ANOVA

Tutorial day 7

ANOVA

= Analysis of variance

* all explanatory variables are categorical

Terminology

* explanatory variables are called factors

* one-way anova

= one factor with two or more levels

* two-way/three-way anova

= two or more factors with two or more levels

* factorial design

= replication of each combination of levels in a multi-way ANOVA (allows to study interaction of variables)

* split-plot design

= experiments have different treatments applied to plots of different sizes

Example of one way ANOVA

* y= yield
* x= soil
* k=three soil types
* n = 10

* **degrees of freedom:**

= k * (n – 1)

= 3 * (10 – 1)

= 27

Factor: soil

n=10

n=10

n=10

Response: yield

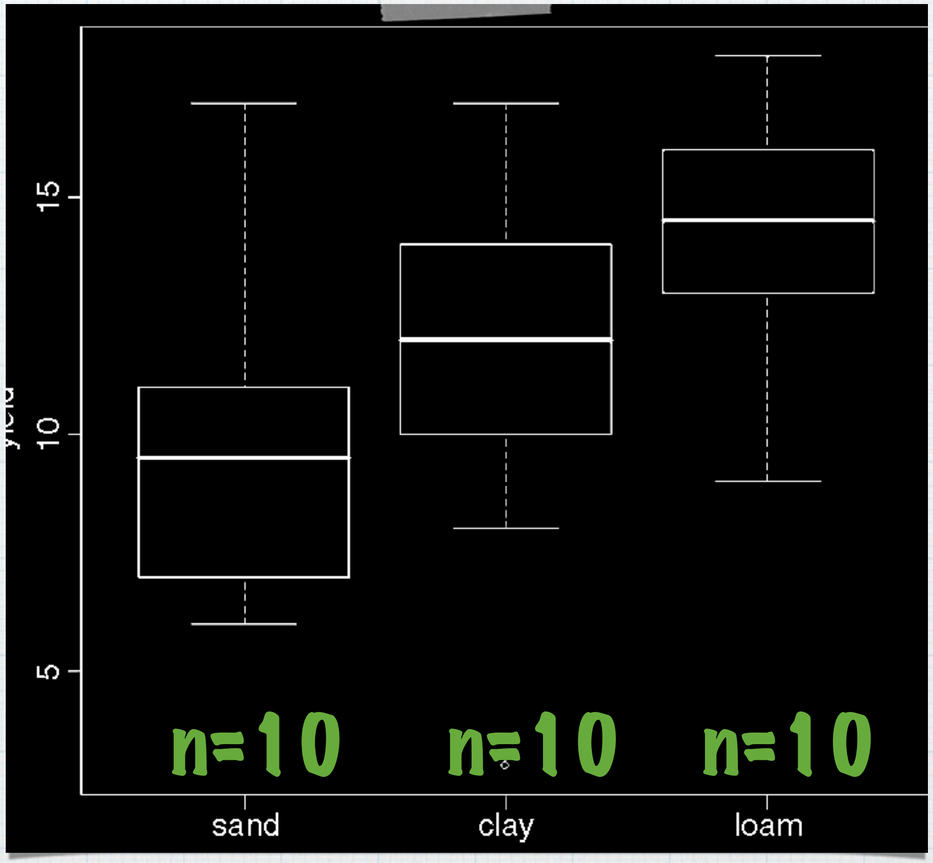
Balanced design & constancy of variance (homoscedasticity)

* fligner.test(yield~soil)

Fligner-Killeen test of homogeneity of variances

data: yield by soil

Fligner-Killeen:med
chi-squared = 0.3651, df = 2, p-value = 0.8332

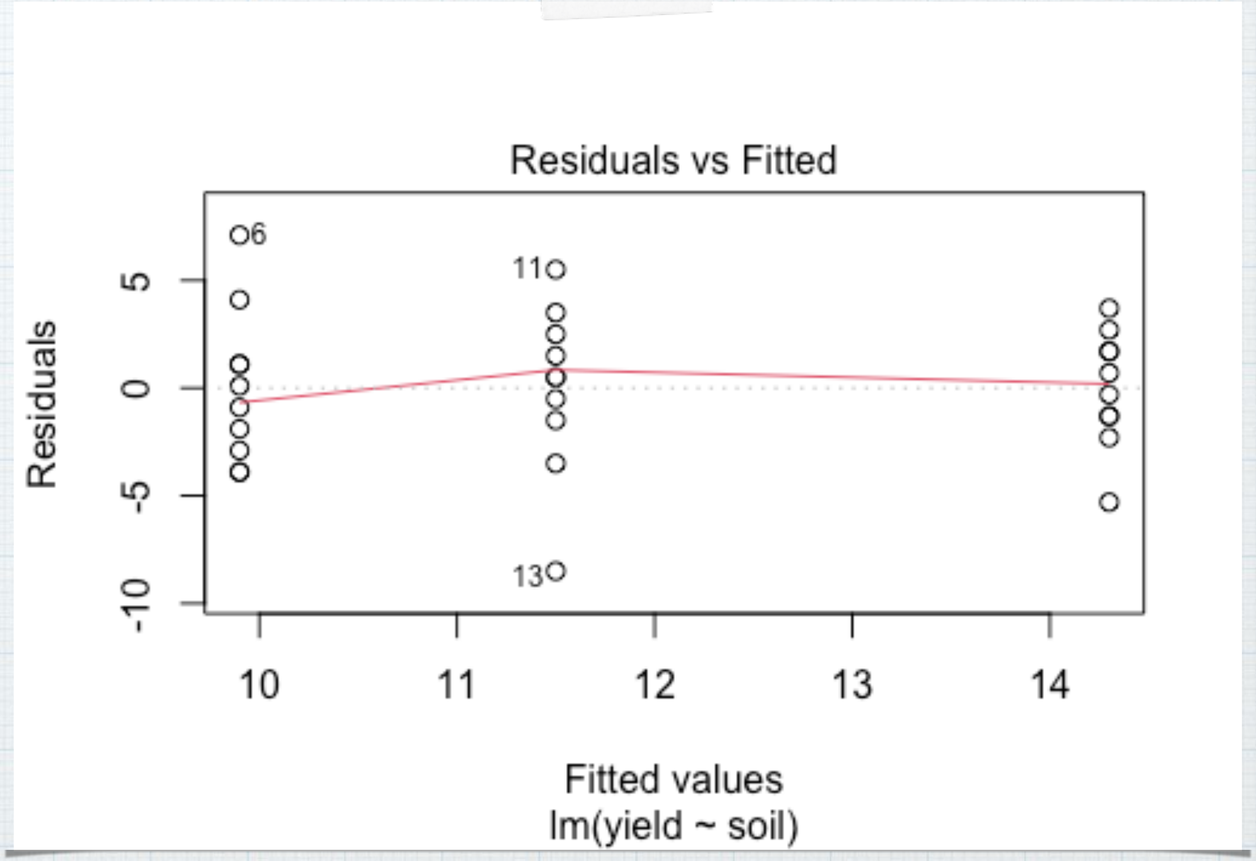


Constancy of variance (homoscedasticity)

* visual inspection

plot(model)

* there should be no pattern in the residuals against the fitted values

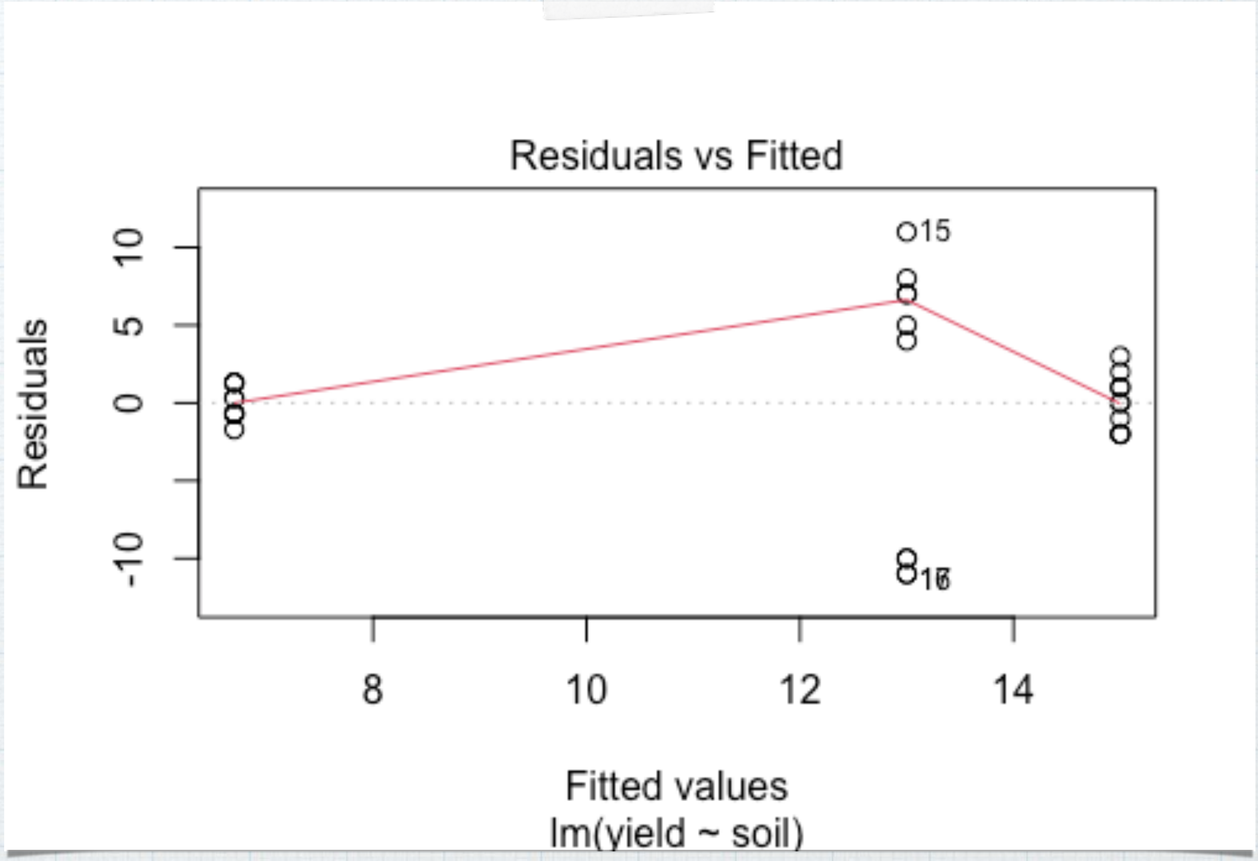


Variance not constant

* visual inspection

plot(model)

* this is a clear pattern



Example of factorial design

Response: growth

Factor: fodder

Factor: supplement

control

supergain

agrimore

supersupp

Example of factorial design

- control
- supergain
- agrimore
- supersupp

- control
- supergain
- agrimore
- supersupp

- control
- supergain
- agrimore
- supersupp

how many factor levels?

control
supergain
agrimore
supersupp

control
supergain
agrimore
supersupp

control
supergain
agrimore
supersupp

Example of factorial design

* allows to test for interactions
* how does fodder effect weight gain of farm animals?
* how does supplement effect weight gain?
* Do the factors interact?

Interaction

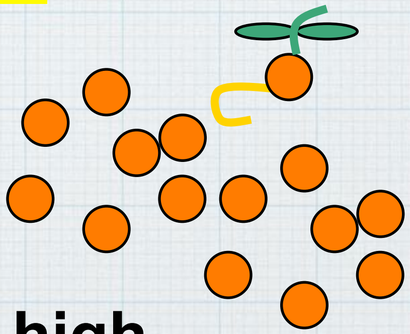
* the response of one factor depends upon the level of the other factor
* in our example: an interactive effect could be that our farm animals grow faster with supersupp and oat, but not with supersupp and wheat or barley

Example of split plot experiment

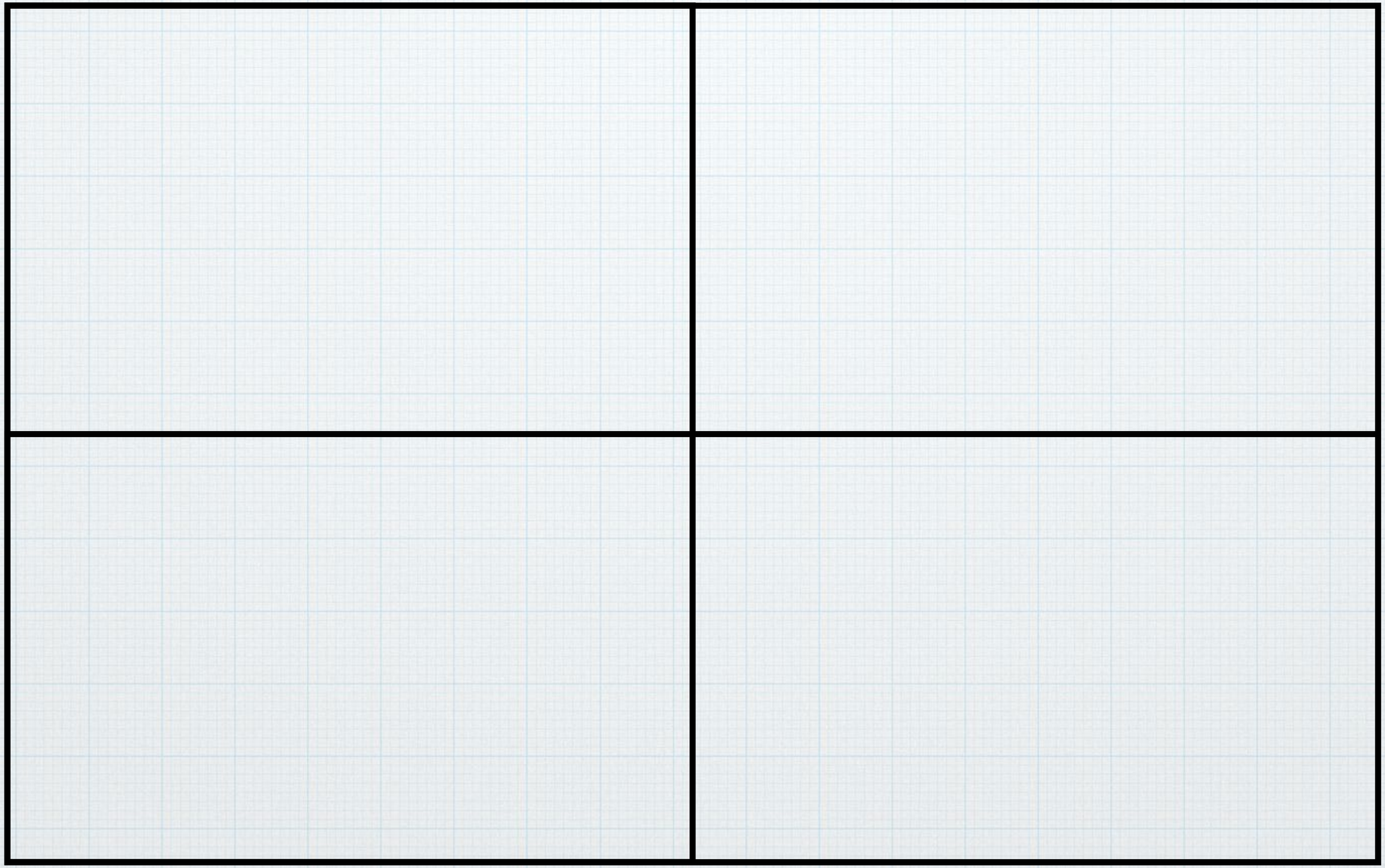
Response: yield

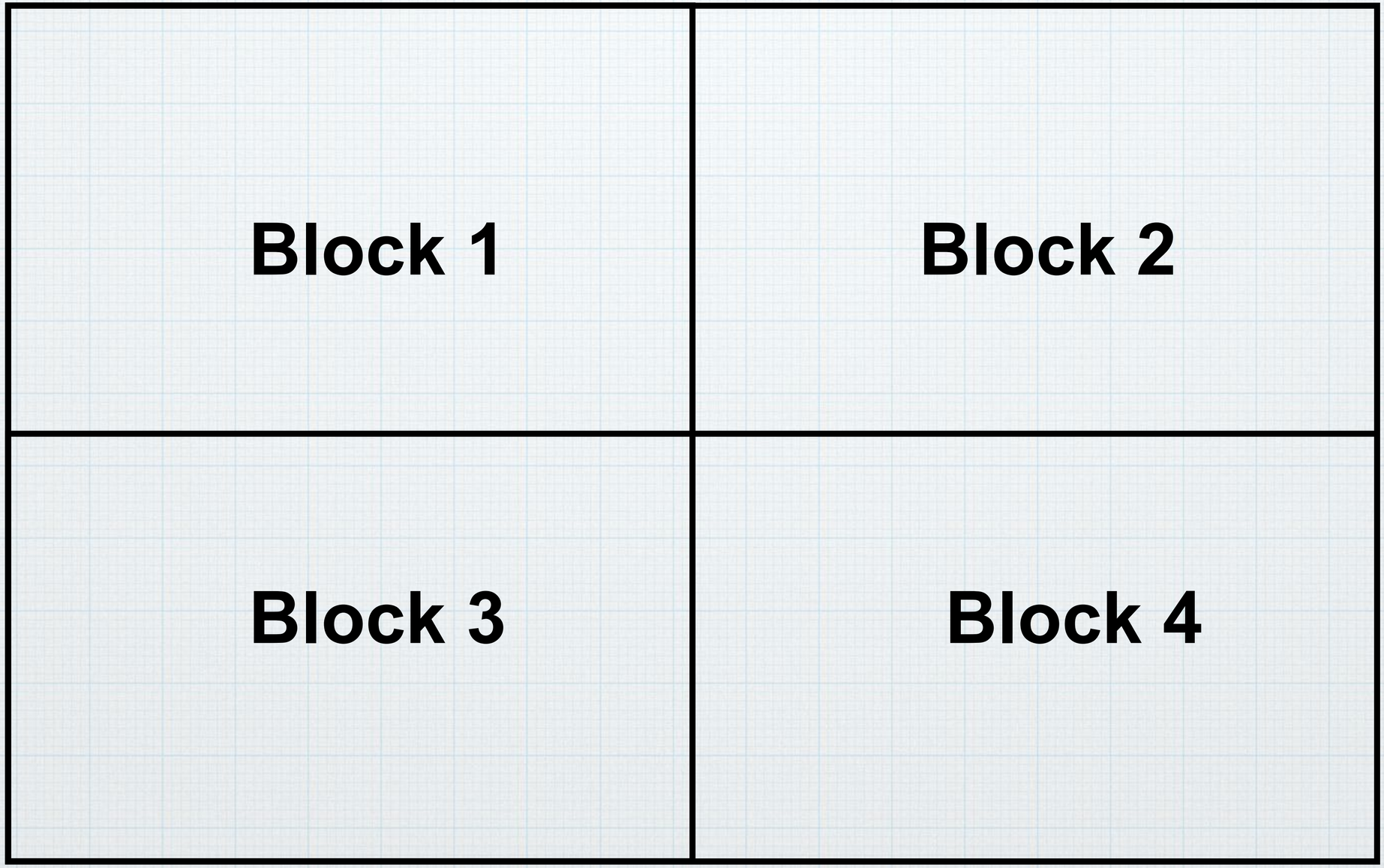
yes
no

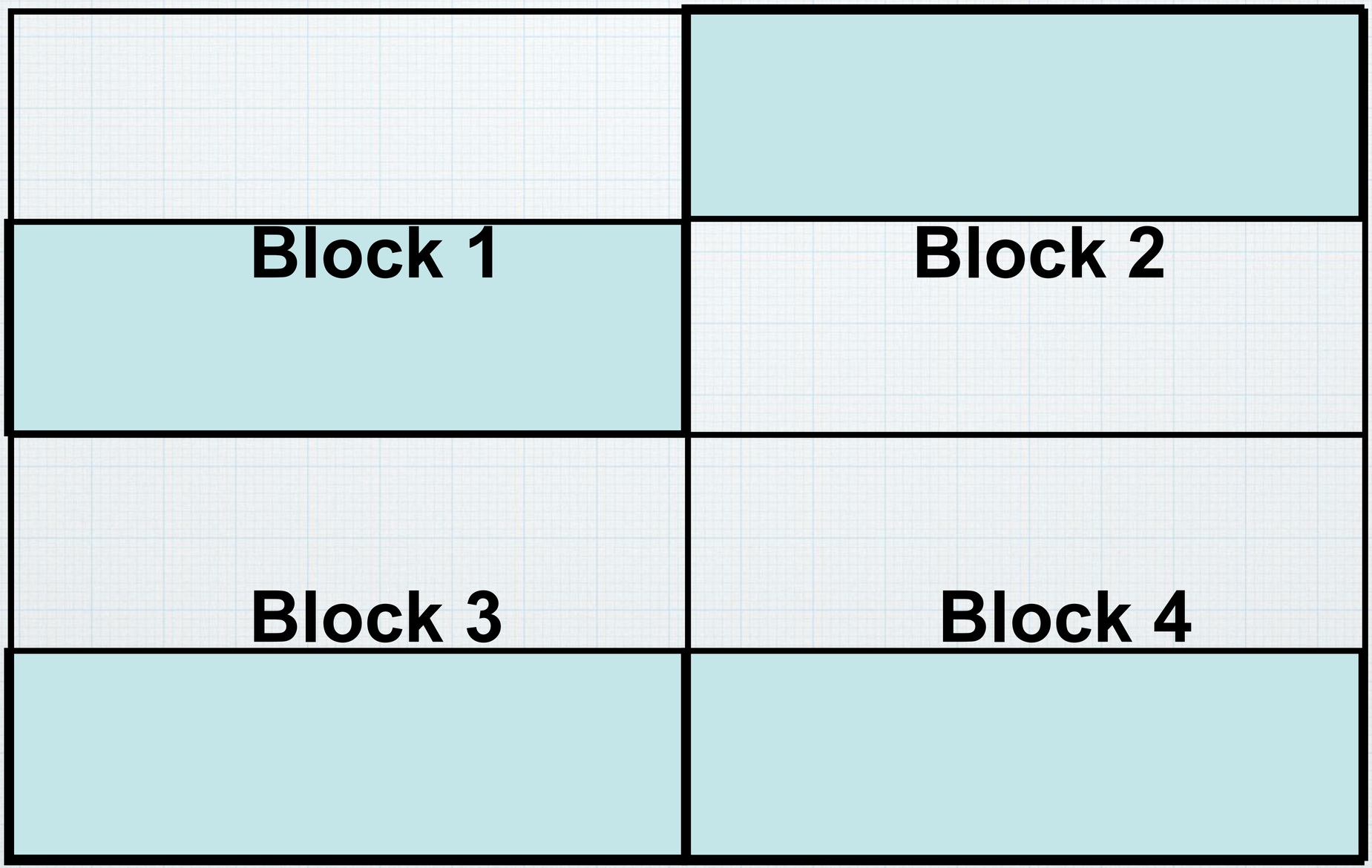
N, P, NP

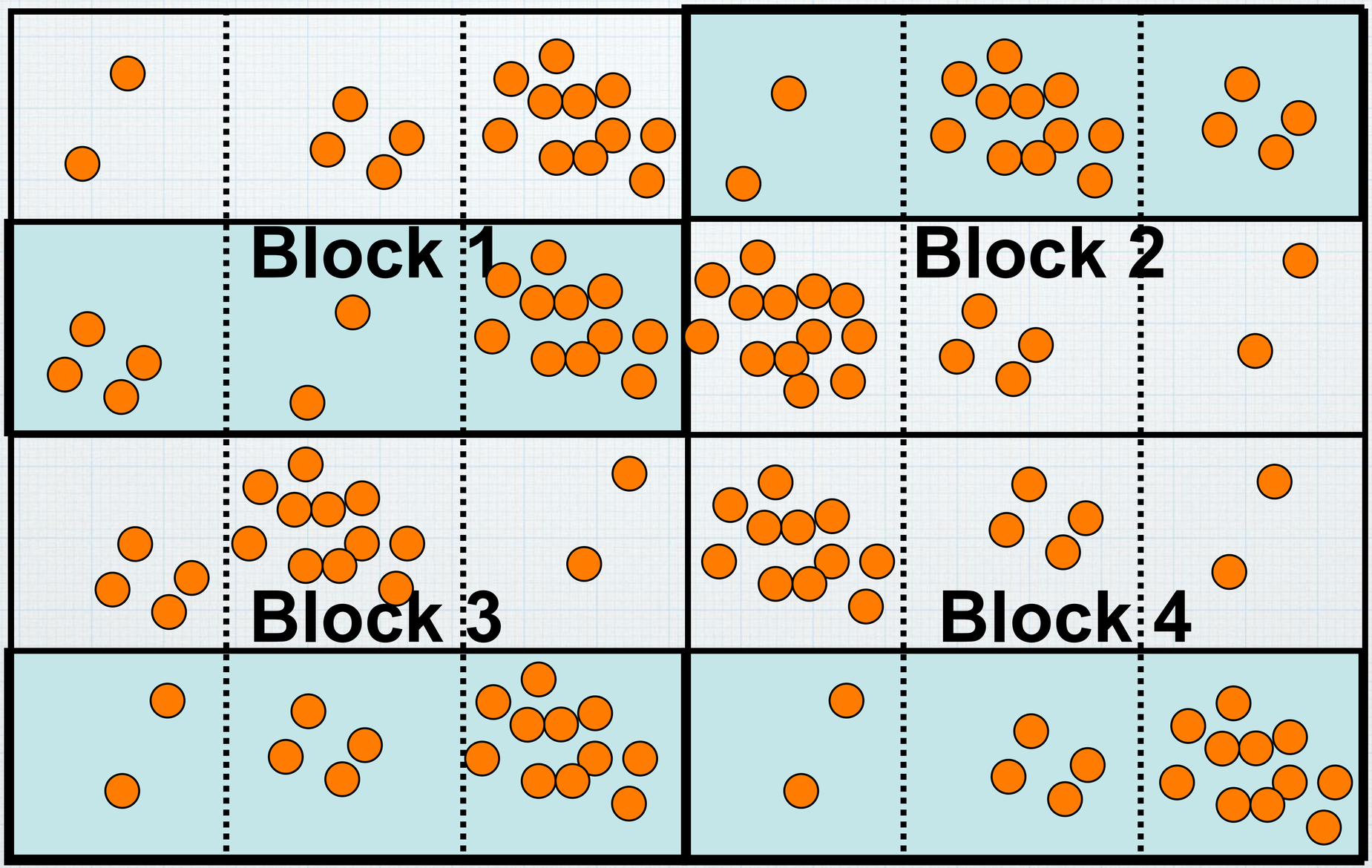


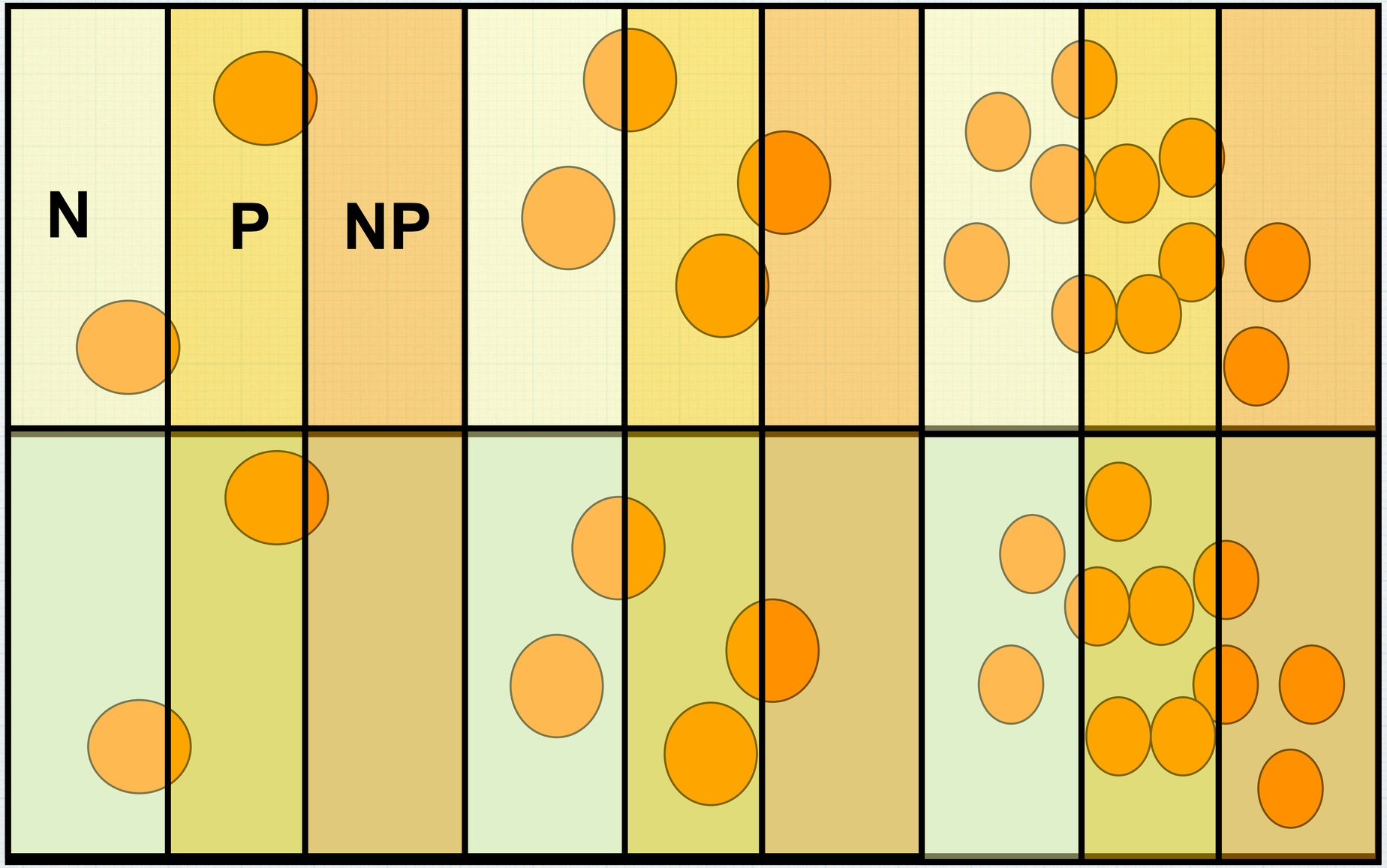
low, medium and high seed density




Block 1
Block 2
Block 3
Block 4


Block 1
Block 2
Block 3
Block 4


Block 1
Block 2
Block 3
Block 4


N
P
NP

Split-plot experiment

* build the full model with the Error structure

* each different plot size is associated with its own error variance

*
```
model<-aov(yield~irrigation*density*
  fertilizer+Error(block/
  irrigation/density))
summary(model)
```

Take home message

* ANOVA has its own terminology (one-way, two-way....)
* all explanatory variables are categorical
* with an interactive effect the response of one factor depends upon the level of the other factor
* preconditions for an ANOVA: sufficient variance on each factor level, homoscedasticity, balanced design (same number of samples for each factor level), normal error structure (residuals)