

10/28 Essay: To give you a better idea of what data science is and how it compares to other areas in the world of work read [this 10-15 min article](#). Taxonomically, in academia, data science is a subfield of computer science. Personally, I see data science as a tool or a set of tools and methods that a researcher or practitioner of data science applies to data in a particular domain. This combination is their "craft", in essence. I, for example, apply data science to geography in the context of international development. This means I use tools such as machine learning on data such as satellite imagery to help solve problems in the international development community. Data science is a domain that can be surprisingly hard to define given its broad applications and newness. I don't necessarily think there's one correct definition and, whatever it is, the field will continue to evolve with new technologies, new industries, and/or new issues. There's also a question of skill ranges. Should the title of "data scientist" be reserved for those with doctoral degrees in the same way it is in physics, for example, and anyone below that is basically something else (like an analyst)?

Think about an area of study or field you care about (outside of your formal assignment topic), maybe the major you intend to choose. Based on what you read and know about data science and its related areas (data analysis, stats, and ML), briefly (1-2 paragraphs) write about how you think these disciplines could possibly be used in your field (it's okay to think ambitiously). If methods are already being used, what are they and to what extent (if you know)? If you are a prospective data science major, what ideas do you have for how data science could be used in ways you believe they are currently not? You have until 10:10.

Medium

What is Data Science?

Each data science project has one of these three goals

Reading time

10 min read

In my opinion, the title of a data scientist should not be reserved just for those with doctoral degrees. In our world today, there are so many different formats and ways people can use data science, in addition to the wide variety of professions it can be applied to. Likewise, there are so many different servers and technologies that one can use to implement or develop algorithms for this data. Data science has such a wide range of applications and implications throughout different fields, so as I see it, it is challenging to categorize it into one subject group. In my opinion, being considered a data scientist should not be based on simply having a doctoral degree - data science is subjective and can apply to many different scenarios

In relation to a field I really care about, I have always really been fascinated by meteorology, the study of the weather and climate. Some of the methods I read about in the article are already being used. Meteorology is all about statistics, probability, and models. As the article explains, the most common technology tools for data analysis are SQL, Excel, Tableau in that order. Programming with Python and R are the next most common. While departments like the Weather Channel or Wunderground may not be using these exact platforms, they are

definitely using similar platforms to make generalizations about the weather. In addition, the article then goes into talking about different statistical methods and values one can use, like Chi-squared tests, ANOVA, linear regression, and logistic regression, p-values, and confidence intervals. These are all different statistical tools I am learning right now in my Business Statistics class, and while watching the weather, I have seen direct implications of these topics.

In contrast, some disciplines like deep learning and machine learning engineering may be useful in fields like Meteorology. For example, machine learning engineers have to ensure models will be continually updated and that they will reliably make fast, high quality predictions at scale. Some models among different weather stations are definitely outdated and could use more updates. This discipline could surely be used in the meteorology field to improve the technology. Likewise, looking towards the future for meteorology, deep learning will become that much more important. As climate change continues to alter weather patterns and increase the size and quantity of storms, we will need to look towards new creation through general adversarial networks. This could include completely changing our current systems, models, and technologies with a new algorithm, or looking towards an entirely new way of thinking with the idea of new creation. This will all depend on how drastically weather patterns change and over what quantity of time. Weather platforms are also working intensively on increasing data gathering via those who are actually being impacted by the weather. Platforms such as the Weather Channel are able to track people through their phones when and where they report damaging weather or a power outage. This will continue to be super important as our world becomes more advanced.