Proyecto lavado de Manos

Morales Torres

2024-06-22

## Foto en blanco y negro de un hombre con traje y corbata Descripción generada automáticamenteDesarrollo en R

## Médico húngaro Dr. Ignaz Semmelweis trabajó en el Hospital General de Viena con la fiebre de parto de las pacientes. La fiebre de parto es una enfermedad mortal que afecta a las mujeres que acaban de dar a luz, y en la década de 1840, como el 10% de las mujeres de dar a luz murió en el Hospital General de Viena. El Dr. Semmelweis descubrió que se trataba de la contaminación de las manos de los doctores de la entrega de los bebés, y el 1 de junio de 1847, decretó que todas las personas deben lavarse las manos, un poco ortodoxo y controvertido solicitud; nadie en Viena sabía acerca de las bacterias.

## Usted va a volver a analizar los datos que hizo Semmelweis descubrir la importancia del lavado de manos y su impacto en el hospital.

## Los datos se almacenan como dos archivos CSV dentro de la data carpeta.

yearly\_deaths\_by\_clinic.csv contiene el número de mujeres que dan a luz en las dos clínicas del Hospital General de Viena entre los años 1841 y 1846.

|  |  |
| --- | --- |
| **Columna** | **Descripción** |
| year | Años (1841-1846) |
| births | Número de nacimientos |
| deaths | Número de muertes |
| clinic | Clínica 1 o clínica 2 |

## monthly\_deaths.csv contiene los datos de "Clínica 1' de los hospitales, donde la mayoría de las muertes.

|  |  |
| --- | --- |
| **Columna** | **Descripción** |
| date | Fecha (AAAA-MM-DD) |
| births | Número de nacimientos |
| deaths | Número de muertes |

# Cargar bibliotecas necesarias  
library(tidyverse)

# Leer los archivos CSV desde GitHub y suprimir el mensaje de tipo de columna  
yearly <- read\_csv('https://raw.githubusercontent.com/shukkkur/Analyzing-The-Discovery-of-Handwashing/main/datasets/yearly\_deaths\_by\_clinic.csv', show\_col\_types = FALSE)  
monthly <- read\_csv('https://raw.githubusercontent.com/shukkkur/Analyzing-The-Discovery-of-Handwashing/main/datasets/monthly\_deaths.csv', show\_col\_types = FALSE)  
  
# Recuperar la especificación completa de las columnas de los datos anuales  
yearly\_spec <- spec(yearly)  
print(yearly\_spec)

## cols(  
## year = col\_double(),  
## births = col\_double(),  
## deaths = col\_double(),  
## clinic = col\_character()  
## )

# Recuperar la especificación completa de las columnas de los datos mensuales  
monthly\_spec <- spec(monthly)  
print(monthly\_spec)

## cols(  
## date = col\_date(format = ""),  
## births = col\_double(),  
## deaths = col\_double()  
## )

# Inspeccionar los datos mensuales  
head(monthly)

## # A tibble: 6 × 3  
## date births deaths  
## <date> <dbl> <dbl>  
## 1 1841-01-01 254 37  
## 2 1841-02-01 239 18  
## 3 1841-03-01 277 12  
## 4 1841-04-01 255 4  
## 5 1841-05-01 255 2  
## 6 1841-06-01 200 10

summary(monthly)

## date births deaths   
## Min. :1841-01-01 Min. :190.0 Min. : 0.00   
## 1st Qu.:1843-02-08 1st Qu.:242.5 1st Qu.: 8.00   
## Median :1845-02-15 Median :264.0 Median :16.50   
## Mean :1845-02-11 Mean :267.0 Mean :22.47   
## 3rd Qu.:1847-02-22 3rd Qu.:292.8 3rd Qu.:36.75   
## Max. :1849-03-01 Max. :406.0 Max. :75.00

str(monthly)

## spc\_tbl\_ [98 × 3] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ date : Date[1:98], format: "1841-01-01" "1841-02-01" ...  
## $ births: num [1:98] 254 239 277 255 255 200 190 222 213 236 ...  
## $ deaths: num [1:98] 37 18 12 4 2 10 16 3 4 26 ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. date = col\_date(format = ""),  
## .. births = col\_double(),  
## .. deaths = col\_double()  
## .. )  
## - attr(\*, "problems")=<externalptr>

# Añadir columna proportion\_deaths a los datos anuales  
yearly <- yearly %>%  
 mutate(proportion\_deaths = deaths / births)  
  
# Añadir columna proportion\_deaths a los datos mensuales  
monthly <- monthly %>%  
 mutate(proportion\_deaths = deaths / births)  
  
# Verificar las nuevas columnas  
head(yearly)

## # A tibble: 6 × 5  
## year births deaths clinic proportion\_deaths  
## <dbl> <dbl> <dbl> <chr> <dbl>  
## 1 1841 3036 237 clinic 1 0.0781  
## 2 1842 3287 518 clinic 1 0.158   
## 3 1843 3060 274 clinic 1 0.0895  
## 4 1844 3157 260 clinic 1 0.0824  
## 5 1845 3492 241 clinic 1 0.0690  
## 6 1846 4010 459 clinic 1 0.114

head(monthly)

## # A tibble: 6 × 4  
## date births deaths proportion\_deaths  
## <date> <dbl> <dbl> <dbl>  
## 1 1841-01-01 254 37 0.146   
## 2 1841-02-01 239 18 0.0753   
## 3 1841-03-01 277 12 0.0433   
## 4 1841-04-01 255 4 0.0157   
## 5 1841-05-01 255 2 0.00784  
## 6 1841-06-01 200 10 0.05

# Añadir columna handwashing\_started a los datos mensuales  
monthly <- monthly %>%  
 mutate(handwashing\_started = date >= as.Date("1847-06-01"))  
  
# Calcular la media de la proporción de muertes antes y después del lavado de manos  
monthly\_summary <- monthly %>%  
 group\_by(handwashing\_started) %>%  
 summarise(mean\_proportion\_deaths = mean(proportion\_deaths))  
#presentar tabla  
monthly\_summary

## # A tibble: 2 × 2  
## handwashing\_started mean\_proportion\_deaths  
## <lgl> <dbl>  
## 1 FALSE 0.105   
## 2 TRUE 0.0211

## Gráficos

A continuación, se Muestra la programación de los gráficos:

# Graficar la proporción de muertes anual a lo largo del tiempo por clínica

ggplot(yearly, aes(x = year, y = proportion\_deaths, color = clinic)) +

geom\_line() +

geom\_point() +

labs(title = "Proporción de muertes anual por clínica",

x = "Año",

y = "Proporción de muertes",

color = "Clínica") +

theme\_minimal()

#Graficar la proporción de muertes mensual a lo largo del tiempo

ggplot(monthly, aes(x = date, y = proportion\_deaths)) +

geom\_line () +

geom\_point () +

labs(title = "Proporción de muertes mensual",

x = "Fecha",

y = "Proporción de muertes") +

theme\_minimal()

ggplot(monthly, aes(x = date, y = proportion\_deaths, color = handwashing\_started)) +

geom\_line () +

geom\_point () +

labs(title = "Proporción de muertes mensual con lavado de manos",

x = "Fecha",

y = "Proporción de muertes",

color = "Lavado de manos iniciado") +

theme\_minimal()

