Stephanie Nelson

March 28th, 2023

Professor Bemley

**Title**: “Introduction: What Is Data Science”

**Author**: O’Reilly

I. **Introduction**

In this article, O’Reilly discusses the misconceptions and confusion tied to data science. The article begins by breaking down data science and big data into specific categories. O’Reilly addresses how statistics and statisticans play a role in data science. To clear th confusion around data science, O’Reilly address confusion between learning data science in an academic setting versus in the industry. This article provides you with insight on why data science is emerging and how it is useful to our everyday lives.

II. **Basic Components of Dataset**

This article dives into a couple of concepts such as datafication, data science and visualization. Datafication is “taking all aspects of life and turning them into data,.” (O’Reilly, 2014). This is useful as we learn how to think and approach situations like data scientists. You must first learn that data is all around us and is comprised in everything that we do. The specific definition of data science is the civil engineering of data. To break this down this is the use of resources to create an understanding of data. Visulization is the process of turning data into graphs and tools.

III. **Statistics is the Grammar of Data Science**

Statistics is tied into data science due to statistics being the process of collecting and analyzing large sums of data. According to O’Reilly, statistics and data science are connected due to it being statistics taken a step further. Data science takes large sums of data and then incorporates machine learning, data analysis and other techniques to provide a deliverable. Data Scientists encompass the skills of statisticans, computer scientists, communication professionals and more to be useful members of their teams.

V. **Conclusion**

In all, in this article, Provost and Fawcett introduces the academic approach to data science as well as debunks the media definitions. This begins our journey to learn how to properly think and approach data as data scientists would. The use of diagrams and outside sources help strengthen this argument as well as resonate with the audience.

Vi. **References**

O’Reilly (2014). Introduction: What Is Data Science, <https://bsuonline.blackboard.com/bbcswebdav/pid-17985968-dt-content-rid-28398463_1/xid-28398463_1>