Syntax:

Exploring the Anxiety variable at 3 different times.

```
> summary(moms.anxietywide$anxiety0)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
   0.00    5.00    9.50   11.69   16.00   51.00 
> summary(moms.anxietywide$anxiety1)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
  0.000   2.000   5.000   8.005  10.000  43.000      33 
> summary(moms.anxietywide$anxiety2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
  0.000   2.000   5.000   7.954  11.000  47.000      64 
> summary(indanxietyaverages$anxiety)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  0.000   3.000   7.000   9.437  13.000  51.000 
```

```
## Taking out the NAs for the anxiety variable in the wide dataset.
> widereset <-
moms.anxietywide[complete.cases(moms.anxietywide[c("anxiety0","anxiety1","anx
iety2")]),]
> summary(widereset)
       id                rmarital        educ            white       
 Min.   :  1.0   Married    :125   Min.   : 2.000   Min.   :0.0000  
 1st Qu.:301.2   Not Married: 19   1st Qu.: 6.000   1st Qu.:1.0000  
 Median :516.5   NA's       :  2   Median : 7.000   Median :1.0000  
 Mean   :495.9                     Mean   : 7.048   Mean   :0.7808  
 3rd Qu.:740.8                     3rd Qu.: 8.000   3rd Qu.:1.0000  
 Max.   :928.0                     Max.   :10.000   Max.   :1.0000  

     income         bmtrisk         disrisk         abeartox     
 Min.   :1.00   Min.   :1.000   Min.   :1.000   Min.   : 0.000  
 1st Qu.:4.00   1st Qu.:2.000   1st Qu.:2.000   1st Qu.: 1.000  
 Median :6.00   Median :2.000   Median :3.000   Median : 2.000  
 Mean   :6.24   Mean   :1.993   Mean   :2.486   Mean   : 2.671  
 3rd Qu.:9.00   3rd Qu.:2.000   3rd Qu.:3.000   3rd Qu.: 3.750  
 Max.   :9.00   Max.   :3.000   Max.   :4.000   Max.   :10.000  
                                NA's   :2                        
      aicu              cancer         time0       time1       time2  
 Min.   :0.00000   Cancer    :117   Min.   :1   Min.   :2   Min.   :3  
 1st Qu.:0.00000   No cancer: 29   1st Qu.:1   1st Qu.:2   1st Qu.:3  
 Median :0.00000                   Median :1   Median :2   Median :3  
 Mean   :0.08904                   Mean   :1   Mean   :2   Mean   :3  
 3rd Qu.:0.00000                   3rd Qu.:1   3rd Qu.:2   3rd Qu.:3  
 Max.   :1.00000                   Max.   :1   Max.   :2   Max.   :3  

     fears0           fears1           fears2          anxiety0    
 Min.   : 0.000   Min.   : 0.000   Min.   : 0.000   Min.   : 0.00  
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.: 4.00  
 Median : 4.000   Median : 4.000   Median : 4.000   Median : 9.00  
 Mean   : 4.548   Mean   : 4.848   Mean   : 4.685   Mean   :11.21  
 3rd Qu.: 6.000   3rd Qu.: 6.000   3rd Qu.: 6.000   3rd Qu.:15.00  
 Max.   :12.000   Max.   :12.000   Max.   :12.000   Max.   :51.00  
                  NA's   :1                                        
    anxiety1         anxiety2         depress0        depress1     
 Min.   : 0.000   Min.   : 0.000   Min.   : 0.00   Min.   : 0.000  
 1st Qu.: 2.000   1st Qu.: 2.000   1st Qu.: 6.00   1st Qu.: 4.000  
 Median : 5.000   Median : 5.000   Median : 9.00   Median : 7.000  
 Mean   : 8.322   Mean   : 8.021   Mean   :10.25   Mean   : 8.814  
 3rd Qu.:11.000   3rd Qu.:11.000   3rd Qu.:13.00   3rd Qu.:13.000  
 Max.   :43.000   Max.   :47.000   Max.   :38.00   Max.   :46.000  
                                                   NA's   :1       
    depress2          gvhd0            gvhd1            gvhd2      
```

```
 Min.   : 0.000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000  
 1st Qu.: 3.000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000  
 Median : 5.500   Median :0.0000   Median :0.0000   Median :0.0000  
 Mean   : 7.938   Mean   :0.1986   Mean   :0.1644   Mean   :0.1049  
 3rd Qu.:12.000   3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:0.0000  
 Max.   :45.000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000  
                                                    NA's   :3       
    infect0          infect1          infect2      
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000  
 Median :0.0000   Median :0.0000   Median :0.0000  
 Mean   :0.5479   Mean   :0.3904   Mean   :0.3077  
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.0000  
 Max.   :5.0000   Max.   :3.0000   Max.   :4.0000  
                                   NA's   :3       
## Exploring variables in the wide dataset.

> summary(widereset$anxiety1)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  0.000   2.000   5.000   8.322  11.000  43.000 
> summary(wideset$anxiety0)
Error in summary(wideset$anxiety0) : object 'wideset' not found
> summary(widereset$anxiety0)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
   0.00    4.00    9.00   11.21   15.00   51.00 
> summary(widereset$anxiety2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  0.000   2.000   5.000   8.021  11.000  47.000 
> mean(widereset$anxiety2)
[1] 8.020548

##Finding the average of individual anxiety averages.

> moms.anxietylong <- read.csv("C:/Users/James/Desktop/moms.anxietylong.csv")
>   View(moms.anxietylong)
> indanxietyaverages <-aggregate( anxiety~id+time, moms.anxietylong, mean )
> mean(indanxietyaverages$anxiety)
[1] 9.437387
> summary(indanxietyaverages$anxiety)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  0.000   3.000   7.000   9.437  13.000  51.000 
```

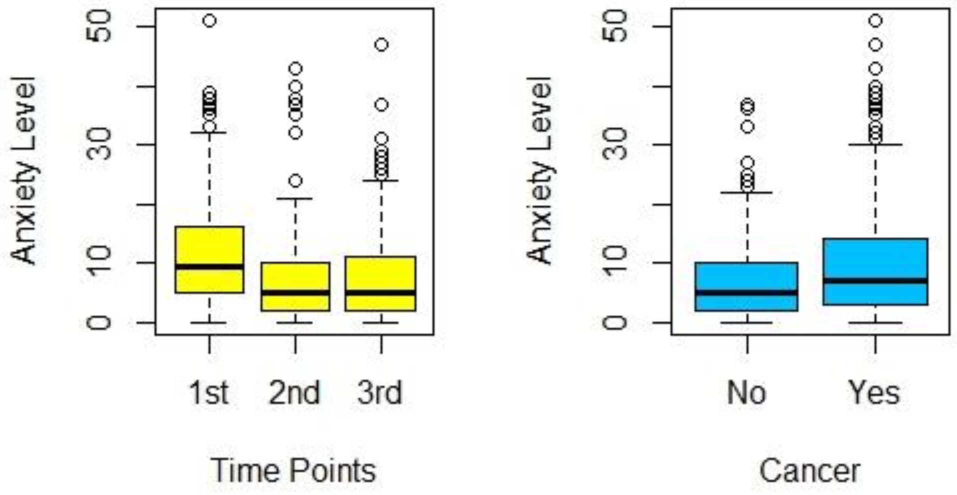
Creating a Table:

```
> dayaverages<- cbind(Baseline=mean(widereset$anxiety0),
FirstDay=mean(widereset$anxiety1), Secondday=mean(widereset$anxiety2),
Patientaverages=mean(indanxietyaverages$anxiety))
> dayaverages
     Baseline FirstDay Secondday Patientaverages
[1,] 11.21233 8.321918  8.020548        9.437387
> write.csv(dayaverages, "C:/Users/James/Desktop/Dayaverages.csv")
```

Table 1. Anxiety Levels of Mothers with Bone Marrow Transplant Children at 3 Different Times and the Average of All Participants

	Baseline	First Day	Second Day	Patient Averages
Time Average	11.212	8.322	8.021	9.437

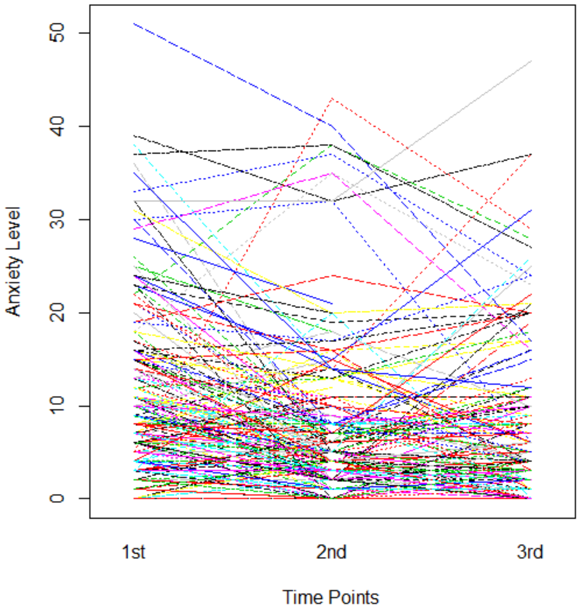
```
## 2-Way Anova Boxplots:
##Problem: Does not account for individual differences!!

## Categorizing time variable.
> is.factor(moms.anxietylong$time)
[1] FALSE
> moms.anxietylong$time <-cut(moms.anxietylong$time, br=c(0,1,2,3))
> summary(moms.anxietylong$time)
(0,1] (1,2] (2,3]
  216   216   216

> levels(moms.anxietylong$time) <- c("1st", "2nd", "3rd")

## Categorizing Cancer Variable.
> is.factor(moms.anxietylong$cancer)
[1] FALSE
> moms.anxietylong$cancer <-cut(moms.anxietylong$cancer, br=c(-1,0,1))
> summary(moms.anxietylong$cancer)
(-1,0]  (0,1]   NA's
   126     498    24
## Naming the cancer variable.
> levels(moms.anxietylong$cancer) <- c("No", "Yes")

## Creating boxplots that do not account for individual differences.
> par(mfrow=c(1,2))

> plot(moms.anxietylong$anxiety ~ moms.anxietylong$time, xlab="Time Points", ylab="Anxiety Level", col="yellow")
> plot(moms.anxietylong$anxiety ~ moms.anxietylong$cancer, xlab="Cancer", ylab="Anxiety Level", col="deepskyblue1")
```



##Spaghetti Plot

##Spaghetti Plot is not very useful either because there is a lot of overlap and we can’t really see if there are any patterns between the individuals for the three days.

```
> interaction.plot(moms.anxietylong$time, moms.anxietylong$id,
moms.anxietylong$anxiety, xlab="Time Points",ylab="Anxiety Level",
col=c(1:10), legend=F)
```



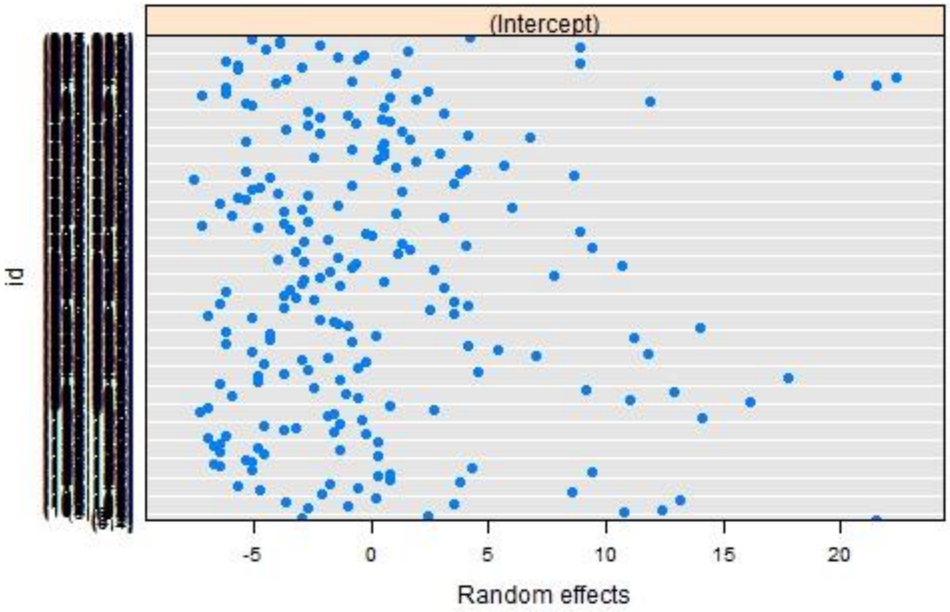
```
##Categorizing IDs.
##When I tried categorizing the IDs without labeling using the
factor() function without ##labeling, I could only get the first
fifty participants to show up in the output.
##So, I decided to use the cut() function and listed all the
intervals manually so that the ID ##variable would be
categorical.

>moms.anxietylong$id <-cut(moms.anxietylong$id,
br=c(0,1,2,3,5,101,102,103,105,106,107,109,110,112,113,114,115,1
16,117,120,121,123,126,127,128,129,131,132,134,135,136,137,138,1
39,140,141,143,144,145,146,147,148,149,150,151,152,153,154,155,1
56,157,158,301,302,304,305,306,307,308,309,310,311,312,313,314,3
16,317,318,320,322,324,326,329,330,332,333,334,335,337,339,340,3
41,343,345,346,347,348,349,350,352,355,501,504,505,508,510,511,5
13,514,516,517,519,521,522,523,524,525,526,529,532,533,535,536,5
38,539,540,541,542,543,544,701,702,703,704,705,706,707,708,709,7
10,711,712,713,717,719,720,721,722,723,724,725,726,727,728,729,7
30,731,732,733,734,736,737,739,740,741,743,744,747,748,750,751,7
```

```
52,753,755,756,757,758,759,760,761,762,763,764,765,766,767,768,7
69,770,771,772,773,774,775,901,902,903,905,908,909,910,911,912,9
13,914,915,917,918,919,920,921,922,923,924,925,926,928,929,930,9
31,932,933,934,935,936,937,938))
```

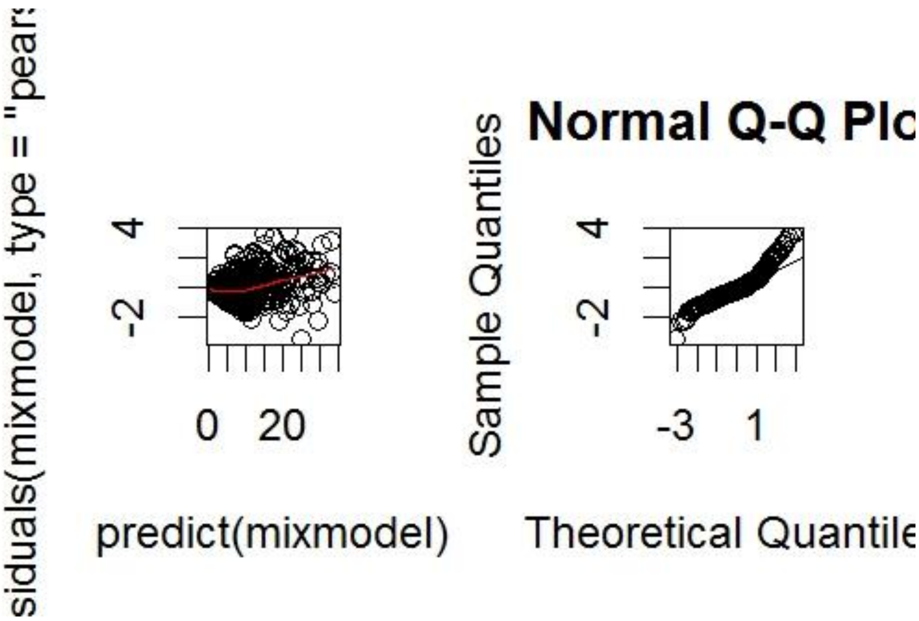
```
##Use mixed effects model:
## This is the correct method to calculate ICC because it does
not overestimate it like 2-Way Anova!
> library(nlme)

## Take out NAs from response variable.
> anxietylongreset <-
moms.anxietylong[complete.cases(moms.anxietylong[c("anxiety","cancer”)]),]
> summary(anxietylongreset$anxiety)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  0.000   3.000   7.000   9.437  13.000  51.000 
> summary(anxietylongreset$id)
    (0,1]     (1,2]     (2,3]     (3,5] (101,102] (103,105] (106,107] 
        3         3         3         3         3         3         3 
(113,114] (114,115] (116,117] (120,121] (121,123] (123,126] (126,127] 
        3         3         3         3         3         3         3 
(127,128] (128,129] (129,131] (131,132] (132,134] (134,135] (135,136] 
        3         3         3         3         3         3         3 
(136,137] (137,138] (138,139] (139,140] (140,141] (141,143] (144,145] 
        3         3         3         3         3         3         3 
(145,146] (146,147] (147,148] (148,149] (149,150] (150,151] (153,154] 
        3         3         3         3         3         3         3 
(154,155] (158,301] (301,302] (302,304] (304,305] (305,306] (306,307] 
        3         3         3         3         3         3         3 
(307,308] (308,309] (309,310] (310,311] (311,312] (312,313] (313,314] 
        3         3         3         3         3         3         3 
(314,316] (316,317] (317,318] (318,320] (320,322] (322,324] (324,326] 
        3         3         3         3         3         3         3 
(326,329] (329,330] (330,332] (333,334] (334,335] (337,339] (339,340] 
        3         3         3         3         3         3         3 
(340,341] (345,346] (346,347] (347,348] (501,504] (504,505] (508,510] 
        3         3         3         3         3         3         3 
(511,513] (513,514] (514,516] (516,517] (517,519] (519,521] (521,522] 
        3         3         3         3         3         3         3 
(523,524] (524,525] (525,526] (529,532] (532,533] (536,538] (544,701] 
        3         3         3         3         3         3         3 
(701,702] (702,703] (703,704] (704,705] (706,707] (708,709] (710,711] 
        3         3         3         3         3         3         3 
(711,712] (712,713] (713,717] (717,719] (720,721] (721,722] (722,723] 
        3         3         3         3         3         3         3 
(723,724]   (Other) 
        3       254 


##Mixed effects model output, where id is the random variable and part of
calculating the random intercept.
> mixmodel <- lme(anxiety ~ time, data=anxietylongreset, random=~1|id)
> mixmodel
Linear mixed-effects model fit by REML
  Data: anxietylongreset 
  Log-restricted-likelihood: -1901.313
  Fixed: anxiety ~ time 
(Intercept)    time2nd    time3rd 
  11.694444  -3.545332  -3.568249 
```

```
Random effects:
 Formula: ~1 | id
        (Intercept) Residual
StdDev:    6.750332 5.729195

Number of Observations: 551
Number of Groups: 216
```

```
##Mix effects model plot.
> par(cex=2)
> plot(ranef(mixmodel))
```



```
## Making the Normal Q-Q Plot to see if there is normal distribution for the
individual means (Residuals).
> par(mfrow=c(1,2), cex=1.5)
> plot(residuals(mixmodel, type="pearson") ~ predict(mixmodel))
> abline(h=0, lty=3)
> lines(lowess(x=predict(mixmodel), y=resid(mixmodel, type="pearson")),
col="red")
> qqnorm(residuals(mixmodel, type="pearson"))
> qqline(residuals(mixmodel, type="pearson"))
```



```
## Calculate ICC using the numbers in green below.
> summary(mixmodel)
Linear mixed-effects model fit by REML
 Data: anxietylongreset
       AIC      BIC    logLik
  3812.625 3834.157 -1901.313

Random effects:
 Formula: ~1 | id
        (Intercept) Residual
StdDev:    6.750332 5.729195

Fixed effects: anxiety ~ time
                Value Std.Error  DF   t-value p-value
(Intercept) 11.694444 0.6024282 333 19.412180       0
time2nd     -3.545332 0.5882570 333 -6.026843       0
time3rd     -3.568249 0.6293873 333 -5.669401       0
 Correlation:
        (Intr) tim2nd
time2nd -0.429
time3rd -0.401  0.443

Standardized Within-Group Residuals:
       Min         Q1        Med         Q3        Max
-3.4687975 -0.4814763 -0.1363871  0.3446362  3.7827155

Number of Observations: 551
Number of Groups: 216

##ICC: (6.750332)^2 / (6.750332)^2 + (5.729195)^2
##ICC: 0.5812807
## Proportion of total variation between mothers.


## Question #3.
> intervals(mixmodel)
Approximate 95% confidence intervals

 Fixed effects:
                lower      est.     upper
(Intercept) 10.509400 11.694444 12.879489
```

```
time2nd     -4.702501 -3.545332 -2.388164
time3rd     -4.806325 -3.568249 -2.330173
attr(,"label")
[1] "Fixed effects:"

 Random Effects:
  Level: id
                   lower     est.    upper
sd((Intercept)) 5.961013 6.750332 7.644168

 Within-group standard error:
   lower     est.    upper
5.312753 5.729195 6.178280
```

Intercept is positive, meaning that there is an increase in anxiety for individuals given the independent variables included in the model.

##Intercept is nowhere near zero and the confidence intervals are also do not contain zero, which means that the difference between variances is statistically significant at the p <0.05 level.

Add malignancy into the model for question 3b.

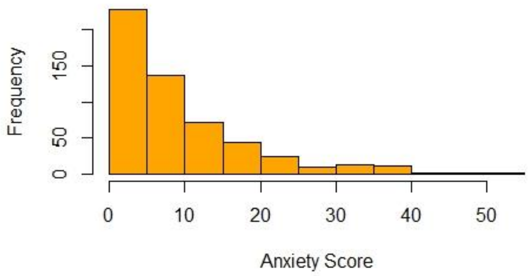
```
> mixmodel2 <- lme(anxiety ~ time + cancer, data=anxietylongreset,
random=~1|id)
> summary(mixmodel2)
Linear mixed-effects model fit by REML
 Data: anxietylongreset
       AIC      BIC    logLik
  3742.577 3768.304 -1865.288

Random effects:
 Formula: ~1 | id
        (Intercept) Residual
StdDev:    6.748217 5.698216

Fixed effects: anxiety ~ time + cancer
                Value Std.Error  DF   t-value p-value
(Intercept)  9.548133 1.2250260 332  7.794229  0.0000
time2nd     -3.512998 0.5885160 332 -5.969248  0.0000
time3rd     -3.566740 0.6281376 332 -5.678277  0.0000
cancerYes    2.783062 1.3294038 206  2.093466  0.0375
 Correlation:
          (Intr) tim2nd tim3rd
time2nd   -0.207
time3rd   -0.206  0.447
cancerYes -0.866 -0.011  0.004

Standardized Within-Group Residuals:
       Min         Q1        Med         Q3        Max
-3.5120761 -0.4783503 -0.1338862  0.3398523  3.7731861

Number of Observations: 542
Number of Groups: 208

> intervals(mixmodel2)
Approximate 95% confidence intervals

 Fixed effects:
                 lower      est.     upper
(Intercept)  7.1383416  9.548133 11.957925
```

```
time2nd     -4.6706888 -3.512998 -2.355308
time3rd     -4.8023711 -3.566740 -2.331108
cancerYes    0.1620803  2.783062  5.404043
attr(,"label")
[1] "Fixed effects:"

 Random Effects:
  Level: id
                   lower     est.    upper
sd((Intercept)) 5.946948 6.748217 7.657446

 Within-group standard error:
   lower     est.    upper 
5.281953 5.698216 6.147285 


## Adding time into the random effect.
## Close R and restart R. Convert ID into categorical. Repeat everything
abov##e before running “mixmodel” Do not convert time into categorical. Leave
tim##e continuous.

> mixmodel3 <- lme(anxiety ~ time, data=anxietylongreset, random=~time|id,
na.action=na.omit)

> summary(mixmodel3)
Linear mixed-effects model fit by REML
 Data: anxietylongreset 
       AIC      BIC    logLik
  3752.014 3777.764 -1870.007

Random effects:
 Formula: ~time | id
 Structure: General positive-definite, Log-Cholesky parametrization
            StdDev   Corr  
(Intercept) 9.425403 (Intr)
time        2.792382 -0.685
Residual    5.079775       

Fixed effects: anxiety ~ time 
                Value Std.Error  DF   t-value p-value
(Intercept) 13.287528 0.8663399 333 15.33755       0
time        -1.971096 0.3471212 333 -5.67841       0
 Correlation: 
     (Intr)
time -0.786

Standardized Within-Group Residuals:
        Min          Q1         Med          Q3         Max 
-2.97484063 -0.42484952 -0.09971079  0.29550039  3.95325574 

Number of Observations: 542
Number of Groups: 208 

> intervals(mixmodel3)
Approximate 95% confidence intervals

 Fixed effects:
                lower      est.    upper
(Intercept) 11.583339 13.287528 14.99172
time        -2.653923 -1.971096 -1.28827
attr(,"label")
[1] "Fixed effects:"

 Random Effects:
```

```
  Level: id
                              lower       est.       upper
sd((Intercept))           7.8331257  9.4254028 11.3413498
sd(time)                  1.9730117  2.7923823  3.9520288
cor((Intercept),time) -0.8100394 -0.6850355 -0.5004739

 Within-group standard error:
   lower     est.    upper 
4.564888 5.079775 5.652737 


## Using Time*cancer (interaction).

> mixmodel4 <- lme(anxiety ~ time*cancer, data=anxietylongreset,
random=~time|id, na.action=na.omit)
> summary(mixmodel4)
Linear mixed-effects model fit by REML
 Data: anxietylongreset 
       AIC      BIC    logLik
  3746.552 3780.855 -1865.276

Random effects:
 Formula: ~time | id
 Structure: General positive-definite, Log-Cholesky parametrization
            StdDev   Corr  
(Intercept) 9.460254 (Intr)
time        2.801113 -0.7  
Residual    5.080944       

Fixed effects: anxiety ~ time * cancer 
                   Value Std.Error  DF   t-value p-value
(Intercept)    12.354417 1.9268552 332  6.411700  0.0000
time           -2.646124 0.7655042 332 -3.456707  0.0006
cancerYes       1.156483 2.1583074 206  0.535829  0.5927
time:cancerYes  0.850169 0.8589233 332  0.989807  0.3230
 Correlation: 
               (Intr) time   cncrYs
time           -0.790              
cancerYes      -0.893  0.705       
time:cancerYes  0.704 -0.891 -0.790

Standardized Within-Group Residuals:
       Min         Q1        Med         Q3        Max 
-2.9625233 -0.4303656 -0.1079389  0.3150975  3.9552983 

Number of Observations: 542
Number of Groups: 208 
> intervals(mixmodel4)
Approximate 95% confidence intervals

 Fixed effects:
                    lower       est.     upper
(Intercept)     8.5640325 12.3544168 16.144801
time           -4.1519739 -2.6461238 -1.140274
cancerYes      -3.0987205  1.1564833  5.411687
time:cancerYes -0.8394496  0.8501686  2.539787
attr(,"label")
[1] "Fixed effects:"

 Random Effects:
  Level: id
                              lower       est.       upper
sd((Intercept))           7.8607531  9.4602539 11.3852200
sd(time)                  1.9783751  2.8011131  3.9659994
```

```
cor((Intercept),time) -0.8192784 -0.6996491 -0.5216523

 Within-group standard error:
    lower     est.    upper 
4.565601 5.080944 5.654456 
```

```
##Histogram for anxiety score frequency.
> hist(anxietylongreset$anxiety,  col="orange", main="", xlab="Anxiety Score", ylab="Frequency", breaks=10)
```



```
##Bar graph for Malignancy Status.
> counts <- table(anxietylongreset$cancer)
> barplot(counts, xlab="Number of Mothers Whose Children have Cancer", col=c("darkblue","red"))
```