UPRIGHT PUBLICATIONS
https://upright.pub/index.php/ijrsps/

Original Contribution

Data Science: The Sexiest Job in this Century

Mahesh Babu Pasupuleti [1]

Keywords: Data Science, Data Scientists, Sexiest Job

International Journal of Reciprocal Symmetry and Theoretical Physics

Vol. 2, Issue 1, 2015 [Pages 8-11]

Data science is the study of huge amounts of data in order to uncover insights that may be used to assist companies in making strategic decisions. There are several routes to a job in data science; the majority, but not all, entail a little math, a little science, and a great deal of interest about the subject matter of data. New data scientists must be inquisitive, critical, and persuasive in their reasoning. The research focuses on why data science is regarded the sexiest job in the twenty-first century, as seen by the high compensation paid to qualified people.

INTRODUCTION

Data Science is a process rather than a single event. It is the process of using data to better understand diverse things, and ultimately, the world. When you have a model or hypothesis about an issue and you try to validate that hypothesis or model with evidence, it is what I consider to be "validation." It is the art and science of unearthing the insights and trends that are hidden inside large amounts of data. It is the process of converting data into a narrative. As a result, storytelling may be used to develop understanding. In addition, you may use these insights to make strategic decisions for a firm or a group of institutions. Data science is a discipline that deals with the procedures and technologies that are used to extract information from diverse data sources, whether they are in an unstructured or an organized format. The study of data is referred to as data science. Physical sciences are similar to biological sciences in that they are concerned with the study of biology. It is real, data has genuine properties, and we must understand these aspects if we are to be successful in our endeavors. Data Science is a combination of data and a little science. When some academics were looking at the statistics curriculum in the 1980s and 1990s, they came up with the definition of the word "data science," and they decided it would be more appropriate to call it that. However, what exactly is Data Science? Data science, in my opinion, is an individual's endeavor to work with data in order to find answers to problems that they are studying. Briefly said, it is more about data than it is about science in this case. You can call what you do with data science "data science." If you have information and you have a burning desire to learn, and you're dealing with it, changing it, and studying it, the sheer act of going through the data and attempting to obtain some answers from it is data science (Cleveland & Hafen, 2014). The field of data science is vital now because we have an abundance of data at our disposal. We used to be concerned about a lack of information. We are presently in the midst of a data flood; in the past, we lacked algorithms; now, algorithms are available. Previously, the program was prohibitively expensive; today, however, it is open source and free. Although we couldn't store massive volumes of data in the past, we can now

[1]Student, Masters of Science in Computer Science, Northwestern Polytechnic University, 47671 Westinghouse Dr, Fremont, CA 94539, USA

do so for a fraction of the cost, allowing us to have zillions of datasets at a very cheap cost. In other words, the tools for working with data, the availability of data itself, and the capacity to store and analyze data are all inexpensive, readily available, and widely distributed. It is all here. In the field of data science, there has never been an easier time to make a living.

FUNDAMENTALS OF DATA SCIENCE

No matter who you ask, everyone will have a somewhat different definition of what Data Science is, but the majority of people will agree that it involves a large amount of data analysis. Data analysis is not a new concept. In contrast, the large amount of data accessible from a wide range of sources is new: from log files to email to social media to sales data to medical information databases to sports performance data to sensor data to security cameras, and many other sources in between. Meanwhile, we have the computer capability to do valuable analyses and uncover new information at a time when there is more data available than ever before. Data science may assist companies in better understanding their surroundings, identifying and analyzing current difficulties, and identifying previously undiscovered possibilities. Using data analysis, data scientists contribute to the company's expertise by studying data and determining the most effective method to use it to create value to the organization. So, what exactly is the data science methodology? Many businesses may utilize data science to narrow their attention on a specific problem, thus it's critical to define the question that the company wishes to have addressed before beginning. This is the first and most important phase since it determines how the data science project will proceed. Good data scientists are inquisitive individuals who ask probing questions in order to define the business requirement. What follows are the following questions: "What data do we require to address the problem, and from whence will that data come?" In addition to structured and unstructured data from a variety of sources, data scientists can pick from a variety of approaches to evaluate the data, depending on the nature of the problem they are trying to solve. Exploring data using several models exposes trends and outliers; sometimes this may corroborate what the company already suspects, but other times it can disclose wholly new information, prompting the organization to take a different approach to the problem. When the data has disclosed its insights, the data scientist's function shifts to that of a storyteller, in which he or she communicates the findings to the various project stakeholders. Data scientists may employ strong data visualization tools to assist stakeholders in understanding the nature of the results as well as the suggested course of action to follow (Hazen et al., 2014). Data Science is altering the way we work; it is altering the way we utilize data, and it is altering the way organizations perceive the world.

ADVICE FOR NEW DATA SCIENTISTS

The study advice to a budding data scientist is to be inquisitive, opinionated, and critical. Curiosity is a must-have. You wouldn't know what to do with the info if you weren't intrigued. Because you wouldn't know where to begin if you didn't have predetermined beliefs about things. Argumentative because if you can argue and make a case, you can at least get started, and then you learn from data, adjust your assumptions and hypotheses, and your data will assist you in learning. It's also possible that you'll begin at the incorrect location. You may argue that I thought I believed it, but facts has proven otherwise. As a result, you might engage in a learning process. Curiosity about being able to take a firm stance and then move forward with it. The data scientist should also have some familiarity and flexibility with analytics platforms, like as software and computer platforms, but this is secondary (Reed, 2014). Curiosity and the capacity to take positions are the most crucial factors. You'll have some answers once you've done that, once you've analyzed. And the capacity to tell a story is the very last thing a data scientist requires. That after you have your statistics and tabulations, you should be able to convey a compelling story with them. Because if you don't make a good tale out of it, your discoveries will remain concealed, buried, and nobody will know. Your success is mostly dependent on your ability to deliver compelling stories. Finding out what your competitive edge is would be a good place to start. Do you want to be a data scientist in general or in a particular field? Because, for example, if you want to work as a data scientist for an IT business or a web-based or Internet-based organization, you'll need a distinct set of abilities. And if you want to work as a data scientist in the health business, for example, you'll need a distinct set of talents. So first figure out what you're passionate about and what your competitive edge is. Analytical talents aren't always going to be your competitive edge (Lupia & Elman, 2014). Your competitive advantage is the knowledge of a certain facet of life

that you know better than others. Perhaps it's movies, perhaps it's retail, perhaps it's health, perhaps it's computers. After you've determined your area of expertise, you may begin to develop analytical abilities. What platforms to study on, and the platforms and tools to use, are relevant to the sector in which you're interested? Then, once you've gained some proficiency with the tools, you can apply your knowledge to real-world problems and show the rest of the world what you've learned (Jifa & Zhang, 2014).

THE SEXIEST JOB IN THIS CENTURY

In today's data-driven environment, data scientists have risen to the top of the demand list. The race is on to identify the top data science talent available today. Already, analysts predict that millions of positions in data science may go unfilled due to a scarcity of qualified candidates in the market. The global quest for qualified data scientists is more than just a search for statisticians or computer scientists, as many people believe. In fact, the companies are looking for individuals who are well-rounded, who have subject matter expertise, some experience in software programming and analytics, and exceptional communication skills, among other qualities (Provost & Fawcett, 2013).

Over the past ten years, our digital footprint has grown at an alarming rate. In 1995, the digital world had a total storage capacity of around 130 billion gigabytes. Companies will compete for hundreds of thousands, if not millions, of new employees who will be required to navigate the digital age. It's no surprise that the esteemed Harvard Business Review named data science the most exciting job of the twenty-first century.

According to a research by the McKinsey Global Institute, there is a severe scarcity of data and analytics skills. Approximately 140,000 to 190,000 professionals with strong analytical abilities, as well as 1.2 million managers and analysts with the know-how to apply the analysis of big data to make efficient choices, might be in short supply in the United States alone by 2015.

Because the digital revolution has permeated every part of our life, there is now more potential than ever before to get insight into our habits and actions. The right data can allow marketers to get a sneak peek into our habit formation. Research in neuroscience and psychology is uncovering the mechanisms through which habits and preferences are formed, and businesses such as Target are looking to capitalize on this information. Customers will be able to shop more efficiently if merchants have data scientists working for them. "As a result, hiring statisticians has become something of an arms race in recent years," said Andreas Weigend, a former chief scientist at Amazon.com.

It is still necessary to persuade the C-suite executives of the importance of data and analytics in their organizations. According to the evidence, senior management may be a step or two behind middle management when it comes to being aware of the possibilities of analytics-driven planning. In his role as director of the Customer Analytics Initiative at Wharton, Professor Peter Fader understands that CEOs may reach the C-suite without having to interface with data. He believes that when leaders are well-versed in data and analytics, meaningful transformation will occur in their organizations (Mangilli, 2014).

SAP, a pioneer in data and analytics, released the results of a survey that revealed that 92 percent of the responding companies in its sample had witnessed a considerable rise in their data assets. At the same time, three-quarters of respondents stated that their companies were in need of fresh data science talents. Accenture predicts that demand for data scientists would surpass supply by 250,000 people in the first half of 2015. In 2014, KPMG conducted a similar study of 150 executives and discovered that 85 percent of those who participated did not know how to interpret data. According to Alwin Magimay, head of digital and analytics at KPMG UK, most businesses are unable to connect the dots because they do not completely comprehend how data and analytics may alter their business. Magimay stated this in an interview with the Financial Times in May 2015.

Bernard Marr, writing for Forbes, expresses worry about a scarcity of qualified analytics professionals as well. This is due to a lack of qualified individuals who can evaluate and understand this information, changing it from raw numerical (or other types of) data into actionable insights, which is the ultimate goal of any Big Data-driven endeavor, according to the author of the article. Bernard cites a Gartner survey of corporate leaders, in which more than half of those polled acknowledged a dearth of in-house data science competence.

Bernard wrote about Walmart, which relied on crowdsourcing to meet its data analytics requirements. Walmart approached Kaggle with the idea of hosting a competition for evaluating private data that they had collected. The company gave sales data from a selection of outlets and requested the rivals to generate improved sales estimates based on marketing strategies based on the data they provided (Fox & Hendler, 2014).

The scarcity of data scientists has resulted in organizations being prepared to pay high cash for

top-tier expertise. Michael Chui, a principal at McKinsey & Company, is all too familiar with this. "Data science is now applicable to every type of organization... There is a fierce competition for this sort of ability "In an interview, he explained. Take the case of Paul Minton, for instance. He was earned $20,000 a year waiting tables in a restaurant, according to his earnings statement. He had studied mathematics as a major in college. Mr. Minton completed a three-month programming course that completely transformed his life. In 2014, he earned more than $100,000 working as a data scientist for a San Francisco-based online firm. Starting at six figures, it's an easy sell... Mr Minton described the experience as "astonishing." Is it possible that Mr Minton has been extraordinarily lucky, or are such generous incomes the norm? The New York Times stated that the average base wage of a software engineer was $100,000, while the average base compensation of a data scientist was $112,000, indicating that luck had little role.

CONCLUSION

Using data to comprehend the world is a method. It's when you use data to test a problem model or hypothesis. Data science involves finding hidden patterns and trends. When data tells a tale. So narrate to learn. Then make organizational strategic decisions. Unstructured or structured data can be extracted using data science. Data science. Biochemistry and physics examine physical responses. Data is real, and we need to examine it. It's all about data. The 1980s and 1990s saw the statistics program as data science. But DS? It's data-driven question-answering. Not science, but numbers. Data science is the study of data. There's enough data now to examine. The data was scarce. Algorithms have replaced the data flood. Previously expensive software is now free. We couldn't store enormous volumes of data before, but now we can. Yes. Data tools are cheap, widespread, and easy to use. Data scientists are in a golden age.

REFERENCES

Cleveland, W. S., & Hafen, R. (2014). Divide and recombine (D&R): Data science for large complex data. *Statistical Analysis and Data Mining, 7*(6), 425-433. https://dx.doi.org/10.1002/sam.11242

Fox, P., & Hendler, J. (2014). The science of data science. *Big Data, 2*(2), 68-70. https://dx.doi.org/10.1089/big.2014.0011

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics, 154*, 72-80. https://dx.doi.org/10.1016/j.ijpe.2014.04.018

Jifa, G., & Zhang, L. (2014). Some comments on big data and data science. *Annals of Data Science, 1*(3-4), 283-291. https://dx.doi.org/10.1007/s40745-014-0021-9

Lupia, A., & Elman, C. (2014). Openness in political science: Data access and research transparency: Introduction. *PS, Political Science & Politics, 47*(1), 19-42. https://dx.doi.org/10.1017/S1049096513001716

Mangilli, A. (2014). From data to science: Planck data and the CMB non-gaussianity. *International Astronomical Union.Proceedings of the International Astronomical Union, 10*, 131-134. https://dx.doi.org/10.1017/S1743921314013465

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data, 1*(1), 51-59. https://dx.doi.org/10.1089/big.2013.1508

Reed, D. (2014). Data smart: Using data science to transform information into insight. *Journal of Direct, Data and Digital Marketing Practice, 15*(4), 354-355. https://dx.doi.org/10.1057/dddmp.2014.33

--0--