



# Importar, validar y explorar datos

**Clase N° 3**

23 agosto 2021

**Análisis de datos estadísticos en R**

**Profesora** Valentina Andrade de la Horra  
**Ayudantes** Dafne Jaime y Nicolás Godoy

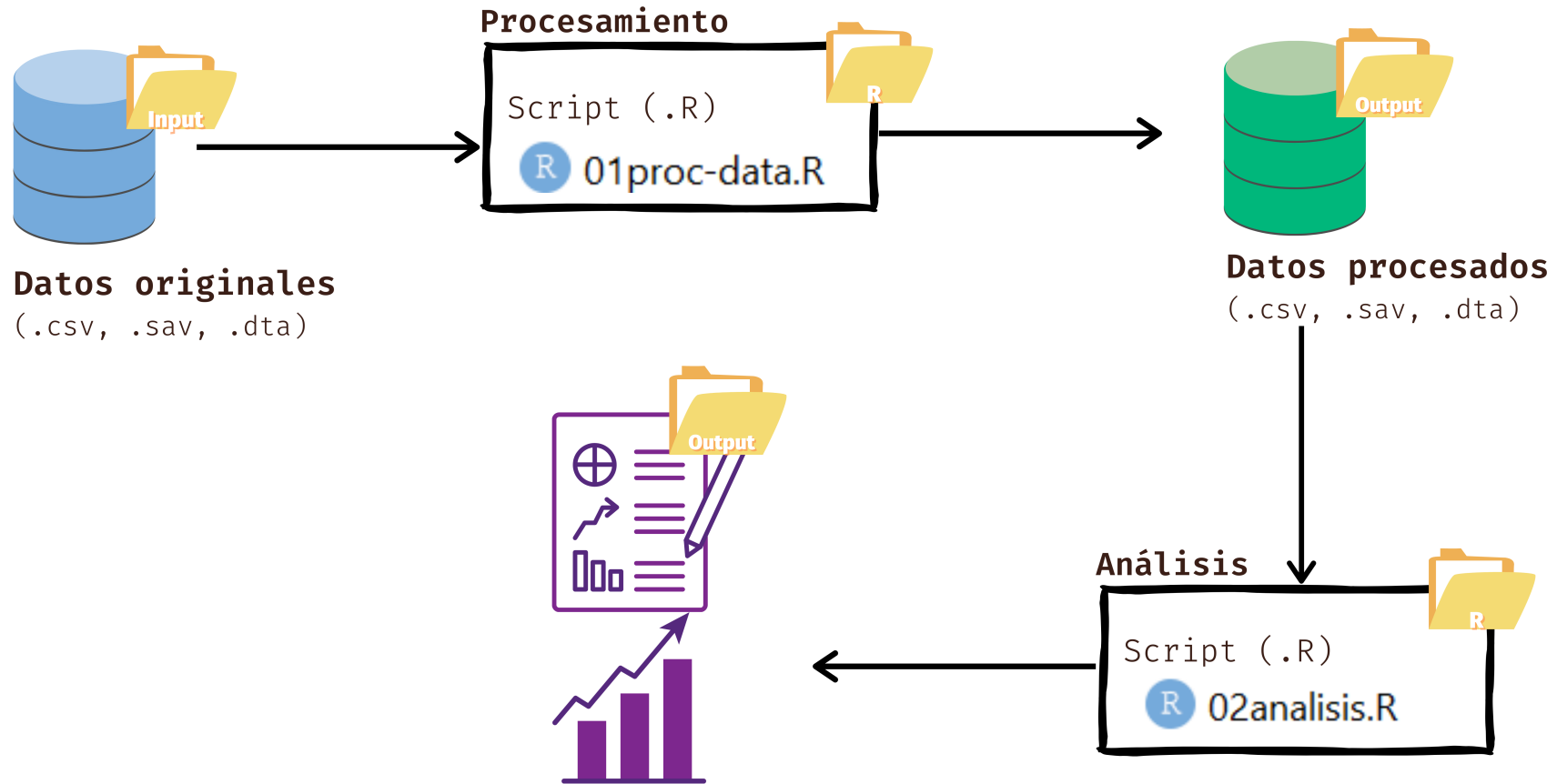
# Contenidos Sesión N°3



**Importar y exportar datos**

**Validación y limpieza datos**

**Selección de variables**

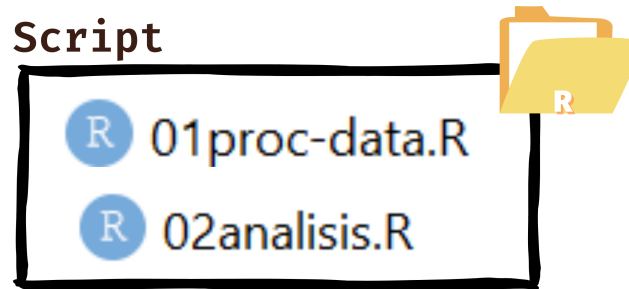




# Orden de un script



Script



```
# 1. Cargar paquetes -----  
# 2. Cargar datos -----  
# 3. Procesar datos -----  
# 4. Guardar datos -----
```



# Paso 1: Cargar paquetes



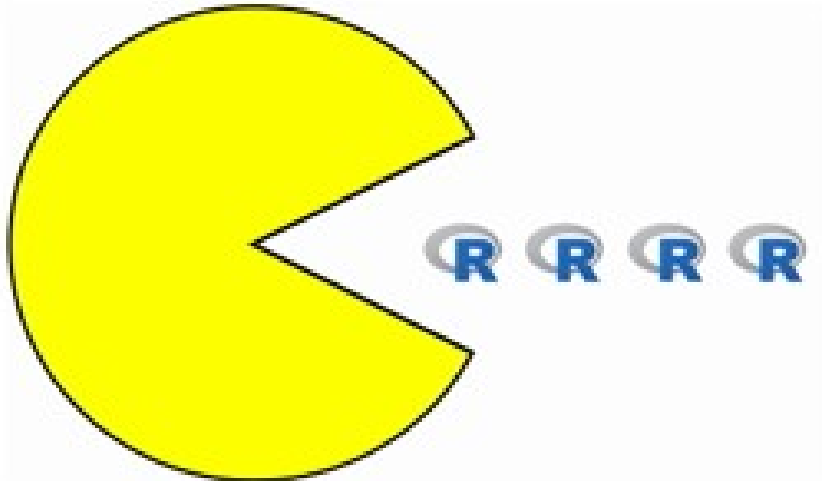
# Cargar paquetes

- Este proceso se compone de dos partes esenciales
  - (1) Instalar paquetes (`install.packages()`)
  - (2) Llamar paquetes (`library()`). Veamos su paso a paso

# Una forma fácil de cargar paquetes:



- `pacman` es un paquete que *literalmente* se comió procesos de R base y las simplificó en funciones únicas y más intuitivas.





# p\_load()



- Resume las funciones `library()` e `install.packages()`
  - optimiza esta relación entre ambas pues solo las aplica cuando son necesarias (`if requiere()`), es decir, **¡no te reinicia R si ya está instalada la librería!**

# Paso 1: Cargar paquetes



```
pacman::p_load(tidyverse,  
               magrittr,  
               car,  
               sjmisc)
```



# Paso 2: Importar datos

# Importar datos



Una vez que hemos decidido qué datos queremos trabajar

Un paso evidente es cargar los datos


**¡Y en R no es la excepción! De hecho, hablamos que R facilita el almacenamiento de datos**

# Consideraciones antes de importar



Para **importar** los datos en R debemos tener en consideración tres cosas:

1. Cómo se llaman los datos (en nuestro caso Casen en Pandemia 2020 STATA)
2. El formato de nuestros datos (en nuestro caso .sav)
3. El lugar de donde están alojados nuestros datos.

**¡Vamos a este segundo**   
**paso en el script: Cargar**  
**datos!**

# ¿Dónde?



**Descargar el zip de la clase N°3 el sitio del curso**

# 1. Recursos de la clase N°3 y práctica



- Datos: *Encuesta de Caracterización Socioeconómica (CASEN)* (2020).
- Para ello, deben dirigirse al **siguiente enlace** y descargar los zip.
- **Libro de códigos** antes de trabajar una base de datos.



# Paquetes



1. `pacman`: este facilita y agiliza la lectura de los paquetes a utilizar en R
2. `sjmisc`: explorar datos
3. `tidyverse`: colección de paquetes, del cuál utilizaremos `dplyr` y `haven`
4. `haven`: cargar y exportar bases de datos en formatos `.sav` y `.dta`
5. `readxl` y `writexl`: para cargar y exportar bases de datos en formato `.xlsx` y `.xls`
6. `dplyr`: nos permite seleccionar variables de un set de datos

**¡No fue tan difícil!**



**Repasemos algunos  
elementos**

# Consideraciones antes de importar datos



Para **importar** los datos en R debemos tener en consideración tres cosas:

1. 2. 3.

# Tipos de variables



<b>Clase</b>	<b>Tipo de variable</b>
numeric	Cuantitativa
character	Categórica
Logic	Lógica (TRUE, FALSE, NA)
factor	Categórica con niveles y etiquetas

A repasar con dos materiales complementarios: **Tipos de variables** y **Tipos de variables**

# En síntesis



**Importar y exportar datos**

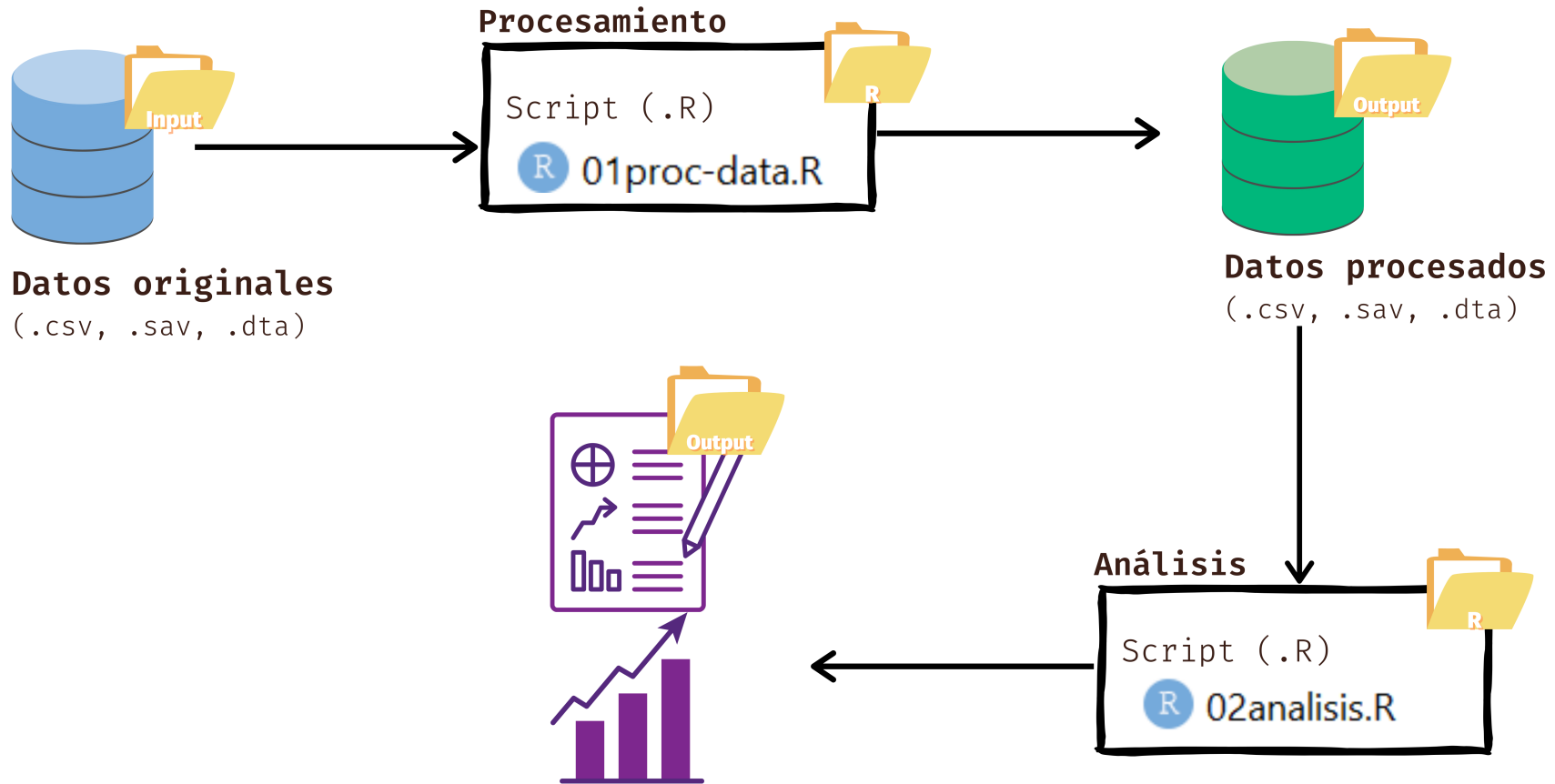
**Validación y limpieza datos**

**Selección de variables**

**¡Y a no olvidar el flujo  
para el análisis!**



**Nos permite hacernos amigas/os más rápido del programa**



¿Y eso era?

¡Antes! **Dos bonus track**





# Dos bonus track



- ¿Cómo hacer mi propio repositorio en GitHub?
- ¡Vamos a verlo!

# Dos bonus track: cargar datos en url



- ¿Cómo cargar datos desde una **url**?
  - Seguir los tres pasos para importar datos
  - **Hint:** `funcion(url("ruta_web"))`
- Quienes averiguen **como** cargar datos con url tendrán +0.5 décimas para la tarea N°2.
  - Deben dejar en slack un ejemplo reproducible en el canal #tareas
  - Pueden hacerlo **hasta antes del lunes 30**

¡Ahora si que si! Nos vemos el próximo **lunes** en la clase N°



**No olvides hacer el práctico antes**





# Importar, validar y explorar datos

**Clase N°3**

23 agosto 2021

**Análisis de datos estadísticos en R**

**Profesora** Valentina Andrade de la Horra

**Ayudantes** Dafne Jaime y Nicolás Godoy