RStudio PBC seeks to augment R and Python analytic workflows in Tableau, illuminates data science business

December 2 2021

by Krishna Roy

The company is looking to tap into the ubiquity of Tableau as a business intelligence platform and the popularity of R and Python open source programming languages by delivering integration between its wares and Tableau. RStudio has also shed light on its data science business since it was certified as a Public Benefit Corp in 2020.
Introduction

Data science isn't just for data scientists – other individuals require it as predictive decision-making becomes increasingly prevalent. Indeed, 30% of respondents to 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2021 survey said predictive analytics might be used by their organization for decision-making over the next two years. RStudio is looking to tap into this growth – as well as to harness the popularity of Tableau as a business intelligence (BI) platform – by providing Tableau analysts and other user personas that wish to code data science in R and Python within Tableau with improved analytical workflows for doing so. Furthermore, RStudio has provided visibility into its business model as a Public Benefit Corp. RStudio became a PBC in May 2019 and officially announced it in January 2020.

THE TAKE

RStudio appears to be continuing to handle the tricky balance between adhering to the ESG mandates necessary to legally operate as a PBC while running a profitable growing business, which is no mean feat. Furthermore, the company is better positioned to harness the growth in predictive decision-making by integrating with Tableau – its first integration with a BI platform – because Tableau has a loyal fan base, which RStudio could ultimately sell into more broadly. That said, RStudio's integration needs clear and careful articulation to avoid confusion with Tableau's own 'business science' strategy.

Details

Founded in 2009, RStudio's raison d'être is to provide open source and commercial data science for code-loving R and Python users. While it is keen to retain its mantle as one of the biggest contributors to the open source ecosystem for R by continuing to develop and sell free open source tools, RStudio started making forays into proprietary products five years after the company was established. RStudio began selling commercial products in 2014 as part of a strategic shift to support enterprise data science using R. Furthermore, the firm expanded this strategy to support Python in early 2020. RStudio notes that 51% of deals in Q3 2021 involved R and Python, compared with 43% in Q1 2021. The vendor also cites a paying customer base of more than 5,200 paying accounts as of Q2 2021.

RStudio now seeks to bring its R and Python-based data science capabilities to Tableau's BI platform to provide better analytical workflows when coding data science jobs in R and Python. RStudio's objective is twofold: to give Tableau users a better experience when coding predictions and recommendations by introducing supplemental data management and modeling capabilities, as well as to make R and Python programmers more comfortable using Tableau for data science. Tableau already integrates and visualizes data from R and Python models. However, RStudio contends that Tableau's approach, while good for communicating results through visualization and handling certain data-wrangling tasks, doesn't comprehensively address the entire analytic workflow for Tableau users, as well as R and Python coders – two distinct audiences with different needs, albeit united by a common requirement for R- and Python-based data science.

Tableau's approach involves the execution of arbitrary code, which RStudio contends poses a security threat because arbitrary code is vulnerable. RStudio also argues that it creates code debugging issues due to the paucity of debugging tools available, as well as risking a high degree of failure as environment and package management is also absent. The company is seeking to remedy these perceived shortfalls by integrating its RStudio Connect commercial platform – designed to be a content hub and execution engine for R and Python code – into Tableau. Existing RStudio Connect licenses work with the new Tableau integration, which is available in the latest version of RStudio Connect. RStudio Connect essentially hosts the data wrangling and modeling code as well as hooks into Tableau's Analytical Extension API – the vendor’s longtime interface to third-party data science tools.
RStudio's Tableau integration also draws on Plumber – one of many open source R packages the firm supports – to convert existing R code into a web API, ready for RStudio Connect to automatically manage the dependent files and packages associated with the API. RStudio Connect also enables deployment, management and scaling of Python APIs built with FastAPI. Furthermore, Shiny, another open source tool RStudio supports, is also on hand if R and Python coders require more sophisticated and interactive capabilities. RStudio has created shinytableau to make it possible for Tableau users to extend Tableau dashboards using Shiny and R.

Over on the business front, RStudio reports that it continues to operate as a profitable growing entity. Moreover, the company notes that is tracking ahead of its financial plan for FY 2021, which is already a substantial increase over 2020. RStudio continues to recruit to expand beyond the engineering-heavy organization it was a few years ago. The firm notes that it has 20 positions currently open and cites a headcount of 220 employees, compared with 185 in early 2021 when it announced PBC certification.