

Introducción a R

Act. Eric Daniel Hernández Jardón

UNAM

Motivación

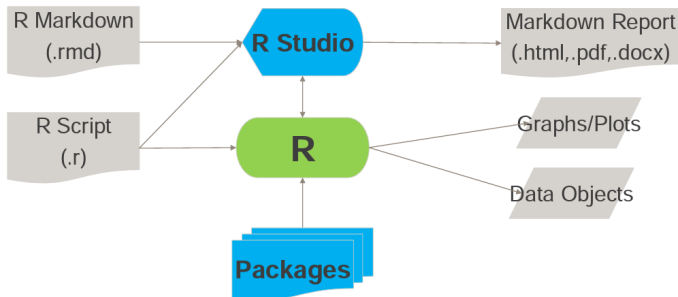
Flujo de trabajo en un proyecto

Motivación

Flujo de trabajo en un proyecto

Motivación

Flujo de trabajo en un proyecto



Motivación

Flujo de trabajo en un proyecto

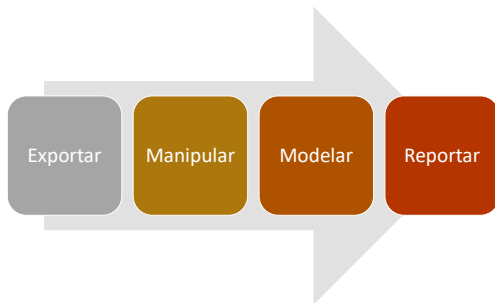
Pasos a seguir

Introducción a R

**Act. Eric Daniel
Hernández Jardón**

Motivación

Flujo de trabajo en un proyecto



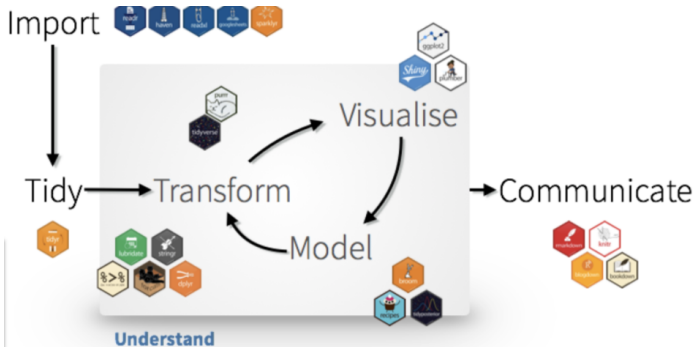
Como hacerlo en R

Introducción a R

Act. Eric Daniel
Hernández Jardón

Motivación

Flujo de trabajo en
un proyecto



Exportar

- ▶ Librerías.

Exportar

- ▶ Librerías.
 - ▶ Readr

Exportar

- ▶ Librerías.
 - ▶ Readr
 - ▶ Readxl

Exportar

Introducción a R

Act. Eric Daniel
Hernández Jardón

Motivación

Flujo de trabajo en
un proyecto

- ▶ Librerías.
 - ▶ Readr
 - ▶ Readxl
 - ▶ foregin.

- ▶ Librerías.
 - ▶ Readr
 - ▶ Readxl
 - ▶ foregin.
 - ▶ odbc.

- ▶ Librerías.
 - ▶ Readr
 - ▶ Readxl
 - ▶ foregin.
 - ▶ odbc.
- ▶ ¿Qué cuidar?

- ▶ Librerías.
 - ▶ Readr
 - ▶ Readxl
 - ▶ foregin.
 - ▶ odbc.
- ▶ ¿Qué cuidar?
 - ▶ Codificación

- ▶ Librerías.
 - ▶ Readr
 - ▶ Readxl
 - ▶ foregin.
 - ▶ odbc.
- ▶ ¿Qué cuidar?
 - ▶ Codificación
 - ▶ Tipos de datos

- ▶ Librerías.
 - ▶ Readr
 - ▶ Readxl
 - ▶ foregin.
 - ▶ odbc.
- ▶ ¿Qué cuidar?
 - ▶ Codificación
 - ▶ Tipos de datos
 - ▶ Nombres

Manipular

► Librerías.

► ¿Cómo?

- ▶ Librerías.
- ▶ dplyr

- ▶ ¿Cómo?
- ▶ Select, filter, mutate

Manipular

- ▶ Librerías.
 - ▶ dplyr
 - ▶ stringr
-
- ▶ ¿Cómo?
 - ▶ Select, filter, mutate
 - ▶ Manipular caracteres

- ▶ Librerías.
 - ▶ dplyr
 - ▶ stringr
 - ▶ lubridate
-
- ▶ ¿Cómo?
 - ▶ Select, filter, mutate
 - ▶ Manipular caracteres
 - ▶ Manipular fechas

Pipe %>%

Permite encadenar llamadas a funciones para así realizar de forma sencilla transformaciones de datos complejas.

Pipe %>%

Permite encadenar llamadas a funciones para así realizar de forma sencilla transformaciones de datos complejas.

```
data %>%  
  do_something(.) %>%  
  do_another_thing(.) %>%  
  do_last_thing(.)
```

- ▶ **Fácil lectura:** Permite pasar un elemento como argumento de la siguiente función.
- ▶ No anidar funciones como en Excel
`do_last_thing(do_another_thing(do_something(data))))`

Pipe %>%

Permite encadenar llamadas a funciones para así realizar de forma sencilla transformaciones de datos complejas.

```
data %>%  
  do_something(.) %>%  
  do_another_thing(.) %>%  
  do_last_thing(.)
```

- ▶ **Fácil lectura:** Permite pasar un elemento como argumento de la siguiente función.
- ▶ No anidar funciones como en Excel
`do_last_thing(do_another_thing(do_something(data)))`
- ▶ **Permite no guardar datos extras.**

Componentes de un gráfico

ggplot2

Introducción a R

Act. Eric Daniel
Hernández Jardón

Motivación

Flujo de trabajo en
un proyecto

Componentes de un gráfico

Introducción a R

Act. Eric Daniel
Hernández Jardón

Motivación

Flujo de trabajo en
un proyecto

ggplot2

- ▶ **Data:** Conjunto de datos que se quiere resumir.
- ▶ `ggplot()`

ggplot2

- ▶ **Data:** Conjunto de datos que se quiere resumir.
- ▶ `ggplot()`
- ▶ **Mapeo estético:** El gráfico usa varias señales visuales para representar la información proveída por el set de datos. Las dos señales más importantes suelen ser el eje-x y el eje-y. El color es otra señal visual que asignamos a la región. Nos referimos a esto como el componente de mapeo estético. La forma en que definimos el mapeo depende de qué geometría estamos usando.

ggplot2

- ▶ **Data:** Conjunto de datos que se quiere resumir.
- ▶ `ggplot()`
- ▶ **Mapeo estético:** El gráfico usa varias señales visuales para representar la información proveída por el set de datos. Las dos señales más importantes suelen ser el eje-x y el eje-y. El color es otra señal visual que asignamos a la región. Nos referimos a esto como el componente de mapeo estético. La forma en que definimos el mapeo depende de qué geometría estamos usando.
- ▶ `aes()`

Componentes de un gráfico

ggplot2

Introducción a R

Act. Eric Daniel
Hernández Jardón

Motivación

Flujo de trabajo en
un proyecto

ggplot2

- ▶ **Geometría:** El tipo gráfico que se quiere realizar.
Posibles geometrías son gráficos de dispersión, diagrama de barras, histograma, densidades suaves (smooth densities en inglés), gráfico Q-Q y diagrama de cajas.
- ▶ `geom_point()`, `geom_bar()`, `geom_histogram()`, etc.

ggplot2

- ▶ **Geometría:** El tipo gráfico que se quiere realizar. Posibles geometrías son gráficos de dispersión, diagrama de barras, histograma, densidades suaves (smooth densities en inglés), gráfico Q-Q y diagrama de cajas.
- ▶ `geom_point()`, `geom_bar()`, `geom_histogram()`, etc.
- ▶ **Personalización del gráfico.** Títulos, leyendas, títulos de los ejes, escala de los ejes, fondo del gráfico.

ggplot2

- ▶ **Geometría:** El tipo gráfico que se quiere realizar. Posibles geometrías son gráficos de dispersión, diagrama de barras, histograma, densidades suaves (smooth densities en inglés), gráfico Q-Q y diagrama de cajas.
- ▶ `geom_point()`, `geom_bar()`, `geom_histogram()`, etc.
- ▶ **Personalización del gráfico.** Títulos, leyendas, títulos de los ejes, escala de los ejes, fondo del gráfico.
- ▶ `labs (xlab, ylab, ggtitle)`, `theme`, etc.

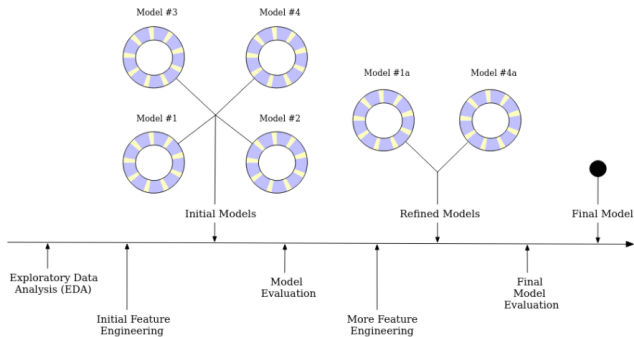
Proceso de modelación de datos

Introducción a R

Act. Eric Daniel
Hernández Jardón

Motivación

Flujo de trabajo en
un proyecto



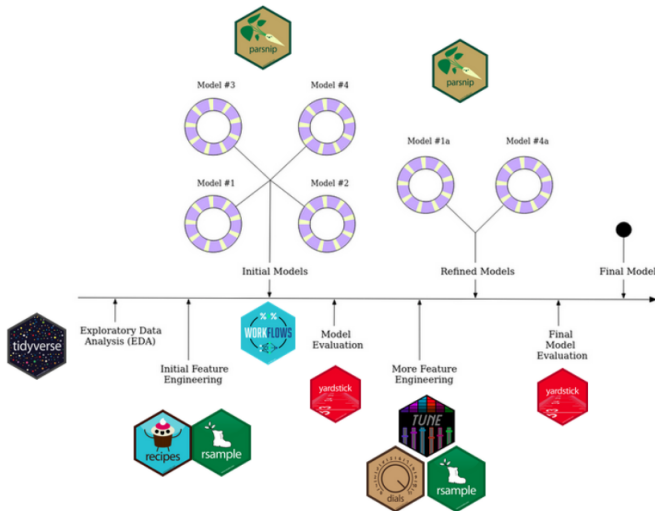
Entorno dinámico

Introducción a R

Act. Eric Daniel
Hernández Jardón

Motivación

Flujo de trabajo en
un proyecto



Actualizar R

Introducción a R

Act. Eric Daniel
Hernández Jardón

Motivación

Flujo de trabajo en
un proyecto

```
install.packages("installr")  
library(installr)  
updateR(TRUE)
```