

StatPREP: Instructor Notes

Daniel Kaplan

2018-06-06

Contents

1	Pocket guide to StatPREP commands	5
2	Signing up for cloud services	7
2.1	Google	7
2.2	GitHub	7
2.3	RStudio cloud	8
3	StatPrep Annie	9
3.1	Copying Annie’s StartUsingR project	9
3.2	Editing docs/index.Rmd	10
3.3	Writing your own tutorial	10
4	Your course web site	11
4.1	Starting point	11
4.2	Creating a GitHub repository for your project	11
4.3	Connecting your RStudio project to GitHub	14
5	Class data using Google Sheets	17
5.1	Resources	17
5.2	Setting up a new spreadsheet	18
6	Using RStudio.cloud	21
7	Publishing tutorials on shinyapps.io	23
7.1	Authorizing rstudio.cloud to publish	23

Chapter 1

Pocket guide to StatPREP commands

goal (**formula** , data = **table** , **options**)

```

counts(      ~ cat1
props(       ~ cat1
props( cat1 ~ cat2
df_stats(    ~ x
df_stats( x ~ cat1
gf_point(    y ~ x
gf_jitter(   y ~ x
gf_smooth(   y ~ x
gf_lm(       y ~ x
gf_density(  ~ y
gf_histogram( ~ x
gf_boxplot(x~ cat1
gf_violin(x ~ cat1
gf_col(      x ~ cat1
facet: x ~ y | cat1

```

x	y	cat1	cat2
3.6	6	A	red
1.2	9	A	blue
12.0	2	B	blue

Quick facts on data frames
names(D) nrow(D) head(D)

Functions for randomization
Trials←do(100)*{task}
 Confidence intervals
data=resample(df)
 Permutation test
x ~ shuffle(cat1)

optional arguments
 graphics
fill = "blue"
fill = ~ cat1
color = ~ cat1
alpha = 0.5
jitter
width = 0.2

for df_stats
 mean ci.mean
 median ci.median
 sd ci.sd
 coverage
 min Q1 Q3 max
 length

Pocket guide to StatPREP R

rev 5-25-18

Figure 1.1:

Chapter 2

Signing up for cloud services

You don't need to install any special software on your own computer. Instead, we'll use services in the cloud that work through a standard web browser.

It's helpful if you set up a personal account on the cloud services listed below. That will save time on the day of the workshop and you can be a resource for your neighbor if they haven't had a chance to do so. (Do remember to keep track of your user ID and password. Writing it down is a good idea; you can change the password after the workshop if you are worried about security.)

2.1 Google

You may already have a Google account: many people have an account already or work at an institution that provides email and other services through Google. If so, you are all set.

If you don't already have an account, follow this link to sign up.

Setting up a Google account is entirely to streamline authentication to other services that we use with StatPREP. You do not need to change anything about your existing email service or how you use it.

2.2 GitHub

Funny name, huh? GitHub is a free service with tens of millions of users. It's most closely associated with software development, but our main use for it will be to give you a way easily to create a web page to give your classes access to whichever StatPREP tutorials, lessons, or Little Apps you choose to use with them.

Your institution may already provide you with a web site or a system such as Moodle or Blackboard that gives you a class-specific web page. If so, the point of setting up a GitHub account to use at the workshop is to make it easier for us to avoid having to figure out how to upload documents to a multiplicity of different web platforms.

In selecting your user ID for GitHub, keep in mind that the ID is something that will be visible to students. So, **ProfJones** or something of that kind is probably better than **Red_hot_pepper_dude**.

Follow this link to GitHub's account creation page. And *don't be intimidated* by the "Built for developers" label.

Once you've set up your account, GitHub will email a confirmation letter. **Make sure to click the confirm button** in the email.

2.3 RStudio cloud

We want you to have access to RStudio so that you can use it *if you decide you want to*. We'll show everyone some basics at the workshop so you can make an informed decision.

You can sign in using either your Google or your GitHub credentials; there's no need to set up a separate ID or password. Go to `rstudio.cloud`.

Chapter 3

StatPrep Annie

StatPrep Annie is a persona created to depict a real-world StatPREP instructor who is setting up their statistics course.

She’s got a website for her course, a couple of interactive lessons, and so on.

One way to get a quick start on your own website, lessons, etc. is to copy Annie’s project to your own accounts. Then customize it for your own purposes.

3.1 Copying Annie’s StartUsingR project

1. Using your browser, login to your account on `rstudio.cloud`. The main page for your “workspace” will be displayed.
2. Open a new tab in the browser. Cut and paste exactly this URL into that new tab.

`https://rstudio.cloud/project/40418`

Annie’s template will be copied into your workspace. It will open with a red “Temporary” in the top line.

3. Press “Save a permanent copy” so that you have your own, fully independent copy of Annie’s StartUsingR project.



Figure 3.1: StatPrep Annie

To transform StatPrep Annie’s web site to your own, just edit the `docs/index.Rmd` file in your copy of the project. Chapter [Your_course_web_site] shows how to publish the website.

3.1.1 Note for those using desktop RStudio

Copying Annie’s `StartUsingR` project can be done with the “new project” menu in RStudio. Choose to create a new project from a GitHub repo using this address for cloning:

<https://github.com/StatPrep-Annie/StartingR.git>

3.2 Editing `docs/index.Rmd`

This will have instructions for editing `docs/index.Rmd`

3.3 Writing your own tutorial

The file `tutorials/first-tutorial/first-tutorial.Rmd` has a template for a tutorial lesson.

Chapter 4

Your course web site

As statistics instructors start using data in their classes, they find that they need to make data files available to students. An excellent way to do that is to put the files on a web site, so that the students can access them with a URL.

If your institution uses course support software such as Blackboard or Moodle, you may want to take advantage of those resources.

Many instructors don't have a web server available to them and aren't sure how to set up a web site. (And, warranted or not, many instructors grumble about Blackboard and Moodle) The point of this repository is to help you set up your own course web site on which you can place data files, etc. so that your students can easily get to them.

4.1 Starting point ...

We'll assume that you have already created an RStudio project, perhaps simply by copying StatPrep Annie's "StartUsingR" project to your own rstudio.cloud account, as described in Chapter [StatPrep_Annie]

4.2 Creating a GitHub repository for your project

Leaving your RStudio.cloud tab for a few moments, you're going to create a new repository on GitHub to use for publishing web pages from your project.

1. Login to GitHub. Once you have done this, access the +v dropdown menu in the upper right of the GitHub display:

Select "new repository"

2. In response to your selecting "new repository," GitHub will display a set-up page:

- Choose a suitable name for the repo. For instance, if this is to be a course site, you might use the name of the course, e.g. **Stat101**.
- Once you have set the new repository's name, skip directly on down to the green "Create repository" button. Press it

3. GitHub will now display a "Quick setup page." Near the top will be a section that looks like this:

Note the repo URL that appears in the editing box. It is composed from your GitHub user name and the name you selected for the repository. Keep that handy. Later, you're going to paste that URL into a command.

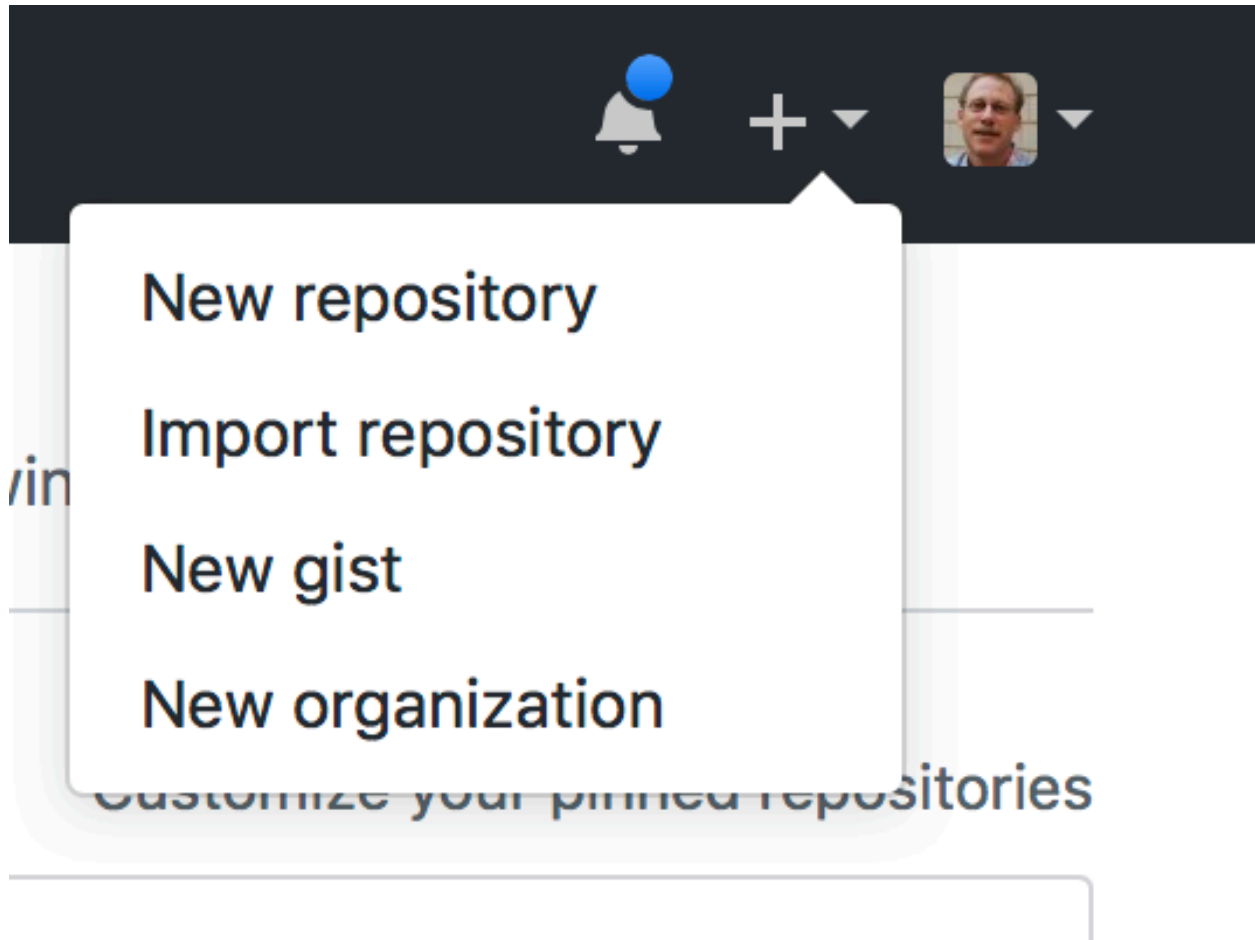
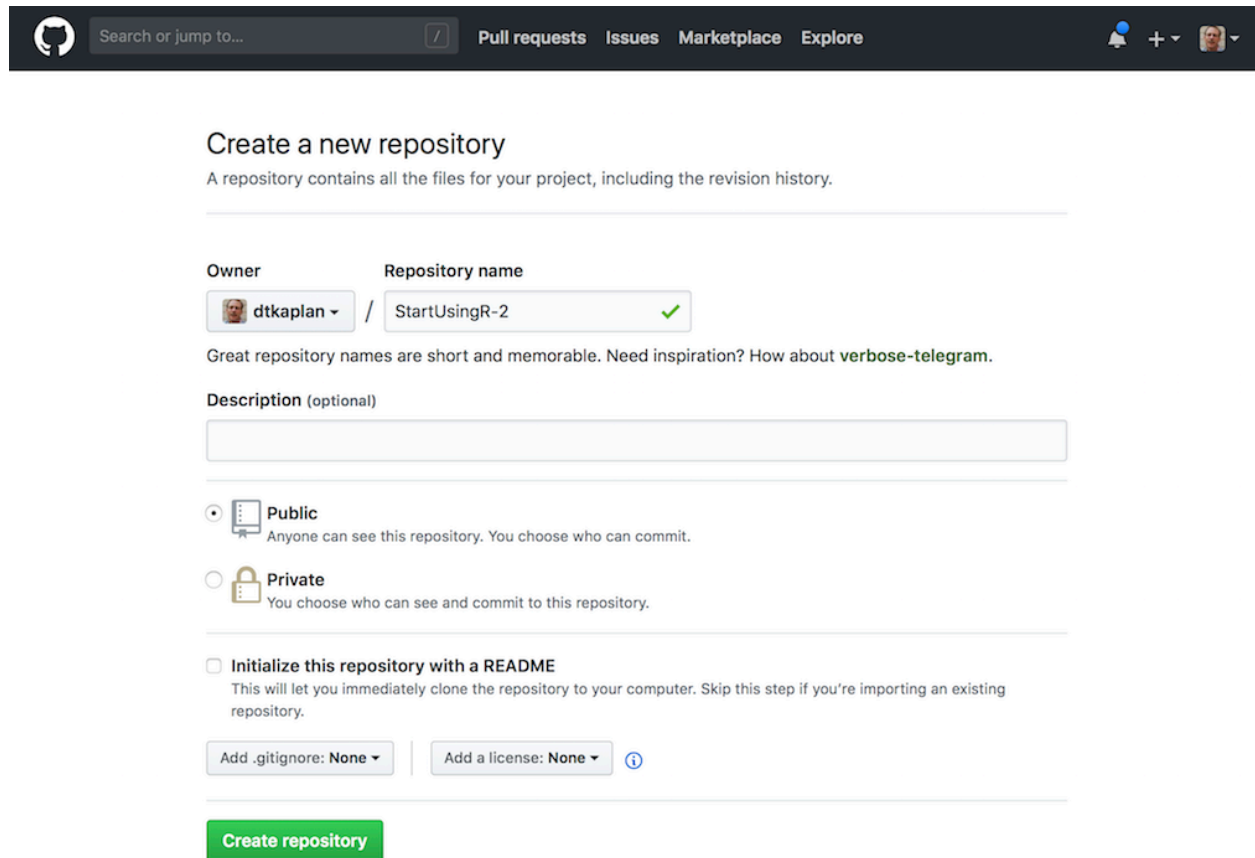
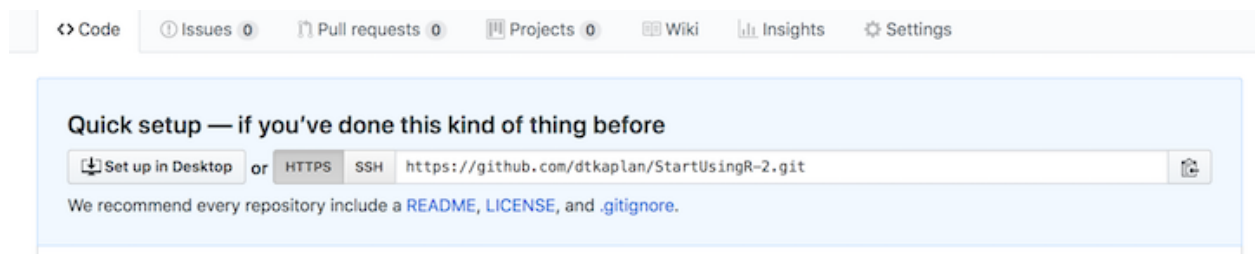


Figure 4.1:



The screenshot shows the GitHub interface for creating a new repository. At the top is a dark navigation bar with the GitHub logo, a search bar, and links for Pull requests, Issues, Marketplace, and Explore. Below this, the main heading is 'Create a new repository' with a subtext: 'A repository contains all the files for your project, including the revision history.' The form includes fields for 'Owner' (set to 'dtkaplan') and 'Repository name' (set to 'StartUsingR-2' with a green checkmark). A note suggests repository names should be short and memorable. There is an optional 'Description' text area. Under 'Visibility', the 'Public' option is selected, with a description: 'Anyone can see this repository. You choose who can commit.' The 'Private' option is also available. An unchecked checkbox for 'Initialize this repository with a README' is present, with a note about cloning to a computer. At the bottom of the form are dropdowns for 'Add .gitignore: None' and 'Add a license: None', followed by a green 'Create repository' button.

Figure 4.2:

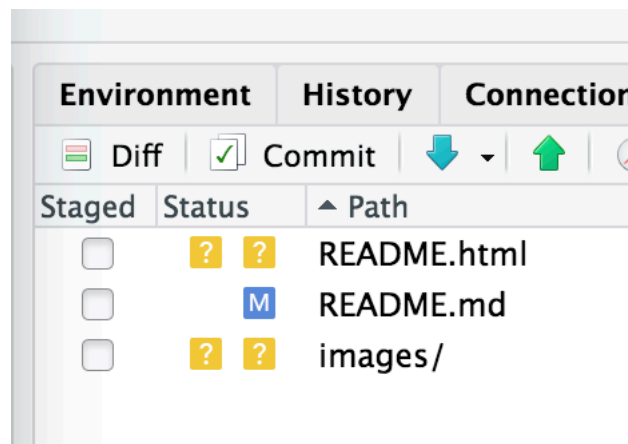


The screenshot shows the top of a GitHub repository page. A navigation bar at the top contains links for Code, Issues (0), Pull requests (0), Projects (0), Wiki, Insights, and Settings. Below this is a light blue box titled 'Quick setup — if you've done this kind of thing before'. Inside the box, there are buttons for 'Set up in Desktop' and 'HTTPS', followed by 'SSH' and a text input field containing the URL 'https://github.com/dtkaplan/StartUsingR-2.git'. A small icon for cloning is on the right. Below the input field, a note states: 'We recommend every repository include a README, LICENSE, and .gitignore.'

Figure 4.3:

4.3 Connecting your RStudio project to GitHub

Your task now is to connect your own copy of the StartUsingR project to GitHub. To do this, go back to your rstudio.cloud tab displaying your project.



1. In the Git tab in RStudio, select the “gear” menu and then “shell.”

This will open up a new tab called “Terminal”, next to the console.

2. In the Terminal tab, cut and paste these commands, making sure to **provide your own email address and name** rather than StatPrep Annie’s. (If you have multiple email addresses, or multiple names, you can use any of them!)

```
git config --global user.email "StatPrep.Annie@gmail.com"
git config --global user.name "StatPrep Annie"
```

Press enter. There will be no response by the computer. You’re going to be using the terminal tab later, as well.

3. Give the command, in the terminal tab, that will instruct RStudio to refer to your GitHub repository. The command will look like the following, but **you must** change USERNAME to be your own GitHub username, and change REPOSITORY to be your own repository, set up in Step (1) of this section.

```
git remote set-url origin <paste_your_repo_URL_here>
```

Again, when you press “enter”, the computer will not respond.

4. In RStudio, open up some file, say, README.Rmd, and make some trivial change, such as adding a space after the document title. Then save the file.
5. In RStudio, go to the Git tab. You will see the name of the file you just edited (and maybe some others). Check the little box under “Staged” to the left of the file name. Then, press Commit.
6. A new window will open that looks like a bigger version of the Git tab.

Write some short message in the box and press the “commit” button underneath the message box.

7. Almost done ... Press the green upward pointing arrow above the message box. You will be prompted to enter your GitHub account ID and password. Do so.
8. Now back to your GitHub account. Go to the repository you set up and press “settings.” Under gh-pages in settings, select “master branch docs folder.” In response, a github.io URL will appear in the gh-pages section. That’s the address of the web site.

In a few minutes, you should be able to access your new web page using your own GitHub.io address.

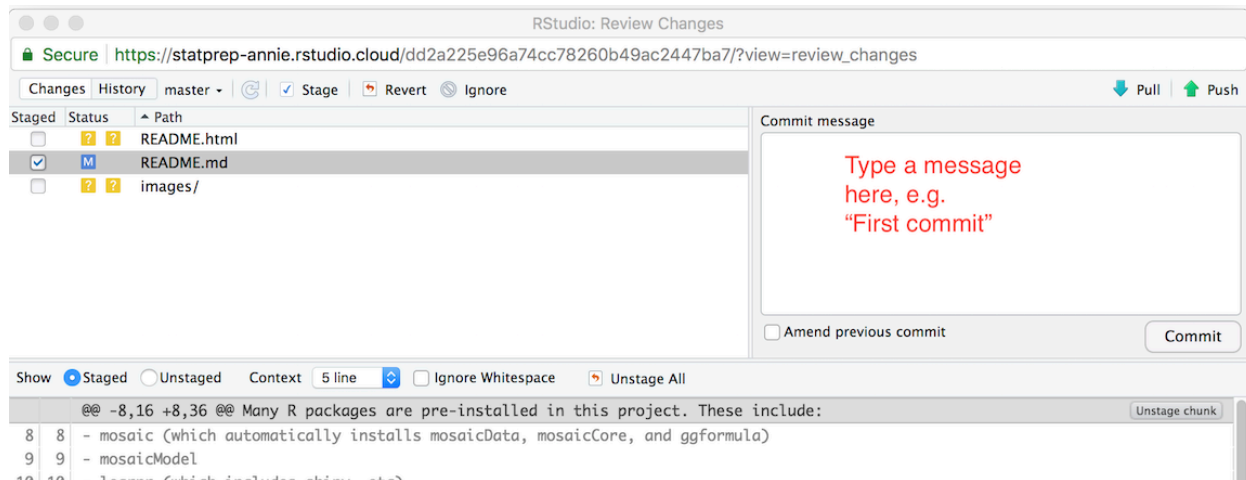


Figure 4.4:

4.3.0.1 Putting links to data files on your own course web site

If you are going to use your site to provide student access to data sets of particular interest to you, you will want to put links and instructions on your course web site.

The markup that you include in your `index.md` file (in the `docs/` directory) might look like this:

```
## Google files used in class
```

```
- `Survey1 <- gs_read(gs_key("1ucevNh7wKLtOukyEpacUKi5_-KZUQGtI00NhWRnnnQ4"))`
```

```
## Data files
```

```
Data files for this week:
```

```
- `https://dtkaplan.github.io/stat101/test.csv`
```

To create the data table in your R session, copy and paste this command into your console:

```
```r
My_data <- read.csv("https://dtkaplan.github.io/stat101/test.csv")
```
```

4.3.1 Customizing your site with RStudio

Outline:

- clone the repo
- open a new project in RStudio, choosing the option for a GitHub repository.
- Edit as needed. Every file you edit should be “knitted” to HTML.
- State, commit, push, and pull.

Chapter 5

Class data using Google Sheets

Collecting data interactively with students has several benefits:

1. Students see the connection between the data and their own actions. Students are especially motivated when the data is about them or something they are doing.
2. The inevitable imperfections in user-entered data serve as a lesson in coding factors and the measuring variables in a standard way.
3. The instructor (or individual students) can analyze the data and see how the analysis changes in real time, as more data rows are added.

5.1 Resources

Two example lessons, from the StatPREP 2018 Workshops are:

- Globe toss
- Riverboat card trick

You're welcome to use those interactive documents. Each document has a link to the spreadsheet for entering data. It's your own choice whether to start by clearing out any existing rows or add on to the rows from another session. **NOTE:** No guarantee that your class's data will be there later on, since someone else might have erased it. And so you might want to set up your own spreadsheet exclusively for your class's use.

The section below on *Setting up a new spreadsheet* shows how to create a spreadsheet for your class. Note that you should post the link to your spreadsheet on your course website, so that students can access it to enter data.

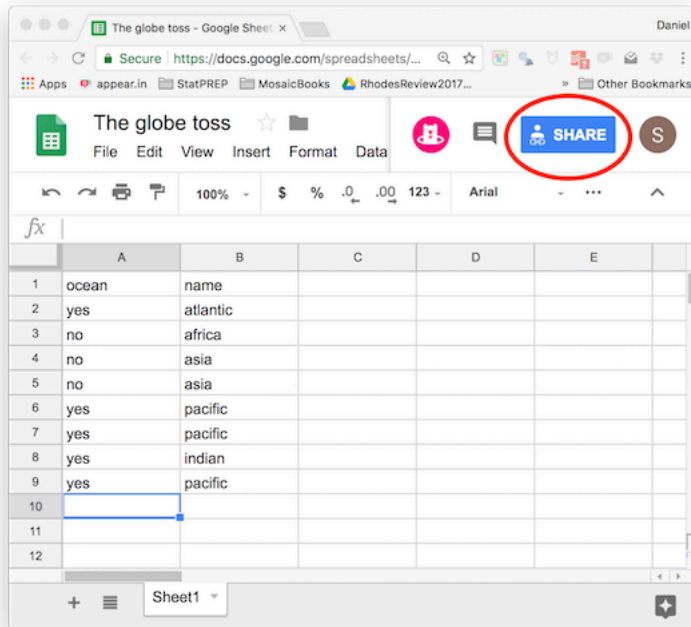
For data analysis, you have a choice of options:

1. Use the Read class data activity. The user will have to paste the command to read in your spreadsheet at the start of every command block. (See *Setting up a new spreadsheet*.)
2. Use your own R session. Again, you will need to give the command to read in your spreadsheet into the console.
3. Create your own tutorial document in Rmd. This presumes that you are comfortable editing Rmd documents and, if you want to give your students access, publishing them on a server. A working template is available via the StatPREP Workshops2018 project on rstudio.cloud. (Follow this link. to the file `Google_data_template.Rmd`.)

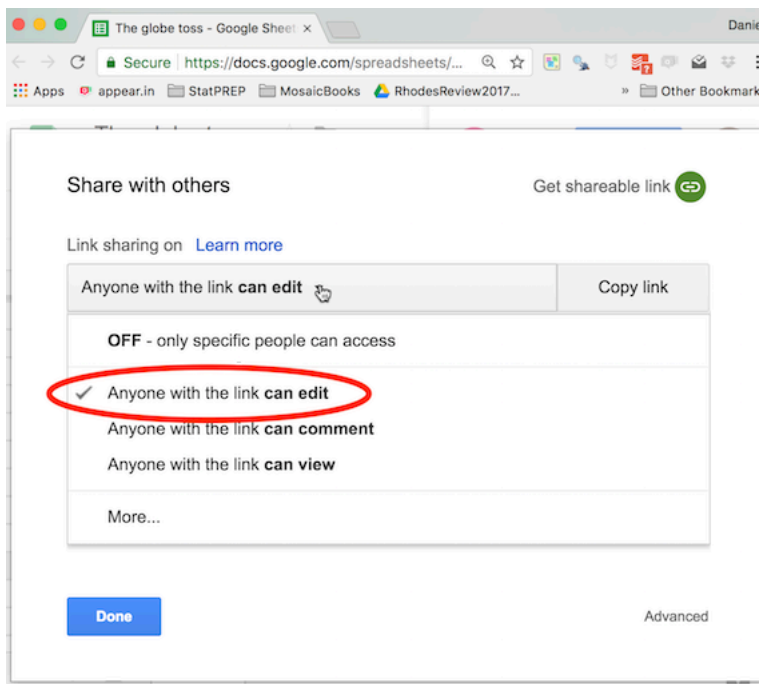
5.2 Setting up a new spreadsheet

You can modify this document to work with a spreadsheet of your own. Here's how.

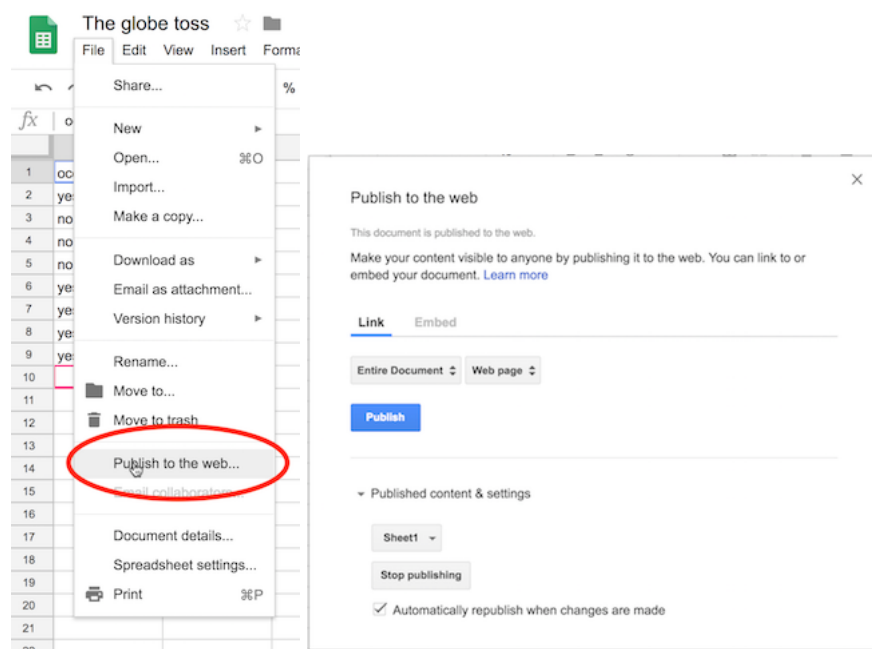
1. Set up a Google spreadsheet. It's a good idea to populate it with some variable names and a few values. This will let you test to make things are working before you start the activity in class.
2. *Within* the Google spreadsheet document press the “Share” button.



3. After you press “Share,” you will see a dialog box.



- Pull down the menu to select “Anyone with the link **can edit.**” (This is what lets your students add data.)
 - Copy the link and paste it somewhere you can get to it again. You’ll need it. The link will look like this: https://docs.google.com/spreadsheets/d/1ucevNh7wKLtOukyEpacUKi5_-KZUQGtIOONhWRnnnQ4/edit?u
 - You will also need to copy the *key* that’s contained in the link. The key is just the central gibberish in the link, like this: **1ucevNh7wKLtOukyEpacUKi5_-KZUQGtIOONhWRnnnQ4**
4. Put the link on your course web site so that your students can get to it. That’s how they will access the spreadsheet for entering data.
 5. Back in your Google sheet, select the File/Publish_to_the_web menu item. Use the resulting dialog box to publish the entire document.



6. Create the R command that will load the spreadsheet data into your R session.

```
Globe <- gs_read(gs_key("1ucevNh7wKLtOukyEpacUKi5_-KZUQGtIOONhWRnnnQ4"))
```

In forming the command, replace the quoted string in the above with your own key. The key is located in the center of the link to the spreadsheet: an incomprehensible set of characters similar to that highlighted in **bold** in the example in step (3). You might also replace **Globe** with a name that’s more suited to your own activity.

Paste the command you’ve created someplace handy. You’ll need it. (Suggestion: Paste it next to the link in step (4).)

Time for class!

7. Once you reach the point in your class where you want to do statistics on your data, bring up the lesson document provided for this purpose by StatPREP located here. That document has several R command chunks, all of which are blank. You can put any R commands in those chunks, **but make sure** that the command from step (5) always is the first command in any chunk that you use. That way, whenever you run the code in the chunk, the data will be read in from Google. Keep in mind that the chunks are all independent of one another, so you’ll need to read in the data in any chunk you use.

Try it out in the following command chunk:

- a. Paste in the command you created in step (5).

- b. Below that, add any R commands you like.

For instructors who want to write their own tutorial, you will find that this simplifies things since you can arrange to have the spreadsheet data read-in globally and not have to put the data-reading command in every chunk. Use the template `.Rmd` document provided by StatPREP [here](#). Modify the chunk named `read_data` the top of the document by inserting your own command (with your own key). Notice that in any new chunk you create, you'll have to reference the `read_data` chunk as the `exercise-setup`. The chunks already in the template document do this, so you can just copy (and rename!) an existing chunk.

Chapter 6

Using RStudio.cloud

Just some preliminary notes ...

Under preferences/Rmd, arrange to have the preview opened in the Viewer tab. It doesn't seem to work to leave it as opening in a web page.

Chapter 7

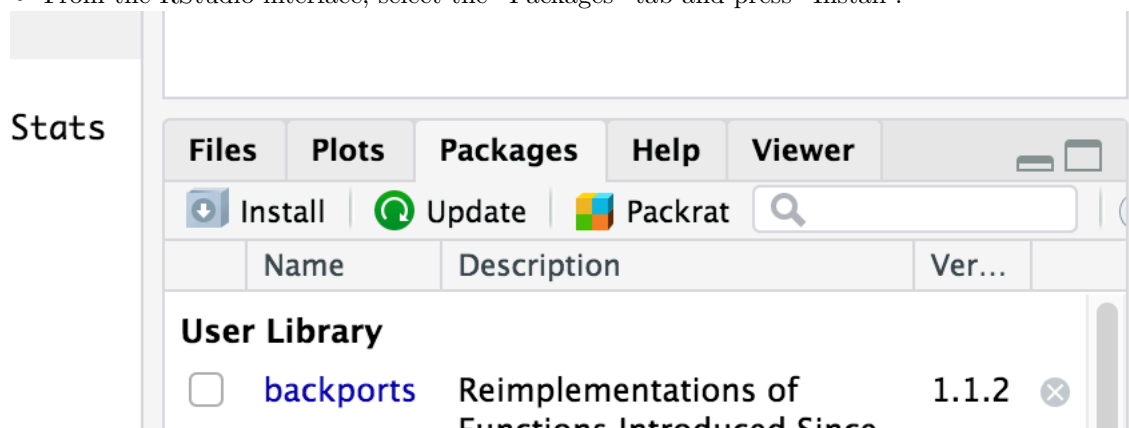
Publishing tutorials on shinyapps.io

This chapter is not yet complete.

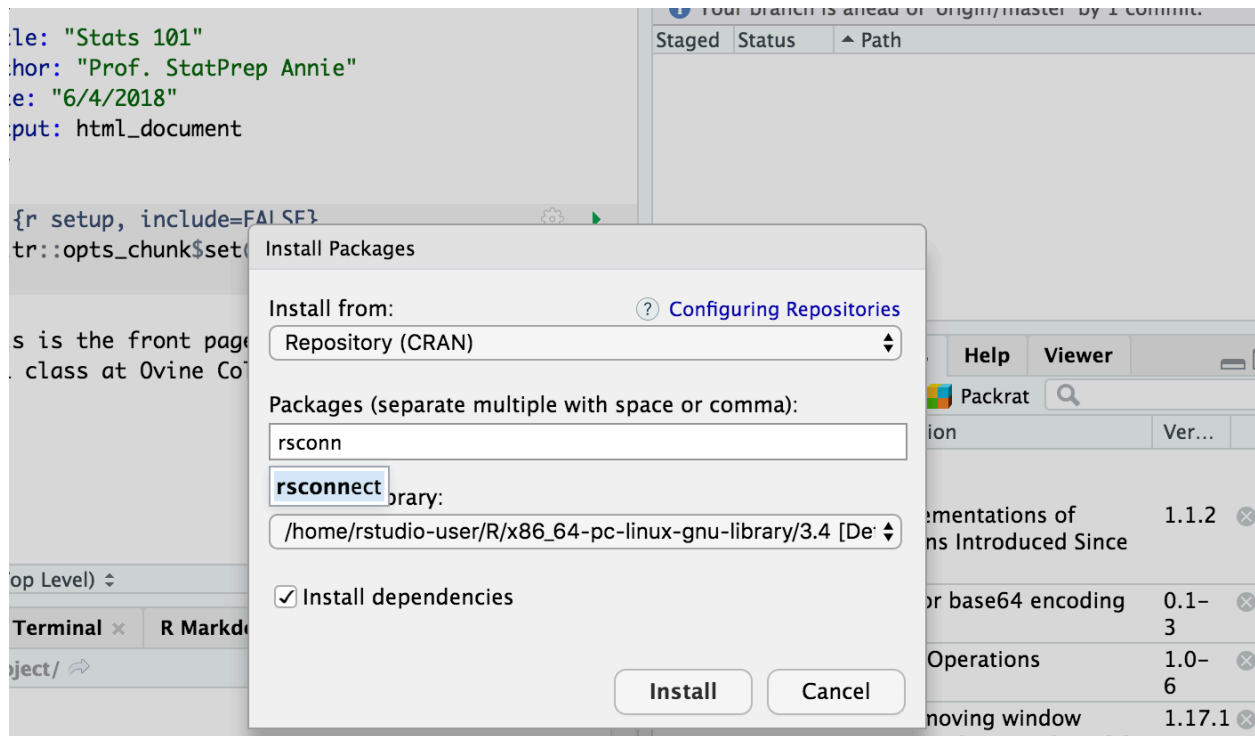
7.1 Authorizing rstudio.cloud to publish

You need to authorize `rstudio.cloud` to publish to your `shinyapps.io` account. How do you do this? The general idea is that you login to `shinyapps.io` and ask it to tell you a secret. Then, from `rstudio.cloud`, you send that secret back to `shinyapps.io`. Once `shinyapps.io` knows that `rstudio.cloud` is in on the secret, shinyapps will accept future commands from you `rstudio.cloud` account.

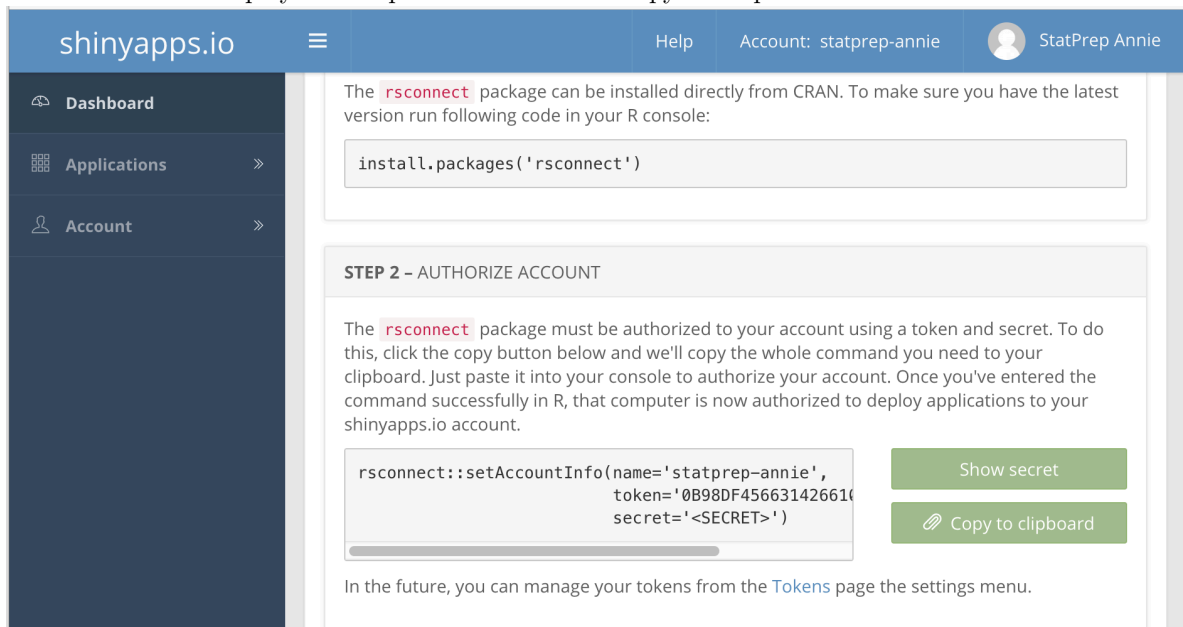
1. Login to your account on `rstudio.cloud`.
2. Open any project on `rstudio.cloud`.
 - From the RStudio interface, select the “Packages” tab and press “Install”.



- In the dialog box that appears, start typing `rsconnect`. At some point, the dropdown menu will show that choice. Click on that and press “Install”.



3. Login to shinyapps.io. Select the “Dashboard” tab. You’ll see a section entitled “Authorize account” with a display of computer code and a “copy to clipboard” button. Press that button.



Depending on your browser, you may be asked to press CNTR-C to copy the code.

4. Return to the console in `rstudio.cloud` and paste in the code you copied in (3).

Press the “Publish” button. - Select “RStudio Connect” - Select “Publish finished document only” - In the dialog box titled “Select RStudio Connect Account”, type the address shinyapps.io.