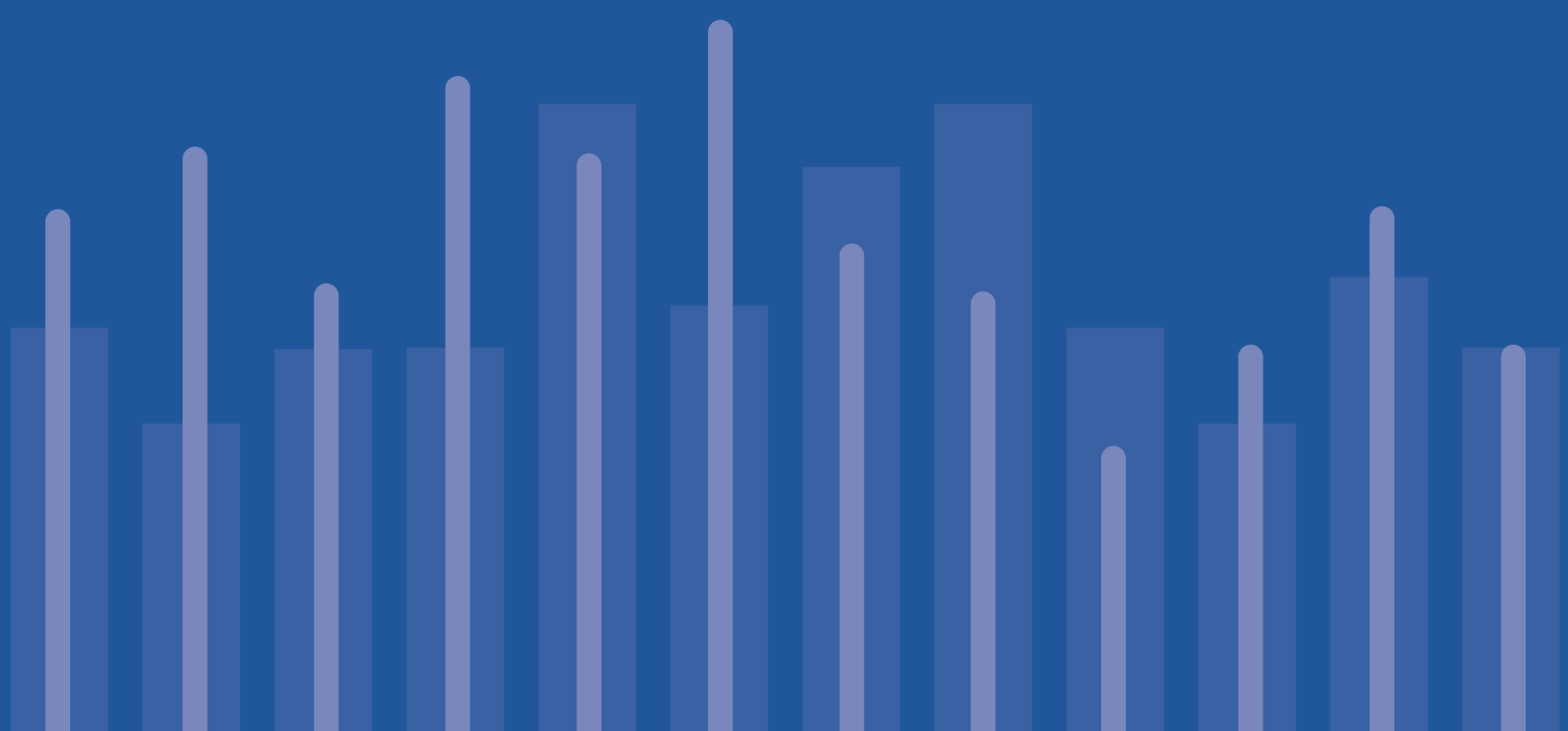




The Power of R and Visual Analytics

Partnering Advanced Statistics with Data Visualization

Nathan Mannheimer, Product Manager, Advanced Analytics, Tableau
Tyler Martin, Senior Software Engineer, Advanced Analytics, Tableau



Data scientists and statisticians know a powerful analytical tool simply by the name “R”—an open-source, free programming language and environment for statistical computing. It is historically one of the fastest-growing statistical programming languages and a top tool of choice in the data scientist’s toolkit.

Statisticians, data scientists, and analysts are an integral part of the business, tasked with sifting through the vast amount of data available today to discover valuable new insights. Whether it’s identifying complex patterns and relationships or using the data to predict future activities, uncovering these insights requires the ability to build statistical models, then analyze and share the results—a job suited perfectly for Tableau’s integration with R.

While R handles the “heavy lifting” of statistical analysis, Tableau allows you to explore and visualize the results in a simple drag-and-drop environment, turning your statistical findings into a beautiful dashboard. With **Tableau Server**, you can securely share these visualizations across the business, putting the hard work of data scientists into the hands of the business user. When it’s easy to see and communicate the results of advanced analytics, these analyses can play a more integral role in driving decision-making across the business.

Together, Tableau and R offer the best of both worlds, making it quick and easy to take advantage of data to solve today’s increasingly complex business challenges.

Contents

- Sift Through the Noise to Learn What Customers Are Saying
Sentiment Analysis3

- Automate the Decision-Making Process
Prescriptive Analytics.....4

- Analyze Complex Data Sets with Machine Learning
Decision Trees5

- Show it Wasn’t Chance
Statistical Testing6

- See How Data Changes Over Time
Time-series Analysis7

- Test Different Scenarios
What-if Analysis8

Sift Through the Noise to Learn What Customers Are Saying

Sentiment Analysis

Online opinions, such as comments, reviews, and discussions across popular sites like Twitter, Facebook, YouTube, Pinterest, and Amazon, are like buried treasure for companies wanting to manage their relationships with customers. But finding the hidden gems requires cutting through a lot of noise. From product development to reputation management, organizations need to systematically sift through online conversations in order to understand them and take action.

To do that, many are turning to sentiment analysis.

R offers functionality that makes it easy to parse text and determine sentiment, but people need a way to view the output to use it for decision-making. By integrating R with Tableau, you can pull that output into a visualization, drawing attention to positive and negative comments with color (Figure 1). You can also take it a step further, grouping these comments to analyze how sentiment has changed over time.



Figure 1: By visualizing the output of R's sentiment analysis, it's easy for anyone to spot customer problems.

In no time at all, sentiment analysis can cut through the noise and show business users what customers are saying, as they are saying it. This lets your company quickly identify specific issues or changing trends and respond with targeted public relations or marketing strategies.

Automate the Decision-Making Process

Prescriptive Analysis

Have you ever wondered how airlines determine their ever-changing ticket prices? Airlines sort through a complex set of travel variables, customer demand, and timing to offer ticket prices that will entice sales without hurting profits. Given a set of constraints, they're hunting for the optimal ticket price. This is a task for prescriptive analytics.

Where descriptive analytics describes what has already happened, and predictive analytics describes what could happen, prescriptive analytics helps decide what should happen. When faced with a number of potential decisions, it analyzes for the best possible outcome. Prescriptive analytics thrives on big data.

R handles prescriptive analytics quite easily using optimization functions, and with Tableau, you can visualize and communicate those results, empowering the business to make decisions based on the statistical horsepower of R. For instance, below you'll find an example of R's `optim()` output visualized as a treemap. In this case, we're visualizing optimal portfolio allocation for four stocks.

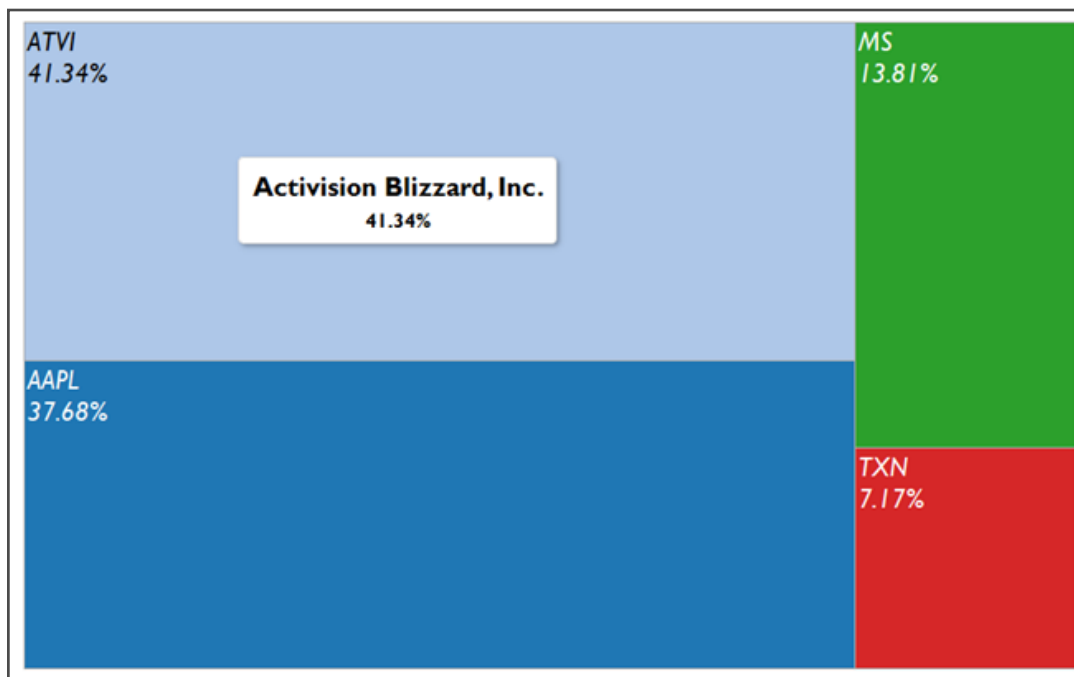


Figure 2: Tableau makes it easy to visualize the output of R's prescriptive analytics functionality, like in this example of portfolio allocation.

Analyze Complex Data Sets with Machine Learning

Decision Trees

Sometimes data is so complex or fast moving that it is difficult to group or make decisions for individual instances. One solution is to use a machine learning model to segment your data into like subsets for easier exploration and further modeling. R allows a user to quickly build decision trees to classify complex data based on previous examples.

Once your team has written an R script to classify data using a decision tree model, you can use that classification in any Tableau data visualization to drive color, shape, or grouping, simply by dragging and dropping. In fact, once you've created the R script to classify the data, any business user can leverage the results in a data visualization.

In the visualization below (Figure 3), a decision tree was used to classify new accounts into groups by the predicted lifetime spend of the account based on factors like location, size, revenue, and growth. To forecast revenue for the following year, we applied a second regression model to each new account, specially tailored to companies of each size. Using Tableau with R, any business user could explore data in this way, building on the decision tree functionality created in R by a data scientist or data science team.

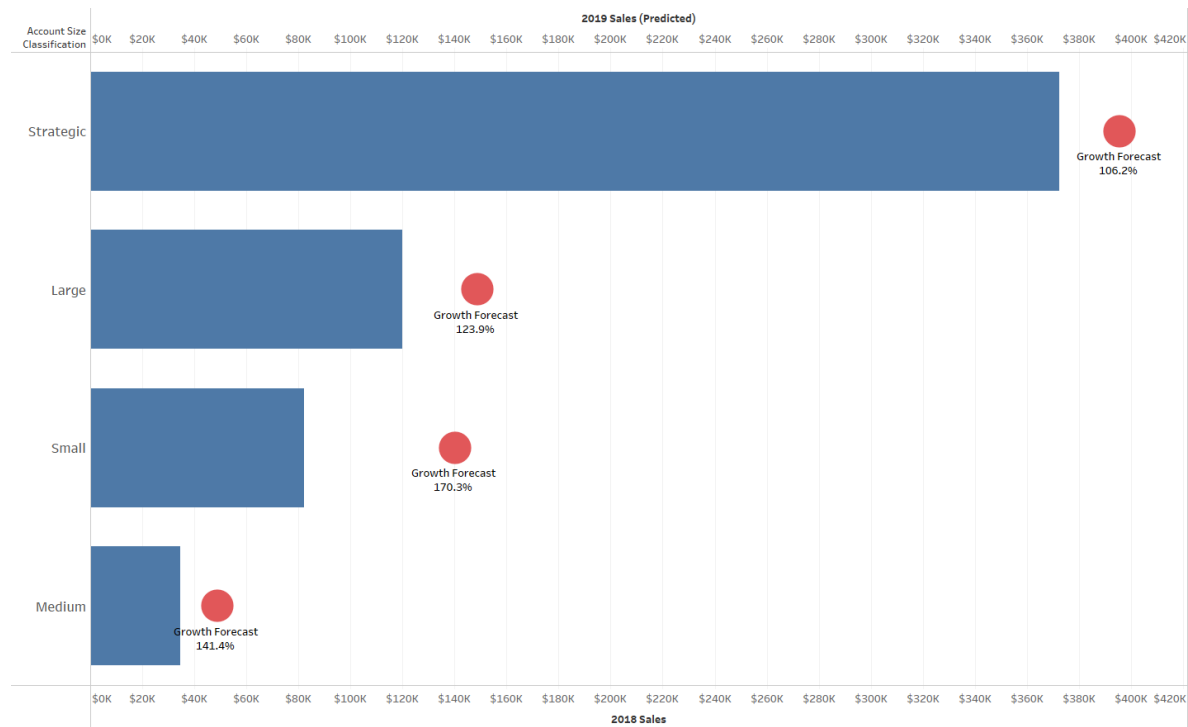


Figure 3: This visualization shows customer segment classifications created by an R decision tree. A business user can easily combine that data with real and projected sales data (in this case, data from 2017 and 2018) to understand how each segment's spend varied between years.

Show it Wasn't Chance

Statistical Testing

Statistical testing aims to show that it is unlikely that we would see the data we are seeing, purely as a result of random chance. For example, in pharmaceutical research, researchers give one group (group A) a placebo and one group (group B) the actual drug. They may see both groups improve, but let's say group B improved more than group A. They use statistical testing to determine if the difference in improvement between the two groups is big enough to be statistically significant.

Your sample may include categorical data such as gender or ethnicity, or it may include measures such as weight, age, or even amount of time. You can use Tableau to group your data into as many categories as are relevant and you can compute significance (p-values) for each of these with only one line of code.

R offers a wide variety of statistical testing functions and Tableau makes it easy to see the results of these tests. For example, Tableau's parameters make it simple for any user to adjust options in tests like the [Welch's two-sample t-test](#) and visualize the result.

Chi-square test of independence

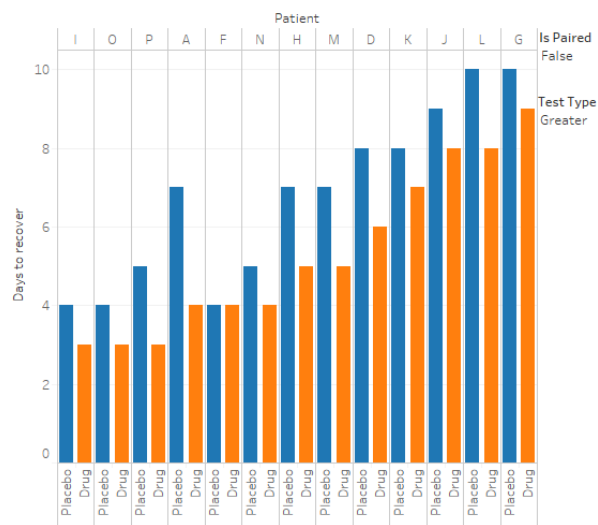
Contingency Table

Order Priority	Delivery Truck	Express Air	Regular Air	Totals
Critical	228	200	1,180	1,608
High	248	212	1,308	1,768
Medium	205	201	1,225	1,631
Low	250	190	1,280	1,720
Not Specified	215	180	1,277	1,672
	1,146	983	6,270	8,399

P value

0.4205

Welch Two Sample t-test



P value

0.05050

Figure 4: Tableau makes it easy to visualize the outputs of R's statistical testing functions and provide filters that allow users to adjust test options.

See How Data Changes Over Time

Time-series Analysis

Much of the world’s data contains an element of time—everything from weather to sales to stock portfolios. Analyzing data over time has a variety of benefits, including the ability to see important trends and seasonality.

R offers a wide range of tools for analyzing and forecasting time series data. Once you’ve created a time-series model in R, you can leverage that model to visualize the data in Tableau. In Tableau, you can drag and drop other fields into the view to see how other factors impacted those trends in the data. Let’s say there’s a spike in a particular month. In Tableau, you can drill into that spike and determine the underlying cause, without any additional coding.

Tableau also offers a drag-and-drop forecast feature based on the forecast package in R. If you are using the forecast package, simply click on Forecast in Tableau’s Analytics Tab to create an interactive visualization. With the combination of R and Tableau, you have the power to customize your model along with the ability to interact with it using filters and parameters.

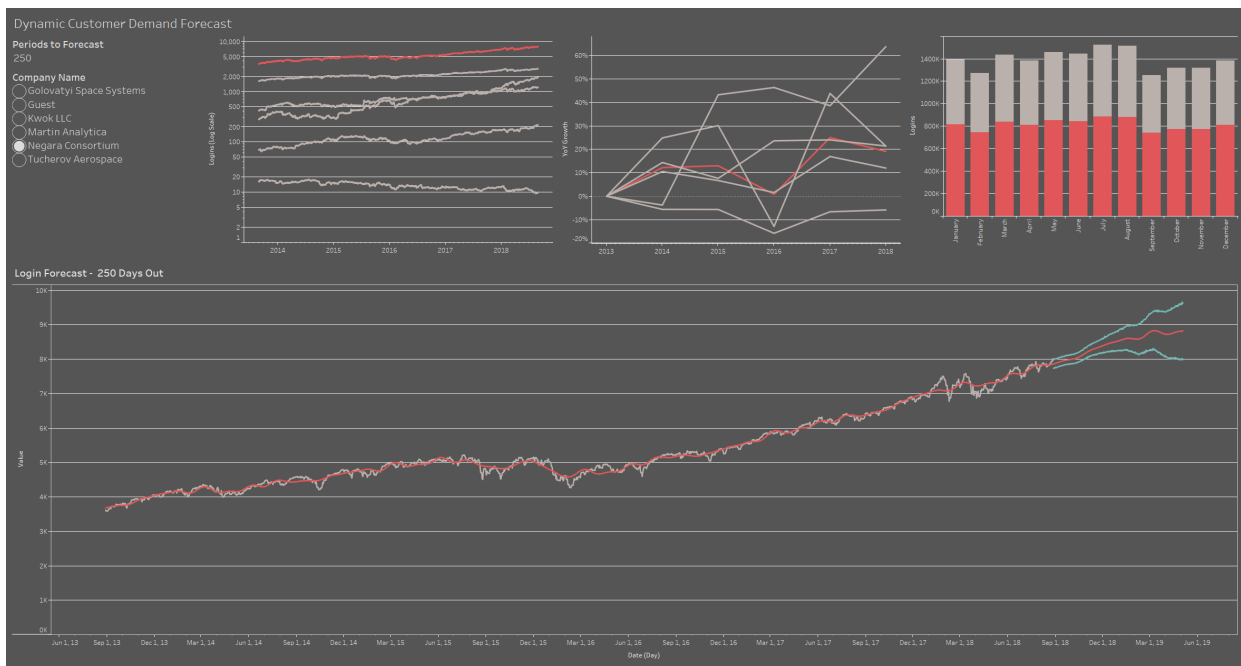


Figure 5: In this example, teams can explore a customer demand forecast, with the ability to drill down into the visualization and filter by company name.

Test Different Scenarios

What-if Analysis

“What-if analysis” allows you to test hypotheses for better decision making. To test different business scenarios, you may want to explore how changes to a particular value or set of values affect the output of your analysis. The combination of Tableau and R allows you to perform this testing in a visual way, without additional coding.

Let’s say you’ve built a model in R to see how different production indicators affect overall sales. You can bring that model into a Tableau dashboard to add parameters for each of these indicators. This allows your team to adjust these indicators and quickly determine how these changes affect the overall output. In the visualization below (Figure 6), teams could adjust start and end dates to see how production timelines affect the predicted sales for the year. This is yet another way to open up the power of an R model across an organization.

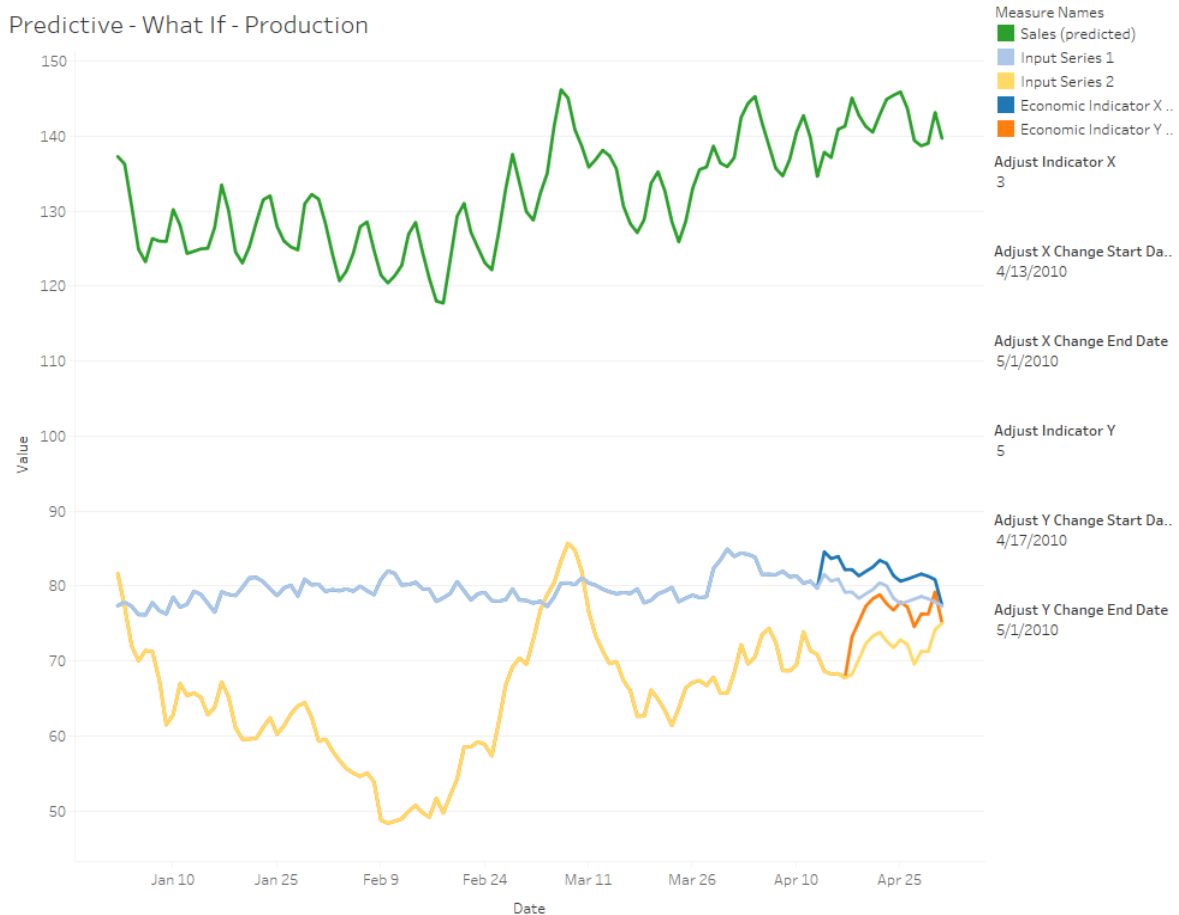


Figure 6: Tableau parameters allow teams to leverage a model to test different scenarios in the context of a visualization.

Combine the statistical power of R with Tableau to visually explore your data and make adjustments on the fly. Share your results with people of all skill levels to drive value from findings across an organization.

About the authors

Nathan Mannheimer

Nathan is the Product Manager for Advanced Analytics at Tableau, working on capabilities and integrations to support data science and statistical analysis. The Advanced Analytics team is focused on delivering powerful tools that let people answer deeper questions, including features like k-means clustering and Python integration. Outside of Tableau, Nathan is a lecturer in the Data Science and Human Centered Design and Engineering masters programs at the University of Washington.

Tyler Martin

Tyler is a Senior Software Engineer for Advanced Analytics at Tableau, working on building features for data science and statistical analysis. Before coming to Tableau, Tyler received his M.Sc. from the University of Washington in Applied Mathematics, where he studied fluid dynamics.

Related content

[Learn more about Tableau's R integration](#)

[How to use R with Tableau \(and when you should\)](#)

[Advanced Analytics with Tableau](#)

[Tableau and R: FAQ](#)

Explore other resources

[Product Demo](#)

[Training & Tutorials](#)

[Community & Support](#)

[Customer stories](#)

[Solutions](#)

