

RenR 480/711

Golden rules of data tables

Golden rules of data tables

Each row represents a unit

- All attributes (measurements) of a unit are usually in the same row
- Different units **must be** in different rows
- What is the unit?

Golden rules of data tables

Each row represents a unit

- All attributes (measurements) of a unit are usually in the same row
- Different units **must be** in different rows
- What is the unit?

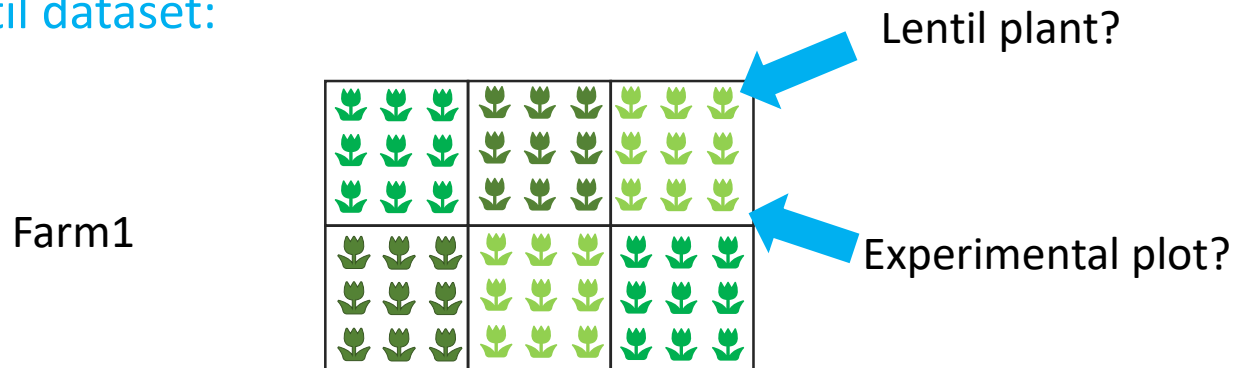
In the assignment, we assumed trees to be independent units

Golden rules of data tables

What is the unit?

Depends on design:

E.g. Lentil dataset:



The unit is the smallest independent entity you collect data for

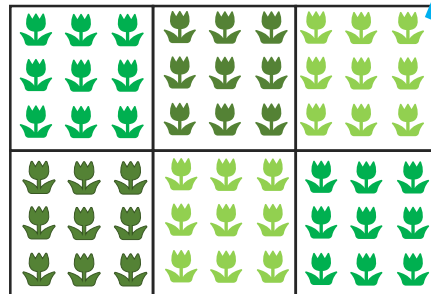
Golden rules of data tables

What is the unit?

Depends on design:

E.g. Lentil dataset:

Farm1



Lentil plant? No

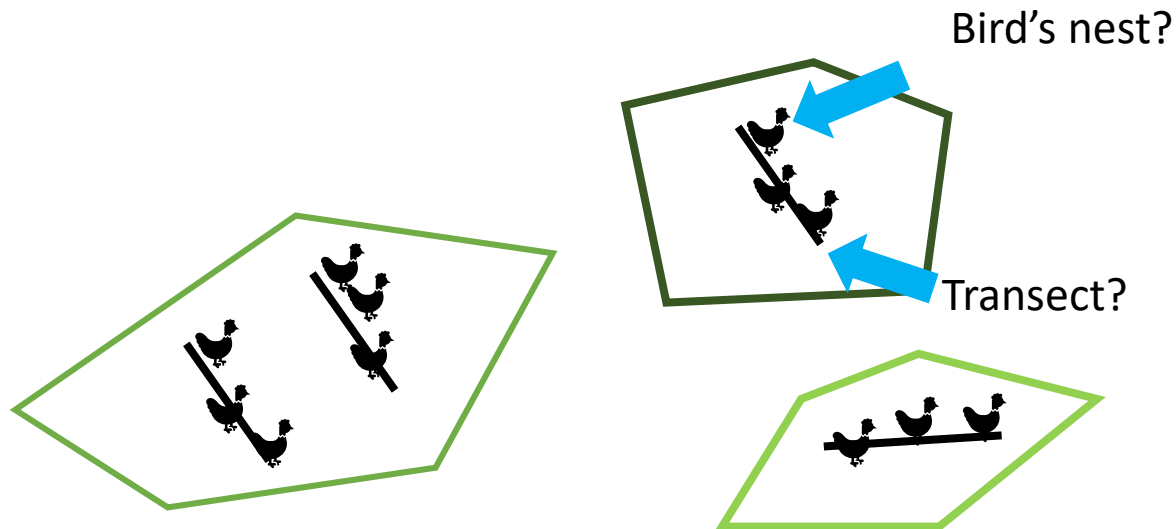
Experimental plot? Yes

The unit is the smallest independent entity you collect data for

Golden rules of data tables

What is the unit?

E.g. Forest transects

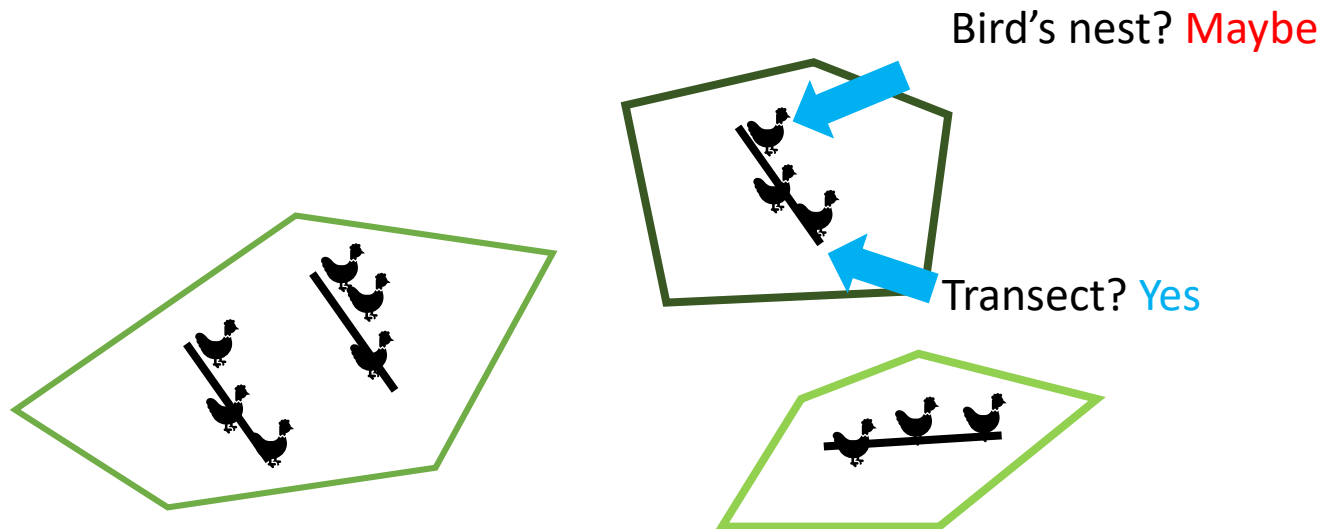


The unit is the smallest independent entity you collect data for

Golden rules of data tables

What is the unit?

E.g. Forest transects

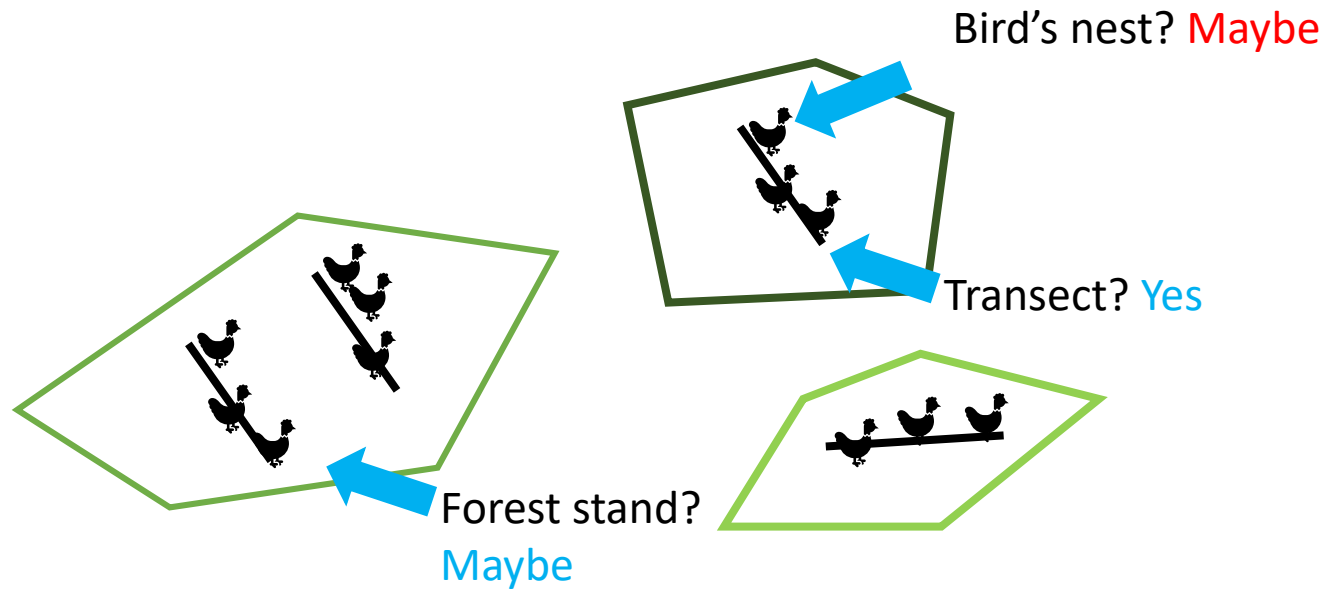


The unit is the smallest independent entity you collect data for

Golden rules of data tables

What is the unit?

E.g. Forest transects



The unit is the smallest independent entity you collect data for

Golden rules of data tables

UnitID	Crop Variety	Block	Yield	Resistance
1	A	x	5	0
2	A	y	3	1
3	A	z	6	0
4	B	x	12	0
5	B	y	15	0
6	B	z	14	1

Categorical descriptors

Data

Use letters not numbers

(combined letters and numbers are ok)

Golden rules of data tables

Use multiple concise, descriptive IDs

Forest_ID	Transect_ID	BA
AB_CR_PC	AB_CR_PC_T1	15
AB_CR_PC	AB_CR_PC_T2	13
AB_CR_PM	AB_CR_PM_T1	35
AB_CR_PM	AB_CR_PM_T2	41

No special characters, no spaces, no units