HQSC\_Cancer\_Data

Kyle Gibson

25/09/2022

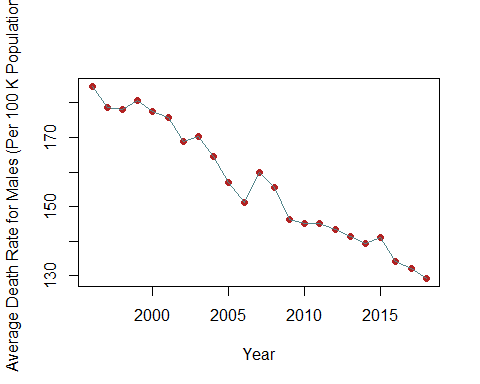
library(readxl)  
library(ggplot2)  
library(gridBase)  
library(gridExtra)  
library(gridGraphics)

## Loading required package: grid

getwd()

## [1] "C:/HQSC"

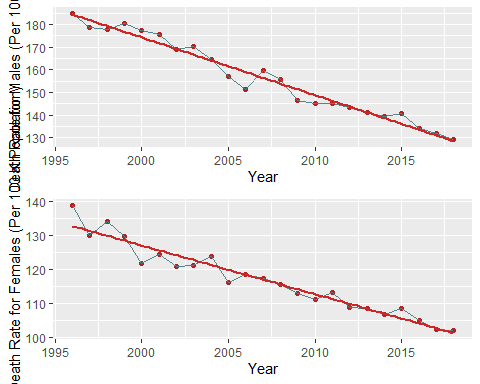
setwd("C:\\HQSC")  
  
avg\_death\_rate\_male <- read.csv("avg\_death\_rate\_male.csv") # Reading in table from SQL query.  
View(avg\_death\_rate\_male) # Viewing data.  
  
plot(avg\_death\_rate\_male$Year, avg\_death\_rate\_male$Average.Death.Rate.for.Males..Per.100.K., xlab = "Year", ylab = "Average Death Rate for Males (Per 100 K Population)", pch = 19, col = "firebrick")  
lines(avg\_death\_rate\_male, col = "cadetblue4") # Traditional graphics used but I have decided to use ggplot going forward.



p1 <- ggplot(data = avg\_death\_rate\_male, mapping = aes(x = Year, y = Average.Death.Rate.for.Males..Per.100.K.)) +   
 geom\_point(col = "firebrick") +   
 labs(y = "Average Death Rate for Males (Per 100 K Population)", x = "Year") + geom\_line(col = "cadetblue4")+  
 stat\_smooth(method = "lm", se = FALSE, col = "firebrick3")   
   
  
avg\_death\_rate\_female <- read.csv("avg\_death\_rate\_female.csv")  
  
p2 <- ggplot(data = avg\_death\_rate\_female, mapping = aes(x = Year, y = Average.Death.Rate.for.Females..Per.100.K.)) +   
 geom\_point(col = "firebrick") +   
 labs(y = "Average Death Rate for Females (Per 100 K Population)", x = "Year") + geom\_line(col = "cadetblue4")+  
 stat\_smooth(method = "lm", se = FALSE, col = "firebrick3")  
  
  
grid.arrange(p1, p2, nrow = 2) # Viewing male and female death rate plots.

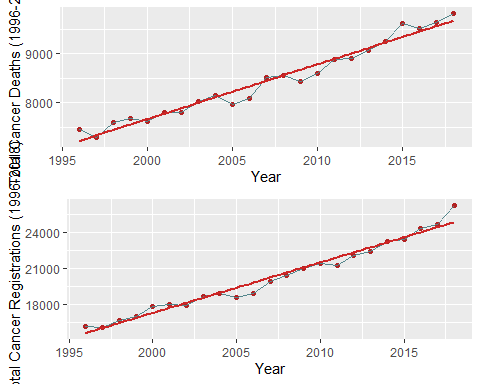
## `geom\_smooth()` using formula 'y ~ x'

## `geom\_smooth()` using formula 'y ~ x'



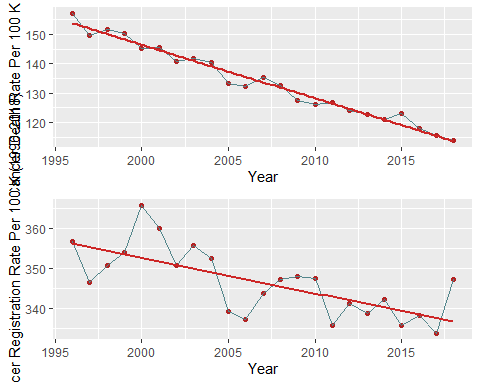
total\_cancer\_deaths <- read.csv("total\_cancer\_deaths.csv")  
View(total\_cancer\_deaths)  
  
  
p3 <- ggplot(data = total\_cancer\_deaths, mapping = aes(x = Year, y = Total.Cancer.Deaths)) +   
 geom\_point(col = "firebrick") +   
 labs(y = "Total Cancer Deaths (1996-2018)", x = "Year") + geom\_line(col = "cadetblue4")+  
 stat\_smooth(method = "lm", se = FALSE, col = "firebrick3")  
  
total\_cancer\_regs <- read.csv("total\_cancer\_regs.csv")  
  
p4 <- ggplot(data = total\_cancer\_regs, mapping = aes(x = Year, y = Total.Cancer.Registrations)) +   
 geom\_point(col = "firebrick") +   
 labs(y = "Total Cancer Registrations (1996-2018)", x = "Year") + geom\_line(col = "cadetblue4")+  
 stat\_smooth(method = "lm", se = FALSE, col = "firebrick3")  
  
grid.arrange(p3, p4, nrow = 2)

## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'



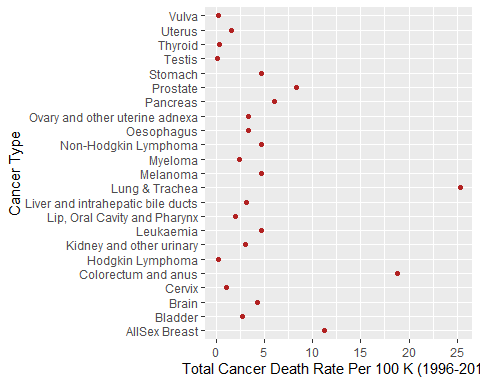
avg\_death\_rate\_over\_time <- read.csv("avg\_death\_rate\_over\_time.csv")  
View(avg\_death\_rate\_over\_time)  
  
p5 <- ggplot(data = total\_cancer\_regs, mapping = aes(x = Year, y = avg\_death\_rate\_over\_time$Total.Cancer.Death.Rate.Per.100.K..1996.2018.)) +   
 geom\_point(col = "firebrick") +   
 labs(y = "Total Cancer Death Rate Per 100 K (1996-2018)", x = "Year") + geom\_line(col = "cadetblue4")+  
 stat\_smooth(method = "lm", se = FALSE, col = "firebrick3")  
  
total\_cancer\_regs\_rate <- read.csv("total\_cancer\_regs\_rate.csv")  
  
p6 <- ggplot(data = total\_cancer\_regs\_rate, mapping = aes(x = Year, y = Total.Cancer.Registration.Rate.Per.100.K..1996.2018.)) +   
 geom\_point(col = "firebrick") +   
 labs(y = "Total Cancer Registration Rate Per 100 K (1996-2018)", x = "Year") + geom\_line(col = "cadetblue4")+  
 stat\_smooth(method = "lm", se = FALSE, col = "firebrick3")  
  
grid.arrange(p5, p6, nrow = 2)

## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'



total\_death\_rate\_cancer\_type <- read.csv("total\_death\_rate\_cancer\_type.csv")  
  
ggplot(total\_death\_rate\_cancer\_type, mapping = aes(x = Total.Cancer.Death.Rate..Per.100.K.Population., y = Cancer.Type)) +  
 geom\_point(col = "fire brick") +  
 labs(y = "Cancer Type", x = "Total Cancer Death Rate Per 100 K (1996-2018)")+  
 stat\_smooth(method = "lm")

## `geom\_smooth()` using formula 'y ~ x'



avg\_reg\_rate\_male <- read.csv("avg\_reg\_rate\_male.csv")   
  
p7 <- ggplot(data = avg\_reg\_rate\_male, mapping = aes(x = Year, y = Average.Registration.Rate.for.Males..Per.100.K.)) +   
 geom\_point(col = "firebrick") +   
 labs(y = "Average Registration Rate for Males", x = "Year") + geom\_line(col = "cadetblue4")+  
 stat\_smooth(method = "lm", se = FALSE, col = "firebrick3")   
  
  
avg\_reg\_rate\_female <- read.csv("avg\_reg\_rate\_female.csv")  
  
p8 <- ggplot(data = avg\_reg\_rate\_female, mapping = aes(x = Year, y = Average.Registration.Rate.for.Females..Per.100.K.)) +   
 geom\_point(col = "firebrick") +   
 labs(y = "Average Registration Rate for Females", x = "Year") + geom\_line(col = "cadetblue4")+  
 stat\_smooth(method = "lm", se = FALSE, col = "firebrick3")  
  
grid.arrange(p7, p8, nrow = 2)

## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'

