

Humpback Whales and Ship Noise

Fabian Blasch

03/30/2022



1 Data and Descriptive Statistics

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

# first 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", "atem")]

# 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

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

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

# over szenarios
invisible(Map(\(y){

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

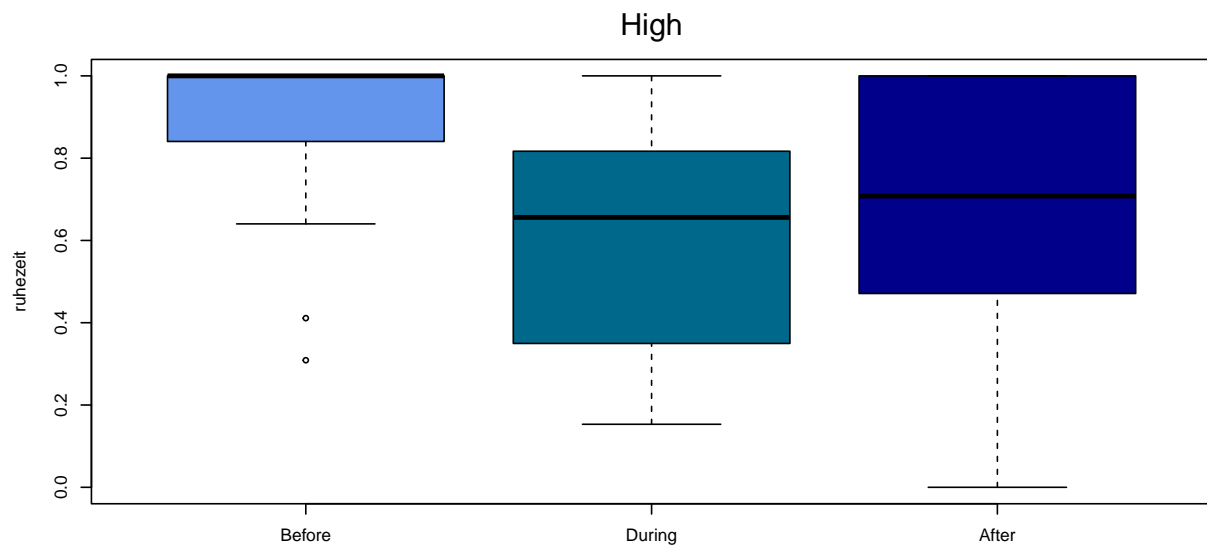
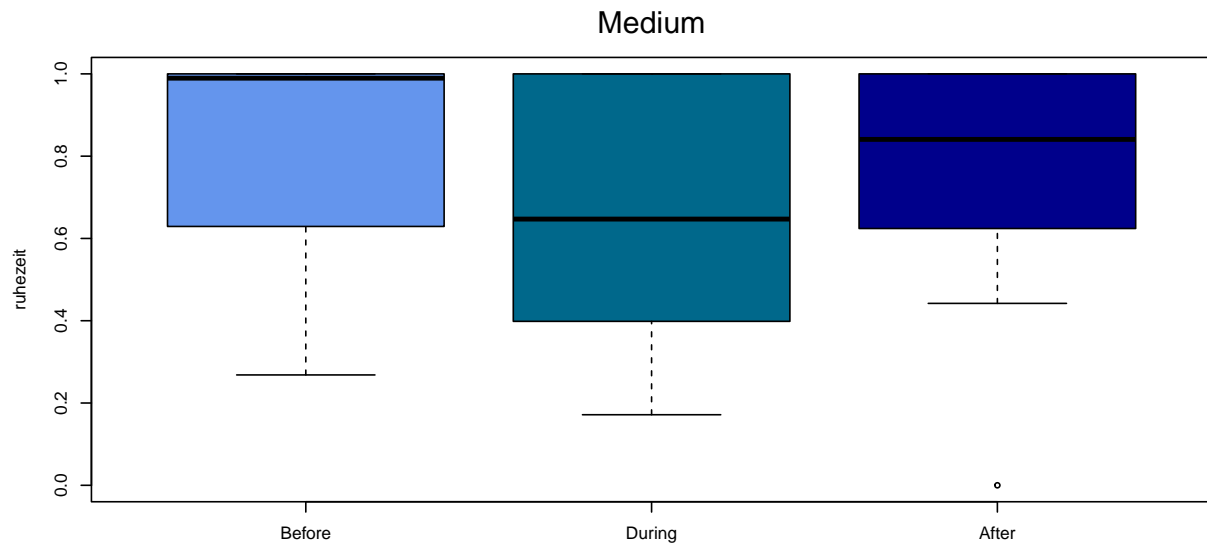
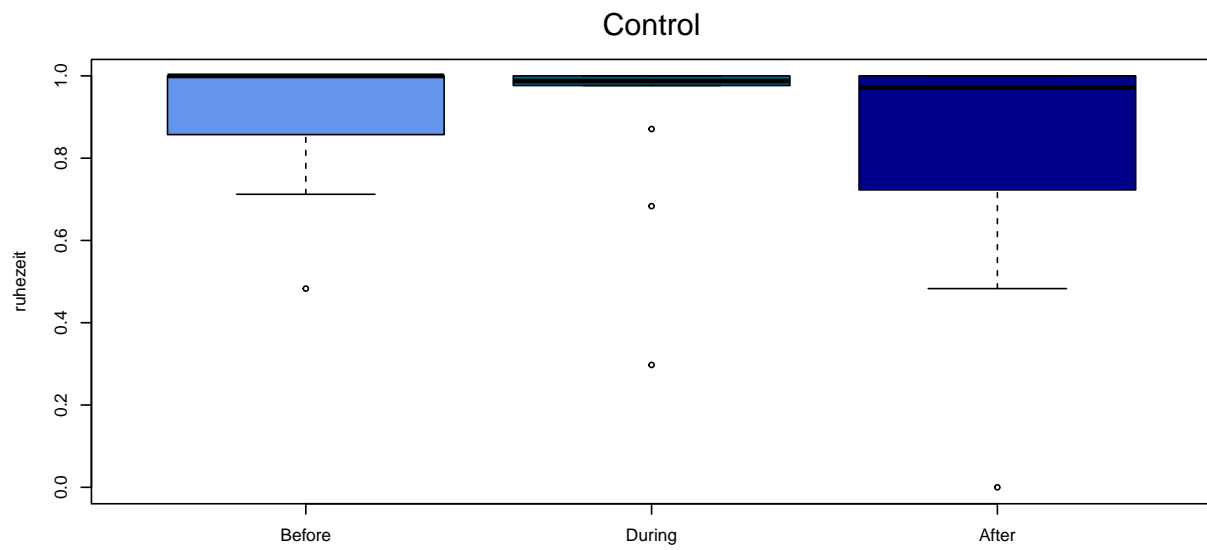
  # over treatment
  invisible(Map(\(x, nom){
```

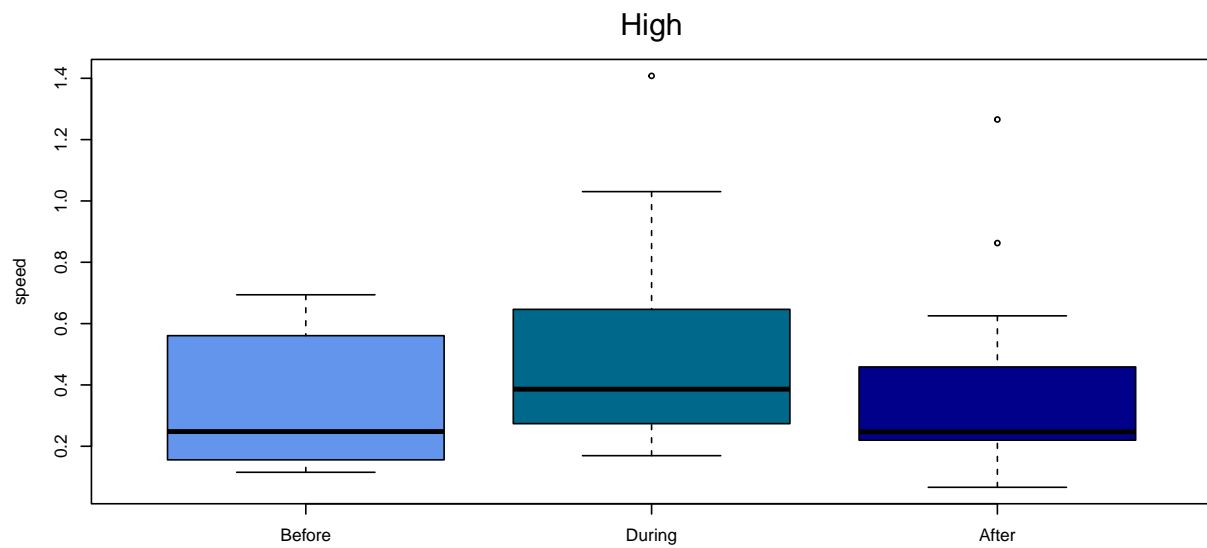
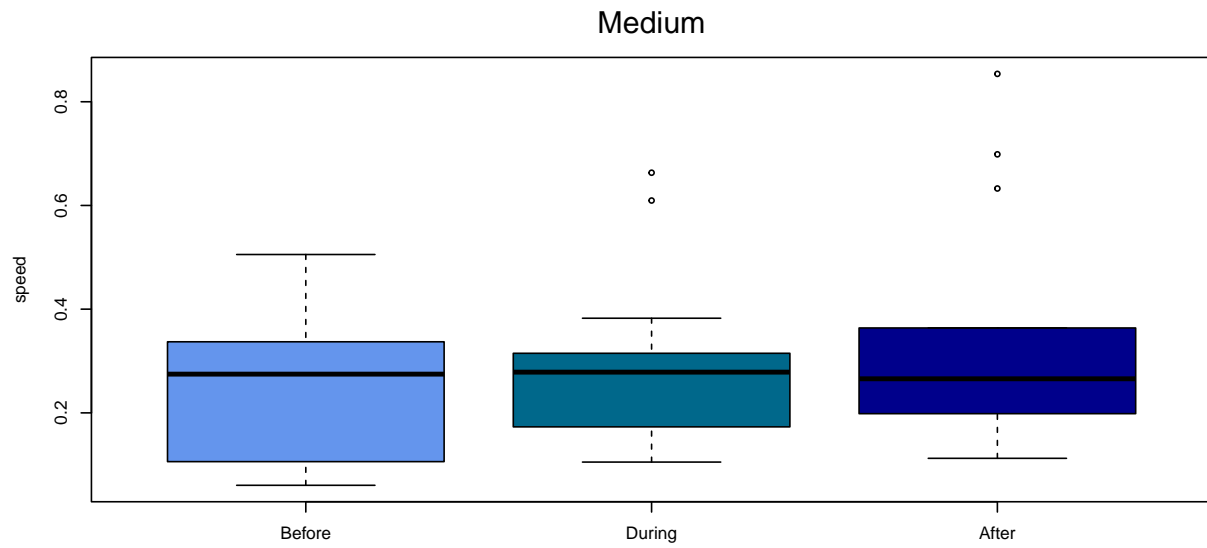
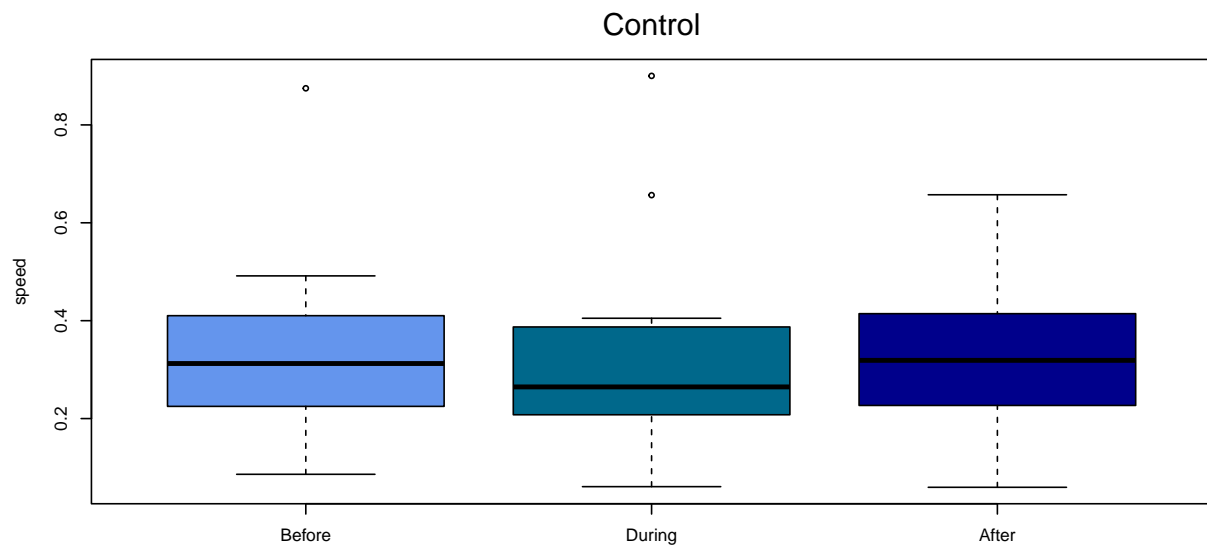
```
# boxplots
boxplot(as.formula(y), data = x,
        col = c("cornflowerblue", "deepskyblue4", "darkblue"))

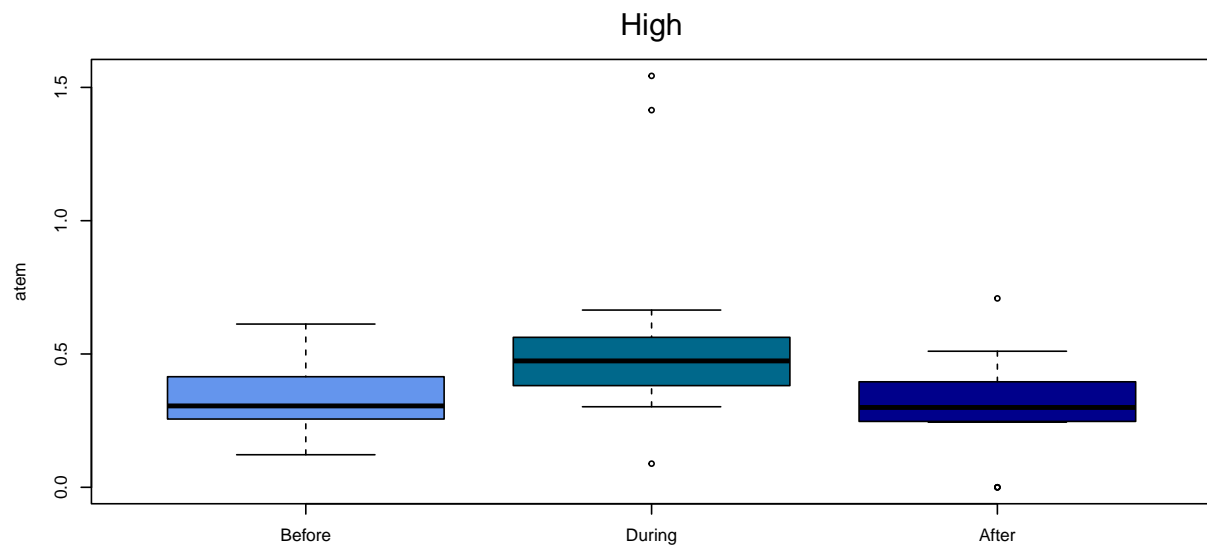
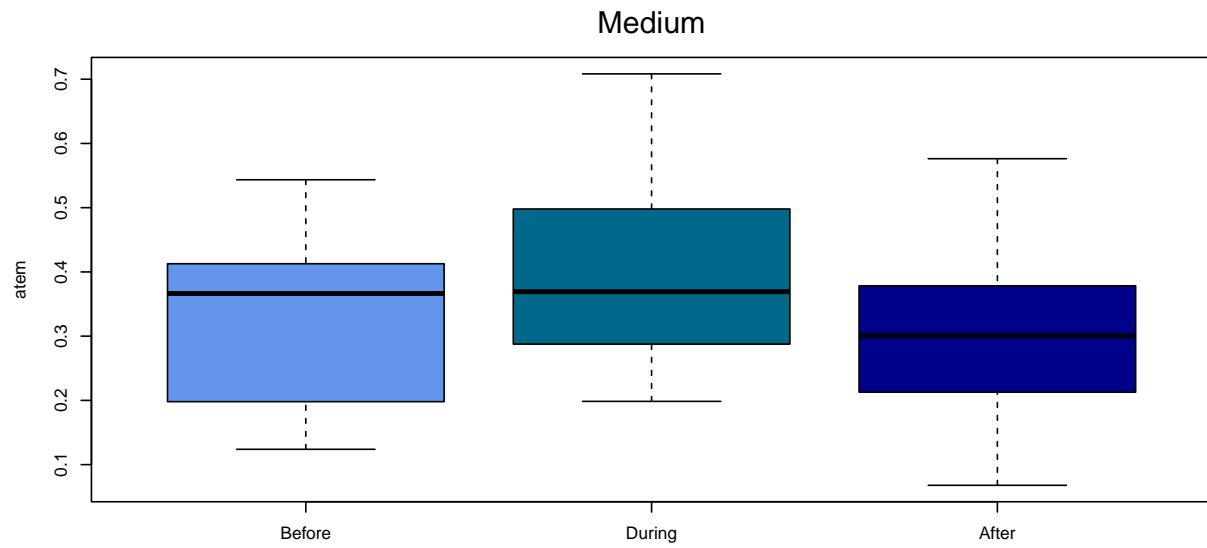
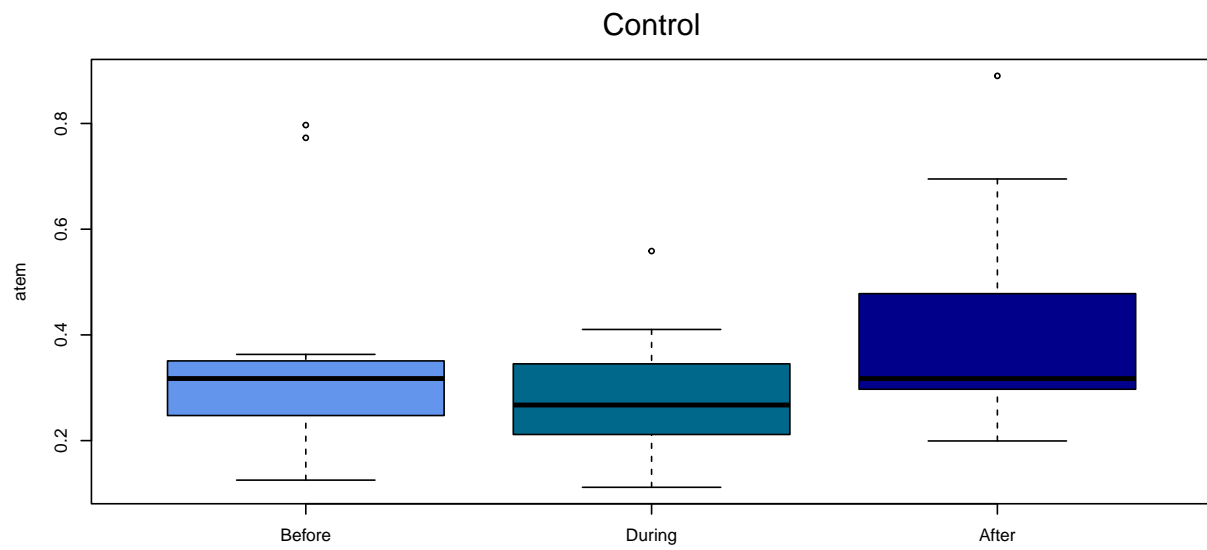
# add label
mtext(nom, side = 3, line = 1, cex = 1.2)

}, dat_whale_intens, names(dat_whale_intens))

}, formulae)
```







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 breather 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)) +
  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)) +
  geom_point(size = 2) +
  geom_smooth(method = "lm", se = FALSE) +
  theme_bw() -> plot2

# display side by side
plot1 + plot2
```

