

Final Project

Ryan Chiang, Jaxi Cohen, Dahlia Friis, Emily Lieu

Amy Kim

Replicating Statistical Analysis in the Paper “Behavioral Syndromes in Paper Wasps: Links Between Social and Non-social Personality in Paper Wasps”

Introduction The paper investigates the relationships between social (aggression, affiliation, and antennation) and non-social (exploration and activity) personality traits of paper wasps. To model the social traits, researchers placed wasps in compartments with dead “dummy” wasps for 10 minutes, and recorded their behavior. To model the non-social traits, researchers placed wasps in a plexiglass arena consisting of 10 built in wooden chambers. The wasps were free to roam for 10 minutes. Aggression is measured by a count of aggressive behaviors displayed by the wasp. Affiliation is measured by the time spent in non-aggressive, stationary contact with the dummy wasp. Antennation is measured by taps by the wasp onto the dummy wasp with its antennae. Exploration is measured through a count of the number of chambers by the wasp in the maze. Lastly, activity is measured by the time spent by the wasp actively moving. These traits can be modeled and predicted with the variables of ID number, trial number, color of the dummy, and weight of the wasp. We chose to reproduce this paper because we are interested in the application of statistics for predicting animal behavior. Understanding the connection between social and non-social traits can help us understand behavioral syndromes, which are groups of personality traits that are phenotypically related with each other. Understanding this connection can help us learn about how traits respond to selection and what combinations of traits are better for different species and environments. **Response variables:** Affiliation (bodily contact time): time spent by wasp in stationary, non aggressive contact with dummy, measured in seconds. Antennation: number of wasp taps on dummy with extended antennae, used for chemical assessment of other wasps; a neutral social behavior. Aggression (log aggression): count of aggressive behaviors displayed by wasp, including dart, dart with mandibles, antenna drumming, bite, mount; log-transformed data before analysis to normalize data. Exploration (chambers entered): number of individual chambers entered by wasp. Activity (seconds moving): time wasp spent actively moving inside the area, measured in seconds. **Explanatory variables:** Wasp.ID: identifying number assigned to each wasp. Trial: trial number per individual wasp. Dummy Color: Color of dummy wasp. Weight: Body weight of wasp, measured in grams.

Summary Stats We did not look for outliers in our exploratory data analysis. Removing outliers from each model would quickly become complicated, as we would have to create a new data set excluding the outliers. Behavior is also highly variable and complex, so removing outliers might impact our data in ways we cannot predict. The wasps participated in four trials for each behavioral test. When the wasp was placed in a clear box with a dead dummy wasp for 10 minutes, the non-aggressive bodily contact time spent between the wasp and the dummy averaged 183.5 seconds. The maximum was 589 seconds and the minimum was zero. The average time the wasp spent exploring the dummy via antennation averaged 5.6 seconds. The maximum was 16 and the minimum was zero. In terms of aggressive behavior, the authors measured how long the wasp spent being aggressive to the dummy in seconds and log-transformed it to normalize the data. The log-transformed average for aggressive behavior is 0.69 seconds. The maximum was three seconds and the minimum was zero. When the wasps were placed in the maze and given 10 minutes to roam, the average amount of chambers entered was five. The maximum was 10 chambers and the minimum was zero. The average amount of time wasps spent moving in the maze was 230 seconds. The maximum was 591 seconds and the minimum was five seconds. The authors recorded the weight of the wasps, though it is unclear how the weight of the wasps impacted their behavior. The average weight of the wasps was

0.15 grams, with the maximum being 0.19 grams and the minimum being 0.09 grams. Model fitting The behaviors we are modeling are affiliation, aggression, activity, antennation, and exploration. We started with a multiple regression model for each behavior that has all the remaining behaviors and factors as the predictor variables. We performed stepwise selection on each model to pick the best predictor variable(s). The results are five fixed effects, linear regression models (lm) where one behavior is the response variable and a different behavior, and sometimes other factors, are the predictor variables. Then we found the BIC of each final lm model. Next, we constructed additional models containing all the same explanatory variables as our final lm models but with an added random effect variable for wasp ID. The wasp ID is used to check if there is a significant difference in behavior between individual wasps in the study. We used the lmer function to create our mixed effect models. We obtained their BIC's. The final lm and corresponding lmer functions were compared to each other. Whichever model had the lower BIC was selected to be our final model for a given behavior. This results in a combination of linear and mixed models. Affiliation and antenation fit better as mixed effect models, aggression and exploration fit better as fixed effect, linear regression models.

Antennation ~ log aggression + wasp ID Aggression ~ dummy color + antennation Exploration ~ activity Affiliation ~ trial number + activity + wasp ID Activity ~ exploration + wasp ID Model diagnostics The assumptions for linear mixed effects models are the same for multiple regression models: constant and normal distribution of errors, a linear model is appropriate, and a random sample is taken. We assume a random sample is taken and a linear model must be appropriate. For normality of errors we used the Shapiro-Wilks (SW) test and qq plots. Our p-values for all of the tests were very low. Therefore we reject the null hypothesis at $\alpha = 0.05$ and support the alternative that our models do not have normality of errors. To test for constant variance of errors we created residual plots and performed the Fligner-Killeen (FK) test on our models when available. We were unable to do FK tests on our mixed models. For exploration and aggression we reject the null hypothesis at $\alpha = 0.05$, indicating there is not constant variance of errors. For all our residual plots, except antenation, we found the average residual to be far from zero and/or a pattern in the plot. Therefore, there is not constant variance of errors.

FK p-values: Exploration: 0.001585; Aggression: $1.807e^{-14}$ SW p-values: Affiliation: $1.777e^{-8}$; Exploration: 0.01837; Antennation: 0.0001166; Aggression: $1.49e^{-9}$, Act: 0.0006617 Interpretation Antennation: When log aggression increases by 1, we expect antennation to increase by 0.7655 taps on the dummy. This relationship goes against expectations, because antenna tapping is a neutral, non-aggressive behavior of wasps. There is a positive relationship between antennation and aggression. Affiliation: When trial number increases by 1, we expect bodily contact time to increase by 24.67 seconds, holding seconds moving constant. A possible explanation of this coefficient is that as the wasp spends more time with the dummy throughout the trials, it is more willing to touch the dummy wasp. When the number of seconds moving increases by 1, we expect bodily contact time to decrease by 0.1134 seconds, holding trial number constant. This is implied, as when the wasp increases time moving around, it has less time to interact with the dummy. There is a negative relationship between affiliation and activity. Activity: When the number of chambers entered increases by 1, we expect the time spent moving to increase by 47.95 seconds. This is implied, as it takes time to enter a chamber. There is a positive relationship between activity and exploration. Exploration: When the time spent moving increases by 1 second, we expect the number of chambers entered to increase by 0.0152. There is a positive relationship between exploration and activity. Aggression: When compared to the baseline color gold, gold-silver and silver have the greatest effect on log aggression; they increase aggression. The colors green, red, white, and yellow decrease aggression relative to gold. This was further supported by our confidence interval for these coefficients. Gold-silver and silver were the only two confidence intervals that did not include 0, meaning none of the other colors had a significant effect on aggression, compared to the gold dummy color. We found through ANOVA testing that gold, the baseline color, has little effect on aggression. This indicates that gold-silver and silver specifically aggravate wasps, versus all metallic colors. When antennation increases by 1 antenna tap, the log aggression increases by 0.03224 movements. There is a positive relationship between aggression and antennation. The majority of the behavior trait models include wasp ID, the random variable, as a predictor variable. While the coefficients of the wasp ID variable have no practical application, the inclusion of it in these models indicates that the individual wasp has a significant effect on its behavior. The individual personalities of these wasps, separate from dummy color, trial number, and body weight, have an effect on antennation, affiliation, and activity. For example, an individual wasp with a curious personality may be more inclined to interact with the dummy. For aggression

and exploration, the individual wasp was not relevant in predicting these traits. Our models suggest that how aggressive a wasp is is not dependent on the wasp's personality. The same can be said for exploration.

Comparison (the paper vs us) The main difference in our methods and those used in the paper is that the paper utilized the rpt function, while we used the lmer function. The rpt function estimates the repeatability of the data from a linear mixed effects model fitted by restricted maximum likelihood. Repeatability is a measure of the partitioning of variance into within-group and between-group sources of variance. Repeatability is often used in behavioral research because it quantifies individual differences, in this case, in different wasp personalities. The paper generated two sets of models. There are five models in each set, one for each behavior trait. Their first set only includes one random effect, wasp ID. Their second set includes body mass and trial number as fixed effects, as well as wasp ID as a random effect. We created two sets of models, but our first set includes only fixed effects. Our second set of models includes the fixed effects and random effect, wasp ID. Our methods align with the authors' in using mixed models for the second set. Our methods differ in the first set, where we focused on fixed effects and they focused on random effects. For each of the behavior traits, we compared the two models' BICs to determine which was a better predictor of the trait. The models created in the paper isolated each behavior trait as the response variable, and no behavior traits were included as predictor variables in each other's models. We found that certain behavior traits could help predict other behaviors in our models. For antennation, our final model includes wasp ID and log aggression as predictors. This differs from the paper's corresponding mixed model, which includes wasp ID, dummy ID, body weight, and trial number. For aggression, our final model includes dummy color and antennation as predictors. This differs from the paper, as they used a mixed model including wasp ID, dummy ID, body weight, and trial number. For exploration, our final model includes seconds moving as a predictor. This differs from the paper, as they used a mixed model including wasp ID, dummy ID, body weight, and trial number. For affiliation, our final model includes trial number, seconds moving, and wasp ID as predictors. The paper's mixed model includes wasp ID, dummy ID, body weight, and trial number. For activity, we found that the mixed model was a better fit than the simple model including only chambers entered as a predictor. The mixed model from the paper included individual ID, trial number, and body weight. A major component of the paper uses Spearman's Rank Correlation Coefficients and the associated p-values to determine whether a meaningful correlation between different behaviors exists. Their statistically significant results were a positive correlation ($r_s = 0.89$) between exploration and activity, a negative correlation ($r_s = -0.35$) between affiliation and activity, and a positive correlation ($r_s = 0.33$) between aggression and activity. With our models we can interpret the relationships between behaviors. Similar to the paper, we found a negative relationship between activity and affiliation and a positive relationship between exploration and activity. Unlike the paper, we found a positive relationship between log aggression and antennation. Our models also differed because the rpt package utilizes MCMCglmm to fit linear mixed models, while we used lmer. MCMCglmm uses a Bayesian approach, in which people make inferences based on priors about the parameters and the gathered data. Priors are assumptions about the values of the parameters before seeing the data. This method also uses Markov Chain Monte Carlo sampling, which estimates the distribution of parameters by generating many random samples. On the other hand, lmer uses a frequentist approach, in which inferences are based solely on the data at hand without having previous assumptions, or priors. lmer also uses a likelihood approach to estimate parameters by finding the values that make the observed data most probable. As a result, we can make conclusions based on this likelihood approach, which measures how likely the observed data is given certain parameter values. To simplify, using MCMCglmm method is like making a guess about an outcome, then using your gathered data to repeatedly improve the guess, keeping your prior beliefs in mind. Using lmer is like looking at the data directly to find the best-fitting solution without any prior assumptions. For the authors of the paper, using MCMCglmm is more fitting since they likely have prior knowledge about the behavior of wasps or experience from previous analysis of wasp behavior. Since we are working only with the scientific article and data the authors provided, lmer makes more sense for us.

Conclusion Most of the difficulties we encountered during this project were centered around working with the data and fitting our model. Random and fixed effect models were not covered in the class until late so we struggled to interpret what they were doing up in their paper. Navigating through building the mixed effects models was difficult because we were not familiar with the package. Running diagnostics on our mixed effects models also caused us some difficulties. We realized we had to prepare the model in a different way in order to run tests on it.

Important Questions for Future Research To improve this work on the relationship between social and non-social personality traits in paper wasps, we considered factors such

as diversity, testing over time, environment, and expanding predictor variables. In terms of diversity, the authors could sample wasps from different regions or sample different types of wasps to examine potential variations in personality. Testing the wasps over time would allow us to see how or if a wasp's personality would change over its life cycle. Alternatively, testing the same wasp individuals at different times would allow us to see the stability of the wasp's personality. Changing the wasp environment would also allow us to further study personality through methods such as observing wasp behavior in the presence of a predator by exposing the wasps to the scent of a predator. We also suggest including more predictor variables or including only relevant predictor variables to better assess wasp personality. For example, body weight did not have a significant effect on any personality traits. Instead, the study can expand and include wasp rank or different gender wasps. Overall, an increased sample size would also benefit this study and allow the authors to have more generalizable findings.

Some lingering questions we have are (1) why does bodily contact increase every time the experiment is repeated (increasing trial number)? (2) Why is wasp ID an important factor in some models?

Ryan: Much of the R work, building models and running tests Emily: Background research of statistical methods Dahlia: Summarizing in our report and slides Jaxi: Summarizing in our report and slides

Appendix:

```
#install.packages("Matrix")
#install.packages("MCMCglmm")
library(Matrix)
#remove.packages("Matrix")
#install.packages("/Users/apple/Downloads/Matrix_1.2-7.1.tar", repos = "NULL", type = "source")
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(rptR)
#library(lme4)
library(MCMCglmm)
```

```
## Loading required package: coda
```

```
## Loading required package: ape
```

```
##
## Attaching package: 'ape'
```

```
## The following object is masked from 'package:dplyr':
##
##   where
```

```
library(MPV)
```

```
## Loading required package: lattice

## Loading required package: KernSmooth

## KernSmooth 2.23 loaded
## Copyright M. P. Wand 1997-2009

## Loading required package: randomForest

## randomForest 4.7-1.1

## Type rfNews() to see new features/changes/bug fixes.

##
## Attaching package: 'randomForest'

## The following object is masked from 'package:ggplot2':
##
##     margin

## The following object is masked from 'package:dplyr':
##
##     combine
```

```
library(leaps)
library(MASS)
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:MPV':
##
##     cement

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

```
library(lme4)
```

```
wasp <- read.csv("../final_project_STA101/wasp.csv")
```

```
summary(wasp)
```

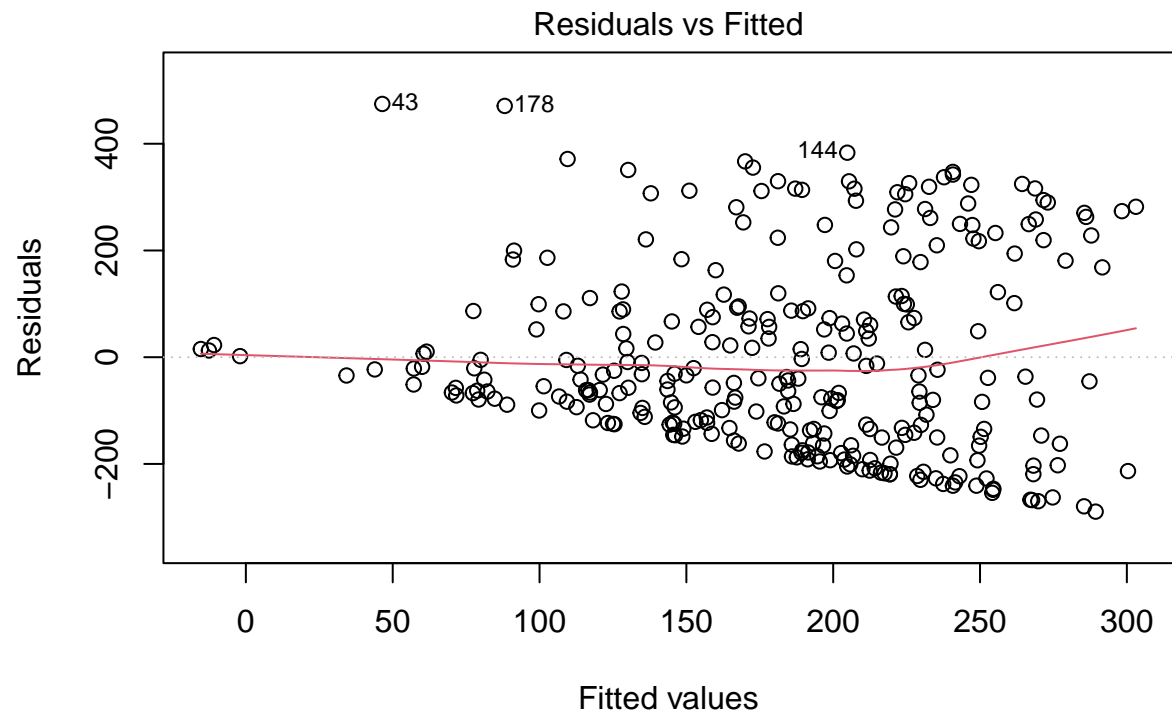
```
##           X           Wasp.ID           Trial           bodily_contact_time
## Min.      : 1.00   Length:296   Min.      :1.00   Min.      : 0.0
## 1st Qu.: 74.75   Class :character 1st Qu.:1.75   1st Qu.: 23.0
## Median :148.50   Mode  :character  Median :2.50   Median :115.5
## Mean      :148.50           Mean      :2.50   Mean      :183.5
## 3rd Qu.:222.25           3rd Qu.:3.25   3rd Qu.:284.5
## Max.      :296.00           Max.      :4.00   Max.      :589.0
## anntenation   Log_aggression   Dummy_Color           weight
## Min.      : 0.000   Min.      :0.0000   Length:296   Min.      :0.0900
## 1st Qu.: 3.000   1st Qu.:0.0000   Class :character 1st Qu.:0.1400
## Median : 5.000   Median :0.6931   Mode  :character  Median :0.1500
## Mean      : 5.642   Mean      :0.6991           Mean      :0.1514
## 3rd Qu.: 8.000   3rd Qu.:1.0986           3rd Qu.:0.1700
## Max.      :16.000   Max.      :3.0910           Max.      :0.1900
## chambers_entered seconds_moving
## Min.      : 0.000   Min.      : 5.00
## 1st Qu.: 3.000   1st Qu.: 81.75
## Median : 4.500   Median :160.50
## Mean      : 5.206   Mean      :230.79
## 3rd Qu.: 8.000   3rd Qu.:385.50
## Max.      :10.000   Max.      :591.00
```

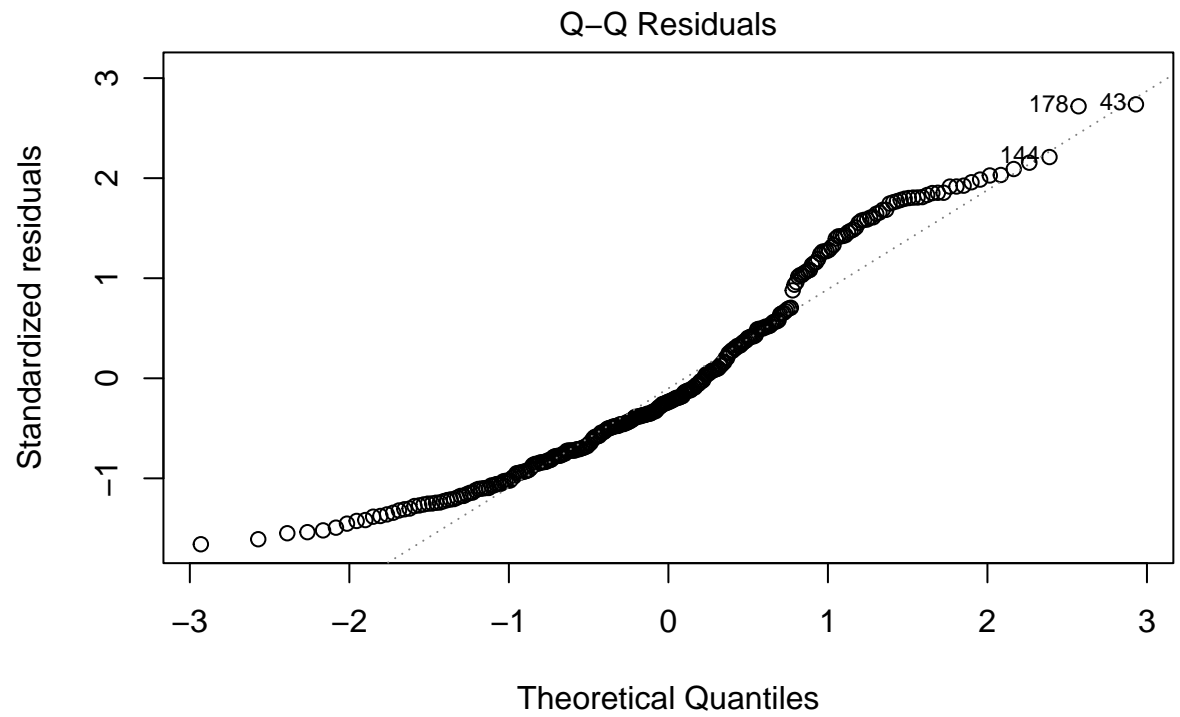
```
lm_model_affiliation <- lm(bodily_contact_time ~ Dummy_Color + Trial + weight + Log_aggression + seconds_moving + anntenation + chambers_entered, data = wasp)
lm_model_affiliation
```

```
##
## Call:
## lm(formula = bodily_contact_time ~ Dummy_Color + Trial + weight +
##     Log_aggression + seconds_moving + anntenation + chambers_entered,
##     data = wasp)
##
## Coefficients:
##      (Intercept) Dummy_Colorgoldsilver Dummy_Colorgreen
##             155.2345             52.2334             28.2654
## Dummy_Colorred Dummy_Colorsilver Dummy_Colorwhite
##          -107.7908             6.1998             37.4223
## Dummy_Coloryellow Trial weight
##          -26.5071      21.3410      291.4209
## Log_aggression seconds_moving anntenation
##          -12.7849          -0.2476          -3.7196
## chambers_entered
##             2.3193
```

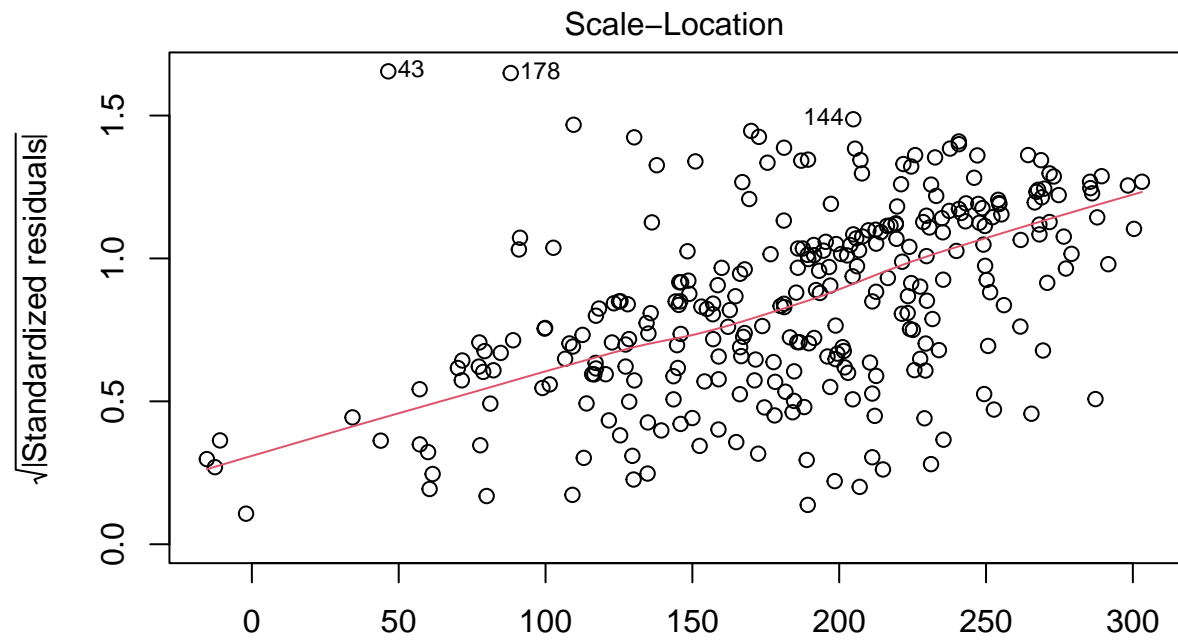
```
lm_empty_affiliation_model <- lm(bodily_contact_time ~ 1, data = wasp)
```

```
plot(lm_model_affiliation)
```

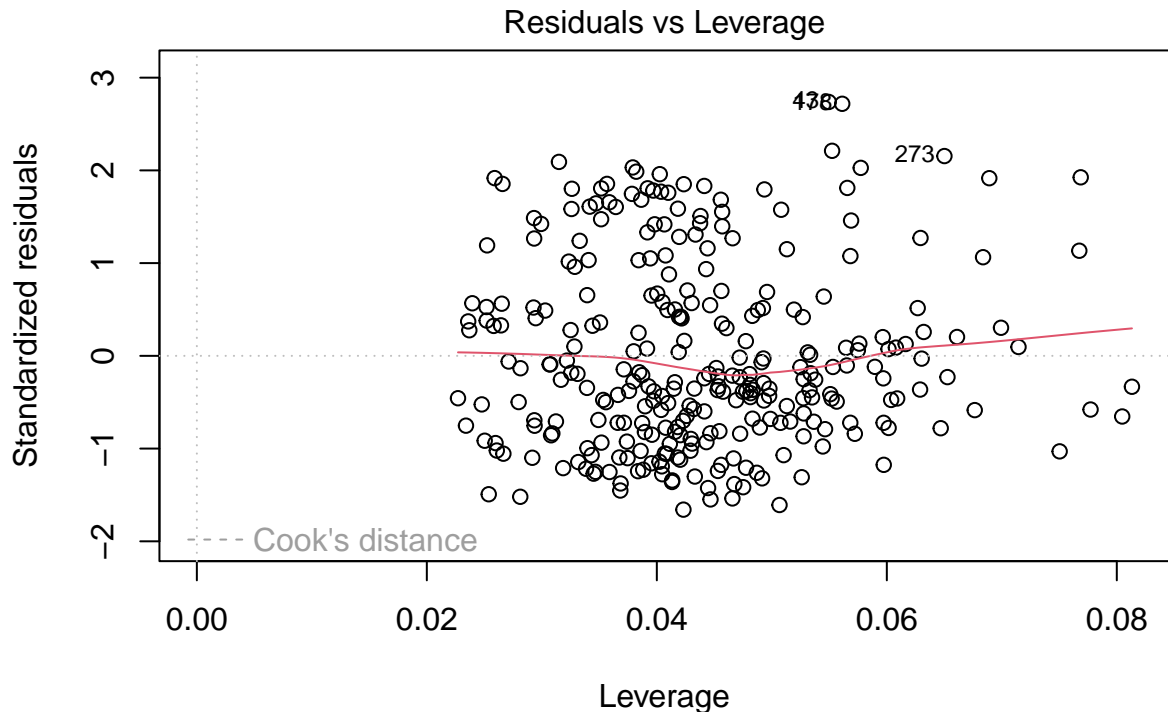




lm(bodily_contact_time ~ Dummy_Color + Trial + weight + Log_aggression + se ...



Fitted values
`lm(bodily_contact_time ~ Dummy_Color + Trial + weight + Log_aggression + se ...`



lm(bodily_contact_time ~ Dummy_Color + Trial + weight + Log_aggression + se ...)

```
n = nrow(wasp)
```

```
aff_backward.model.BIC = stepAIC(lm_model_affiliation, scope = list(lower = lm_empty_affiliation_model
aff_forward.model.BIC = stepAIC(lm_empty_affiliation_model, scope = list(lower = lm_empty_affiliation_model
```

```
## Start: AIC=3096.71
## bodily_contact_time ~ 1
##
##           Df Sum of Sq    RSS    AIC
## + seconds_moving  1    291403  9858807 3093.8
## <none>                        10150210 3096.7
## + chambers_entered  1    183378  9966832 3097.0
## + Trial              1    177237  9972973 3097.2
## + anntenation       1     43558 10106652 3101.1
## + Log_aggression    1     21167 10129043 3101.8
## + weight            1      8887 10141323 3102.1
## + Dummy_Color       6    556433  9593776 3114.2
##
## Step: AIC=3093.78
## bodily_contact_time ~ seconds_moving
##
##           Df Sum of Sq    RSS    AIC
## + Trial      1    261705  9597101 3091.5
## <none>      1    261705  9597101 3091.5
## + anntenation  1     36697  9822110 3098.4
```

```
## + weight          1      10036 9848770 3099.2
## + Log_aggression  1       5211 9853596 3099.3
## + chambers_entered 1       4165 9854642 3099.3
## + Dummy_Color     6      638176 9220630 3108.1
##
## Step: AIC=3091.51
## bodily_contact_time ~ seconds_moving + Trial
##
##              Df Sum of Sq    RSS    AIC
## <none>                9597101 3091.5
## + antennation      1      58302 9538799 3095.4
## + Log_aggression    1      14556 9582545 3096.8
## + weight            1      10214 9586887 3096.9
## + chambers_entered  1       4212 9592889 3097.1
## + Dummy_Color       6      500220 9096881 3109.8
```

```
aff_FB.model.BIC = stepAIC(lm_empty_affiliation_model, scope = list(lower = lm_empty_affiliation_model, upper = lm_model_affiliation))
aff_BF.model.BIC = stepAIC(lm_model_affiliation, scope = list(lower = lm_empty_affiliation_model, upper = lm_model_affiliation))
BIC(aff_backward.model.BIC)
```

```
## [1] 3937.21
```

```
BIC(aff_forward.model.BIC)
```

```
## [1] 3937.21
```

```
BIC(aff_FB.model.BIC)
```

```
## [1] 3937.21
```

```
BIC(aff_BF.model.BIC)
```

```
## [1] 3937.21
```

```
aff_BF.model.BIC
```

```
##
## Call:
## lm(formula = bodily_contact_time ~ Trial + seconds_moving, data = wasp)
##
## Coefficients:
##      (Intercept)          Trial  seconds_moving
##      163.6369         26.9257         -0.2055
```

```
confint(aff_BF.model.BIC)
```

```
##              2.5 %      97.5 %
## (Intercept) 109.0359340 218.23782494
## Trial        8.1782302  45.67314211
## seconds_moving -0.3249099 -0.08611324
```

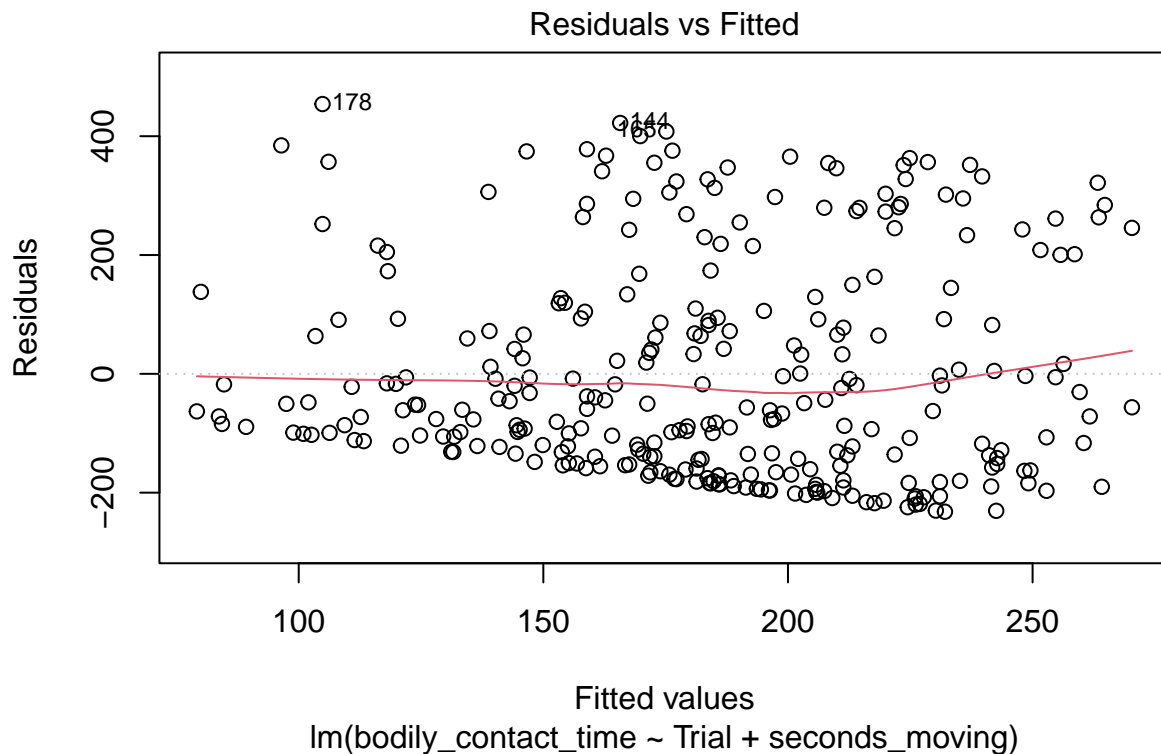
```
lmer_model_affiliation <- lmer(bodily_contact_time ~ Trial + seconds_moving + (1|Wasp.ID), data = wasp)
BIC(lmer_model_affiliation)
```

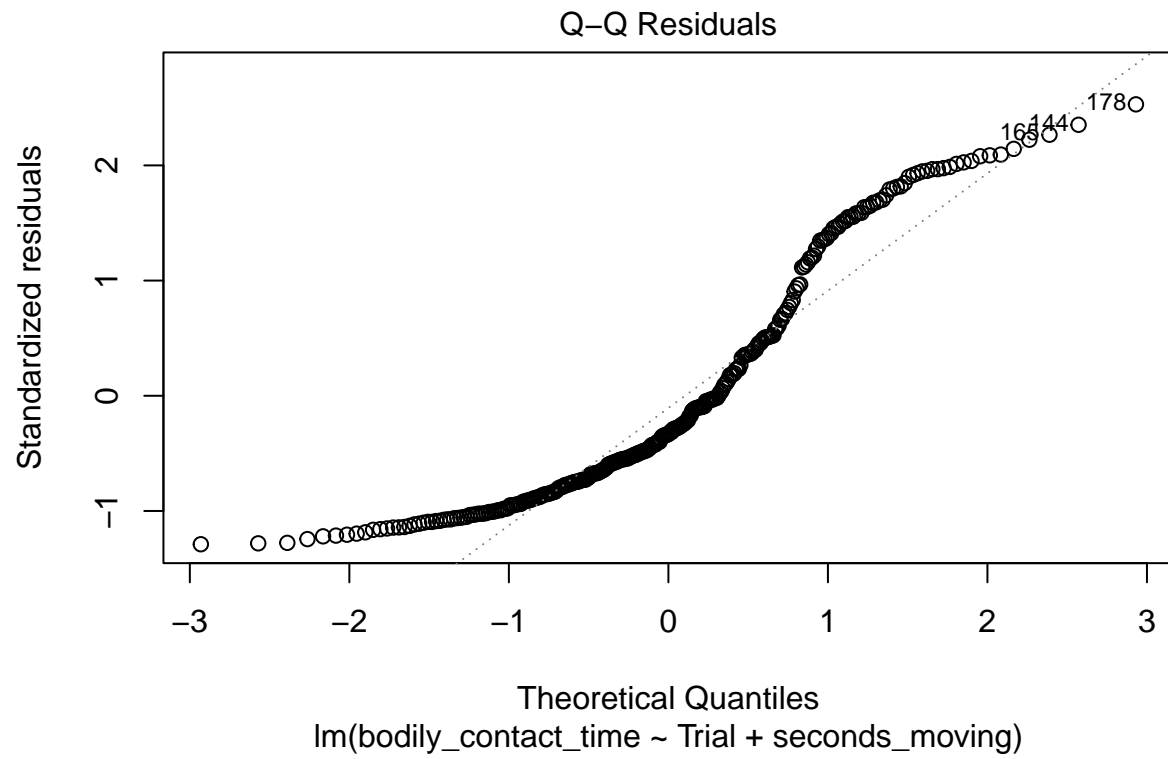
```
## [1] 3898.763
```

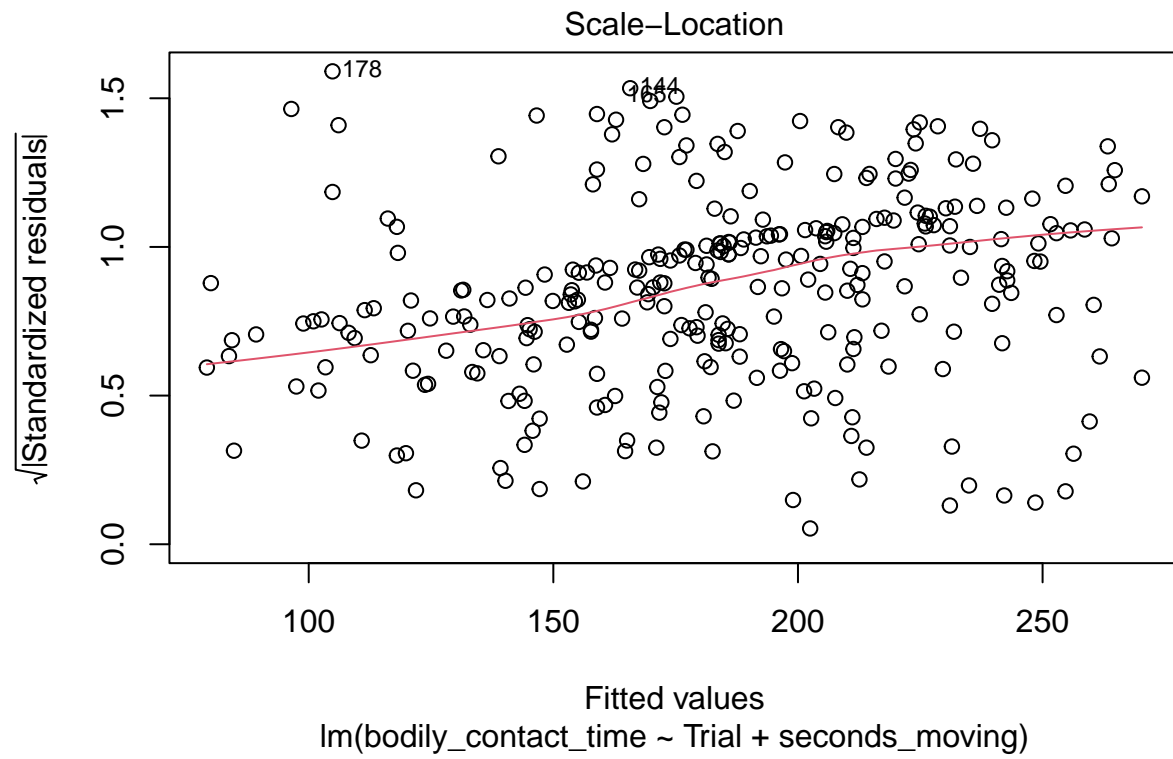
```
lmer_model_affiliation
```

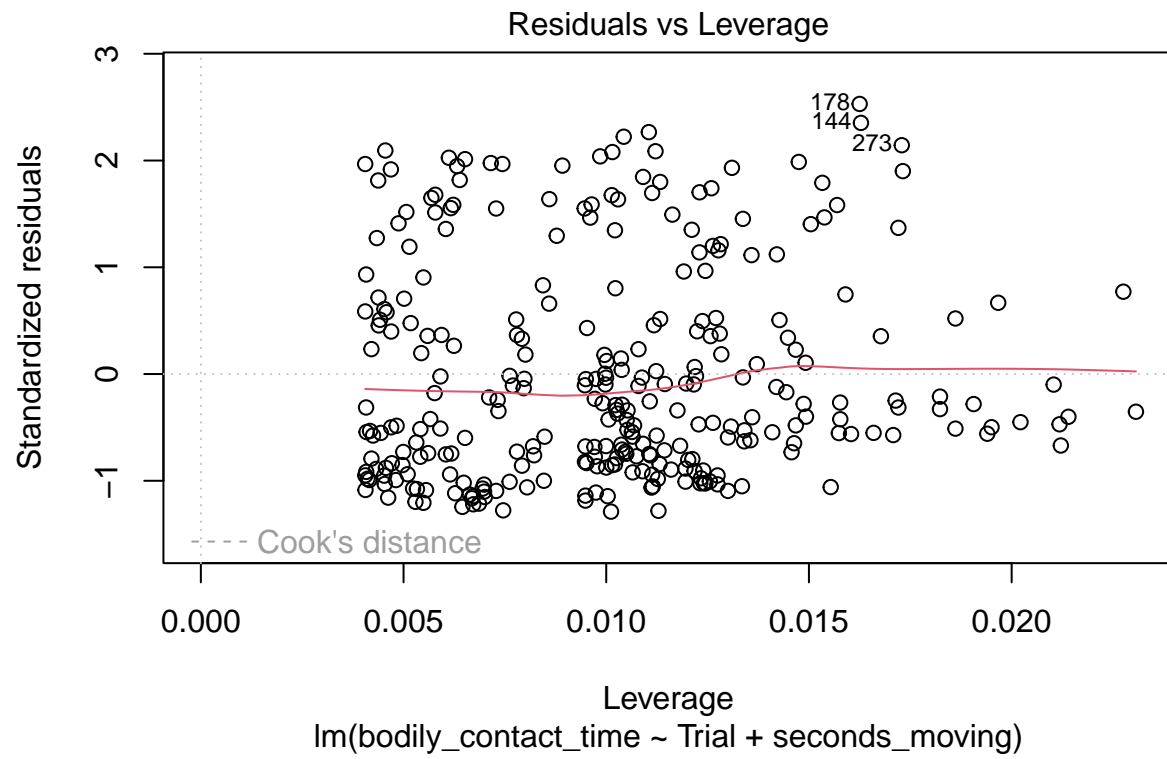
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: bodily_contact_time ~ Trial + seconds_moving + (1 | Wasp.ID)
## Data: wasp
## REML criterion at convergence: 3870.311
## Random effects:
## Groups Name Std.Dev.
## Wasp.ID (Intercept) 104.3
## Residual 149.0
## Number of obs: 296, groups: Wasp.ID, 74
## Fixed Effects:
## (Intercept) Trial seconds_moving
## 148.0218 24.6667 -0.1134
```

```
plot(aff_backward.model.BIC)
```

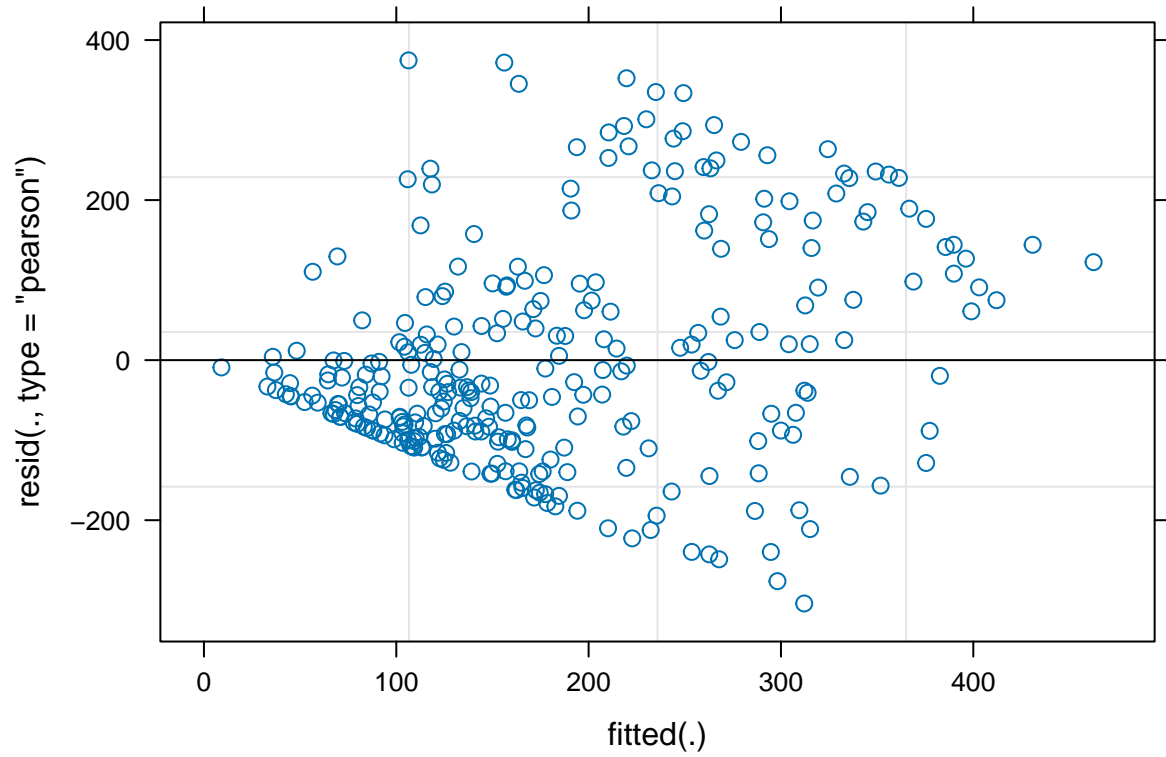






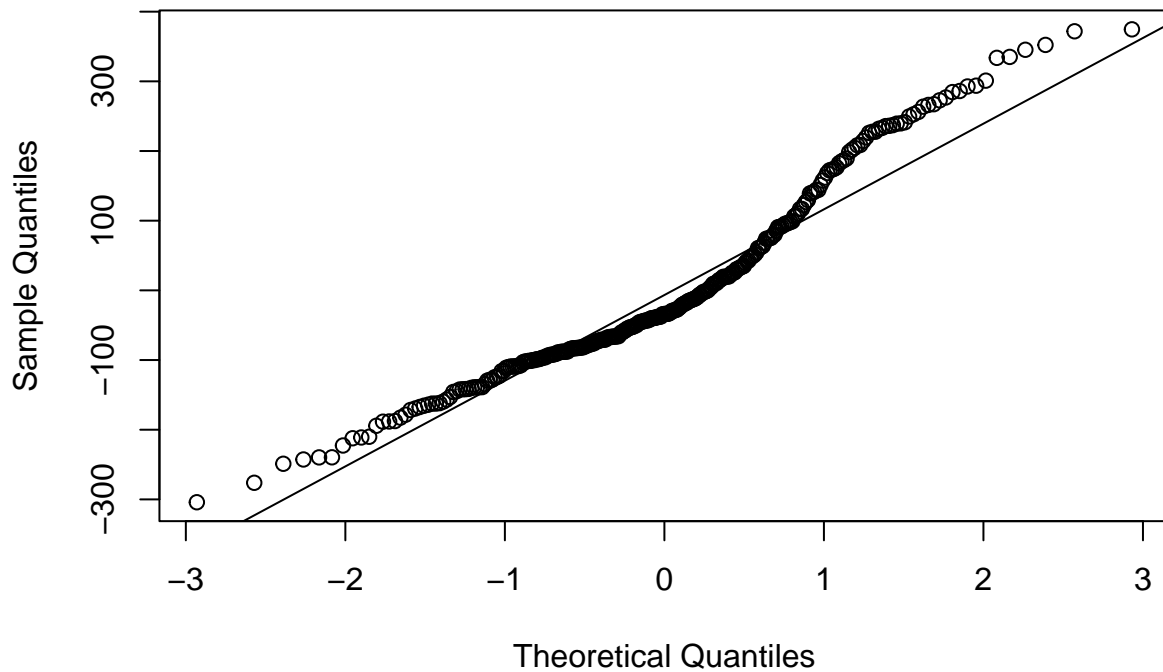


```
plot(lmer_model_affilitation)
```



```
qqnorm(resid(lmer_model_affilitation))  
qqline(resid(lmer_model_affilitation))
```


Normal Q-Q Plot

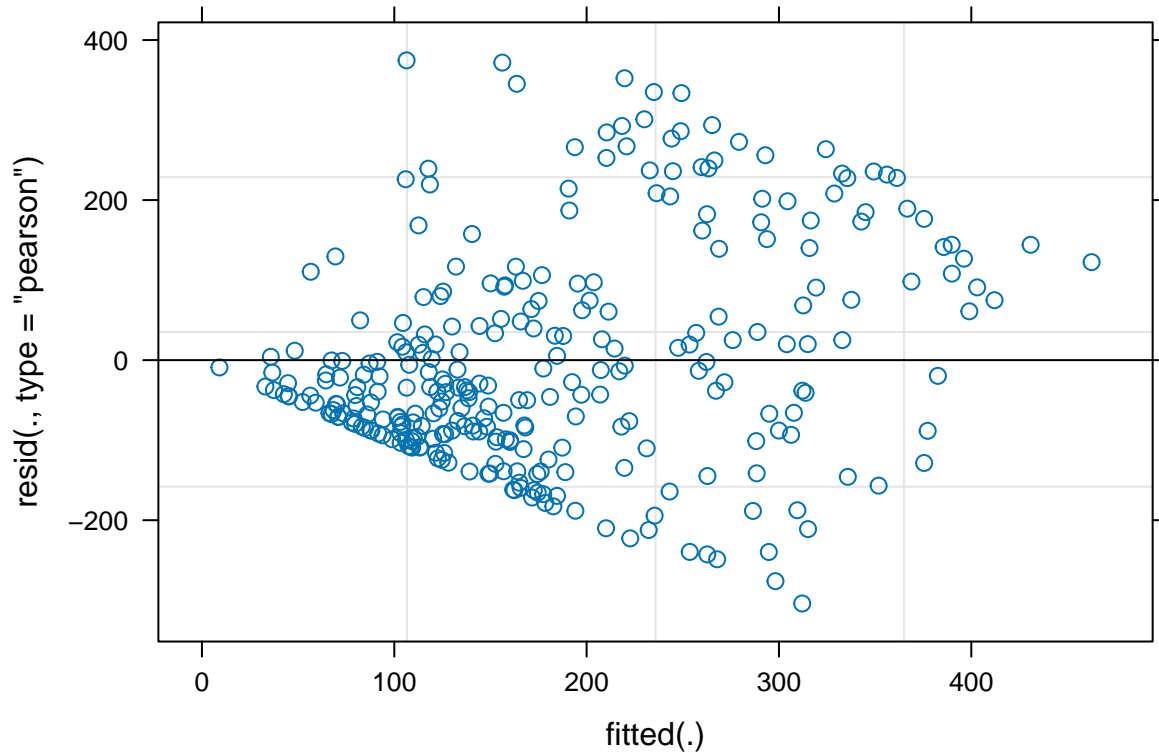


```
shapiro.test(resid(lmer_model_affilitation))
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  resid(lmer_model_affilitation)  
## W = 0.95022, p-value = 1.777e-08
```

```
#Group.aff = rep("Lower",nrow(wasp)) #Creates a vector that repeats "Lower" n times  
#Group.aff[wasp$bodily_contact_time < median(wasp$bodily_contact_time)] = "Upper" #Changing the appropriate  
#Group.aff = as.factor(Group.aff) #Changes it to a factor, which R recognizes as a grouping variable.  
#fligner.test(lmer_model_affilitation$residuals, Group.aff)
```

```
plot(lmer_model_affilitation)
```



```
lm_model_ant <- lm(anntenation ~ Dummy_Color + Trial + weight + Log_aggression + seconds_moving + bodily_contact_time + chambers_entered, data = wasp)
lm_model_ant
```

```
##
## Call:
## lm(formula = anntenation ~ Dummy_Color + Trial + weight + Log_aggression +
##     seconds_moving + bodily_contact_time + chambers_entered,
##     data = wasp)
##
## Coefficients:
##      (Intercept)  Dummy_Colorgoldsilver  Dummy_Colorgreen
##             3.842039             -0.213593             0.546385
##      Dummy_Colorred  Dummy_Colorsilver  Dummy_Colorwhite
##             0.499150             0.257893             0.920762
##      Dummy_Coloryellow      Trial      weight
##             -0.909314      0.340920      6.853397
##      Log_aggression      seconds_moving  bodily_contact_time
##             0.681158      0.002366      -0.001398
##      chambers_entered
##             -0.188734
```

```
lm_empty_ant <- lm(anntenation ~ 1, data = wasp)
```

```
ant_backward.model.BIC = stepAIC(lm_model_ant, scope = list(lower = lm_empty_ant, upper= lm_model_ant))
ant_forward.model.BIC = stepAIC(lm_empty_ant, scope = list(lower = lm_empty_ant, upper= lm_model_ant),
```

```
## Start: AIC=748.14
## anntenation ~ 1
##
##           Df Sum of Sq  RSS   AIC
## + Log_aggression    1   103.378 3532.7 745.30
## + <none>                3636.0 748.14
## + Trial              1    35.743 3600.3 750.91
## + bodily_contact_time 1    15.603 3620.4 752.56
## + weight             1     6.094 3629.9 753.34
## + seconds_moving     1     3.708 3632.3 753.53
## + chambers_entered   1     1.808 3634.2 753.69
## + Dummy_Color        6    91.455 3544.6 774.75
##
## Step: AIC=745.3
## anntenation ~ Log_aggression
##
##           Df Sum of Sq  RSS   AIC
## + <none>                3532.7 745.30
## + Trial              1    23.620 3509.0 749.00
## + bodily_contact_time 1    12.176 3520.5 749.97
## + weight             1     6.480 3526.2 750.44
## + chambers_entered   1     5.970 3526.7 750.49
## + seconds_moving     1     0.289 3532.4 750.96
## + Dummy_Color        6    69.089 3463.6 773.59
```

```
ant_FB.model.BIC = stepAIC(lm_empty_ant, scope = list(lower = lm_empty_ant, upper= lm_model_ant), k = 1)
ant_BF.model.BIC = stepAIC(lm_model_ant, scope = list(lower = lm_empty_ant, upper= lm_model_ant), k = 1)
```

```
BIC(ant_backward.model.BIC)
```

```
## [1] 1590.999
```

```
BIC(ant_forward.model.BIC)
```

```
## [1] 1590.999
```

```
BIC(ant_BF.model.BIC)
```

```
## [1] 1590.999
```

```
BIC(ant_FB.model.BIC)
```

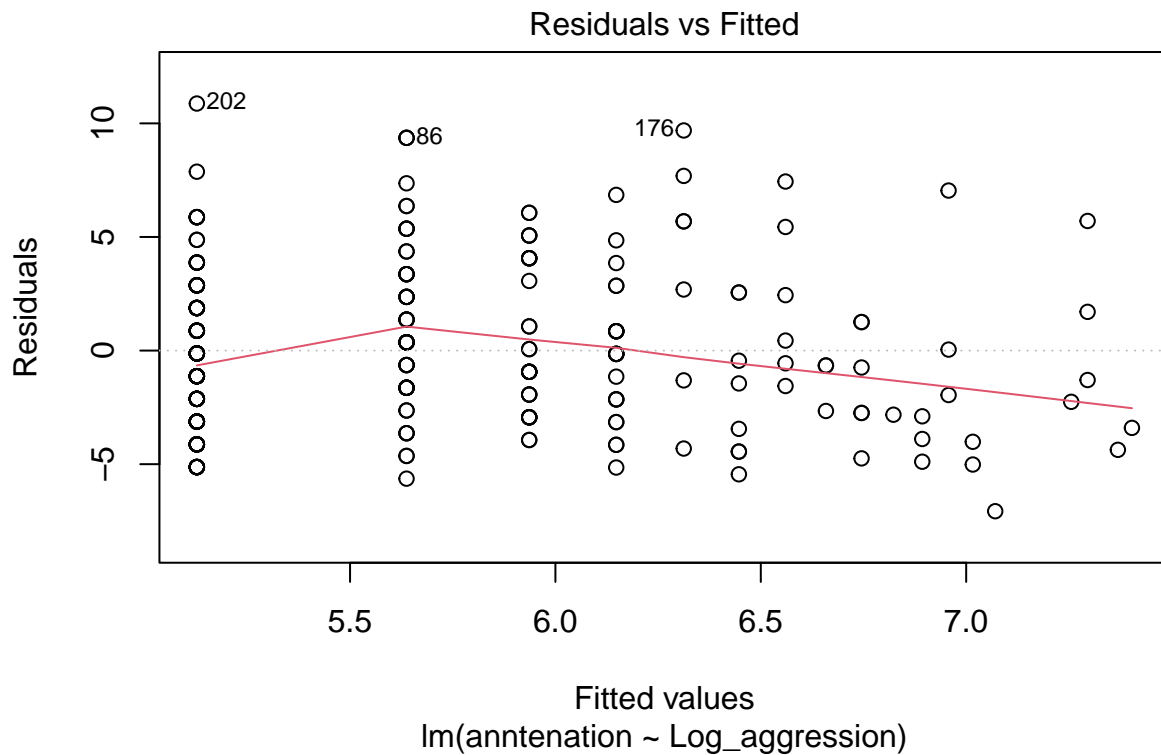
```
## [1] 1590.999
```

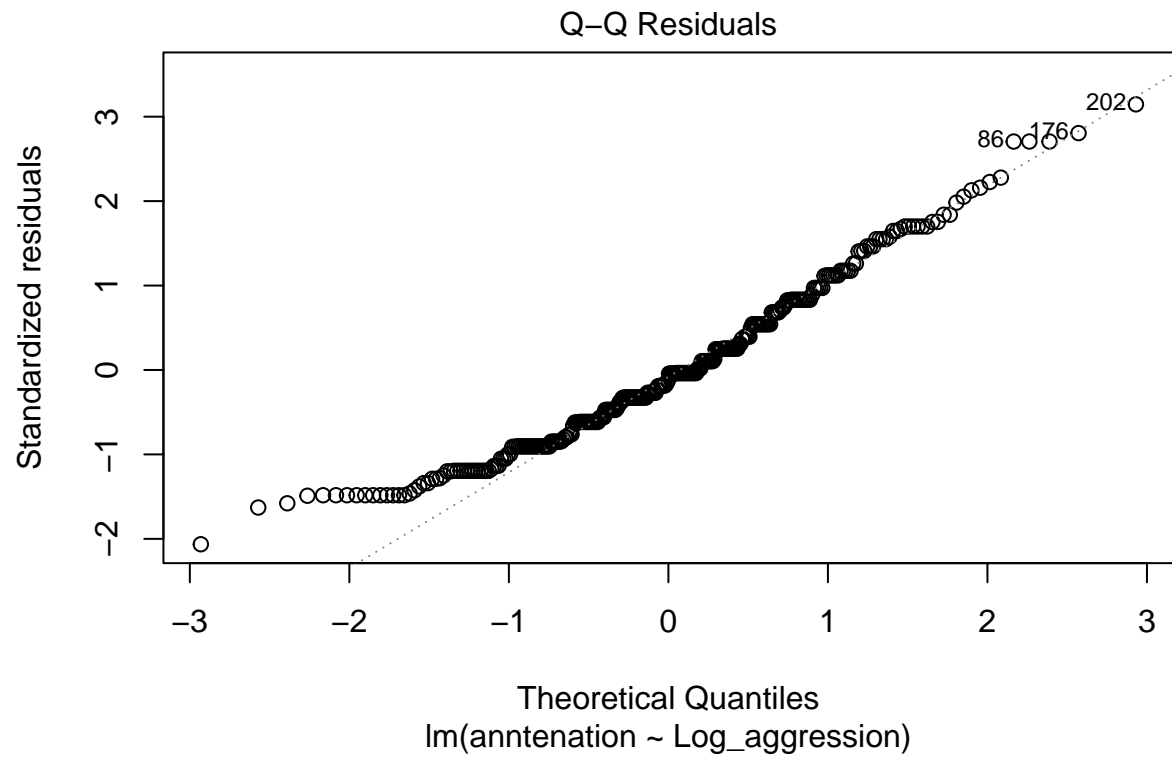
```
ant_BF.model.BIC
```

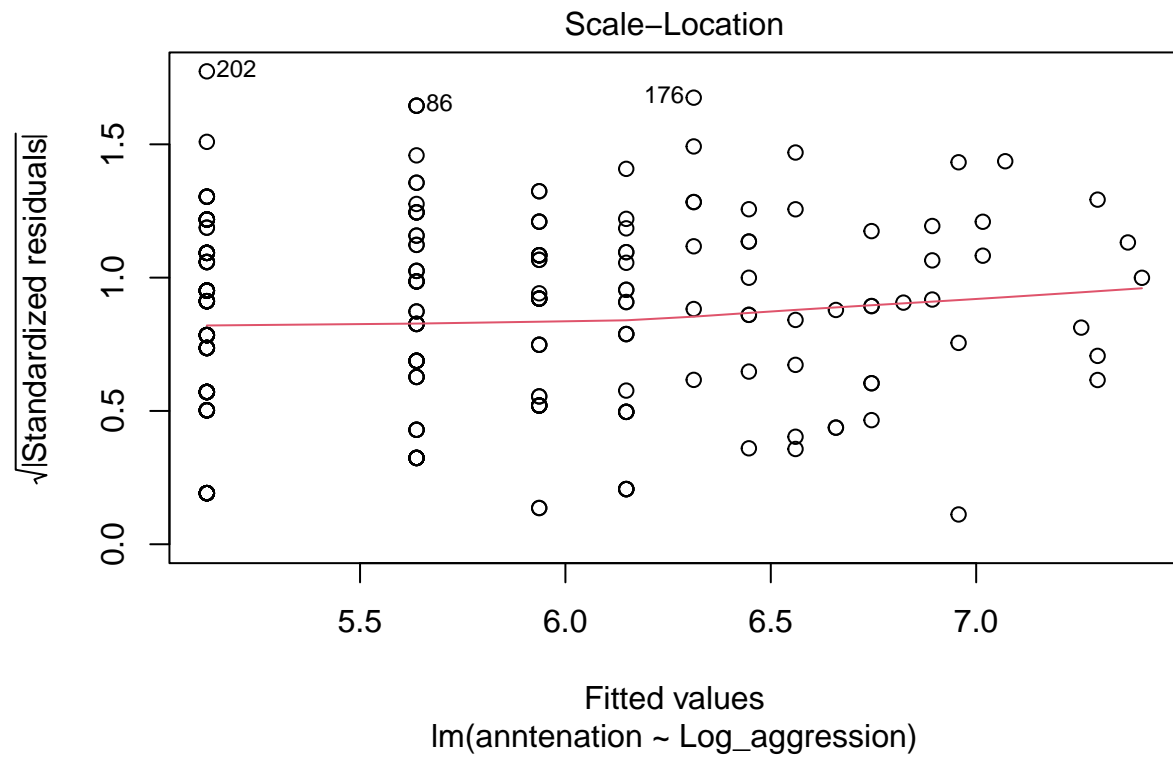
```
##  
## Call:  
## lm(formula = anntenation ~ Log_aggression, data = wasp)  
##  
## Coefficients:  
##      (Intercept)      Log_aggression  
##          5.1270           0.7365
```

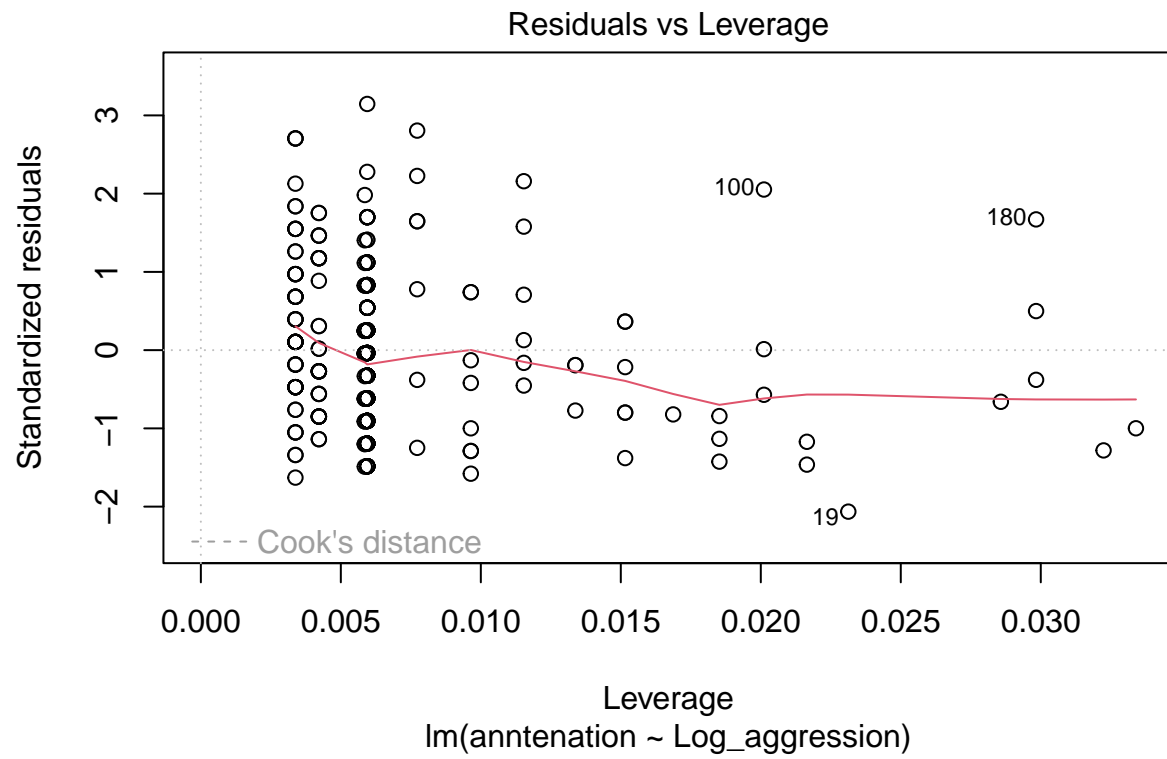
```
lmer_model_ant <- lmer(anntenation ~ Log_aggression + (1|Wasp.ID), data = wasp)
```

```
plot(ant_BF.model.BIC)
```

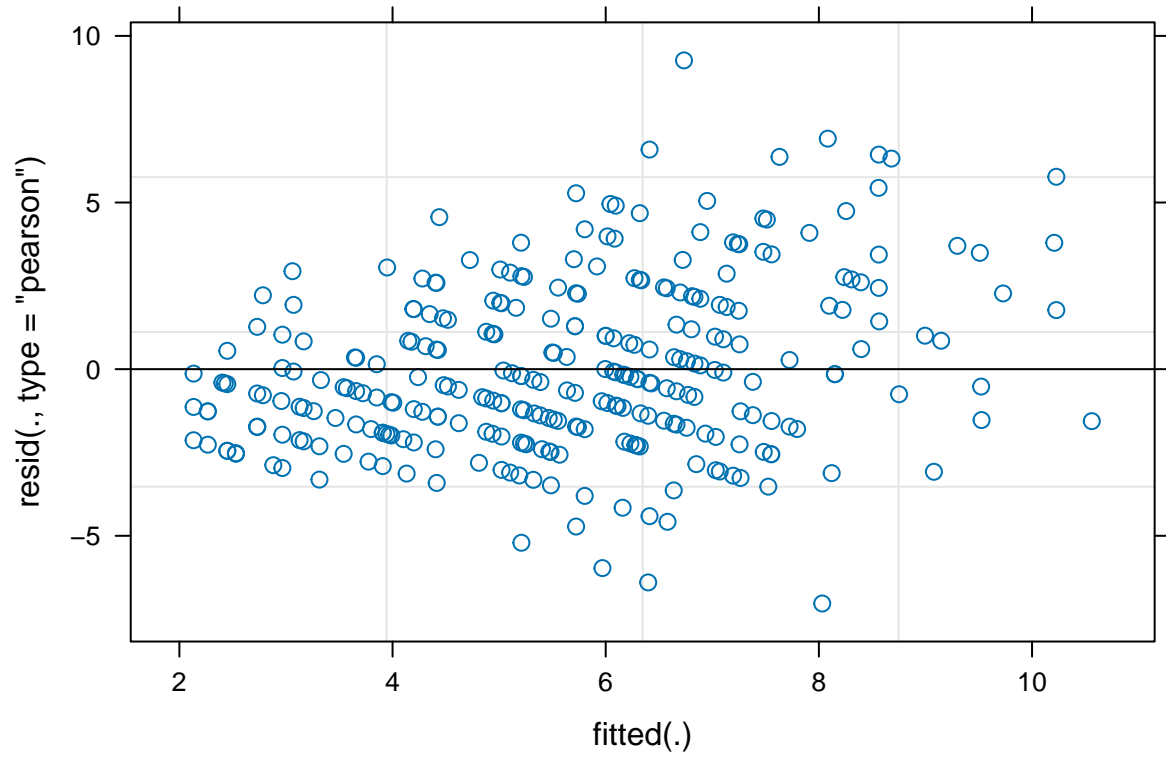






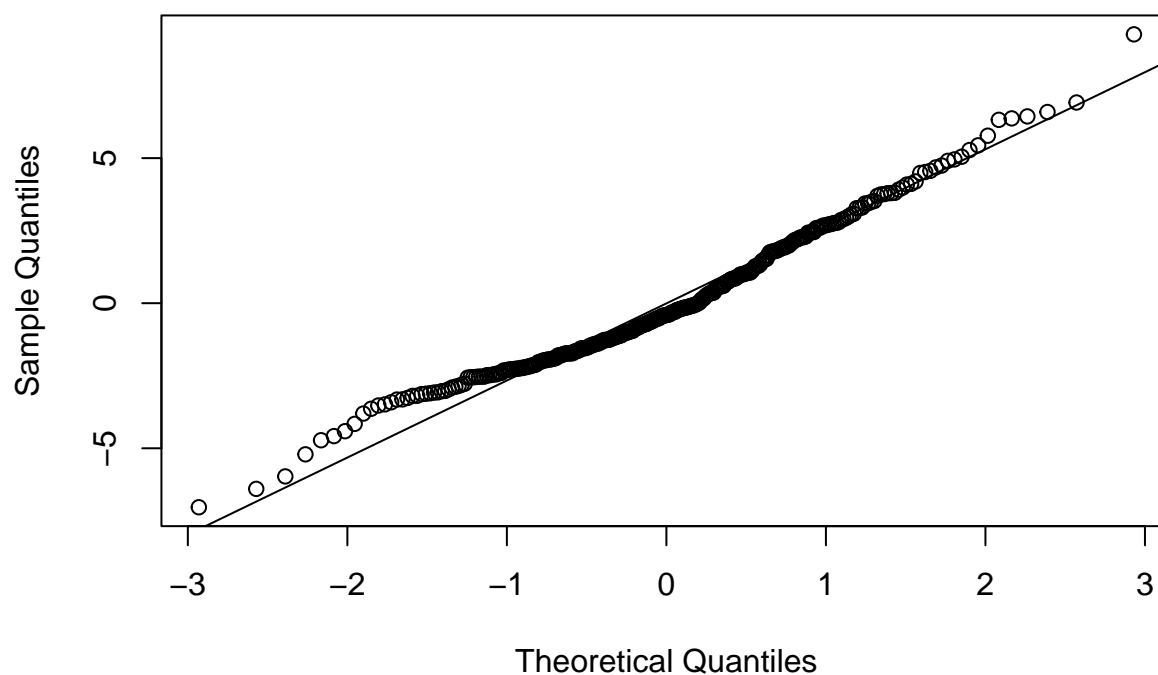


```
plot(lmer_model_ant)
```



```
qqnorm(resid(lmer_model_ant))  
qqline(resid(lmer_model_ant))
```


Normal Q-Q Plot



```
shapiro.test(resid(lmer_model_ant))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  resid(lmer_model_ant)
## W = 0.9772, p-value = 0.0001166
```

```
###mixed effect model better.
```

```
BIC(lmer_model_ant)
```

```
## [1] 1555.338
```

```
lm_model_agg <- lm(Log_aggression ~ Dummy_Color + Trial + weight + bodily_contact_time + seconds_moving
lm_model_agg
```

```
##
## Call:
## lm(formula = Log_aggression ~ Dummy_Color + Trial + weight +
##     bodily_contact_time + seconds_moving + anntenation + chambers_entered,
##     data = wasp)
##
```

```
## Coefficients:
##      (Intercept) Dummy_Colorgoldsilver Dummy_Colorgreen
##      0.3166691      0.5499208      -0.0585080
##      Dummy_Colorred      Dummy_Colorsilver      Dummy_Colorwhite
##      -0.0641304      0.5887456      -0.1366981
##      Dummy_Coloryellow      Trial      weight
##      -0.1920834      -0.0390310      0.5698611
##      bodily_contact_time      seconds_moving      anntenation
##      -0.0002190      0.0007041      0.0310535
##      chambers_entered
##      -0.0144865
```

```
lm_empty_agg <- lm(Log_aggression ~ 1, data = wasp)
```

```
agg_backward.model.BIC = stepAIC(lm_model_agg, scope = list(lower = lm_empty_agg, upper= lm_model_agg))
agg_forward.model.BIC = stepAIC(lm_empty_agg, scope = list(lower = lm_empty_agg, upper= lm_model_agg),
```

```
## Start: AIC=-124.63
## Log_aggression ~ 1
##
##           Df Sum of Sq  RSS    AIC
## + Dummy_Color      6   29.3719 161.21 -140.03
## + anntenation      1    5.4185 185.16 -127.48
## <none>              190.58 -124.63
## + seconds_moving    1    3.5797 187.00 -124.55
## + Trial              1    2.4368 188.15 -122.75
## + chambers_entered  1    2.1703 188.41 -122.33
## + bodily_contact_time 1    0.3974 190.18 -119.56
## + weight            1    0.0109 190.57 -118.96
##
## Step: AIC=-140.03
## Log_aggression ~ Dummy_Color
##
##           Df Sum of Sq  RSS    AIC
## + anntenation      1    3.6845 157.53 -141.19
## <none>              161.21 -140.03
## + seconds_moving    1    2.3637 158.85 -138.71
## + bodily_contact_time 1    1.3528 159.86 -136.84
## + chambers_entered  1    1.0123 160.20 -136.21
## + Trial              1    0.1200 161.09 -134.56
## + weight            1    0.0778 161.13 -134.49
##
## Step: AIC=-141.19
## Log_aggression ~ Dummy_Color + anntenation
##
##           Df Sum of Sq  RSS    AIC
## <none>              157.53 -141.19
## + seconds_moving    1    2.24455 155.28 -139.74
## + chambers_entered  1    1.15309 156.37 -137.67
## + bodily_contact_time 1    1.04936 156.48 -137.47
## + Trial              1    0.26916 157.26 -136.00
## + weight            1    0.03459 157.49 -135.56
```

```
agg_FB.model.BIC = stepAIC(lm_empty_agg, scope = list(lower = lm_empty_agg, upper= lm_model_agg), k = 1)
agg_BF.model.BIC = stepAIC(lm_model_agg, scope = list(lower = lm_empty_agg, upper= lm_model_agg), k = 1)
```

```
BIC(agg_backward.model.BIC)
```

```
## [1] 704.5164
```

```
BIC(agg_forward.model.BIC)
```

```
## [1] 704.5164
```

```
BIC(agg_BF.model.BIC)
```

```
## [1] 704.5164
```

```
BIC(agg_FB.model.BIC)
```

```
## [1] 704.5164
```

```
agg_BF.model.BIC
```

```
##
## Call:
## lm(formula = Log_aggression ~ Dummy_Color + anntenation, data = wasp)
##
## Coefficients:
##          (Intercept) Dummy_Colorgoldsilver Dummy_Colorgreen
##             0.33147      0.53597             -0.02882
## Dummy_Colorred      Dummy_Colorsilver Dummy_Colorwhite
##        -0.02245             0.59191             -0.09271
## Dummy_Coloryellow      anntenation
##        -0.15931             0.03224
```

```
confint(agg_BF.model.BIC)
```

```
##              2.5 %      97.5 %
## (Intercept)  0.088915645 0.57401911
## Dummy_Colorgoldsilver 0.254389237 0.81754111
## Dummy_Colorgreen    -0.347165274 0.28952096
## Dummy_Colorred      -0.367823461 0.32293161
## Dummy_Colorsilver    0.315677715 0.86813840
## Dummy_Colorwhite    -0.414073162 0.22865845
## Dummy_Coloryellow   -0.484804522 0.16619211
## anntenation         0.007790959 0.05669036
```

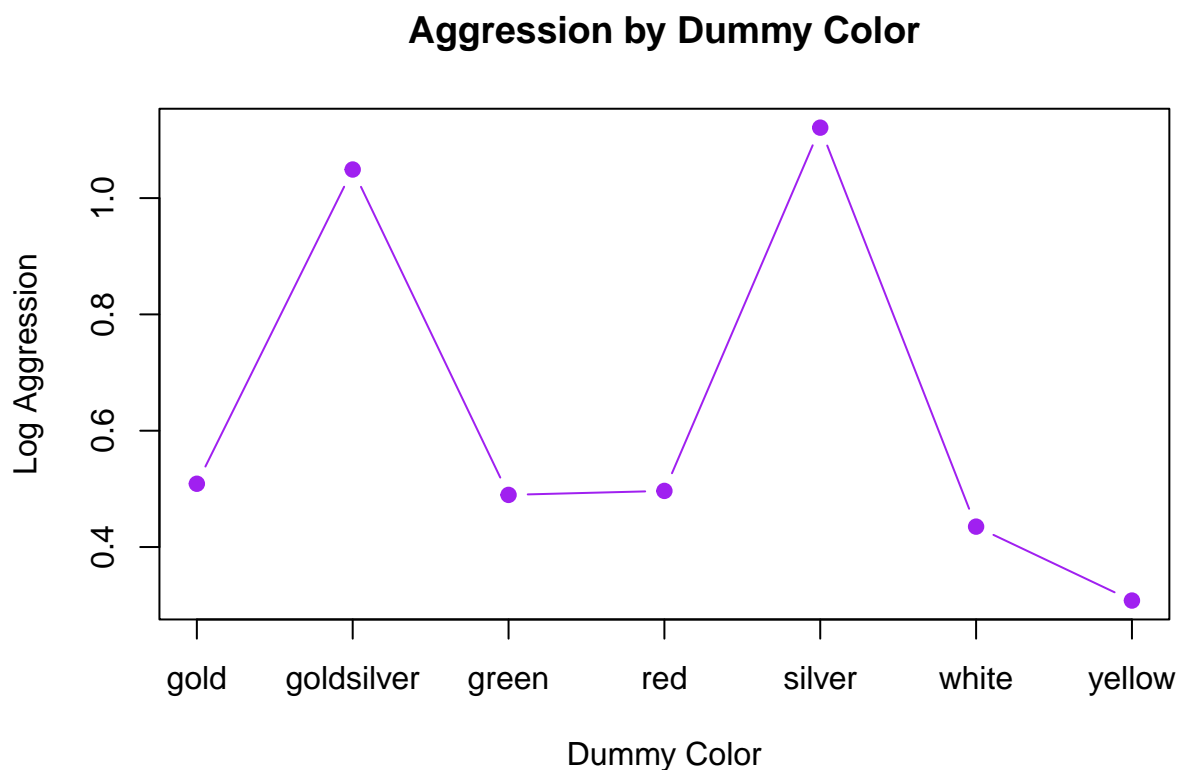
```
###anova on color
```

```
library(asbio)
```

```
## Loading required package: tcltk
```

```
anova_empty_model_aggression = lm(Log_aggression ~ 1, data =wasp)
anova_full_model_aggression = lm(Log_aggression ~ Dummy_Color, data=wasp)
```

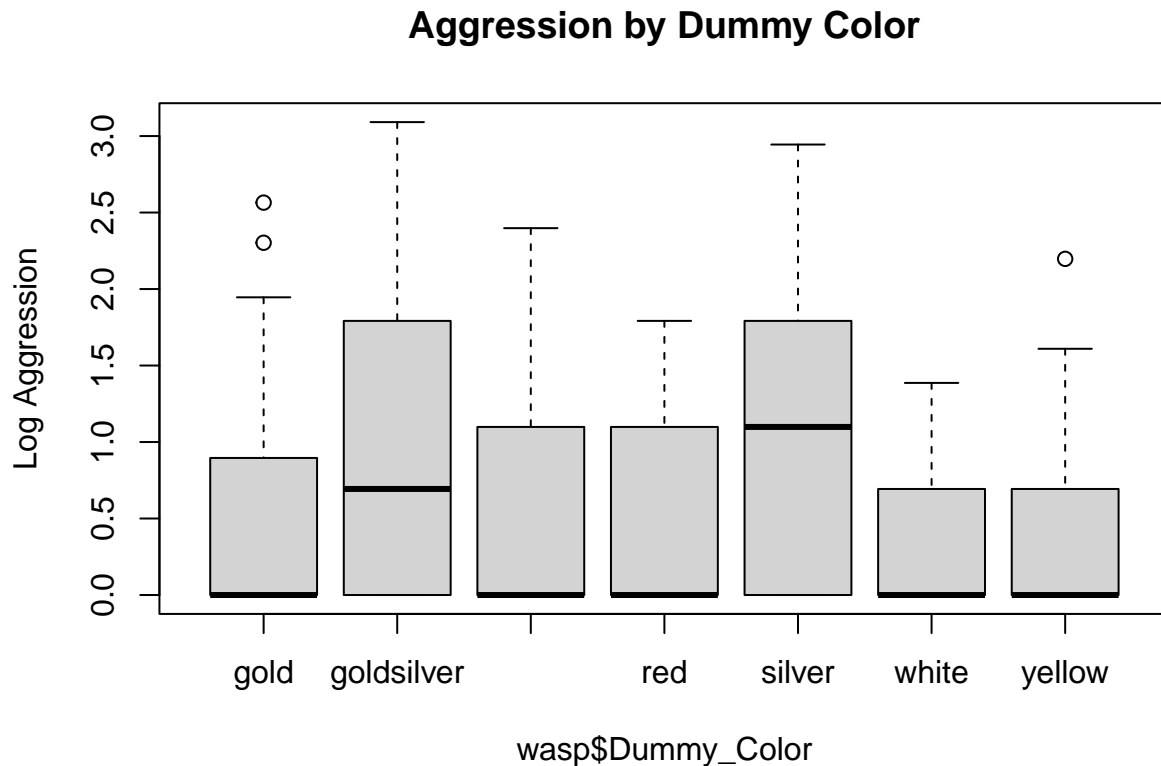
```
group.means = by(wasp$Log_aggression,wasp$Dummy_Color,mean) # First argument is Y, second is grouping
plot(group.means,xaxt = "n",pch = 19,col = "purple",xlab = "Dummy Color",ylab = "Log Aggression",main =
axis(1,1:length(group.means),names(group.means)) #Adding in our own X axis names
```



```
anova(anova_empty_model_aggression, anova_full_model_aggression)
```

```
## Analysis of Variance Table
##
## Model 1: Log_aggression ~ 1
## Model 2: Log_aggression ~ Dummy_Color
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1     295 190.58
## 2     289 161.21   6    29.372 8.7758 8.56e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
boxplot(wasp$Log_aggression ~ wasp$Dummy_Color, main = "Aggression by Dummy Color", ylab = "Log Aggression")
```



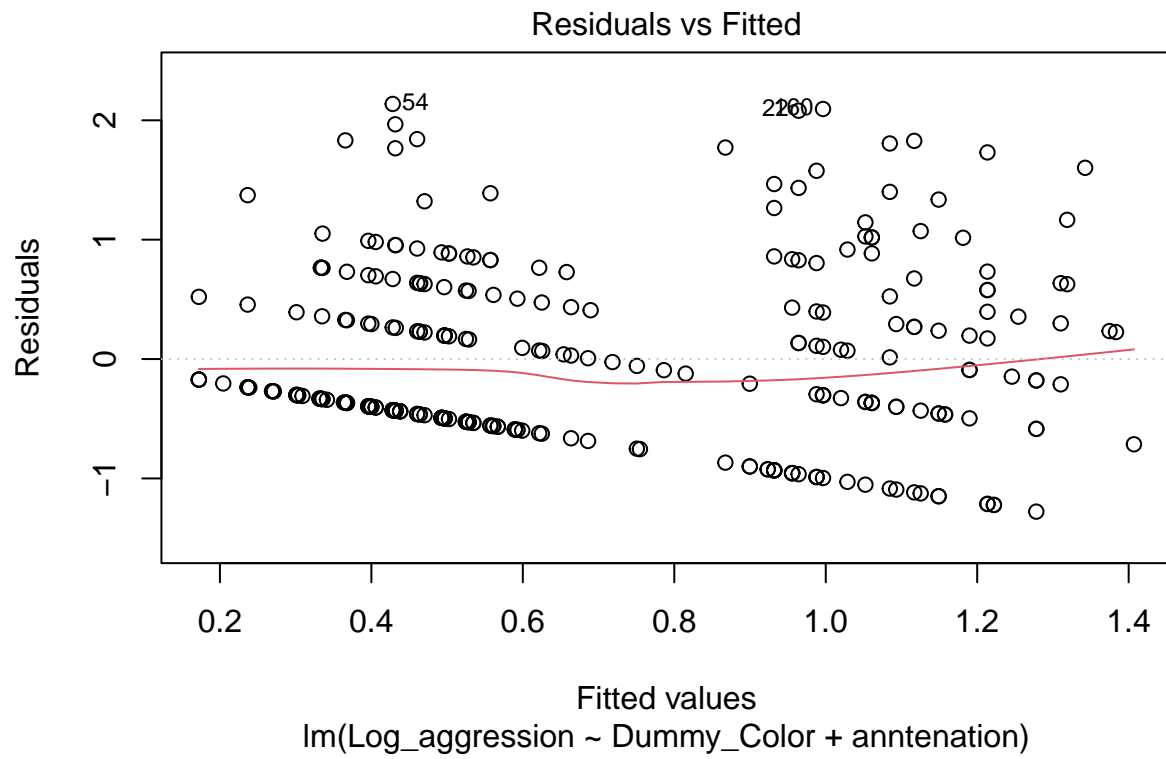
```
shapiro.test(resid(agg_BF.model.BIC))
```

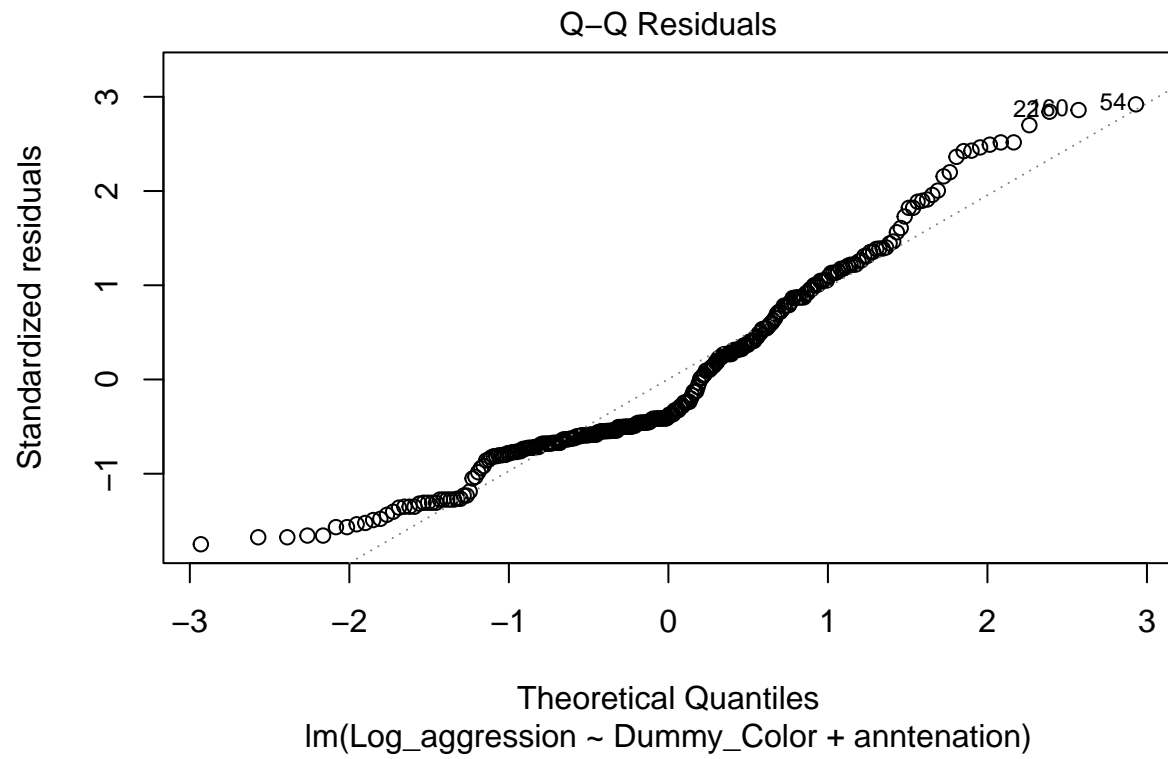
```
##
## Shapiro-Wilk normality test
##
## data:  resid(agg_BF.model.BIC)
## W = 0.94043, p-value = 1.49e-09
```

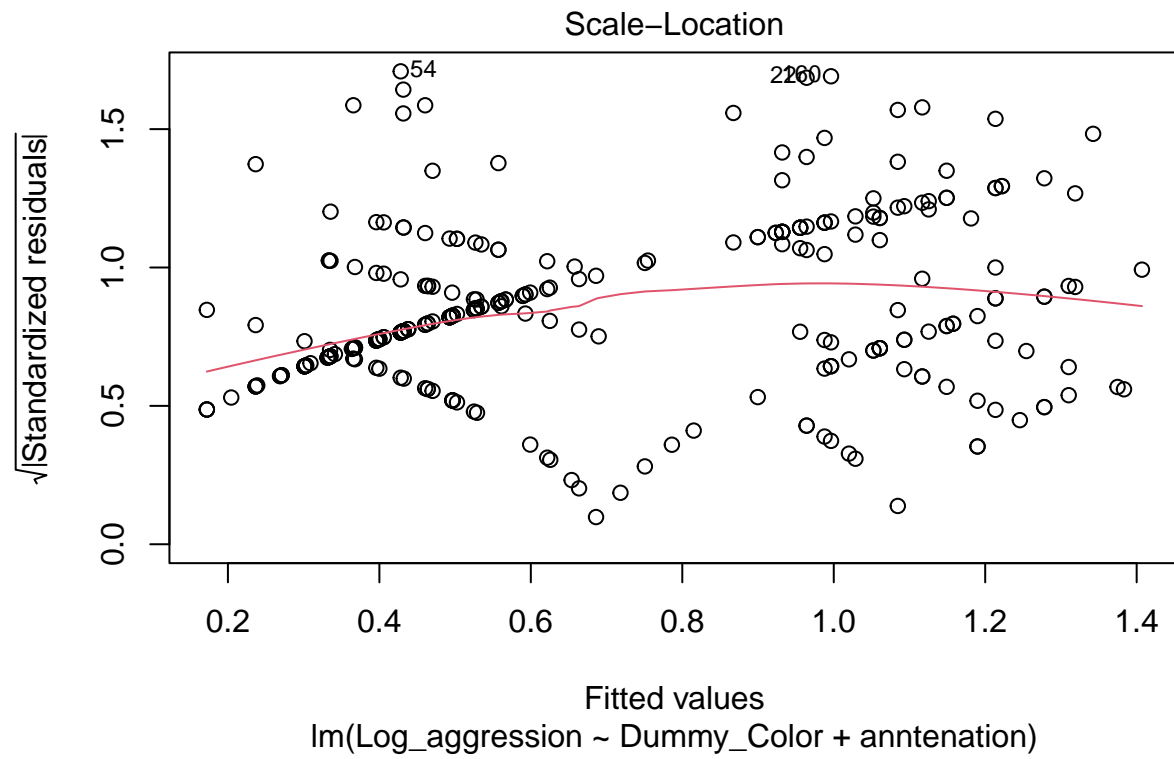
```
Group.agg = rep("Lower", nrow(wasp)) #Creates a vector that repeats "Lower" n times
Group.agg[wasp$Log_aggression < median(wasp$Log_aggression)] = "Upper" #Changing the appropriate values
Group.agg = as.factor(Group.agg) #Changes it to a factor, which R recognizes as a grouping variable.
fligner.test(agg_BF.model.BIC$residuals, Group.agg)
```

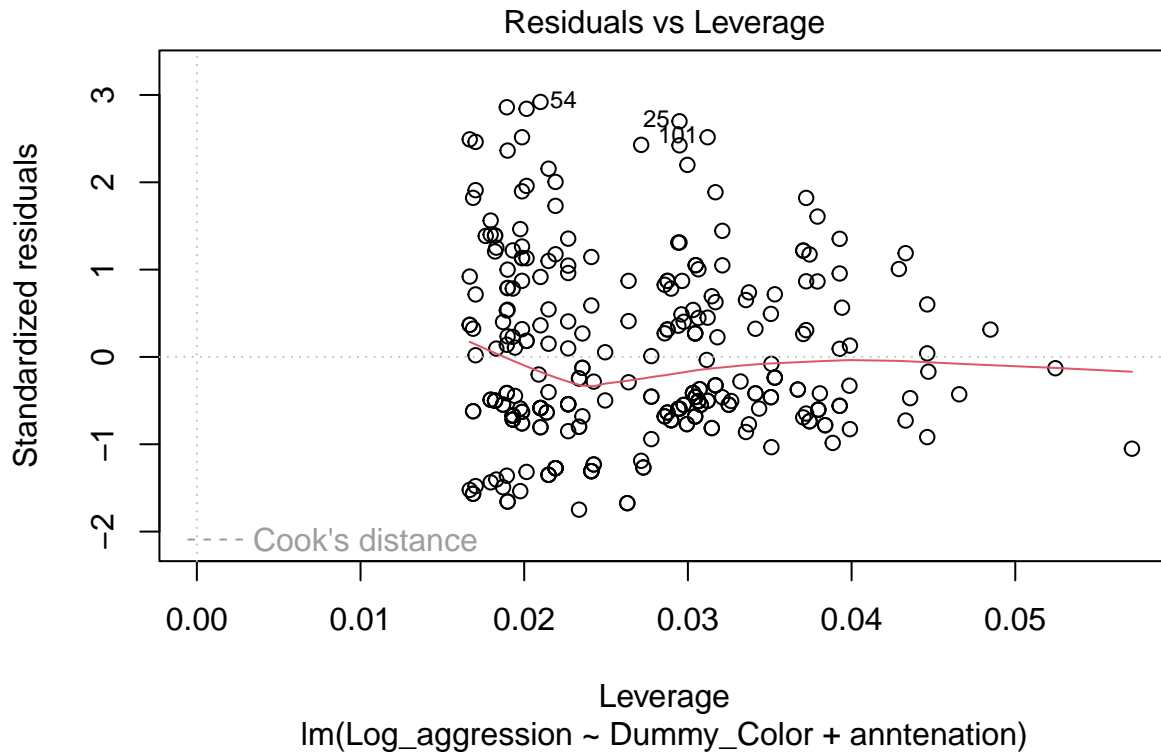
```
##
## Fligner-Killeen test of homogeneity of variances
##
## data:  agg_BF.model.BIC$residuals and Group.agg
## Fligner-Killeen:med chi-squared = 58.732, df = 1, p-value = 1.807e-14
```

```
plot(agg_BF.model.BIC)
```









```
lmer_model_agg <- lmer(Log_aggression ~ Dummy_Color + anntenation + (1|Wasp.ID), data = wasp)
lmer_model_agg
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Log_aggression ~ Dummy_Color + anntenation + (1 | Wasp.ID)
## Data: wasp
## REML criterion at convergence: 668.824
## Random effects:
## Groups Name Std.Dev.
## Wasp.ID (Intercept) 0.2920
## Residual 0.6798
## Number of obs: 296, groups: Wasp.ID, 74
## Fixed Effects:
## (Intercept) Dummy_Colorgolddsilver Dummy_Colorgreen
## 0.318735 0.527531 -0.009458
## Dummy_Colorred Dummy_Colorsilver Dummy_Colorwhite
## 0.027020 0.584500 -0.046199
## Dummy_Coloryellow anntenation
## -0.125362 0.032218
```

```
BIC(lmer_model_agg)
```

```
## [1] 725.7275
```

```
lm_model_exp <- lm(chambers_entered ~ Dummy_Color + Trial + weight + Log_aggression + seconds_moving +
lm_model_exp
```

```
##
## Call:
## lm(formula = chambers_entered ~ Dummy_Color + Trial + weight +
##     Log_aggression + seconds_moving + bodily_contact_time + anntenation,
##     data = wasp)
##
## Coefficients:
##             (Intercept) Dummy_Colorgolddsilver      Dummy_Colorgreen
##                2.587045          0.054953          0.127885
##      Dummy_Colorred      Dummy_Colorsilver      Dummy_Colorwhite
##                0.112001          0.222856          -0.474804
##      Dummy_Coloryellow          Trial          weight
##                -0.240149          -0.041701          -3.678774
##      Log_aggression      seconds_moving      bodily_contact_time
##                -0.070734          0.015302          0.000194
##      anntenation
##                -0.042012
```

```
lm_empty_exp <- lm(chambers_entered ~ 1, data = wasp)
```

```
exp_backward.model.BIC = stepAIC(lm_model_exp, scope = list(lower = lm_empty_exp, upper= lm_model_exp)
exp_forward.model.BIC = stepAIC(lm_empty_exp, scope = list(lower = lm_empty_exp, upper= lm_model_exp),
```

```
## Start:  AIC=679.19
## chambers_entered ~ 1
##
##              Df Sum of Sq    RSS    AIC
## + seconds_moving      1    2106.82  773.61 295.75
## <none>                                2880.43 679.19
## + bodily_contact_time  1      52.04 2828.39 679.48
## + Trial                 1      51.10 2829.33 679.58
## + Log_aggression       1      32.80 2847.63 681.49
## + anntenation          1       1.43 2879.00 684.73
## + weight               1       0.93 2879.50 684.78
## + Dummy_Color          6      45.26 2835.17 708.64
##
## Step:  AIC=295.75
## chambers_entered ~ seconds_moving
##
##              Df Sum of Sq    RSS    AIC
## <none>                                773.61 295.75
## + anntenation          1      7.0960 766.51 298.71
## + weight               1      2.1437 771.46 300.62
## + bodily_contact_time  1      0.3268 773.28 301.31
## + Log_aggression       1      0.3235 773.28 301.32
## + Trial                 1      0.0004 773.60 301.44
## + Dummy_Color          6     10.1387 763.47 325.99
```

```
exp_FB.model.BIC = stepAIC(lm_empty_exp, scope = list(lower = lm_empty_exp, upper= lm_model_exp), k = 1)
exp_BF.model.BIC = stepAIC(lm_model_exp, scope = list(lower = lm_empty_exp, upper= lm_model_exp), k = 1)
```

```
BIC(exp_backward.model.BIC)
```

```
## [1] 1141.451
```

```
BIC(exp_forward.model.BIC)
```

```
## [1] 1141.451
```

```
BIC(exp_BF.model.BIC)
```

```
## [1] 1141.451
```

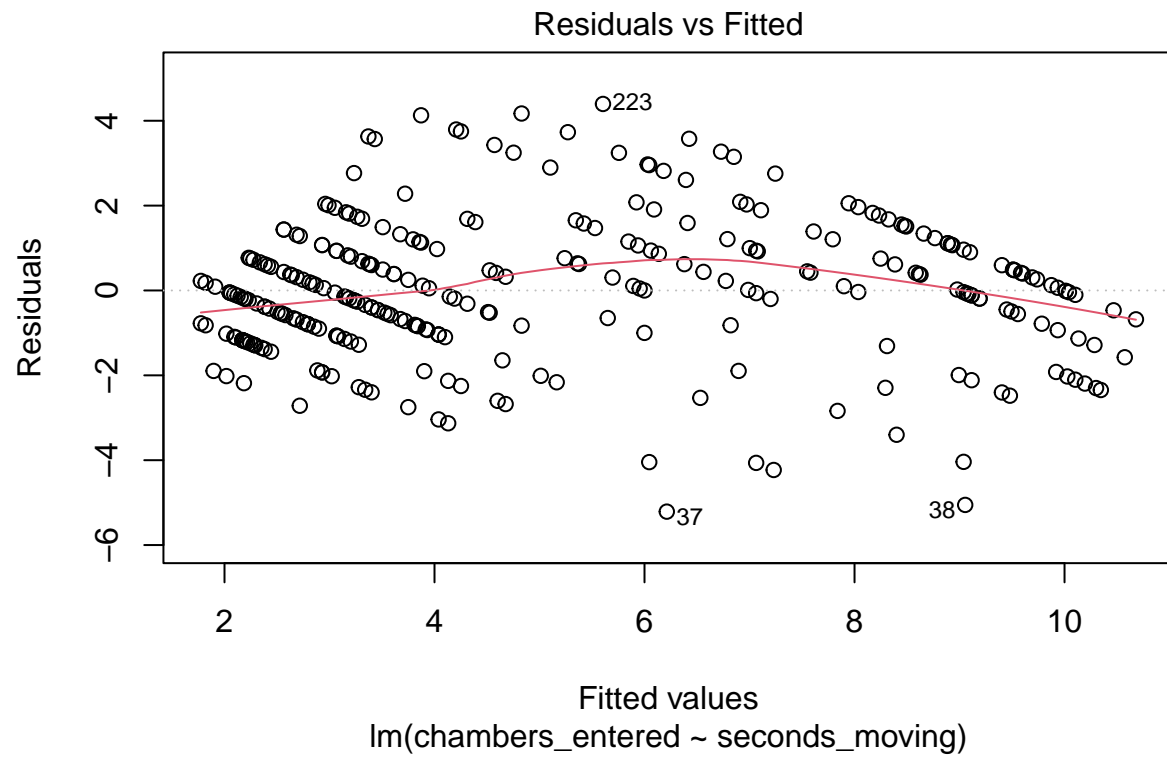
```
BIC(exp_FB.model.BIC)
```

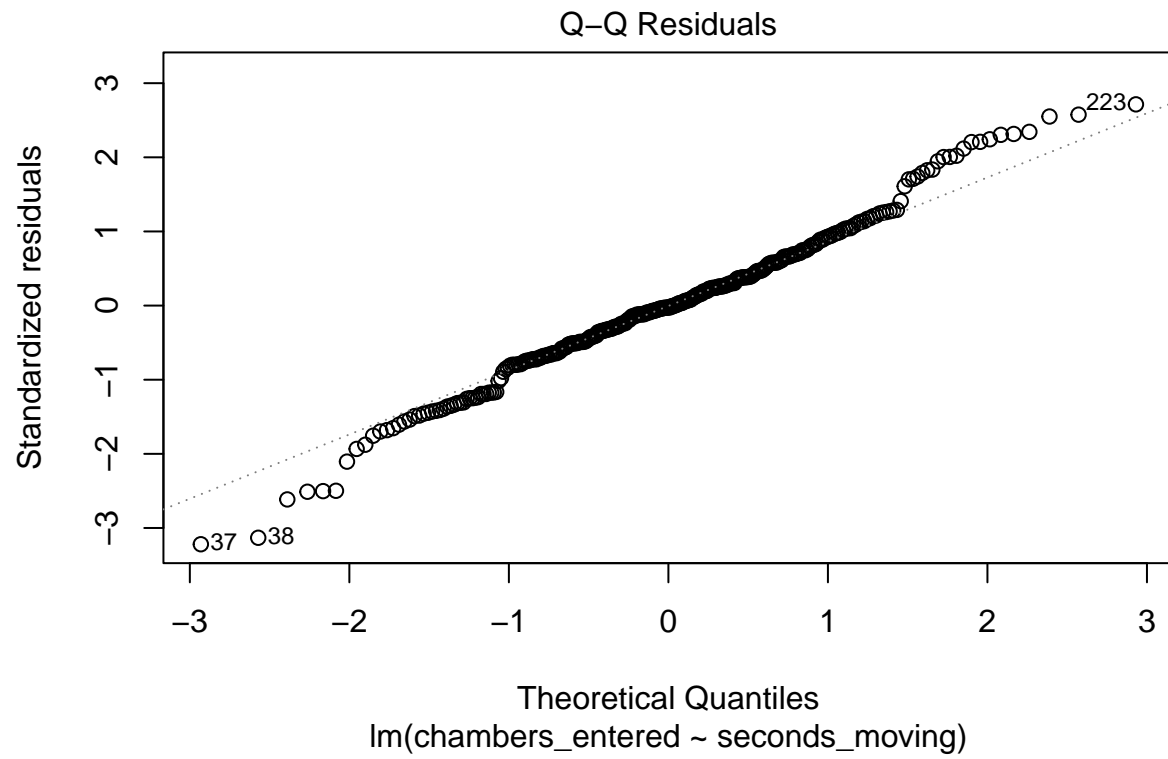
```
## [1] 1141.451
```

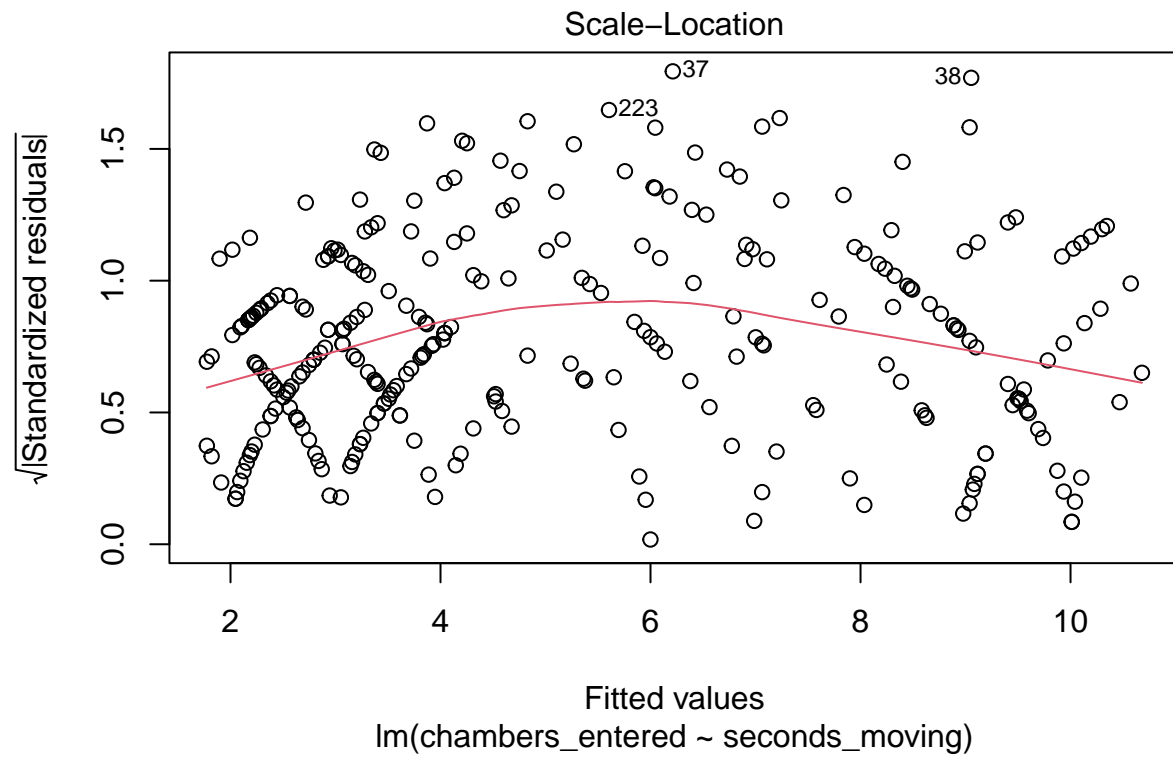
```
exp_BF.model.BIC
```

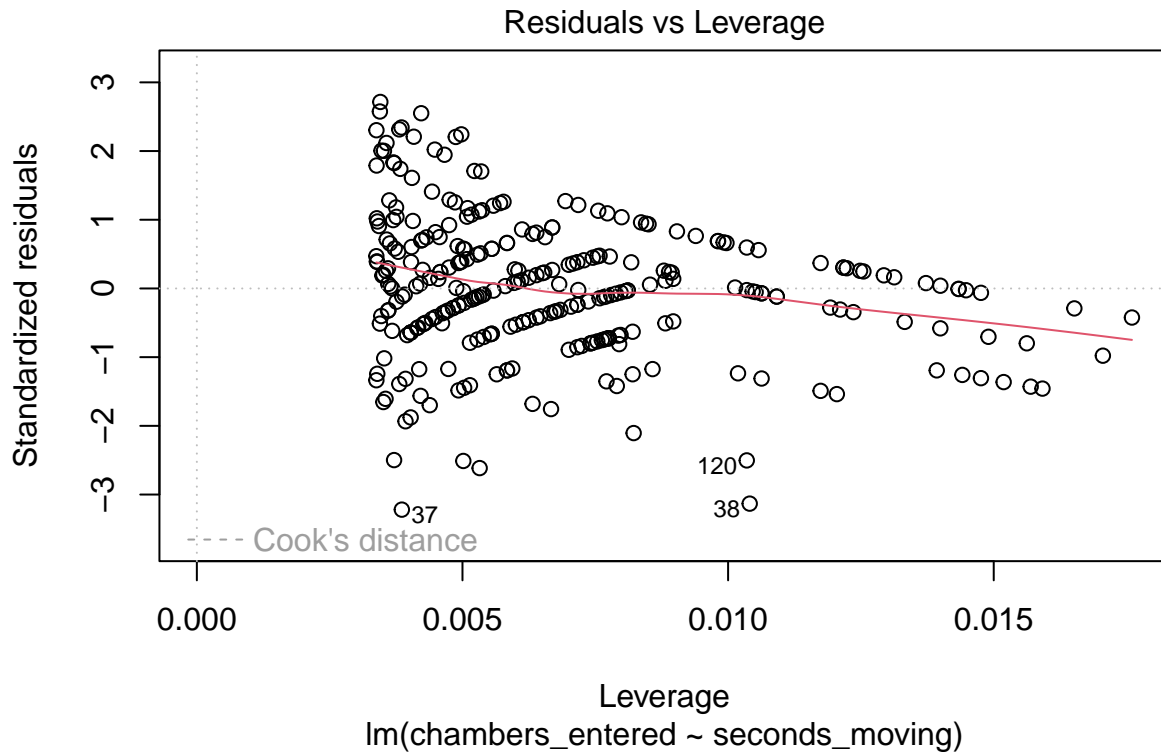
```
##
## Call:
## lm(formula = chambers_entered ~ seconds_moving, data = wasp)
##
## Coefficients:
##      (Intercept)  seconds_moving
##           1.6986           0.0152
```

```
plot(exp_BF.model.BIC)
```









```
shapiro.test(resid(exp_BF.model.BIC))
```

```
##
## Shapiro-Wilk normality test
##
## data: resid(exp_BF.model.BIC)
## W = 0.98843, p-value = 0.01837
```

```
Group.exp = rep("Lower", nrow(wasp)) #Creates a vector that repeats "Lower" n times
Group.exp[wasp$chambers_entered < median(wasp$chambers_entered)] = "Upper" #Changing the appropriate va
Group.exp = as.factor(Group.exp) #Changes it to a factor, which R recognizes as a grouping variable.
fligner.test(exp_BF.model.BIC$residuals, Group.exp)
```

```
##
## Fligner-Killeen test of homogeneity of variances
##
## data: exp_BF.model.BIC$residuals and Group.exp
## Fligner-Killeen:med chi-squared = 9.9773, df = 1, p-value = 0.001585
```

```
##LMER MODEL NOT BETTER.
```

```
lmer_model_exp <- lmer(chambers_entered ~ seconds_moving + (1|Wasp.ID), data = wasp)
BIC(lmer_model_exp)
```

```
## [1] 1142.717
```

```
lm_model_act <- lm(seconds_moving ~ Dummy_Color + Trial + weight + bodily_contact_time + Log_aggression  
lm_model_act
```

```
##  
## Call:  
## lm(formula = seconds_moving ~ Dummy_Color + Trial + weight +  
##     bodily_contact_time + Log_aggression + anntenation + chambers_entered,  
##     data = wasp)  
##  
## Coefficients:  
##             (Intercept) Dummy_Colorgolddsilver      Dummy_Colorgreen  
##             -73.74424      -3.63268              1.86132  
##             Dummy_Colorred      Dummy_Colorsilver      Dummy_Colorwhite  
##             -16.70804      -13.08140              30.69797  
##             Dummy_Coloryellow      Trial              weight  
##             5.65166              9.62748              201.88969  
##             bodily_contact_time      Log_aggression      anntenation  
##             -0.06383              10.59644              1.62328  
##             chambers_entered  
##             47.16311
```

```
lm_empty_act <- lm(seconds_moving ~ 1, data = wasp)
```

```
act_backward.model.BIC = stepAIC(lm_model_act, scope = list(lower = lm_empty_act, upper= lm_model_act)  
act_forward.model.BIC = stepAIC(lm_empty_act, scope = list(lower = lm_empty_act, upper= lm_model_act),
```

```
## Start:  AIC=3065.1  
## seconds_moving ~ 1  
##  
##           Df Sum of Sq    RSS    AIC  
## + chambers_entered      1  6672074 2449921 2681.7  
## + bodily_contact_time    1   261884 8860110 3062.2  
## + Trial                   1   222460 8899534 3063.5  
## <none>                      9121994 3065.1  
## + Log_aggression         1   171337 8950657 3065.2  
## + anntenation            1    9303 9112691 3070.5  
## + weight                 1    1092 9120902 3070.8  
## + Dummy_Color            6   120978 9001017 3095.3  
##  
## Step:  AIC=2681.66  
## seconds_moving ~ chambers_entered  
##  
##           Df Sum of Sq    RSS    AIC  
## <none>                      2449921 2681.7  
## + bodily_contact_time      1   27577.3 2422343 2684.0  
## + anntenation              1   23741.8 2426179 2684.5  
## + Log_aggression           1   19342.8 2430578 2685.0  
## + Trial                    1   16581.3 2433339 2685.3  
## + weight                   1    6296.7 2443624 2686.6  
## + Dummy_Color              6   26088.0 2423833 2712.6
```



```
act_FB.model.BIC = stepAIC(lm_empty_act, scope = list(lower = lm_empty_act, upper= lm_model_act), k = 1)
act_BF.model.BIC = stepAIC(lm_model_act, scope = list(lower = lm_empty_act, upper= lm_model_act), k = 1)
```

```
BIC(act_backward.model.BIC)
```

```
## [1] 3527.36
```

```
BIC(act_forward.model.BIC)
```

```
## [1] 3527.36
```

```
BIC(act_BF.model.BIC)
```

```
## [1] 3527.36
```

```
BIC(act_FB.model.BIC)
```

```
## [1] 3527.36
```

```
act_BF.model.BIC
```

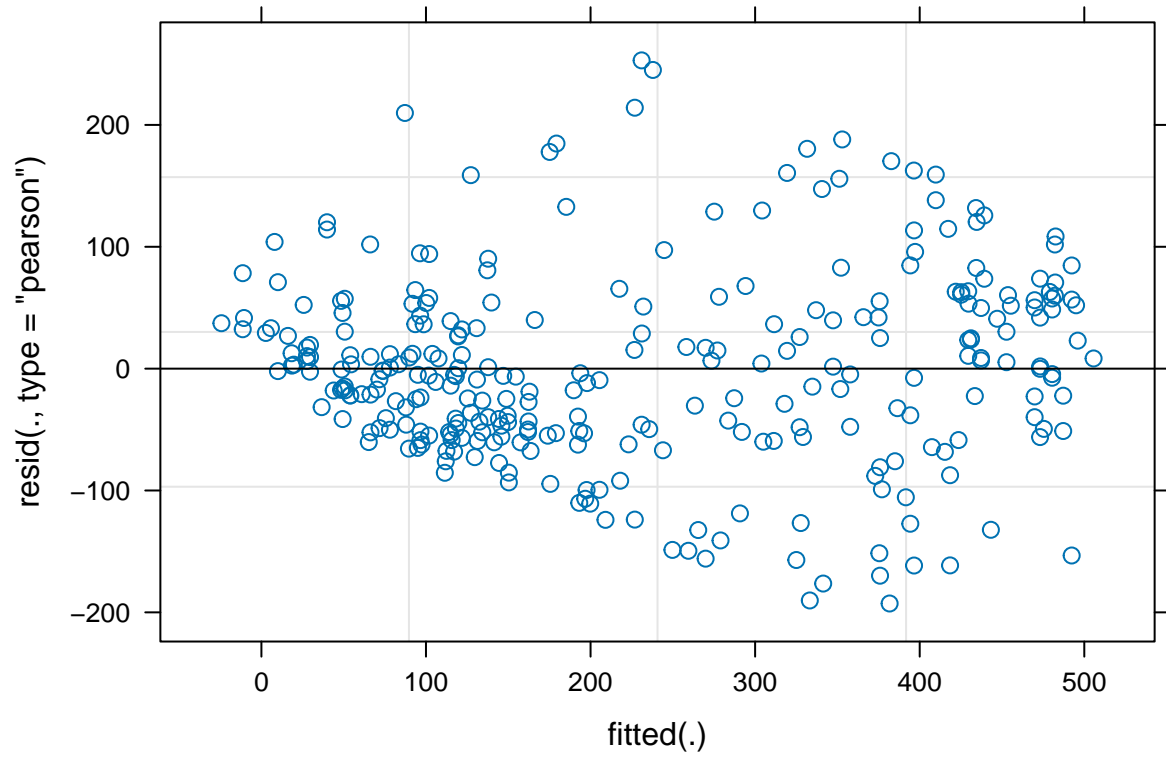
```
##
## Call:
## lm(formula = seconds_moving ~ chambers_entered, data = wasp)
##
## Coefficients:
##      (Intercept)  chambers_entered
##           -19.77           48.13
```

```
lmer_model_act <- lmer(seconds_moving ~ chambers_entered + (1|Wasp.ID), data = wasp)
```

```
BIC(lmer_model_act)
```

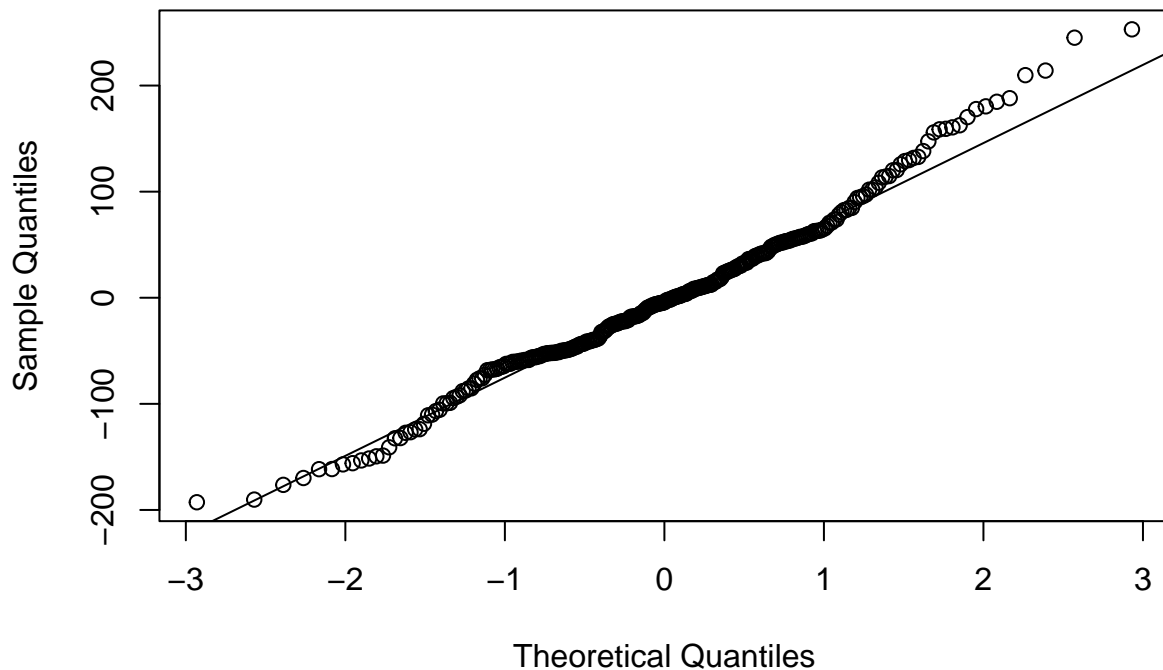
```
## [1] 3513.757
```

```
plot(lmer_model_act)
```



```
qqnorm(resid(lmer_model_act))  
qqline(resid(lmer_model_act))
```

Normal Q-Q Plot



```
shapiro.test(resid(act_BF.model.BIC))
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  resid(act_BF.model.BIC)  
## W = 0.98133, p-value = 0.0006617
```

```
#install.packages("MPV")  
#install.packages("leaps")  
library(MPV)  
library(leaps)  
library(MASS)  
#install.packages("lmerTest")
```

```
#lmerTest::step(lmer_model_affilitation_full)
```

```
lmer_model_affilitation
```

```
## Linear mixed model fit by REML ['lmerMod']  
## Formula: bodily_contact_time ~ Trial + seconds_moving + (1 | Wasp.ID)  
## Data: wasp  
## REML criterion at convergence: 3870.311  
## Random effects:
```

```
## Groups   Name          Std.Dev.
## Wasp.ID   (Intercept) 104.3
## Residual              149.0
## Number of obs: 296, groups: Wasp.ID, 74
## Fixed Effects:
##      (Intercept)          Trial   seconds_moving
##      148.0218          24.6667          -0.1134
```

```
lmer_model_ant
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: anntenation ~ Log_aggression + (1 | Wasp.ID)
## Data: wasp
## REML criterion at convergence: 1532.576
## Random effects:
## Groups   Name          Std.Dev.
## Wasp.ID   (Intercept) 2.073
## Residual              2.785
## Number of obs: 296, groups: Wasp.ID, 74
## Fixed Effects:
##      (Intercept) Log_aggression
##      5.1067          0.7655
```

```
lmer_model_act
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: seconds_moving ~ chambers_entered + (1 | Wasp.ID)
## Data: wasp
## REML criterion at convergence: 3490.996
## Random effects:
## Groups   Name          Std.Dev.
## Wasp.ID   (Intercept) 38.11
## Residual              83.06
## Number of obs: 296, groups: Wasp.ID, 74
## Fixed Effects:
##      (Intercept) chambers_entered
##      -18.81          47.95
```

```
agg_BF.model.BIC
```

```
##
## Call:
## lm(formula = Log_aggression ~ Dummy_Color + anntenation, data = wasp)
##
## Coefficients:
##      (Intercept) Dummy_Colorgolddsilver Dummy_Colorgreen
##      0.33147          0.53597          -0.02882
##      Dummy_Colorred Dummy_Colorsilver Dummy_Colorwhite
##      -0.02245          0.59191          -0.09271
##      Dummy_Coloryellow anntenation
##      -0.15931          0.03224
```

```
exp_BF.model.BIC
```

```
##  
## Call:  
## lm(formula = chambers_entered ~ seconds_moving, data = wasp)  
##  
## Coefficients:  
##      (Intercept)  seconds_moving  
##          1.6986          0.0152
```