

Conceptos Clase 4:

- Definir directorio en R
- Crear R Projects
- Importar y exportar datos desde distintas fuentes
 - o .txt
 - o .csv
 - o .xlsx / .xls
 - o .dta / .sav / .sas
- Importar datos desde pdf
 - o Librería tabulizer
- Web scraping
 - o Librería htmltab

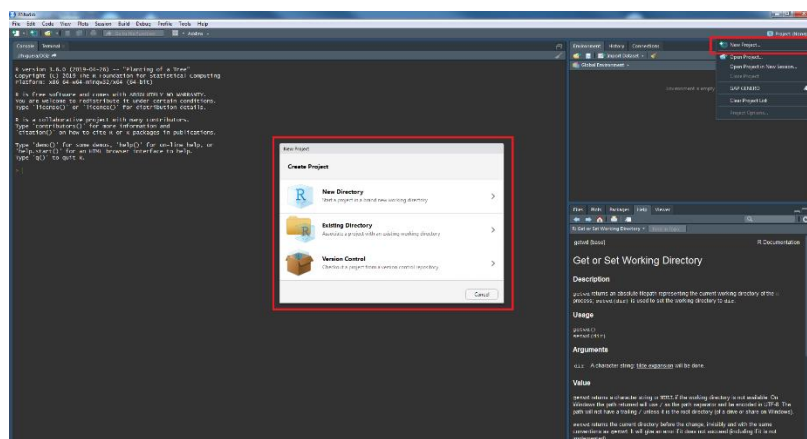
Resumen Clase 4:

Directorio de trabajo

Definir “*working directory*” (directorio): Por defecto, R no sabe dónde buscar los archivos que se quieren importar.

<code>getwd()</code>	Entrega la ruta actual en la que R está trabajando. Esta función no utiliza argumentos
<code>setwd(“ruta”)</code>	Permite cambiar la ruta en la cual trabajar. Notar que la ruta debe usar <code>/</code> y no <code>\</code> . Ej no usar “ <code>C:\Users\nombre.apellido</code> ” y si “ <code>C:/Users/nombre.apellido</code> ”

Otra forma de trabajar con rutas es a través de R Projects, los cuales también facilitan poder compartir el trabajo con terceros.



En el directorio que habremos especificado se nos creará un objeto de extensión .RProj. Cuando abrimos este objeto se nos abre una sesión de RStudio con el directorio donde

tenemos el objeto de extensión .RProj. Esta singularidad añade una gran ventaja a la hora de compartir carpetas de proyectos porque el directorio es mutable según qué usuario abre el objeto .RProj.

Importar y exportar datos

- Importar datos

Tipo de datos	Función	Librería
Separados por espacio	read_tsv	readr
Separado por coma y con punto como decimal (.csv)	read_csv	readr
Separado por ; y con coma como decimal (.csv)	read_csv2	readr
Excel (.xlsx o .xls)	read_excel	readxl
Stata (.dta)	read_stata	haven
SPSS (.sav)	read_spss	haven
SAS (.sas)	read_sas	haven
Tablas en PDF	extract_tables	tabulizer
Tablas en páginas web	htmltab	htmltab

- Exportar datos

Tipo de datos	Función	Librería
Separado por coma y con punto como decimal (.csv)	write_csv	readr
Excel (.xlsx)	write.xlsx	openxlsx
Stata (.dta)	write_dta	haven
SPSS (.sav)	write_sav	haven
SAS (.sas)	write_sas	haven