

5. Repeatability estimation & other random effects

- Learn how to estimate how much variance is accounted for by different random effects
- Learn the difference between crossed and nested random effects
 - Where do their errors come from?
 - How to properly specify?

Estimating repeatability

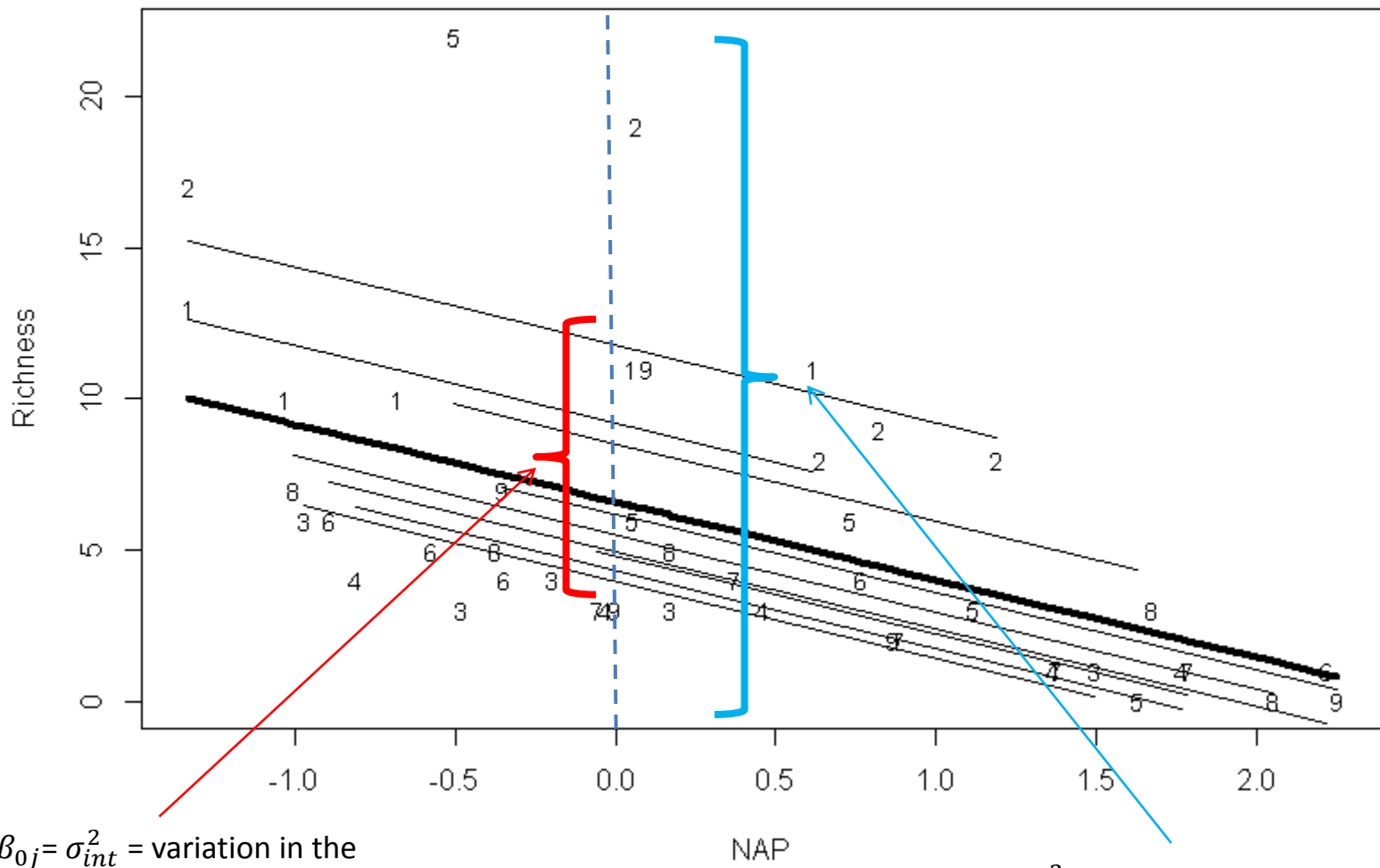
- Repeatability is the proportion of variance that is due to between-subject variance (i.e. intercepts)
- Also called the “intraclass correlation coefficient (ICC)”

$$y_{ij} = (\beta_0 + \beta_{0j}) + \beta_1 \times x_{ij} + \epsilon_{ij}$$

$$\beta_{0j} \sim N(0, \sigma_{int}^2)$$

$$\epsilon_{ij} \sim N(0, \sigma_e^2)$$

$$R = \frac{\sigma_{int}^2}{(\sigma_{int}^2 + \sigma_e^2)}$$



$\beta_{0j} = \sigma_{int}^2$ = variation in the intercepts among the beaches

$\epsilon_{ij} = \sigma_e^2$ = residual variation (everything not explained by fixed & random effects)

Estimating repeatability

```
> mod.final <- lme(Richness ~ NAP + fExposure, random = ~1|Beach,  
+                 method = "REML", data = RIKZ)  
> summary(mod.final)
```

Linear mixed-effects model fit by REML

Data: RIKZ

	AIC	BIC	logLik
	240.5538	249.2422	-115.2769

Random effects:

Formula: ~1 | Beach

(Intercept) Residual

StdDev: 1.907175 3.059089

Fixed effects: Richness ~ NAP + fExposure

	Value	Std.Error	DF	t-value	p-value
(Intercept)	8.601088	1.0594875	35	8.118158	0.0000
NAP	-2.581708	0.4883901	35	-5.286160	0.0000
fExposure11	-4.532777	1.5755610	7	-2.876929	0.0238

Correlation:

(Intr) NAP

NAP	-0.136
fExposure11	-0.655 -0.037

Standardized within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-1.5163203	-0.4815106	-0.1218700	0.2922854	3.8777562

Number of Observations: 45

Number of Groups: 9

$$R = \frac{\sigma_{int}^2}{(\sigma_{int}^2 + \sigma_e^2)}$$

Nested versus crossed effects

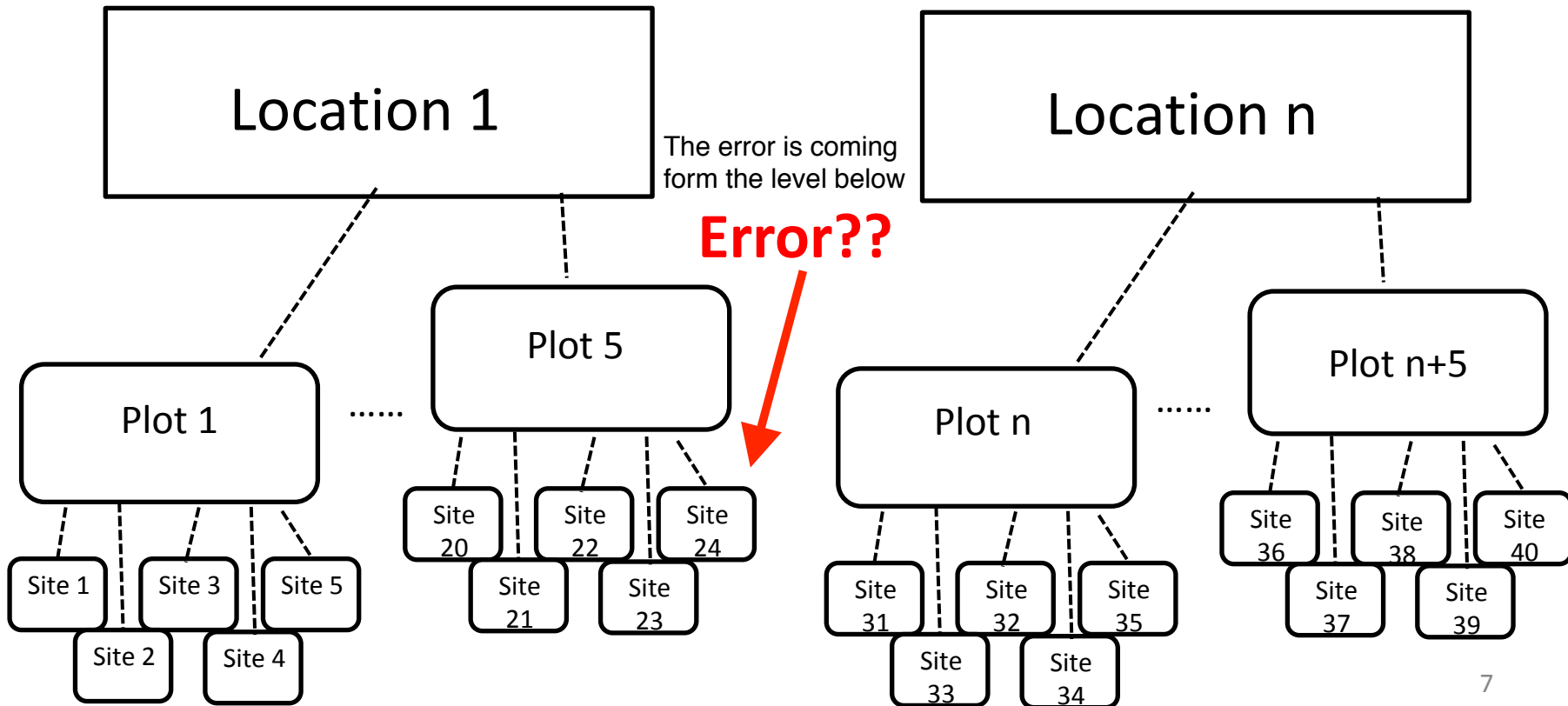
- Understand conceptually the difference between nested and crossed random effects
- Learn how to specify each in R (through proper labeling)
- Learn how to write models with nested or crossed effects in R using the nlme or lme4 packages

More than one random effect

- Sometimes you have more than one random effect in your model
- It is important to properly specify how these effects are related to each other in order to properly estimate your error for each effect
 - The good (and bad) thing is that most of the specification just comes in how you label the observations in R

Nested random effects

- Nested: each level of a random effect is ONLY present in one other level of a different effect



Crossed random effects

- Crossed: random effects that are not nested
 - Completely crossed: have an observation for each combination of the levels of each of the effects

The error/replicates from "A" is coming from the tanks
The error/replicates from "tanks" is coming from the different clones

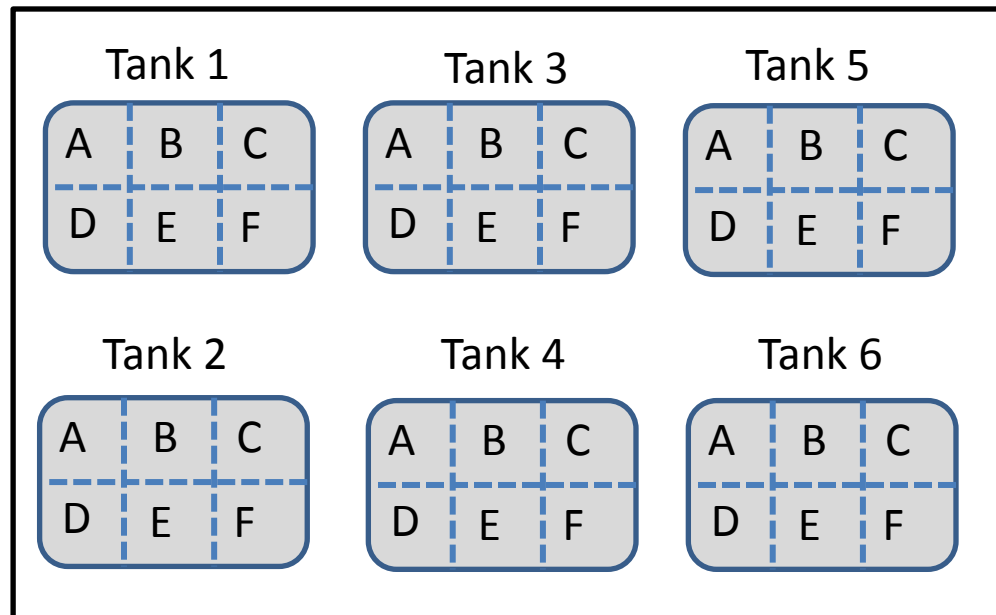
Clone is a random effect

Error???

Incubator

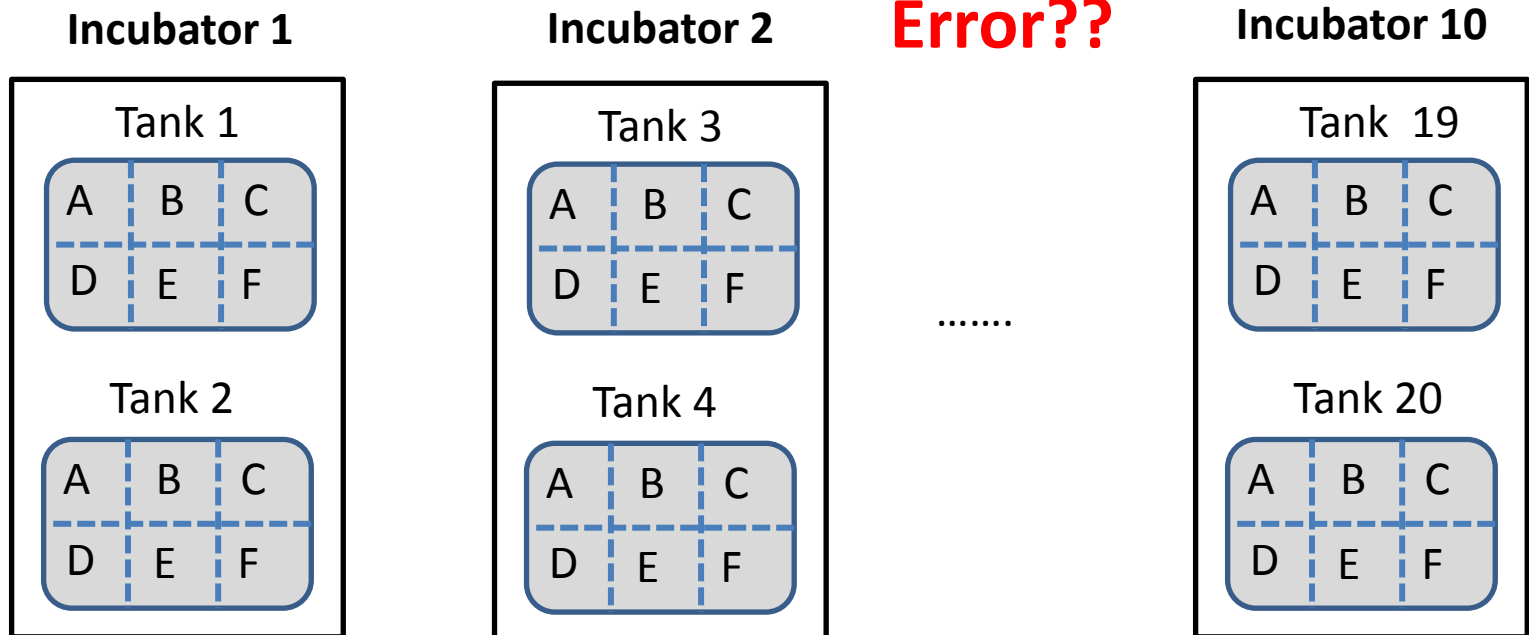
You want to assess variability in life history traits in 6 randomly selected Daphnia clones. Each clone is placed into a separate compartment in an aquarium and then 6 aquariums are placed in an incubator

Clone + Tank are CROSSED



Many random effects

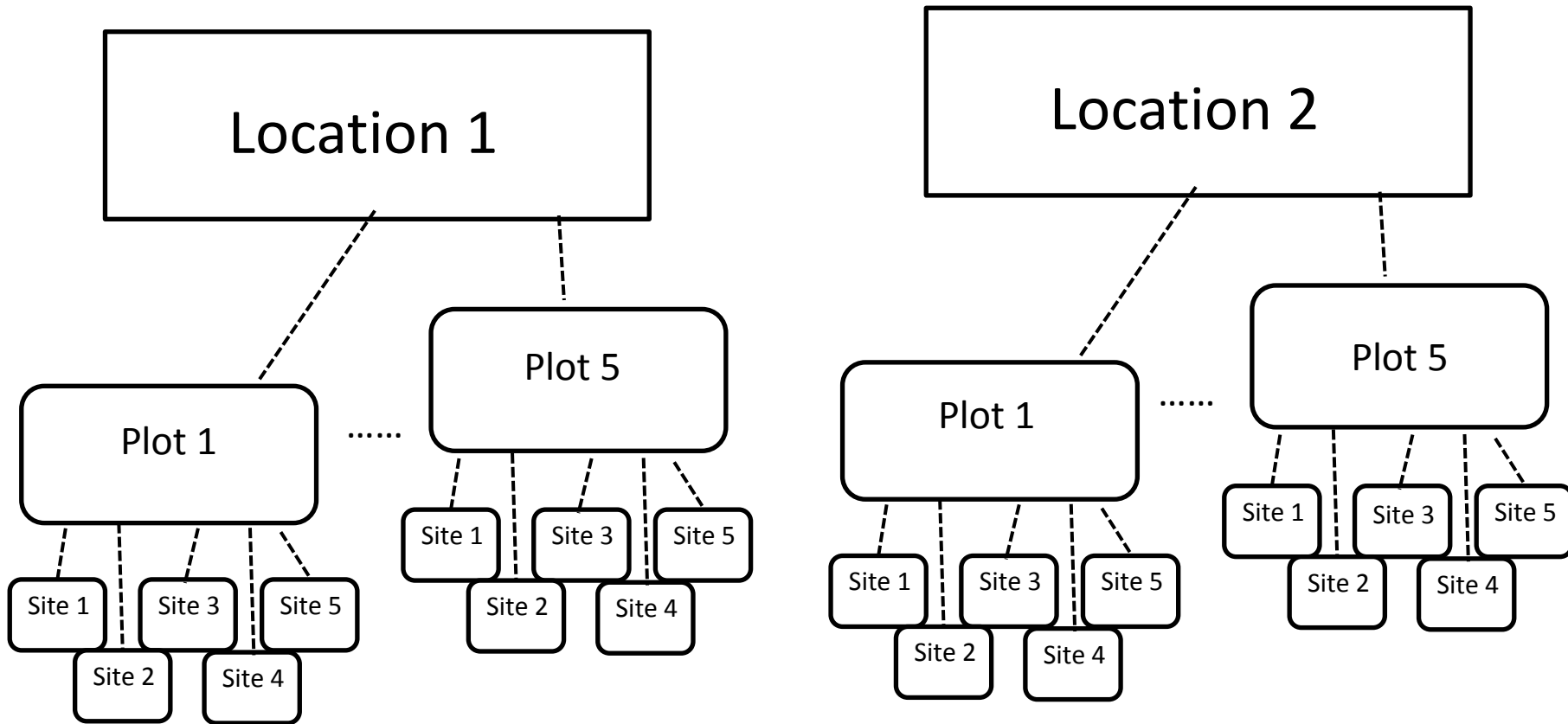
- Now you have multiple incubators into which you place the tanks
- Tank is nested within Incubator, but Clone is crossed with both of these
 - Each Clone is present in every Tank and Incubator, but each Tank is present in only one Incubator



Types of random effects

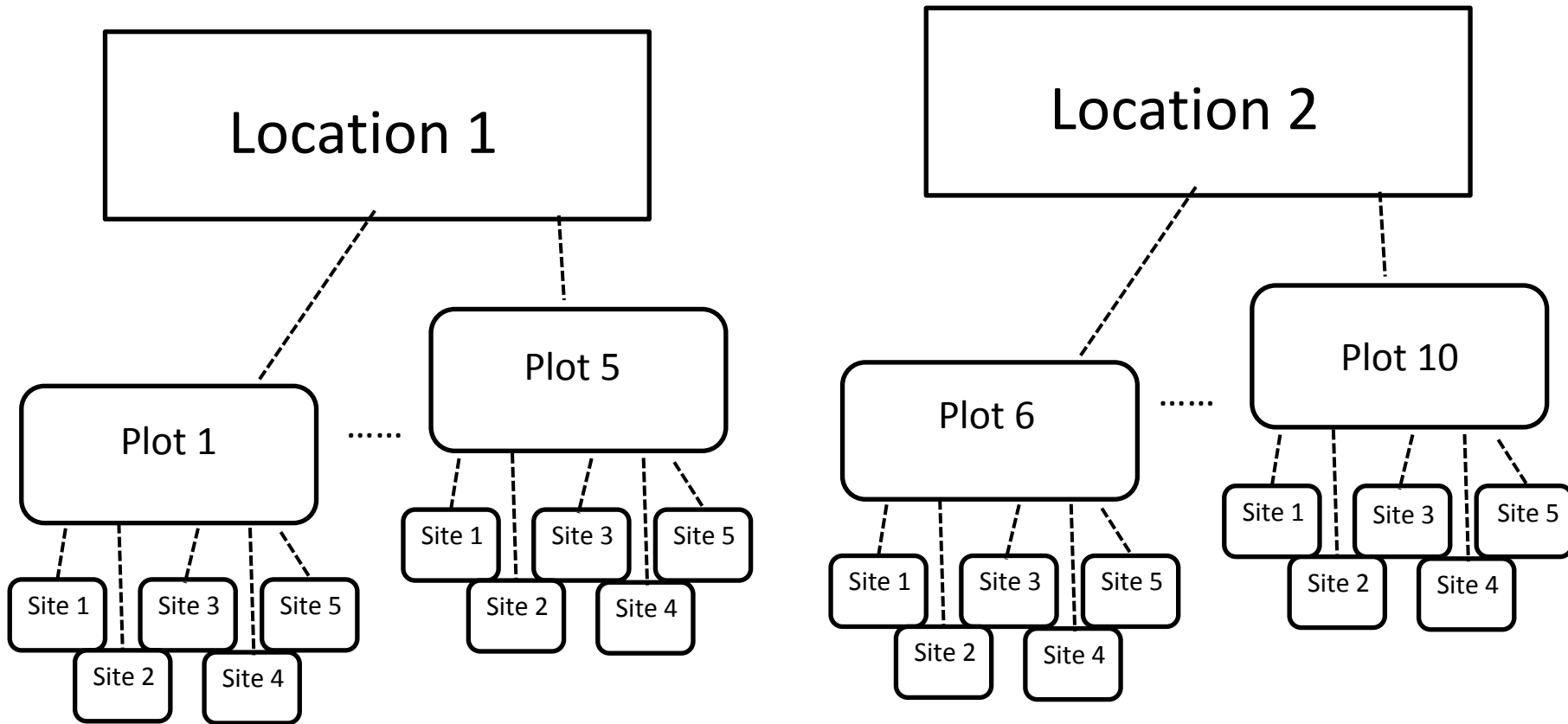
- R will only “know” whether something is nested/crossed based on the labeling!
 - Use unique labels for nested effects
 - Use the same labels for crossed effects

Labeling matters!



R will assume that Plots are CROSSED with Location because “Plot 1” is present in both Locations

Labeling matters!



R will assume that Plots are NESTED with Location because "Plot 1" is not present in both Locations

Labeling matters

Plots is CROSSED with location

Location	Plot	y
1	1	10.3
1	2	23.1
1	3	12.7
1	4	7.0
1	5	4.6
2	1	7.9
2	2	13.0
2	3	15.1
2	4	11.6
2	5	6.8

Plots are NESTED in Locations

Location	Plot	y
1	1	10.3
1	2	23.1
1	3	12.7
1	4	7.0
1	5	4.6
2	6	7.9
2	7	13.0
2	8	15.1
2	9	11.6
2	10	6.8

Draw the graph, write the model

- ‡ Water pollution is a major problem. Cities and states differ in the enforcement of water pollution limits. A group of researchers measured the effluent coming out of 20 randomly selected sewer drains in 35 cities across 8 states. They wanted to see whether the number of Waste Management Controllers employed by each city influenced the average effluent level. The variables are: effluent (liters per second), average education level and size (km

Draw the graph, write the model

- ‡ A researcher is interested in measuring how dams impact total aquatic vegetation biomass. The researcher picks 6 rivers that are dammed and measures along 10 transects above, and 10 transects below the dam. The researcher is not interested in particular plants and chooses 5 species at random to measure along each transect. She finds the following biomass (g/m²) along every transect.

Dealing with crossed effects in R

- lme function (in nlme package) can only handle nested effects
- Lmer function (in lme4 package) can handle crossed (and nested) effects
 - BUT – lmer does not return p-values, and there are no options to alter variance weights or correlation structures (which we will get to later)
 - Use lmerTest instead to get p-values

Age determination in cetaceans

‡ We wanted to test how comparable 3 different (tooth) staining techniques are for aging various species. We picked 6 species of whales, and sampled multiple individuals (25) per species. The species were collected from two different locations. From each individual we stained different teeth.

‡ What does our model look like?

‡ Draw the random effects schematic

Draw random effects here

Nested versus crossed effects

- Understand conceptually the difference between nested and crossed random effects
- Learn how to specify each in R (through proper labeling)
- Learn how to write models with nested or crossed effects in R using the nlme or lme4 packages