

Introduction to R Workshop

Joscelin Rocha Hidalgo - she/her - @JoscelinRocha

Grace Goodwin - she/her - @gracejgoodwin

Yolanda Yang - she/her - @yexinyu_yang

Slides adapted from David Keyes (@dgkeyes), inspired by Danielle Navarro (@djnavarro) and Paul Campbell (@paulcampbell91)

**Welcome to our
workshop!**

Agenda

Agenda

- Intro to R & RStudio (Yolanda)

Agenda

- Intro to R & RStudio (Yolanda)
- Intro to R Markdown (Grace)

Agenda

- Intro to R & RStudio (Yolanda)
- Intro to R Markdown (Grace)
- Data Wrangling (Joscelin)

Agenda

- Intro to R & RStudio (Yolanda)
- Intro to R Markdown (Grace)
- Data Wrangling (Joscelin)
- Data Visualization (Grace)

Stuck? Ask your partner



Still stuck? Raise your hand



Intro to R and Importing Datasets

Download and Install R

- Visit: <http://www.Rproject.org/>, choose the CRAN that is the closest with you physically.
- Visit: <https://www.rstudio.com/products/rstudio/download/>



[Home]

Download
CRAN

R Project

About R
Logo
Contributors
What's New?
Reporting Bugs
Conferences
Search
Get Involved: Mailing Lists
Get Involved: Contributing
Developer Pages
R Blog

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).

If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

News

- **R version 4.2.1 (Funny-Looking Kid)** has been released on 2022-06-23.
- **R version 4.2.0 (Vigorous Calisthenics)** has been released on 2022-04-22.
- **R version 4.1.3 (One Push-Up)** was released on 2022-03-10.
- Thanks to the organisers of useR! 2020 for a successful online conference. Recorded tutorials and talks from the conference are available on the [R Consortium YouTube channel](#).
- You can support the R Foundation with a renewable subscription as a [supporting member](#)

RStudio Desktop
Open Source License

Free

DOWNLOAD

[Learn more](#)

RStudio Desktop Pro
Commercial License

\$995

/year

BUY

[Learn more](#)

RStudio Server
Open Source License

Free

DOWNLOAD

[Learn more](#)

RStudio Workbench
Commercial License

\$4,975

/year

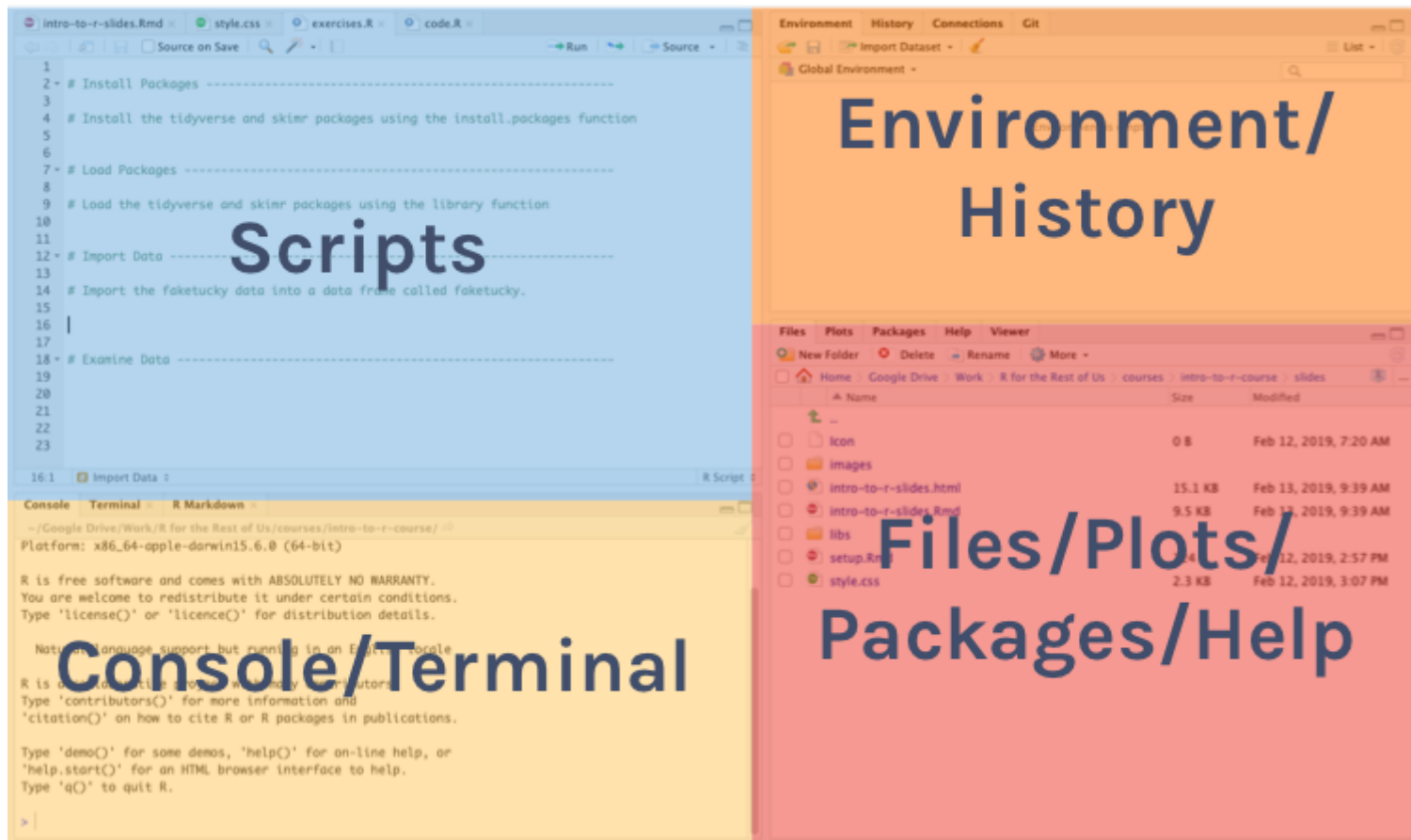
(5 Named Users)

BUY

[Evaluation](#) | [Learn more](#)

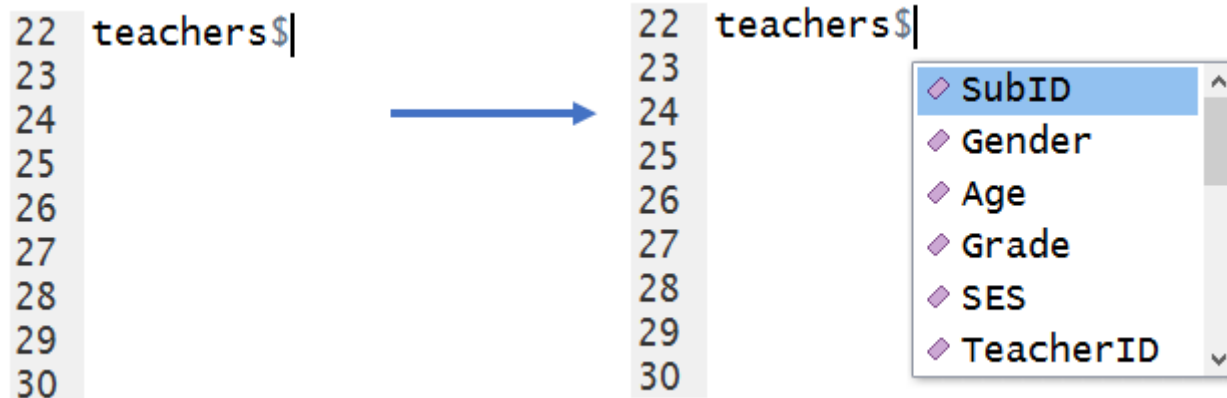
Integrated Tools for R	✓	✓	✓	✓
Priority Support		✓		✓
Access via Web Browser			✓	✓
RStudio Professional Drivers		✓		✓
Connect to RStudio Workbench remotely		✓		
Enterprise Security				✓
Project Sharing				✓
Manage Multiple R Sessions & Versions				✓

Tour of RStudio



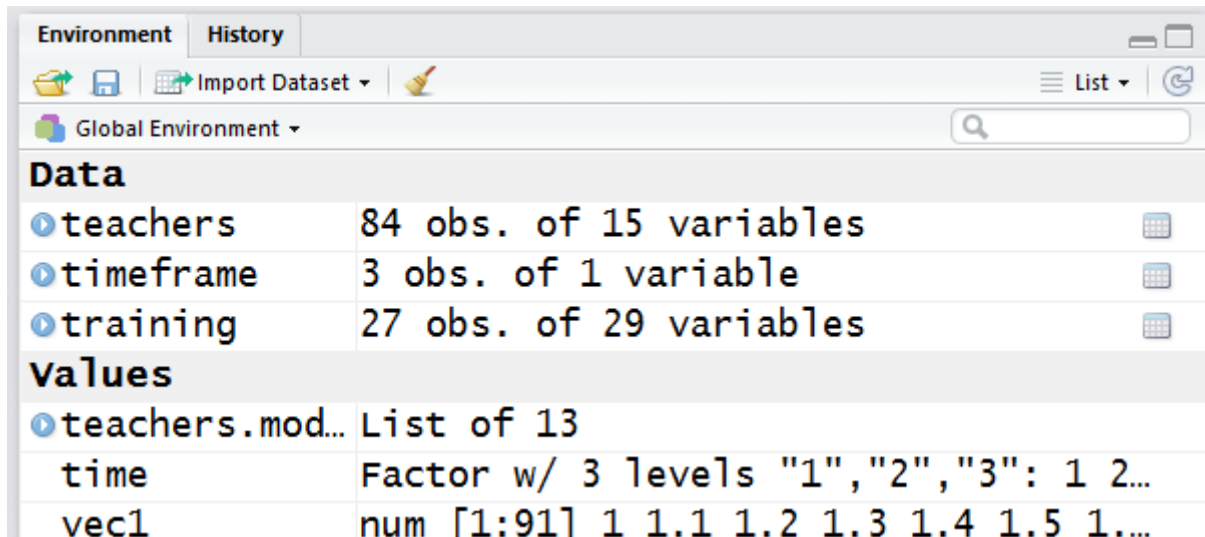
Why do we recommend using RStudio

1) RStudio makes **script writing** easier.



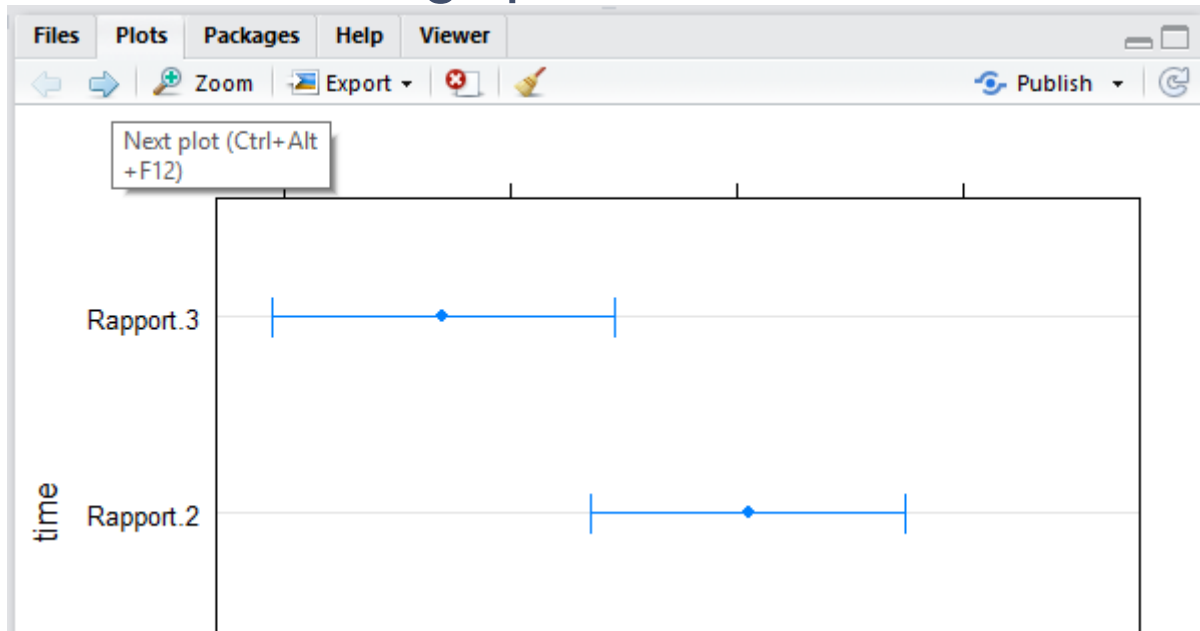
Why do we recommend using RStudio

2) RStudio makes it convenient to **view your environment** and interact with stored objects.



Why do we recommend using RStudio

3) RStudio makes **graphics** accessible.



Two File Types in R

There are **two main file types** that you'll work with:

R scripts (.R)

Text is assumed to be executable R code unless you comment it (more on this soon)

```
# This is a comment  
data <- read_csv("data.csv")
```

Two File Types in R

There are **two main file types** that you'll work with:

R scripts (.R)

Text is assumed to be executable R code unless you comment it (more on this soon)

```
# This is a comment  
data <- read_csv("data.csv")
```

RMarkdown files (.Rmd)

Text is assumed to be text unless you put it in a code chunk (more on this soon)

R Scripts

Create new script file: File -> New File -> R Script

R Start-up Kit: 4 Essential Skills

- Case Sensitivity
- Run Codes
- Write Comments
- Handling packages



Skill 1 - Case Sensitivity

Skill 1 - Case Sensitivity

R is **case sensitive** so choose one of the following for all objects and **be consistent**.

Skill 1 - Case Sensitivity

R is **case sensitive** so choose one of the following for all objects and **be consistent**.

Options

snake_case

camelCase

periods.in.names

Options

snake_case

camelCase

periods.in.names

Examples

student_data

studentData

student.data

Art by @allison_horst

Skill 2 - How to Run Code

Run the code:

Skill 2 - How to Run Code

Run the code:

If you have Windows: control + enter



Skill 2 - How to Run Code

Run the code:

If you have Windows: control + enter



If you have IOS: command + enter



Skill 3 - Make Comments

Comments help others understand your codes & help yourself keep track of your coding.

```
# Show the first 5 rows of my data  
head(data, n = 5)
```



Skill 4.1 - Install Packages

To install packages, use the "install.packages" command. See following syntax:

```
install.packages("tidyverse")  
install.packages("skimr")
```

The package name must be in quotation marks when you install them.

Skill 4.2 - Load Packages

To load packages, use the "library" command. See following syntax:

```
library(tidyverse)  
library(skimr)
```


Skill 4.2 - Load Packages

To load packages, use the "library" command. See following syntax:

```
library(tidyverse)  
library(skimr)
```

Packages should be loaded **once per session**.

Skill 4.2 - Load Packages

To load packages, use the "library" command. See following syntax:

```
library(tidyverse)  
library(skimr)
```

Packages should be loaded **once per session**.

Packages should be installed **once per computer**.

Import Datasets

CSV

```
chds6162_data <- read_csv(here("data", "chds6162_data.csv"))  
chds6162_data <- read_csv("data/chds6162_data.csv")
```

Excel

```
library(readxl)  
  
chds6162_data <- read_excel("data/chds6162_data.xls")
```

Excel

```
library(readxl)

chds6162_data <- read_excel("data/chds6162_data.xls")
```

SPSS

```
library(haven)

chds6162_data <- read_sav("data/chds6162_data.sav")
```

Set Working Directories

If the data file is in the working directory, you only need to specify its name.

```
chds6162_data <- read_csv("chds6162_data.csv")
```

Set Working Directories

If the data file is in the working directory, you only need to specify its name.

```
chds6162_data <- read_csv("chds6162_data.csv")
```

If the data file is not in the working directory, you need to specify the full path name.

```
chds6162_data <- read_csv("data/chds6162_data.csv")
```

Set Working Directories

If the data file is in the working directory, you only need to specify its name.

```
chds6162_data <- read_csv("chds6162_data.csv")
```

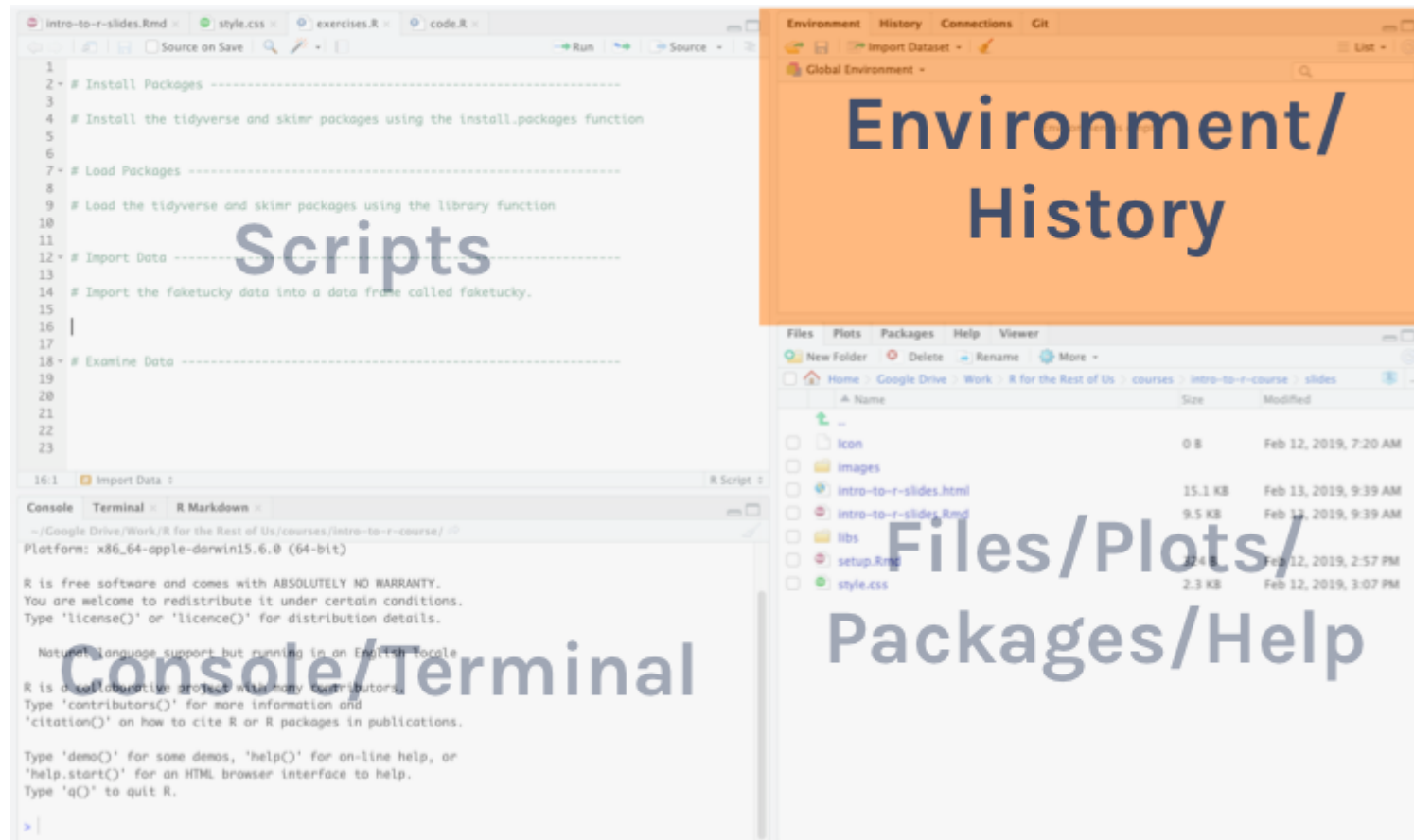
If the data file is not in the working directory, you need to specify the full path name.

```
chds6162_data <- read_csv("data/chds6162_data.csv")
```

Using an RStudio project sets your working directory to the folder where your project lives so you only need to specify the location relative to that

Where Does Our Dataset Live?

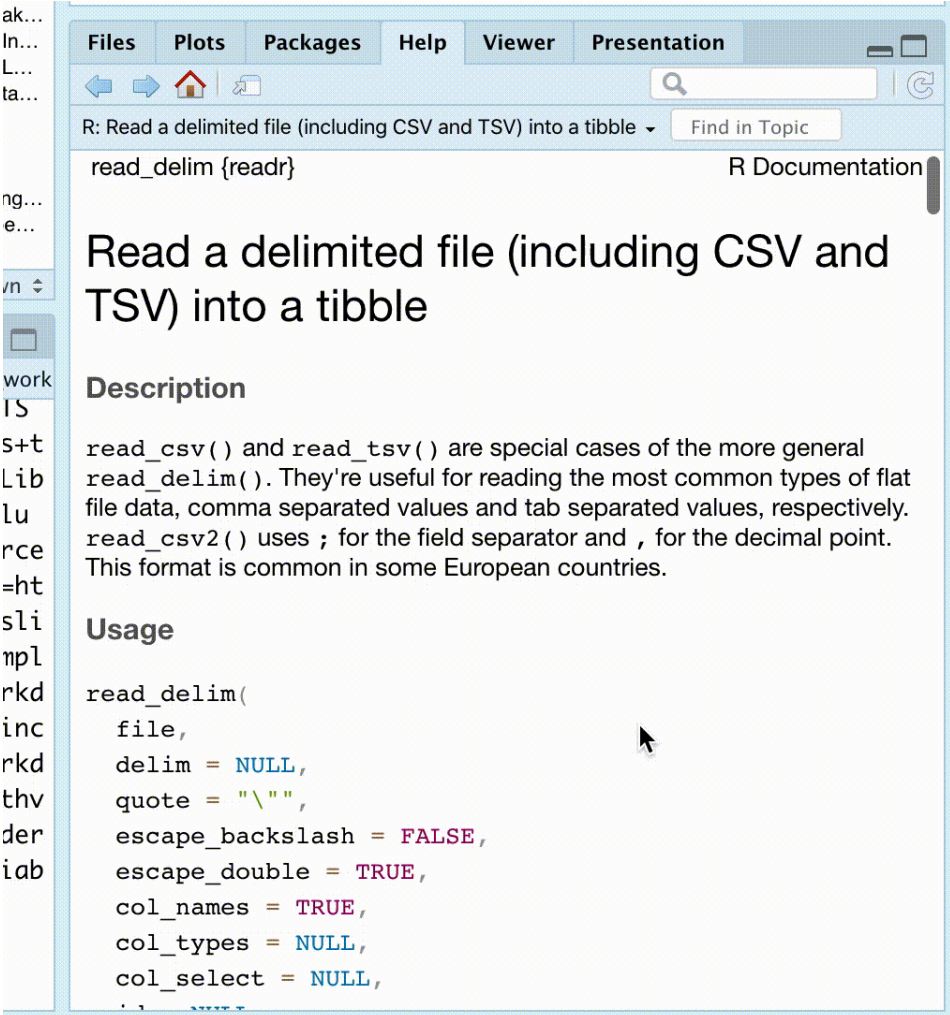
Data we have imported is available in the environment/history pane.



?function

Use the ? to get help about anything
you're confused about

```
?read_csv
```



The screenshot shows the R Documentation window for the `read_delim` function. The window has a menu bar with 'Files', 'Plots', 'Packages', 'Help', 'Viewer', and 'Presentation'. Below the menu bar is a search bar and a 'Find in Topic' button. The main content area displays the title 'Read a delimited file (including CSV and TSV) into a tibble' and a 'Description' section. The 'Usage' section shows the function signature `read_delim(file, delim = NULL, quote = "\"", escape_backslash = FALSE, escape_double = TRUE, col_names = TRUE, col_types = NULL, col_select = NULL)`. The left sidebar shows a list of files and folders, including 'ak...', 'In...', 'L...', 'ta...', 'ng...', 'e...', 'vn', 'work', 'IS', 's+t', 'Lib', 'lu', 'rce', '=ht', 'sli', 'mpl', 'rkd', 'inc', 'rkd', 'thv', 'der', and 'iab'.

R: Read a delimited file (including CSV and TSV) into a tibble

read_delim {readr}

R Documentation

Read a delimited file (including CSV and TSV) into a tibble

Description

`read_csv()` and `read_tsv()` are special cases of the more general `read_delim()`. They're useful for reading the most common types of flat file data, comma separated values and tab separated values, respectively. `read_csv2()` uses `;` for the field separator and `,` for the decimal point. This format is common in some European countries.

Usage

```
read_delim(  
  file,  
  delim = NULL,  
  quote = "\"",  
  escape_backslash = FALSE,  
  escape_double = TRUE,  
  col_names = TRUE,  
  col_types = NULL,  
  col_select = NULL,  
  ...  
)
```