

R: manipulación de datos

Un proyecto de datos tiene una gran cantidad de componentes. Sin embargo, en básicamente todos se necesita iterar sobre el ciclo que se muestra en la figura ??.

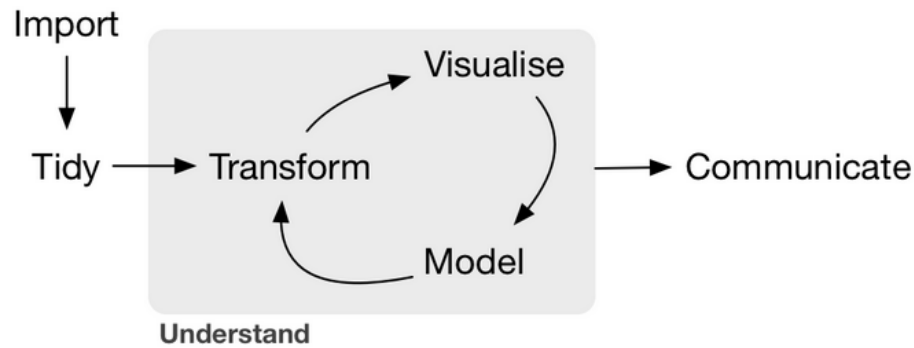


Figura 1: Modelo de las herramientas que se necesitan en un proyecto de datos según Golemund y Wickham (2016, Introducción).

Importación de datos

Para aplicar las herramientas de R a nuestro trabajo, es necesario poder importar nuestros datos a R. R tiene conectores ya implementados para casi cualquier tipo y formato de datos. Entre los más comunes están:

- Archivos de texto plano (como csvs) con `readr`.
- Datos desde una API con `httr`.
- Binarios (como excel o sas) con `haven` y `readxl`.

1

Leer y escribir archivos de datos

Lectura

Desde archivo

Datos de muestra

¹Los materiales para `readr`, `httr`, `haven` y `readxl` están basados en el capítulo de importación de datos de Golemund y Wickham (2016).

Extracción de partes de los objetos

Escritura

Do's y don'ts de R

Vectoriza y pre-asigna estructuras de datos

Which

La familia apply

Dplyr

La manipulación de datos es parte del ciclo de análisis de datos.

1.