At most organizations, data science projects are typically overseen by three types of managers:

Business managers: These managers work with the data science team to define the problem and develop a strategy for analysis. They may be the head of a line of business, such as marketing, finance, or sales, and have a data science team reporting to them. They work closely with the data science and IT managers to ensure that projects are delivered.

IT managers: Senior IT managers are responsible for the infrastructure and architecture that will support data science operations. They are continually monitoring operations and resource usage to ensure that data science teams operate efficiently and securely. They may also be responsible for building and updating IT environments for data science teams.

Data science managers: These managers oversee the data science team and their day-to-day work. They are team builders who can balance team development with project planning and monitoring.

But the most important player in this process is the data scientist.

As a specialty, data science is young. It grew out of the fields of statistical analysis and data mining. The Data Science Journal debuted in 2002, published by the International Council for Science: Committee on Data for Science and Technology. By 2008 the title of data scientist had emerged, and the field quickly took off. There has been a shortage of data scientists ever since, even though more and more colleges and universities have started offering data science degrees.

A data scientist's duties can include developing strategies for analyzing data, preparing data for analysis, exploring, analyzing, and visualizing data, building models with data using programming languages, such as Python and R, and deploying models into applications.

The data scientist doesn't work solo. In fact, the most effective data science is done in teams. In addition to a data scientist, this team might include a business analyst who defines the problem, a data engineer who prepares the data and how it is accessed, an IT architect who oversees the underlying processes and infrastructure, and an application developer who deploys the models or outputs of the analysis into applications and products.

Despite the promise of data science and huge investments in data science teams, many companies are not realizing the full value of their data. In their race to hire talent and create data science programs, some companies have experienced inefficient team workflows, with different people using different tools and processes that don't work well together. Without more disciplined, centralized management, executives might not see a full return on their investments.

This chaotic environment presents many challenges.

Data scientists can't work efficiently. Because access to data must be granted by an IT administrator, data scientists often have long waits for data and the resources they need to analyze it. Once they have access, the data science team might analyze the data using different—and possibly incompatible—tools. For example, a scientist might develop a model using the R language, but the application it will be used in is written in a different language. Which is why it can take weeks—or even months—to deploy the models into useful applications.

Application developers can't access usable machine learning. Sometimes the machine learning models that developers receive are not ready to be deployed in applications. And because access points can be inflexible, models can't be deployed in all scenarios and scalability is left to the application developer.

IT administrators spend too much time on support. Because of the proliferation of open source tools, IT can have an ever-growing list of tools to support. A data scientist in marketing, for example, might be using different tools than a data scientist in finance. Teams might also have different workflows, which means that IT must continually rebuild and update environments.

Business managers are too removed from data science. Data science workflows are not always integrated into business decision-making processes and systems, making it difficult for business managers to collaborate knowledgeably with data scientists. Without better integration, business managers find it difficult to understand why it takes so long to go from prototype to production—and they are less likely to back the investment in projects they perceive as too slow.

Many companies realized that without an integrated platform, data science work was inefficient, unsecure, and difficult to scale. This realization led to the development of data science platforms. These platforms are software hubs around which all data science work takes place. A good platform alleviates many of the challenges of implementing data science, and helps businesses turn their data into insights faster and more efficiently.

With a centralized, machine learning platform, data scientists can work in a collaborative environment using their favorite open source tools, with all their work synced by a version control system.

A data science platform reduces redundancy and drives innovation by enabling teams to share code, results, and reports. It removes bottlenecks in the flow of work by simplifying management and incorporating best practices.

In general, the best data science platforms aim to:

Data science platforms are built for collaboration by a range of users including expert data scientists, citizen data scientists, data engineers, and machine learning engineers or specialists. For example, a data science platform might allow data scientists to deploy models as APIs, making it easy to integrate them into different applications. Data scientists can access tools, data, and infrastructure without

having to wait for IT.

The demand for data science platforms has exploded in the market. In fact, the platform market is expected to grow at a compounded annual rate of more than 39 percent over the next few years and is projected to reach US$385 billion by 2025.

If you're ready to explore the capabilities of data science platforms, there are some key capabilities to consider:

Choose a project-based UI that encourages collaboration. The platform should empower people to work together on a model, from conception to final development. It should give each team member self-service access to data and resources.

Prioritize integration and flexibility. Make sure the platform includes support for the latest open source tools, common version-control providers, such as GitHub, GitLab, and Bitbucket, and tight integration with other resources.

Include enterprise-grade capabilities. Ensure the platform can scale with your business as your team grows. The platform should be highly available, have robust access controls, and support a large number of concurrent users.

Make data science more self-service. Look for a platform that takes the burden off of IT and engineering, and makes it easy for data scientists to spin up environments instantly, track all of their work, and easily deploy models into production.

Ensure easier model deployment. Model deployment and operationalization is one of the most important steps of the machine learning lifecycle, but it's often disregarded. Make sure that the service you choose makes it easier to operationalize models, whether it's providing APIs or ensuring that users build models in a way that allows for easy integration.

Your organization could be ready for a data science platform, if you've noticed that:

A data science platform can deliver real value to your business. Oracle's data science platform includes a wide range of services that provide a comprehensive, end-to-end experience designed to accelerate model deployment and improve data science results.