

ADSA Data Science Community Newsletter

Data Science Community Newsletter features journalism, research papers and tools/software for July 27, 2022.

Please let us ([Micaela Parker](#), [Brad Stenger](#), [Laura Norén](#)) know if you have something to add to next week's newsletter. We are grateful for the generous financial support from the [Academic Data Science Alliance](#).

We hope everyone has enjoyed a relaxing, enjoyable summer with still more sunny days and festive nights ahead.

STARGAZING + HIKING

Introducing night-hiking, the practice of walking through natural areas free of light-pollution, usually in order to see astronomical entities. This [list has nine stargazing hikes](#) in Idaho, Arkansas, South Dakota, Maine, Colorado, Florida, and Washington. But any hike away from city lights will reveal a stellar cornucopia.

If you're into astronomy, but find walking around in the woods after dark creepy or foolish (all those roots to trip over and disease-carrying ticks to avoid), you may prefer to read about the exciting [first two weeks](#) of the new James Webb telescope.

VISIT YOSEMITE FROM YOUR DESK

Yosemite is one of the US's most popular national parks, but you don't have to be in the park to attend the monthly [Yosemite Forum](#). In August **Michael Goss** from **NOAA** will be presenting "Climate change is increasing the likelihood of extreme autumn wildfire conditions across California" and in September **Crystal Kolden** of **University of California, Merced** presents a more optimistic talk on wildfires "How I learned to love wildfire: a guide to surviving in a burning world."

NORTH AMERICAN SUPERVOLCANOS AND YELLOWSTONE

These past couple weeks have had many people sweltering in record-breaking heat. This got us thinking about the relatively narrow temperature ranges humans can tolerate compared to geological heat scales (the Earth's core is 9,392° Fahrenheit) and united our fondness for National Parks with the most dramatic geological 'heatwaves': volcanic activity.

The second largest supervolcano crater in North America [sits under Yellowstone National Park](#) where, "magma may be as little as 3-8 miles beneath Sour Creek Dome and 8-12 miles beneath Mallard Lake Dome, and both domes inflate and subside as the volume of magma or hydrothermal fluids changes

beneath them." It's unlikely there will be a magma-producing event within our lifetime and evidence suggests that the next resurgence is more likely to be a lava flow than an explosion.

If you're wondering what the largest supervolcano in North America is, **Brigham Young University** scientists **discovered** the Wah Wah supervolcano on the border of Utah and Nevada in 2013. It's dormant, but significantly bigger than the one under Yellowstone. There's also a smaller, but more active supervolcano in California — **the Long Valley caldera**. Unlikely to blow in our lifetime, it appears to be the only North American volcano that will produce a Hollywood-style volcanic eruption for its next act. In the meantime, "a robust geothermal system inside the caldera fuels the Casa Diablo power plant, part of the **Mammoth Geothermal Complex**, just east of Mammoth Lakes, California, which generates enough power for 40,000 homes."

WHALES AND SENSORS (BEACH READ!)

Despite being beloved and large, whales are not well understood. Two attempts to learn more about these majestic ocean-dwellers use sensors and data science. First, sensors originally designed to detect seismic shifts along the Juan de Fuca tectonic plate — which could generate a massive earthquake in or near Seattle — turned out to pick up the whale calls of fin whales. First considered a data collection problem to the geologists interested in plate tectonics, the sensors dutifully recording whale calls were **eventually treated** as important primary data collectors for understanding the swimming routes and migratory patterns of fin whales. Now there's a large 70-sensor array of ocean-based sensors tracking multiple whale species and some dolphins called the **Cascadia Research Collective**.

But there's another whale project out of the **Cascadia Initiative** and collaborators at **University of California, Santa Cruz** in Monterey Bay that suction cups sensors right onto blue and humpback whales for 12-24 hours at a time [explainer **video**, including suction team]. These sensors help scientists understand whale's feeding, migration, and biological baselines (heart rate, temp), but also to gather data about the ocean.

"TOO OLD" IN SOCCER IS A NUMBER MUCH HIGHER THAN 30

Many football clubs start to twilight players when they reach ages from 30-32, but **new data science** suggests this approach is wasting talent. In an important counter-attack to coaches and GMs who refuse to renew contracts for players over 30, evidence suggests that players over 34 do not have less playing time than younger players (e.g. they can keep up just fine, thank you), and that these longer careers have not come "at any notable cost to their performance." In fact, "older players win more aerial duels, complete more dribbles, pass with greater accuracy — if they are central midfielders — and score more goals." The players note that their performance past age 30 is less likely to be rattled by nerves.

The evidence among soccer players is consistent with what is happening among world-class marathon runners. **Eliud Kipchoge**, considered the best mens marathoner in the world, is 37 years old. Kipchoge **opted not to run** at the recent World Championship races held in Oregon, but many of the male and female elites who did run are older than 35, including Americans **Sara Hall** (39) and **Keira D'Amato** (37). Hall, D'Amato and another American, Emma Bates (30), each **placed in the Top-10** of the Women's Marathon World Championship. As **University of Oregon** physiologist **Damien Callahan explained** on **KUOW** public radio, that's because more parts of our bodies contribute to peak performance during endurance runs, such as heart, lungs and fuel delivery. Some of those body systems age slower than the raw muscle power necessary for sprints.

We're not going to be that person, the one who tells you that "life's a marathon not a sprint." That person is unbearable, no matter what the science says.

SUMMER FUN FACTS

- + According to simulated data, "turf grasses, occupying 1.9% of the surface of the continental United States, would be the single largest irrigated crop in the country" and are carbon sequestering. [**Milesi et al**]
- + The **@chuckumentary** Twitter account is all photos and short videos of Lake Superior. Worth it.
- + **The Ocean Cleanup** project has **removed** 1/1000th of the plastic from the Great Pacific Garbage

Patch

- + Speaking of plastic clean-up, if you see plastic in a US national park, there's a [new citizen science/activism project](#) asking you to photograph it to help catalog plastic's impact on the parks.
- + **Uber paid** \$100,000 to get the esteemed **Alan Krueger** to publish a paper with the **National Bureau of Economic Research**. Krueger did disclose his consulting relationship with Uber, but not the total \$\$\$.
- + Scientists [located](#) a living example of a rare oak species thought to be extinct in **Big Bend National Park**, the tardifolia. It has a fungal infection and scarring from wildfires, but it's alive.
- + Grocery and takeout delivery robots run by Russian-based **Yandex** are [driving](#) residents off the sidewalks of Tel Aviv.

NEW PROGRAMS, FOLLOW THE MONEY

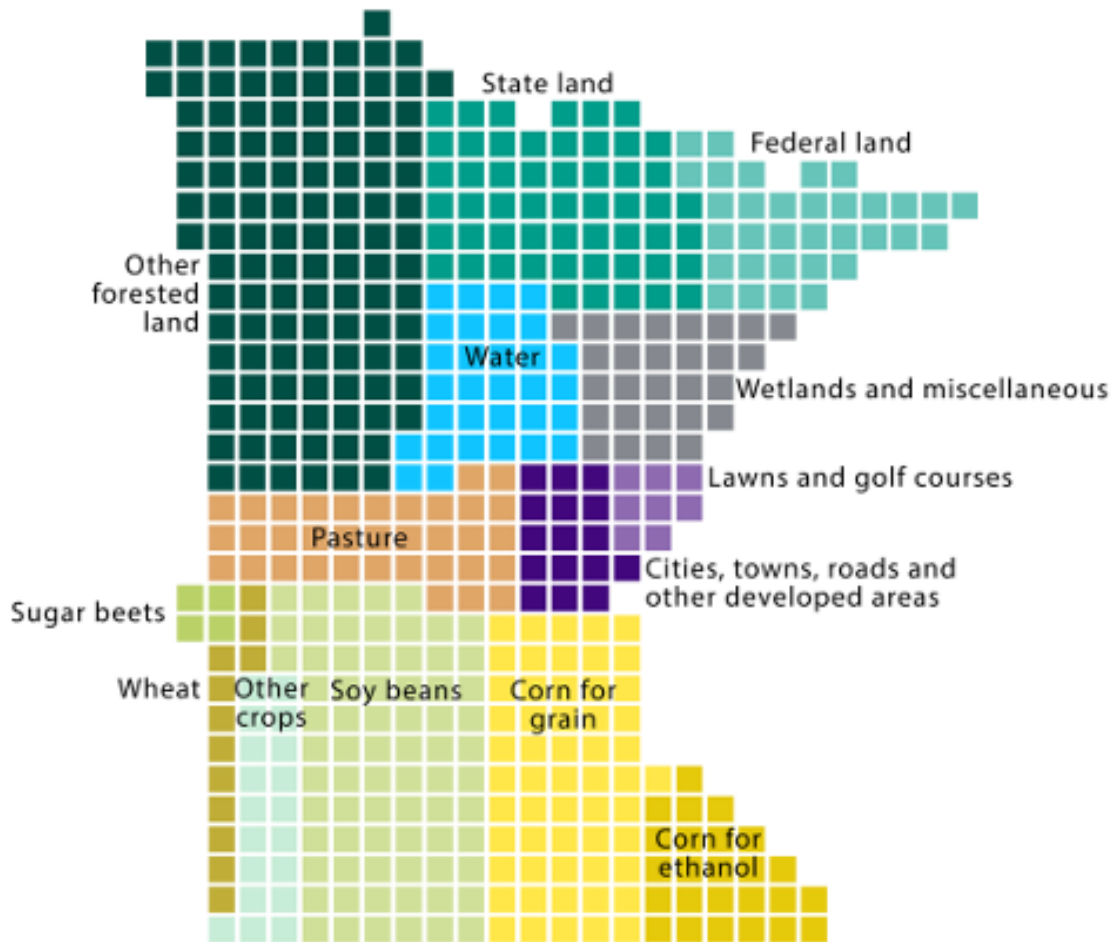
Click through to access [a structured spreadsheet](#) of New Programs and money moving around in academic data science.

DATA VISUALIZATION OF THE WEEK

by Christopher Ingraham for [The Minnesota Reformer](#) using data from the U.S. Department of Agriculture, Minnesota Department of Natural Resources, and the University of Minnesota from July 26, 2022

How Minnesota's land is used

Estimated proportions of Minnesota surface area, by land cover and use



Sources: USDA, MN DNR, University of Minnesota, Milesi et al. (2005)

Minnesota Reformer

Deadlines

Education Opportunities

[Applications to become a Fall 2022 Google Public Policy Fellow are now open!](#)

"If you're a student interested in the Internet and tech policy, apply today"

Contests/Award

[@lena_maierhein, @annika_re, @mjorgecardoso et al. present the results of a biomedical image segmentation challenge](#)

"Showing that a method capable of performing well on multiple tasks will generalize well to a previously unseen task @MICCAI_Society." The competition referenced, the Medical Segmentation Decathlon, is still ongoing.

Tools & Resources

[Announcing a new tool for exploring federal grant data: <http://GrantExplorer.org>.](#)

Twitter, Jevin West from July 15, 2022

With support from [the **Moore Foundation**], **Cole Chamberlin**, [@jportenoy](#) & I built a tool for answering

questions like, how much @NSF funding has supported “data science”? Feedback welcome.

The Caltech Fish Counting Dataset: A Benchmark for Multiple-Object Tracking and Counting

DeepAI, Justin Kay, et al. from July 19, 2022

"We present the **Caltech** Fish Counting Dataset (CFC), a large-scale dataset for detecting, tracking, and counting fish in sonar videos. We identify sonar videos as a rich source of data for advancing low signal-to-noise computer vision applications and tackling domain generalization in multiple-object tracking (MOT) and counting."

Inside the effort to refine one of the world’s most popular programming languages

TechRadar.Pro, Joel Khalili from July 23, 2022

"For almost three decades, all revisions and additions to the language were vetted personally by its creator, **Guido van Rossum**, who was affectionately known as the Benevolent Dictator for Life (BDFL). But since 2018, this role has been fulfilled by the five members of the **Python Steering Council** instead."

About Us: The Data Science Community Newsletter was founded in 2015 in the Moore-Sloan Data Science Environment at NYU's Center for Data Science. We continue to be supported by the Gordon and Betty Moore Foundation and the Alfred P. Sloan Foundation through the [Academic Data Science Alliance](#). Our archive of newsletters is at <https://academicdatascience.org/resources/newsletter>. Our mailing address is [1037 NE 65th St #316; Seattle, WA 98115](#).

OPT OUT: If you do not want to receive this newsletter, please email brad.stenger@gmail.com with the word 'unsubscribe' in the subject line.

OPT IN: Feel free to forward the Data Science newsletter to colleagues. They can sign up for the newsletter [using this web form](#).