

Author: Bora Beran, PhD

The Power of R and Visual Analytics

Partnering Advanced Statistics
with Data Visualization

Data scientists and statisticians know a powerful analytical tool simply by the name “R.” R is an open-source, free programming language and environment for statistical computing. It is one of the fastest-growing statistical programming languages and is becoming the tool of choice for most data scientists. In fact, R overtook SPSS Statistics and SAS to become the tool used by the most data scientists and is now being used by nearly half of all data scientists.^{1 2}

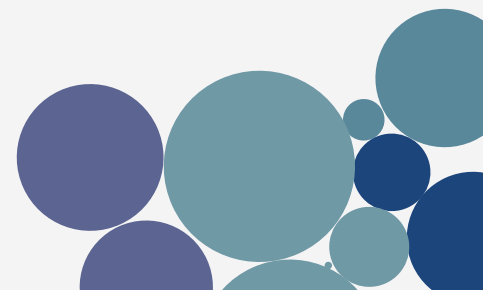
Statisticians, data scientists, and analysts are becoming an integral part of the business, tasked with sifting through the vast amount of data available today to discover valuable new insights. Whether it’s identifying complex patterns and relationships or using the data to predict future activities, 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.

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

Table of Contents

Sift Through the Noise to Learn What Customers Are Saying.....	3
Automate the Decision-Making Process.....	4
Make Sense of Large, Complex Datasets.....	5
Prove It Wasn’t Chance	5
About the Author	6



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 the 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 users 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.

CommentText	Sentiment
Crisp and bright screen, well worth the money.	positive
Customer service was unresponsive. They wouldn't allow me to ret..	negative
Do not buy this product. It only lasts 2 weeks. It is garbage. And th..	neutral
Don't waste your time on this product. One word. Bad.	negative
I had a pleasant experience. Responsive customer service, great pr..	positive
I want a refund. It does not work! Ridiculous.	neutral
I would recommend this product to anyone. I will buy another one ..	positive
It is OK for what it costs. Not great but works as expected.	positive
Much better than the competing products for a much more reasona..	positive
My experience was terrible. The product was dead on arrival.	negative
Product description is misleading. It is much smaller than what's s..	neutral
This is an awesome product. I enjoy using it every day.	positive
Took 4 weeks to receive it even though I paid for 2 day delivery. W..	positive
What a joy to use.	positive

Caption: 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 the business users what their 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.

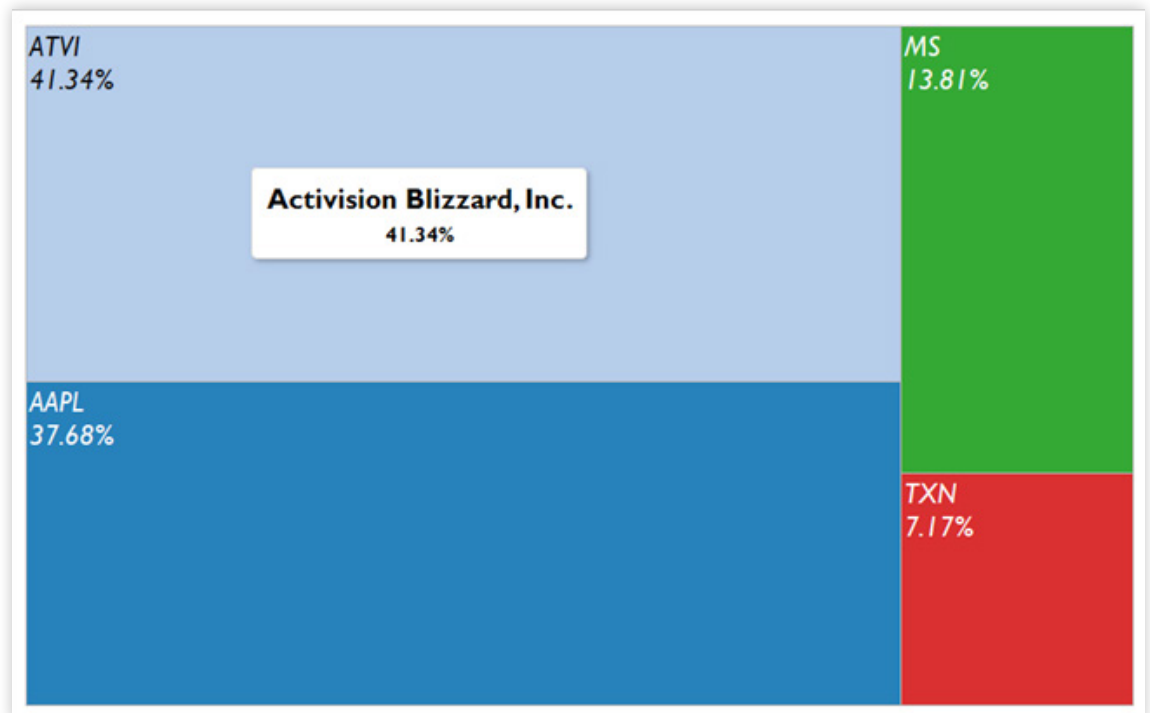
Automate the Decision-Making Process

Prescriptive Analytics

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 in order to offer ticket prices in real time 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 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.



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

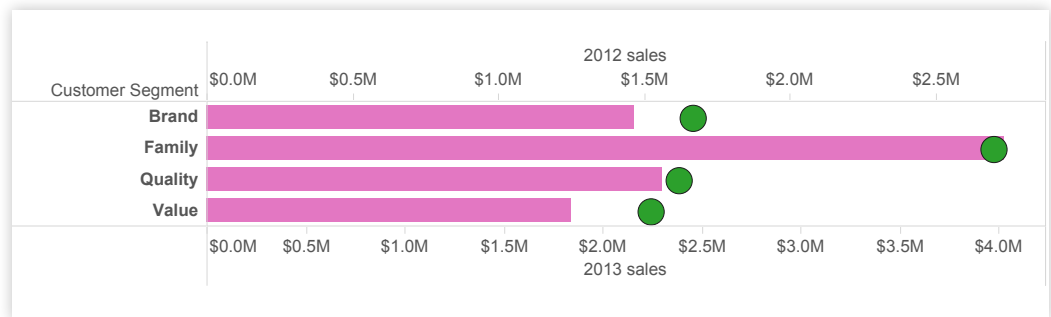
Make Sense of Large, Complex Datasets

Decision Trees

Sometimes data is so complex, it is difficult to find an overall regression model or single predictive formula that can describe it. One solution is to break up your data into smaller subsets until you can explain each subset with a simple model. R allows a user to quickly build decision trees to classify complex data.

Once your team has written an R script to classify data using decision trees, you can use that classification in any 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 its results in a data visualization.

In the visualization below, we've built a decision tree to classify a retail brand's customers based on their buying habits. We've then used that classification in a simple bar chart comparing 2012 sales to 2013 sales for those groups. 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. In this way, you can give business users meaningful, everyday access to the work done by a data science team.



Caption: This visualization shows customer segment classifications created by an R decision tree. A business user can easily combine that data with sales data (in this case, data from 2012 and 2013) to understand how each segment's spend varied between years.

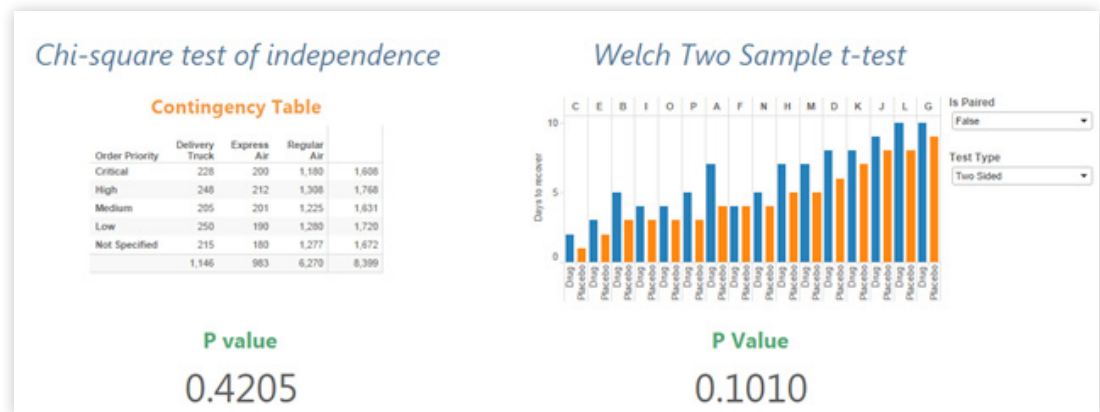
Prove It Wasn't Chance

Statistical Testing

Like in the courts, where a person is innocent until proven guilty, statistical testing aims to prove something is true by ruling out that it is not false. In most cases, this means proving something did not happen by chance alone. This is often described as being "statistically significant." For example, pharmaceutical research aims to prove a particular medicine made a difference – that the results prove success was statistically significant and not a result of chance.

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. Using that data, you may wish to test the probability that a particular drug had a positive effect on members with a particular age distribution versus those in the control set.

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



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

About the Author

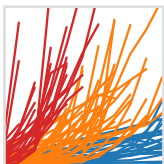
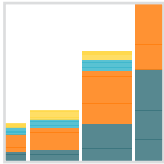
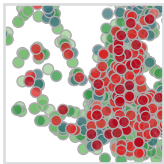
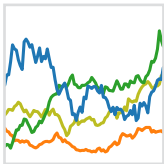
Bora Beran is a Program Manager at Tableau Software. His team focuses on statistics & calculations, technical partnerships and query generation pipeline. Prior to joining Tableau, Bora worked at Microsoft Research, SQL Server and Windows HPC Server on tools and infrastructure for large-scale computational modeling and simulations, data mining, visualization and knowledge representation. He holds a PhD from Drexel University and is a reviewer for several scholarly journals in the area of technical computing.

¹Karl Rexer, Heather Allen, & Paul Gearan (2011); [2011 Data Miner Survey Summary](#), presented at Predictive Analytics World, Oct. 2011.

²Karl Rexer, Heather Allen, & Paul Gearan (2010); [2010 Data Miner Survey Summary](#), presented at Predictive Analytics World, Oct. 2010.

About Tableau

Tableau Software helps people see and understand data. Tableau helps anyone quickly analyze, visualize and share information. More than 19,000 customer accounts get rapid results with Tableau in the office and on-the-go. And tens of thousands of people use Tableau Public to share data in their blogs and websites. See how Tableau can help you by downloading the free trial at www.tableausoftware.com/trial.



Additional Resources

[Download Free Trial](#)

Related Resources

[See the R Project for Statistical Computing](#)

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

[Check out Bora Beran's blog](#)

Explore Other Resources

- [Product Demo](#)
- [Training & Tutorials](#)
- [Community & Support](#)
- [Customer Stories](#)
- [Solutions](#)