

DATA608 Final Project Proposal

Banu Boopalan

10/23/2022

BANU BOOPALAN PROPOSAL: DATA 608

##DATA SOURCES:

NCHS - Leading Causes of Death: United States <https://data.cdc.gov/NCHS/NCHS-Leading-Causes-of-Death-United-States/bi63-dtpu>

Summary of the dataset : This dataset presents the age-adjusted death rates for the 10 leading causes of death in the United States beginning in 1999. Data are based on information from all resident death. . .

Global health estimates: Leading causes of death

<https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death>

Cause-specific mortality, 2000–2019

Final Project Proposal :

Use : R, Shiny, Plotly Example: Some of these graphs on plotly's page are interesting so I would like to build similar graphs to make the visualization dashboard interactive for the user. Reference website : <https://blog.plotly.com/post/127132206092/analyzing-data-eighteen-graphs-about-the-death>

Steps to present data

1. Data Preparation
2. Exploratory Analysis:

Top ten causes of death by Year, Death percentage change between Years, Mortality by State, Identify Leading causes of death.

3. I would like build visualizations and summarize data to allow the user to use R, Shiny and Plotly to understand and confirm summary findings such as referenced by WHO below about causes of death within the WHO dataset across countries.

4. Represent data within small multiples and design graphs for using small multiples. Maybe grouping by similar sized countries on WHO dataset and comparing the causes of death similarities.
5. Create map based visualization for US states and the select countries of the world showing causes of death across different years.

Citing the reference findings below from WHO site

"Summary findings

Noncommunicable diseases have become more prominent with Alzheimer's disease and diabetes entering whereas communicable diseases are on the decline with both HIV and tuberculosis dropping out of the top 10. Ischaemic heart disease was the top cause of death in both 2000 and 2019. It is responsible for the largest increase in deaths – more than 2 million – over the last two decades. Lower respiratory infections were responsible for the most deaths in the communicable disease category in both 2000 and 2019, although the total number of deaths from lower respiratory infections has decreased. "