L2 Cultural sharing - practical

File 1: Plant knowledge

- File 1: *plant_knowledge.xlsx*
- This is an incidence matrix:
  - X individuals by Y plants

- Entry (i, j) indicates whether individual i knows plant j

- Exercise:

- Think of other examples of incidence matrices

File 2: Plant participants

- File 2: *plant_participants.xlsx*

- File with info on individuals
  - ID
  - Camp of residence
  - Sex
  - Age (cohort)
  - Place of birth (*born, born_cluster*)

- Exercise: based on the two files, which questions about cultural knowledge or transmission could you ask?

Individual and group knowledge

- Plants:
  - Are plants equally known?
  - Distribution of plant knowledge

- Subjects:
  - Is there significant variation in plant knowledge across individuals?

- Relation between individuals and plant knowledge
  - Is there an effect of sex, age, birth place on plant knowledge?
    - Group comparisons
    - Total score: linear regression
    - One plant at a time (subject knows/doesn’t know a plant: logistic regression
    - All plants: mixed effects logistic regression controlling for subject ID

Shared knowledge

- In addition to analysing individual knowledge and group differences, we can investigate shared knowledge as a proxy for knowledge transmission

- Analysing shared knowledge at the simplest level requires creation of dyads (individual pairs)
  - advantages?
  - limitations?

- Dyads become datapoints
  - Dyadic characteristics must be defined based on the two individuals

Exercise:

Based on file *plant_participants*, how can you define

i) dyadic sex?

ii) dyadic age?

iii) dyadic birth place?

Defining dyadic characteristics

- We must create variables with dyadic
  - ID
  - sex
  - etc

Exercise:

- Think of a case where dyadic characteristics may differ between/among
  - female-female vs. female-male vs. male-male
  - same sex (male-male or female-female) vs. different sex (female-male)

Plant knowledge

- We can encode shared knowledge by dyads based on individual knowledge

- We start with an *incidence* matrix
  - rows: subjects S
  - columns: characteristics C

	C1	C2	C3	C4
S1	1	0	0	0
S2	1	1	0	0
S3	1	1	1	0
S4	0	1	1	0
S5	0	0	1	1
S6	0	0	0	1

- Let's create a fake plant knowledge matrix where
  - Rows are individuals
  - Columns are medicinal plant species
  - Values: 1 (subject i knows plant j) or 0 (doesn't know it)

Shared (dyadic) knowledge

- We want to create all possible dyads
  - avoiding repetition
    - Links are not directional
    - (a,b) = (b,a)

- See code

Exercise: run very simple and preliminary analyses on the *m.dyads* file

- Using proportion tests, is there a difference in shared knowledge as a function of
  - dyadsex
  - samesex
- Run a logistic regression to predict shared knowledge as a function of
  - dyadsex
  - samesex
- Which other ways could you analyse this dataset?

Task for this week:

Start working on a dataset of shared plant knowledge using real data from files

- *plant_knowledge*
- *plant_participants*

- (This is the dataset you will require for Report 1!)