We live in an age when everyone understands the importance and pervasiveness of data—the pictures we take, the activities we track, the transactions we make, and the websites we visit. Data can often be contentious in terms of ownership and access. Today, businesses must deal with the availability, pervasiveness, control, and usage of data for their day-to-day workings. Now data and related aspects touch everyone’s life in one way or another. This portends a sea-change that needs to occur in the way we view technology jobs.
As an increasing number of jobs focus on data and require data analysis capabilities, we need to train a new data-capable workforce to respond to this demand. There is a need to train a diverse population to fill these jobs, and create an inclusive and equitable playing field that nurtures and retains that population.

As two women with careers that intersect with the burgeoning landscape of data science, we are frequently among the few women in the room. This experience hints at some of the challenges women face in data science, and in computing fields more broadly. Some of the challenges we face as women are not unique to computing. For example, women tend to be interrupted more than men (https://books.google.com/books?hl=en&lr=&id=TXQ8DwAAQBAJ&oi=fnd&pg=PA231&dq=women%20interrupted%20more%20frequently&ots=CHWEGAkA9y&sig=MocRjjV4jfs1v99sE94JmVUNtQ#v=onepage&q=women%20interrupted%20more%20frequently&f=false); women are viewed as less credible, less expert (https://journals.sagepub.com/doi/abs/10.2466/pr0.1984.55.2.483) and less competent (https://www.pnas.org/content/early/2012/09/14/1211286109#aff-1) than men; and despite speaking less in mixed-gender settings (https://timedotcom.files.wordpress.com/2017/06/d3375-genderandlanguageintheworkplace.pdf), women are perceived as dominating the conversation (https://pure.mpg.de/rest/items/item_68785_7/component/file_506904/content).

However, many of these challenges may be more pronounced in computing fields. Additionally, stereotypes (both explicit and implicit) exist regarding women’s abilities and perceptions of their belonging in computing fields, which have been demonstrated to impair performance in testing situations. (https://www.ncbi.nlm.nih.gov/pubmed/19656335) Some of these challenges are likely shared across multiple underrepresented minority groups. Although the field of data science appears on a surface level to be more diverse than computing, given its inherent connections to multiple disciplines, we believe that this is a misconception based on our experiences attending and participating in data science conferences and workshops. Despite the perceived diversity in data science, all-male panels and non-diverse program and organizing committees and journal editorial boards are still unfortunately prevalent (https://www.forbes.com/sites/metabrown/2016/08/29/meet-9-women-leading-the-pack-in-data-analytics/#6946297255dd).

Given the expansion happening in the field of data science, and the multidisciplinary views needed to tackle data science problems, the issue of gender diversity—and diversity more broadly—is becoming an imminent problem.

Why should we care about diversity, you might ask? Diversity is important, not only because it is socially just, but also because there is a business case for it. According to some reports, (e.g. McKinsey 2015), diversity in executive management teams has been associated with higher profit margins (https://www.mckinsey.com/business-functions/organization/our-insights/why-diversity-matters).

We not only need people who can discover and execute data science algorithms, but also people who can think of data in different ways and manage data enterprises differently. One possibility is that a diverse set of views can lead to innovative new ideas and products, such as those offered by the company ThirdLove (https://www.businesswire.com/news/home/20181018005166/en/ThirdLove-Named-Forbes%E2%80%99-Billion-Dollar-Startups-List), which has been named in Forbes’ Next Billion-Dollar Startups List for 2018 (https://www.businesswire.com/news/home/20181018005166/en/ThirdLove-Named-Forbes%E2%80%99-Billion-Dollar-Startups-List) and incorporates personal survey data into fit recommendations for women’s bras, or the collaboration between ClickMedix and DC Greens (https://www.forbes.com/sites/ashoka/2016/06/28/when-female-social-entrepreneurs-
that created a mobile platform for low-income residents to redeem vouchers for fruits and vegetables at farmers markets. This project is also helping farmers set up shop in locations where data patterns reveal food deserts to increase access to healthy produce.

In addition to such examples, which unfortunately are rare, there is a great need for people to grapple with the deeper questions and processes that need to be resolved prior to the data crunching. Given the complexity of these issues, greater diversity in teams can only be beneficial to solving some of the pernicious problems arising in data science.

The pervasiveness of data has also made data science an area of study that crosses into many silos. Most data studies require data science plus the “X” factor, where X comes from any of the fields where the data are being generated and where research questions may arise. Indeed a lot of CS+X programs are popping up (e.g., at Stanford University; at University of Illinois, Urbana-Champaign; at Caltech) not just in education but also research. This opens opportunities for the influx of women from other areas that have traditionally been more diverse than computing and information systems.

Opening up these pathways for cross-pollination between computing and other fields can broaden participation in data science. This could encourage more multidisciplinary, humanistic, and even empathetic thinking, where we start to view data—not just as a resource to test an algorithm, or gold nuggets, or the new oil—but also as people, because, increasingly, data capture deeply personal aspects of real individuals. There is also an opportunity for data scientists to work in more team-based settings where the impact from the entire team is bigger than the sum of its parts. It is a well-known phenomenon that women constitute profitable components of team-based environments and important and although it may not be possible to discern whether individual characteristics play a role, some suggest that female scientists are more likely to engage in interdisciplinary research projects.

A call to action

To foster broader engagement of women in data science, not only do we still need to fight the stereotypes we see in everyday society and particularly in computing fields, but we also have to open up paths for cross-pollination between computing and other scientific fields driving data science studies. Some such fields (for example biology, humanities and arts), which have historically been more diverse than computing fields, may be inherently more conducive for broadening participation of women and minorities. Areas of study such as CS+X, where X might refer to an area within the humanities, to biology, or to human-centered computing, should be leveraged to expand the horizon of data science work.

Broad systemic change through peer mentoring and other supportive approaches should be explored as well. A lack of sense of belonging can be harmful. Peer support through local and even virtual networks may mitigate some of these feelings, in addition to providing instrumental support in tackling technical and experiential challenges. Systemic changes in culture need to be made at each step of
the way, starting in infancy and early childhood when gender socialization and stereotypes about
ability begin to be instilled, such that by age 6, girls believe they are intellectually inferior to boys.
Yet another example from Girls who code is that roughly 74% girls express interest in STEM fields in middle school,
but only about 0.4% take up computer science courses in high school. This is also supported by other pieces of evidence,
for example an NSF study indicates that female students enroll in AP computer science courses at a much lower rate (19%) than male
students (81%).
This trickles down to data professionals where only 26% of data professionals are women.
Thus, culture change is also needed during secondary and post-secondary education.

At the faculty level, the trends indicate a more difficult scenario for women and minorities
than for individuals from majority groups. Although women and minorities join in similar numbers at
the assistant professor rank, they often are not promoted at similar levels to full professor. Microaggressions
are often experienced by women and minority faculty throughout their careers, even after tenure. In addition, intersectionality of factors creates
complex challenges for individuals with multiple minority criteria (e.g., women of color). These need
to be addressed through faculty support and retention efforts.

Multidisciplinary data science teams can be leveraged as a new tool to expand participation.
Departmental units and workplaces can increase diversity and provide environments conducive to
individual and departmental growth by creating partnerships and networks across disciplines that
provide support structures. Since data science is increasingly used as a tool to extract meaning from
data in multiple scientific disciplines, teams can be built that capture diversity from these partnering
disciplines to solve not only data science challenges but also diversity challenges. In these teams,
women and underrepresented minorities who join from various scientific domains can share their
expertise, and in the process, perhaps cultivate additional interests and strengths in data science to
bring back to their respective disciplines.

Recommended further reading

1. A Growing Demand for Data Expertise, February 06, 2018, Harvard Business Analytics Program Blog

2. 5 Ways To Support Diversity in Data Science, Meredith M. Lee Posted on October 5, 2017
3. Data Science Education Lags Behind in Diversity. September 25th, 2017, General Assembly


Note: The two authors contributed equally to this piece.

Image: Pixabay

Source URL:
https://www.aaaspolicyfellowships.org/blog/increasing-participation-women-computing-fields-data-science-perspective

List of links present in page

- https://books.google.com/books?hl=en&lr=&id=TXQ8DwAAQBAJ&oi=fnd&pg=PA231&dq=women%20interrupted%20more%20frequently&ots=CHWEgAKA9y&sig=MocRj7V4jI5v99sE94JmVUnTQ#v=onepage&q=women%20interrupted%20more%20frequently&f=false
- https://journals.sagepub.com/doi/abs/10.2466/pr0.1984.55.2.483
- https://www.pnas.org/content/early/2012/09/14/1211286109#aff-1
- https://timedotcom.files.wordpress.com/2017/06/d3375-genderandlanguagetheworkplace.pdf
- https://pure.mpg.de/rest/items/item_68785_7/component/file_506904/content