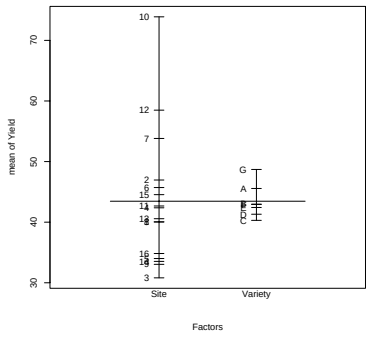
Linear Models: Unreplicated two-way analysis of variance, and comparison with one-way analysis

The `niab` data set records yields of potatoes of 7 different varieties, each grown at 16 Sites in the UK.

```
> niab
    Yield Site Variety
1    36.6    1       A
2    39.2    1       B
3    38.2    1       C
4    37.4    1       D
...
111  36.5   16       F
112  38.8   16       G
```

A simple display of the mean response for each level of each factor is obtained by `plot(niab)`:



We fit a factorial model without interactions:

```
> fit<-lm(Yield~Site+Variety)
> anova(fit)
Analysis of Variance Table

Response: Yield
          Df  Sum Sq Mean Sq F value    Pr(>F)
Site      15 12892.3   859.5 24.2712 < 2.2e-16 ***
Variety    6   769.1   128.2  3.6197  0.002913 **
Residuals 90  3187.1    35.4
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

This is an example of the ANOVA table on slide 64, except that the Total row is omitted from the **R** output.

Both `Site` and `Variety` effects are highly significant.

If we had mistakenly thought that we could fit interactions as well, the ANOVA table would have clearly indicated the error:

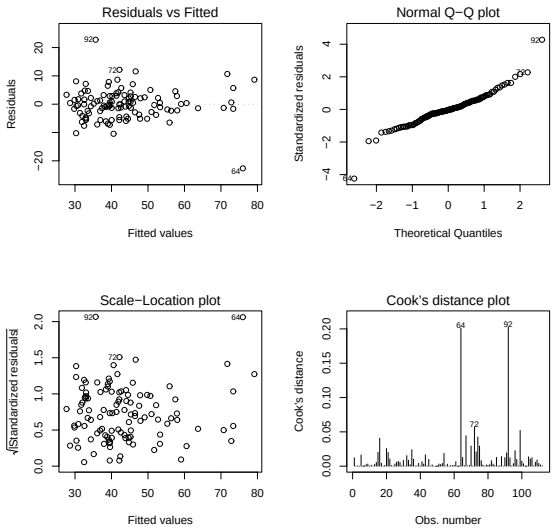
```
> fit2<-lm(Yield~Site*Variety)
> anova(fit2)
Analysis of Variance Table

Response: Yield
             Df  Sum Sq Mean Sq F value Pr(>F)
Site         15 12892.3   859.5
Variety       6   769.1   128.2
Site:Variety 90  3187.1    35.4
Residuals     0     0.0
```

A diagnostic plot to check all is well:

```
> par(mfrow=c(2,2))
> plot(fit)
```

There is no evidence of pattern in the fitted values/residuals plot. Finally, let us see



what happens if we think that because we are really only interested in the effects of the varieties we could do a one-way analysis instead:

```
> fit3<-lm(Yield~Variety)
> anova(fit3)
Analysis of Variance Table

Response: Yield
           Df  Sum Sq Mean Sq F value Pr(>F)
Variety     6   769.1   128.2   0.837  0.544
Residuals 105 16079.3   153.1
```

Variety effects are no longer significant – the (incorrect) one-way analysis is less sensitive, as explained in the notes.