# Data Science Careers and Additional

*19*

Case Studies

Learning Objectives

* Know about careers in Data Science
* Analyze some additional data analytics case studies

### DATA SCIENCE CAREERS

A data scientist works through the entire BIDM cycle (see Figure 1.1) to help grow the business. A good data scientist knows how to identify and focus on high-priority problems. He/she finds the right angle to approach the problem, dives into the data, and seeks new insights and models that the data may throw up that can solve the problem.

A data scientist works through the entire data processing chain from gathering, preparing, modeling, warehousing, analyzing, and visualizing the data for easy consumption by business experts. An ideal data scientist is a multidisciplinary person, persistent in pursuing the solution, and well-versed in business management, databases, statistics, artificial intelligence, and communication skills. The data scientists find bliss in immersing into the data and finding patterns that ultimately align with the unified field of all the laws of nature.

Data science is an emerging and growing field. A useful parallel is to compare its rise with the rise of computer science. Just like computer science took a couple of decades before it became mainstream, so also data science is becoming mainstream now. Computer science had its first course offerings in the early 1960s. At this moment, every university is already offering or developing, a data science curriculum.

* + A data science major takes a wide range of jobs such as Data Engineer, Data Analyst, Data Visualizer, Data Warehousing specialist, Machine learning specialist, and so on.
  + Data scientists are earning high salaries.
  + Fortunately, many online data science courses are available for just in time learning. This field evolves very rapidly, so one should continue to enhance one’s skills.

### DATA SCIENCE APTITUDE

There is no one defined skill for a data scientist. The real test is how strongly you want to do data science. Are you willing to put your heart and effort into becoming good at it?

The first self-test is your relationship with data. Reflect upon yourself and see how much you love working with data to solve problems or develop insights? What have you done with data lately? Collect a small dataset and start playing with it.

Analyze and visualize a new dataset. See what kind of insights you can generate out of the data. Can you visualize and communicate those insights? Does that process excite you? Visit kaggle.com to see free datasets, and find out the kind of data science challenges being offered for prize money.

Data preparation takes about 70-80 percent of the time of a data science project. It can be tedious work but very critical. Without good data, one gets GIGO (garbage in garbage out). It requires patience and attention to detail.

Executives will not trust the data-based insights immediately. They will ask for the data to be looked at in many more ways before they will accept the results. Curiosity, persistence, and creativity are needed in any work, and more so in data science.

Download R and RStudio. Use it to complete the exercises from the tutorial in the Appendix. Does that process look easy and joyful to you? Do you enjoy programming in R?

### POPULAR SKILLS

Data science is a truly multidisciplinary field.

The popular programming languages for data analytics are open-source languages such as R, Python, and Scala.

There are many other platforms from major providers such as IBM SPSS, SAS, and others. There are also other GUI based platforms such as Weka.

There are many good sources on the web for advice on data science careers and education opportunities. Here is a good one:  [www.mastersindatascience.org/careers/data-](http://www.mastersindatascience.org/careers/data-) scientist/

### CASE STUDIES

Here are some additional case studies to understand how data science helps grow the business.

#### Caselet A1: Marketing Analytics for Indian Telecom

*MTS is a medium-sized telecom company that serves more than 10 million mobile phone customers in the Indian subcontinent. They desired to grow their business by delighting their customers and encouraging increased usage of their mobile and broadband services. They decided to invest in a near real-time data analytics solution to help retain existing customers and suitably upselling those who had reached certain levels of usage with attractive marketing promotions. As a result, the acceptance of their usage and retention promotions increased from an industry average of 3 percent to 6.5 percent.*

*MTS primarily gathers usage data from its prepaid service customers from core IT and network elements for marketing campaigns. This data staging/loading solution provides a business-critical platform for the development of online, close to real-time MTS marketing campaigns for direct one-to-one customer outreach. The MTS data analytics solutions helped them identify customers who had reached a certain level of usage and provide an instant customer gratification offer around that usage.*

*Using tools from HP, SAS, and others, they created a semi-automated, resilient, and high-performance analytics engine that was easily accessible by marketing to enable the rapid roll-out of competitive incentive campaigns. They loaded campaign data loads at 2 gigabits per second and processed it every 30 minutes. Their targeted campaigns became so accurate that they were able to more than double the conversion rate.*

*(Source: *www8.hp.com/h20195/V2/GetPDF.aspx/4AA5-2844ENW.pdf*)*

1. *What data mining would you use for identifying what campaign to pitch to which customers?*
2. *Can the learnings from this case also apply to the credit card industry in your country?*

#### Caselet A2: Banking Industry in India

*A small and new bank ran into financial trouble during the 2008 financial crisis. They had captured customer sentiments through surveys at their branches to assess the overall functioning of the bank. It also analyzed data from 20000 credit card customers to determine their service preferences.*

*All credit and debit transactions for bank accounts, credit cards, and other channels are logged for auditable. These data can be correlated with important other phenomena like salary days, and festive spending, etc., to determine consumer preferences. That in turn can lead to the creation of attractive financial products for those customers.*

*Addressing the rising expectations of customers and improving customer satisfaction is a top priority for Indian firms. This bank analyzed consumer behavior based on channel usage and consumption patterns and segmented consumers and identified potential customers for selling relevant financial products. Data analytics is thus helping this and other banks deliver better service to the customers and also grow their own business profitably.*

*(Source: *[www.sciencedirect.com/science/article/pii/S1877050915005992*)*](http://www.sciencedirect.com/science/article/pii/S1877050915005992))

1. *What kind of a data strategy will be needed for this bank?*
2. *Can data analytics help small banks move faster and survive against big banks?*

#### Caselet A3: Bharti Airtel Uses IBM to Get Closer to Customers

*IBM led data analytics solutions to help Bharti Airtel understand their business and optimize its end-to-end for a complete information-led transformation. They are concerned with the diversity of customers; their reach into the smallest of towns and villages; and the range and scale of their services and usage. Analytics are used to provide the right offer for the customer and help the customer discover services. Personalized offers can be sent directly to mobile phones. Analytics help provides accurate information to more than 1 million agents and over 10,000 call center agents in real-time to sell and service better. Analytics help with revenue assurance, revenue planning, ERP, supply chain, and more, and how to deploy all their resources more intelligently in real-time. Analytics also help identify infrastructure shortfalls and plan investment needs in a prioritized manner.*

*(Source: *[https://www.youtube.com/watch?v=SlktJCYIjHc](http://www.youtube.com/watch?v=SlktJCYIjHc))*)*

1. *What kind of a data analytics goal should the telecom companies pursue?*
2. *Can data analytics help smaller telecoms move faster and survive against bigger telecoms?*

#### Caselet A4: Driverless Cars, Big Data, and Law

*Modern automobiles are mobile computer centers with 1.5 km of cables, more than 50 control units, and the computing power of 20 highly advanced PCs. Now the challenge is that of making mobility more intelligent and more networked together with the IT industry. Automakers have developed technologies to alert drivers that they’re straying into another lane, or following another car too closely, or seem to*

*be getting sleepy. Auto manufacturers are teaming up with computer and legal specialists and politicians to ensure the privacy and security of drivers’ information as cars become increasingly digitalized and data get collected continuously. There is a need to protect the customers against the abuse of their data. I clearly say yes to Big Data, yes to greater security and convenience, but no to paternalism and Big Brother. And as hackers get more sophisticated and nimble, they will pose a threat to the data.*

*Volkswagen is displaying the ‘James 2025’, a study of how a virtual cockpit of the future could look. When the automatic driving mode is activated, the steering wheel and seating position change, and a large central screen give all the vehicle occupants details of any planned driving maneuvers. A second screen with touchpad operation on the console provides infotainment functions. Automated or ‘driverless’ cars must however be developed within a framework establishing the required infrastructure and legal aspects. For instance, there’s a question that who would be liable in an accident involving a highly automated car – the manufacturer or the person monitoring the controls?*

1. *What are the benefits of automated driving? What are the downsides? Could hackers completely take over the car and the driver will be the one to blame for any mishap?*
2. *What other machines in other industries could become similarly automated?*