

Midterm-Paper

Team 4

2023-06-08

Data Preparation

```
# Prepare dataframe for analysis
testing_dat <- data.frame(
  Id=integer(),
  SitP=double(),
  SitNP=double(),
  WalkP=double(),
  WalkNP=double(),
  stringsAsFactors=FALSE
)

for (index in c(1:10)){
  # Imported preprocessed data
  dat <- read.csv(sprintf("../Desktop/statistics_healthcare-main_github/Data_clean/log_p%s.txt", index), sep="|")

  # Acceleration
  dat$acceleration <- sqrt(dat$x^2 + dat$y^2 + dat$z^2)

  # "Stratify" data
  sitting_phone <- dat[dat$statusId == 2, ] # this is the id for sitting on phone
  sitting_no_phone <- dat[dat$statusId == 1, ] # this is the id for sitting
  walking_phone <- dat[dat$statusId == 4, ] # this is the id for walking on phone
  walking_no_phone <- dat[dat$statusId == 3, ] # this is the id for walking

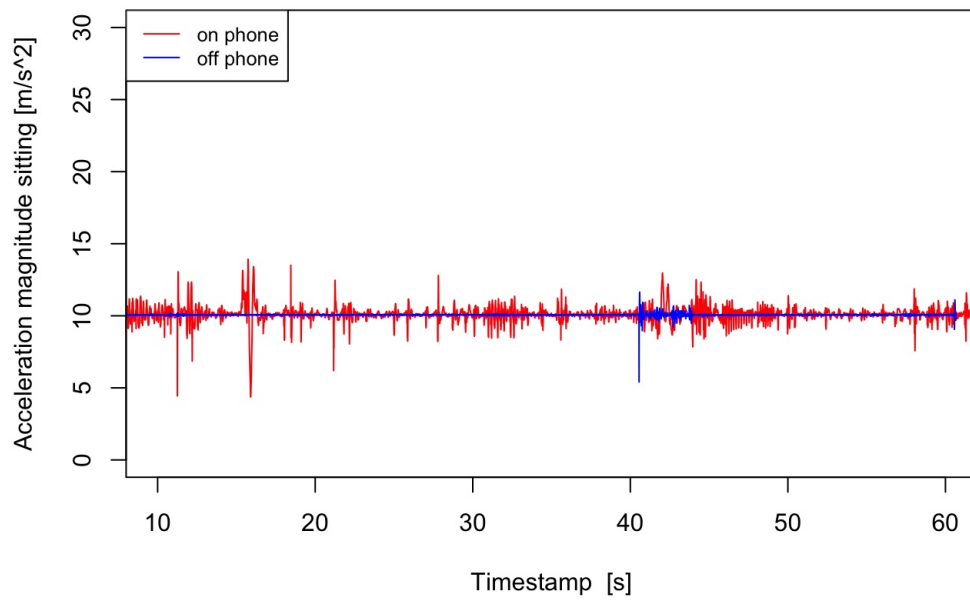
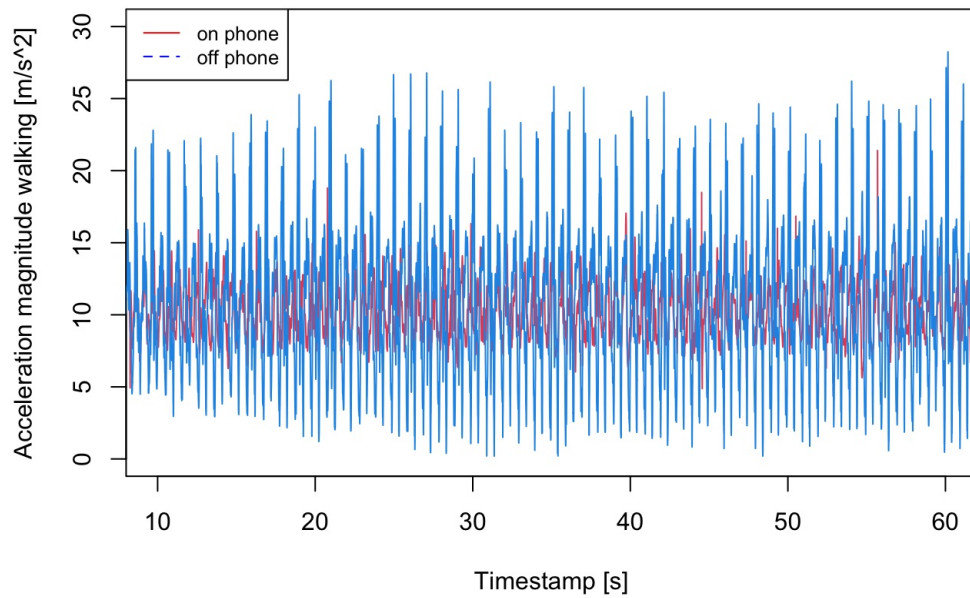
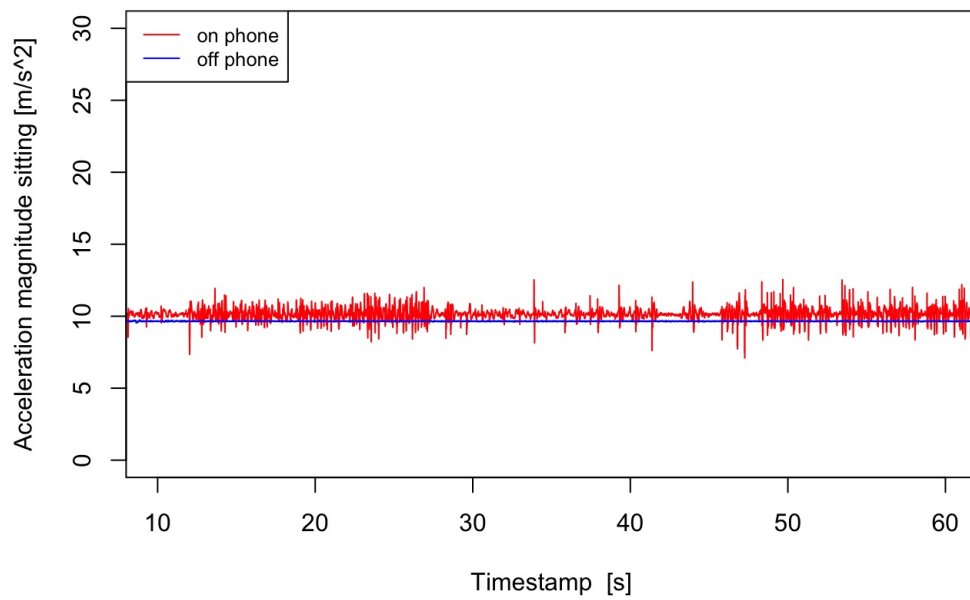
  # Normalize timestamps to beginnings of recordings
  min_sp <- min(sitting_phone$timestamp)
  min_snp <- min(sitting_no_phone$timestamp)
  min_wp <- min(walking_phone$timestamp)
  min_wnp <- min(walking_no_phone$timestamp)
  sitting_phone$timestamp <- (sitting_phone$timestamp - min_sp)
  sitting_no_phone$timestamp <- (sitting_no_phone$timestamp - min_snp)
  walking_phone$timestamp <- (walking_phone$timestamp - min_wp)
  walking_no_phone$timestamp <- (walking_no_phone$timestamp - min_wnp)

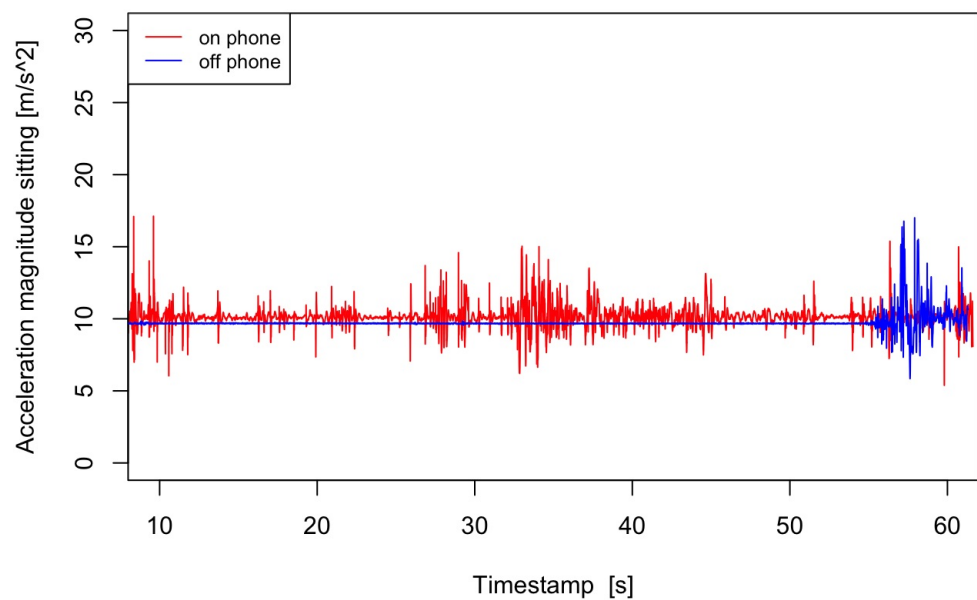
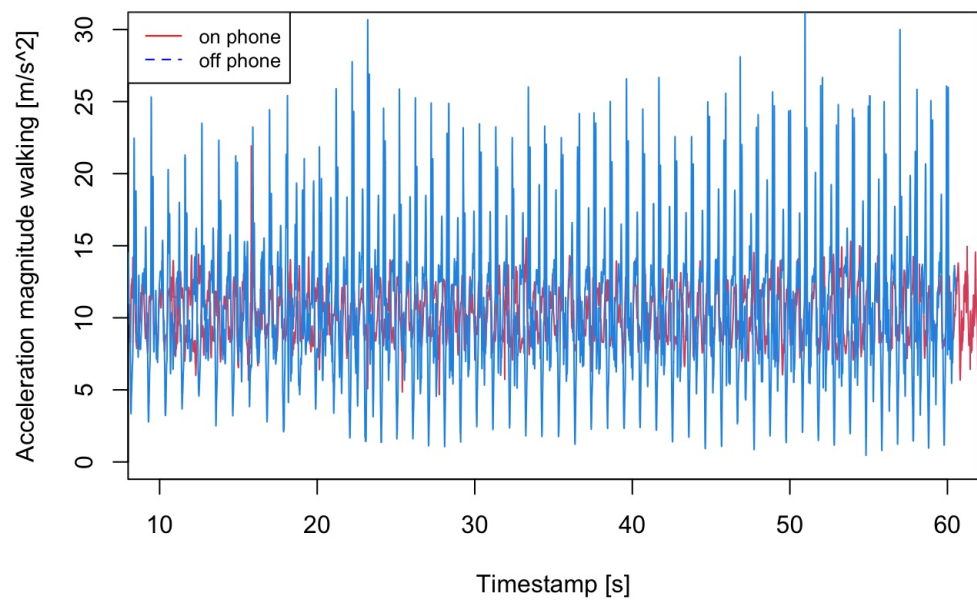
  # Plot sitting data
  {plot(sitting_phone$timestamp/1000, sitting_phone$acceleration,type = "l",col = "red",xlab = "Timestamp [s]",x
lim = c(10, 60), ylim=c(0,30),ylab = "Acceleration magnitude sitting [m/s^2]")
  lines(sitting_no_phone$timestamp/1000, sitting_no_phone$acceleration,col = "blue")
  legend("topleft", legend=c("on phone", "off phone"),col=c("red", "blue"), lty=1, cex=0.8)}

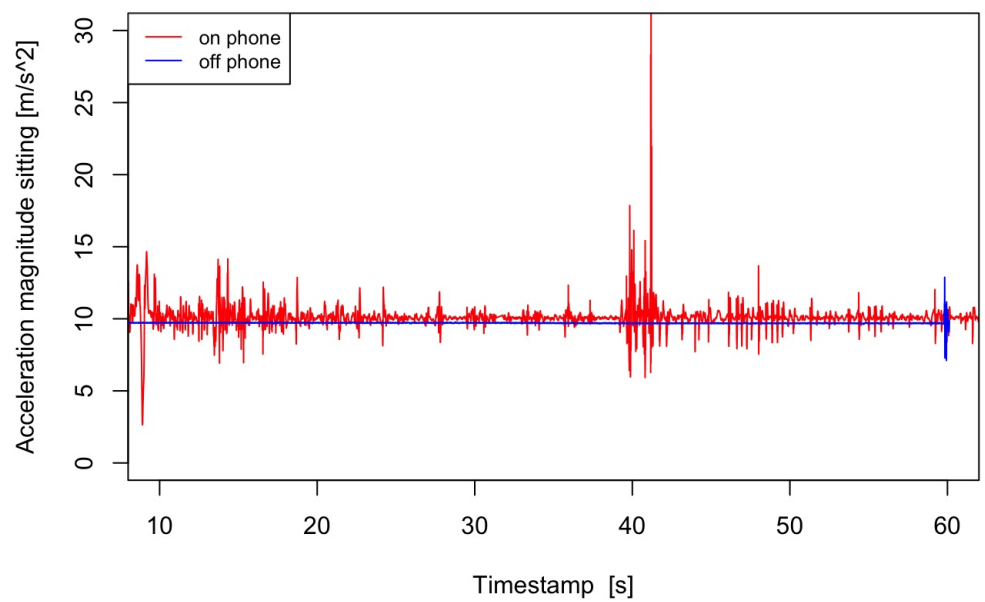
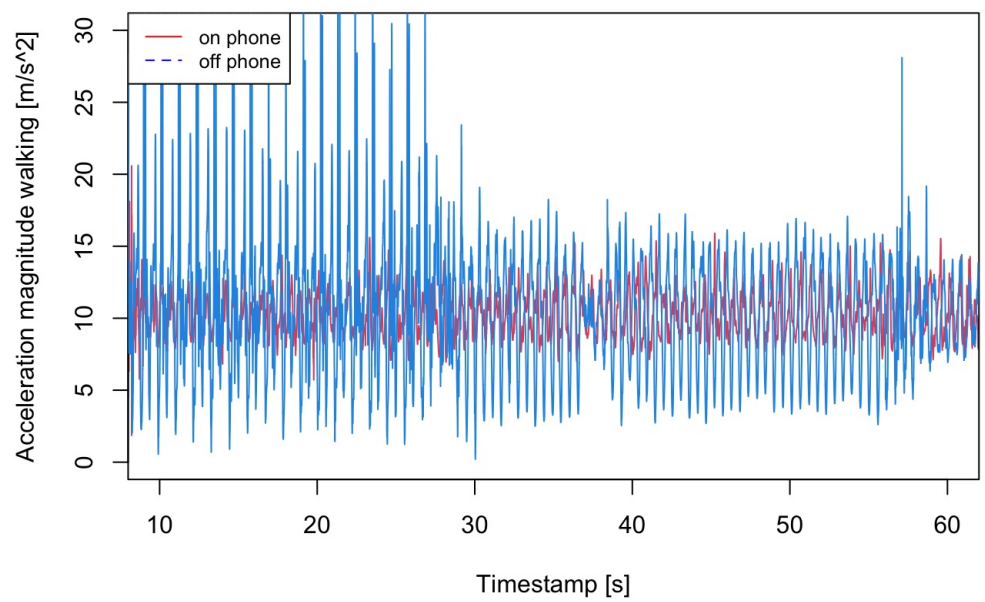
  # Plot walking phone data
  {plot(walking_phone$timestamp/1000, walking_phone$acceleration,type = "l",col = 2,xlab = "Timestamp [s]",xlim =
c(10, 60), ylim=c(0,30),ylab = "Acceleration magnitude walking [m/s^2]")
  lines(walking_no_phone$timestamp/1000, walking_no_phone$acceleration,col = 4)
  legend("topleft", legend=c("on phone", "off phone"),col=c("red", "blue"), lty=1:2, cex=0.8)}

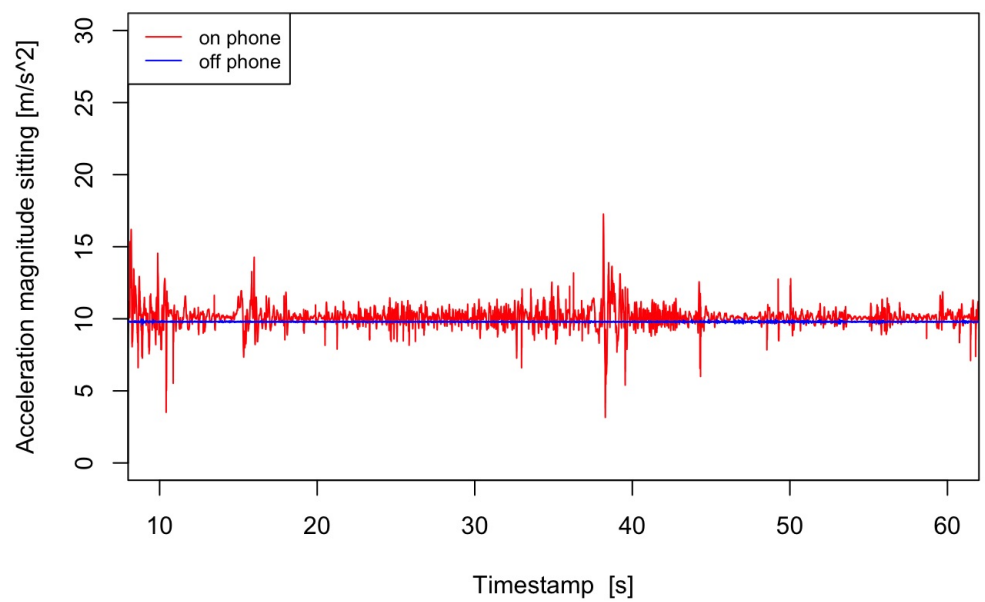
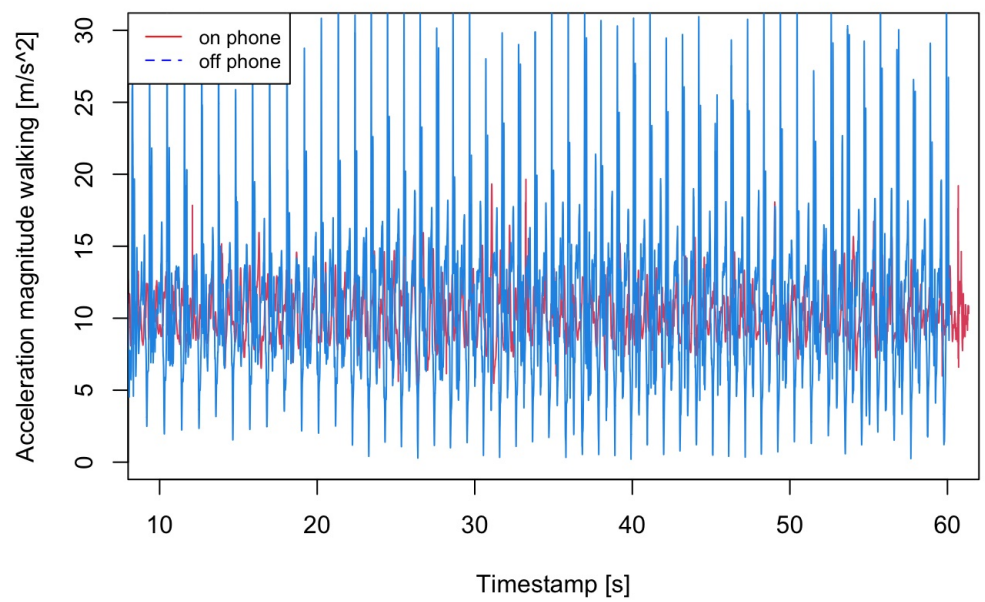
  # We will minimize the effects of starting phase on the statistical analysis by
  # eliminating the first 3000 timestamps:
  sitting_phone <- sitting_phone[-(1:3000),]
  sitting_no_phone <- sitting_no_phone[-(1:3000),]
  walking_phone <- walking_phone[-(1:3000),]
  walking_no_phone <- walking_no_phone[-(1:3000),]

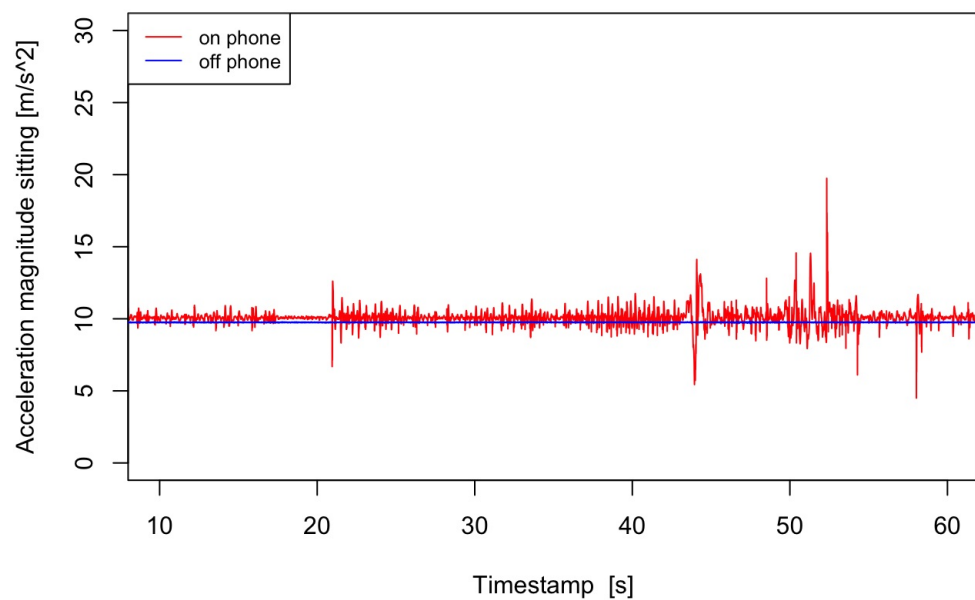
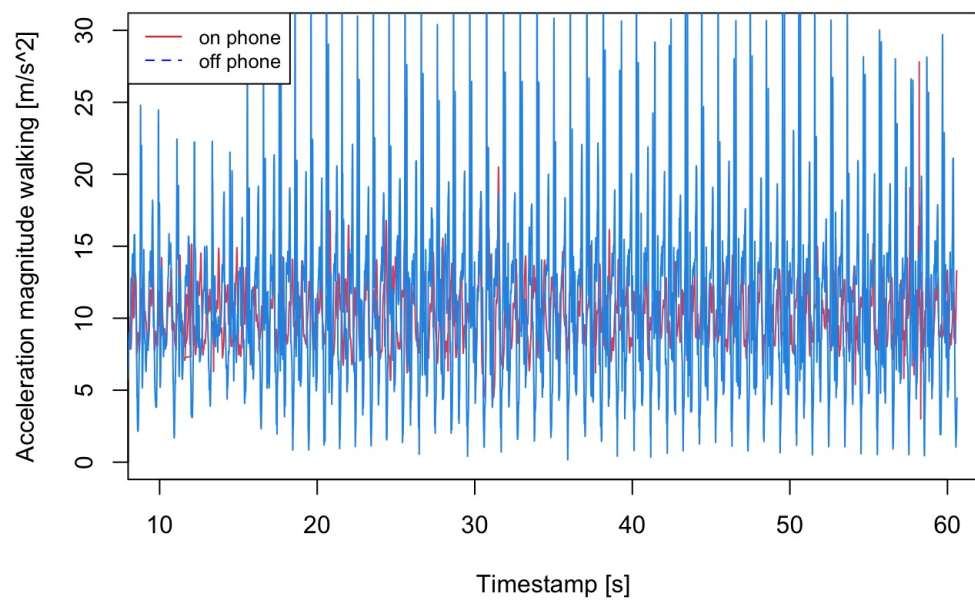
  # Add entry for this specific participant to testing data
  testing_dat[nrow(testing_dat) + 1,] = c(
    index,
    mean(sitting_phone$acceleration),
    mean(sitting_no_phone$acceleration),
    mean(walking_phone$acceleration),
    mean(walking_no_phone$acceleration)
  )
}
```

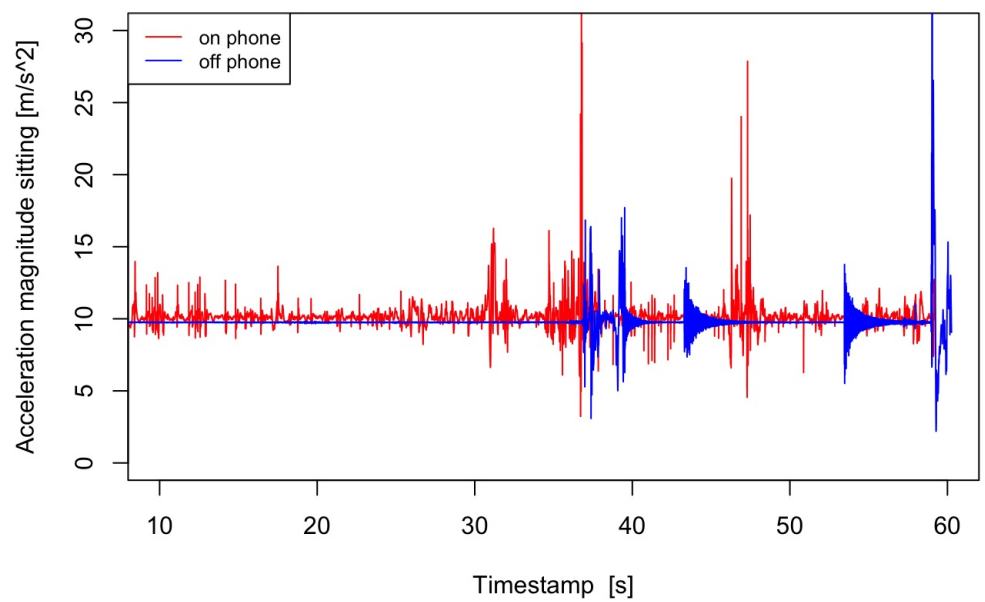
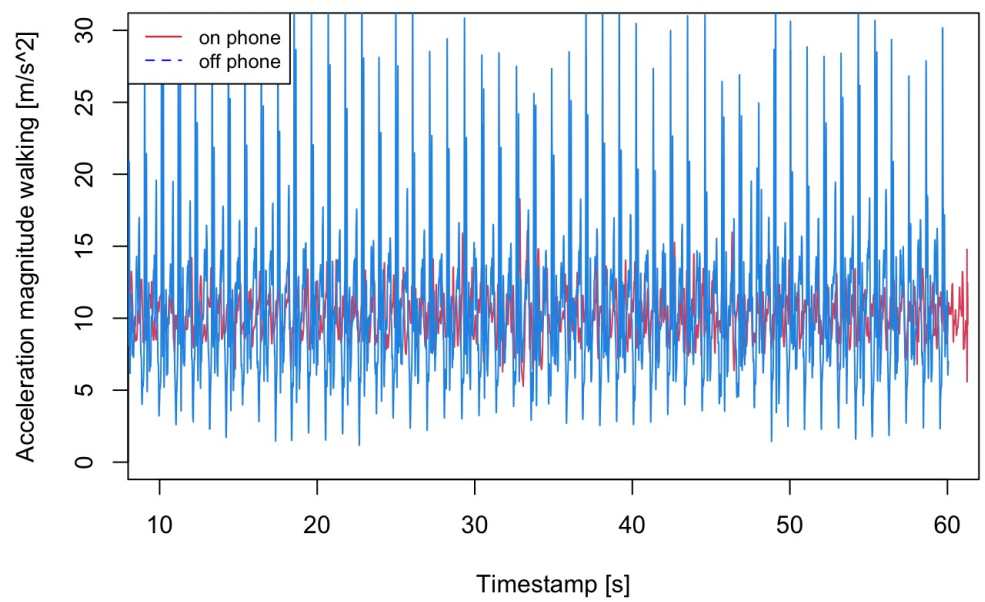


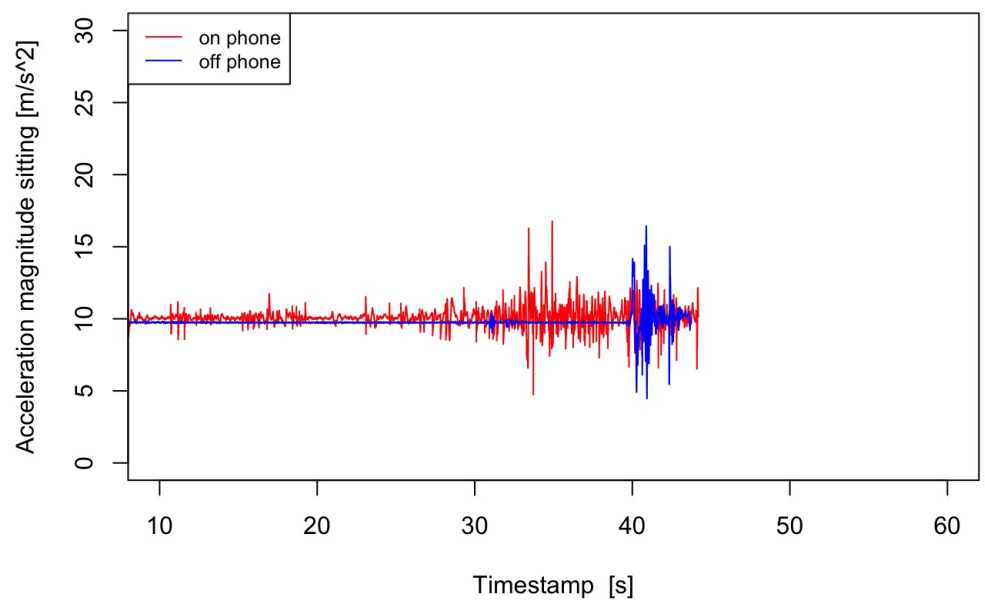
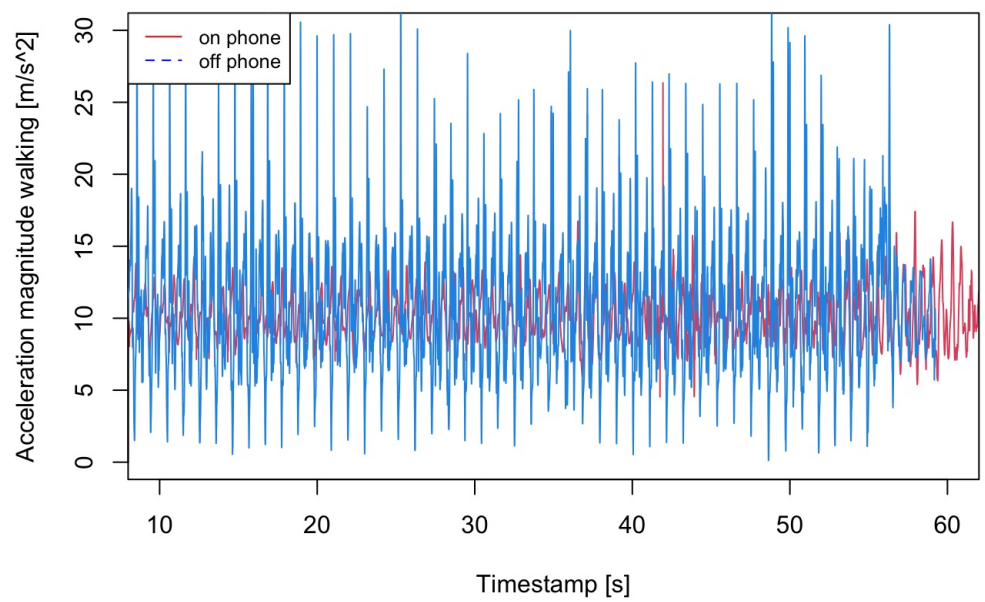


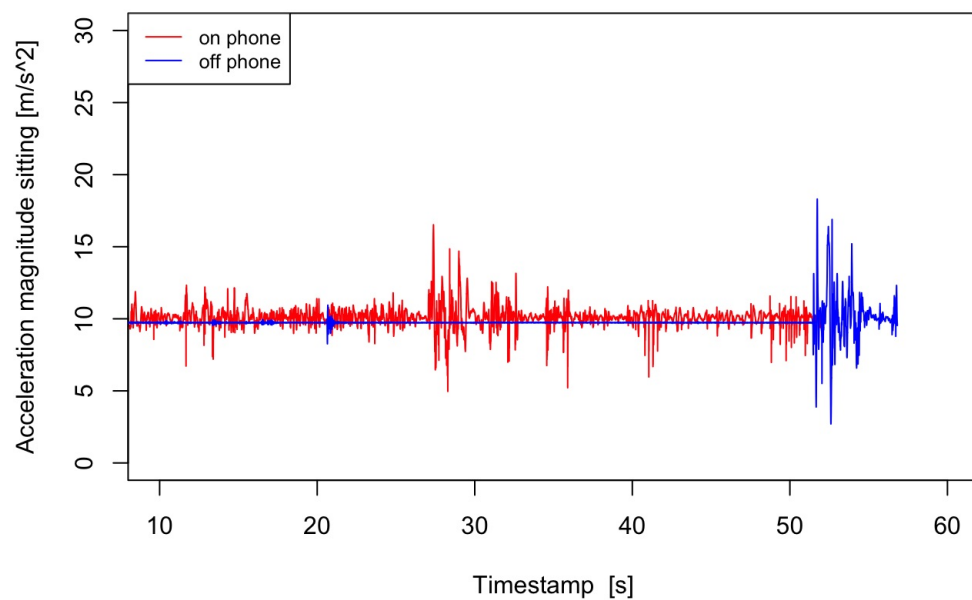
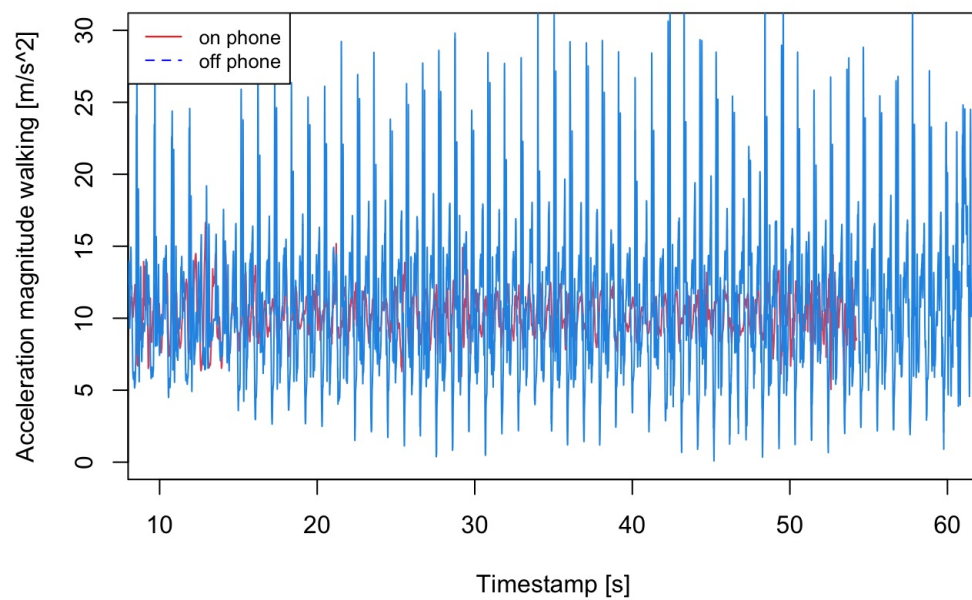


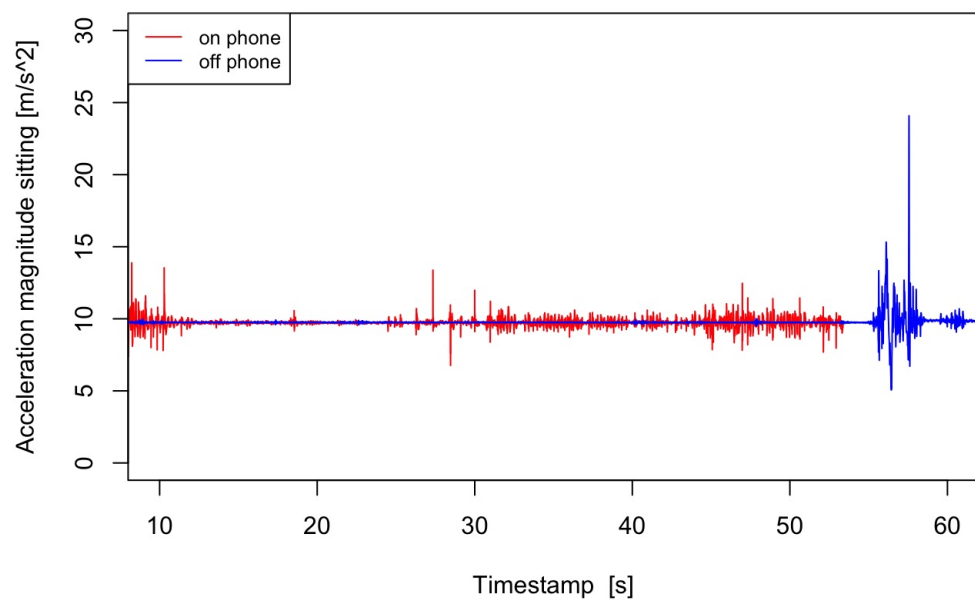
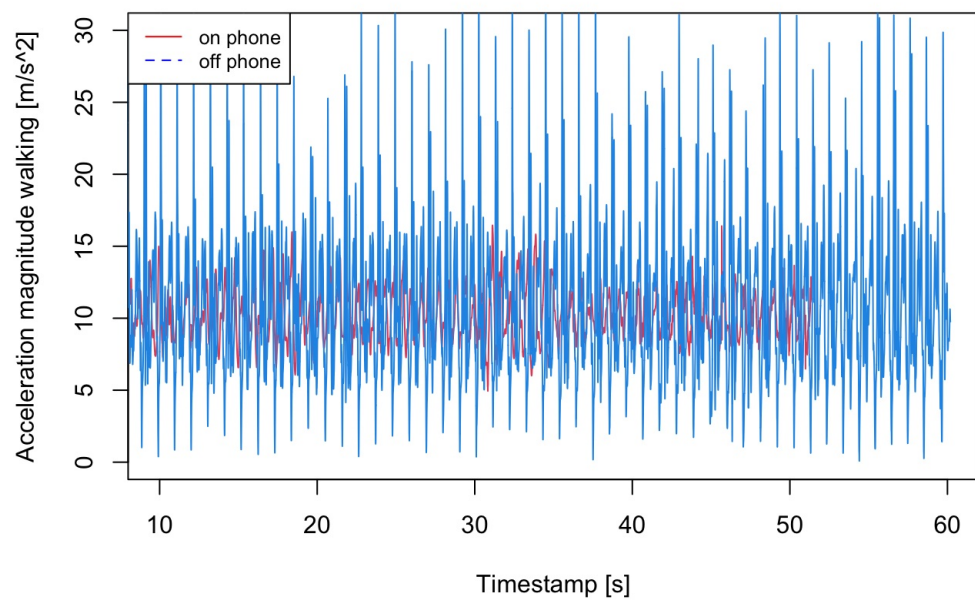


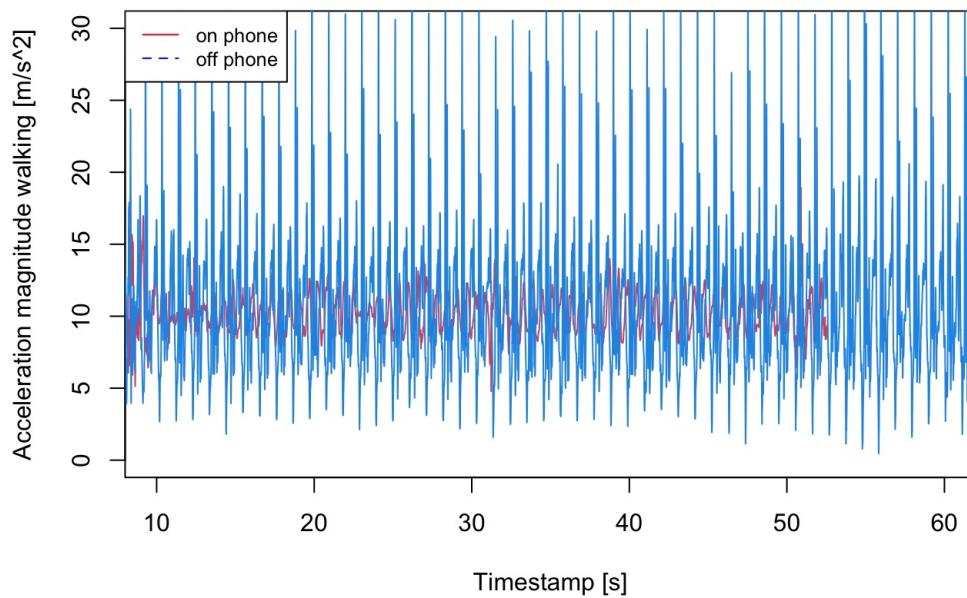










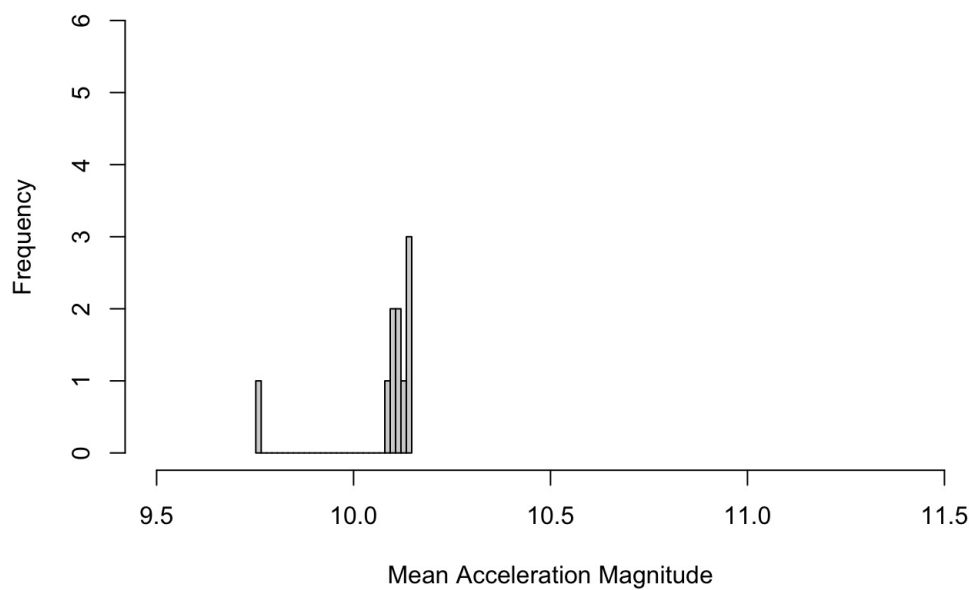


Normal Distribution Check for 10

Participants

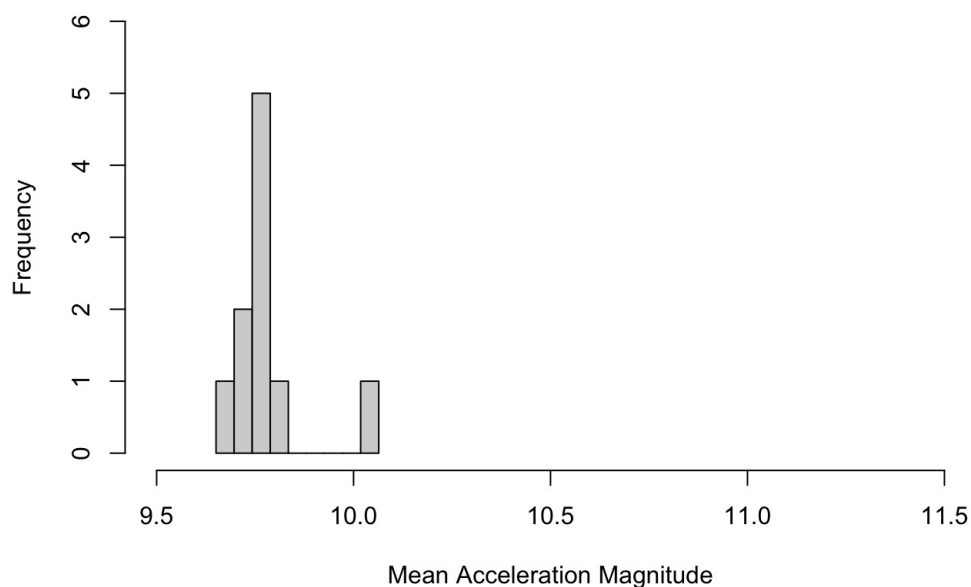
```
# Histogram Sit with phone
hist(testing_dat$SitP, breaks = seq(min(testing_dat$SitP), max(testing_dat$SitP), length.out = 30), ylim = c(0, 6), main = "Histogram of Sitting with Phone", xlab = "Mean Acceleration Magnitude", ylab = "Frequency", xlim = c(9.5, 11.5))
```

Histogram of Sitting with Phone



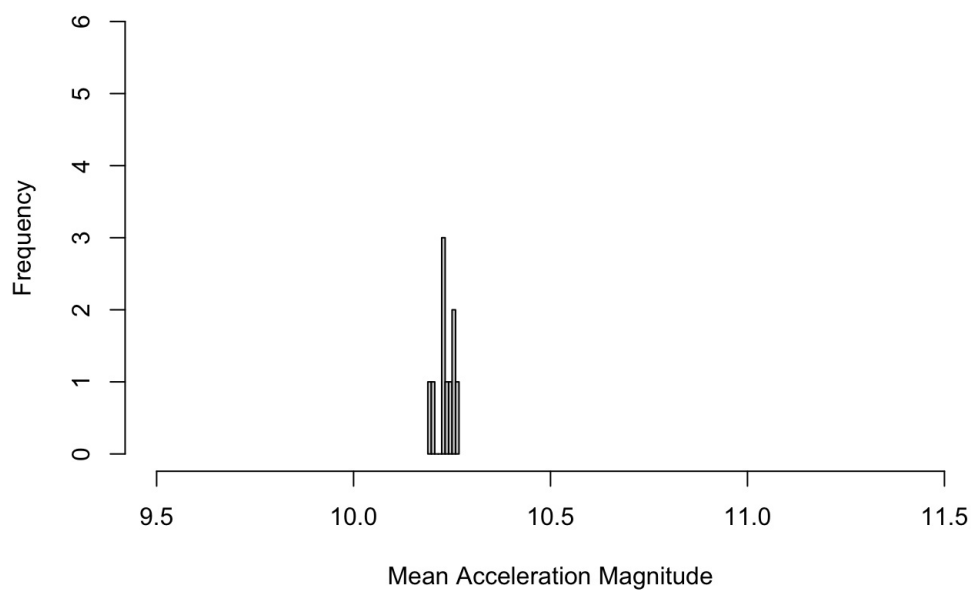
```
# Histogram Sit without phone
hist(testing_dat$SitNP, breaks = seq(min(testing_dat$SitNP), max(testing_dat$SitNP), length.out = 10), ylim = c(0, 6), main = "Histogram of Sitting without Phone", xlab = "Mean Acceleration Magnitude", ylab = "Frequency", xlim = c(9.5, 11.5))
```

Histogram of Sitting without Phone



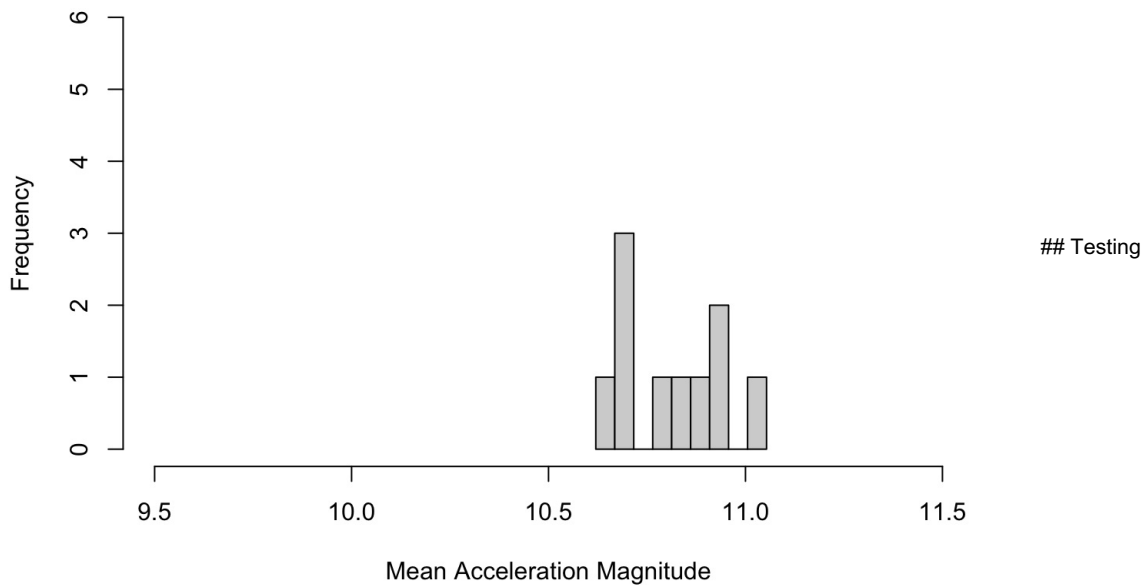
```
# Histogram Walk with phone
hist(testing_dat$WalkP, breaks = seq(min(testing_dat$WalkP), max(testing_dat$WalkP), length.out = 10), ylim = c(0, 6), main = "Histogram of Walking with Phone", xlab = "Mean Acceleration Magnitude", ylab = "Frequency", xlim = c(9.5, 11.5))
```

Histogram of Walking with Phone



```
#Histogram Walk without phone
hist(testing_dat$WalkNP, breaks = seq(min(testing_dat$WalkNP), max(testing_dat$WalkNP), length.out = 10), ylim = c(0, 6), main = "Histogram of Walking without Phone", xlab = "Mean Acceleration Magnitude", ylab = "Frequency", xlim = c(9.5, 11.5))
```

Histogram of Walking without Phone



Hypotheses: > When sitting, people induce more movement in their phone while using it than while not using it. > When walking, people induce less movement in their phone while using it than while not using it.

```
# For none of the categories we can really assume normality of accelerations
# from this small data. This very likely might change when the experiment is
# conducted with more people

# But for now this means we use Mann-Whitney test
# Test first hypothesis: Different movement sitting while using than not
wilcox.test(testing_dat$SitP, testing_dat$SitNP, alternative = "two.sided",paired=TRUE)
```

```
##
## Wilcoxon signed rank exact test
##
## data: testing_dat$SitP and testing_dat$SitNP
## V = 54, p-value = 0.003906
## alternative hypothesis: true location shift is not equal to 0
```

```
# We can reject Null-Hypothesis, and confirm there is substantially different
# movement in the phone while using it sitting than not using it sitting

wilcox.test(testing_dat$WalkP, testing_dat$WalkNP, alternative="less",paired=TRUE)
```

```
##
## Wilcoxon signed rank exact test
##
## data: testing_dat$WalkP and testing_dat$WalkNP
## V = 0, p-value = 0.0009766
## alternative hypothesis: true location shift is less than 0
```

```
# We can reject Null-Hypothesis, and confirm there is substantially higher
# movement in the phone while using it walking than not using it walking
```