Multilevel Modeling Formative

Anonymous Marking Code: Z0195806

2024-02-28

Part 0: Prepare for the Runtime and Dataset

gridExtra is a good library for combining multiple ggplots graphs in one graph, while par() can't make it.

```
# ------------------------------------------------------------
## clear the environment var area
# rm(list = ls())
## clear all plots
# graphics.off()
## clear the console area
# cat("\014")
# ------------------------------------------------------------
# install.packages("gridExtra")
# ------------------------------------------------------------
require(lme4)
require(lmerTest)
require(ggplot2)
require(sjPlot)
require(gridExtra)
```

Download data set from GitHub and show the first lines.

```
CRT <-
  read.csv("https://andygolightly.github.io/teaching/MATH43515/CRT.csv",
           header = TRUE)
head(CRT)
```

```
##   Pupil School Posttest Intervention Pretest FSM class
## 1     1      1       16            1       1   0     1
## 2     2      1       13            1       4   1     1
## 3     3      1       18            1       5   1     1
## 4     4      1       14            1       4   1     1
## 5     5      1       25            1       5   1     1
## 6     6      1       13            1       2   0     1
```

```
# dim(CRT)
```

Part 1: Introduction

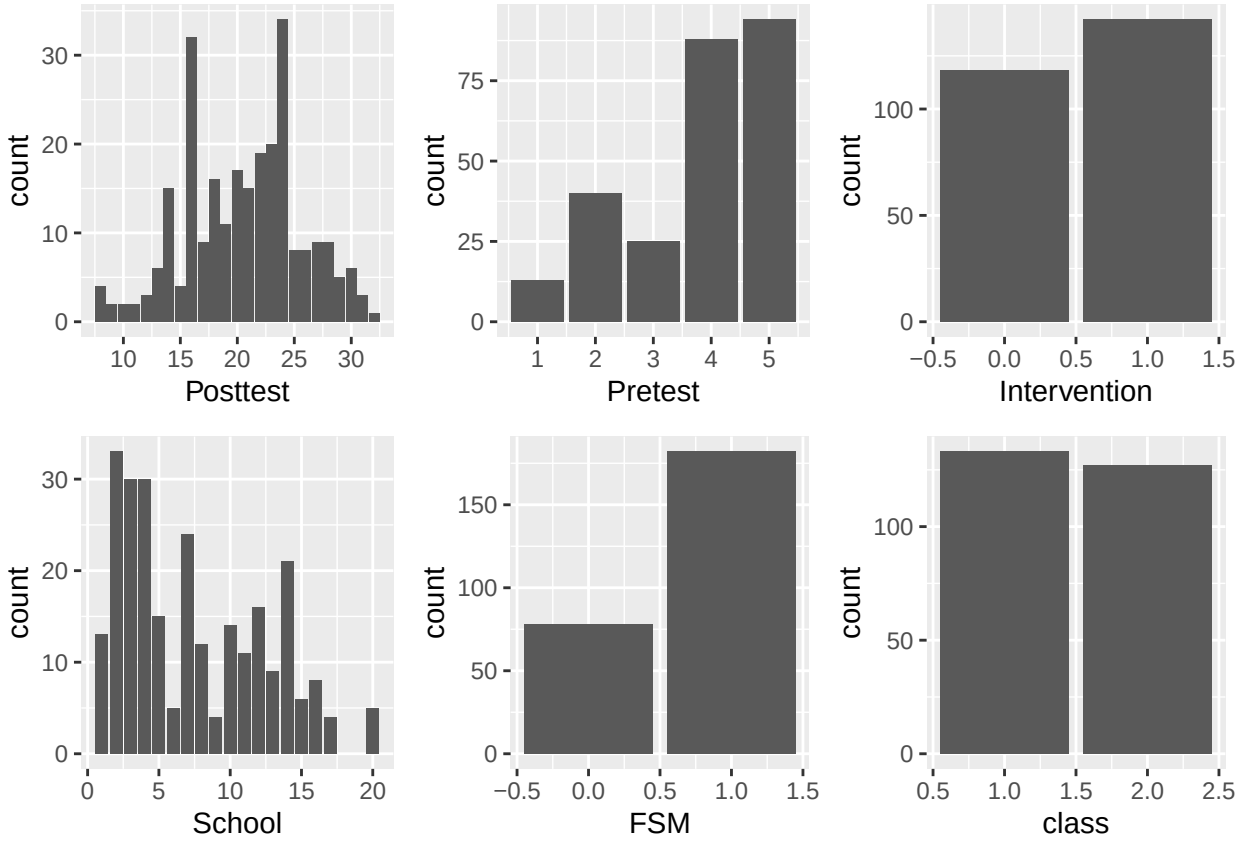
- randomized controlled trials and cluster randomized trials
- Intro to the CRT data set
  - `Pupil`: anonymized student ID;
  - `School`: anonymized school ID $\{1, \ldots, 20\}$;
  - `Class`: anonymized class ID $\{1, 2\}$;
  - `Intervention`: Intervention indicator (0=control; 1=treatment);
  - `FSM`: pupil eligible for free school meal (0=not eligible; 1 =eligible)
  - `Pretest`: A pre-test score for each pupil;
  - `Posttest`: A post-test score for each pupil (**response**).
- Histograms / Bar Plots / EDA with ggplots
- Check missing values
- Target: predict the Post-test

```
summary(CRT)
```

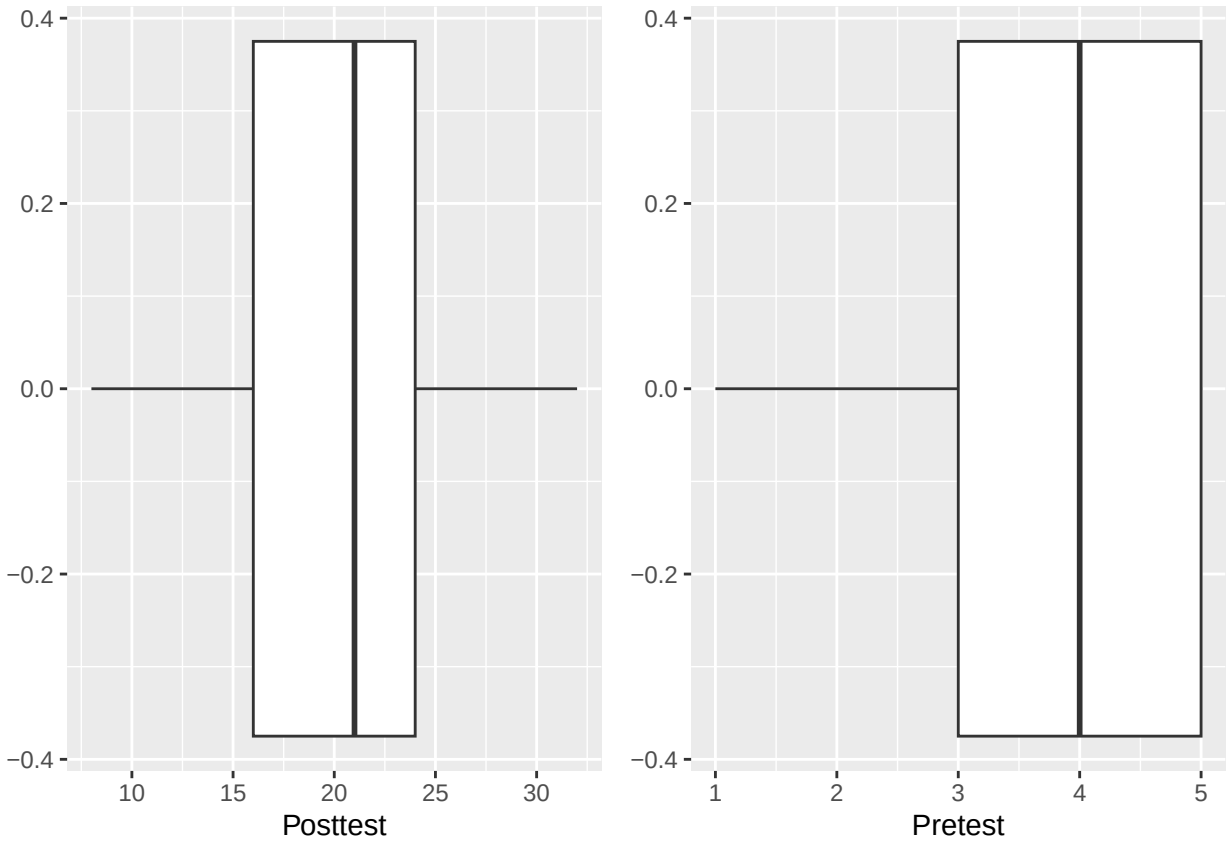
```
##      Pupil            School          Posttest      Intervention
##  Min.   :  1.00   Min.   : 1.000   Min.   : 8.00   Min.   :0.0000
##  1st Qu.: 65.75   1st Qu.: 3.000   1st Qu.:16.00   1st Qu.:0.0000
##  Median :130.50   Median : 7.000   Median :21.00   Median :1.0000
##  Mean   :130.56   Mean   : 7.477   Mean   :20.54   Mean   :0.5462
##  3rd Qu.:195.25   3rd Qu.:12.000   3rd Qu.:24.00   3rd Qu.:1.0000
##  Max.   :263.00   Max.   :20.000   Max.   :32.00   Max.   :1.0000
##     Pretest           FSM          class
##  Min.   :1.000   Min.   :0.0   Min.   :1.000
##  1st Qu.:3.000   1st Qu.:0.0   1st Qu.:1.000
##  Median :4.000   Median :1.0   Median :1.000
##  Mean   :3.808   Mean   :0.7   Mean   :1.488
##  3rd Qu.:5.000   3rd Qu.:1.0   3rd Qu.:2.000
##  Max.   :5.000   Max.   :1.0   Max.   :2.000
```

```
# ------------------------------------------------------
bar.Posttest = ggplot(data = CRT, aes(Posttest)) +
  geom_bar()
bar.Pretest = ggplot(data = CRT, aes(Pretest)) +
  geom_bar()
bar.Intervention = ggplot(data = CRT, aes(Intervention)) +
  geom_bar()
bar.School = ggplot(data = CRT, aes(School)) +
  geom_bar()
bar.FSM = ggplot(data = CRT, aes(FSM)) +
  geom_bar()
bar.class = ggplot(data = CRT, aes(class)) +
  geom_bar()
# ------------------------------------------------------
# boxplot
boxplot.Posttest = ggplot(data = CRT, aes(Posttest)) +
  geom_boxplot(outlier.colour = "red", outlier.shape = 1)
boxplot.Pretest = ggplot(data = CRT, aes(Pretest)) +
```

```
  geom_boxplot(outlier.colour = "red", outlier.shape = 1)
# ------------------------------------------------------------
# put these bar charts together
grid.arrange(bar.Posttest,
             bar.Pretest,
             bar.Intervention,
             bar.School,
             bar.FSM,
             bar.class,
             ncol = 3)
```

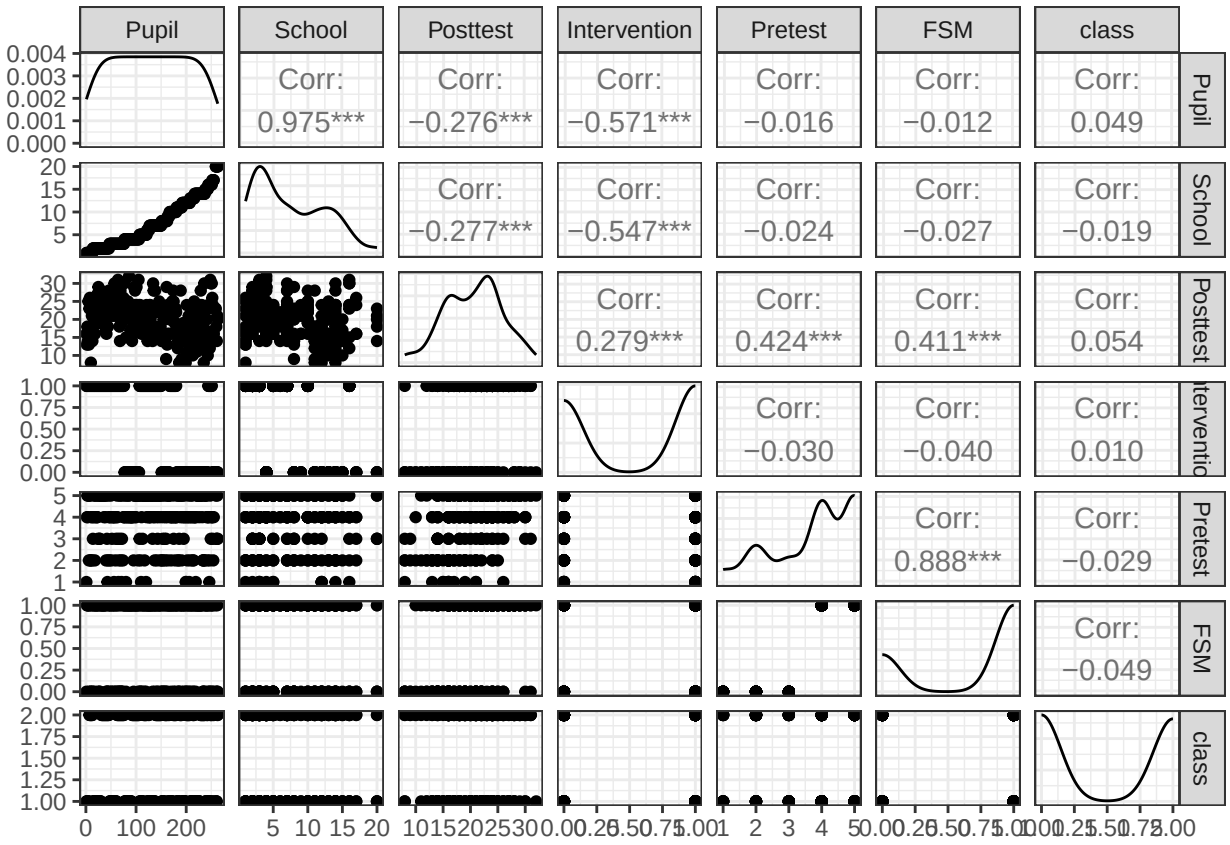



```
grid.arrange(boxplot.Posttest,
             boxplot.Pretest,
             ncol = 2)
```




```
# ------------------------------------------------------
# remove the plot cache memory
remove(bar.Posttest)
remove(bar.Pretest)
remove(bar.Intervention)
remove(bar.School)
remove(bar.FSM)
remove(bar.class)
remove(boxplot.Posttest)
remove(boxplot.Pretest)
```

```
# Correlation between
library("GGally")
ggpairs(CRT)+theme_bw()
```




```
## Not good to read and analyse
# tapply(CRT$School, CRT$class, table)
```

Part 2: Methods

Part 3: Analysis

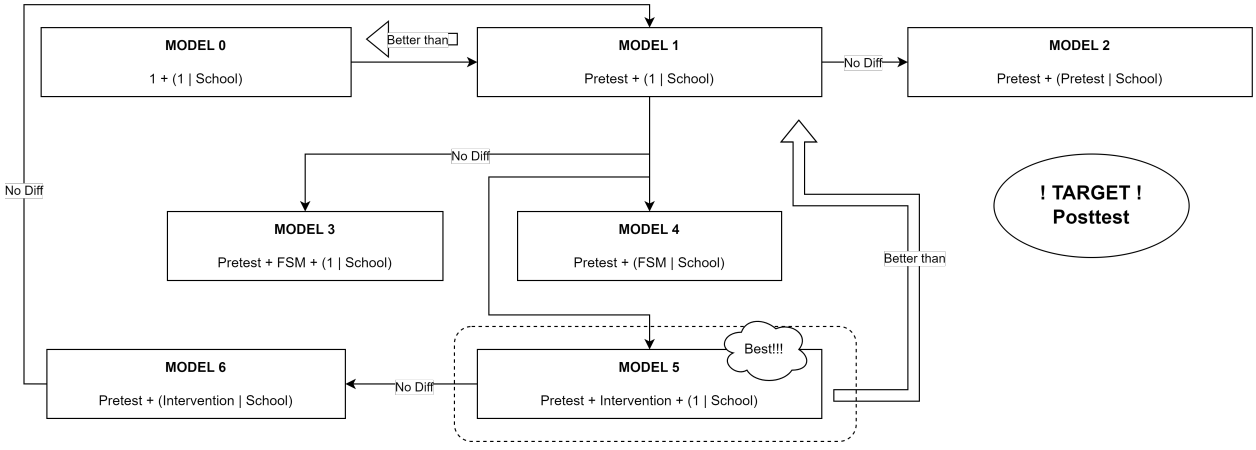


Figure 1: Create Models Bottom-up

```r
# empty model / intercept-only mode
Model.0.2level = lmer(Posttest ~ 1 + (1 | School),
                      data = CRT)
Model.0.3level = lmer(Posttest ~ 1 + (1 | School)
                      + (1 | School:class),
                      data = CRT)
ranova(Model.0.2level)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Posttest ~ (1 | School)
##               npar  logLik    AIC   LRT Df Pr(>Chisq)
## <none>           3 -768.93 1543.9
## (1 | School)     2 -792.14 1588.3 46.43  1  9.494e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
ranova(Model.0.3level)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Posttest ~ (1 | School) + (1 | School:class)
##                      npar  logLik    AIC    LRT Df Pr(>Chisq)
## <none>                  4 -768.73 1545.5
## (1 | School)            3 -772.30 1550.6 7.1322  1   0.007571 **
## (1 | School:class)      3 -768.93 1543.9 0.3933  1   0.530550
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Model.0.2level is better
Model.0 = Model.0.2level
```

```r
# summary(Model.0)
# ------------------------------------------------------------
# REsummary <- as.data.frame(VarCorr(Model.0))
# REsummary
# summary(Model.0)$varcor
# ------------------------------------------------------------
Model.1 = lmer(Posttest ~ Pretest + (1 | School),
               data = CRT)
# summary(Model.1)
# ------------------------------------------------------------
ranova(Model.1)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Posttest ~ Pretest + (1 | School)
##               npar  logLik    AIC    LRT Df Pr(>Chisq)
## <none>           4 -733.93 1475.9
```

```
## (1 | School)    3 -767.00 1540.0 66.121  1  4.241e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(Model.1, Model.0)
```

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

```
## Data: CRT
## Models:
## Model.0: Posttest ~ 1 + (1 | School)
## Model.1: Posttest ~ Pretest + (1 | School)
##         npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## Model.0    3 1544.9 1555.6 -769.44   1538.9
## Model.1    4 1475.6 1489.8 -733.78   1467.6 71.33  1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Model 1 is better than Model 0
```

```
# Model 2 is not better than model 1
Model.2 = lmer(Posttest ~ Pretest + (Pretest | School),
               data = CRT)
ranova(Model.2)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Posttest ~ Pretest + (Pretest | School)
##                               npar  logLik    AIC     LRT Df Pr(>Chisq)
## <none>                           6 -733.77 1479.5
## Pretest in (Pretest | School)    4 -733.93 1475.9 0.33308  2     0.8466
```

```
anova(Model.2, Model.1)
```

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

```
## Data: CRT
## Models:
## Model.1: Posttest ~ Pretest + (1 | School)
## Model.2: Posttest ~ Pretest + (Pretest | School)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## Model.1    4 1475.6 1489.8 -733.78   1467.6
## Model.2    6 1479.4 1500.7 -733.68   1467.4 0.1946  2     0.9073
```

```
# Model 3 (add FSM to Fix Effects) is not better than Model 1
Model.3 = lmer(Posttest ~ Pretest + FSM + (1 | School),
               data = CRT)
ranova(Model.3)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Posttest ~ Pretest + FSM + (1 | School)
##              npar  logLik    AIC    LRT Df Pr(>Chisq)
## <none>          5 -732.03 1474.1
## (1 | School)    4 -764.90 1537.8 65.739  1  5.147e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(Model.3, Model.1)
```

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

## Data: CRT
## Models:
## Model.1: Posttest ~ Pretest + (1 | School)
## Model.3: Posttest ~ Pretest + FSM + (1 | School)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## Model.1    4 1475.6 1489.8 -733.78   1467.6
## Model.3    5 1475.9 1493.7 -732.95   1465.9 1.6475  1     0.1993
```

```
# Model 4 (add FSM to Random Effects) is not better than Model 1
Model.4 = lmer(Posttest ~ Pretest + (FSM | School),
               data = CRT)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
ranova(Model.4)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Posttest ~ Pretest + (FSM | School)
##                        npar  logLik    AIC     LRT Df Pr(>Chisq)
## <none>                    6 -733.83 1479.7
## FSM in (FSM | School)     4 -733.93 1475.9 0.20514  2     0.9025
```

```
anova(Model.4, Model.1)
```

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

## Data: CRT
## Models:
## Model.1: Posttest ~ Pretest + (1 | School)
## Model.4: Posttest ~ Pretest + (FSM | School)
##         npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## Model.1    4 1475.6 1489.8 -733.78   1467.6
## Model.4    6 1479.3 1500.7 -733.67   1467.3 0.211  2     0.8999
```

```r
Model.5 = lmer(Posttest ~ Pretest + Intervention + (1 | School),
               data = CRT)
ranova(Model.5)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Posttest ~ Pretest + Intervention + (1 | School)
##              npar  logLik    AIC    LRT Df Pr(>Chisq)
## <none>          5 -730.06 1470.1
## (1 | School)    4 -752.58 1513.2 45.041  1   1.93e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(Model.5, Model.1)
```

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

## Data: CRT
## Models:
## Model.1: Posttest ~ Pretest + (1 | School)
## Model.5: Posttest ~ Pretest + Intervention + (1 | School)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## Model.1    4 1475.6 1489.8 -733.78   1467.6
## Model.5    5 1471.6 1489.4 -730.81   1461.6 5.9407  1     0.0148 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
Model.6 = lmer(Posttest ~ Pretest + Intervention + (Intervention | School),
               data = CRT)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## unable to evaluate scaled gradient

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge: degenerate Hessian with 1 negative eigenvalues

## Warning: Model failed to converge with 1 negative eigenvalue: -4.2e-05
```

```r
ranova(Model.6)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Posttest ~ Pretest + Intervention + (Intervention | School)
##                                          npar  logLik    AIC      LRT Df
## <none>                                      7 -730.03 1474.1
## Intervention in (Intervention | School)     5 -730.06 1470.1 0.062275  2
##                                          Pr(>Chisq)
## <none>
## Intervention in (Intervention | School)     0.9693
```

```r
anova(Model.6, Model.5)
```

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

## Data: CRT
## Models:
## Model.5: Posttest ~ Pretest + Intervention + (1 | School)
## Model.6: Posttest ~ Pretest + Intervention + (Intervention | School)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## Model.5    5 1471.6 1489.4 -730.81   1461.6
## Model.6    7 1475.6 1500.5 -730.79   1461.6 0.0429  2     0.9788
```

```r
require(performance)
```

```
##      performance
```

```r
# https://easystats.github.io/performance/reference/icc.html
icc(Model.0)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.244
##   Unadjusted ICC: 0.244
```

```r
icc(Model.1)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.324
##   Unadjusted ICC: 0.263
```

Part 4: Discussion of results

References

Evaluating Intervention Programs with a Pretest-Posttest Design: A Structural Equation Modeling Approach

Word count

```r
# install.packages("devtools")
# devtools::install_github("benmarwick/wordcountaddin",
#                          type = "source", dependencies = TRUE)
require(wordcountaddin)
word_count()
```

```
## [1] 151
```

```
text_stats()
```

Method	koRpus	stringi
Word count	151	133
Character count	1032	1143
Sentence count	24	Not available
Reading time	0.8 minutes	0.7 minutes