Humpback Whales and Ship Noise

Fabian Blasch

04/29/2022

1 Data and Desciptive Statistics

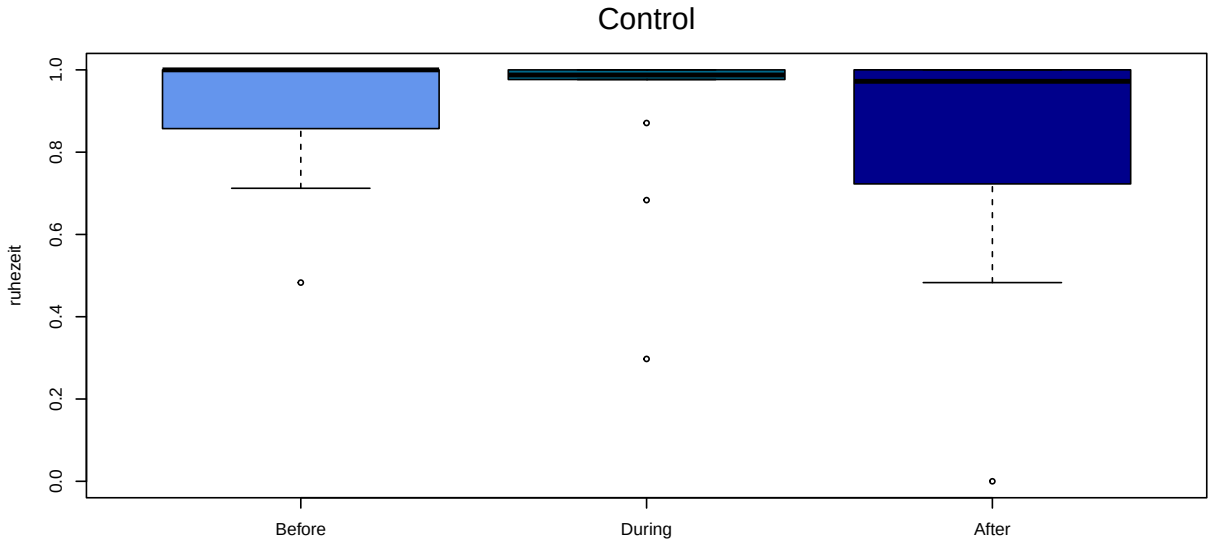
```
# import data (this dataset is unfortunatelly not public)
openxlsx::read.xlsx("./../Data/Humpback_Whales_Data.xlsx") -> dat_whale

# fist a quick look at the missing values in the data
sapply(dat_whale,  \(x) sum(is.na(x))) |> knitr::kable(col.names = "NAs")
```

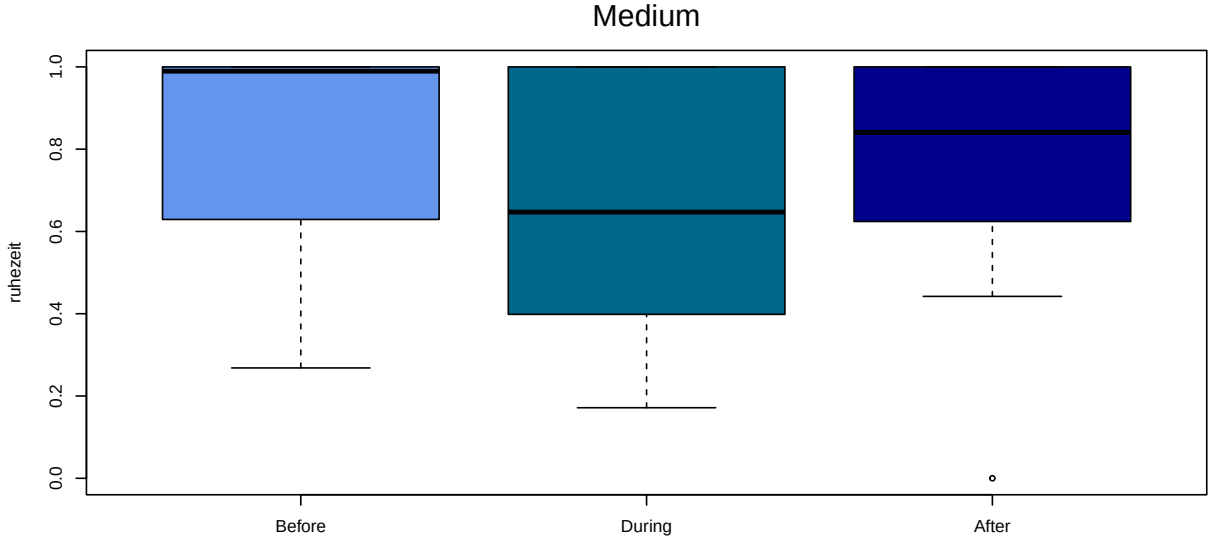
	NAs
Individuum	0
Treatment	0
Szenario	0
ruhezeit	0
speed	5
Atem	0

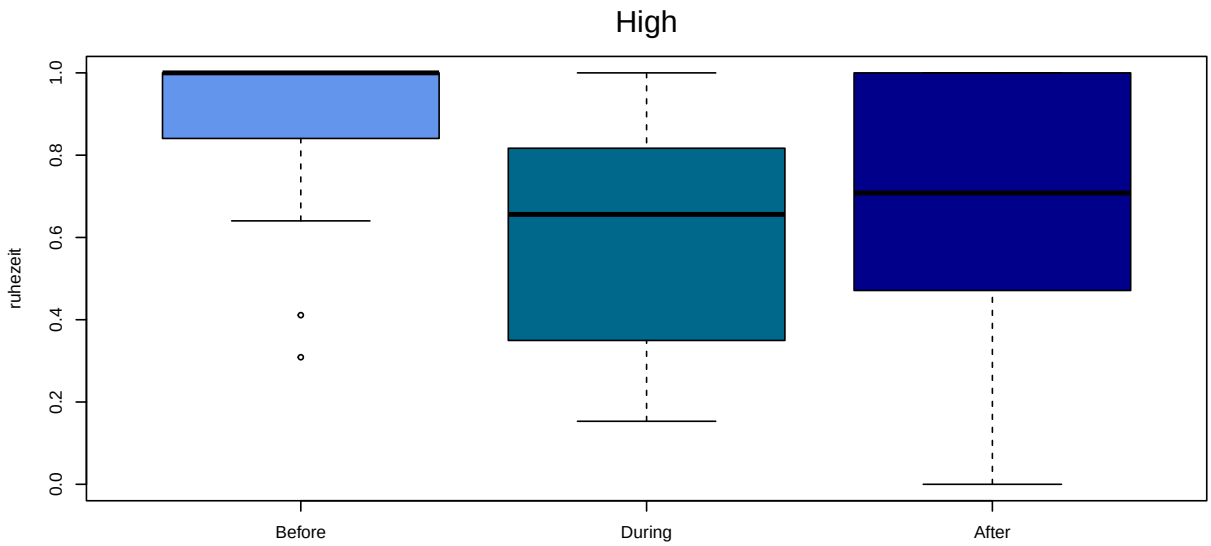
```
# harmonize names
colnames(dat_whale) <- tolower(colnames(dat_whale))

# to numeric
lapply(dat_whale[, c("ruhezeit", "speed", "atem")], as.numeric) -> dat_whale[, c("ruhezeit", "speed", "a

# to factor
lapply(dat_whale[ ,!(colnames(dat_whale) %in%  c("ruhezeit", "speed", "atem"))],
       as.factor) -> dat_whale[ ,!(colnames(dat_whale) %in%  c("ruhezeit", "speed", "atem"))]

# relevel
factor(dat_whale[, "szenario"],
       levels = c("Before", "During", "After")) -> dat_whale[, "szenario"]
factor(dat_whale[, "treatment"], c("Control", "Medium", "High")) -> dat_whale[, "treatment"]

# add log
within(dat_whale,{
   logspeed <- log(speed)
   logatem <- log(atem)
}) -> dat_whale


# frist split into different intensities
dat_whale_intens <- split(dat_whale, dat_whale[, "treatment"])
```

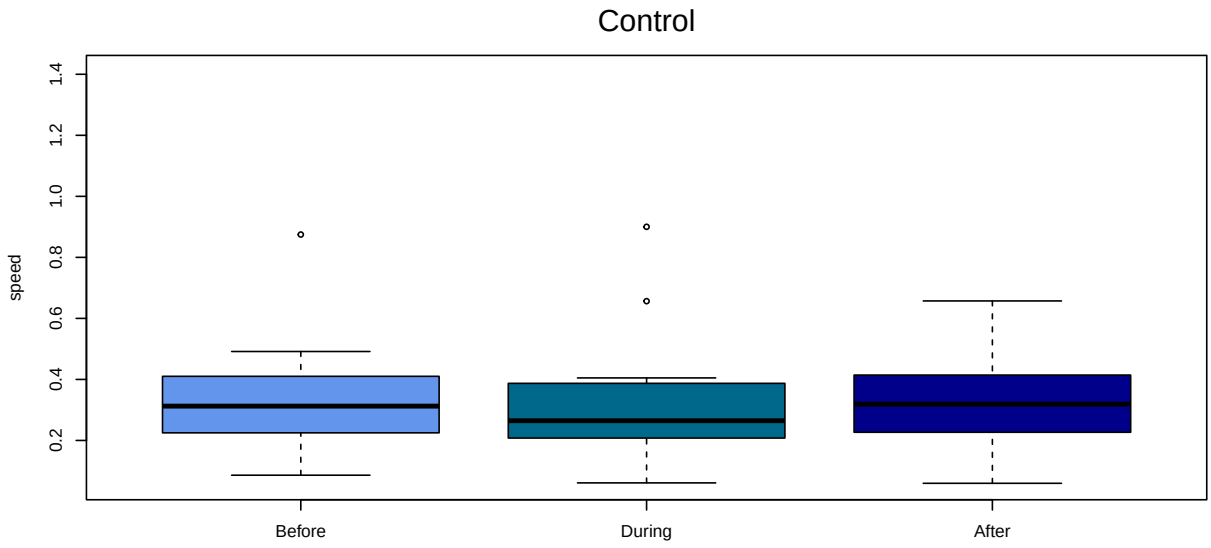
```
# build formulas
formulae <- paste(c("ruhezeit", "speed", "atem"), "~", "szenario")

# max and min for plot y-axis
sapply(c(min, max), \(x){

   sapply(dat_whale[, c("ruhezeit", "speed", "atem")], \(y) x(y, na.rm = TRUE))


}) -> ylims

# over szenarios
```

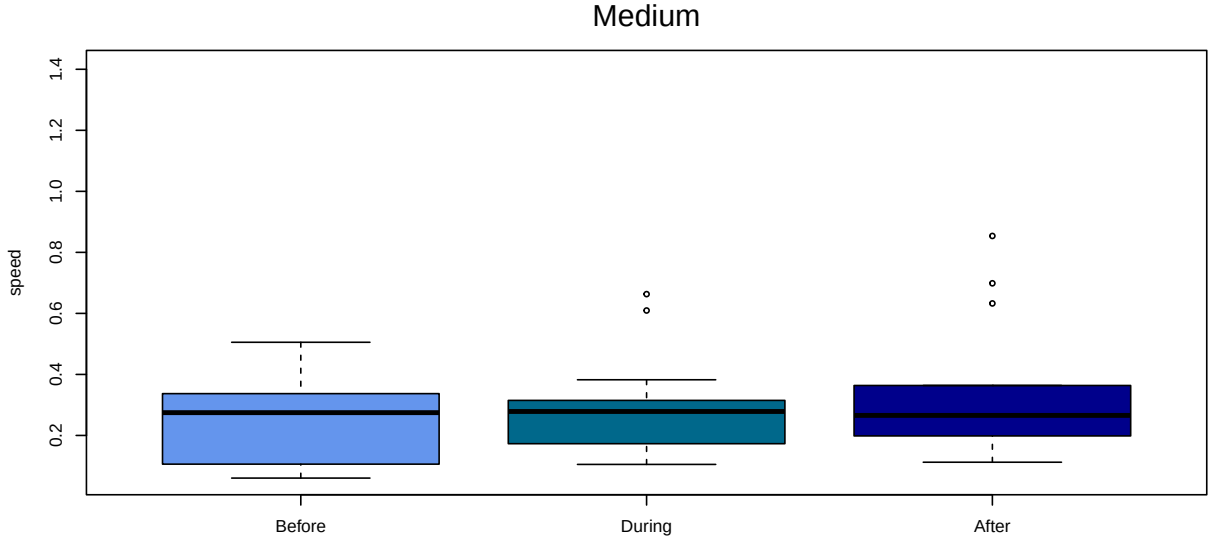
```
invis.Map(\(y, nom, lims){

   # safe for presentation
   # pdf(paste0("./../Presentation/", nom, ".pdf"))

   # align
   par(mfrow = c(3, 1), mar = c(2, 4, 4, 2) + 0.1)

   # over treatment
   invis.Map(\(x, nom){

      # boxplots
      boxplot(as.formula(y), data = x,
              col = c("cornflowerblue", "deepskyblue4", "darkblue"),
              ylim = c(lims[1], lims[2]))

      # add label
      mtext(nom, side = 3, line = 1, cex = 1.2)

   }, dat_whale_intens, names(dat_whale_intens))

   # close graph. device
   # dev.off()

}, formulae, c("resting", "speed", "respatory"), ylims |> t() |> as.data.frame())
```

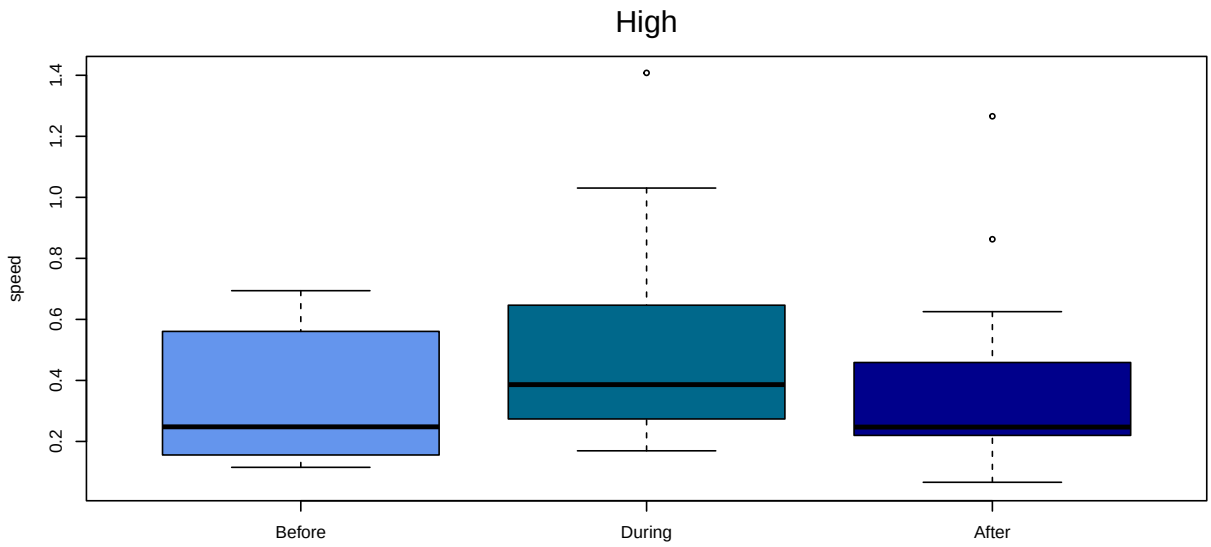

Control
ruhezeit
0.0
0.2
0.4
0.6
0.8
1.0
Before
During
After

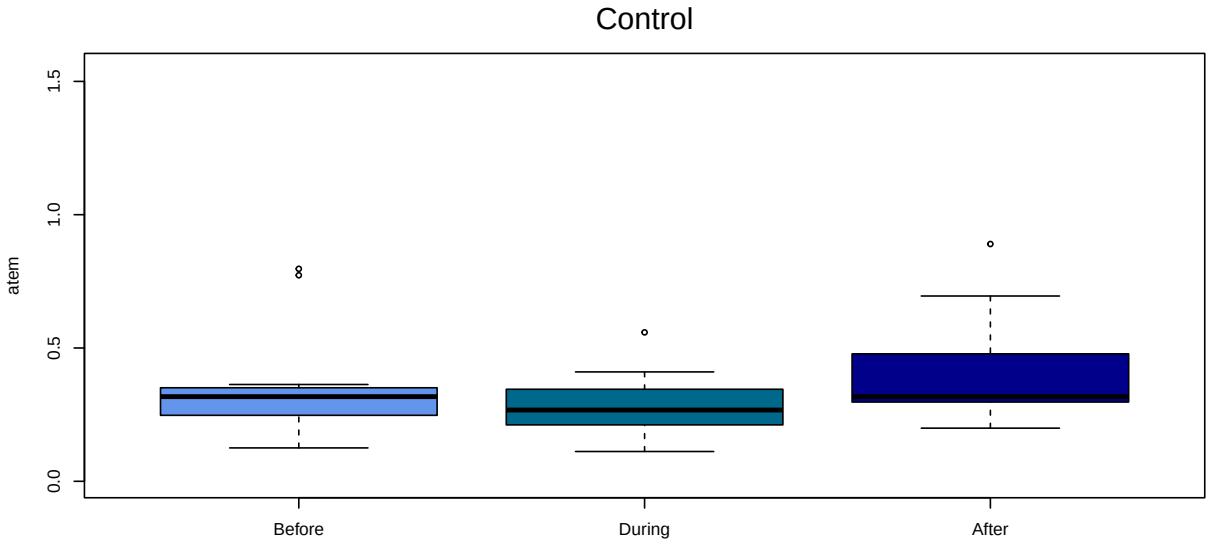

Medium
ruhezeit
0.0
0.2
0.4
0.6
0.8
1.0
Before
During
After

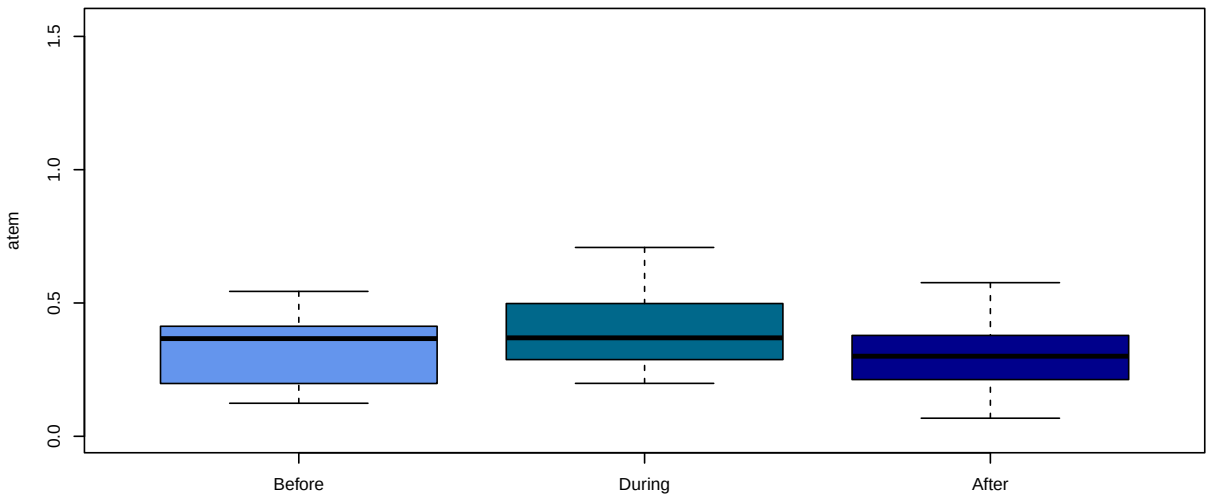

High
ruhezeit
0.0
0.2
0.4
0.6
0.8
1.0
Before
During
After

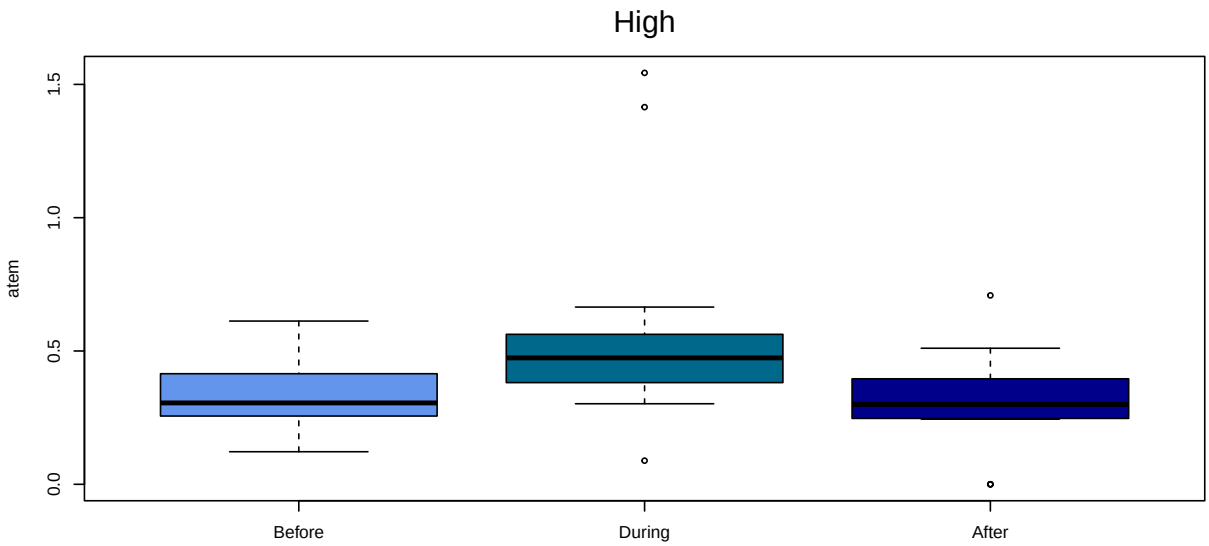

Control
speed
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Before
During
After


Medium
speed
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Before
During
After


High
speed
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Before
During
After


Control
atem
0.0
0.5
1.0
1.5
Before
During
After


Medium
atem
0.0
0.5
1.0
1.5
Before
During
After


High
atem
0.0
0.5
1.0
1.5
Before
During
After

2 Speed and Breath

The basic idea behind looking at a scatterplot of the speed of the animal and the breath frequency for different groups is, that we would expect the breathe frequency to increase with speed, but if the whale is additionally stressed because of external influence, then the breathe should increase above the normally expected level of breathing frequency for a given speed level.

```
# scatter 1
ggplot(dat_whale, aes(x = speed, y = atem, color = treatment,
                      shape = szenario, linetype = szenario)) +
  geom_point(size = 2) +
  geom_smooth(method = "lm", se = FALSE) +
  theme_bw() +
  theme(legend.position = "none") -> plot1


# scatter 2
ggplot(dat_whale, aes(x = logspeed, y = logatem, color = treatment,
                      shape = szenario, linetype = szenario)) +
  geom_point(size = 2) +
  geom_smooth(method = "lm", se = FALSE) +
  theme_bw() -> plot2

# display side by side
plot1 + plot2
```