Data Report of Multiple Causes of Mortality in the US in the past decade

by

M.S., Wenxiao Zhou

University of Connecticut, 2021

Project Plan

STAT 5225—Data Management and Programming in R&SAS

Advisor: Haim Bar

Spring 2021

Motivation

With the progress of medical treatment and science and technology, people can control diseases, avoid risks and accidents, and reduce mortality through effective and reasonable ways. These advances have been made thanks to the efficient use of mortality statistics for scientific data analysis. Mortality statistics contain information on individuals' life and health problems that can help improve the public's health by using systematic methods to monitor the whole community. Although the abrupt increase in COVID-19 mortality in 2020 affects the overall trend of mortality, a long-term analysis of mortality trends over the past decade is of great significance for long-term prevention and monitoring.

In my study, a work basing on the Data of Mortality Multiple Causes in the US within the year 2010-2019 aims to make a review and summary of the mortality rate variation trend, personal characteristics of the death cases, the main factors of causing mortality within different groups are being discussed. What's more, a specific analysis on year 2019 is focusing on the main causes of mortality in both infants and adults. A comparison of death manners' ratio is considered to measure trends in natural mortality and suicide rates. All the work is mainly the combination of descriptive analysis by EDA and frequency tables. Further data visualization basing on the geographical dataset will be included.

Outline of Project

- **Overall analysis**
  1. Age-adjusted Mortality Rate for Year 2010-2019 among overall group and category by sex.
  2. Descriptive analysis: construct frequency tables of marital status, resident status, education level for Year 2010-2019 respectively.
  3. Descriptive analysis: Determine the relationship between Sex, Race, Education among years, and how factors influence mortality rates. (Bar plot of Contingency tables to determine first, Bubble Charts to make visualization)
  4. Make further discussion and conclusions on the findings in 2-3.

- **Year 2019 analysis**
  1. Finding the top find mortality causes of infant cases as well as of adult cases in year 2019.
  2. Determine the average frequencies of month mortality and day mortality in year 2019 and summarize the reasons. (Coxcombs chart)
  3. Draw the map basing on the mortality in states of the country, includes details of outcome levels on the plot for further analysis. (Circular Barchart & Bubble/ Choropleth/ Interactive map plot)
  4. Finding the relationships between latitude/longitude division and the effect of cause of death on mortality (Further geographical data to support analysis, Bubble plot)

- **Comparison analysis**
  1. Comparison of mortality rate on age groups, race groups, leading causes in the year 2018 and 2019 (Bar plot)
  2. Discuss the trend of the manner of death in year 2015-2019 to see if there is a trend towards a lower natural mortality ratio and higher suicide ratio. Conclude the findings. (Dual-line chart)