

Reading Data in R

Module 3

Reading a File

From a CSV file

`read.csv()` reads a comma separated file and returns a data frame.

```
penguins <- read.csv("penguins.csv")
str(penguins)
```

The `str()` call prints the structure of the object, so you can check the column types before going further.

Two arguments matter more than the others.

- `stringsAsFactors` controls whether character columns become factors.
- `na.strings` lists the values to read as missing.

From an Excel file

R has no built-in reader for Excel files, so you need a package.

```
readxl::read_excel("penguins.xlsx", sheet = "data")
```

Common Pitfalls

Separators and decimal marks

Function	Separator	Decimal mark
<code>read.csv()</code>	,	.
<code>read.csv2()</code>	;	,
<code>read.delim()</code>	tab	.

Pick the reader that matches the file you were given, rather than reading the file and repairing it afterwards.

Checking your work

Always print the dimensions of what you read.

```
dim(penguins)
```