

# STAT 9210 Final Project

Pragyat (Prag) Agrawal and Sahil Kulkarni

2023-04-25

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Study</b>	<b>1</b>
2.1	Participants . . . . .	1
2.2	Treatment . . . . .	2
2.3	Data Visualization . . . . .	2
2.4	Mixed-effect Model . . . . .	5
2.5	Matching . . . . .	14
2.6	Footnotes and Sources . . . . .	22

## 1 Introduction

Breathing practices are vital for humans as they influence how we feel and shape human behavior. But scientific evidence on breathing and its effect on mind-body connection have been limited. Structured breathing practices have their own share of benefits as they help with enhancing mood and reduce physiological arousal in humans. There have been several studies which have established this. We seek to study this same effect and replicate its result by applying the concepts we learned in class to find if long breathing exercises can help people reach mindfulness and reduce stress and anxiety.

## 2 Study

We looked into a study by Cell Reports Medicine which according to their website cell.com is “premium open-access journal from Cell Press publishing cutting-edge research in translational and clinical biomedical sciences that inform and influence human health and medicine.”

### 2.1 Participants

Throughout the study, 140 potential targets were consented but 114 were invited to participate in the study. There was some attrition because of pandemic related reasons or loss of contact with the target. According to the study, The general ease of following instructions and performing the interventions and subjective experience of the interventions were assessed by an optional debriefing questionnaire at the end of the study.

## 2.2 Treatment

The experiment put the 114 participants to a breathing and meditation protocol and measured the affect of that on their state anxiety. Specifically, it measured positive affect (Positive And Negative Affect Schedule (PANAS)) and state anxiety (State-Trait Anxiety Inventory (STAI)) scores on each participant before and after the experiment. Mindfulness meditation and breathwork groups both experienced an increase in daily positive affect.

## 2.3 Data Visualization

```
attach(filename)
filename$`Sleep Score` <- as.numeric(filename$`Sleep Score`)
```

```
## Warning: NAs introduced by coercion
```

```
filename$`Hours of Sleep` <- as.numeric(filename$`Hours of Sleep`)
```

```
## Warning: NAs introduced by coercion
```

```
# create new variables with different names
```

```
my_hours <- filename$`Hours of Sleep`
```

```
my_score <- filename$`Sleep Score`
```

```
# use the original variables with the `::` operator
```

```
ggplot(filename, aes(x = my_hours, y = my_score)) +
```

```
  geom_point() +
```

```
  geom_smooth(method = lm, color = 'red', se = F) +
```

```
  ylim(NA, 100) +
```

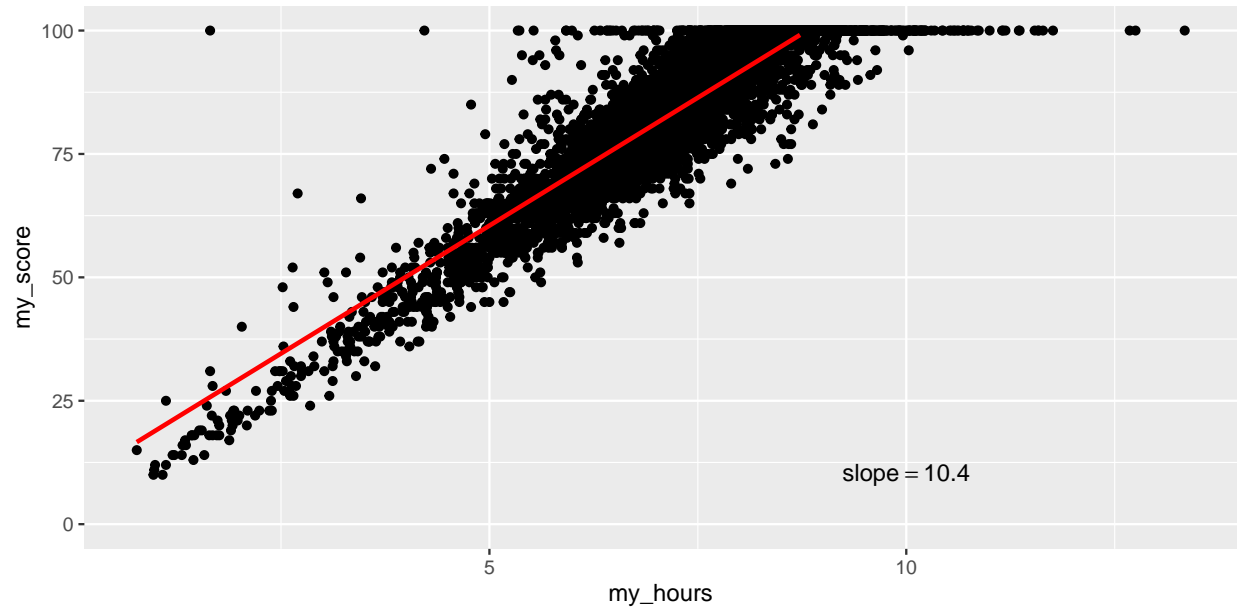
```
  annotate("text", x=10, y=10, label=paste0("slope==", coef(lm(filename$`Sleep Score`~filename$`Hours of S
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

```
## Warning: Removed 597 rows containing non-finite values ('stat_smooth()').
```

```
## Warning: Removed 597 rows containing missing values ('geom_point()').
```

```
## Warning: Removed 29 rows containing missing values ('geom_smooth()').
```



This graph shows that sleep score is positively related to the number of hours a person sleeps. This means a higher number of hours sleeping, corresponds to a higher sleep score. This makes sense as the quality of sleep increases with the quantity of sleep a person gets.

```
colnames(filename)
```

```
## [1] "ScrSubjectID"
## [2] "Days from Round1 Day1"
## [3] "Round 1 Exercise"
## [4] "Pre PANAS Positive Affect"
## [5] "Pre PANAS Negative Affect"
## [6] "Post PANAS Positive Affect"
## [7] "Post PANAS Negative Affect"
## [8] "Pre STAI State Anxiety"
## [9] "Post STAI State Anxiety"
## [10] "BW Exercise Minutes (timer)"
## [11] "Did you complete BW exercises?"
## [12] "Sleep Q Score"
## [13] "Full STAI State Anxiety"
## [14] "Full STAI Trait Anxiety"
## [15] "RHR"
## [16] "HRV"
## [17] "Sleep Score"
## [18] "Hours of Sleep"
## [19] "Sleep Efficiency"
## [20] "Respiration Rate"
## [21] "Sleep T Score"
## [22] "Overall, how easy was it to use the instructions/videos for doing the intervention?"
## [23] "Overall, how easy was it do do the daily intervention?"
## [24] "subjective experience: grounding"
## [25] "subjective experience: focusing"
## [26] "subjective experience: calming"
## [27] "subjective experience: energizing"
## [28] "subjective experience: brief and simple"
```

```
## [29] "subjective experience: other positive"
## [30] "subjective experience: some challenges"
```

```
variables <- variables[!apply(is.na(variables) | variables == "", 1, all),]
variables_df <- as.data.frame(variables)

group1 <- variables_df[, c(1, 2, 5)]
group2 <- variables_df[, c(3, 4, 6)]

group1_melted <- gather(group1, key = "variable", value = "value")
group2_melted <- gather(group2, key = "variable", value = "value")

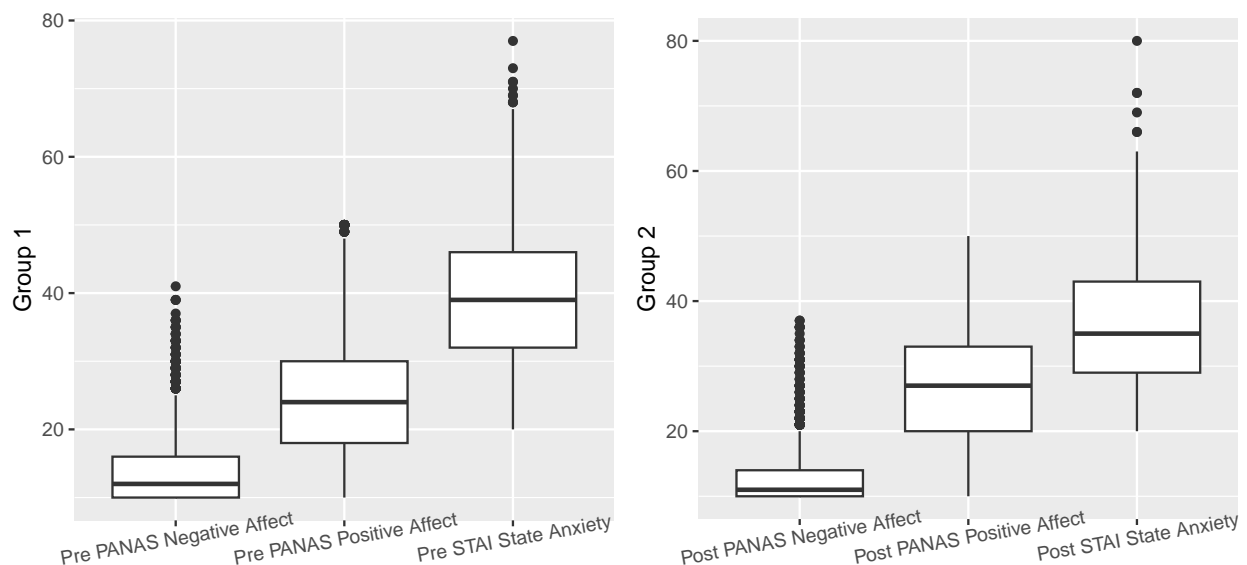
p1 <- ggplot(group1_melted, aes(x = variable, y = value)) +
  geom_boxplot(position = position_dodge(0.75)) +
  xlab("") +
  ylab("Group 1") +
  theme(axis.text.x = element_text(angle = 10))

p2 <- ggplot(group2_melted, aes(x = variable, y = value)) +
  geom_boxplot(position = position_dodge(0.75)) +
  xlab("") +
  ylab("Group 2")+
  theme(axis.text.x = element_text(angle = 10))

plot_grid(p1, p2, ncol = 2)
```

```
## Warning: Removed 10 rows containing non-finite values ('stat_boxplot()').
```

```
## Warning: Removed 393 rows containing non-finite values ('stat_boxplot()').
```



Looking at the results, we see the breathwork group had a notably higher increase in daily positive affect, a reduction in daily negative affect and a reduction in the state anxiety. All of these indicate that there

is some positive effect because of the breathing and meditation protocol that has been assigned to the participants. This is proven by the regression below which shows Mindful Meditation as a significant indicator for a reduction in anxiety and an increase in positive effect.

## 2.4 Mixed-effect Model

Here we use Linear mixed-effect to show the positive effect change based upon the type of exercise, and number of days over which the exercise was performed. Each participant id is treated as a random effect to account for inter-participant variability over the number of days.

Linear mixed-effect models are a type of regression model that are used when there are both fixed (exercise type) and random(participant id) effects in the data. The mixed-effect model provides estimates of the fixed effects and the variance components of the random effects, as well as goodness-of-fit measures and other diagnostic statistics. The lmerTest package was used to obtain p-values and confidence intervals for the fixed effects.

We found that Days from Round 1 (i.e. increase in number of days following a practice) helps in increase the positive effect on the participant (indicator for goodness/wellness).

```
library(readxl)
library(lme4)

library(lmerTest)
library(tidyverse)
filename <- read_excel("BWPilot_CombinedData_20210803_fordryad_addvars_cleaned_noround2_v3.xlsx")

filename[filename=='.'] <- NA

data_set <- filename[, c(1, 2, 3, 4, 6)]

data_set <- data_set %>% drop_na()

data_set$`Post PANAS Positive Affect` <- as.numeric(data_set$`Post PANAS Positive Affect`)
data_set$`Pre PANAS Positive Affect` <- as.numeric(data_set$`Pre PANAS Positive Affect`)
data_set$`Round 1 Exercise` <- as.factor(data_set$`Round 1 Exercise`)
data_set$`Days from Round1 Day1` <- as.numeric(data_set$`Days from Round1 Day1`)

data_set$change <- data_set$`Post PANAS Positive Affect` - data_set$`Pre PANAS Positive Affect`

model = lmer(change ~ `Days from Round1 Day1` + `Round 1 Exercise` * `Days from Round1 Day1` +
              `Round 1 Exercise` + (1|ScrSubjectID),
              data=data_set,
              REML=TRUE)
summary(model)

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## change ~ 'Days from Round1 Day1' + 'Round 1 Exercise' * 'Days from Round1 Day1' +
##       'Round 1 Exercise' + (1 | ScrSubjectID)
## Data: data_set
##
## REML criterion at convergence: 11440
```

```

##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.958 -0.461  0.013  0.534  4.029
##
## Random effects:
##      Groups          Name          Variance Std.Dev.
## ScrSubjectID (Intercept)  9.36      3.06
## Residual                12.91      3.59
## Number of obs: 2066, groups: ScrSubjectID, 110
##
## Fixed effects:
##
##                                     Estimate
## (Intercept)                        1.3510
## 'Days from Round1 Day1'            0.0435
## 'Round 1 Exercise'Mindful Meditation -0.2773
## 'Round 1 Exercise'Slow Breathing    -0.2543
## 'Round 1 Exercise'Super Oxygenation  0.3521
## 'Days from Round1 Day1':'Round 1 Exercise'Mindful Meditation -0.0393
## 'Days from Round1 Day1':'Round 1 Exercise'Slow Breathing    0.0281
## 'Days from Round1 Day1':'Round 1 Exercise'Super Oxygenation -0.0211
##                                     Std. Error
## (Intercept)                        0.7476
## 'Days from Round1 Day1'            0.0208
## 'Round 1 Exercise'Mindful Meditation 1.0389
## 'Round 1 Exercise'Slow Breathing    0.9772
## 'Round 1 Exercise'Super Oxygenation  0.9523
## 'Days from Round1 Day1':'Round 1 Exercise'Mindful Meditation 0.0299
## 'Days from Round1 Day1':'Round 1 Exercise'Slow Breathing    0.0277
## 'Days from Round1 Day1':'Round 1 Exercise'Super Oxygenation 0.0270
##                                     df t value
## (Intercept)                        136.9884    1.81
## 'Days from Round1 Day1'            1969.0147    2.09
## 'Round 1 Exercise'Mindful Meditation 142.7319   -0.27
## 'Round 1 Exercise'Slow Breathing    140.4811   -0.26
## 'Round 1 Exercise'Super Oxygenation  137.7327    0.37
## 'Days from Round1 Day1':'Round 1 Exercise'Mindful Meditation 1979.0557   -1.31
## 'Days from Round1 Day1':'Round 1 Exercise'Slow Breathing    1978.4564    1.01
## 'Days from Round1 Day1':'Round 1 Exercise'Super Oxygenation 1975.6388   -0.78
##                                     Pr(>|t|)
## (Intercept)                        0.073 .
## 'Days from Round1 Day1'            0.037 *
## 'Round 1 Exercise'Mindful Meditation 0.790
## 'Round 1 Exercise'Slow Breathing    0.795
## 'Round 1 Exercise'Super Oxygenation  0.712
## 'Days from Round1 Day1':'Round 1 Exercise'Mindful Meditation 0.189
## 'Days from Round1 Day1':'Round 1 Exercise'Slow Breathing    0.311
## 'Days from Round1 Day1':'Round 1 Exercise'Super Oxygenation 0.434
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) 'DfR1D 'R1E'M 'R1E'B 'R1E'O 'fRD1EM 'fRD1EB
## 'DfRnd1Dy1' -0.365

```

```
## 'R1Exrcs'MM -0.720 0.263
## 'R1Exrcs'SB -0.765 0.279 0.551
## 'R1Exrcs'SO -0.785 0.287 0.565 0.601
## 'DfR1D1'1EM 0.254 -0.696 -0.372 -0.194 -0.200
## 'DfR1D1'1EB 0.274 -0.751 -0.197 -0.366 -0.215 0.523
## 'DfR1D1'1EO 0.282 -0.773 -0.203 -0.216 -0.365 0.538 0.581
```

```
anova(model)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##
## Sum Sq Mean Sq NumDF DenDF F value
## 'Days from Round1 Day1' 170.3 170.3 1 1983 13.19
## 'Round 1 Exercise' 8.6 2.9 3 143 0.22
## 'Days from Round1 Day1': 'Round 1 Exercise' 86.8 28.9 3 1983 2.24
## Pr(>F)
## 'Days from Round1 Day1' 0.00029 ***
## 'Round 1 Exercise' 0.88135
## 'Days from Round1 Day1': 'Round 1 Exercise' 0.08150 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
effect_trait <- as.numeric(trait$`Post Trait`) - as.numeric(trait$`Pre Trait`)

trait['delta trait'] <- effect_trait

b <- trait$Group == "Mindful Meditation"
trait_medi <- trait[b,]

b <- trait$Group == "Box Breathing"
trait_box <- trait[b,]

b <- trait$Group == "Super Oxygenation"
trait_superox <- trait[b,]

b <- trait$Group == "Slow Breathing"
trait_slow <- trait[b,]
####

effect_superox <- as.numeric(trait_superox$`Post Promis`) - as.numeric(trait_superox$`Pre Promis`)
effect_medi <- as.numeric(trait_medi$`Post Promis`) - as.numeric(trait_medi$`Pre Promis`)
effect_box <- as.numeric(trait_box$`Post Promis`) - as.numeric(trait_box$`Pre Promis`)
effect_slow <- as.numeric(trait_slow$`Post Promis`) - as.numeric(trait_slow$`Pre Promis`)

wilcox.test(trait_superox$`Post Trait`, trait_superox$`Pre Trait`, paired = TRUE)

## Warning in wilcox.test.default(trait_superox$`Post Trait`, trait_superox$`Pre
## Trait`, : cannot compute exact p-value with ties

## Warning in wilcox.test.default(trait_superox$`Post Trait`, trait_superox$`Pre
## Trait`, : cannot compute exact p-value with zeroes
```

```

##
## Wilcoxon signed rank test with continuity correction
##
## data: trait_superox$'Post Trait' and trait_superox$'Pre Trait'
## V = 38, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0

wilcox.test(trait_slow$`Post Trait`,trait_slow$`Pre Trait`, paired = TRUE)

## Warning in wilcox.test.default(trait_slow$'Post Trait', trait_slow$'Pre Trait',
## : cannot compute exact p-value with ties

## Warning in wilcox.test.default(trait_slow$'Post Trait', trait_slow$'Pre Trait',
## : cannot compute exact p-value with zeroes

##
## Wilcoxon signed rank test with continuity correction
##
## data: trait_slow$'Post Trait' and trait_slow$'Pre Trait'
## V = 61, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0

wilcox.test(trait_box$`Post Trait`,trait_box$`Pre Trait`, paired = TRUE)

## Warning in wilcox.test.default(trait_box$'Post Trait', trait_box$'Pre Trait', :
## cannot compute exact p-value with ties

## Warning in wilcox.test.default(trait_box$'Post Trait', trait_box$'Pre Trait', :
## cannot compute exact p-value with zeroes

##
## Wilcoxon signed rank test with continuity correction
##
## data: trait_box$'Post Trait' and trait_box$'Pre Trait'
## V = 34, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0

wilcox.test(trait_medi$`Post Trait`,trait_medi$`Pre Trait`, paired = TRUE)

## Warning in wilcox.test.default(trait_medi$'Post Trait', trait_medi$'Pre Trait',
## : cannot compute exact p-value with ties

## Warning in wilcox.test.default(trait_medi$'Post Trait', trait_medi$'Pre Trait',
## : cannot compute exact p-value with zeroes

##
## Wilcoxon signed rank test with continuity correction
##
## data: trait_medi$'Post Trait' and trait_medi$'Pre Trait'
## V = 27, p-value = 1
## alternative hypothesis: true location shift is not equal to 0

```



```

kruskal.test(`delta trait`~Group, data=trait)

##
## Kruskal-Wallis rank sum test
##
## data: delta trait by Group
## Kruskal-Wallis chi-squared = 2, df = 3, p-value = 0.7

wilcox.test(trait_slow$`Post Trait`, trait_slow$`Pre Trait`, paired = TRUE) #WILCOX IS 2 SAMPLE ONLY, KR

## Warning in wilcox.test.default(trait_slow$`Post Trait`, trait_slow$`Pre Trait`,
## : cannot compute exact p-value with ties

## Warning in wilcox.test.default(trait_slow$`Post Trait`, trait_slow$`Pre Trait`,
## : cannot compute exact p-value with zeroes

##
## Wilcoxon signed rank test with continuity correction
##
## data: trait_slow$`Post Trait` and trait_slow$`Pre Trait`
## V = 61, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0

#round 1
trait <- trait [complete.cases(trait),]
trait$Group<- factor(trait$Group, levels = c('Mindful Meditation', 'Slow Breathing', 'Box Breathing', 'V
e <- ggplot(trait, aes(x = Group, y = `delta trait`))+geom_boxplot(alpha = 0.3)+geom_point (aes(colour =

## by type t test

traitmed<- subset(trait, Type == 'Mindfulness')
traitbreath <- subset(trait, Type == 'Breathwork')
t.test(traitmed$`delta trait`, y =traitbreath$`delta trait`,
       alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = FALSE, var.equal = FALSE,
       conf.level = 0.95)

##
## Welch Two Sample t-test
##
## data: traitmed$`delta trait` and traitbreath$`delta trait`
## t = 1, df = 16, p-value = 0.2
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.10 5.96
## sample estimates:
## mean of x mean of y
## 0.75 -1.68

```

```

t.test(traitmed$`Pre Trait`, y =traitmed$`Post Trait`,
       alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = TRUE, var.equal = FALSE,
       conf.level = 0.95)

##
## Paired t-test
##
## data: traitmed$`Pre Trait` and traitmed$`Post Trait`
## t = -0.5, df = 11, p-value = 0.6
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -4.06 2.56
## sample estimates:
## mean of the differences
## -0.75

t.test(traitbreath$`Pre Trait`, y =traitbreath$`Post Trait`,
       alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = TRUE, var.equal = FALSE,
       conf.level = 0.95)

##
## Paired t-test
##
## data: traitbreath$`Pre Trait` and traitbreath$`Post Trait`
## t = 2, df = 52, p-value = 0.02
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.228 3.131
## sample estimates:
## mean of the differences
## 1.68

traitmed<- traitmed[complete.cases(traitmed),]
mean (traitmed$`delta trait`)/sd(traitmed$`Pre Trait`)

## [1] 0.0802

traitbreath<- traitbreath[complete.cases(traitbreath),]
mean (traitbreath$`delta trait`)/sd(traitbreath$`Pre Trait`)

## [1] -0.149

wilcox.test(traitbreath$`Post Trait`,traitbreath$`Pre Trait`, paired = TRUE)

## Warning in wilcox.test.default(traitbreath$`Post Trait`, traitbreath$`Pre
## Trait`, : cannot compute exact p-value with ties

## Warning in wilcox.test.default(traitbreath$`Post Trait`, traitbreath$`Pre
## Trait`, : cannot compute exact p-value with zeroes

```

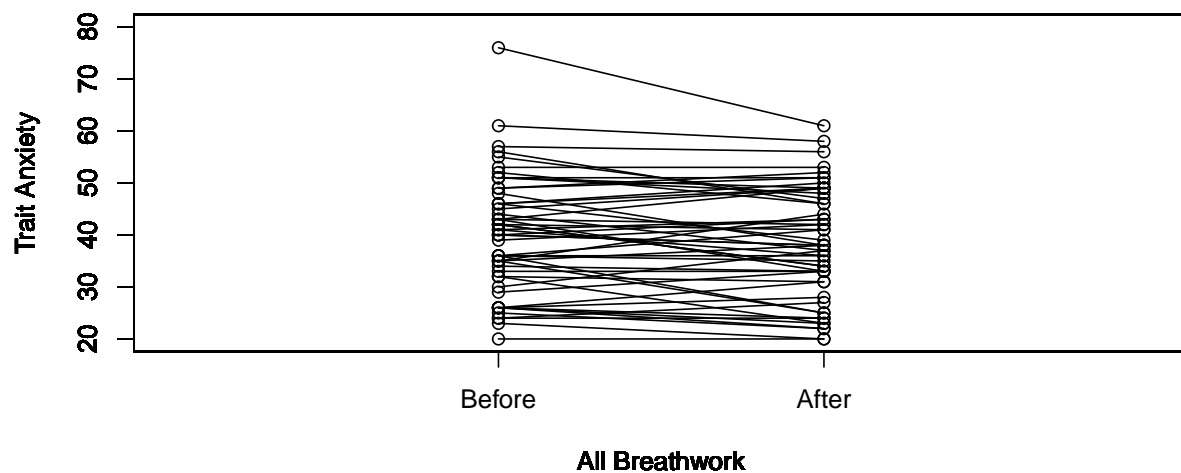
```
##
## Wilcoxon signed rank test with continuity correction
##
## data: traitbreath$'Post Trait' and traitbreath$'Pre Trait'
## V = 392, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
```

```
wilcox.test(traitbreath$`delta trait`,traitmed$`delta trait`, paired = FALSE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: traitbreath$'delta trait' and traitmed$'delta trait'
## W = 262, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
```

```
plot.new()
for (i in 1:nrow(traitbreath)){
  if (is.na(traitbreath$`Pre Trait`[i])|is.na(traitbreath$`Post Trait`[i])) next
  par(new=T)
  plot(c(traitbreath$`Pre Trait`[i],traitbreath$`Post Trait` [i]), type= 'o', ylab = 'Trait Anxiety', x=
}

axis(1, at=c(1,2), labels=c('Before', 'After'))
```



```
trait$Type<- factor(trait$Type, levels = c('Mindfulness', 'Breathwork'))
trait<- trait[complete.cases(trait),]
e <- ggplot(trait, aes(x = Type, y = `delta trait`))+geom_boxplot(alpha = 0.3)+geom_point (aes(colour =
kruskal.test(`delta trait`~Type, data = trait)
```

```
##
```

```
## Kruskal-Wallis rank sum test
##
## data: delta trait by Type
## Kruskal-Wallis chi-squared = 0.9, df = 1, p-value = 0.3
```

```
#ttest by group
```

```
traitbreath <- subset(trait, Group == 'Box Breathing')
traitmed <- subset(trait, Group == 'Super Oxygenation')

t.test(traitmed$`delta trait`, y = traitbreath$`delta trait`,
       alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = FALSE, var.equal = FALSE,
       conf.level = 0.95)
```

```
##
## Welch Two Sample t-test
##
## data: traitmed$`delta trait` and traitbreath$`delta trait`
## t = -0.4, df = 25, p-value = 0.7
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -4.18 2.80
## sample estimates:
## mean of x mean of y
## -1.82 -1.13
```

We have conducted Wilcoxin and Kruskal-Wallis rank sum test, whose output is above. The wilcoxin test for the super-oxygenation group is significant at the 0.10 level. The Kruskal-Wallis test is not significant which may indicate that there is no evidence that the medians of the pre and post groups are significantly different. However, it does not say that there are no differences in the groups. It may be the case that the differences are too small to be detected.

The paired t-test for pre and post traitbreath shows that the difference in factor is significant.

This graphs shows the before and after trait anxiety. This graph does show clearly that there is a reduction in before and after state anxiety.

```
#remove any empty columns
```

```
slopes <- slopes[,colSums(is.na(slopes))<nrow(slopes)]
slopes_rhrhrvRR_excludinglessthan8datapoints<- slopes # we are not excluding data.
```

```
#statistical tests
```

```
#non-parametric
```

```
slopes_rhrhrvRR_excludinglessthan8datapoints<- slopes_rhrhrvRR_excludinglessthan8datapoints[complete.cases(slopes_rhrhrvRR_excludinglessthan8datapoints)]
kruskal.test(slopeRR~Group, data = slopes_rhrhrvRR_excludinglessthan8datapoints)
```

```
##
## Kruskal-Wallis rank sum test
##
## data: slopeRR by Group
## Kruskal-Wallis chi-squared = 1, df = 3, p-value = 0.7
```

```
pairwise.wilcox.test(slopes_rhrhrvRR_excludinglessthan8datapoints$slopeRR, slopes_rhrhrvRR_excludingles
```

```
##
## Pairwise comparisons using Wilcoxon rank sum exact test
##
## data: slopes_rhrhrvRR_excludinglessthan8datapoints$slopeRR and slopes_rhrhrvRR_excludinglessthan8da
##
##           Box Breathing Mindful Meditation Slow Breathing
## Mindful Meditation 1           -           -
## Slow Breathing      1           1           -
## Super Oxygenation   1           1           1
##
## P value adjustment method: holm
```

```
for (i in 1:nrow(slopes_rhrhrvRR_excludinglessthan8datapoints)){
  if (slopes_rhrhrvRR_excludinglessthan8datapoints$Group[i] == "Mindful Meditation"){ slopes_rhrhrvRR_ex
# else if (c$`Round 1 Exercise` == 'Box Breathing'){ weekly$Type[i] <- 'Box Breathing'}
else {slopes_rhrhrvRR_excludinglessthan8datapoints$Type[i] <- 'Breathing'}
}
slopes_rhrhrvRR_excludinglessthan8datapoints$Type <- factor(slopes_rhrhrvRR_excludinglessthan8datapoin

# kruskal.test(slopeRR~Type, data = slopes_rhrhrvRR_excludinglessthan8datapoints)

rrm <- subset(slopes_rhrhrvRR_excludinglessthan8datapoints, Type == "Mindfulness")
rrb <- subset(slopes_rhrhrvRR_excludinglessthan8datapoints, Type == "Breathing")
rrsb <- subset(slopes_rhrhrvRR_excludinglessthan8datapoints, Group == "Slow Breathing")
rrbb <- subset(slopes_rhrhrvRR_excludinglessthan8datapoints, Group == "Box Breathing")
rrso <- subset(slopes_rhrhrvRR_excludinglessthan8datapoints, Group == "Super Oxygenation")

deltapanasposm <- subset(slopes_rhrhrvRR_excludinglessthan8datapoints, Type == "Mindfulness")
deltapanasposb <- subset(slopes_rhrhrvRR_excludinglessthan8datapoints, Type == "Breathing")

# t test

t.test(rrm$slopeRR, y =rrb$slopeRR,
        alternative = c ("two.sided", "less", "greater"),
        mu = 0, paired = FALSE, var.equal = FALSE,
        conf.level = 0.95)
```

```
##
## Welch Two Sample t-test
##
## data: rrm$slopeRR and rrb$slopeRR
## t = 1, df = 35, p-value = 0.3
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.00224 0.00674
## sample estimates:
## mean of x mean of y
## 0.002666 0.000415
```

```

signal<- slopes_rhrhrvRR_excludinglessthan8datapoints
signal <- signal [complete.cases(signal),]
signal$Type<- factor(signal$Type, levels = c('Mindfulness', 'Breathing'))
signal$Group<- factor(signal$Group, levels = c('Mindful Meditation', 'Slow Breathing', 'Box Breathing', 'Super Oxygenation'))
print(i)

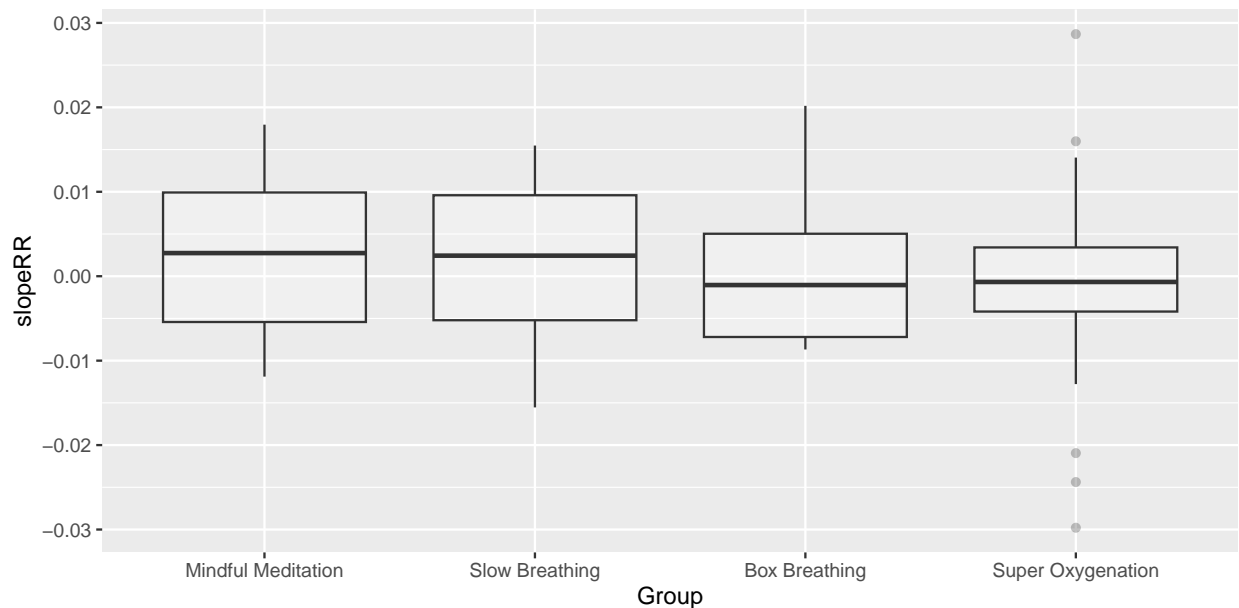
```

```
## [1] 105
```

```

e <- ggplot(signal, aes(x = Group, y = slopeRR))+geom_boxplot(alpha = 0.3)
e

```



Here we have boxplots behind the mindfulness and breathing exercises. The boxplot is slopeRR vs the different kinds of breathing exercises. It shows that there are four groups and two of the groups (except box breathing and super oxygenation) show that there is a positive effect between them and the slopeRR. SlopeRR corresponds to the effect of the number of days we stick to a regime. Non-zero slopes show that all exercises are beneficial in long run (corresponding to how long we stick with a regime).

## 2.5 Matching

Here we consider matching participants from a group where group is defined by type of exercise performed by the participants to, a control group the control group is defined by all the participants who enrolled in the study. The control group is chosen as such because we have data for all the participants before they started doing any of the recommended exercises. We can match each treatment group to all the participants. The control group and the treatment group will be matched based on: RHR, HRV, Sleep Score, Hours of sleep, sleep efficacy, hours of sleep and respiration rate Outcomes before and after the exercises will be chosen for determining treatment effect

```

#-----
#Preparing dataset
library(tidyverse)
library(readxl)
library(dplyr)

```

```

filename <- read_excel("BWPilot_CombinedData_20210803_fordryad_addvars_cleaned_noround2_v3.xlsx")
filename[filename=="."] <- NA
data_set <- filename[, c(1,2,15,16,17,18,19,20)]

data_set <- data_set %>% drop_na()

data_set$`Days from Round1 Day1` <- as.numeric(data_set$`Days from Round1 Day1`)

data_set <- data_set[-20 < data_set$`Days from Round1 Day1`, ]

data_set$RHR <- as.numeric(data_set$RHR)
data_set$HRV <- as.numeric(data_set$HRV)
data_set$`Sleep Score` <- as.numeric(data_set$`Sleep Score`)
data_set$`Hours of Sleep` <- as.numeric(data_set$`Hours of Sleep`)
data_set$`Sleep Efficiency` <- as.numeric(data_set$`Sleep Efficiency`)
data_set$`Respiration Rate` <- as.numeric(data_set$`Respiration Rate`)

t.first <- filename[match(unique(data_set$ScrSubjectID), data_set$ScrSubjectID),c(3)]
t.sec <- ifelse(t.first$`Round 1 Exercise` == "Box Breathing", 1,
               ifelse(t.first$`Round 1 Exercise` == "Slow Breathing", 2,
                      ifelse(t.first$`Round 1 Exercise` == "Super Oxygenation", 3, 4))
               )

a <- rep(0, 113)

treatment <- c(rbind(a, t.sec))

treatment <- treatment[-122]
treatment <- treatment[-137]
data_set <- data_set %>%
  mutate(pos = `Days from Round1 Day1`>0)

data_set_matching <- data_set %>% group_by(ScrSubjectID, pos) %>%
  summarise_all(mean)

data_set_matching <- data.frame(data_set_matching, treatment)
#-----
library(tidyverse)
library(ggplot2)
library(MASS)

##
## Attaching package: 'MASS'

## The following object is masked from 'package:plotly':
##
##     select

## The following object is masked from 'package:dplyr':
##
##     select

```

```

#-----
smahal=
function(z,X){
  X<-as.matrix(X)
  n<-dim(X)[1]
  rownames(X)<-1:n
  k<-dim(X)[2]
  m<-sum(z)
  for (j in 1:k) X[,j]<-rank(X[,j])
  cv<-cov(X)
  vuntied<-var(1:n)
  rat<-sqrt(vuntied/diag(cv))
  cv<-diag(rat)%*%cv%*%diag(rat)
  out<-matrix(NA,m,n-m)
  Xc<-X[z==0,]
  Xt<-X[z==1,]
  9
  rownames(out)<-rownames(X)[z==1]
  colnames(out)<-rownames(X)[z==0]
  icov<-ginv(cv)
  for (i in 1:m) out[i,]<-mahalanobis(Xc,Xt[i,],icov,inverted=T)
  out
}

addcaliper=function(dmat,z,logitp,calipersd=.5,penalty=1000){
  # Pooled within group standard deviation
  sd.logitp=sqrt((sd(logitp[z==1])^2+sd(logitp[z==0])^2)/2)
  adif=abs(outer(logitp[z==1],logitp[z==0],"-"))
  adif=(adif-(calipersd*sd.logitp))*(adif>(calipersd*sd.logitp))
  dmat=dmat+adif*penalty
  dmat
}
#-----
subset <- data_set_matching[data_set_matching$treatment %in% c(0,1), ]
datatemp=subset

subset$Hours.of.Sleep

```

```

## [1] 7.45 7.22 7.41 7.36 6.23 7.22 6.69 8.69 6.75 6.66 7.38 6.81 7.95 7.24 7.06
## [16] 8.39 6.67 6.61 6.25 6.64 7.67 6.63 6.91 5.63 6.14 9.56 8.42 7.08 6.76 7.15
## [31] 6.82 5.77 7.23 7.13 5.79 7.01 7.42 8.14 7.06 8.45 8.45 7.17 7.15 7.29 6.87
## [46] 8.07 6.37 7.15 7.39 8.08 7.62 4.72 6.20 8.40 7.54 7.10 7.67 8.30 5.99 6.47
## [61] 7.04 6.40 6.31 7.42 6.67 8.10 8.07 7.11 8.41 8.02 6.38 6.50 7.38 7.43 7.21
## [76] 7.43 6.78 6.44 7.24 7.35 7.01 3.57 6.59 6.57 3.82 7.27 7.70 7.75 7.88 6.17
## [91] 5.87 7.08 7.06 7.22 6.86 6.21 7.80 7.62 6.66 7.66 7.81 7.50 7.48 7.28 7.38
## [106] 6.93 6.98 7.59 7.42 6.86 6.38 6.63 6.87 7.36 7.12 7.62 6.21 4.80 7.49 7.07
## [121] 7.16 7.81 8.20 6.51 8.33 7.25 6.40 6.70 7.24 6.08 6.79 7.60 7.76 7.01 8.78
## [136] 6.82 7.46 7.26

```

```
subset$Sleep.Efficiency
```

```
## [1] 89.3 92.1 93.0 92.7 90.6 91.6 86.3 94.0 91.7 93.1 95.5 94.8 94.6 93.7 93.4
```



```
## [16] 92.5 88.2 83.0 93.2 89.7 93.8 94.1 92.7 72.8 90.2 92.2 96.1 91.4 89.1 94.2
## [31] 89.6 91.0 87.8 86.0 92.5 92.2 85.6 96.1 94.0 94.8 95.0 88.9 90.4 91.1 86.3
## [46] 95.2 93.4 93.2 94.5 95.7 95.3 88.2 93.2 92.8 93.4 90.5 86.6 92.0 85.9 89.8
## [61] 90.3 92.8 94.5 91.9 85.8 92.8 91.0 95.5 90.0 92.4 93.9 93.3 92.3 96.2 93.9
## [76] 93.5 92.1 94.5 90.4 91.2 85.8 97.0 82.1 91.1 71.8 89.9 93.7 95.2 94.2 93.2
## [91] 94.8 95.8 95.6 90.8 93.8 90.3 94.0 93.2 88.5 90.3 91.6 89.2 94.0 93.4 93.2
## [106] 93.3 93.7 88.2 87.4 92.7 93.6 88.1 87.8 94.7 94.2 94.3 93.9 67.7 93.4 93.7
## [121] 93.6 96.1 94.4 84.6 89.8 85.5 77.4 82.5 88.7 86.0 85.6 90.5 90.9 94.0 93.5
## [136] 86.8 91.8 91.9
```

```
subset$Respiration.Rate
```

```
## [1] 13.8 13.9 17.0 16.9 17.4 17.1 17.0 16.0 17.5 17.6 13.3 13.6 15.0 15.1 15.4
## [16] 15.0 17.5 16.2 17.5 16.7 14.9 15.6 15.1 17.6 14.8 16.2 16.4 18.1 15.4 13.7
## [31] 14.1 16.1 15.1 14.7 17.0 17.1 14.2 15.5 17.2 15.5 15.2 14.5 15.6 15.5 16.6
## [46] 16.3 15.4 17.2 16.7 17.7 17.5 19.2 17.3 16.2 15.0 15.5 14.2 14.6 17.8 16.0
## [61] 16.1 17.0 15.3 19.3 17.2 16.2 16.7 16.1 15.9 14.5 16.6 16.8 15.0 18.0 14.3
## [76] 16.7 15.7 18.8 16.4 16.7 14.2 16.9 16.1 17.7 12.7 12.3 16.7 13.9 16.4 17.2
## [91] 17.3 13.4 13.5 18.1 16.5 14.7 18.9 18.0 14.3 16.5 15.3 15.9 16.2 14.9 14.2
## [106] 13.8 13.7 13.6 13.7 14.7 16.6 13.3 13.6 17.5 17.9 14.9 15.7 15.1 15.8 15.5
## [121] 15.5 14.3 14.0 13.5 17.7 13.9 13.1 12.8 17.2 15.4 17.1 15.9 16.1 14.6 13.2
## [136] 14.4 14.2 14.5
```

```
subset$Days.from..Round1.Day1
```

```
## [1] -9.50 15.50 -9.50 15.50 -9.63 19.56 -9.50 -9.50 -9.50 14.30
## [11] -9.50 16.50 -9.50 -9.50 15.50 -9.50 -8.50 -9.50 -9.50 -9.50
## [21] -9.50 -9.50 -9.50 -9.50 -2.00 -9.50 19.39 -10.24 -9.58 -9.50
## [31] -9.26 -6.93 -9.72 14.55 -8.00 14.71 -9.50 -9.26 -9.50 -9.50
## [41] 15.50 -9.50 -9.50 15.50 -9.50 -9.50 -3.00 -10.47 -12.15 -9.74
## [51] -9.50 -9.61 -9.50 -9.50 -10.79 -9.50 -9.33 -9.50 -8.54 -9.47
## [61] -10.06 -8.79 -9.50 -9.89 -9.50 -9.63 16.06 -9.47 -9.50 -9.05
## [71] -9.50 15.50 -9.50 -11.50 -4.50 -10.53 -9.00 -1.50 -9.71 -9.50
## [81] -9.50 -6.00 -7.00 -9.50 -16.50 -9.62 -8.94 -9.50 -1.50 -9.50
## [91] 5.00 -9.50 15.97 -4.00 -9.41 -9.50 -1.50 -7.50 -9.50 -9.89
## [101] -10.75 -11.46 -7.50 -9.76 5.89 -9.50 18.00 -9.50 15.50 -5.00
## [111] -7.50 -8.00 16.94 -5.50 18.18 -3.17 12.27 -8.50 -8.00 -3.29
## [121] 12.39 -4.00 19.50 -3.00 -2.50 -2.50 -3.29 9.00 -2.50 -3.00
## [131] -2.80 -5.50 -3.57 -2.00 -1.50 -2.00 -1.50 15.00
```

```
colnames(subset)[3] <- "days"
```

```
#Score model
```

```
propscore.model=glm(pos ~ RHR + HRV +Sleep.Score+
  Respiration.Rate,
  family=binomial,x=TRUE,y=TRUE,
  data=subset)
```

```
# #-----
datatemp$outcome = datatemp$Sleep.Efficiency
datatemp$treated=propscore.model$y
datatemp$treatment=datatemp$treated
```



```

library(optmatch)
matchvec=fullmatch(distmat,data=datatemp)
datatemp$matchvec=matchvec
# Number the strata
matchedset.index=substr(matchvec,start=3,stop=10)
matchedset.index.numeric=as.numeric(matchedset.index)
# Have matchedset.index.numeric.full append 0 to matchedset.index.numeric for
# the removed subjects
if(no.control.lack.overlap+no.treated.lack.overlap==0){
  matchedset.index.numeric.full=matchedset.index.numeric
}
if(no.control.lack.overlap+no.treated.lack.overlap>0){
  matchedset.index.numeric.full=c(matchedset.index.numeric,rep(0,no.control.lack.overlap+no.treated.lack.overlap))
}

standardized.diff.harmonic.func=function(x,treatment,stratum.myindex,missing=rep(0,length(x))){
  xtreated=x[treatment==1 & missing==0];
  xcontrol=x[treatment==0 & missing==0];
  var.xtreated=var(xtreated);
  var.xcontrol=var(xcontrol);
  combinedsd=sqrt(.5*(var.xtreated+var.xcontrol));
  std.diff.before.matching=(mean(xtreated)-mean(xcontrol))/combinedsd;
  nostratum=length(unique(stratum.myindex))-1*max(stratum.myindex==0);
  if(max(stratum.myindex==0)==0){
    stratumlist=sort(unique(stratum.myindex))
  }
  if(max(stratum.myindex==0)==1){
    templist=sort(unique(stratum.myindex))
    stratumlist=templist[-1]
  }
  diff.in.stratum=rep(0,nostratum);
  number.in.stratum=rep(0,nostratum);
  harmonic.weight=rep(0,nostratum)
  for(i in 1:nostratum){
    if(sum(stratum.myindex==stratumlist[i] & treatment==1 & missing==0)==0 | sum(stratum.myindex==stratumlist[i] & treatment==0 & missing==0)==0){
      number.in.stratum[i]=0
    }
    if(sum(stratum.myindex==stratumlist[i] & treatment==1 & missing==0)>0 & sum(stratum.myindex==stratumlist[i] & treatment==0 & missing==0)>0){
      diff.in.stratum[i]=mean(x[stratum.myindex==stratumlist[i] & treatment==1 & missing==0])-mean(x[stratum.myindex==stratumlist[i] & treatment==0 & missing==0])
      number.in.stratum[i]=sum(stratum.myindex==stratumlist[i])
      harmonic.weight[i]=1/((.5/sum(stratum.myindex==stratumlist[i] & treatment==1)+.5/sum(stratum.myindex==stratumlist[i] & treatment==0 & missing==0)))
    }
  }
  std.diff.after.matching=(sum(harmonic.weight*diff.in.stratum)/sum(harmonic.weight))/combinedsd;
  list(std.diff.before.matching=std.diff.before.matching,std.diff.after.matching=std.diff.after.matching)
}

# Also compute balance on logit propensity score
# Xmatmahal$logit.ps=datatemp$logit.ps
# Calculate the standardized differences
std.diff.before=rep(0,ncol(Xmatmahal));
std.diff.after=rep(0,ncol(Xmatmahal));
names(std.diff.before)=names(Xmatmahal[1,]);

```

```

names(std.diff.after)=names(Xmatmahal[1,]);
for(i in 1:ncol(Xmatmahal)){
  missing.temp=is.na(Xmatmahal[,i])
  temp.stand.diff=standardized.diff.harmonic.func(Xmatmahal[,i],datatemp$treated,matchedset.index.number)
  std.diff.before[i]=temp.stand.diff$std.diff.before.matching;
  std.diff.after[i]=temp.stand.diff$std.diff.after.matching;
}

sd.bf=std.diff.before
sd.af=std.diff.after
options(digits=2)
cbind(sd.bf,sd.af)

```

```

##           sd.bf  sd.af
## RHR        -0.131 -0.091
## HRV         0.124  0.041
## Sleep.Score -0.011  0.107
## Respiration.Rate -0.296 -0.030

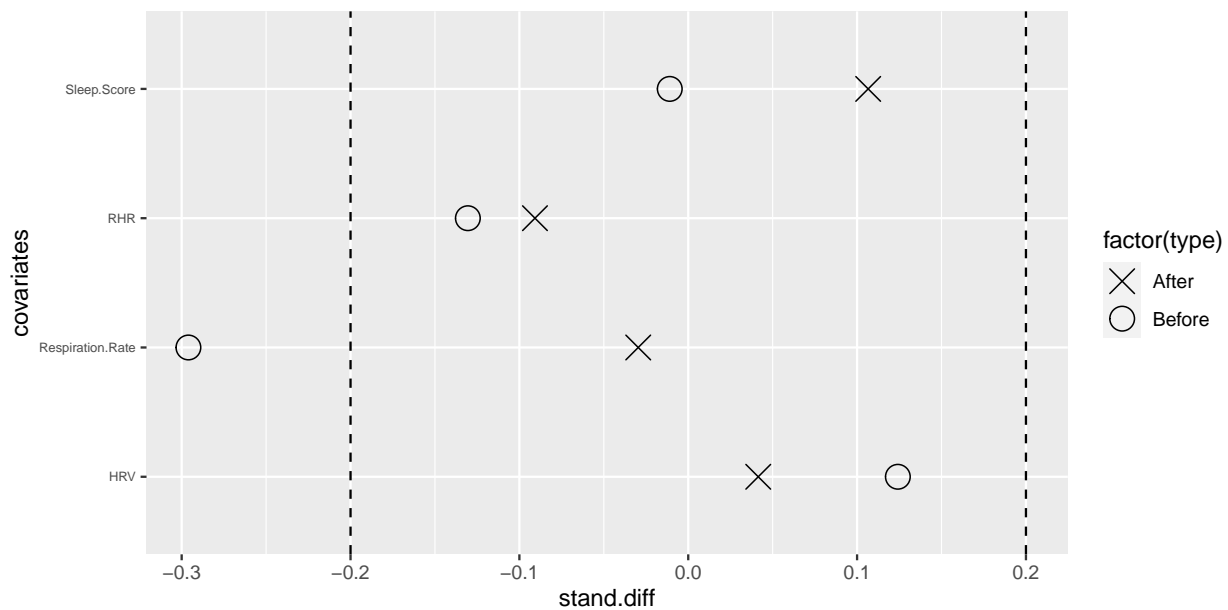
```

```

covariates=names(sd.bf)

plot.dataframe2=data.frame(stand.diff=c(sd.bf,sd.af),covariates=rep(covariates,2),type=c(rep("Before",1),rep("After",1)))
ggplot(plot.dataframe2,aes(x=stand.diff,y=covariates))+geom_point(size=5,aes(shape=factor(type)))+scale_x_continuous(limits=c(-0.3,0.2))
theme(axis.text.y = element_text(size=6))

```



```

ggsave(
  filename = "l5.png",
  device = "png", width = 6, height = 4)

#Point estimate for optimal matching and Confidence interval.
reg.formula=update(propscore.model$formula,outcome~treated+matchvec+. )

```

```
matched.reg.model=lm(reg.formula,data=datatemp)
summary(matched.reg.model)
```

```
##
## Call:
## lm(formula = reg.formula, data = datatemp)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-16.869	-1.493	0.542	2.190	9.854

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	68.71316	14.60001	4.71	7.7e-06 ***
treated	1.68585	0.87248	1.93	0.0560 .
matchvec1.10	0.55462	3.01477	0.18	0.8544
matchvec1.11	-0.75081	3.20074	-0.23	0.8150
matchvec1.12	-1.41057	4.43997	-0.32	0.7513
matchvec1.13	0.06147	3.66895	0.02	0.9867
matchvec1.14	1.66528	5.29452	0.31	0.7537
matchvec1.15	1.20051	3.21156	0.37	0.7093
matchvec1.16	0.57763	3.39334	0.17	0.8652
matchvec1.17	-0.18505	3.25848	-0.06	0.9548
matchvec1.18	-2.03379	3.35302	-0.61	0.5454
matchvec1.19	-1.17249	3.26901	-0.36	0.7206
matchvec1.2	-1.67335	3.89235	-0.43	0.6681
matchvec1.20	-3.23396	5.08877	-0.64	0.5265
matchvec1.21	-3.41764	3.99087	-0.86	0.3937
matchvec1.22	-2.45271	3.36870	-0.73	0.4682
matchvec1.23	-0.51125	3.24921	-0.16	0.8753
matchvec1.24	-10.32070	3.46212	-2.98	0.0036 **
matchvec1.25	-2.28907	3.17919	-0.72	0.4731
matchvec1.3	-1.32347	3.67470	-0.36	0.7194
matchvec1.4	-6.26546	4.80445	-1.30	0.1950
matchvec1.5	2.05082	2.99904	0.68	0.4956
matchvec1.6	-1.23991	4.06625	-0.30	0.7610
matchvec1.7	-1.61127	3.51187	-0.46	0.6473
matchvec1.8	-0.03670	2.81944	-0.01	0.9896
matchvec1.9	-1.49280	4.18927	-0.36	0.7223
RHR	-0.01485	0.09022	-0.16	0.8695
HRV	-0.00335	0.02099	-0.16	0.8736
Sleep.Score	0.11957	0.04752	2.52	0.0134 *
Respiration.Rate	0.94247	0.87621	1.08	0.2845

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.8 on 106 degrees of freedom
## Multiple R-squared:  0.445, Adjusted R-squared:  0.294
## F-statistic: 2.94 on 29 and 106 DF, p-value: 3.13e-05
```

```
# Point estimate of treatment effect
coef(matched.reg.model)[2]
```

```
## treated
##      1.7
```

```
# Confidence interval
confint(matched.reg.model)[2,]
```

```
##  2.5 % 97.5 %
## -0.044  3.416
```

The love plot depicts standard differences before and after matching the treatment group of people who did box breathing vs control group comprising of all people before they started exercises. We performed regression analysis to determine significance of treatment effect of performing box breathing exercise. We compared used sleep efficacy of both groups as the outcome metric. The p-value suggests that box breathing improves sleep efficacy, we did a similar analysis for other breathing techniques using full matching but did not find any significant treatment effect.

## 2.6 Footnotes and Sources

1. Stat 9210 Notes [Matching Code]
2. Brief structured respiration practices enhance mood and reduce physiological arousal (Melis Yilmaz Balban, Eric Neri, Manuela M. Kogon, . . . , Jamie M. Zeitzer, David Spiegel, Andrew D. Huberman) [Boxplots, T-test, Kruskal, Wilcoxon for comparing all the breathworks were generated using the data and code provided in this paper]
3. The American Journal of Medicine [<https://www.amjmed.com/>]