Homework 2

Brent Rappaport

September 12, 2017

Question 1

```
model1 <- lm(PPeerScale ~ age, data=data_long)
summary(model1)
```

```
##
## Call:
## lm(formula = PPeerScale ~ age, data = data_long)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -39.708  -4.254   2.502   6.677  10.641
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 61.01512    0.63836  95.581   <2e-16 ***
## age          0.11097    0.07297   1.521    0.129
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.904 on 1680 degrees of freedom
##   (766 observations deleted due to missingness)
## Multiple R-squared:  0.001375,   Adjusted R-squared:  0.0007804
## F-statistic: 2.313 on 1 and 1680 DF,  p-value: 0.1285
```

Average intercept is 61.02 and average slope is 0.11.

Question 2

```
model2 <- lmer(PPeerScale ~ 1 + (1 | ID), data=data_long)
summary(model2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: PPeerScale ~ 1 + (1 | ID)
##    Data: data_long
##
## REML criterion at convergence: 11498.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.0271 -0.4336  0.1407  0.5496  3.0506
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 42.93    6.552
```

```
##  Residual             38.33    6.191
## Number of obs: 1684, groups:  ID, 302
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  61.7003     0.4114     150
```

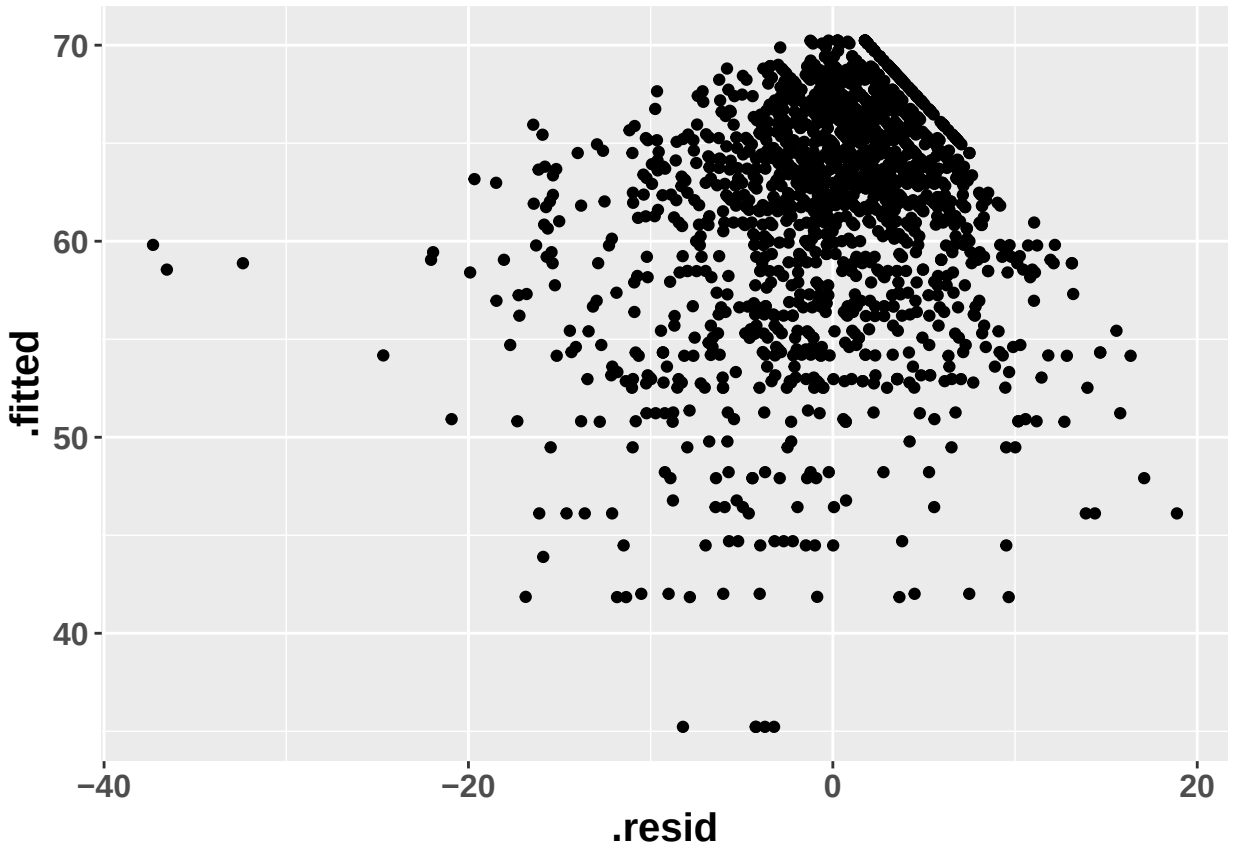
```
# Calculate ICC
42.93/(42.93+38.33)
```

```
## [1] 0.5283042
```

```
model2.aug <- augment(model2, data_long)
```

```
## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning in indices[which(stats::complete.cases(original))] =
## seq_len(nrow(x)): number of items to replace is not a multiple of
## replacement length
```

```
ggplot(model2.aug,aes(.resid,.fitted,group=ID)) +
  geom_point(alpha=1) +
  theme(text=element_text(lineheight=1, face="bold", size=15),
        legend.position="none")
```



Question 3

```
model3 <- lmer(PPeerScale ~ age + (1 | ID), data=data_long)
summary(model3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: PPeerScale ~ age + (1 | ID)
##    Data: data_long
##
## REML criterion at convergence: 11476.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.0643 -0.4159  0.1398  0.5449  3.1692
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 43.02    6.559
##  Residual             38.00    6.164
## Number of obs: 1682, groups:  ID, 302
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 60.51297    0.59831  101.14
## age          0.14905    0.05429    2.75
```

```
##
## Correlation of Fixed Effects:
##     (Intr)
## age -0.726
```

The esimates of the fixed effect for the Intercept doesn't change much, however now that age is being included as a fixed effect, we can get a sense of the average change over time in parent reported Peer scores. That is, for ever 1 year, participants increase an average of 0.15 in their score.

```
print(c(sigma(model2),sigma(model3)))
```

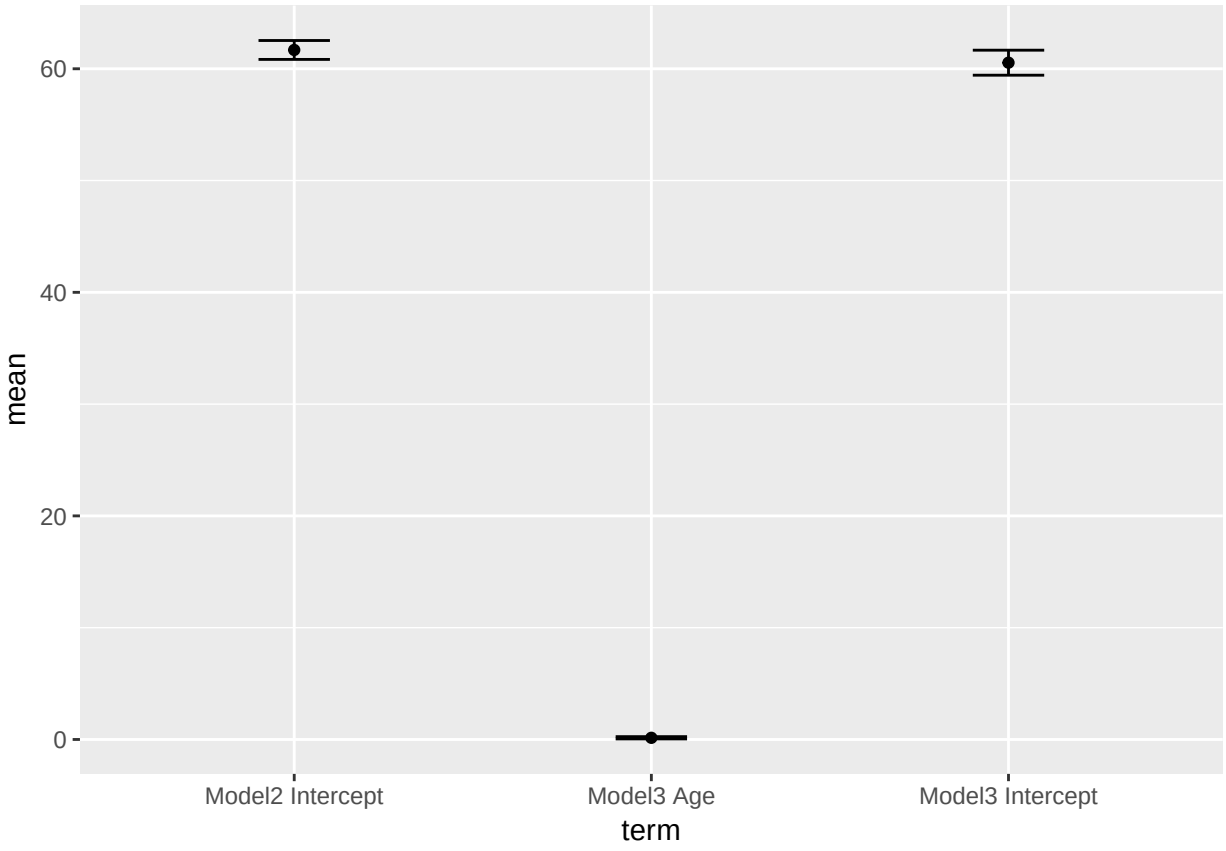
```
## [1] 6.190725 6.164179
```

The residual standard error (sigma here), doesn't change much between the two models, but does decrease when age is introduced as a fixed effect.

```
plot3_table <- FEsim(model2)
plot3_table[2:3,] <- FEsim(model3)
```

```
## Warning in `[<-.factor`(`*tmp*`, iseq, value = structure(1:2, .Label =
## c("(Intercept)", : invalid factor level, NA generated
```

```
plot3_table[1] <- c("Model2 Intercept","Model3 Intercept","Model3 Age")

plot3_table$ci <- 1.96*plot3_table$sd

ggplot(plot3_table,aes(x=term,y=mean)) +
  geom_point() +
  geom_errorbar(aes(ymin=mean-ci,ymax=mean+ci),width=0.2)
```



Question 4

```
model4 <- lmer(PPeerScale ~ age + (age | ID), data=data_long)
print(c(sigma(model3),sigma(model4)))
```

```
## [1] 6.164179 5.511522
```

```
anova(model3,model4)
```

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

```
## Data: data_long
## Models:
## model3: PPeerScale ~ age + (1 | ID)
## model4: PPeerScale ~ age + (age | ID)
##        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model3  4 11480 11502 -5736.1    11472
## model4  6 11386 11418 -5686.8    11374 98.722      2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Adding the random slope improves fit, according to the AIC, BIC, and likelihood ratio test, while also considerably lowering the residual standard error (from 6.1641795 to 5.5115224).

Question 5

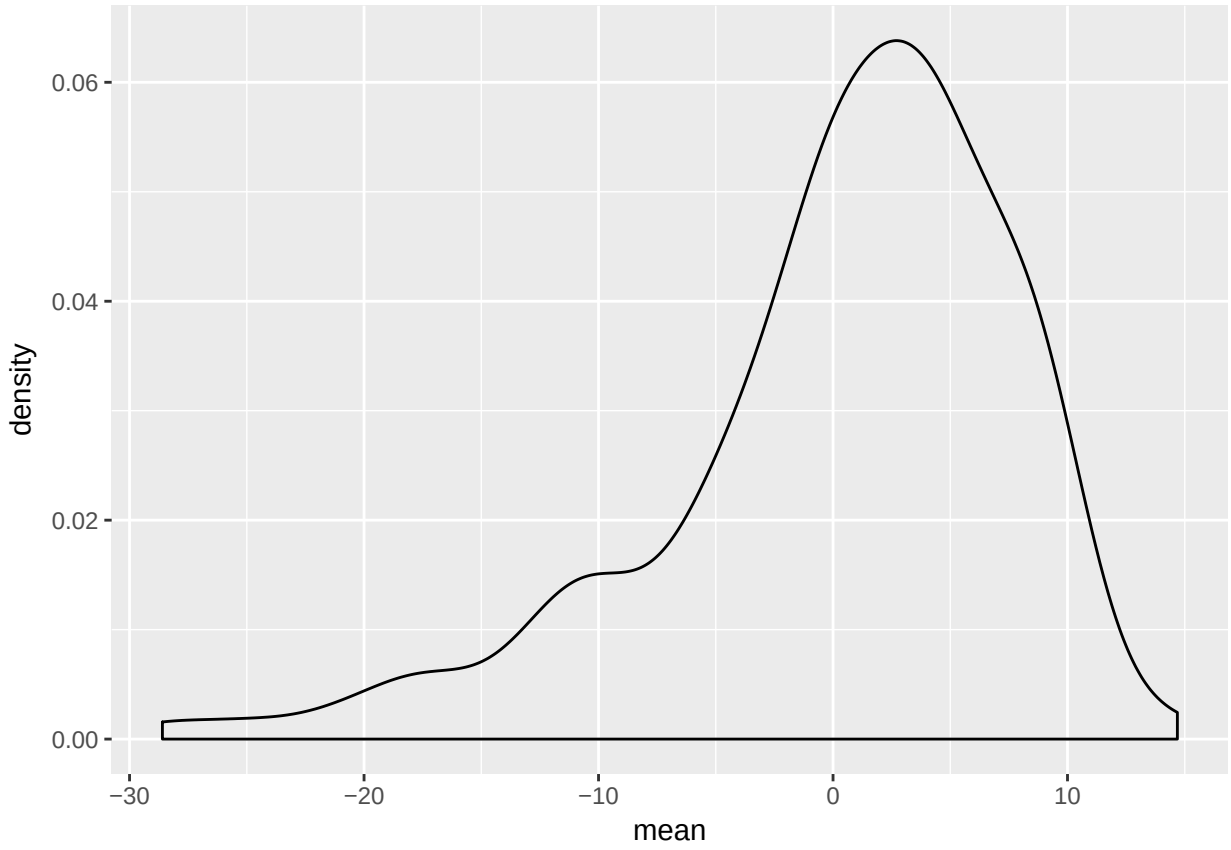
```
summary(model4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: PPeerScale ~ age + (age | ID)
##    Data: data_long
##
## REML criterion at convergence: 11376.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.8222 -0.3835  0.1106  0.4980  2.7892
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  ID       (Intercept) 88.0332  9.3826
##           age          0.8576  0.9261   -0.72
##  Residual             30.3769  5.5115
## Number of obs: 1682, groups:  ID, 302
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 60.58395    0.69693   86.93
## age          0.13846    0.07578    1.83
##
## Correlation of Fixed Effects:
##     (Intr)
## age -0.812
```

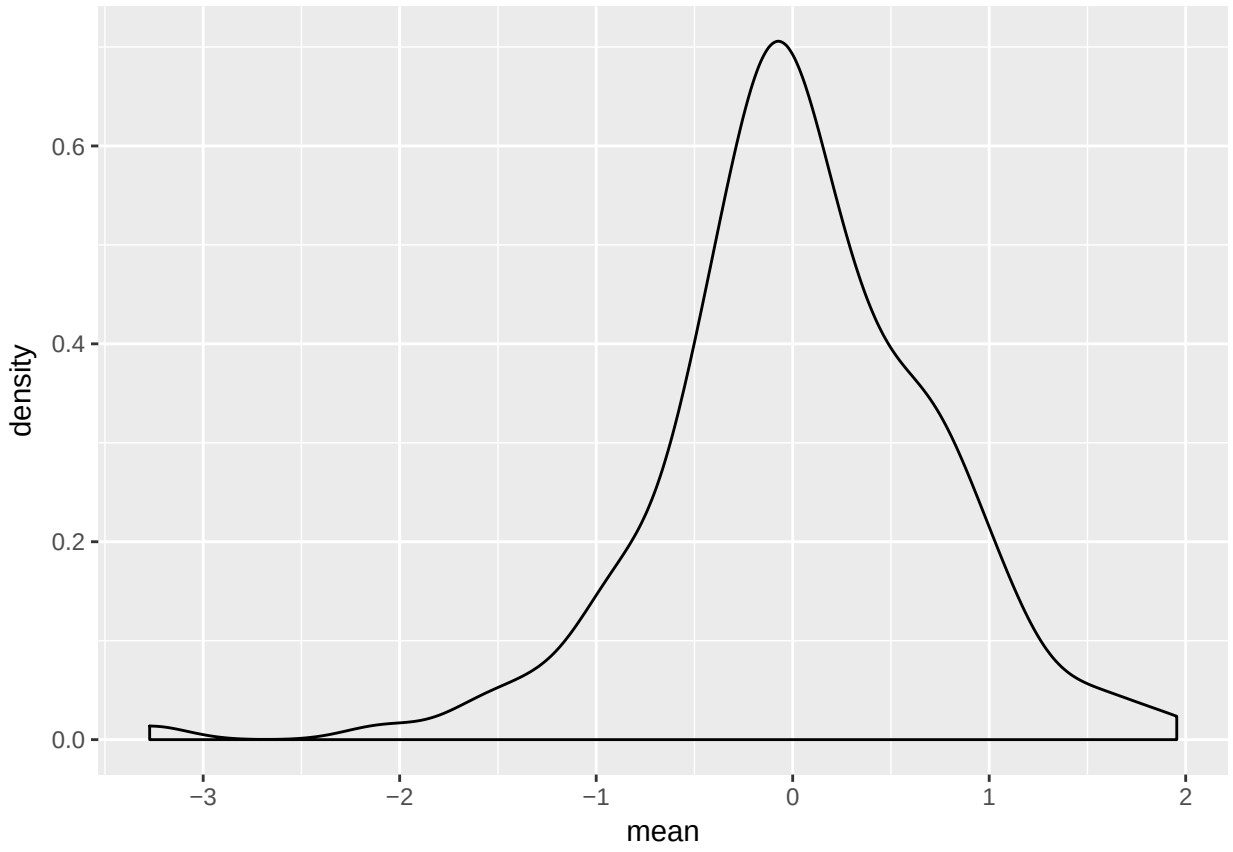
The correlation means that there is a negative correlation between the intercept and the slope, such that the lower the intercept, the higher the slope will be.

Question 6

```
model4_re.sim <- REsim(model4)
p1.gg <- model4_re.sim %>%
  filter(term == c("(Intercept)","age"))

# Intercept
ggplot(p1.gg[p1.gg$term %in% "(Intercept)",], aes(mean)) +
  geom_density()
```

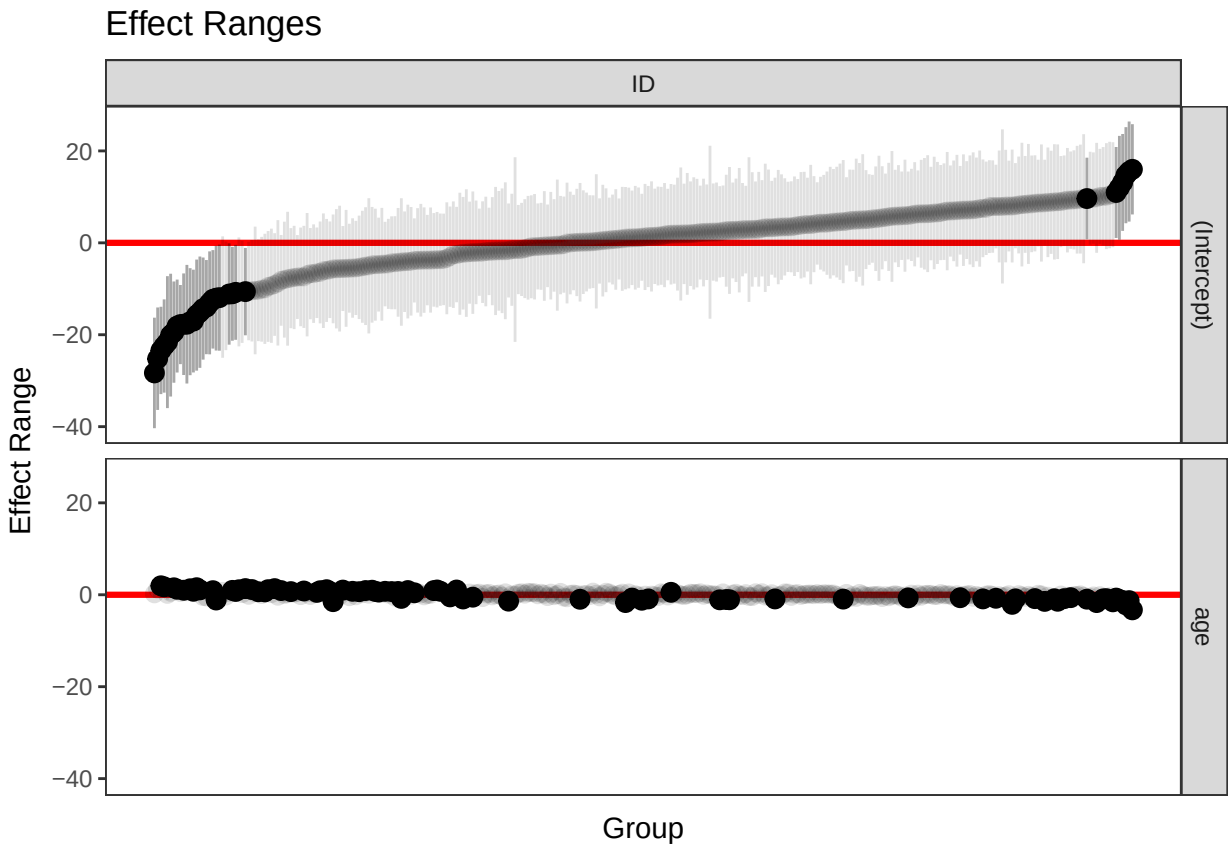



```
# Age
ggplot(p1.gg[p1.gg$term %in% "age",], aes(mean)) +
  geom_density()
```



Question 7

```
plotREsim(model4_re.sim)
```



No, there are no obvious outliers in terms of unusually large standard errors around the intercept or slopes.

Question 8

```
PT <- predictInterval(merMod=model4, newdata=data_long, level=0.9, n.sims=100, stat="mean",
                      include.resid.var=TRUE)
```

```
## Warning: executing %dopar% sequentially: no parallel backend registered
```

```
## Warning:      The following levels of ID from newdata
##  -- 86 -- are not in the model data.
##      Currently, predictions for these values are based only on the
##  fixed coefficients and the observation-level error.
```

```
## Warning:      The following levels of ID from newdata
##  -- 163, 165 -- are not in the model data.
##      Currently, predictions for these values are based only on the
##  fixed coefficients and the observation-level error.
```

```
## Warning:      The following levels of ID from newdata
##  -- 196 -- are not in the model data.
##      Currently, predictions for these values are based only on the
##  fixed coefficients and the observation-level error.
```

```
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
```

```
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
```

```
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
```

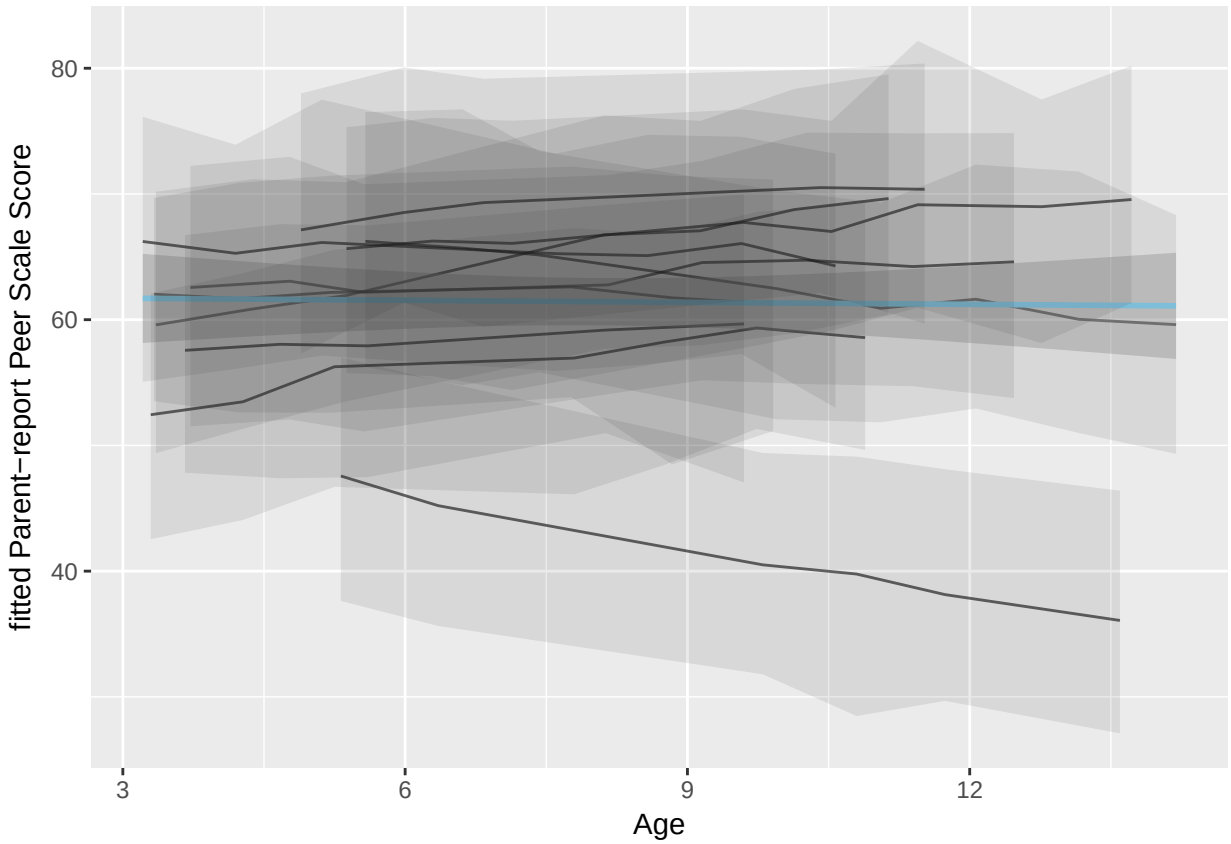
```
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
```

```
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
```

```
## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced

## Warning in rnorm(N, yhat[, x], sigmahat[x]): NAs produced
```

```
PT_plot <- data_long; PT_plot[,15:17] <- PT

#example graph using ONLY FIRST 10 participants and fitted scores with confidence bands
ggplot(PT_plot[1:80,],aes(age,fit,group=ID)) +
  geom_line(alpha=0.6) +
  stat_smooth(aes(group=1),method="lm", color="sky blue") +
  geom_ribbon(aes(ymin=lwr,ymax=upr),linetype=2,alpha=0.1) +
  labs(x="Age",y="fitted Parent-report Peer Scale Score")
```

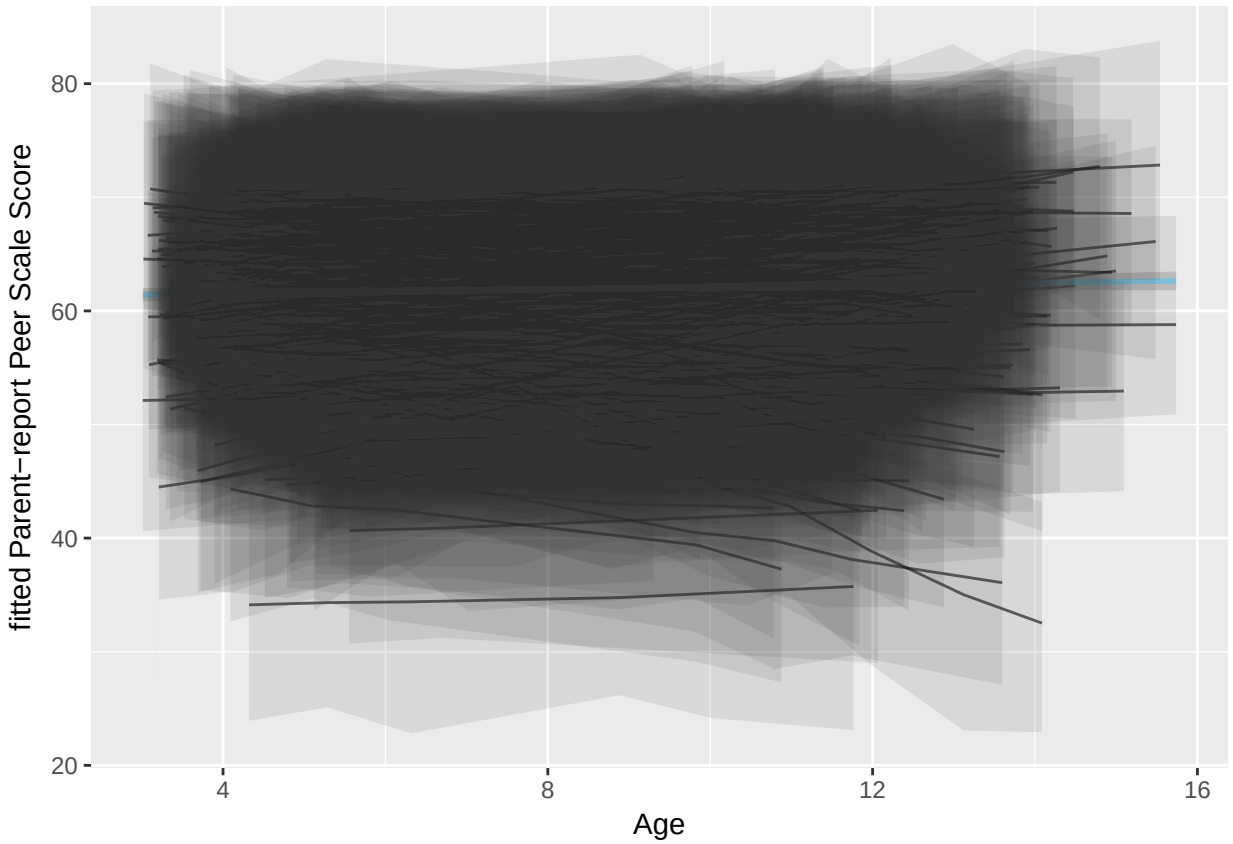
```
## Warning: Removed 13 rows containing non-finite values (stat_smooth).

## Warning: Removed 13 rows containing missing values (geom_path).
```




```
#graphed using ALL participants and fitted scores with confidence bands
ggplot(PT_plot,aes(age,fit,group=ID)) +
  geom_line(alpha=0.6) +
  stat_smooth(aes(group=1),method="lm", color="sky blue") +
  geom_ribbon(aes(ymin=lwr,ymax=upr),linetype=2,alpha=0.1) +
  labs(x="Age",y="fitted Parent-report Peer Scale Score")
```

```
## Warning: Removed 685 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 685 rows containing missing values (geom_path).
```


fitted Parent−report Peer Scale Score
80
60
40
20
4
8
12
16
Age