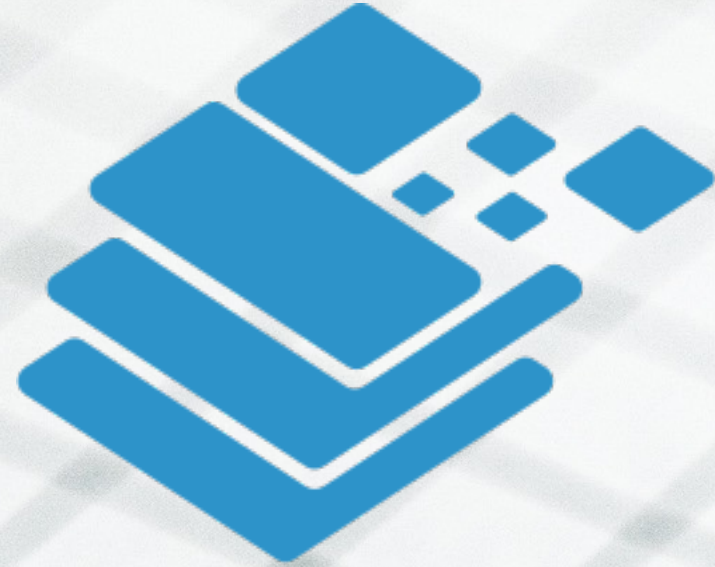




**Spatial
Lab
Analytics**

partimos pronto



**Spatial
Lab
Analytics**

ANÁLISIS DE DATOS CON R

Manipulación de datos con R

Bastían Olea Herrera - baolea@uc.cl

Exploración, manipulación, transformación y limpieza de datos con el lenguaje de programación R, apoyados por la colección de paquetes Tidyverse

01 Cargar datos

02 Exploración

03 Manipulación

04 Combinación

Tidyverse

- Ecosistema de paquetes diseñados para la ciencia de datos en R
- Focalizados en producir código legible y sencillo de comprender



Cargar datos

- **{readr}** es un paquete que lee y escribe datos de múltiples formatos, usualmente de forma más veloz y flexible
- **{readxl}** es uno de los paquetes de lectura de planillas Excel. Para escribir un archivo excel, usar **{writexl}**
- **{haven}** permite cargar archivos *spss*, *stata*, y *sas* de forma consistente
- **tibble** y **tribble**



Explorar datos

- **head, tail, glimpse:** vistazos a los datos
- **select:** seleccionar columnas o variables de una tabla
 - **starts_with, contains, where, any_of**
- **slice:** extraer filas de una tabla por su posición o distintos criterios
 - **slice_max, slice_sample**
- **pull:** convertir una columna en vector
- **filter:** filtrar observaciones en base a condiciones personalizadas
- **distinct:** filtrar observaciones repetidas a través de una o más columnas

Manejar datos

- **arrange:** ordenar las observaciones en base al valor de una o más columnas
- **relocate:** cambiar la posición de una columna
- **rename:** cambiar el nombre de una columna
- **mutate:** crear o modificar una columna
- **case_match:** retorna valores si una variable coincide con un valor
- **if_else:** retorna valores de acuerdo a si una condición es verdadera o falsa
- **case_when:** retorna valores de acuerdo a múltiples condiciones
- **group_by:** operaciones agrupadas
 - **arrange**
 - **mutate**
 - **summarize**



**Spatial
Lab
Analytics**

Soluciones en análisis de datos

www.spatiallab.cl