To: Dave Garhselis

From: John Fieberg

Date: 7/3/2015

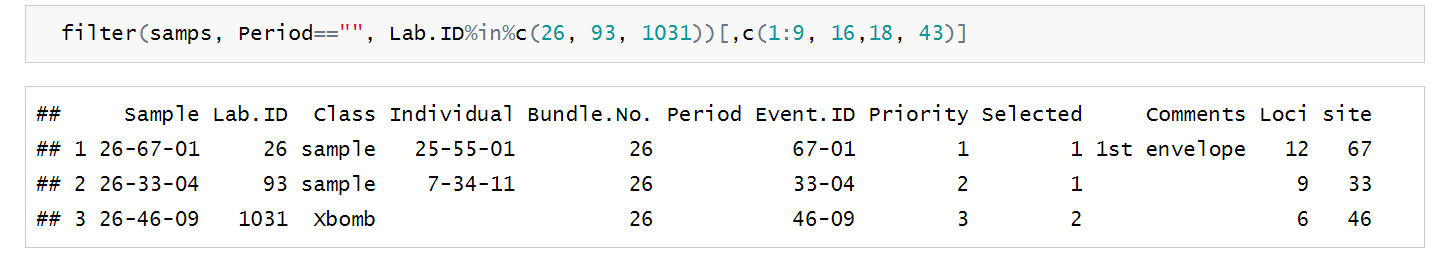
RE: Bear Hair Snare Data Questions

This memo accompanies two html files containing R code and associated output, and also a .csv file.

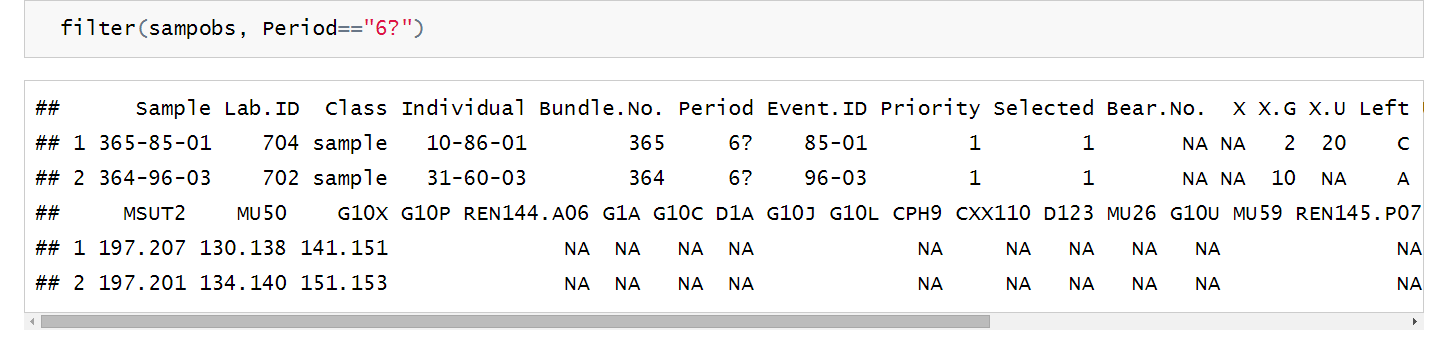
* Data Cleaning/Data Development.htlm: data quality checks, raises some questions regarding missing values, etc (summarized below).
* Data Summary.html: looks at the capture histories, summarized in different ways
* MissingPeriod.csv: data set containing all observations for which Period was missing.

Below, I list a few initial issues with the data.

*Observations that are missing Period:* Several observations are missing Period (the time of the event). I have not been able to figure out how to determine this field from the other columns. The MissingPeriod.csv data set contains all of these observations. A few of these are from collared bears (meant to be dropped?). Several are associated with samples that were not sent to the lab. These latter observations may not be problematic, provided that raw summaries of the number of observations per (site x period) combination are correct and available. The 3 below are likely the most problematic:



*Observations with period = 6?* Two observations had “6?” for period. These should probably be verified.



Lastly, we’ll of course need the locations of the sites before we can start using SECR to estimate abundance. At that point, we will have a few different things we can try:

1. Estimate abundance using the full data set, with the counts at each site x period combination (assumed to be Poisson distributed).
2. Estimate abundance using the full data set, with presence/absence at each site x period combination (rather than the counts).
3. Estimates obtained by subsampling the data in multiple ways, repeated many times (e.g., sampling the pool of observations randomly vs. sampling to maximize the number of site x period combinations).

At the end of the day, I went back and found the Research Summary that Karen and you wrote summarizing the data in 2013. I tried duplicating these results with the attached data. I was generally able to match things up.

* Report: “Of 1113 samples that were analyzed, 14 appeared to be mixtures of >1 bear and 80 failed to amplify.” (I found 14 with Class Xmixed and 80 with Xbomb), matching the numbers here
* Report: “Genotyping identified 43 different individuals: 26 males and 17 females.” Checks out.
* Report: “Individual bears were detected up to 132 times each and up to 32 times in a single sampling session.” Checks out.
* I was also able to reproduce Table 1 and Table 2

Things that don’t appear to line up perfectly (at least yet):

* Thus 1019 samples (92%) were successfully genotyped; these were from 96 different sites and 333 site-sessions. I found 100 unique sites had samples (not 96), and 346 unique site x session combinations (not 333).