NYU Data Science Community Newsletter features journalism, research papers, events, tools/software, and jobs for August 28, 2015.

Please Let us (Laura Noren, Brad Stenger) know if you have something to add to the newsletter. We are grateful for the generous financial support from the Moore-Sloan data Science Environment and NYU’s Center for Data Science.

NYU Data Science Community Newsletter Issue 016.

Data Science News

Is mass authorship destroying the credibility of papers?
Times Higher Education, UK from August 24, 2015
The rise in ‘kilo-authors’ and ‘gift authorship’ is causing the academy to rethink how it assesses the worth of academic publications.

Budding UW Data Scientists Use Their Powers for Social Good
Xconomy from August 24, 2015
Earn a degree in the field of data science these days and your ticket is punched: Google, Amazon, Facebook, leading-edge academic research, a well-funded startup—they’re all clamoring for people proficient in the tools and techniques needed to sift through today’s endless streams of digital data in search of something valuable.

Social service organizations and local governments are confronting the data deluge, too, often without the capacity to pay the salaries that profit-driven companies can offer these sought-after experts.

Enter the University of Washington’s just-concluded Data Science for Social Good summer internship.

duecredit/duecredit · GitHub
GitHub, duecredit from August 27, 2015
Automated collection and reporting of citations for used software/methods/datasets

Why the Stock Market Is So Turbulent
... What’s fascinating is that there is no clear, simple story about what is different about the outlook now for interest rates, for United States and European corporate profits or for economic growth compared with one week ago, when the S.&P. 500 index was 10 percent higher.

Here’s how to make sense of what is a truly global story, stretching from the streets of Shanghai, where stock investing has become a middle-class sport in recent years, to the oil fields of both the Middle East and Middle America, to the hallways of power in the Federal Reserve in Washington.
Cohort Analysis with Python
Greg Reda from August 23, 2015
Despite having done it countless times, I regularly forget how to build a cohort analysis with Python and pandas. I've decided it's a good idea to finally write it out - step by step - so I can refer back to this post later on. Hopefully others find it useful as well.

I'll start by walking through what cohort analysis is and why it's commonly used in startups and other growth businesses. Then, we'll create one from a standard purchase dataset.

I am here to talk about the science behind visualization. I am Prof. Tamara Munzner from the University of British Columbia. Ask Me Anything! : dataisbeautiful
reddit.com/r/dataisbeautiful from August 26, 2015
Are there any instances you know of where alternative data visualization methods have led to breakthroughs in fundamental science?

Let me start at the top - it's both a good question and a hard one!

I have multiple answers along a continuum of speculation (which is arguably a big word for weaselling) depending on just how high the bar is for 'breakthrough'.

The most rock-solid answer is the least satisfying: no, I don't.

Rethinking Data: How to Create a Holistic View of Students
KQED, MindShift from August 26, 2015
For at least a decade now, the driving force behind education reform has been data. We talk about collecting data, analyzing data, and making data-driven decisions. All of this data can certainly be useful, helping us notice patterns we might not have seen without aggregating our numbers in some way, looking for gaps and dips and spikes, allowing us to figure out where we are strong and where we need help. In terms of certain academic behaviors, we can quantify student learning to some extent and improve our practice as a result.

Fashion Metric Builds a Strong Data Science Team in Austin
SiliconHills from August 27, 2015
Fashion Metric is working to solve the problem of ill-fitting clothes by employing data science.

The startup created software, which it calls a virtual tailor, to help online shoppers find clothes that fit by calculating their detailed body measures. It does this by simply asking a few questions online and then processing that information through its proprietary algorithm and databases to come up with the correct size and fit.

Behind the Scenes of “What’s Really Warming The World?” with the Bloomberg Team (DS#59)
Data Stories from August 20, 2015
Hi folks! We have Blacki Migliozzi and Eric Roston from Bloomberg on the show to talk about their recent data graphic piece on climate change called “What’s Really Warming The World?”.

The graphics shows, through a "scrollytelling", what factors may influence the world’s temperature according to well established climate models and it guides your through a
series of questions and visuals to see with your eyes what does correlate (spoiler: carbon emissions) and what does not. [audio, 44:46]

**Chicago’s Data Science for Social Good project trying to predict police misconduct before it happens**
*WBEZ 91.5 Chicago from August 21, 2015*

Every day there are thousands of interactions between police officers and citizens across the country. While most are uneventful, a small number leave a member of the public disrespected, unprotected, harassed or — in all too many cases seen recently — hurt or even killed.

This summer, fellows with Data Science for Social Good — a program at the University of Chicago that connects data scientists with governments and nonprofits — are working to predict when officers are at risk of misconduct, the goal being to prevent problems before they happen. [audio, 10:03]

**The Science Of Love In The 21st Century**
*The Huffington Post, Highline magazine from August 19, 2015*

... Over decades, John [Gottman] has observed more than 3,000 couples longitudinally, discovering patterns of argument and subtle behaviors that can predict whether a couple would be happily partnered years later or unhappy or divorced. He has won awards from the National Institute of Mental Health and the National Council of Family Relations and has become the subject of increasing public fascination. He went on Oprah and the “Today” show. A book he co-authored that summarizes his findings, Seven Principles for Making Marriage Work, is a New York Times best-seller. His work took off because the consistency of his predictions is astonishing. One 1992 experiment found that certain indicators in how couples talked about their relationship could forecast— with 94 percent accuracy—which pairs would stay together.

**Interview with Sherri Rose and Laura Hatfield**
*simplify stats from August 21, 2015*

Laura Hatfield and Sherri Rose are Assistant Professors specializing in biostatistics at Harvard Medical School in the Department of Health Care Policy. Laura received her PhD in Biostatistics from the University of Minnesota and Sherri completed her PhD in Biostatistics at UC Berkeley. They are developing novel statistical methods for health policy problems.

**SimplyStats: Do you consider yourselves statisticians, data scientists, machine learners, or something else?**

**Rose:** I’d definitely say a statistician. Even when I’m working on things that fall into the categories of data science or machine learning, there’s underlying statistical theory guiding that process, be it for methods development or applications. Basically, there’s a statistical foundation to everything I do.

**Events**

**6th NYU Data Science Showcase**

Featured Faculty talk by Vasant Dhar, Professor and Head, Information Systems Group, Co-Director, Center for Business Analytics, Talk Title: *Big Data Driven Innovation*

Wednesday, September 16, at NYU Torch Club, Tap Room, starting at 4 p.m.
NYC Media Lab 2015 Annual Summit
NYC Media Lab's Annual Summit is a snapshot of the best thinking, projects, and talent in digital media from universities in NYC and beyond.

This is an opportunity for media executives, technologists, and decision makers to see more than 100 university prototypes and demonstrations that explore interesting technologies and applications related to the future of media. Attendees will be also invited to roll up their sleeves during a series of interactive workshops led by NYC faculty.

Friday, September 25, at 9 a.m., NYU Skirball Center for the Performing Arts

hackNY Fall Hackathon
Every spring and fall, hundreds of students from scores of universities around the country flock to hackNY’s Student Hackathons, where they participate in collaborative and creative coding challenges in a 24-hour coding sprint. The events open with API demos from New York City startups selected by the student organizing committee, after which students work in teams to build projects. Students work around the clock, and the event culminates in a presentation before a panel of judges the next day. hackNY is excited to announce its 12th student hackathon from September 26 and 27 at New York University.

Saturday-Sunday, September 26-27

PyData NYC
PyData conferences are a gathering of users and developers of data analysis tools in Python. The goals are to provide a place to share ideas and learn from each other about how best to apply the Python language and tools to ever-evolving challenges in the realm of data management, processing, analytics, and visualization.

Monday-Wednesday, November 9-11, at Bank of America Merrill Lynch, 250 Vesey Street

Deadlines
UrbanGIS 2015
The First International Workshop on Smart Cities and Urban Analytics (UrbanGIS 2015), in conjunction with ACM SIGSPATIAL 2015, has announced a call for papers ahead of its November 3, 2015, workshop in Seattle, WA.

Paper submission deadline: Tuesday, September 1

Grants for Social Science Meta Analysis and Research Transparency (SSMART) Berkeley Initiative for Transparency in the Social Sciences
The Berkeley Initiative for Transparency in the Social Sciences (BITSS) invites academic researchers to submit proposals to the 2015 Grants for Social Science Meta-Analysis and Research Transparency (SSMART). The main objective of the SSMART grants is to support and encourage important meta-research in the social sciences, with the ultimate goal of strengthening reliability and validity of social science research findings.

Application Deadline: Sunday, September 6
Vizzies Visualization Challenge
Do you love animating data, creating science apps, or taking macrophotographs? In the Visualization Challenge, sponsored by the National Science Foundation and Popular Science, your handiwork can receive its due glory and win you cash prizes. Contest winners will be announced in February 2016, and will be featured in the March issue of Popular Science, on popsci.com and nsf.gov/news/vizzies.

Deadline for entries is Tuesday September 15.

CDS News
Political Polarization on Twitter Depends on the Issue
Association for Psychological Science from August 27, 2015
Twitter offers a public platform for people to post and share all sorts of content, from the serious to the ridiculous. While people tend to share political information with those who have similar ideological preferences, new research from NYU’s Social Media and Political Participation Lab demonstrates that Twitter is more than just an “echo chamber.”

This is an illustration of various political symbols coming out of a megaphone. The research is published in Psychological Science, a journal of the Association for Psychological Science.

A Glimpse into the Future of Deep Learning Hardware
The Platform from August 25, 2015
While many recognize Yann LeCun as the inventor of convolutional neural networks, the momentum of which has ignited artificial intelligence at companies like Google, Facebook, and beyond, LeCun has not been strictly rooted in algorithms. Like others who have developed completely new approaches to computing, he has an extensive background in hardware, specifically, chip design, and this recognition of specialization of hardware, movement of data around complex problems, and ultimately, core performance, has proven handy.

Click here to receive the NYU Data Science Community Newsletter 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 and brad.stenger@nyu.edu by 9 pm Eastern time on Thursday evenings for inclusion in Friday’s newsletter. We retain curatorial discretion.

NYU Data Science Community Newsletter Issue 016.