

# Concurso Rladies

Maria Dermit

25/07/2020

## Florence Nightingale

El objetivo de este analysis es conmemorar los 200 años del nacimiento de Florence Nightingale una enfermera y estadística británica pionera en visualización de datos. Para ello se propone analizar el conjunto de datos que ella utilizó para determinar los factores asociados a la alta mortalidad.

Florence Nightingale fue la primera mujer que utilizando las conclusiones de sus análisis logro convencer al gobierno británico para que impusiera medidas higiénicas que llevaron a reducir significativamente las causas de mortalidad del ejército británico durante la guerra de Crimea.

```
library(tidyverse) #para realizar los pasos de DS (data science)
library(here) #para organizar inputs y outputs
library(readxl) #para leer la hoja de excel
library(knitr) # para ver tablas con la funcion kable
library(magrittr) #para hacer pipe assignment
library(scales) #para por cientos.
library(showtext) #para fuentes de letra especificas
library(ggplot2) #para colores pomológicos
library(gridExtra) #para poner titulos con colores
```

Voy a crear un proyecto de RStudio donde voy a ordenar mis carpetas en:

1. Data
2. Plots
3. Rscript

Voy a utilizar la library `here` que hace muy facil leer y poner los outputs en order.

Voy a escribir mi código en un markdown que genera un documento pdf de facil lectura.

## Empezamos leyendo los datos

```
datos_florence_raw<-read_excel(here("data/datos_florence.xlsx"))
```

Parece que los datos tienen un encabezado adicional que clasifica muertes y tasas de mortalidad (1000‰) en 3 categorías: enfermedades infecciosas (relacionadas con infección, prevenibles o mitigables), heridas y lesiones y otras causas.

Tenemos que hacer un limpieza inicial para poner los encabezados de forma correcta:

```
datos_florence_raw<-read_excel(here("data/datos_florence.xlsx"), skip = 1)
names(datos_florence_raw)[3:5]<-paste0(names(datos_florence_raw)[3:5],"_", "death")
names(datos_florence_raw)[6:8]<-paste0(names(datos_florence_raw)[6:8],"_", "rate")
```

En realidad solo estamos interesados en las columnas de tasas.

Vamos a limpiar los datos un poco mas:

```
datos_florence_clean<-datos_florence_raw
datos_florence_clean<-datos_florence_clean %>%
  rename(fecha=Month,
         army='Average size of army',
         zymotic_death='Zymotic diseases...3_death',
         wounds_death='Wounds & injuries...4_death',
         other_death='All other causes...5_death',
         zymotic_rate='Zymotic diseases...6_rate',
         wounds_rate='Wounds & injuries...7_rate',
         other_rate='All other causes...8_rate') %>%
  select(-zymotic_death,-wounds_death,-other_death) %>%
  mutate_if(is.character,as.factor)

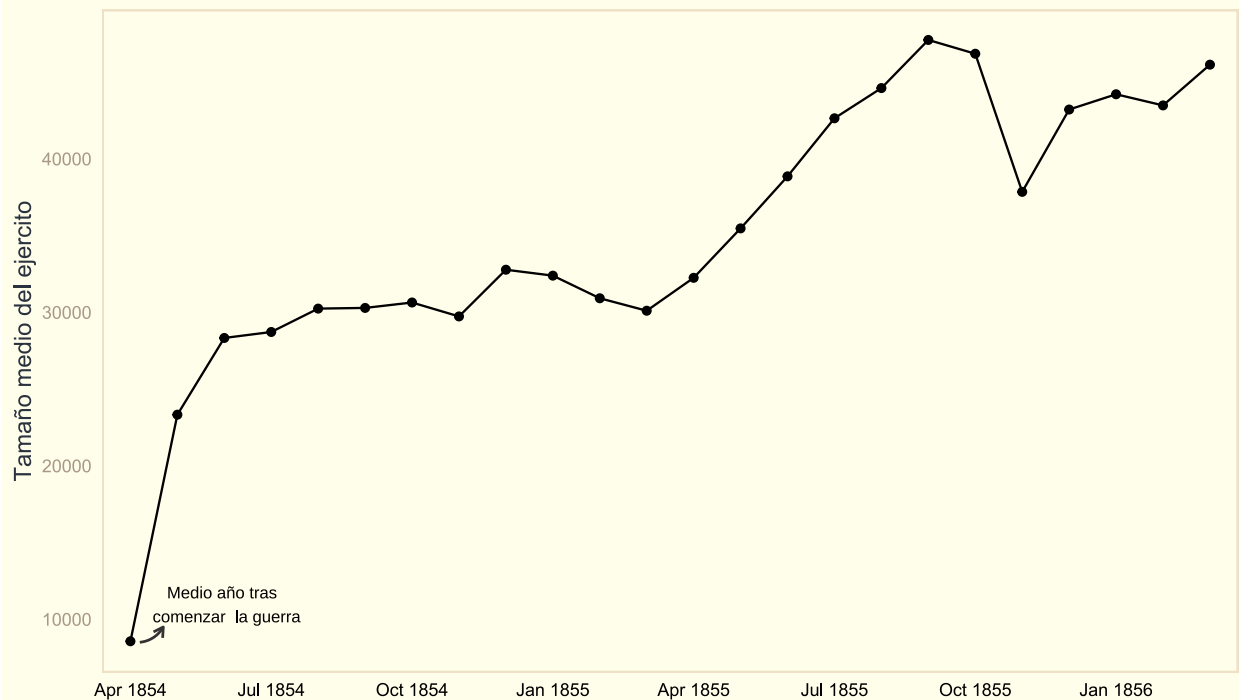
datos_florence_clean_final<-datos_florence_clean %>%
mutate(fecha = str_replace(fecha, '_ ', " "))
```

Ahora ya tenemos los datos limpios.

## EDA

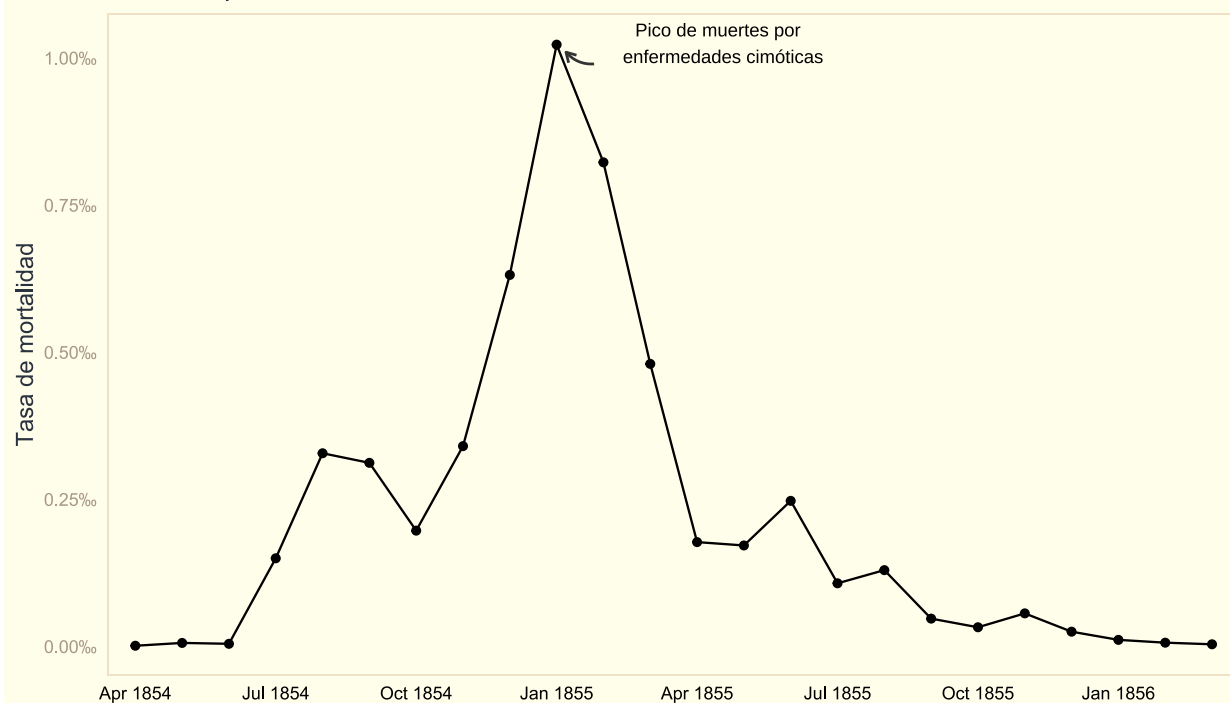
Inicialmente hacemos unos primeros pasos de EDA (exploratory data analysis)

### Tamaño medio del ejercito



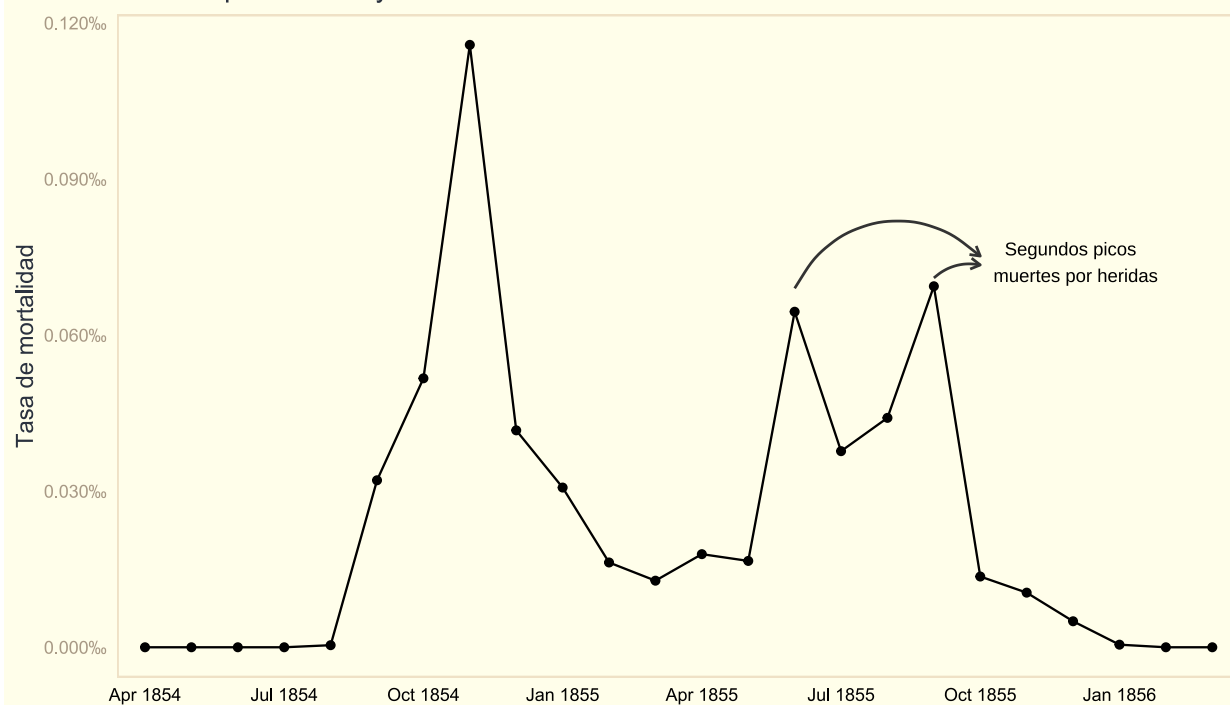
El aumento del tamaño del ejército aumentó hasta cuatro veces con el avance de la guerra.

### Tasa de muerte por enfermedades cimotoicas



Las muertes por enfermedades cimoticas disminuyeron considerablemente despues del comienzo del segundo año de la guerra.

### Tasa de muerte por heridas y lesiones

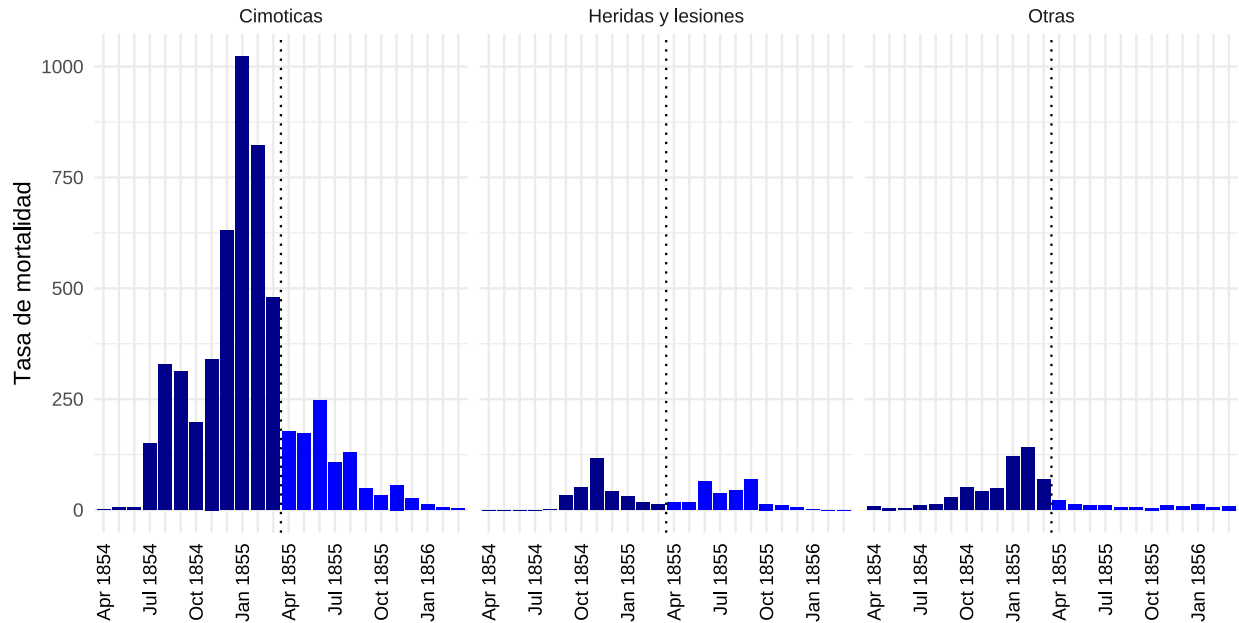


Las muertes por heridas y lesiones tuvieron un repunte en la 2/3 de la guerra, potencialmente por personas que desgraciadamente terminaron muriendo como consecuencia de las heridas de guerra.

La tasa de mortalidad por enfermedades de causa cimotica contribuia mucho mas al elevado numero de muertes. Vamos a visualizar esto mejor:

## Tasa de muerte por enfermedad

Enfermedades cimoticas, heridas y lesiones y otras enfermedades



Momento en el que se introducen las medidas de limpieza y mejoraras de ventilacion

Sin duda las muertes por enfermedades cimoticas eran las que contribuian mas al numero total de muertes. Esto resulta en alguna medida sorprendente ya que estas enfermedades eran clasificadas como prevenibles o mitigables. Florence entendio que esa tasa de muertes se podia reducir, lo cual le llevo a desarrollar el Diagrama de la Rosa para representar estos datos. Fueron estos graficos los que posteriormente ella enseño al gobierno britanico y consiguio que en marzo de 1855 se tomaran medidas de higiene que llevaron a reducir la tasa de muerte por enfermedades infecciosas.

Estos graficos se encuentran en la coleccion “Florence Nightingale Museum Collection” en el hospital de St Thomas de Londres

Vamos a representar estos datos con R:

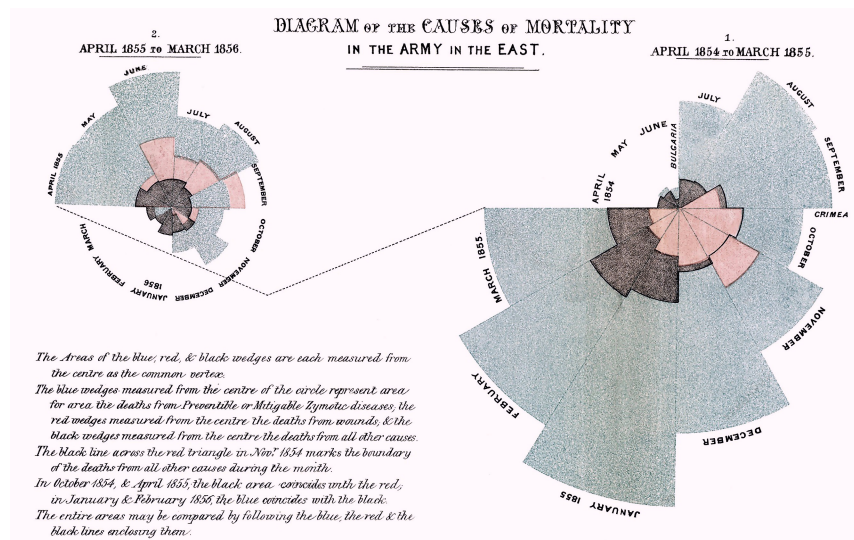
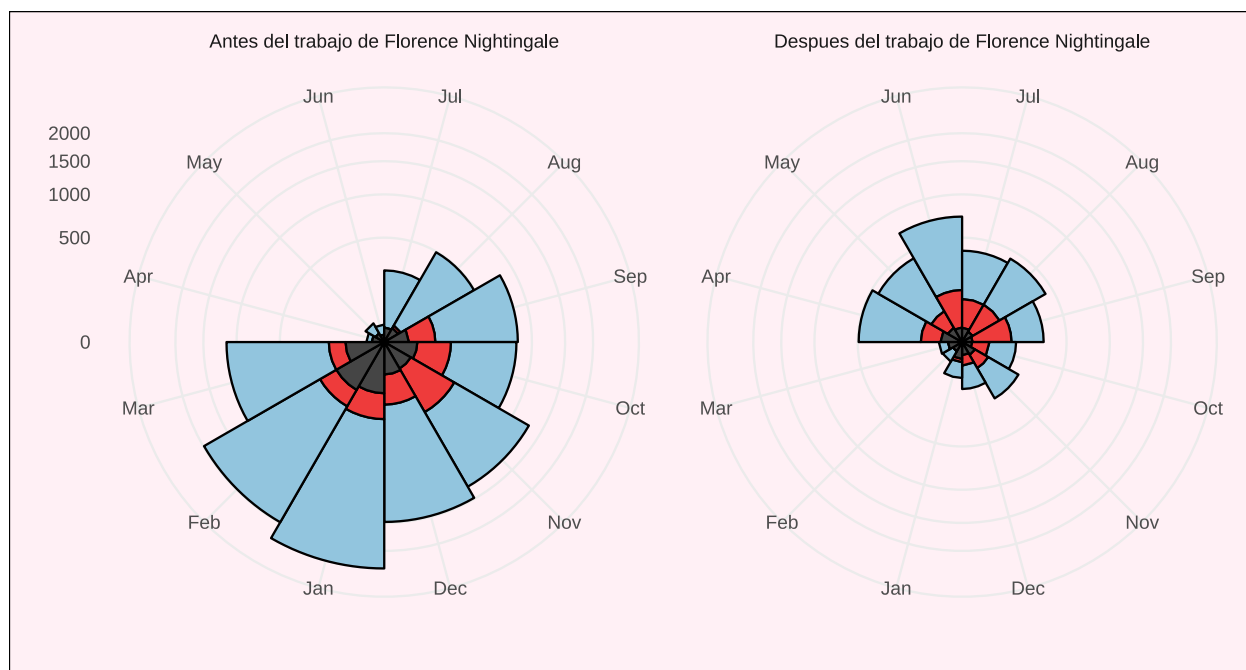


Figure 1: Causas de la mortalidad en la armada en el este por Florence Nightingale, (1820 – 1910)

Tasa de muerte por enfermedad: **cimoticas** , **heridas y lesiones** y otras enfermedades



```
datos_modeling %>% filter(Regime=="Before") %>% select(-Regime) %>% map_dbl(sum,na.rm = TRUE)/
datos_modeling %>% filter(Regime=="After") %>% select(-Regime) %>% map_dbl(sum,na.rm = TRUE)
```

```
##      army zymotic_rate  wounds_rate  other_rate
##    0.6676587    4.2211746    1.0775554    4.6704645
```

La media de tasa de muertes por enfermedades cimoticas se redujo en mas de cuatro veces, mientras que la media de tasa de muertes por heridas se mantuvo estable.

## Conclusiones

Florence Nightingale fue la primera teorica en el campo de la enfermeria observando el elevado numero de muertes a causa de infecciones en los hospitales durante la guerra de Crimea. Su trabajo de analisis estadistico tuvo consecuencias directas mejorando la practica de la enfermeria y estableciendo gran parte de las bases de la enfermeria moderna.