



## Data Science Community Newsletter

20 March 2020

Data Science Community Newsletter Issue 194

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

### DATA SCIENCE NEWS

#### University Data Science News

Let's dive right in, lots to talk about this week with new studies coming out.

The biggest story this week are the following two papers:

1. [Impact of non-pharmaceutical interventions \(NPIs\) to reduce COVID-19 mortality and healthcare demand](#) otherwise known as the [The Imperial College COVID19 Study](#) [Update: lead author of the Imperial Study, Neil Ferguson, [has contracted COVID-19](#)]
2. [Severe Outcomes Among Patients with Coronavirus Disease 2019 \(COVID-19\) — United States, February 12–March 16, 2020](#) by the U.S. Centers for Disease Control and Prevention (CDC)

In [the Imperial College Study](#) researchers looked at the impact of mitigation and suppression efforts on the spread of the coronavirus in the US and the UK. The paper is not terribly long and should be read in full. To sum up their findings: neither the US nor the UK have the hospital capacity to handle the number of patients likely to need hospitalization if neither suppression nor mitigation tactics are pursued. In order to keep the number of cases at or below the threshold in both countries schools and universities must close and remain closed, patients who have any symptoms must be isolated, people over 70 must be quarantined, and the entire population of both countries must practice social distancing. Under their simulation assumptions, this will usually keep the strain on hospitals right at the threshold. Local situations may have breakthroughs during which hospitals cannot handle the patient load; fatalities will inevitably spike when hospitals are at or above capacity. The paper also factored in the difficulty of maintaining social distancing for months on end. Even if the population is able to keep a stiff upper lip for five months, as soon as social distancing habits weaken, we're more or less back to where we are today because we won't have meaningfully increased herd immunity. Their message: Figure out how to avoid contact with other human beings, aside from those with whom you live. Epidemiologically, the safest route is to sit at home and wait for a vaccine or for a combination of rigorous testing and drugs that shrink the contagion window. (Economic analyses were not included in the Imperial College Study; neither were mental health measures.)

Now let's move on to the CDC paper.

The [CDC paper](#) looked at COVID-19 hospitalization and mortality data from China and the currently available data from the United States (N=4,226 in the US). The US data do not have information on serious underlying health conditions nor are all records complete with hospitalization length, age, ICU

admission or death (it can take weeks to die from COVID19). While the data continue to support the fact that higher age is correlated with higher likelihood of hospitalization and death, those in the 20-64 age group still have about a one in five chance of ending up in the hospital. That number suggests it is unwise to protect only those 65 or older. Companies and cities that have encouraged work from home appear to have had a good idea. Please take care of yourselves, readers. If we all have to learn how to relate to each other via screens and exercise at home, so be it. Again, the economic toll of job loss was not addressed.

A preprint (not peer reviewed) [published](#) by a team from **Notre Dame** estimated the case load in the U.S. They were inspired to do so because the U.S. is frustratingly unable to get its shit together and test broadly for COVID already. They estimate that fewer than 10% of community acquired cases were detected due to the restrictive rules that strongly favored testing those who had recently traveled even after community transmission had evidently begun. They also estimate that by Friday, March 13, a median of 22,876 people were already infected (95% posterior predictive interval: 7,451 - 53,044) compared to the CDC "official" count of 1,215 cases in the U.S. on that Friday. That's an ~18X difference in estimated versus reported.

If you're interested in jumping in and contributing your own skills to producing useful knowledge, you might want to check out the COVID-19 Open Research Dataset (CORD), a [collection of 29,000 articles \(13,000 full text\)](#) covering coronaviruses, related viruses, and research on our very own novel SARS-CoV-2. The project is a collaborative effort led by **The Chan Zuckerberg Initiative** with **Allen Institute for Artificial Intelligence**, **Georgetown University's Center for Security and Emerging Technology**, **Microsoft Research**, **Cold Spring Harbor Laboratory**, and the **National Library of Medicine of the National Institutes of Health**. The papers are all machine-readable.

The **University of Pittsburgh** has also [launched a portal](#) for sharing papers and data about COVID-19: the [MIDAS Online Portal for COVID-19 Modeling Research](#). This effort has somewhat less material, but efforts have been made to create rich metadata which is likely to aid in the process of efficiently discovering relevant information.

With the restrictions against local labs testing coronavirus samples loosening, the **University of Washington Virology** lab tested [a couple thousand last weekend](#). Their lab is on the list of [approved testing facilities](#) with a capacity of 2000 tests/day. Very impressive when the biggest private for-profit lab companies in the US — **Lab Corp** and **Quest** — are both rated at 5000 tests per day.

Speaking of testing, Professor **Pardis Sabeti** of **Harvard's TH Chan School of Public Health** calculated that Massachusetts will likely [need 1.4 million tests](#). They had done 475 as of last Sunday. Professor Sabeti's lab has had working diagnostics since January, "within days of receiving the virus's genome from Chinese researchers" and finds it absolutely unacceptable that testing has not ramped up. Los Angeles Mayor **Eric Garcetti** was reporting [continued testing delays of 4-6 days](#) in his late-Wednesday press briefing.

No longer at **UW Medicine Pathology**, professor and esteemed vascular biologist, **Steven Schwartz**, died of COVID-19.

The 80 **University of California-Santa Cruz** graduate students who had their contracts terminated and their health insurance cut off during the labor dispute between grad students and UC Chancellor **Janet Napolitano** a month ago have had their **health insurance re-instated**. Thank goodness for small favors. Of course, participating in protests is a violation of social distancing that raises interesting questions about the right of assembly.

Both **NYU** and **Columbia University** — probably many other universities — have suspended all in-person interactions with research subjects. Fail to follow this and you'll violate your IRB authorization.

Black hole astronomy — which seems like it could avoid COVID complications — is also looking at a cancellation. According to **Geoff Bower**, a scientist onsite, the **Event Horizon Telescope** in Hawaii will be **unable to continue most of its planned observations** because researchers have had their travel canceled.

**Stanford** population health statistician **John Ioannidis** thinks any leader who has asked workers to work from home is likely making a huge mistake. **His argument for this outlier opinion** is that there is insufficient data to react strongly to COVID because on the Diamond Princess Cruise, the reports suggest only 1% of the ship had it. They skewed older so that would make 1% an upper bound on fatality rates, according to Ioannidis. His viewpoint contradicts just about every other published piece written by epidemiologists or clinical doctors. **Hilda Bastian** took the time to **rebut five of Ioannidis' arguments**.

**Harvard students are upset** about their grading options, despite new pass-fail alternatives. **University of Virginia** Provost **Liz Magill** has defaulted to **a credit/no-credit grading system** for spring semester. Some students and faculty will get sick or have sick family members; this is an act of grace.

In a similar vein, several schools including **The Ohio State University**, have **announced** that they will treat AY2019-2020 as a non-year as far as the tenure clock is concerned.

2020 could turn out to be **the year without conferences**, raising incredibly important — and now accessibly empirical — questions about the value of the academic conference. Are they simply indulgent professional vacations that happen to involve anxiety-inducing presentations put together on the flight there? Do they catalyze problem discovery and solution finding? We'll now have the data to find out. (We will not, however, be able to answer one of the questions I want to answer: how much does conference attendance matter to a particular person's career? If all conferences are canceled, we cannot understand the value of conferences for one scholar relative to another extremely similar scholar who did not attend conferences.) In more practical matters, conference organizers are scrambling to try to predict when travel bans will be lifted.

In proof that the world still turns on its axis, another college has gotten into the data science game. The college of **William & Mary** has **received state approval** to offer a Bachelor of Science degree in Data Science.

And **UC Santa Barbara** won a three-year, \$920,000 NSF grant to **start up the Central Coast Data Science Initiative**.

Astroparticle physicist **Dr. Eleonora Presani** **will be the first** Executive Director of the **arXiv**. Based in New York City she will lead fundraising and operations efforts for the first and biggest open access preprint publisher.

### Company Data Science News

**Ford** and **GM** have ordered their plants to stop making cars, but they may be **retooled to make ventilators**.

Craft distilleries have **retooled to manufacture chemical sanitizers** in **Vermont, Michigan, Houston, Cincinnati and Western New York**.

An **Amazon** warehouse worker at a facility in Queens, New York **tested COVID+**. Research suggests COVID can **live on cardboard for up to 24 hours**.

**Bill Gates** is **stepping down** from the executive boards of **Microsoft** and **Berkshire Hathaway** to focus on The Gates Foundation. **Eric Horvitz** will **become** Microsoft's Chief Scientific Office. **Peter Lee** will take over running **Microsoft Research**.

**Starsky Robotics**, a self-driving truck company, compared itself to Shackleton trying to cross Antarctica, and has closed up shop. Why? In the CEO's own words: "**supervised machine learning does not live up to the hype**." With quotes like that, my snarky job here is easy.

Canceling the **NBA** season — but preserving the playoffs somehow — is projected to **cost the league and owners \$1 billion** according to new analysis by **Neil Paine** of **FiveThirtyEight**.

**Thermo Fisher** will be able to produce **5 million tests** per month? as of April 3rd. **Roche** is able to manufacture 400,000 test per week soon, too, which will greatly help identify cases, quarantine people, and quarantine close contacts. I wish I were more optimistic but I'll believe it when these tests are being deployed and coming back to providers within 24 hours.

**LinkedIn**, owned by **Microsoft** since 2016, is continuing to **fight a Ninth Circuit Court ruling** in favor of a small startup called **hiQ** that used bots to scrape personal data from LinkedIn in violation of LinkedIn's terms of service and probably users' privacy. The lower court argued that because the data hiQ scraped was publicly available without a login, it was fair to scrape it. LinkedIn disagrees and has appealed to the **Supreme Court of the United States**. This case will *define* the boundary between public and private on the internet (at least in the US) and we will continue to bring you updates.

In another **case to watch** — this one against **Google** in the **EU** — **Dr. Johnny Ryan** has sued the company "to force Google to reveal what it does with everybody's personal data." Brave, the entity sponsoring the case, believes that Google's sharing of internal data across its many divisions from Gmail to YouTube to Google Maps may violate the GDPR's "purpose limitation" that "Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes..." We shall see how far this case gets.

Speaking of Google's many products, the Google Translate app for Android now offers **real-time transcription** for "any combination of eight currently supported languages: English, French, German, Hindi, Portuguese, Russian, Thai, and Spanish."

**GitHub**, once itself an open source project acquired by **Microsoft**, has **acquired the open source project npm**, the package manager for shareable computer programs written in JavaScript.

### Government Data Science News

If you would like to donate, the **U.S. Center for Disease Control and Prevention's CDC Foundation** is **accepting donations** to support efforts to treat and contain the coronavirus in the US. It feels bizarre to ask people to donate to the government, but we are in unprecedented times.

The already beleaguered **2020 Census** is facing another problem. Social distancing rules prevent Census workers from going door to door. At the moment, **field operations have been delayed** until April 1st, but that seems wildly optimistic.

Before coronavirus turned New York into an eerily quiet hellscape waiting to burn millions with fever dreams, the **New York City Council** passed legislation outlawing bias in algorithmic talent recruiting processes. The law requires that any company selling a tool to aid in employment decisions "governed by statistical theory" must conduct a bias audit proving the tool is compliant with employment law. There are **many open questions** remaining such as where the training set of CVs would come from in these audits and who gets to perform the bias audits.

**Israel** is **monitoring its citizen's locations** and the social networks of those they've been in contact with to mandate person-specific quarantines. If an Israeli's phone gets close to the phone of someone who eventually tests positive, that person will receive a text telling them to stay home for 14 days.

**Prime Minister Netanyahu** admitted that this approach involves, "a certain degree of violation of the privacy." Yeah, that is certainly the truth. As I've said elsewhere protecting public health and finding missing children are generally the sheep's clothing that the wolves of surveillance and authoritarian control hide under in democratic societies.

**NIH** has **begun Phase 1** of an investigational vaccine for COVID-19. Healthy adult volunteers can enroll. Any vaccine ready to be used across the whole human population is still 12-18 months off.

### Extra Extra

This story of a couple who had almost solved the two-body problem by getting an offer for one with a promise of a spousal hire for the other at the **University of Michigan** turned sinister and sociopathic when a jealous colleague **made up a series of sexual harassment complaints**. Spoiler: **Roxane Gay's tweeps** seem to know who he is.

Or for something *completely different* try watching **sixty free science movies** underwritten by the **Alfred P. Sloan Foundation**. There's also **Niall Ferguson's new three-part documentary series** on network science — even if you know the subject it's a master how-to on communicating complex information to a broad audience.

Oh, and **Jstor made serious fools of themselves**, putting out an article that implied they had made a bunch of new, content free when really...they made no changes in their free access policy that puts most content behind a paywall.

Social isolation is **generally bad for human health**. Take walks to expand your horizons and stretch your legs. Be in touch with family and/or friends. With all of our apps, phones, and cameras let's look after one another, shall we?

#### **Tweet of the Week**

*Twitter, Dana Jay Bein from March 18, 2020*

 Dana Jay Bein (#DJB)  
@danajaybein

I've lost my mind.

I wrote Coronavirus Rhapsody:

Is this a sore throat?  
Is this just allergies?  
Caught in a lockdown  
No escape from reality.

9:02 AM - Mar 18, 2020 · TweetDeck

37.4K Retweets 127.4K Likes

  
Dana Jay Bein (#DJB) @danajaybein · 10h  
Replying to @danajaybein  
Don't touch your eyes  
Just hand sanitize quicklyyyyy  
I'm just a poor boy, no job security  
Because of easy spread, even though  
washed your hands, laying low  
I look out the window, the curve doesn't look flatter to me, to me

  
Dana Jay Bein (#DJB) @danajaybein · 10h  
mama, just killed a man  
i didn't stay inside in bed  
I walked by him, now he's dead  
mama, life was so much fun  
but now I've caught this unforgiving plague

  
Dana Jay Bein (#DJB) @danajaybein · 10h  
mama, oooohh  
didn't mean to make them die  
if I'm not back to work this time tomorrow  
carry on, carry on as if people didn't matter

  
Dana Jay Bein (#DJB) @danajaybein · 10h  
oo late, my time has come  
sends shivers down my spine  
body's aching all the time  
goodbye everybody, I've got the flu  
gotta leave you all behind and face the truth

  
Dana Jay Bein (#DJB) @danajaybein · 10h  
mama, oooohh  
I don't wanna die  
I sometimes wish I never went out at all

## Video of the Week

YouTube, CDC from March 16, 2020



## EVENTS

### [Join us for the digital Google for Games Developer Summit](#)

**Online** March 23, starting at 9 a.m. Pacific time. "Last month, **Game Developers Conference** (GDC) organizers made the difficult decision to postpone the conference. We understand this decision, as we have to prioritize the health and safety of our community. GDC is one of our most anticipated times of the year to connect with the gaming industry. Though we won't be bringing the news in-person this year, we're hosting the **Google for Games Developer Summit**, a free, digital-only experience where developers can watch the announcements and session content that was planned for GDC." [free]

### [COVID-19 and AI: A Virtual Conference](#)

**Online** April 1. The virtual conference "will convene experts from **Stanford** and beyond to advance the understanding of the virus and its impact on society. It will be livestreamed to engage the broad research community, government and international organizations, and civil society." [Event details and agenda will be posted soon.]

## DEADLINES

### Contests/Award

#### [COVID-19 Open Research Dataset Challenge \(CORD-19\)](#)

"A list of our initial key questions can be found under the Tasks section of this dataset. These key scientific questions are drawn from the NASEM's SCIED (**National Academies of Sciences, Engineering, and Medicine's Standing Committee on Emerging Infectious Diseases and 21st Century Health Threats**) research topics." ... "Many of these questions are suitable for text mining, and we encourage researchers to develop text mining tools to provide insights on these questions." Round 1 submission deadline is April 16.

#### [Announcing the 2020 John Hunter Excellence in Plotting Contest](#)

"Participants are invited to submit scientific plots to be judged by a panel. The winning entries will be announced and displayed at **SciPy 2020** or announced in the John Hunter Excellence in Plotting Contest website and youtube channel." Deadline for submissions is June 1.

#### [ACM RecSys Challenge 2020](#)

"The challenge focuses on a real-world task of tweet engagement prediction in a dynamic environment. The goal is to predict the probability for different types of engagement (Like, Reply, Retweet, and Retweet with Comment) of a target user for a set of tweets, based on heterogeneous input data." RecSys Challenge ends on June 7.

## Education Opportunities

### [BIDS Data Science Fellowships](#)

"The **Berkeley Institute for Data Science** (BIDS) announces a call for applications for the BIDS Data Science Fellowship Program. Successful applicants will be offered two-year appointments at BIDS beginning in the 2020 Fall Semester." Deadline for applications is April 1.

### [Apply for Frontier Development Lab's 2020 Summer Program](#)

"Frontier Development Lab (FDL) is looking for researchers at the PHD and Post Doc level with

expertise in machine learning/artificial intelligence/data science as well as those with expertise in science domains including heliophysics, planetary science, Earth science, disaster management, astrophysics and astronaut health." Deadline to apply is April 6.

### **Studies/Surveys**

**are you a #DataScience practitioner who has had trouble finding suitable venues for publishing your work? we'd like to know more about these challenges.**

share your stories here! DMs also open! [Academic Data Science Alliance]

### **RFP**

**Given shift to virtual classes and events, am interested to explore Sloan support for research on outcomes/experience as well as new models beyond Zoom/Slack paradigms.**

Can turn around funding at small scale fast. Email me Letters of Inquiry; DMs open for leads/pointers.

[Josh Greenberg at Sloan Foundation]

### **Funding for COVID-19 Projects - Sam Altman**

"I'm trying to fund startups/projects helping with COVID-19, because it's basically the one thing I know how to do that can help. I think we will soon have enough testing capacity, so now I'd like to start funding more startups working on:"

"Producing a lot of ventilators or masks/gowns very quickly. This will require a lot of repurposing and creativity but thankfully is an engineering problem not a scientific ones."

"Screening existing drugs for effectiveness."

"Novel approaches to vaccines (i.e., not doing what the big pharma companies are already doing)."

"Novel therapeutics that the big pharma companies are unlikely to work on."

### **TOOLS & RESOURCES**

#### **Google Cloud for Student Developers: G Suite APIs (intro & overview)**

*Google Developers Blog*, Wesley Chun from March 09, 2020

"Students graduating from STEM majors at universities with development experience using industry APIs (application programming interfaces) have real-world practice that can prove valuable in terms of career readiness."

"To that end, the **Google Cloud** team is creating a "Google Cloud for Student Developers" YouTube video series crafted specifically for the student developer audience."

#### **Introducing the TensorFlow Developer Certificate!**

*The TensorFlow Blog*, Alina Shinkarsky, from March 12, 2020

"We are excited to launch the TensorFlow Developer Certificate, which provides developers around the world the opportunity to showcase their skills in ML in an increasingly AI-driven global job market. This certificate in TensorFlow development is intended as a foundational certificate for students, developers, and data scientists who want to demonstrate practical machine learning skills through building and training of basic models using TensorFlow."

#### **Election News Pathways Methodology - American Trends Panel February 2020 survey methodology**

*Pew Research Center, Journalism & Media* from March 11, 2020

"The American Trends Panel (ATP), created by **Pew Research Center**, is a nationally representative

panel of randomly selected U.S. adults. Panelists participate via self-administered web surveys. Panelists who do not have internet access at home are provided with a tablet and wireless internet connection. The panel is being managed by **Ipsos**."

"Data in the Election News Pathways project is drawn from the panel wave conducted Feb. 18 to March 2, 2020. A total of 10,300 panelists responded out of 11,036 who were sampled, for a response rate of 93%. This does not include six panelists who were removed from the data due to extremely high rates of refusal or straightlining. The cumulative response rate accounting for nonresponse to the recruitment surveys and attrition is 5%. The break-off rate among panelists who logged onto the survey and completed at least one item is 1.3%. The margin of sampling error for the full sample of 10,300 respondents is plus or minus 1.5 percentage points."

### **Announcing Bokeh 2.0**

*Medium, Bokeh from March 09, 2020*

"Bokeh 2.0 brings enhanced security options, new widgets, improved Jupyter integration, and more!" "This is a major release that includes some breaking changes, with the goal of improving the usability and maintainability of the library. The most important change to be aware of is that Python 3.6 is now the minimum Python version. Python 2 is no longer supported."

### **pycovid: Python package to easily access the most updated Coronavirus data using Python/PIP**

*GitHub - sudharshan-ashok from March 14, 2020*

"The PyCOVID package provides a Pandas Dataframe of the 2019 Novel Coronavirus COVID-19 (2019-nCoV) epidemic based on Rami Krispin's 'coronavirus' package in R. The raw data pulled from the **Johns Hopkins University Center for Systems Science and Engineering** (JHU CCSE) Coronavirus"

### **[2002.09405] Learning to Simulate Complex Physics with Graph Networks**

*arXiv, Computer Science > Machine Learning; Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, Peter W. Battaglia from February 21, 2020*

"Here we present a general framework for learning simulation, and provide a single model implementation that yields state-of-the-art performance across a variety of challenging physical domains, involving fluids, rigid solids, and deformable materials interacting with one another."

### **COVID-19 Open Research Dataset (CORD-19)**

*The Allen Institute for Artificial Intelligence, Semantic Scholar from March 16, 2020*

"In response to the COVID-19 pandemic, the **Allen Institute for AI** has partnered with leading research groups to prepare and distribute the COVID-19 Open Research Dataset (CORD-19), a free resource of over 29,000 scholarly articles, including over 13,000 with full text, about COVID-19 and the coronavirus family of viruses for use by the global research community."

### **Exploring helpful uses for BERT in your browser with Tensorflow.js**

*The TensorFlow Blog; Philip Bayer, Ping Yu and Jason Mayes from March 13, 2020*

"We've prototyped a Chrome Extension using the MobileBERT Q&A model that allows us to ask a question of any webpage. Using TensorFlow.js, the extension returns an answer based on the contents of the page. The model runs entirely on-device in the browser session, so nothing is ever sent to a server, maintaining privacy."

## **CAREERS**

### **Tenured and tenure track faculty positions**

#### **Associate Professor/Professor/Full Professor Research Position in Data Management and Data Wrangling**

Hasselt University, Data Science Institute; Hasselt, Belgium

#### **Assistant/Associate Professor and Director of the Spatial Analysis & Visualization Initiative**

Pratt Institute, School of Information

### **Full-time, non-tenured academic positions**

#### **OHI/O Program Coordinator**

The Ohio State University, Computer Science and Engineering Department; Columbus, OH

#### **Research Director**

University of Chicago, Urban Crime Labs; Chicago, IL

#### **Program Analyst, Ocean Team**

University of California-Berkeley, Haas School of Business; David and Lucile Packard Foundation; Berkeley, CA

#### **Visiting Assistant Professor in Learning Engineering**

Boston College, Lynch School of Education; Chestnut Hill, MA

## **Postdocs**

#### **Knight Postdoctoral Research Fellow - "algorithms, misinformation, polarization, propaganda, political institutions, and platform companies"**

University of North Carolina, School of Information and Library Science; Chapel Hill, NC

#### **Hewlett Postdoctoral Research Fellow - "misinformation, partisanship, polarization, propaganda, political institutions, and journalism"**

University of North Carolina, School of Information and Library Science; Chapel Hill, NC

#### **AI/Machine Learning/Encryption Postdoctoral Research Associate (CS Position)**

University of North Carolina, School of Information and Library Science; Chapel Hill, NC

#### **AI/Machine Learning/Encryption Postdoctoral Research Associate (Social Science PhD)**

University of North Carolina, School of Information and Library Science; Chapel Hill, NC

#### **Postdoctoral Associate - Open Science**

Carnegie Mellon University, University Libraries; Pittsburgh, PA

#### **Postdocs (2) Computational Sociology - Science of Science**

University of Copenhagen, Department of Sociology; Copenhagen, Denmark

#### **NEON Postdoctoral Fellows**

Battelle, National Ecological Observatory Network (NEON); Boulder, CO

#### **Postdoc positions available in the Duffy Lab at Michigan!**

University of Michigan, Department of Ecology & Evolutionary Biology; Ann Arbor, MI

## **Full-time positions outside academia**

#### **Director of Technology Projects**

Electronic Frontier Foundation; San Francisco, CA

**Diversity and Inclusion Research Fellow**

Partnership on AI; San Francisco, CA

**Research and Analytics Director, Chatbots**

National Domestic Workers Alliance, NDWA Labs; New York, NY, or Remote

**Data Science Manager**

Center for Policing Equity; Washington, DC

**Research Analyst (Health Policy)**

The Urban Institute; Washington, DC

**Senior Researcher**

Girl Scouts; New York, NY

**Research Analyst, Africa Growth Initiative**

The Brookings Institution, Global Economy and Development program; Washington, DC

**Data Analysis Team Manager**

Space Telescope Science Institute; Baltimore, MD

**Senior Data Visualization Editor**

McKinsey, Global Editorial Services; Atlanta, Boston, New York City or Waltham, MA

**Research Engineer, AllenNLP**

The Allen Institute for Artificial Intelligence; Seattle, WA

**Policy Fellow or Policy Counsel (Youth & Education)**

Future of Privacy Forum; Washington, DC

**Quantitative Researcher & Developer**

Hudson and Thames Quantitative Research; New York, NY

**Data Scientist**

Observable; San Francisco, CA

**Machine Learning Engineer - High-Performance Deep Learning and Neural Computing**

Target Data Sciences; Sunnyvale, CA

**Director, IARPA Office of Collections**

Office of the Deputy Director of National Intelligence for Strategy and Engagement (S&E); Bethesda, MD

**Director, IARPA Office of Analysis**

Office of the Deputy Director of National Intelligence for Strategy and Engagement (S&E); Bethesda, MD

**Internships and other temporary positions****Basketball Analytics Summer Intern**

Los Angeles Lakers; El Segundo, CA

**Adelie Penguin Population Ecology Internship**

Point Blue Conservation Science; Ross Island, Antarctica

**part-time grad student work**

New York University, Marron Institute; New York, NY, or Berlin, Germany

**Data Science Intern - Merch Product Development**

Stitch Fix; San Francisco, CA

**Click here to receive the Data Science Community Newsletter** and/or to have us follow your twitter feed so that our data science twitter bot can easily grab links from your tweets.

To send us an announcement for the newsletter, please email [laura.noren@nyu.edu](mailto:laura.noren@nyu.edu) and [brad.stenger@gmail.com](mailto:brad.stenger@gmail.com). We retain curatorial discretion.

**Data Science Community Newsletter Issue 194.**