



2. Introduction to the Tidyverse

Jonathan Hersh, PhD (Chapman Argyros School of Business)

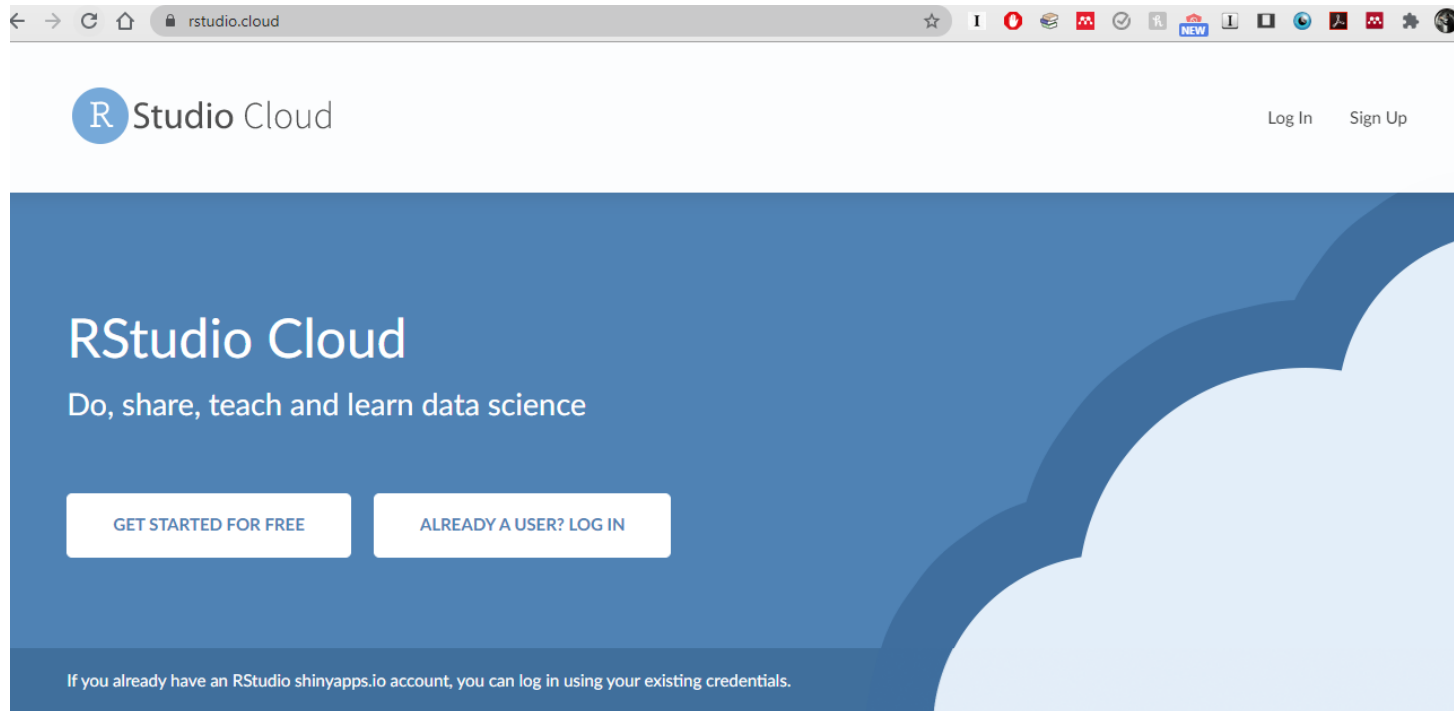
12/2/20

Section 2: Outline

1. Data Analysis

- Loading data
- Glimpse to view
- Pipe operator
- `slice()` to select rows
- `arrange()` to order data frame
- `select()` to choose variables
- `rename()` to rename variables
- `filter()` to select rows matching characteristics
- `mutate()` to create new variables
- `group_by()` and `summarize()` to create group

R Studio Cloud



- Go to rstudio.cloud if your version of R is ever not working

Data science without the hardware hassles

RStudio Cloud is a lightweight, cloud-based solution that allows anyone to do, share, teach and learn data science online.

- Analyze your data using the RStudio IDE, directly from your browser.
- Share projects with your team, class, workshop or the world.
- Teach data science with R to your students or colleagues.
- Learn data science in an instructor-led environment or with interactive tutorials.

[\\$ AVAILABLE PRICING PLANS](#)

[RSTUDIO CLOUD GUIDE](#)

[RSTUDIO.COM](#)

R Studio Cloud

The screenshot displays the RStudio Cloud web interface in a browser. The address bar shows the URL `rstudio.cloud/project/118272`. The main workspace is titled "Your Workspace / Untitled Project". The interface is divided into several panels:

- Source Editor:** Contains R code for installing packages and analyzing data. The code is as follows:

```
1 library('ggplot2')
2 install.packages('glmnet')
3 install.packages('glmnetUtils')
4 library('glmnet')
5 library('glmnetUtils')
6 library(ggplot2)
7 data(mpg)
8
9 movies <- read.csv("movie_metadata.csv")
10 head(movies)
11 ridge <- cv.glmnet(gross ~ duration + num_voted_users +
12                   budget, data = movies, alpha = 0)
13 plot(ridge)
14 coef(ridge, s = ridge$lambda.min)
15
16 lasso_mod <- cv.glmnet(gross ~ duration + num_voted_users +
17                       budget, data = movies, alpha = 1)
```
- Environment Panel:** Shows the "Global Environment" and states "Environment is empty".
- Files Panel:** Displays a file tree with the following items:

Name	Size
..	
.Rhistory	24.2 KB
movie_metadata.csv	1.4 MB
project.Rproj	205 B
- Console:** Shows the output of the R session, including package installation status and a warning about the R version change from 3.6.0 to 3.6.3.

```
* installing *binary* package 'ipred' ...
* DONE (ipred)
* installing *binary* package 'recipes' ...
* DONE (recipes)
* installing *binary* package 'caret' ...
* DONE (caret)

The downloaded source packages are in
'/tmp/RtmpPdRM1A/downloaded_packages'
R version change [3.6.0 -> 3.6.3] detected when restoring session;
search path not restored
> |
```

- R Studio Cloud is a full featured version of R in your browser!