

## Data Import to R

- Read “*R Data Import/Export*”
- Import from SPSS, SAS *etc.*: Standard library foreign
- CEP data: Function `read.cep` in library `vegan`
- Spreadsheet:
  1. Edit for R
  2. Save as *comma separated values* (csv)
  3. Read with `read.csv` (decimal points, separators “,”) or `read.csv2` (decimal commas, separators “;”)

## Preparing a spreadsheet

The screenshot shows a spreadsheet window titled 'dune.xls : Gnumeric'. The data table is as follows:

	A	B	C	D	E	F	G	H
1		Belper	Empnig	Junbuf	Junart	Airpra	Elepal	Rumace
2	2	3	0	0	0	0	0	0
3	13	0	0	0	3	0	0	0
4	4	2	0	0	0	0	0	0
5	16	0	0	0	3	0	0	8
6	6	0	0	0	0	0	0	0
7	1	0	0	0	0	0	0	0
8	8	0	0	0	4	0	0	4
9	5	2	0	0	0	0	0	0

## Comma separated values

```
Belper,Empnig,Junbuf,Junart,Airpra,Elepal,Rumace
2,3,0,0,0,0,0,0
13,0,0,3,0,0,0,0
4,2,0,0,0,0,0,0
16,0,0,0,3,0,8,0
6,0,0,0,0,0,0,6
1,0,0,0,0,0,0,0
8,0,0,0,4,0,4,0
5,2,0,0,0,0,0,5
```

**NB** you may have to remove a comma (,) before the first species name

## Community data and Environmental data

- Best to keep in two separate files
- The order of sites must be identical in both data sets
- Factor variables best coded with informative names
- Do **not** use uninformative numbers or “dummy” variables
- Species data must be numeric: use zero (0) for missing species, since blanks default to missing values (NA).

## Names

- It is best to name both variables (columns) and observations (rows), or R will make up uninformative names for both
- Names should be informative but short to avoid congestion
- Convention for species names: 4 + 4 from generic and specific names (*Asiootus*, *Tytoalba*, *Bubobubo*)
- Avoid gaps (blanks), mathematical symbols (+-\*:\*^), accented characters (øãñ) and underscore (\_) — but “.” is OK
- R will change non-conforming names (`make.names`)