

# The Burtch Works Study

## *Salaries of Data Scientists*

*May 2017*

**Linda Burtch**  
Managing Director



**Burtch Works**  
Executive Recruiting

## Defining Data Scientists

---

There are numerous definitions of what a data scientist is, and we thought it appropriate to share our definition up front, in order to provide clarity and transparency throughout this report.

A data scientist is a specific type of predictive analytics professional who applies sophisticated quantitative and computer science skills to both structure and analyze massive stores or continuous streams of unstructured data, with the intent to derive insights and prescribe action.

The depth and breadth of data scientists' coding skills distinguishes them from other predictive analytics professionals and allows them to exploit data regardless of its source, size, or format. Through the use of one or more general-purpose coding languages and data infrastructures, data scientists can tackle problems that are made very difficult by the size and disorganization of the data.

For more information about how we identify data scientists, see **Identifying Data Scientists** on pages 35.

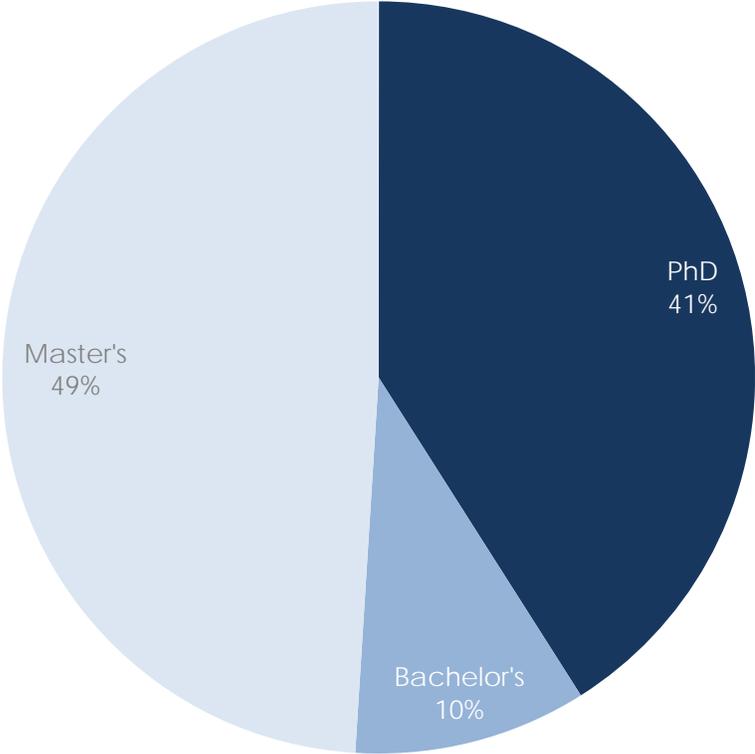
# Demographics & Compensation | Education

---

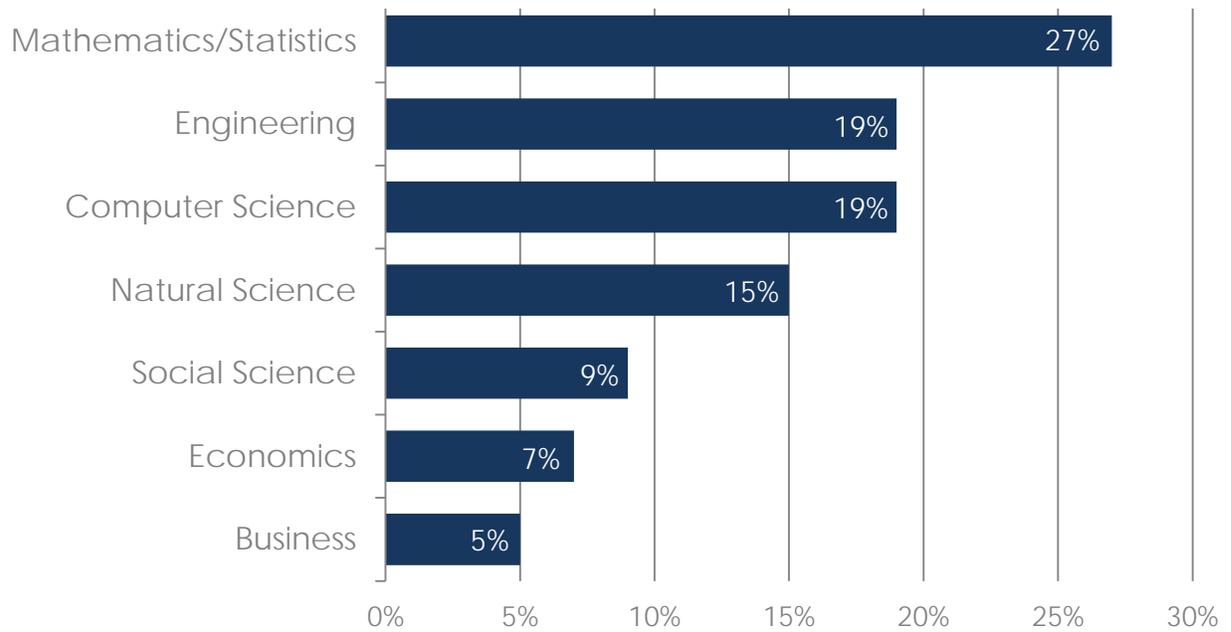
- 90% of data scientists have an advanced degree: 49% hold a Master's degree, and 41% hold a PhD.
- 27% of data scientists hold a degree in statistics or mathematics, while 19% have an engineering degree, and an additional 19% hold a computer science degree.
- In every job category, data scientists who have a PhD earn median base salaries higher than those with only a Master's degree.
- Although the overall number of PhD holders among data scientists declined this year, from 48% to 41%, data scientists are still much more likely to have a PhD when compared to predictive analytics professionals, of whom only 17% have a doctorate.

Figure 8 Distribution of Data Scientists by Education

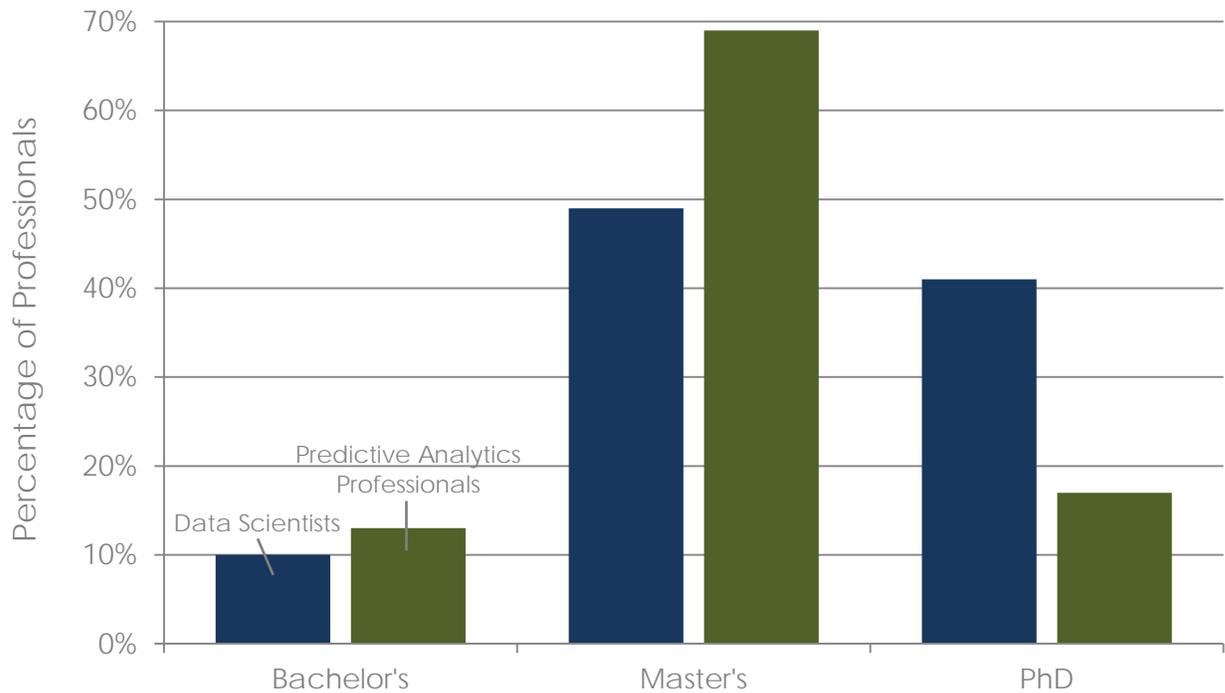
---



**Figure 9** Distribution of Data Scientists by Area of Study



**Figure 10** Distribution of Data Scientists & Other Predictive Analytics Professionals by Education



**Figure 11** Distribution of Base Salaries of Individual Contributors by Job Level & Education

Job Level	Education	Base Salary			
		25%	Median	Mean	75%
Individual Contributor Level 1	Master's	\$78,000	\$90,000	\$91,245	\$100,000
	PhD	\$92,620	\$100,000	\$101,250	\$110,000
Individual Contributor Level 2	Master's	\$110,000	\$120,000	\$123,908	\$140,000
	PhD	\$120,000	\$128,500	\$129,429	\$142,200
Individual Contributor Level 3	Master's	\$150,000	\$155,000	\$165,640	\$175,000
	PhD	\$146,500	\$165,000	\$164,148	\$181,000

**Figure 12** Distribution of Base Salaries of Managers by Job Level & Education

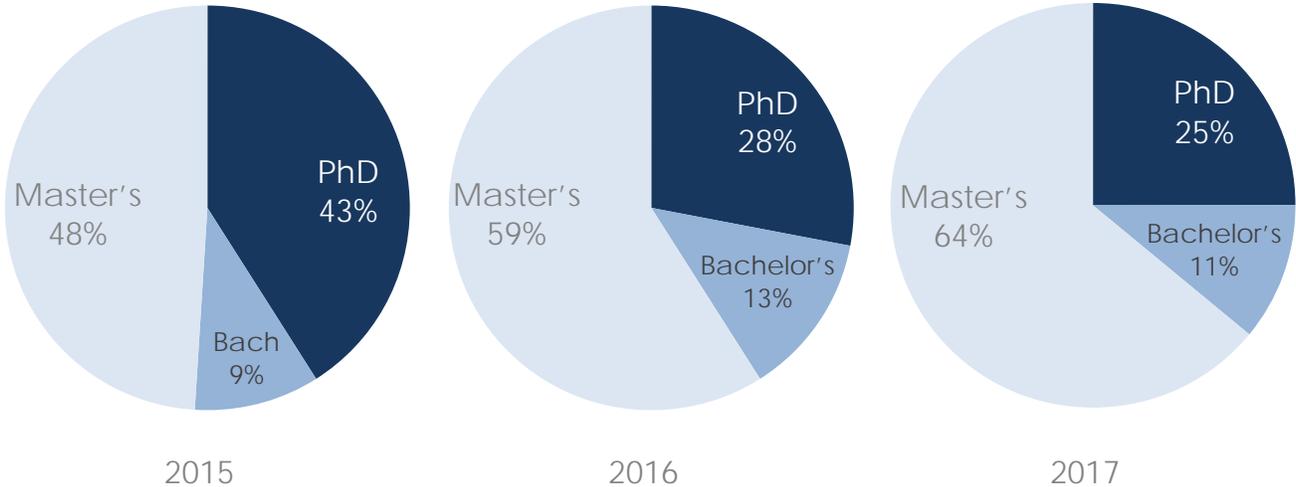
Job Level	Education	Base Salary			
		25%	Median	Mean	75%
Manager Level 1	Master's	\$130,000	\$145,000	\$141,790	\$155,000
	PhD	\$140,000	\$156,200	\$153,139	\$163,800
Manager Level 2	Master's	\$166,200	\$182,500	\$188,342	\$210,000
	PhD	\$175,000	\$190,000	\$194,634	\$210,000
Manager Level 3	Master's	\$223,800	\$240,000	\$262,286	\$263,500
	PhD	\$247,500	\$252,000	\$282,700	\$281,200

# Deeper Dive | Education of Early-Career Data Scientists

Last year, we provided an in-depth look into the educational backgrounds of early-career data scientists (level 1 individual contributors), which showed a trend away from a large number of young data scientists holding a PhD, in favor of Master's degrees. In this year's data, that trend continues.

- This year, 64% of level 1 individual contributors' highest level of education is a Master's degree, up from 59% last year.
- The proportion of PhD holders decreased again as well, from 28% in the 2016 report, to 25% this year.
- As we stated last year, the data suggests that aspiring data scientists are choosing terminal Master's degree programs as, perhaps, a faster route to the workplace.

Figure 13 Comparison of Education of Level 1 Individual Contributors



## How Changes in Compensation Were Measured

---

While some of the 412 data scientists in this sample were also in the samples for our previous studies (published in 2014, 2015, and 2016), others were not. Therefore, changes in compensation were *not* measured by differencing current compensation and compensation reported for the previous study and then taking medians (and other percentiles) of the differences. Instead, changes were measured by comparing medians (and other percentiles) of current compensation to those reported in last year's study.

## Identifying Data Scientists

---

Data scientists apply sophisticated quantitative and computer science skills to both structure and analyze massive stores or continuous streams of unstructured data, with the intent to derive insights and prescribe action. The depth and breadth of their coding skills distinguishes them from other predictive analytics professionals, and allows them to exploit data regardless of its source, size, or format. Through the use of one or more general-purpose coding languages and data infrastructures, data scientists can tackle problems that are made very difficult by the size and disorganization of the data.

To identify data scientists, Burtch Works uses the following criteria:

- **Educational Background** – Data scientists typically have an advanced degree, such as a Master's or PhD, in a quantitative discipline, such as Applied Mathematics, Statistics, Computer Science, Engineering, Economics, or Operations Research. New educational options include data science degree programs, MOOCs (massive open online courses), and bootcamps which continue to take hold in the quantitative community. Some professionals from related careers or fields of study have successfully pivoted into entry-level data science roles through premier bootcamps and mid-career Master's programs.
- **Skills** – Data scientists have expert knowledge of statistical and machine learning methods using tools such as Python and R, with predictive analytics still at the core of the discipline. Data scientists are usually proficient users of relational databases such as SQL, Big Data infrastructures like Hadoop and Spark, related tools like Pig and Hive, and, frequently, AWS. Apache Spark has quickly gained the traction to rival Hadoop in the data science toolbox. Data scientists may use languages such as Python, Java, and Scala (among others) to write programs to wrangle and manage data, automate analysis, and, at times, build these functions into production level code for SaaS companies. Many also use other methods to derive useful information from data, including pattern recognition, signal processing, and visualization.
- **Dataset Size** – Data scientists typically work with datasets that are measured in gigabytes or larger increments, usually too large to be housed in local memory, and may work with continuously streaming data.
- **Job Responsibilities** – Data scientists are equipped to work on every stage of the analytics life cycle which includes:

*Data Acquisition* – This may involve scraping data, interfacing with APIs, querying relational and non-relational databases, building ETL pipelines, or defining strategy in relation to what data to pursue.

*Data Cleaning/Transformation* – This may involve parsing and aggregating messy, incomplete, and unstructured data sources to produce datasets that can be used in analytics and/or predictive modeling.

*Analytics* – This involves statistical and machine learning-based modeling in order to understand, describe, or predict patterns in the data.

*Prescribing Actions* – This involves interpreting analytical results through the lens of business priorities, and using data-driven insights to inform strategy.

*Programming/Automation* – In many cases, data scientists are also responsible for creating libraries and utilities to operationalize or simplify various stages of this process. Often, they will contribute production-level code for a firm's data products.

Professionals whose jobs are described as predictive analytics, analytics management, business intelligence, and operations research were not classified as data scientists. This is because they either do not work with exceptionally large datasets, or do not work with unstructured data. In the specific case of operations researchers, their function is to optimize well-described processes rather than predict and prescribe insights towards more nebulous problems like customer behavior. Predictive analytics professionals were the subject of their own study, *The Burtch Works Study: Salaries of Predictive Analytics Professionals*, released in September 2016.

## Completeness & Age of Data

---

A data scientist was included in the sample only if Burtch Works has complete data about their compensation, and demographic and job characteristics.

All of the 412 data scientists in the sample were interviewed over the 12-month period ending March 2017, which is the year immediately following the period of interviews for the 2016 study. All were interviewed by Burtch Works recruiters while executing searches for clients.



# Burtch Works

## Executive Recruiting

Your source for data science opportunities, talent, and industry trends.

[info@burtchworks.com](mailto:info@burtchworks.com) | 847-440-8555 | [www.burtchworks.com/study](http://www.burtchworks.com/study)