**Data Science in Small and Big Companies**

**A comprehensive comparison of a data scientist’s role as a function of company size**

Data science has become an indispensable part for businesses and has penetrated into almost every domain, from e-commerce to autonomous driving and healthcare. Within each domain, data science teams are working hard to produce valuable, data-driven insights that can subsequently be used to make informed business decisions. The underlying process that governs the derivation of those insights, however, can vary extensively.

When thinking about data science, the concepts that come to mind are typically confined within a narrow scope, such as data exploration, experimentation, model training, and data visualization. However, the role of a data scientist can be extremely versatile, depending on the size of the organization.

This article provides an outline on what it’s like to be a data scientist in a small versus a big company by evaluating a set of criteria including tasks and responsibilities, impact and innovation, learning and skill development, available resources, mentorship opportunities, and upward mobility.

It should be noted that the criteria discussed here will vary from company to company and will also depend on the maturity of the data science division that has been established there.

**Tasks and Responsibilities**

Within smaller companies, the nature of tasks and the scope of responsibilities of a data scientist can vary widely depending on how well-established the current data science division is.

If it has just been created or only consists of a handful of people, you could potentially find yourself as a *jack of all trades* tasked with the building of the entire infrastructure that includes data collection, processing, cleaning, exploration, model training, deployment, monitoring, and so on. In some cases, responsibilities may also include the active involvement in and implementation of data-driven strategies for the company, such as identifying new business opportunities or increasing operational efficiency.

On the upside, the visibility of one’s work is typically much higher in smaller organizations and data scientists tend to enjoy greater autonomy and project ownership. This, however, also comes with increased responsibility and accountability.

When I was working at a small biotech startup comprised of about 30 employees, I was reporting directly to the CTO of the company. This, of course, was a fantastic way to have my work noticed and receive credit for the successful and timely delivery of features. However, expectations from C-suite executives can often be quite high, which increases the pressure a data scientist faces when it comes to delivering results on a tight schedule.

By contrast, the role of a data scientist in larger organizations with a well-established and mature data science division tends to be much more narrow and specialized. For instance, if you join Apple’s Siri team, you will most likely work on a niche area within the field of natural language processing. Likewise, joining a data science division that works on Microsoft Teams will likely require you to become an expert in areas such as noise suppression or anomaly detection.

A major advantage of working in a bigger corporation is the support one receives in the form of collaboration and domain knowledge. Larger companies often have more experienced and diverse teams that can provide valuable support and expertise. Receiving information from a broader range can help data scientists gain a more holistic understanding of the underlying business, which not only provides a great learning opportunity, but also helps with the identification of data-driven solutions that can help solve inefficiencies within the company.

**Impact and Innovation**

In smaller companies, data scientists often have a more noticeable impact on the overall success of the business since their work can directly affect decision-making and strategy. However, due to the lack of readily available resources such as compute power or domain knowledge, data scientists may have to work more creatively and efficiently to produce impactful solutions, thus requiring them to display greater versatility and adaptability.

A considerable advantage of smaller companies is a faster decision-making process. If you have a new and innovative idea, obtaining approval from leadership to implement and execute it tends to happen much faster due to a much smaller organizational hierarchy.

On the other hand, in a big company, data scientists may have less direct impact on business outcomes due to the scale and complexity of the organization. However, they may have access to more resources and infrastructure to store, process, and analyze data, which can enable them to tackle more complex and ambitious projects. Additionally, the increased domain knowledge and expertise found in big companies can create more opportunities for collaboration and innovation, enabling data scientists to work alongside other experts to develop state-of-the-art solutions.

**Learning and Skill Development**

The need for a data scientist to be familiar with the entire data science workflow in a smaller company offers a great opportunity to develop a broad and well-rounded set of skills. This exposure to various aspects of the data science cycle further enables them to determine which areas particularly resonate with them, which in turn can be leveraged to make a more informed decision on their next career step.

By contrast, the increased focus on a specific niche that is typical for data scientists at larger companies may limit the breadth of skills they develop, but lead to considerable in-depth knowledge and expertise in their area of focus. They may also benefit from greater company resources such as specialized training programs or the free use of online learning platforms such as LinkedIn Learning, Coursera, or Pluralsight. Within Microsoft, for instance, we also have a biannual, internal conference called MLADS (Machine Learning, AI, and Data Science) in which the state-of-the-art is shared across a wide range of disciplines.

**Available Resources**

While resource restrictions are usually in place at companies across all sizes, they tend to be more pronounced in small organizations. These restrictions can range from smaller datasets, limited compute power, fewer specialized tools and software to fewer opportunities to learn and grow, characterized by limited funding for training programs, academic conferences, and fewer sources for domain knowledge within the company. This makes it more challenging for data scientists to conduct large-scale analyses and tackle complex projects that require a large amount of domain expertise, data, and compute power.

For instance, during my startup tenure, we only had two GPUs at our disposal, which significantly limited our models’ capacity, increased their training time, and restricted us from conducting experiments at scale. As a result, we had to allocate considerable time and effort to finding more efficient solutions and, at times, delay the initial date of the feature delivery.

The opposite is generally true for larger companies. They often provide their data scientists with access to large datasets, more compute power, and a wide range of tools and software that enables them to implement a solution in a variety of different ways. Efficiency still matters as it can reduce latency and bring about cost savings; however, compared to small companies, it is often not a primary concern. In addition, larger companies tend to have more funding for learning and career growth through internal training programs, bootcamps, online learning resources, or academic conferences.

**Mentorship Opportunities**

Mentorship is a critical aspect of career development and can help guide you in the right direction and make more informed decisions when it comes to determining your next career step.

Due to the limited number of data scientists and senior employees at smaller companies, they may not always provide the opportunity for mentorship, particularly of technical nature. For instance, the data science division I worked in during my startup stint consisted of only three people, which meant that in order to retrieve technical information, we had to consult the scientific literature and figure out a way to translate novel methodologies from papers to code. On the upside, however, you may find yourself having more direct access to senior executives or even founders, who could still provide you with valuable and potentially more high-level advice with respect to business operations and strategies.

Many large corporations, by contrast, have well-developed mentorship programs that pair data scientists to experienced mentors and enable them to receive guidance on technical skill development, career growth, and leadership. At Microsoft, for instance, employees are offered a variety of such mentorship programs that are extremely well received. I have been receiving active mentorship myself and it has been an incredibly rewarding experience, allowing me to not only get help on technical issues but also learn more about navigating the career landscape of a data scientist and have stimulating discussions about the latest trends in the field.

However, based on my own experience and that of colleagues, finding a suitable mentor with respect to relevant expertise in a specific domain remains challenging, particularly given the growing pool of data scientists.

**Upward Mobility**

In smaller companies, upward mobility highly depends on the availability of senior positions and the maturity of the data science division. If the latter is still in its infancy, it may increase the likelihood of being assigned to a senior position more quickly with the right skills and experience. This way, one has more opportunities to take on greater project ownership and responsibilities, which can have a direct impact on shaping the direction of the company.

Conversely, large companies tend to have more well-defined career paths for data scientists in which they outline the specific criteria that need to be met in order to advance to the next level. While they may offer a wider range of opportunities to specialize in specific areas within the data sciences or to take on more responsibilities, the hierarchical structure of larger corporations can make it more challenging to have a significant impact on the direction of the company, and promotions may be more dependent on meeting some predefined criteria rather than taking on a wider range of responsibilities.

**Conclusion**

We have seen that smaller organizations typically require data scientists to have a basic understanding of the entire data science cycle, whereas at larger corporations the role is typically more focused on a specific niche. While resource restrictions are present across all companies irrespective of size, they tend to be more pronounced at smaller companies and can limit one’s access to large datasets, compute resources, software, and career development programs. More direct access to senior executives is definitely a plus at smaller companies; however, this may come at the expense of fewer opportunities for technical mentorship due to fewer employees possessing the relevant expertise.

Deciding whether to join a small or big company as a data scientist is ultimately a personal decision that should be based on your career goals, priorities, and preferences. I hope that this article provides you with some guidance and gives you an overall idea of what it’s like to be a data scientist at a small versus a big company.