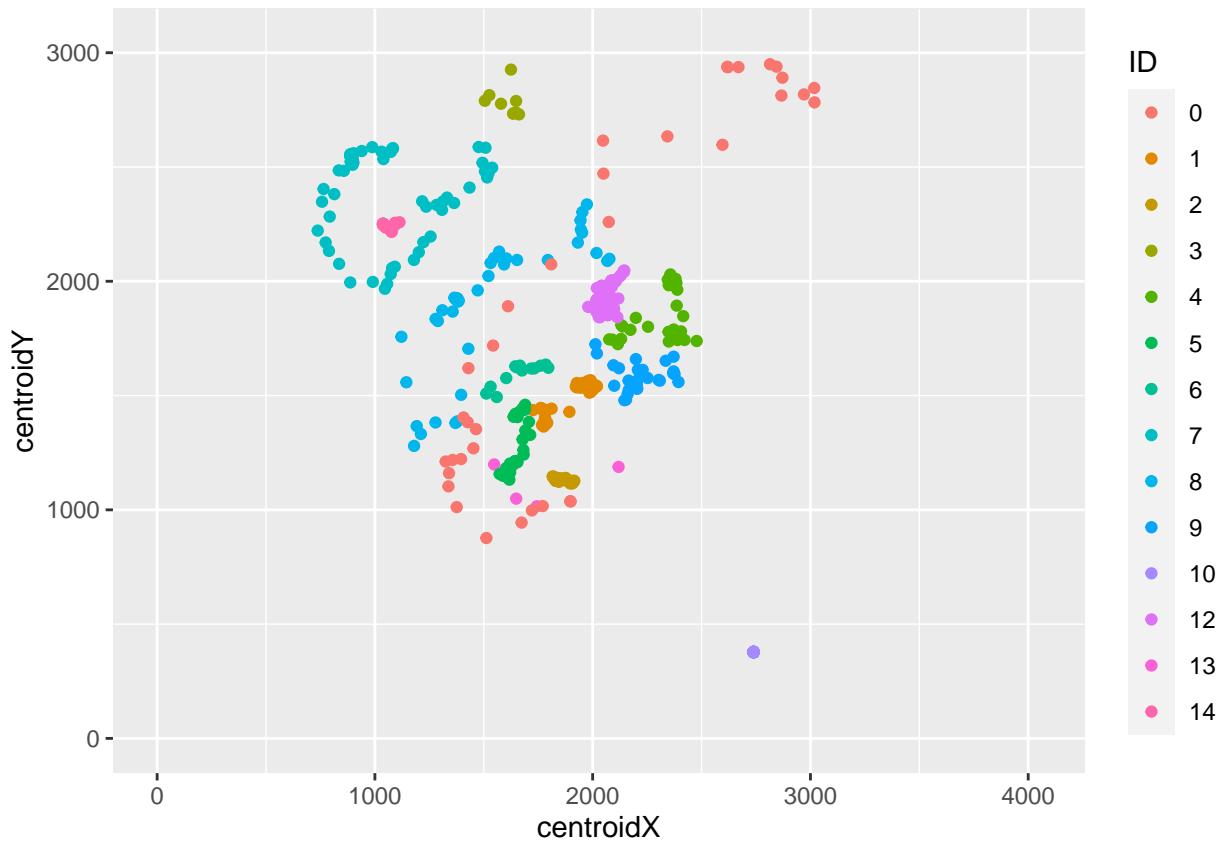


# BumbleboxExampleAnalysis

2023-11-07

## Part 1: Load and plot data tracking data from single video



```
##               filename ID average.distance.from.center
## 1 bumblebox-01_2023-10-28_00-55-03  0                      783
## 2 bumblebox-01_2023-10-28_00-55-03  1                      194
## 3 bumblebox-01_2023-10-28_00-55-03  2                      492
## 4 bumblebox-01_2023-10-28_00-55-03  3                     1060
## 5 bumblebox-01_2023-10-28_00-55-03  4                      497
## 6 bumblebox-01_2023-10-28_00-55-03  5                      433
## 7 bumblebox-01_2023-10-28_00-55-03  6                      177
## 8 bumblebox-01_2023-10-28_00-55-03  7                     1086
## 9 bumblebox-01_2023-10-28_00-55-03  8                      571
## 10 bumblebox-01_2023-10-28_00-55-03 9                      334
## 11 bumblebox-01_2023-10-28_00-55-03 10                     1522
## 12 bumblebox-01_2023-10-28_00-55-03 12                     413
```

```

## 13 bumblebox-01_2023-10-28_00-55-03 13          537
## 14 bumblebox-01_2023-10-28_00-55-03 14          973
##   average.speed frames.tracked.in.video
## 1           130          35
## 2            20          48
## 3             6          45
## 4            52           9
## 5            49          28
## 6            20          48
## 7            43          13
## 8            47          52
## 9            75          40
## 10           63          28
## 11           1          52
## 12           32          52
## 13          140           4
## 14            9          23

```

## Part II: Combine data across all trials

```

#Make sure you don't have a 'comb_data' object in memory
rm('tracking_data')

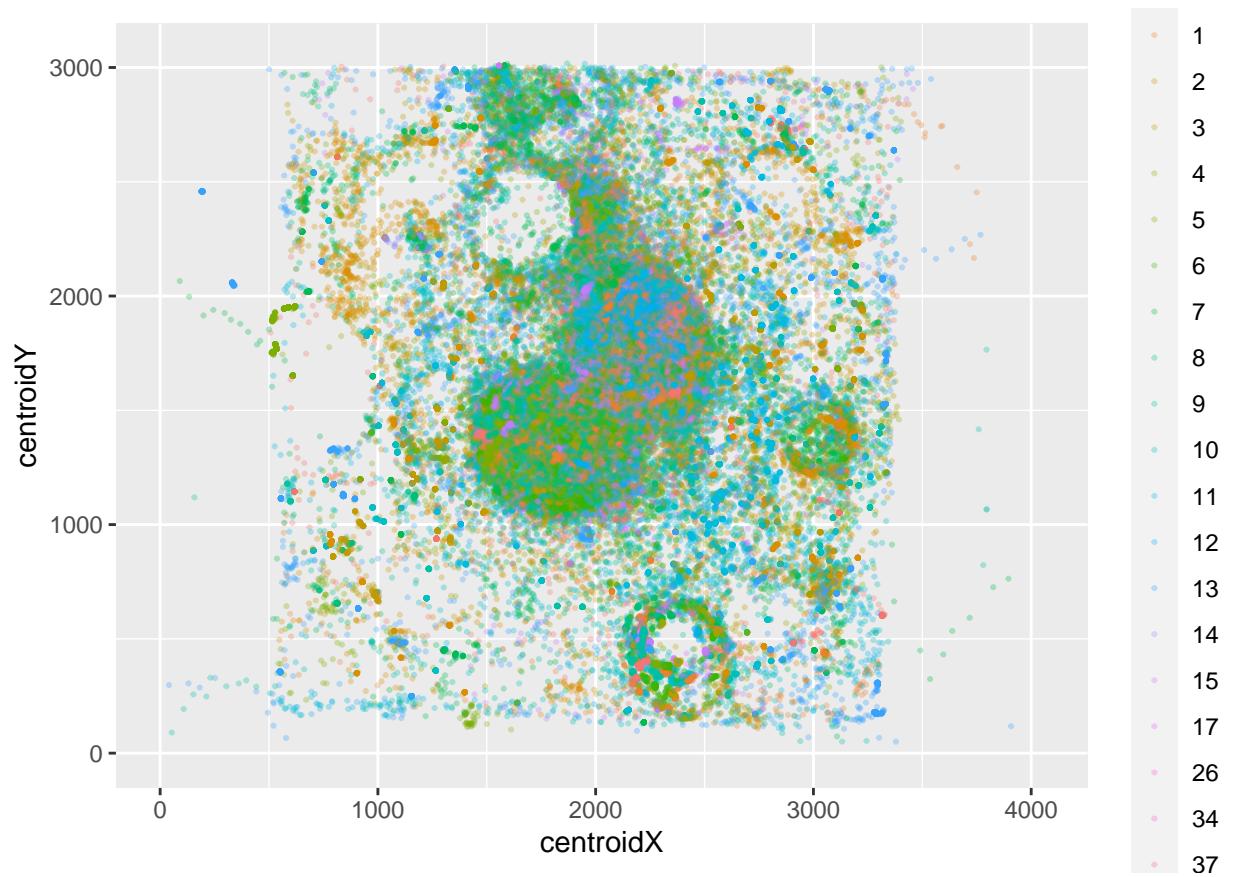
for(file in trials){

  data <- read.csv(str_replace(file, '_averages.csv', '_raw.csv'))
  ## Compile average data
  #Check to see if output data frame exists
  if(!exists('tracking_data')){
    tracking_data <- data #If not, create it from the first loop
  } else {
    tracking_data <- rbind(tracking_data, data) #Otherwise append new data to the end
  }

}

tracking_data$ID <- as.factor(tracking_data$ID)
ggplot(tracking_data, aes(x = centroidX, y = centroidY, colour = ID)) + geom_point(alpha = 0.3, cex = 0.5)

```



```

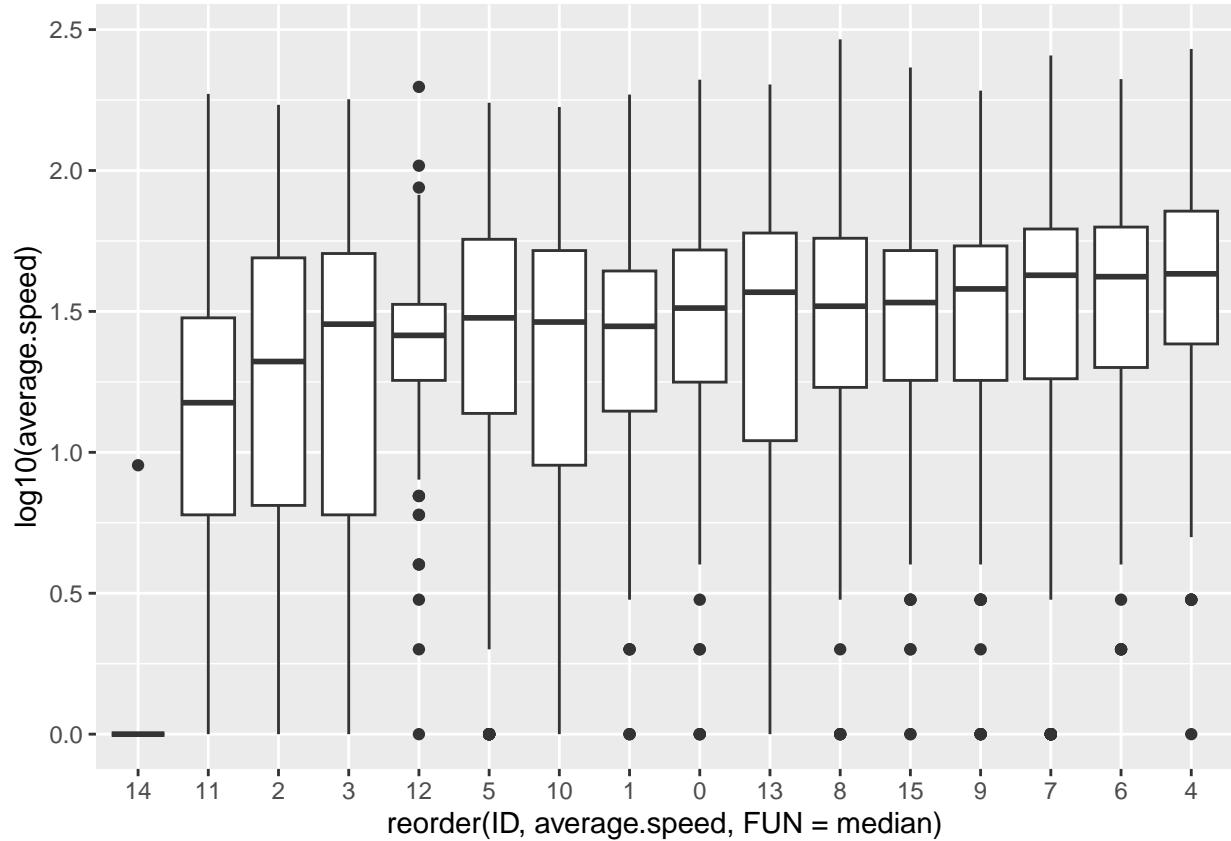
## Warning in rm("comb_data"): object 'comb_data' not found

#Clean out missing data
comb_data <- comb_data[complete.cases(comb_data),]

#Boxplot of speed vs. individual
ggplot(comb_data, aes(x = reorder(ID, average.speed, FUN = median), y = log10(average.speed)))+geom_boxp

## Warning: Removed 103 rows containing non-finite values ('stat_boxplot()').

```

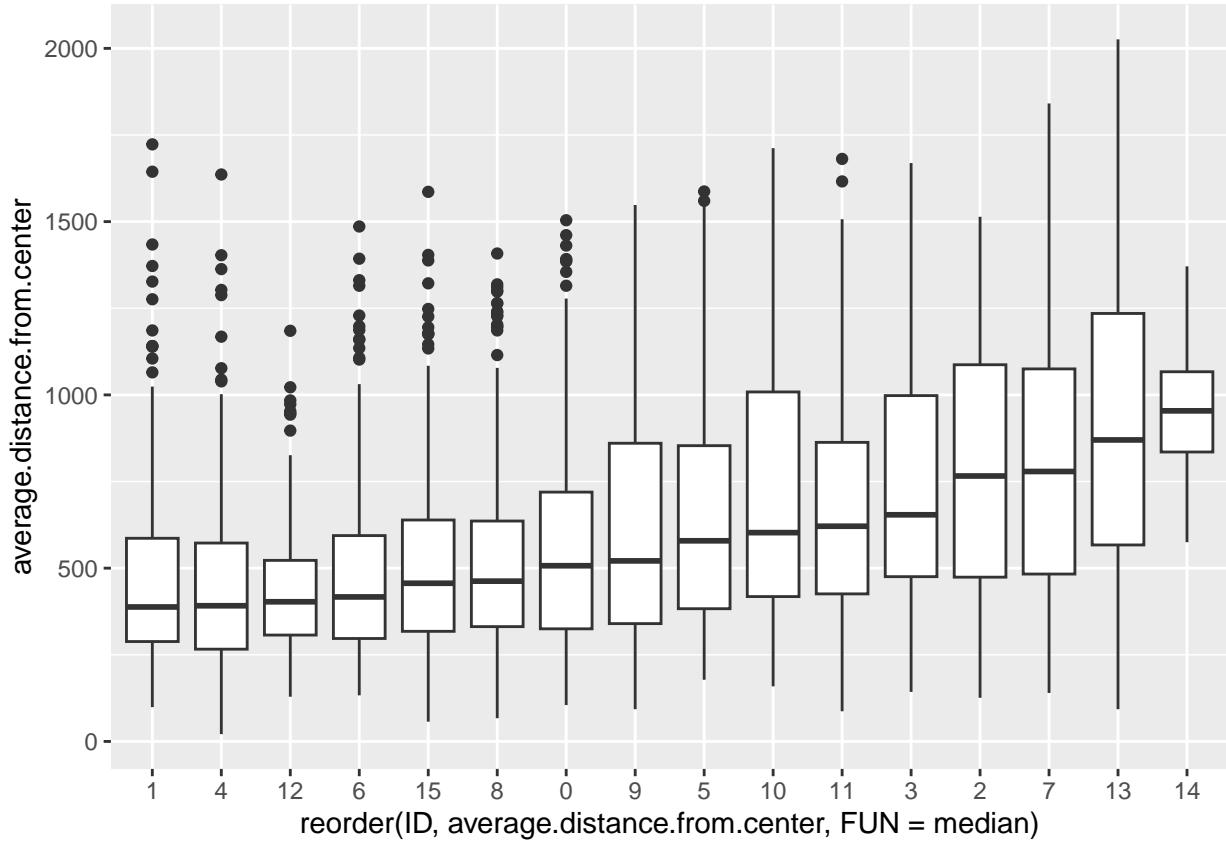


```
#Check for evidence of individual variation
model <- lmer(log10(1+average.speed)~ID + (1|filename), data = comb_data)
model.1 <- lmer(log10(1+average.speed)~1 + (1|filename), data = comb_data)
anova(model, model.1)

## refitting model(s) with ML (instead of REML)

## Data: comb_data
## Models:
## model.1: log10(1 + average.speed) ~ 1 + (1 | filename)
## model: log10(1 + average.speed) ~ ID + (1 | filename)
##          npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## model.1     3 4660.9 4678.9 -2327.5    4654.9
## model      18 4199.4 4307.3 -2081.7    4163.4 491.56 15 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#Show distance from center by individual
ggplot(comb_data, aes(x = reorder(ID, average.distance.from.center, FUN = median), y = average.distance
```



```
#Get individual averages
individual_averages <- aggregate(cbind(average.distance.from.center, average.speed, frames.tracked.in.v),
colnames(individual_averages) <- c('ID', 'dist', 'speed', 'frames_tracked')

# Remove outlier with zero velocity
individual_averages <- subset(individual_averages, speed > 1)

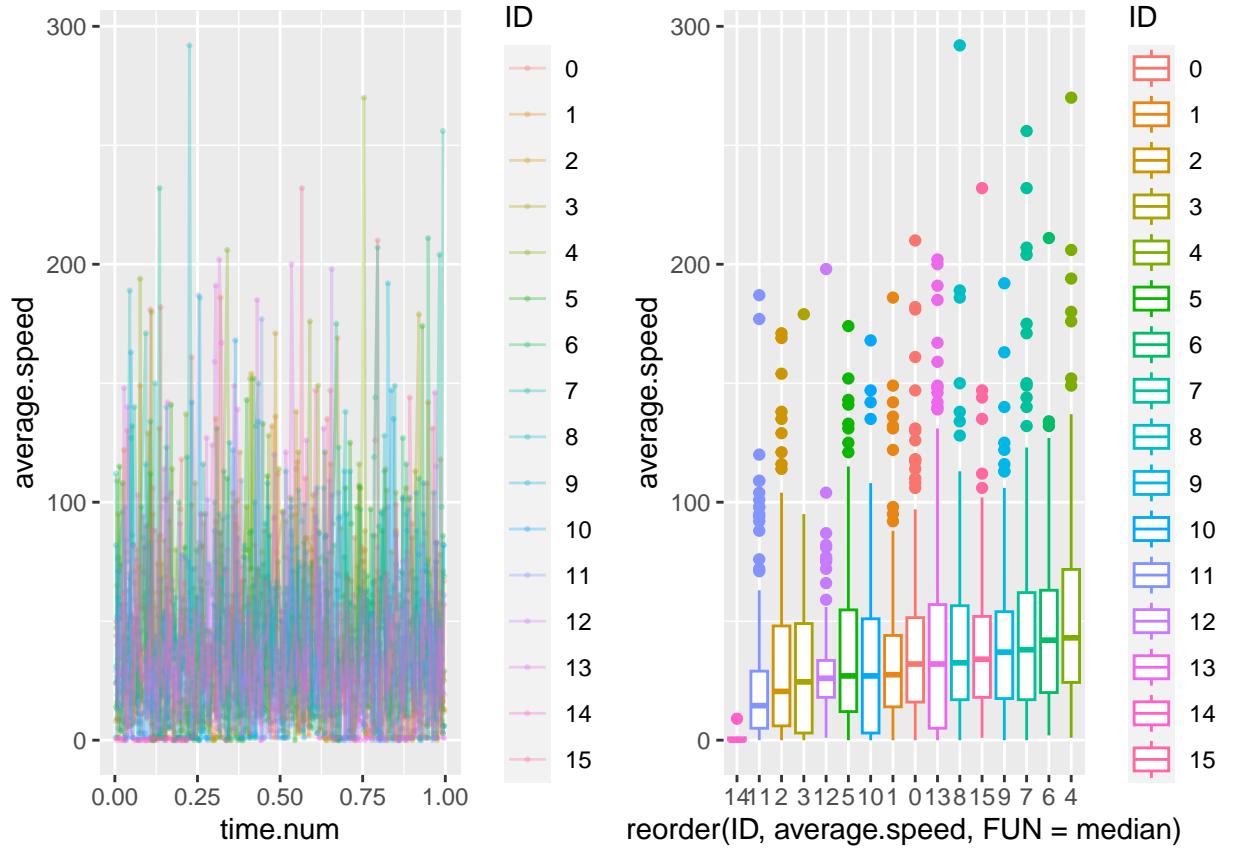
#Visualize individual speed/distance over time

#Set up timestamps
start.date <- parse_date_time('2023-10-28 00:00:00', "%Y-%m-%d %H:%M:%S")
comb_data$datetime <- parse_date_time(substr(comb_data$filename, 14, 32), "%Y-%m-%d_%H-%M-%S")
comb_data$time.num <- as.numeric(difftime(comb_data$datetime, start.date, units = 'days'))

comb_data <- comb_data[order(comb_data$ID, comb_data$time.num),]

p1 <- ggplot(comb_data, aes(x = time.num, y = average.speed, colour = ID)) + geom_point(alpha = 0.3, cex = 1)
p2 <- ggplot(comb_data, aes(x = reorder(ID, average.speed, FUN = median), y = average.speed, colour = ID))

grid.arrange(p1, p2, ncol=2)
```



```

p1 <- ggplot(comb_data, aes(x = time.num, y = average.distance.from.center, colour = ID)) + geom_point()
p2 <- ggplot(comb_data, aes(x = reorder(ID, average.distance.from.center, FUN = median), y = average.distance.from.center, colour = ID)) + geom_point()

grid.arrange(p1, p2, ncol=2)

```

