



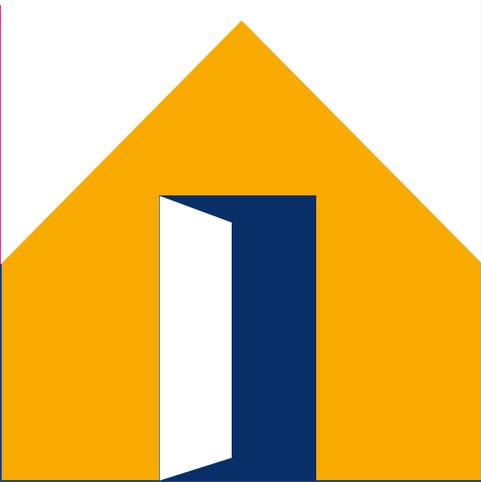
Importación de datos

Contenidos

- 01.** Importación de datos
- 02.** readxl
- 03.** readr I
- 04.** readr II
- 05.** haven
- 06.** feather y read_rds()



01.



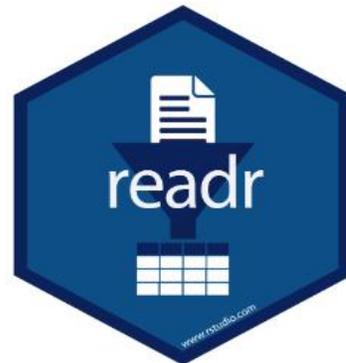
Importación de datos



Queremos traer a R datos que muchas veces provienen de fuentes externas

- De softwares ofimáticos como Excel (.xlsx, xls)
- Formatos genéricos (.csv, .tsv, etc)
- De softwares estadísticos: Stata, SPSS, SAS.
- En formatos eficientes: .rds, .feather

Veremos alternativas para importar cada uno de estos tipos de archivos.





02.

readxl





Tablas Excel

Escribimos código y ejecutamos:

```
library(readxl)
```

```
tabla_excel <- read_excel(path = "data/paises.xlsx")
```



País	Capital
Chile	Santiago
Perú	Lima
Bolivia	La Paz



03.

readr I





Tablas csv

Escribimos código y ejecutamos:

```
library(readr)
```

```
tabla_csv <- read_csv(file = "data/hospital.csv")
```

Number	City	Gender	Age	Income	Illness
1	Dallas	Male	41	40367	No
2	Dallas	Male	54	45084	No
3	Dallas	Male	42	52483	No
4	Dallas	Male	40	40941	No
5	Dallas	Male	46	50289	No



04.

readr II





Tablas delimitadas por otro caracter

Escribimos código y ejecutamos:

```
tabla_delim <- read_delim(file = "data/tabla_delim.txt")
```

nombre	%	apellido
carlos	arturo	% lehmann
jonathan	alejandro	% pizarro
camila	amaranta	% agloni
felipe	gabriel	% concha

¿Qué pasó?

```
> tabla_delim <- read_delim(file = "data/tabla_delim.txt")
Rows: 4 Columns: 3
-- column specification -----
Delimiter: " "
chr (3): nombre, %, apellido

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

 tabla_delim.txt: Bloc de notas

Archivo Edición Formato Ver Ayuda

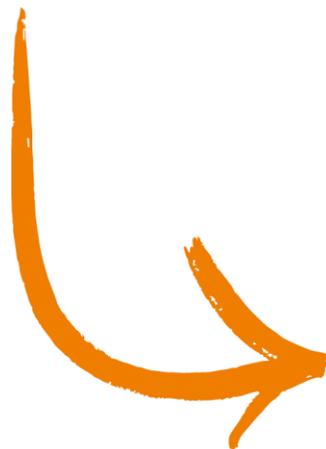
```
nombre % apellido
carlos arturo % lehmann
jonathan alejandro % pizarro
camila amaranta % agloni
felipe gabriel % concha
```



Tablas delimitadas por otro caracter

Escribimos código y ejecutamos:

```
tabla_delim <- read_delim(file = "data/tabla_delim.txt", delim = '%')
```



nombre	apellido
carlos arturo	lehmann
jonathan alejandro	pizarro
camila amaranta	aglioni
felipe gabriel	concha



05.

haven





Este paquete permite principalmente la carga de tablas de datos provenientes de otros softwares estadísticos. En particular:

- **Stata (archivos .dta):**

```
library(haven)
```

```
esi_dta = read_dta('https://www.ine.gob.cl/docs/default-source/encuesta-suplementaria-de-ingresos/bbdd/stata_esi/2021/esi-2021---personas.dta?sfvrsn=b04b324c_4&download=true')
```



- **SPSS (archivos .sav):**

```
esi_sav = read_sav('https://www.ine.gob.cl/docs/default-source/encuesta-suplementaria-de-  
ingresos/bbdd/spss_esi/2021/esi-2021---personas.sav?sfvrsn=1d2ec611_4&download=true')
```

- **SAS (archivos .sas7bdat):**

```
tabla_sas = read_sas('data/airline.sas7bdat')
```

¡Noten que también es posible cargar tablas de datos desde una URL en vez de un archivo en nuestro directorio de trabajo!



06.

`feather y
read_rds()`





El paquete **feather** y el comando `read_rds()` de R base permiten la carga eficiente de tablas de datos que se encuentren en los formatos que soportan (`.feather` el primero y `.rds` el último).

- **Feather**

```
library(feather)
```

```
tabla_feather = read_feather('data/tabla_feather.feather')
```

- **Objetos de R (.rds)**

```
copia_tabla_sas = read_rds('data/tabla_sas.rds')
```

¡Muchas gracias!





www.ine.gob.cl