Death by Data

Mutaz Alhennawi, mutaz.alhennawi@utah.edu, u1463385

Amsa Khan, asma.khan@utah.edu, u6011009

Project repository: https://github.com/asmakhans/DataVis

Background and Motivation

Understanding the causes of death is one of the most important challenges in public health. Mortality statistics shape national health priorities, influence international funding decisions, and guide prevention campaigns. Organizations such as IHME and WHO release detailed datasets with global, national, and subnational estimates of deaths by cause, sex, and age. These datasets contain millions of records across decades, making them a rich evidence base.

Despite their richness, the datasets present a major gap. Policymakers, educators, journalists, and the public often struggle to extract insights from them because of their size and complexity. For example, a policymaker might want to compare deaths from cardiovascular disease in India versus the United States, but the raw tables require several filtering and aggregation steps before such a comparison becomes interpretable. Similarly, a teacher might want to illustrate to students how injury-related deaths in adolescents differ from deaths in older adults, but static tables cannot highlight these patterns without additional processing.

The objective of this project is to transform the dataset into an interactive visualization that highlights the most important patterns: which causes of death dominate in specific regions, how these patterns differ across sex and age, and how they evolve over time. The project aims to simplify access without sacrificing detail, allowing users to explore health outcomes through an intuitive interface rather than raw numbers.

The implications of such a tool extend beyond academic analysis. A policymaker could use the visualization to identify urgent needs for intervention. A journalist could use it to communicate

trends to a wide audience. An educator could use it to spark classroom discussions about global health inequalities. By bridging the gap between large-scale datasets and human understanding, the project has the potential to support evidence-based decision-making, improve teaching, and enhance public awareness of global health challenges.

Project Objectives

- Identify the leading causes of death across countries, genders, and age groups.

- Show trends over time (e.g., 1990 to 2021) for major causes.

- Enable users to compare mortality rates across populations (for example, comparing the United States to India or men to women).

- Provide a resource for educators, policymakers, and the public to explore mortality patterns. General users can explore health risks in their region.

Data

- **Source**: Institute for Health Metrics and Evaluation (IHME) Global Burden of Disease (GBD) 2021 dataset. (https://www.healthdata.org/)

- **Variables**: deaths, DALYs (loss of the equivalent of one year of full health), age-standardized rates, by location, age group, sex, cause, and year.

- The dataset provided includes columns for location, cause, metric type (number, percent, rate), age group, sex, and uncertainty bounds.

Data Processing

- Filter for *measure_name* = *Deaths* to focus on mortality.

- Aggregate data by country, cause, and year.

- Normalize death counts by population to generate mortality rates.

- Handle missing or undefined values by excluding rows with no numeric estimates.

- Derived quantities: age-standardized mortality rate, ranking of causes of death by location.

Visualization Design

- **Option 1**: Interactive world map where users can select a cause and year to see death rates by country

Sketch #1: Interactive world map where users can select a cause and year to see death rates by country.
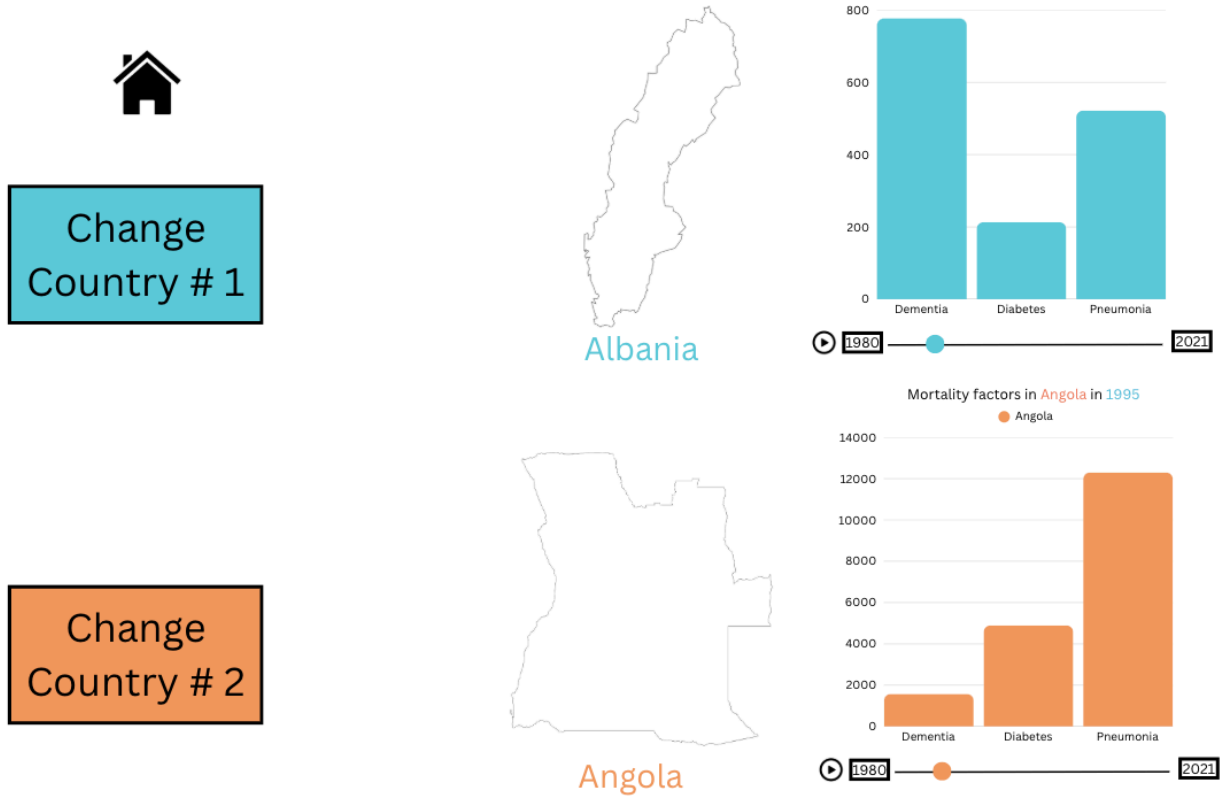
Select a country to visualize mortality variables

Select mortality variables to visualize:

- Dementia
- Malaria
- Hypertension
- Cardiovascular disease
- ....

- **Option 2**: Time-series plots comparing causes of death across decades for a singular selected country.

Sketch #2: Time-series plot comparing causes of death across decades for a singular selected country.

Mortality factors in Albania in 1995

Change Country

Albania
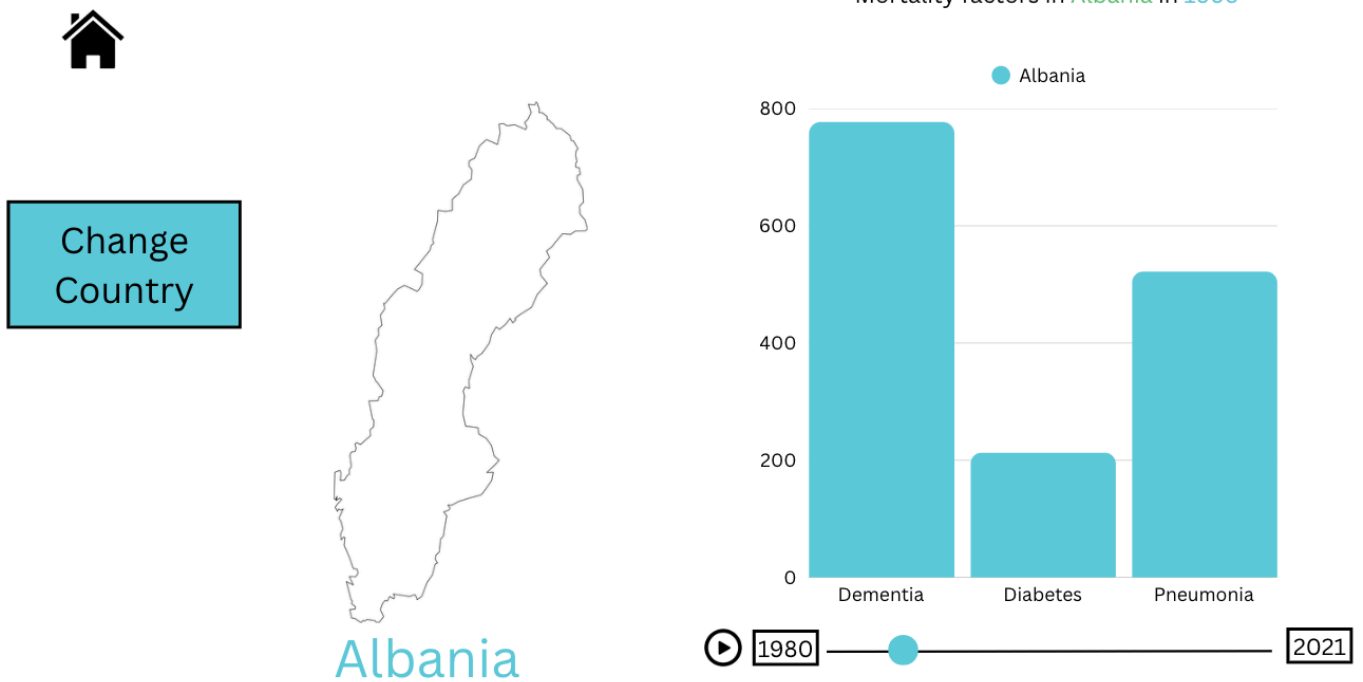
- **Option 3**: Time-series plots comparing causes of death across decades for two selected countries.

Sketch #3: Time-series plot comparing causes of death across decades for two selected countries.

Change Country # 1

Albania

Change Country # 2

Angola

Mortality factors in Albania in 1995
● Albania

Dementia   Diabetes   Pneumonia

1980 ——————— 2021

Mortality factors in Angola in 1995
● Angola

Dementia   Diabetes   Pneumonia

1980 ——————— 2021

- **Option 4:** Time-series plots comparing causes of death across decades for two selected countries according to population scale to depict misrepresentations
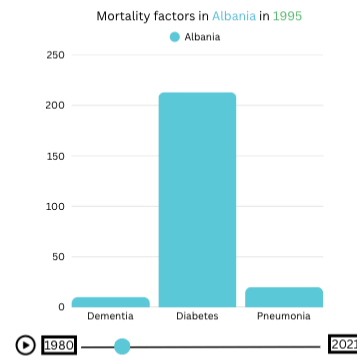
Sketch #4: Time-series plot comparing causes of death across decades for two selected countries according to population scale to depict misrepresentations
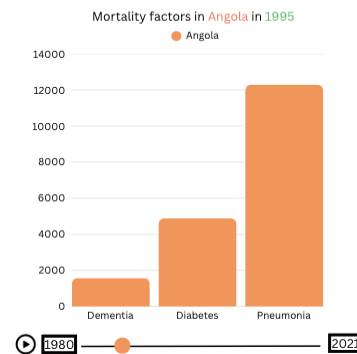
Change Country # 1

Albania

Change Country # 2

Angola

Mortality factors in Albania in 1995
● Albania

Dementia   Diabetes   Pneumonia

1980   2021

Mortality factors in Angola in 1995
● Angola

Dementia   Diabetes   Pneumonia

1980   2021

Data representations can be misrepresented due to population size!!!
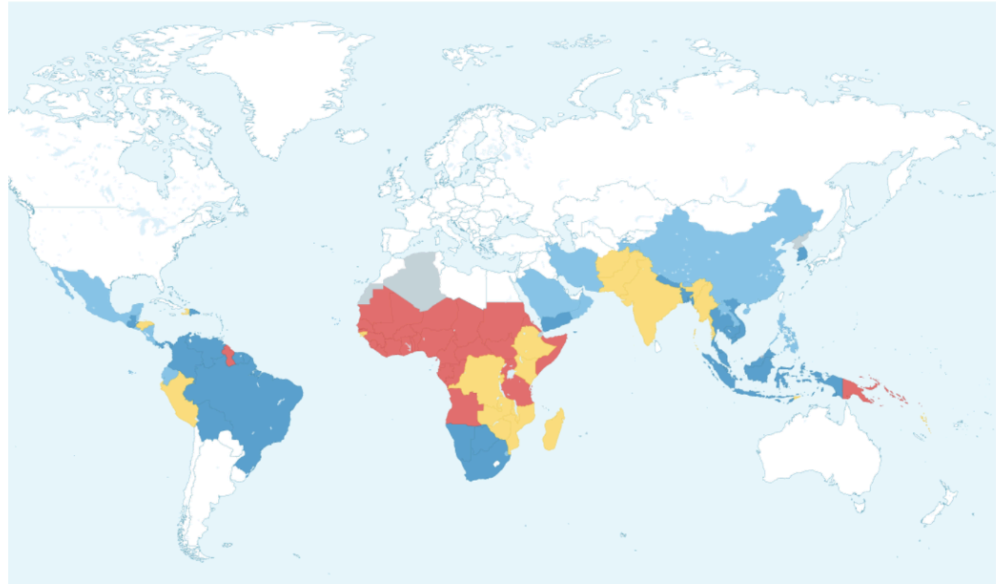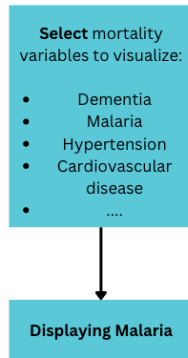
Population size in Albania in 1995: 3,413,904

Population size in Angola in 1995: 11,000,904

- **Option 5:** Interactive global heatmap comparing a single selected mortality variable.

Sketch #5: Interactive global heat map comparing a single selected mortality variable.

Final Design Choice:

Combining an interactive world map with various options. Users select a country, then view time trends and cause breakdowns. Color encodings will represent death rates, while hover tooltips will display absolute numbers with uncertainty ranges.

Features

- Interactive map of mortality by country.

- Filtering by cause, year, sex, and age group.

- Ability to compare two or more countries and scale the population to get a more accurate representation of comparisons.

Optional Features

- Downloadable charts for reports.
- Animated trends over time.

Project Schedule

Week 5: Project review and discussions
Week 6: Set up the framework environment and create the page workflows
Week 7: Exam week
Week 9: Data cleaning and preprocessing\Integrate features for filtering and comparison.
Week 10: Milestone

Week 11: Peer feedback

Week 12: Work on must-have features

Week 13: Work on must-have features

Week 14: User testing and refinement/try to implement optional features

Week 15: Finalize visualization and documentation

Week 16: Submission