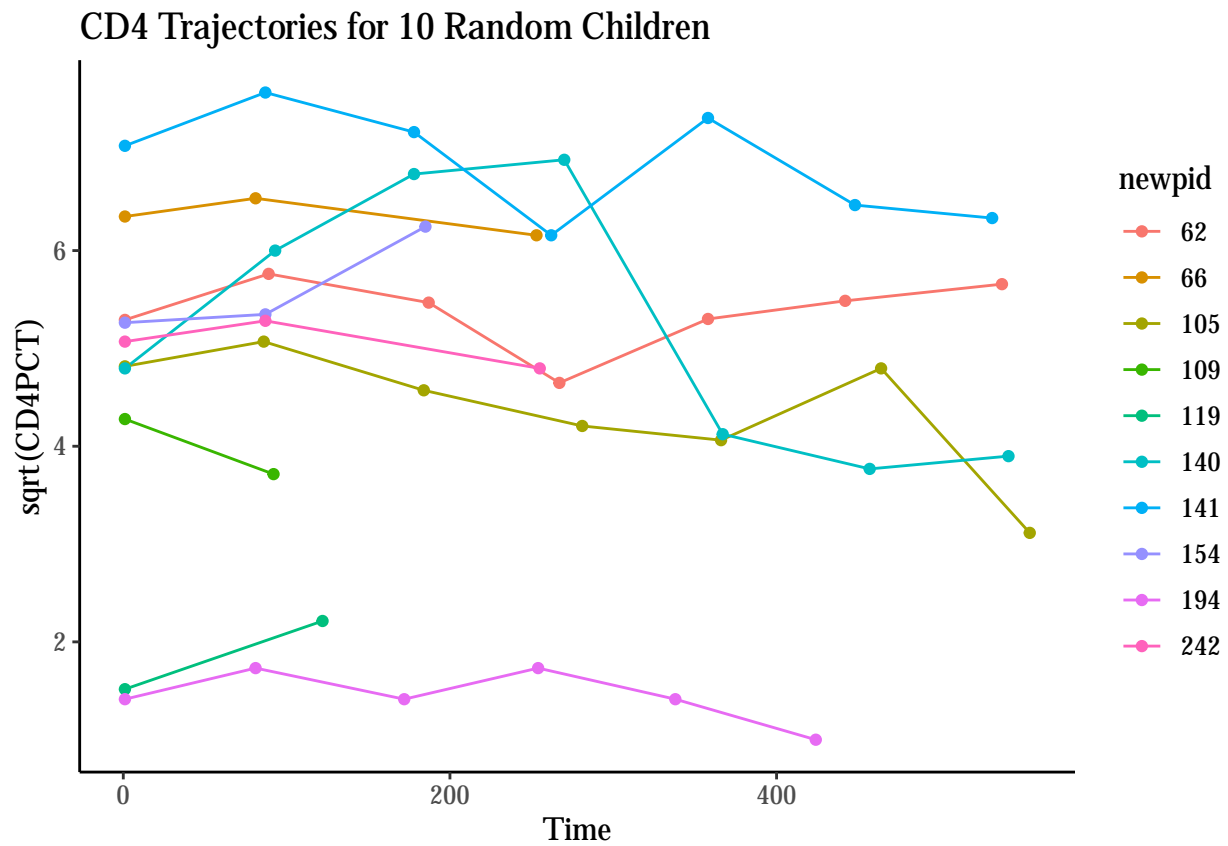# PHP 2517 Homework #1

Blain Morin
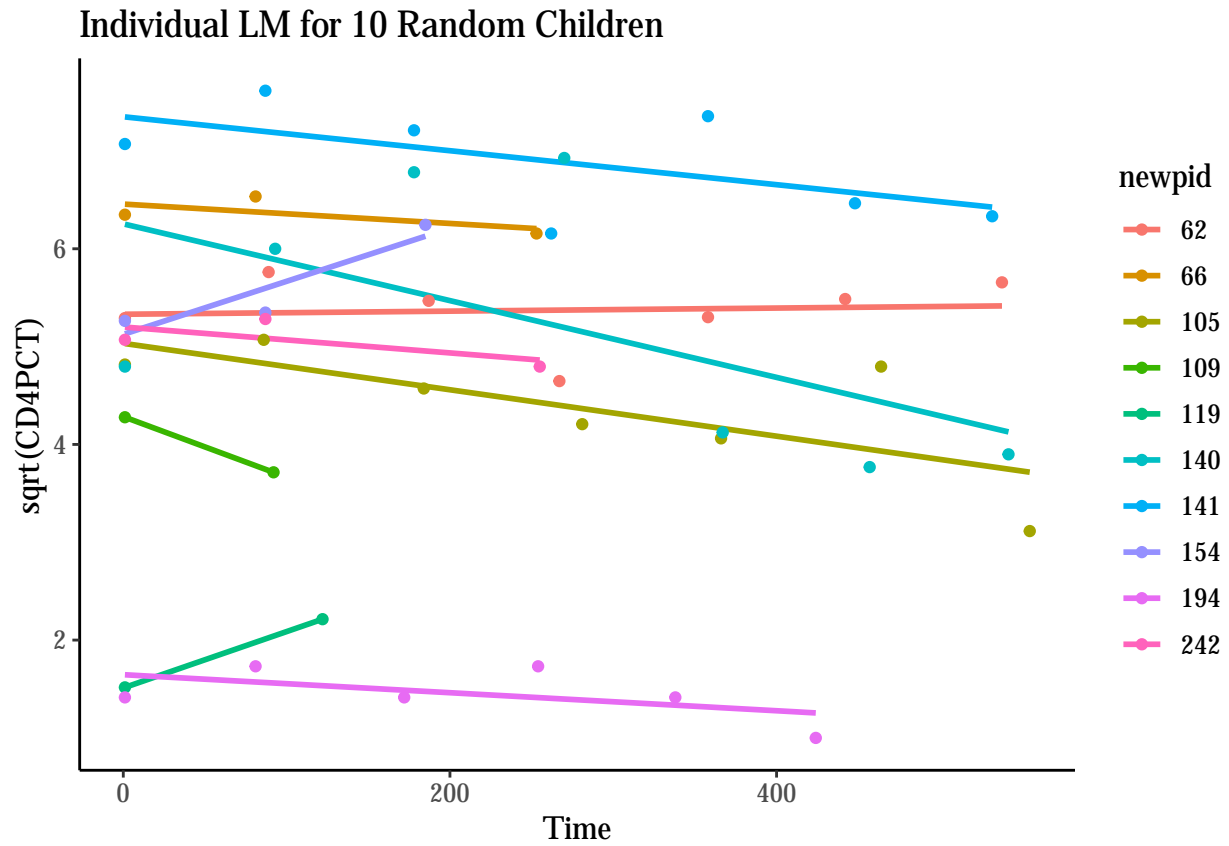
February 11, 2019

Q1 GH Chapter 11: Exercise 4

a.) Graph the outcome (the CD4 percentage, on the square root scale) for 10 children as a function of time.

CD4 Trajectories for 10 Random Children

b.) Each child's data has a time course that can be summarized by a linear fit. Estimate these lines and plot them for 10 children.

Individual LM for 10 Random Children

c.) Set up a model for the children's slopes and intercepts as a function of the treatment and age at baseline. Estimate this model using the two-step procedure–first estimate the intercept and slope separately for each child, then fit the between-child models using the point estimates from the first step.
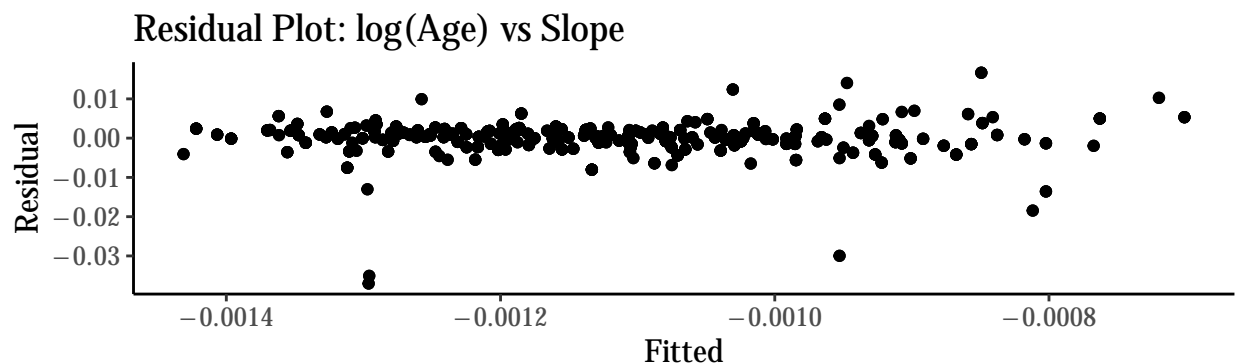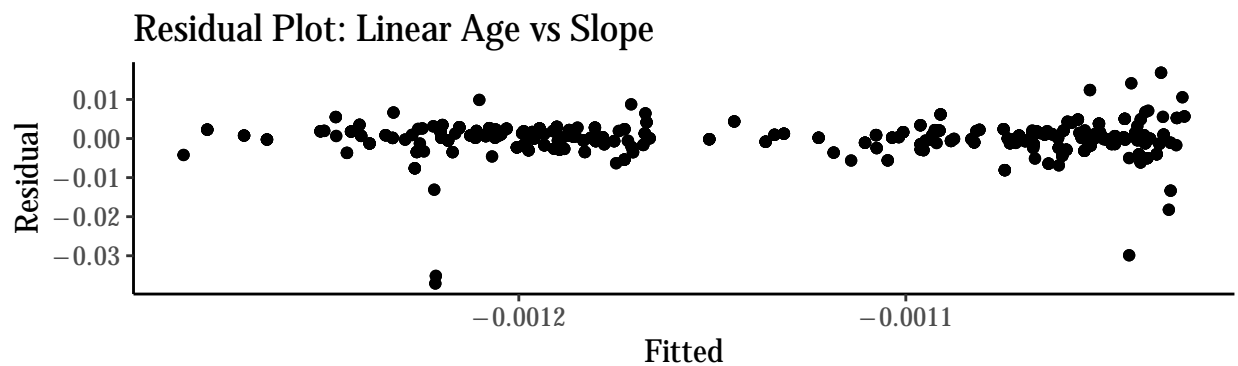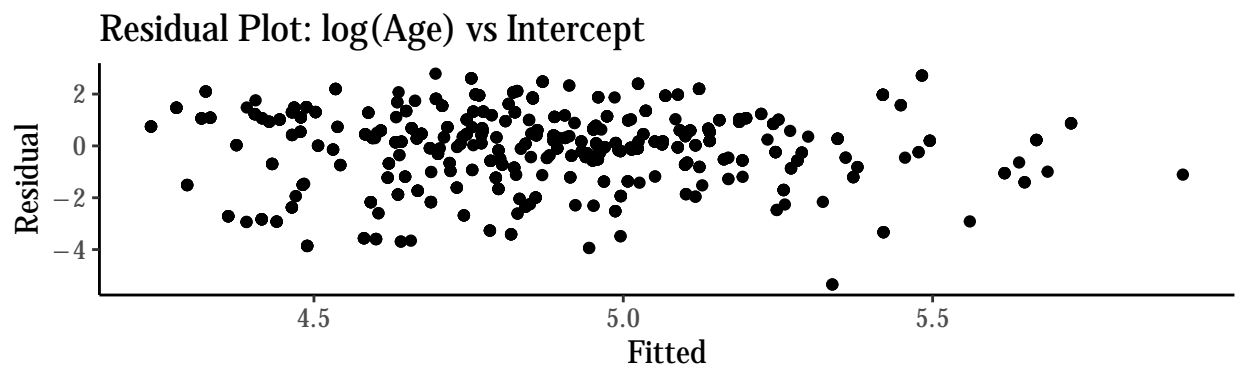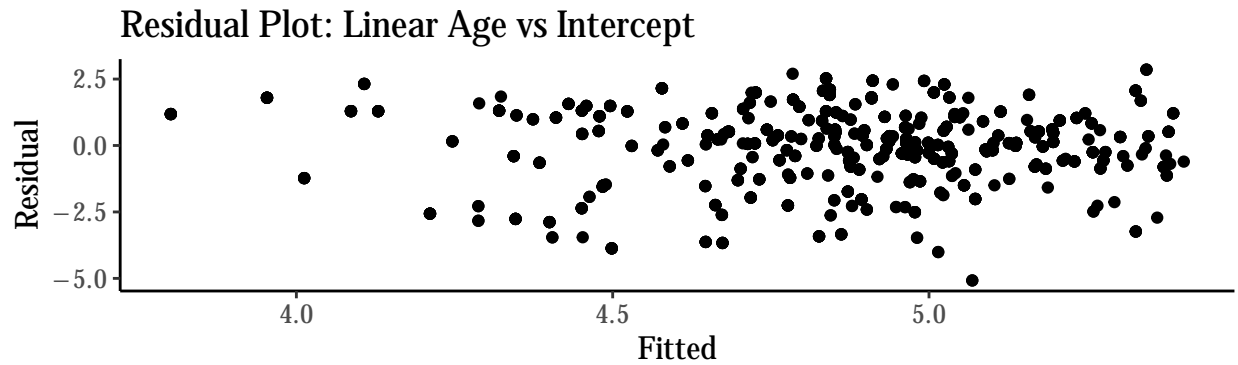
Residual Plot: Linear Age vs Intercept

Residual Plot: log(Age) vs Intercept

Residual Plot: Linear Age vs Slope

Residual Plot: log(Age) vs Slope

Table 1: Regression on Intercept and Slope

	Dependent variable:	
	beta0	beta1
	(1)	(2)
Treatment $= 2$	0.319***	-0.0001
	(0.088)	(0.0003)
log(baseage)	-0.335***	-0.0002
	(0.061)	(0.0002)
Constant	5.051***	-0.001***
	(0.087)	(0.0003)
Observations	1,055	1,028
R^2	0.040	0.001
Adjusted R^2	0.038	-0.001
Residual Std. Error	1.424 (df = 1052)	0.004 (df = 1025)
F Statistic	21.734*** (df = 2; 1052)	0.645 (df = 2; 1025)

Note: *p<0.1; **p<0.05; ***p<0.01

Individuals with only one observation do not have a slope estimate.

Q2 GH Chapter 12: Exercise 2

a.) Write a model predicting CD4 percentage as a function of time with varying intercepts across children. Fit using lmer() and interpret the coefficient for time.

$$Level1 : CD4PCT_{ij} \sim \mathcal{N}(\alpha_j + \beta time_{ij}, \sigma^2)$$

$$Level2 : \alpha_j \sim \mathcal{N}(\mu, \tau^2)$$

```
rand.int.a = lmer(sqrt(CD4PCT) ~ time + (1 | newpid), data = cd4)
texreg(rand.int.a, caption.above = TRUE,
       caption = "Random Intercept Model",
       custom.model.names = "sqrt(CD4PCT)",
       float.pos = 'H')
```

Table 2: Random Intercept Model

	sqrt(CD4PCT)
(Intercept)	4.75***
	(0.10)
time	−0.00***
	(0.00)
AIC	3120.89
BIC	3140.73
Log Likelihood	-1556.45
Num. obs.	1054
Num. groups: newpid	245
Var: newpid (Intercept)	1.99
Var: Residual	0.60

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

b.) Extend the model in (a) to include child-level predictors (that is, group-level predictors) for treatment and age at baseline. Fit using lmer() and interpret the coefficients on time, treatment, and age at baseline.

$$Level1 : CD4PCT_{ij} \sim \mathcal{N}(\alpha_j + \beta time_{ij}, \sigma^2)$$

$$Level2 : \alpha_j \sim \mathcal{N}(\mu + \gamma_0 Treat + \gamma_1 Age, \tau^2)$$

```
rand.int.b = lmer(sqrt(CD4PCT) ~ time + treatmnt + baseage + (1|newpid), data = cd4)
texreg(rand.int.b, caption.above = TRUE,
       caption = "Random Intercept Model with Child Level Predictors",
       custom.model.names = "sqrt(CD4PCT)",
       float.pos = 'H')
```

Table 3: Random Intercept Model with Child Level Predictors

	sqrt(CD4PCT)
(Intercept)	5.07***
	(0.19)
time	−0.00***
	(0.00)
treatmnt2	0.21
	(0.19)
baseage	−0.12**
	(0.04)
AIC	3121.15
BIC	3150.91
Log Likelihood	-1554.58
Num. obs.	1054
Num. groups: newpid	245
Var: newpid (Intercept)	1.92
Var: Residual	0.61

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$