

SUMMARIZING FROG AND TOAD COUNT DATA

This set of protocols will take you through all the steps necessary for summarizing the frog and toad data for each NAAMP route that was been assigned to you.

BEFORE YOU START, HERE ARE SOME IMPORTANT DATA REMINDERS, BOTH FOR THIS PROJECT AND FOR YOUR LIFE IN GENERAL.

- Don't leave data entries blank – if they're blank someone (including you) might just think you forgot about them. If a variable comes out as zero, then enter 0. If you can't calculate something for whatever reason, enter NA for "Not Available."
- Be careful and consistent with formatting. Data analysis software can do amazing things, but it can't read your mind (at least not yet). For example, a software package won't know that "N" means the same thing as "no" or even the same as "n". So, please follow the templates exactly and use consistent notation with things like capitalization.
- Never, ever guess! I know this goes counter to all the student skills you've accumulated over your life. But, at least for now, think of yourself as a scientist, not a student. If you have some extra NAs, no problem, we can drop these and still have plenty of data to find out what we want to find out. But incorrect entries can really mess things up as far as the data analysis goes.
- If you have questions, please ask me (marsh@nceas.ucsb.edu) or ask your instructor. This project is challenging and we don't expect anyone to understand everything at the beginning. Asking questions is part of the process.

Start by opening each file in Excel or other spreadsheet software. Within the folder for each route ([LINK](#)), a data sheet (.csv) contains all NAAMP frog and toad data collected for that route since the volunteer survey began.

A sample data sheet will look something like this:

Run ID	Survey Year	Route Number	Run Number	Stop Num	SiteID	Skipped Stop	Noise (Y/T)	Car Count	Acris crep	Anaxyrus	Anaxyrus	Gastrophyr	Hyla chrys	Hyla ciner	Lithobate	Lithobate
8915	2006	520102	1	1	13572	N	N	1								
9464	2006	520102	2	1	13572	N	N	0	2	2						
9465	2006	520102	3	1	13572	N	Y	0	1							
10962	2007	520102	1	1	13572	N	N	1		2						
11088	2007	520102	2	1	13572	N	N	0	3							2
11319	2007	520102	3	1	13572	N	N	0	1							
12590	2008	520102	2	1	13572	N	Y	2	3							
12870	2008	520102	3	1	13572	N	N	0	3							
14039	2009	520102	1	1	13572	N	N	1								
14177	2009	520102	2	1	13572	N	N	1	3	1						
14706	2009	520102	3	1	13572	N	N	0	2							
17663	2011	520102	1	1	13572	N	N	0								
17887	2011	520102	2	1	13572	N	N	0	3	3						
8915	2006	520102	1	4	13575	N	N	1								
9464	2006	520102	2	4	13575	N	N	1		2						
9465	2006	520102	3	4	13575	N	N	0	3							
10962	2007	520102	1	4	13575	N	N	0								
11088	2007	520102	2	4	13575	N	N	0	3							
11319	2007	520102	3	4	13575	N	N	0	3							
12590	2008	520102	2	4	13575	N	N	0	3	1						2
12870	2008	520102	3	4	13575	N	N	1	3							
14039	2009	520102	1	4	13575	N	N	0								
14177	2009	520102	2	4	13575	N	N	0	3	1						2
14706	2009	520102	3	4	13575	N	N	1	3							
17663	2011	520102	1	4	13575	N	N	0								
17887	2011	520102	2	4	13575	N	N	1	3	3						
8915	2006	520102	1	7	13578	N	N	0								
9464	2006	520102	2	7	13578	N	N	0		1						2
9465	2006	520102	3	7	13578	N	N	0	1							1
10962	2007	520102	1	7	13578	N	N	0		3						2
11088	2007	520102	2	7	13578	N	N	0		1						1
11319	2007	520102	3	7	13578	N	N	0	3							1
12590	2008	520102	2	7	13578	N	N	0	1	1						1
12870	2008	520102	3	7	13578	N	N	0		1						1
14039	2009	520102	1	7	13578	N	N	0								
14177	2009	520102	2	7	13578	N	N	1	3	1						2
14706	2009	520102	3	7	13578	N	N	0	1							
17663	2011	520102	1	7	13578	N	N	0								
17887	2011	520102	2	7	13578	N	N	0		3						
8915	2006	520102	1	10	13581	N	N	0								
9464	2006	520102	2	10	13581	N	N	0		1						1
9465	2006	520102	3	10	13581	N	N	0	3							1
10962	2007	520102	1	10	13581	N	N	0		3						3
11088	2007	520102	2	10	13581	N	N	1	3	1						

I. FIRST, WE'LL TAKE YOU THROUGH A SUMMARY OF EACH OF THESE VARIABLES:

A) **Run ID.** A run is a survey of the 10 stops within a route. *Each time an individual does a NAAMP survey, this survey is assigned a unique Run ID.* Typically, there are 3 runs for each route in each year, corresponding to three different time periods in which amphibians might be calling. Sometimes there are fewer than 3 runs when an individual can't do all three surveys. On occasion, an individual surveys his/her route twice or even three times within each survey period (or two different people survey the same route). These surveys will have the same route number and the same run number, but they will still have different Run IDs.

B) **Survey Year.** This is just the year in which the survey took place. We won't be using the year data, but it's useful to keep track of where you are in the datasheet.

C) **Route number.** Each route has a route number. The first two digits of this number are unique to whichever state contains the route.

D) **Run number.** Again, this is just the number of the survey (1,2, or 3) within each year. These numbers correspond to the time period for the survey, so if a survey is missing it could be 1,2 or 3 (i.e. you could have a run 2 and 3 in a given year, but no run 1).

E) **Stop number.** There are 10 stops along each route at which an individual listens for frogs. However, we have included only stops 1, 4, 7, and 10 for each route. The reason for this is that stops are a minimum of 0.5 km apart. Since you will be analyzing the landscape within 1 km of each stop, adjacent stops will include many of the same landscape features. By analyzing stops that are at least 2 km apart (e.g. 1, 4, 7, and 10), we ensure that stops within a route are both ecologically and statistically independent from one another. *Every once in a while, a stop may be missing from a route. These stops were surveyed, but we do not have latitude/longitude information for them that can be used to study the landscape around them. If you're missing a stop, show your instructor.*

F) **SiteID.** Each stop has a unique SiteID. We will use these IDs to relate the frog data at each site to the landscape data at each site. In theory, it's possible that SiteIDs would differ for a given stop if that stop had been moved at some point. If you find one of these, please let your instructor know.

G) **Skipped Stop.** Occasionally a stop is skipped for one reason or another, even though most of the route is completed. In these cases, surveyors will enter a "Y" to indicate that the stop was skipped and there should be no amphibian counts for that row. *When you count up the number of surveys for each stop (see below), make sure to exclude any stops that were skipped.*

H) **Noise (Y/N).** Sometimes during a survey there will be noise that could interfere with one's ability to hear frogs (e.g. traffic, people talking, etc.). Surveyors note when such noise is present, and we'll likely test analyze our data both including and excluding sites that tend to be noisy.

I) **Car Count.** In many (but not all) of the surveys, surveyors recorded the number of cars that went by during the survey of each stop. This may end up being important data, as it provides an assessment of nocturnal traffic at the stop that could have negative consequences for frogs and toads (i.e. they could get squished).

J, K, L etc.) Call intensity data for each species. From here on out, the datasheet lists each species and then a blank (not detected) or a 1,2, or 3 for the intensity of calling. Here's what the numbers mean:

- 1: Individuals can be counted; there is space between calls
- 2: Calls of individuals can be distinguished but there is some overlapping of calls
- 3: Full chorus, calls are constant, continuous and overlapping

You can think of these as corresponding to "a few", "some", and "many" if that makes it easier to remember. In some cases, we'll consider the call intensity, but in others we'll just look at whether or not a given species is detected.

II. NOW WE'LL TAKE YOU THROUGH WHAT YOU NEED TO CALCULATE FROM THESE DATA.

You will need to summarize the frog and toad data for each stop (1,4,7 and 10). Following are all the variables you need to summarize for each stop. The data template is provided at [\(LINK TO TEMPLATE\)](#)

The first four columns are the identifying information for each stop. These columns are: A) Route number, B) Stop number, C) SiteID, and D) EnteredBy (your initials, so that we know who to ask if we have questions). Then come the columns with actual calculations.

E) TOTAL SURVEYS (TOT_SURV). This is just the number of surveys (i.e. rows) that correspond to each stop/ SiteID. *This total should exclude any surveys that note that the stop was skipped (i.e. Skipped Stop = "Y").*

F) SURVEYS WITH NOISE (NOISE_NUM). This is the number of surveys for each stop in which Noise was noted (Noise = "Y"). You'll include surveys with noise in all your calculations, but when we're all done, we're planning to do an analysis with that excludes stops with frequent noise.

G) CAR_MEAN. This is the mean (average) of the car count data for each stop. *Make sure to exclude empty cells from any skipped stops.*

H) RICH1. This is species richness for each stop (i.e. number of species present at each stop) with a species considered present if it was recorded in at least one survey. This is a high end estimate of species richness since a single frog might have just been passing through or a species might have been recorded once in error.

I) RICH3. This is species richness for each stop (i.e. number of species present at each stop) but with a species only counted as present if it was recorded in three separate surveys. This is a low end estimate of species richness but may be less prone to error since it only includes species heard regularly at a site.

J-L), M-O), P-R) etc...

For each species, we'll record three summary statistics, so you'll have three columns for each species. First, we'll record whether a species is present at a stop based on whether it was ever heard at that stop. Presence will be recorded as a "1" and absence as a "0". For *Acris crepitans* (for example), this column would be called ACCR1; for *Lithobates clamitans* this column would be called LICL1.

Second, we'll record whether a species is present at a stop based on them having been heard in at least three surveys of a stop. Again, presence will be recorded as a "1" and absence as a "0". These columns will be labeled with the first two letters of the genus name, then the first two letters of the species name, then 3 (e.g. ACCR3, LICL3).

Finally, we'll record the mean calling index for each species across for each stop. For example, if *Acris crepitans* was recorded in 8 surveys as: 1, 3, 0, 0, 0, 2, 3, 1, then the mean calling index (ANAM_MEAN) = 10 (the sum of the calling index values) divided by 8 (the number of surveys of the stop) = 1.25

Hyla gratiosa
0
0
0
0
0
0
0
1
0
0
0
2

So, for the data column to the left, the three columns would be called “HYGR1”, “HYGR3”, and “HYGR_MEAN”. The values entered into each of these would be “1” (since the species was heard in at least one survey), “0” (since the species was not heard in at least 3 surveys), and “0.25” (since the mean calling index is 3/12).

A couple of point to remember when putting these data together:

- Include columns for all species listed in your raw data, even if they aren’t recorded in any of your surveys. Although these species aren’t in your surveys, they will be in other surveys, and all the datasheets need to be the same to be combined.
- For each of these three summary statistics, remember to exclude any surveys where a “skipped stop” is indicated.

Your final data sheet should look something like this:

Route Number	Stop Number	SiteID	Entered_by	TOT_SURV	NOISE_NUM	CAR_MEAN	RICH_1	RICH_3	ANAM_1	ANAM_3	ANAM_MEAN	HYCH_1	HYCH_3	HYCH_MEAN	HYVE_1	HYVE_3	HYVE_MEAN
720107	1	20372 NFG		36	21	0.579	6	3	1	1	0.278	1	0	0.056	1	0	
720107	4	20375 NFG		36	6	3.188	8	7	1	1	0.278	1	1	0.361	1	1	
720107	7	20378 NFG		36	13	2.345	3	2	0	0	0	1	1	0.139	1	0	
720107	10	20381 NFG		36	8	1.633	8	7	1	1	0.167	1	1	0.194	1	1	

Congratulations – you’ve now summarized the frog and toad calling data. Now move on to the protocols for the landscape data.