Getting Started with STAT 961

August 24, 2021

1 Setup

R and RStudio setup.

1. If you do not have R and/or RStudio installed, download and install these softwares by following the instructions here. If you do have RStudio installed, update it to the latest version by opening RStudio and clicking `Help` -> `Check for Updates`. If you do have R installed, update it to the latest version by following the instructions here. Open RStudio.
2. Install packages for compiling reports by entering the following commands at the console (labeled B in Figure 1):

```
install.packages("knitr")        # for integrating LaTeX with R
install.packages("kableExtra")   # for creating tables
install.packages("tinytex")      # to compile LaTeX documents
tinytex::install_tinytex()       # to compile LaTeX documents
tinytex:::install_yihui_pkgs()   # to compile LaTeX documents
```

3. Go to `Tools` -> `Global Options` -> `Sweave` on Windows or `RStudio` -> `Preferences` -> `Sweave` on Mac and select the options pictured in Figure 2.

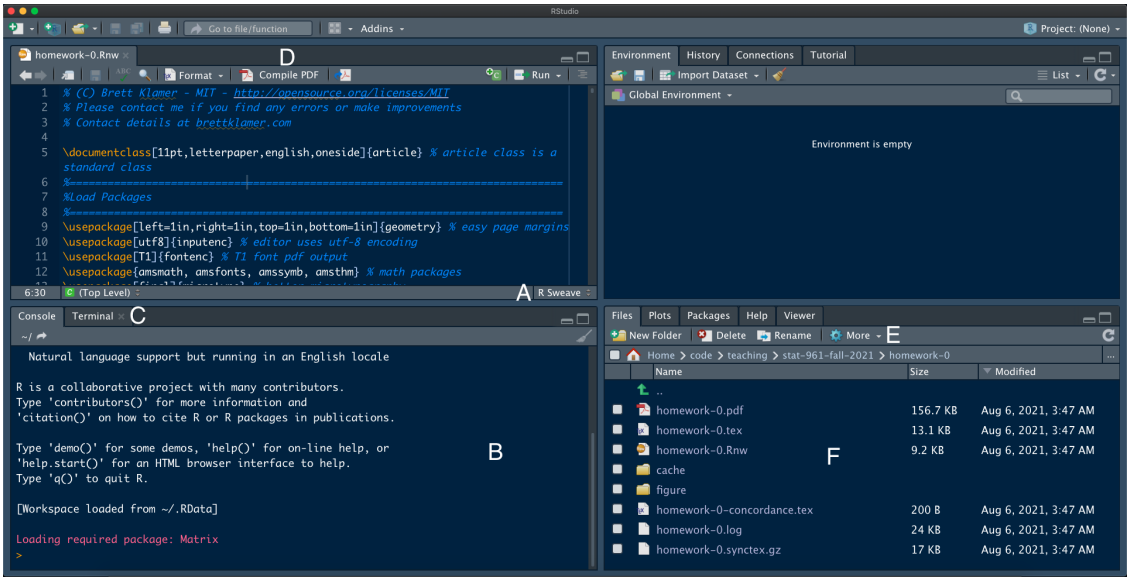

Figure 1: RStudio interface. Relevant parts are marked with white letters: A (`R Sweave` option), B (`Console`), C (`Terminal`), D (`Compile PDF` button), E (`More` button), F (`Files` pane).

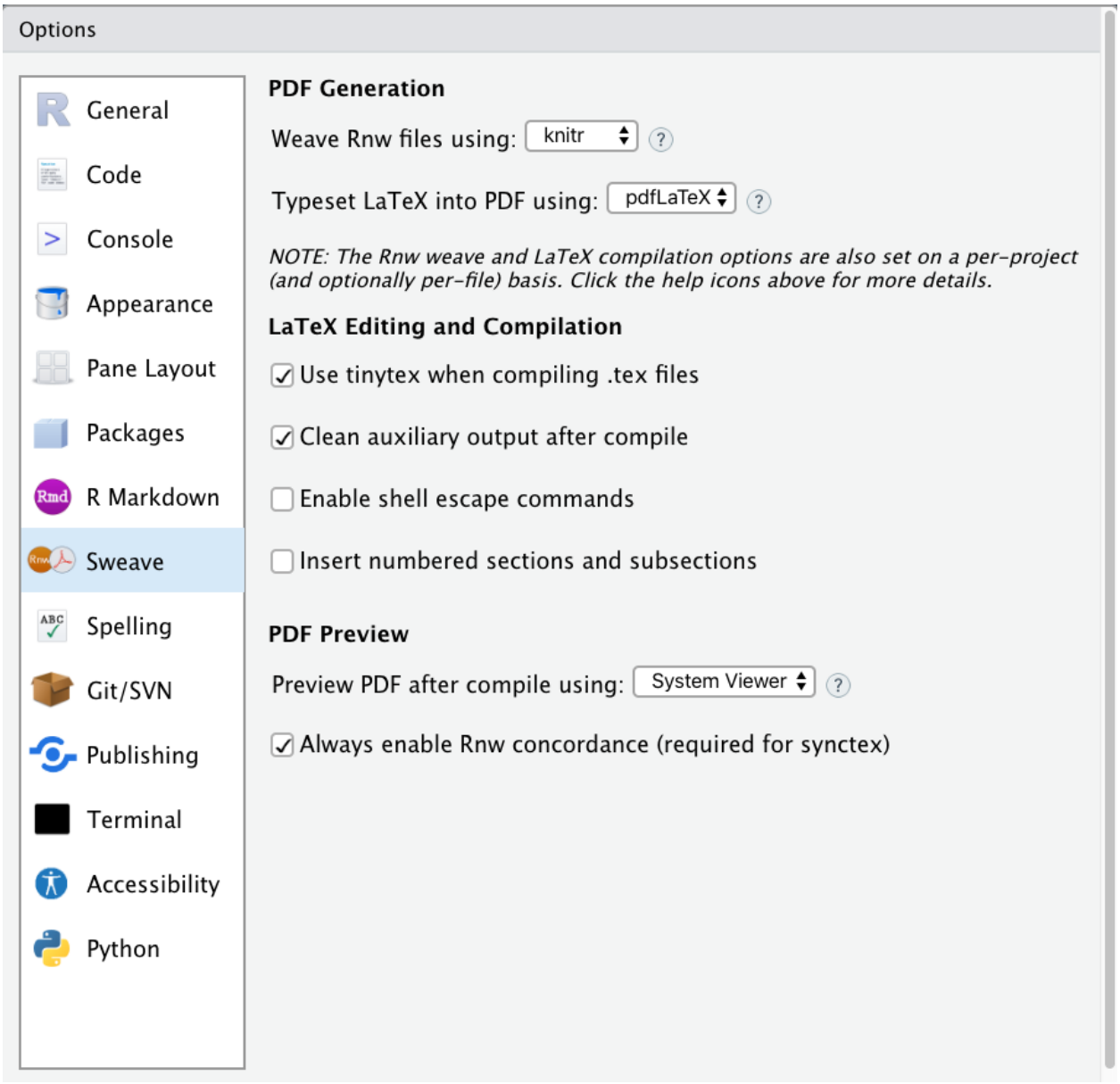

Figure 2: Options for PDF generation.

Git and Github setup. Git and Github are version control and code sharing tools used ubiquitously in applied statistics and data science. In STAT 961, you will use these tools to obtain course materials as well as to complete assignments. If you are unfamiliar with these tools (familiarity is not a prerequisite for this course), please read the first three sections of this webpage.

Course materials are stored in a Github repository (course repository) called `stat-961-fall-2021`, which also serves as the primary course webpage (the repository is private; see instructions below to gain access). You have read-only access to this repository. You will create a separate Github repository (personal repository) with the same name but under your personal Github account, which will contain your work in addition to all of the course materials available in the course repository. It is also private and other students will not have access to it. Both of these repositories will be linked to a folder named `stat-961-fall-2021` on your computer. See Figure 3 to visualize the relationship among these three repositories.

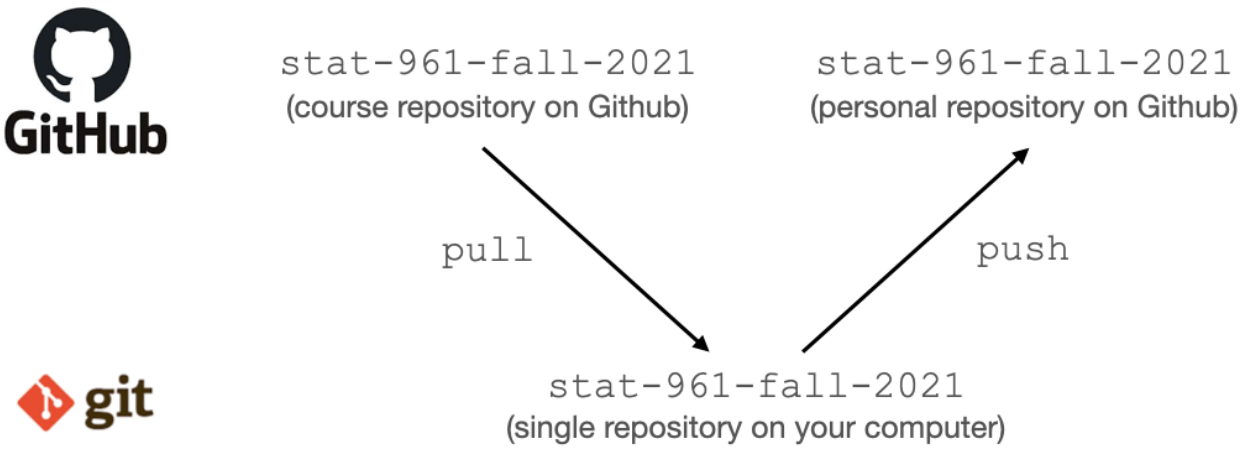

Figure 3: Relationship among Git repositories on Github and on your computer.

Follow the instructions below to set up these repositories.

1. Create an account on Github if you do not already have one, and log into your account.
2. Create a personal access token, if you do not already have one, by following these instructions. In step 6 of those instructions, name the token `Github access`. In step 7, choose the option for the token not to expire. In step 8, check the boxes next to `repo`, `notifications`, `user`, and `delete_repo`. When the token is created, copy it to your clipboard and/or a temporary file on your computer.
3. Click the course invitation link (please do not share this link with students not in the class). Then, follow the prompts to click `Authorize Github Classroom`, join `STAT 961 Students (Fall 2021)`, and then navigate to the class Github repository.
4. Navigate back to your Github profile (e.g. by clicking your icon in the upper-right corner and then clicking `Your profile`). Create a repository named `stat-961-fall-2021` on your personal Github account by clicking `Repositories`, clicking the green `New` button, choosing the options pictured in Figure 4, and clicking the green `Create repository` button. **Please make this repository private.**
5. Open a Terminal either via RStudio (see Figure 1C; works on any operating system) or by opening the Terminal application (works on Mac and Linux systems).

6. Check if you have Git installed by typing

```
git --version
```

   in the Terminal. If you get an error message indicating that git does not exist, install it by following the instructions here.

7. If you are on a Windows machine, go to `Tools` -> `Global Options` -> `Terminal`. In the drop down box for `New terminals open with`, select `Git Bash`. Restart RStudio.

8. To set up Git (if not already done so), type the following commands in the Terminal, using your name and email address:

```
git config --global user.name "Your name here"
git config --global user.email "your_email@example.com"
```

   Note that no messages will be printed as a result of these operations.

9. (This step is not necessary for students with Windows systems who installed Git in step 6.) Set up credential caching by following the instructions here, so you don't need to enter your personal access token every time. Make sure to follow the instructions for your operating system. If prompted for authentication during the installation of GCM Core, please enter your computer password (as opposed to your Github password or personal access token).

10. In the Terminal, change the directory to one within which you would like the `stat-961-fall-2021` directory to reside. For example, on my computer this directory is `/Users/ekatsevi/code/teaching`. To do so, you can directly type a `cd` command in the terminal, e.g. `cd /Users/ekatsevi/code/teaching`. Alternatively, in RStudio, you can do this by navigating to the directory in the `Files` pane (Figure 1E), clicking `More` -> `Set As Working Directory` (Figure 1D), clicking the down arrow next to the `Terminal` tab, and clicking `Go to Current Directory`.

11. In the Terminal, type

```
git clone https://github.com/katsevich-teaching/stat-961-fall-2021.git
```

    You will be prompted for your Github username and password. In place of the password, paste the personal access token you created in step 2 above. This `clone` operation will copy the STAT 961 Github repository onto your local computer. It will reside in an automatically-created folder called `stat-961-fall-2021`.

12. Change your directory to `stat-961-fall-2021`, similarly to step 7 above.

13. In the Terminal, type

```
git remote set-url --push origin https://github.com/[USERNAME]/stat-961-fall-2021.git
```

    Here, `[USERNAME]` should be replaced by your Github username. This command will tell Github that any changes you make will be synchronized to your personal repository as opposed to the course repository.

14. In the Terminal, type

```
git push
```

    If you successfully set up credential caching in step 8 above, you will not be asked for authentication. In this case, you can delete the temporary file you may have created in step 2 above to store your token. The `push` operation will have the effect of copying the contents of the repository to

your personal Github account. Go to `https://github.com/[USERNAME]/stat-961-fall-2021` and check that this operation succeeded.

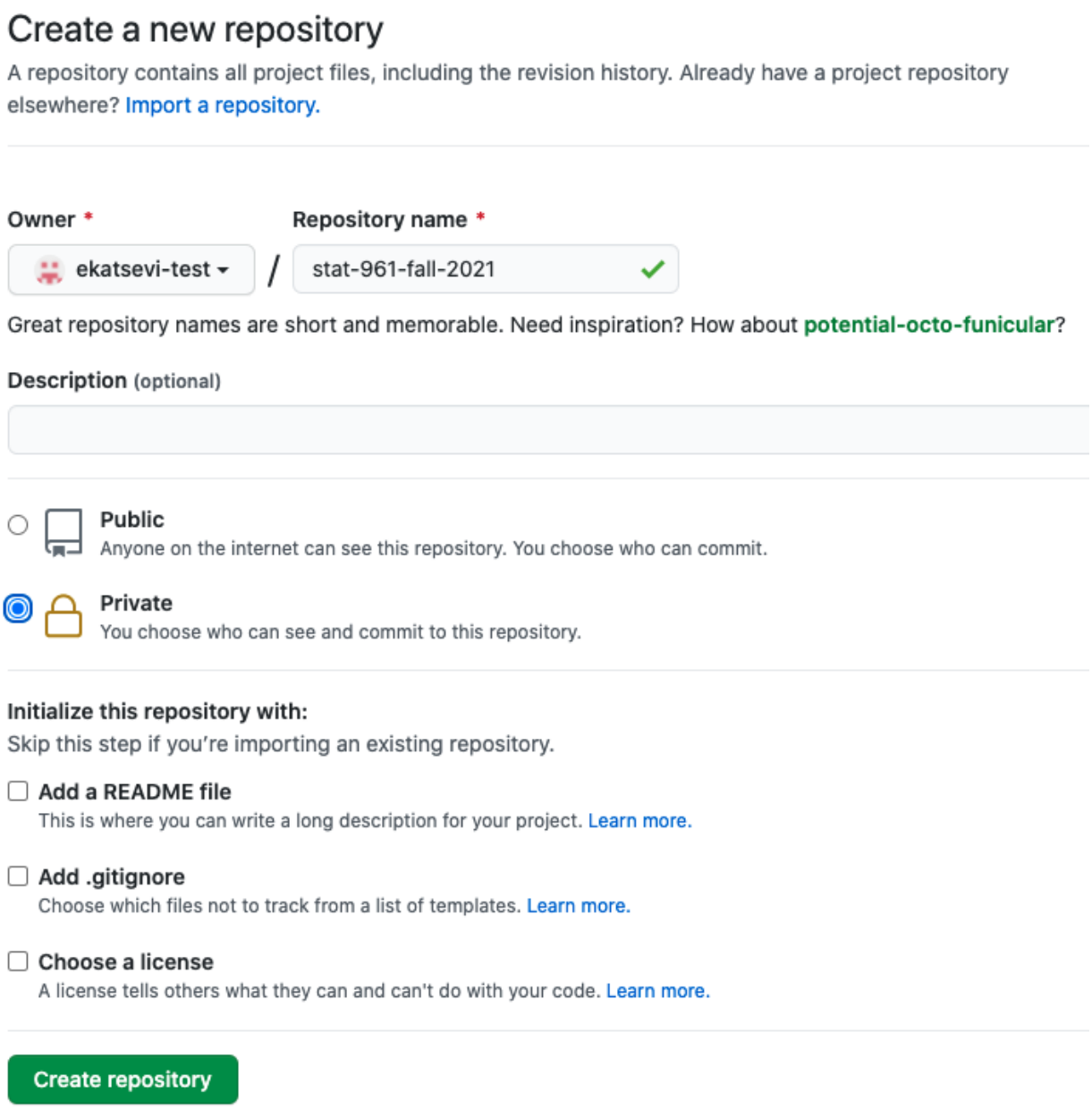

Figure 4: Creating a new Github repository for the class. Please make sure to check `Private` to make the repository private! The `Owner` at the top will be your username instead of `ekatsevi-test`.

Gradescope and Piazza setup. Assignment submission will be through Gradescope. Please verify that you can log into this page. If not, please sign yourself up using the entry code `ZRBXED`. Please also sign yourself up for the Piazza discussion forum.

2 Homework assignments

Assignment format. The homework assignment will be distributed as an Rnw document; see homework-0.Rnw for an example. This format facilitates the integration of LaTeX (for typesetting mathematical formulas) with R code (for statistical programming), producing a PDF report; see homework-0.pdf for an example. Read more about Rnw format here. For those unfamiliar with LaTeX, please browse this webpage. Useful articles on this page are those under the heading `mathematics` as well as the article on cross-referencing. For those unfamiliar with R, please read Chapters 1-5, 7, 9, 10, 12, 15, 17-21 of R for Data Science. Those familiar with R should also browse these chapters to fill in any gaps.

Assignment download, completion and submission. The following steps should be carried out for each homework assignment. Work through these steps for Homework 0.

- **Download.** Homework assignments will be made available on the class Github repository (under the homework folder). To download the assignment, navigate to the `stat-961-fall-2021` directory on your computer in the Terminal (either through RStudio or a separate Terminal window), as in step 5 of the section `Git and Github setup` above. Then, pull the latest version of the class repository from Github by typing

  ```
  git pull
  ```

  If the repository has changed on Github and on your computer, then Git will automatically merge these two sets of changes. In this case, Git will automatically open an editor to allow you to type an informative message about this merge. However, a default message is already present and you do not need to change it, so simply close this editor. The default editor Git uses is Vim, and to close it you can type `:q` and then press enter. *Note:* Other course materials (e.g. lecture notes) will also be distributed via the course Github. These materials should be downloaded using the same process as above.
- **Complete.** Open RStudio. Navigate to the directory (e.g. `homework-0/`) containing the homework assignment in the files pane (Figure 1F). Click `More` -> `Set As Working Directory` (Figure 1E). Then, click on the Rnw file (e.g. `homework-0.Rnw`) to open it in the RStudio editor. Select `R Sweave` from the menu in the bottom right-hand corner of the editor pane (Figure 1A). Change `FirstName LastName` to your first and last name. Work through the problems in the assignment, keeping the problem statements as they are and updating only the solutions between `\begin{sol}` and `\end{sol}`. When you make changes to the Rnw file, click `Compile PDF` near the top of the editor pane (Figure 1D) to update the PDF output file. Consult the log file if any issues arise with compilation; if issues persist then post on Piazza. It is convenient to place the RStudio and PDF preview windows side by side. To save your work, periodically *commit* it to Git by typing

  ```
  git add --all
  git commit -m "[commit message briefly describing the changes]"
  ```

  at the Terminal.
- **Submit.** When you have completed your assignment, compile to PDF and commit one last time. Then, push the changes to your Github repository by typing the following command in the Terminal:

```
git push
```

This will have the effect of synchronizing your files with your *personal* Github repository (but not the class repository). Submit the compiled PDF to Gradescope. If you are unfamiliar with Gradescope, please watch this brief video.