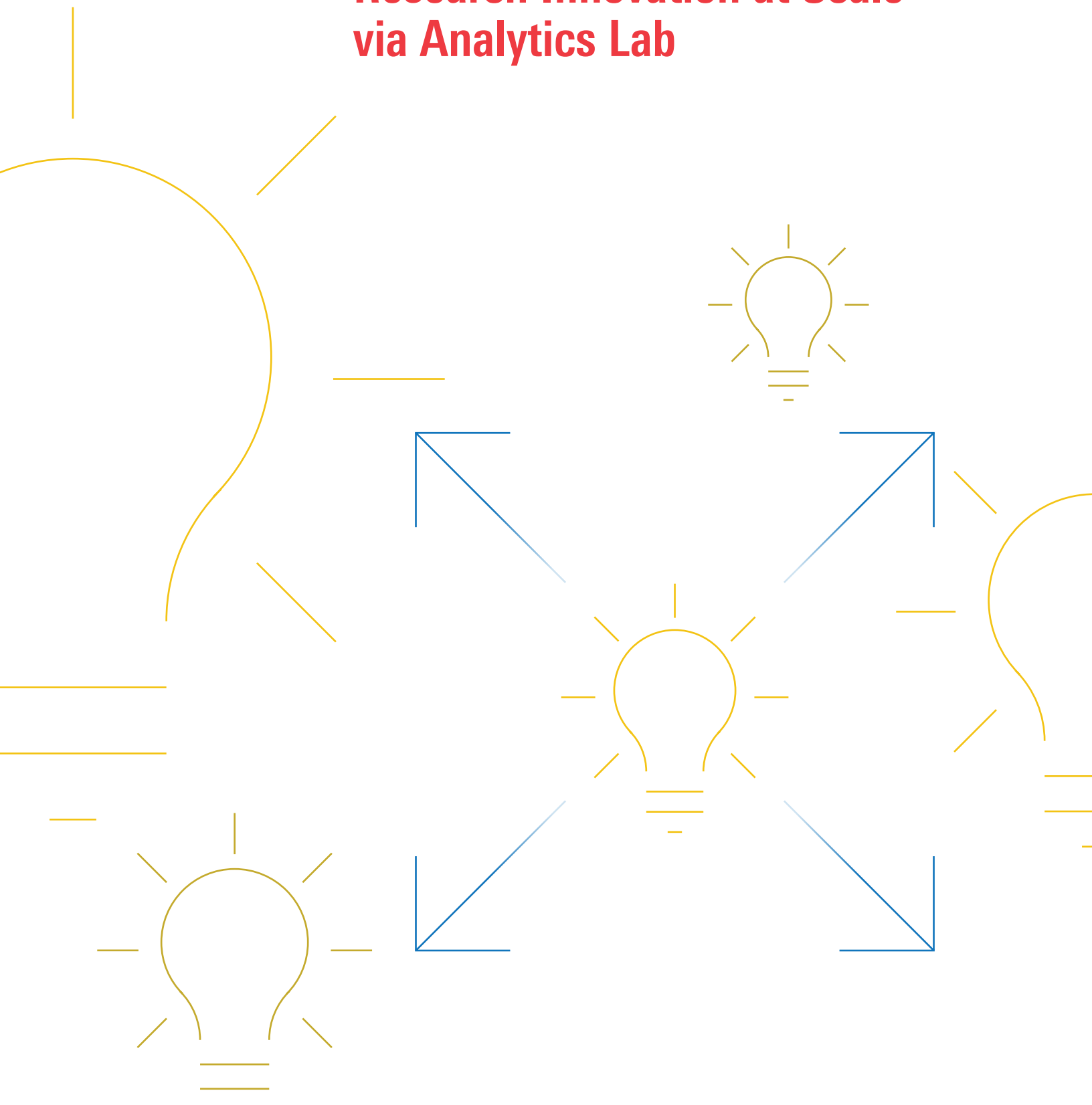
MORNINGSTAR®

Research Innovation at Scale via Analytics Lab



Analytics Lab: The Future of Innovation

Data and research have always been foundational to Morningstar's efforts to empower investor success. With the release of Analytics Lab in Morningstar Direct, we've raised the bar on both.

Analytics Lab streamlines a data scientist or quant analyst's most common daily tasks – data discovery, analysis, and insight creation, and integration of that analysis into day-to-day business operations. It's built on top of an open-source data ecosystem of tools and cloud-based "big data" technology. This allows researchers to use Analytics Lab to access all of Morningstar's data in a scaleable environment using the tools they are already familiar with.

Facing the Challenges of Modern Data Analytics

The inception of Analytics Lab began with an internal effort to support our quantitative analysts.

They were facing difficulties getting access to all the relevant data they needed for their research. If the research led to publishable insights/models, significant time and effort was required to build production data pipelines for the needed datasets and implement systems to publish their research.

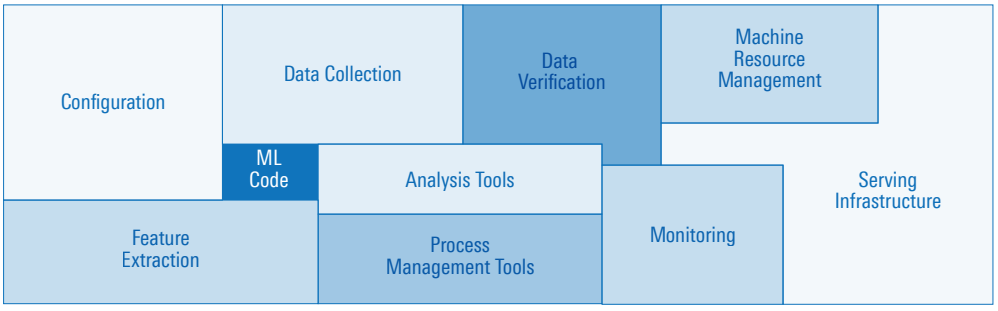
In addition to the headwinds they faced in getting access to data, some analysts needed more computing power to properly analyze the large amounts of data available. When they needed more compute resources, they either had to learn more technical skills like Amazon Web Services (AWS), Spark, and Kubernetes, or work with various technology teams and become beholden to their roadmaps.

*"For years, **our best ideas were landlocked**. They were slow to prototype. And once unearthed, they sat still stuck in time as of a publication date within the confines of white papers.*

*We needed a channel to **deliver our insights to investors soon after we discovered them**. And allow investors to subscribe to these insights if they wanted to."*

—**Lee Davidson**
Morningstar Head of Manager and Quantitative Research

Value Add vs. Relative Effort of Researchers' Work



The larger the box, the greater the Relative Effort.
Value Add: High – Low

As the old saying in data science goes, you spend 80% of our time cleaning and prepping the data, 5% doing the science, and 15% explaining what you did. The above diagram illustrates all the work that goes into creating value, represented by the dark blue box. We wanted to find a way to allow our researchers to focus primarily on the dark blue box, by having the engineering team create infrastructure to accommodate the other work in a simple, automated and intuitive fashion.

Since our analysts have similar objectives and workflows to others in positions across asset and wealth management, we knew that if we could solve the problem for our internal users, we'd have a solution that would eliminate pain points for the broader industry as well.

Our Solution: Analytics Lab

Our goal was simple: create a system in which analysts could bypass cumbersome data management tasks and focus on developing insights. We wanted to provide easy programmatic access to all of Morningstar data and a cloud development environment where users could leverage all the computing/querying power they needed.

"Effective research and model development requires access to vast quantities of data, which takes significant effort to collect, move and curate.

Analytics Lab eliminates the expense of moving data, making it easy to model and publish results; ultimately reducing the complexity required to deliver research that empowers investor success."

—**Alex Golbin**
Chief Data Officer at Morningstar

Leveraging Open-Source Jupyter Technology

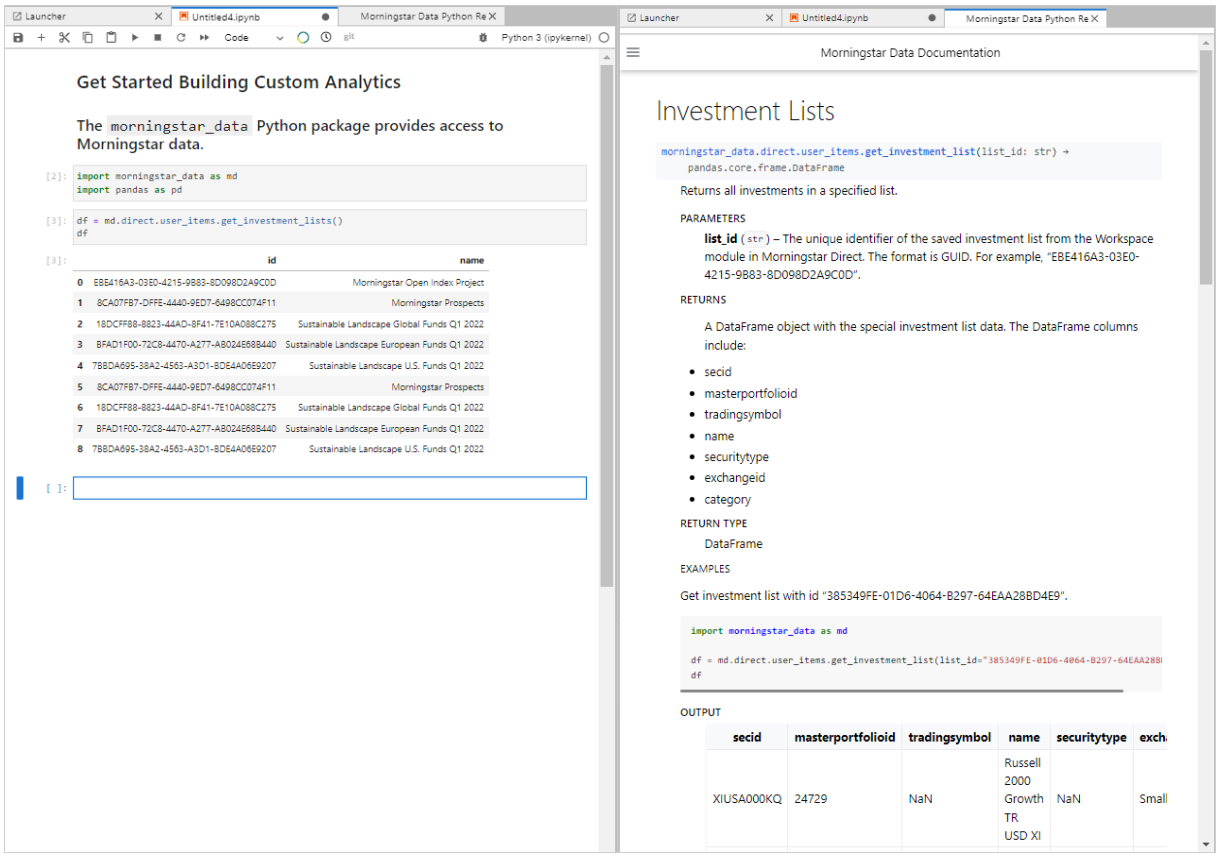
We built a platform on top of JupyterHub and JupyterLab to provide researchers a place to create and share Jupyter notebooks. Jupyter notebooks allow an individual to develop an analytical idea iteratively while simultaneously narrating how and why they constructed their analytics the way they did. Thus, with Jupyter notebooks, analysts can create one comprehensive document that queries data, generates insights, and documents the analysis and results. Analytics Lab also offers firm-level sharing of notebooks, so users can collaborate seamlessly and share valuable insights with others at their firm.

Seamless Data Access

Along with a cloud hosted platform to build notebooks, we created a Python package called morningstar_data. Morningstar_data provides an intuitive way to programmatically query Morningstar's large collection of data sets within a notebook. Once analysts have data in a Pandas dataframe, they can use the Python open-source data tools they are familiar with to clean and analyze the data. In addition, the morningstar_data Python library makes it easy to retrieve saved lists or data sets from Morningstar Direct and use them as inputs to customize analysis within a notebook.

With data, the whole is typically worth more than the sum of the parts. While each Morningstar dataset is powerful enough by itself, Analytics Lab allows analysts and data scientists to combine these datasets together more flexibly than ever to generate new and meaningful insights.

Using the morningstar_data python library to access investment list data



Morningstar Research

Our research teams have already been using Analytics Lab to conduct research and uncover insights. We'll be making this research available to all users of Analytics Lab so that clients can understand how our experts are thinking about industry problems and use our analysis as a jumping off point to customize as they need.

In addition, we'll be publishing sample notebooks and will make common calculations and methodologies easily accessible via the morningstar_data Python library. We're also supplying a set of visual components and charts built on top of libraries commonly used in the Notebook ecosystem. The result is a user experience that allows clients to focus exclusively on research by leveraging common analytics libraries to help with insight creation, as well as visual tools to help with communicating to readers in a digestible fashion.

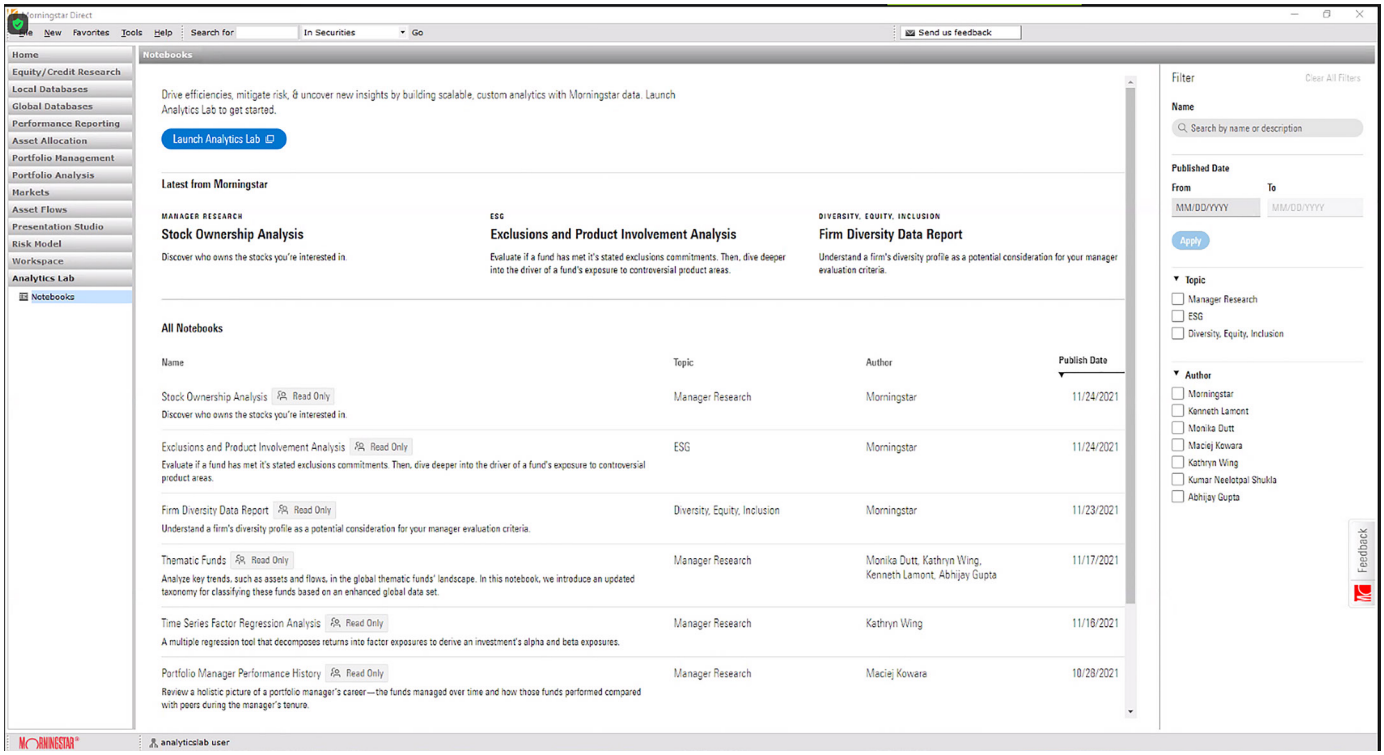
Leveraging the Results

Now that our analysts have a way to create insights quickly and efficiently, we wanted to make sure that the results could be shared broadly, and easily consumable by those without any programming knowledge.

We've leveraged open-source technology to transform notebooks into interactive dashboards and created a publishing pipeline to bring those user-friendly dashboards into Direct for all users to access. This means all Direct clients, whether proficient in Python or not, can take advantage of the latest insights from our research teams.

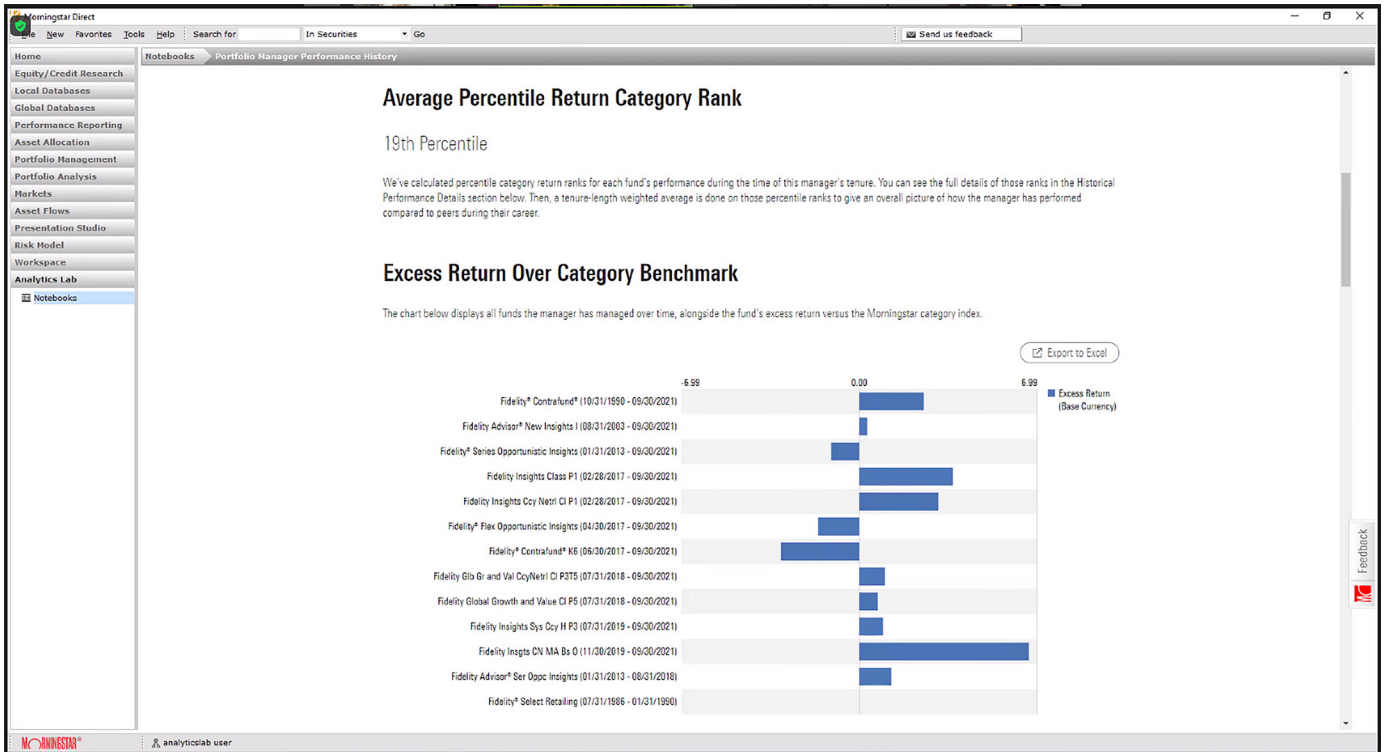
Our vision is to open this publishing channel to users of Direct as well, so that data scientists, quantitative researchers, or code savvy analysts can develop analytics and publish them to other members of their firm to leverage without ever needing to touch code.

Catalog of published Morningstar Notebooks in Direct



Whether published results are consumed through an interactive dashboard, report, or data feed system, our mission is to allow users to take their analysis to the final destination: the hands of investors.

Morningstar Notebook in Direct



Summary

With the release of Analytics Lab, we've combined Morningstar's data and research into a tightly integrated experience and automated the data discovery and insight generation processes significantly. With Analytics Lab, users can take ideas through the entire lifecycle from analysis to presentation.

"The objective of most data analysis is to communicate a non-obvious insight to a broader group. By integrating Analytics Lab with our existing systems, we enable data discovery, analysis and communication in a single platform.

The power of enabling this cohesive experience from analysis to communication is monumental."

—**James Rhodes**
Chief Technology Officer at Morningstar

See how Analytics Lab can help you build custom analytics using Morningstar data. Sign up for a free 14-day trial.