**Data Analysis and Visualisation Of HMD data**

**Created By:** Aparna Krishnakumar (2023)

**Data:**

The data is taken from The Human Mortality Database (HMD). The data contains information about mortality rates by age group throughout the world over the

past 100-200 years (approximately) (www.mortality.org, n.d.).

Each data set from the project contains the following information:

1. Country Name

2. The time span for which the data is collected

3. Age range/intervals

4. Number of female deaths

5. Number of male deaths

**Data Preparation:**

Based on the countries given, here are twelve countries were selected that can be compared in terms of population and mortality rates, namely:

* Australia (AUS)
* Austria (AUT)
* Belgium (BEL)
* Canada (CAN)
* Chile (CHL)
* Denmark (DNK)
* Finland (FIN)
* Italy (ITA)
* Japan (JPN)
* United States of America (USA)
* New Zealand
* Israel

These countries were chosen because they represent a mix of different regions, population sizes, and economic development levels.

In terms of population, the countries range from smaller ones like Austria and Chile, to larger ones like the United States and Japan. Population statistics can provide insight into a country's demographic structure, as well as potential factors that contribute to health outcomes such as income level, education level, and access to healthcare (World health statistics, 2021).

A population table, death rate table and birth rate table was created from the data containing the details of the selected countries for years from 1920-2023. These three were then combined and pre-processed using python to create a combined table containing Country name, Region (Column newly added group countries), Year, Population, Births, Deaths.

**Visualizations:**

**A screenshot of a computer screen

Description automatically generated with low confidence**

**Visual Design Type:** Treemap

**Name of Tool:** Tableau

**Country:** Australia (AUS), Austria (AUT), Belgium (BEL), Canada (CAN), Chile (CHL), Denmark (DNK), Finland (FIN), Italy (ITA), Japan (JPN), United States of America (USA), New Zealand, Israel

**Years:** 1920 -2023

**Visual Mappings:**

* Color: Color is mapped to deaths in each year as the number of deaths increase it becomes more green and as it decreases it moves to yellow and red indicates lowest.
* Shape: Each bottom level rectangle represents years in decreasing order of deaths and each country is grouped together.
* Size: Size is mapped to deaths in each year as biggest rectangle in a country shows year with most deaths.
* Position: The position of each rectangle depends on the number of deaths as it reduces it moves further left.
* Hierarchy: The years are grouped by country, Country -> Years

**Unique Observation:** We can clearly see that the highest death among all the recorded data is for Japan, 1947 which is shown is the only dark green rectangle in the graph. This is due to the aftereffects of World War II and sudden outbreak of infectious diseases during that time period (Cambill et al., n.d.).

More visualizations on the countries chosen:

A map of the world

Description automatically generated with low confidence

A screenshot of a computer

Description automatically generated with medium confidence

Sources:

World Health Organization. (2021). World health statistics 2021: Monitoring health for the SDGs, sustainable development goals. Geneva: World Health Organization.  
  
Sugiura, Y., Ju, Y.-S., Yasuoka, J. and Jimba, M. (2007). Rapid increase in Japanese life expectancy after World War II. [online] Available at: <https://cdn1.sph.harvard.edu/wp-content/uploads/sites/114/2012/10/RP245.pdf>.

Cambill, B., Yashin, A., Vaupel, J., Nanjo, T. and Shigematsu (n.d.). Available at: <https://user.demogr.mpg.de/jwv/pdf/Vaupel-IIASA-WP-86-078.pdf>.

www.mortality.org. (n.d.). Zipped Data Files. [online] Available at: https://www.mortality.org/Data/ZippedDataFiles [Accessed 9 May 2023].