

# Baboon Teeth Scott Scores

Gregory J. Matthews

## Intro

Baboon teeth

## Questions from Department Chair Krueger

Lastly, here are the questions: - Do wear scores increase with age? (Yes, duh, but we want to test it since we modified the method) - Do wear scores differ by sex? If so, how? (Is this by a specific quadrant? Is this by a side of tooth like buccal side or lingual side?) - Are there any patterns for rate of wear? That is, is there a standard wear through time (by month or year), or are some baboons wearing their teeth down faster than what we might expect?

## Exploratory Data Analysis

### Variable summaries

```
summary(scott)
```

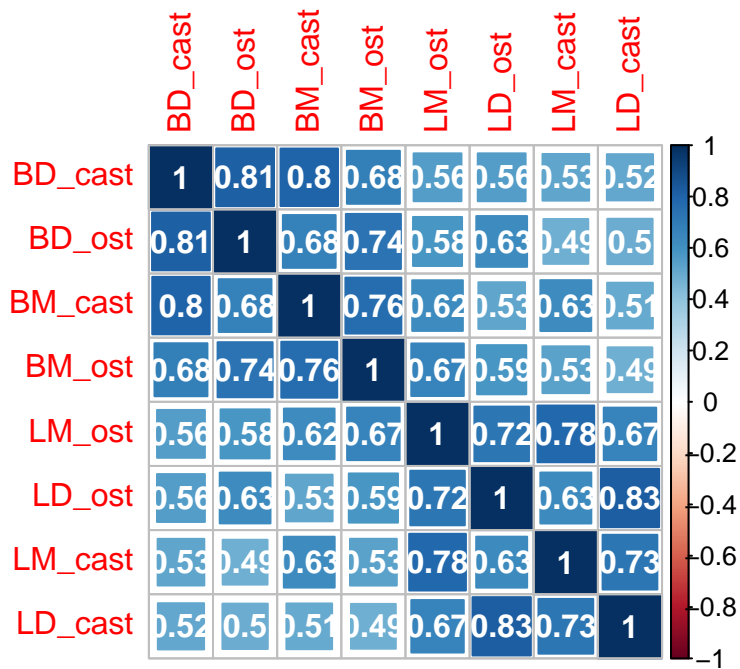
id	sex	age_cast	age_death
Min. : 3.0	Length:202	Min. : 85.0	Min. :102.0
1st Qu.:171.2	Class :character	1st Qu.:185.0	1st Qu.:215.2
Median :262.0	Mode :character	Median :232.5	Median :260.0
Mean :285.1		Mean :224.2	Mean :253.2
3rd Qu.:419.8		3rd Qu.:256.0	3rd Qu.:289.0
Max. :621.0		Max. :401.0	Max. :404.0
sfbrid	leftright	BM_cast	BD_cast
Length:202	Length:202	Min. :2.000	Min. :1.000

Class :character	Class :character	1st Qu.:5.000	1st Qu.:5.000
Mode :character	Mode :character	Median :6.000	Median :5.000
		Mean :5.856	Mean :5.317
		3rd Qu.:6.000	3rd Qu.:6.000
		Max. :9.000	Max. :9.000

LM_cast	LD_cast	BM_ost	BD_ost
Min. :1.000	Min. :1.000	Min. :4.000	Min. :1.000
1st Qu.:2.000	1st Qu.:2.000	1st Qu.:6.000	1st Qu.:5.000
Median :4.000	Median :4.000	Median :6.000	Median :6.000
Mean :4.094	Mean :3.614	Mean :6.505	Mean :6.064
3rd Qu.:5.000	3rd Qu.:5.000	3rd Qu.:7.000	3rd Qu.:7.000
Max. :9.000	Max. :9.000	Max. :9.000	Max. :9.000

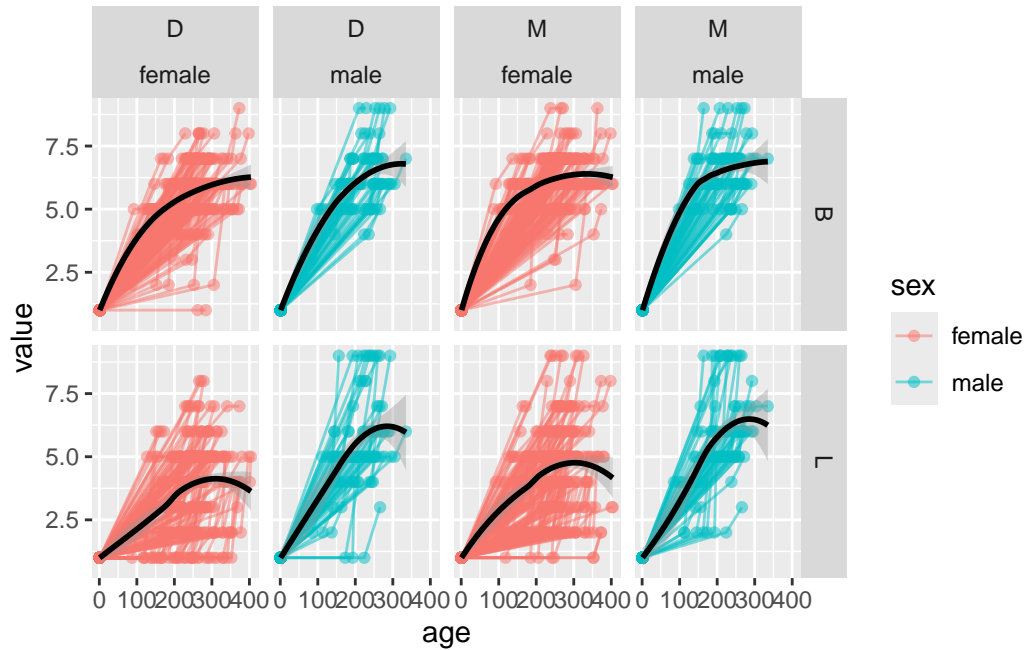
LM_ost	LD_ost
Min. :2.000	Min. :1.000
1st Qu.:4.000	1st Qu.:3.000
Median :5.000	Median :5.000
Mean :5.347	Mean :4.688
3rd Qu.:6.000	3rd Qu.:6.000
Max. :9.000	Max. :9.000

## Correlations



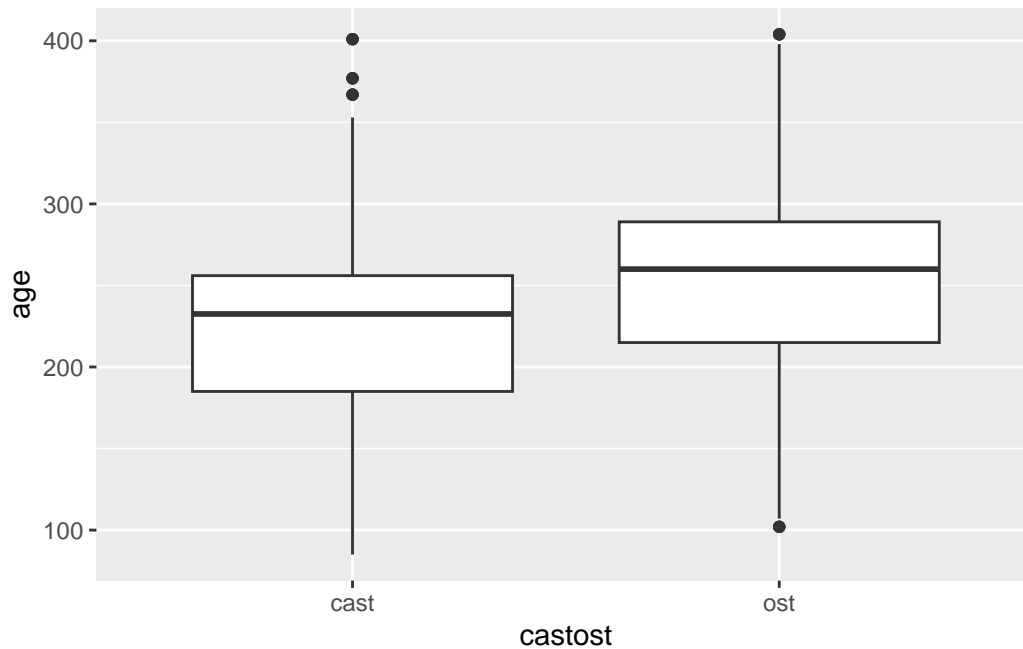
## Data Viz: Age vs wear by sex

``geom_smooth()`` using method = 'loess' and formula = 'y ~ x'



Age distributions by measurement time.

```
ggplot(aes(x = castost, y = age), data = scott_long %>% filter(age > 1)) + geom_boxplot() +
```



**Do wear scores increase with age? (Yes, duh, but we want to test it since we modified the method)**

The tests below test the following hypothesis:  $H_0 : \mu_{cast} = \mu_{ost}$  vs  $H_0 : \mu_{cast} < \mu_{ost}$  for each location of the tooth individually. The test used here is a two sample dependent t-test. In all cases, the null hypothesis is rejected with a p-value less than  $2.2 \times 10^{-16}$ . There is statistical evidence that the mean at ost is larger than the mean at cast.

```
t.test(scott$BM_cast, scott$BM_ost, paired = TRUE, alternative = "less")
```

Paired t-test

```
data:  scott$BM_cast and scott$BM_ost
t = -13.771, df = 201, p-value < 2.2e-16
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
    -Inf -0.5706953
sample estimates:
mean difference
    -0.6485149
```

```
t.test(scott$BD_cast, scott$BD_ost, paired = TRUE, alternative = "less")
```

Paired t-test

```
data:  scott$BD_cast and scott$BD_ost
t = -15.363, df = 201, p-value < 2.2e-16
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
    -Inf -0.6671212
sample estimates:
mean difference
    -0.7475248
```

```
t.test(scott$LM_cast, scott$LM_ost, paired = TRUE, alternative = "less")
```

Paired t-test

```
data:  scott$LM_cast and scott$LM_ost
t = -15.299, df = 201, p-value < 2.2e-16
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
    -Inf -1.117192
sample estimates:
mean difference
    -1.252475
```

```
t.test(scott$LD_cast, scott$LD_ost, paired = TRUE, alternative = "less")
```

Paired t-test

```
data:  scott$LD_cast and scott$LD_ost
t = -14.154, df = 201, p-value < 2.2e-16
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
    -Inf -0.9488388
sample estimates:
mean difference
    -1.074257
```

## Modeling effect of age

Here are pearson correlation values for you. But the relationship clearly isn't linear so these probably aren't what you want. I've also computed Spearman correlation for you, which tests for any kind of monotonic relationship. This is much closer to what you are looking for than Pearson's coefficient.

```
#Pearson correlation  
cor.test(scott$BM_cast,scott$age_cast)
```

Pearson's product-moment correlation

```
data:  scott$BM_cast and scott$age_cast  
t = 3.1828, df = 200, p-value = 0.001691  
alternative hypothesis: true correlation is not equal to 0  
95 percent confidence interval:  
 0.08406264 0.34709569  
sample estimates:  
      cor  
0.2195656
```

```
cor.test(scott$BD_cast,scott$age_cast)
```

Pearson's product-moment correlation

```
data:  scott$BD_cast and scott$age_cast  
t = 4.6562, df = 200, p-value = 5.863e-06  
alternative hypothesis: true correlation is not equal to 0  
95 percent confidence interval:  
 0.1825572 0.4321221  
sample estimates:  
      cor  
0.3127268
```

```
cor.test(scott$LM_cast,scott$age_cast)
```

Pearson's product-moment correlation

```
data:  scott$LM_cast and scott$age_cast
t = 1.5552, df = 200, p-value = 0.1215
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.02918248  0.24368267
sample estimates:
      cor
0.1093089
```

```
cor.test(scott$LD_cast,scott$age_cast)
```

Pearson's product-moment correlation

```
data:  scott$LD_cast and scott$age_cast
t = 1.2261, df = 200, p-value = 0.2216
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.05229916  0.22178189
sample estimates:
      cor
0.08637546
```

```
cor.test(scott$BM_ost,scott$age_death)
```

Pearson's product-moment correlation

```
data:  scott$BM_ost and scott$age_death
t = 2.2238, df = 200, p-value = 0.02728
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
  0.01766402  0.28722782
sample estimates:
      cor
0.1553362
```

```
cor.test(scott$BD_ost,scott$age_death)
```

Pearson's product-moment correlation

```
data:  scott$BD_ost and scott$age_death
t = 4.0605, df = 200, p-value = 7.025e-05
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1433828 0.3988276
sample estimates:
      cor
0.2759712
```

```
cor.test(scott$LM_ost,scott$age_death)
```

Pearson's product-moment correlation

```
data:  scott$LM_ost and scott$age_death
t = 0.9911, df = 200, p-value = 0.3228
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.06880546 0.20597272
sample estimates:
      cor
0.06990958
```

```
cor.test(scott$LD_ost,scott$age_death)
```

Pearson's product-moment correlation

```
data:  scott$LD_ost and scott$age_death
t = 1.2046, df = 200, p-value = 0.2298
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.05381057 0.22034029
sample estimates:
      cor
0.08487091
```



```
#Spearman correlation  
cor.test(scott$BM_cast,scott$age_cast, method = "spearman")
```

Warning in cor.test.default(scott\$BM\_cast, scott\$age\_cast, method = "spearman"): Cannot compute exact p-value with ties

Spearman's rank correlation rho

```
data:  scott$BM_cast and scott$age_cast  
S = 980622, p-value = 3.654e-05  
alternative hypothesis: true rho is not equal to 0  
sample estimates:  
      rho  
0.2861458
```

```
cor.test(scott$BD_cast,scott$age_cast, method = "spearman")
```

Warning in cor.test.default(scott\$BD\_cast, scott\$age\_cast, method = "spearman"): Cannot compute exact p-value with ties

Spearman's rank correlation rho

```
data:  scott$BD_cast and scott$age_cast  
S = 884530, p-value = 1.98e-07  
alternative hypothesis: true rho is not equal to 0  
sample estimates:  
      rho  
0.3560972
```

```
cor.test(scott$LM_cast,scott$age_cast, method = "spearman")
```

Warning in cor.test.default(scott\$LM\_cast, scott\$age\_cast, method = "spearman"): Cannot compute exact p-value with ties

Spearman's rank correlation rho

```
data:  scott$LM_cast and scott$age_cast
S = 1179306, p-value = 0.04455
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.1415119
```

```
cor.test(scott$LD_cast,scott$age_cast, method = "spearman")
```

```
Warning in cor.test.default(scott$LD_cast, scott$age_cast, method =
"spearman"): Cannot compute exact p-value with ties
```

Spearman's rank correlation rho

```
data:  scott$LD_cast and scott$age_cast
S = 1234954, p-value = 0.1526
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.1010022
```

```
cor.test(scott$BM_ost,scott$age_death, method = "spearman")
```

```
Warning in cor.test.default(scott$BM_ost, scott$age_death, method =
"spearman"): Cannot compute exact p-value with ties
```

Spearman's rank correlation rho

```
data:  scott$BM_ost and scott$age_death
S = 1163676, p-value = 0.02983
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.1528901
```

```
cor.test(scott$BD_ost,scott$age_death, method = "spearman")
```

```
Warning in cor.test.default(scott$BD_ost, scott$age_death, method =  
"spearman"): Cannot compute exact p-value with ties
```

Spearman's rank correlation rho

```
data:  scott$BD_ost and scott$age_death  
S = 961396, p-value = 1.426e-05  
alternative hypothesis: true rho is not equal to 0  
sample estimates:  
      rho  
0.3001418
```

```
cor.test(scott$LM_ost,scott$age_death, method = "spearman")
```

```
Warning in cor.test.default(scott$LM_ost, scott$age_death, method =  
"spearman"): Cannot compute exact p-value with ties
```

Spearman's rank correlation rho

```
data:  scott$LM_ost and scott$age_death  
S = 1264346, p-value = 0.2601  
alternative hypothesis: true rho is not equal to 0  
sample estimates:  
      rho  
0.07960586
```

```
cor.test(scott$LD_ost,scott$age_death, method = "spearman")
```

```
Warning in cor.test.default(scott$LD_ost, scott$age_death, method =  
"spearman"): Cannot compute exact p-value with ties
```

Spearman's rank correlation rho

```
data:  scott$LD_ost and scott$age_death  
S = 1231301, p-value = 0.1421  
alternative hypothesis: true rho is not equal to 0  
sample estimates:  
      rho  
0.1036615
```

```
Warning in cor.test.default(scott$BM_cast, scott$age_cast, method =  
"spearman"): Cannot compute exact p-value with ties
```

```
Warning in cor.test.default(scott$BD_cast, scott$age_cast, method =  
"spearman"): Cannot compute exact p-value with ties
```

```
Warning in cor.test.default(scott$LM_cast, scott$age_cast, method =  
"spearman"): Cannot compute exact p-value with ties
```

```
Warning in cor.test.default(scott$LD_cast, scott$age_cast, method =  
"spearman"): Cannot compute exact p-value with ties
```

```
Warning in cor.test.default(scott$BM_ost, scott$age_death, method =  
"spearman"): Cannot compute exact p-value with ties
```

```
Warning in cor.test.default(scott$BD_ost, scott$age_death, method =  
"spearman"): Cannot compute exact p-value with ties
```

```
Warning in cor.test.default(scott$LM_ost, scott$age_death, method =  
"spearman"): Cannot compute exact p-value with ties
```

```
Warning in cor.test.default(scott$LD_ost, scott$age_death, method =  
"spearman"): Cannot compute exact p-value with ties
```

```
out %>% pivot_wider(names_from = "BDLM", values_from = "pvalue")
```

```
# A tibble: 2 x 5  
  castost      BM      BD      LM      LD  
  <chr>      <dbl>    <dbl>  <dbl> <dbl>  
1 cast    0.0000365 0.000000198 0.0445 0.153  
2 ost     0.0298    0.0000143  0.260  0.142
```

## Do wear scores differ by sex? If so, how? (Is this by a specific quadrant? Is this by a side of tooth like buccal side or lingual side?)

If we ignore age entirely, there are significant differences in the mean vector of wear for males vs females at both cast and ost. If we then do individual tests for wear differences by region at cast and ost separately, there are significant differences in LM and LD at cast and LM at ost. (I used a 0.05 level of significant and a Bonferroni correction with 8 tests to get a cut off for significance of 0.00625. This is, as always, conservative.)

```
#Look for differences in male vs female with a hotelling t test
library(Hotelling)
```

Loading required package: corpcor

Attaching package: 'Hotelling'

The following object is masked from 'package:dplyr':

summarise

```
results_cast <- hotelling.test(~sex, data = scott %>% select(sex, BM_cast, BD_cast, LM_cast,
results_cast
```

Test stat: 43.503  
Numerator df: 4  
Denominator df: 86.6737395211063  
P-value: 4.217e-07

```
results_ost <- hotelling.test(~sex, data = scott %>% select(sex, BM_ost, BD_ost, LM_ost, LD_
results_ost
```

Test stat: 46.955  
Numerator df: 4  
Denominator df: 77.1019743929347  
P-value: 2.096e-07

```
#Now do individual t-tests on each component  
t.test(BM_cast~sex,data = scott)
```

Welch Two Sample t-test

```
data: BM_cast by sex  
t = -1.0677, df = 89.469, p-value = 0.2885  
alternative hypothesis: true difference in means between group female and group male is not equal to 0  
95 percent confidence interval:  
 -0.4743964  0.1427512  
sample estimates:  
mean in group female    mean in group male  
      5.814570           5.980392
```

```
t.test(BD_cast~sex,data = scott)
```

Welch Two Sample t-test

```
data: BD_cast by sex  
t = -2.1367, df = 88.856, p-value = 0.03537  
alternative hypothesis: true difference in means between group female and group male is not equal to 0  
95 percent confidence interval:  
 -0.75133061 -0.02726958  
sample estimates:  
mean in group female    mean in group male  
      5.218543           5.607843
```

```
t.test(LD_cast~sex,data = scott)
```

Welch Two Sample t-test

```
data: LD_cast by sex  
t = -5.9193, df = 81.962, p-value = 7.24e-08  
alternative hypothesis: true difference in means between group female and group male is not equal to 0  
95 percent confidence interval:  
 -2.267221 -1.126624  
sample estimates:
```

mean in group female	mean in group male
3.185430	4.882353

```
t.test(LM_cast~sex,data = scott)
```

Welch Two Sample t-test

```
data: LM_cast by sex
t = -4.4066, df = 91.756, p-value = 2.843e-05
alternative hypothesis: true difference in means between group female and group male is not 0
95 percent confidence interval:
 -1.7201029 -0.6512774
sample estimates:
mean in group female mean in group male
      3.794702      4.980392
```

```
#Now do individual t-tests on each component for ost
t.test(BM_ost~sex,data = scott)
```

Welch Two Sample t-test

```
data: BM_ost by sex
t = -2.1975, df = 72.846, p-value = 0.03116
alternative hypothesis: true difference in means between group female and group male is not 0
95 percent confidence interval:
 -0.71266152 -0.03477387
sample estimates:
mean in group female mean in group male
      6.410596      6.784314
```

```
t.test(BD_ost~sex,data = scott)
```

Welch Two Sample t-test

```
data: BD_ost by sex
t = -2.6646, df = 78.037, p-value = 0.009362
alternative hypothesis: true difference in means between group female and group male is not 0
```

```
95 percent confidence interval:
-0.8578023 -0.1241481
sample estimates:
mean in group female    mean in group male
      5.940397           6.431373
```

```
t.test(LD_ost~sex,data = scott)
```

Welch Two Sample t-test

```
data: LD_ost by sex
t = -6.2685, df = 78.183, p-value = 1.856e-08
alternative hypothesis: true difference in means between group female and group male is not 0
95 percent confidence interval:
-2.312318 -1.197615
sample estimates:
mean in group female    mean in group male
      4.245033           6.000000
```

```
t.test(LM_ost~sex,data = scott)
```

Welch Two Sample t-test

```
data: LM_ost by sex
t = -4.636, df = 76.593, p-value = 1.434e-05
alternative hypothesis: true difference in means between group female and group male is not 0
95 percent confidence interval:
-1.8871464 -0.7530302
sample estimates:
mean in group female    mean in group male
      5.013245           6.333333
```

## Modeling without adding in age 1.

But we should really be controlling for sex and age and id all at once. So we do that here using the observed data. I'm assuming linearity of the relationship between age and wear here.



## BM model

```
library(lme4)
```

Loading required package: Matrix

Attaching package: 'Matrix'

The following objects are masked from 'package:tidyr':

expand, pack, unpack

```
mod_bm <- lmer(value ~ age * sex + (1|id), data = scott_long %>% filter(BLMD == "BM" & age > 1)
mod_bm_red <- lmer(value ~ age + (1|id), data = scott_long %>% filter(BLMD == "BM" & age > 1)
#test stat
-2*(logLik(mod_bm_red) - logLik(mod_bm))
```

'log Lik.' 8.388506 (df=4)

```
#p-value
1 - pchisq(-2*(logLik(mod_bm_red) - logLik(mod_bm)),2)
```

'log Lik.' 0.015082 (df=4)

```
summary(mod_bm)
```

Linear mixed model fit by REML ['lmerMod']

Formula: value ~ age \* sex + (1 | id)

Data: scott\_long %>% filter(BLMD == "BM" & age > 1)

REML criterion at convergence: 1029.9

Scaled residuals:

	Min	1Q	Median	3Q	Max
	-3.5985	-0.4909	-0.0181	0.4744	2.3902

Random effects:

Groups	Name	Variance	Std.Dev.
id	(Intercept)	0.6855	0.8280
	Residual	0.2986	0.5464

Number of obs: 404, groups: id, 202

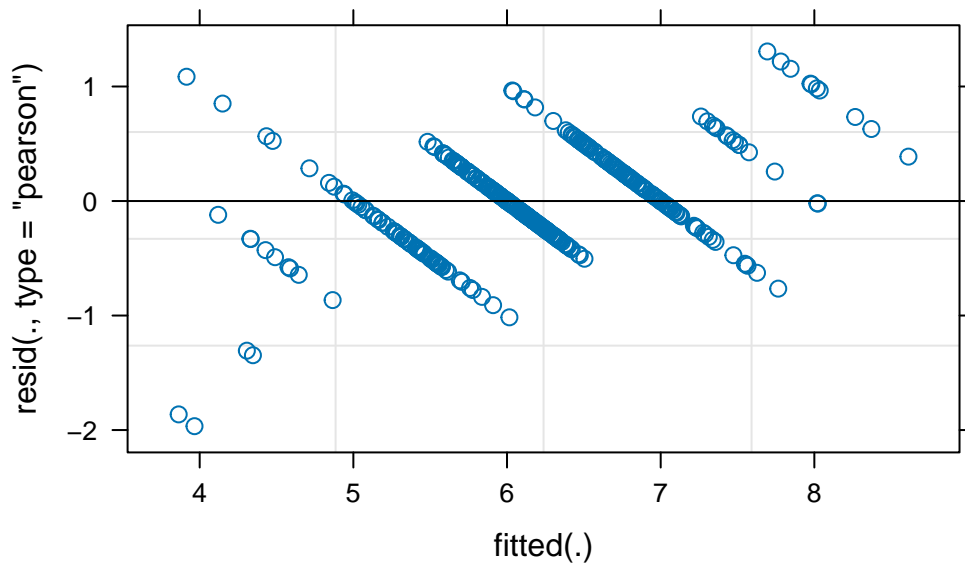
Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	4.297378	0.271336	15.838
age	0.007310	0.001051	6.956
sexmale	-0.643460	0.529005	-1.216
age:sexmale	0.005668	0.002324	2.439

Correlation of Fixed Effects:

	(Intr)	age	sexmal
age	-0.962		
sexmale	-0.513	0.493	
age:sexmale	0.435	-0.452	-0.958

```
plot(mod_bm)
```



```
confint(mod_bm)
```

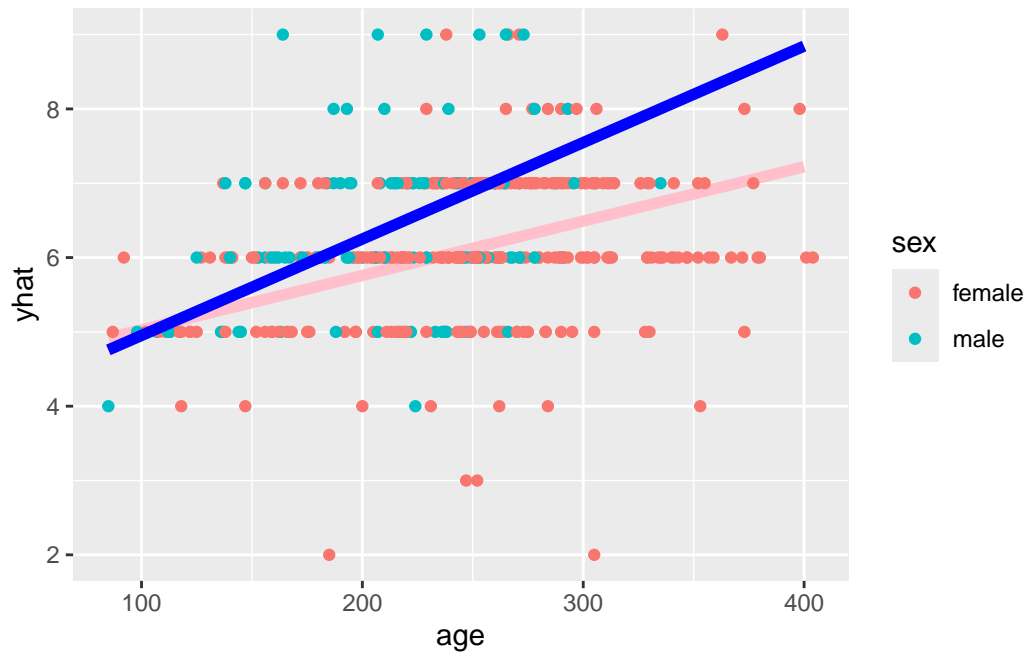
Computing profile confidence intervals ...

	2.5 %	97.5 %
.sig01	0.717071853	0.934476579
.sigma	0.493535385	0.607612666
(Intercept)	3.736882609	4.864515720
age	0.005100206	0.009488054
sexmale	-1.679204458	0.398340561
age:sexmale	0.001060733	0.010232615

```
newx <- data.frame(age = 85:400, sex = "female")
newx$yhat <- (predict(mod_bm, newdata = newx, re.form = ~0))

newxm <- data.frame(age = 85:400, sex = "male")
newxm$yhat <- (predict(mod_bm, newdata = newxm, re.form = ~0))

ggplot(aes(x = age, y = yhat), data = newx) + geom_path(col = "pink", lwd = 2) + geom_point(col = "red", lwd = 2)
```



## BD Model

```
library(lme4)
mod_bd <- lmer(value ~ age*sex + (1|id), data = scott_long %>% filter(BLMD == "BD"& age > 1))
```

```
mod_bd_red <- lmer(value ~ age + (1|id), data = scott_long %>% filter(BLMD == "BD"& age > 1))

#Test stat
-2*(logLik(mod_bd_red) - logLik(mod_bd))
```

```
'log Lik.' 20.38624 (df=4)
```

```
#p-value
1 - pchisq(-2*(logLik(mod_bd_red) - logLik(mod_bd)),2)
```

```
'log Lik.' 3.742702e-05 (df=4)
```

```
summary(mod_bd)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: value ~ age * sex + (1 | id)
Data: scott_long %>% filter(BLMD == "BD" & age > 1)
```

```
REML criterion at convergence: 1085
```

```
Scaled residuals:
```

Min	1Q	Median	3Q	Max
-3.1920	-0.4849	0.0213	0.5208	2.1544

```
Random effects:
```

Groups	Name	Variance	Std.Dev.
id	(Intercept)	0.8460	0.9198
Residual		0.3257	0.5707

Number of obs: 404, groups: id, 202

```
Fixed effects:
```

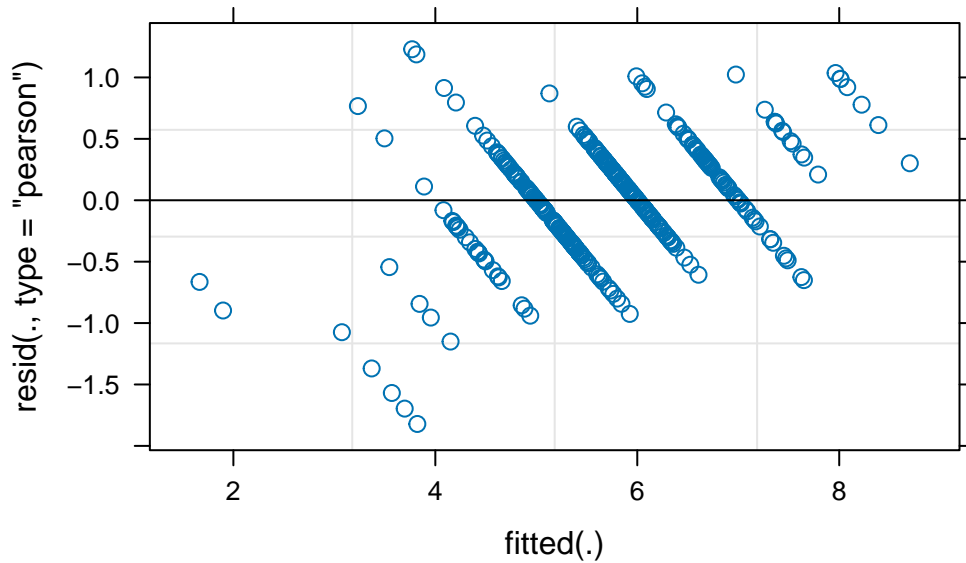
	Estimate	Std. Error	t value
(Intercept)	2.953393	0.293663	10.057
age	0.010575	0.001136	9.310
sexmale	-0.376460	0.571225	-0.659
age:sexmale	0.005800	0.002505	2.316

```
Correlation of Fixed Effects:
```

	(Intr) age	sexmal
age	-0.960	

```
sexmale      -0.514  0.494
age:sexmale   0.436 -0.453 -0.956
```

```
plot(mod_bd)
```



```
confint(mod_bd)
```

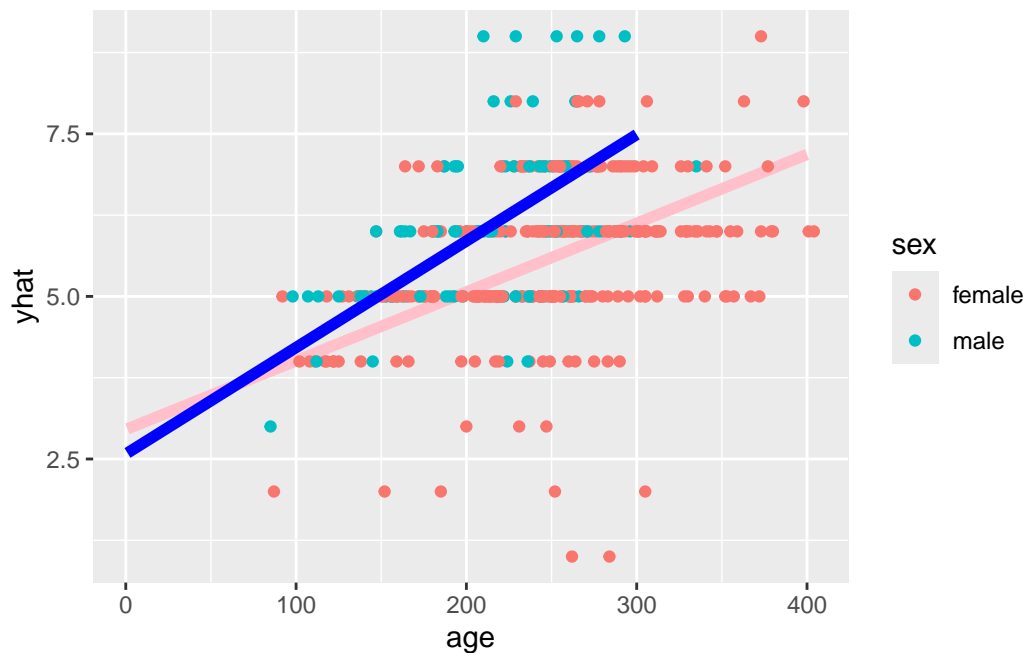
Computing profile confidence intervals ...

	2.5 %	97.5 %
.sig01	0.8027379060	1.03322393
.sigma	0.5161327884	0.63335972
(Intercept)	2.3468412675	3.56769773
age	0.0081841638	0.01292951
sexmale	-1.4921059944	0.73938156
age:sexmale	0.0008982572	0.01069068

```
newx <- data.frame(age = 1:400, sex = "female")
newx$yhat <- (predict(mod_bd, newdata = newx, re.form = ~0))

newxm <- data.frame(age = 1:300, sex = "male")
newxm$yhat <- (predict(mod_bd, newdata = newxm, re.form = ~0))

ggplot(aes(x = age, y = yhat), data = newx) + geom_path(col = "pink", lwd = 2) + geom_point(
```



## LM model

```
library(lme4)
mod_lm <- lmer(value ~ age * sex + (1|id), data = scott_long %>% filter(BLMD == "LM" & age > 100))
mod_lm_red <- lmer(value ~ age + (1|id), data = scott_long %>% filter(BLMD == "LM" & age > 100))

#test stat
-2*(logLik(mod_lm_red) - logLik(mod_lm))
```

'log Lik.' 36.76324 (df=4)

```
#p-value
1 - pchisq(-2*(logLik(mod_lm_red) - logLik(mod_lm)),2)
```

'log Lik.' 1.039833e-08 (df=4)

```
summary(mod_lm)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: value ~ age * sex + (1 | id)
Data: scott_long %>% filter(BLMD == "LM" & age > 1)
```

REML criterion at convergence: 1498.8

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.0781	-0.4614	0.0396	0.4585	3.1828

Random effects:

Groups	Name	Variance	Std.Dev.
id	(Intercept)	2.127	1.4584
	Residual	0.991	0.9955

Number of obs: 404, groups: id, 202

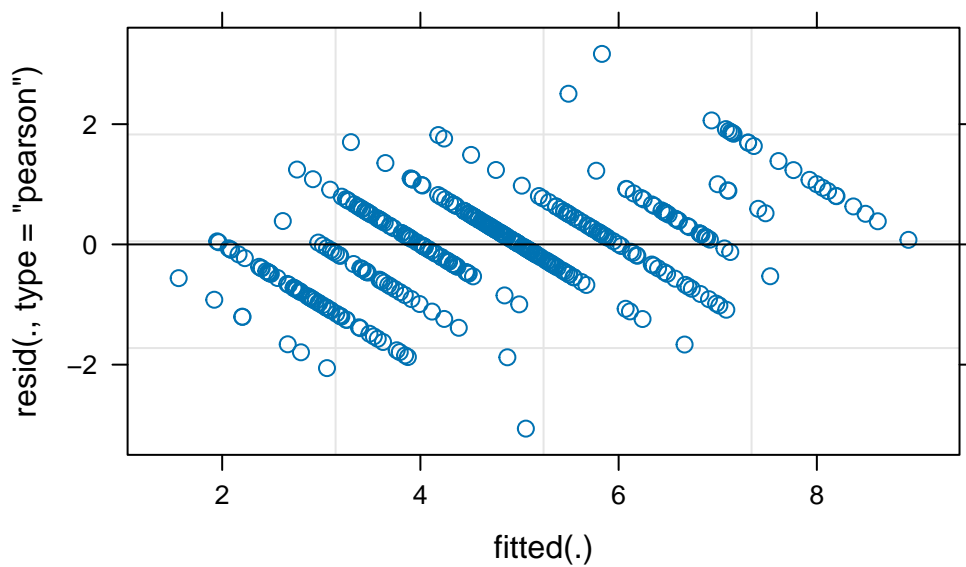
Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	1.075884	0.484758	2.219
age	0.013402	0.001879	7.134
sexmale	-0.028880	0.946246	-0.031
age:sexmale	0.008525	0.004160	2.049

Correlation of Fixed Effects:

	(Intr)	age	sexmal
age	-0.962		
sexmale	-0.512	0.493	
age:sexmale	0.435	-0.452	-0.958

```
plot(mod_lm)
```



```
confint(mod_lm)
```

Computing profile confidence intervals ...

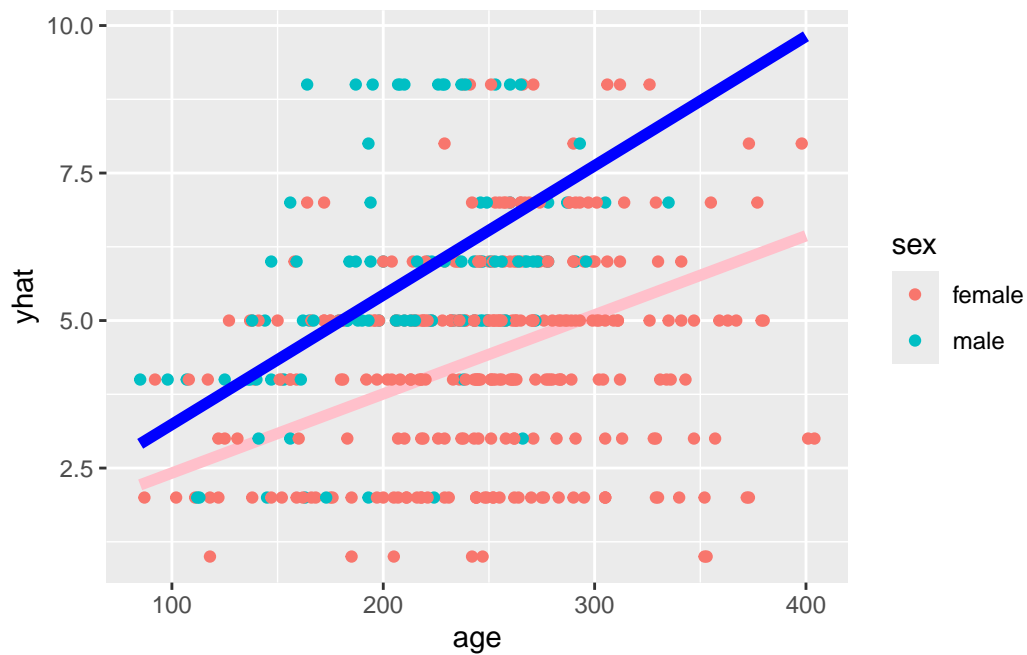
	2.5 %	97.5 %
.sig01	1.2513442405	1.65650448
.sigma	0.8965131755	1.11095339
(Intercept)	-0.0124080211	2.17636729
age	0.0090896745	0.01765623
sexmale	-1.8941560587	1.82481653
age:sexmale	0.0003987641	0.01669083

```
newx <- data.frame(age = 85:400, sex = "female")
newx$yhat <- (predict(mod_lm, newdata = newx, re.form = ~0))

newxm <- data.frame(age = 85:400, sex = "male")
newxm$yhat <- (predict(mod_lm, newdata = newxm, re.form = ~0))

ggplot(aes(x = age, y = yhat), data = newx) + geom_path(col = "pink", lwd = 2) + geom_point(
```





## LD Model

```
library(lme4)
mod_ld <- lmer(value ~ age*sex + (1|id), data = scott_long %>% filter(BLMD == "LD"& age > 1))

mod_ld_red <- lmer(value ~ age + (1|id), data = scott_long %>% filter(BLMD == "LD"& age > 1))

#Test stat
-2*(logLik(mod_ld_red) - logLik(mod_ld))
```

'log Lik.' 60.42077 (df=4)

```
#p-value
1 - pchisq(-2*(logLik(mod_ld_red) - logLik(mod_ld)),2)
```

'log Lik.' 7.582823e-14 (df=4)

```
summary(mod_ld)
```

```

Linear mixed model fit by REML ['lmerMod']
Formula: value ~ age * sex + (1 | id)
Data: scott_long %>% filter(BLMD == "LD" & age > 1)

```

REML criterion at convergence: 1442.5

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.37959	-0.41518	0.04341	0.41282	2.83646

Random effects:

Groups	Name	Variance	Std.Dev.
id	(Intercept)	2.2376	1.496
	Residual	0.7517	0.867

Number of obs: 404, groups: id, 202

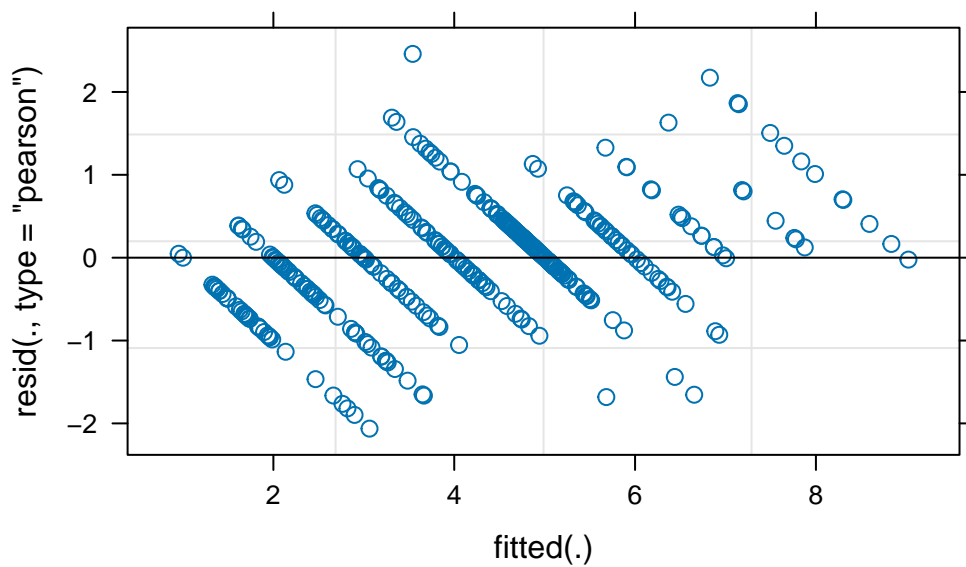
Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	0.225817	0.463850	0.487
age	0.014052	0.001791	7.845
sexmale	0.493513	0.899906	0.548
age:sexmale	0.008408	0.003937	2.135

Correlation of Fixed Effects:

	(Intr)	age	sexmal
age	-0.959		
sexmale	-0.515	0.494	
age:sexmale	0.436	-0.455	-0.954

```
plot(mod_ld)
```



```
confint(mod_lm)
```

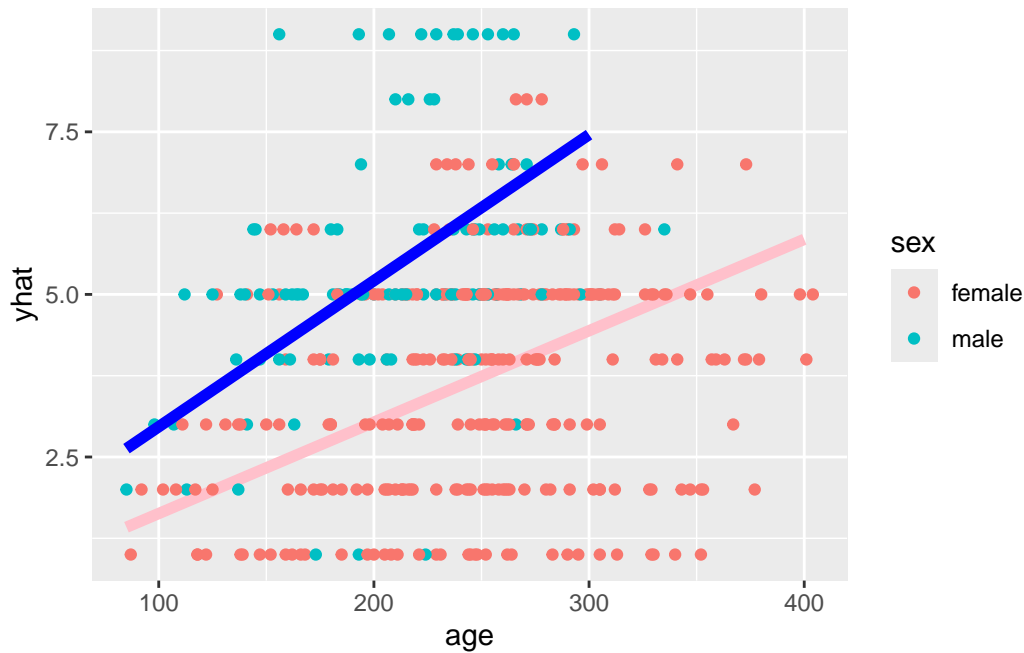
Computing profile confidence intervals ...

	2.5 %	97.5 %
.sig01	1.2513442405	1.65650448
.sigma	0.8965131755	1.11095339
(Intercept)	-0.0124080211	2.17636729
age	0.0090896745	0.01765623
sexmale	-1.8941560587	1.82481653
age:sexmale	0.0003987641	0.01669083

```
newx <- data.frame(age = 85:400, sex = "female")
newx$yhat <- (predict(mod_ld, newdata = newx, re.form = ~0))

newxm <- data.frame(age = 85:300, sex = "male")
newxm$yhat <- (predict(mod_ld, newdata = newxm, re.form = ~0))

ggplot(aes(x = age, y = yhat), data = newx) + geom_path(col = "pink", lwd = 2) + geom_point(
```



## Modeling

But we should really be controlling for sex and age and id all at once and I think we should be looking at age 0.

### BM model

```
library(lme4)
mod_bm <- lmer(value ~ (age+I(age^2)+ I(age^3))*sex + (1|id), data = scott_long %>% filter(BL == 1))
```

Warning: Some predictor variables are on very different scales: consider rescaling

```
mod_bm_red <- lmer(value ~ (age+I(age^2)+ I(age^3)) + (1|id), data = scott_long %>% filter(BL == 1))
```

Warning: Some predictor variables are on very different scales: consider rescaling

```
#test stat
-2*(logLik(mod_bm_red) - logLik(mod_bm))
```

```
'log Lik.' -43.27445 (df=6)
```

```
#p-value
1 - pchisq(-2*(logLik(mod_bm_red) - logLik(mod_bm)),4)
```

```
'log Lik.' 1 (df=6)
```

```
summary(mod_bm)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: value ~ (age + I(age^2) + I(age^3)) * sex + (1 | id)
Data: scott_long %>% filter(BLMD == "BM")
```

```
REML criterion at convergence: 1509.1
```

```
Scaled residuals:
```

Min	1Q	Median	3Q	Max
-4.9601	-0.4293	-0.0483	0.5005	3.2601

```
Random effects:
```

Groups	Name	Variance	Std.Dev.
id	(Intercept)	0.1967	0.4435
Residual		0.4227	0.6502

```
Number of obs: 606, groups: id, 202
```

```
Fixed effects:
```

	Estimate	Std. Error	t value
(Intercept)	9.548e-01	6.467e-02	14.764
age	5.031e-02	3.192e-03	15.759
I(age^2)	-1.614e-04	2.289e-05	-7.050
I(age^3)	1.779e-07	4.084e-08	4.356
sexmale	-2.007e-02	1.289e-01	-0.156
age:sexmale	1.172e-02	7.407e-03	1.582
I(age^2):sexmale	-8.112e-05	6.284e-05	-1.291
I(age^3):sexmale	1.667e-07	1.342e-07	1.242

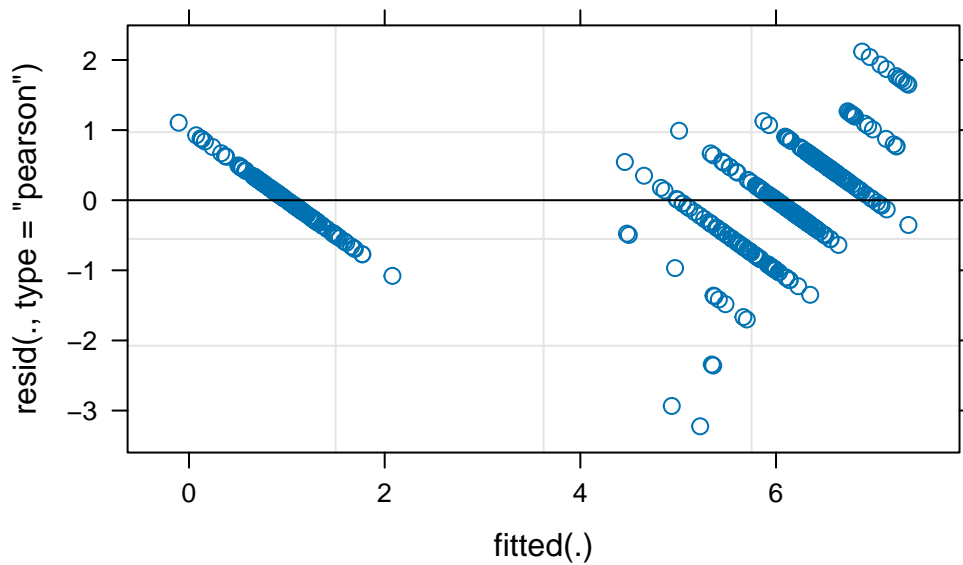
```
Correlation of Fixed Effects:
```

```

      (Intr) age      I(g^2) I(g^3) sexmal ag:sexm I(^2):
age      -0.230
I(age^2)   0.152 -0.984
I(age^3)  -0.122  0.951 -0.989
sexmale   -0.502  0.115 -0.076  0.061
age:sexmale 0.099 -0.431  0.424 -0.410 -0.231
I(g^2):sexml -0.055  0.359 -0.364  0.360  0.152 -0.981
I(g^3):sexml  0.037 -0.289  0.301 -0.304 -0.121  0.941 -0.988
fit warnings:
Some predictor variables are on very different scales: consider rescaling

```

```
plot(mod_bm)
```



```
confint(mod_bm)
```

Computing profile confidence intervals ...

Warning in profile.merMod(object, which = parm, signames = oldNames, ...):  
non-monotonic profile for I(age^3)

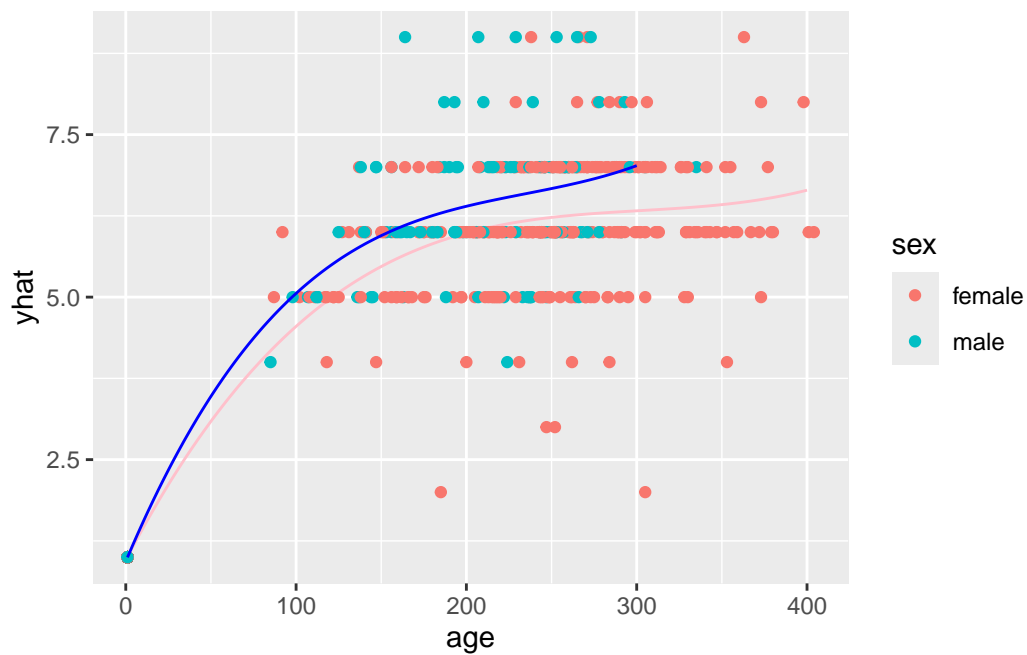
Warning in confint.thpr(pp, level = level, zeta = zeta): bad spline fit for  
I(age^3): falling back to linear interpolation

	2.5 %	97.5 %
.sig01	3.620505e-01	5.189840e-01
.sigma	6.042450e-01	6.939545e-01
(Intercept)	8.287261e-01	1.080873e+00
age	4.408148e-02	5.654561e-02
I(age^2)	-2.060484e-04	-1.167261e-04
I(age^3)	9.825674e-08	2.575734e-07
sexmale	-2.713796e-01	2.312428e-01
age:sexmale	-2.725521e-03	2.616369e-02
I(age^2):sexmale	-2.036417e-04	4.145212e-05
I(age^3):sexmale	-9.510296e-08	4.282120e-07

```
newx <- data.frame(age = 1:400, sex = "female")
newx$yhat <- (predict(mod_bm, newdata = newx, re.form = ~0))

newxm <- data.frame(age = 1:300, sex = "male")
newxm$yhat <- (predict(mod_bm, newdata = newxm, re.form = ~0))

ggplot(aes(x = age, y = yhat), data = newx) + geom_path(col = "pink") + geom_point(aes(x = age, y = yhat, sex = sex))
```



## BD Model

```
library(lme4)
mod_bd <- lmer(value ~ (age+I(age^2)+ I(age^3))*sex + (1|id), data = scott_long %>% filter(BLMD == "BD"))
```

Warning: Some predictor variables are on very different scales: consider rescaling

```
mod_bd_red <- lmer(value ~ (age+I(age^2)+ I(age^3)) + (1|id), data = scott_long %>% filter(BLMD == "BD"))
```

Warning: Some predictor variables are on very different scales: consider rescaling

```
#Test stat
-2*(logLik(mod_bd_red) - logLik(mod_bd))
```

'log Lik.' -16.54741 (df=6)

```
#p-value
1 - pchisq(-2*(logLik(mod_bd_red) - logLik(mod_bd)),4)
```

'log Lik.' 1 (df=6)

```
summary(mod_bd)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: value ~ (age + I(age^2) + I(age^3)) * sex + (1 | id)
Data: scott_long %>% filter(BLMD == "BD")
```

REML criterion at convergence: 1609

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.2191	-0.4719	0.0327	0.4445	2.7899

Random effects:

Groups	Name	Variance	Std.Dev.
id	(Intercept)	0.2484	0.4984



```

Residual          0.4915   0.7011
Number of obs: 606, groups:  id, 202

```

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	9.622e-01	7.067e-02	13.616
age	3.837e-02	3.472e-03	11.052
I(age^2)	-1.063e-04	2.490e-05	-4.268
I(age^3)	1.128e-07	4.444e-08	2.538
sexmale	-6.353e-03	1.409e-01	-0.045
age:sexmale	4.916e-03	8.048e-03	0.611
I(age^2):sexmale	-1.080e-05	6.829e-05	-0.158
I(age^3):sexmale	1.954e-08	1.457e-07	0.134

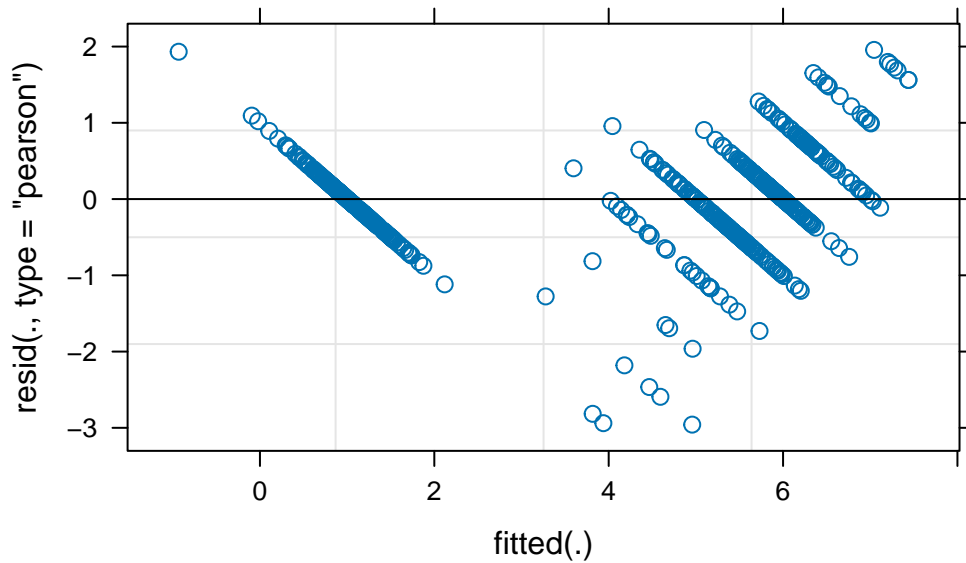
Correlation of Fixed Effects:

	(Intr)	age	I(g^2)	I(g^3)	sexmal	ag:sexm	I(^2):
age		-0.226					
I(age^2)	0.150	-0.984					
I(age^3)	-0.120	0.951	-0.989				
sexmale	-0.502	0.113	-0.075	0.060			
age:sexmale	0.097	-0.431	0.425	-0.410	-0.227		
I(g^2):sxm1	-0.055	0.359	-0.365	0.361	0.149	-0.981	
I(g^3):sxm1	0.037	-0.290	0.302	-0.305	-0.119	0.941	-0.988

fit warnings:

Some predictor variables are on very different scales: consider rescaling

```
plot(mod_bd)
```



```
confint(mod_bm)
```

Computing profile confidence intervals ...

Warning in profile.merMod(object, which = parm, signames = oldNames, ...):  
non-monotonic profile for I(age^3)

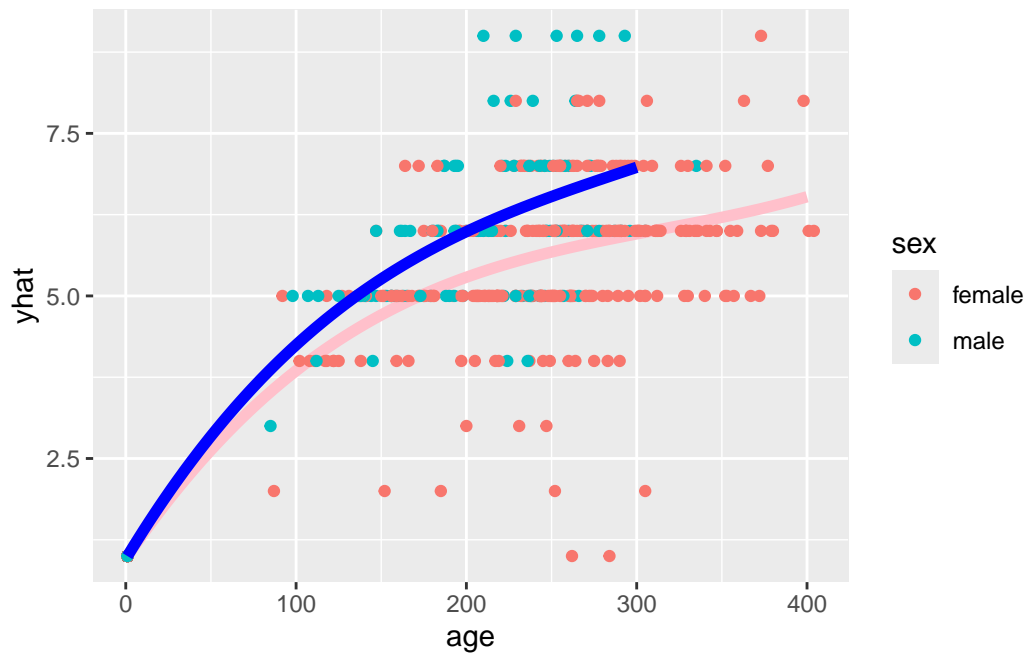
Warning in confint.thpr(pp, level = level, zeta = zeta): bad spline fit for  
I(age^3): falling back to linear interpolation

	2.5 %	97.5 %
.sig01	3.620505e-01	5.189840e-01
.sigma	6.042450e-01	6.939545e-01
(Intercept)	8.287261e-01	1.080873e+00
age	4.408148e-02	5.654561e-02
I(age^2)	-2.060484e-04	-1.167261e-04
I(age^3)	9.825674e-08	2.575734e-07
sexmale	-2.713796e-01	2.312428e-01
age:sexmale	-2.725521e-03	2.616369e-02
I(age^2):sexmale	-2.036417e-04	4.145212e-05
I(age^3):sexmale	-9.510296e-08	4.282120e-07

```
newx <- data.frame(age = 1:400, sex = "female")
newx$yhat <- (predict(mod_bd, newdata = newx, re.form = ~0))

newxm <- data.frame(age = 1:300, sex = "male")
newxm$yhat <- (predict(mod_bd, newdata = newxm, re.form = ~0))

ggplot(aes(x = age, y = yhat), data = newx) + geom_path(col = "pink", lwd = 2) + geom_point(col = "pink", lwd = 2) + geom_point(col = "teal", lwd = 2)
```



## LM model

```
library(lme4)
mod_lm <- lmer(value ~ (age+I(age^2)+ I(age^3))*sex + (1|id), data = scott_long %>% filter(BL < 100))
```

Warning: Some predictor variables are on very different scales: consider rescaling

```
mod_lm_red <- lmer(value ~ (age+I(age^2)+ I(age^3)) + (1|id), data = scott_long %>% filter(BL < 100))
```

Warning: Some predictor variables are on very different scales: consider rescaling

```
#test stat  
-2*(logLik(mod_lm_red) - logLik(mod_lm))
```

'log Lik.' 18.82511 (df=6)

```
#p-value  
1 - pchisq(-2*(logLik(mod_lm_red) - logLik(mod_lm)),4)
```

'log Lik.' 0.0008506206 (df=6)

```
summary(mod_lm)
```

Linear mixed model fit by REML ['lmerMod']  
Formula: value ~ (age + I(age^2) + I(age^3)) \* sex + (1 | id)  
Data: scott\_long %>% filter(BLMD == "LM")

REML criterion at convergence: 2190.2

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.68172	-0.45551	0.03389	0.46949	2.83204

Random effects:

Groups	Name	Variance	Std.Dev.
id	(Intercept)	0.5773	0.7598
Residual		1.3398	1.1575

Number of obs: 606, groups: id, 202

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	9.882e-01	1.138e-01	8.684
age	1.392e-02	5.638e-03	2.469
I(age^2)	2.961e-05	4.041e-05	0.733
I(age^3)	-1.096e-07	7.209e-08	-1.521
sexmale	-6.299e-03	2.268e-01	-0.028
age:sexmale	3.324e-03	1.309e-02	0.254
I(age^2):sexmale	5.081e-05	1.111e-04	0.457
I(age^3):sexmale	-1.392e-07	2.372e-07	-0.587

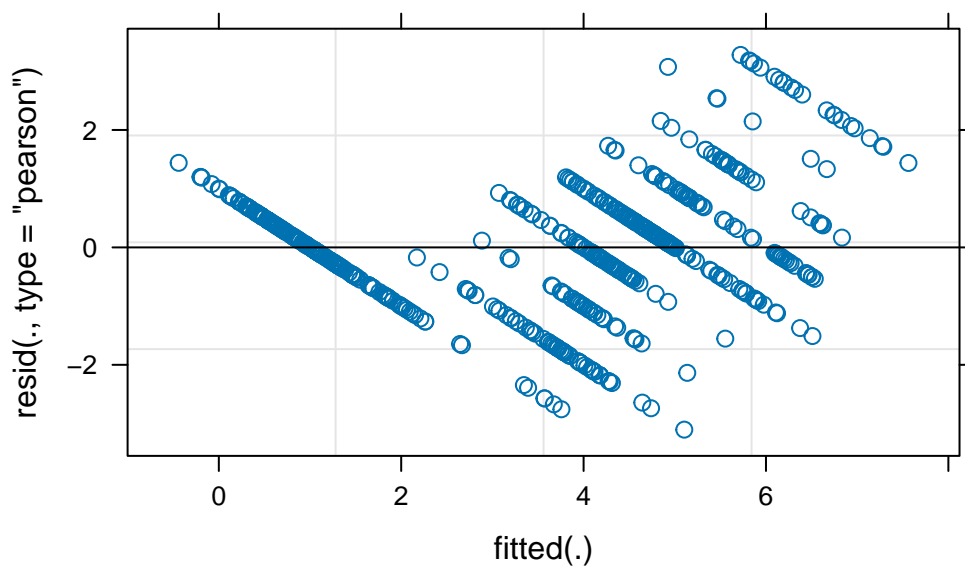
Correlation of Fixed Effects:

```

      (Intr) age      I(g^2) I(g^3) sexmal ag:sexm I(^2):
age      -0.234
I(age^2)   0.154 -0.984
I(age^3)  -0.124  0.951 -0.989
sexmale   -0.502  0.117 -0.077  0.062
age:sexmale 0.101 -0.431  0.424 -0.409 -0.235
I(g^2):sexml -0.056  0.358 -0.364  0.360  0.154 -0.981
I(g^3):sexml  0.038 -0.289  0.301 -0.304 -0.122  0.941 -0.988
fit warnings:
Some predictor variables are on very different scales: consider rescaling

```

```
plot(mod_lm)
```



```
confint(mod_lm)
```

Computing profile confidence intervals ...

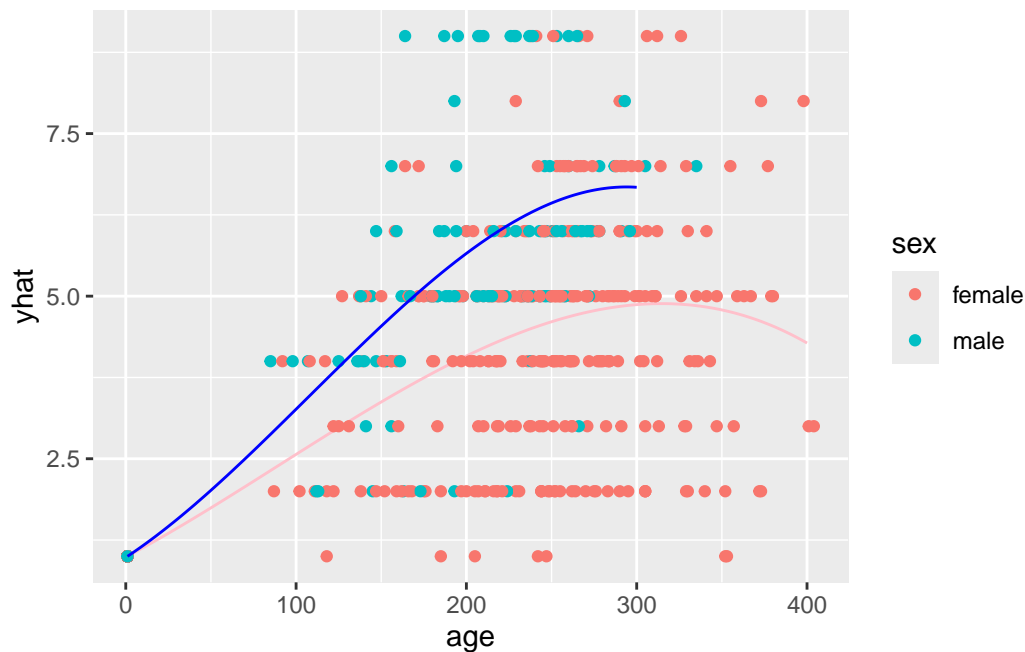
	2.5 %	97.5 %
.sig01	6.136642e-01	8.936871e-01
.sigma	1.075742e+00	1.235637e+00
(Intercept)	7.663453e-01	1.210014e+00
age	2.914712e-03	2.496790e-02
I(age^2)	-4.943002e-05	1.084477e-04
I(age^3)	-2.501810e-07	3.119453e-08
sexmale	-4.485078e-01	4.359092e-01

```
age:sexmale      -2.223166e-02  2.884756e-02
I(age^2):sexmale -1.657130e-04  2.675826e-04
I(age^3):sexmale -6.021448e-07  3.231202e-07
```

```
newx <- data.frame(age = 1:400, sex = "female")
newx$yhat <- (predict(mod_lm, newdata = newx, re.form = ~0))

newxm <- data.frame(age = 1:300, sex = "male")
newxm$yhat <- (predict(mod_lm, newdata = newxm, re.form = ~0))

ggplot(aes(x = age, y = yhat), data = newx) + geom_path(col = "pink") + geom_point(aes(x = age, y = yhat, sex = sex), col = "pink")
```



## BD Model

```
library(lme4)
mod_ld <- lmer(value ~ (age+I(age^2)+ I(age^3))*sex + (1|id), data = scott_long %>% filter(BL > 0))
```

Warning: Some predictor variables are on very different scales: consider rescaling

```
mod_ld_red <- lmer(value ~ (age+I(age^2)+ I(age^3)) + (1|id), data = scott_long %>% filter(BLMD == "LD"))
```

Warning: Some predictor variables are on very different scales: consider rescaling

```
#Test stat  
-2*(logLik(mod_ld_red) - logLik(mod_ld))
```

```
'log Lik.' 73.99111 (df=6)
```

```
#p-value  
1 - pchisq(-2*(logLik(mod_ld) - logLik(mod_ld)),4)
```

```
'log Lik.' 1 (df=10)
```

```
summary(mod_ld)
```

Linear mixed model fit by REML ['lmerMod']

Formula: value ~ (age + I(age^2) + I(age^3)) \* sex + (1 | id)

Data: scott\_long %>% filter(BLMD == "LD")

REML criterion at convergence: 2149.9

Scaled residuals:

	Min	1Q	Median	3Q	Max
	-2.91388	-0.53420	0.01674	0.54409	2.92409

Random effects:

Groups	Name	Variance	Std.Dev.
id	(Intercept)	0.6172	0.7856
	Residual	1.2127	1.1012

Number of obs: 606, groups: id, 202

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	9.938e-01	1.111e-01	8.942
age	4.706e-03	5.457e-03	0.862
I(age^2)	6.451e-05	3.915e-05	1.648
I(age^3)	-1.476e-07	6.985e-08	-2.113

```
sexmale          -1.221e-02  2.215e-01  -0.055
age:sexmale       1.201e-02  1.265e-02   0.950
I(age^2):sexmale  1.176e-05  1.073e-04   0.110
I(age^3):sexmale -9.301e-08  2.291e-07  -0.406
```

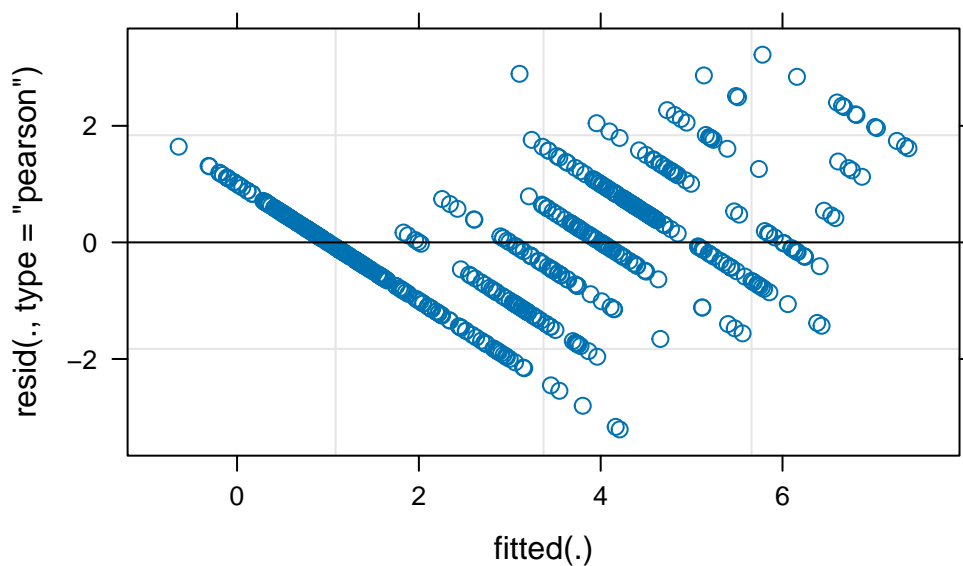
Correlation of Fixed Effects:

```
(Intr) age    I(g^2) I(g^3) sexmal ag:sexm I(^2):
age          -0.225
I(age^2)      0.149 -0.984
I(age^3)     -0.120  0.951 -0.989
sexmale      -0.502  0.113 -0.075  0.060
age:sexmale   0.097 -0.431  0.425 -0.410 -0.227
I(g^2):sexml -0.055  0.359 -0.365  0.361  0.149 -0.981
I(g^3):sexml  0.037 -0.290  0.302 -0.305 -0.119  0.941 -0.988
```

fit warnings:

Some predictor variables are on very different scales: consider rescaling

```
plot(mod_ld)
```



```
confint(mod_bm)
```

Computing profile confidence intervals ...

```
Warning in profile.merMod(object, which = parm, signames = oldNames, ...):
non-monotonic profile for I(age^3)
```



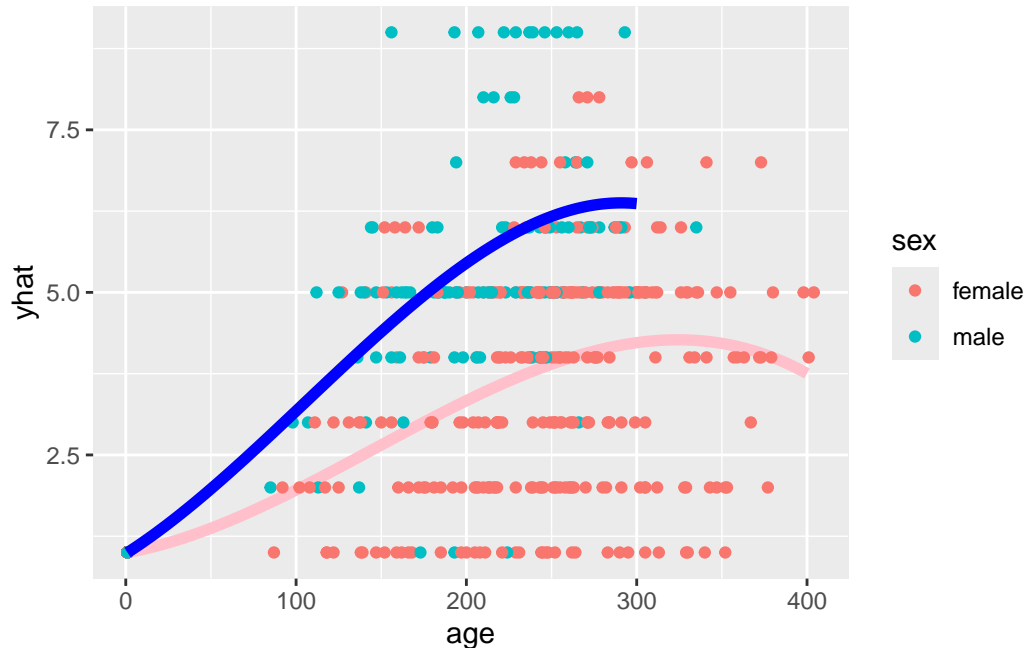
Warning in confint.thpr(pp, level = level, zeta = zeta): bad spline fit for I(age^3): falling back to linear interpolation

	2.5 %	97.5 %
.sig01	3.620505e-01	5.189840e-01
.sigma	6.042450e-01	6.939545e-01
(Intercept)	8.287261e-01	1.080873e+00
age	4.408148e-02	5.654561e-02
I(age^2)	-2.060484e-04	-1.167261e-04
I(age^3)	9.825674e-08	2.575734e-07
sexmale	-2.713796e-01	2.312428e-01
age:sexmale	-2.725521e-03	2.616369e-02
I(age^2):sexmale	-2.036417e-04	4.145212e-05
I(age^3):sexmale	-9.510296e-08	4.282120e-07

```
newx <- data.frame(age = 1:400, sex = "female")
newx$yhat <- (predict(mod_ld, newdata = newx, re.form = ~0))

newxm <- data.frame(age = 1:300, sex = "male")
newxm$yhat <- (predict(mod_ld, newdata = newxm, re.form = ~0))

ggplot(aes(x = age, y = yhat), data = newx) + geom_path(col = "pink", lwd = 2) + geom_point(aes(sex = "male", y = yhat), data = newxm, col = "teal", lwd = 2)
```



**Are there any patterns for rate of wear? That is, is there a standard wear through time (by month or year), or are some baboons wearing their teeth down faster than what we might expect?**