

# Lectura de datos

## Clase 4

**Joaquin Cavieres G.**

Estudiante doctorado

`j.cavieres.g@gmail.cl`

11 de diciembre de 2019

Existen diferentes formas de leer datos en R pero las más comunes son:

- `scan()`: Flexible, lee vectores y sirve para gran cantidad de datos y para datos desordenados.
- `read.table()`: Fácil de utilizar y lee `data.frames`.
- `read.csv()`: Lee hojas de cálculos de excel y datos separados por “,”.

Estas últimas dos son las comunmente utilizadas.

- `read.table()`
  - `header`: Primera fila con nombres para las columnas de los datos (`FALSE`).
  - `sep`: Como las entradas son separadas ("" , espacio).
  - `na.strings`: Que valores son interpretados como NA (NA).
  - `nrows`: Máximo de filas a ser leídas.
  - `col.names`: Nombre de las columnas (V)

- ❶ Cargar archivos “.dat o .txt”.

- Especificar el “path” del directorio donde tenemos los datos.

```
setwd("C:/Users/joaquin.cavieres/Desktop/data")
```

- Comando para leer desde un archivo fuera del entorno de R.

```
read.table('nombre.dat', header=TRUE)
```

- Leer archivos .txt o .dat desde la web.

Algunas observaciones:

- Si no existen los nombres de las variables en el archivo:

```
> read.table("nombre.dat")
```

	V1	V2	V3
1	31	12	M
...	...	...	...

- Ahora, si defino los nombres:

```
> read.table("nombre.dat", col.names=c('id',  
    'age', 'sex'))
```

	id	age	sex
1	31	12	M