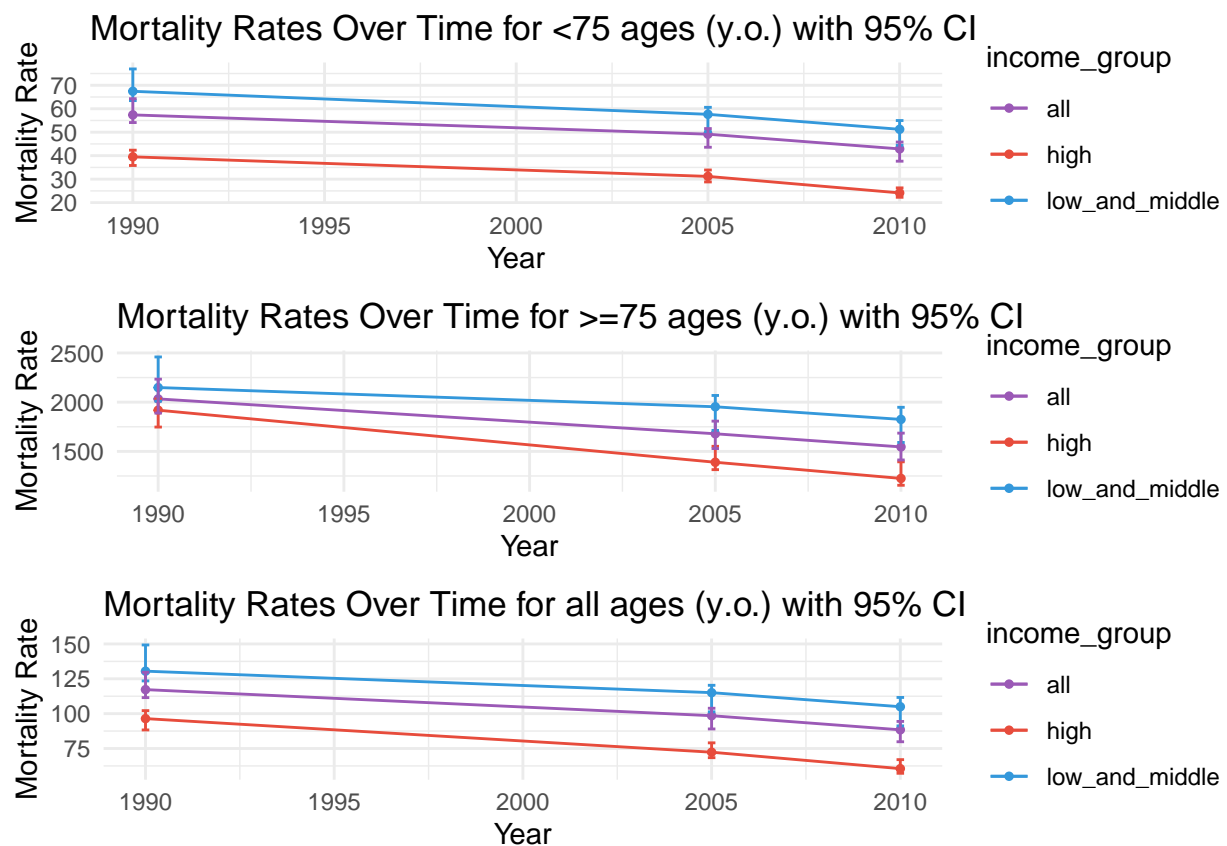


HW1

Yujia Cai

2023-10-19

Table 1, Feigin et al. (2014) Lancet 383:245-255



The plot here I choose is basically a line chart. I consider it the best to reflect a trend of mortality rate along with years, with error bars to indicate the range of 95% CI. We have three lines indicating the different behaviors in varying income groups. The subplots then also show the the difference in varying age groups.

If we are interest in making a comprehensive graph of all the information in the table, we may need to give more subplot, which is nine in this case. For each age and income combination, we can then draw lines for each of the variables of interest, e.g.,incidence, prevalence, mortality rate.

Code Appendix

```
knitr::opts_chunk$set(echo = TRUE)
library(ggplot2)
library(gridExtra)
data <- read.csv('feigin2014_table1_mortality.csv')
plots <- list()
for (age_category in unique(data$age_group)) {
  subset_data <- data[data$age_group == age_category, ]
  p <- ggplot(subset_data, aes(x = year, y = mortality_rate, color = income_group)) +
    geom_line() +
    geom_point(size = 1) +
    geom_errorbar(aes(ymin = interval_low, ymax = interval_high), width = 0.2) +
    labs(title = paste("Mortality Rates Over Time for", age_category,
                      "ages (y.o.) with 95% CI"), x = "Year",
         y = "Mortality Rate") +
    scale_color_manual(values = c("#9B59B6", "#E74C3C", "#3498DB")) +
    theme_minimal()
  plots[[age_category]] <- p
}

grid.arrange(grobs = plots, nrow = 3)
```