# Oats

*Dr Richard Wilkinson*

*8 February 2016*

## jsp data

```
library(faraway)
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(ggplot2)
data(jsp)
head(jsp)
```

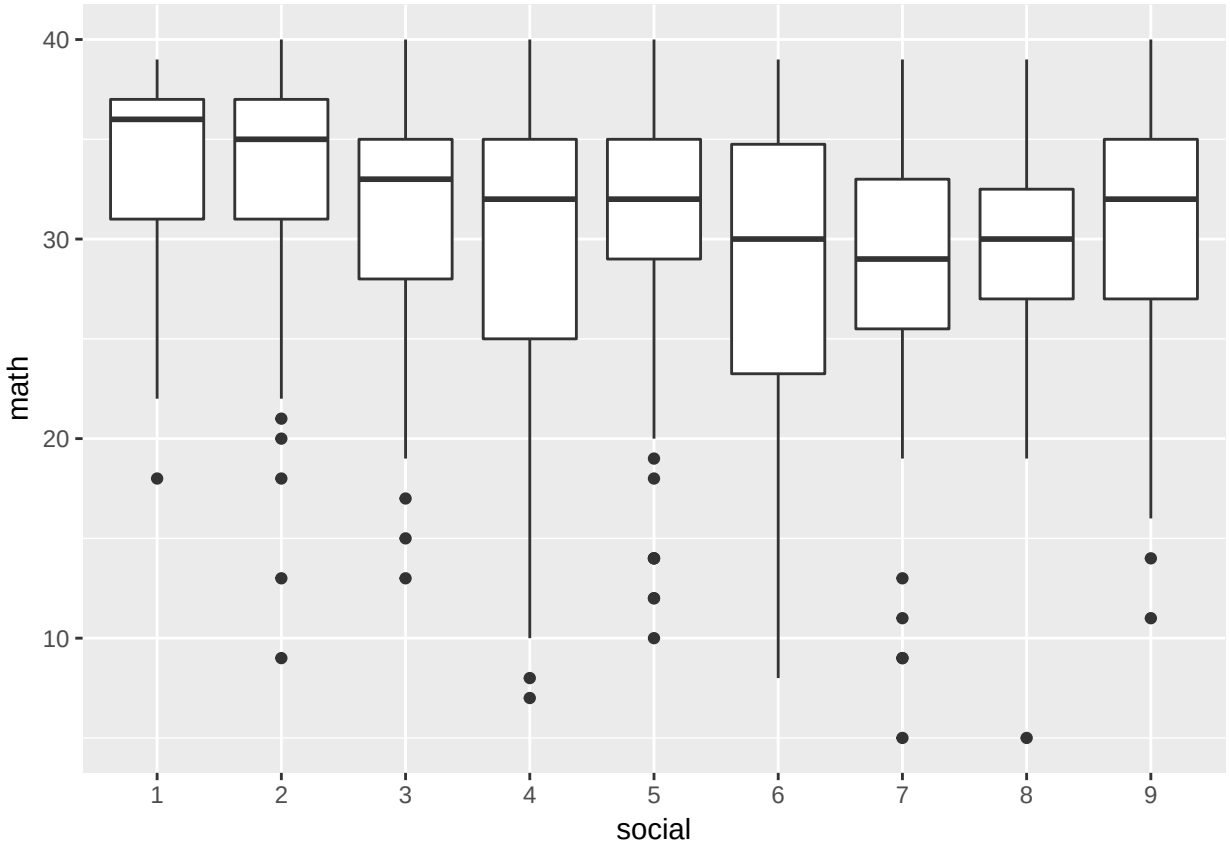
```
##   school class gender social raven id english math year
## 1      1     1   girl      9    23  1      72   23    0
## 2      1     1   girl      9    23  1      80   24    1
## 3      1     1   girl      9    23  1      39   23    2
## 4      1     1    boy      2    15  2       7   14    0
## 5      1     1    boy      2    15  2      17   11    1
## 6      1     1    boy      2    22  3      88   36    0
```

```
dim(jsp)
```

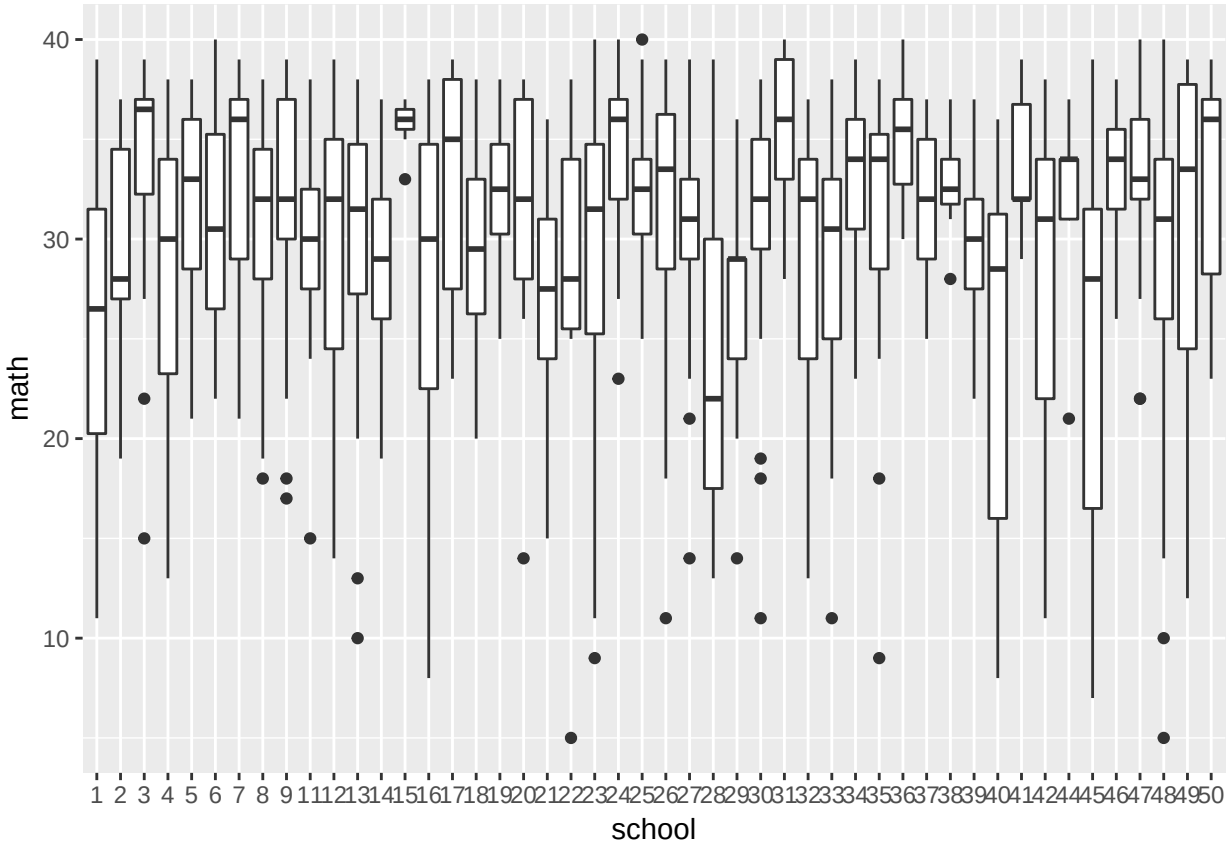
```
## [1] 3236    9
```

Select final year only

```
jspr <- jsp[jsp$year==2,]
qplot(social, math, data = jspr, geom='boxplot')
```




```
qplot(school, math, data = jspr, geom='boxplot')
```




Lets account for school

```
lmer(math~ social + (1|school), data=jspr, contrasts=list(social=contr.sum))
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: math ~ social + (1 | school)
##    Data: jspr
## REML criterion at convergence: 6210.064
## Random effects:
##  Groups   Name        Std.Dev.
##  school   (Intercept) 2.133
##  Residual             6.166
## Number of obs: 953, groups:  school, 48
## Fixed Effects:
## (Intercept)      social1      social2      social3      social4
##     30.5112       2.9684       2.6650       0.9707      -0.8156
##     social5      social6      social7      social8
##     -0.2525      -1.4180      -2.1362      -2.2664
```

To account for class we need to nest effects

```
lmer(math~ social + (1|school)+(1| school:class), data=jspr, contrasts=list(social=contr.sum))
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: math ~ social + (1 | school) + (1 | school:class)
##    Data: jspr
## REML criterion at convergence: 6207.032
## Random effects:
```

```
##  Groups       Name        Std.Dev.
##  school:class (Intercept) 1.273
##  school       (Intercept) 1.864
##  Residual                 6.111
## Number of obs: 953, groups:  school:class, 90; school, 48
## Fixed Effects:
## (Intercept)      social1      social2      social3      social4
##     30.5005       2.9751       2.6323       0.9643      -0.7690
##     social5      social6      social7      social8
##     -0.2694      -1.3992      -2.2050      -2.2331
```

```
lmer(math~ social + (1| school/class), data=jspr, contrasts=list(social=contr.sum))
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: math ~ social + (1 | school/class)
##    Data: jspr
## REML criterion at convergence: 6207.032
## Random effects:
##  Groups       Name        Std.Dev.
##  class:school (Intercept) 1.273
##  school       (Intercept) 1.864
##  Residual                 6.111
## Number of obs: 953, groups:  class:school, 90; school, 48
## Fixed Effects:
## (Intercept)      social1      social2      social3      social4
##     30.5005       2.9751       2.6323       0.9643      -0.7690
##     social5      social6      social7      social8
##     -0.2694      -1.3992      -2.2050      -2.2331
```

# Chapter 3 Section 3.1.1

```
load('MAS473.RData')
```

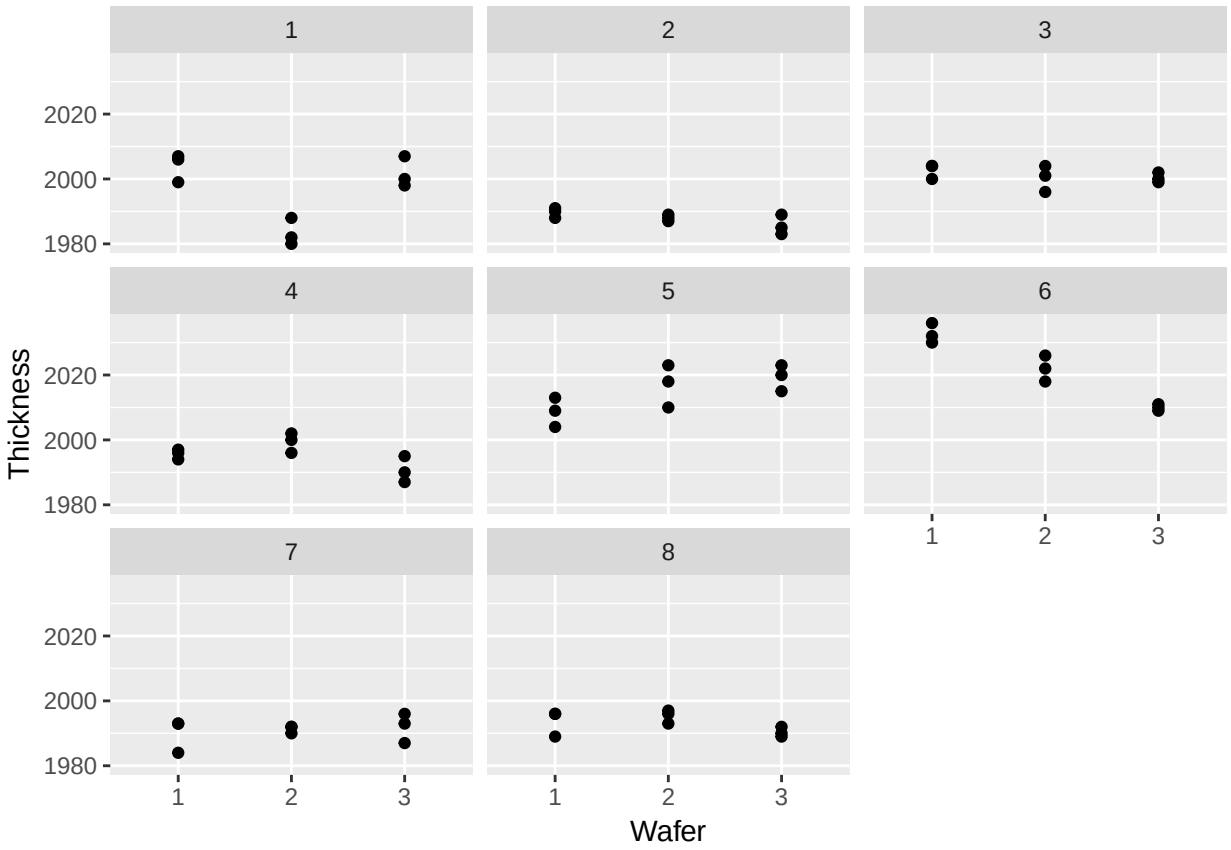
## Inspect the data

```
head(Oxide)
```

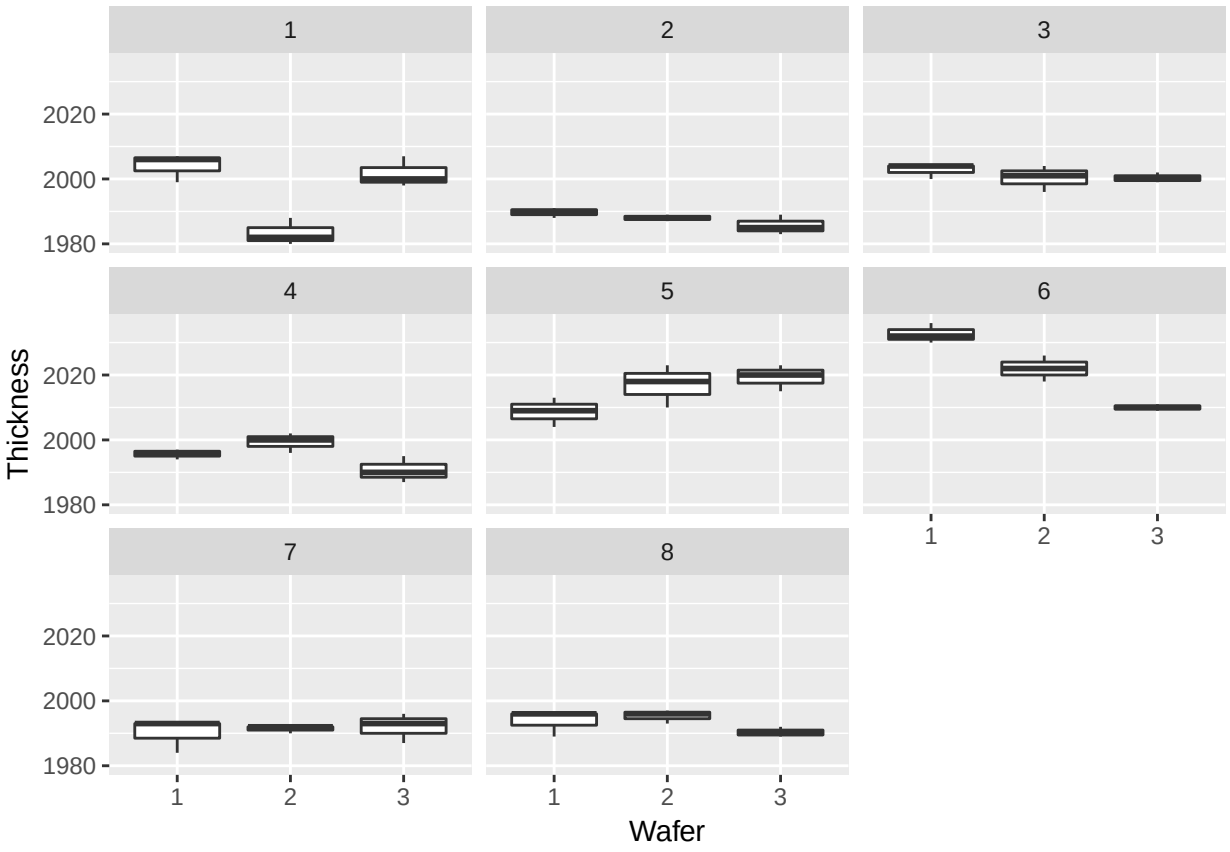
```
##   Lot Wafer Site Thickness
## 1   1     1    1      2006
## 2   1     1    2      1999
## 3   1     1    3      2007
## 4   1     2    1      1980
## 5   1     2    2      1988
## 6   1     2    3      1982
```

Note that Wafer isn't a 3-level factor; Wafer is nested within Lot eg $Y_{121}$ and $Y_{221}$ do not refer to the same Wafer.

```
library(ggplot2)
qplot(data=Oxide, Wafer, Thickness, facets=~Lot)
```

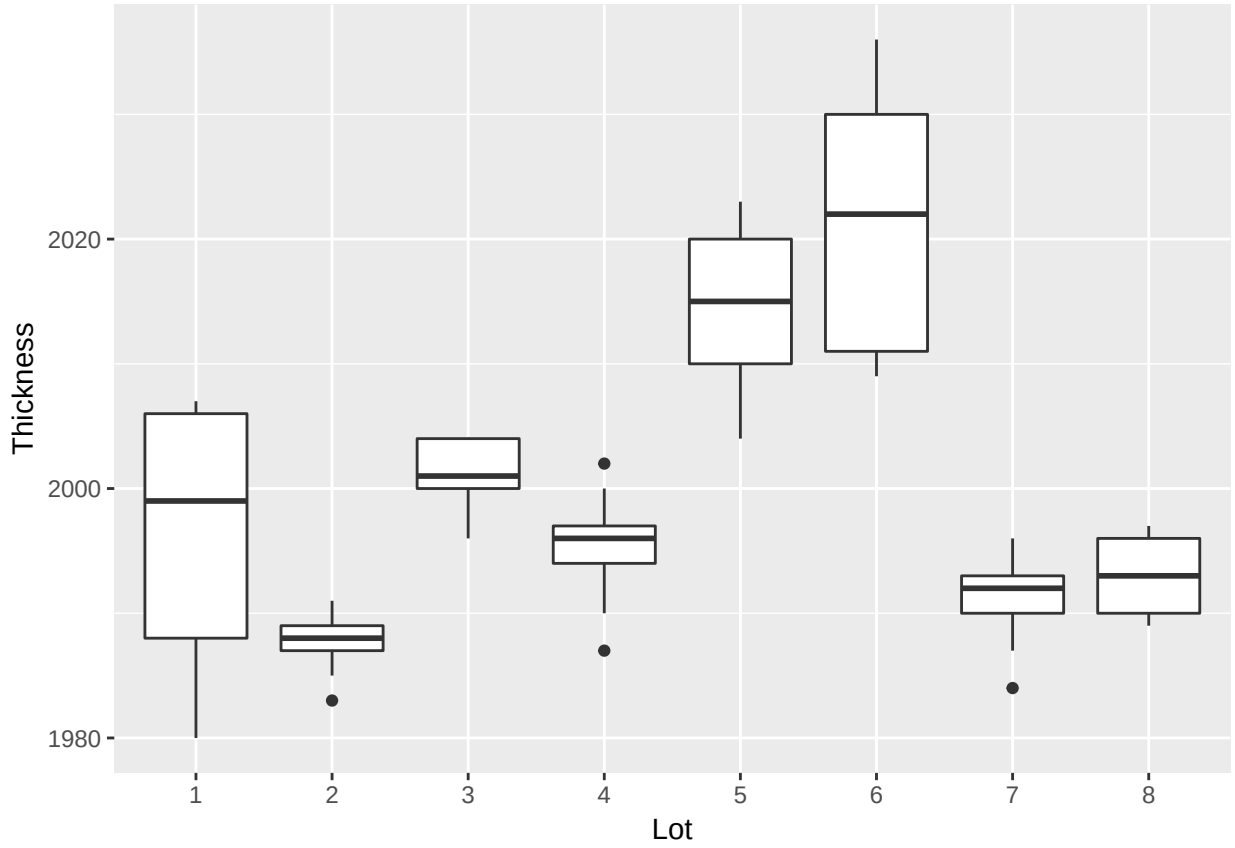



```
qplot(data=Oxide, Wafer, Thickness, facets=~Lot, geom='boxplot')
```




which do you prefer?

```
qplot(data=Oxide,Lot,Thickness,geom='boxplot')
```




We can see variation between Wafers in the same Lot and variation between Lots

```
fm1<-lmer(Thickness~1+(1|Lot/Wafer),data=Oxide)
summary(fm1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Thickness ~ 1 + (1 | Lot/Wafer)
##    Data: Oxide
##
## REML criterion at convergence: 454
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8746 -0.4991  0.1047  0.5510  1.7922
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Wafer:Lot (Intercept)  35.87    5.989
##  Lot       (Intercept) 129.91   11.398
##  Residual               12.57    3.545
## Number of obs: 72, groups:  Wafer:Lot, 24; Lot, 8
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 2000.153      4.232   472.7
```

equivalent command

```r
fm2 <- lmer(Thickness~1 +(1|Lot)+(1|Lot:Wafer), data =Oxide)
summary(fm2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Thickness ~ 1 + (1 | Lot) + (1 | Lot:Wafer)
##    Data: Oxide
##
## REML criterion at convergence: 454
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8746 -0.4991  0.1047  0.5510  1.7922
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Lot:Wafer (Intercept)  35.87    5.989
##  Lot       (Intercept) 129.91   11.398
##  Residual               12.57    3.545
## Number of obs: 72, groups:  Lot:Wafer, 24; Lot, 8
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 2000.153      4.232   472.7
```

```r
# Estimated random effects
ranef(fm1)
```

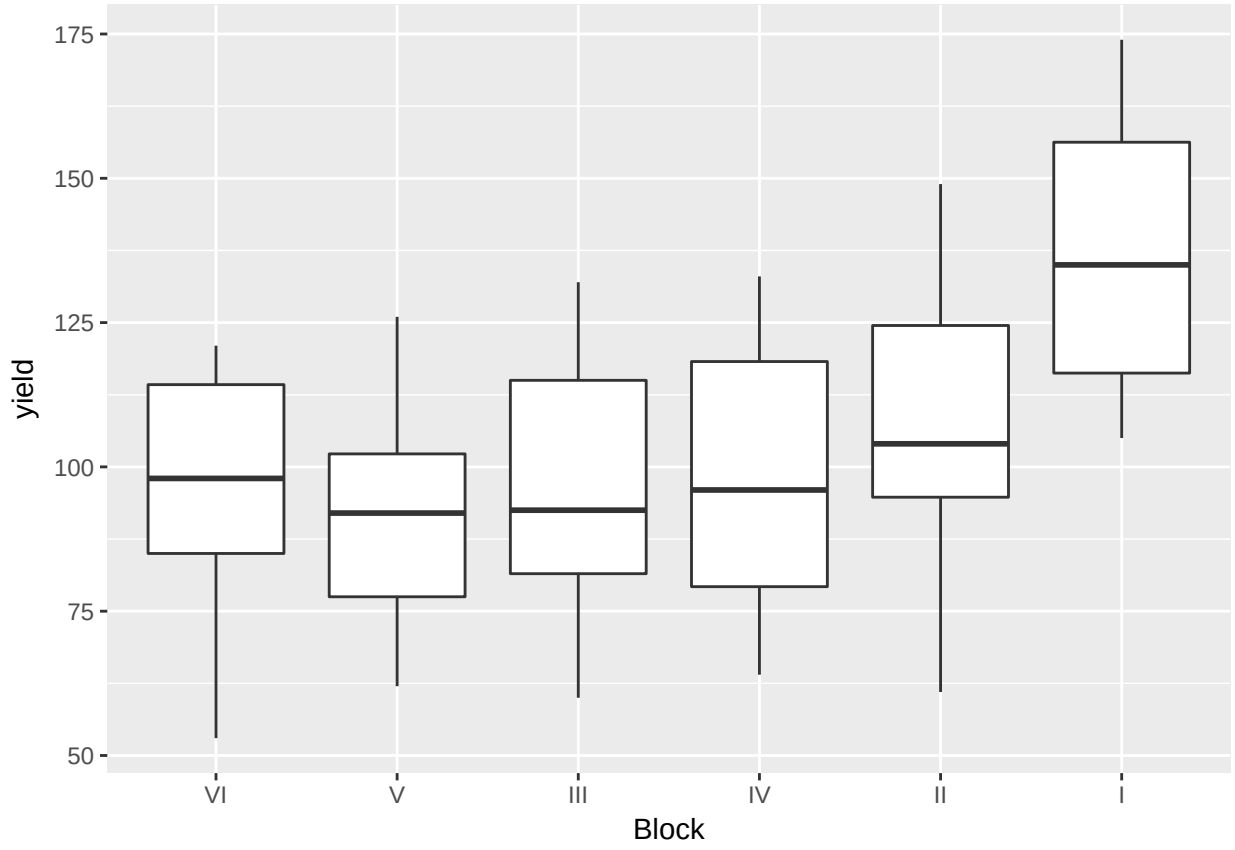
```
## $`Wafer:Lot`
##      (Intercept)
## 1:1   6.54599243
## 1:2   0.65859294
## 1:3   1.47281908
## 1:4  -0.01350901
## 1:5  -4.43183625
## 1:6  11.73499147
## 1:7  -1.74943356
## 1:8  -0.09019648
## 2:1 -11.95893879
## 2:2  -0.83374023
## 2:3  -0.61644735
## 2:4   3.26962395
## 2:5   3.02982956
## 2:6   2.18405923
## 2:7  -0.55556703
## 2:8   1.40213668
## 3:1   4.45672600
## 3:2  -2.92300666
## 3:3  -0.61644735
## 3:4  -4.49050850
## 3:5   5.11909599
## 3:6  -8.56073955
## 3:7   0.04136623
## 3:8  -3.07486281
##
## $Lot
```

```
##   (Intercept)
## 1  -3.4634693
## 2 -11.2216405
## 3   0.8690159
## 4  -4.4710240
## 5  13.4634497
## 6  19.4080225
## 7  -8.1989764
## 8  -6.3853779
```

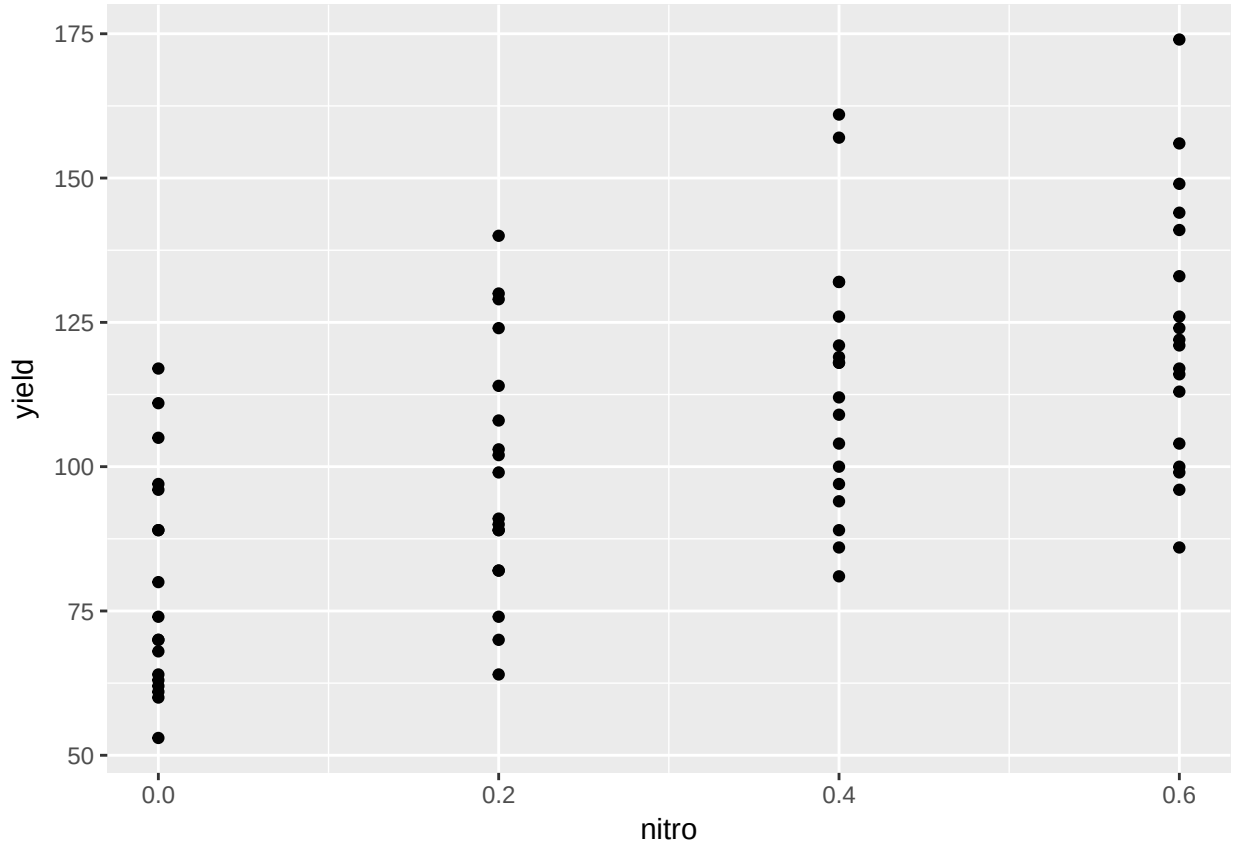
## Section 3.1.2

Split plot example

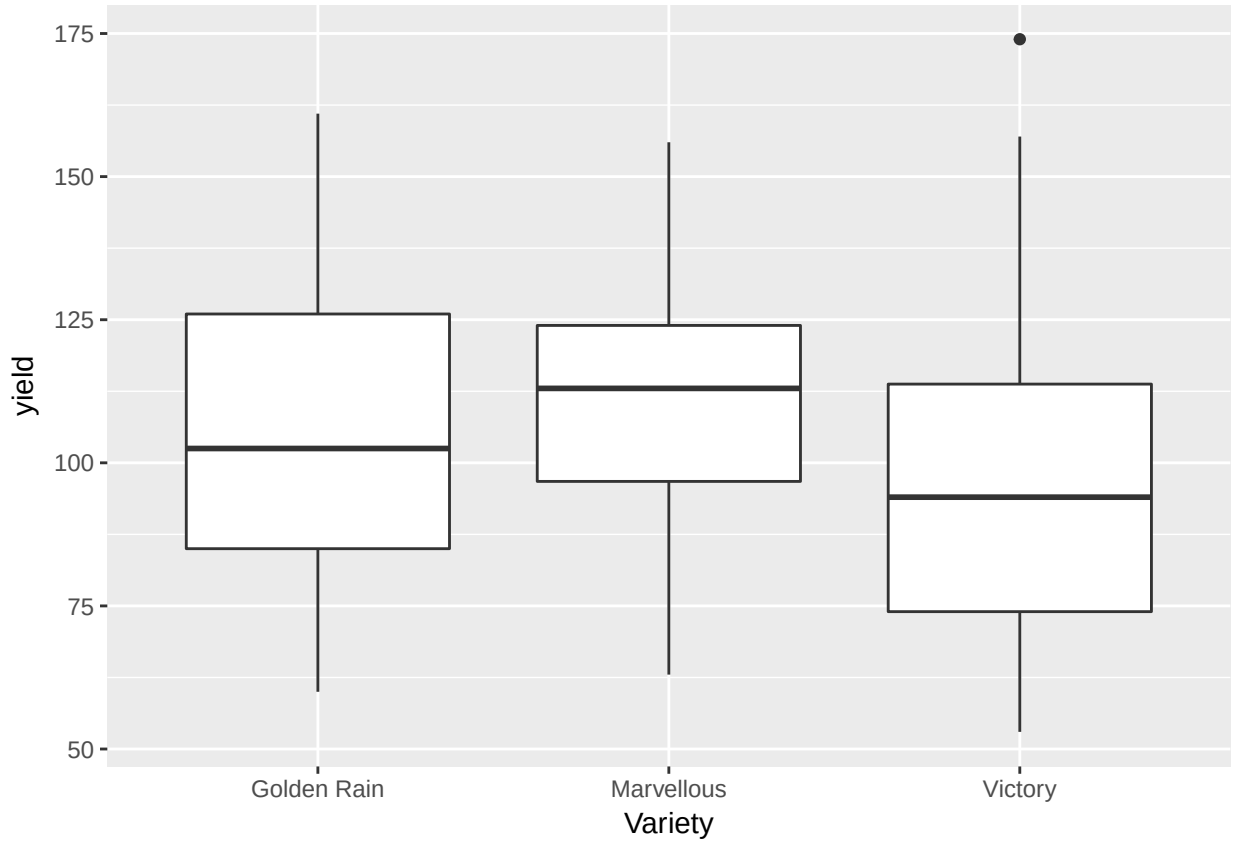
```
attach(Oats)
qplot(Block, yield, geom='boxplot')
```



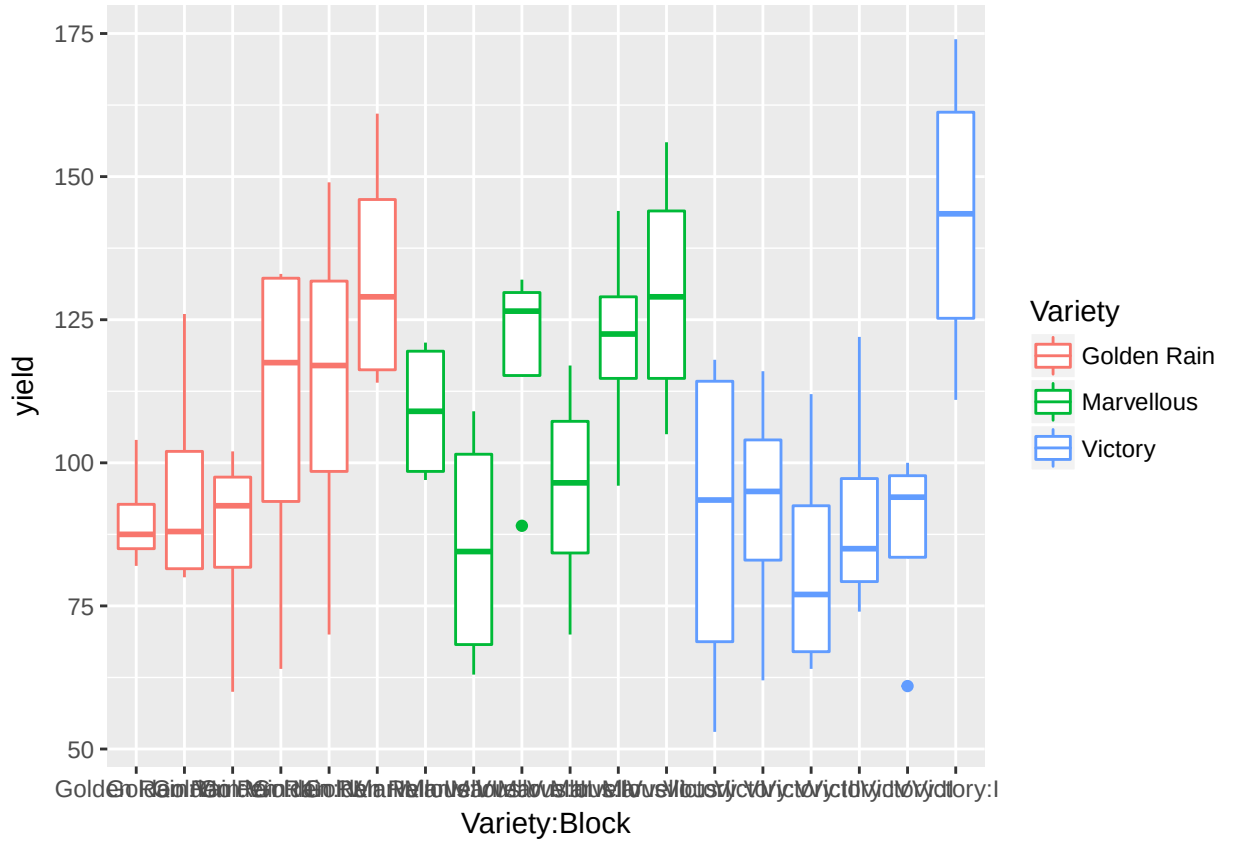

```
qplot(nitro, yield)
```




```
qplot(Variety,yield, geom='boxplot')
```




```
qplot(Variety:Block, yield, geom='boxplot', col=Variety)
```




Can get the plot random effect via Block/Variety Variety included as a fixed effect too

```
(fm1<-lmer(yield~ nitro + Variety + (1|Block/Variety),Oats))
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: yield ~ nitro + Variety + (1 | Block/Variety)
##    Data: Oats
## REML criterion at convergence: 578.8918
## Random effects:
##  Groups        Name        Std.Dev.
##  Variety:Block (Intercept) 10.44
##  Block         (Intercept) 14.65
##  Residual                  12.87
## Number of obs: 72, groups:  Variety:Block, 18; Block, 6
## Fixed Effects:
##       (Intercept)              nitro  VarietyMarvellous
##            82.400             73.667              5.292
##    VarietyVictory
##            -6.875
```

doesn't mean that variety is a random effect, just that it indicates the different plots within the block note that we don't need a random effect for the subplot - only 1 observation in each so represented by

$$\epsilon_{ijk}$$

If this confuses you, can specify a plot factor:

```
Plt<-gl(18,4)
Oats<-data.frame(Oats,Plt)
(Oats)
```

```
##    Block     Variety nitro yield Plt
## 1      I     Victory   0.0   111   1
## 2      I     Victory   0.2   130   1
## 3      I     Victory   0.4   157   1
## 4      I     Victory   0.6   174   1
## 5      I Golden Rain   0.0   117   2
## 6      I Golden Rain   0.2   114   2
## 7      I Golden Rain   0.4   161   2
## 8      I Golden Rain   0.6   141   2
## 9      I  Marvellous   0.0   105   3
## 10     I  Marvellous   0.2   140   3
## 11     I  Marvellous   0.4   118   3
## 12     I  Marvellous   0.6   156   3
## 13    II     Victory   0.0    61   4
## 14    II     Victory   0.2    91   4
## 15    II     Victory   0.4    97   4
## 16    II     Victory   0.6   100   4
## 17    II Golden Rain   0.0    70   5
## 18    II Golden Rain   0.2   108   5
## 19    II Golden Rain   0.4   126   5
## 20    II Golden Rain   0.6   149   5
## 21    II  Marvellous   0.0    96   6
## 22    II  Marvellous   0.2   124   6
## 23    II  Marvellous   0.4   121   6
## 24    II  Marvellous   0.6   144   6
## 25   III     Victory   0.0    68   7
## 26   III     Victory   0.2    64   7
## 27   III     Victory   0.4   112   7
## 28   III     Victory   0.6    86   7
## 29   III Golden Rain   0.0    60   8
## 30   III Golden Rain   0.2   102   8
## 31   III Golden Rain   0.4    89   8
## 32   III Golden Rain   0.6    96   8
## 33   III  Marvellous   0.0    89   9
## 34   III  Marvellous   0.2   129   9
## 35   III  Marvellous   0.4   132   9
## 36   III  Marvellous   0.6   124   9
## 37    IV     Victory   0.0    74  10
## 38    IV     Victory   0.2    89  10
## 39    IV     Victory   0.4    81  10
## 40    IV     Victory   0.6   122  10
## 41    IV Golden Rain   0.0    64  11
## 42    IV Golden Rain   0.2   103  11
## 43    IV Golden Rain   0.4   132  11
## 44    IV Golden Rain   0.6   133  11
## 45    IV  Marvellous   0.0    70  12
## 46    IV  Marvellous   0.2    89  12
## 47    IV  Marvellous   0.4   104  12
## 48    IV  Marvellous   0.6   117  12
## 49     V     Victory   0.0    62  13
```

```
## 50     V     Victory   0.2    90  13
## 51     V     Victory   0.4   100  13
## 52     V     Victory   0.6   116  13
## 53     V Golden Rain   0.0    80  14
## 54     V Golden Rain   0.2    82  14
## 55     V Golden Rain   0.4    94  14
## 56     V Golden Rain   0.6   126  14
## 57     V  Marvellous   0.0    63  15
## 58     V  Marvellous   0.2    70  15
## 59     V  Marvellous   0.4   109  15
## 60     V  Marvellous   0.6    99  15
## 61    VI     Victory   0.0    53  16
## 62    VI     Victory   0.2    74  16
## 63    VI     Victory   0.4   118  16
## 64    VI     Victory   0.6   113  16
## 65    VI Golden Rain   0.0    89  17
## 66    VI Golden Rain   0.2    82  17
## 67    VI Golden Rain   0.4    86  17
## 68    VI Golden Rain   0.6   104  17
## 69    VI  Marvellous   0.0    97  18
## 70    VI  Marvellous   0.2    99  18
## 71    VI  Marvellous   0.4   119  18
## 72    VI  Marvellous   0.6   121  18
```

Same model as fm1

```
(fm1b<-lmer(yield~nitro+Variety+(1|Block)+(1|Plt),Oats))
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: yield ~ nitro + Variety + (1 | Block) + (1 | Plt)
##    Data: Oats
## REML criterion at convergence: 578.8918
## Random effects:
##  Groups   Name        Std.Dev.
##  Plt      (Intercept) 10.44
##  Block    (Intercept) 14.65
##  Residual             12.87
## Number of obs: 72, groups:  Plt, 18; Block, 6
## Fixed Effects:
##       (Intercept)              nitro  VarietyMarvellous
##            82.400             73.667              5.292
##    VarietyVictory
##            -6.875
```

Model is

$$Y_{ijk} = \mu + \tau_{v(i,j)} + \beta x_{ijk} + b_i + b_{ij} + \epsilon_{ijk}$$

where

$$\begin{aligned} i &= 1, ..., 6 : \text{block} \\ j &= 1, 2, 3 : \text{plot} \\ k &= 1, 2, 3, 4 : \text{subplot} \\ v(i,j) &= 1, 2, 3 \text{ corresponding to variety (eg. v(1,2)=2 for "Golden Rain")} \\ x_{ijk} &\text{ nitro level} \end{aligned}$$

```r
summary(fm1b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: yield ~ nitro + Variety + (1 | Block) + (1 | Plt)
##    Data: Oats
##
## REML criterion at convergence: 578.9
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.62948 -0.65841 -0.07207  0.55785  1.71463
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Plt      (Intercept) 108.9    10.44
##  Block    (Intercept) 214.5    14.65
##  Residual             165.6    12.87
## Number of obs: 72, groups:  Plt, 18; Block, 6
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)         82.400      8.059  10.225
## nitro               73.667      6.781  10.863
## VarietyMarvellous    5.292      7.079   0.748
## VarietyVictory      -6.875      7.079  -0.971
##
## Correlation of Fixed Effects:
##             (Intr) nitro  VrtyMr
## nitro       -0.252
## VartyMrvlls -0.439  0.000
## VarityVctry -0.439  0.000  0.500
```

```r
ranef(fm1b)
```

```
## $Plt
##    (Intercept)
## 1    14.459851
## 2     2.412050
## 3    -3.959093
## 4    -9.444021
## 5     4.415479
## 6     6.378153
## 7    -6.228697
## 8    -8.130981
## 9    11.042837
## 10   -1.028299
## 11    5.946743
## 12   -7.308857
## 13    3.592916
## 14    1.147122
## 15  -10.115609
## 16   -1.351749
## 17   -5.790413
## 18    3.962569
```

```
##
## $Block
##     (Intercept)
## VI    -6.259694
## V    -10.582936
## III   -6.529897
## IV    -4.706029
## II     2.656993
## I     25.421564
```

Note: can't have an ordinary linear model with plot effect, as confounded with Variety

```
(lm1<-lm(yield~nitro+Variety+Plt ,Oats, contrasts=list(Plt=contr.sum)))
```

```
##
## Call:
## lm(formula = yield ~ nitro + Variety + Plt, data = Oats, contrasts = list(Plt = contr.sum))
##
## Coefficients:
##       (Intercept)              nitro  VarietyMarvellous
##            68.150             73.667             41.917
##    VarietyVictory               Plt1               Plt2
##            -0.750             53.500             43.000
##              Plt3               Plt4               Plt5
##            -2.417             -2.250             23.000
##              Plt6               Plt7               Plt8
##           -10.917             -7.000             -3.500
##              Plt9              Plt10              Plt11
##           -13.667              2.000             17.750
##             Plt12              Plt13              Plt14
##           -37.167              2.500              5.250
##             Plt15              Plt16              Plt17
##           -46.917                 NA                 NA
```

The 'equivalent' fixed effects model can be achieved by having a Variety:Block interaction Gives same estimates for Variety effects, but with smaller standard errors.

```
lm1<-lm(yield~nitro+Variety*Block ,Oats, contrasts=list(Block=contr.sum))
summary(lm1)
```

```
##
## Call:
## lm(formula = yield ~ nitro + Variety * Block, data = Oats, contrasts = list(Block = contr.sum))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.9000  -8.1875   0.6417   6.6083  22.6167
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 82.400      3.322  24.803  < 2e-16 ***
## nitro                       73.667      6.781  10.863  4.3e-15 ***
## VarietyMarvellous            5.292      3.714   1.425  0.16012
## VarietyVictory              -6.875      3.714  -1.851  0.06976 .
## Block1                     -14.250      5.873  -2.426  0.01868 *
## Block2                      -9.000      5.873  -1.532  0.13136
```

```
## Block3                    -17.750      5.873  -3.022  0.00386 **
## Block4                      3.500      5.873   0.596  0.55374
## Block5                      8.750      5.873   1.490  0.14218
## VarietyMarvellous:Block1   13.458      8.306   1.620  0.11108
## VarietyVictory:Block1       6.125      8.306   0.737  0.46410
## VarietyMarvellous:Block2  -15.542      8.306  -1.871  0.06684 .
## VarietyVictory:Block2       3.375      8.306   0.406  0.68612
## VarietyMarvellous:Block3   26.458      8.306   3.186  0.00242 **
## VarietyVictory:Block3       2.625      8.306   0.316  0.75320
## VarietyMarvellous:Block4  -18.292      8.306  -2.202  0.03201 *
## VarietyVictory:Block4      -9.625      8.306  -1.159  0.25171
## VarietyMarvellous:Block5    2.708      8.306   0.326  0.74564
## VarietyVictory:Block5     -19.125      8.306  -2.303  0.02525 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.87 on 53 degrees of freedom
## Multiple R-squared:  0.8312, Adjusted R-squared:  0.7739
## F-statistic:  14.5 on 18 and 53 DF,  p-value: 1.676e-14
```