

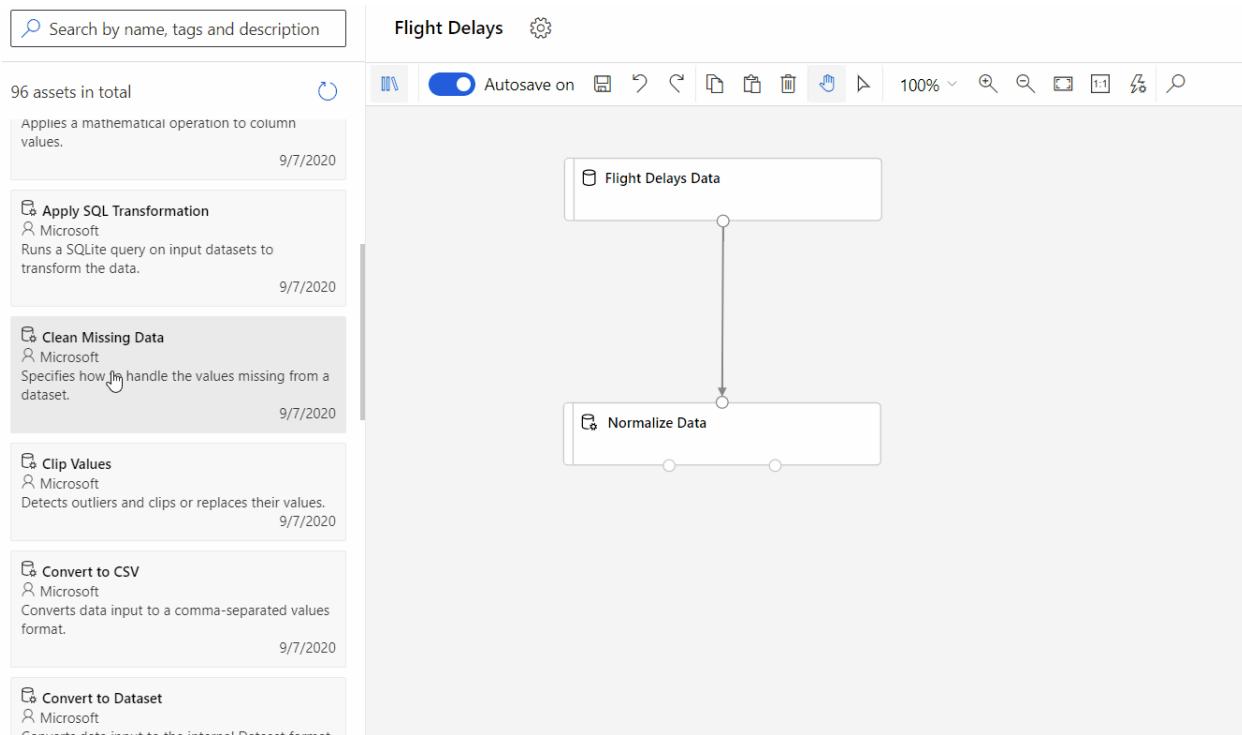
Introducción:

Azure Machine Learning es un servicio en la nube que permite acelerar y administrar el ciclo de vida de los proyectos de aprendizaje automático. Los profesionales de aprendizaje automático, científicos de datos e ingenieros pueden usarlo en sus flujos de trabajo diarios: entrenamiento e implementación de modelos y administración de MLOps.

Puede crear un modelo en Azure Machine Learning o usar un modelo creado a partir de una plataforma de código abierto, como Pytorch, TensorFlow o Scikit-learn. Las herramientas de MLOps le ayudan a supervisar, volver a entrenar y volver a implementar modelos.

¿Qué es el diseñador de Azure Machine Learning?

El diseñador de Azure Machine Learning es una interfaz de arrastrar y soltar que se usa para entrenar e implementar modelos en Azure Machine Learning



El diseñador usa su área de trabajo de Azure Machine Learning para organizar recursos compartidos como:

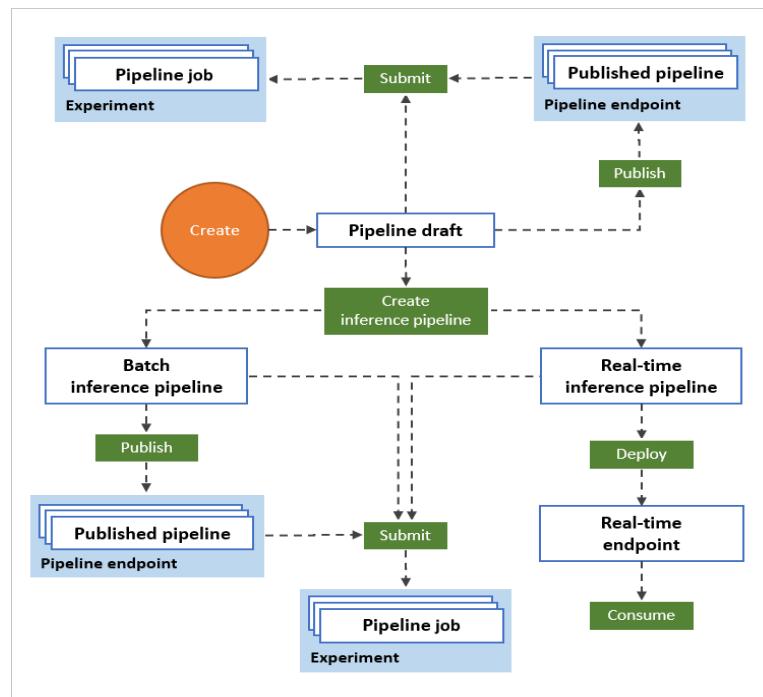
- Tuberías
- conjuntos de datos
- Recursos informáticos

- Modelos registrados
- Canalizaciones publicadas
- Puntos finales en tiempo real

Entrenamiento e implementación de modelos

Utilice un lienzo visual para crear un flujo de trabajo de aprendizaje automático de extremo a extremo. Entrene, pruebe e implemente modelos, todo en el diseñador:

- Arrastre y suelte conjuntos de datos y componentes en el lienzo.
- Conecte los componentes para crear un borrador de canalización .
- Envíe una ejecución de canalización con los recursos informáticos de su área de trabajo de Azure Machine Learning.
- Convierta sus canalizaciones de entrenamiento en canalizaciones de inferencia .
- Publique sus canalizaciones en un extremo de canalización REST para enviar una nueva canalización que se ejecute con diferentes parámetros y conjuntos de datos.
- Publique una canalización de entrenamiento para reutilizar una sola canalización para entrenar múltiples modelos mientras cambia parámetros y conjuntos de datos.
- Publique una canalización de inferencia por lotes para hacer predicciones sobre nuevos datos mediante el uso de un modelo previamente entrenado.
- Implemente una canalización de inferencia en tiempo real en un punto final en línea para hacer predicciones sobre nuevos datos en tiempo real.



Tubería

Una canalización consta de conjuntos de datos y componentes analíticos, que usted conecta. Las canalizaciones tienen muchos usos: puede crear una canalización que entrene un solo modelo o una que entrene varios modelos. Puede crear una canalización que haga predicciones en tiempo real o por lotes, o crear una canalización que solo limpie los datos. Las canalizaciones le permiten reutilizar su trabajo y organizar sus proyectos.

Borrador de tubería

A medida que edita una canalización en el diseñador, su progreso se guarda como un borrador de canalización . Puede editar un borrador de canalización en cualquier momento agregando o eliminando componentes, configurando objetivos de proceso, creando parámetros, etc.

Una canalización válida tiene estas características:

- Los conjuntos de datos solo pueden conectarse a componentes.
- los componentes solo pueden conectarse a conjuntos de datos u otros componentes.
- Todos los puertos de entrada de los componentes deben tener alguna conexión con el flujo de datos.
- Se deben establecer todos los parámetros requeridos para cada componente.
- Cuando esté listo para ejecutar su borrador de canalización, envíe una ejecución de canalización.

Ejecución de canalización

Cada vez que ejecuta una canalización, la configuración de la canalización y sus resultados se almacenan en su espacio de trabajo como una ejecución de canalización . Puede volver a cualquier ejecución de canalización para inspeccionarla en busca de solución de problemas o auditoría. Clone una ejecución de tubería para crear un nuevo borrador de tubería para que lo edite.

Las ejecuciones de canalización se agrupan en experimentos para organizar el historial de ejecución. Puede configurar el experimento para cada ejecución de canalización.

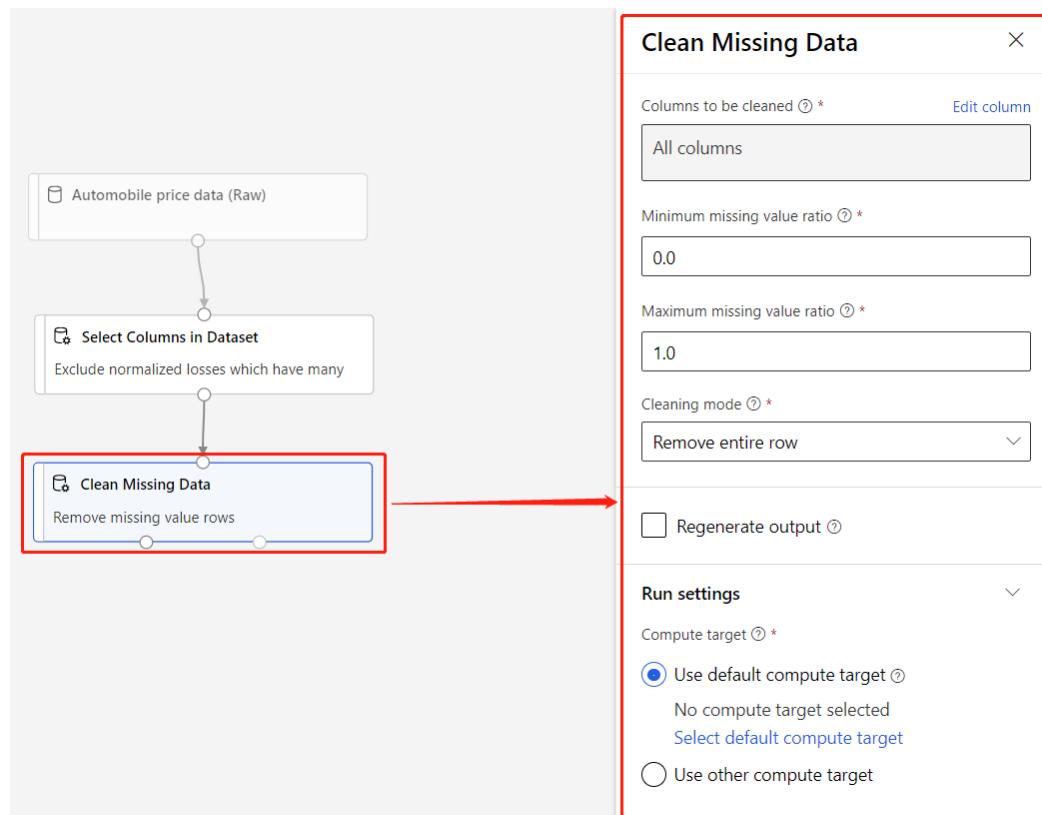
Conjuntos de datos

Un conjunto de datos de aprendizaje automático facilita el acceso y el trabajo con sus datos. Se incluyen varios conjuntos de datos de muestra en el diseñador para que experimente. Puede registrar más conjuntos de datos a medida que los necesite.

Componente

Un componente es un algoritmo que puede realizar en sus datos. El diseñador tiene varios componentes que van desde funciones de ingreso de datos hasta procesos de capacitación, calificación y validación.

Un componente puede tener un conjunto de parámetros que puede usar para configurar los algoritmos internos del componente. Cuando selecciona un componente en el lienzo, los parámetros del componente se muestran en el panel Propiedades a la derecha del lienzo. Puede modificar los parámetros en ese panel para ajustar su modelo. Puede configurar los recursos informáticos para componentes individuales en el diseñador.



Para obtener ayuda para navegar a través de la biblioteca de algoritmos de aprendizaje automático disponibles, consulte Introducción a la referencia de algoritmos y componentes . Para obtener ayuda con la elección de un algoritmo, consulte la hoja de referencia del algoritmo de Azure Machine Learning .

Recursos informáticos

Utilice los recursos informáticos de su espacio de trabajo para ejecutar su canalización y alojar sus modelos implementados como puntos finales en línea o puntos finales de canalización (para inferencia por lotes). Los destinos de proceso admitidos son:

Calcular objetivo	Capacitación	Despliegue
Cómputo de Azure Machine Learning	✓	
Servicio Azure Kubernetes		✓

Los destinos de proceso se adjuntan a su área de trabajo de Azure Machine Learning . Administre sus objetivos de cómputo en su área de trabajo en el estudio Azure Machine Learning .

Desplegar

Para realizar inferencias en tiempo real, debe implementar una canalización como punto final en línea . El punto final en línea crea una interfaz entre una aplicación externa y su modelo de puntuación. Una llamada a un punto final en línea devuelve resultados de predicción a la aplicación en tiempo real. Para realizar una llamada a un punto final en línea, pasa la clave de API que se creó cuando implementó el punto final. El punto final se basa en REST, una opción de arquitectura popular para proyectos de programación web.

Los puntos de conexión en línea deben implementarse en un clúster de Azure Kubernetes Service.

Publicar

También puede publicar una canalización en un extremo de canalización . Al igual que un punto final en línea, un punto final de canalización le permite enviar nuevas ejecuciones de canalización desde aplicaciones externas mediante llamadas REST. Sin embargo, no puede enviar ni recibir datos en tiempo real mediante un extremo de canalización.

Las canalizaciones publicadas son flexibles, se pueden usar para entrenar o volver a entrenar modelos, realizar inferencias por lotes , procesar nuevos datos y mucho más. Puede publicar varias canalizaciones en un solo punto de enlace de canalización y especificar qué versión de canalización ejecutar.

Una canalización publicada se ejecuta en los recursos informáticos que define en el borrador de la canalización para cada componente.

El diseñador crea el mismo objeto PublishedPipeline que el SDK.

Tutorial: Diseñador: entrenar un modelo de regresión sin código

Entrene un modelo de regresión lineal que prediga los precios de los automóviles mediante el diseñador de Azure Machine Learning. Este tutorial es la primera parte de una serie de dos partes.

En la primera parte del ejercicio, aprenderá a:

- Crear una nueva canalización.
- Importar Datos
- Preparar datos.
- Entrenar un modelo de aprendizaje automático.
- Evaluar un modelo de aprendizaje automático.

En la segunda parte del se implementará su modelo como punto final de inferencia en tiempo real para predecir el precio de cualquier automóvil en función de las especificaciones técnicas que le envíe.

Crear una nueva canalización

Las canalizaciones de Azure Machine Learning organizan varios pasos de aprendizaje automático y procesamiento de datos en un único recurso. Las canalizaciones le permiten organizar, administrar y reutilizar flujos de trabajo de aprendizaje automático complejos entre proyectos y usuarios.

Para crear una canalización de Azure Machine Learning, necesita un área de trabajo de Azure Machine Learning. En esta sección, aprenderá a crear estos dos recursos.

The screenshot shows the Microsoft Azure portal homepage. At the top, there's a navigation bar with links like 'Acceso tutor', 'Banca web - Transa...', 'Android code gene...', 'Por el alimento imp...', 'Pet food machinery...', 'Cómo alcanzar t...', 'Review Berkshire H...', 'Como configurar c...', 'Aula virtual de IMF...', and 'Otros favoritos'. The main header says 'Microsoft Azure' and has a search bar. Below the header is a row of service icons: 'Create a resource', 'Machine Learning Stud...', 'Language understanding', 'Subscriptions', 'Key vaults', 'Power Platform', 'Azure Cosmos DB', 'SQL databases', 'Applied AI services', and 'More services'. The main content area is titled 'Resources' and shows a table of recent resources. The table has columns for 'Name', 'Type', and 'Last Viewed'. The resources listed are: 'Ejercicio1-ml' (Azure Machine Learning, 11 minutes ago), 'Labz' (Resource group, 2 weeks ago), 'APLICACIONAZUREMAPS' (Azure Maps Account, 2 months ago), 'Suscripción de Azure 1' (Subscription, 2 months ago), 'Azure_Databricks1' (Azure Databricks Service, 5 months ago), 'apps1' (App Configuration, 6 months ago), 'unoprueba2' (Azure Cosmos DB account, 10 months ago), 'iot-hub-eg001' (IoT Hub, 11 months ago), 'NetworkWatcherRG' (Resource group, a year ago), 'nosql' (Resource group, a year ago), and 'LanLabz' (Virtual network, a year ago). There are also 'Recent' and 'Favorite' tabs at the top of the table, and a 'See all' link at the bottom.

Crear un nuevo espacio de trabajo

Materia: Minería de Datos

Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning

Necesita un área de trabajo de Azure Machine Learning para usar el diseñador. El área de trabajo es el recurso de nivel superior para Azure Machine Learning, proporciona un lugar centralizado para trabajar con todos los artefactos que crea en Azure Machine Learning.

The screenshot shows the Microsoft Azure portal at <https://portal.azure.com/#create/hub>. The search bar contains "Search resources, services, and docs (G+)" and the user email "estevan.gomez@iot-bd...". The main content area is titled "Create a resource" and lists various Azure services. Under the "Analytics" category, "Azure Machine Learning" is highlighted with a blue border. Other listed services include Computer Vision, Face, Language Understanding, QnA Maker, Speech, Language service, Translator, and Azure Cognitive Search. To the right, there are sections for "1) Starter Plan" (NVIDIA Image for AI using GPUs - v21.06.0), "NVIDIA GPU-Optimized TensorFlow Image - v21.06.0", "NVIDIA GPU-Optimized PyTorch Image - v21.06.0", "NVIDIA GPU-Optimized Image for AI & HPC - v21.06.0", "NVIDIA PyTorch Image for AI on GPUs - v21.06.0", "Azure Data Science Hub - DSVM", "Free Plan", and "Dataiku DSS trial".

The screenshot shows the Microsoft Azure portal at <https://portal.azure.com/#create/Microsoft.MachineLearningServices>. The search bar contains "Search resources, services, and docs (G+)" and the user email "estevan.gomez@iot-bd...". The main content area is titled "Azure Machine Learning" and shows the "Create a machine learning workspace" step. It includes fields for "Subscription" (selected "Suscripción de Azure 1") and "Resource group" (selected "(New) lab2022"). Below this, "Workspace details" are configured with "Workspace name" set to "lab2022", "Region" set to "South Central US", "Storage account" set to "(new) lab20225717818378", "Key vault" set to "(new) lab20221830028567", "Application insights" set to "(new) lab202221611078778", and "Container registry" set to "None". At the bottom, there are "Review + create" and "Next : Networking" buttons.

Materia: Minería de Datos

Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning

The screenshot shows the 'Create a machine learning workspace' wizard in the Azure portal. The 'Review + create' step is active. A progress bar indicates 'Submitting deployment...' with the message 'Submitting the deployment template for resource group 'lab2022''. The configuration details include:

Category	Setting
Subscription	Suscripción de Azure 1
Resource group	(New) lab2022
Region	South Central US
Workspace name	lab2022
Storage account	(new) lab20225717818378
Key vault	(new) lab20221830028567
Application insights	(new) lab20221611078778
Container registry	None
Networking	Enable public access from all networks
Advanced	Identity type: System assigned Encryption type: Microsoft-managed keys Enable HBI Flag: Disabled

At the bottom, there are 'Create', '< Previous', 'Next >', and 'Download a template for automation' buttons.

The screenshot shows the 'Microsoft.MachineLearningServices | Overview' page. It displays a message 'Deployment is in progress'. Deployment details are listed:

Resource	Type	Status	Operation details
lab20225717818378	Microsoft.Storage/storageAccounts	Accepted	Operation details
lab20221830028567	Microsoft.KeyVault/vaults	OK	Operation details
lab20221611078778	Microsoft.Insights/components	OK	Operation details

Para información adicional revisar: Administración de grupos de recursos de Azure con Azure Portal:

[**Administración de grupos de recursos: Azure Portal - Azure Resource Manager | Microsoft Docs**](#)

Materia: Minería de Datos

Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning

Your deployment is complete

Deployment name: Microsoft.MachineLearningServices
Subscription: Suscripción de Azure 1
Resource group: lab2022

Start time: 7/7/2022, 6:11:17 PM
Correlation ID: c56273be-4788-496d-b1de-126f84aa58cf

Deployment details (Download)

Next steps

Go to resource

Cost Management
Get notified to stay within your budget and prevent unexpected charges on your bill.
Set up cost alerts >

Microsoft Defender for Cloud
Secure your apps and infrastructure
Go to Microsoft Defender for Cloud >

Free Microsoft tutorials
Start learning today >

Work with an expert
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.
Find an Azure expert >

Escogemos ir al recurso y nos presenta la siguiente pantalla:

lab2022 ★ ...

Overview

Download config.json Delete

Essentials

Resource group	: lab2022
Location	: South Central US
Subscription	: Suscripción de Azure 1
Subscription ID	: dc647937-5ec1-4e9e-9172-ce7f4e28aabe
Storage	: lab20225717818378
Studio web URL	: https://ml.azure.com/?tid=89601dd9-bee3-4422-a04-228b1d60596d&wsid=...
Registry	: ...
Key Vault	: lab20221830028567
Application Insights	: lab202211078778
MLflow tracking URI	: azurerm://southcentralus.api.azurerm.ms/mlflow/v1.0/subscriptions/dc647937-5ec1-4e9e-9172-ce7f4e28aabe

Work with your models in Azure Machine Learning Studio

The Azure Machine Learning Studio is a web app where you can build, train, test, and deploy ML models. Launch it now to start exploring, or learn more about the Azure Machine Learning Studio.

Launch studio

Escogemos ir a Studio, entonces estamos en Azure Machine Learning Studio

Al seleccionar + aparecen todas las opciones que podemos crear

Materia: Minería de Datos

Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning

The screenshot shows the Microsoft Azure Machine Learning Studio interface. On the left, there's a sidebar with navigation links like Home, Notebooks, Automated ML, Designer, Data, Jobs, Components, Pipelines, Environments, Models, Endpoints, Compute, Datasources, Linked Services, and Data Labeling. The main area has four large cards: 'Create new' (Notebook, Automated ML job, Pipeline, Component), 'Notebooks' (with 'Start now' button), 'Automated ML' (with 'Start now' button), and 'Designer' (with 'Start now' button). Below these is a 'Recent' section showing items like Data asset, Compute instance, Training cluster, Datastore, Data labeling project, Environment, Endpoint, and Job (preview). A 'Data' section follows, with a table header for Experiment, Status, Submitted time, Submitted..., and Job type, and a message 'No jobs to display'. At the top, the URL is https://ml.azure.com/?tid=89601dd9-bec3-4422-af04-228b1d60596d&wsid=/subscriptions/dc647937-5ec1-4e9e-9172-ce7f4e28aabe/resourcegr... and the title bar says 'Microsoft Azure Machine Learning Studio'.

Crear la canalización

1. Inicie sesión en ml.azure.com y seleccione el espacio de trabajo con el que desea trabajar.
2. Selecciona **Diseñador**.

The screenshot shows the Microsoft Azure Machine Learning Studio Designer page. The sidebar on the left is identical to the Home page. The main area has a 'New pipeline' section with a large plus sign icon and the text 'Easy-to-use prebuilt components'. To the right is a diagram titled 'Image Classification using DenseNet' showing a flow from input images through several processing steps to output categories. The 'Designer' link in the sidebar is highlighted with a red box.

3. Seleccione **Componentes preconstruidos fáciles de usar**.
4. En la parte superior del lienzo, seleccione el nombre de canalización predeterminado **Pipeline-Created-on**. Cámbiele el nombre a *Predicción de precios de automóviles*. El nombre no necesita ser único.

5.

Establecer el destino de proceso predeterminado

Una canalización se ejecuta en un destino informático, que es un recurso informático adjunto a su área de trabajo. Después de crear un destino de proceso, puede reutilizarlo para ejecuciones futuras.

Importante

El cómputo adjunto no es compatible, use [instancias de cómputo o clústeres](#) en su lugar.

Puede establecer un **destino de proceso predeterminado** para toda la canalización, lo que indicará a cada componente que utilice el mismo destino de proceso de forma predeterminada. Sin embargo, puede especificar objetivos de proceso por módulo.

1. Junto al nombre de la canalización, seleccione el **ícono de engranaje** en la parte superior del lienzo para abrir el panel **Configuración**.
2. En el panel **Configuración** a la derecha del lienzo, seleccione **Seleccionar destino de proceso**.

Si ya tiene un destino de proceso disponible, puede seleccionarlo para ejecutar esta canalización.

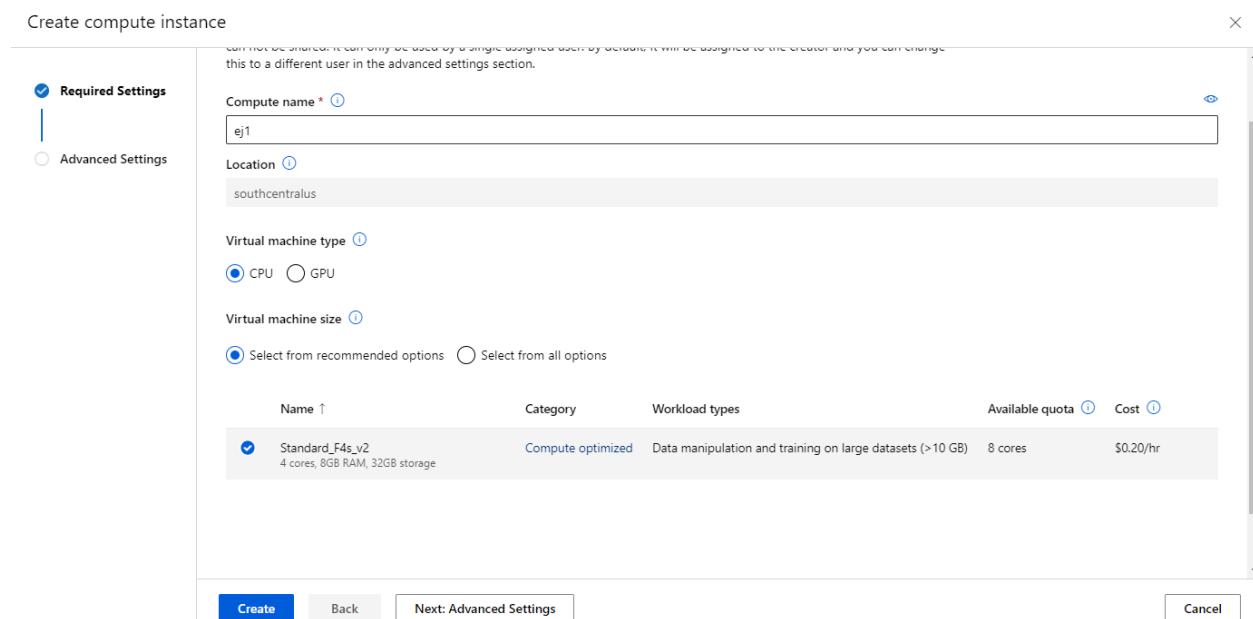
3. Introduzca un nombre para el recurso informático.
4. Seleccione **Guardar**.

Nota

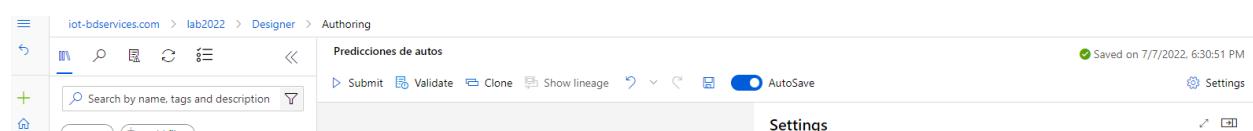
Se tarda aproximadamente cinco minutos en crear un recurso informático. Una vez creado el recurso, puede reutilizarlo y omitir este tiempo de espera para ejecuciones futuras.

El recurso informático escala automáticamente a cero nodos cuando está inactivo para ahorrar costos. Cuando lo vuelva a usar después de un retraso, es posible que experimente aproximadamente cinco minutos de tiempo de espera mientras se vuelve a escalar.

Configuramos la instancia de computo (tomar en cuenta que luego vamos a apagarla)



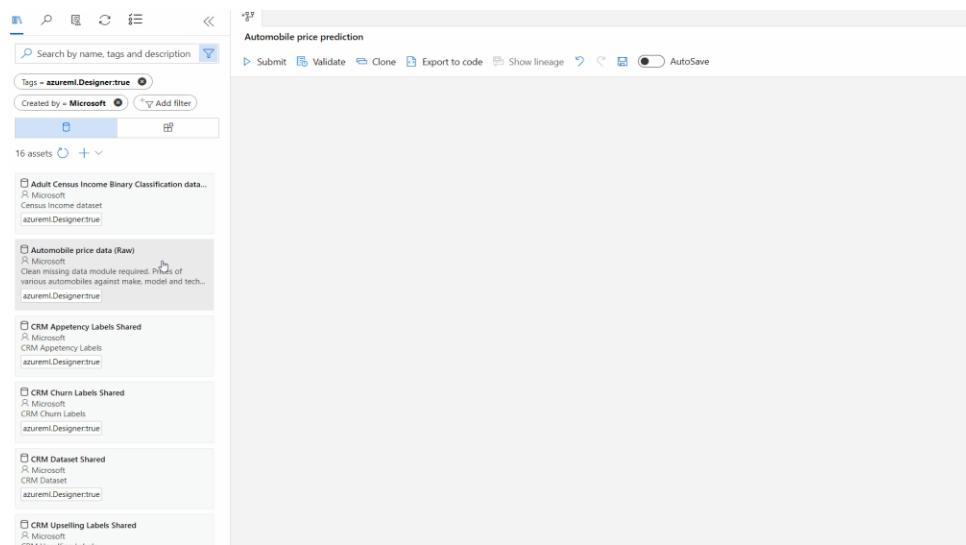
Activar el autosave



Datos de importación

Hay varios conjuntos de datos de muestra incluidos en el diseñador para que experimente. Para este tutorial, use **Datos de precios de automóviles (sin procesar)**.

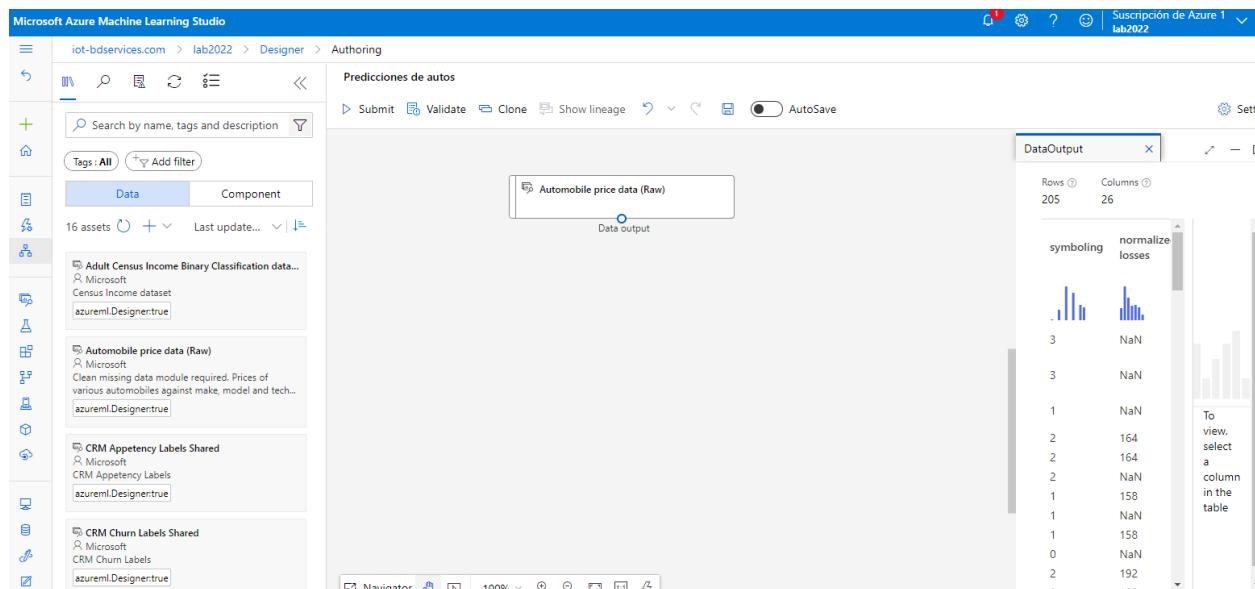
1. A la izquierda del lienzo de canalización hay una paleta de conjuntos de datos y componentes. Seleccione Conjuntos de **datos de muestra** para ver los conjuntos de datos de muestra disponibles.
2. Seleccione el conjunto de **datos Datos de precios de automóviles (sin procesar)** y arrástrelo al lienzo.



Visualiza los datos

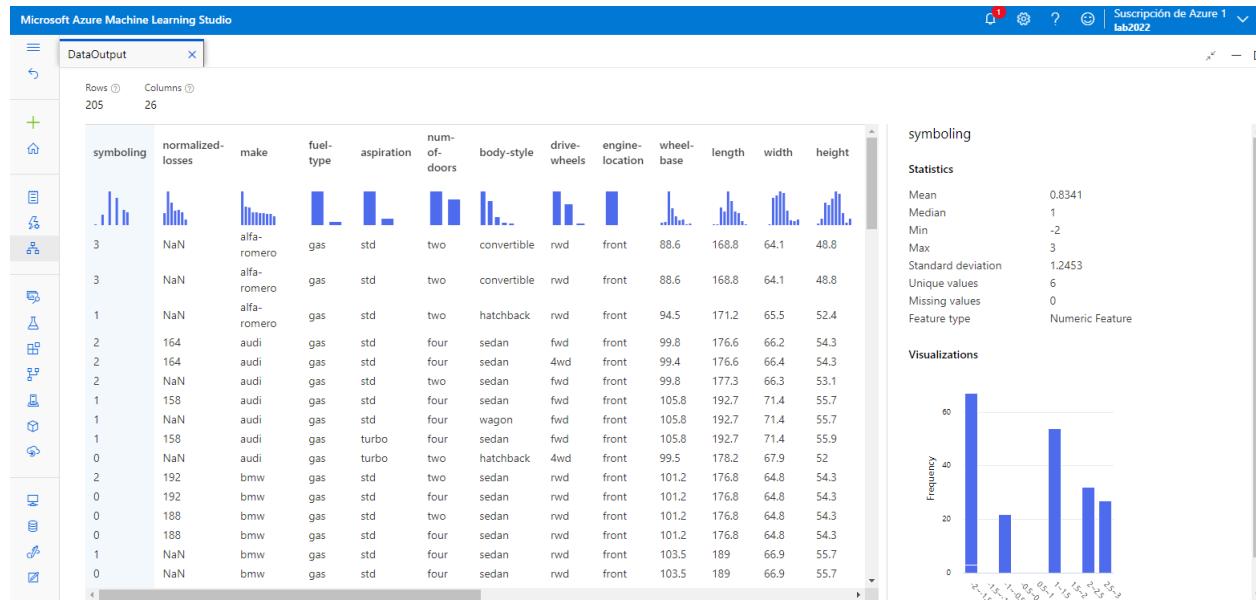
Puede visualizar los datos para comprender el conjunto de datos que utilizará.

1. Haga clic con el botón derecho en **Datos de precios de automóviles (sin procesar)** y seleccione **Vista previa de datos**.



2. Seleccione las diferentes columnas en la ventana de datos para ver información sobre cada una.

Cada fila representa un automóvil y las variables asociadas a cada automóvil aparecen como columnas. Hay 205 filas y 26 columnas en este conjunto de datos.



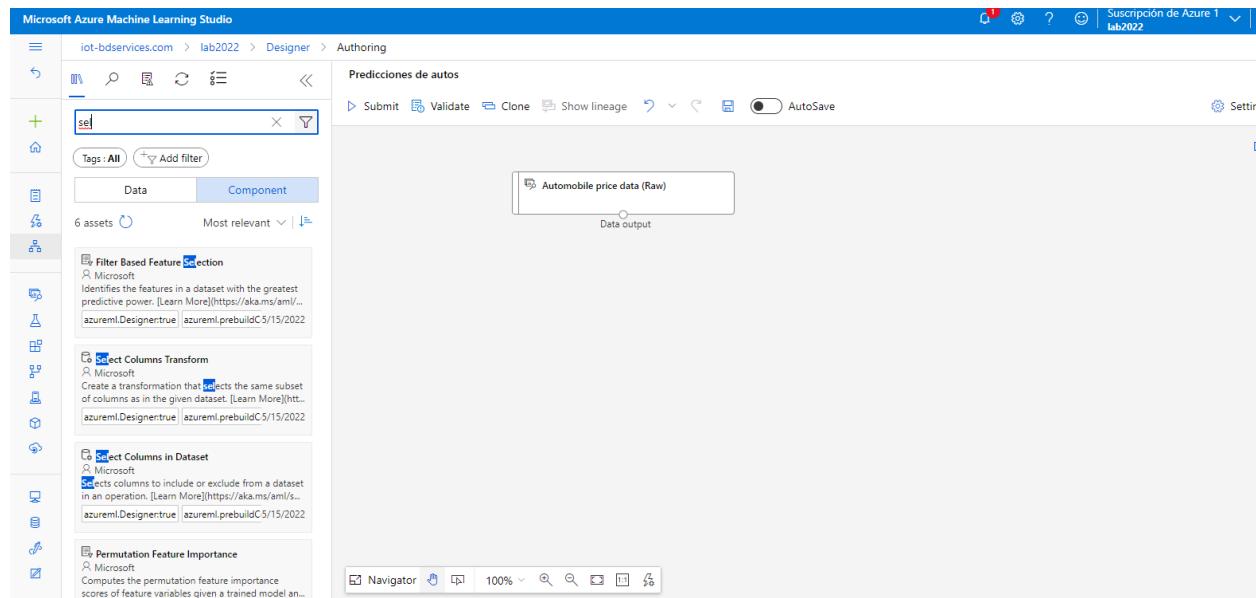
Preparar datos

Los conjuntos de datos generalmente requieren algún procesamiento previo antes del análisis. Es posible que haya notado que faltan algunos valores cuando inspeccionó el conjunto de datos. Estos valores faltantes deben limpiarse para que el modelo pueda analizar los datos correctamente.

Quitar una columna

Cuando entrena un modelo, tiene que hacer algo con los datos que faltan. En este conjunto de datos, a la columna **de pérdidas normalizadas** le faltan muchos valores, por lo que excluirá esa columna del modelo por completo.

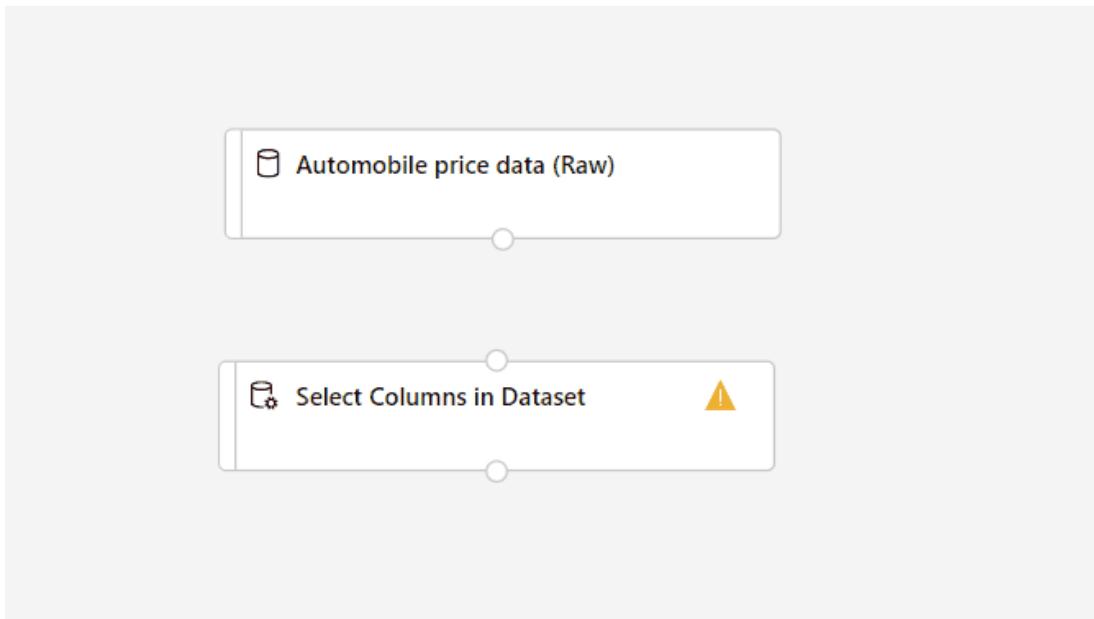
1. En la paleta de componentes a la izquierda del lienzo, expanda la sección **Transformación de datos** y busque el componente **Seleccionar columnas en conjunto de datos**.



2. Arrastre el componente **Seleccionar columnas en conjunto de datos** al lienzo. Suelte el componente debajo del componente del conjunto de datos.
3. Conecte el conjunto de datos de **datos de precios de automóviles (sin procesar)** al componente **Seleccionar columnas en conjunto de datos**. Arrastre desde el puerto de salida del conjunto de datos, que es el círculo pequeño en la parte inferior del conjunto de datos en el lienzo, hasta el puerto de entrada de **Seleccionar columnas en el conjunto de datos**, que es el círculo pequeño en la parte superior del componente.

Consejo

Usted crea un flujo de datos a través de su tubería cuando conecta el puerto de salida de un componente a un puerto de entrada de otro.

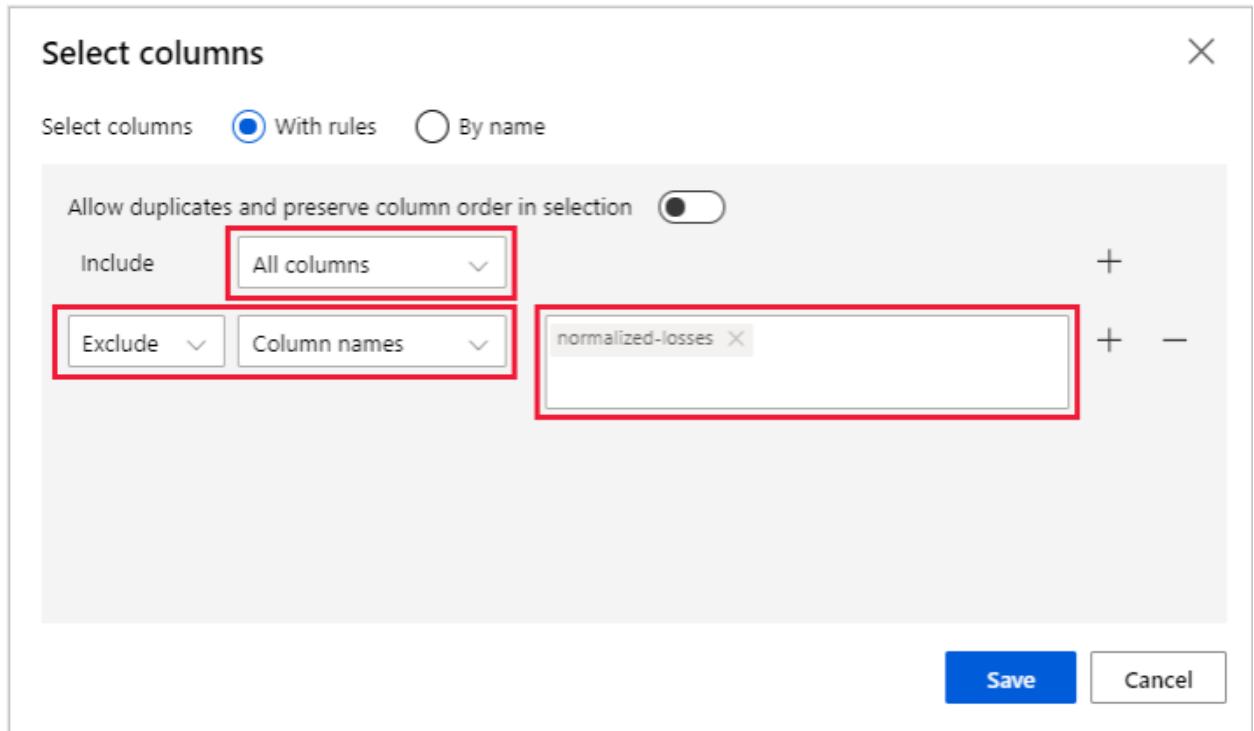


4. Seleccione el componente **Seleccionar columnas en conjunto de datos**.
5. En el panel de detalles del componente a la derecha del lienzo, seleccione Editar columna .

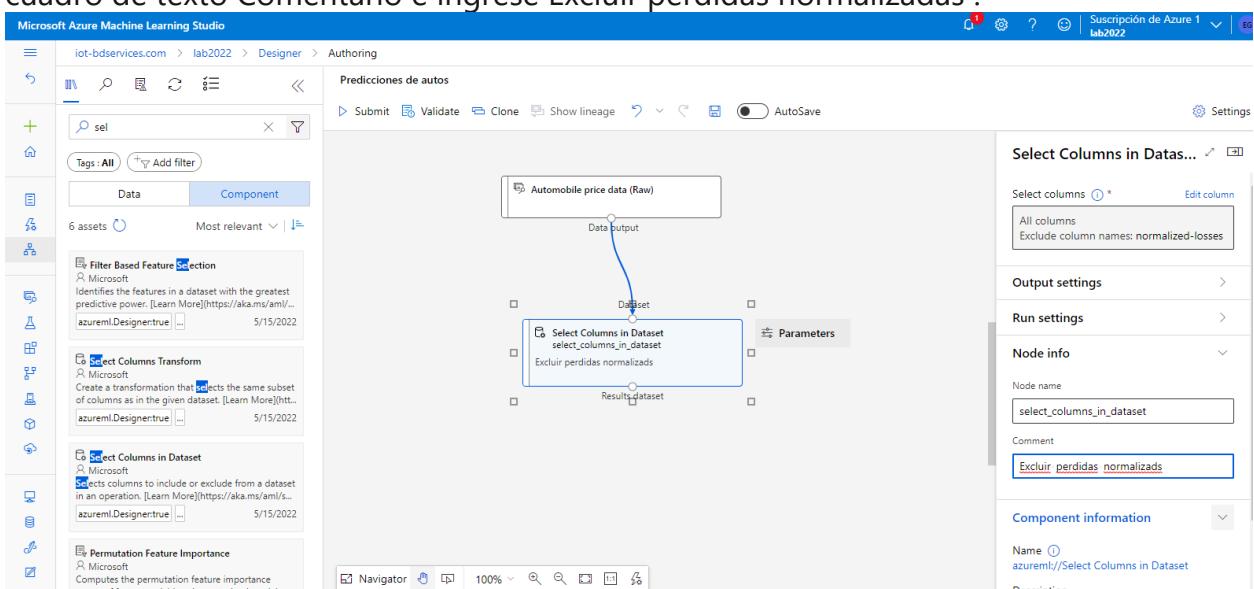
The screenshot shows the Microsoft Azure Machine Learning Studio interface. On the left, there is a sidebar with various components listed under "Data" and "Component". In the center, there is a canvas with a "Automobile price data (Raw)" dataset and a "Select Columns in Dataset" component. A connection line from the dataset leads to the transformation component. On the right, there is a detailed view of the "Select Columns in Dataset" component. The "Select columns" section is expanded, showing a dropdown menu with the message "A value is required." Below it, there are sections for "Output settings", "Run settings", "Node info", and "Component information". The overall interface is professional and user-friendly.

6. Expanda el menú desplegable Nombres de columnas junto a Incluir y seleccione Todas las columnas .
7. Seleccione el + para agregar una nueva regla.
8. En los menús desplegables, seleccione Excluir y Nombres de columna .

9. Ingrese las pérdidas normalizadas en el cuadro de texto.
10. En la parte inferior derecha, seleccione Guardar para cerrar el selector de columnas.



11. Seleccione el componente **Seleccionar columnas en conjunto de datos**.
12. En el panel de detalles del componente a la derecha del lienzo, seleccione el cuadro de texto Comentario e ingrese Excluir pérdidas normalizadas .



- 13.

Los comentarios aparecerán en el gráfico para ayudarlo a organizar su embudo.

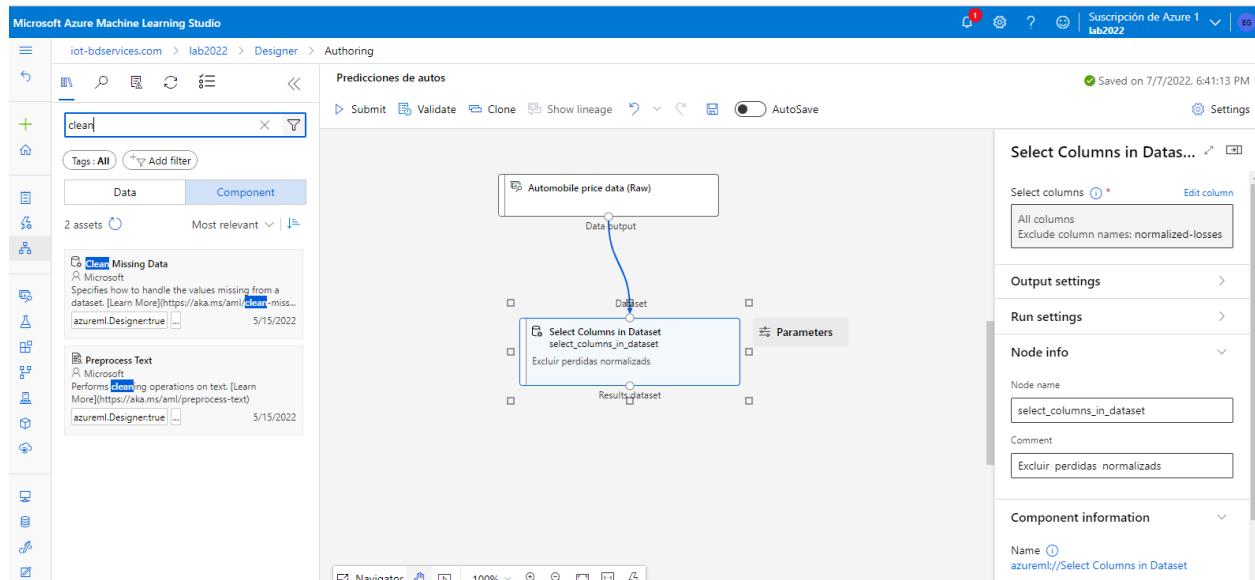
Limpiar datos faltantes

Su conjunto de datos aún tiene valores faltantes después de eliminar la columna **de pérdidas normalizadas**. Puede eliminar los datos faltantes restantes mediante el componente **Limpiar datos faltantes**.

Consejo

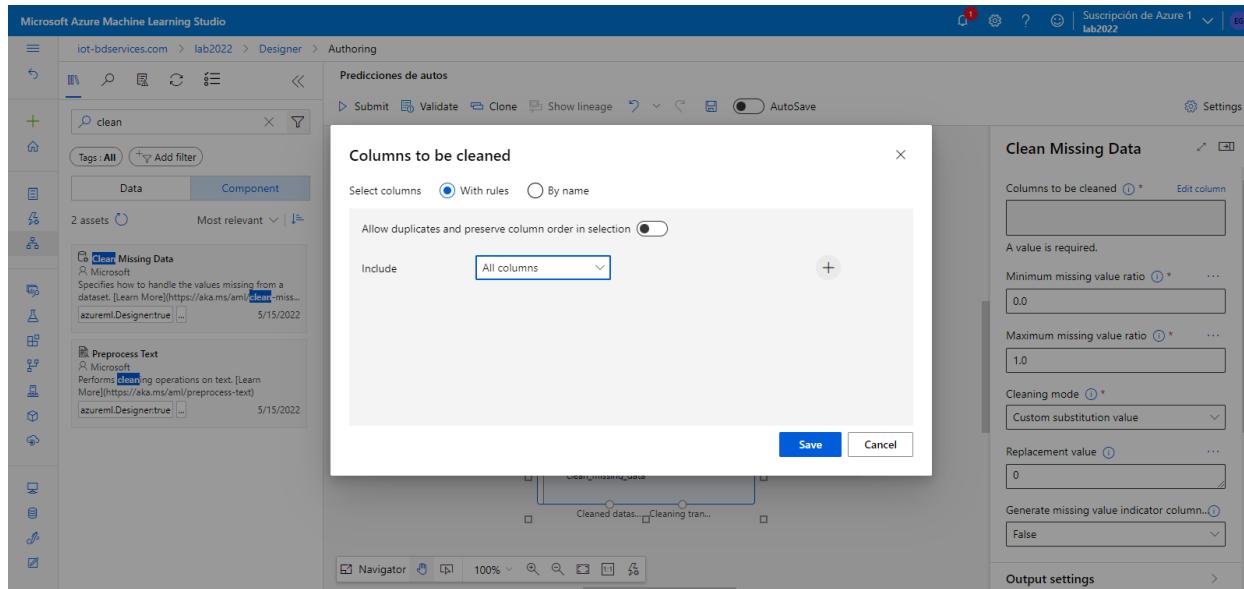
Limpiar los valores faltantes de los datos de entrada es un requisito previo para usar la mayoría de los componentes en el diseñador.

1. En la paleta de componentes a la izquierda del lienzo, expanda la sección **Transformación** de datos y busque el componente **Limpiar datos que faltan**.



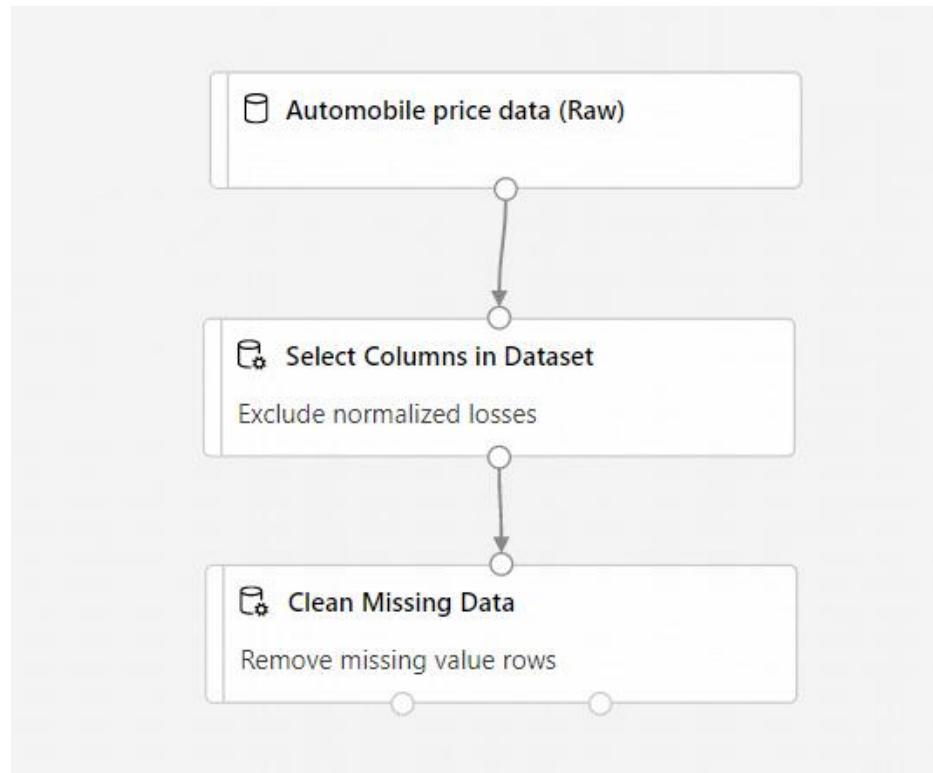
2. Arrastre el componente **Limpiar datos que faltan** al lienzo de canalización. Conéctelo al componente **Seleccionar columnas en conjunto de datos**.
3. Seleccione el componente **Limpiar datos que faltan**.
4. En el panel de detalles del componente a la derecha del lienzo, seleccione **Editar columna**.

5. En la ventana **Columnas para limpiar** que aparece, expanda el menú desplegable junto a **Incluir**. Seleccionar, **Todas las columnas**
6. Seleccione **Guardar**



7. En el panel de detalles del componente a la derecha del lienzo, seleccione **Eliminar fila completa** en **Modo de limpieza**.
8. En el panel de detalles del componente a la derecha del lienzo, seleccione el cuadro **Comentario** e ingrese *Eliminar filas de valores faltantes*.

Su tubería ahora debería verse así:



Entrenar un modelo de aprendizaje automático

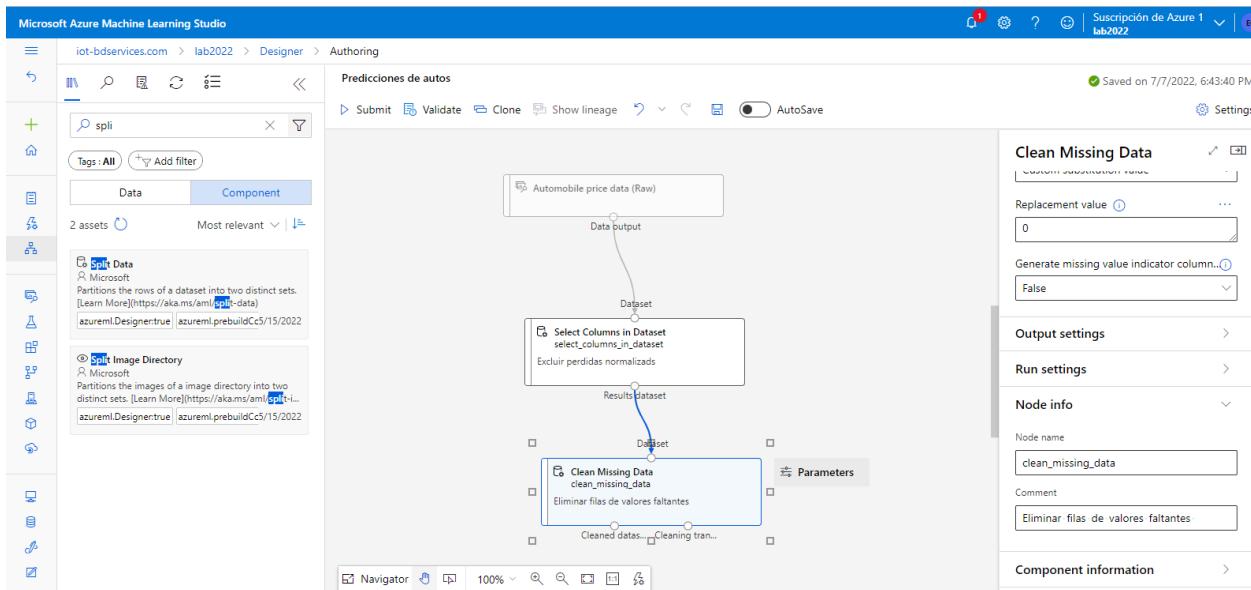
Ahora que tiene los componentes para procesar los datos, puede configurar los componentes de entrenamiento.

Como desea predecir el precio, que es un número, puede usar un algoritmo de regresión. Para este ejemplo, utiliza un modelo de regresión lineal.

dividir los datos

La división de datos es una tarea común en el aprendizaje automático. Dividirá sus datos en dos conjuntos de datos separados. Un conjunto de datos entrenará el modelo y el otro probará qué tan bien se desempeñó el modelo.

1. En la paleta de componentes, expanda la sección **Transformación de datos** y busque el componente **Dividir datos**.



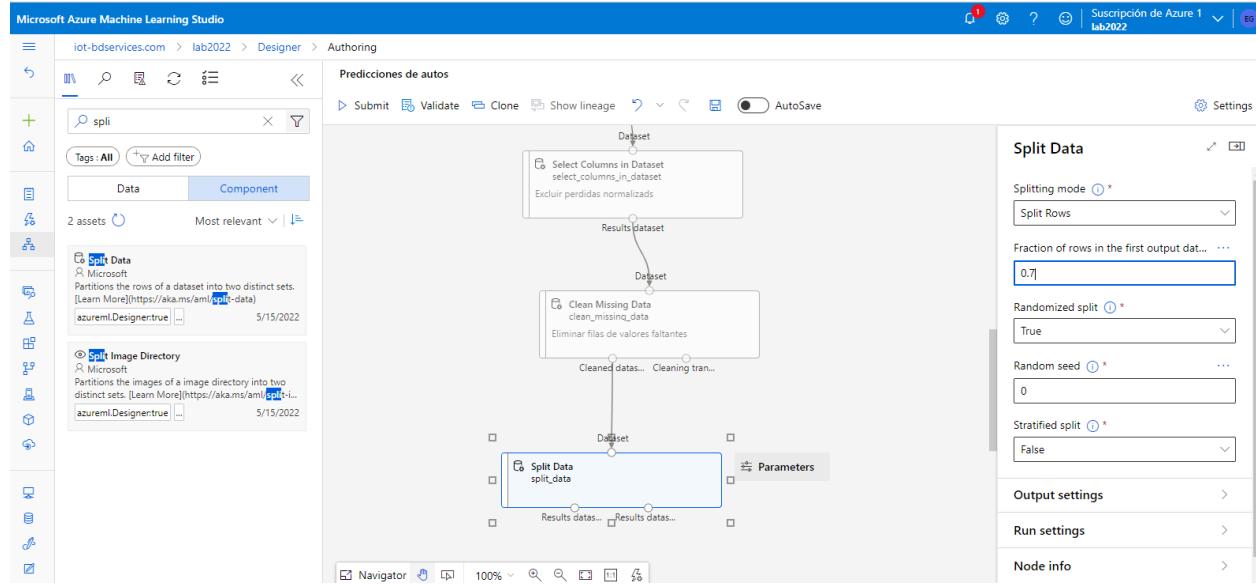
2. Arrastre el componente **Dividir datos** al lienzo de canalización.
3. Conecte el puerto izquierdo del componente **Limpiar datos faltantes** al componente **Dividir datos**.

Importante

Asegúrese de que los puertos de salida izquierdos de **Limpiar datos faltantes** se conecten a **Dividir datos**. El puerto izquierdo contiene los datos limpios. El puerto derecho contiene los datos descartados.

4. Seleccione el componente **Dividir datos**.
5. En el panel de detalles del componente a la derecha del lienzo, establezca la **Fracción de filas en el primer conjunto de datos de salida** en 0,7.

Esta opción divide el 70 por ciento de los datos para entrenar el modelo y el 30 por ciento para probarlo. Se podrá acceder al conjunto de datos del 70 por ciento a través del puerto de salida izquierdo. Los datos restantes estarán disponibles a través del puerto de salida derecho.



- En el panel de detalles del componente a la derecha del lienzo, seleccione el cuadro **Comentario** e ingrese *Dividir el conjunto de datos en conjunto de entrenamiento (0.7) y conjunto de prueba (0.3)*.

entrenar al modelo

Entrene el modelo dándole un conjunto de datos que incluya el precio. El algoritmo construye un modelo que explica la relación entre las características y el precio tal como lo presentan los datos de entrenamiento.

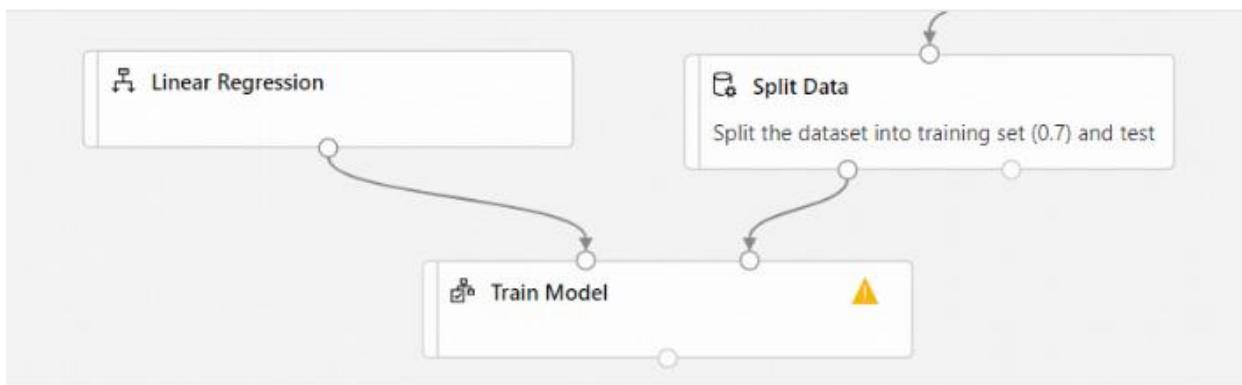
- En la paleta de componentes, expanda **Algoritmos de aprendizaje automático**.

Esta opción muestra varias categorías de componentes que puede utilizar para inicializar algoritmos de aprendizaje.

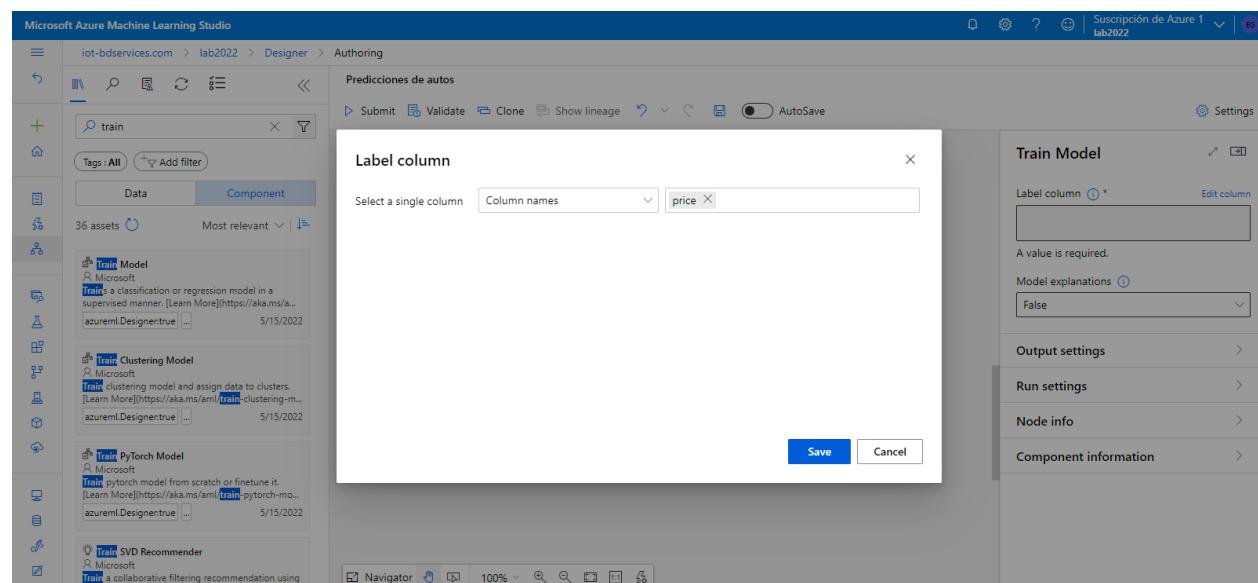
- Seleccione **Regresión > Regresión lineal** y arrástrelo al lienzo de canalización.
- En la paleta de componentes, expanda la sección **Entrenamiento de módulos** y arrastre el componente **Entrenar modelo** al lienzo.
- Conecte la salida del componente **Regresión lineal** a la entrada izquierda del componente **TarinModel**
- Conecte la salida de datos de entrenamiento (puerto izquierdo) del componente **Split Data** a la entrada derecha del componente **Train Model**.

Importante

Asegúrese de que los puertos de salida izquierdos de **Split Data** se conecten a **Train Model**. El puerto izquierdo contiene el conjunto de entrenamiento. El puerto derecho contiene el conjunto de prueba.



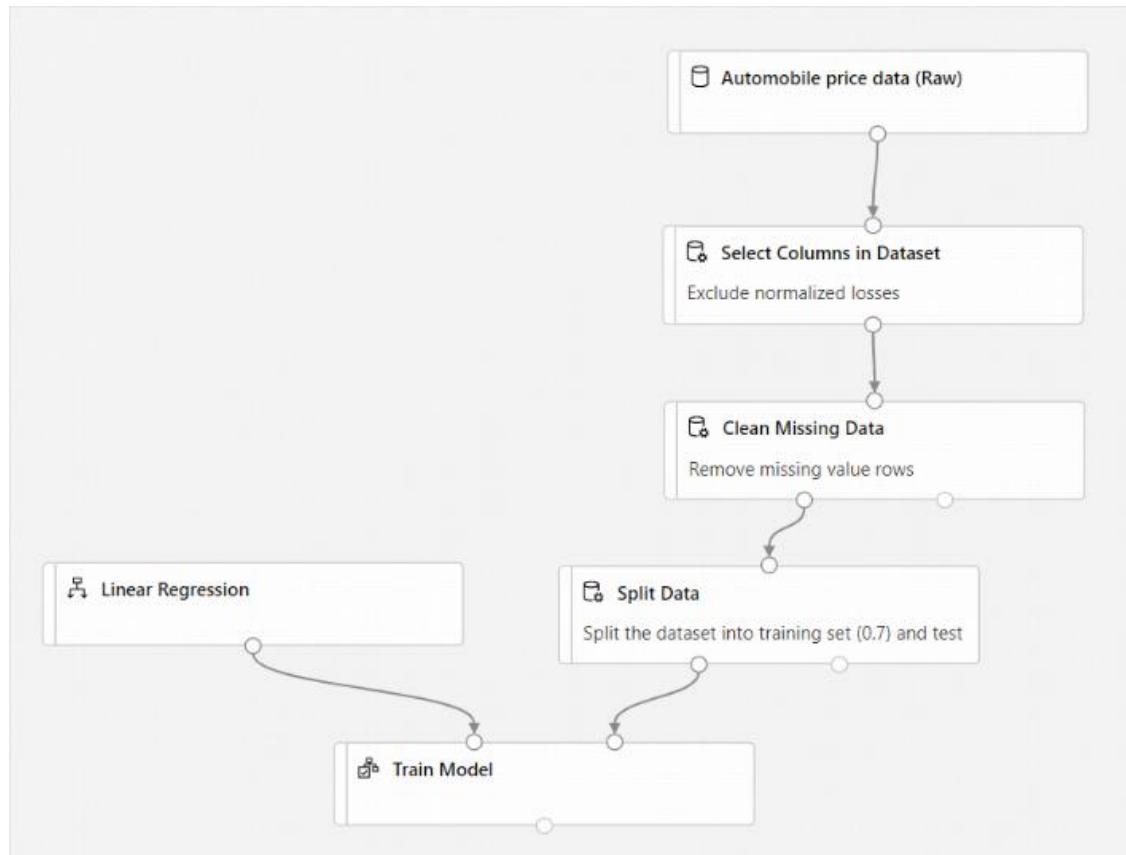
6. Seleccione el componente **Modelo de Entrenamiento**.
7. En el panel de detalles del componente a la derecha del lienzo, seleccione **Editar** selector de columna.
8. En el cuadro de diálogo **Columna de etiqueta**, expanda el menú desplegable y seleccione **Nombres de columna**.
9. En el cuadro de texto, ingrese el *precio* para especificar el valor que su modelo va a predecir.



Importante

Asegúrese de ingresar el nombre de la columna exactamente. No capitalice **el precio**.

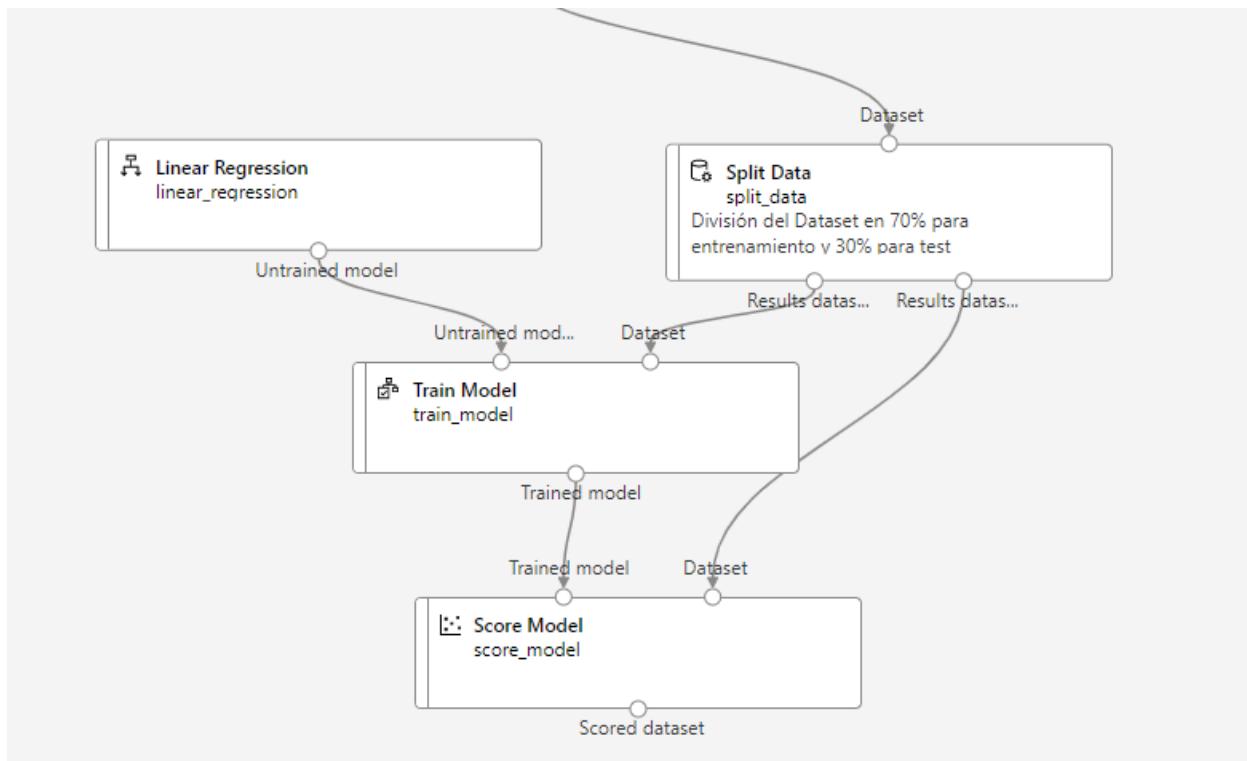
Su tubería debería verse así:



Agregar el componente Modelo de puntuación

Después de entrenar su modelo usando el 70 por ciento de los datos, puede usarlo para calificar el otro 30 por ciento para ver qué tan bien funciona su modelo.

1. Introduzca el *modelo de puntuación* en el cuadro de búsqueda para encontrar el componente **Modelo de puntuación**. Arrastre el componente al lienzo de canalización.
2. Conecte la salida del componente **Train Model** al puerto de entrada izquierdo de **Score Model**. Conecte la salida de datos de prueba (puerto derecho) del componente **Split Data** al puerto de entrada derecho de **Score Model**.

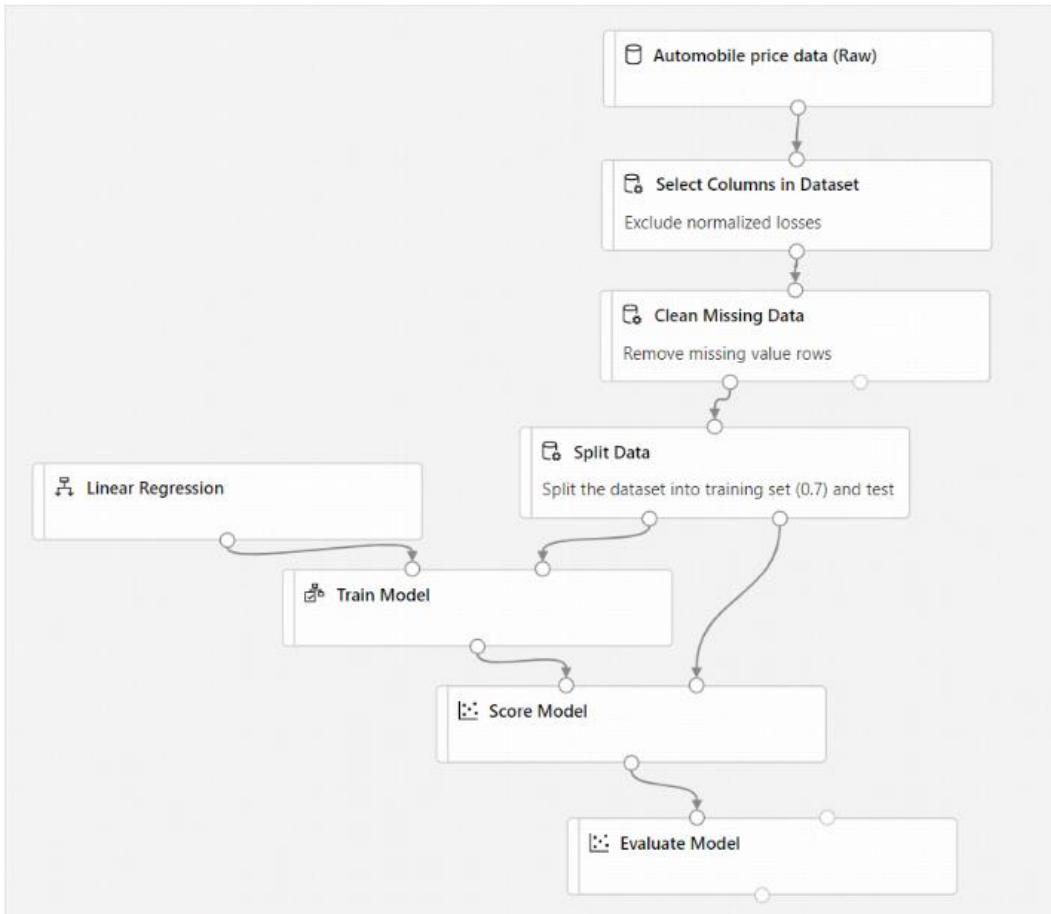


Agregar el componente **Evaluar modelo**

Utilice el componente **Evaluar modelo** para evaluar qué tan bien puntuó su modelo en el conjunto de datos de prueba.

1. Ingrese *evaluar* en el cuadro de búsqueda para encontrar el componente **Evaluar modelo**. Arrastre el componente al lienzo de canalización.
2. Conecte la salida del componente **Score Model** a la entrada izquierda de **Evaluate Model**.

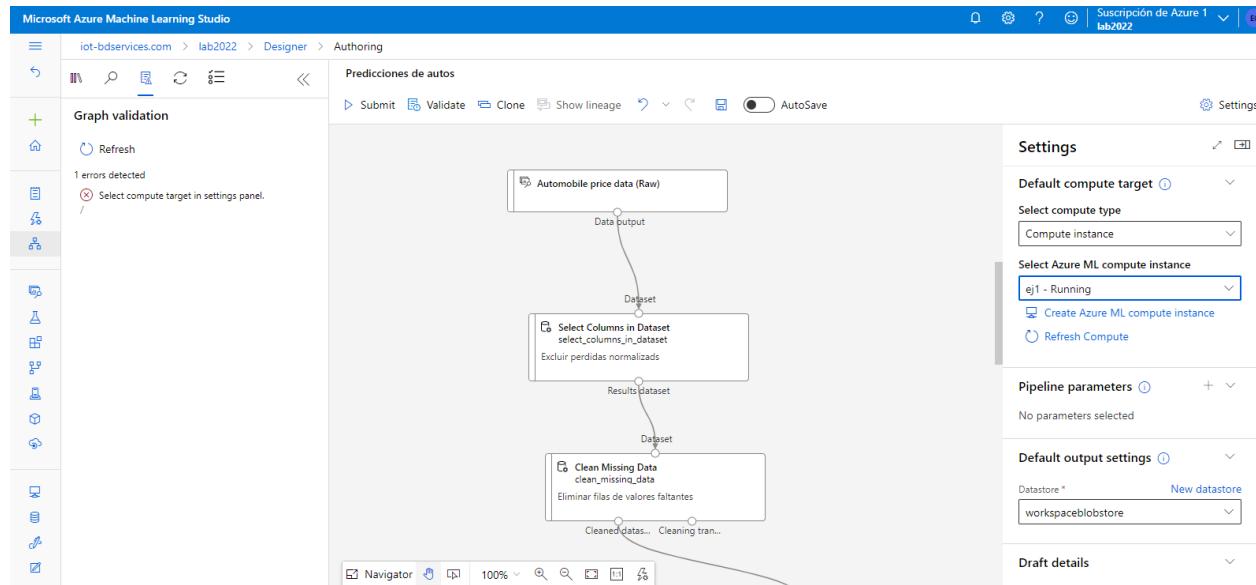
La tubería final debería verse así:



Enviar la canalización

Ahora que su canalización está configurada, puede enviar una ejecución de canalización para entrenar su modelo de aprendizaje automático. Puede enviar una ejecución de canalización válida en cualquier momento, que se puede usar para revisar los cambios en su canalización durante el desarrollo.

1. En la parte superior del lienzo, seleccione **Enviar**. Configurar la Instancia computacional



2. En el cuadro de diálogo **Configurar trabajo de canalización , seleccione Crear nuevo .**

Nota

Los experimentos agrupan ejecuciones de canalizaciones similares. Si ejecuta una canalización varias veces, puede seleccionar el mismo experimento para ejecuciones sucesivas.

0. Para **Nuevo nombre de experimento** , ingrese **Tutorial-CarPrices** .

Set up pipeline job

Experiment

Select existing Create new

New experiment name *

Job display name

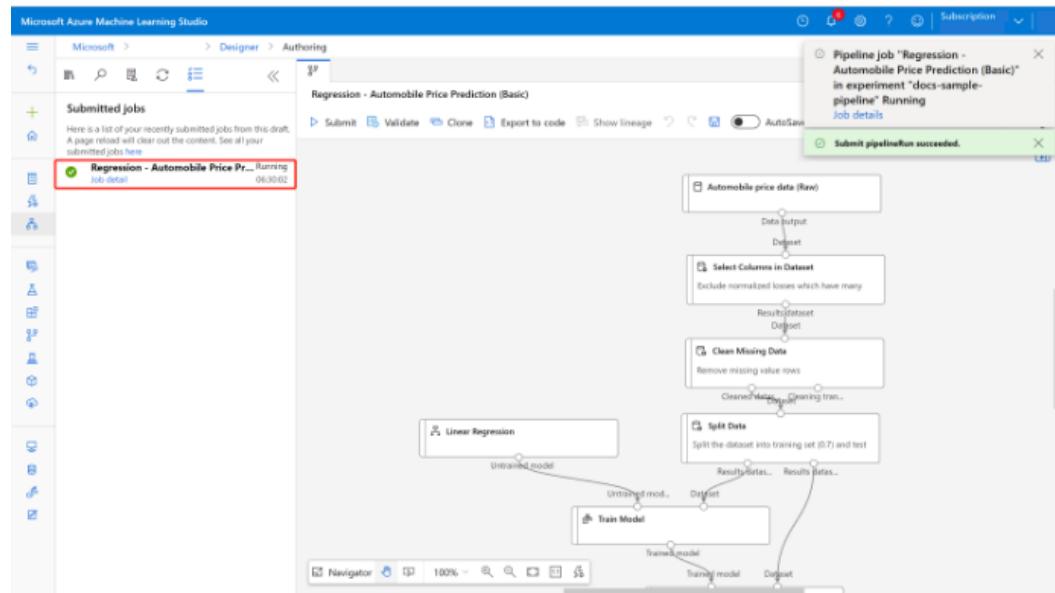
Job description

Job tags : + Delete

Continue on failure step

Submit Cancel

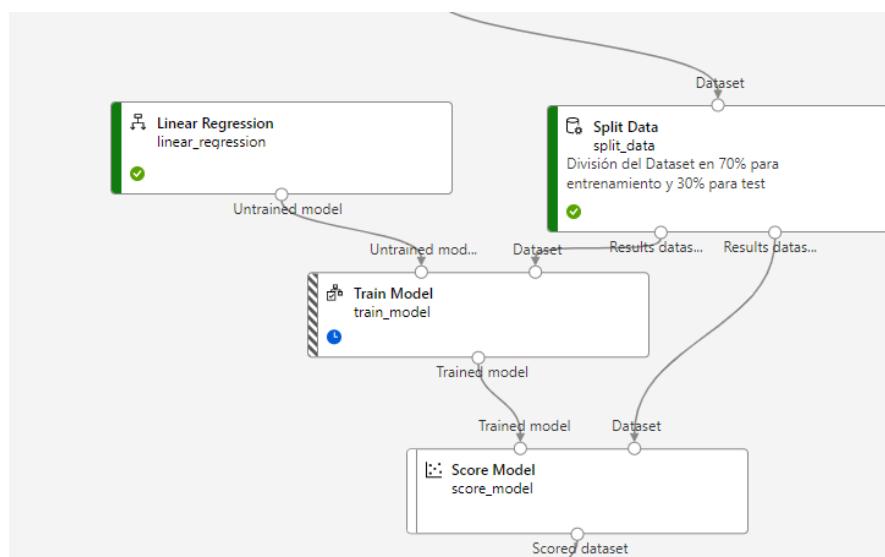
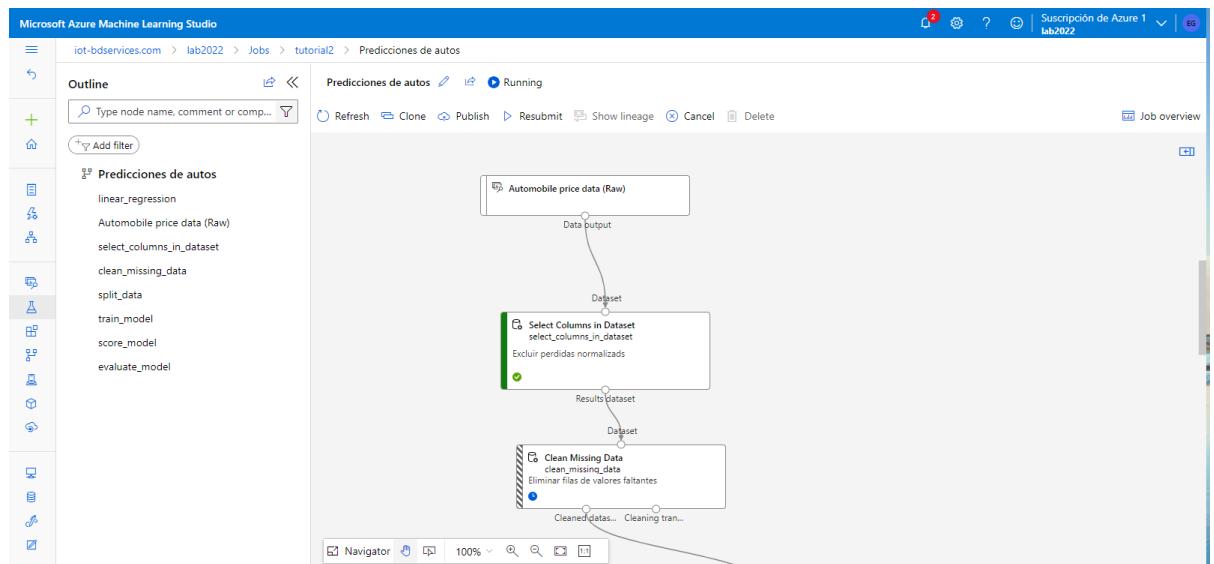
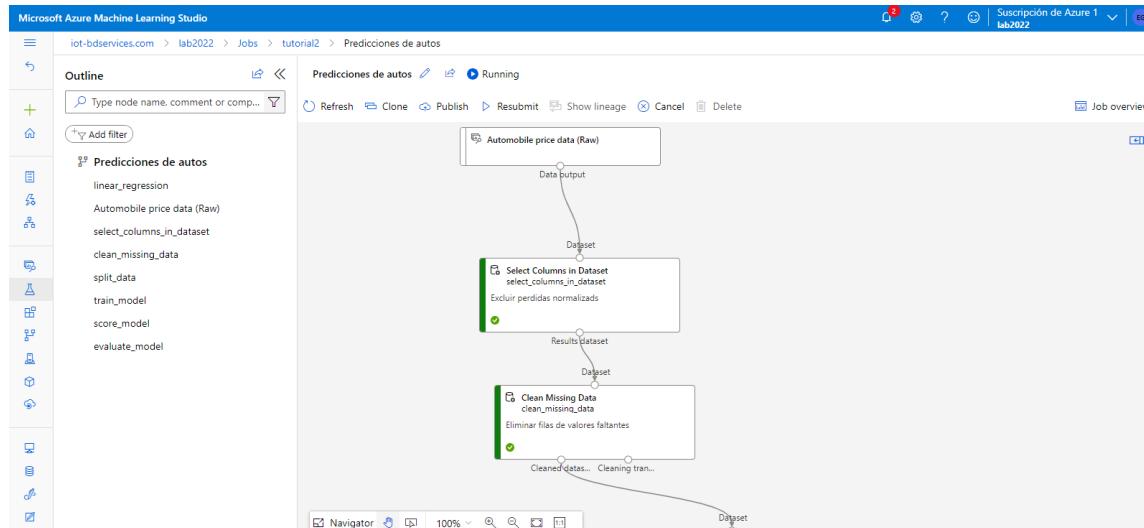
1. Seleccione **Enviar**.
2. Verá una lista de envíos en el panel izquierdo del lienzo y aparecerá una notificación en la esquina superior derecha de la página. Puede seleccionar el enlace Detalles del **trabajo** para ir a la página de detalles del trabajo para la depuración.



1. Si esta es la primera ejecución, la canalización puede tardar hasta 20 minutos en terminar de ejecutarse. La configuración de proceso predeterminada tiene un tamaño de nodo mínimo de 0, lo que significa que el diseñador debe asignar recursos después de estar inactivo. Las ejecuciones de canalización repetidas llevarán menos tiempo, ya que los recursos informáticos ya están asignados. Además, el diseñador utiliza resultados almacenados en caché para cada componente para mejorar aún más la eficiencia.
2. Para ver su ejecución hacemos click en Job Detail, lo que nos llevará a la siguiente pantalla, donde se puede verificar su ejecución y seguir los trabajos (los finalizados los coloca en verde)

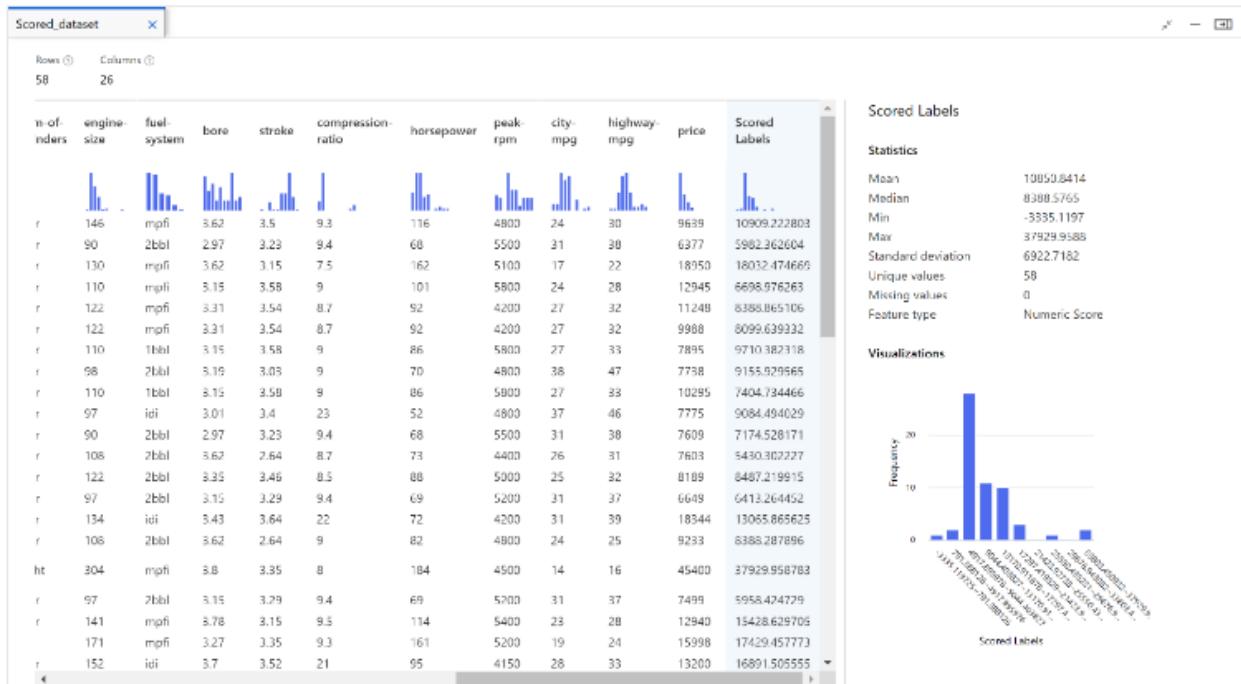
Materia: Minería de Datos

Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning



Ver etiquetas puntuadas

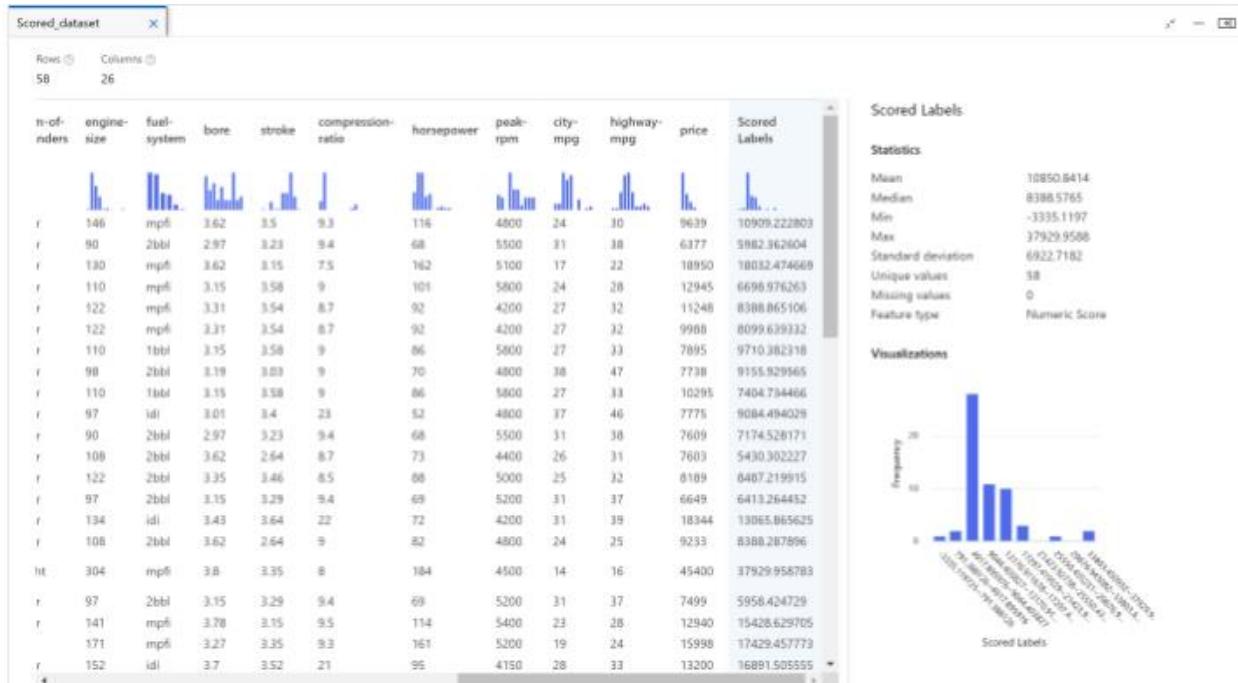
En la página de detalles del trabajo, puede verificar el estado, los resultados y los registros del trabajo de canalización.



Una vez completada la ejecución, puede ver los resultados de la ejecución de la canalización. Primero, mire las predicciones generadas por el modelo de regresión.

1. Haga clic con el botón derecho en el componente **Modelo de puntuación** y seleccione **Vista previa de datos** > **Conjunto de datos puntuado** para ver su salida.

Aquí puede ver los precios pronosticados y los precios reales de los datos de prueba.



Evaluar modelos

Utilice **Evaluar modelo** para ver qué tan bien se desempeñó el modelo entrenado en el conjunto de datos de prueba.

1. Haga clic con el botón derecho en el componente **Evaluar modelo y seleccione Vista previa de datos > Resultados de la evaluación** para ver su salida.

Se muestran las siguientes estadísticas para su modelo:

- **Error Absoluto Medio (MAE)** : El promedio de errores absolutos. Un error es la diferencia entre el valor predicho y el valor real.
- **Error cuadrático medio de la raíz (RMSE)** : la raíz cuadrada del promedio de los errores cuadráticos de las predicciones realizadas en el conjunto de datos de prueba.
- **Error absoluto relativo** : el promedio de errores absolutos en relación con la diferencia absoluta entre los valores reales y el promedio de todos los valores reales.
- **Error cuadrático relativo** : el promedio de los errores cuadráticos en relación con la diferencia cuadrática entre los valores reales y el promedio de todos los valores reales.
- **Coeficiente de determinación** : también conocido como valor R cuadrado, esta métrica estadística indica qué tan bien se ajusta un modelo a los datos.

Para cada una de las estadísticas de error, cuanto más pequeño, mejor. Un valor más pequeño indica que las predicciones están más cerca de los valores reales. Para el coeficiente de determinación, cuanto más cercano sea su valor a uno (1,0), mejores serán las predicciones.

Limpiar recursos

Omita esta sección si desea continuar con la parte 2 del tutorial, [implementación de modelos](#).

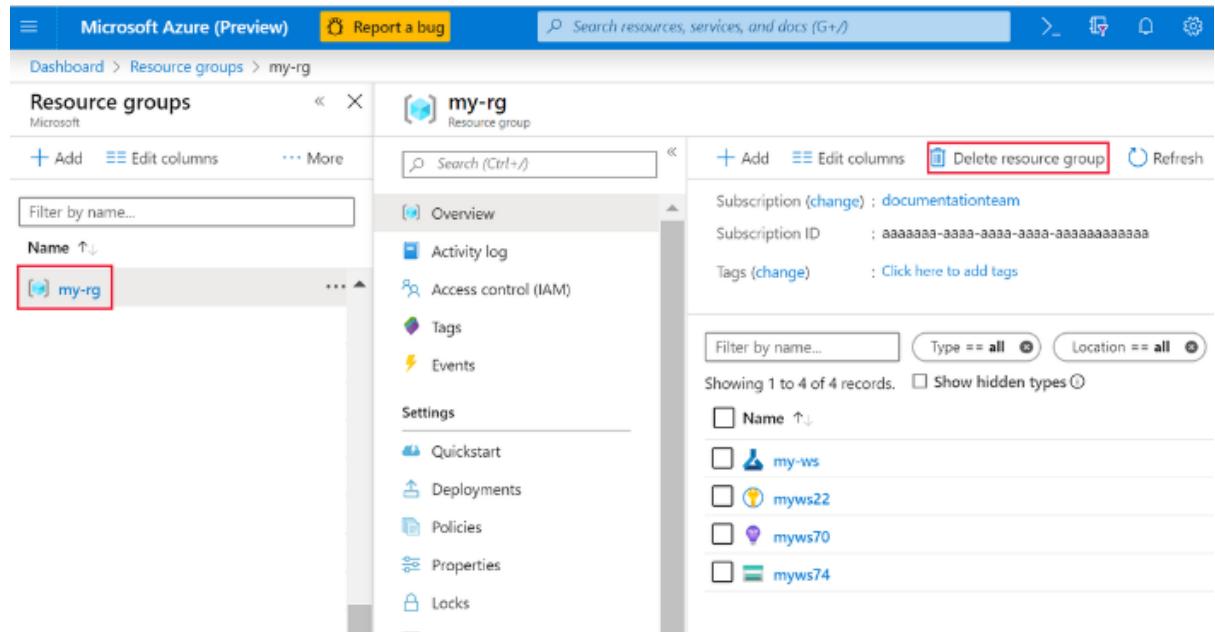
Importante

Puede usar los recursos que creó como requisitos previos para otros tutoriales y artículos de procedimientos de Azure Machine Learning.

Elimina todo

Si no planea usar nada que haya creado, elimine todo el grupo de recursos para no incurrir en ningún cargo.

1. En Azure Portal, seleccione **Grupos de recursos** en el lado izquierdo de la ventana.



The screenshot shows the Azure Portal interface with the 'Resource groups' blade open. On the left, there's a sidebar with a search bar and a 'Resource groups' section. In the main area, there's a list of resource groups, with 'my-rg' selected and highlighted by a red box. To the right of the list, there's an 'Overview' panel with details like 'Subscription (change) : documentationteam', 'Subscription ID : aaaaaaaaaaaaaaaa-aaaa-aaaa-aaaa-aaaaaaaaaaaa', and 'Tags (change) : Click here to add tags'. At the top right of the main content area, there's a red box around the 'Delete resource group' button.

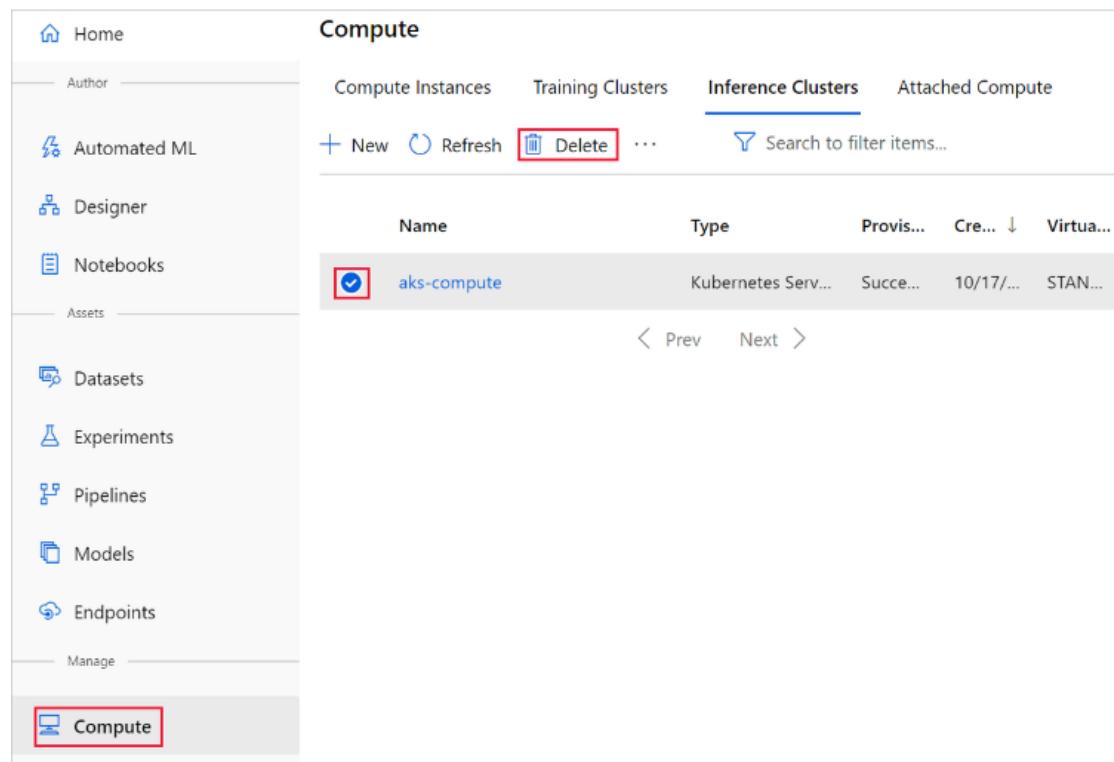
2. En la lista, seleccione el grupo de recursos que creó.
3. Seleccione **Eliminar grupo de recursos**.

Al eliminar el grupo de recursos, también se eliminan todos los recursos que creó en el diseñador.

Eliminar activos individuales

En el diseñador donde creó su experimento, elimine recursos individuales seleccionándolos y luego seleccionando el botón **Eliminar**.

El destino de proceso que creó aquí *se ajusta automáticamente* a cero nodos cuando no se usa. Esta acción se toma para minimizar los cargos. Si desea eliminar el destino de proceso, siga estos pasos:



The screenshot shows the 'Compute' blade in the Azure Machine Learning studio. The left sidebar has sections for Home, Author, Automated ML, Designer, Notebooks, Assets (Datasets, Experiments, Pipelines, Models, Endpoints), Manage, and Compute. The Compute section is highlighted with a red box. The main area is titled 'Compute' and shows tabs for Compute Instances, Training Clusters, Inference Clusters (which is underlined), and Attached Compute. Below the tabs are buttons for New, Refresh, Delete (which is also highlighted with a red box), and There is a search bar: 'Search to filter items...'. A table lists an inference cluster: Name (aks-compute), Type (Kubernetes Serv...), Provisioning Status (Success), Creation Date (10/17/...), and Status (STAN...). Navigation arrows < Prev and Next > are at the bottom of the table.

Puede anular el registro de conjuntos de datos de su espacio de trabajo seleccionando cada conjunto de datos y seleccionando **Anular registro**.

Materia: Minería de Datos

Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning

The screenshot shows the Azure Machine Learning Studio interface. On the left, a sidebar lists various assets: New, Home, Author, Notebooks, Automated ML, Designer, Assets (with Datasets highlighted), Experiments, Pipelines, Models, Endpoints, Manage, Compute, Environments, Datastores, and Data labeling. The main content area displays a dataset named 'TD-Sample_1: Regression - Automobile_Price_Prediction_(Basic)-Clean_Missing_Data-Cleaning_transformation-f6dc0eb1'. The 'Datasets' tab is selected. In the top navigation bar, there are buttons for Refresh, Generate profile, Unregister (highlighted with a red box), and New version. The dataset details include:

- Properties:** File
- Description:** This is a dataset promoted by inference graph generation automatically on 11/12/2018.
- Datastore:** workspaceblobstore
- Relative path:** azureml/4393076b-19ff-4e41-81d9-a1146d905096/Cleaning_transformation
- Profile:** No profile generated
- Files in dataset:** 4
- Current version:** 1
- Latest version:** 1

Tags:

- CreatedByAMLStudio
- true

Sample usage:

```
# azureml-core of version 1.0.72 or higher is required
from azureml.core import Workspace, Dataset

subscription_id = 'ec085ed72-2b26-48f6-aee8-ccb5bcf98fb9'
resource_group = 'test-like'
workspace_name = 'like_test'

workspace = Workspace(subscription_id, resource_group, workspace_name)

dataset = Dataset.get_by_name(workspace, name='TD-Sample_1: Regression - Aut
dataset.download(target_path='.', overwrite=False)
```

Para eliminar un conjunto de datos, vaya a la cuenta de almacenamiento mediante Azure Portal o Azure Storage Explorer y elimine manualmente esos activos.

Materia: Minería de Datos

Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning

Fuente:

<https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-designer-automobile-price-train-score>