

# Evaluacion del comportamiento de una vacuna en centros de mar.

Gustavo Bustos Yáñez

30 October 2021

## Exploración de datos.

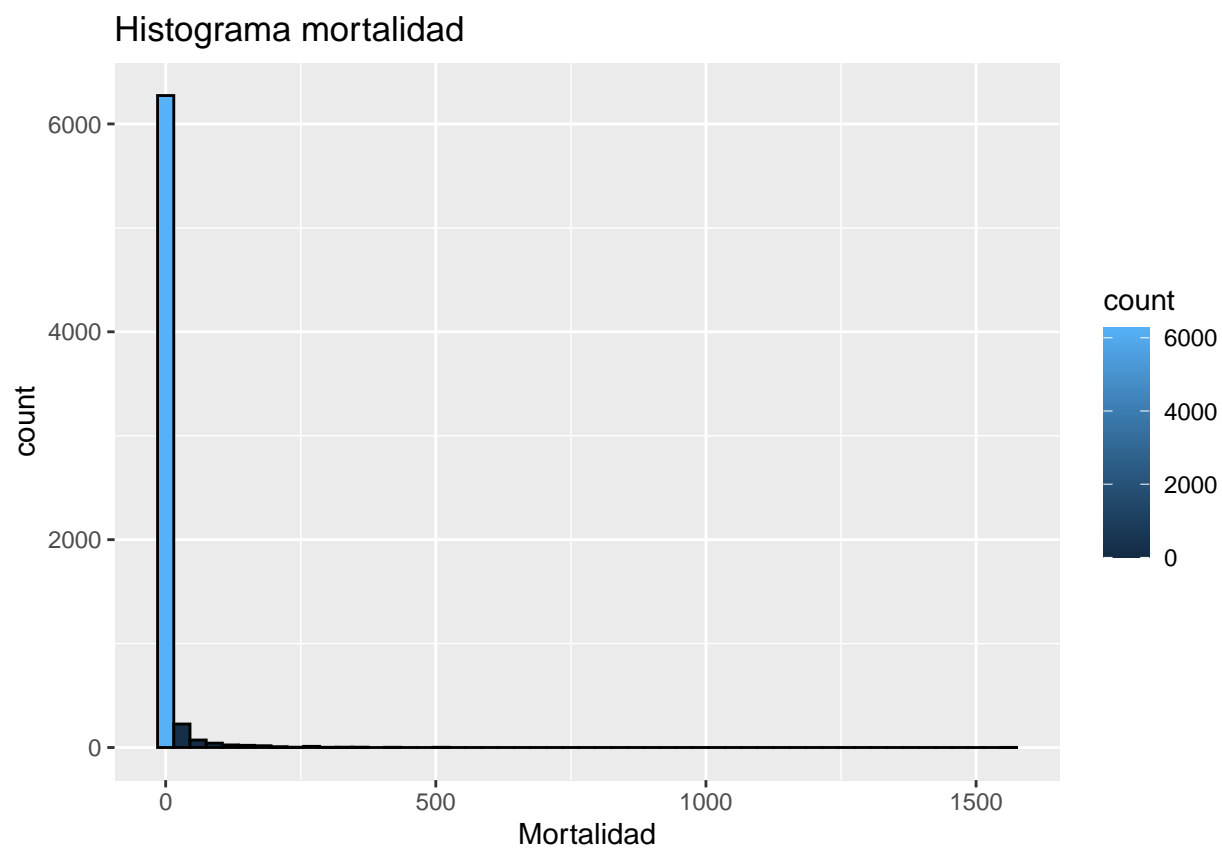
EL set de datos se compone de un data frame conformado por 6 columnas y 6.711 observaciones. De los cuales, 2 columnas corresponden a los tratamientos a analizar, identificados como Tratamiento y Vacuna.

```
## tibble [6,711 x 6] (S3: tbl_df/tbl/data.frame)
##   $ Unidad      : num [1:6711] 3 5 2 3 4 5 2 3 4 5 ...
##   $ Fecha       : Factor w/ 493 levels "2019-12-11","2019-12-12",...: 1 1 2 2 2 2 3 3 3 3 ...
##   $ Origen      : chr [1:6711] "P1" "P1" "P1" "P1" ...
##   $ Tratamiento: Factor w/ 2 levels "No","Si": 2 2 2 2 2 2 2 2 2 2 ...
##   $ Vacuna      : Factor w/ 2 levels "No","Si": 1 1 2 1 2 1 2 1 2 1 ...
##   $ Mortalidad : num [1:6711] 0 0 0 0 0 0 0 0 0 0 ...
```

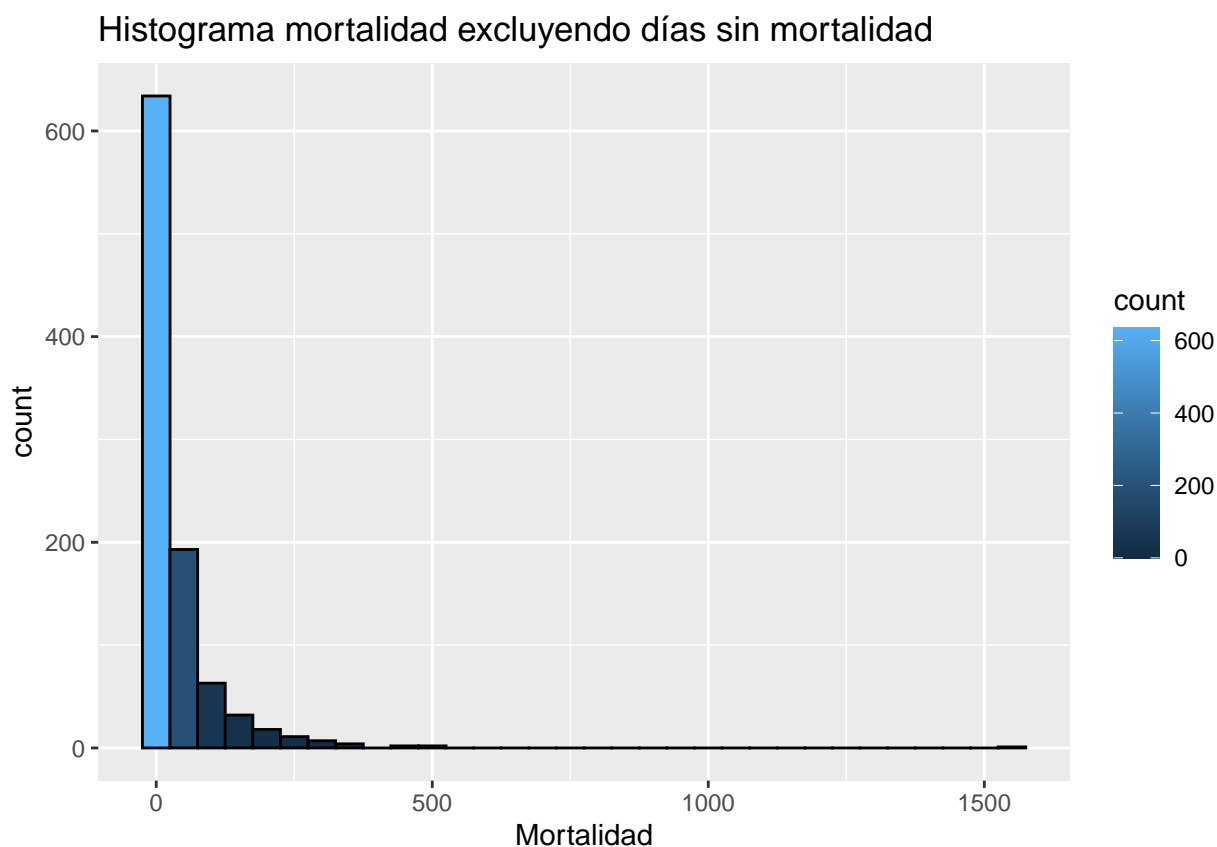
Table: Estructura del set de datos.

|| || || ||

## Histograma



Para ayudar a la visualización del histograma se eliminan los datos en 0, que representan días sin extracción de mortalidad con características que indican la presencia de la patología en los ejemplares extraídos en la mortalidad.



El histograma demuestra que la distribución de los datos no evidencia que sea del tipo normal.

#### Tabla de frecuencias:

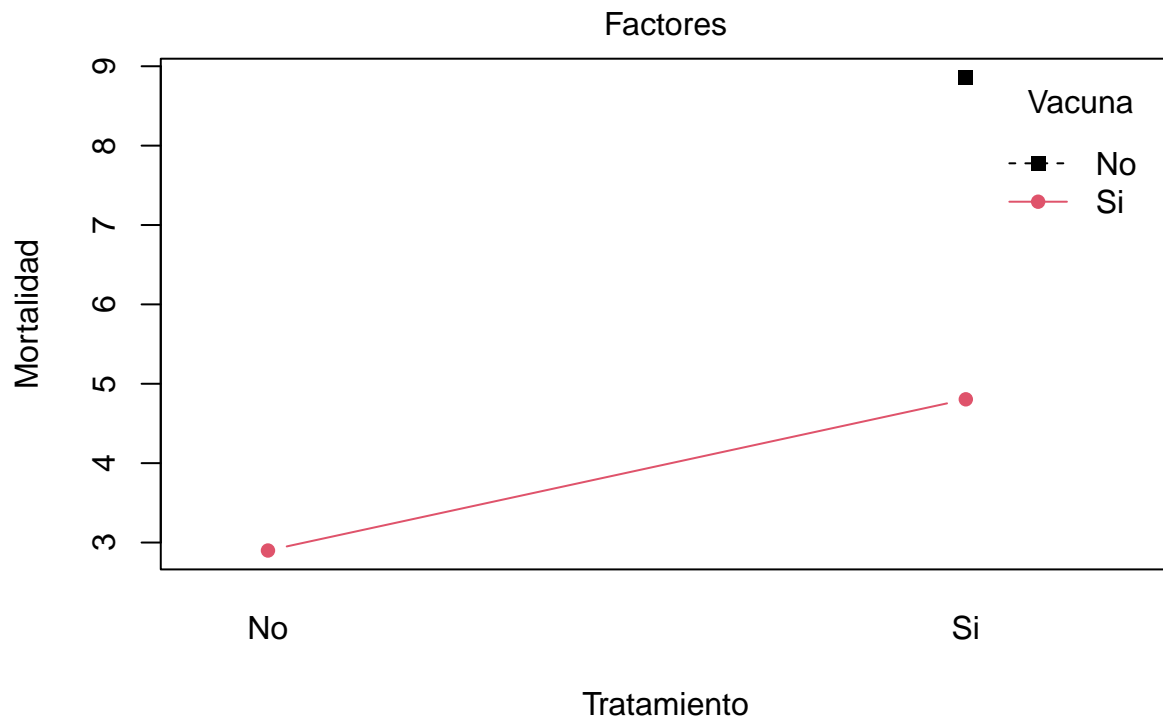
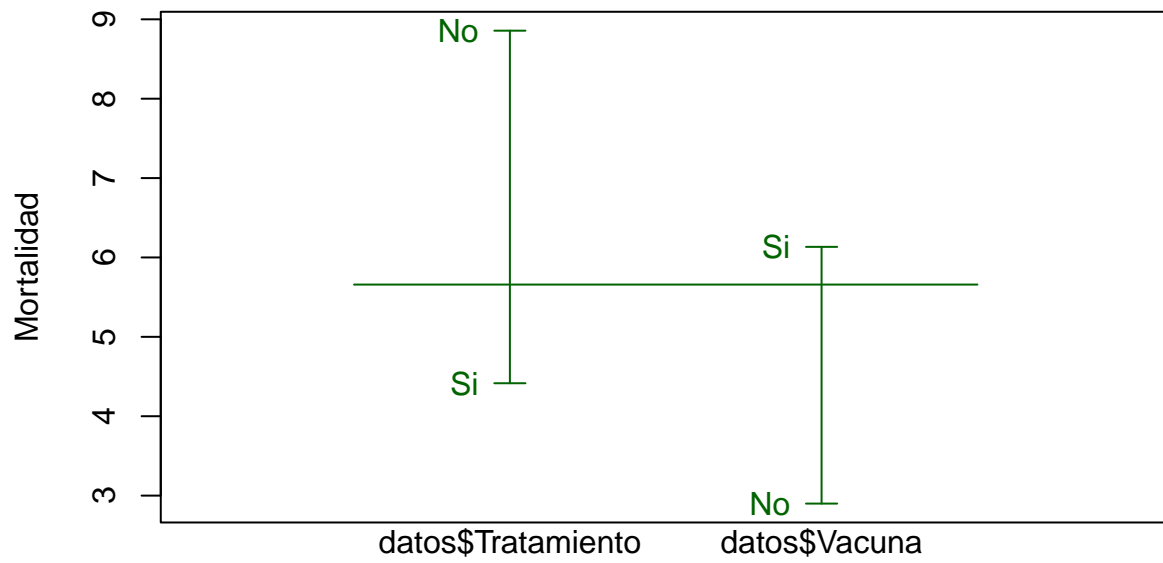
Table 1: Frecuencias variables Tratamiento y Vacuna.

	No	Si
No	0	1878
Si	986	3847

La tabla de frecuencias demuestra que los datos no se encuentran balanceados, esto puede deberse a que existen unidades con más tiempo de producción que otras y menor cantidad de unidades solamente con el tratamiento.

**Establece relación entre variables cuantitativas y factores usando gráficas de correlación, boxplot, interacción o de tamaño de los efectos (4 puntos).**

```
## geom_point: na.rm = FALSE
## stat_identity: na.rm = FALSE
## position_identity
```



Identifica si existen errores, datos faltantes o valores at picos (4 puntos).

Resume los datos usando tablas y estad stica descriptiva (4 puntos).