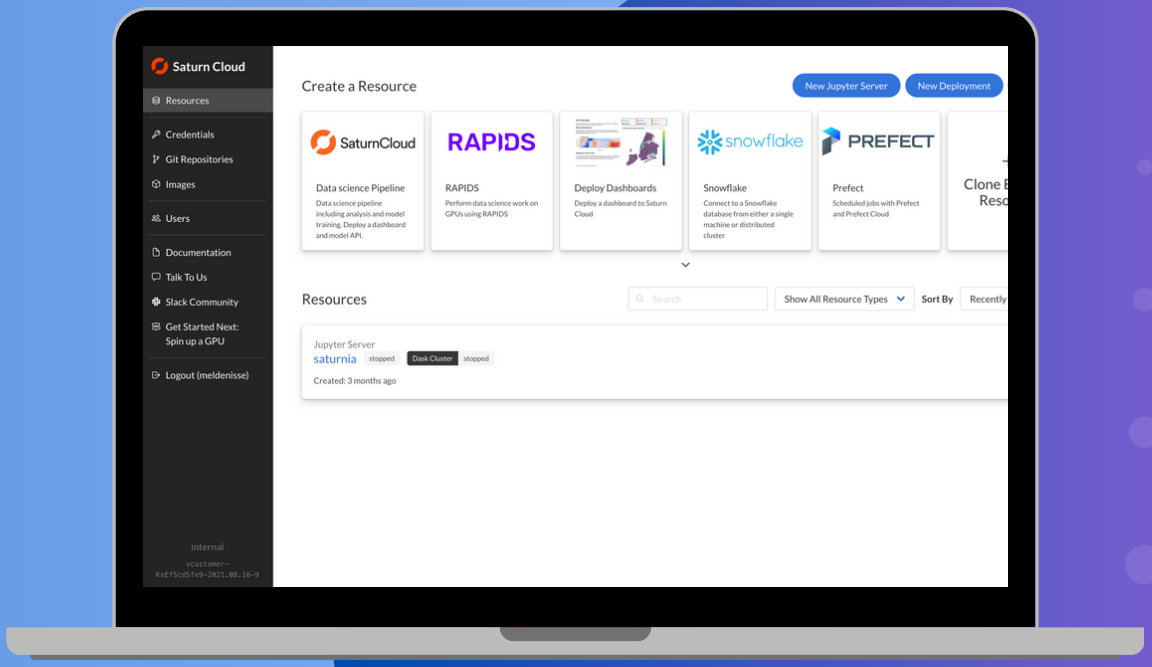
Saturn Cloud

Comparing Saturn Cloud and Amazon SageMaker

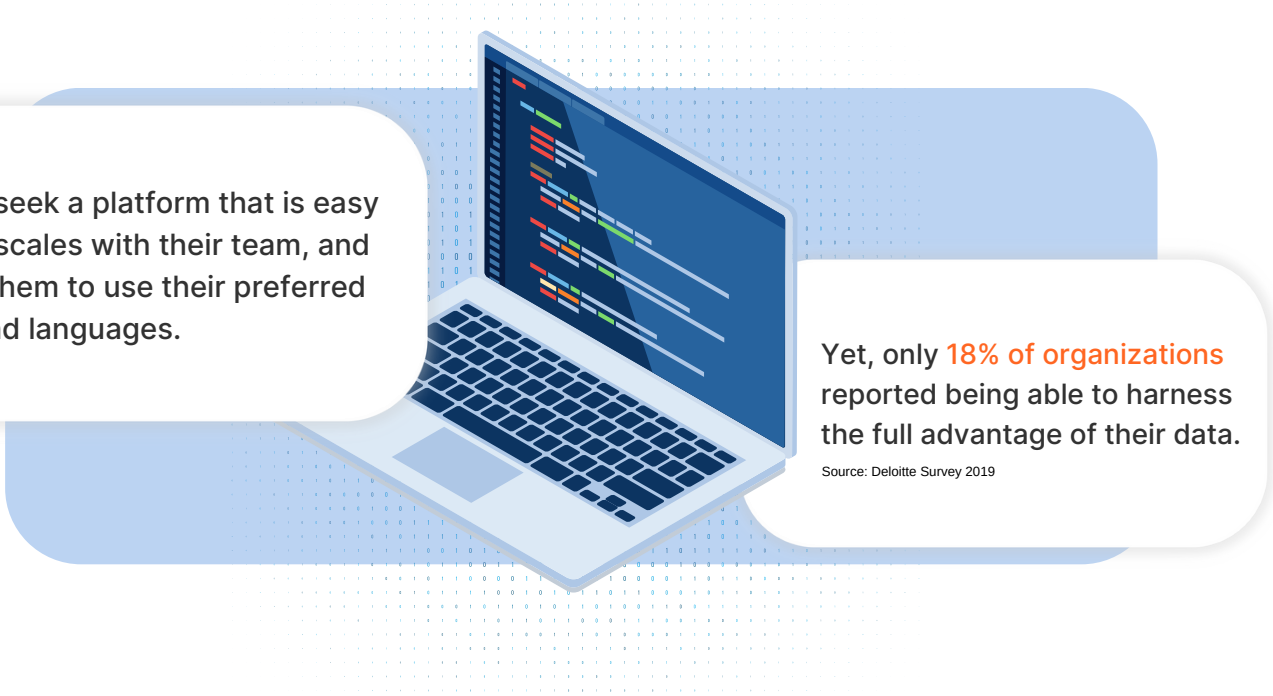


Sections:

Introduction

Comparing Saturn Cloud and Amazon SageMaker

When choosing a data science platform, data scientists look for several qualities and capabilities to ensure their work is productive and valuable.



Teams seek a platform that is easy to use, scales with their team, and allows them to use their preferred tools and languages.

Yet, only 18% of organizations reported being able to harness the full advantage of their data.

Source: Deloitte Survey 2019

Saturn Cloud and SageMaker are two popular platforms among data scientists. Both allow them to work in the cloud using hosted notebooks, but they differ significantly in their features and ease of use.

In this comparison, we'll walk through what differentiates the two platforms and how teams can best make decisions on what can benefit their team the most.

Overview

Saturn Cloud

Saturn Cloud allows data scientists and teams to collaborate with scalable resources in the cloud easily.

With just a few clicks, users can access computing resources with customizable amounts of memory and power, including GPUs and Dask distributed computing clusters, in a wholly hosted environment.

Data scientists can use their preferred languages, IDEs, and machine learning libraries. It offers full Git integration, shared custom images, and secure credential storage, making it easy to scale and build any data science team in the cloud.

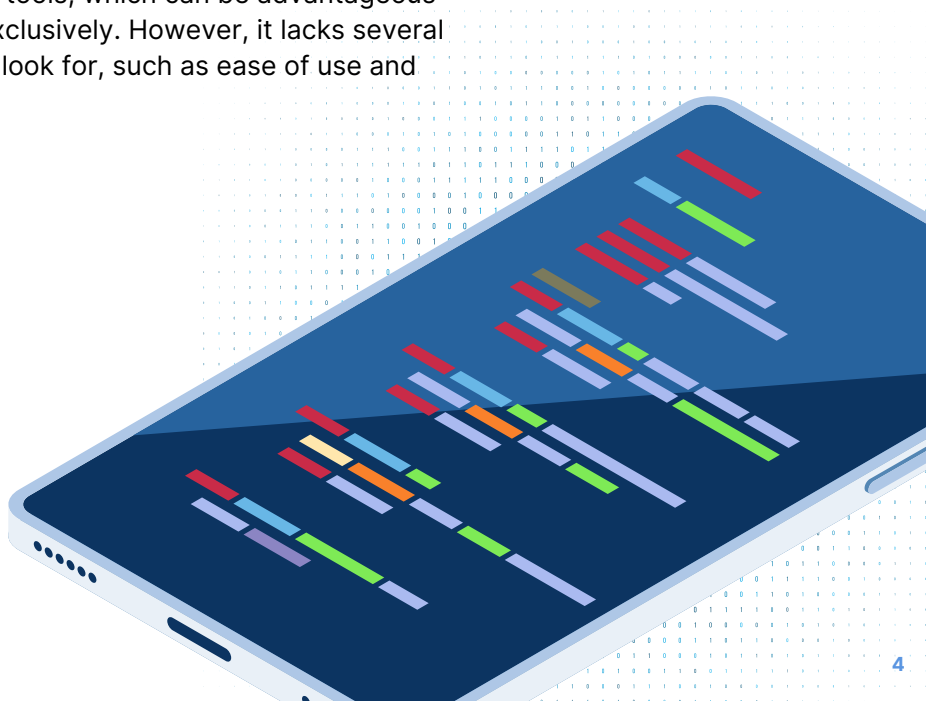
With features like jobs and deployments, it supports the entire machine learning lifecycle from experimentation to production at the lowest costs.

Amazon SageMaker

Amazon SageMaker allows data scientists to build, train, and deploy machine learning models in the cloud.

It gives users a Jupyter Notebook interface that supports Python and other programming languages (e.g., R, Scala, Julia). In addition to model development and deployment, SageMaker features data labeling tools, custom algorithms designed or implemented by AWS, model interpretability tooling, and feature storage products.

It is tightly integrated with other AWS tools, which can be advantageous for a team that uses AWS products exclusively. However, it lacks several key features that data science teams look for, such as ease of use and interoperability.



See How Saturn Cloud Stacks Up Against SageMaker

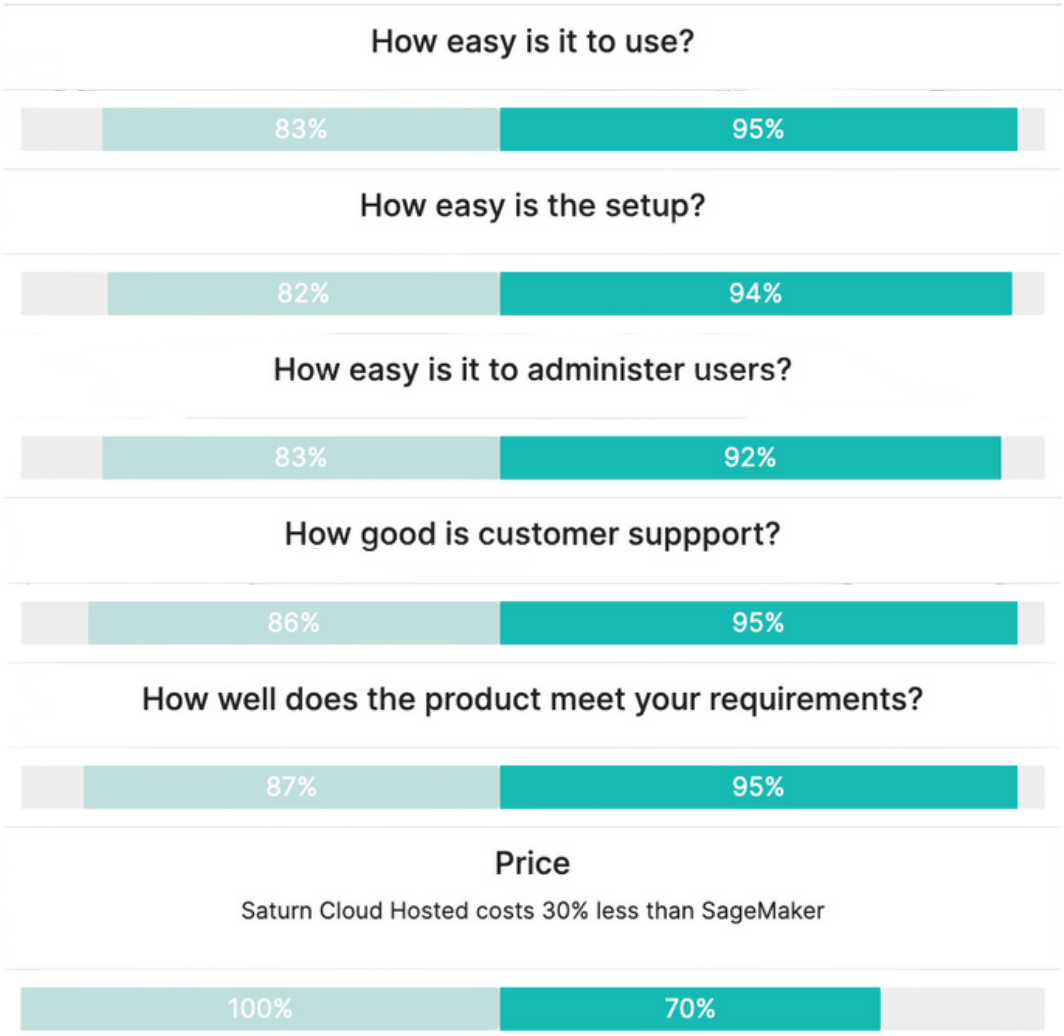
As you prepare to choose a data science platform for your team, consider how the following features align with your business priorities and needs.

Platform Features	SageMaker	Saturn Cloud
ML Apps and Dashboards	SageMaker has no application or dashboard capability built in.	Saturn Cloud supports flexible application and dashboard deployment including fine grained access control and collaboration for editors and viewers. All web frameworks are supported, including Streamlit, Panel, Bokeh, Plotly and Shiny.
Compute Clusters	Only supports training and inference jobs.	Supports parallel training, inference jobs, and advanced compute clusters like Dask and Ray1.
Security	SageMaker is only secure if you set it up correctly.	Secure out of the box with many additional upgrades depending on your needs: • Installation within your AWS account and VPN. • limiting egress to the internet. • integrating with transparent proxies. • Connecting to your existing monitoring and threat detection and response stack.
Admin and Cost Controls	No admin or cost management tools.	Saturn Cloud has the ability to constraint instance types, budgets and capabilities for users. Advanced cost management tools give you transparency into your organizations usage. Automatic shutoff prevents wasted spend.
Customer Support	AWS support is acceptable.	Saturn Cloud support is truly comprehensive. We will help you figure out conda solve errors, dask timeout issues, debug Pandas issues and more.

*Ray support coming in Q2 of 2024

What Users Are Saying

SageMaker	Saturn Cloud
★★★★☆ (4.2)	★★★★★ (4.9)
37 Reviews	255 Reviews



Ease of Use

The computing power of any data science platform is only helpful if it is easily accessible.

Saturn Cloud

Saturn Cloud's user interface is intuitive and straightforward, providing the capabilities and configurability that data scientists need.

The start-up process for new users is only two steps: create an account, which can use Github or Google credentials, and, with one click each, create and start a template project.

From there, data scientists have access to a Jupyter environment and straightforward instructions for adding their code, customizing their workspace, scaling up compute resources to Dask clusters, and more. The Saturn Cloud environment is as familiar to the users as their local environment.

SageMaker

SageMaker's start-up process is much more complex for new users. First, they must have an AWS account and the appropriate permissions. To get started, they need to understand concepts such as VPC creation, subnets, and AWS credentialing just to start up an instance.

Once they get started in SageMaker, more challenges await them. In particular, they are confronted with a UI experience in the Jupyter workspace that can be busy, confusing, and stressful, especially if users already have a perfect setup in a local environment.

What's more, AWS advocates separating machine learning training and inference from its hosted notebook, which can provide flexibility in deployment but requires additional user training and experience. Frequently, extra assistance from DevOps professionals is needed to complete the set-up processes and get them working securely.

Open and Interoperable Computing

Users should be able to choose their own tools and workflows in a data science platform.

IDEs

Data science teams looking at new software often question how well it will mesh with their existing, preferred tools. Few teams are interested in scrapping their current workflow altogether, but they do have a gap or many gaps in functionality that need to be filled. For example, users may like developing in a local environment but require computing resources in the cloud to power that work.

Saturn Cloud

Saturn Cloud makes SSH access both versatile and easy. Users can use SSH to connect a local IDE to a Saturn Cloud Jupyter instance and work from there, or they can use a local Jupyter installation and just SSH to a Dask cluster hosted in Saturn Cloud. This means that users can access powerful resources without leaving the development environment where they do their best work.

SageMaker

SageMaker does not natively support this type of connection. Notably, SSH connectivity from local to cloud development settings require workarounds or are impossible.

Images

Consistent, robust development images make data science practice easier, reproducible, and testable. If images are challenging to make or edit, users will likely struggle to reproduce analyses or bugs, slowing development and potentially introducing avoidable errors. analyses or bugs, slowing development and potentially introducing preventable errors.

Saturn Cloud

In Saturn Cloud, the default images are designed to be as slim as possible and regularly updated to keep users' tools current. Custom images may be created from a simple list of packages, YAML, or conda environment specifications, among other options. If desired, the images can even be based on Docker containers that a data science team already has available. Images may be shared with one or many users and can be viewed without starting a project.

SageMaker

In SageMaker, several default images are provided to users as kernels, but these images often have outdated versions of libraries. For example, at the time of this writing, the Python 3 Data Science kernel in SageMaker contains outdated versions of pandas, numpy, scikit-learn, ipykernel, jupyterlab, and other standard libraries. Outdated images require more time investment for updates, reducing productivity. Using a custom image in SageMaker is a series of lengthy multi-step processes: first, to create the image and then to attach the image to a project. Custom images must all be stored as a container image in Amazon ECR, and they may not be created in the UI.

Libraries

Data scientists and analysts work best using the tools they want to use.

Saturn Cloud

In Saturn Cloud, users can use their preferred machine learning libraries. Saturn Cloud explicitly and actively supports different machine learning frameworks, including scikitlearn, and deep learning tools such as PyTorch and TensorFlow. Moreover, it's designed so data scientists can adapt their local or single node code to the product. Here, code is transferable: Code written outside Saturn Cloud can be brought in, and code written inside Saturn Cloud can run in other workspaces.

SageMaker

While SageMaker passively supports various machine learning approaches, users are strongly encouraged to utilize AWS internal libraries, which abstract away large and essential pieces of the modeling process. These abstractions might be acceptable for users who do not have existing workflows in their local environment, but they might not be ideal for users who already do. Moreover, if users or teams ever depart from the SageMaker ecosystem, the code they wrote using the AWS libraries might become obsolete, resulting in re-work and wasted time.

Takeaways

In the rapidly evolving landscape of cloud computing, choosing the right platform is pivotal. This whitepaper illustrates how businesses can leverage Saturn Cloud or SageMaker based on their unique features, efficiency, and scalability to make an informed decision.

Explore how Saturn Cloud can transform your data science capabilities and distinguish your services from the competition.

Learn more about Saturn Cloud at saturncloud.io

> Saturn Cloud makes my work so much easier. When I sit down at the beginning of the day, I just want my environment to work. I want my favorite packages installed and available on demand. I want it to be easy to scale my workspace and have it shut down automatically when I'm done. Saturn Cloud solves all of that. Their customer service is also top-notch.

Daniel Burkhardt
Machine Learning Scientist

Cellarity
A Flagship Pioneering Company

Read Our Reviews