Charles Hwang

Christie Asif, Meg Seth

UNIV 102-001

9 April 2020

Occupational Project Paper

I chose the career field of data science for my occupational project. I first became aware of this career path through coursework in my Predictive Analytics course in which machine learning was an often-discussed topic as part of my statistics major. Data science is one of the fastest growing fields worldwide and is used by companies in nearly every sector. As advancements in technology increase exponentially, companies collect large amounts of data pertaining to their operations, sales, and development, known as "big data" ("SAS", n.d.). Because different people record these data in different ways, they are not uniform and some observations may be missing or neglected. (Many people in the industry often use the term "data" as a plural noun.) Data scientists are needed to "clean" and transform the data and model them so that they can be presented for practical interpretation. Because of this increasing demand for data scientists, the average salary can be higher than many other jobs. According to "Master's in Data Science" (n.d.), the average salary of a data scientist is $118,370 per year.

Because of the nature of the job, there is not a "typical" day in the position of a data scientist. The job description can vary widely, as different companies will look to extract different conclusions from their data depending on their sector, intent, and expected timeline. Additionally, it is difficult to describe to the average person what data scientists do for a living because of the complex nature of the programming and modeling, as well as the statistical vocabulary used that can only be fully understood with further statistical knowledge, which is

why data interpretation is a large part of the role to help others understand how the data are being transformed and what is being analyzed. As described by data scientist Hugo Bowne-Anderson (n.d.), "data scientists lay a solid data foundation in order to perform robust analytics" and use experimental methods like machine learning to optimize certain functions.

Common job qualifications for a data science position include at least bachelor's degree in statistics, data analytics, computer science, mathematics, or a similar field (although a master's degree can be beneficial); considerable prior experience with common statistical programming languages and applications (R and SAS are two universally-used languages, while SQL, SPSS, Tableau, Python, Java, MATLAB, Stata, Maple, and Minitab are also used frequently); and an inclination to statistical and quantitative problem-solving. Entry-level positions usually do not require previous experience other than academic coursework, although some employers may ask for an academic transcript or list of relevant courses taken. Although not explicitly listed, it is very beneficial to have some public speaking skills in order to effectively interpret and explain models for others in practical terms.

The field of data science was incredibly different ten to fifteen years ago—nearly unrecognizable from the field it is today. Computing technology was obviously much more simple back then, and the programming languages like the ones I listed earlier were much less user-friendly and inefficient because they were based on lower-powered software. The field itself was also considered much smaller and less important than it is today. In September 2005, the National Science Board Committee on Programs and Plans wrote a recommendation to the larger federal National Science Foundation that they "should act to develop and mature the career path for data scientists and to ensure that the research enterprise includes a sufficient number of high-quality data scientists" (National Science Foundation, 2005), sparking an increase in interest and

funding to the field that has influenced its continued and consistent growth over time. Being closely related to technology has also brought increasing amounts of attention to the field.

It is difficult to foresee how the field of data science will change within the next five to ten years for several reasons. The size of the field itself is difficult to project because it is unknown how long it will sustain its currently astronomical rate of growth with potential future trends and demands in other career fields. Technology is developing at an exponential pace, with currently unknown technologies possibly being discovered and reshaping not only the field of data science but also society as a whole and what it demands of data scientists. New programming software may be developed to suit these transformations. All else being equal, the one thing that seems certain is that the field will be much larger in the short-term future. How much larger will depend on economic and technological factors.