

If You're Serious About Data Scientists, There Are Ways to Get Them
Thomas H. Davenport

One of the greatest difficulties that companies wishing to become more analytical have encountered over the last several years is finding good analysts and data scientists. A considerable amount of printer's ink has been spilled over this issue. Many of them mention consultants' or analyst firms' projections about how many quantitative analysts or data scientists would be needed in our society, and conclude that it will be incredibly difficult to find them.

I always thought this issue was a bit overblown, and I didn't find too many organizations that had huge problems with hiring unless they are based in Dubuque. But now I have concluded that if you are serious about analytics people, there are ways to ensure that they are available within your organization when they are needed.

The two main approaches to this issue, of course, are hiring people in the external labor market, and training or retraining them yourself. The former approach is much more common, though I'm not sure it should be. Perhaps the best exemplar of this approach is GE, which set out four years ago to hire four hundred or so data scientists. I don't know the exact number they've hired (Bill Ruh, the Chief Digital Officer of GE, told me a couple of years ago they had hired more than 200), but by all accounts the program has been successful. It wasn't easy, of course—GE hired several recruiters who specialized in data science recruiting, the company revised its compensation guidelines for people in the role, and it even created a set of fun TV ads (see one of the famous "Owen" ads [here](#)) to publicize its digital/industrial orientation.

GE's apparent success to the contrary, I am convinced that many companies—particularly those that already employ a lot of technical people—could retrain substantial numbers of people to become data scientists. For this approach, the best example I have seen by far is Cisco Systems. The telecommunications equipment maker has been expanding for several years into advanced services that analyze the data thrown off by devices like routers and switches. In addition, Cisco has been using analytics extensively for internal purposes, such as sales propensity modeling and demand/production forecasting.

However, these analytics demand processes were less the occasion for Cisco's retraining approach than supply-side issues. Desmond Murray, who is a Senior Director for IT at Cisco, was running Enterprise Data and Data Security for the company in 2012. His team was adopting new big data technologies (Hadoop, SAP HANA, etc.) for the company to use, but demand within the business was limited. He concluded that a set of educational offerings around data science would build awareness and stoke demand for these new technologies.

Murray began to think and talk about an education initiative, and he didn't do it on a small scale. He spec'd-out a distance education (obviously, given Cisco's business) program on data science, initially with the Information School of the University of Washington, and later with North Carolina State. The program would last for six months (it now lasts nine) and would conclude with a certificate in data science from the university. Students attended virtual classes a couple of nights a week, and of course had homework as well. Cisco started by enrolling 28 people from all around the organization in the course, and is now on the fifth cohort with 40 students in each. More than 200 data scientists have been trained and certified, and are now based in a variety of different functions and business units at Cisco.

But Murray, by now head of the Enterprise Data Science Office within the IT organization, didn't stop there. He realized that the newly-trained data scientists needed some "air cover" from their managers if they were going to be satisfied in their new roles. So Cisco also created a two-day executive program led by business school professors on what analytics and data science are and how they are typically applied to business problems. The program also covers how to manage a workforce that includes data scientists, and how to know whether their products are any good.

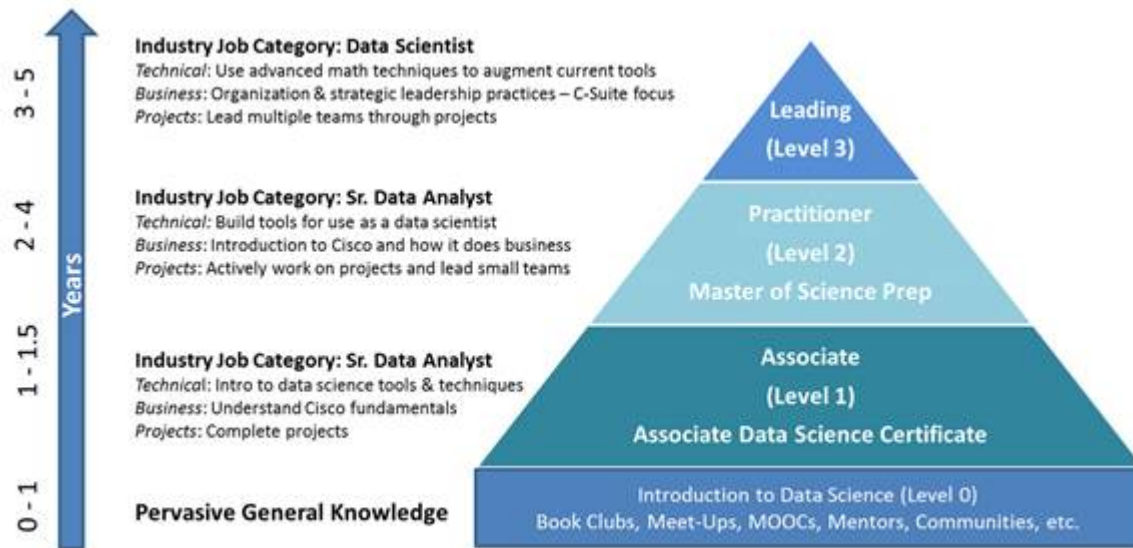
Murray and the Enterprise Data Science team also began to reach out to universities to begin attracting the best data science-oriented graduates. They made an important observation that rather than dealing with universities as a whole, the knowledge of students is resident in individual faculty. So Murray and his team identified the faculty they wanted to work with, told them about Cisco's needs, and supplied them with data that students could use in course projects. They're now working with faculty at ten different universities in N. America, and are planning to expand the outreach to Asian faculty.

One other problem that the Enterprise Data Science Office identified is that there was no standard at Cisco—or at many other firms, for that matter—for who is a serious data scientist and who isn't. So they created a "Pyramid of Analytic Knowledge" (see graphic below) to classify different levels of expertise and establish a career track. Murray and his successor have been working with Cisco's HR organization to incorporate these into official job classifications and compensation bands.

Murray has moved on to another job in IT—he's now running IT for Cisco's supply chain—but he's been succeeded by Kristen Burton. They've moved the reporting relationship for Enterprise Data Science out of IT and into Centralized Business Operations. This move recognizes that data science isn't just about IT, and is increasingly at the core of how Cisco does its work.

Cisco's initiatives in this area are unusual, but they don't have to be. Any company that is serious about analytics and data science could undertake similar steps. The only mystery is why more companies haven't done so. As with the weather, a lot of

people talk about the scarcity and importance of data scientists, but not many do anything about it.



Source: Cisco, August 2015