**Brown Pelican Data Management Protocols – April 2018**

Care should be taken to ensure that data are easily accessible and interpretable to all collaborators at all stages of the project. Please make efforts to follow these guidelines whenever possible.

*In the Field*

* To ensure that reliable and unbiased data are collected, make sure that appropriate field gear is used consistently in all bouts of data collection. Checklists of required equipment should be created for each project, reviewed for completion by at least one collaborator, and consulted before departing for field sites.
* Pre-print all data sheets before leaving for the field, as printers are not reliably available at field sites.
* Whenever available, use Rite-in-the-Rain field paper to record data, store in covered clipboards until back from the field site. Avoid bringing recorded data back into the field on subsequent days. Digitize data collected in the field as quickly as possible to minimize risk of data loss due to lost documents. Copy or transfer electronic data from associated devices.
* All data sheet headings should include the type of data being recorded (i.e. Brown Pelican Nest Success), date (Gregorian and Julian is preferred for later convenience), location, and identities of both data collector (i.e. who is taking/dictating data) and recorder (i.e. who is writing data down). Data sheets should also be signed by the person who digitized data, and again when proofed.

*Data Formatting*

* Raw data should be organized so that headings are easily understood. All file names should include authors’ last names, descriptive info, no spaces (in case whitespace creates issues with software to which the data might be uploaded) and date of creation. Day-month-year format is preferred (most common internationally).
	+ Example: “Geary\_RaccoonNestSuccess\_23apr2018.csv”
* Create readme files with file names clearly associated with the main data if detailed explanations of file contents are necessary. Keep both files in the same directory.
* When statistical methods are settled upon, copies of data should also be saved that are appropriately formatted for analysis (i.e. for R, save a csv with no spaces, no starting column names that start with numbers, etc.)
* Make sure that all analytical programming scripts are clearly annotated and written to ensure reproducible results/output. Good rules of thumb (general points extend beyond R) can be found here: swcarpentry.github.io/r-novice-inflammation/06-best-practices-R/
* To protect against weather- or theft-related data loss, SD cards containing camera footage should never be re-deployed in the field without removal of all previously collected data. Video or image files should be labeled with their colony locations, dates, and associated cameras (e.g. QueenBess\_21-23May2014\_camera1.avi) and stored in multiple locations for ease of access (see next section).

*Version Control and Backup*

* If important changes are made to shared documents, separate copies should be created with updated date and author info (e.g. Geary\_doc\_LebergEdits\_24apr2018.docx).
* Backup files should be saved in at least three locations to prevent data loss. These files should be updated at a reasonable interval (i.e. every 2 weeks). These locations include:
	+ Personal computer (day-to-day work)
	+ Leberg lab cloud storage (ULL Sharepoint)
	+ Leberg lab hard drive (offsite storage ensures data retention in event of local disasters)
* For additional security, copies can also be saved on:
	+ Personal external hard drives (for immediate recovery in case of personal computer failure)
	+ Third-party online databases (e.g. Movebank for telemetry data)