STAT 847: Midterm Project

Chris Binoi Verghese ID:21092999

Q1. (20 points) Use the elements column of the **detailed dataset** to infer about the differences in blinking and jaw clenching between resting and non-resting states. Specifically, find differences in the distribution of the number of blinks and jaw clenches per minute during resting and active sessions.

```
dat_all = read.csv("C:/Users/chris/Desktop/MDSAI/STAT847/Midterm/Mind Monitor detailed data 2024-01-21.
dat_summary = ddply(dat_all, "sessionnum", summarise,
                    activity = activity[1],
                    session_time = session_time[1],
                    mean_alpha = mean(total_alpha, na.rm=TRUE),
                    mean_beta = mean(total_beta, na.rm=TRUE),
                    mean_gamma = mean(total_gamma, na.rm=TRUE),
                    mean_delta = mean(total_delta, na.rm=TRUE),
                    mean_theta = mean(total_theta, na.rm=TRUE),
                    var_alpha = var(total_alpha, na.rm=TRUE),
                    var_beta = var(total_beta, na.rm=TRUE),
                    var_gamma = var(total_gamma, na.rm=TRUE),
                    var_delta = var(total_delta, na.rm=TRUE),
                    var_theta = var(total_theta, na.rm=TRUE),
                    blinks_minute = Nblinks[1]/session_time[1]*60,
                    jaws_minute = Njaw[1]/session_time[1]*60,
                    mean_pos_xy = mean(sqrt(Accelerometer_X^2 + Accelerometer_Y^2),
                                       na.rm=TRUE),
                    mad_accel = mean(abs(Accelerometer_X^2 + Accelerometer_Y^2 + Accelerometer_Z^2 - 1)
                                     na.rm=TRUE),
                    rmse_gyro = mean(sqrt(Gyro_X^2 + Gyro_Y^2 + Gyro_Z^2), na.rm=TRUE)
)

dat_rest = dat_summary %>% filter(activity == 'resting')
dat_active = dat_summary %>% filter(activity != 'resting')
summary(dat_active[c("blinks_minute","jaws_minute")])
```

```
##  blinks_minute     jaws_minute
##  Min.   : 3.678   Min.   : 0.02834
##  1st Qu.:10.493   1st Qu.: 1.38188
##  Median :14.741   Median : 3.88198
##  Mean   :20.064   Mean   : 5.54173
##  3rd Qu.:23.947   3rd Qu.: 6.22323
##  Max.   :62.433   Max.   :27.55847
```

```
summary(dat_rest[c("blinks_minute","jaws_minute")])
```

```
##  blinks_minute      jaws_minute
##  Min.   : 0.1890   Min.   : 0.2835
##  1st Qu.: 0.5381   1st Qu.: 0.6015
##  Median : 1.1120   Median : 1.9553
##  Mean   : 9.7638   Mean   : 9.5730
```

```
##  3rd Qu.: 1.4153   3rd Qu.: 4.1449
##  Max.   :45.5648   Max.   :40.8800
```

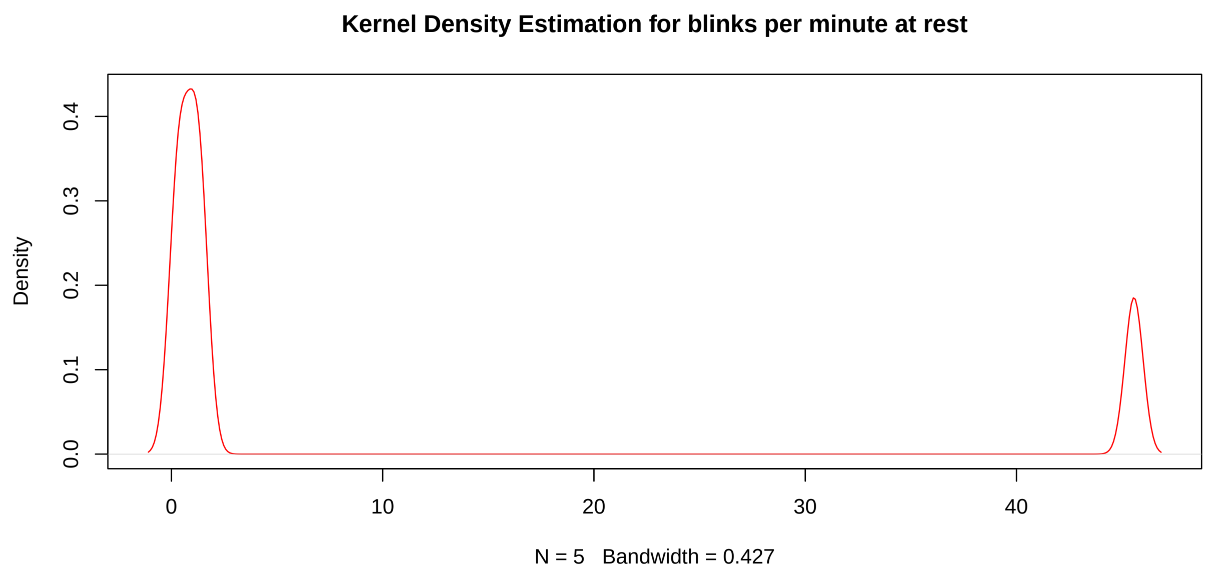
The summary analysis of the number of blinks and jaw clenches per minute reveals interesting patterns between resting and active sessions.

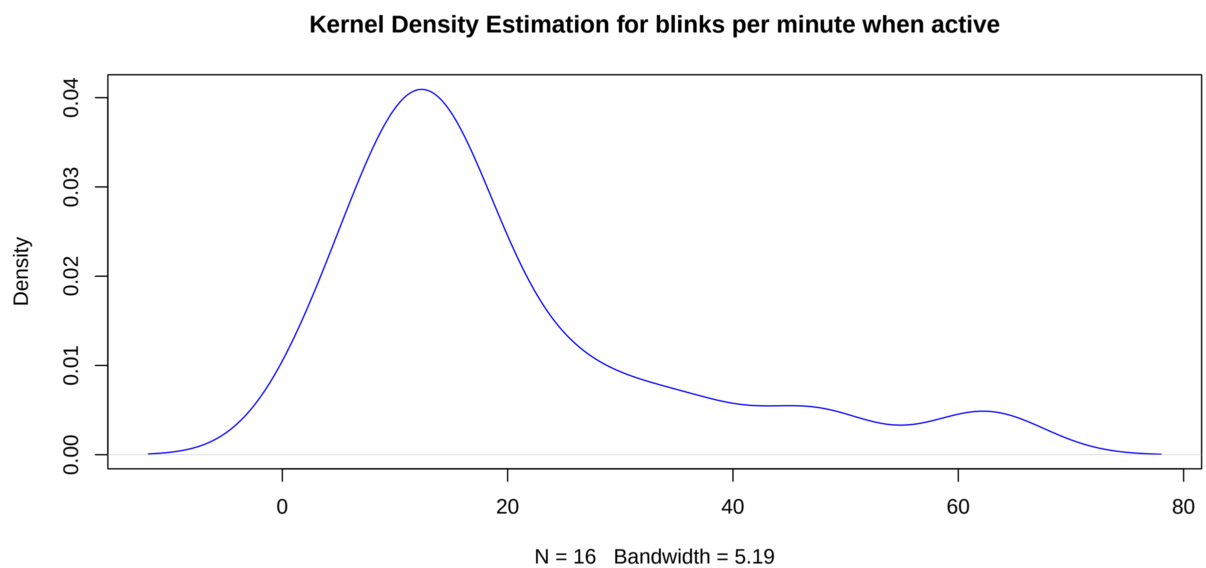
For the number of blinks per minute, the summary statistics show that during active sessions, there is a significantly higher frequency of blinks. This is evident from the smaller minimum, maximum, median, and the extremely small 3rd quantile value during resting sessions compared to the active sessions. The resting sessions exhibit a much narrower range of blink frequencies, suggesting a more consistent pattern during activity.

In contrast, the number of jaw clenches per minute follows a different trend. While the minimum and maximum values are higher during resting sessions, the overall median and 3rd quantile values are higher during active sessions. This indicates that the distribution of jaw clenches per minute is skewed by an extremely high maximum value during one resting session, while the distribution during active sessions is more stable.

To visualize these distributions, Kernel Density Estimation (KDE) is used, providing a smooth estimate of the probability density function of the data. This non-parametric method helps in understanding the distribution patterns more clearly, especially in cases where standard summary statistics might not capture the nuances of the data distribution.

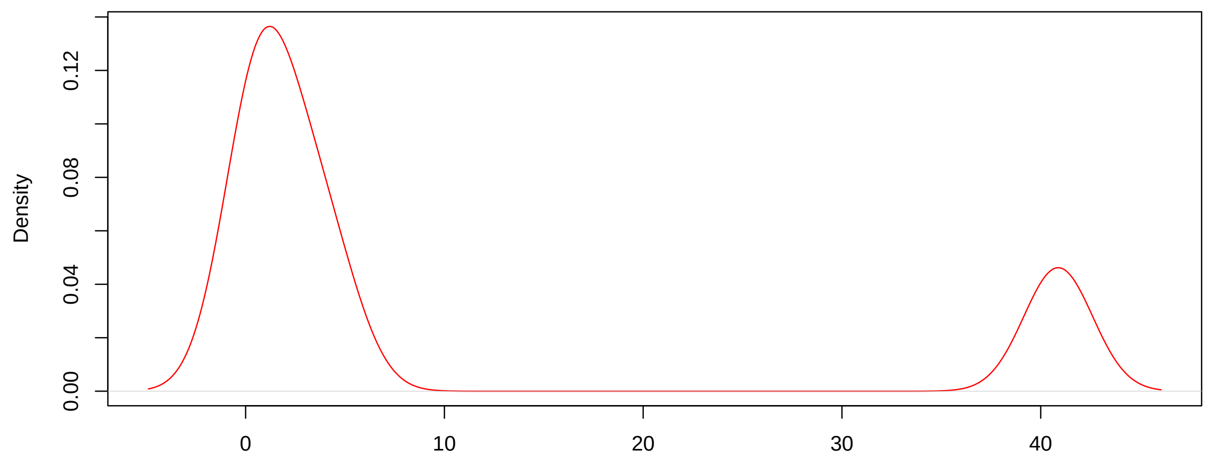
```r
# Set up a 2x2 layout for the plots
par(mfrow = c(2, 1))
plot(density(dat_rest$blinks_minute),
     main = "Kernel Density Estimation for blinks per minute at rest",
     col = "red")
plot(density(dat_active$blinks_minute),
     main = "Kernel Density Estimation for blinks per minute when active",
     col = "blue")
```





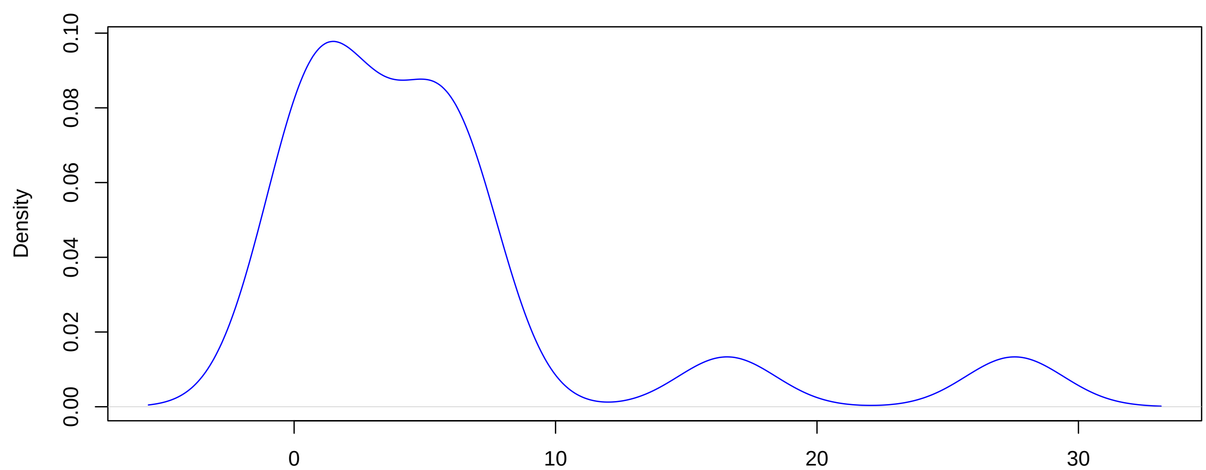

```
plot(density(dat_rest$jaws_minute),
     main = "Kernel Density Estimation for jaw clenches per minute at rest",
     col = "red")
plot(density(dat_active$jaws_minute),
     main = "Kernel Density Estimation for jaw clenches per minute when active",
     col = "blue")
```

Kernel Density Estimation for jaw clenches per minute at rest



N = 5   Bandwidth = 1.725

Kernel Density Estimation for jaw clenches per minute when active



N = 16   Bandwidth = 1.868

The Kernel Density Estimation (KDE) graphs provide insights into the frequency distribution of jaw clenches and blinks per minute during resting and active sessions. In resting sessions, both the number of jaw clenches and blinks per minute exhibit a wider range of values, with variations between very high and low frequencies.

In contrast, during active sessions, the KDE for blinks is more bell-shaped, indicating a more concentrated distribution around a specific frequency, while the distribution for jaw clenches during an active session resembles the distribution of blinks, suggesting a similar behavior pattern.

This analysis suggests that people tend to blink more frequently during active sessions compared to resting sessions. Additionally, the behavior of jaw clenching, though more defined in rest sessions, shows similarities between resting and active states.

Q2.(15 points) Use the data in the **summary dataset** to infer about the differences in brainwave activity between resting and non-resting states. Specifically, conduct a t-test of each of the ten brainwave variables to find any statistically significant differences. To get the p-value from a test, use `t.test(...)$p.value`.

```
#Read Summary dataset
dat_sum = read.csv("C:/Users/chris/Desktop/MDSAI/STAT847/Midterm/Mind Monitor session summaries 2024-01-
dat_rest = dat_sum %>% filter(activity == 'resting')
dat_active = dat_sum %>% filter(activity != 'resting')
# List of features for which you want to perform t-tests
features <- c("mean_alpha", "mean_beta", "mean_gamma", "mean_delta", "mean_theta",
              "var_alpha", "var_beta", "var_gamma", "var_delta", "var_theta")

# Create an empty data frame to store results
p_values_df <- data.frame(Feature = character(), P_Value = numeric(), stringsAsFactors = FALSE)

# Loop through each feature and perform t-test
for (feature in features) {
  p_value <- t.test(dat_rest[[feature]], dat_active[[feature]])$p.value
  p_values_df <- rbind(p_values_df, data.frame(Feature = feature, P_Value = p_value))
}

knitr::kable(p_values_df, col.names = c("brainwave variable", "p-value"),
             caption = "p-Values for brain wave means and variances")
```

Table 1: p-Values for brain wave means and variances

brainwave variable	p-value
mean_alpha	0.0686528
mean_beta	0.2228271
mean_gamma	0.0003942
mean_delta	0.4138743
mean_theta	0.5947658
var_alpha	0.2536182
var_beta	0.6361448
var_gamma	0.9870243
var_delta	0.6019243
var_theta	0.3555589

In our attempt to assess the similarity of data distribution for various brainwave variables (including "mean_alpha," "mean_beta," "mean_gamma," "mean_delta," "mean_theta," "var_alpha," "var_beta," "var_gamma," "var_delta," and "var_theta") derived from the summary dataset, which represents means and variances of different brainwaves over each session, we employed a Welch two-sample t-test. This test aimed to determine whether there are any statistically significant differences between the brainwave values during resting and active sessions.

Upon conducting the t-test, we discovered that the mean value of the total gamma waves is the only brainwave variable exhibiting a statistically significant difference. The calculated p-value for the t-test on the mean of gamma waves during active and resting sessions is 0.00039, an exceptionally low value. Generally, a p-value of 0.05 is considered a suitable threshold to assess whether a null hypothesis of similar means can be rejected. In comparison, the next closest p-value is for the mean of alpha waves, which is 0.0687—slightly lower than the other variables and in proximity to the 0.05 threshold.

Therefore, the brainwave variable that distinctly demonstrates statistically significant differences between resting and non-resting states is **mean gamma**.

Q3. (10 points) Are there any notable differences between the first 20% and the last 20% (by time_in) of active sessions? Look in the **detailed dataset** and isolate the first active session to find out. (The answer may be 'no'. In which case, mention some similarities). Compare at least three variables.

```
dat_active = dat_all %>% filter(activity != 'resting')
min_sessionnum <- min(dat_active$sessionnum)
# Subset the data frame for rows where sessionnum is the minimum and activity is resting
isolated_values <- dat_active[dat_active$sessionnum == min_sessionnum, ]
max_time = max(isolated_values$time_in)
first_20 <- isolated_values[isolated_values$time_in < max_time*0.2, ]
last_20 <- isolated_values[isolated_values$time_in > max_time*0.8, ]
p_values <- c(
  t.test(first_20$total_alpha, last_20$total_alpha)$p.value,
  t.test(first_20$total_beta, last_20$total_beta)$p.value,
  t.test(first_20$total_gamma, last_20$total_gamma)$p.value,
  t.test(first_20$total_delta, last_20$total_delta)$p.value,
  t.test(first_20$total_theta, last_20$total_theta)$p.value
)

# Create a data frame to display the p-values
results_df <- data.frame(
  Brainwave_Band = c("total_alpha", "total_beta", "total_gamma", "total_delta", "total_theta"),
  P_Value = p_values
)
knitr::kable(results_df, caption = "p-Values for data")
```

Table 2: p-Values for data

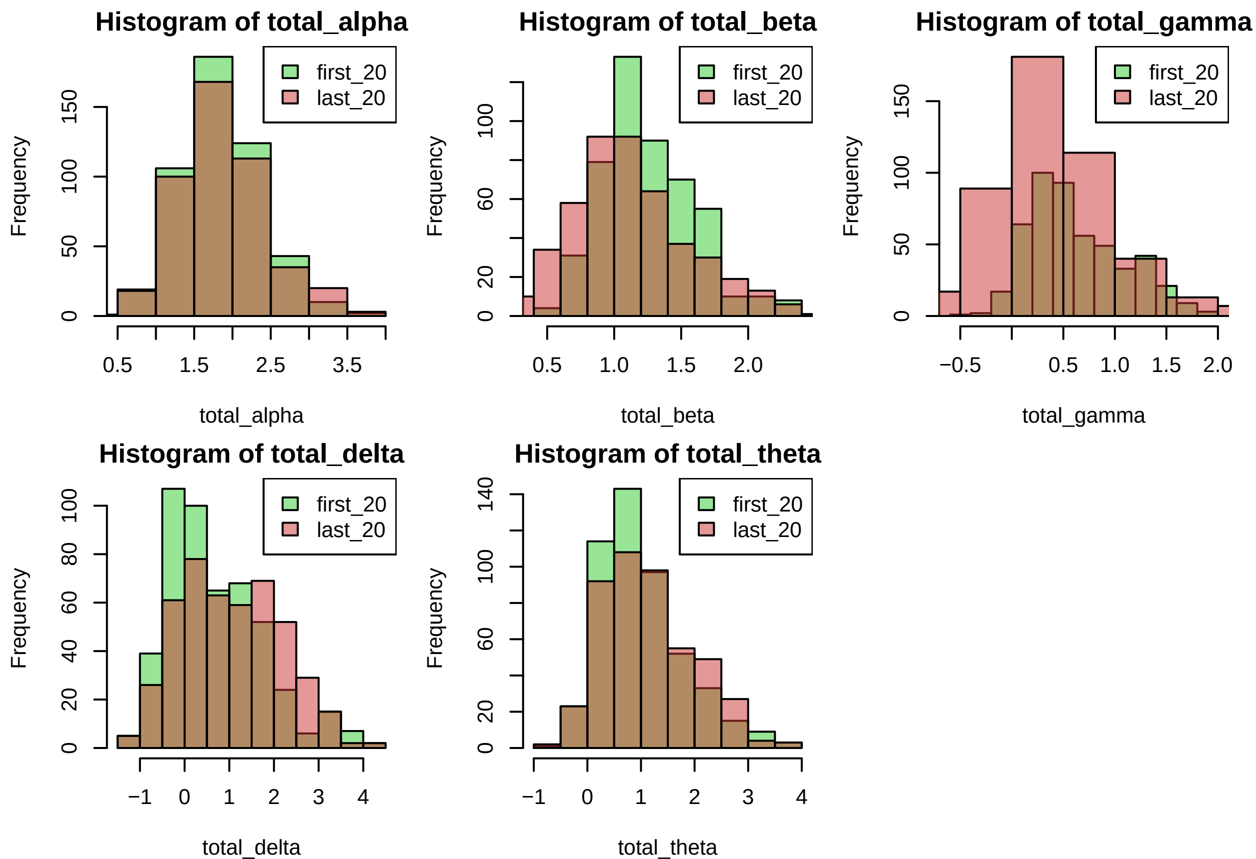
Brainwave_Band	P_Value
total_alpha	0.8515147
total_beta	0.0000001
total_gamma	0.0000000
total_delta	0.0000021
total_theta	0.0114208

The p-values obtained from a Welch two-sample t-test provide insights into the statistically significant differences between the first 20% and last 20% of the first active session, considering the variables 'total_alpha', 'total_beta', 'total_gamma', 'total_delta', and 'total_theta'. Notably, there is a discernible difference between the datasets for all brainwaves, except for total alpha brain waves.

The standard p-value threshold considered here is 0.05. However, most of the brainwaves do not approach this value, except for total alpha and theta brainwaves. Total theta waves have a p-value of 0.0114, while total alpha waves exhibit a reasonably higher p-value. To visualize these changes more clearly, histograms of these values can be examined.

```
column_names <- c('total_alpha', 'total_beta', 'total_gamma', 'total_delta', 'total_theta')

# Set up the layout for the plot
par(mfrow = c(2, 3), mar = c(4, 4, 2, 1))

# Loop through each column and create overlaid histograms
for (col in column_names) {
  hist(first_20[[col]], col = rgb(0.2, 0.8, 0.2, alpha = 0.5),
       main = paste("Histogram of", col), xlab = col,
```

```r
       ylim = c(0, max(hist(first_20[[col]], plot = FALSE)$counts,
                       hist(last_20[[col]], plot = FALSE)$counts)))
  hist(last_20[[col]], col = rgb(0.8, 0.2, 0.2, alpha = 0.5), add = TRUE)
  legend("topright", legend = c("first_20", "last_20"),
         fill = c(rgb(0.2, 0.8, 0.2, alpha = 0.5), rgb(0.8, 0.2, 0.2, alpha = 0.5)))
}
```



The plotted histograms reveal that the distribution of total alpha waves exhibits an extremely good overlap, indicating very similar means. However, this is not the case for the rest of the total values, except for total theta waves, which approach a similar mean. The distribution of total beta waves for the last 20% is slightly to the left and has a lower frequency compared to the first 20%. In contrast, the first 20% has a distribution with very similar but much smaller frequency peaks compared to the last 20%. Total delta waves show almost two max frequency peaks, suggesting a very different distribution, while total theta has a distribution where the mean of the last 20% is slightly to the right of the first 20%.

In summary, the most statistically similar distributions are observed in the first and last 20% of **total alpha waves** for the first active session.

Q4. (10 points) Are there any notable differences between the first 20% and the last 20% (by time_in) of resting sessions? Look in the **detailed dataset** and isolate the first resting session to find out. (The answer may be 'no'. In which case, mention some similarities). Compare at least three variables.

```r
dat_rest = dat_all %>% filter(activity == 'resting')
min_sessionnum <- min(dat_rest$sessionnum)
# Subset the data frame for rows where sessionnum is the minimum and activity is resting
isolated_values <- dat_rest[dat_rest$sessionnum == min_sessionnum, ]
max_time = max(isolated_values$time_in)
first_20 <- isolated_values[isolated_values$time_in < max_time*0.2, ]
last_20 <- isolated_values[isolated_values$time_in > max_time*0.8, ]
p_values <- c(
  t.test(first_20$total_alpha, last_20$total_alpha)$p.value,
  t.test(first_20$total_beta, last_20$total_beta)$p.value,
  t.test(first_20$total_gamma, last_20$total_gamma)$p.value,
  t.test(first_20$total_delta, last_20$total_delta)$p.value,
  t.test(first_20$total_theta, last_20$total_theta)$p.value
)

# Create a data frame to display the p-values
results_df <- data.frame(
  Brainwave_Band = c("total_alpha", "total_beta", "total_gamma", "total_delta", "total_theta"),
  P_Value = p_values
)
knitr::kable(results_df, caption = "p-Values for data")
```

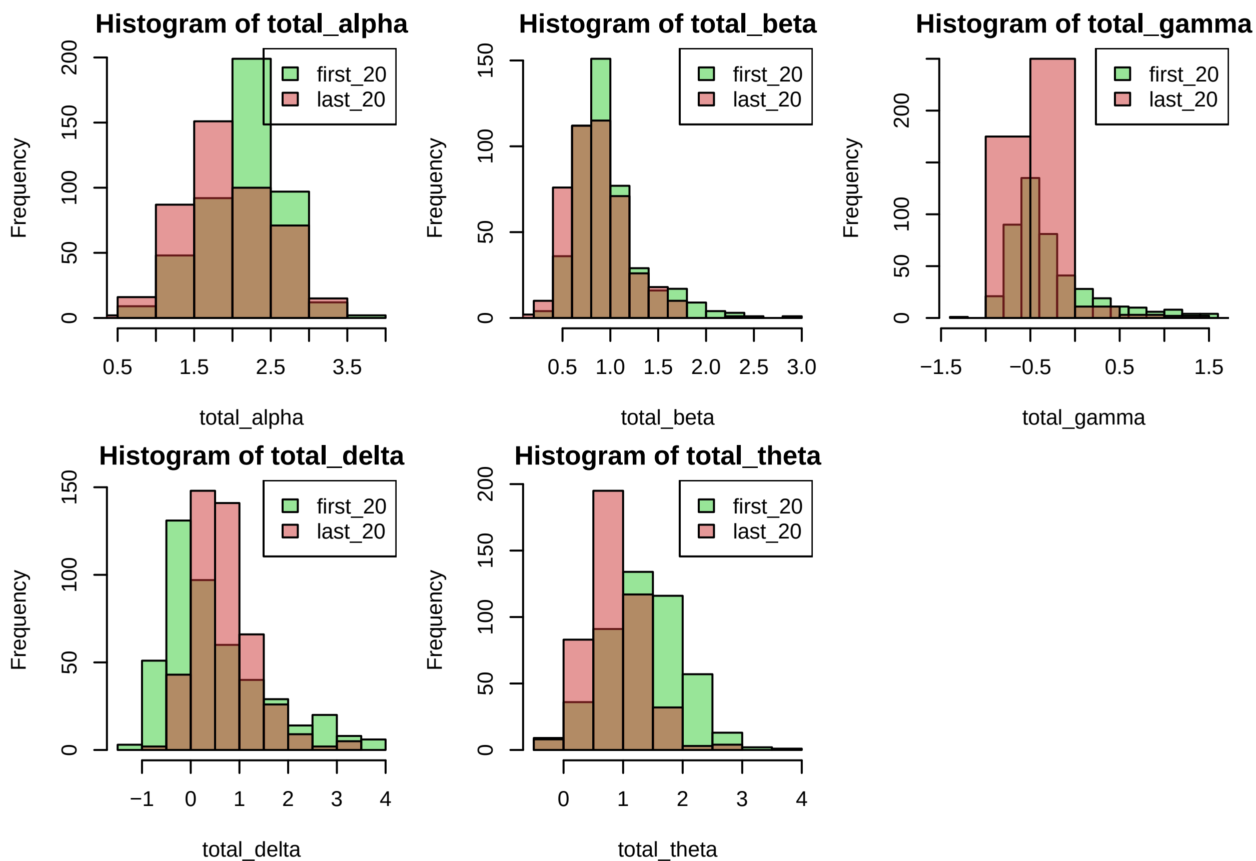
Table 3: p-Values for data

Brainwave_Band	P_Value
total_alpha	0.0000000
total_beta	0.0000096
total_gamma	0.0000022
total_delta	0.0015565
total_theta	0.0000000

In resting sessions, the Welch two-sample t-test reveals significant differences in the distributions of "total_alpha," "total_beta," "total_gamma," "total_delta," and "total_theta" values. None of the p-values approach the set threshold of 0.05, with the closest being for total delta waves at 0.00156. Histograms visually illustrate these distinctions in wave distributions.

```r
column_names <- c('total_alpha', 'total_beta', 'total_gamma', 'total_delta', 'total_theta')

# Set up the layout for the plot
par(mfrow = c(2, 3), mar = c(4, 4, 2, 1))

# Loop through each column and create overlaid histograms
for (col in column_names) {
  hist(first_20[[col]], col = rgb(0.2, 0.8, 0.2, alpha = 0.5),
       main = paste("Histogram of", col), xlab = col,
       ylim = c(0, max(hist(first_20[[col]], plot = FALSE)$counts,
                       hist(last_20[[col]], plot = FALSE)$counts)))
  hist(last_20[[col]], col = rgb(0.8, 0.2, 0.2, alpha = 0.5), add = TRUE)
  legend("topright", legend = c("first_20", "last_20"),
```

```
        fill = c(rgb(0.2, 0.8, 0.2, alpha = 0.5), rgb(0.8, 0.2, 0.2, alpha = 0.5)))
}
```



The distributions of total alpha and beta waves show a slight leftward shift in the last 20% of the session compared to the first 20%. Similarly, total gamma waves exhibit a more spread-out distribution in the first 20%, concentrating between 0 and -0.75 in the last 20%. Total delta waves display a rightward shift in frequency distribution in the last 20%, while total theta waves reach a maximum frequency at a lower total gamma wave value in the last 20% compared to the first 20%.

In summary, the total brainwave distributions in the first and last 20% of the resting session exhibit various shifts and differences, and there are no statistically significant similarities between the two segments.

Q5. (5 points) Use the **detailed dataset** Make a variable in the summary data that gets the Pearson correlation between Alpha_TP9 and Alpha_TP10 during a session. (That is, the alpha waves above the left and right ear, respectively). Compare the TP9-TP10 correlations during in resting and non-resting activities.

```r
cor_summary <- dat_all %>%
  na.omit() %>%  # Remove rows with NA values
  group_by(sessionnum) %>%
  summarise(
    activity = first(activity),
    cor_value = {
      # Check for zero standard deviations before calculating correlation
      if (sd(Alpha_TP9) != 0 & sd(Alpha_TP10) != 0) {
        cor(Alpha_TP9, Alpha_TP10)
      } else {
        NA  # Set correlation to NA if standard deviation is zero
      }
    }
  )
knitr::kable(cor_summary,
             col.names = c("session", "activity", "correlation"),
             caption = "Data Summary with Correlation of Alpha")
```

Table 4: Data Summary with Correlation of Alpha

session	activity	correlation
1	active	0.3867726
2	resting	0.6026397
3	active	0.4108271
4	gaming	NA
5	reading	0.2833611
6	active	0.3209793
7	resting	0.6506334
8	active	0.4457501
9	watchingtv	0.4022244
10	active	0.0078191
11	resting	0.5979053
12	meeting	0.3796727
13	active	0.4638552
14	resting	0.6646884
15	resting	0.7028927
16	active	0.4064373
17	active	0.5041046
18	music	0.4039509
19	active	0.4928117
20	active	0.3792405
21	active	0.1861603

The correlation between alpha brainwaves over the left and right ears provides insights into their directional relationship at any given time. During resting sessions, there is a notable increase in correlation compared to active sessions. It is important to note that during active sessions, the correlation does not follow a clear trend, as the nature of the activities during these sessions impacts the direction of the alpha waves.

Notably, the presence of NA values in the correlation during active sessions, particularly when the user was gaming, suggests instances where the sensor was unable to capture the alpha brainwaves around the ears accurately. This indicates potential challenges in reading alpha brainwave data during certain activities.

Q6. (5 points) The accelerometer always shows some acceleration because of gravity. (When completely upright and still, it shows accelerometer Z = 1, X = 0, Y = 0) You can use this to determine the relative position of the head, as well as how much the head moves. Use the acceleration data in the **detailed dataset** to infer about the differences in both the position of the head, and its movements, between resting and non-resting states.

e.g., you final answer should be something like "Head position was more __________ when resting than when active, and we know this because the accelerometer ___________ instead of ______________", with code and supporting summary information.

```
dat_accelerometer <- dat_all %>%
  na.omit() %>%
  group_by(sessionnum) %>%
  summarise(
    activity = first(activity),
    mean_X = mean(Accelerometer_X),
    mean_Y = mean(Accelerometer_Y),
    mean_Z = mean(Accelerometer_Z)
  )
knitr::kable(dat_accelerometer,
             caption = "Accelerometer Information for every session")
```

Table 5: Accelerometer Information for every session

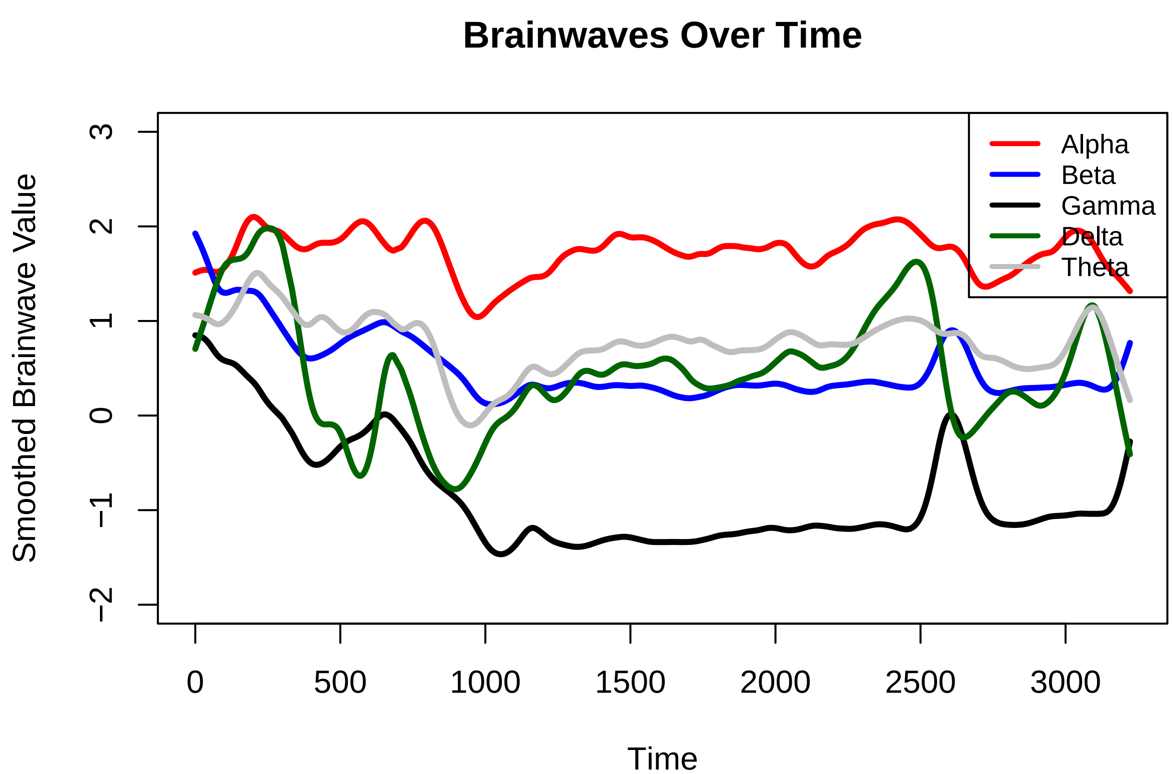
sessionnum	activity	mean_X	mean_Y	mean_Z
1	active	-0.1452983	-0.0102457	0.9660075
2	resting	-0.6368529	0.1014819	0.7065703
3	active	0.0491192	-0.0608379	0.9802723
4	gaming	0.0948194	-0.0337791	0.9794287
5	reading	-0.0504814	0.0675906	0.9499806
6	active	0.0107961	-0.0371245	0.9704976
7	resting	-0.7220977	0.2166379	0.6322272
8	active	0.0581812	-0.0979612	0.9698474
9	watchingtv	0.0173756	-0.0254898	0.9641957
10	active	0.0303966	-0.0344628	0.9776636
11	resting	-0.8536742	0.2964188	0.2342812
12	meeting	-0.0808475	-0.1754744	0.9684092
13	active	0.0490723	-0.0020953	0.9709905
14	resting	-0.8100607	0.4076965	0.3938879
15	resting	-0.6054954	0.5831397	0.4872128
16	active	0.0344518	-0.0405797	0.9490575
17	active	-0.0369643	-0.0188115	0.9797044
18	music	0.1412012	0.0600831	0.9593866
19	active	-0.0051881	-0.0033064	0.9750845
20	active	-0.1331992	-0.0192631	0.9731560
21	active	0.0996670	-0.0895152	0.9676931

Head position was more towards the bottom right while being lower when resting than when active, and we know this because the accelerometer readings had a mean X value of about negative -0.6 while the mean z value was lower and the mean y value is always positive and more than 0.1 instead of in the case of being active where the z-value tends to be about 0.9 most of the time while the x-value tends to have a minimum of -0.145 which is more than any other x-value when active while the y value has a maximum right tilt of 0.068 while active.

From these observations, we can infer that the user's resting position involves tilting the head to the right and lowering it. This is supported by the lower Z values and more negative X values and more positive y values recorded by the accelerometer during rest sessions while compared to active sessions.

Q7. (10 points) In the **detailed dataset**, plot the smoothed brainwaves for all five bands (alpha, beta, gamma, delta, and theta) session number 11 (resting, 3222 seconds long). Use any smoothing technique we discussed in class. Make a guess at when the patient fell asleep.

```r
# Subset data for session number 11
dat_11 <- dat_all %>% filter(sessionnum == 11  & !is.na(Delta_TP9))

## Smooth over time
ss_alpha = smooth.spline(dat_11$time_in, dat_11$total_alpha, spar=0.6)
ss_beta = smooth.spline(dat_11$time_in, dat_11$total_beta, spar=0.6)
ss_gamma = smooth.spline(dat_11$time_in, dat_11$total_gamma, spar=0.6)
ss_delta = smooth.spline(dat_11$time_in, dat_11$total_delta, spar=0.6)
ss_theta = smooth.spline(dat_11$time_in, dat_11$total_theta, spar=0.6)
plot(ss_alpha, type='n', col="White", lwd=3, ylim=c(-2,3),
     main="Brainwaves Over Time", xlab = "Time", ylab = "Smoothed Brainwave Value")
lines(ss_alpha, type='l', col="Red", lwd=3)
lines(ss_beta, type='l', col="Blue", lwd=3)
lines(ss_gamma, type='l', col="Black", lwd=3)
lines(ss_delta, type='l', col="Darkgreen", lwd=3)
lines(ss_theta, type='l', col="Grey", lwd=3)
# Add legend
legend("topright", legend = c("Alpha", "Beta", "Gamma", "Delta", "Theta"),
       col = c("Red", "Blue", "Black", "Darkgreen", "Grey"), lty = 1, lwd = 2.5, cex = 0.8)
```



Here is the smoothed representation of the patient's brainwaves during session 11. A notable observation is a significant dip in all brainwave frequencies, except for gamma waves, occurring at approximately **750 seconds**. This distinctive pattern strongly suggests that the patient entered a sleep state around this time.

Another supporting factor for this assumption is the remarkably stable brainwave activity observed shortly after the onset of sleep, in contrast to the fluctuations observed before falling asleep.

Q8.(10 points) Note that total_alpha and its similar functions have a flaw - measurements that were missing because of a bad connection are marked as zero, while true values are rarely zero. Find another measure using the **detailed dataset**, mean_alpha, that takes the average of **all the non-zero alpha measurements**. Take the mean of this new measurement for each session. For example, the average of mean_alpha for all of session 1 as a single value, all of session 2 as a single value, and so on.

```r
detailed_data <- dat_all %>%
  mutate(
    counter = (Alpha_TP9 != 0) + (Alpha_AF7 != 0) + (Alpha_AF8 != 0) + (Alpha_TP10 != 0),
    mean_alpha = (Alpha_TP9 + Alpha_AF7 + Alpha_AF8 + Alpha_TP10) / counter
  )

results <- detailed_data %>%
  group_by(sessionnum) %>%
  summarise(
    activity  = first(activity),
    average_alpha = mean(mean_alpha, na.rm = TRUE)
  )

knitr::kable(results,
             col.names = c("session", "activity", "Mean Alpha"),
             caption = "Data Summary for total means")
```

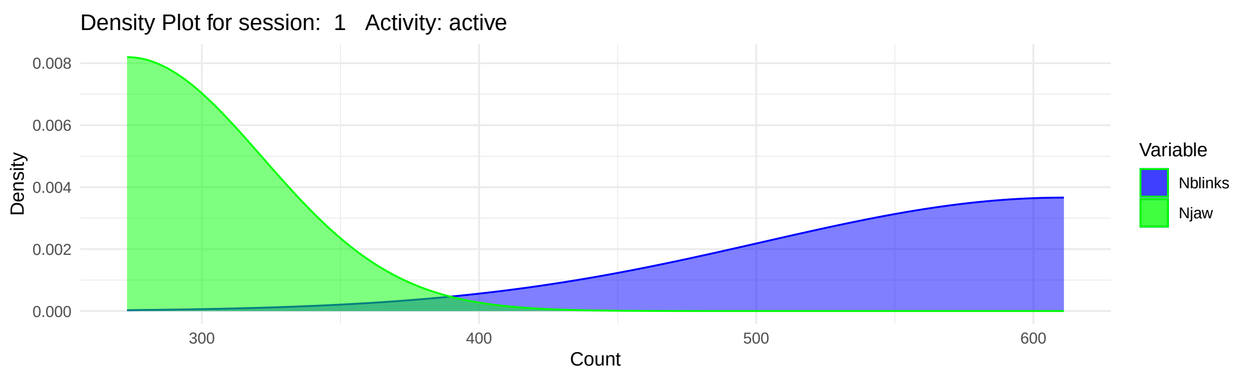
Table 6: Data Summary for total means

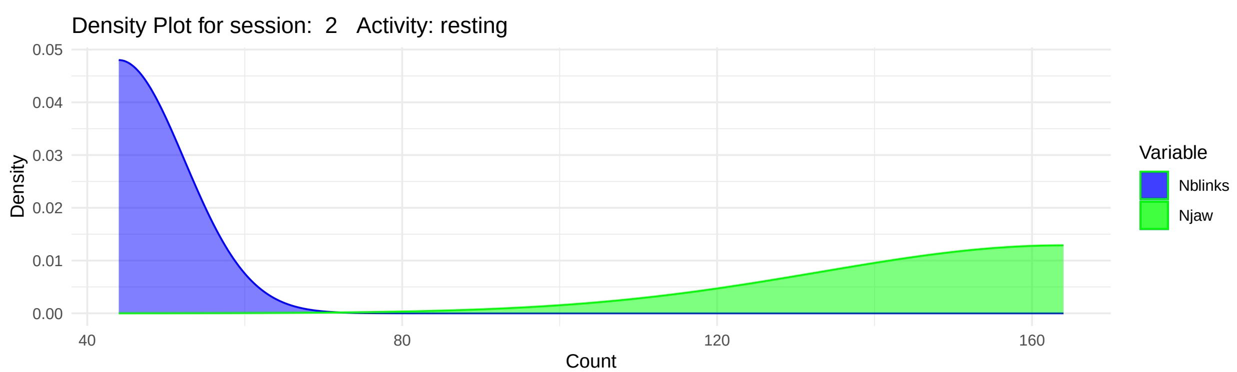
session	activity	Mean Alpha
1	active	0.6280909
2	resting	0.6741808
3	active	0.4848200
4	gaming	0.3979257
5	reading	0.5274557
6	active	0.9114997
7	resting	0.6000797
8	active	0.4842123
9	watchingtv	0.6282106
10	active	0.4183184
11	resting	0.5784362
12	meeting	0.6802712
13	active	0.6920264
14	resting	0.7764909
15	resting	0.7199625
16	active	0.6499894
17	active	0.5108848
18	music	0.5266167
19	active	0.6533709
20	active	0.6452325
21	active	0.6132322

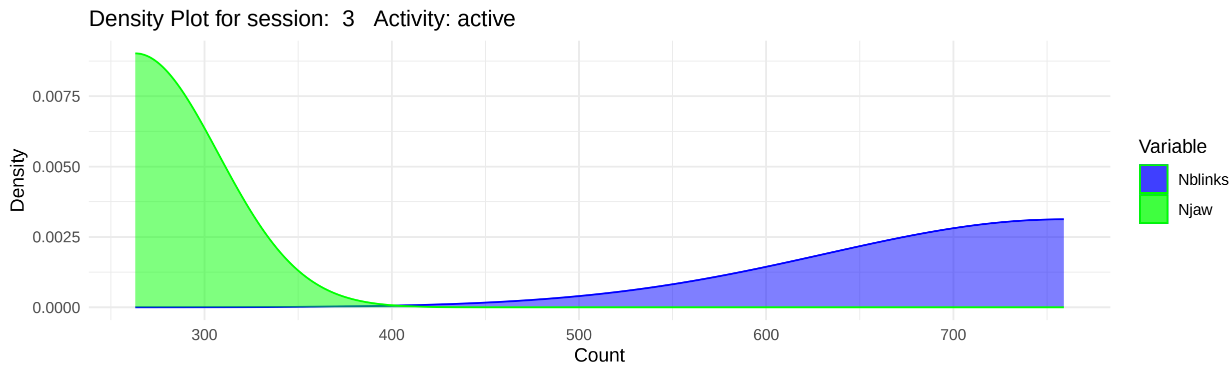
Q9.(15 points) Find two interesting patterns in the data. These could be almost anything. You can use **either the detailed or the summary datasets**.

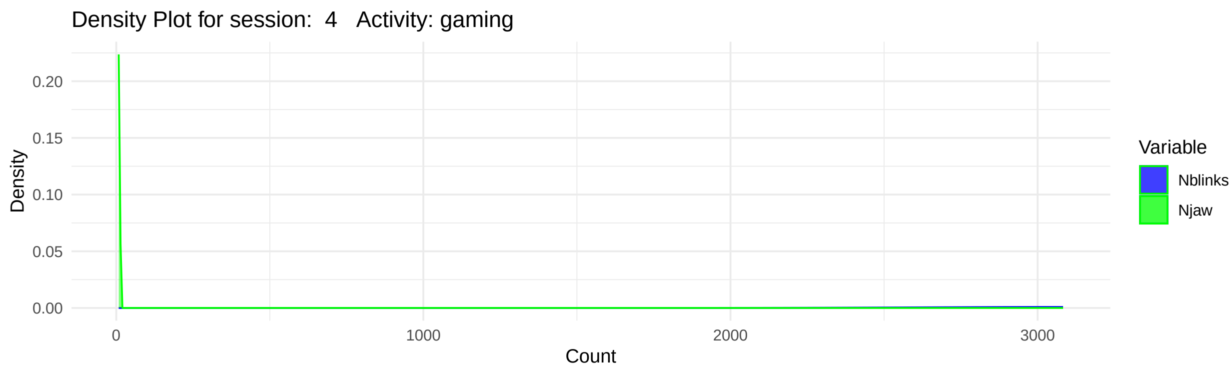
The first interesting pattern of data involved the frequency of number of blinks and number of jaw clenches in each session in relation to its activity and each other.

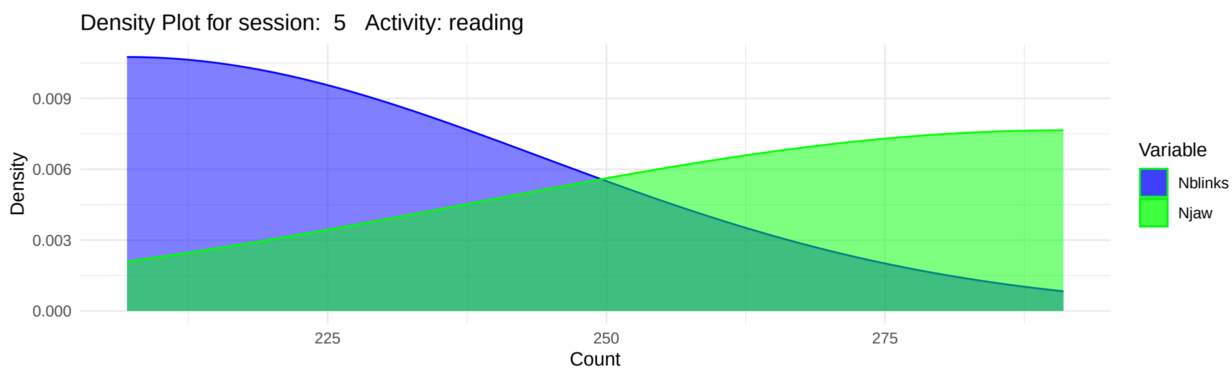
```
# Get unique sessionnum values
unique_sessionnums <- unique(dat_all$sessionnum)

# Loop through each sessionnum
for (k in unique_sessionnums) {
  session_data <- dat_all %>% filter(sessionnum == k)

  # Create a density plot for N-blinks and N-jaws
  plot <- ggplot(session_data, aes(x = `Nblinks`, fill = "Nblinks")) +
    geom_density(alpha = 0.5, color = "blue") +
    geom_density(aes(x = `Njaw`, fill = "Njaw"), alpha = 0.5, color = "green") +
    labs(title = paste("Density Plot for session: ",
                       k, "  Activity:", unique(session_data$activity)),
         x = "Count",
         y = "Density",
         fill = "Variable") +
    scale_fill_manual(values = c("blue", "green")) +
    theme_minimal()

  # Print each plot individually
  print(plot)
}
```

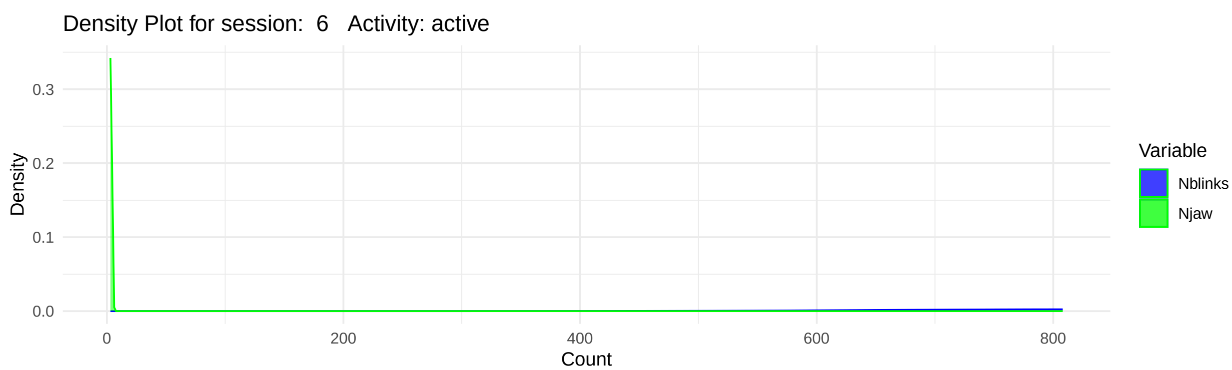


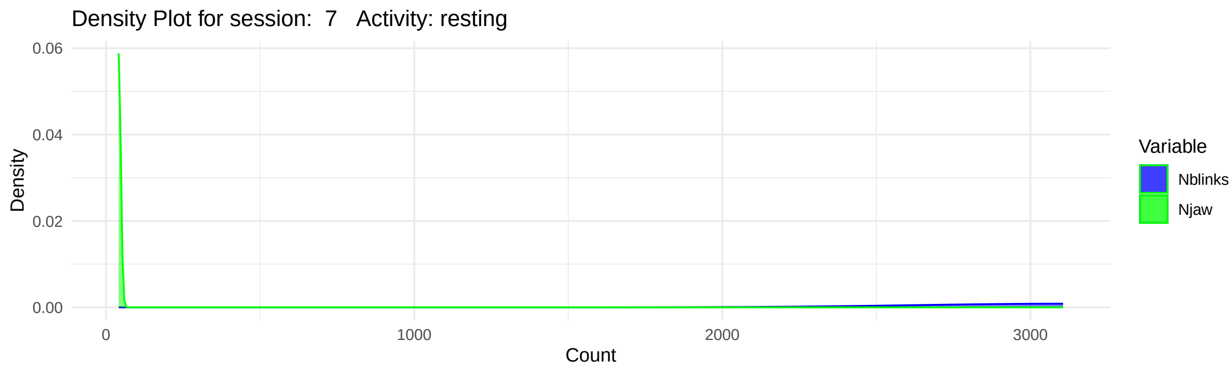


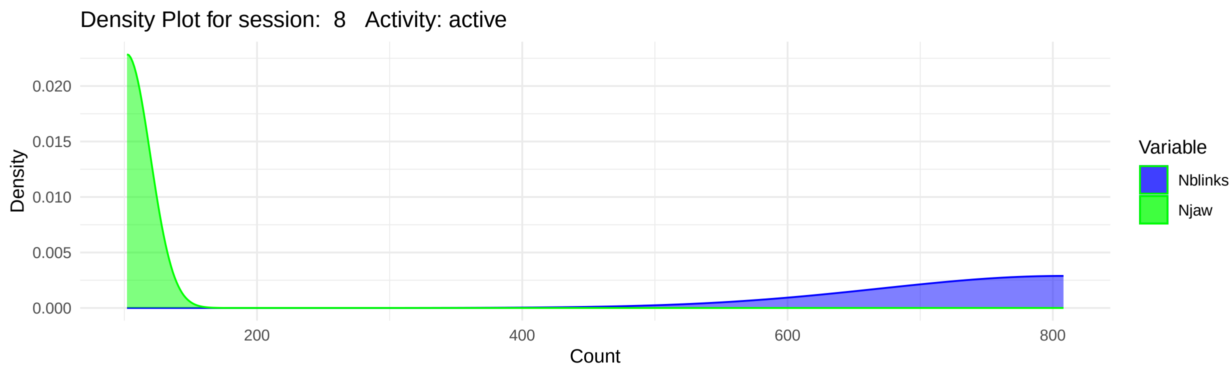

Density Plot for session: 3 Activity: active
Density
0.0075
0.0050
0.0025
0.0000
300
400
500
600
700
Count
Variable
Nblinks
Njaw

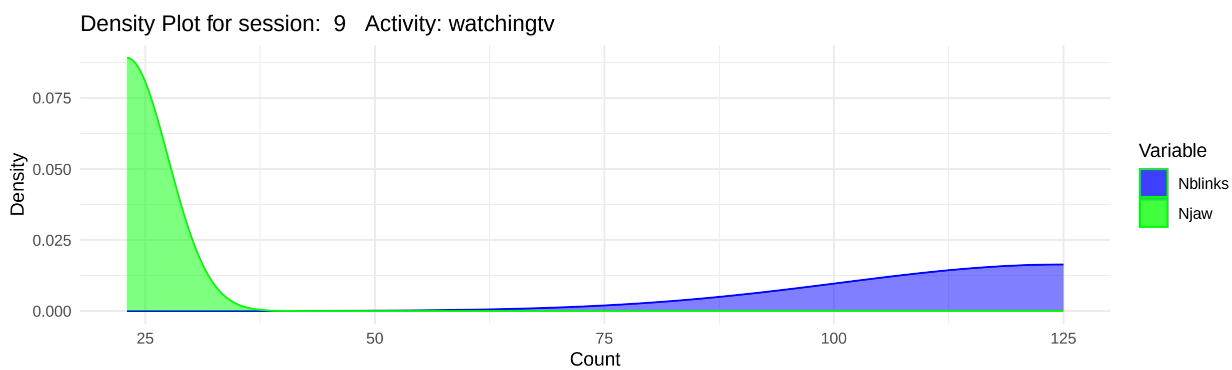

Density Plot for session: 4 Activity: gaming
Density
0.20
0.15
0.10
0.05
0.00
0
1000
2000
3000
Count
Variable
Nblinks
Njaw

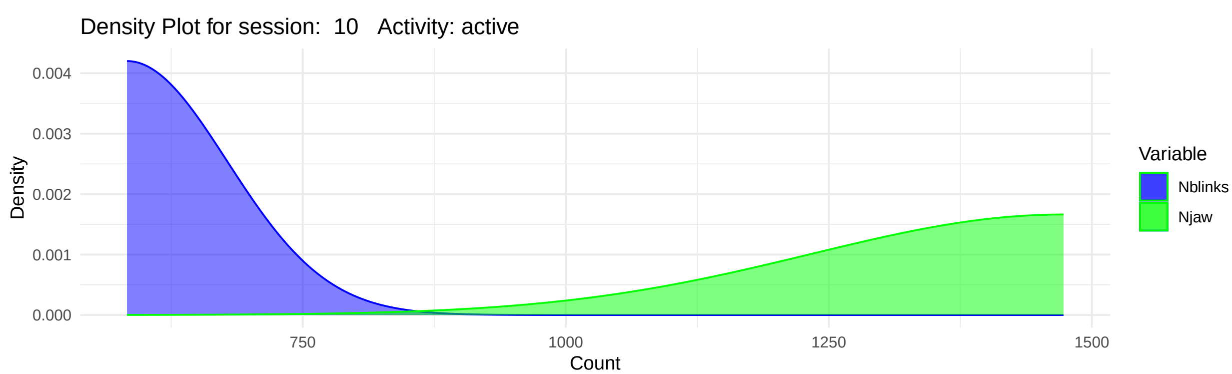

Density Plot for session: 5 Activity: reading
Density
0.009
0.006
0.003
0.000
225
250
275
Count
Variable
Nblinks
Njaw

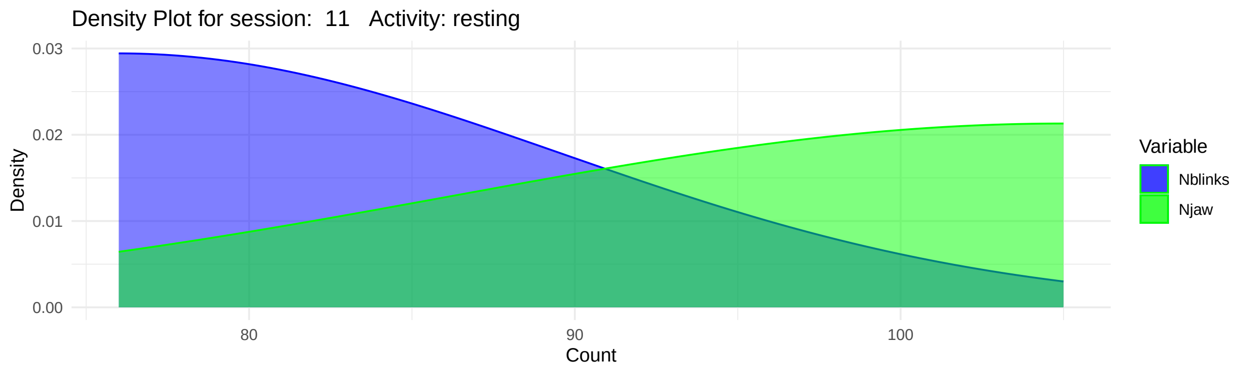

Density Plot for session: 6 Activity: active
Density
0.3
0.2
0.1
0.0
0
200
400
600
800
Count
Variable
Nblinks
Njaw


Density Plot for session: 7 Activity: resting
Density
0.06
0.04
0.02
0.00
0
1000
2000
3000
Count
Variable
Nblinks
Njaw

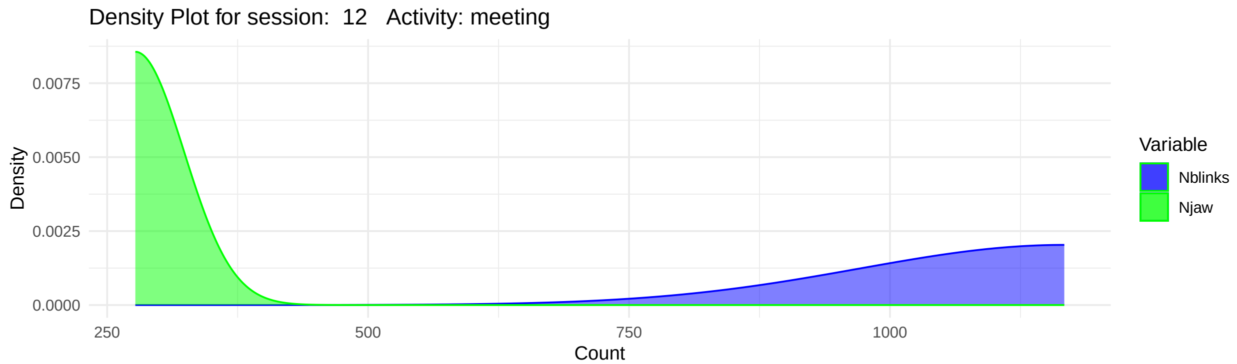

Density Plot for session: 8 Activity: active
Density
0.020
0.015
0.010
0.005
0.000
200
400
600
800
Count
Variable
Nblinks
Njaw


Density Plot for session: 9 Activity: watchingtv
Density
0.075
0.050
0.025
0.000
25
50
75
100
125
Count
Variable
Nblinks
Njaw

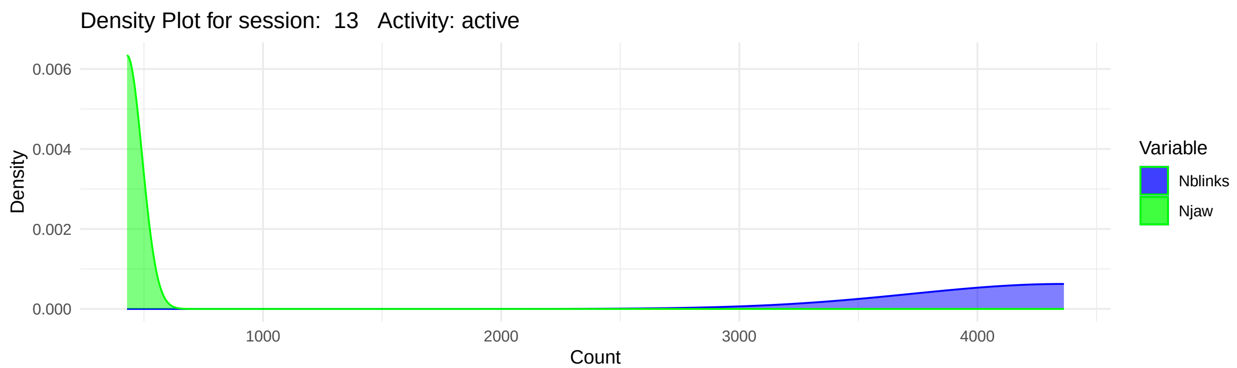

Density Plot for session: 10 Activity: active
Density
0.004
0.003
0.002
0.001
0.000
750
1000
1250
1500
Count
Variable
Nblinks
Njaw



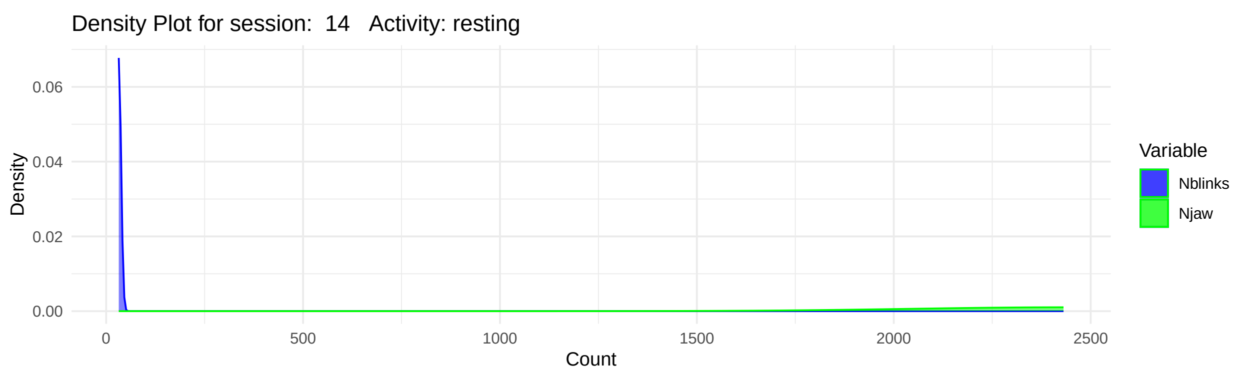
Density Plot for session: 11 Activity: resting
0.03
0.02
0.01
0.00
Density
80
90
100
Count
Variable
Nblinks
Njaw



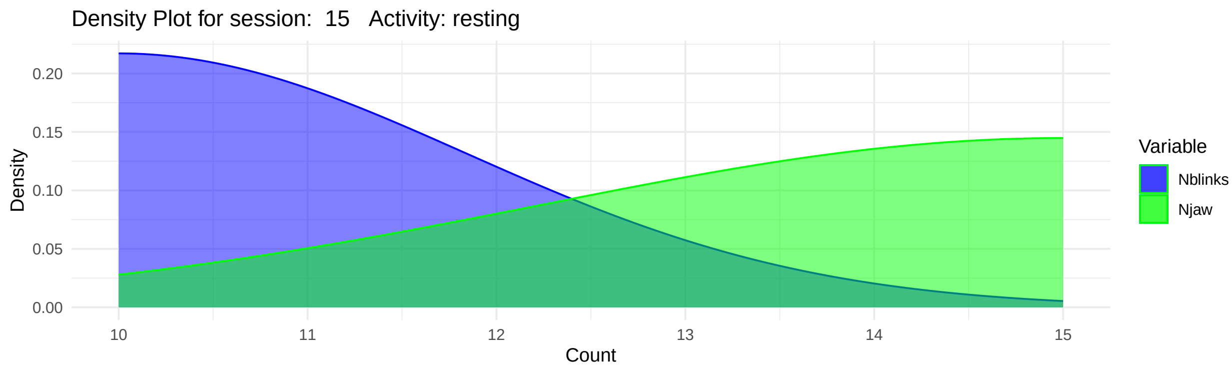
Density Plot for session: 12 Activity: meeting
0.0075
0.0050
0.0025
0.0000
Density
250
500
750
1000
Count
Variable
Nblinks
Njaw

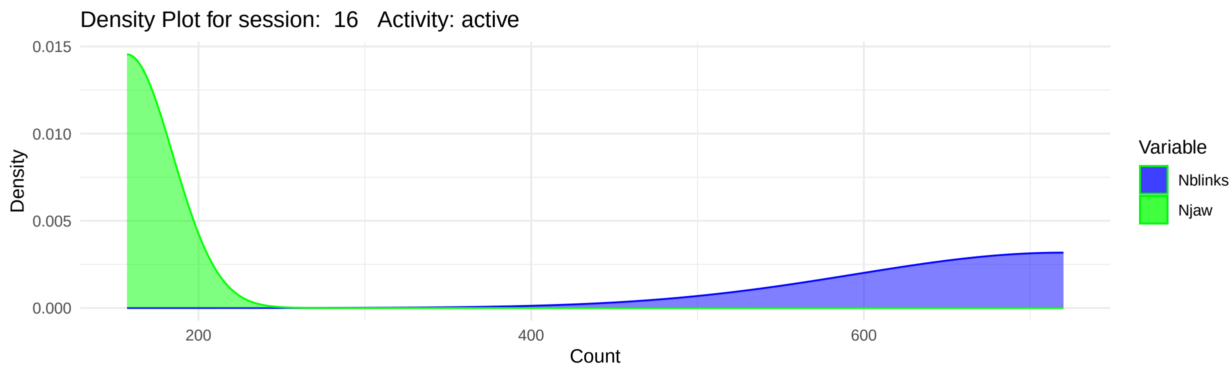


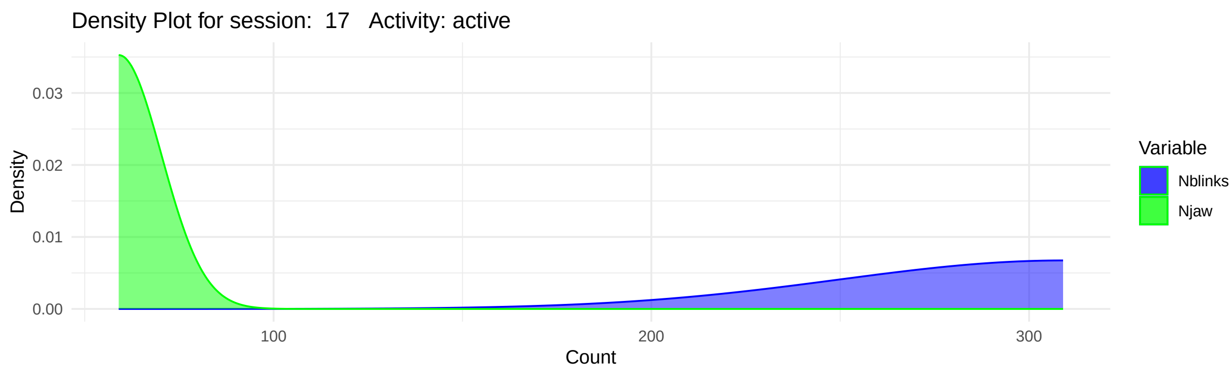
Density Plot for session: 13 Activity: active
0.006
0.004
0.002
0.000
Density
1000
2000
3000
4000
Count
Variable
Nblinks
Njaw

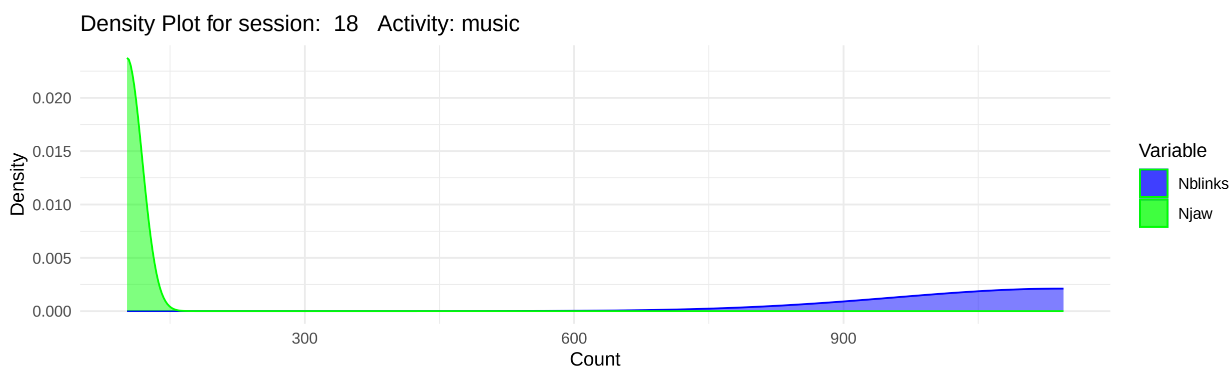


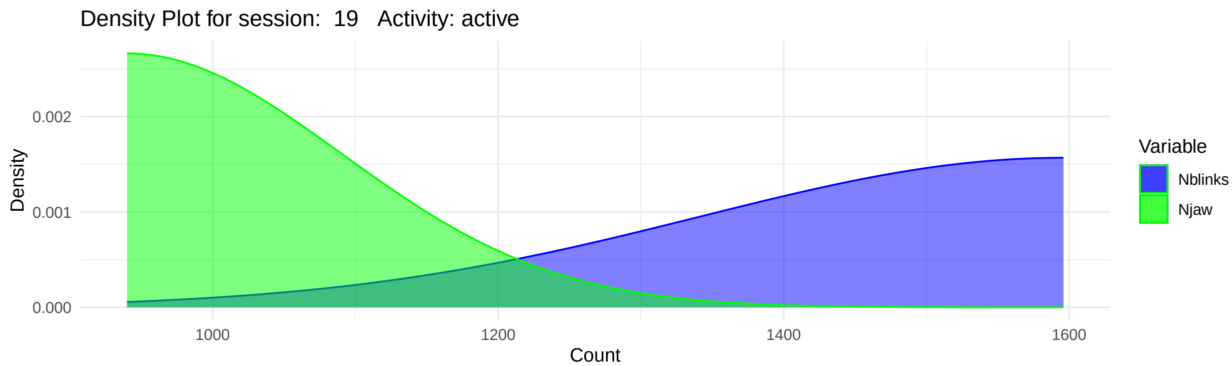
Density Plot for session: 14 Activity: resting
0.06
0.04
0.02
0.00
Density
0
500
1000
1500
2000
2500
Count
Variable
Nblinks
Njaw

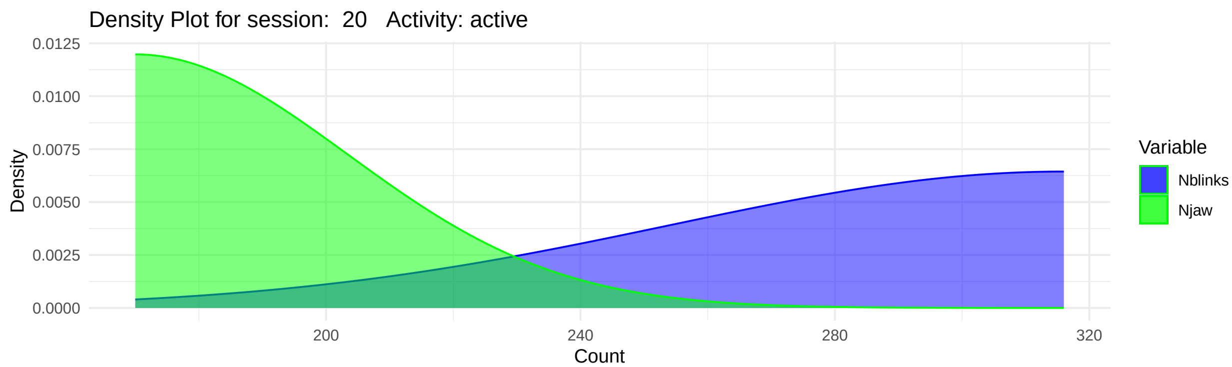

Density Plot for session: 15 Activity: resting
Density
0.20
0.15
0.10
0.05
0.00
10
11
12
13
14
15
Count
Variable
Nblinks
Njaw

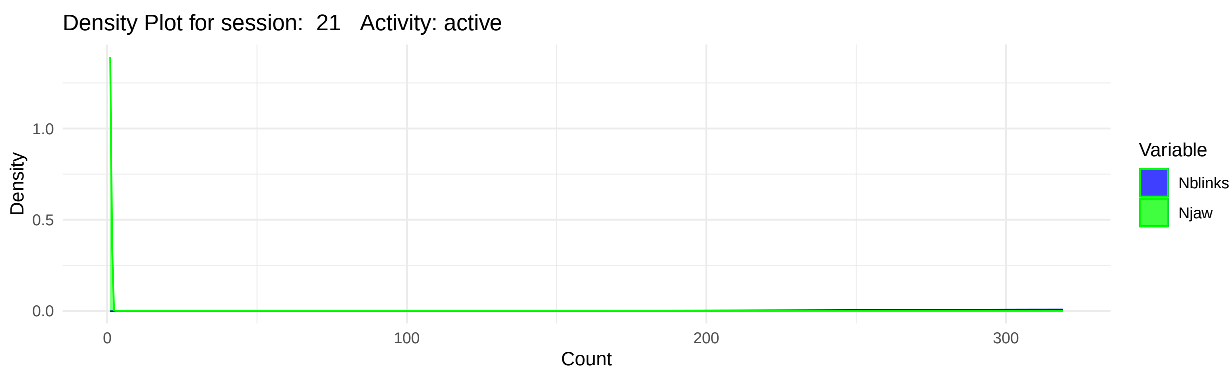

Density Plot for session: 16 Activity: active
Density
0.015
0.010
0.005
0.000
200
400
600
Count
Variable
Nblinks
Njaw


Density Plot for session: 17 Activity: active
Density
0.03
0.02
0.01
0.00
100
200
300
Count
Variable
Nblinks
Njaw


Density Plot for session: 18 Activity: music
Density
0.020
0.015
0.010
0.005
0.000
300
600
900
Count
Variable
Nblinks
Njaw







In general, when the patient is engaged in active tasks, there is a tendency for a higher frequency of blinks compared to jaw clenches. Conversely, during resting sessions, there is a prevalence of jaw clenches compared to blinks. However, the nature of the activity significantly influences this pattern.

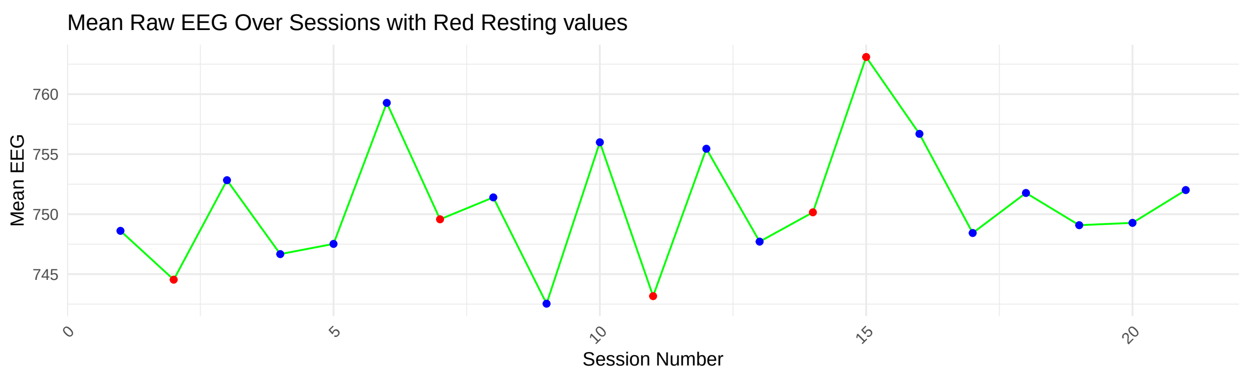
During reading, despite the patient being actively engaged, the trend deviates. Jaw clenches surpass blinks in number, and their distribution is more evenly spread rather than concentrated around extremes observed in typical resting and active sessions. This suggests a unique pattern during reading, where the patient exhibits a moderate and evenly distributed jaw-clenching activity.

In session 4, characterized by gaming, an extreme occurrence of both jaw clenches and blinks is observed. Interestingly, instances of no blinking at all are present, contributing to the distinctive trend during gaming sessions.

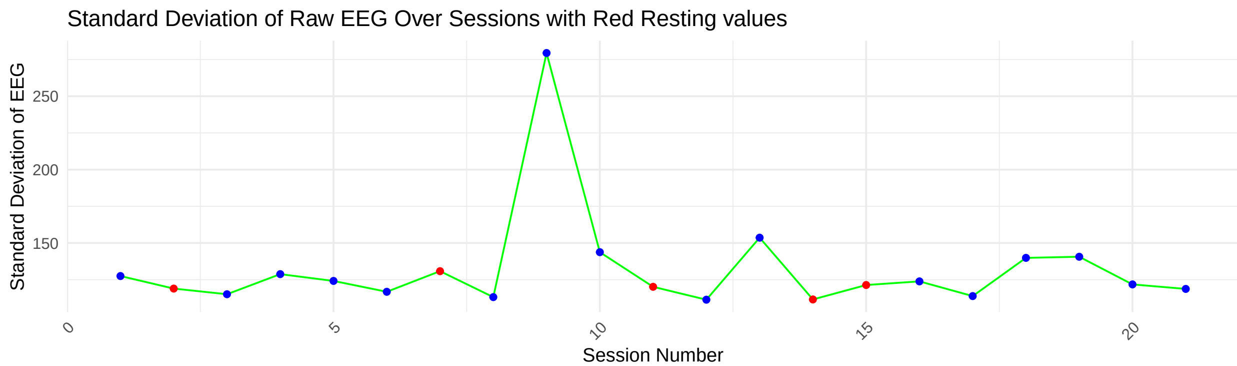
Watching TV generally follows the pattern of regular active sessions, with a higher number of blinks compared to jaw clenches. However, in the session immediately following TV watching and preceding resting, the trend aligns more with active sessions. Here, the number of jaw clenches significantly exceeds the number of blinks, implying that the patient might be reflecting on or processing the content viewed, rather than engaging in a specific activity.

During session number 18, where the patient is listening to music, a more relaxed state is apparent. Fewer jaw clenches are observed, and there is a higher frequency of blinks spread over a broader range. This suggests a calmer and less physically active state during music listening compared to other active sessions.

```
dat_EEG <- dat_all %>%
  na.omit() %>%
  group_by(sessionnum) %>%
  summarise(
    sessionnum = first(sessionnum),
    activity = first(activity),
    mean_eeg = mean(AUX_RIGHT),
    std_eeg = sd(AUX_RIGHT)
  )
ggplot(dat_EEG, aes(x = sessionnum, y = mean_eeg)) +
  geom_line(color = 'green') +
  geom_point(aes(color = ifelse(activity == 'resting', 'red', 'blue'))) +
  labs(title = "Mean Raw EEG Over Sessions with Red Resting values",
       x = "Session Number",
       y = "Mean EEG") +
  scale_color_identity() +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```
ggplot(dat_EEG, aes(x = sessionnum, y = std_eeg)) +
  geom_line(color = 'green') +
  geom_point(aes(color = ifelse(activity == 'resting', 'red', 'blue'))) +
  labs(title = "Standard Deviation of Raw EEG Over Sessions with Red Resting values",
       x = "Session Number",
       y = "Standard Deviation of EEG") +
  scale_color_identity() +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



The first graph depicts the mean Raw EEG sensor values in microvolts from AUX_RIGHT in the detailed dataset. The red points correspond to mean EEG values during resting, while the blue points represent active periods.

It is noteworthy that the mean EEG consistently decreases during each resting session across all sessions. Session 9, during which the patient watched TV, exhibits a noticeable dip in mean EEG. Conversely, in session 13, characterized as an active session, there is a spike in EEG before transitioning to resting. Notably, two subsequent resting sessions feature elevated EEG levels, standing out from the established pattern. Conversely, session 13, following a meeting, emerges as an outlier, deviating from the usual trend observed in prior sessions.

The second plot illustrates the standard deviation of Raw EEG sensor values in microvolts from AUX_RIGHT in the detailed dataset. Red points represent mean EEG values during resting, while blue points signify active periods.

In this plot, the standard deviation of Raw EEG from the mean appears relatively uniform, with exceptions observed at session 9 and session 13. Session 9's variation can be attributed to the patient watching TV, where EEG dynamics are likely influenced by the content viewed. Session 13 stands out as an anomalous session, as noted in the previous graph. This active session, following a meeting, displays distinctive EEG variability, suggesting a different activity level compared to previous sessions.