Haussmann GLM

Ken Field

Note that I loaded quite a few new packages. These packages, like `performance` and `interactions` are useful for advanced model selection and might be worth exploring on your own. The `jtools` package provides plot_summs(), a better way to vizulize the coefficients of a model. The `skimr` package is a great tool for peeking at your data during exploration. The package `MASS` is essential for `glm.nb()`. The package `foreign` is needed to read in the .dta file.

Background

TRF is dependent. All the rest are independent factors.

Data Loading and Wrangling

Load in the data and then convert the appropriate columns to factors.

```
original.dat <- read_excel("MixedEffects KF 2023.03.01.xlsx")
dat <- original.dat %>%
  mutate(Family = as.factor(Family),
         Pen = as.factor(Pen),
         `Wing Band No.` = as.factor(`Wing Band No.`))
# Some different summary functions
summary(dat)
```

```
 Wing Band No. Family Pen         CORT          TRF age          TRF
 50     :  4   A:24   1:24   Min.   :0.040   Min.   :1.00   Min.   :17.1
 88     :  4   B:24   2:24   1st Qu.:1.113   1st Qu.:1.75   1st Qu.:18.6
 90     :  4   C:24   3:24   Median :2.015   Median :2.50   Median :19.1
 91     :  4   D:24   4:24   Mean   :2.502   Mean   :2.50   Mean   :19.1
 92     :  4   E:24   5:24   3rd Qu.:3.183   3rd Qu.:3.25   3rd Qu.:19.7
 94     :  4   F:24   6:24   Max.   :7.700   Max.   :4.00   Max.   :20.9
```

```
 (Other):120                 NA's   :78                           NA's   :42
```

```r
glimpse(dat)
```

```
Rows: 144
Columns: 6
$ `Wing Band No.` <fct> 395, 412, 422, 429, 431, 433, 424, 444, 410, 413, 94, ~
$ Family          <fct> A, B, C, D, E, F, A, B, C, D, E, F, A, B, C, D, E, F, ~
$ Pen             <fct> 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, ~
$ CORT            <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
$ `TRF age`       <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ~
$ TRF             <dbl> 19.5, 19.7, 20.1, 18.6, 19.9, 20.3, 18.8, 20.5, 19.4, ~
```

```r
skim(dat)
```

Table 1: Data summary

Name	dat
Number of rows	144
Number of columns	6
Column type frequency:	
factor	3
numeric	3
Group variables	None

Variable type: factor

skim_variable	n_missing	complete_rate	ordered	n_unique	top_counts
Wing Band No.	0	1	FALSE	36	50: 4, 88: 4, 90: 4, 91: 4
Family	0	1	FALSE	6	A: 24, B: 24, C: 24, D: 24
Pen	0	1	FALSE	6	1: 24, 2: 24, 3: 24, 4: 24

Variable type: numeric

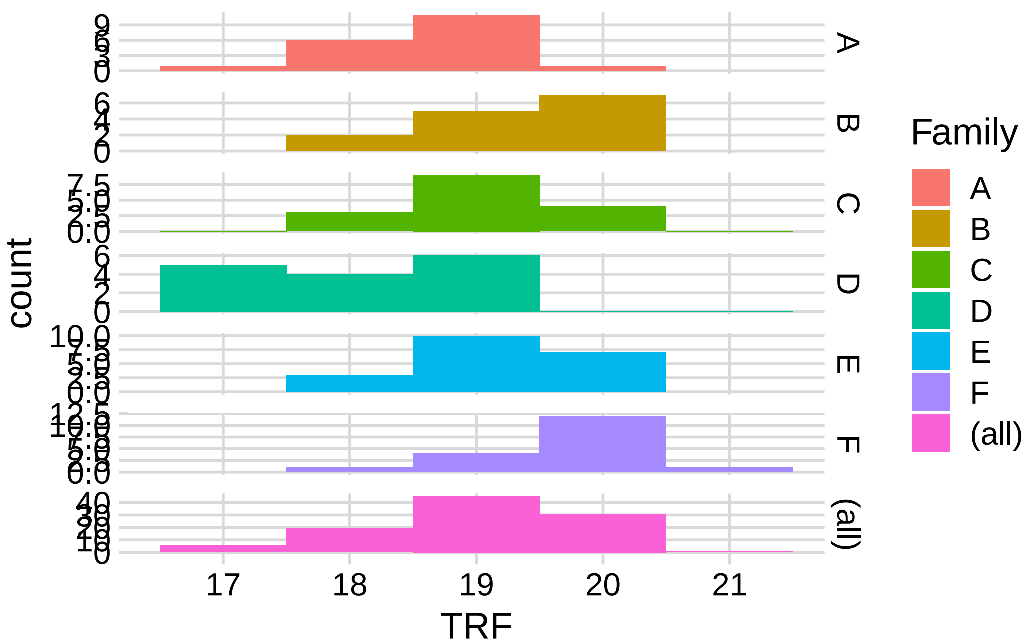
skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
CORT	78	0.46	2.5	1.94	0.04	1.11	2.02	3.18	7.7	
TRF age	0	1.00	2.5	1.12	1.00	1.75	2.50	3.25	4.0	
TRF	42	0.71	19.1	0.85	17.10	18.60	19.10	19.70	20.9	

Data Exploration

Generate histograms.

```
ggplot(dat, aes(TRF, fill = Family)) +
  geom_histogram(binwidth = 1) +
  facet_grid(Family ~ ., margins = TRUE, scales = "free") +
  theme_minimal_grid()
```

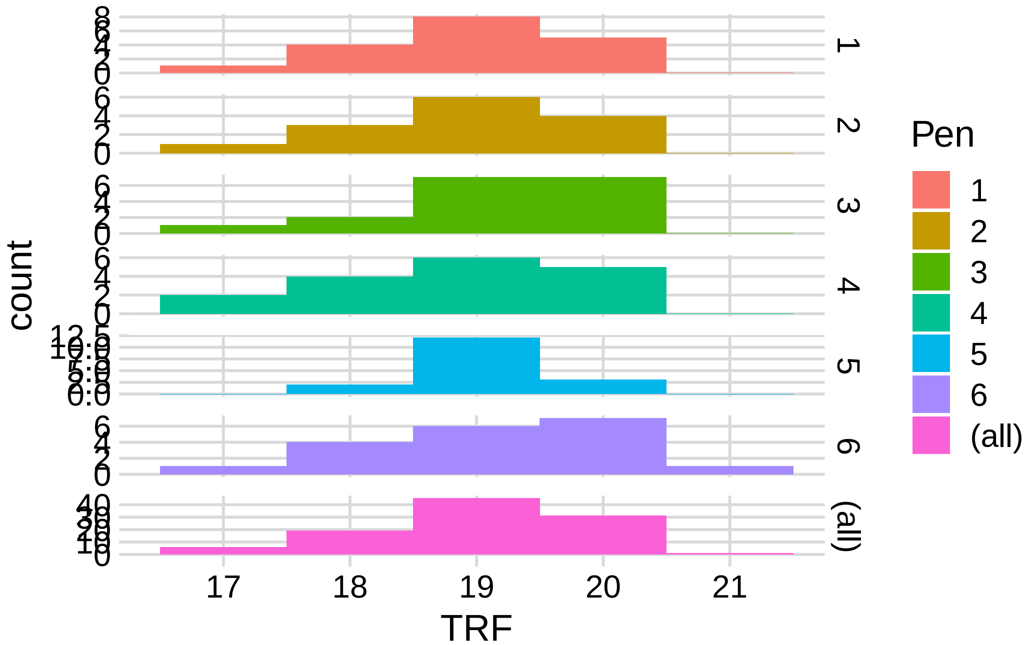
```
Warning: Removed 84 rows containing non-finite values (`stat_bin()`).
```




```
ggplot(dat, aes(TRF, fill = Pen)) +
  geom_histogram(binwidth = 1) +
  facet_grid(Pen ~ ., margins = TRUE, scales = "free") +
```

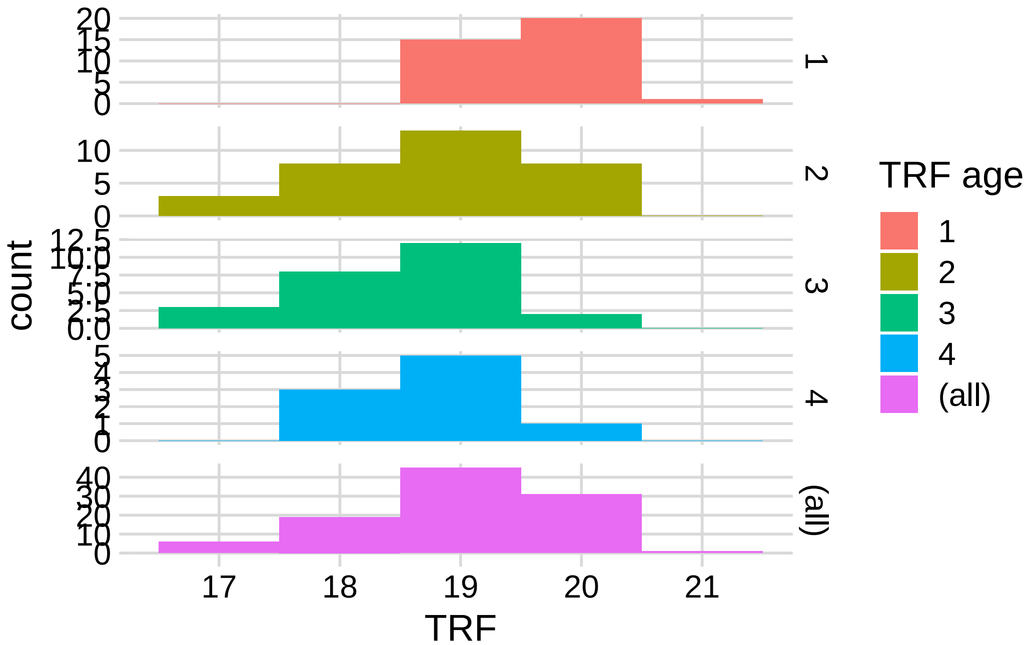
```
  theme_minimal_grid()
```

```
Warning: Removed 84 rows containing non-finite values (`stat_bin()`).
```




```
ggplot(dat, aes(TRF, fill = `TRF age`)) +
  geom_histogram(binwidth = 1) +
  facet_grid(`TRF age` ~ ., margins = TRUE, scales = "free") +
  theme_minimal_grid()
```

```
Warning: Removed 84 rows containing non-finite values (`stat_bin()`).
```



Possible statistical models

Below is a list of some analysis methods you may have encountered. Some of the methods listed are quite reasonable, while others have either fallen out of favor or have limitations.

- Negative binomial regression -Negative binomial regression can be used for over-dispersed count data, that is when the conditional variance exceeds the conditional mean. It can be considered as a generalization of Poisson regression since it has the same mean structure as Poisson regression and it has an extra parameter to model the over-dispersion. If the conditional distribution of the outcome variable is over-dispersed, the confidence intervals for the Negative binomial regression are likely to be wider as compared to those from a Poisson regression model.
- Poisson regression – Poisson regression is often used for modeling count data. Poisson regression has a number of extensions useful for count models.
- Zero-inflated regression model – Zero-inflated models attempt to account for excess zeros. In other words, two kinds of zeros are thought to exist in the data, "true zeros" and "excess zeros". Zero-inflated models estimate two equations simultaneously, one for the count model and one for the excess zeros.
- OLS regression – Count outcome variables are sometimes log-transformed and analyzed using OLS regression. Many issues arise with this approach, including loss of data due to undefined values generated by taking the log of zero (which is undefined), as well as the lack of capacity to model the dispersion.

GLM

Make a glm and check its summary() and performance()

```
m1 <- glm(TRF ~ ., data = dat)
summary(m1)
```

```

Call:
glm(formula = TRF ~ ., data = dat)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.47000  -0.11177   0.00000   0.09014   0.47000

Coefficients: (10 not defined because of singularities)
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)        20.42042    0.32280  63.260  < 2e-16 ***
`Wing Band No.`88   1.03986    0.26801   3.880 0.000758 ***
`Wing Band No.`92  -0.33889    0.36085  -0.939 0.357415
`Wing Band No.`95  -1.23136    0.27716  -4.443 0.000187 ***
`Wing Band No.`99   0.22466    0.28828   0.779 0.443749
`Wing Band No.`276 -0.81391    0.28509  -2.855 0.008961 **
`Wing Band No.`277 -0.26178    0.26655  -0.982 0.336281
`Wing Band No.`281  0.70120    0.27108   2.587 0.016493 *
`Wing Band No.`290 -1.30201    0.37897  -3.436 0.002255 **
`Wing Band No.`327 -0.17365    0.27555  -0.630 0.534773
`Wing Band No.`354 -1.26126    0.36671  -3.439 0.002235 **
`Wing Band No.`395 -0.25553    0.27184  -0.940 0.356989
`Wing Band No.`412 -0.26416    0.28516  -0.926 0.363875
`Wing Band No.`413 -0.91088    0.26792  -3.400 0.002460 **
`Wing Band No.`418 -0.09774    0.33984  -0.288 0.776223
`Wing Band No.`420 -1.59647    0.26975  -5.918 4.94e-06 ***
`Wing Band No.`422 -0.24607    0.27065  -0.909 0.372693
`Wing Band No.`424 -1.20541    0.29763  -4.050 0.000497 ***
`Wing Band No.`429 -1.85709    0.34897  -5.322 2.11e-05 ***
`Wing Band No.`431  0.75020    0.24942   3.008 0.006274 **
`Wing Band No.`433 -0.83740    0.36048  -2.323 0.029379 *
`Wing Band No.`438 -2.28213    0.37275  -6.122 3.03e-06 ***
`Wing Band No.`439 -0.44248    0.25646  -1.725 0.097880 .
`Wing Band No.`440 -1.63144    0.35904  -4.544 0.000145 ***
`Wing Band No.`441 -0.01148    0.27322  -0.042 0.966858
```

```
`Wing Band No.`445  0.17997    0.26466   0.680 0.503289
`Wing Band No.`492  0.51049    0.29252   1.745 0.094305 .
`Wing Band No.`494 -0.40944    0.28390  -1.442 0.162724
`Wing Band No.`499  0.43808    0.28455   1.540 0.137324
FamilyB                  NA         NA      NA       NA
FamilyC                  NA         NA      NA       NA
FamilyD                  NA         NA      NA       NA
FamilyE                  NA         NA      NA       NA
FamilyF                  NA         NA      NA       NA
Pen2                     NA         NA      NA       NA
Pen3                     NA         NA      NA       NA
Pen4                     NA         NA      NA       NA
Pen5                     NA         NA      NA       NA
Pen6                     NA         NA      NA       NA
CORT               -0.04970    0.03614  -1.375 0.182333
`TRF age`          -0.50821    0.07083  -7.175 2.63e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.08255374)

    Null deviance: 33.1683  on 53  degrees of freedom
Residual deviance:  1.8987  on 23  degrees of freedom
  (90 observations deleted due to missingness)
AIC: 36.464

Number of Fisher Scoring iterations: 2
```

```r
performance(m1)
```

```
Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type == :
prediction from a rank-deficient fit may be misleading

# Indices of model performance

AIC    |    AICc |     BIC |    R2 |  RMSE | Sigma
--------------------------------------------------
36.464 | 137.036 | 100.112 | 0.943 | 0.188 | 0.287
```

Another approach would be to use model selection to confirm the minimal model. The stepAIC() function will take a model and iteratively perform model selection in either the backward or forward direction.

```r
stepAIC(m1, direction = "backward")
```

```
Start:  AIC=36.46
TRF ~ `Wing Band No.` + Family + Pen + CORT + `TRF age`


Step:  AIC=36.46
TRF ~ `Wing Band No.` + Family + CORT + `TRF age`


Step:  AIC=36.46
TRF ~ `Wing Band No.` + CORT + `TRF age`

                  Df Deviance     AIC
<none>                 1.899  36.464
- CORT             1    2.055  38.731
- `TRF age`        1    6.148  97.914
- `Wing Band No.` 28   31.646 132.389


Call:  glm(formula = TRF ~ `Wing Band No.` + CORT + `TRF age`, data = dat)

Coefficients:
       (Intercept)   `Wing Band No.`88   `Wing Band No.`92   `Wing Band No.`95
          20.42042             1.03986            -0.33889            -1.23136
 `Wing Band No.`99  `Wing Band No.`276  `Wing Band No.`277  `Wing Band No.`281
           0.22466            -0.81391            -0.26178             0.70120
`Wing Band No.`290  `Wing Band No.`327  `Wing Band No.`354  `Wing Band No.`395
          -1.30201            -0.17365            -1.26126            -0.25553
`Wing Band No.`412  `Wing Band No.`413  `Wing Band No.`418  `Wing Band No.`420
          -0.26416            -0.91088            -0.09774            -1.59647
`Wing Band No.`422  `Wing Band No.`424  `Wing Band No.`429  `Wing Band No.`431
          -0.24607            -1.20541            -1.85709             0.75020
`Wing Band No.`433  `Wing Band No.`438  `Wing Band No.`439  `Wing Band No.`440
          -0.83740            -2.28213            -0.44248            -1.63144
`Wing Band No.`441  `Wing Band No.`445  `Wing Band No.`492  `Wing Band No.`494
          -0.01148             0.17997             0.51049            -0.40944
`Wing Band No.`499                CORT           `TRF age`
           0.43808            -0.04970            -0.50821

Degrees of Freedom: 53 Total (i.e. Null);  23 Residual
```

```
  (90 observations deleted due to missingness)
Null Deviance:      33.17
Residual Deviance: 1.899    AIC: 36.46
```

Selected model: TRF ~ `Wing Band No.` + CORT + `TRF age`

```r
m2 <- glm(TRF ~ `Wing Band No.` + CORT + `TRF age`, data = dat)
summary(m2)
```

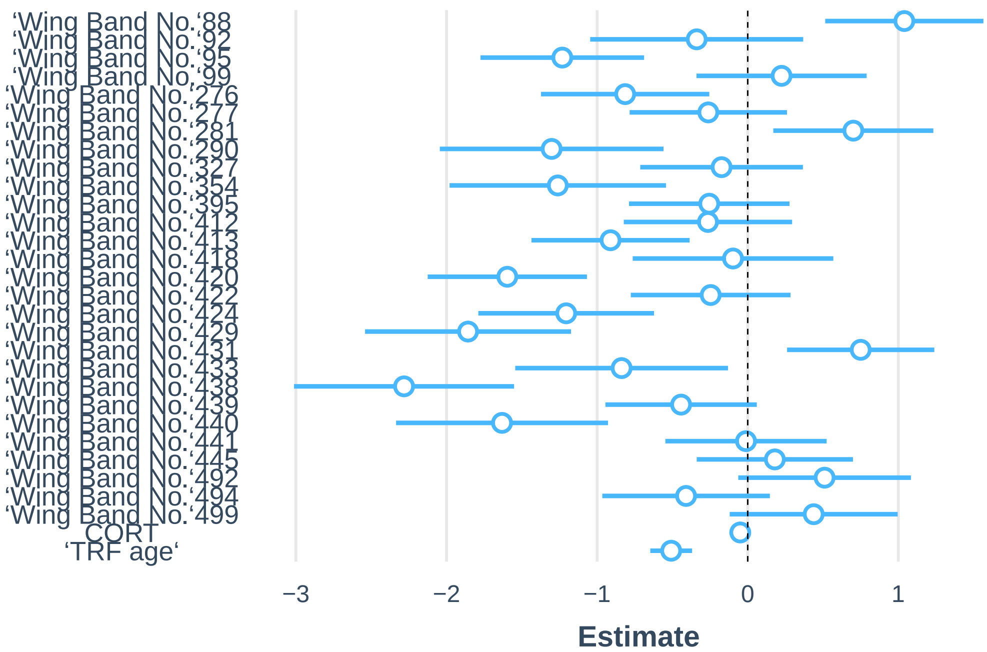
```

Call:
glm(formula = TRF ~ `Wing Band No.` + CORT + `TRF age`, data = dat)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.47000  -0.11177   0.00000   0.09014   0.47000

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)        20.42042    0.32280  63.260  < 2e-16 ***
`Wing Band No.`88   1.03986    0.26801   3.880 0.000758 ***
`Wing Band No.`92  -0.33889    0.36085  -0.939 0.357415
`Wing Band No.`95  -1.23136    0.27716  -4.443 0.000187 ***
`Wing Band No.`99   0.22466    0.28828   0.779 0.443749
`Wing Band No.`276 -0.81391    0.28509  -2.855 0.008961 **
`Wing Band No.`277 -0.26178    0.26655  -0.982 0.336281
`Wing Band No.`281  0.70120    0.27108   2.587 0.016493 *
`Wing Band No.`290 -1.30201    0.37897  -3.436 0.002255 **
`Wing Band No.`327 -0.17365    0.27555  -0.630 0.534773
`Wing Band No.`354 -1.26126    0.36671  -3.439 0.002235 **
`Wing Band No.`395 -0.25553    0.27184  -0.940 0.356989
`Wing Band No.`412 -0.26416    0.28516  -0.926 0.363875
`Wing Band No.`413 -0.91088    0.26792  -3.400 0.002460 **
`Wing Band No.`418 -0.09774    0.33984  -0.288 0.776223
`Wing Band No.`420 -1.59647    0.26975  -5.918 4.94e-06 ***
`Wing Band No.`422 -0.24607    0.27065  -0.909 0.372693
`Wing Band No.`424 -1.20541    0.29763  -4.050 0.000497 ***
`Wing Band No.`429 -1.85709    0.34897  -5.322 2.11e-05 ***
`Wing Band No.`431  0.75020    0.24942   3.008 0.006274 **
`Wing Band No.`433 -0.83740    0.36048  -2.323 0.029379 *
`Wing Band No.`438 -2.28213    0.37275  -6.122 3.03e-06 ***
`Wing Band No.`439 -0.44248    0.25646  -1.725 0.097880 .
```

```
`Wing Band No.`440 -1.63144    0.35904  -4.544 0.000145 ***
`Wing Band No.`441 -0.01148    0.27322  -0.042 0.966858
`Wing Band No.`445  0.17997    0.26466   0.680 0.503289
`Wing Band No.`492  0.51049    0.29252   1.745 0.094305 .
`Wing Band No.`494 -0.40944    0.28390  -1.442 0.162724
`Wing Band No.`499  0.43808    0.28455   1.540 0.137324
CORT               -0.04970    0.03614  -1.375 0.182333
`TRF age`          -0.50821    0.07083  -7.175 2.63e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.08255374)

    Null deviance: 33.1683  on 53  degrees of freedom
Residual deviance:  1.8987  on 23  degrees of freedom
  (90 observations deleted due to missingness)
AIC: 36.464

Number of Fisher Scoring iterations: 2
```

```r
performance(m2)
```

```
# Indices of model performance

AIC    |    AICc |     BIC |    R2 |  RMSE | Sigma
--------------------------------------------------
36.464 | 137.036 | 100.112 | 0.943 | 0.188 | 0.287
```

```r
plot_summs(m2)
```




```r
anova(m1, m2)
```

```
Analysis of Deviance Table

Model 1: TRF ~ `Wing Band No.` + Family + Pen + CORT + `TRF age`
Model 2: TRF ~ `Wing Band No.` + CORT + `TRF age`
  Resid. Df Resid. Dev Df Deviance
1        23     1.8987
2        23     1.8987  0        0
```

```r
anova(m2, m1)
```

```
Analysis of Deviance Table

Model 1: TRF ~ `Wing Band No.` + CORT + `TRF age`
Model 2: TRF ~ `Wing Band No.` + Family + Pen + CORT + `TRF age`
  Resid. Df Resid. Dev Df Deviance
1        23     1.8987
2        23     1.8987  0        0
```

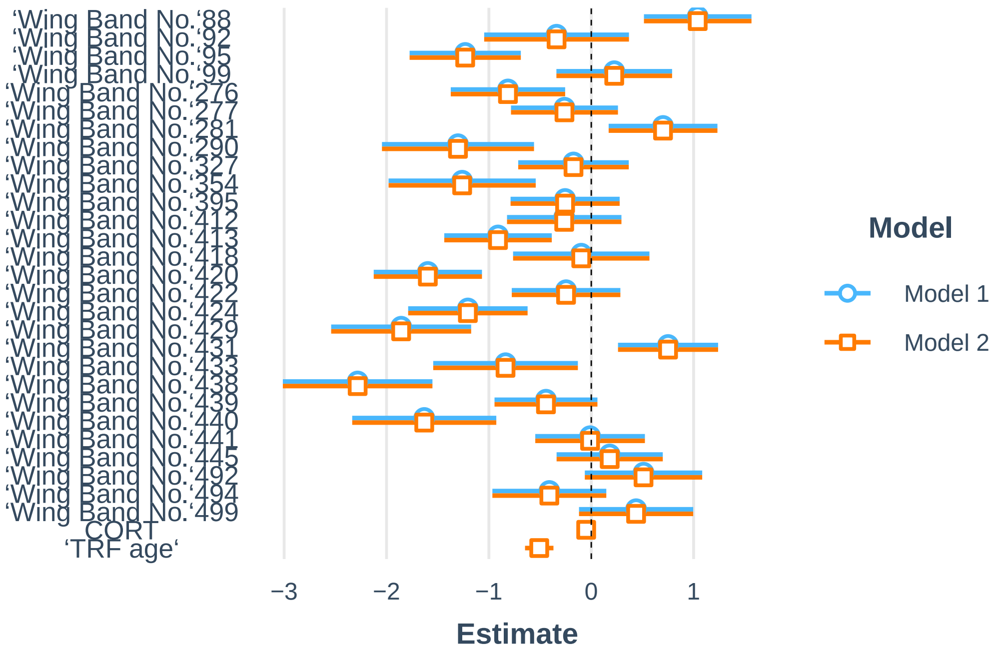
Those results indicate that there is essentially no difference in the models (presumably because Family and Pen are fully nested with Wing Band.

We can visualize the effect sizes and confidence intervals using the plot_summs() function from jtools.

```
plot_summs(m1, m2)
```

```
Warning in base$statistic[!is.na(base$statistic)] <- x$coeftable[, stat_col]:
number of items to replace is not a multiple of replacement length
```

```
Warning in base[["p.value"]][!is.na(base$statistic)] <- x$coeftable[, "p"]:
number of items to replace is not a multiple of replacement length
```



Check for interactions

```
m4 <- glm(TRF ~ `Wing Band No.` * CORT * `TRF age`, data = dat)
compare_performance(m1, m4)
```

```
Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type == :
prediction from a rank-deficient fit may be misleading

Warning: Model has zero degrees of freedom!

Warning in logLik.glm(x, ...): extra arguments discarded

Warning: Model has zero degrees of freedom!

Warning in logLik.glm(x, ...): extra arguments discarded

Warning: Model has zero degrees of freedom!

Warning in logLik.glm(x, ...): extra arguments discarded

Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type == :
prediction from a rank-deficient fit may be misleading

# Comparison of Model Performance Indices

Name | Model |   AIC (weights) |  AICc (weights) |   BIC (weights) |    R2 |      RMSE | Sigm
-----------------------------------------------------------------------------------------------
m1   |   glm |    36.5 (<.001) |   137.0 (<.001) |   100.1 (<.001) | 0.943 |     0.188 | 0.28
m4   |   glm | -3164.4 (>.999) | -6244.4 (>.999) | -3055.0 (>.999) | 1.000 | 1.647e-14 |   In
```

```r
stepAIC(m4, direction = "backward")
```

```
Start:  AIC=-3164.35
TRF ~ `Wing Band No.` * CORT * `TRF age`


Step:  AIC=-3164.35
TRF ~ `Wing Band No.` + CORT + `TRF age` + `Wing Band No.`:CORT +
    `Wing Band No.`:`TRF age` + CORT:`TRF age`


Step:  AIC=-3164.35
TRF ~ `Wing Band No.` + CORT + `TRF age` + `Wing Band No.`:CORT +
    `Wing Band No.`:`TRF age`
```

```
                            Df Deviance     AIC
<none>                         0.000000 -3164.4
- `Wing Band No.`:`TRF age`  3 0.011071  -201.3
- `Wing Band No.`:CORT       3 0.019322  -171.3


Call:  glm(formula = TRF ~ `Wing Band No.` + CORT + `TRF age` + `Wing Band No.`:CORT +
    `Wing Band No.`:`TRF age`, data = dat)

Coefficients:
                 (Intercept)              `Wing Band No.`88
                    19.24901                        1.73987
           `Wing Band No.`92              `Wing Band No.`95
                     0.16245                       -3.82831
           `Wing Band No.`99             `Wing Band No.`276
                     0.84060                       -0.58069
          `Wing Band No.`277             `Wing Band No.`281
                    -1.12734                       -0.77668
          `Wing Band No.`290             `Wing Band No.`327
                    -0.72082                        0.29126
          `Wing Band No.`354             `Wing Band No.`395
                    -0.73162                        0.80099
          `Wing Band No.`412             `Wing Band No.`413
                     3.09851                        0.50482
          `Wing Band No.`418             `Wing Band No.`420
                     0.22503                       -1.62700
          `Wing Band No.`422             `Wing Band No.`424
                     0.92478                        0.22471
          `Wing Band No.`429             `Wing Band No.`431
                    -1.71225                        0.65391
          `Wing Band No.`433             `Wing Band No.`438
                    -0.33794                       -1.72609
          `Wing Band No.`439             `Wing Band No.`440
                    -0.08357                       -1.13953
          `Wing Band No.`441             `Wing Band No.`445
                    -0.93943                        1.27495
          `Wing Band No.`492             `Wing Band No.`494
                     2.48363                       -0.20968
          `Wing Band No.`499                           CORT
                    -0.88501                        0.01318
                   `TRF age`         `Wing Band No.`88:CORT
```

```
                  -0.21436                      -0.10803
      `Wing Band No.`92:CORT        `Wing Band No.`95:CORT
                        NA                       1.20913
      `Wing Band No.`99:CORT       `Wing Band No.`276:CORT
                  -0.18232                       0.05479
     `Wing Band No.`277:CORT       `Wing Band No.`281:CORT
                   0.28121                       0.56695
     `Wing Band No.`290:CORT       `Wing Band No.`327:CORT
                        NA                      -0.07395
     `Wing Band No.`354:CORT       `Wing Band No.`395:CORT
                        NA                      -0.84651
     `Wing Band No.`412:CORT       `Wing Band No.`413:CORT
                  -1.75904                      -0.23122
     `Wing Band No.`418:CORT       `Wing Band No.`420:CORT
                        NA                       0.01367
     `Wing Band No.`422:CORT       `Wing Band No.`424:CORT
                  -0.30005                      -1.17275
     `Wing Band No.`429:CORT       `Wing Band No.`431:CORT
                        NA                       0.01929
     `Wing Band No.`433:CORT       `Wing Band No.`438:CORT
                        NA                            NA
     `Wing Band No.`439:CORT       `Wing Band No.`440:CORT
                  -0.03935                            NA
     `Wing Band No.`441:CORT       `Wing Band No.`445:CORT
                   0.42715                      -0.20519
     `Wing Band No.`492:CORT       `Wing Band No.`494:CORT
                  -1.32736                       0.06736
     `Wing Band No.`499:CORT   `Wing Band No.`88:`TRF age`
                   0.93140                            NA
 `Wing Band No.`92:`TRF age`   `Wing Band No.`95:`TRF age`
                        NA                            NA
 `Wing Band No.`99:`TRF age`  `Wing Band No.`276:`TRF age`
                        NA                            NA
`Wing Band No.`277:`TRF age`  `Wing Band No.`281:`TRF age`
                        NA                            NA
`Wing Band No.`290:`TRF age`  `Wing Band No.`327:`TRF age`
                        NA                            NA
`Wing Band No.`354:`TRF age`  `Wing Band No.`395:`TRF age`
                        NA                      -0.04397
`Wing Band No.`412:`TRF age`  `Wing Band No.`413:`TRF age`
                        NA                            NA
`Wing Band No.`418:`TRF age`  `Wing Band No.`420:`TRF age`
                        NA                            NA
```

```
`Wing Band No.`422:`TRF age`  `Wing Band No.`424:`TRF age`
                          NA                            NA
`Wing Band No.`429:`TRF age`  `Wing Band No.`431:`TRF age`
                          NA                       0.06663
`Wing Band No.`433:`TRF age`  `Wing Band No.`438:`TRF age`
                          NA                            NA
`Wing Band No.`439:`TRF age`  `Wing Band No.`440:`TRF age`
                    -0.03669                            NA
`Wing Band No.`441:`TRF age`  `Wing Band No.`445:`TRF age`
                          NA                            NA
`Wing Band No.`492:`TRF age`  `Wing Band No.`494:`TRF age`
                          NA                            NA
`Wing Band No.`499:`TRF age`
                          NA

Degrees of Freedom: 53 Total (i.e. Null);  0 Residual
  (90 observations deleted due to missingness)
Null Deviance:      33.17
Residual Deviance: 1.465e-26    AIC: -3164
```

Some of the interactions are kept, although many of the individual factors don't seem to have any statistical power (NA in the coefficient table)

TRF ~ `Wing Band No.` + CORT + `TRF age` + `Wing Band No.`:CORT + `Wing Band No.`:`TRF age`

```
m5 <- lm(TRF ~ `Wing Band No.` + CORT + `TRF age` + `Wing Band No.`:CORT + `Wing Band No.`
summary(m5)
```

```
Call:
lm(formula = TRF ~ `Wing Band No.` + CORT + `TRF age` + `Wing Band No.`:CORT +
    `Wing Band No.`:`TRF age`, data = dat)

Residuals:
ALL 54 residuals are 0: no residual degrees of freedom!

Coefficients: (33 not defined because of singularities)
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                 19.24901        NaN     NaN      NaN
`Wing Band No.`88            1.73987        NaN     NaN      NaN
`Wing Band No.`92            0.16245        NaN     NaN      NaN
```

```
`Wing Band No.`95           -3.82831        NaN     NaN      NaN
`Wing Band No.`99            0.84060        NaN     NaN      NaN
`Wing Band No.`276          -0.58069        NaN     NaN      NaN
`Wing Band No.`277          -1.12734        NaN     NaN      NaN
`Wing Band No.`281          -0.77668        NaN     NaN      NaN
`Wing Band No.`290          -0.72082        NaN     NaN      NaN
`Wing Band No.`327           0.29126        NaN     NaN      NaN
`Wing Band No.`354          -0.73162        NaN     NaN      NaN
`Wing Band No.`395           0.80099        NaN     NaN      NaN
`Wing Band No.`412           3.09851        NaN     NaN      NaN
`Wing Band No.`413           0.50482        NaN     NaN      NaN
`Wing Band No.`418           0.22503        NaN     NaN      NaN
`Wing Band No.`420          -1.62700        NaN     NaN      NaN
`Wing Band No.`422           0.92478        NaN     NaN      NaN
`Wing Band No.`424           0.22471        NaN     NaN      NaN
`Wing Band No.`429          -1.71225        NaN     NaN      NaN
`Wing Band No.`431           0.65391        NaN     NaN      NaN
`Wing Band No.`433          -0.33794        NaN     NaN      NaN
`Wing Band No.`438          -1.72609        NaN     NaN      NaN
`Wing Band No.`439          -0.08357        NaN     NaN      NaN
`Wing Band No.`440          -1.13953        NaN     NaN      NaN
`Wing Band No.`441          -0.93943        NaN     NaN      NaN
`Wing Band No.`445           1.27495        NaN     NaN      NaN
`Wing Band No.`492           2.48363        NaN     NaN      NaN
`Wing Band No.`494          -0.20968        NaN     NaN      NaN
`Wing Band No.`499          -0.88501        NaN     NaN      NaN
CORT                         0.01318        NaN     NaN      NaN
`TRF age`                   -0.21436        NaN     NaN      NaN
`Wing Band No.`88:CORT      -0.10803        NaN     NaN      NaN
`Wing Band No.`92:CORT            NA         NA      NA       NA
`Wing Band No.`95:CORT       1.20913        NaN     NaN      NaN
`Wing Band No.`99:CORT      -0.18232        NaN     NaN      NaN
`Wing Band No.`276:CORT      0.05479        NaN     NaN      NaN
`Wing Band No.`277:CORT      0.28121        NaN     NaN      NaN
`Wing Band No.`281:CORT      0.56695        NaN     NaN      NaN
`Wing Band No.`290:CORT           NA         NA      NA       NA
`Wing Band No.`327:CORT     -0.07395        NaN     NaN      NaN
`Wing Band No.`354:CORT           NA         NA      NA       NA
`Wing Band No.`395:CORT     -0.84651        NaN     NaN      NaN
`Wing Band No.`412:CORT     -1.75904        NaN     NaN      NaN
`Wing Band No.`413:CORT     -0.23122        NaN     NaN      NaN
`Wing Band No.`418:CORT           NA         NA      NA       NA
`Wing Band No.`420:CORT      0.01367        NaN     NaN      NaN
```

```
`Wing Band No.`422:CORT      -0.30005        NaN     NaN      NaN
`Wing Band No.`424:CORT      -1.17275        NaN     NaN      NaN
`Wing Band No.`429:CORT            NA         NA      NA       NA
`Wing Band No.`431:CORT       0.01929        NaN     NaN      NaN
`Wing Band No.`433:CORT            NA         NA      NA       NA
`Wing Band No.`438:CORT            NA         NA      NA       NA
`Wing Band No.`439:CORT      -0.03935        NaN     NaN      NaN
`Wing Band No.`440:CORT            NA         NA      NA       NA
`Wing Band No.`441:CORT       0.42715        NaN     NaN      NaN
`Wing Band No.`445:CORT      -0.20519        NaN     NaN      NaN
`Wing Band No.`492:CORT      -1.32736        NaN     NaN      NaN
`Wing Band No.`494:CORT       0.06736        NaN     NaN      NaN
`Wing Band No.`499:CORT       0.93140        NaN     NaN      NaN
`Wing Band No.`88:`TRF age`        NA         NA      NA       NA
`Wing Band No.`92:`TRF age`        NA         NA      NA       NA
`Wing Band No.`95:`TRF age`        NA         NA      NA       NA
`Wing Band No.`99:`TRF age`        NA         NA      NA       NA
`Wing Band No.`276:`TRF age`       NA         NA      NA       NA
`Wing Band No.`277:`TRF age`       NA         NA      NA       NA
`Wing Band No.`281:`TRF age`       NA         NA      NA       NA
`Wing Band No.`290:`TRF age`       NA         NA      NA       NA
`Wing Band No.`327:`TRF age`       NA         NA      NA       NA
`Wing Band No.`354:`TRF age`       NA         NA      NA       NA
`Wing Band No.`395:`TRF age` -0.04397        NaN     NaN      NaN
`Wing Band No.`412:`TRF age`       NA         NA      NA       NA
`Wing Band No.`413:`TRF age`       NA         NA      NA       NA
`Wing Band No.`418:`TRF age`       NA         NA      NA       NA
`Wing Band No.`420:`TRF age`       NA         NA      NA       NA
`Wing Band No.`422:`TRF age`       NA         NA      NA       NA
`Wing Band No.`424:`TRF age`       NA         NA      NA       NA
`Wing Band No.`429:`TRF age`       NA         NA      NA       NA
`Wing Band No.`431:`TRF age`  0.06663        NaN     NaN      NaN
`Wing Band No.`433:`TRF age`       NA         NA      NA       NA
`Wing Band No.`438:`TRF age`       NA         NA      NA       NA
`Wing Band No.`439:`TRF age` -0.03669        NaN     NaN      NaN
`Wing Band No.`440:`TRF age`       NA         NA      NA       NA
`Wing Band No.`441:`TRF age`       NA         NA      NA       NA
`Wing Band No.`445:`TRF age`       NA         NA      NA       NA
`Wing Band No.`492:`TRF age`       NA         NA      NA       NA
`Wing Band No.`494:`TRF age`       NA         NA      NA       NA
`Wing Band No.`499:`TRF age`       NA         NA      NA       NA

Residual standard error: NaN on 0 degrees of freedom
```

```
  (90 observations deleted due to missingness)
Multiple R-squared:      1, Adjusted R-squared:    NaN 
F-statistic:   NaN on 53 and 0 DF,  p-value: NA
```

There is not enough data in the dataset to test all these interactions.

I am going to use m2 moving forward, but it might be worth looking at the interactions further.

Check model assumptions

First use the plot function to see if this model seems to be a reasonable fit to the data. Let's use a better function than just plot(), try out the check_model() function of `performance`

```
# check_model(m2) Does not work: Error in data.frame(x = fitted_, y = res_) : arguments im
check_autocorrelation(m2)
```

```
OK: Residuals appear to be independent and not autocorrelated (p = 0.560).
```

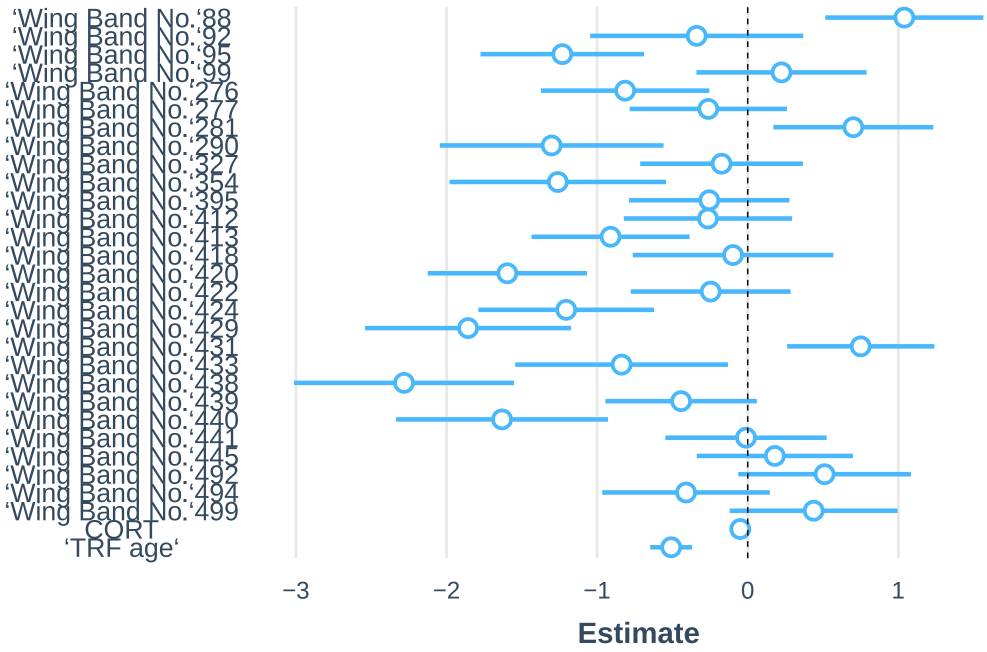
```
check_collinearity(m2)
```

```
# Check for Multicollinearity

Low Correlation

          Term  VIF   VIF 95% CI Increased SE Tolerance Tolerance 95% CI
 Wing Band No. 4.59 [3.82, 5.56]         2.14      0.22     [0.18, 0.26]
          CORT 3.53 [2.97, 4.26]         1.88      0.28     [0.23, 0.34]
       TRF age 1.30 [1.17, 1.53]         1.14      0.77     [0.65, 0.85]
```

```
check_outliers(m2)
```

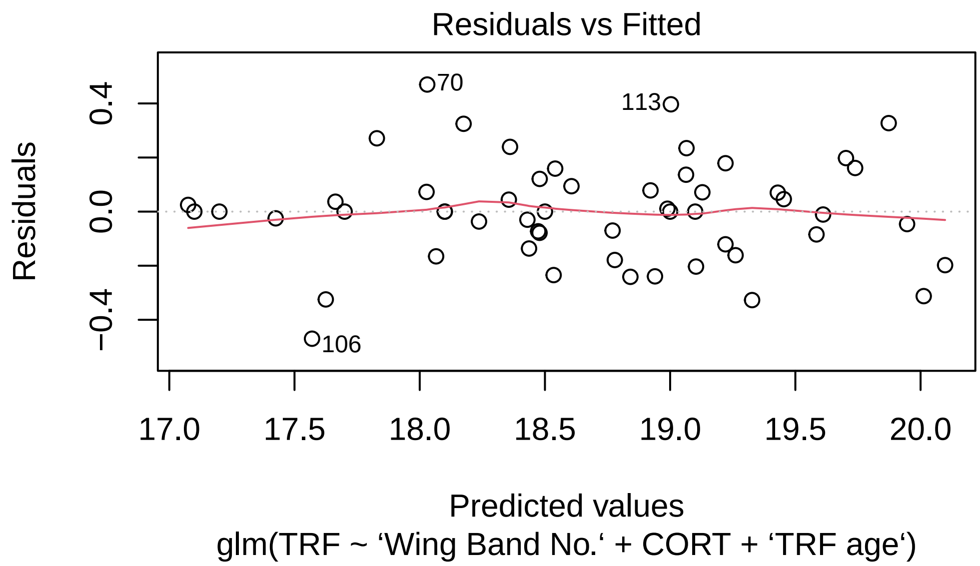
```
OK: No outliers detected.
- Based on the following method and threshold: cook (1.01).
- For variable: (Whole model)
```

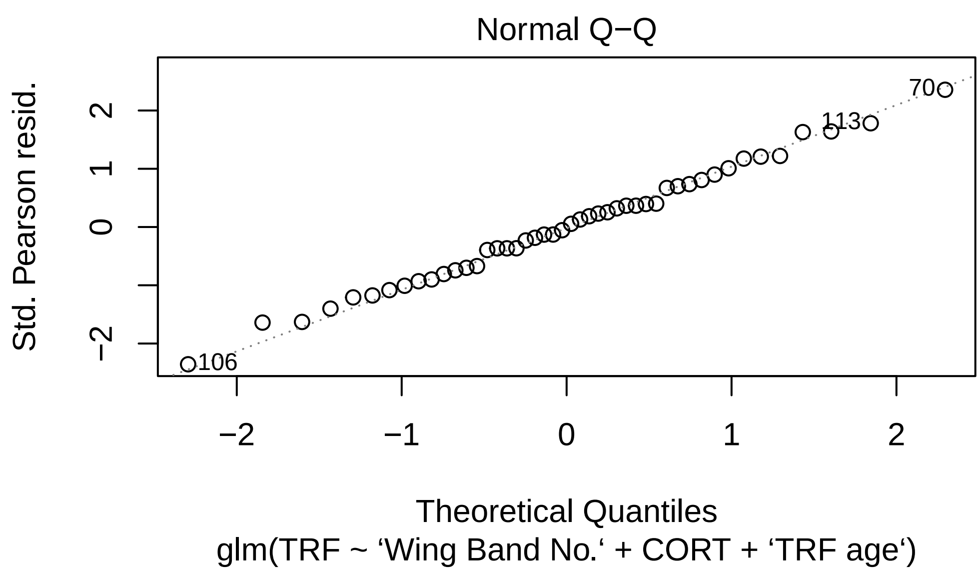
```r
plot_summs(m2)
```

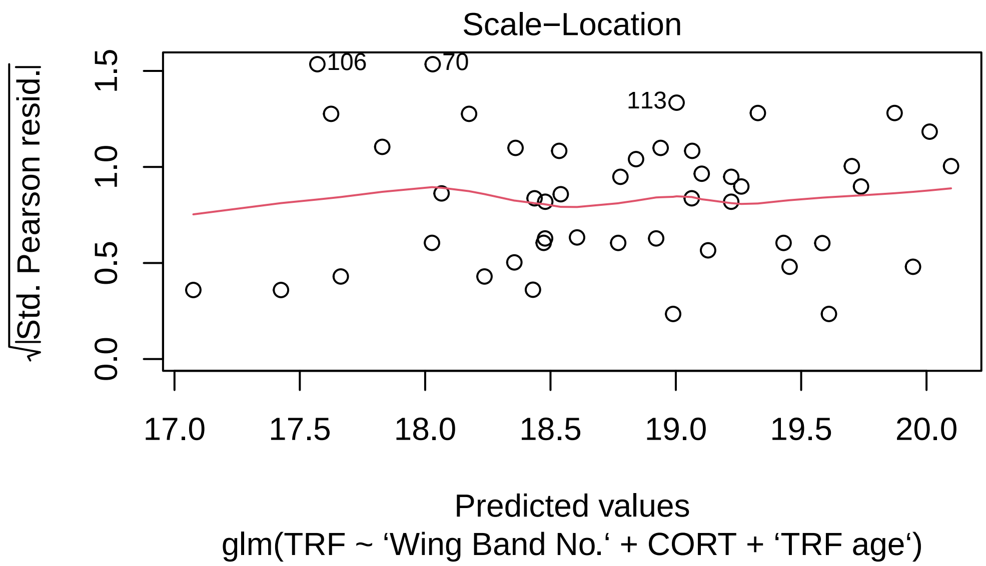


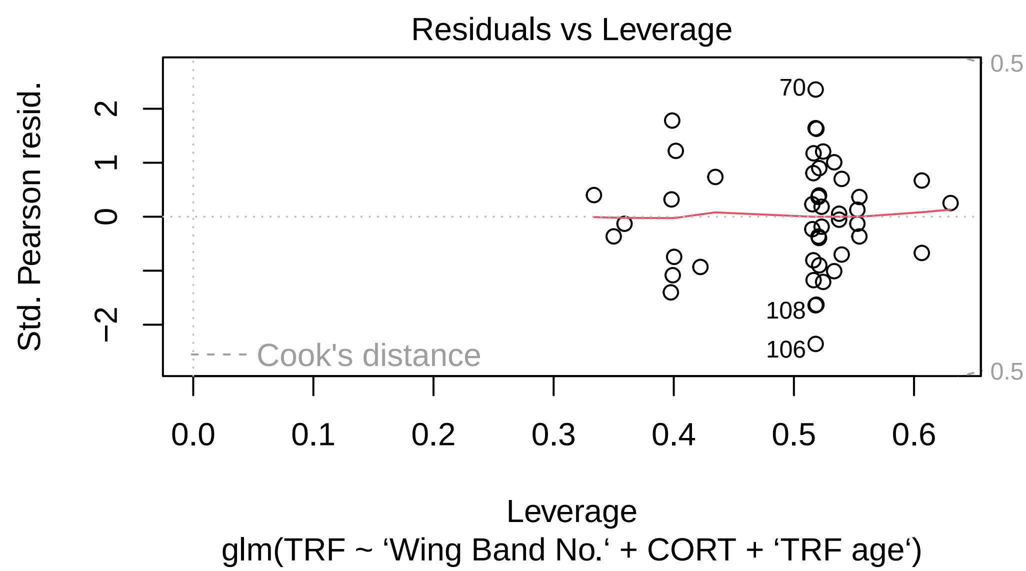

```r
plot(m2)
```

```
Warning: not plotting observations with leverage one:
  4, 6, 9, 11, 12, 16, 18, 21
```


Residuals vs Fitted
Residuals
0.4
0.0
−0.4
70
113
106
17.0
17.5
18.0
18.5
19.0
19.5
20.0
Predicted values
glm(TRF ~ ‘Wing Band No.‘ + CORT + ‘TRF age‘)


Normal Q−Q
Std. Pearson resid.
2
1
0
−2
70
113
106
−2
−1
0
1
2
Theoretical Quantiles
glm(TRF ~ ‘Wing Band No.‘ + CORT + ‘TRF age‘)



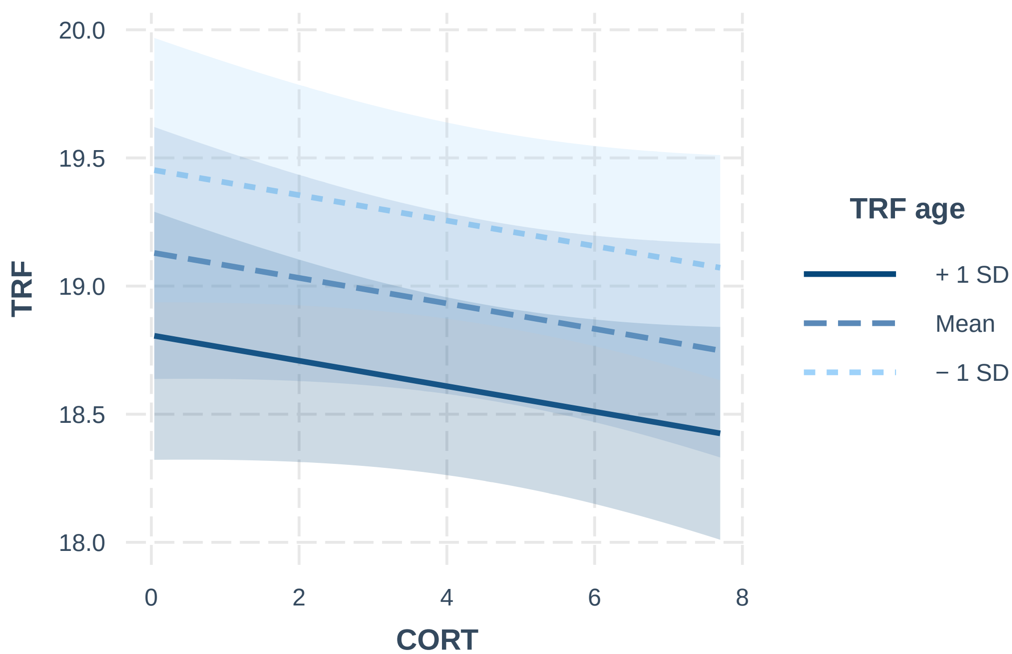


I'm not sure why check_model() isn't working. But the data look pretty good with plot(m2). Although note that 8 observations are being identified as having excessive leverage (but they are not identified as outliers using Cook's).

Interaction Plots

```
interact_plot(m1, pred = CORT,
              modx= `TRF age`,
              interval = TRUE)
```

```
Warning: CORT and TRF age are not included in an interaction with one another in the
model.

Warning: 1.90122043351102 is outside the observed range of TRF age

Warning in predict.lm(object, newdata, se.fit, scale = residual.scale, type =
if (type == : prediction from a rank-deficient fit may be misleading

Warning in predict.lm(object, newdata, se.fit, scale = residual.scale, type =
if (type == : prediction from a rank-deficient fit may be misleading

Warning in predict.lm(object, newdata, se.fit, scale = residual.scale, type =
if (type == : prediction from a rank-deficient fit may be misleading
```

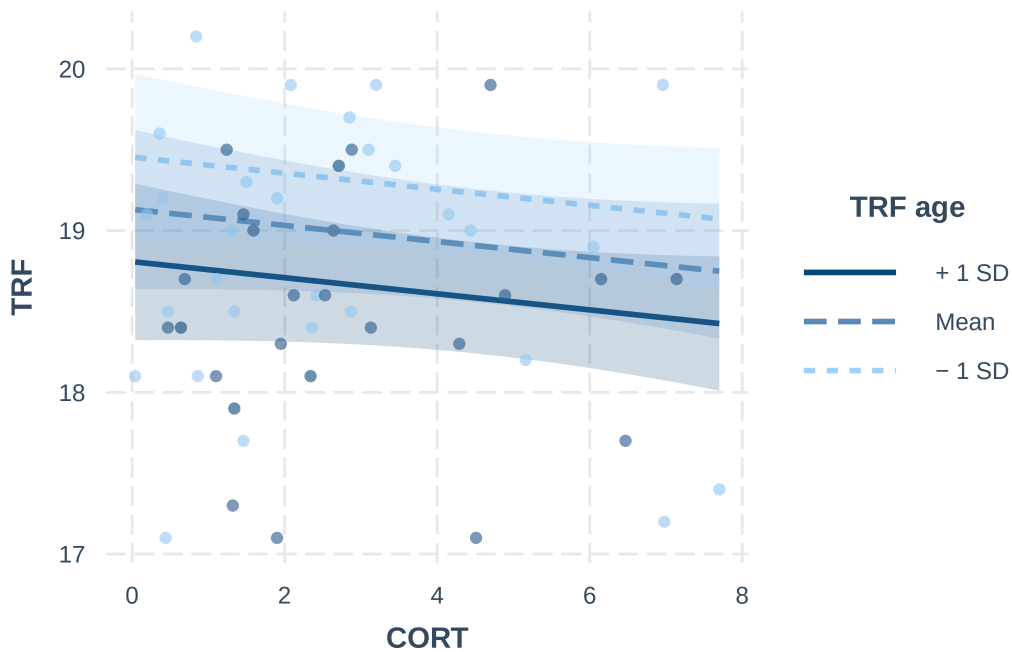


```
interact_plot(m1, pred = CORT,
              modx= `TRF age`,
              interval = TRUE, plot.points = TRUE)
```

```
Warning: CORT and TRF age are not included in an interaction with one another in the
model.
```

```
Warning: 1.90122043351102 is outside the observed range of TRF age
```

```
Warning in predict.lm(object, newdata, se.fit, scale = residual.scale, type =
if (type == : prediction from a rank-deficient fit may be misleading

Warning in predict.lm(object, newdata, se.fit, scale = residual.scale, type =
if (type == : prediction from a rank-deficient fit may be misleading

Warning in predict.lm(object, newdata, se.fit, scale = residual.scale, type =
if (type == : prediction from a rank-deficient fit may be misleading
```



Just for fun, I decided to try taking out the ID and using family and pen instead:

```r
m6 <- lm(TRF ~ Family + Pen + CORT + `TRF age`, data = dat)
summary(m6)
```

```
Call:
lm(formula = TRF ~ Family + Pen + CORT + `TRF age`, data = dat)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.09973 -0.27708  0.05947  0.33680  0.87430 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 19.26842    0.39508  48.771  < 2e-16 ***
FamilyB      0.49712    0.26609   1.868 0.068892 .  
FamilyC      0.41414    0.26945   1.537 0.131974    
FamilyD     -0.98278    0.25840  -3.803 0.000467 ***
FamilyE      0.40386    0.23846   1.694 0.097924 .  
FamilyF      0.90513    0.27366   3.308 0.001965 ** 
Pen2        -0.19425    0.33513  -0.580 0.565336    
Pen3        -0.13194    0.24668  -0.535 0.595632    
Pen4        -0.31270    0.26801  -1.167 0.250052    
Pen5        -0.03397    0.23688  -0.143 0.886680    
Pen6         0.11300    0.22273   0.507 0.614644    
CORT         0.03542    0.04072   0.870 0.389439    
`TRF age`   -0.31064    0.12069  -2.574 0.013770 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.539 on 41 degrees of freedom
  (90 observations deleted due to missingness)
Multiple R-squared:  0.6409,	Adjusted R-squared:  0.5357 
F-statistic: 6.097 on 12 and 41 DF,  p-value: 5.593e-06
```

Interesting. The effect size is a little smaller and the p value is much higher. ID definitely looks like a more important confounder than pen and family together.