For Instructors: Instructions to the Lewis Acid–Base Predictive Model

Download and install R and RStudio

Download and install R: https://cran.r-project.org
Download and install RStudio: https://rstudio.com/products/rstudio/download/

RStudio is an integrative development environment for R. We recommend using RStudio and will provide images in this tutorial in the RStudio environment.

Preparation of student response data file

A .csv file is necessary for evaluation of student responses. This .csv file should contain a minimum of two (2) columns titled "ID" and "response" (case-sensitive). Each student response should be in its own row; there is no limit on how many rows there are.

You may find it useful to have an "ID" column for student name or student ID.

The following files should be in a single folder:
    (1) .csv file student responses
    (2) demo-data.csv
    (3) custom-stopwords.txt
    (4) replacements.csv
    (5) lewis-acid-base-instructors.R
    (6) lewis-acid-base-model.RData

Help: Setting the working directory

On Mac: it is easiest if you right-click on the demo-data.csv file, hold down the "option" key on your keyboard, then select: Copy "demo-data.csv" as Pathname. This will be the file pathname, so delete the ending part "demo-data.csv" to obtain the folder pathname.

On PC: it is easiest if you right-click on the demo-data.csv file and select "Copy as path". This will be the file pathname, so delete the ending part "demo-data.csv" to obtain the folder pathname.

Help: rJava

On Mac: if you obtain an error: Error: package or namespace load failed for 'rJava':

    1) Update and install the latest version of R/RStudio. If that doesn't work try:
   2) Refer to the file called "rJava-RStudio-macOS.pdf" for troubleshooting

Step 1. Open RStudio. Go to File → Open File → Select "lewis-acid-base-instructors.R". Your screen should resemble the one below.
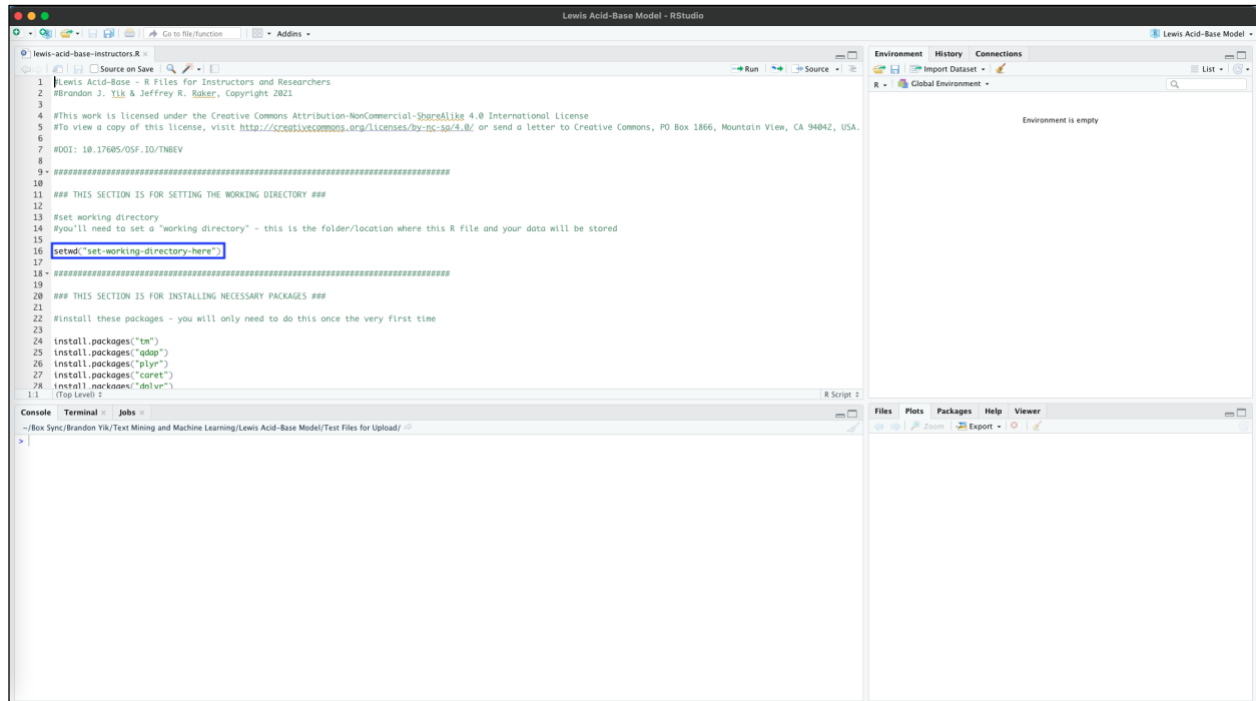
Step 2. In line 16, you will need to set your working directory. Currently it says: setwd("set-working-directory-here"). You will need to locate the directory of your folder containing your files and replace the set-working-directory-here portion of the text; make sure that the quotation marks remain. See page 1 of this tutorial if you need help locating the directory for your files. Highlight line 16 and click Run (blue box, below).
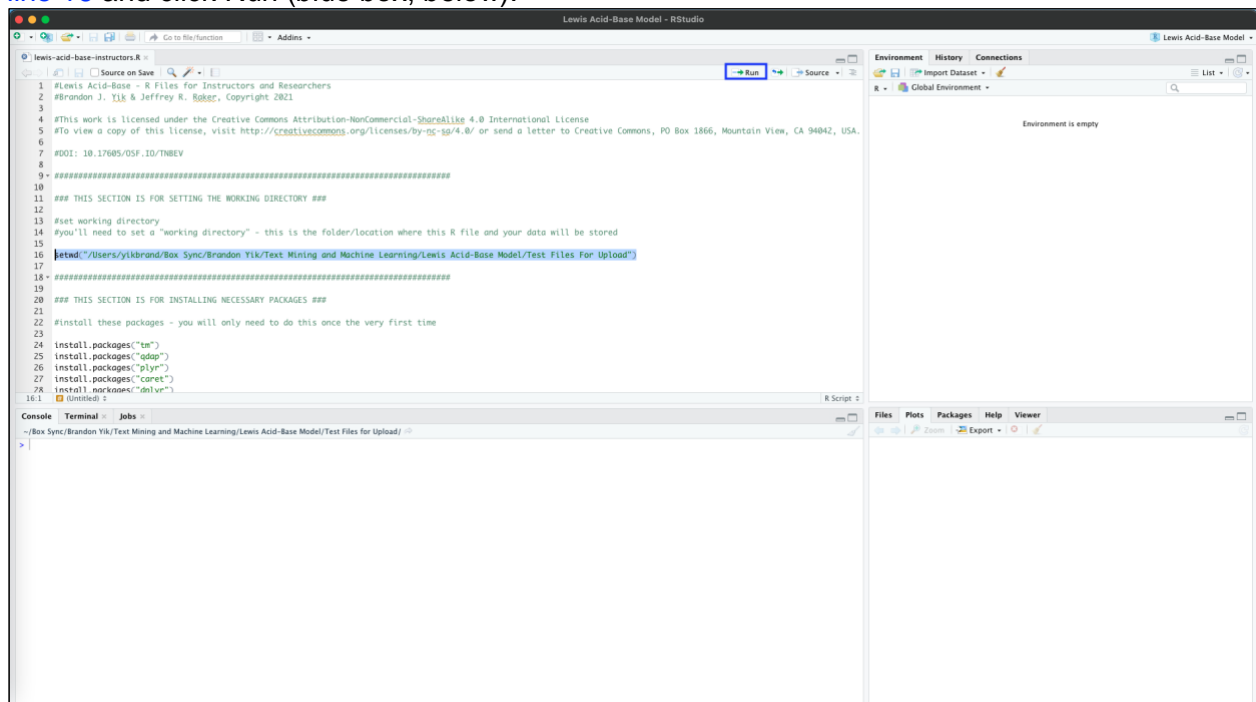
Step 3. In Finder (Mac) or Windows Explorer (PC), double-click on "lewis-acid-base-model.RData" to open. You may need to confirm loading the RData; select "Yes". Now, you should see to lines appear under the Environment/Data tab on the upper-right section of your screen (blue box, below).
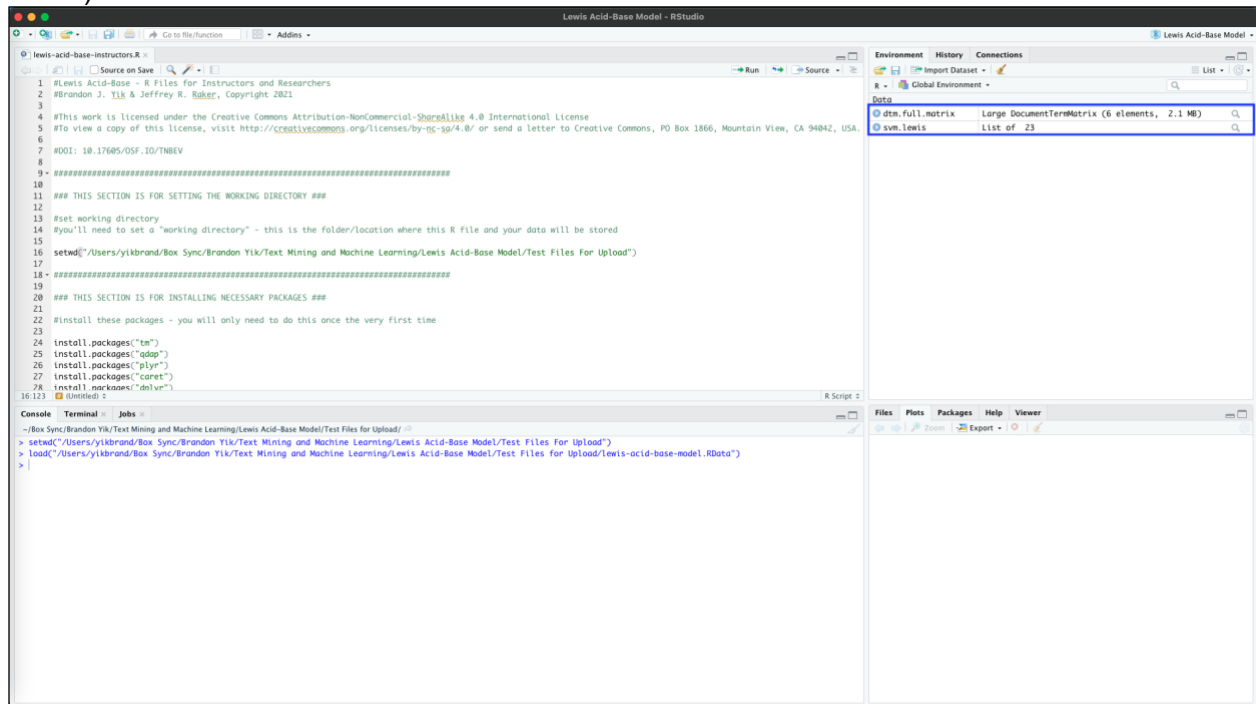

Step 4.
If this is your first time running this predictive model, you will need to install R packages that are necessary for this process. Highlight lines 24-29 and click Run (shown below). Some lines in black and blue will appear in the Console section on the lower-left part of the screen; this is normal.
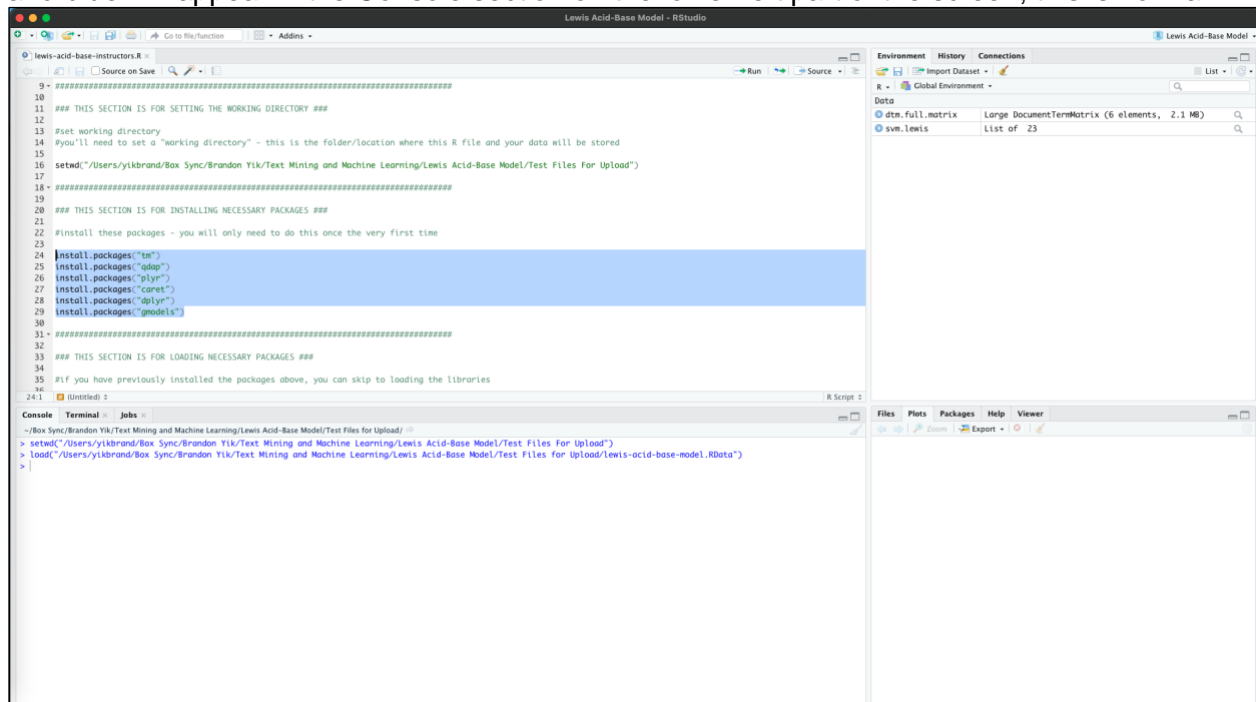
If this is not your first time running this predictive model or after the first time installing, highlight lines 38-43 and click Run (below).
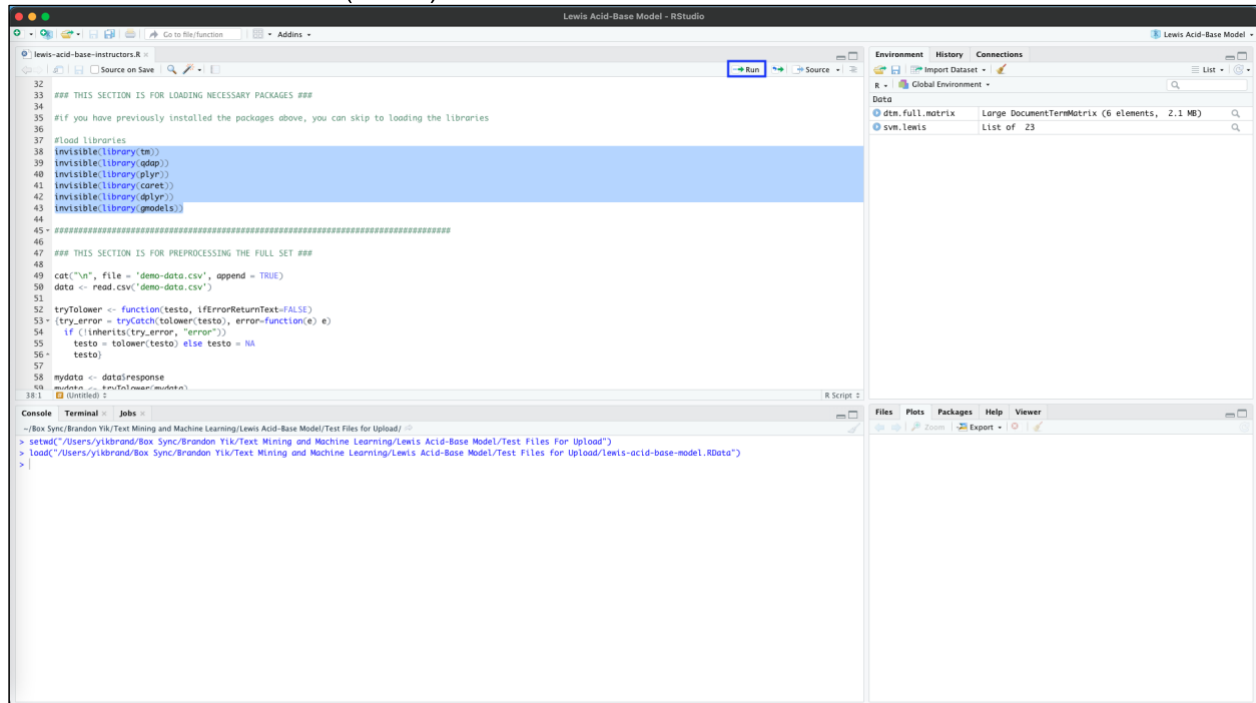
Step 5. Highlight lines 49-120 and click Run (below).
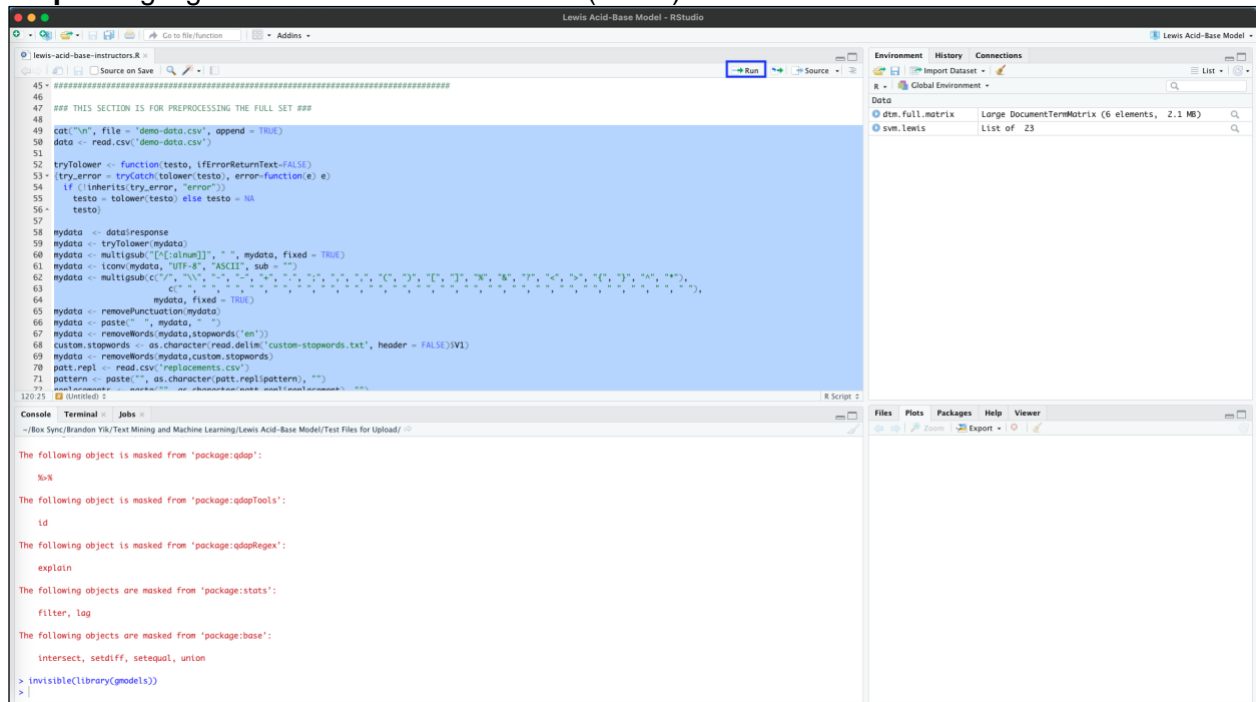
After this runs (it may take a few seconds/minutes), you should see a screen similar to the one below. The blue box shows the predictive model summary: N is the sample size, i.e., the number of responses; also called Total Observations in Table. The number of incorrect/non-use and correct and the fraction of the total responses (multiply by 100 for a percentage) are also given.

If you wish to obtain a .csv file of the responses with columns of the prediction (incorrect/non-use or correct) as well as a binary (0/1, respectively), highlight line 126 (green box, above) and click Run. This will generate a new .csv file named demo-data-predicted.csv in your folder located on your computer.

After opening your .csv file on your computer (it is shown in MS Excel, above), you will see two new columns compared to your original file: predict and predict.binary. The "predict" column will give the model predictions as correct or incorrect/non-use and the "predict.binary" column will give the model predictions as 1 or 0, respectively, for correct or incorrect/non-use.

Step 6. You are now ready to run your own students' responses!

The only difference between running the predictive model with the demo data and your own data is that you need to replace demo-data.csv with the filename of your own .csv file in lines 49-50. In line 126, you may rename the output filename of the newly generated .csv file with the model predictions.

Then, you may highlight lines 38-126 and click Run.