INTRODUCTIONS

1. Name
2. In what city were you born?
3. What you consider your "hometown"?
4. Program and Degree?
5. Hobbies or what you do to relax.

DISCUSS

1. Motivation for taking the class
2. Concerns about this class (in particular) and this semester (in general)?
3. Big Question on the next slide

What are "DATA?"

INTRODUCTIONS

1. Name
2. In what city were you born?
3. What you consider your "hometown"?
4. Program and Degree?
5. Hobbies or what you do to relax.

DISCUSS

1. Motivation for taking the class
2. Concerns about this class (in particular) and this semester (in general)?
3. What are "Data'?
4. Record all this on the slides to report back to the group

Record all this on the flip charts and report back to the group

What are "DATA?"

FEDERAL REGISTER

Vol. 78 Thursday,
No. 248 December 26, 2013

Part III

Office of Management and Budget

2 CFR Chapter I, Chapter II, Part 200, et al.
Uniform Administrative Requirements, Cost Principles, and Audit Requirements for Federal Awards; Final Rule

OFFICE OF MANAGEMENT AND BUDGET

2 CFR Chapter I, and Chapter II, Parts 200, 215, 220, 225, and 230

Uniform Administrative Requirements, Cost Principles, and Audit Requirements for Federal Awards

AGENCY: Executive Office of the President, Office of Management and Budget (OMB).

ACTION: Final guidance.

SUMMARY: To deliver on the promise of a 21st-Century government that is more efficient, effective and transparent, the Office of Management and Budget (OMB) is streamlining the Federal government's guidance on Administrative Requirements, Cost Principles, and Audit Requirements for Federal awards. These modifications are a key component of a larger Federal effort to more effectively focus Federal resources on improving performance and outcomes while ensuring the financial integrity of taxpayer dollars in partnership with non-Federal stakeholders. This guidance provides a

§ 200.315 Intangible property.

(a) Title to intangible property (see § 200.59 Intangible property) acquired under a Federal award vests upon acquisition in the non-Federal entity. The non-Federal entity must use that property for the originally-authorized purpose, and must not encumber the property without approval of the Federal awarding agency. When no

(d) The Federal government has the right to:

(1) Obtain, reproduce, publish, or otherwise use the data produced under a Federal award; and

(2) Authorize others to receive, reproduce, publish, or otherwise use such data for Federal purposes.

(e) Freedom of Information Act (FOIA).

(1) In addition, in response to a Freedom of Information Act (FOIA) request for research data relating to published research findings produced under a Federal award that were used by the Federal government in developing an agency action that has the force and effect of law, the Federal awarding agency must request, and the non-Federal entity must provide, within a reasonable time, the research data so that they can be made available to the public through the procedures established under the FOIA. If the

(3) Research data means the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This "recorded" material excludes physical objects (e.g., laboratory samples). Research data also do not include:

(i) Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and

(ii) Personnel and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.

OECD

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

OECD Principles and Guidelines for Access to Research Data from Public Funding

In the context of these *Principles and Guidelines*, "research data" are defined as factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated.

This term does not cover the following: laboratory notebooks, preliminary analyses, and drafts of scientific papers, plans for future research, peer reviews, or personal communications with colleagues or physical objects (*e.g.* laboratory samples, strains of bacteria and test animals such as mice). Access to all of these products or outcomes of research is governed by different considerations than those dealt with here.

"Anything you perform analysis upon."

Kristin Briney, p. 6

With that in mind….

1. Identify different kinds of data collected in different disciplines

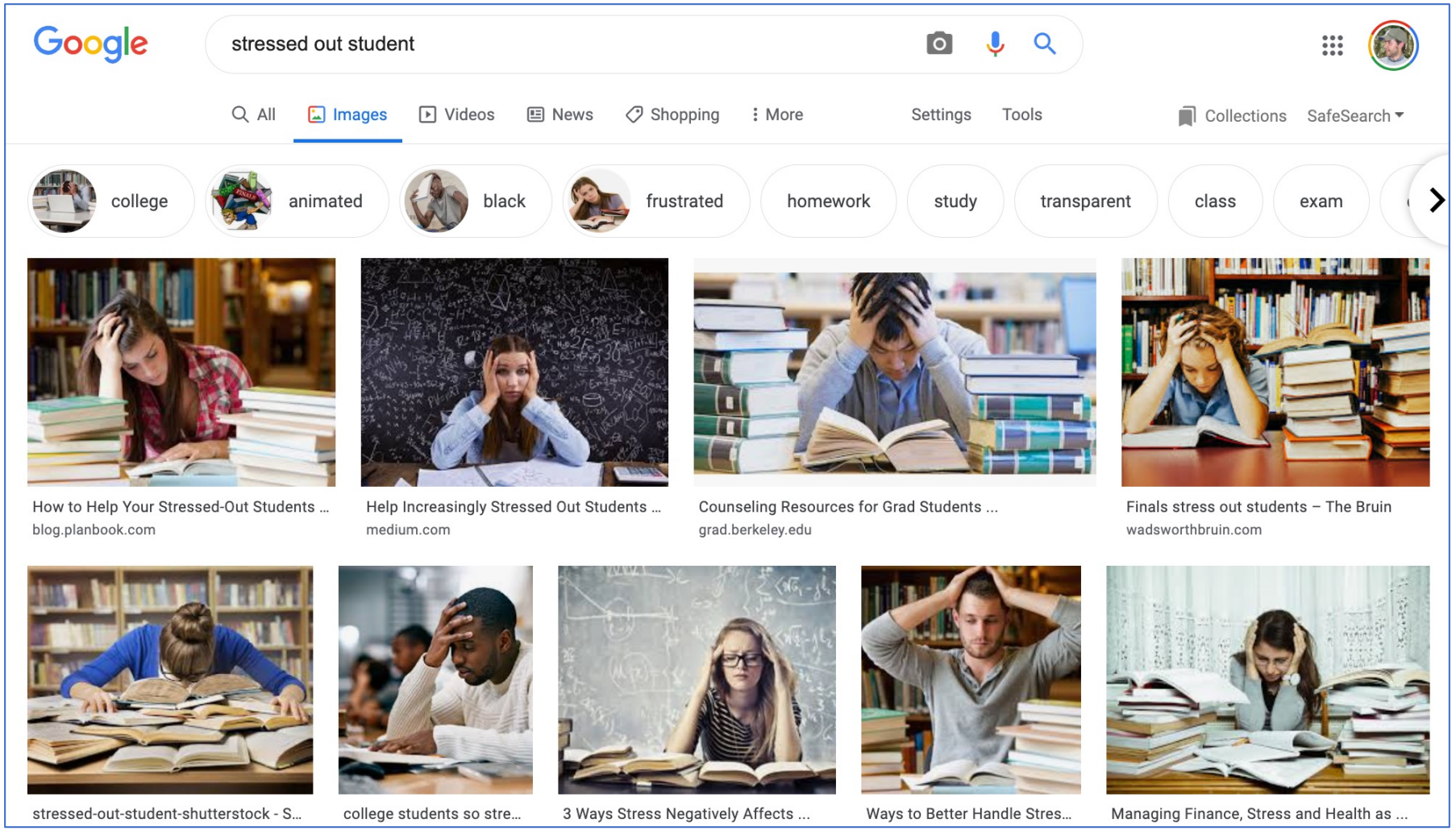
2. How are these data collected & recorded?

But first….

WHAT IS *MY* MOTIVATION?

short answer: data management might be **the** most important skill you learn at UF

Longer answer: Reason 1



CTV NEWS NEWS ▾ VIDEO ▾ WEATHER

PhD student offering $5,000 reward after car thief steals all his research

CTV Montreal

Published Wednesday, November 29, 2017 10:43PM EST

"The thief stole a backpack containing his laptop, a notebook, and a thumb drive with all his backup material... With all his work gone, he will likely have to head back to the lab and redo experiments, delaying graduation by months."

https://tinyurl.com/yxbwo7qw

Final year student has laptop stolen in Pret, with entire dissertation due in two weeks

Her dissertation research was also in the laptop case

Saffron Sibthorp
14 mins ·

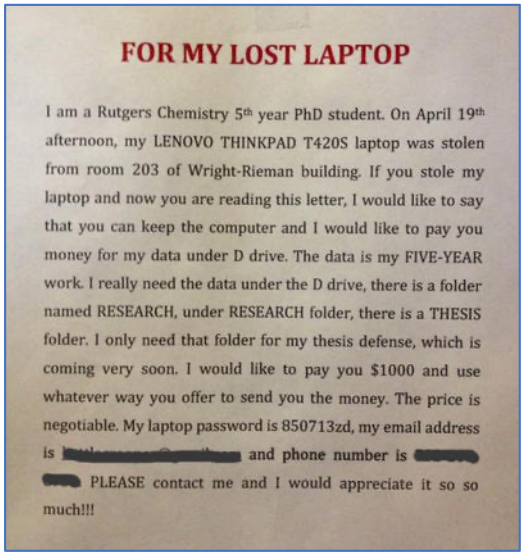
PLEASE SHARE

I had my backpack stolen today from Pret in Oxford Street, containing my laptop with all my university work on and my final year dissertation which is due in two weeks time! Literally the worst thing that could happen, I'm really desperate for my laptop back, without it I could fail my degree with nothing to submit. Also had all of my research inside my laptop case. Know I'm probably wishing for the impossible but please if you hear anything or are offered a stolen Apple MacBook Pro, please let me know! I'm waiting for CCTV images from Pret, as the police can't do anything without any evidence, and I'll try and upload images of the suspect ASAP. Would really appreciate any help!

Incident happened around 6pm today (29/01/18), at Pret A Manger, 298 Regent Street, W1B 3AP.

https://tinyurl.com/y5amkf6m

Ouch!: Thief Steals Laptop Containing 5 Years of Ph.D. Student's Research Data



FOR MY LOST LAPTOP

I am a Rutgers Chemistry 5th year PhD student. On April 19th afternoon, my LENOVO THINKPAD T420S laptop was stolen from room 203 of Wright-Rieman building. If you stole my laptop and now you are reading this letter, I would like to say that you can keep the computer and I would like to pay you money for my data under D drive. The data is my FIVE-YEAR work. I really need the data under the D drive, there is a folder named RESEARCH, under RESEARCH folder, there is a THESIS folder. I only need that folder for my thesis defense, which is coming very soon. I would like to pay you $1000 and use whatever way you offer to send you the money. The price is negotiable. My laptop password is 850713zd, my email address is [redacted] and phone number is [redacted] PLEASE contact me and I would appreciate it so so much!!!

https://tinyurl.com/yyosxmeb

How long do hard drives last?

How long do hard drives last?

PAUL BISCHOFF - TECH JOURNALIST, PRIVACY ADVOCATE AND VPN EXPERT
@pabischoff August 27, 2017

Hard drive failure is unpredictable, so answering the question of how long hard drives last will inherently comes with a *lot* of caveats.

Short answer: That being said, if you just want a quick rule of thumb for how long you can expect the hard drive in your laptop should last, we'd say you should be prepared for disk failure after three years of use.

Long answer: A handful of studies on the lifespan of hard drives might give you some clearer indication, but they still aren't very helpful. Many of the hard drives tested do not fail at all. These drives are also kept in controlled environments and don't undergo the same conditions as, say, your laptop drive.

Factor in the following questions:

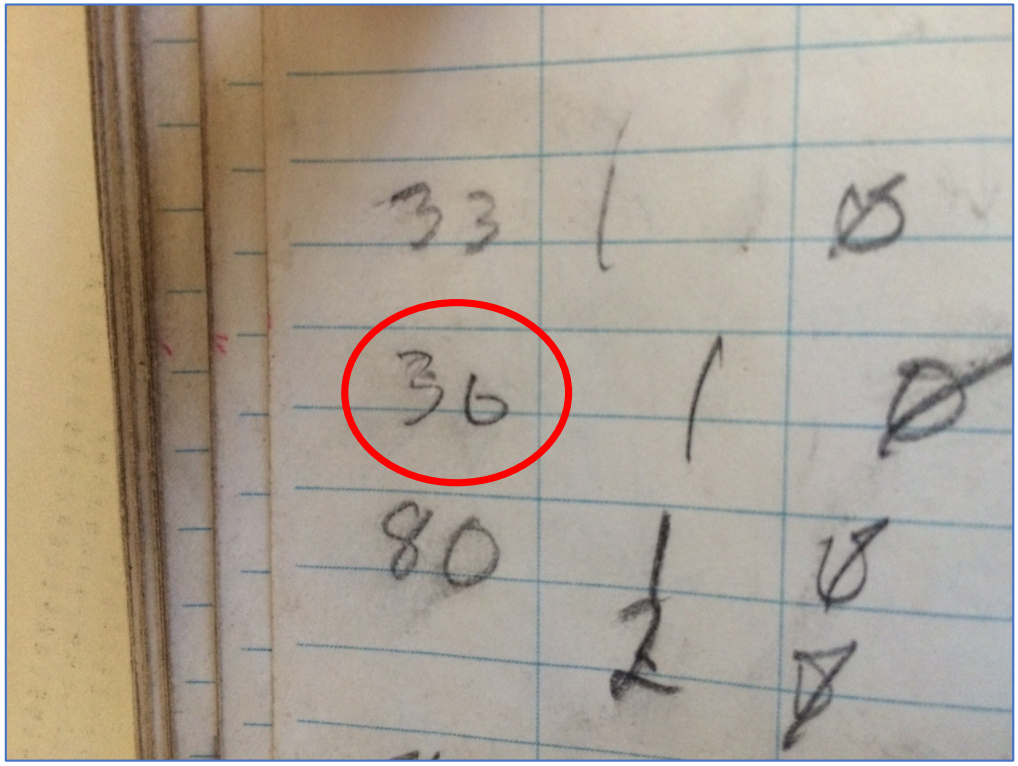
- What brand is the hard drive?
- What do you use it for? Running applications, viewing media, or storage?
- How often do you use your computer?
- Is it frequently shaken, vibrated, or bumped?
- How hot does your computer get?

What causes hard drive failure?

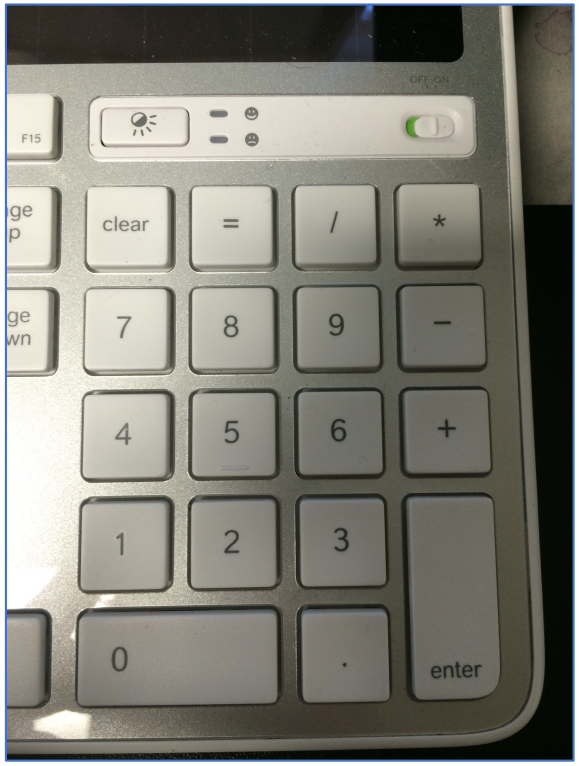
Answer: Factory defects & vibration (mostly)

Photo: Zane Selvans (CC BY-NC-SA 2.0)

Uhhh….

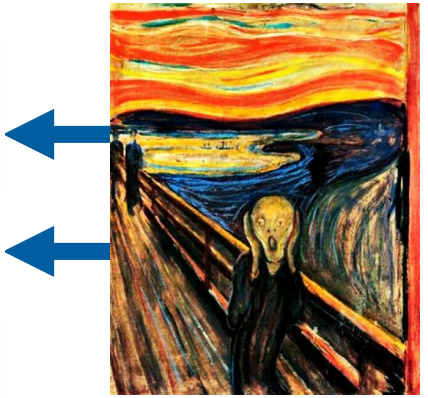


The 6/3 Problem



UF Graduate School
UNIVERSITY of FLORIDA

Deadlines: Electronic Thesis and Dissertation (ETD)	Spring 2023	Summer 2023
Classes Start	Jan 9	May 15
Degree Application	Feb 3	July 6*
Doctoral Dissertation Submission	Feb 8	June 16*
Master's Thesis Submission	March 3	July 14
Final Submission (all ETD students)	April 7	July 28
Final Clearance (Approval)	April 26	Aug 11
Clear Prior (to the upcoming term)	May 12	Aug 22



Longer answer: Reason 2

Animal personality researcher Jonathan Pruitt is under fire about his data on social spiders.

Spider biologist denies suspicions of widespread data fraud in his animal personality research

By **Elizabeth Pennisi** | Jan. 31, 2020 , 6:35 PM

doi:10.1126/science.abb1258

Science

Microplastics killing fish before they reach reproductive age, study finds

Tiny particles of plastic litter in oceans causing deaths, stunted growth and altering behaviour of some fish that feed on them, research shows

▲ A pike (*Esox lucius*) feeds on perch that have ingested microplastic particles. Photograph: Oona Lönnstedt/Science

Search US edition The Guardian

LETTERS

Editorial expression of concern

Jeremy Berg

+ See all authors and affiliations

Science 09 Dec 2016:
Vol. 354, Issue 6317, pp. 1242
DOI: 10.1126/science.aah6990

Article Info & Metrics eLetters PDF

In the 3 June issue, *Science* published the Report "Environmentally relevant concentrations of microplastic particles influence larval fish ecology" by Oona M. Lönnstedt and Peter Eklöv (*1*). The authors have notified *Science* of the theft of the computer on which the raw data for the paper were stored. These data were not backed up on any other device nor deposited in an appropriate repository. *Science* is publishing this Editorial Expression of Concern to alert our readers to the fact that no further data can be made available, beyond those already presented in the paper and its supplement, to enable readers to understand, assess, reproduce, or extend the conclusions of the paper.

Reference

1. ↵ O. M. Lonnstedt, P. Eklov, *Science* **352**, 1213 (2016). Abstract/FREE Full Text Google Scholar

J. MCDONALD/GETTY

THE HOTTEST YEAR

The release of climate-science e-mails last November ripped apart Phil Jones's life. He's now trying to patch it back together.

BY DAVID ADAM

"I like to think the worst is over, but it's coming up to the first anniversary and it's something I'll always remember at this time of year, when the nights close in. This is the time it happened."

Twelve months ago, Phil Jones was a productive, if not particularly outspoken, climate scientist. That was the way he liked it. Head of the Climatic Research Unit (CRU) at the University of East Anglia (UEA), UK, Jones worked with the Met Office to compile data from weather stations around the world into a monthly series showing global average temperature. He had much on his mind — not least a puzzling drop in North Atlantic sea surface temperatures during the mid-twentieth century that he had recently helped to discover. It was a curious finding, but Jones would soon have bigger things to ponder.

On 19 November 2009, someone released roughly 1,000 e-mail messages and documents stolen from a server at the CRU. Many of them contained Jones's private correspondence, which sometimes showed him in an unflattering light.

He gloated about the death of a prominent climate sceptic, and suggested to colleagues they should delete e-mails to keep sceptics from gaining access to information. Most famously, he boasted that he had used a "trick" to "hide the decline" in a temperature chart.

Very soon, members of the sceptic community had pounced on these messages as evidence that Jones and others had concealed flaws in their temperature data and abused the peer-review system to gag critics of climate researchers. Jones faced a storm of accusations that ranged from scientific misconduct to plans to install an autocratic world government through the spread of false hysteria about

guardian.co.uk

News | Sport | Comment | Culture | Business | Money | Life & style

Environment › Hacked climate science emails

Climategate scientists cleared of manipulating data on global warming

Muir Russell report says scientists did not fudge data, but they should have been more open about their work

Longer answer: Reason 3

The New York Times

By Sheryl Gay Stolberg, Sheila Kaplan and Sarah Mervosh

Published May 22, 2020 Updated June 3, 2020

C.D.C. Test Counting Error Leaves Epidemiologists 'Really Baffled'

The Centers for Disease Control has been lumping together tests for active coronavirus with tests for recovered patients, boosting testing totals but muddying the pandemic's course.

future tense

An Outdated Version of Excel Led the U.K. to Undercount COVID-19 Cases

By WHITNEY TESI OCT 07, 2020 • 3:47 PM

NEWS

Audit finds U.S. anti-terror statistics inflated

Feb. 20, 2007, 10:38 PM EST / Source: The Associated Press

Federal prosecutors counted immigration violations, marriage fraud and drug trafficking among anti-terror cases in the four years after 9/11 even though no evidence linked them to terror activity, a Justice Department audit said Tuesday.

Fine's office took care to say the flawed data appear to be the result of "decentralized and haphazard" methods of collection or disagreement over how the numbers are reported, and do not appear to be intentional.

GLOBAL FOREST WATCH

DATA AND RESEARCH PEOPLE

Geospatial Data Brings Indigenous and Community Lands to the Forefront of Forest Management

Geospatial data raises the profile of forested communities

Transparent and accurate geospatial data can play a key role in the acknowledgement of indigenous and local communities and their inclusion in forest management decisions.

ISSUES IN SCIENCE AND TECHNOLOGY

VOL. XXXIV, NO. 3, SPRING 2018

Helping Communities Use Data to Make Better Decisions

BY SALLIE KELLER, SARAH NUSSER, STEPHANIE SHIPP, CATHERINE E. WOTEKI

Communities, especially small and rural ones, need to take advantage of new techniques for collecting and analyzing data to better serve their residents. Here's a plan to help them succeed.

Longer answer: Reason 4

ZME SCIENCE

How scientists are using climate records made by 15th century Japanese monks

If you're old enough, you might remember how some flowers around where you live blossom earlier or that summers and winters are unusually harsh. In short, freak weather is more common to the point it's becoming the new norm. Human memory is fallible, which is why we keep records of things like temperature, humidity, concentration of gases in the atmosphere and so on. These record don't go back that long though -- maybe only a century. Some, however, go way back and scientists are using these to keep track of climate change over the centuries.

by **Tibi Puiu** — April 30, 2016 in **Climate, News**

https://www.zmescience.com/ecology/climate/climate-change-japanese-monks/

PNAS Proceedings of the National Academy of Sciences of the United States of America

Ecological contingency in the effects of climatic warming on forest herb communities

Susan Harrison, Ellen I. Damschen, and James B. Grace

Here we couple historical and contemporary data to compare 60 y of change in adjacent communities differing in elevation and land use. In 2007–2009 we resampled, as closely as possible, 185 sites in the Siskiyou Mountains (Oregon) that first were studied by ecologist Robert H. Whittaker in 1949–1951 with the goal of assessing community variation along steep environmental gradients (25). The region is a hotspot of botanical diversity with

We obtained Whittaker's data from the Cornell University Library's Rare Document Division and entered it into a database with tree and shrub counts by species, herb cover by species, and site locations (road, elevation, slope, and aspect).

https://www.pnas.org/content/107/45/19362

Putz, F.E., Biotropica, 1983

(courtesy of FE Putz)

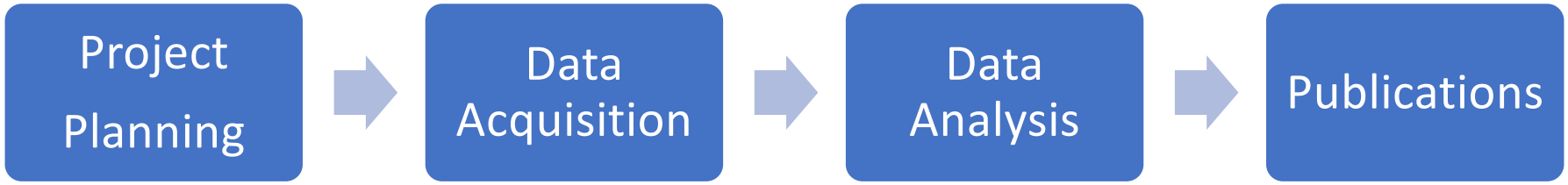
With that in mind….

1. Identify different kinds of data collected in different disciplines

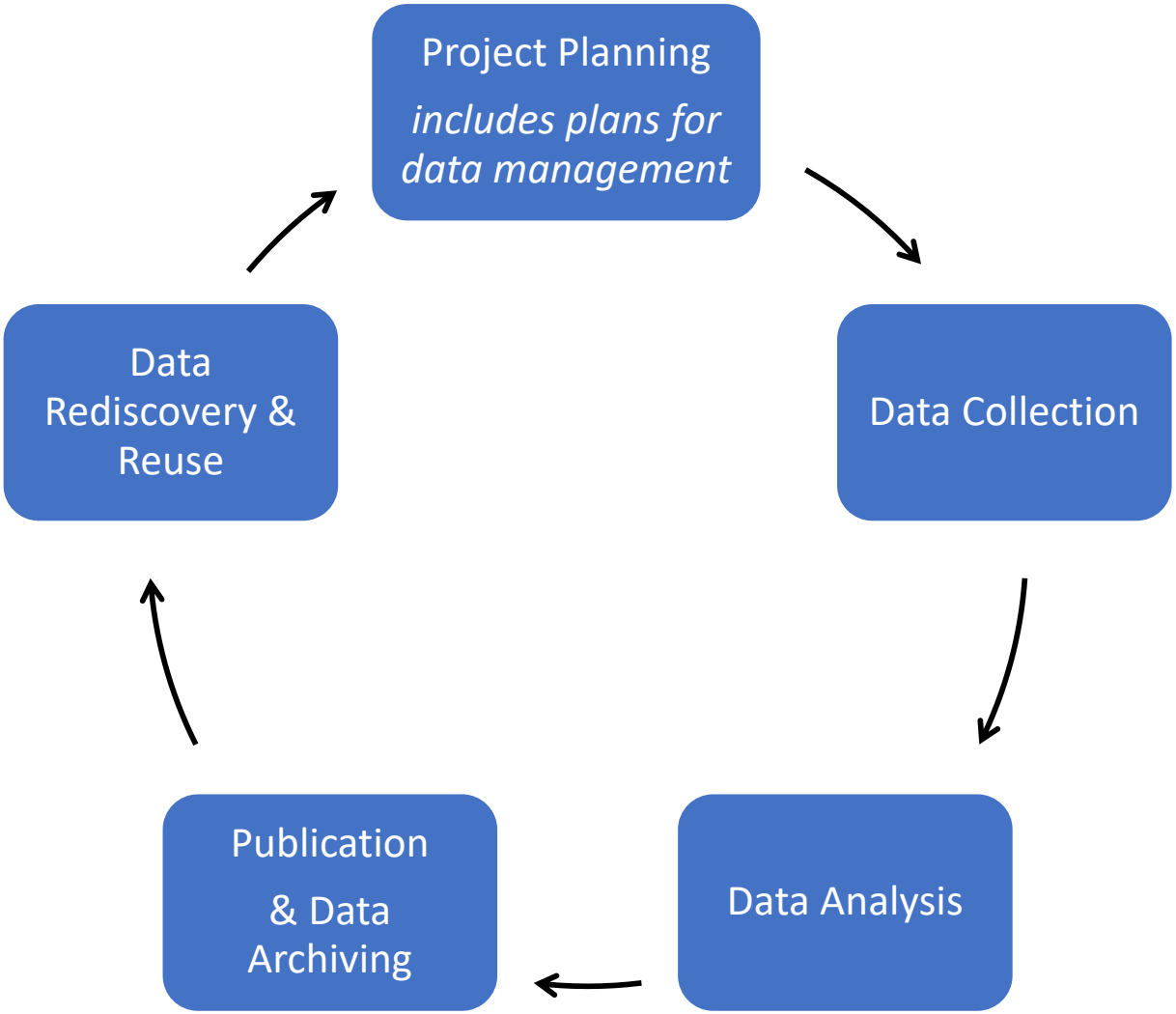
2. How are these data collected & recorded?

(and time for a snack…)

“Typical” (old-school) Research Flow



Contemporary “Research Data Life-Cycle”



Week	Dates	Topic
1	1/10	'Data' across disciplines and the Research Data Life Cycle
2	1/17	File Formats, Naming Conventions, Data Storage & Security
3	1/24	Structure & Format of Data & Datasets
4	1/31	Reproducible Data (Re)organization
5	2/7	QA/AC 1: Data Entry & Validation
6	2/14	QA/QC 2: Correction & Synthesis with Open Refine
7	2/21	QA/QC 3: Visualizing Data (to Find Mistakes)
8	2/28	Documentation: Metadata & Codebooks
9	3/7	Data Management Plans
10	3/14	No Class - Spring Break
11	3/21	Efficient Data Collection
12	3/28	Transcription & Translation
13	4/4	Paperless Data Collection
14	4/11	Automated Data Extraction
15	4/18	Legal & Ethical Issues
16	4/25	Data Sharing, Reuse, & Archives
Finals Week	5/1	Submission of Final Projects by 5 pm

My data are a mess…

…and I can't remember what some of these things mean.

If only I had planned...

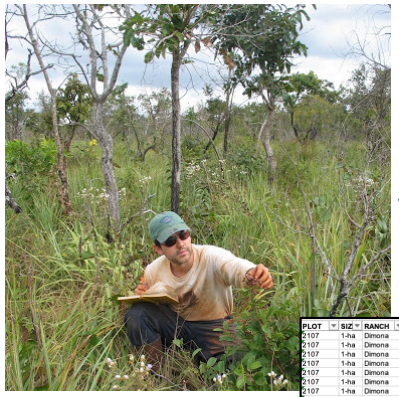
…how I was going to collect & archive my data.

Before Class

Read, Watch, Reflect

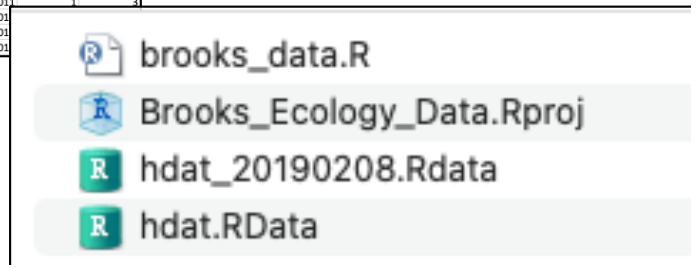
During Class

Exercises to Reinforce
Discuss Readings
Work on Individual Project



This class is an excuse to do what you have to do anyway.

key	number	source	prop.rem.1	prop.rem.2	prop.rem.3	plot	destination	year	replicate	replicate2
6	12591	High-N	0.489	0.489	0.439	2	Control	2009	1	1
7	12258	High-N	0.687	0.687	0.637	2	Control	2009	1	2
8	12295	High-N	0.76	0.76	0.71	2	Control	2009	1	3
9	9387	High-N	0.827	0.827	0.777	2	Control	2009	1	4
80	12352	Low-N	0.429	0.429	0.379	2	Control	2009	1	1
81	9403	Low-N	0.656	0.656	0.606	2	Control	2009	1	2
82	9466	Low-N	0.775	0.775	0.725	2	Control	2009	1	3
83	9481	Low-N	0.884	0.884	0.834	2	Control	2009	1	4
84	12609	Low-N	0.892	0.892	0.842	2	Control	2009	1	5
155	12771	Control	0.607	0.607	0.557	2	Control	2009	1	1
156	12563	Control	0.736	0.736	0.686	2	Control	2009	1	2
157	12114	Control	0.772	0.772	0.722	2	Control	2009	1	3
158	12601	Control	0.776	0.776	0.726	2	Control	2009	1	4
159	12598	Control	0.824	0.824	0.774	2	Control	2009	1	5
229	9371	High-N	0.84739379	0.84739379	0.797	2	Control	2010	1	1
230	9126	High-N	0.85311909	0.85311909	0.803	2	Control	2010	1	2
231	9280	High-N	0.87127609	0.87127609	0.821	2	Control	2010	1	3
232	12831	High-N	0.89589064	0.89589064	0.846	2	Control	2010	1	4
233	12375	High-N	0.91290673	0.91290673	0.863	2	Control	2010	1	5
305	12409	Low-N	0.84452011	0.84452011	0.795	2	Control	2010	1	1
306	12048	Low-N	0.87297216	0.87297216	0.823	2	Control	2010	1	2
307	12807	Low-N	0.88332576	0.88332576	0.833	2	Control	2010	1	3
308	12564	Low-N	0.88733004	0.88733004	0.837	2	Control	2010	1	4
309	9088	Low-N	0.90896663	0.90896663	0.859	2	Control	2010	1	5
381	12630	Control	0.85197849	0.85197849	0.802	2	Control	2010	1	1
382	12388	Control	0.85995751	0.85995751	0.81	2	Control	2010	1	2
383	12760	Control	0.89649557	0.89649557	0.846	2	Control	2010	1	3
384	12131	Control	0.90523975	0.90523975	0.855	2	Control	2010	1	4
385	9078	Control	1.00626173	0.95	0.956	2	Control	2010	1	5
454	12659	High-N	0.66091574	0.66091574	0.611	2	Control	2011	1	1
455	12257	High-N	0.75657844	0.75657844	0.707	2	Control	2011	1	2
456	12119	High-N	0.81600997	0.81600997	0.766	2	Control	2011	1	3
457	12587	High-N	0.90932914	0.90932914	0.859	2	Control	201		
458	12347	High-N	0.94207819	0.94207819	0.892	2	Control	201		
529	12721	Low-N	0.81127107	0.81127107	0.761	2	Control	201		



Let’s find out a little more…

In-Class Exercise Wk. 1