An Analysis of the Data Science Job Market

A major topic of conversation brought up during this boot camp is where our fellow students are working now and what they would like to do with their newly acquired skills in the future. While many of us are completing this course for our current jobs, it is not surprising that another big group of the students would like to make a career change into the data science field where most of the responsibilities will require the use of the concepts learned in this class. Our group felt that it was not only relevant but significant to share data we found on the data science job market for our project. Collected from job postings categorized under Data Science published on Indeed.com, this data set includes 6,959 available positions across the nation from August 2018. Thorough brainstorming allowed the creation of several pertinent questions we think will be the most informative to the class and will hopefully aid in the process of landing a job post-bootcamp.

We can all agree that location is a very important component when deciding on a job. Another is the company, the kind of work they do and how their employees like working for them. This data set supplies some information on this, as company names and review numbers are included. Again, this data set does not surprise us as the companies with the most job postings on this list include Amazon, Facebook, Google, and Microsoft. However, the top 15 companies also include some other big names like KPMG, several universities (NYU, University of Washington) and research centers (Broad Institute, Lab126). It is important to note here that the size of the company isn't always indicative of more positions available. The numbers here could be due to high turnover of employees at the company, recent growth or expansion in the company, or recent growth of a particular field. A take away here is that people should decide what size company they'd like to work for first before they start sending out applications. It is likely that the big name companies will always be hiring but if one is looking to work for a smaller, growing company they might need to do a little more research. It is also key to do some research into the ownership of the company you're applying to. Not only is Amazon in the number one spot of available positions, but it also has ownership of Lab216 which takes the eighth spot on this list. If you've decided you wouldn't want to work for a big company like Amazon or Google, than you may or may not want to cross off their companies from your list as well.

One of the important parameters to set when searching for a job is location. Location is extremely important when considering a job as it dictates an individual's quality of life, lifestyle, and many other variables. When analyzing the data science field, it is important to figure out where the most available positions are at the moment. This data set included many descriptive columns, one of which includes the city and state where the position is available. It is not a surprise that the majority of the postings in this data set are located in our nation's biggest cities- New York, Seattle, Cambridge, Boston, San Francisco and Chicago. While at first glance the inclusion of several cities in Massachusetts might be a surprise, one must remember the high density of elite universities in the area as well as hospitals and the rise of tech jobs in all major metropolitan areas. San Francisco comes in at number five and it is not a shock that the majority of the cities ranked in the top 30 are located just south in Silicon Valley, which is known as the epicenter of cutting edge technology. An important point to emphasize here is that when

setting parameters for job postings, one might want to search by state or area rather than a specific city. Many companies might be moving out of the downtown areas for more space or cheaper rent prices like we've seen in San Francisco. By searching via a county or mile parameter, more postings will pop up in neighboring cities that still might be considered in the desirable area.

The high number of companies hiring in the bay area led to further analysis of job postings specific to California. Not surprisingly, some of the big names were consistent across both lists. The list contains Google, Walmart and Amazon with some other research or media companies. While many of these companies are located in the bay area, a top five on this list named Cymer is located down in San Diego showing that the tech field is truly expanding. This could also explain why Google replaced Amazon as the biggest company in California as Amazon is also trying to expand outside of this beautiful yet expensive state. This top 15 list contains companies focused on everything from developing microchips, digital media, eCommerce, pharmaceuticals, biotechnology, to e-cigarettes. A takeaway from this analysis is that California has a diverse range of companies looking to hire more data scientists up and down the coast. It is not longer necessary to move to silicon valley to get into this field.

According to further analysis of the descriptions included in the job postings, a significant number of job listings do not specifically call out whether they are looking specifically for an analyst, scientist or engineer. Rather, it seems the large majority of employers use general job titles. This analysis searched by three titles - Data Scientist, Data Analyst, Engineer. Only 17.9% of the postings specified that they were looking for a Data Scientist, 12.6% specified for a Data Analyst, 19.3% specified for an Engineer, and the other 50% did not specify. While this is surprising at first because this data set is described as a compiling of data scientist positions in August 2018, it does make more sense from the companies point of view. By not specifying what type of position it is, the company might get a wider range of applications who have a variety of different skills. People should not shy away from titles that don't contain either of the three words mentioned above but apply anyways as they could have the skills necessary for the job.

The last chunk of information from this data set to be analyzed is the list of skills mentioned throughout the different descriptions. Through individual experience as well as prior research into the field, we created a list of 14 skills to search for. This significance was to not only count how many times these skills were mentioned but to than assume that the higher counts ment that more postings called for the applicants to have these skills. The analysis ranked the skills as follows- Excel, Python, Machine Learning, SQL, Data analysis, Java, C++, Statistical analysis, Data Visualization, Tableau and so on. Individuals can take this list and use it to decide which skills and tools they definitely need to master if they'd like to get a job in this field. This style of analysis was also completed for type of degrees mentioned. Overall, Ph. D was mentioned the most followed by Bachelors and then Masters degrees. This does not mean that one must need a Ph. D to get a great job in this field, but that most postings probably won't consider someone without a Bachelor's degree. It is also important to note that this list probably changes over time as the field evolves and individuals should be aware of what the latest trends are or developments that could alter this list.

The analysis of data science job postings led to the creation of some significant tips on how to navigate this field. Applicants should decide first what type company they'd like to work for whether it be small and growing or large and powerful. Next step is to decide where they'd like to work, whether it be which state or what size city. This field is expanding nationwide and there are cities with many open positions one might not have thought about. The job title to search by is not as important as many postings will not mention analyst or scientist. Most positions in the field are looking for applicants who have mastered Excel and Python, but to stick out one should consider adding a couple more skills to their tool belts. To be a competitive application, one should have at least a Bachelor's degree while a Masters or Phd would help the applicant stick out of the crowd. However prior experience in the field and networking should not be forgotten, these are also very valuable for applicants.