

Auto-graded formative coding assessments for life science students

eCOTS 2026: Sparking Joy and Discovery In a World of AI
June 15-18th, 2026

Dr. Gracia Dong

Assistant Professor, Teaching Stream
Human Biology Program and Department of Statistical Sciences
University of Toronto

Amanda Ng, MSc

Department of Statistical Sciences
University of Toronto



UNIVERSITY OF
TORONTO

DEFY
GRAVITY

STA288H1: Statistics and Scientific Inquiry in the Life Sciences

- Required introductory statistics course for students in all Human Biology specialist and major programs at the University of Toronto
 - Applied Genetics and Biotechnology, Global Health, Health and Disease, Human Biology, and Neuroscience
- Hybrid course, 36L/18P: ~2 hr asynchronous video lecture + 2 hr in-person session
- 2 sections x 200 students

R Coding in STA288H1

Course Assessments in Syllabus

- Weekly quizzes, marked for completion (12)
- In-person class activities, marked for completion (8)
- **R coding lab assignments**
 - (1 warm up + 3 “real” assignments)
- Midterm
- Group Research Project
- Final Exam

Concepts

- Data manipulation
- Summary statistics
- Bootstrap
- Simulation-based models
- Mathematical models
- Statistical inference: confidence intervals, hypothesis tests

R Coding in STA288H1

1. LearnR Modules

Self-guided interactive tutorials
Ungraded
Using publicly available datasets

3. Coding Lab Assignment

Using datasets from published literature
Students reproduce results from paper, create summary statistics, perform statistical inference
Written answers for justifying choices and interpretations
Graded by TAs

2. *NEW* Optional MarkUs Autograded Assessments for bonus marks

Autograded assignments marked for correctness
Students have multiple submission attempts that regenerate over time
Questions with numeric or multiple-choice answers
Using same datasets as LearnR Modules
Similar questions/concepts as Lab Assignments

4. Group Project, Final Exam, ...

MarkUs Assessments

- In the Winter 2026 offering of STA288H1, we trialed using MarkUs, an autograder (Magnin et al., 2012) to let students to check understanding and get immediate feedback before 1 of their 3 Lab Assignments
 - Goal is to bridge the gap between the modules and the formal Lab Assignments to improve student confidence
 - Unit tests are set up using the **RMarkus** package, which is built on top of the **testthat** package (Wickham, 2011)
 - Expected completion time, including reading instructions, submitting, running tests, correcting answers, resubmitting: 2 hours

Developing MarkUs Assessments (Unit Tests)

1. What Concepts Do We Want to Test?

Core statistical skills
Data wrangling steps
Simulation understanding
Plot construction
Interpretation

3. Modify Questions for Auto-Grading

Closed-form answers (logical, numeric, character, MCQ, plot object)
Explicit answer requirements: rounding, capitalization, axis labels, object names
Specify object names clearly so tests can reference them
Provide code hints
Set seed for simulations

5. Provide Meaningful Feedback

Targeted messages (“Your x-axis label is incorrect”)
Actionable hints (“Round to 3 decimals”)
Encouraging tone

2. Design Questions *Without* Worrying About Auto Grading

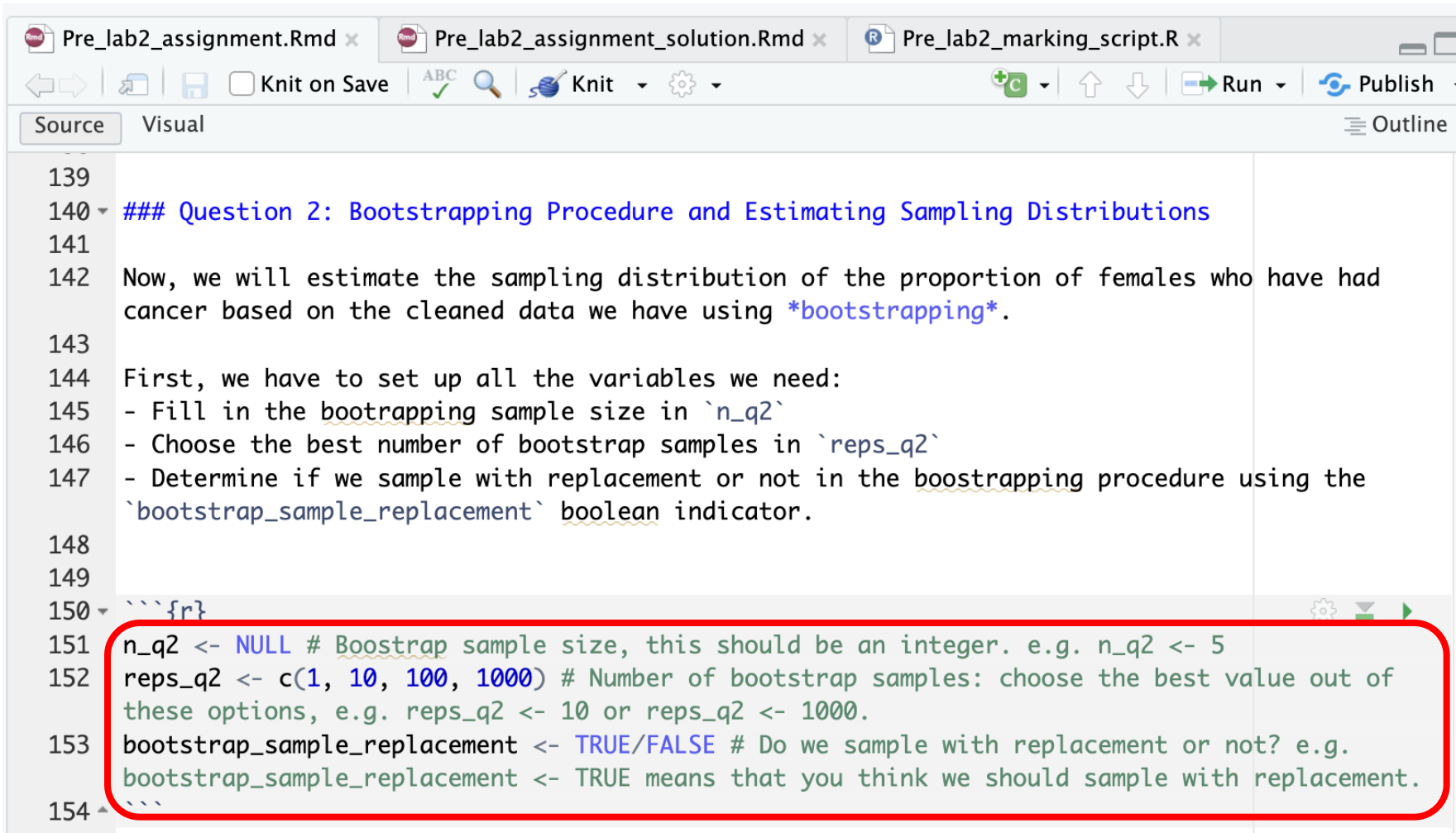
Write pedagogically sound questions first
Avoid open-ended tasks
Decompose the problem into smaller components

4. What We Test in the Auto grader

Correct answer (numeric, logical, character)
Data type (numeric vs. character vs. logical)
Intermediate objects (cleaned data, summary tables)
Plot layers + aesthetics

- correct geom
- correct aes mappings
- correct axis labels
- correct number of layers

Demo (question Rmd)

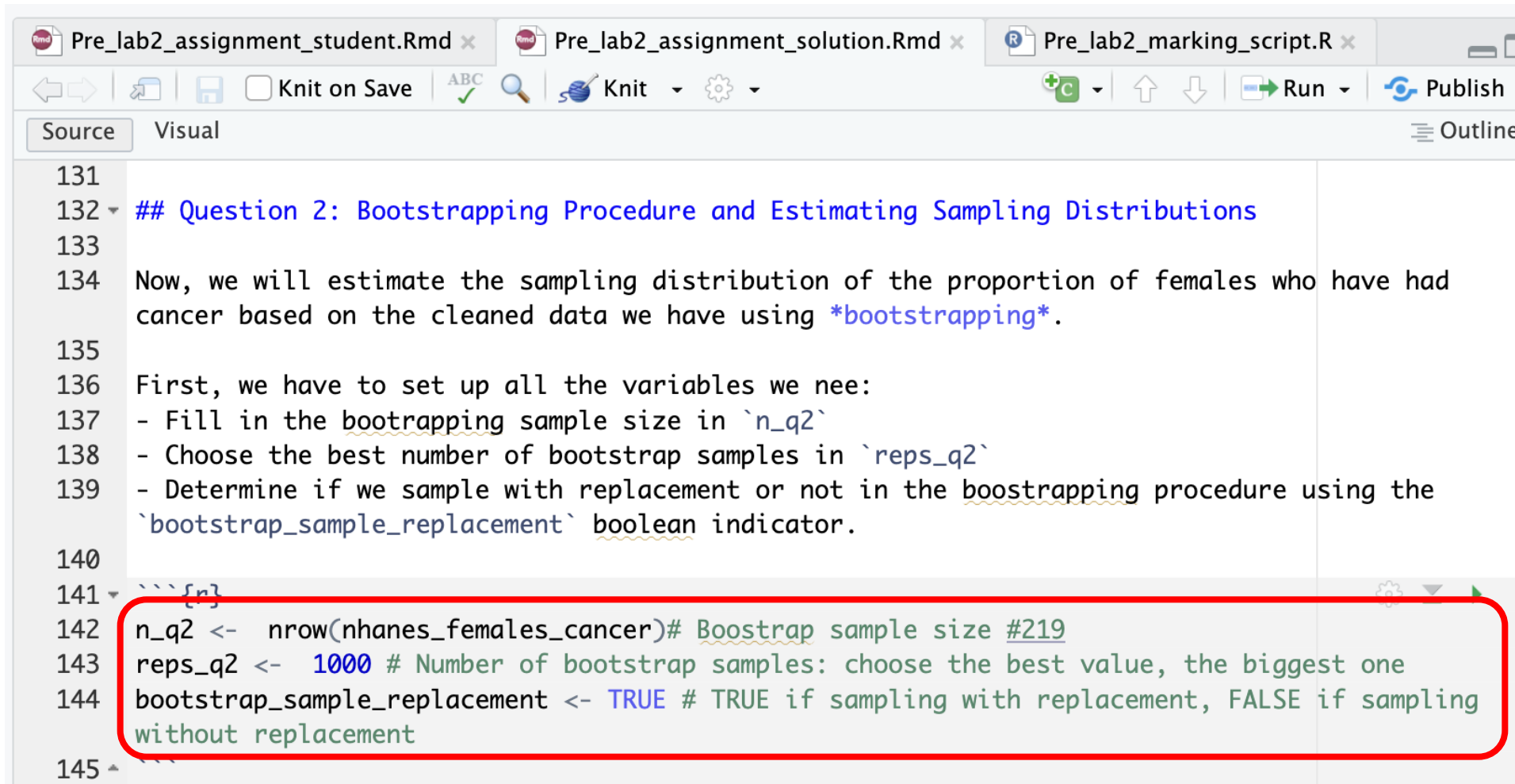


```
139
140 ### Question 2: Bootstrapping Procedure and Estimating Sampling Distributions
141
142 Now, we will estimate the sampling distribution of the proportion of females who have had
143 cancer based on the cleaned data we have using *bootstrapping*.
144
145 First, we have to set up all the variables we need:
146 - Fill in the bootstrapping sample size in `n_q2`
147 - Choose the best number of bootstrap samples in `reps_q2`
148 - Determine if we sample with replacement or not in the bootstrapping procedure using the
149 `bootstrap_sample_replacement` boolean indicator.
150
151 n_q2 <- NULL # Bootstrap sample size, this should be an integer. e.g. n_q2 <- 5
152 reps_q2 <- c(1, 10, 100, 1000) # Number of bootstrap samples: choose the best value out of
153 these options, e.g. reps_q2 <- 10 or reps_q2 <- 1000.
154 bootstrap_sample_replacement <- TRUE/FALSE # Do we sample with replacement or not? e.g.
155 bootstrap_sample_replacement <- TRUE means that you think we should sample with replacement.
```

Clear indication of
objects that will be
checked

Instructions on how to
input final answer

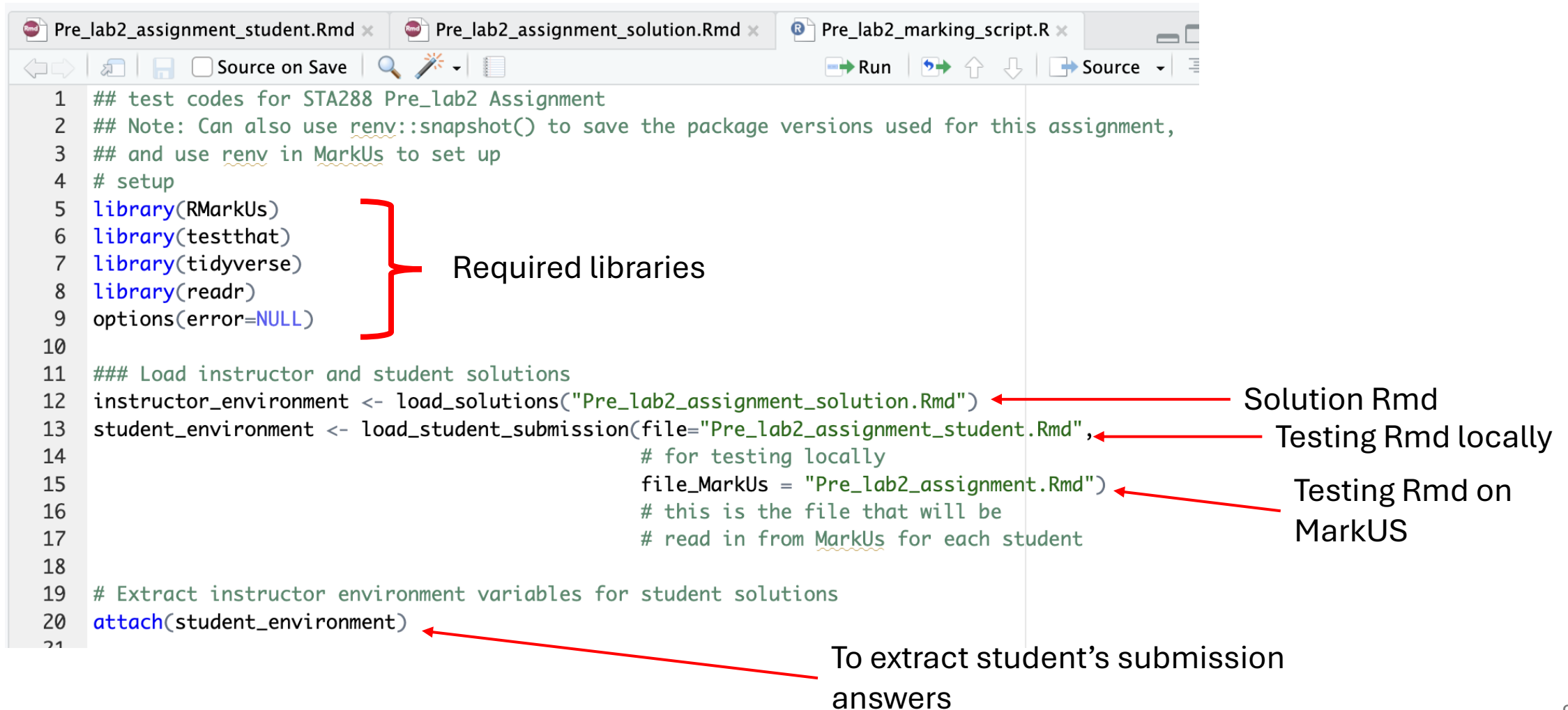
Demo (solution Rmd)



```
131
132 ## Question 2: Bootstrapping Procedure and Estimating Sampling Distributions
133
134 Now, we will estimate the sampling distribution of the proportion of females who have had
135 cancer based on the cleaned data we have using *bootstrapping*.
136
137 First, we have to set up all the variables we need:
138 - Fill in the bootstrapping sample size in `n_q2`
139 - Choose the best number of bootstrap samples in `reps_q2`
140 - Determine if we sample with replacement or not in the bootstrapping procedure using the
141 `bootstrap_sample_replacement` boolean indicator.
142
143 n_q2 <- nrow(nhanes_females_cancer) # Bootstrap sample size #219
144 reps_q2 <- 1000 # Number of bootstrap samples: choose the best value, the biggest one
145 bootstrap_sample_replacement <- TRUE # TRUE if sampling with replacement, FALSE if sampling
146 without replacement
```

Solutions to check

Demo (marking R script)



The screenshot shows an RStudio window with three tabs: `Pre_lab2_assignment_student.Rmd`, `Pre_lab2_assignment_solution.Rmd`, and `Pre_lab2_marking_script.R`. The `Pre_lab2_marking_script.R` tab is active, displaying the following R code:

```
1 ## test codes for STA288 Pre_lab2 Assignment
2 ## Note: Can also use renv::snapshot() to save the package versions used for this assignment,
3 ## and use renv in MarkUs to set up
4 # setup
5 library(RMarkUs)
6 library(testthat)
7 library(tidyverse)
8 library(readr)
9 options(error=NULL)
10
11 ### Load instructor and student solutions
12 instructor_environment <- load_solutions("Pre_lab2_assignment_solution.Rmd")
13 student_environment <- load_student_submission(file="Pre_lab2_assignment_student.Rmd",
14                                               # for testing locally
15                                               file_MarkUs = "Pre_lab2_assignment.Rmd")
16 # this is the file that will be
17 # read in from MarkUs for each student
18
19 # Extract instructor environment variables for student solutions
20 attach(student_environment)
```

Annotations with red arrows point to specific lines of code:

- A red bracket groups lines 5 through 9, labeled "Required libraries".
- A red arrow points from line 12 to the text "Solution Rmd".
- A red arrow points from line 13 to the text "Testing Rmd locally".
- A red arrow points from line 15 to the text "Testing Rmd on MarkUS".
- A red arrow points from line 20 to the text "To extract student's submission answers".

Demo (marking R script)

```
37 # Question 2
38
39 testScalar(variableName = "n_q2",
40             student_environment = student_environment,
41             instructor_environment = instructor_environment,
42             correct_error_msg = "Incorrect. Is your bootstrap sample size the same as the actual",
43             check_datatype = FALSE)
44 testScalar(variableName = "reps_q2",
45             student_environment = student_environment,
46             instructor_environment = instructor_environment,
47             correct_error_msg = "Incorrect. Recall that we need the number of bootstrap simulati",
48             check_datatype = FALSE)
49 testScalar(variableName = "bootstrap_sample_replacement",
50             student_environment = student_environment,
51             instructor_environment = instructor_environment,
52             correct_error_msg = "Incorrect. We need to sample with replacement for a bootstrap c",
53             check_datatype = FALSE)
```

For more testing functions, visit: <https://github.com/RAutoGrading/RMarkUs>

Demo (MarkUs)

MarkUs Statistics and Scientific Inquiry in the Life Sciences

Dashboard	Assignments	Users	Notes	Grades Summary	Course Settings
M2_assignment	Settings	Groups	Graders	Submissions	
Properties	Tags	Criteria	Starter Files	Annotations	Automated Testir

Edit Assignment: M2_assignment

Required Fields*

Properties

Short Identifier* M2_assignment

Assignment description* M2 testing

Message

Due date* 2026-02-28 12:00 PM

Visibility ☐ Hidden ☒ Visible ☐ Visible during scheduled time

(If the section-specific settings option is checked, the value for individual sections will override this setting.)

Allow students to submit through the web interface ☒

Allow students to submit through a version control system ☐

Allow students to submit URLs ☐

Allow students to submit through the MarkUs API ☐

Display assignment median to students ☐

Display grader names to students ☐

Only allow students to view released marks via a unique URL ☐

Required files

You may specify file(s) that students are required to submit.

[Add a Required File](#)

Filename* M2_assignment_student (Delete ☐)

Filename* M2_assignment_student (Delete ☐)

Only allow students to submit the required files (this also restricts the file extensions) ☒

Group Properties

Students can work in groups ☐

Late Submission Policy

- ☒ Accept no late submissions
- ☐ Automatically deduct grace credits
- ☐ Use penalty decay formula
- ☐ Set manual penalty periods

Group work
Late submission
Regrade request

Google Chrc

Demo (MarkUs)

MarkUs Statistics and Scientific Inquiry in the Life Sciences Wai Yu Amanda Ng ▼

Dashboard

Assignments

Users

Notes

Grades Summary

Course Settings

M2_assignment (hidden)

Settings

Groups

Graders

Submissions

Grades

Test Runs Status

Properties

Tags

Criteria

Starter Files

Annotations

Automated Testing

Groups already have been created for this assessment.
If any changes are made, students will need to log in to MarkUs to download a new version of their starter files.

Starter File Groups

New Starter File Group

Filename

M2_assignment.Rmd
nhanes_data_dictionary.csv
nhanes_subset.csv



+

Download All Upload File Add Folder Delete

Submitted At

Revised By

Monday, February 23, 2026, 05:52:45 PM EST
Monday, February 23, 2026, 05:38:28 PM EST
Monday, February 23, 2026, 05:38:36 PM EST

Assignment Rmd,
data files

Starter File Assignment Rules

Download assigned starter files csv | Download starter files example

- ☒ Assign the default starter file group to all students
- ☐ Assign starter file groups by section
- ☐ Randomly select one starter file group
- ☐ Randomly select one top level file or directory from each starter file group

Default Starter File Group New Starter File Group ▼

☒ Starter files are available to students after the due date has passed

Save

Demo (MarkUs)

MarkUs Statistics and Scientific Inquiry in the Life Sciences Wai Yu Amanda Ng ▼

Dashboard

Assignments

Users

Notes

Grades Summary

Course Settings

M2_assignment (hidden)

Settings

Groups

Graders

Submissions

Grades

Test Runs Status

Properties

Tags

Criteria

Starter Files

Annotations

Automated Testing

Automated Testing

Groups already have been created for this assessment.
If any changes are made, students will need to log in to MarkUs to download a new version of their starter files.

Tests have been run for this assignment. Please re-run tests once the settings have been updated.

☒ Enable tests for this assignment

Test Files

Filename

-  M2_assignment_solution.Rmd 
-  M2_marking_script.R 
-  nhanes_data_dictionary.csv 
-  nhanes_subset.csv 
-  renv.lock 

} solution Rmd,
Marking R script,
data files,
Renv.lock

 Download All  Upload File  Add Folder  Delete

Submitted At

Revised By

Monday, February 23, 2026, 05:37:04 PM EST	
Monday, February 23, 2026, 05:37:16 PM EST	
Monday, February 23, 2026, 05:37:23 PM EST	
Monday, February 23, 2026, 05:37:29 PM EST	
Monday, February 23, 2026, 05:37:43 PM EST	

Demo (MarkUs)

Testers

Tester type*

R environment*

Use renv to set up environment ☒

Package requirements

Test Groups*

Test group name

Display test output to students?*

Criterion

Run by

Feedback files

Test files*

Timeout (s)*

Student-Run Tests

Allow students to run tests ☒

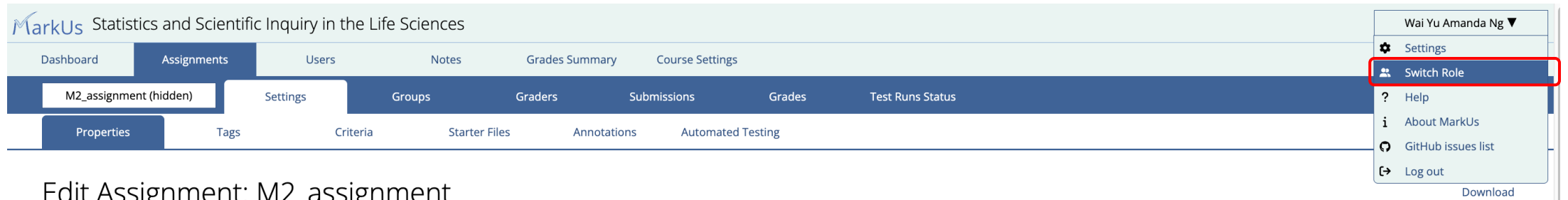
Tokens per group (Unlimited tokens ☒)

Tokens available on

Tokens available until

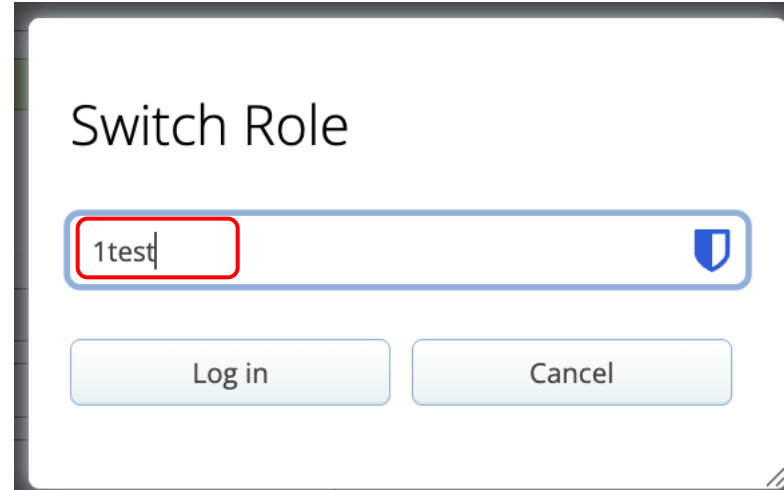
Tokens regenerate after hour(s) (Tokens do not regenerate ☐)

Demo (MarkUs Testing)



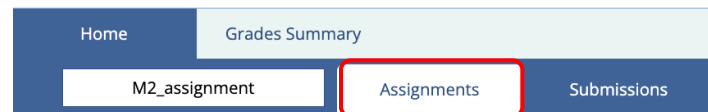
The screenshot shows the MarkUs web interface. The top navigation bar includes the MarkUs logo and the text "Statistics and Scientific Inquiry in the Life Sciences". Below this is a horizontal menu with tabs: Dashboard, Assignments, Users, Notes, Grades Summary, and Course Settings. The "Assignments" tab is active, showing a sub-menu with "M2_assignment (hidden)", Settings, Groups, Graders, Submissions, Grades, and Test Runs Status. The "Settings" sub-tab is selected, showing further options: Properties, Tags, Criteria, Starter Files, Annotations, and Automated Testing. On the right side, a user profile dropdown for "Wai Yu Amanda Ng" is open, with a red box highlighting the "Switch Role" option. Other options in the dropdown include Settings, Help, About MarkUs, GitHub issues list, and Log out. A "Download" link is visible at the bottom right of the dropdown.

Edit Assignment: M2 assignment



The "Switch Role" dialog box is shown. It has a title "Switch Role" and a text input field containing "1test". The input field is highlighted with a red box. Below the input field are two buttons: "Log in" and "Cancel".

Demo (MarkUs Testing)



M2_assignment: M2 testing

Assignment information

Due date: Saturday, February 28, 2026, 12:00:00 PM EST

Policies

You must complete this assignment individually.

Starter files

Starter files have been provided for this assignment.

Starter files were last updated on Monday, February 23, 2026, 05:52:45 PM EST

[Download Starter Files](#)

Submissions

- 2 files submitted
- Last revised Monday, February 23, 2026, 05:41:42 PM EST

Required files

- M2_assignment_student.Rmd
- M2_assignment_student.pdf

0 missing required files



M2_assignment: Submissions

The due date for this assignment has passed. Any changes you make to your submitted files will be recorded, but not graded.

Required files

Good job! All required files have been submitted.

☒ M2_assignment_student.Rmd

☒ M2_assignment_student.pdf

You may only submit the required files listed above.

Maximum file upload size: 10 MB

[Download All](#) [Submit File](#) [Add Folder](#) [Delete](#)

Filename

M2_assignment_student.pdf

Submitted At

Revised By

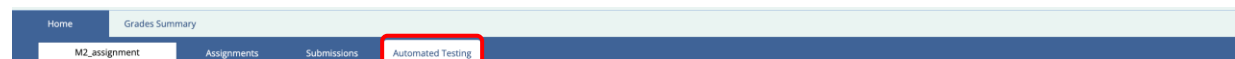
Monday, February 23, 2026, 05:41:42 PM EST

1test

M2_assignment_student.Rmd

Monday, February 23, 2026, 05:41:29 PM EST

1test



Automated Testing

The due date for this assignment has passed.

Tokens available on: Monday, February 23, 2026, 05:36:36 PM EST

Tokens available until: Saturday, February 28, 2026, 12:00:00 PM EST

Tokens available: Unlimited tokens

You must submit at least one file in order to run tests.

Test Results

[Run tests](#) [Cancel tests](#)

Monday, February 23, 2026, 05:51:28 PM EST		Run by 1test		complete
Test Name		Status	Mark	
		All		
Test group				45 / 45
[M2_marking_script.R] q1a_random_sample Exists		pass	1 / 1	
[M2_marking_script.R] q1a_random_sample datatype test		pass	1 / 1	
[M2_marking_script.R] q1a_random_sample correct length test		pass	1 / 1	
[M2_marking_script.R] q1a_random_sample correct value		pass	1 / 1	

Student Feedback

General information

Participants: 2026 Winter STA288H1 students (330 students enrolled at the end of term, 231 students completed the MarkUs assignment; 285 survey responses, 212 after data cleaning)

We collected student feedback in their weekly check-in survey

Student Feedback

Engagement & Navigation



- > 60% of respondents ran the MarkUs tests 3 or less times.
- Moderate ease of navigation with room for improvement.

Time to Completion



- 55% completed the MarkUs assignment, including the submission process, within 2 hours.
- 85% completed within 3 hours → indicates reasonable workload for a formative activity.

Perceived Learning Value



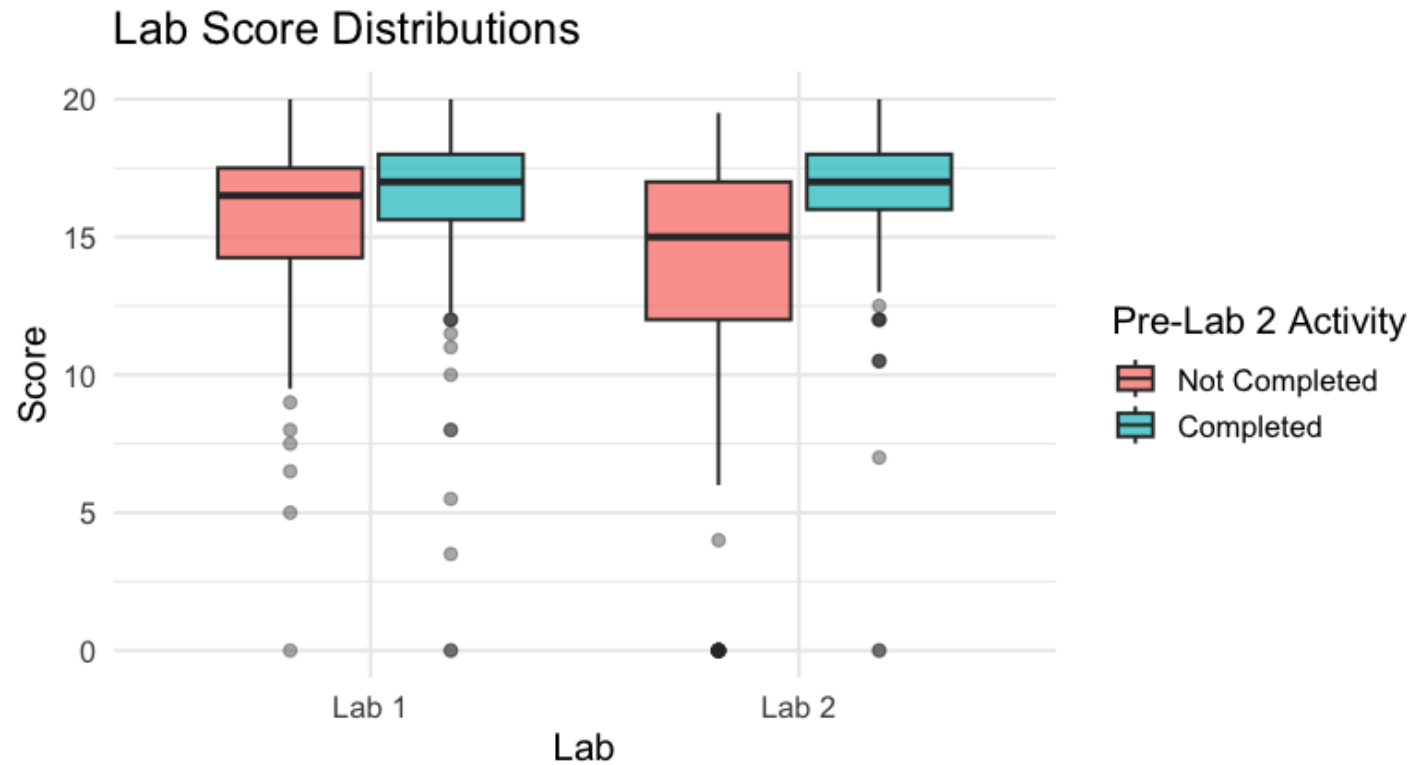
- 95% found the module helpful for preparing for the graded assignment.
- 93% reported increased confidence with STA288H1 knowledge and skills.

Future Development



- > 80% rated the idea of developing similar graded Pre-Lab activities as moderately to highly beneficial.
- Sentiment analysis of open-ended comments was generally positive → clear interest in expansion

Students' Results Comparison



- Students who completed the pre-lab activity scored roughly 2.8 points more on Lab 2 (compared with Lab 1) than students who did not
- This difference is significant ($p\text{-value} < 4.67\text{e-}06$).

Next steps

- Create cohesive sets of LearnR module + MarkUs + Lab Assignments to deploy in future offerings of STA288H1 and other courses
- Expansion to other quantitative courses for life science students: e.g. HMB325H1: Biostatistics and Data Science for Human Biology



Thank you!

Questions or Comments?

- Post a comment in the discussion at the bottom of <https://www.causeweb.org/cause/ecots/ecots26/program/posters/28>
- Synchronous Session: Tuesday, June 16th, 2:45-3:30 EST



Contact: gracia.dong@utoronto.ca

References

- Magnin, Morgan, et al. "Markus: An open-source web application to annotate student papers on-line." *Engineering Systems Design and Analysis*. Vol. 44878. American Society of Mechanical Engineers, 2012.
- Wickham. *testthat: Testthat code. Tools to make testing fun :)*, . R package version 0.5, 2011. URL <http://CRAN.R-project.org/package=testthat>.
- RMarkUs Github Repository:
<https://github.com/RAutoGrading/RMarkUs>