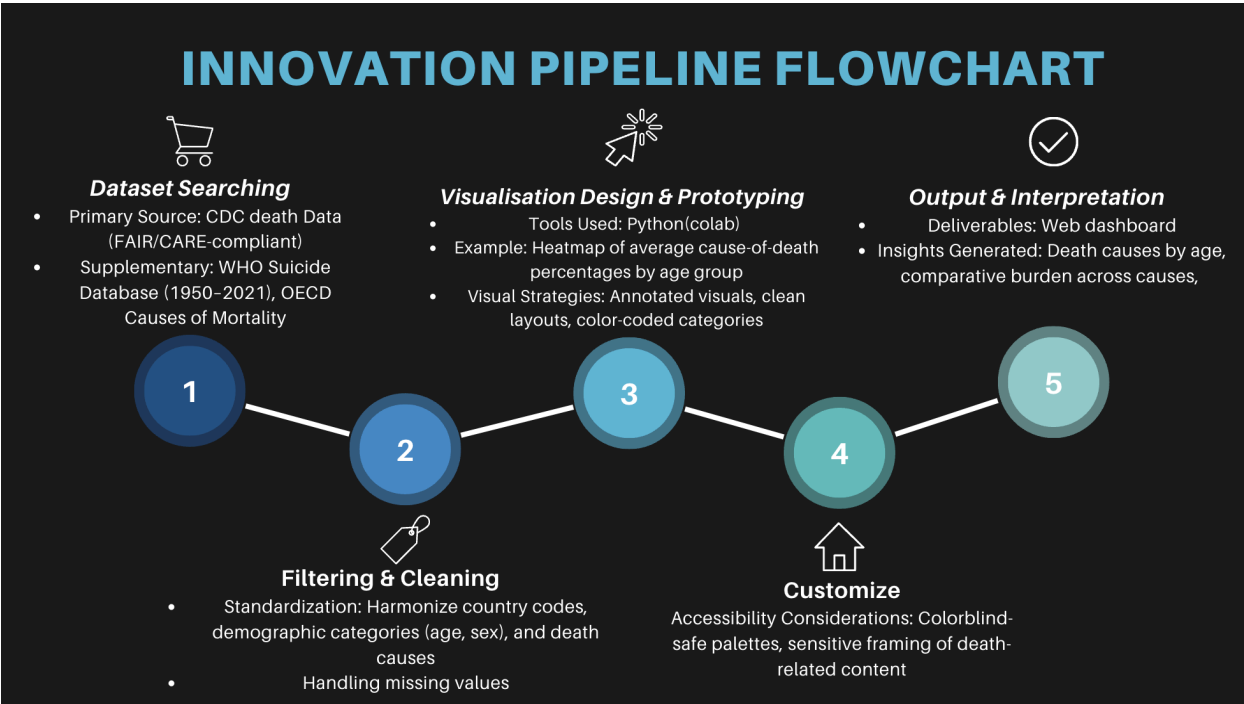
What are the top causes of death for each age group, and how do they shift as people age?

Rationale:

Understanding how leading causes of death vary across age groups is essential for designing age-sensitive healthcare strategies. This sub-question aims to identify which causes of death are more prevalent among younger populations, such as accidents or mental health-related issues, and which dominate among older adults, such as chronic illnesses, falls, or dementia. By mapping these patterns, policymakers can allocate resources more effectively, prioritizing disease prevention and management programs that align with age-specific health burdens. Such insights support data-driven policy to improve health outcomes and reduce avoidable mortality among aging populations.

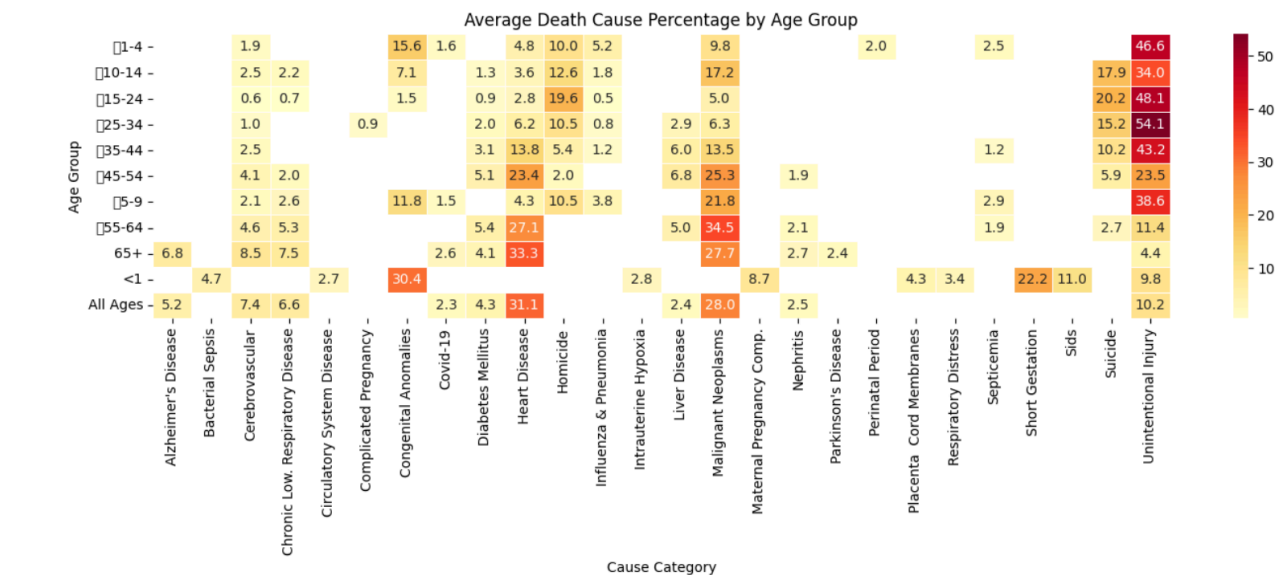
3. Sample Visualizations, Advanced Tools & Innovation Flowchart

Figure 1. Innovation pipeline flowchart



3.1 Prototype Visualisation

Figure 2. Average Death Cause Percentage by Age Group



Source: CDC

3.2 Cite Precedents

Our World in Data. (n.d.). *Self-harm death rates by age*. Retrieved April 20, 2025, from <https://ourworldindata.org/grapher/self-harm-death-rates-by-age>

Qian, K., Xia, F., & Stasko, J. (2022). *Visualizing Suicide Risk Prediction in Clinical Settings*. IEEE VIS Workshop on Visualizations for NLP and Healthcare (NLVis). Retrieved April 20, 2025, from https://content.iiieeevis.org/year/2022/paper_w-nlvis-1009.html

Team Contribution

All team members contributed equally and collaboratively to the successful development of this project. Qianyue Jiao led the data collection and preprocessing phase, ensuring that all datasets adhered to international standards such as FAIR and CARE principles. Yifei Yang focused on exploratory data analysis and the application of potential methods for insight generation, contributing to both the technical depth and interpretability of the project. Zixuan Li was responsible for designing the interactive visualization dashboard and drafting the innovation flowchart, integrating user-centred design principles and technical implementation.