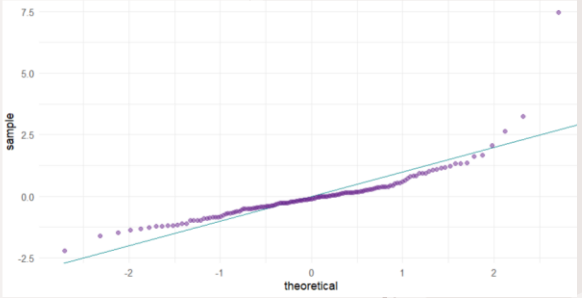
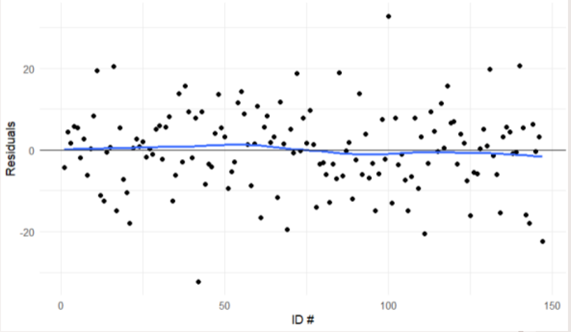
Mixed Effects Models

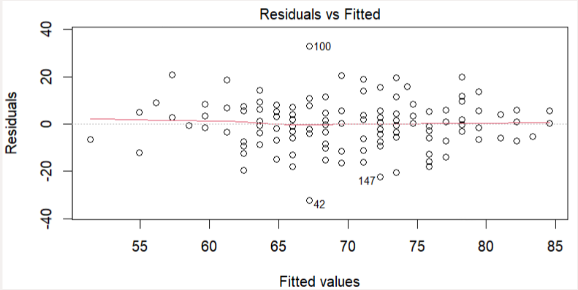
PSY 612

Ordinary Least Squares (OLS) Regression Assumptions

- Errors are **normally distributed**
- Errors are **independent**
- Errors are **equally distributed** across the range of fitted values (i.e., homogeneity of variance)



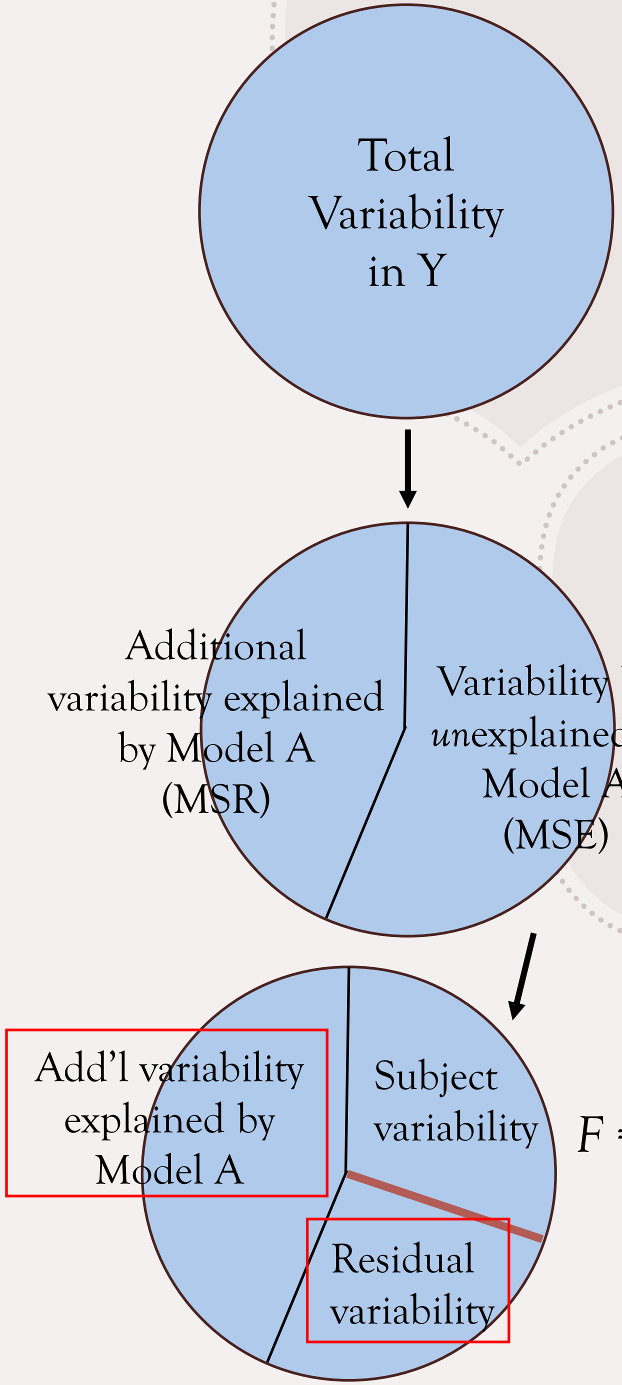




Repeated Measures Design

- A **repeated measures design** in which multiple scores are measured from the same participants *violates the independence assumption*

- Repeated measures designs are more **powerful**
  - But they also require a more sophisticated analysis tool than Ordinary Least Squares (OLS) regression that controls for non-independence among participants' scores



$$F = \frac{MSR}{MS_{Residual}}$$

Within-Subjects & Between-Subjects Predictors

- Within-Subjects Predictor
  - A variable on which scores vary within the same (or related) participants
- Between-Subjects Predictor
  - A variable on which scores vary between different participants

Mixed Effects Models

- Unlike OLS regression models that assumes all scores are independent, **mixed effects models** can be used to fit models that include within-subjects (and between-subjects) predictors

- Using the `lmer( )` function from the `lme4` package
  - Estimates the correct denominator for the *F*-statistic and *t*-statistic

- Estimates two types of effects:
  - Fixed effects
  - Random effects

Example

- Researchers conducted a study in which they deprived participants of sleep for 9 consecutive days and measured their reaction times on each day. The researcher is interested in how days of sleep deprivation predict participants' reaction times.
- The researcher is aware, though, that reaction times also tend to vary with age. They measure age to control for it in the model.
- This way, they can assess the relationship between days of sleep deprivation and reaction times while controlling for age.

ID	Reaction	Days	Age
308	249.56	0	30
308	258.7047	1	30
308	250.8006	2	30
308	321.4398	3	30
308	356.8519	4	30
308	414.6901	5	30
308	382.2038	6	30
308	290.1486	7	30
308	430.5853	8	30
308	466.3535	9	30
309	222.7339	0	24
309	205.2658	1	24
309	202.9778	2	24
309	204.707	3	24
309	207.7161	4	24
309	215.9618	5	24
309	213.6303	6	24
309	217.7272	7	24
309	224.2957	8	24
309	237.3142	9	24
310	199.0539	0	18
310	194.3322	1	18
310	234.32	2	18

Example

- The variables included in the study are:

- ID = unique ID (18 total participants)
- Reaction = reaction time in milliseconds
- Days = number of days of sleep deprivation
- Age = age of participants

- Days is a **within-subjects predictor** because it varies within participants across measurement occasions. Age is a **between-subjects predictor** because it varies between different participants.

ID	Reaction	Days	Age
308	249.56	0	30
308	258.7047	1	30
308	250.8006	2	30
308	321.4398	3	30
308	356.8519	4	30
308	414.6901	5	30
308	382.2038	6	30
308	290.1486	7	30
308	430.5853	8	30
308	466.3535	9	30
309	222.7339	0	24
309	205.2658	1	24
309	202.9778	2	24
309	204.707	3	24
309	207.7161	4	24
309	215.9618	5	24
309	213.6303	6	24
309	217.7272	7	24
309	224.2957	8	24
309	237.3142	9	24
310	199.0539	0	18
310	194.3322	1	18
310	234.32	2	18

Setting up the Data

Data must be in **long format** to analyze using the `lmer( )` function

Long format data:

- Every observation is a unique row
- Multiple rows per participant for every unique observation
- For between-subjects variables that don't change per participant across measurement occasions, list the constant score in every row

ID	Reaction	Days	Age
308	249.56	0	30
308	258.7047	1	30
308	250.8006	2	30
308	321.4398	3	30
308	356.8519	4	30
308	414.6901	5	30
308	382.2038	6	30
308	290.1486	7	30
308	430.5853	8	30
308	466.3535	9	30
309	222.7339	0	24
309	205.2658	1	24
309	202.9778	2	24
309	204.707	3	24
309	207.7161	4	24
309	215.9618	5	24
309	213.6303	6	24
309	217.7272	7	24
309	224.2957	8	24
309	237.3142	9	24
310	199.0539	0	18
310	194.3322	1	18
310	234.32	2	18

Coding & Centering the Predictors

- The same concepts we've applied to fitting models with categorical and continuous predictors throughout the course apply here.

- **Code categorical predictors**
  - Contrast codes, or
  - Dummy codes (make interpreting interaction difficult)

- **Mean-center continuous predictors**
  - For a meaningful y-intercept, and
  - To reduce multicollinearity with continuous by continuous interactions

Fixed and Random Effects

- **Fixed effects** are the same conceptually as the parameter estimates from the models we've been discussing up to this point
  - They test, on average (across all participants), how a 1-unit change in a predictor is associated with changes in the outcome while controlling for the other predictors in the model.

Model A: $Reaction_i = \beta_0 + \beta_1 * Days\_C + \beta_2 * Age\_C$

- b_0 = y-intercept; predicted reaction time for a participant with average number of days of sleep deprivation and average age
- b_1 = predicted change in reaction time per 1-unit increase in days of sleep deprivation, controlling for age
- b_2 = predicted change in reaction time per 1-unit increase in age, controlling for days of sleep depreivation

ID	Reaction	Days	Age
308	249.56	0	30
308	258.7047	1	30
308	250.8006	2	30
308	321.4398	3	30
308	356.8519	4	30
308	414.6901	5	30
308	382.2038	6	30
308	290.1486	7	30
308	430.5853	8	30
308	466.3535	9	30
309	222.7339	0	24
309	205.2658	1	24
309	202.9778	2	24
309	204.707	3	24
309	207.7161	4	24
309	215.9618	5	24
309	213.6303	6	24
309	217.7272	7	24
309	224.2957	8	24
309	237.3142	9	24
310	199.0539	0	18
310	194.3322	1	18
310	234.32	2	18

Fixed and Random Effects

- **Random effects** measure variability in the model's intercept and slopes due to individual differences among participants.
  - Individual differences: characteristic differences between participants
  - Can be measured when we have multiple scores from the same participants on a variable

- For example, we could measure how much of the variation in participants' reaction times is associated with differences in the average reaction times of each participant.
  - Regardless of sleep deprivation, some participants will have slower, or faster, reaction times compared to others

ID	Reaction	Days	Age
308	249.56	0	30
308	258.7047	1	30
308	250.8006	2	30
308	321.4398	3	30
308	356.8519	4	30
308	414.6901	5	30
308	382.2038	6	30
308	290.1486	7	30
308	430.5853	8	30
308	466.3535	9	30
309	222.7339	0	24
309	205.2658	1	24
309	202.9778	2	24
309	204.707	3	24
309	207.7161	4	24
309	215.9618	5	24
309	213.6303	6	24
309	217.7272	7	24
309	224.2957	8	24
309	237.3142	9	24
310	199.0539	0	18
310	194.3322	1	18
310	234.32	2	18

Random Effects

- There are two types of random effects that we can specify we would like the mixed effects model to estimate
  - Random intercepts
  - Random slopes

- If we were to fit a separate model representing the relationship between days of sleep deprivation and reaction times for each participant…
  - How would the intercept of each individual model vary from the average intercept across all participants?
  - How would the slopes of each individual model vary from the average slope across all participants?

ID	Reaction	Days	Age
308	249.56	0	30
308	258.7047	1	30
308	250.8006	2	30
308	321.4398	3	30
308	356.8519	4	30
308	414.6901	5	30
308	382.2038	6	30
308	290.1486	7	30
308	430.5853	8	30
308	466.3535	9	30
309	222.7339	0	24
309	205.2658	1	24
309	202.9778	2	24
309	204.707	3	24
309	207.7161	4	24
309	215.9618	5	24
309	213.6303	6	24
309	217.7272	7	24
309	224.2957	8	24
309	237.3142	9	24
310	199.0539	0	18
310	194.3322	1	18
310	234.32	2	18

Specifying Random Intercepts & Slopes in lmer()

- The lmer( ) function for fitting mixed models follows the general format:

Variable leading to non-independence among participants
- For repeated measures, this is unique participant ID

$$\text{model} \leftarrow \text{lmer}(\underbrace{Y \sim X1 + X2 + \dots X_p}_{\text{Same as lm() function}} + (\underbrace{1 + X1 \dots X_p}|\overbrace{\text{GroupingVariable}}), \text{data} = \text{data})$$

1 = random intercept for the grouping variable
$X_1 \dots X_p$ = random slopes for within-subjects predictors

Parameter Estimation

- So far in this course, we have discussed estimating the parameters of a model using ordinary least squares (OLS) regression
  - Based on OLS estimation, the best-fitting model is the one that does best at minimizing the sum of squared errors

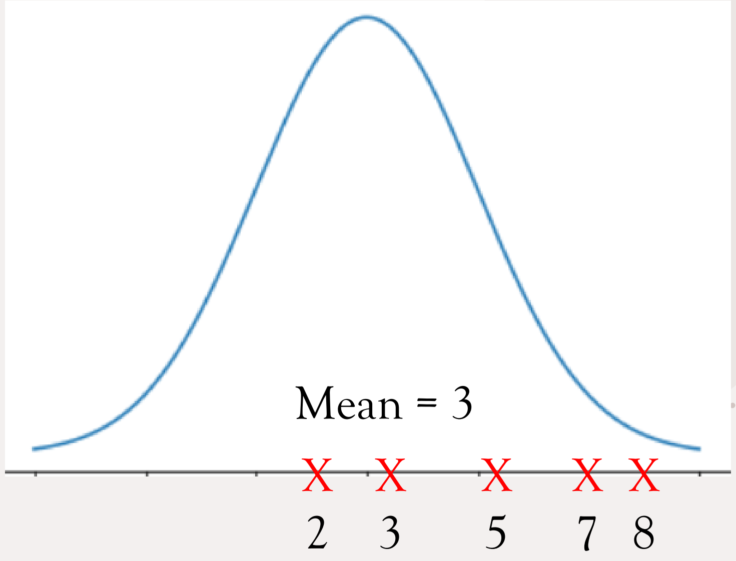
- The mixed effects model function, lmer( ), uses a different tool for finding the best-fitting model
  - **Maximum likelihood estimation (ML)**: finds the model that does best at maximizing the likelihood of the data
    - Uses an iterative process to find the best-fitting model
    - This estimation method became more accessible with advancements in computing

A Simple Example: Maximum Likelihood Estimation

- For example, say we have a sample made up of the scores: 2, 3, 5, 7, 8

- What are the chances this sample comes from a normal distribution with a mean of 3 and a SD of 1?
  - The likelihood of each data point is the height of the normal distribution with these characteristics at each value.
  - $L(2) = 0.24$, $L(3) = 0.40$, $L(5) = 0.05$, $L(7) = 0.0001$, $L(8) = 0.000001$

- Maximum likelihood estimation finds the distribution, or the model, that results in the greatest likelihood across all of the data points.



Performing Analysis in R

- Recommended to start with a model that specifies **both** random slopes and random intercepts
  - If model doesn't converge, take out the most complex random terms one at a time

- Random intercept + slopes model:

```
model_slopes <- lmer(Reaction ~ days_c + age_c + (1 + days_c|ID), data = sleep)
```

- Random intercept-only model:

```
model_ints <- lmer(Reaction ~ days_c + age_c + (1|ID), data = sleep)
```

But let's examine the random intercept-only model output first.

Random Intercept Model

- summary( ) output:

```
Random effects:
 Groups   Name        Variance Std.Dev.
 ID       (Intercept) 1036.5   32.19
 Residual              960.5   30.99
Number of obs: 180, groups:  ID, 18

Fixed effects:
            Estimate Std. Error       df t value Pr(>|t|)
(Intercept) 298.5079     7.9320  16.0000  37.633   <2e-16
days_c       10.4673     0.8042 161.0000  13.015   <2e-16
age_c         2.1910     0.8850  16.0000   2.476   0.0249
```

The model output is broken down into two sections

- **Fixed Effects**
  - We interpret these the same as we would the normal lm() output
- **Random Effects**
  - Variation in the model's intercepts and slopes driven by differences among participants

Random Intercept Model

- summary( ) output:

```
Random effects:
 Groups   Name        Variance Std.Dev.
 ID       (Intercept) 1036.5   32.19
 Residual              960.5   30.99
Number of obs: 180, groups:  ID, 18

Fixed effects:
            Estimate Std. Error       df t value Pr(>|t|)
(Intercept) 298.5079     7.9320  16.0000  37.633   <2e-16
days_c       10.4673     0.8042 161.0000  13.015   <2e-16
age_c         2.1910     0.8850  16.0000   2.476   0.0249
```

See a description of Satterthwaite's Method for df for a description of how the df are calculated.

- And a note from the author of the lmer() function on why they don't by default report them (we have to load `lmerTest` for df and *p*-values for each fixed effect to show up)

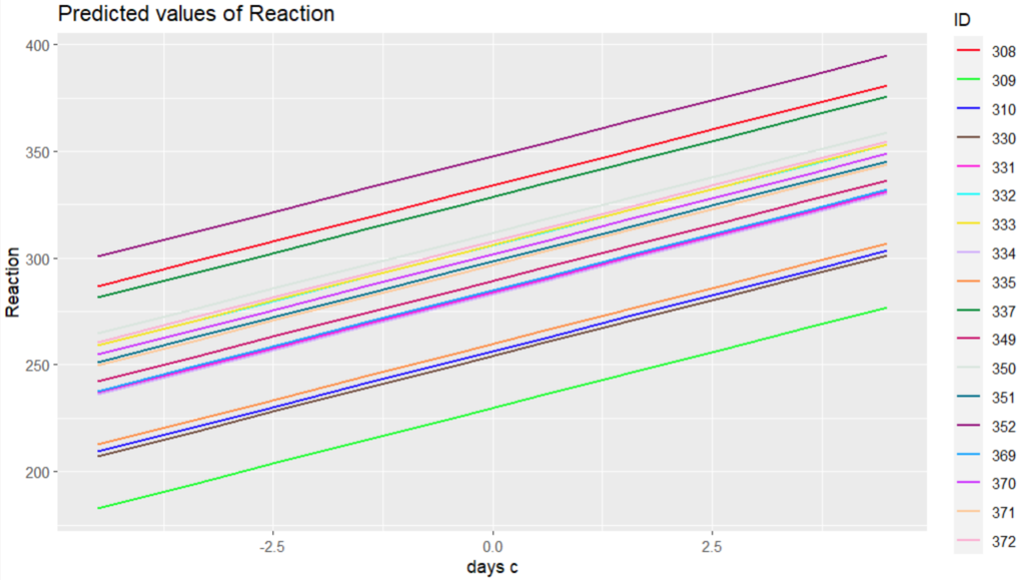
- Fixed effects
  - b_0 = 298.51, predicted reaction time for someone sleep deprived for an average number of days (M_{Days} = 4.50) and average age (M_{Age} = 30.33)
  - b_1 = the model predicts that reaction times increase by 10.47 ms for every 1-unit increase in days of sleep deprivation controlling for age
  - b_2 = the model predicts that reaction times increase by 2.19 ms for every 1-unit increase in age controlling for days of sleep deprivation

Random Intercept Model

- summary( ) output:

```
Random effects:
 Groups   Name        Variance Std.Dev.
 ID       (Intercept) 1036.5   32.19
 Residual              960.5   30.99
Number of obs: 180, groups:  ID, 18

Fixed effects:
            Estimate Std. Error       df t value Pr(>|t|)
(Intercept) 298.5079     7.9320  16.0000  37.633   <2e-16
days_c       10.4673     0.8042 161.0000  13.015   <2e-16
age_c         2.1910     0.8850  16.0000   2.476   0.0249
---
```

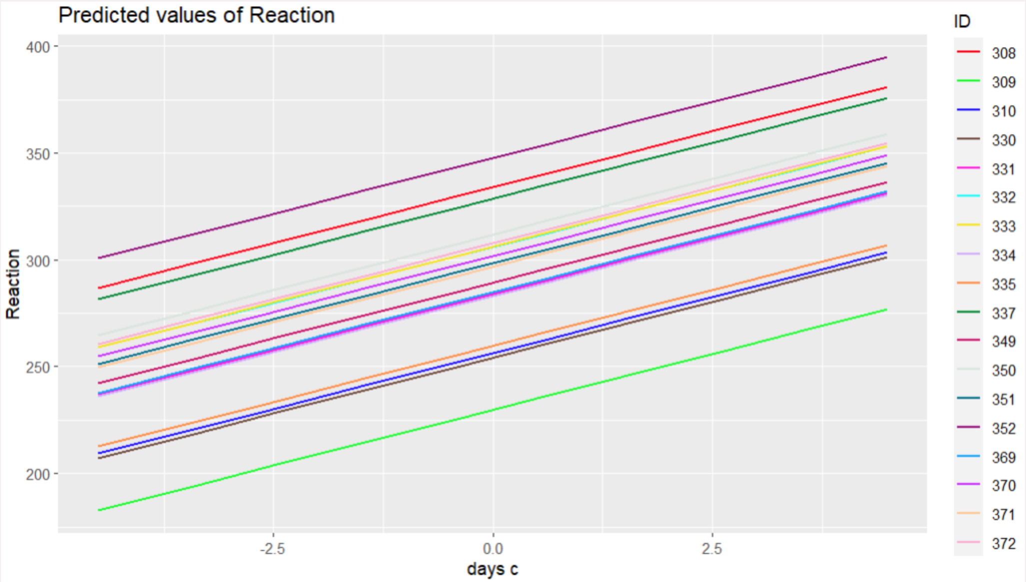


Random Effects

- **ID (Intercept)**: variance and SD of the model intercepts across participants
  - Participants' overall reaction times deviated from the overall slope on average by ± 32.19 ms
- **Residual:** The residual error left unexplained by the model

Extracting Random Intercepts

- We can extract the unique intercept for each participant by passing our model to coef( ):




```
> coef(model_ints)
$ID
    (Intercept)   days_c   age_c
308    339.1024 10.46729 2.19104
309    234.9951 10.46729 2.19104
310    261.4575 10.46729 2.19104
330    259.3752 10.46729 2.19104
331    289.1254 10.46729 2.19104
332    311.2351 10.46729 2.19104
333    311.3194 10.46729 2.19104
334    288.2215 10.46729 2.19104
335    264.8725 10.46729 2.19104
337    333.7471 10.46729 2.19104
349    294.4675 10.46729 2.19104
350    317.0014 10.46729 2.19104
351    303.5107 10.46729 2.19104
352    352.8367 10.46729 2.19104
369    290.0229 10.46729 2.19104
370    306.9839 10.46729 2.19104
371    301.9670 10.46729 2.19104
372    312.9007 10.46729 2.19104
```

Random Intercept + Slopes Model

- summary( ) output

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ID       (Intercept) 1028.26  32.066
          days_c        35.07   5.922   0.86
 Residual              654.94  25.592
Number of obs: 180, groups:  ID, 18

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept) 298.5079     7.7951 16.8033  38.294  < 2e-16
days_c       10.4673     1.5458 16.9996   6.771 3.26e-06
age_c         2.1227     0.5683 16.0002   3.736   0.0018
```

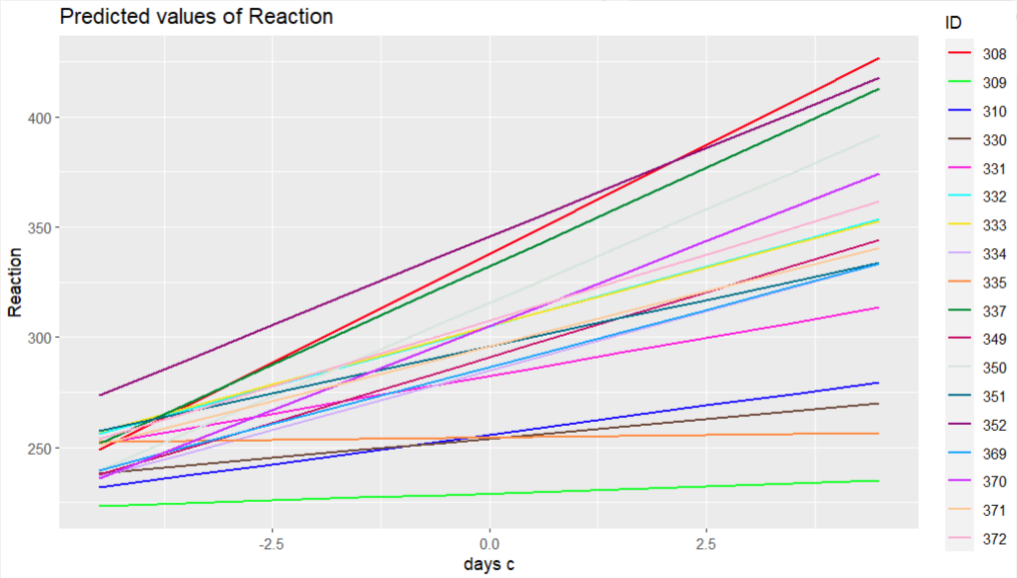
- Fixed effects
  - b_0 = 298.51, predicted reaction time for someone sleep deprived for an average number of days (M_{Days} = 4.50) and average age (M_{Age} = 30.33)
  - b_1 = the model predicts that reaction times increase by 10.47 ms for every 1-unit increase in days of sleep deprivation controlling for age
  - b_2 = the model predicts that reaction times increase by 2.12 ms for every 1-unit increase in age controlling for days of sleep deprivation

Random Intercept + Slopes Model

- summary( ) output

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ID       (Intercept) 1028.26  32.066
          days_c        35.07   5.922   0.86
 Residual              654.94  25.592
Number of obs: 180, groups:  ID, 18

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept) 298.5079     7.7951 16.8033  38.294  < 2e-16
days_c       10.4673     1.5458 16.9996   6.771 3.26e-06
age_c         2.1227     0.5683 16.0002   3.736   0.0018
---
```

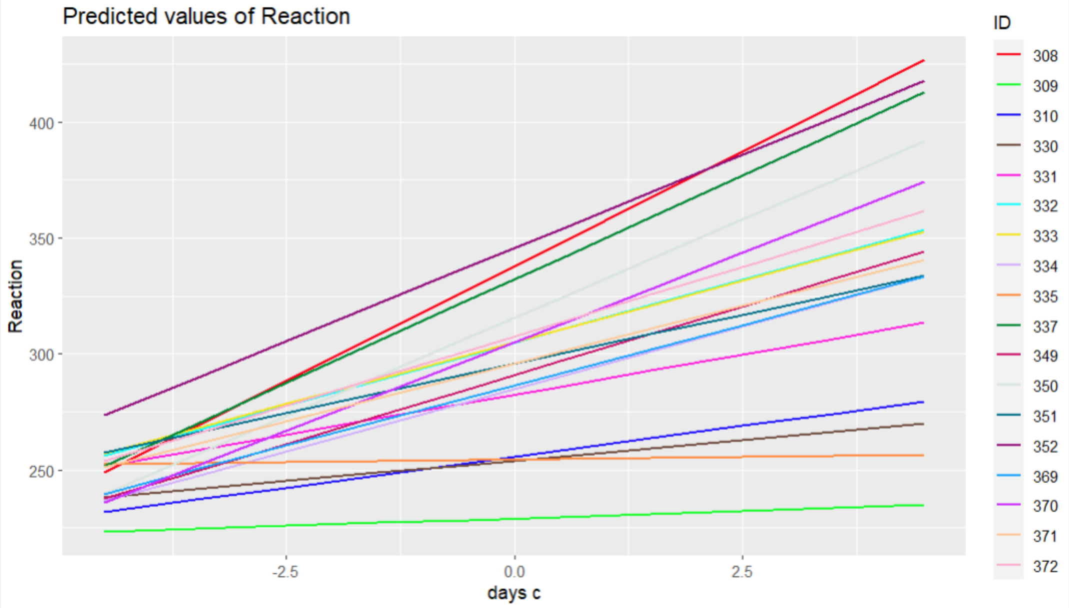


Random Effects

- **ID (Intercept):** variance and SD of the model intercepts across participants
  - Participant's overall reaction times deviated from the overall intercept on average by ± 32.07 ms
- **days_c**: variance and SD of the model slopes across participants
  - The slope for the relationship between days_c and reaction varied from the average slope on average by ± 5.92 ms/days units
- **Corr**: the correlation between the random intercepts and random slopes
  - Positive correlation indicates that people with higher intercepts tend to have larger slopes

Extracting Random Intercepts + Slopes

- We can extract the unique intercept & slope for each participant by passing our model to coef( ):




```
> coef(model_slopes)
$ID
    (Intercept)     days_c    age_c
308    342.6724 19.7577367 2.122705
309    234.0779  1.2770086 2.122705
310    260.6474  5.3143279 2.122705
330    259.0252  3.5538737 2.122705
331    287.4768  6.8922682 2.122705
332    309.8446 10.8483804 2.122705
333    309.9313 10.6310058 2.122705
334    290.0165 10.7677651 2.122705
335    259.4497  0.4810549 2.122705
337    337.2326 17.9031421 2.122705
349    295.8435 11.8618489 2.122705
350    320.5826 16.8981783 2.122705
351    300.6755  8.5042312 2.122705
352    350.6558 16.0382068 2.122705
369    291.4592 10.3995302 2.122705
370    310.1031 15.3436839 2.122705
371    300.8767  9.9718901 2.122705
372    312.5713 11.9670145 2.122705
```

Reporting the Results

- Typically, it's the **fixed effects** that researchers are most interested in testing the pattern and significance of since these evaluate the phenomenon that they are studying

- The random effects are simply included to control for non-independence among the participants' scores in the sample
  - But still need to be clear and what random effects were included in the model

Reporting the Results

- The purpose of this study was to examine how days of sleep deprivation and age predict people's reaction times. Since the same participants' reaction times were measured across nine days, a linear mixed effects model was conducted with scores nested in participant ID. The intercepts across participants were allowed to randomly vary as was the slope for the relationship between days of sleep deprivation and reaction time. Table 1 displays the variance and standard deviation in the model's random intercepts and slopes.

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ID       (Intercept) 1028.26  32.066
          days_c        35.07   5.922   0.86
 Residual              654.94  25.592
Number of obs: 180, groups:  ID, 18

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept) 298.5079     7.7951 16.8033  38.294  < 2e-16
days_c       10.4673     1.5458 16.9996   6.771 3.26e-06
age_c         2.1227     0.5683 16.0002   3.736   0.0018
---
```

- Specifically, days of sleep deprivation was a significant, positive predictor of people's reaction times while controlling for age, $t(17.00) = 6.77$, $p < .001$. With each additional day of sleep deprivation, reaction times tended to increase by 10.47 ms.
- Age was also a significant, positive predictor of people's reaction times while controlling for days of sleep deprivation, $t(16.00) = 3.74$, $p = .002$. With every year increase in age, reaction times tended to increase by 2.12 ms.