ESTP R - Reading and writing data

Reading text files

Reading

- `read.csv`, Comma separated dot is decimal
- `read.csv2`, Semicolon separated, comma is decimal
- `read.table`, all types of text tabular format

Writing

- `write.csv`, Comma separated dot is decimal
- `write.csv2`, Semicolon separated, comma is decimal
- `write.table`, all types of text tabular format

```
dat <- read.csv("myfile.csv")
write.csv2(dat, "yourfile.csv", row.names=FALSE)
```

File names in R

- File names are always between single or double quotes.
- File name can also be an `url`.
- Always use forward slash "/" e.g.

```
dat <- read.csv("C:/users/joe/documents/foo.csv")
```

Tip

Always work in a Rstudio project. The standard working directory of R is the project directory.

- use relative paths ("./")

E.g `C:/projects/foo`. Read data with `C:/projects/foo/data.csv` inlezen met

```
x <- read.csv("data.csv")
```

JSON

- Java script Object Notation: very popular on websites
- Used in Web API, retrieving data via the web
- Useful for reading data from web sources

```
{"age": 34,
 "income": ["2300", "121"]
}
```

JSON (R)

- `jsonlite` package makes it very handy to read and write json files
- allows automatic transformation from and to R data structures.

```
jsonlite::toJSON(iris[1,])
```

```
## [
##   {
##     "Sepal.Length": 5.1,
##     "Sepal.Width": 3.5,
##     "Petal.Length": 1.4,
##     "Petal.Width": 0.2,
##     "Species": "setosa"
##   }
## ]
```

SDMX

- **Format** used by Eurostat and National Banks
- Also defines **webservice** specification
- Both data as well as metadata publication
- uses XML as storage format

```
<document>
  <age>34</age>
  <income>2300</income>
  <income>121</income>
</document>
```

SDMX (R)

- Package `rsdmx` reads sdmx files
- Package `xml2` for generic xml files.

```
library(rsdmx)
readSDMX("<link to sdmx>")
```