Milestone 2 day: get to a finished data file

Eleutherodactylus Research Course · Fall 2026

Tom Devitt

2026-11-02

Today is a work day

Milestone 2 (a clean data file and a notebook that renders from a fresh session) is due now.

Most of the next 50 minutes is yours, at your notebook. I will teach for about 10 minutes, then Tom and the mentors visit each team.

The aim for today is simple: **find the one thing that stops your notebook, and fix it.**

Start now: render from a fresh session

The steward does this first, while the rest of the team looks at the next slide.

1. Quit Session (RStudio, top right)
2. Log back in and open your team notebook
3. Click **Render**

To *render* is to run the whole notebook from the top and produce the finished page. If it stops, the **first error message** is the most useful thing you have. Read it out to your team. Errors after the first one are usually just fallout from it.

Where did it stop?

It stops at	Usual cause	What to do
Setup or Step 2	A file path or folder name is wrong, or a package is missing	Check the path in the error against what is in your team folder
Step 3 (merging selections)	A file is named wrong, a column header is wrong, or two people measured the same recording	Read the message: Step 3 names the file and the problem
Step 4 (measuring)	A very short call, or an old copy of the notebook	Ask a mentor to check that you have the current template
Steps 5 and 6	An earlier step produced no rows, or a measurement is empty	Check the row count after Step 3

If nothing here matches, post the first error to your team and a mentor. Do not guess for more than 10 minutes.

Your next move, by step

Use the four steps from the Oct 26 status card. Pick the highest one you can honestly say yes to.

You are at	Do next
1. Renders through Step 2	Get a clean render first. One error at a time, steward and mentor. Do not measure more calls until it renders
2. Merged, you have a row count	Is the row count what you expected? Find the calls that look wrong and re-measure only those, using your calibration rule
3. Every call has measurements	Look at the biggest and smallest value of each parameter. Check one outlier against its spectrogram. Then build the summary table
4. Clean file and summary table	Write the caveats and the methods paragraph. If *n* is small, say so with the number, or measure more calls from recordings you have not used

Team 1: the same four steps, with your own stages (detector runs, test set built, clips labeled, scores and the TP/FP/FN/TN table).

What "finished" looks like

You can say yes to each of these:

1. The notebook renders from a fresh session, through the last step
2. There is one measurements file, one row per measured call, and you know the row count
3. The numbers in your notebook text match the file (how many males, how many calls)
4. You checked an outlier and wrote down whether you kept or corrected it
5. Your caveats are in the notebook, each with a number: missing files, missing temperatures, mp3 sources, small *n*
6. Someone else on the team could run it without changing a path

(A *caveat* is a limit on what your data can tell a reader.)

Ten minutes: write your methods from the notebook

Your notebook already holds most of your methods section. Writing it now shows you what is missing.

A methods paragraph answers, in order:

- **What recordings?** How many, from how many males, wav or mp3
- **How were calls measured?** Window and boundary rules, the calibration result, window length 512, any filter
- **What was measured?** Parameters, defined as in Kohler et al. 2017
- **What are the limits?** The caveats from the last slide

A skeleton to fill in

Copy this into the end of your notebook and replace every blank with a number or a name from your own data:

```
We measured ___ calls from ___ males of ___ (___ wav and ___ mp3
recordings). Call start and end were selected by hand on the oscillogram
by ___ observers; the observers' agreement on a shared set of recordings was
___ (calibration result). Measurements followed Kohler et al. (2017):
___. Because ___ recordings are mp3, we report only durations, rates,
note counts, and dominant frequency for those. Ambient temperature is
available for ___ of ___ specimens.
```

A blank you cannot fill is a finding. Write it in your caveats.

If your notebook is not finished today

That is normal at this point. What matters is that we can see where it is:

- Leave the notebook as it is, with the first error visible
- Add a short list at the bottom: what works, what is blocking, what you will fix first
- Tell a mentor which step you reached

Wednesday is a work session. We use it to fix what today shows is missing.

Rest of class

Time	What
Now to the end	Work at your notebook. Tom and the mentors visit each team
Wed Nov 4	Work session and stand-ups
Nov 9 and 11	Catch-up week: open work time, no new deadlines
Mon Nov 16	**Milestone 3:** complete figures and results. Your figures come from your clean data file
Wed Nov 18	AmphibiaWeb second draft due

Credits

Prepared with Claude Code (Claude Sonnet 5.5, Anthropic): first draft of the deck from the team notebook template and the Milestone 2 checklist.