GENERAL

Lily liked the hands-on, real-data exercises.

TERM PAPER

See S13: Add back the peer-review process? I like the idea of research presentations (F20) combined with more of a writeup and peer review (S13). Think about how best to combine.

Idea! Have them write four pages for their first draft, then condense it down to two pages for their second draft. Make it tight. Surprise them with the word cut? Recommend that they write long for first draft

READINGS/DISCUSSION

Reading list has too much of my stuff; thin out.

Cut the Williams et al. Model Systems paper from ‘Why Study the Past’ and move it as an alternate reading for no-analog/novel.

**Why Study the Past**: Add a pollution/PIrlA example, not just climate?

**Zoom Lens**: Ditch Delcourts, add a modern pollen-vegetation paper like Dawson or Pirzambein.

**Novelty**: Add the new Pandolfi paper, ditch Finsinger?

sedaDNA: Maybe redundancy between Parducci and Balint. Add a cool data paper like Clarke?

NEOTOMA 2 REBUILD

See my notes and links to online resources in C:\Jack\ResProj\Neotoma\Workshops\20220607\_NeotomaDB\_Rneotoma2

Maybe shift to an online Juypter / binder system? Apparently Geography has a Juypter notebook acct

LAB READINGS

MAT: Add the new Chevalier paper

LAB EXERCISES

Check out Juggins’ resources (in email / teaching / 335 / resources) for microfossil identification and stratigraphic diagramming. Could be a great early exercise.

Slack was a \*great\* idea

Lab 1:

1. Advise searching for just Devil

Lab 2:

1. some confusion about matrix algebra.
2. Some challenges for Apple installation
3. Check Owen: ggplots 🡪 ggplot2
4. Google is your friend!
5. Great resources:
   1. Order of operations: <https://www.mathsisfun.com/operation-order-pemdas.html>
   2. <https://www.mathsisfun.com/algebra/matrix-introduction.html>
   3. <https://www.khanacademy.org/math/precalculus/x9e81a4f98389efdf:matrices/x9e81a4f98389efdf:multiplying-matrices-by-matrices/v/multiplying-a-matrix-by-a-matrix>
   4. <https://www.mathsisfun.com/algebra/matrix-inverse.html>
   5. <https://www.khanacademy.org/math/precalculus/x9e81a4f98389efdf:matrices/x9e81a4f98389efdf:intro-to-matrix-inverses/v/inverse-matrix-part-1>
6. Let students know that in Lab 5, we’ll go into these matrix operations in a bit more detail.

Lab 3:

1. Lecture too long! Spent a long time on foundational concepts, ran out of time. But useful still. So, look for ways to cut less-important slides. Most/all the Neotoma/Geoinformatics stuff can probably go.
2. Age-Depth model: Maybe add / pick a different site, a weirder one? Devil’s Lake pretty well behaved and constrained, so not too much happens with different age models there.

Lab 4:

1. APIs currently pretty fubared, so had to strip out a lot of lab content. See the 523\_Clean version for a earlier version with more API Exercise Questions.
2. Get\_dataset(666) is supposed to expect a siteID (666) but is apparently expecting a datasetID (666, a Delcourt sample).
3. Question 5 is screwing people up.
4. Add some sample questions for browse and get\_publication. Also more sample questions before ‘make a pollen diagram’ – a big jump up there.
5. There is a \*lot\* of advanced code here by Simon. Need to better annotate.
6. See if I can actually run Simon’s code, particularly for Q6 & Q7, and make my own version for Jack Script

Lab 5. MatrixMult/EDA

1. In part 3, Check covariance, order of matrix multiplication seems to be wrong.
2. Possibly move this up, to go right before/after Lab 2 and R Intro.

Lab 6: Dissim/ROC

1. Use rioja:interp.dataset instead of approx. for the linear interpretation and provide model code. Lots of confusion about how to linearly interpolate
2. See ModelCode for Kate for some snippets to put in there.
3. Sample worksheet: Give a full example for ED and have students work out the other two? Add some pointers about Excel formulas?

Lab 7: Novelty

1. In code for mat jant, set k=4, not k=10
2. Should clarify that R2 better for intervariable comparison and rmsep for prediction. Explain/clarify these metrics. From Slack:
   1. R2 and RMSEP are two slightly different measures of goodness of fit...  The differences are subtle, but basically R2 is sensitive to the overall length of the environmental gradient while RMSEP is not; it just assesses the degree to which the fitted curve is close/far from the individual datapoints.  Hence, RMSEP is slightly preferable to R2 as a goodness of fit.  So, if the two differ, go with RMSEP.  Happy to discuss more the next time we chat.
   2. RMSEP is always in the same units as the original variable, and is akin to a standard deviation; it is an estimate of the predictive ability of the model. (RMSEP: root mean standard error of prediction).  So if you predicted 15C for one sample and 17C for another sample, but your RMSEP was 5C, you wouldn't be confident that the 2C difference was meaningful given the high RMSEP.
   3. R2, conversely, is proportional variance explained by your model and always ranges from 0 (terrible model; explains none of the variance in the data) to 1 (perfect model, explains all of the variance in the data).  R2, as a standardized metric, **can** be compared among variables.
3. Add PaleoSig component with a spatial test? See code by Matthias.

Lab 8:

1. Change readings to cut Legendre (keep as supplement), add back Imbrie
2. In presentation (which worked well) add a Tweiten (or other?) example for pollen data and visualization. This was helpful and just had to acknowledge that it was NMDS, not PCA.
3. Typo in Imbrie worksheet – 2 Exercise 36s…
4. Add a broken-stick significance test, based on Borcard? (ADDED TO LECTURE 9(
5. Also note that PCA use with pollen data is a violation of assumptions…

Lab 9: Correspondence Analysis

CUT? Seems not so critical

Lab 10: GAM

1. Add back red and white noise section – couldn’t get past a stupid for-loop bug.
2. Fix GAM code for a better count/% error model

Lab 11: Charcoal: Cut – no R package available

OTHER LABS: Add back a 2-part sequence on Bayes? Boosted Regression Trees? Other…?