

Author: Megan Balcom

As technology becomes more accessible and utilized by people around the world, a variety of industries have decided to capitalize on this technological revolution in the form of data analysis. A common way this data is collected is via websites and mobile apps used by customers of entities within these industries. Customers can do any task ranging from trading stocks to paying an electric bill to ordering groceries all from the palm of their hand anywhere. While these technology advances obviously benefit customers, industries have also realized the value of the data collected via these services. For instance, if Amazon can track a user's click stream history during a session on its website, Amazon can analyze how long a customer was viewing any given item, and learn what items that customer viewed before reaching another. This practice of collecting data about data is commonly referred to as big data. By collecting and analyzing big data, Amazon can alter its marketing strategy to suggest particular items that a customer viewed multiple times or for an extended amount of time.

The beneficial impact that big data analysis can have on an organization's attempt to sell more to its customers is quite obvious. Some data experts speculate that as much as 90 percent of the data that exists today did not exist two years ago – because of this explosion in data there are jobs in high demand of people with the skills to assess this data. These jobs span a wide variety of industries, such as healthcare, finance, government, and telecommunications to name a few.

In order to work with big data successfully, one should have some education in statistics and mathematics, and in-depth knowledge of artificial intelligence, machine learning, and data theory, as well as a background in business fundamentals. Experience with databases, and tools

like SAS, Tableau, and MatLab are also helpful. Experience with the programming languages R and Python would certainly also be beneficial to have.

Three major disciplines of careers in big data include the following: data mining, data science, and data analytics. While these all sound similar, there are differences in the nature of work amongst them. Data mining as a whole is concerned with discovering relations in data, not necessarily to hypothesize possible predictions nor produce explanations for those relations. Typical goals of a data miner might be to uncover clustering or to discover outliers in the data set. On the other hand, the field of data analytics uses tools such as SQL and regular query expressions, as well as Business Intelligence packages, and requires some intermediate knowledge of statistics. Conversely, data science requires advance statistics, programming skills, and knowledge of how to acquire, move, and manipulate data. Data scientists have also been referred to jokingly as "someone who is better at statistics than any software engineer and better at software engineering than any statistician."

While these three disciplines are slightly different in their end goals, common traits of successful professionals across all of these disciplines include the ability to handle failure, being curiosity driven, communication skills (both written and verbal), an aptitude at recognizing patterns, working well with a team, and a sense of creativity.

Some common job titles found in the big data field include: Big Data Analytics Consultant, Big Data Engineer, Metrics and Analytics Specialist, Business Intelligence and Analytics Consultant, and Big Data Solution Architect. Even within the three disciplines focused on in this paper, there are more subsets within each one, for example, depending on the data environment there may exist three types of data analytics: predictive analytics, prescriptive analytics, or descriptive analytics.

As far as salary is concerned, working with data can pay very well. According to one source, in the United States the entry-level data analyst salary is \$52,188 per year. Particular skills such as data modeling, SAS, R, and knowledge of data mining and data warehouses are helpful in increasing salary. According to another source however, the national average for a data mining analyst is \$66,594 per year, with the average in Louisville, KY being \$61,806 per year. For an entry-level data scientist in the United States, the average salary is \$85,975 per year – skills and experience with tools such as Apache Spark, Apache Hive, and statistical analysis affect salaries. In Louisville, KY the average data scientist salary is \$63,000. For data scientists in particular, there will continue to be considerable growth in the field, which bodes well for job security. In fact, according to one source by 2018 there will be a projected shortage of 190,000 data scientists in the United States alone. As with most jobs, location and level of experience can impact wages.

In the field of big data, like a large portion of the tech industry right now, women are still a minority of the workforce, with only about 18% of data scientists being women, however for data analysts the balance is much closer with 41% of analysts being women.

As mentioned previously, these jobs exist in almost every industry and all over the world. A local politician from Rhode Island might be interested in analyzing data from online polls in his or her district as he or she runs for reelection and wants to optimize campaign resources and efforts. Or, a multinational corporation such as the big four accounting firm PwC may want a data scientist to help formulate models for their forensic accounting department. A person wanting to work with data will have no shortage of job demand domestically or internationally.

Within the Louisville market, one of the major employers hiring for talent in these data disciplines is Humana. There are several other healthcare companies hiring as well, including Hosparus and KentuckyOne Health. Another big name hiring for data mining talent in the

surrounding Louisville area is Amazon, both its Jeffersonville, IN office and Shepherdsville, KY office have job postings right now for financial analysts, which requires data mining. While there are other smaller companies hiring as well, there are definitely more positions available in nearby cities such as Nashville, TN and Indianapolis, IN. For any given job, particular certifications may be preferred depending on the nature of the job, as well as the industry, however they are not always required.

Outside of this region, the variety of jobs available in this field is extremely diverse. California in particular is famous for Silicon Valley, which is home to many major names in technology, such as Facebook and Amazon, but also home to many startups that use big data in a variety of ways depending on the startup's specialization. Washington DC also has numerous job postings due to the nature of various government intelligence agencies located in that area.

A general career path of someone working in big data may go something along the lines of starting out working with databases in some capacity, then move into a role with data warehouses and possibly business intelligence tools. Next, one might move into data mining, and from there go into a predictive analysis role. This is only one example however, and would also vary based on particular industry requirements.

In the long run, people who work with data in these three disciplines also have the ability to move up after some years of experience in their respective industry. Some data analysts may work their way up to being a chief information officer, where they would construct policy for data analysts, in addition to supervising programmers or other related workers. Becoming a postsecondary educator is another possibility for those who enjoy teaching newcomers to the field of big data about data analysis, mining, and science.

References

<http://www.bls.gov/careeroutlook/2013/fall/art01.pdf>

<http://www.kdnuggets.com/2014/06/data-science-skills-business-problems.html>

<https://www.import.io/post/data-scientists-vs-data-analysts-why-the-distinction-matters/>

<http://www.kdnuggets.com/jobs/index.html>

<http://www.edureka.co/blog/10-reasons-why-big-data-analytics-is-the-best-career-move>