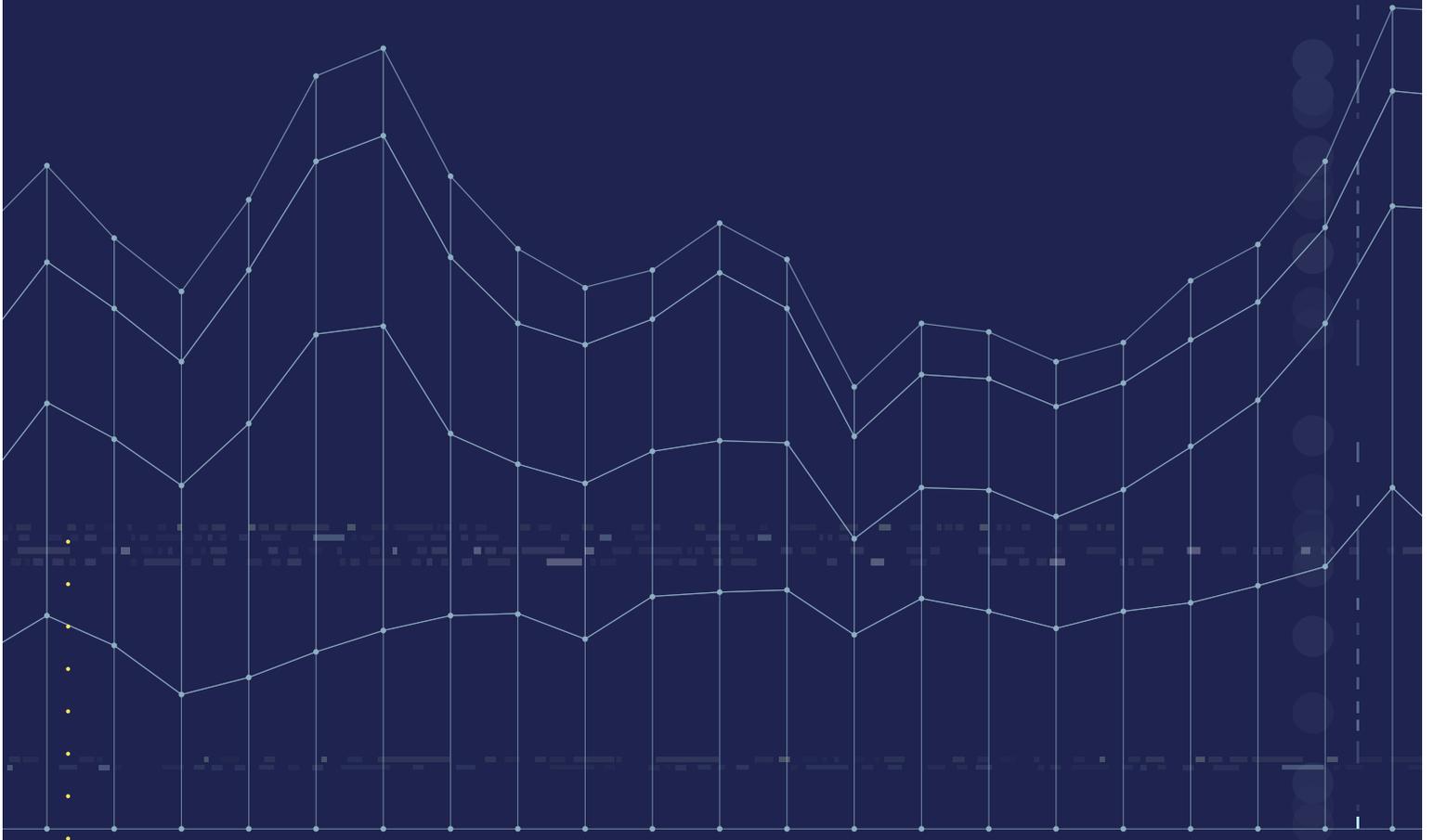




Gestionar el proceso de preparación de datos en la organización

Introducción: Tableau Prep Conductor para una preparación de datos de autoservicio segura y escalable



Este informe está orientado a los profesionales de TI y los responsables de la toma de decisiones de TI. Permite comprender en mayor profundidad las consideraciones que se deben tener en cuenta para la preparación de datos de autoservicio escalable. En él, además, averiguarás cómo Tableau Prep Conductor puede ayudar a gestionar un entorno de preparación de datos gobernado. En este informe, se asume que los usuarios ya cuentan con algún tipo de conocimiento de los productos de Tableau, incluidos Tableau Prep Builder, Tableau Server, Tableau Online y Tableau Desktop.

Contenido

Introducción: ¿Por qué es necesario democratizar la preparación de datos?

| | |
|---|---|
| Desafíos de la preparación de datos de autoservicio en un entorno en constante evolución | 3 |
| Preparar los datos como un experto: Comprender la amplia gama de necesidades de preparación de datos..... | 3 |
| ¿Por qué debemos democratizar la preparación de datos? | 4 |
| Descubrir Tableau Prep Conductor | 5 |

Evaluar la preparación de datos en la organización

| | |
|--|---|
| ¿Quién realiza la preparación de datos? ¿Quién debería realizarla?..... | 7 |
| Preparar datos como un experto: Cubrir las funciones de datos clave dentro de la organización | 8 |
| ¿De dónde provienen los datos? ¿Dónde se almacenan? | 9 |
| ¿Cómo se beneficia la organización con los resultados? | 9 |

Gestionar el proceso de preparación de datos con Tableau Prep Conductor

| | |
|--|----|
| ¿Cómo poner en funcionamiento los flujos de preparación de datos? | 10 |
| Preparar datos como un experto: Automatizar con la API de REST | 11 |
| ¿Cómo se supervisan los flujos de preparación de datos?..... | 11 |
| Preparar datos como un experto: Generar informes personalizados con el repositorio de Tableau Server | 11 |
| ¿Cómo se detectan los inconvenientes? | 12 |
| ¿Cómo aplicar los controles de acceso y mantener los datos seguros de manera escalable? | 13 |

Descubrir el valor de los datos

| | |
|--|----|
| ¿Cómo se identifican las fuentes de datos preparadas en la organización?..... | 14 |
| ¿Cómo ayudar a los miembros de la organización a comprender qué datos tienen a su disposición? | 14 |
| Evaluar la gobernanza escalable | 16 |
| Acerca de Tableau..... | 17 |
| Recursos adicionales | 17 |

Introducción: ¿Por qué es necesario democratizar la preparación de datos?

Desafíos de la preparación de datos de autoservicio en un entorno en constante evolución

En la actualidad, el análisis de autoservicio y las decisiones basadas en datos son la norma en las empresas líder a nivel mundial. Históricamente, la preparación de datos ha sido una función de TI. Pocos miembros de la organización tenían la capacidad de preparar nuevas fuentes de datos y agregarlas al almacén de datos centralizado de la empresa. Sin embargo, el panorama ha cambiado. El volumen y la variedad de datos que las organizaciones recopilan ha aumentado exponencialmente. Los datos no siempre están centralizados, pero, con frecuencia, circulan a través de los procesos de la organización hasta que los casos de uso determinan su obtención e importación, almacenamiento, transformación y distribución. Además, puesto que cada vez más usuarios pueden acceder a las plataformas de análisis, todos ellos tienen la capacidad de tomar decisiones informadas.

Por lo general, los equipos y los miembros de la organización que no pertenecen al departamento de TI o no tienen acceso a los procesos formales de preparación de datos deben esperar que otros equipos preparen los datos o resolver ellos mismos los problemas que tienen con los datos. Con frecuencia, esto significa que los usuarios deben extraer los datos de los sistemas y prepararlos en hojas de cálculo. El resultado, un nuevo conjunto de datos estructurado de tal manera que solo es útil para un único propósito. Por un lado, esto genera una gran cantidad de silos de datos y, por el otro, muchas veces los departamentos duplican sus esfuerzos sin saberlo. Estas soluciones individuales no son eficientes, escalables ni gobernadas.

Dado que el análisis de autoservicio se ha convertido en la nueva norma para las organizaciones basadas en los datos, muchas personas preparan los datos con las herramientas y las funcionalidades que tienen disponibles a fin de que estos resulten útiles. Entre otros métodos, copian y pegan datos o escriben extensos cálculos que no son los adecuados para el servidor. Incluso los analistas indican que no dedican la mayor parte de su tiempo a analizar los datos, sino a limpiarlos y darles forma, ya sea mediante un proceso de ETL, a través de herramientas de preparación de datos de autoservicio o incluso con herramientas de hoja de cálculo como Excel.

Preparar datos como un experto: Comprender la amplia gama de necesidades de preparación de datos

Ciertos factores, como los errores humanos, la diversidad de sistemas y los requisitos empresariales que cambian constantemente, pueden contribuir a generar datos “sucios”. Sin embargo, la preparación de datos, con frecuencia, implica mucho más que simplemente limpiar los datos. Es posible que los usuarios deban ajustar la granularidad de los datos y transformarlos para adecuarlos a otros datos, combinarlos o unirlos. Esto significa que cuando los datos están listos para ser analizados, con frecuencia, son muy diferentes de la fuente de datos original. El proceso para limpiar los datos, darles forma y optimizarlos para el análisis incluye, entre otros, los pasos que se detallan a continuación.

Pivotar: mover un campo de Columnas a Filas o viceversa.

Unir columnas: agregar más campos a la fuente de datos y ampliar la cantidad de campos para el análisis.

Unir filas: combinar dos conjuntos de datos, conservar la estructura de la tabla y agregar más filas.

Filtrar o eliminar: excluir valores o campos para el análisis.

Asignar funciones de datos: validar campos que representan un correo electrónico, una URL o datos geográficos.

Editar valores: modificar manualmente un valor o usar operaciones de limpieza rápida para editar las mayúsculas y minúsculas en el texto o eliminar letras, números, signos de puntuación o espacios, entre otros.

Agrupar y reemplazar: limpiar valores con diferencias de pronunciación, caracteres comunes o escritura.

Dividir valores: dividir unidades de información de un campo en varios campos.

Crear campos calculados: crear campos para el análisis con cálculos basados en otros valores.

Agregar datos: obtener un único valor a partir de varios valores, por ejemplo, suma, promedio, recuento o mínimo.

Obtenga más información acerca de los [datos “sucios” y cómo resolver los problemas más comunes de preparación de datos](#).

¿Por qué debemos democratizar la preparación de datos?

Cada vez más personas colaboran con los datos, desglosan los silos de datos en toda la organización y descubren nueva información útil. Al mismo tiempo, podemos afirmar que cada vez más conjuntos de datos, de la manera en la que están recopilados, no están listos para el análisis o no son los adecuados. En muchos casos, existe una brecha significativa entre el estado óptimo de los datos para una recopilación eficiente y para un análisis eficaz. Es lo que ocurre, por ejemplo, con los datos transaccionales o los datos de transmisión de alta velocidad. Ya sea por la estructura, el formato o la falta de contexto empresarial, es necesario limpiar los datos y, en ocasiones, seleccionarlos. Por ejemplo, cuando se deben agregar las reglas del negocio o los tipos de reclamación a los datos de atención médica antes del análisis.

Muchas organizaciones implementan soluciones de preparación de datos de autoservicio a fin de explorar las nuevas fuentes de datos y los casos de uso de análisis y diseñar prototipos. Las herramientas de preparación de datos de autoservicio ponen los datos en manos de quienes están más familiarizados con ellos. Además, reducen la carga de trabajo de TI para prepararlos. Sin embargo, la preparación de datos de autoservicio continúa siendo un conjunto de habilidades relativamente nuevo. Debe desarrollarse e implementarse de modo que permita a los usuarios comprender las funciones de preparación y realizarlas de manera eficaz, establecer procesos repetibles, automatizarlos para lograr mayor eficiencia y, en última instancia, generar confianza en los datos para un uso más amplio.

¿Vale la pena todo el trabajo de preparación de datos? Según un estudio reciente acerca de la preparación de datos, esta tiene multitud de beneficios que superan ampliamente las expectativas de las empresas. En concreto, obtener una vista única y completa de los datos relevantes de toda la organización, reducir los silos de análisis y mejorar el proceso de toma de decisiones basadas en los datos.

Explora el informe gratuito de BARC, "Preparación de datos: convertir los datos sin procesar en datos útiles" (en inglés).

Descubrir Tableau Prep Conductor

Los avances que Tableau logró en el análisis visual, ahora los está aplicando a la preparación de datos. Tableau Prep Builder, presentado en la primavera de 2018, facilita aún más el acceso a los datos a través de una preparación de datos visual, inteligente y directa. Los analistas y los usuarios corporativos pueden preparar sus propios datos para el análisis. Además, está completamente integrado con Tableau Desktop. De esta manera, los usuarios se mantienen en el flujo del análisis. Con Tableau Prep Conductor, ampliamos las funcionalidades de preparación de datos de la plataforma de Tableau. Así, los usuarios podrán automatizar los flujos de preparación de datos a fin de que se ejecuten sin que sea necesario realizar actualizaciones manuales cuando se modifican los datos. Todos los miembros de la organización, además, podrán descubrir los datos preparados más fácilmente.

Tableau Prep Conductor permite programar los flujos para que se ejecuten en un entorno de servidor centralizado, escalable y fiable. Así, los datos siempre se mantendrán actualizados. Además, ofrece visibilidad a los administradores en relación con la preparación de datos de autoservicio en toda la organización. Con Tableau Prep Conductor, se pueden gestionar y supervisar los flujos y garantizar su seguridad a través del entorno de servidor de Tableau.

Tableau Prep Conductor está integrado en Tableau Server y Tableau Online. De esta manera, se pueden aprovechar las funcionalidades existentes de programación, seguimiento y seguridad. Al igual que las actualizaciones de extracciones, las tareas programadas de los flujos y las ejecuciones solicitadas de los flujos se ponen en cola como tareas en segundo plano. Podrás publicar flujos desde Tableau Prep Builder en Tableau Server o Tableau Online sin inconvenientes con una funcionalidad similar a la que se usa en la actualidad para publicar fuentes de datos y libros de trabajo con Tableau Desktop.

Mantener los datos actualizados automáticamente

Programa los flujos para que se ejecuten cuando lo necesites. Automatiza las tareas de ejecución de los flujos y desarrolla un proceso repetible a fin de ofrecer datos preparados sistemáticamente.

Obtener información con las alertas y el historial de ejecuciones

Explora en una vista el historial de ejecuciones de los flujos de trabajo y determina a simple vista las ejecuciones que se completaron correctamente y las que no. Haz un seguimiento de la calidad de los flujos de preparación y recibe alertas estándar si un flujo no se completa correctamente.

Crear un entorno de preparación de datos gobernado

Determina las reglas y los permisos para compartir los datos y actualizarlos. Aprovecha los permisos y la infraestructura actual de Tableau Server y Tableau Online para controlar qué usuarios pueden publicar, ver y ejecutar los flujos.

Incrementar el descubrimiento de los datos

Usa las funcionalidades de gestión sencillas, como agregar etiquetas con palabras clave, mover los flujos de un proyecto a otro y configurar los permisos de los usuarios, para que los miembros de toda la organización puedan encontrar los datos relevantes y preparados que necesitan.

Home / Default / Superstore Beta Flow

Superstore Beta Flow

FLOW · By Youssef Shoukry · ☆ 0

Overview | Connections | Scheduled Tasks | Run History

Description
Published to Prep Conductor
Tags No tags set on this flow.

| Output Step | Output Name | Status | Schedule | Errors |
|--|---|---------------------------------|----------------|--------|
| Create 'Annual Regional Performance.tde' | Annual Regional Performance Beta | Succeeded: Nov 5, 2018, 9:54 AM | Weekday 2:00AM | |
| Create 'Superstore Sales.tde' | Superstore Sales Beta (not yet published) | Never run | Weekday 2:00AM | |

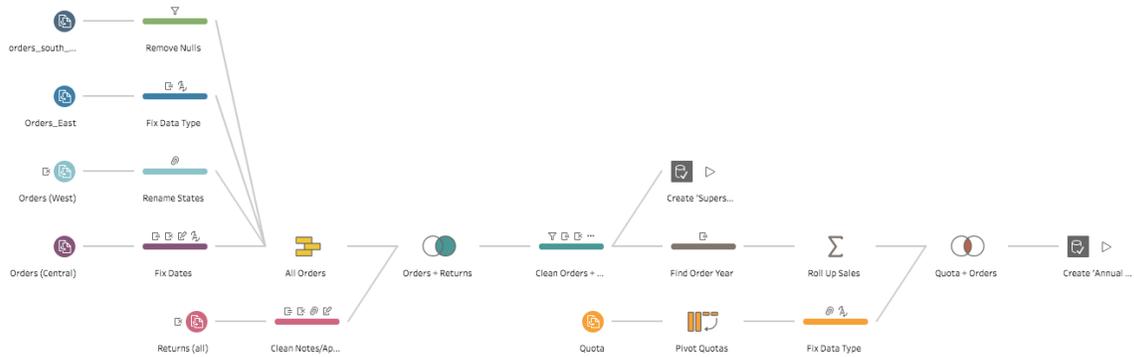


Figura 1 Visualiza los flujos de preparación de datos publicados en Tableau Server con Tableau Prep Conductor.

Evaluar la preparación de datos en la organización

A fin de conocer el contexto completo de los datos, no solo es importante comprender quién usa los datos, sino también quién los prepara, de dónde provienen, dónde están disponibles para el análisis y cómo estos pueden beneficiar a los usuarios finales. Comprender esto es la base de la preparación de datos de autoservicio escalable.

¿Quién realiza la preparación de datos? ¿Quién debería realizarla?

En muchas organizaciones, los analistas preparan los datos a fin de que otros usuarios puedan consumirlos con la misma frecuencia que lo hacen para llevar a cabo sus propios análisis. Las herramientas avanzadas de preparación de datos pueden resultar complejas. Esto implica que, muchas veces, su uso está limitado a un grupo selecto de usuarios avanzados. Sin embargo, que los analistas y los usuarios corporativos no tengan acceso a las herramientas de preparación de datos no significa que no lleven a cabo estas tareas en otras aplicaciones.

Las herramientas de inteligencia de negocios de autoservicio ofrecen funcionalidades de análisis de datos para usuarios con diferentes niveles de conocimientos. No obstante, para obtener información a partir de los datos, estos usuarios aún deben contar con datos estructurados correctamente por los departamentos de TI. Al considerar la preparación de datos de autoservicio escalable, hay que tener en cuenta las funciones existentes y las responsabilidades que se deben compartir o agregar.

La función del administrador de datos puede ser adecuada para una función más tradicional, como el administrador de base de datos. En la actualidad, el administrador de datos está más dedicado a que los datos que se distribuyen en toda la organización (analistas, usuarios corporativos, etc.) sean fiables. De esta manera, los usuarios podrán obtener valor a partir de los datos. Los administradores de bases de datos y los ingenieros de datos priorizan generalmente cómo se almacenan los datos y cómo se accede a ellos. Es posible agregar columnas para que solo las usen las bases de datos, no los usuarios. Cuando un ingeniero crea un almacén de datos específicamente para el análisis, prioriza las principales métricas que responden la mayoría de las preguntas del negocio. Si la información que necesitan los analistas de datos no se encuentra en el conjunto de datos, es posible que deban ajustar las agregaciones o incorporar fuentes de datos externas. Esto puede dar lugar a datos incorrectos o silos de datos.

A medida que se adapta la preparación de datos de autoservicio, es necesario tener en cuenta quiénes serán los responsables de establecer y acordar las prácticas de gobernanza y hacerlas cumplir. Además, es preciso definir quiénes usarán qué herramientas para la preparación de datos, qué formación es necesaria y cómo se medirá el éxito en la organización.

Preparar datos como un experto: Cubrir las funciones de datos clave dentro de la organización

Ciertos factores, como los errores humanos, la diversidad de sistemas y los requisitos empresariales que cambian constantemente, pueden contribuir a generar datos “sucios”. Sin embargo, la preparación de datos, con frecuencia, implica mucho más que simplemente limpiar los datos. Es posible que los usuarios deban ajustar la granularidad de los datos y transformarlos para adecuarlos a otros datos, combinarlos o unirlos. Esto significa que cuando los datos están listos para ser analizados, con frecuencia, son muy diferentes de la fuente de datos original. El proceso para limpiar los datos, darles forma y optimizarlos para el análisis incluye, entre otros, los pasos que se detallan a continuación.

Funciones de los profesionales de TI/BI

El **administrador de bases de datos (DBA)** es el responsable de administrar, supervisar y mantener las bases de datos y garantizar su seguridad en la organización. En coordinación con los ingenieros de datos y los administradores de datos, los DBA conceden el acceso a los datos. Además, ofrecen asistencia y modelan, estructuran y optimizan las fuentes de datos que se conectan a los productos de Tableau.

Los **administradores de sistemas** instalan, configuran, administran y mantienen el hardware y los sistemas operativos en los que se instala Tableau Server, ya sea en el centro de datos o en la nube. Además, aplican las políticas de la empresa en cumplimiento con las estrategias comerciales y de tecnología.

Los **administradores de cliente** configuran el software del cliente. Entre otras tareas, instalan los controladores de bases de datos y los productos de Tableau. Además, **habilitan Tableau Prep Conductor** dentro de Tableau Server o Tableau Online.

Funciones de los administradores de Tableau

El **administrador de Tableau Server** tiene acceso completo a las configuraciones de Tableau Server, todos los sitios del servidor, los usuarios y los grupos y todos los recursos de contenido, como los proyectos, las fuentes de datos y los libros de trabajo. De esta manera, puede supervisar y mantener el funcionamiento general de Tableau Server.

Los **administradores de sitios de Tableau** son los responsables de crear y mantener los usuarios y los grupos de los sitios, generar proyectos para organizar el contenido del sitio y asignar permisos para que los usuarios y los grupos puedan acceder al contenido. Además, promueven y certifican el contenido y miden el uso del análisis en el sitio.

Funciones de los creadores de contenido

Los **administradores de datos** (con licencia de Tableau Creator) comprenden el ámbito empresarial y la interacción de los procesos del negocio con el análisis. Garantizan la disponibilidad de procedimientos documentados y pautas para acceder a los datos y usarlos. Además, trabajan conjuntamente con los administradores de bases de datos o los ingenieros de datos a fin de planificar y ejecutar una implementación empresarial de gobernanza de datos y una política de cumplimiento. Los administradores de datos pueden publicar flujos de preparación de datos o fuentes de datos.

Los **autores de contenido** (con licencia de Tableau Creator) crean y publican dashboards, flujos de preparación de datos o fuentes de datos.

¿De dónde provienen los datos? ¿Dónde se almacenan?

En la actualidad, y aunque parezca sorprendente, muchas organizaciones no tienen visibilidad acerca de cómo se llevan a cabo los procesos de preparación de datos. Sin una metodología estándar y gobernada, el trabajo de preparación de datos y análisis ad hoc puede acabar en esfuerzos duplicados, trabajo manual sin procesos repetibles e incoherencias en las fuentes de datos. Un factor clave para solucionar estos problemas consiste en comprender de dónde provienen los datos y, una vez que están limpios, dónde estarán disponibles. Básicamente, es necesario determinar la relación entre la persona que prepara los datos y los usuarios que llevan a cabo el análisis.

- ¿Cómo se garantiza la seguridad de los datos? ¿Quiénes deben tener permisos para acceder a los datos y darles forma?
- ¿Qué usuarios acceden a las fuentes de datos sin procesar? ¿Qué usuarios acceden a los datos limpios?
- ¿Los usuarios necesitan acceder a fuentes de datos combinadas o datos externos para llegar a la raíz de sus inquietudes o realizar análisis más eficaces?
- ¿Cómo se comparten las fuentes de datos preparadas con otros usuarios para el análisis?

No resulta extraño que los usuarios exporten datos a un archivo CSV u otro archivo de hoja de cálculo para limpiar los datos y llevar a cabo un análisis ad hoc. Sin embargo, la seguridad de los datos puede verse perjudicada, ya que es posible que los archivos no se compartan de manera segura. Las organizaciones que usan Tableau Server o Tableau Online como repositorio de sus fuentes de datos y libros de trabajo pueden beneficiarse con Tableau Prep Conductor. Este facilita la visualización de los procesos de preparación de datos a través de flujos de preparación publicados en Tableau Server o Tableau Online. De esta manera, los usuarios pueden consultar los flujos de preparación de datos y acceder a ellos desde un único lugar. Además, permite explorar la integridad de los flujos y ofrece la oportunidad de obtener más información acerca de las prácticas recomendadas de preparación de datos de autoservicio.

¿Cómo se beneficia la organización con los resultados?

En la actualidad, y aunque parezca sorprendente, muchas organizaciones no tienen visibilidad acerca de cómo se llevan a cabo los procesos de preparación de datos. Sin una metodología estándar y gobernada, el trabajo de preparación de datos y análisis ad hoc puede acabar en esfuerzos duplicados, trabajo manual sin procesos repetibles e incoherencias en las fuentes de datos. Un factor clave para solucionar estos problemas consiste en comprender de dónde provienen los datos y, una vez que están limpios, dónde estarán disponibles. Básicamente, es necesario determinar la relación entre la persona que prepara los datos y los usuarios que llevan a cabo el análisis.

- ¿Cómo se recopilan los requisitos para las fuentes de datos y los informes?
- ¿Qué tipo de preguntas es necesario formular y responder?
- ¿Cuáles son las prioridades empresariales estratégicas de los usuarios que acceden a los datos?
- ¿Está equilibrada la necesidad de ofrecer respuestas inmediatas a las preguntas conocidas con la posibilidad de una exploración más profunda?
- ¿Qué procesos se implementan para garantizar la calidad de los flujos de datos y las fuentes de datos publicadas (por ejemplo, control de calidad o certificación de las fuentes de datos)?

Gestionar el proceso de preparación de datos con Tableau Prep Conductor

A fin de que sea posible adaptar con éxito la preparación de datos de autoservicio, los usuarios y la tecnología deben trabajar conjuntamente bajo un marco de gobernanza que equilibre el control de TI con la flexibilidad y la agilidad que exige el negocio. TI puede así centrar su atención en los procesos de habilitación a través de la puesta en funcionamiento de flujos automatizados y la supervisión del uso, el rendimiento y el acceso. De esta manera, puede garantizar que las prácticas de preparación de datos se adapten de manera eficaz y eficiente.

¿Cómo poner en funcionamiento los flujos de preparación de datos?

En ciertos casos, por ejemplo, para la exploración ad hoc de conjuntos de datos pequeños, sencillos o limpios, las funciones de preparación de datos de Tableau Desktop como pivotar u ocultar columnas pueden ser suficientes. Sin embargo, cuando se trata de conjuntos de datos complejos y de mayor tamaño o flujos que alimentan dashboards clave, las fuentes de datos deben estar actualizadas en todo momento a fin de ofrecer datos fiables para la toma de decisiones. Según el caso de uso, es posible que haya que ir más allá de las actualizaciones de extracciones programadas con Tableau Server. En lugar de actualizar la extracción, automatizar las ejecuciones de los flujos de preparación permitirá que se implementen los pasos de limpieza de datos necesarios. De esta manera, se obtendrán fuentes de datos útiles para el análisis.

Si los usuarios de la empresa ya utilizan Tableau Prep Builder para limpiar los datos, Tableau Prep Conductor permitirá beneficiarse automáticamente de este trabajo ya realizado. Es posible programar los flujos para que se ejecuten en un momento específico o de manera periódica con la funcionalidad de tareas programadas de Tableau Server. También se pueden implementar tareas de flujo a fin de que se ejecuten según un programa predeterminado en Tableau Online. Incluso es posible ejecutar nuevamente los flujos en horas no laborables, sin que sea necesario que un usuario los inicie. Esto permite ahorrar tiempo y costes. Además, los usuarios pueden beneficiarse de un entorno de servidor estable, en lugar de tener que confiar en sus propios recursos de Tableau Desktop para ejecutar los flujos.

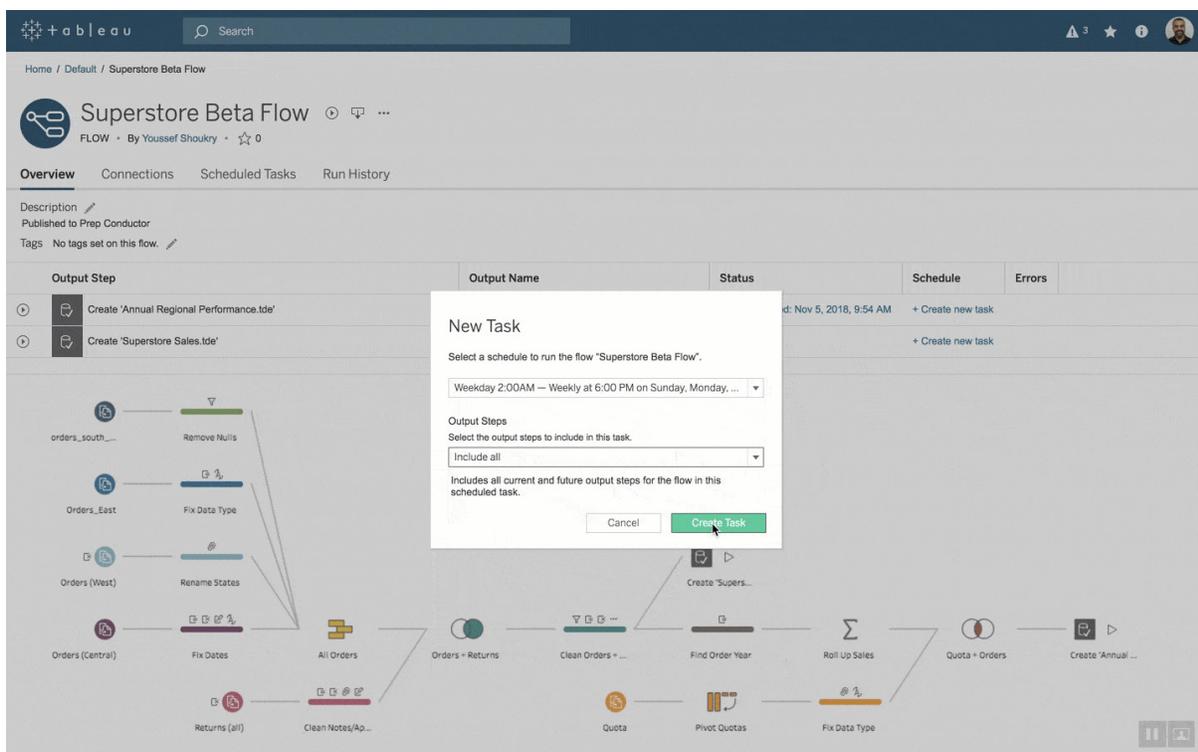


Figura 2 Crear una tarea de flujo a fin de ejecutar los flujos de preparación de datos según un programa establecido.

Preparar datos como un experto: Automatizar con la API de REST

Crear flujos de trabajo con sistemas de terceros y conectar los procesos de datos de una manera eficaz es posible con la API de REST. Esta API permite publicar, programar, descargar y consultar las conexiones de los flujos, ejecutar los flujos y las tareas de flujo según demanda, o gestionar los permisos, entre otros.

Aquí obtendrás más información acerca de las [funcionalidades de la API de REST con Tableau Prep Conductor](#).

¿Cómo se supervisan los flujos de preparación de datos?

La supervisión es un área en la que TI realmente puede ayudar a la organización a medida que se adaptan las prácticas de preparación de datos. El control del uso y el rendimiento, la automatización y la creación de alertas permiten garantizar que los datos de la organización se mantengan actualizados y seguros, ya que se preparan con flujos eficientes.

Con Tableau Prep Conductor, los administradores pueden supervisar los flujos con las mismas herramientas que están disponibles en Tableau Server en la actualidad. Esto incluye Tableau Services Manager (TSM), la página de estado y las vistas administrativas estándar. Estas vistas integradas como parte de Tableau Server y Tableau Online permiten responder preguntas importantes acerca de las prácticas de preparación de datos de la organización. Hay que tener en cuenta que no todas las vistas en Tableau Server están disponibles o son relevantes en Tableau Online.

- **Rendimiento de las ejecuciones de flujo:** comprender qué tareas de flujo están programadas o ejecutándose actualmente, cuál es la duración de las tareas de flujo, qué tareas de flujo se ejecutan con mayor frecuencia y qué flujos se llevan a cabo ad hoc o programados.
- Acciones de **todos los usuarios**, de **usuarios específicos** o de **usuarios recientes**: estas últimas pueden resultar particularmente útiles cuando es necesario realizar el mantenimiento del servidor y se quiere saber qué están haciendo los usuarios en el servidor y cómo podría afectarlos el mantenimiento.
- **Estadísticas de uso de espacio:** identificar cuáles son los resultados de los flujos que ocupan más espacio en el disco del servidor.
- **Retrasos de las tareas del procesador en segundo plano:** esta vista permite identificar cómo es posible mejorar el rendimiento del servidor al optimizar las tareas y distribuir las tareas programadas.

Preparar datos como un experto: Generar informes personalizados con el repositorio de Tableau Server

Además de las vistas administrativas prediseñadas, se puede usar Tableau Desktop para hacer consultas y generar análisis propios de la actividad del servidor. Para hacerlo, tan solo hace falta conectarse a las vistas y realizar las consultas en el repositorio de Tableau Server, una base de datos PostgreSQL.

Aquí obtendrás más información acerca de cómo crear vistas personalizadas para Tableau Server en [Windows](#) o [Linux](#).

¿Cómo se detectan los inconvenientes?

Ya se trate de problemas de conexión o errores en un flujo, hay que estar al corriente de los problemas que pueden retrasar los flujos de preparación de datos y solucionarlos. A fin de minimizar el riesgo de usar datos obsoletos, Tableau Prep Conductor no solo notifica los problemas que se producen durante la ejecución de los flujos, sino que también sugiere posibles soluciones para estos errores.

- **Historial de ejecuciones:** los usuarios pueden consultar las actualizaciones históricas de un flujo a fin de explorar a simple vista las ejecuciones de los flujos que se completaron correctamente y las que no. Esto permite evaluar la calidad de todos los flujos y estar más seguro de que los datos son precisos.
- **Alertas:** las alertas oportunas informarán de si los flujos se están ejecutando correctamente. Si se produce un error, los usuarios recibirán una notificación por correo electrónico y una alerta en la interfaz de Tableau Server. Los enlaces dentro de la notificación de la alerta permiten a los analistas y administradores de datos tomar las medidas necesarias rápidamente y examinar los errores. Además, se ofrecen posibles soluciones a fin de que puedan solucionarlos y continuar con la preparación o el análisis.
- **Páginas de estado:** las páginas de estado de Tableau Server y Tableau Services Manager incluyen los [procesos de Tableau Server](#), junto con enlaces a documentación acerca de cómo solucionar el problema si un proceso no se ejecuta de la manera esperada. Al situar el cursor sobre el indicador de estado de un proceso, se mostrará una descripción emergente con el nombre del nodo y el puerto en el que se está ejecutando el proceso.

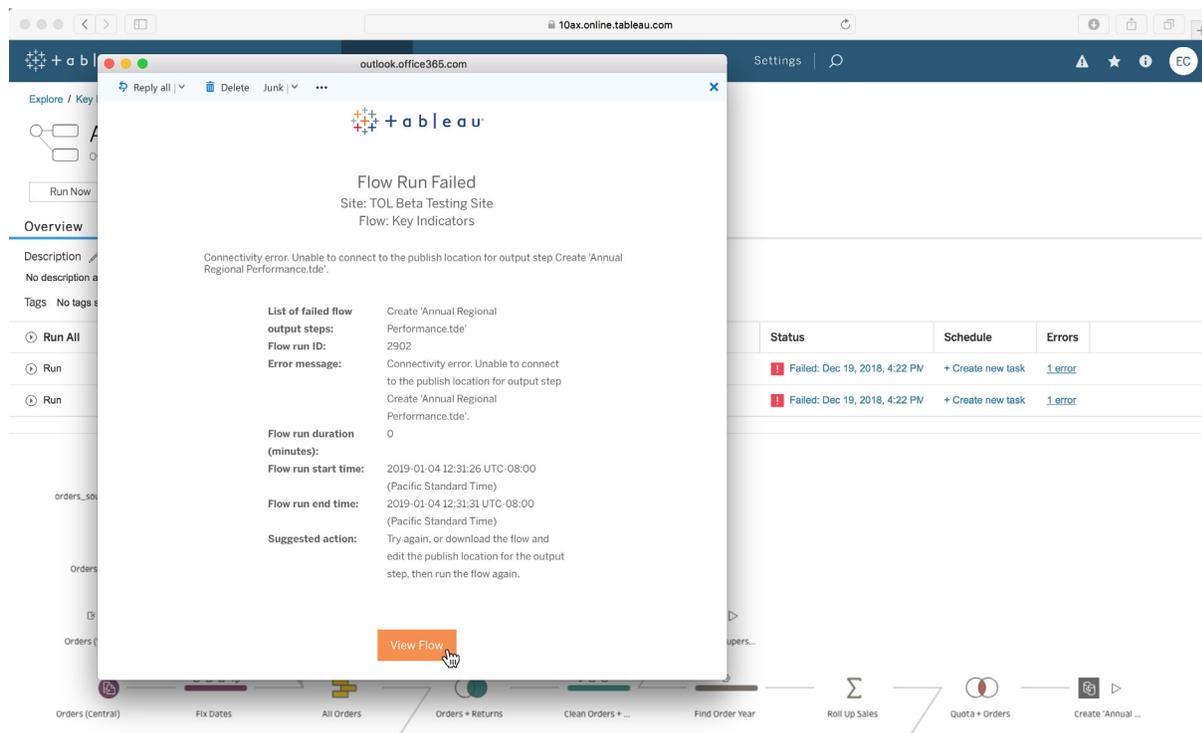


Figura 3 Recibir alertas de una ejecución incorrecta de un flujo por correo electrónico y en Tableau Online.



Una herramienta de preparación de datos debe poder responder las preguntas excepcionales de los analistas y ofrecer al mismo tiempo procesos repetibles.

GORDON STRODEL

CONSULTOR DE ANÁLISIS Y GESTIÓN DE LA INFORMACIÓN, SLALOM

¿Cómo aplicar los controles de acceso y mantener los datos seguros de manera escalable?

Un aspecto clave de la supervisión y la gobernanza de los flujos consiste en garantizar que los usuarios correctos tengan acceso a los flujos de preparación de datos. Los administradores de TI pueden ahorrar tiempo y esfuerzo a través de un control consolidado de los permisos de Tableau Server o Tableau Online. Para los flujos de preparación de datos gestionados por Tableau Prep Conductor, es posible establecer los permisos al publicar el flujo. De esta manera, se puede configurar qué usuarios están autorizados para ver el flujo, editarlo o ejecutarlo, entre otros. Si un flujo se conecta a una base de datos, es posible especificar el tipo de autenticación y configurar las credenciales de acceso a los datos.

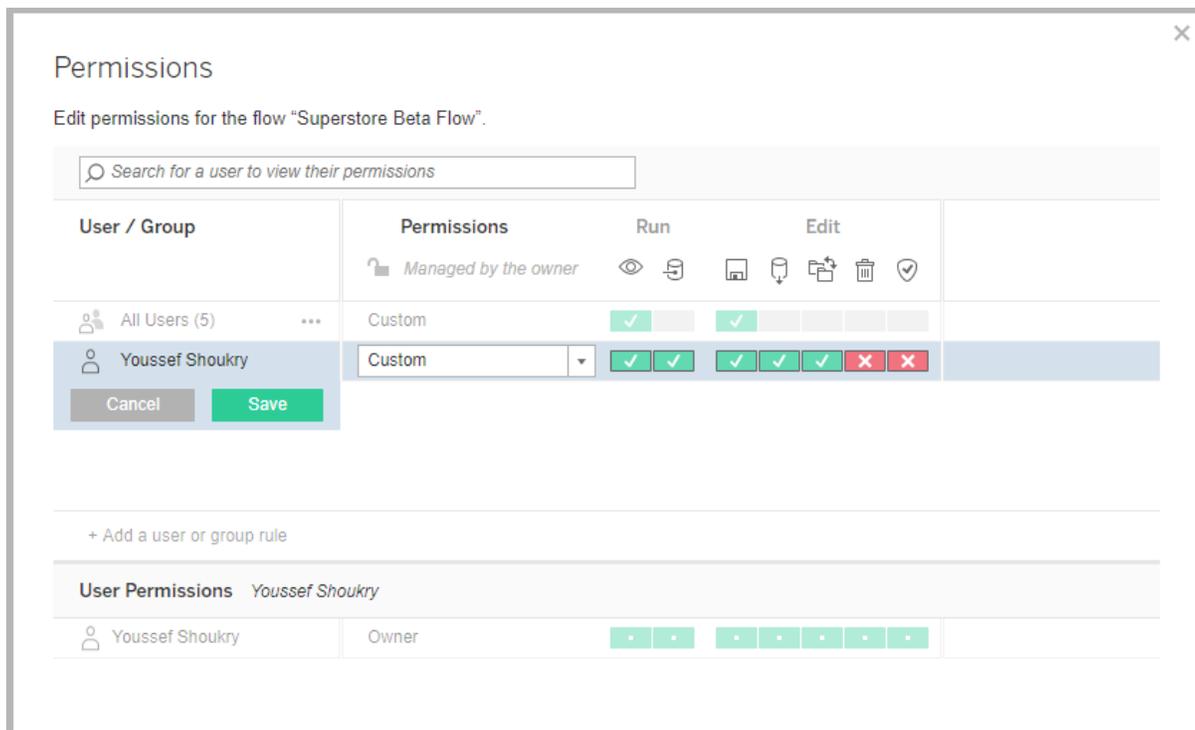


Figura 4 Gestionar y personalizar los permisos de los flujos con una interfaz sencilla y conocida.

Además, las vistas administrativas ofrecen informes en los que se pueden explorar los permisos y las acciones de un usuario específico o de todos los usuarios. De esta manera, se pueden identificar las acciones que están llevando a cabo los usuarios en cada flujo.

Descubrir el valor de los datos

Los miembros de la organización obtendrán más valor de los datos preparados si es fácil acceder a ellos y comprender su relevancia. Además, es importante que estén seguros de que los datos sean fiables, en especial, a medida que se adaptan las prácticas de datos y estos están disponibles para una mayor cantidad de usuarios. En estos casos, la preparación de datos de autoservicio permite potenciar y expandir los beneficios del análisis de autoservicio en un entorno gobernado, como Tableau Server o Tableau Online.

¿Cómo se identifican las fuentes de datos preparadas en la organización?

Muchas de las funcionalidades de gestión de Tableau Prep Conductor permiten a los usuarios descubrir datos relevantes y fiables con mayor facilidad. Para hacerlo, tan solo deben buscar o filtrar los flujos o las fuentes de datos publicados en Tableau Server o Tableau Online.

Etiquetas: agregar palabras clave a los flujos, al igual que lo hace para los libros de trabajo, a fin de que los usuarios puedan encontrar, filtrar y categorizar el contenido. Es posible incluir etiquetas en los flujos individuales o en una selección de varios flujos a la vez.

Organización: implementar los flujos de un proyecto en otro a fin de mantenerlos organizados con las fuentes de datos, los libros de trabajo y otros contenidos relevantes. Los propietarios de los flujos pueden llevar a cabo esta acción de manera predeterminada. Sin embargo, los administradores deben conceder los permisos para mover los flujos a los demás usuarios.

Gestión del ciclo de vida: las funcionalidades de administración adicionales permiten garantizar que los flujos de preparación de datos sean relevantes y se mantengan organizados, e incluyen los permisos para guardar los flujos, cambiarles el nombre o eliminarlos. Los administradores, los propietarios de los flujos y los propietarios de proyectos pueden reasignar la propiedad de los flujos.

La funcionalidad de búsqueda de flujos de datos disponible en Tableau Server y Tableau Online permite evitar la redundancia y las operaciones de preparación de datos duplicadas. Si un usuario es capaz de encontrar una fuente de datos y encuentra un flujo de preparación de datos adecuado para el análisis que desea realizar, no será necesario que dedique tiempo a recrear los pasos de preparación o ejecutar el flujo. Incluso es posible que encuentre un flujo existente que puede descargar y modificar a fin de aplicarlo en un caso de uso similar, sin tener que crear el flujo de preparación desde cero.

¿Cómo ayudar a los miembros de la organización a comprender qué datos tienen a su disposición?

El conocimiento de los datos es una inversión necesaria para cualquier organización que desea adaptar la preparación de datos y el análisis de autoservicio. En muchas organizaciones, los expertos de TI y análisis colaboran en un centro de excelencia. Este integra los recursos y el soporte, un grupo de usuarios interno y la inversión en formación y desarrollo destinados a adquirir las habilidades de análisis. La organización debe evaluar estas necesidades según el tipo y el volumen de usuarios que utilizarán las herramientas como Tableau Prep Builder y Tableau Prep Conductor para la preparación de datos de autoservicio.

Los analistas y los usuarios corporativos están más seguros de que cuentan con datos fiables si pueden averiguar de inmediato cómo se creó la fuente de datos de Tableau. Con Tableau Prep Conductor, cualquier usuario puede consultar el origen de una fuente de datos creada a partir de un flujo y acceder directamente a esta para conocer cómo se generó. Mediante un conocimiento básico de los datos y la capacidad de comprender los pasos del flujo, los usuarios pueden determinar si la fuente de datos les resultará útil o no.

A fin de habilitar a los usuarios con licencias Creator y Explorer, es decir, los usuarios con licencias de Tableau que pueden conectarse a las fuentes de datos, crear contenido nuevo o explorar y personalizar el contenido existente, TI y los administradores de datos deben establecer un proceso para certificar las fuentes de datos que utilizan sistemáticamente los flujos de preparación de datos. Las fuentes de datos certificadas indican a los miembros de la organización que los datos son fiables y están listos para el análisis. Además, estas fuentes de datos certificadas se mostrarán en los resultados de las búsquedas y los filtros con una prioridad más alta en Tableau Server o Tableau Online.



A través de la preparación de datos visual, las personas pueden ver todo el proceso y detectar de manera temprana los posibles errores. Esto incluye, por ejemplo, errores de ortografía, espacios de más o cláusulas de unión incorrectas. Asimismo, ayuda a aumentar la confianza en el análisis final.

JASON HARMER CONSULTOR, NATIONWIDE INSURANCE

Evaluar la gobernanza escalable

Todas las organizaciones tienen necesidades específicas, por lo que no hay una metodología universal para la preparación de datos. Sin embargo, al seleccionar la herramienta de preparación de datos de autoservicio, las empresas deben considerar cómo la herramienta permitirá optimizar los procesos hacia una metodología iterativa y ágil, en lugar de plantear nuevos desafíos. Las personas sentirán un mayor deseo de preparar y comprender los datos si pueden comprobar el impacto que tiene su preparación de datos.

Una colaboración adecuada y una gobernanza eficaz son la clave. Los usuarios pueden intentar resolver sus propios problemas con los datos, pero TI debe ser el responsable de solucionar los inconvenientes con los datos de la organización. Las prácticas de gobernanza permiten que los usuarios correctos accedan a los datos adecuados. Además, garantizan que los datos que consultan los usuarios para tomar decisiones sean precisos y cumplan con las políticas internas o las regulaciones externas.

Un cambio en la gobernanza no implica pedirle a TI que ceda el control, sino permitir que los miembros de la empresa sean más autosuficientes dentro de un entorno centralizado y fiable. Los analistas y los usuarios corporativos son los primeros que pueden identificar problemas con los datos o irregularidades en un modelo de gobernanza acordado por TI y la empresa.

Similar al cambio de paradigma que significó el análisis de autoservicio, expandir la participación de la empresa en la gobernanza a fin de democratizar la preparación de datos también plantea ciertos desafíos, por ejemplo, cambios en la gestión de los procesos y la tecnología, riesgos de seguridad que se deben mitigar y brechas en las habilidades de los usuarios. Sin embargo, es importante recordar que, con una metodología de implementación ágil e iterativa y un método colaborativo de gobernanza, los beneficios de permitir que más usuarios puedan llevar a cabo la preparación de datos superan las expectativas.



Sin Tableau Prep, nuestro equipo empleaba muchas horas en limpiar y organizar las fuentes de datos para garantizar que los análisis fueran precisos y eficaces. Gracias a Tableau Prep, nos hemos ahorrado muchas horas de trabajo y hemos cambiado por completo la forma de observar los datos. Ahora tardamos mucho menos tiempo en poder extraer información útil a partir de los datos recopilados.

GESSICA BRIGGS-SULLIVAN ADMINISTRADORA DE TABLEAU, CHARLES SCHWAB, INC.

Acerca de Tableau

Tableau es una plataforma de inteligencia de negocios visual completa, fácil de usar y preparada para la empresa, que ayuda a las personas a ver y comprender los datos con un análisis de autoservicio rápido y escalable. Ya sea en las instalaciones físicas o en la nube, en Windows o Linux, Tableau aprovecha las inversiones en tecnología existentes y se adapta a medida que el entorno de datos cambia y crece. Libera el poder de tus activos más valiosos: los datos y las personas.

Recursos adicionales

[Más información: Preparación de datos con Tableau](#)

[Más información: Gestión de datos con Tableau](#)

[Ayuda en línea: Tableau Prep Conductor](#)

[Informe: Los datos “sucios” tienen consecuencias. Cómo resolver los problemas más comunes de la preparación de datos.](#)

[Informe: Prácticas recomendadas para ordenar los datos](#)

[Estudio de investigación de BARC: Preparación de datos: convertir los datos sin procesar en datos útiles \(en inglés\)](#)

[Tableau para la empresa: Análisis impulsado por TI](#)

