Repeated Measures (RM-ANOVA)

Andrew Zieffler

Educational Psychology

University of Minnesota

Driven to Discover[SM]

Read in the minneapolis.csv data

```
## Read in the data
> mpls = read.csv("http://www.tc.umn.edu/~zief0002/Data/minneapolis.csv")
```

Packages Needed

- **ez**
- **ggplot2**
- **reshape2**

	studentID	read.5	read.6	read.7	read.8	atRisk	female	minority	ell	sped	att
1	1	172	185	179	194	1	1	1	0	0	0.94
2	2	200	210	209	NA	1	1	1	0	0	0.91
3	3	191	199	203	215	1	0	1	0	0	0.97
4	4	200	195	194	NA	1	1	1	0	0	0.88
5	5	207	213	212	213	1	1	1	0	0	0.85
6	6	191	189	206	195	1	0	1	0	0	0.90
7	7	199	208	213	218	1	0	1	1	0	0.97
8	8	191	194	194	NA	1	1	1	1	1	0.97
9	9	149	154	174	177	1	1	1	0	1	0.97
10	10	200	212	213	NA	1	1	1	0	0	0.96
11	11	218	231	233	239	1	1	0	0	0	0.98
12	12	228	232	248	246	1	1	0	0	1	0.96
13	13	228	236	228	239	0	1	0	0	0	0.99
14	14	199	210	225	235	0	0	1	1	0	0.99
15	15	218	223	236	NA	0	1	0	0	0	1.00
16	16	228	226	234	227	0	0	0	0	0	0.97
17	17	201	210	208	219	0	0	0	0	0	0.98
18	18	218	220	217	221	0	0	0	0	0	1.00
19	19	215	216	221	NA	0	1	0	0	0	0.96
20	20	204	215	219	214	0	1	1	0	0	0.95
21	21	237	241	243	NA	0	0	0	0	0	0.98
22	22	219	233	236	NA	0	1	1	0	0	0.96

Most repeated measures data are entered in the wide-format

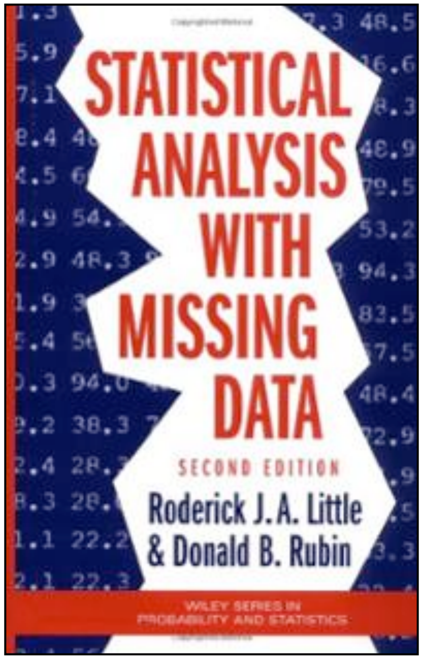
Each of the repeated measures is in its own column.

Missing Data

Missing data is a problem for most data analysis

There are several ways to deal with missing data.

- Remove cases with missing data (*case-wise deletion*)
- Replace missing values with an actual value (*imputation*)



No matter how you treat missing data, there may be a problem for your inferences....Are the cases you deleted an unbiased sample of the full set of cases?

Missing data is problematic for RM-ANOVA.

We will delete *any* cases with missing data.

	studentID	read.5	read.6	read.7	read.8	atRisk	female	minority	ell	sped	att
1	1	172	185	179	194	1	1	1	0	0	0.94
~~2~~	~~2~~	~~200~~	~~210~~	~~209~~	~~NA~~	~~1~~	~~1~~	~~1~~	~~0~~	~~0~~	~~0.91~~
3	3	191	199	203	215	1	0	1	0	0	0.97
~~4~~	~~4~~	~~200~~	~~195~~	~~194~~	~~NA~~	~~1~~	~~1~~	~~1~~	~~0~~	~~0~~	~~0.88~~
5	5	207	213	212	213	1	1	1	0	0	0.85
6	6	191	189	206	195	1	0	1	0	0	0.90
7	7	199	208	213	218	1	0	1	1	0	0.97
~~8~~	~~8~~	~~191~~	~~194~~	~~191~~	~~NA~~	~~1~~	~~1~~	~~1~~	~~1~~	~~1~~	~~0.97~~
9	9	149	154	174	177	1	1	1	0	1	0.97
~~10~~	~~10~~	~~200~~	~~212~~	~~213~~	~~NA~~	~~1~~	~~1~~	~~1~~	~~0~~	~~0~~	~~0.96~~
11	11	218	231	233	239	1	1	0	0	0	0.98
12	12	228	232	248	246	1	1	0	0	1	0.96
13	13	228	236	228	239	0	1	0	0	0	0.99
14	14	199	210	225	235	0	0	1	1	0	0.99
~~15~~	~~15~~	~~218~~	~~223~~	~~236~~	~~NA~~	~~0~~	~~1~~	~~0~~	~~0~~	~~0~~	~~1.00~~
16	16	228	226	234	227	0	0	0	0	0	0.97
17	17	201	210	208	219	0	0	0	0	0	0.98
18	18	218	220	217	221	0	0	0	0	0	1.00
~~19~~	~~19~~	~~215~~	~~216~~	~~221~~	~~NA~~	~~0~~	~~1~~	~~0~~	~~0~~	~~0~~	~~0.96~~
20	20	204	215	219	214	0	1	1	0	0	0.95
~~21~~	~~21~~	~~237~~	~~241~~	~~243~~	~~NA~~	~~0~~	~~0~~	~~0~~	~~0~~	~~0~~	~~0.98~~
~~22~~	~~22~~	~~219~~	~~233~~	~~236~~	~~NA~~	~~0~~	~~1~~	~~1~~	~~0~~	~~0~~	~~0.96~~

Use indexing and `complete.cases()` to remove any cases with NAs

```
> mpls2 = mpls[complete.cases(mpls), ]

> mpls2

   studentID read.5 read.6 read.7 read.8 atRisk female minority ell sped  att
1          1    172    185    179    194      1      1        1   0    0 0.94
3          3    191    199    203    215      1      0        1   0    0 0.97
5          5    207    213    212    213      1      1        1   0    0 0.85
6          6    191    189    206    195      1      0        1   0    0 0.90
7          7    199    208    213    218      1      0        1   1    0 0.97
9          9    149    154    174    177      1      1        1   0    1 0.97
11        11    218    231    233    239      1      1        0   0    0 0.98
12        12    228    232    248    246      1      1        0   0    1 0.96
13        13    228    236    228    239      0      1        0   0    0 0.99
14        14    199    210    225    235      0      0        1   1    0 0.99
16        16    228    226    234    227      0      0        0   0    0 0.97
17        17    201    210    208    219      0      0        0   0    0 0.98
18        18    218    220    217    221      0      0        0   0    0 1.00
20        20    204    215    219    214      0      1        1   0    0 0.95
```

We removed 8 of the original 22 cases (36% were removed!).

Is there an effect of time (i.e., a longitudinal effect) on reading scores?

$$H_0 : \mu_{\text{Grade 5}} = \mu_{\text{Grade 6}} = \mu_{\text{Grade 7}} = \mu_{\text{Grade 8}}$$

Examine this descriptively before any testing...

```
## Examine the means at each measurement wave
> summary(mpls2[2:5])

    read.5          read.6          read.7          read.8     
 Min.   :149.0   Min.   :154.0   Min.   :174.0   Min.   :177.0  
 1st Qu.:193.0   1st Qu.:201.2   1st Qu.:206.5   1st Qu.:213.2  
 Median :202.5   Median :211.5   Median :215.0   Median :218.5  
 Mean   :202.4   Mean   :209.1   Mean   :214.2   Mean   :218.0  
 3rd Qu.:218.0   3rd Qu.:224.5   3rd Qu.:227.2   3rd Qu.:233.0  
 Max.   :228.0   Max.   :236.0   Max.   :248.0   Max.   :246.0  
```

The sample means suggest an increase in reading scores over time, on average

It would be great to plot this as well.

To plot the reading scores over time using **ggplot**, we need to reshape the data from the wide format to the long format

long-format data

In the long-formatted data, each row is not a different student, but a different student/grade combination.

Each student is associated with multiple rows

This is similar to the long format of the NBA data where teams (groups) were associated with multiple rows

In the NBA data, players (each row) were nested in teams (which had multiple rows)

In repeated measures data, time points (each row) are nested in the subjects (having multiple rows)...subjects are the groups in these models!

	studentID	variable	value
1	1	read.5	172
2	1	read.6	185
3	1	read.7	179
4	1	read.8	194
5	3	read.5	191
6	3	read.6	199
7	3	read.7	203
8	3	read.8	215
9	5	read.5	207
10	5	read.6	213
	⋮	⋮	⋮
47	17	read.7	208
48	17	read.8	219
49	18	read.5	218
50	18	read.6	220
51	18	read.7	217
52	18	read.8	221
53	20	read.5	204
54	20	read.6	215
55	20	read.7	219
56	20	read.8	214

one column for time predictor

one column for outcome

Reshape Wide to Long Data

```
## Use the reshape2 package
> library(reshape2)

## Melt the data to the long format
> mplsLong = melt(
    mpls2,
    id = c("studentID"),
    measure = c("grade.5", "grade.6", "grade.7", "grade.8")
    )
```

The id= argument

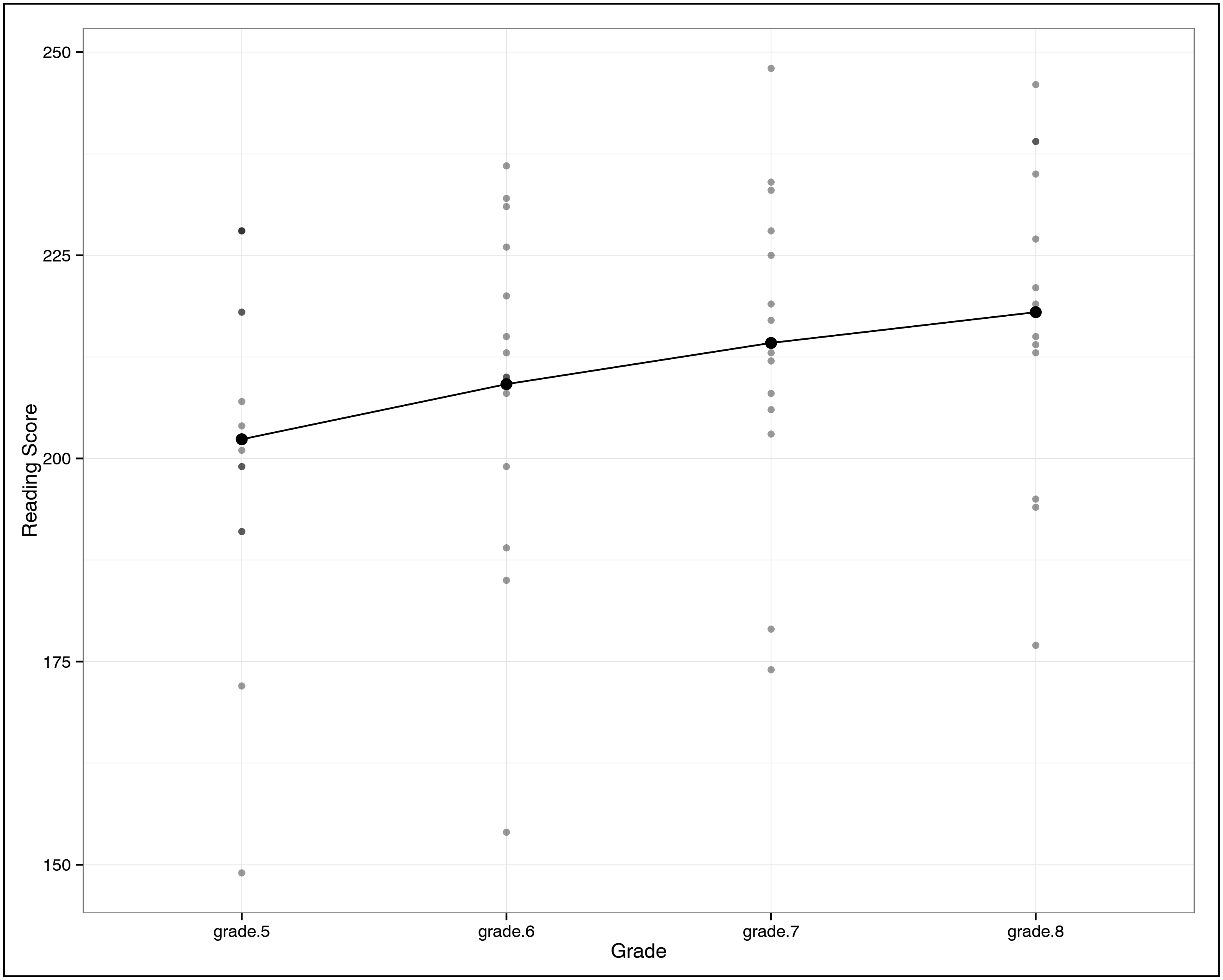
keep these variables 'as is'

The measure= argument

Change these variables into two new ones...variable and value

```
> head(mplsLong)

  studentID variable value
1         1   read.5   172
2         3   read.5   191
3         5   read.5   207
4         6   read.5   191
5         7   read.5   199
6         9   read.5   149
```

```
## Rename the variable and value columns
> names(mplsLong)[2] = "grade"
> names(mplsLong)[3] = "read"

> head(mplsLong)

  studentID  grade read
1         1 read.5  172
2         3 read.5  191
3         5 read.5  207
4         6 read.5  191
5         7 read.5  199
6         9 read.5  149


## Rename the levels of the grade variable
> levels(mplsLong$grade)

[1] "read.5" "read.6" "read.7" "read.8"

> levels(mplsLong$grade)[1] = "grade.5"
> levels(mplsLong$grade)[2] = "grade.6"
> levels(mplsLong$grade)[3] = "grade.7"
> levels(mplsLong$grade)[4] = "grade.8"
```



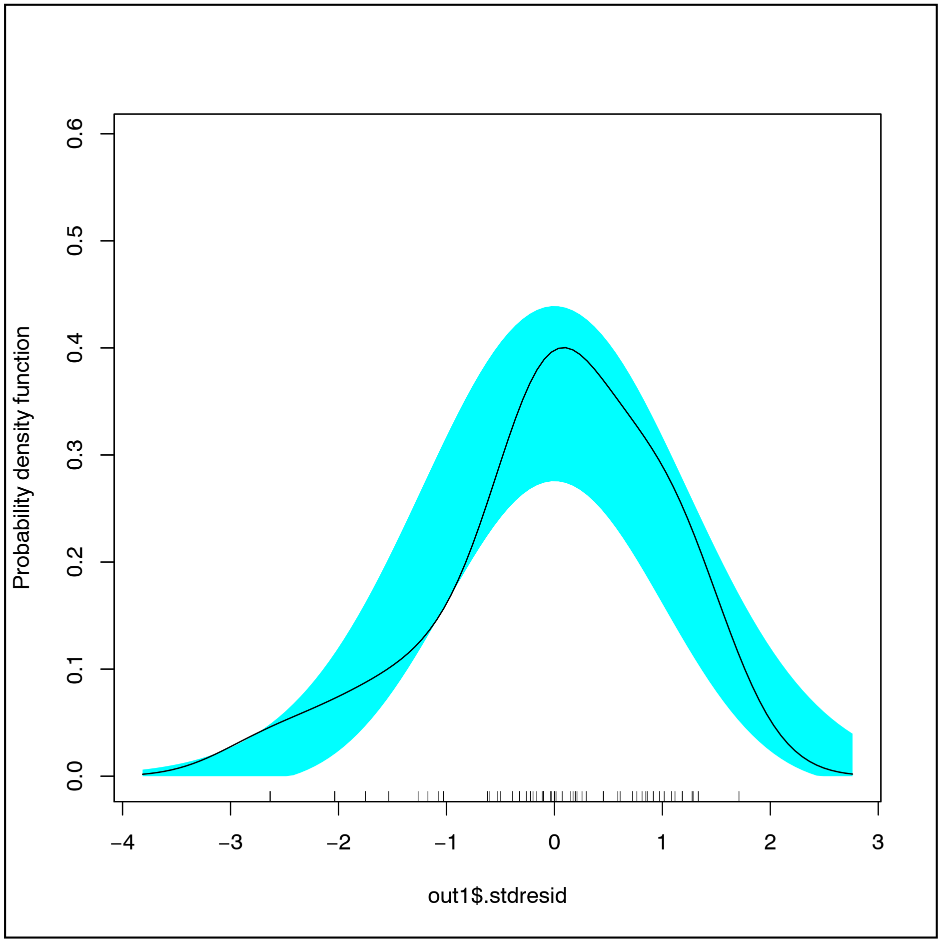
The plot shows the same increasing trend that we observed in the summaries

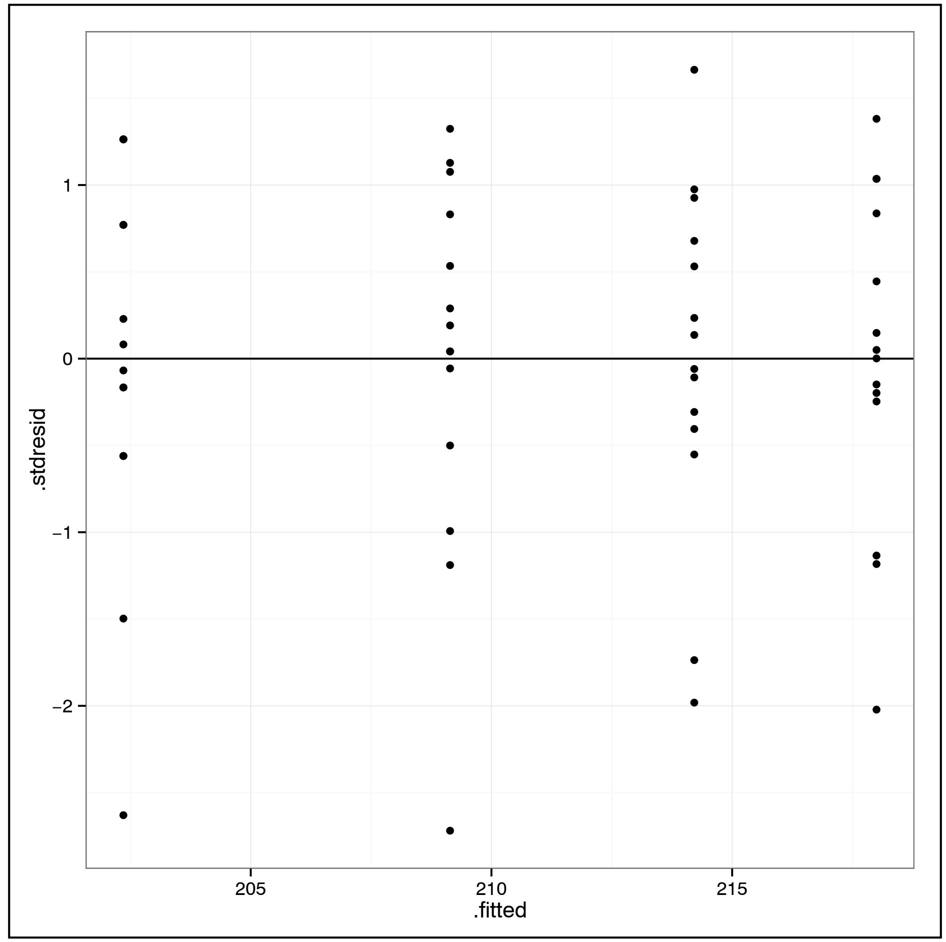
ANALYZING THE DATA UNDER THE ASSUMPTION OF INDEPENDENCE

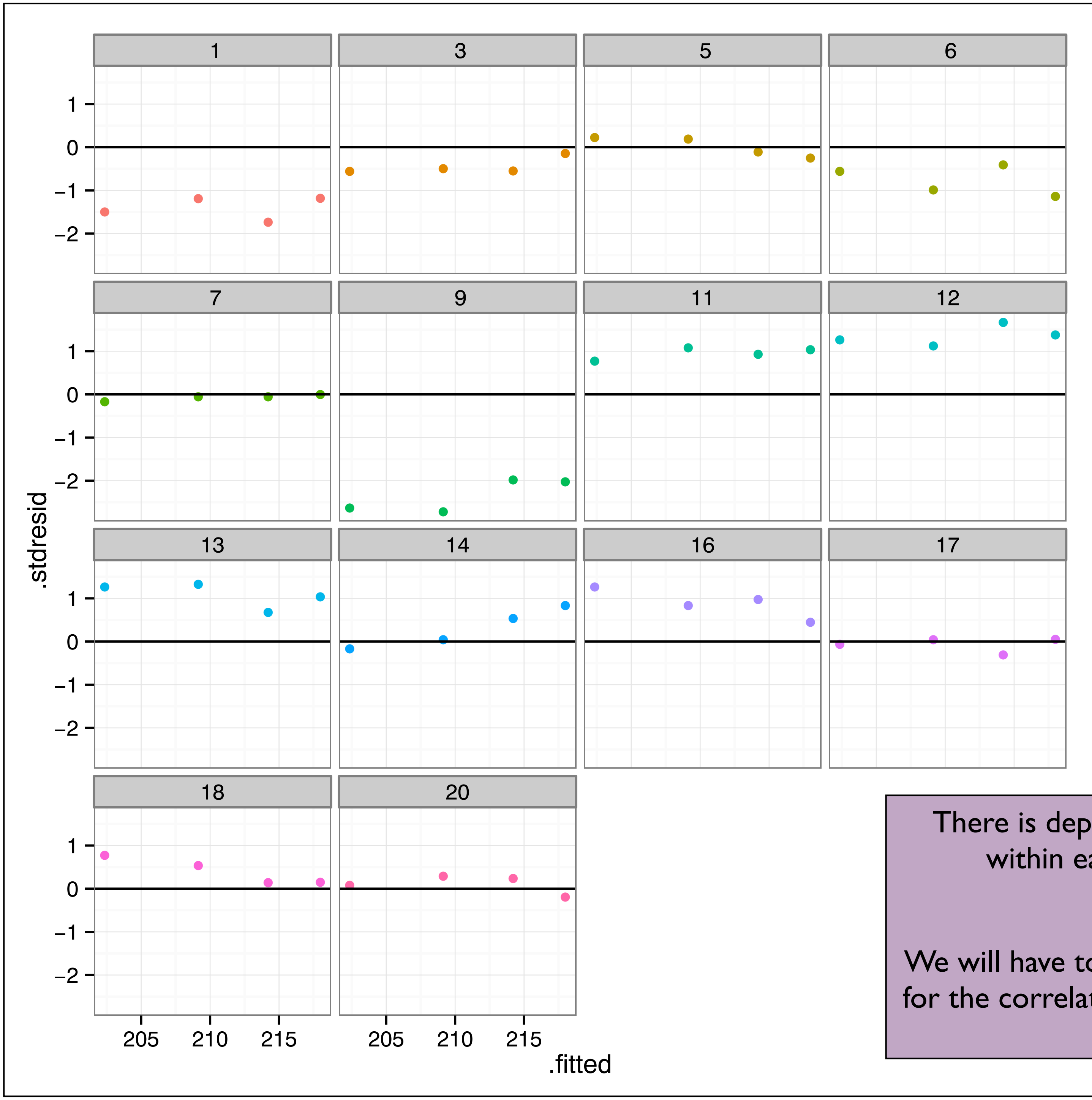
```
## Fit the regression model
> lm.1 = lm(read ~ grade, data = mplsLong)

## Examine anova results
> anova(lm.1)

Analysis of Variance Table

Response: read
          Df  Sum Sq Mean Sq F value Pr(>F)
grade      3  1924.4  641.48  1.4474 0.2397
Residuals 52 23045.3  443.18
```

There does **not** appear to be a main effect of time, $F(3, 52) = 1.45$, $p = 0.240$. This suggests that there are **no differences** in the average reading scores across grades.

Check Assumptions







Remember, when we have groups (nesting), we should check the residuals for each group

There is dependence in the residuals within each group (student).

We will have to fit a model that accounts for the correlation of the residuals within subjects.

ALTERNATIVE ANALYSES

Now What?

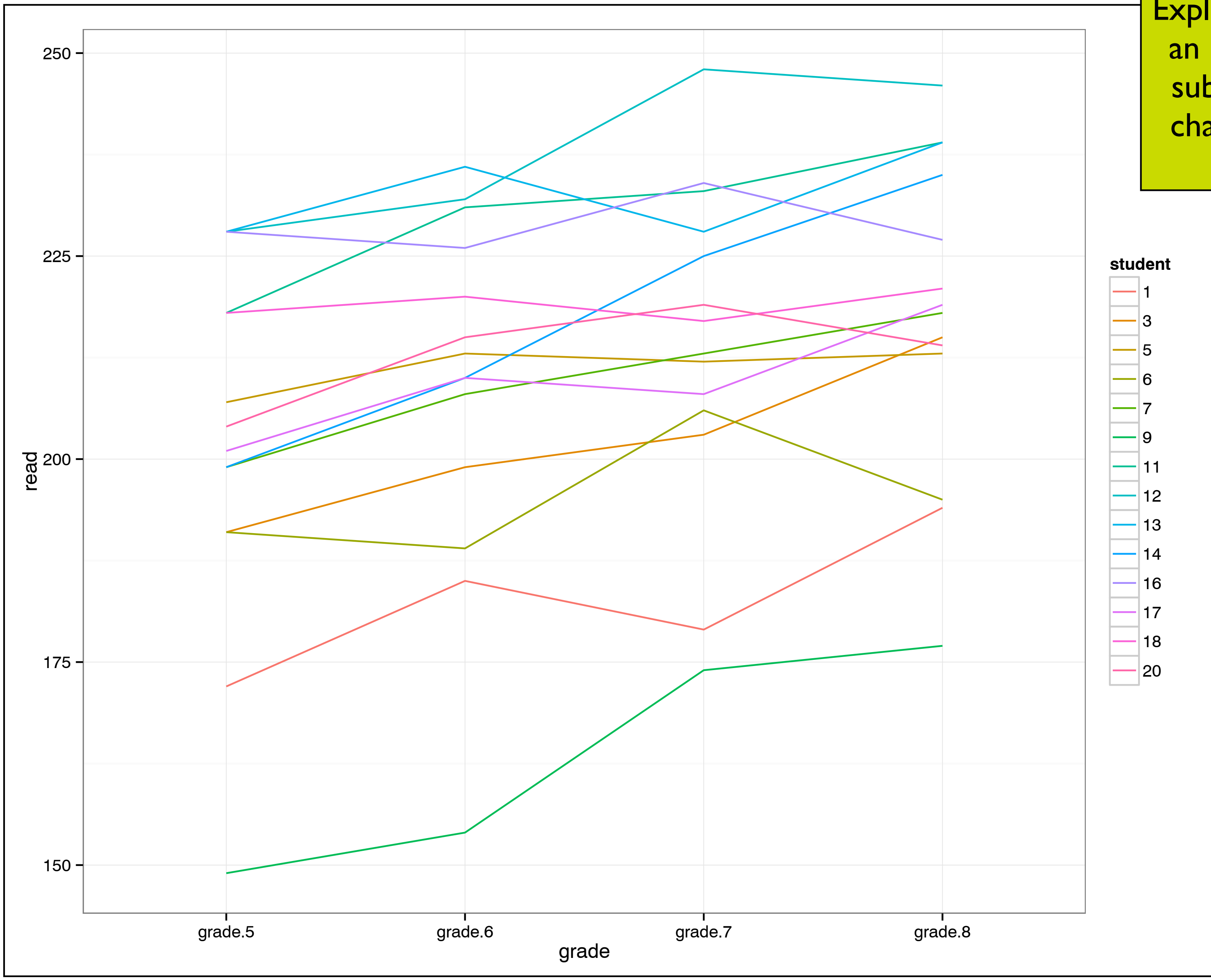
- Remove the correlation by 'transforming' the data
  - ✓ Randomly select one of the two time points for each subject
  - ✓ Compute a composite score
- Or...add subject as a factor into the model explicitly
  - ✓ Two-factor ANOVA
  - ✓ Main effect of time and main effect of subject

Create a subject factor

```
## Coerce studentID into a factor
> mplsLong$student = as.factor(mplsLong$studentID)
```

Should we include an interaction effect between subject and time?

```
## Examine data for possible interaction b/w student and grade
> ggplot(data = mplsLong, aes(x = grade, y = read, group = student)) +
    geom_line(aes(color = student)) +
    theme_bw()
```



Exploratory plot suggests an interaction between subjects and time… the change trajectories are not parallel.

One More Thorn[†,‡]

- All ANOVA analyses to this point have assumed the effect is fixed
  - ✓ Inferences are only drawn to the levels of the factor included in the sample

- In most analyses, we would like to draw inferences to a broader population of subjects (i.e., not just the 22 students in the sample!)
- It is possible to draw inferences to a broader population of levels if we assume that the levels of the factor included in the sample were indeed randomly sampled (or, at least treated a such)
  - ✓ We treat the effect as random
  - ✓ Need to account for the sampling variation that arises in making estimates from a subset of levels

[†] *"Every rose has its thorn" (Michaels, DeVille, Dall & Rockett, 1988).*
[‡] *"Every ANOVA has its thorn" (Zieffler, 2013).*

CONSEQUENCES OF RANDOM-EFFECT IN AN ANOVA: AN EXAMPLE

18 Therapists

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
Psychodynamic	7	6	5	7	6	5	4	4	4	1	2	3	4	4	4	1	2	3	4
Behavioral	4	4	4	1	2	3	7	6	5	7	6	5	1	2	3	4	4	4	4
Rogerian	1	2	3	4	4	4	1	2	3	4	4	4	7	6	5	7	6	5	4
	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	

Three psychotherapeutic modes

Mean ratings of effectiveness

All 18 therapists are, on average, equally effective

All 3 modes are, on average, equally effective

No main effect of therapist nor mode

Is this the case when we randomly sample levels of therapist (rather than using all 18 of them) to use in our study

What if we randomly sample therapists?

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
Psychodynamic	7	6	5	7	6	5	4	4	4	1	2	3	4	4	4	1	2	3	5
Behavioral	4	4	4	1	2	3	7	6	5	7	6	5	1	2	3	4	4	4	3.33
Rogerian	1	2	3	4	4	4	1	2	3	4	4	4	7	6	5	7	6	5	3.67
	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	

All 3 therapists are, on average, equally effective

All 3 modes are **NOT**, on average, equally effective

There appears to be a main effect of mode (fixed-effect) because of the random sampling of the other factor.

Looking at the cell means, it is obvious that the appearance of a main-effect of therapy mode is because therapy mode and therapist interact.

If we use the same *F*-test that are used for testing fixed-effects models, we will reject H_0 far more often than we should (increased type I error rate)!