

Camtrap data harmonization

An examination of replicability across camtrap datasets.

Goal

Consistent, tidy data that are readable directly from data repository into R.

Principles

1. A total of three files. metadata.csv, sites.csv, and observations.csv
2. metadata.csv is a simple file with one row per idea. The purpose of experiment or hypothesis, total number of cameras deployed, number of sites, number of days, how many times cameras were moved, the blocking variables, and other design elements such as camera models. Some of these attributes can be engineered from the observations data or sites data but to make it easy and not to have these elements be joined to every observation, list there here. title, dataset authors, purpose, n_days, n_cams, n_cam_moves, n_sites, join_by, and variables are likely key and replicable elements to mention for every dataset.
3. site.csv is a file the will be left_joined as needed to observations. Use the same camera code such as cam_ID in this file and in the observations.csv dataframe to enable subsequent joins. This file should have the cam_ID, site_name, site_ID, lat, long, shrub x,y,z if relevant, and microsite. If 'rep' is the join element instead of cam_ID, ensure each is unique and reasonable to join.
4. observations.csv are tidy but not compiled or summed data. Each row is an observation or picture from a camera. The first vector should be the key to join to sites.csv such as cam_ID or site_ID. We should all agree on common notation and nomenclature. If a photo took a picture with no animal, code as 0. Use lat and long as vector names for location in degrees decimals. use lower case and snake_case for all vector names.
5. Publish data to knb data. Data are different from papers. List all main participants in collection and design of data. Publish when tidied. List ORCID of each author. Publish as single package - three files (metadata.csv, sites.csv, and observations.csv). Same for each experiment. For secondary authors, just list their names and ORCID - no other deatils. After you hit publish, check it in KNB, but there is one more step, hit publish with DOI.

Carrizo 2019 Test of principles with data from Carrizo, lead by MZ, in 2019.

```
library(tidyverse)

#metadata
metadata <- read_csv("data/Carrizo_2019/metadata.csv")
knitr::kable(metadata)
```

element	metadata
title	The effect of shrub density on animal abundance
dataset_authors	Zuliani, Ghazian, Lortie
purpose	Deployed cameras at different shrub densities to capture animals
n_days	27
n_cams	16
n_cam_moves	0
n_sites	2
join_by	site
variables	Low, medium, and high shrub density plots at shrub and open microsite

```
#sites
sites <- read_csv("data/Carrizo_2019/sites.csv")
knitr::kable(sites)
```

region	ID	site	shrub_density	lat	long	n_shrubs	shrub_ID	shrub_lat	shrub_long	x	y	z	microsite	rep
Carrizo	1	4	high	35.11985	-119.6283	12	1	35.11982	-119.6285	300	240	145	shrub	1
Carrizo	2	4	high	35.11985	-119.6283	12	2	35.11984	-119.6284	434	260	143	shrub	2
Carrizo	3	4	high	35.11985	-119.6283	12	3	35.11988	-119.6285	326	228	156	shrub	3
Carrizo	4	4	high	35.11985	-119.6283	12	4	35.11992	-119.6284	224	158	138	shrub	4
Carrizo	5	4	high	35.11985	-119.6283	12	5	35.11992	-119.6284	222	170	135	shrub	5
Carrizo	6	4	high	35.11985	-119.6283	12	6	35.11994	-119.6284	196	173	151	shrub	6
Carrizo	7	4	high	35.11985	-119.6283	12	7	35.11991	-119.6283	182	157	153	shrub	7
Carrizo	8	4	high	35.11985	-119.6283	12	8	35.11990	-119.6283	185	184	118	shrub	8
Carrizo	9	4	high	35.11985	-119.6283	12	9	35.11993	-119.6283	174	147	102	shrub	9
Carrizo	10	4	high	35.11985	-119.6283	12	10	35.11983	-119.6283	277	223	129	shrub	10
Carrizo	11	4	high	35.11985	-119.6283	12	11	35.11982	-119.6283	281	205	121	shrub	11
Carrizo	12	4	high	35.11985	-119.6283	12	12	35.11980	-119.6283	271	241	125	shrub	12
Carrizo	13	4	medium	35.11920	-119.6295	6	1	35.11910	-119.6284	262	256	148	shrub	13
Carrizo	14	4	medium	35.11920	-119.6295	6	2	35.11917	-119.6284	274	254	159	shrub	14
Carrizo	15	4	medium	35.11920	-119.6295	6	3	35.11923	-119.6285	493	398	144	shrub	15
Carrizo	16	4	medium	35.11920	-119.6295	6	4	35.11929	-119.6284	221	281	131	shrub	16
Carrizo	17	4	medium	35.11920	-119.6295	6	5	35.11929	-119.6283	295	255	123	shrub	17
Carrizo	18	4	medium	35.11920	-119.6295	6	6	35.11926	-119.6282	338	310	148	shrub	18
Carrizo	19	4	low	35.11889	-119.6292	2	1	35.11881	-119.6292	230	200	132	shrub	19
Carrizo	20	4	low	35.11889	-119.6292	2	2	35.11880	-119.6292	269	164	112	shrub	20
Carrizo	21	4	none	35.11845	-119.6286	0	NA	35.11845	-119.6286	NA	NA	NA	open	21
Carrizo	22	3	high	35.16175	-119.6719	11	1	35.16172	-119.6720	230	121	107	shrub	22
Carrizo	23	3	high	35.16175	-119.6719	11	2	35.16174	-119.6720	280	173	117	shrub	23
Carrizo	24	3	high	35.16175	-119.6719	11	3	35.16174	-119.6719	178	165	99	shrub	24

region	ID	site	shrub_density	lat	long	n_shrubs	shrub_ID	shrub_lat	shrub_long	x	y	z	microsite	rep
Carrizo	25	3	high	35.16175	-119.6719	11	4	35.16172	-119.6720	121	108	123	shrub	25
Carrizo	26	3	high	35.16175	-119.6719	11	5	35.16179	-119.6719	239	197	109	shrub	26
Carrizo	27	3	high	35.16175	-119.6719	11	6	35.16184	-119.6719	239	197	111	shrub	27
Carrizo	28	3	high	35.16175	-119.6719	11	7	35.16175	-119.6719	180	173	83	shrub	28
Carrizo	29	3	high	35.16175	-119.6719	11	8	35.16173	-119.6719	240	172	98	shrub	29
Carrizo	30	3	high	35.16175	-119.6719	11	9	35.16171	-119.6719	219	189	107	shrub	30
Carrizo	31	3	high	35.16175	-119.6719	11	10	35.16171	-119.6719	204	203	149	shrub	31
Carrizo	32	3	high	35.16175	-119.6719	11	11	35.16168	-119.6730	235	172	124	shrub	32
Carrizo	33	3	medium	35.16320	-119.6741	7	1	35.16312	-119.6741	344	249	121	shrub	33
Carrizo	34	3	medium	35.16320	-119.6741	7	2	35.16330	-119.6740	187	162	93	shrub	34
Carrizo	35	3	medium	35.16320	-119.6741	7	3	35.16331	-119.6741	358	251	133	shrub	35
Carrizo	36	3	medium	35.16320	-119.6741	7	4	35.16329	-119.6740	257	206	94	shrub	36
Carrizo	37	3	medium	35.16320	-119.6741	7	5	35.16314	-119.6740	208	91	96	shrub	37
Carrizo	38	3	medium	35.16320	-119.6741	7	6	35.16316	-119.6740	175	165	117	shrub	38
Carrizo	39	3	low	35.16252	-119.6729	3	1	35.16223	-119.6730	353	289	145	shrub	39
Carrizo	40	3	low	35.16252	-119.6729	3	2	35.16230	-119.6730	234	154	151	shrub	40
Carrizo	41	3	low	35.16252	-119.6729	3	3	35.16233	-119.6730	243	217	147	shrub	41
Carrizo	42	3	none	35.16256	-119.6743	0	NA	35.16256	-119.6743	NA	NA	NA	open	42

```
#observations
```

```
observations <- read_csv("data/Carrizo_2019/observations.csv")
knitr::kable(head(observations))
```

year	ID	recorder	file	region	method	camera	date	time	plot	density	shrub	level	ID	dominant	BTU	actual	intensity	baseline	observations	notes	microsite	
2019	55	Mar	10	MC	000	camera	17	May	1	1	1	1	1	3	high	1	1	1	Unidentified	standing	might be a k-rat	shrub

year	ID	recorder	file	region	site	tool	camera	date	week	depth	plot	density	shrub	valley	ID	DTU	actual	initial	blat	obs	notes	microsite
2019	555	Mar	IMGC_0008	to camera	17	1	1	1	1	1	3	high	1	8	1	Unide	18.5	100	moving	NA	shrub	
				trap May															not know what it is			
2019	555	Mar	IMGC_0012	to camera	18	2	1	2	1	1	3	high	1	13	1	Unide	2.1	100	unk	something is blocking camera	NA	shrub
				trap May																		
2019	555	Mar	IMGC_0016	to camera	18	2	1	2	1	1	3	high	1	16	1	Unide	5.5	100	standing near camera	either a kit fox or a coyote	shrub	
				trap May																		
2019	555	Mar	IMGC_0018	to camera	18	2	1	2	1	1	3	high	2	58	1	Unide	0.2	100	unk	not acting with camera	may be a coyote or kit fox	shrub
				trap May																		
2019	555	Mar	IMGC_0020	to camera	18	2	1	2	1	1	3	high	2	59	1	Unide	0.3	100	unk	some hit as previous	NA	shrub
				trap May																		