

REPORT REPRINT

RStudio focuses on bilingual data science for coders by embracing Python and R

FEBRUARY 10 2021

By Krishna Roy

The company's raison d'être used to be to provide open source and commercial data science for R coders. RStudio is now looking to expand its addressable audience by embracing a second open source programming language – Python.

THIS REPORT, LICENSED TO RSTUDIO, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, WAS PUBLISHED AS PART OF OUR SYNDICATED MARKET INSIGHT SUBSCRIPTION SERVICE. IT SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR RE-POSTED, IN WHOLE OR IN PART, BY THE RECIPIENT WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.

451 Research

S&P Global

Market Intelligence

Introduction

RStudio began to elucidate a strategy to envelop Python and R open source programming languages, used for data science and machine learning among other use cases, in a bid to provide coders with a single platform for open source and corporate-ready data science in 2019. In 2020, the firm began executing on this strategy in earnest by folding Python into its commercial open source-based portfolio – RStudio IDE, RStudio Server Pro and RStudio Connect – as well as supporting it within open source tools such as the RStudio IDE and Reticulate R interface to Python modules, classes and functions. Here we examine key enhancements to enable this objective, alongside developments in RStudio's business, which became a Public Benefit Corporation in early 2020.

451 TAKE

RStudio knows the user persona that likes to code data science very well. And that knowledge is enabling it to grow sales profitability in a difficult business climate. Furthermore, by enfolding Python within its commercial and free open source portfolio, the firm has opened the door to more coders. However, this opportunity could be expanded by delivering more capabilities as cloud services because this approach remains a top strategy for organizations, with 45% of enterprises surveyed in 451 Research's VoTE: AI & Machine Learning Services, Use Cases, 2021 survey indicating that they use cloud-based services to develop their machine learning apps.

Details

RStudio's Python integration work is all about embracing bilingual data science teams composed of Python and R programmers by providing them with a centralized environment for developing, collaborating, managing and sharing their work. Python integration with the RStudio IDE is noteworthy because the company is arguably still best known as the creator of this development environment. The RStudio IDE used to be exclusively for developing data science in R, using its console and syntax highlighting editor as well as tools for plotting, debugging code and workspace management.

However, native integration with JupyterLab and Jupyter Notebooks (as well as VS Code) have morphed it into a broader offering – the reason the firm now calls it a data science workbench. Indeed, RStudio notes that it has greatly expanded capabilities for native Python coding in the RStudio IDE, including a Python environment and object explorer user interface, for example, in recognition that data science teams don't always standardize on one language, so they need new tools to make it easy to combine R and Python projects. Additionally, coders can launch and manage Jupyter Notebooks, JupyterLab (and VS Code) in RStudio Server Pro, which is designed to provide corporate-ready data science by providing collaboration, centralized management, metrics, security and commercial support in a paid-for offering.

RStudio Connect has been another beneficiary of Python support. RStudio Connect is an open source-based publishing and deployment platform for sharing and scaling the code data scientists have created on their laptops by providing a paid-for server-side environment for it. It is also designed to enable coders to share insights with non-experts, such as decision-makers, via web applications, email or APIs, which is the rationale behind integration with the Dash Python framework for building web apps, Bokeh visualization library and Streamlit open source app framework for machine learning and data science. RStudio Connect users can also now share Python APIs using the Flask micro web framework written in Python – akin to support for Plumber in R introduced within it in 2017.

REPORT REPRINT

Additionally, the vendor has updated RStudio Package Manager with Python coders in mind. Package Manager is designed to be a commercial play for controlling and distributing packages throughout an organization, which will embrace the PyPI third-party software repository for Python. PyPI support is currently in beta.

Over on the business front, RStudio cites a 37% Y/Y growth in revenue in 2020, 5,100 active commercial software accounts and another year of profitability. The firm's goal is to be profitable this year too, noting that it has been profitable since 2016. RStudio is also hiring. It plans to recruit 30 employees in Q1 2021. The company cited a current headcount of 185 employees, which is the same as last October, when it shed light on its cloud strategy. We will fully examine RStudio's cloud strategy in a subsequent report later in the year.