

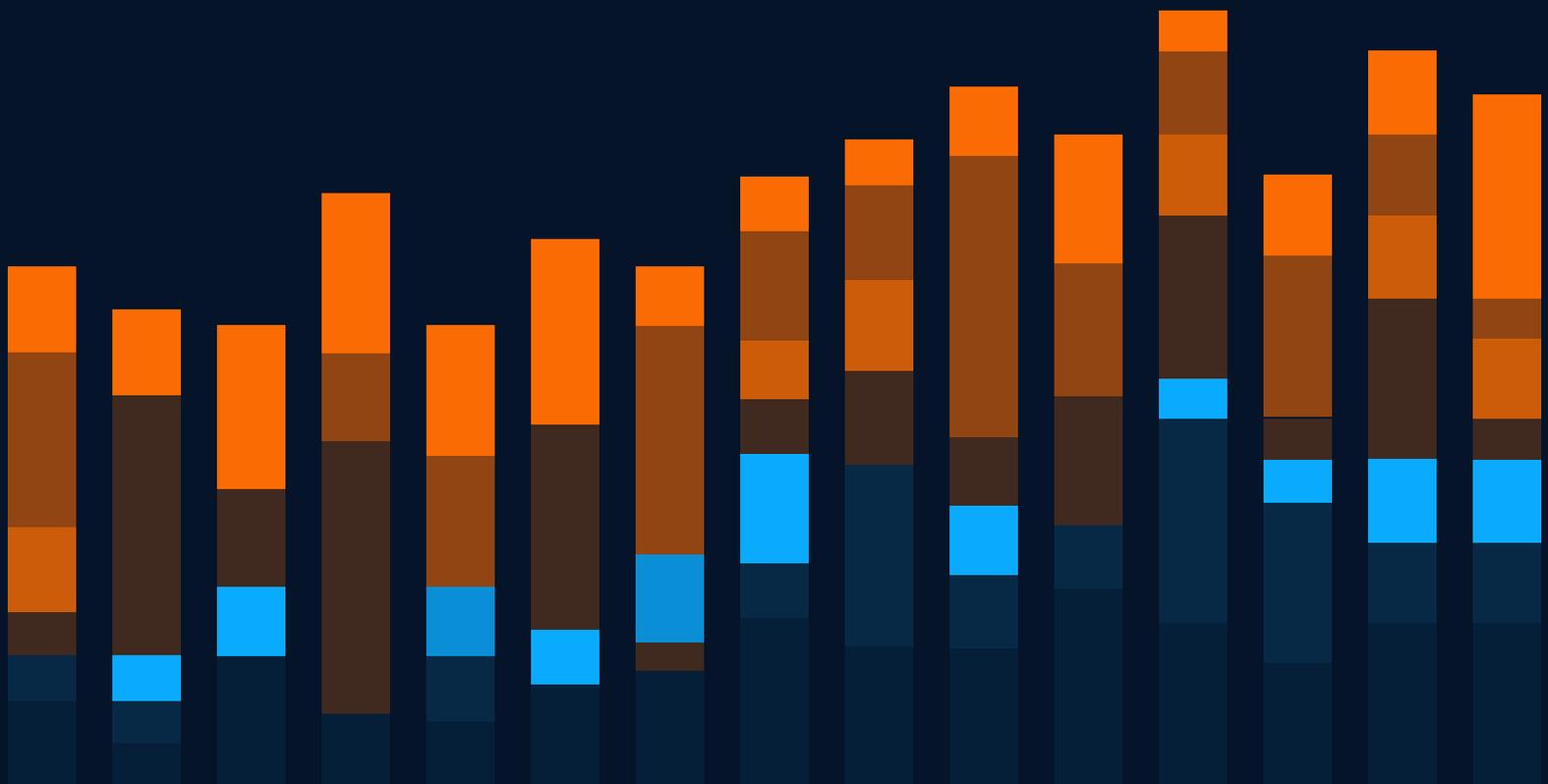


El poder de R y del análisis visual

Asociar estadísticas avanzadas con la visualización de datos

Nathan Mannheimer, administrador de productos, Análisis Avanzado, Tableau

Tyler Martin, ingeniero sénior de software, análisis avanzado, Tableau



Los científicos de datos y los estadísticos conocen una poderosa herramienta analítica que simplemente tiene el nombre de “R”: un lenguaje de programación de código abierto y libre, y un entorno para el cálculo estadístico. Históricamente es uno de los lenguajes de programación estadística de más rápido crecimiento y una de las herramientas preferidas dentro del conjunto de herramientas de los científicos de datos.

Los estadísticos, los científicos de datos y los analistas son parte integral del negocio, su tarea consiste en analizar la gran cantidad de datos disponibles en la actualidad para descubrir nuevas y valiosas informaciones. Ya sea que se trate de identificar patrones y relaciones complejos o usar los datos para predecir actividades futuras, descubrir este tipo de información requiere la capacidad de construir modelos estadísticos, de analizar y de compartir los resultados; un trabajo ideal para la integración de Tableau con R.

Mientras que R se ocupa del "trabajo pesado" del análisis estadístico, Tableau le permite explorar y visualizar resultados en un entorno simple de tipo arrastrar y soltar, para convertir los hallazgos estadísticos en un atractivo dashboard. Con [Tableau Server](#), puede compartir estas visualizaciones de forma segura en todos los niveles del negocio, para lograr poner el trabajo arduo de los científicos de datos en manos del usuario corporativo. Cuando es fácil ver y comunicar los resultados de análisis avanzado, estos análisis pueden desempeñar un papel más integral para impulsar la toma de decisiones en todos los niveles del negocio.

Juntos, Tableau y R, ofrecen lo mejor de ambos mundos, lo que hace que aprovechar los datos con el fin de resolver los retos empresariales cada vez más complejos de hoy en día sea una tarea rápida y fácil.

Contenido

Depurar el ruido para conocer qué es lo que dicen los clientes	
Análisis de sentimientos	3
Automatizar el proceso de toma de decisiones	
Análisis prescriptivo	4
Analizar conjuntos de datos complejos con aprendizaje automático	
Árboles de decisión	5
Mostrar que algo no ha sido casualidad	
Pruebas estadísticas	6
Ver cómo cambian los datos a lo largo del tiempo.	
Análisis de series temporales	7
Probar diferentes escenarios	
Análisis hipotético	8

Depurar el ruido para conocer qué es lo que dicen los clientes

Análisis de sentimientos

Las opiniones en línea, como comentarios, reseñas y discusiones en sitios populares como Twitter, Facebook, YouTube, Pinterest y Amazon, son como un tesoro escondido para las empresas que desean administrar las relaciones con sus clientes. Sin embargo, para encontrar las joyas ocultas necesita depurar el ruido. Desde el desarrollo del producto hasta la administración de la reputación, las organizaciones deben evaluar sistemáticamente las conversaciones en línea para comprenderlas y tomar acción.

Para hacerlo, muchas recurren al análisis de sentimientos.

R ofrece una funcionalidad que facilita la tarea de analizar el texto y de determinar los sentimientos, pero las personas necesitan una forma de ver el resultado para usarlo en la toma de decisiones. Al integrar R a Tableau, puede obtener ese resultado en una visualización, lo que le permite destacar los comentarios positivos y negativos en colores (figura 1). También puede profundizar un poco más, agrupando estos comentarios con el fin de analizar cómo han cambiado los sentimientos a lo largo del tiempo.

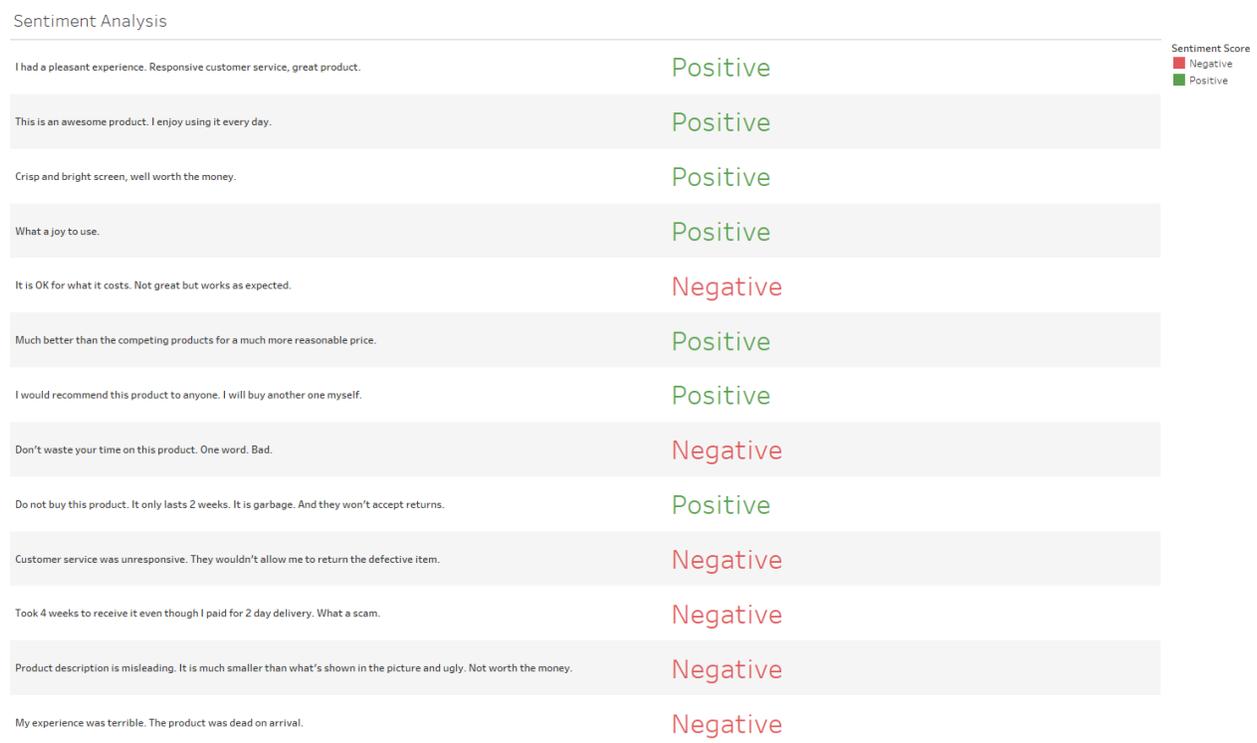


Figura 1: Al visualizar el resultado del análisis de sentimientos de R, es fácil para cualquiera detectar los problemas de los clientes.

En muy poco tiempo, el análisis de sentimientos puede depurar el ruido y mostrar a los usuarios corporativos lo que los clientes dicen, tal como lo dicen. Esto le permite a la empresa identificar rápidamente problemas específicos o tendencias cambiantes y responder con estrategias de marketing o de relaciones públicas específicas.

Automatizar el proceso de toma de decisiones

Análisis prescriptivo

¿Alguna vez se has preguntado cómo determinan las aerolíneas los siempre cambiantes precios de los boletos? Las aerolíneas clasifican un conjunto complejo de variables de viaje, de exigencias de los clientes y de tiempo para ofrecer precios de boletos que atraigan las ventas sin perjudicar las ganancias. Dado un conjunto de restricciones, siempre están a la caza del precio óptimo para los boletos. Esta es una tarea que corresponde al análisis prescriptivo.

Mientras que el análisis descriptivo describe lo que ya ha ocurrido y el análisis predictivo describe lo que podría ocurrir, el análisis prescriptivo ayuda a decidir lo que debería ocurrir. Ante una serie de decisiones potenciales, analiza el mejor resultado posible. El análisis prescriptivo funciona mejor en un entorno de big data.

R maneja el análisis prescriptivo muy fácilmente usando funciones de optimización y, con Tableau, puedes visualizar y comunicar esos resultados, lo cual permite a la empresa tomar decisiones basadas en el poder estadístico de R. Por ejemplo, a continuación encontrará un ejemplo de los resultados óptimos de R visualizados como un diagrama de árbol. En este caso, visualiza una asignación de cartera óptima para cuatro acciones.

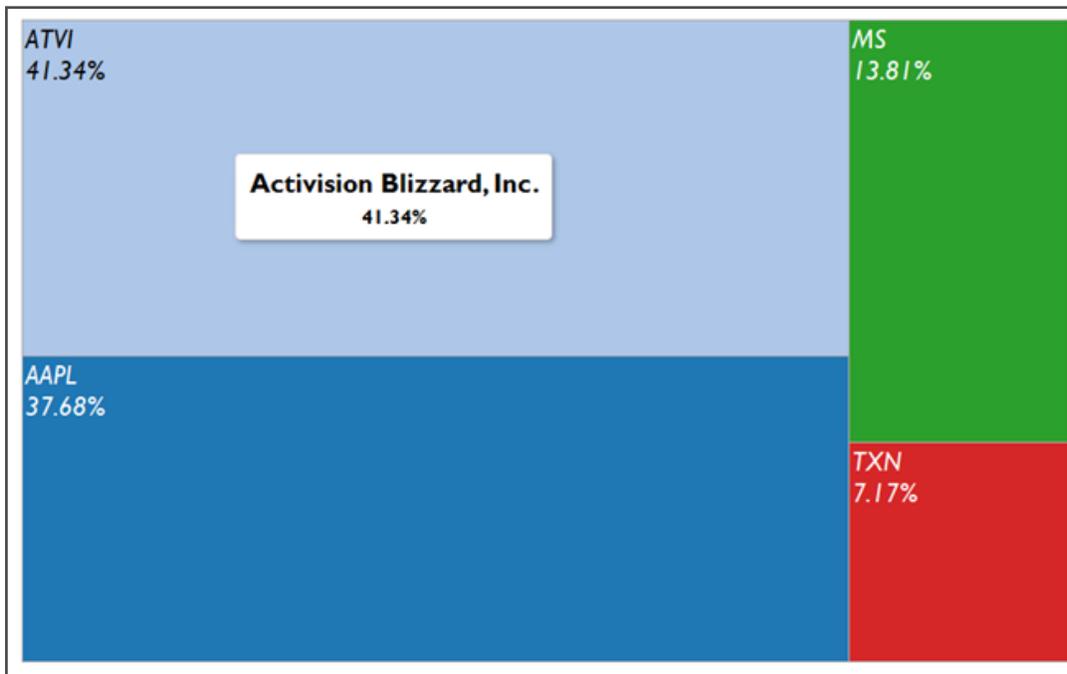


Figura 2: Tableau facilita la visualización del resultado de la funcionalidad del análisis prescriptivo de R, como en este ejemplo de asignación de cartera.

Analizar conjuntos de datos complejos con aprendizaje automático

Árboles de decisión

A veces, los datos son tan complejos o cambian tan rápido que es difícil agruparlos o tomar decisiones para casos individuales. Una solución es usar un modelo de aprendizaje automático para segmentar los datos en subconjuntos similares con el fin de facilitar la exploración y el modelado posterior. R le permite a un usuario construir rápidamente árboles de decisión para clasificar datos complejos en base a ejemplos anteriores.

Una vez que su equipo haya escrito un script de R para clasificar los datos usando un modelo de árbol de decisiones, puede usar esa clasificación en cualquier visualización de datos de Tableau para controlar el color, la forma o la agrupación con solo arrastrar y soltar. De hecho, una vez que hayas creado el script de R para clasificar los datos, cualquier usuario corporativo podrá aprovechar los resultados en una visualización de datos.

En la siguiente visualización (figura 3), se utilizó un árbol de decisiones para clasificar las nuevas cuentas en grupos según los gastos previstos de la cuenta durante su vida útil, en función de factores como la ubicación, el tamaño, los ingresos y el incremento. Con el fin de pronosticar los ingresos para el año siguiente, se aplicó un segundo modelo de regresión a cada cuenta nueva, especialmente diseñado para empresas de cualquier tamaño. Al usar Tableau con R, cualquier usuario corporativo puede explorar los datos de esta manera, con la funcionalidad del árbol de decisiones que un científico de datos o un equipo de científicos de datos crearon en R.

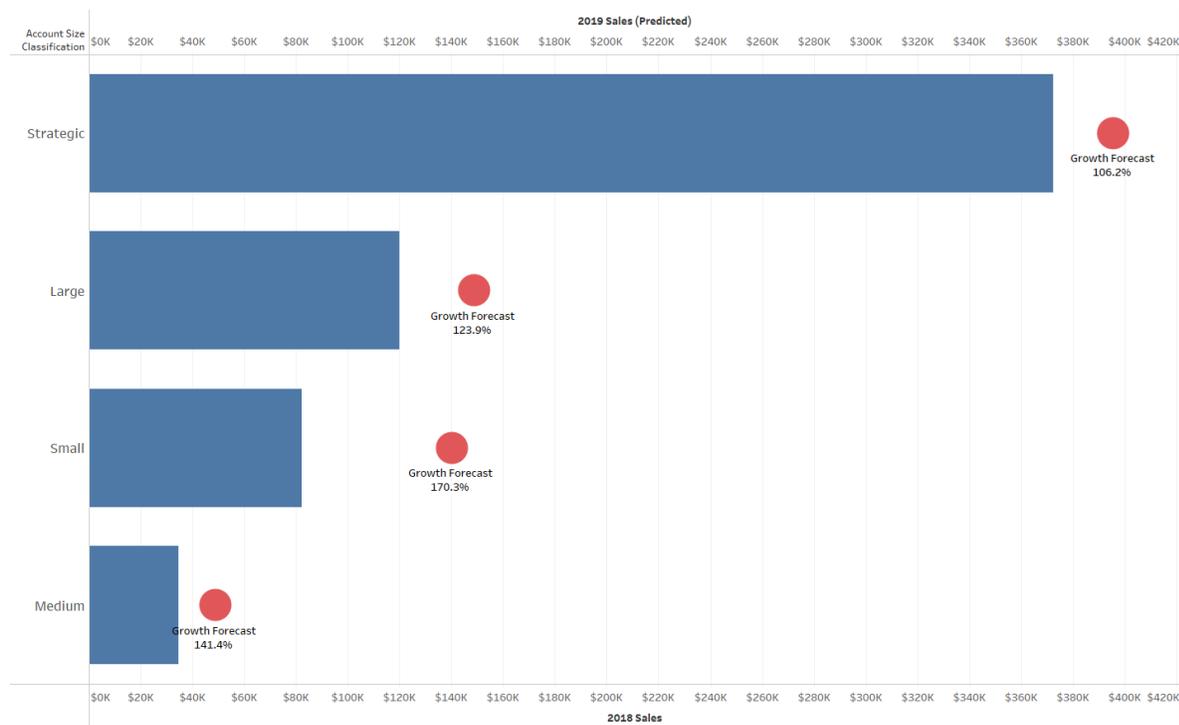


Figura 3: Esta visualización muestra clasificaciones de segmentos de clientes creadas a través de un árbol de decisiones de R. Un usuario corporativo puede combinar fácilmente esos datos con datos de ventas reales y proyectados (en este caso, datos de 2017 y de 2018) para comprender cómo varía el gasto de cada segmento entre los años.

Mostrar que algo no ha sido casualidad

Pruebas estadísticas

Las pruebas estadísticas tratan de mostrar que es poco probable que se obtengan los datos que se ven, simplemente como resultado de una casualidad aleatoria. Por ejemplo, en la investigación farmacéutica, los investigadores le dan a un grupo (grupo A) un placebo y a otro grupo (grupo B) el medicamento real. Si ambos grupos mejoran, pero el grupo B mejora más que el grupo A, usan pruebas estadísticas para determinar si la diferencia en la mejora entre los dos grupos es lo bastante grande como para que sea estadísticamente significativa.

La muestra puede incluir datos de categoría, como el sexo o la etnia, o puede incluir medidas como el peso, la edad o incluso la cantidad de tiempo. Puede usar Tableau para agrupar los datos en tantas categorías como sean relevantes y puede calcular el significado (valores p) para cada una de ellas con solo una línea de código.

R ofrece una amplia variedad de funciones de pruebas estadísticas y Tableau facilita el trabajo de ver los resultados de estas pruebas. Por ejemplo, los parámetros de Tableau hacen que sea sencillo para cualquier usuario ajustar las opciones en pruebas como la **prueba t de Welch de dos muestras** y visualizar el resultado.

Chi-square test of independence

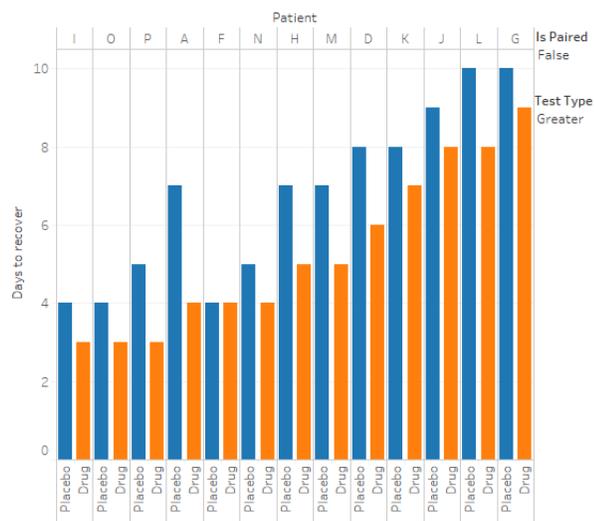
Contingency Table

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

Figura 4: Tableau facilita la visualización de los resultados de las funciones de prueba estadística de R y proporciona filtros que permiten a los usuarios ajustar las opciones de prueba.

Ver cómo cambian los datos a lo largo del tiempo.

Análisis de series temporales

Gran parte de los datos mundiales contienen un elemento temporal: desde la información meteorológica hasta las ventas y las carteras de acciones. El análisis de los datos a lo largo del tiempo tiene una variedad de beneficios, lo que incluye la capacidad de ver tendencias importantes y la estacionalidad.

R ofrece una amplia gama de herramientas para analizar y pronosticar datos de series temporales. Una vez que haya creado un modelo de series temporales en R, puede aprovechar ese modelo para visualizar los datos en Tableau. En Tableau, puede arrastrar y soltar otros campos en la vista para ver cómo otros factores influyeron en las tendencias de los datos. Imagine que hay un repunte en un mes en particular. En Tableau, puede investigar ese repunte y determinar la causa subyacente, sin ninguna codificación adicional.

Tableau también ofrece una función de pronóstico que permite arrastrar y soltar en base al paquete de pronóstico en R. Si usa el paquete de pronóstico, simplemente haga clic en Pronóstico en la pestaña Análisis de Tableau para crear una visualización interactiva. Con la combinación de R y Tableau, tiene la facultad de personalizar el modelo además de la capacidad de interactuar con él mediante filtros y parámetros.

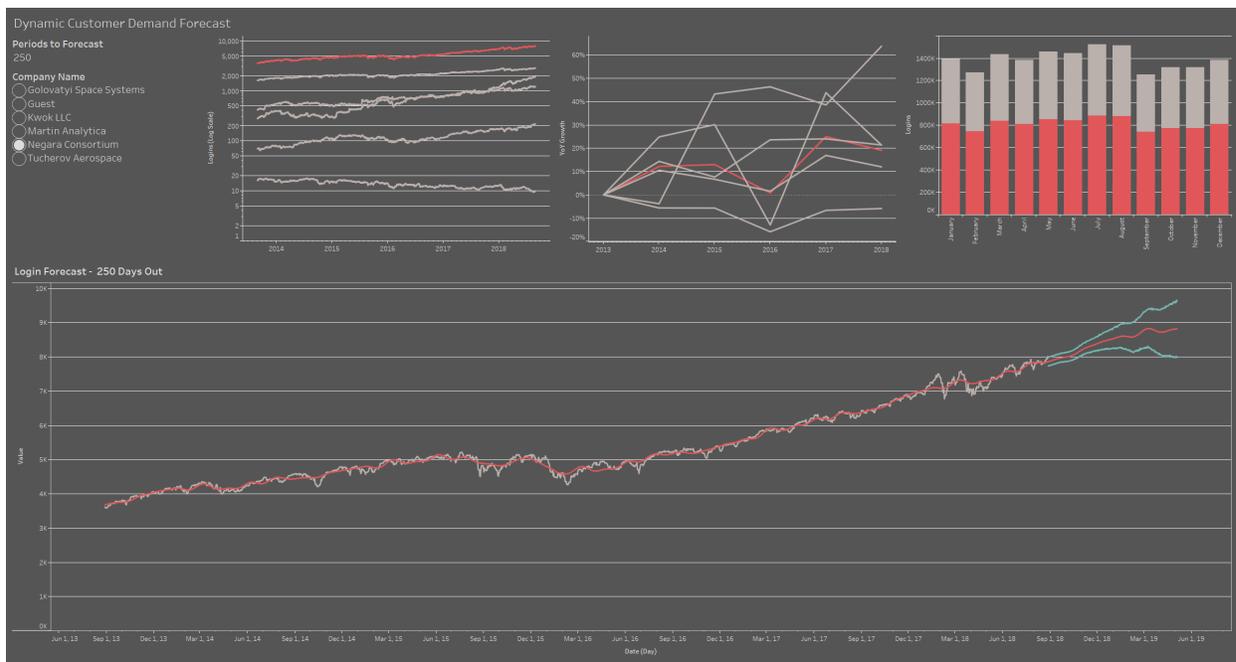


Figura 5: En este ejemplo, los equipos pueden explorar un pronóstico de la demanda de los clientes, con la capacidad de profundizar en la visualización y realizar filtros por nombre de empresa.

Probar diferentes escenarios

Análisis hipotético

El "análisis hipotético" le permite probar las hipótesis para tomar mejores decisiones. Para probar diferentes escenarios de negocios, es posible que desee explorar la manera en que los cambios en un valor particular o en un conjunto de valores afectan al resultado del análisis. La combinación de Tableau y R le permite realizar estas pruebas de manera visual, sin codificación adicional.

Imagine que ha construido un modelo en R para ver la manera en que los diferentes indicadores de producción afectan a las ventas totales. Puedes incluir ese modelo en un dashboard de Tableau con el fin de agregar parámetros para cada uno de estos indicadores. Esto le permite a su equipo ajustar estos indicadores y determinar rápidamente la manera en que estos cambios afectan al resultado general. En la siguiente visualización (figura 6), los equipos pueden ajustar las fechas de inicio y de finalización con el fin de ver la manera en que los cronogramas de producción afectan a las ventas previstas para el año. Esta es otra manera de desplegar el poder de un modelo de R a lo largo de una organización.

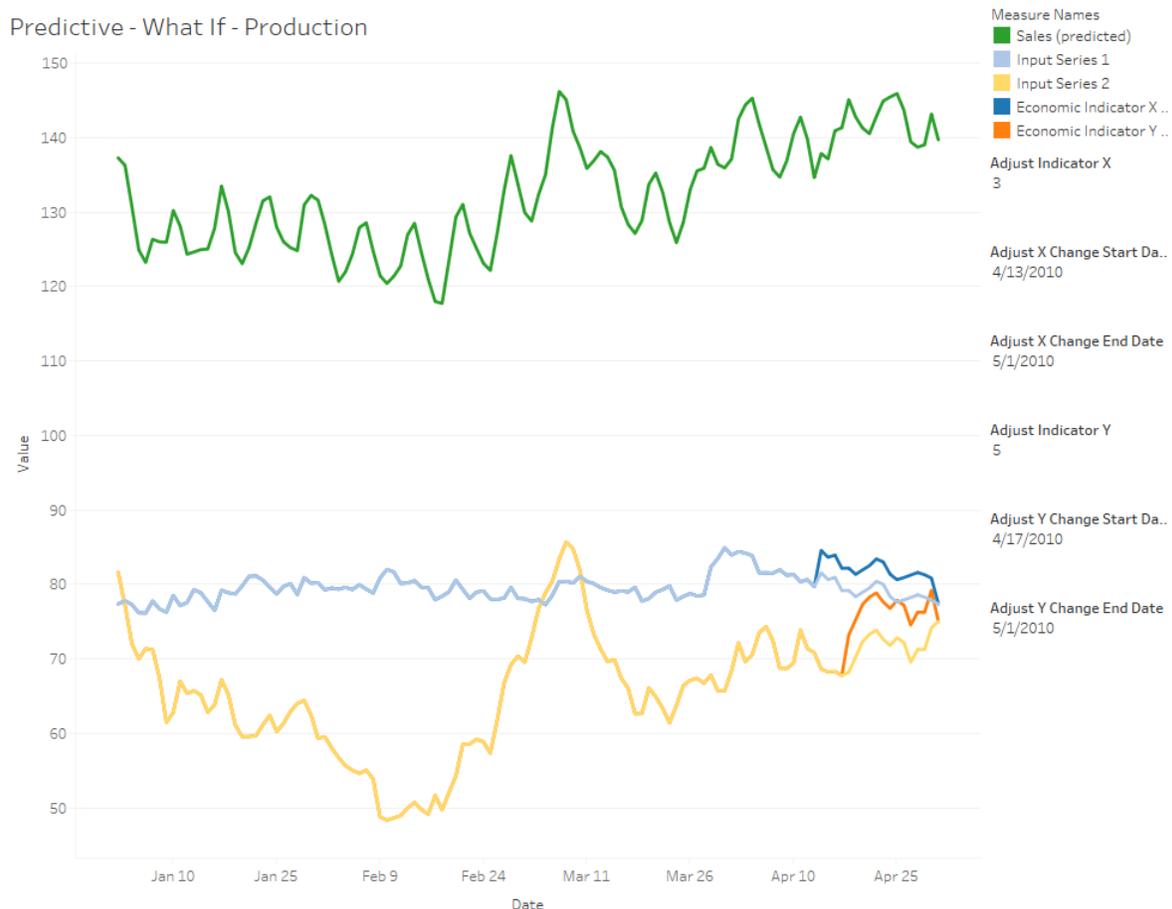


Figura 6: Los parámetros de Tableau permiten a los equipos usar un modelo para probar diferentes escenarios en el contexto de una visualización.

Combine el poder estadístico de R con Tableau para explorar visualmente los datos y realizar ajustes al instante. Comparta los resultados con personas de todos los niveles de aptitud para generar valor a partir de los hallazgos en toda la organización.

Acerca de los autores

Nathan Mannheimer

Nathan es administrador de productos de análisis avanzado de Tableau y trabaja en las aplicaciones e integraciones para apoyar la ciencia de datos y el análisis estadístico. El equipo de análisis avanzado se enfoca en proporcionar herramientas poderosas que permitan a las personas encontrar respuestas a preguntas más profundas, lo que incluye características como la agrupación en clústeres k-means y la integración con Python. Además de Tableau, Nathan es docente en los programas de maestría en Ciencias de Datos y Diseño Centrado en el Humano e Ingeniería de la Universidad de Washington.

Tyler Martin

Tyler es ingeniero sénior de software de análisis avanzado de Tableau y trabaja en el desarrollo de funciones para la ciencia de datos y el análisis estadístico. Antes de incorporarse a Tableau, Tyler obtuvo una maestría de la Universidad de Washington en Matemáticas Aplicadas, donde estudió la dinámica de los fluidos.

Contenido relacionado

[Más información acerca de la integración de R en Tableau](#)

[Cómo usar R con Tableau \(y cuándo debe hacerlo\)](#)

[Análisis avanzado con Tableau](#)

[Tableau y R: preguntas frecuentes](#)

Explorar otros recursos

[Demostraciones de productos](#)

[Capacitación y tutoriales](#)

[Comunidad y soporte](#)

[Historias de clientes](#)

[Soluciones](#)

