Two Factor Split plot design

A: whole plot factor
B: subplot factor

Teaching example:
Three teaching methods (factor A, high, medium, low) assigned to 9 classes. Whole plot unit: class.

Within each class, 2 groups are formed and library use (factor B, yes or no) assigned to each of the 2 groups.
split plot unit: subgroups in each class.

Model: $y_{ijk} = \mu + \alpha_i + \epsilon^w_{k(i)} + \beta_j + \alpha\beta_{ij} + \epsilon^s_{ijk}$.
$i = 1, \cdots, a; j = 1, \cdots, b; k = 1, \cdots, n$, where a is the number of levels of factor A, b is the number of levels of factor B, and n is the number of whole plot units assigned to each level of factor A.
$SST = SSA + SSE_w + SSB + SSAB + SSE_s$.

Test for A effect:
$F = MSA/MSE_w$.
Test for B and AB effect:
$F = MSB/MSE_s, F = MSAB/MSE_s$.

		Baking Time		
Flour Type	Roll	5min	10min	15min
White	1	44	46	47
	2	42	46	48
	3	42	43	43
Wheat	1	40	40	42
	2	40	41	41
	3	40	41	41
Bread	1	43	44	46
	2	43	44	45
	3	41	43	43

Flour example

Flour: whole plot factor
Baking time: subplot factor

Whole plot unit: roll
subplot unit: cookie

data file bake.txt
$>$ out = aov (y$\sim$ Error(roll:flour)+flour*time,data=bake)

```
> summary(out)

Error: flour:roll
          Df  Sum Sq  Mean Sq  F value      Pr(>F)

flour      2  73.407   36.704     9.53      0.0137
Residuals  6  23.1111  3.8519
-------------------------------------------------
Error: Within
           Df  Sum Sq Mean Sq F value    Pr(>F)
time        2 24.9630 12.4815  16.850 0.0003278 ***
flour:time  4  3.4815  0.8704   1.175 0.3700995
Residuals  12  8.8889  0.7407
```

To compare flour type: (95%)

Four Type	mean
White	44.6
Wheat	40.7
Bread	43.6

> qtukey(0.95,3,6)/sqrt(2).

$White - Wheat : 44.6 - 40.7 \pm 3.068\sqrt{3.85}\sqrt{1/9 + 1/9} = (1.1, 6.7)$.

To compare baking time:

Baking Time	mean
5min	41.7
10min	43.1
15min	44.0

qtukey(.95,3, 12)/sqrt(2)=2.668

10min-5min: $43.1 - 41.7 \pm 2.668 * \sqrt{0.74} * \sqrt{1/9 + 1/9} = (0.3, 2.5)$

Get CIs by R

```
  > install.packages("lmerTest")
  > library(lmerTest)

 > fit=lmer(y~flour+time+(1|roll:flour),data=bake)
  # fit a simpler model as interaction is not signific
 > anova(fit)
Type III Analysis of Variance Table with Satterthwaite
      Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
flour 14.747  7.3733     2  6.0049  9.5346 0.0136873 *
time  24.963 12.4815     2 15.9977 16.1401 0.0001456 *
---

> install.packages("pbkrtest")
> library(pbkrtest)
```

```
> confint(emmeans(fit,pairwise~flour))

$emmeans
 flour emmean    SE df lower.CL upper.CL
 bread   43.6 0.654  6     42.0     45.2
 wheat   40.7 0.654  6     39.1     42.3
 white   44.6 0.654  6     43.0     46.2

Results are averaged over the levels of: time
Degrees-of-freedom method: kenward-roger
Confidence level used: 0.95

$contrasts
 contrast       estimate    SE df lower.CL upper.CL
 bread - wheat      2.89 0.925  6    0.051     5.73
 bread - white     -1.00 0.925  6   -3.838     1.84
 wheat - white     -3.89 0.925  6   -6.727    -1.05
```

```
> confint(emmeans(fit,pairwise~time))
$emmeans
 time emmean    SE   df lower.CL upper.CL
 5      41.7 0.447 11.1     40.7     42.6
 10     43.1 0.447 11.1     42.1     44.1
 15     44.0 0.447 11.1     43.0     45.0

Results are averaged over the levels of: flour
Degrees-of-freedom method: kenward-roger
Confidence level used: 0.95

$contrasts
 contrast estimate    SE df lower.CL upper.CL
 5 - 10     -1.444 0.415 16    -2.51   -0.375
 5 - 15     -2.333 0.415 16    -3.40   -1.264
 10 - 15    -0.889 0.415 16    -1.96    0.181
```

Three factor model

	a1			a2			a3			a4		
	b1	b2	b3	b1	b2	b3	b1	b2	b3	b1	b2	b3
c1:												
	4.1	4.6	3.7	4.9	5.2	4.7	5.0	6.1	5.5	3.9	4.4	3.7
	4.3	4.9	3.9	4.6	5.6	4.7	5.4	6.2	5.9	3.3	4.3	3.9
	4.5	4.2	4.1	5.3	5.8	5.0	5.7	6.5	5.6	3.4	4.7	4.0
	3.8	4.5	4.5	5.0	5.4	4.5	5.3	5.7	5.0	3.7	4.1	4.4
	4.3	4.8	3.9	4.6	5.5	4.7	5.4	6.1	5.9	3.3	4.2	3.9
c2:												
	4.8	5.6	5.0	4.9	5.9	5.0	6.0	6.0	6.1	4.1	4.9	4.3
	4.5	5.8	5.2	5.5	5.3	5.4	5.7	6.3	5.3	3.9	4.7	4.1
	5.0	5.4	4.6	5.5	5.5	4.7	5.5	5.7	5.5	4.3	4.9	3.8
	4.6	6.1	4.9	5.3	5.7	5.1	5.7	5.9	5.8	4.0	5.3	4.7
	5.0	5.4	4.7	5.5	5.5	4.9	5.5	5.7	5.6	4.3	4.3	3.8

Read in scan.txt data on course website

y =scan("scan.txt")

a= rep(c(1,1,1,2,2,2,3,3,3,4,4,4),10)

b = rep(c(1,2,3),40)

c = c(rep(1,60),rep(2,60))

a = factor (a); b= factor (b); c= factor (c)

output = lm (y $\sim$ a*b*c)

```
> anova(output)
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq  F value    Pr(>F)
a          3 40.322  13.441 182.4506 < 2.2e-16 ***
b          2  8.821   4.411  59.8722 < 2.2e-16 ***
c          1  4.760   4.760  64.6165 2.356e-12 ***
a:b        6  0.814   0.136   1.8420   0.09895 .
a:c        3  2.351   0.784  10.6376 4.216e-06 ***
b:c        2  0.126   0.063   0.8563   0.42793
a:b:c      6  0.944   0.157   2.1354   0.05616 .
Residuals 96  7.072   0.074
---
```

interaction plot

```
mydata = data.frame(cbind(y,a,b,c))
colnames(mydata) = c("y", "a", "b", "c")
par(mfrow=c(2,1))
with(mydata,interaction.plot(a[c==1],b[c==1],y[c==1]))
with(mydata,interaction.plot(a[c==2],b[c==2],y[c==2]))
```

exercise: check interaction between a and c at three levels of b.

At level c=1 and c=2