

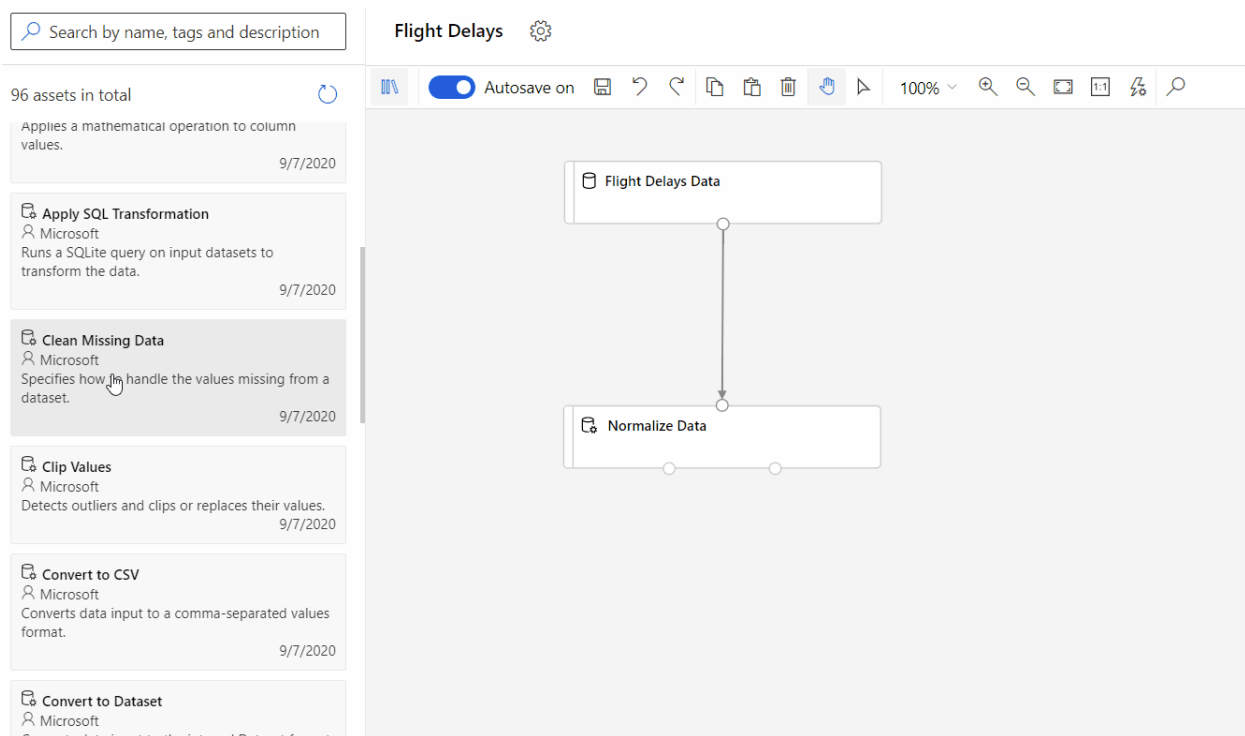
Introducción:

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:

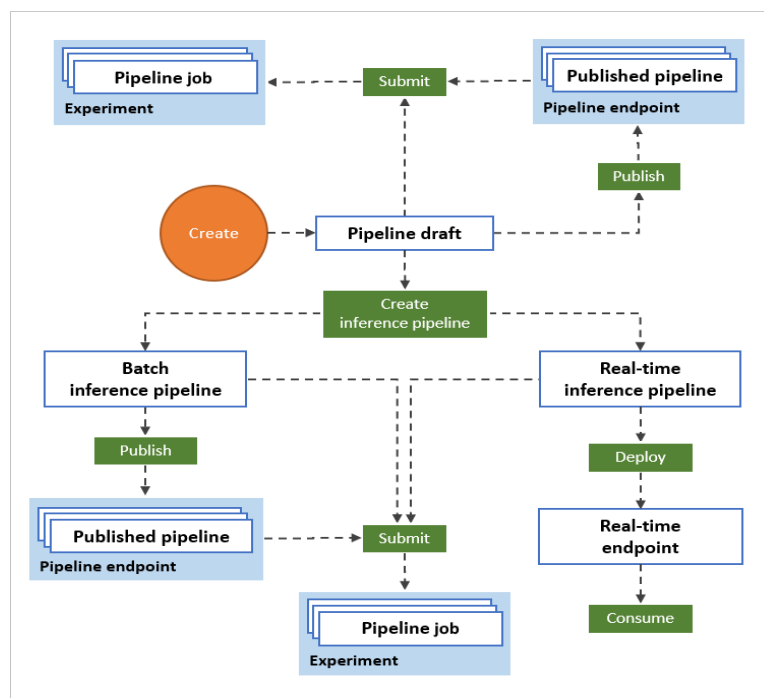
- Tuberías
- conjuntos de datos
- Recursos informáticos

- 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.

The image shows a screenshot of the Azure Machine Learning Designer interface. On the left, a workflow is visible with three components: 'Automobile price data (Raw)', 'Select Columns in Dataset' (with the description 'Exclude normalized losses which have many'), and 'Clean Missing Data' (with the description 'Remove missing value rows'). The 'Clean Missing Data' component is highlighted with a red box, and a red arrow points from it to the configuration panel on the right. The configuration panel, titled 'Clean Missing Data', contains the following settings:

- Columns to be cleaned** *: A dropdown menu set to 'All columns' with an 'Edit column' link.
- Minimum missing value ratio** *: A text input field set to '0.0'.
- Maximum missing value ratio** *: A text input field set to '1.0'.
- Cleaning mode** *: A dropdown menu set to 'Remove entire row'.
- Regenerate output** *: An unchecked checkbox.
- Run settings** *: A section with a dropdown arrow, containing:
 - Compute target** *: A radio button selected for 'Use default compute target', with links for 'No compute target selected' and 'Select default compute target'.
 - A radio button for 'Use other compute target'.

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

Materia: Minería de Datos

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

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:

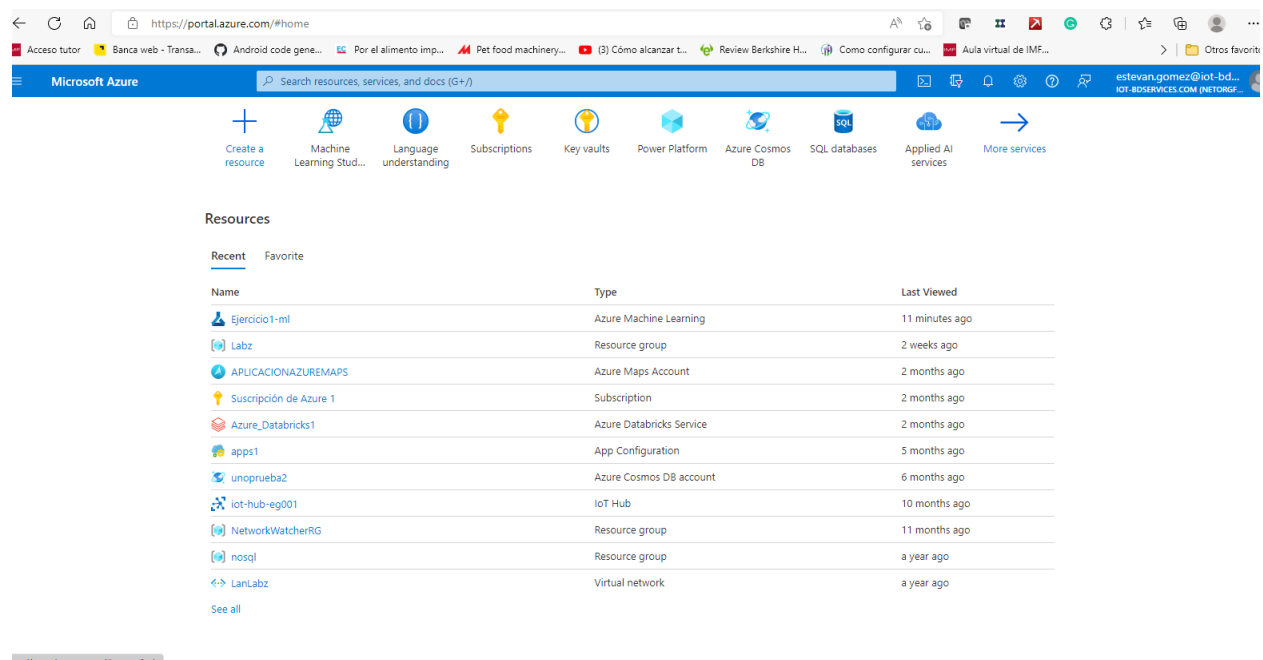
- Crear una nueva canalización.
- Importar Datos
- Preparar datos.
- Entrenar un modelo de aprendizaje automático.
- Evaluar un modelo de aprendizaje automático.

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.



Crear un nuevo espacio de trabajo

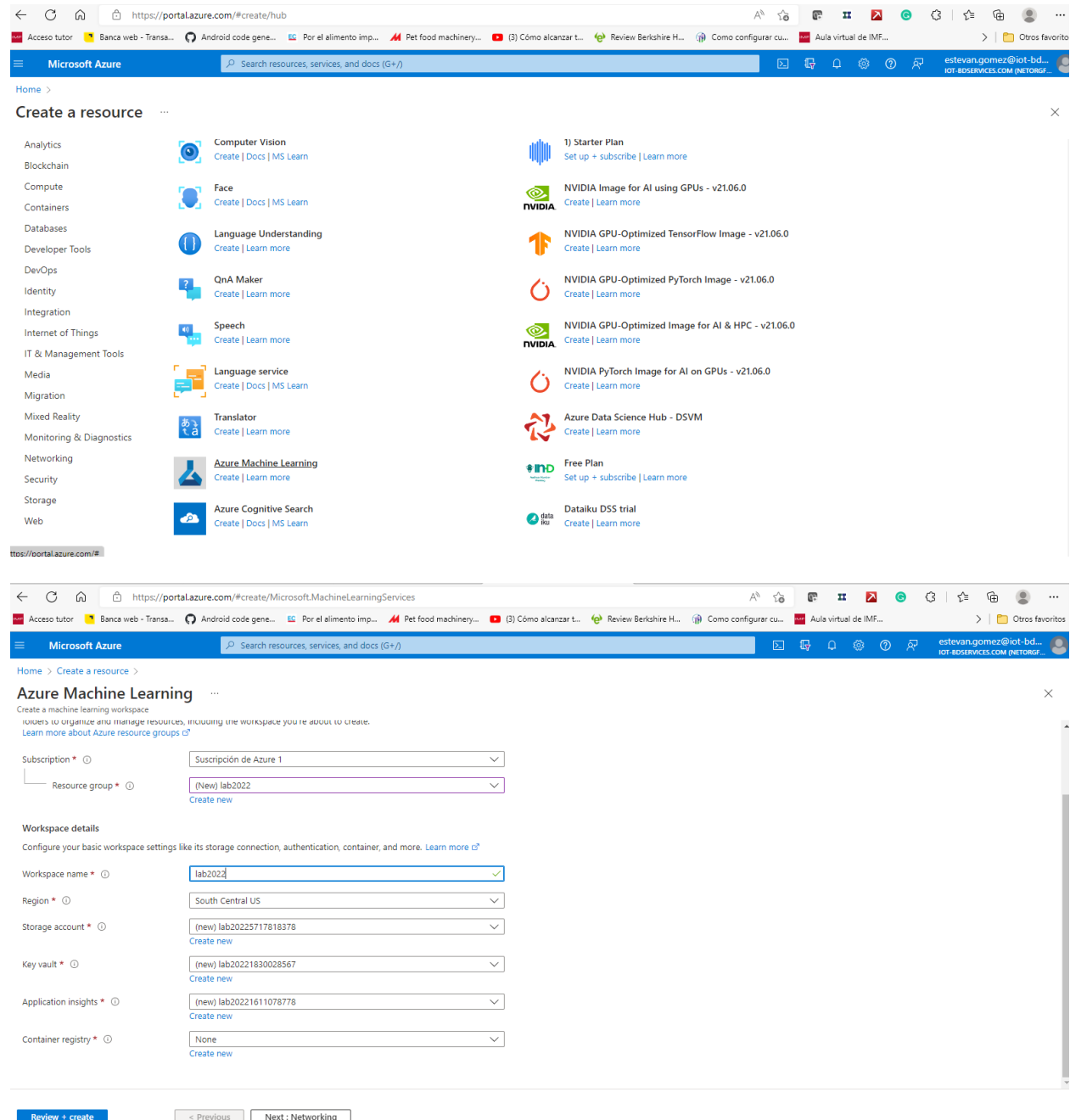
ESPE 2022

Prof: Ing Estevan Gómez, MSc PhD(c)

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 Azure portal interface. The top navigation bar includes the Microsoft Azure logo and a search bar. The main content area is titled 'Create a resource' and displays a grid of service tiles. The 'Azure Machine Learning' tile is highlighted, showing its icon and name. Below the grid, the 'Workspace details' section is visible, containing a form with the following fields:

- Subscription: Suscripción de Azure 1
- Resource group: (New) lab2022
- Workspace name: lab2022
- Region: South Central US
- Storage account: (new) lab20225717818378
- Key vault: (new) lab20221830028567
- Application insights: (new) lab20221611078778
- Container registry: None

At the bottom of the form, there are buttons for 'Review + create', '< Previous', and 'Next: Networking'.

ESPE 2022

Prof: Ing Estevan Gómez, MSc PhD(c)

Materia: Minería de Datos

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

Home > Create a resource >

Azure Machine Learning

Create a machine learning workspace

Basics Networking Advanced Tags Review + create

Basics

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

Connectivity method	Enable public access from all networks
---------------------	--

Advanced

Identity type	System assigned
Encryption type	Microsoft-managed keys
Enable HBI Flag	Disabled

Create < Previous Next > Download a template for automation

Página

Home >

Microsoft.MachineLearningServices | Overview

Deployment

Search (Ctrl+F) Delete Cancel Redeploy Refresh

We'd love your feedback! →

*** Deployment is in progress

Deployment name: Microsoft.MachineLearningServices Start time: 7/7/2022, 6:11:17 PM
Subscription: Suscripción de Azure 1 Correlation ID: c562738e-4788-496d-b1de-126f84aa68cf
Resource group: lab2022

Deployment details (Download)

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

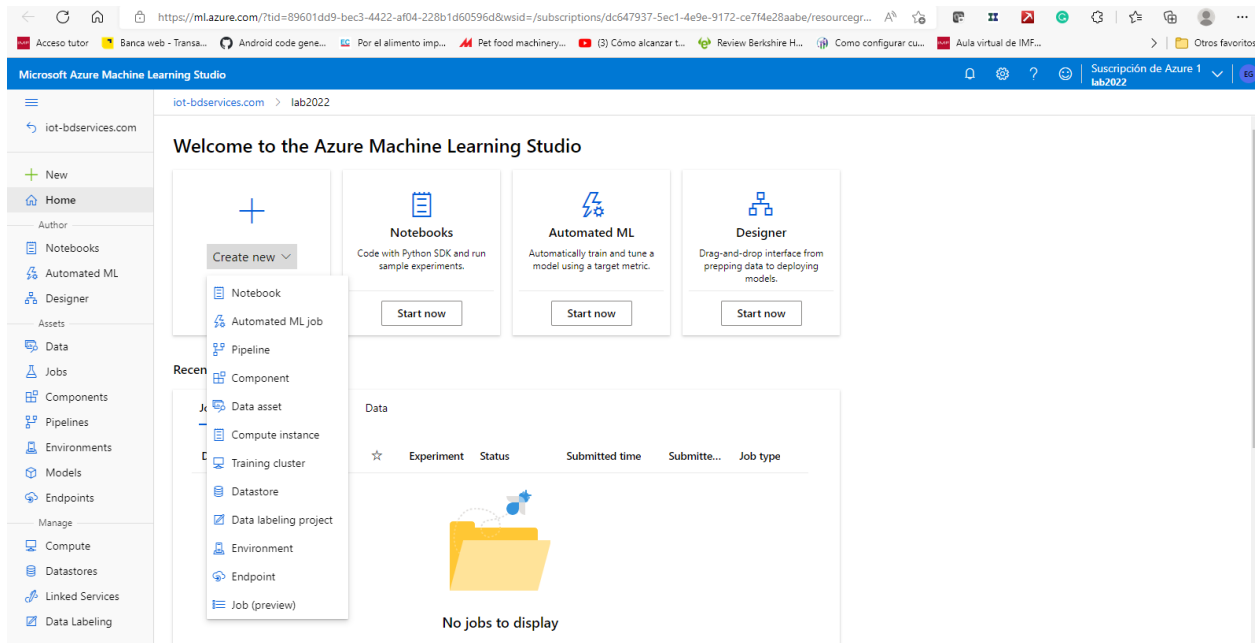
The screenshot shows the Microsoft Azure portal interface. The top navigation bar includes the Azure logo and a search bar. The main content area displays the 'Microsoft.MachineLearningServices | Overview' page. A green checkmark indicates 'Your deployment is complete'. Below this, deployment details are listed: Deployment name: Microsoft.MachineLearningServices, Subscription: Suscripción de Azure 1, Resource group: lab2022, Start time: 7/7/2022, 6:11:17 PM, and Correlation ID: c562738e-4788-496d-b1de-126f84aa68cf. A 'Go to resource' button is visible. On the right, there are sections for 'Cost Management', 'Microsoft Defender for Cloud', and 'Free Microsoft tutorials'.

Escogemos ir al recurso y nos presenta la siguiente pantalla:

The screenshot shows the Microsoft Azure portal interface. The top navigation bar includes the Azure logo and a search bar. The main content area displays the 'lab2022' resource details. The left sidebar shows the 'Overview' tab selected. The main content area shows the 'Essentials' section with various resource details: 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-bec3-4422-af04-228b1d60596d&wsid=js..., Registry: ..., Key Vault: lab20221830028567, Application Insights: lab20221611078778, and MLflow tracking URI: azureml/southcentralus.api.azureml.ms/mlflow/v1.0/subscriptions/dc647937-... Below this, there is a section titled 'Work with your models in Azure Machine Learning Studio' with a 'Launch studio' button.

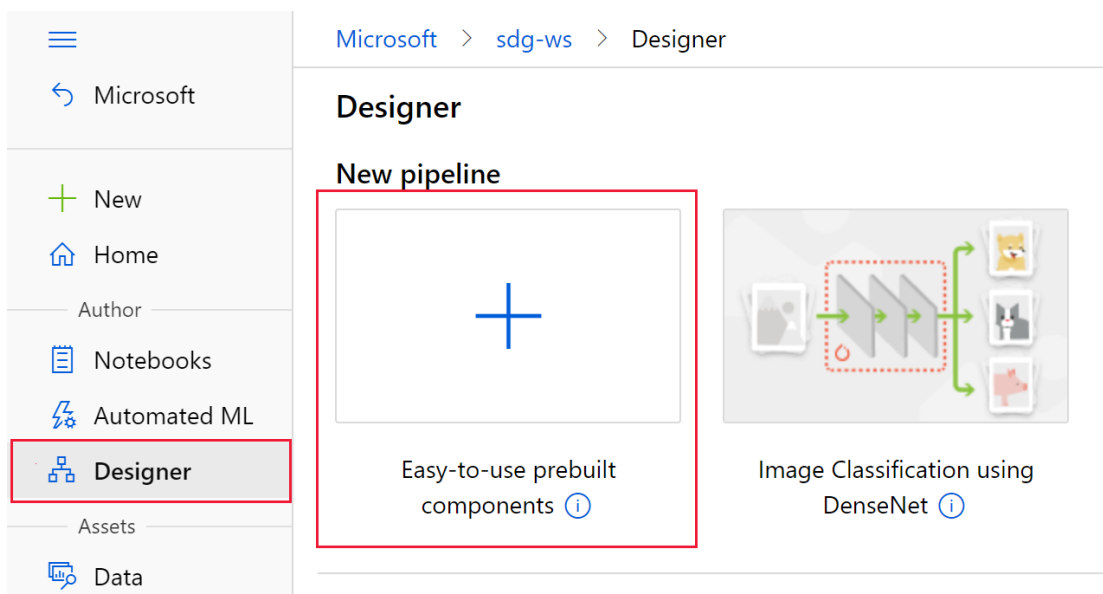
Escogemos ir a Studio, entonces estamos en Azure Machine Learning Studio

Al seleccionar + aparecen todas las opciones que podemos crear

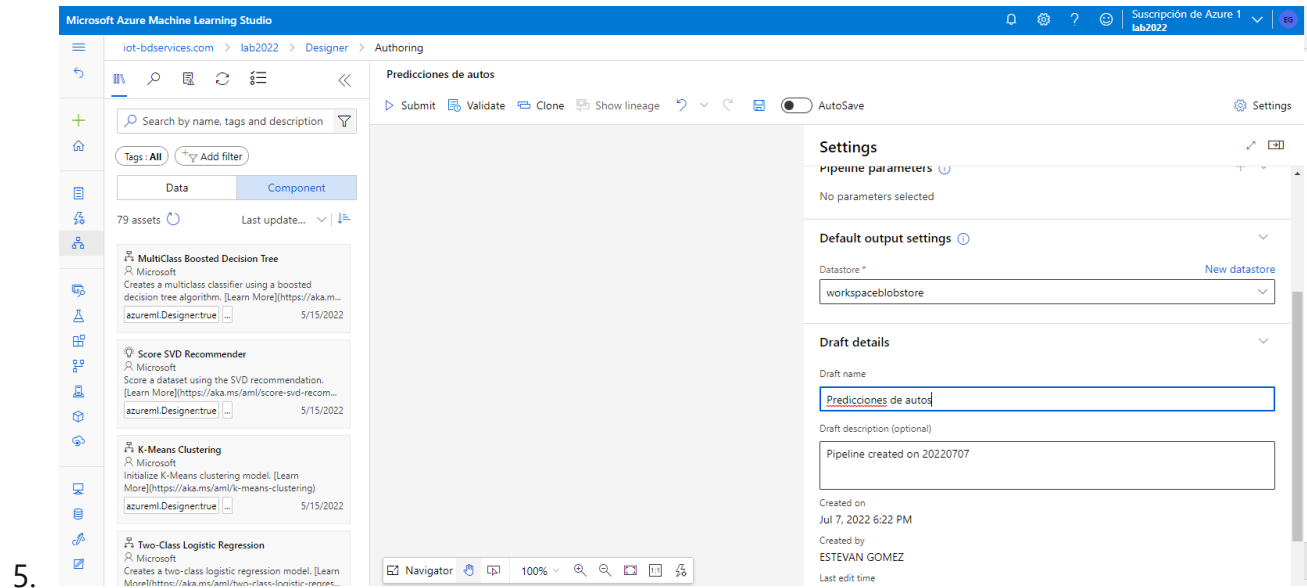


Crear la canalización

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



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.



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 **icono 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)

Can not be shared. It can only be used by a single user. Each user can only be assigned to one creator and you can change this to a different user in the advanced settings section.

Required Settings

Advanced Settings

Compute name *

ej1

Location

southcentralus

Virtual machine type

☒ CPU ☐ GPU

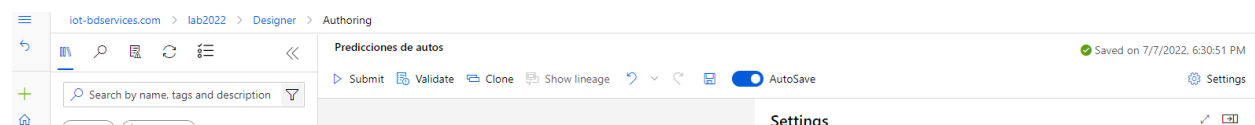
Virtual machine size

☒ Select from recommended options ☐ Select from all options

Name ↑	Category	Workload types	Available quota	Cost
<input checked="" type="checkbox"/> Standard_F4s_v2 4 cores, 8GB RAM, 32GB storage	Compute optimized	Data manipulation and training on large datasets (>10 GB)	8 cores	\$0.20/hr

Create Back Next: Advanced Settings Cancel

Activar el autosave



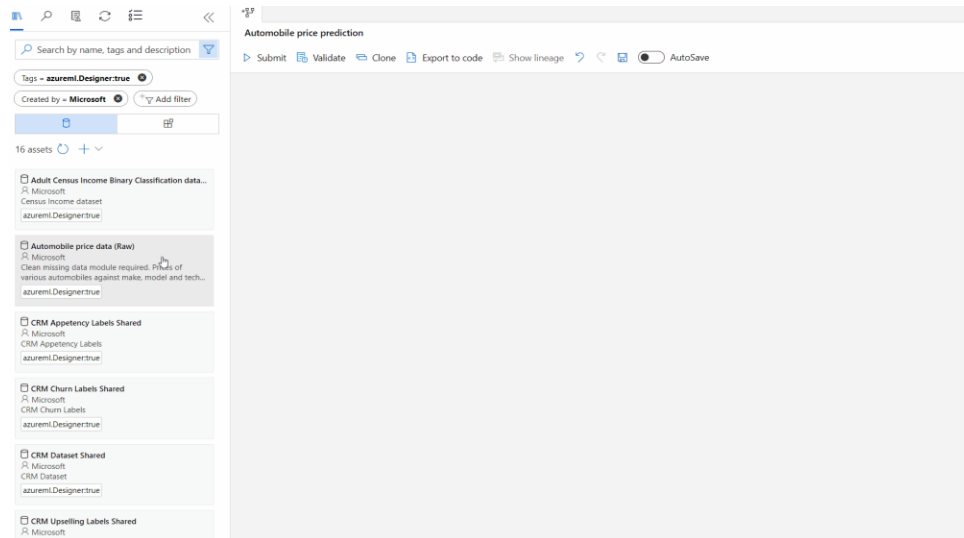
Datos de importación

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)**.

Materia: Minería de Datos

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

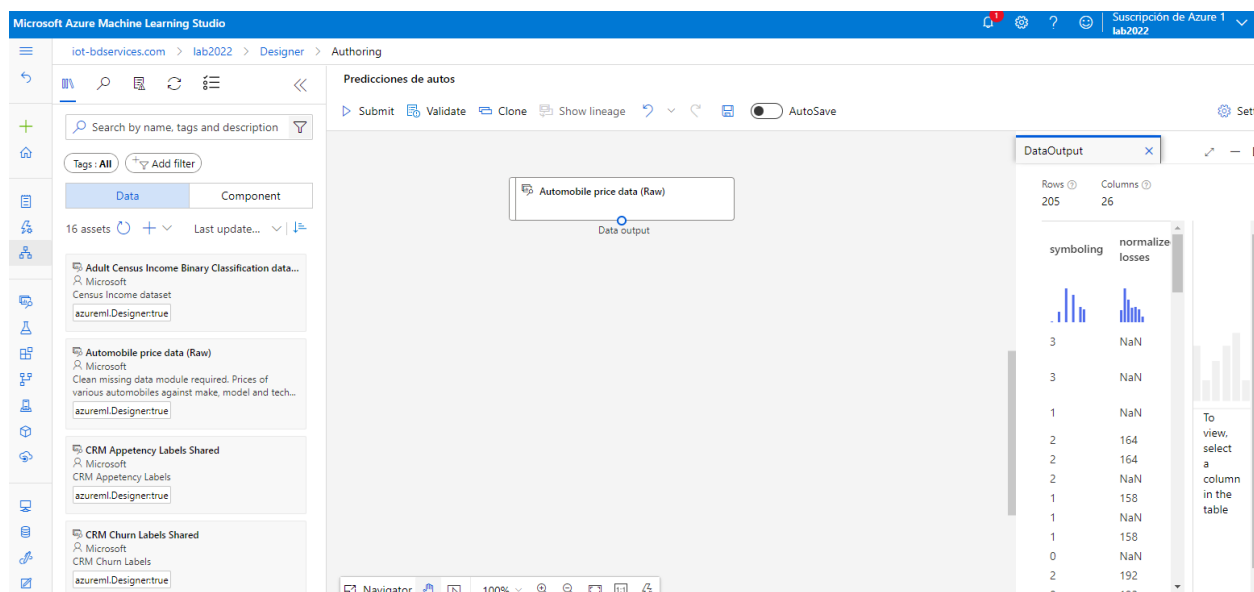
1. A la izquierda del lienzo de canalización hay una paleta de conjuntos de datos y componentes. Seleccione Conjuntos de **datos de muestra** para ver los conjuntos de datos de muestra disponibles.
2. Seleccione el conjunto **de datos Datos de precios de automóviles (sin procesar)** y arrástrelo al lienzo.



Visualiza los datos

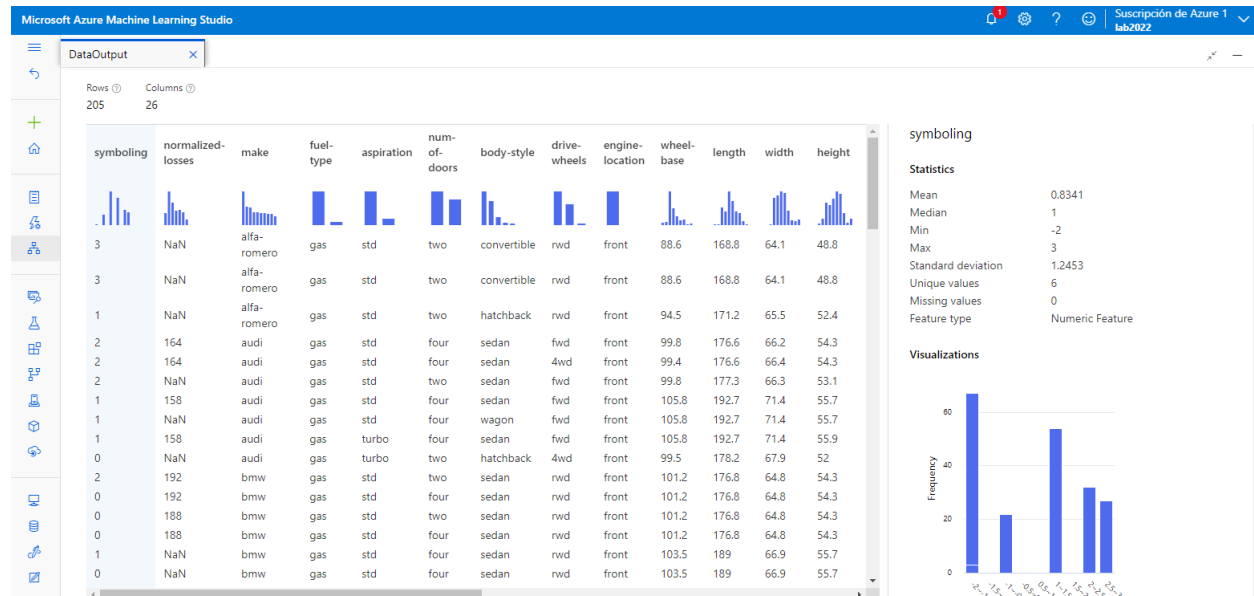
Puede visualizar los datos para comprender el conjunto de datos que utilizará.

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.



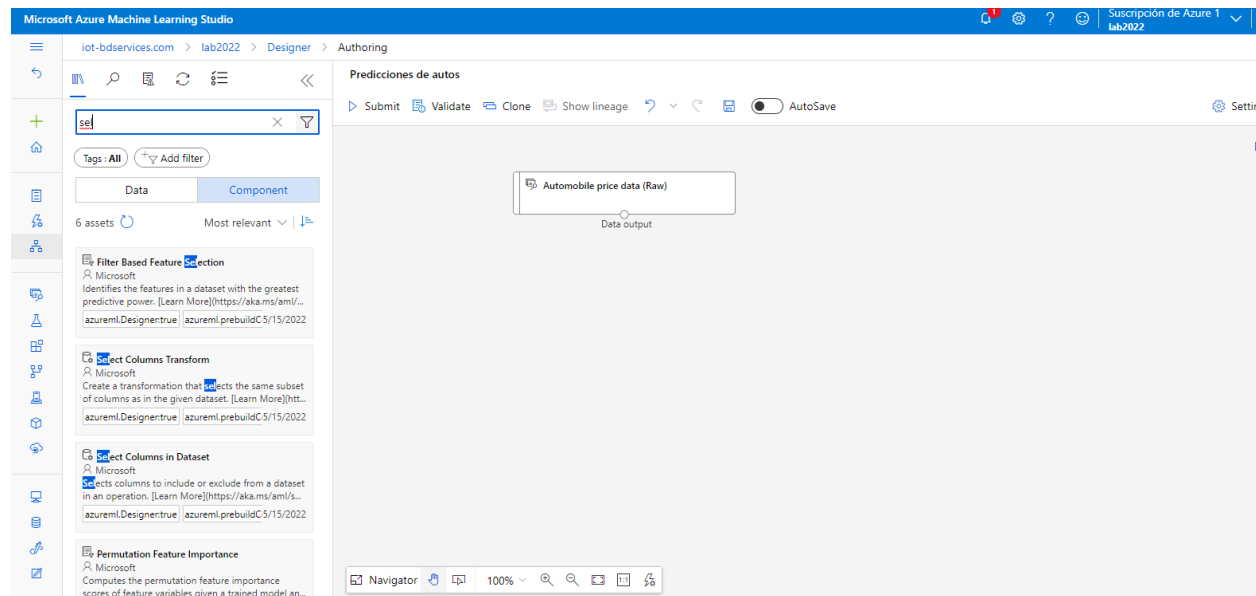
Preparar datos

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.

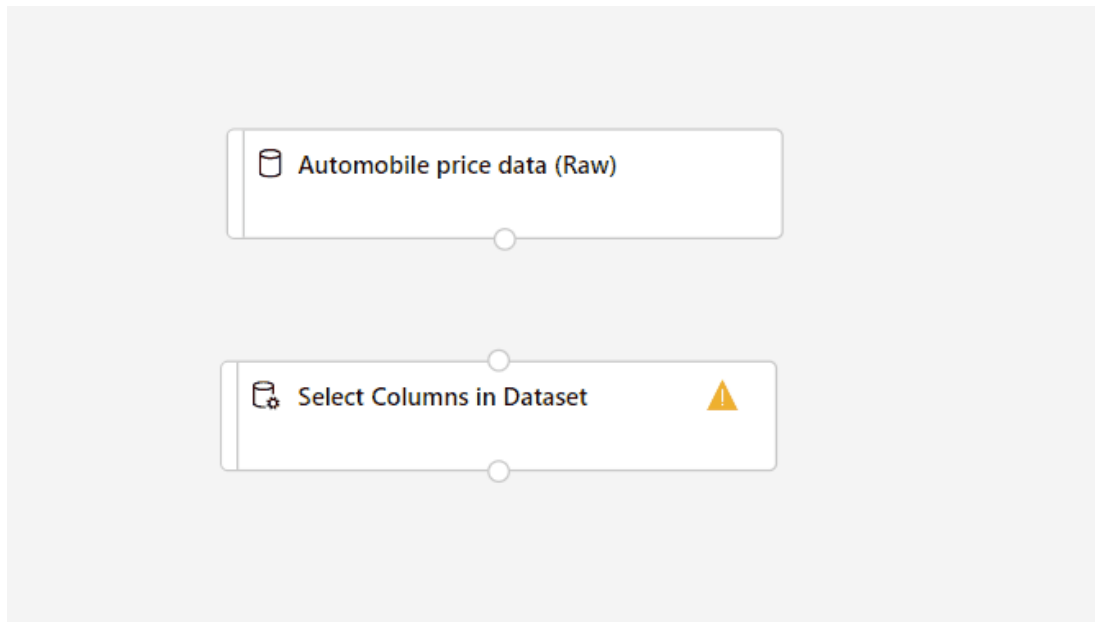
1. En la paleta de componentes a la izquierda del lienzo, expanda la sección **Transformación de datos** y busque el componente **Seleccionar columnas en conjunto de datos**.



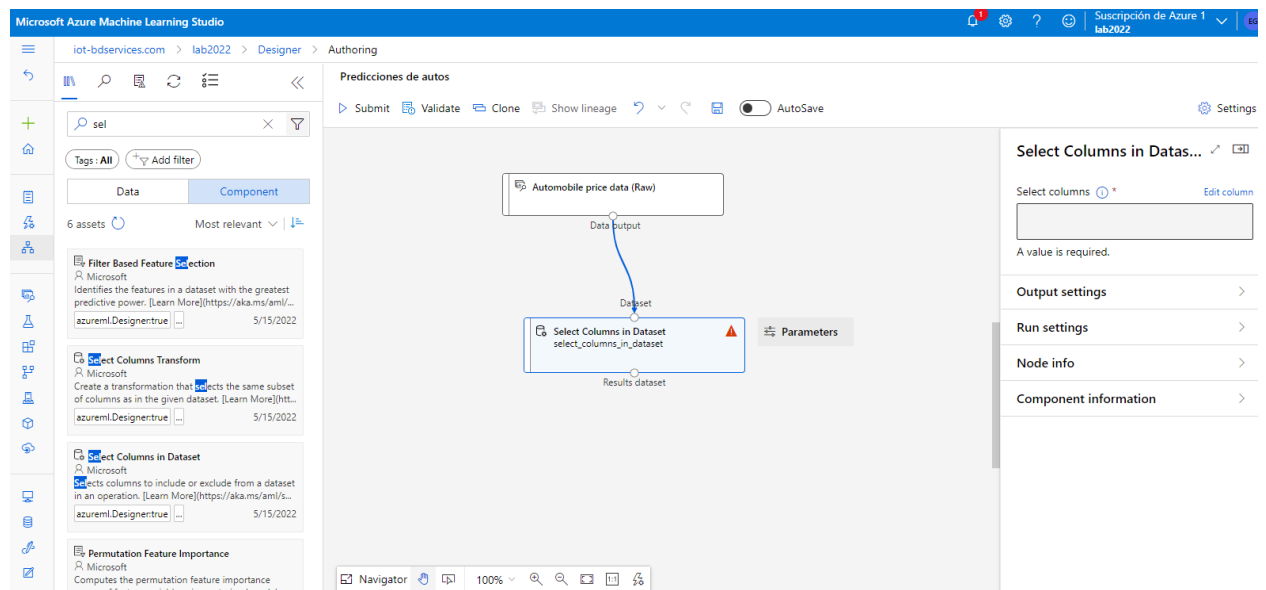
2. Arrastre el componente **Seleccionar columnas en conjunto de datos** al lienzo. Suelte el componente debajo del componente del conjunto de datos.
3. Conecte el conjunto de datos de **datos de precios de automóviles (sin procesar)** al componente **Seleccionar columnas en conjunto de datos**. Arrastre 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 .



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.

Select columns

Select columns ☒ With rules ☐ By name

Allow duplicates and preserve column order in selection ☐

Include All columns

Exclude Column names normalized-losses

Save Cancel

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.

Microsoft Azure Machine Learning Studio

iot-bdservices.com > lab2022 > Designer > Authoring

Predicciones de autos

Submit Validate Clone Show lineage AutoSave

Tags: All Add filter

Data Component

6 assets Most relevant

Filter Based Feature Selection

Select Columns Transform

Select Columns in Dataset

Permutation Feature Importance

Select Columns in Dataset

Automobile price data (Raw)

Data Output

Dataset

Select Columns in Dataset

Excluir pérdidas normalizadas

Results dataset

Parameters

Select Columns in Datas...

Select columns ☒ All columns ☐ Exclude column names: normalized-losses

Output settings

Run settings

Node info

Node name

select_columns_in_dataset

Comment

Excluir pérdidas normalizadas

Component information

Name

azureml/Select Columns in Dataset

Description

13.

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

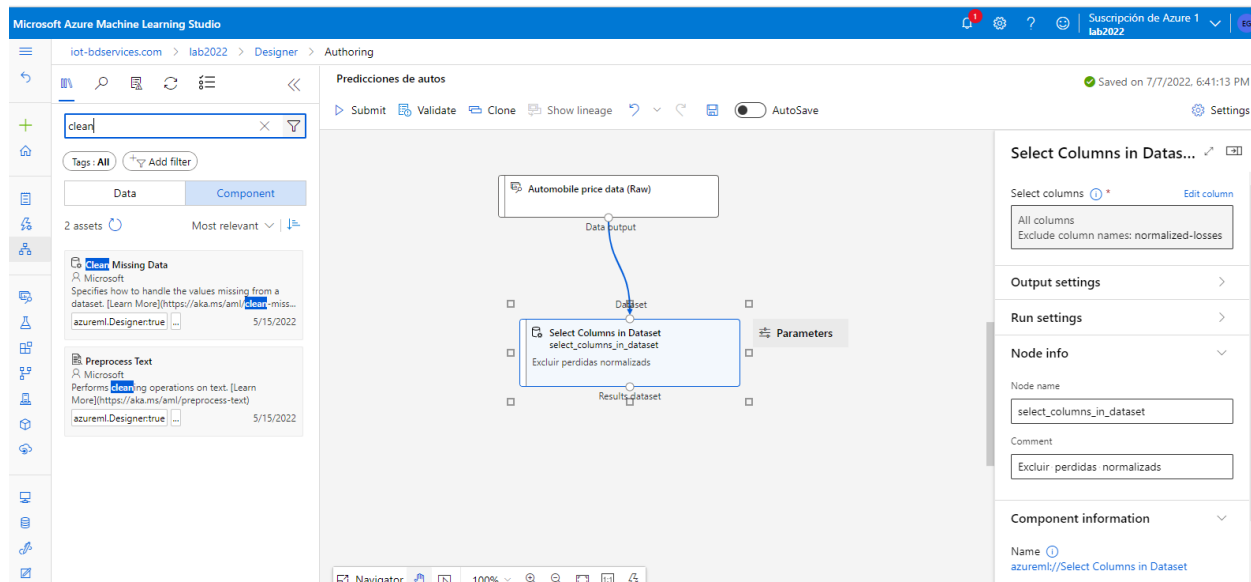
Limpiar datos faltantes

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**.

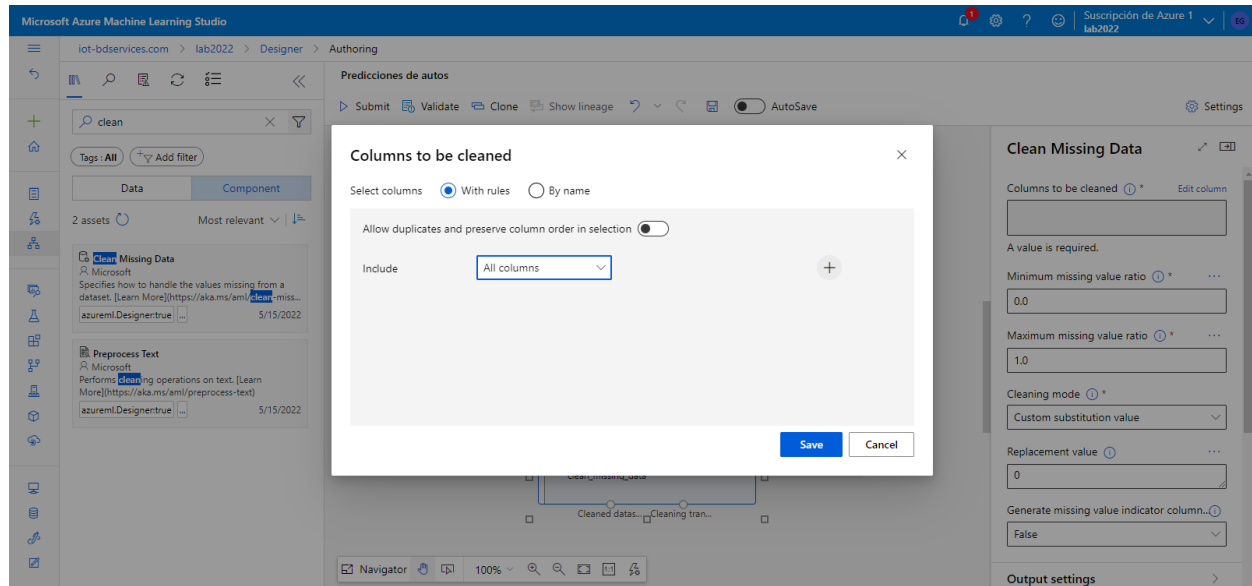


2. Arrastre el componente **Limpiar datos que faltan al lienzo de canalización**. Conéctelo al componente **Seleccionar columnas en conjunto de datos**.
3. Seleccione el componente **Limpiar datos que faltan**.
4. En el panel de detalles del componente a la derecha del lienzo, seleccione **Editar columna**.

Materia: Minería de Datos

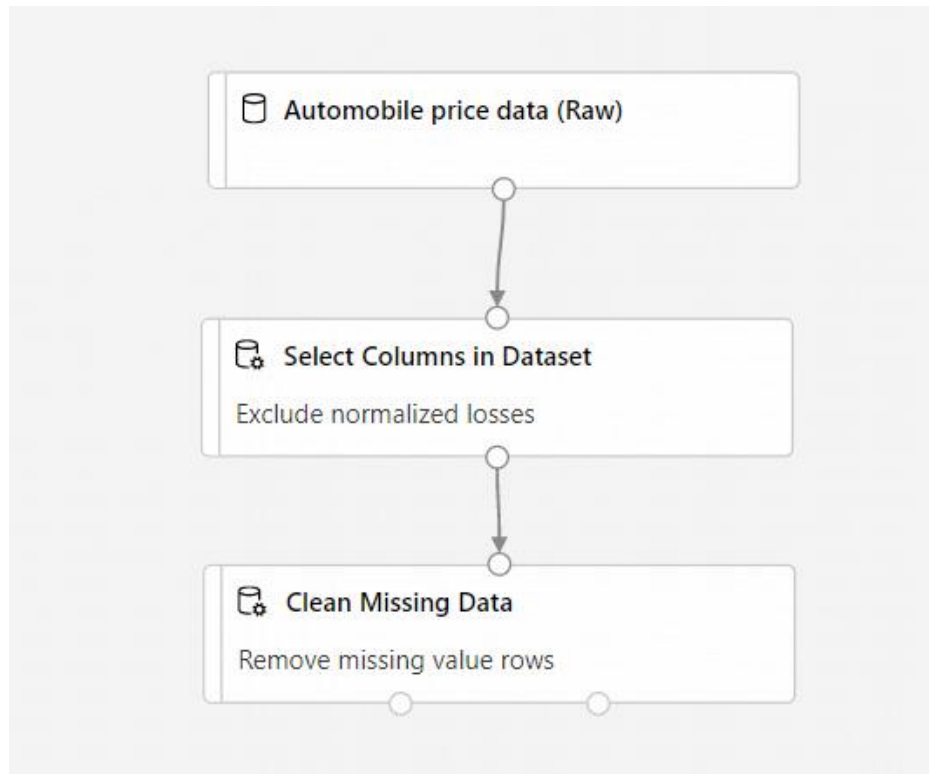
Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning

5. En la ventana **Columnas para limpiar** que aparece, expanda el menú desplegable junto a **Incluir** . Seleccione, **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

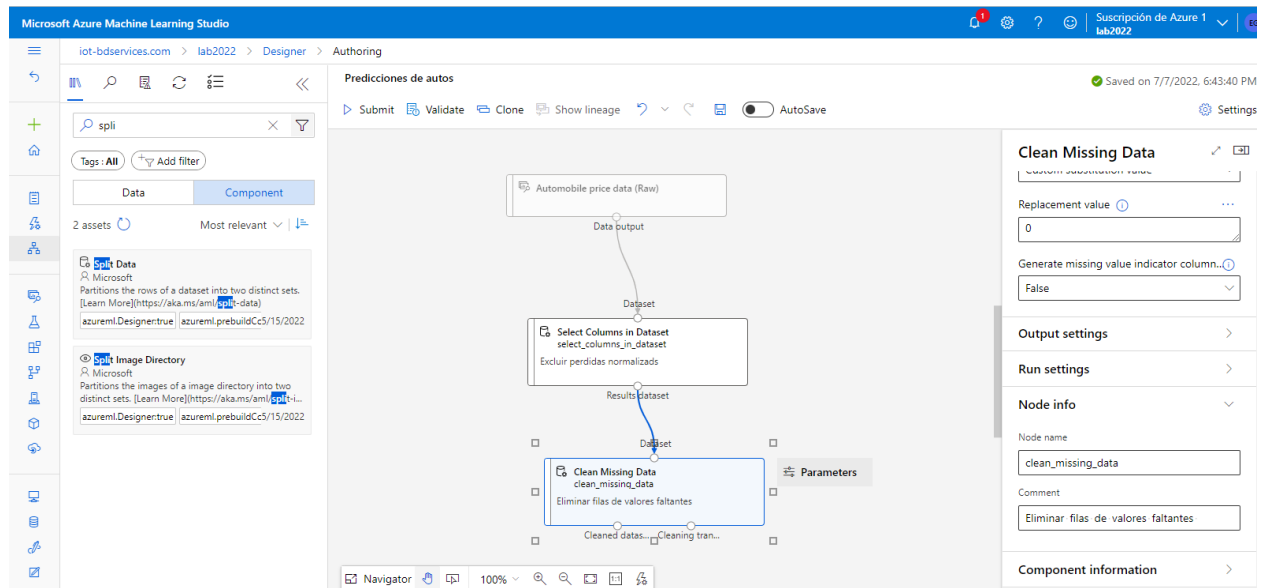
Ahora que tiene los componentes para procesar los datos, puede configurar los componentes de entrenamiento.

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**.



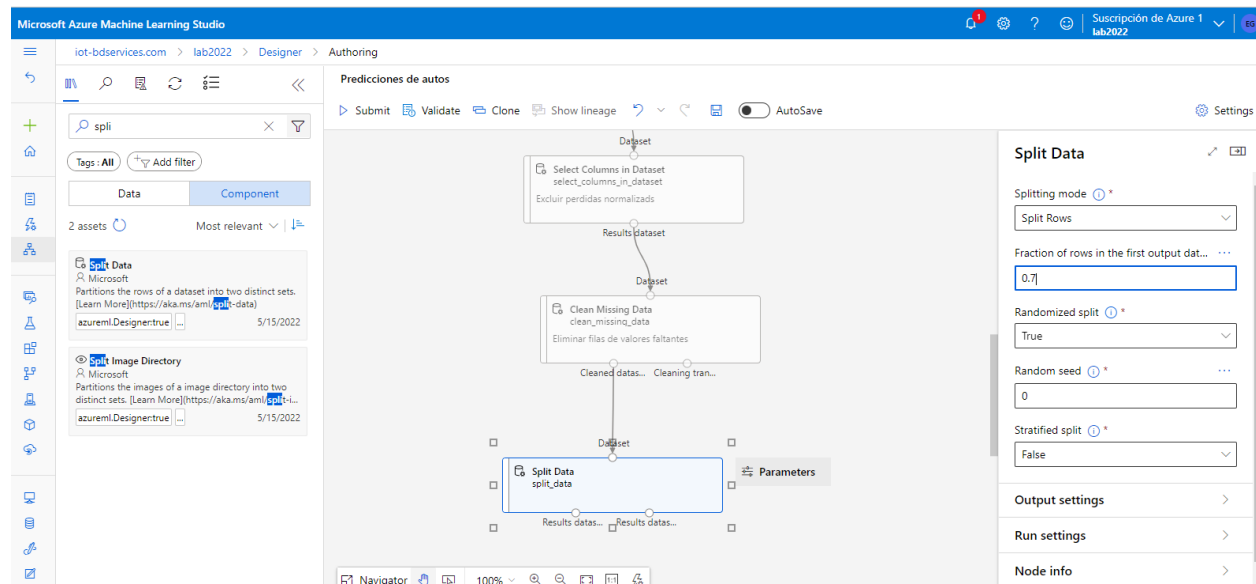
2. Arrastre el componente **Dividir datos** al lienzo de canalización.
3. Conecte el puerto izquierdo del componente **Limpiar datos faltantes** al componente **Dividir datos**.

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.



6. 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.

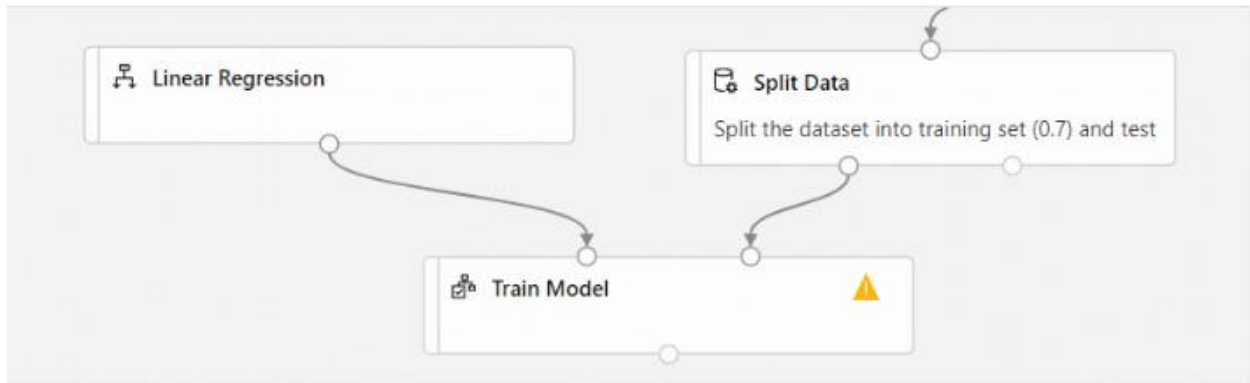
1. En la paleta de componentes, expanda **Algoritmos de aprendizaje automático** .

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

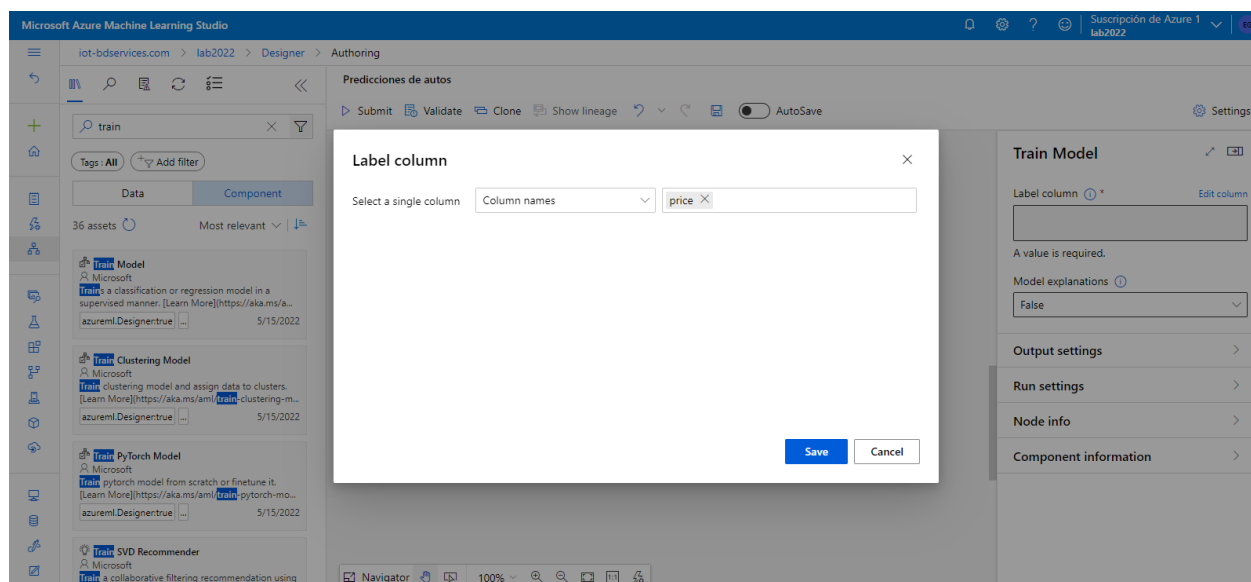
2. Seleccione **Regresión** > **Regresión lineal** y arrástrelo al lienzo de canalización.
3. En la paleta de componentes, expanda la sección **Entrenamiento de módulos** y arrastre el componente **Entrenar modelo** al lienzo.
4. Conecte la salida del componente **Regresión lineal** a la entrada izquierda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

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.



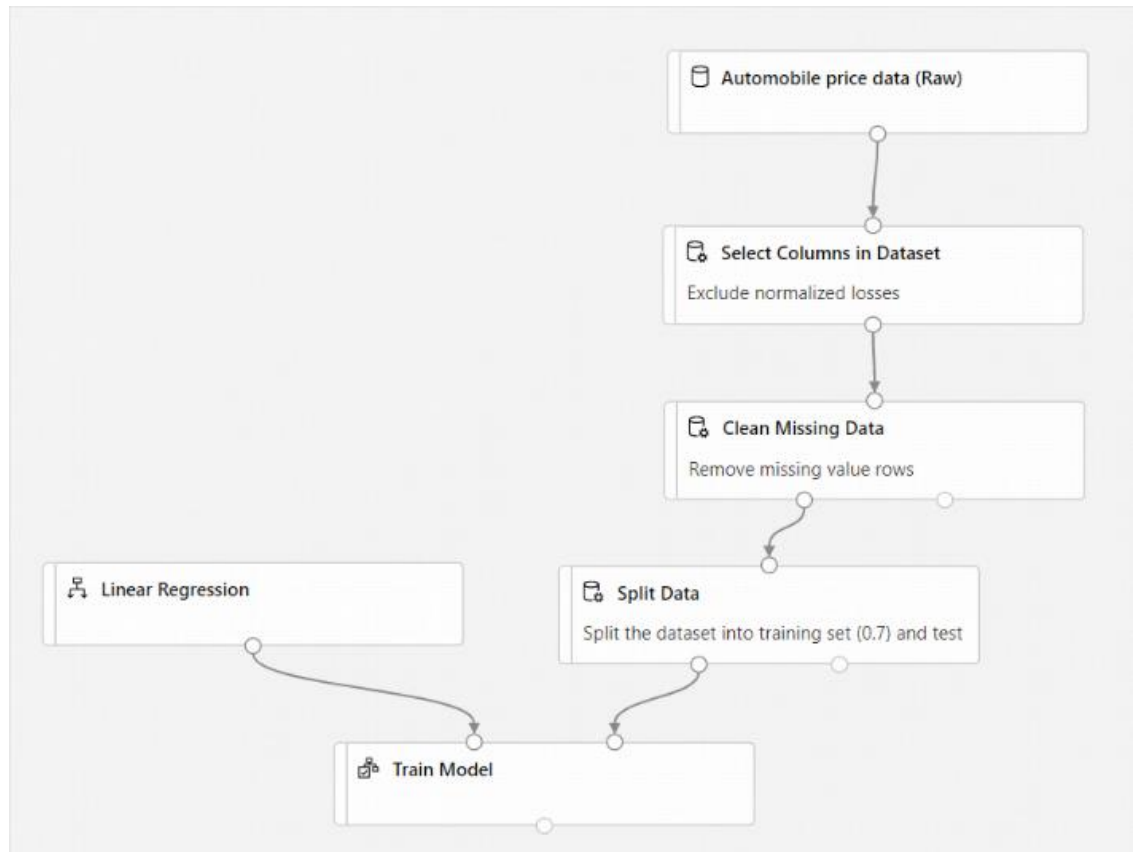
6. Seleccione el componente **Modelo de Entrenamiento**.
7. En el panel de detalles del componente a la derecha del lienzo, seleccione **Editar** selector de columna.
8. En el cuadro de diálogo **Columna de etiqueta**, expanda el menú desplegable y seleccione **Nombres de columna**.
9. En el cuadro de 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** .

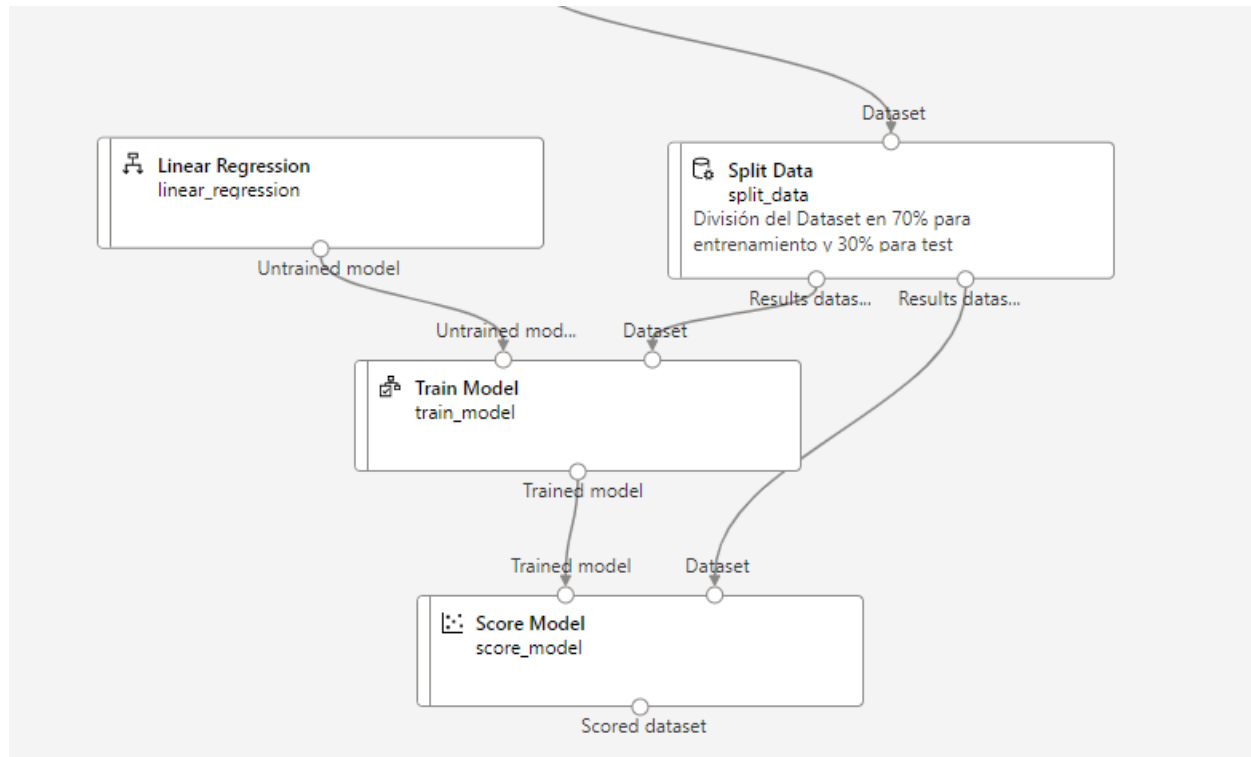
Su tubería debería verse así:



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** .

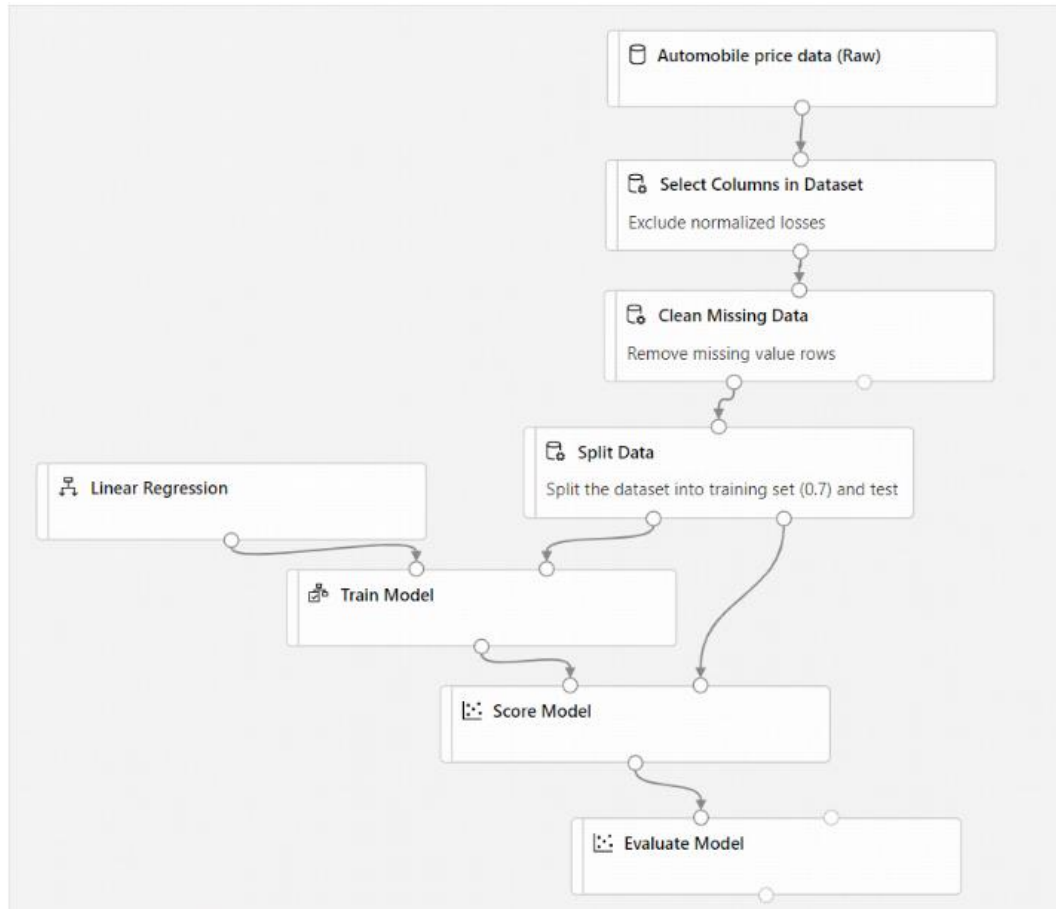


Agregar el componente Evaluar modelo

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** .

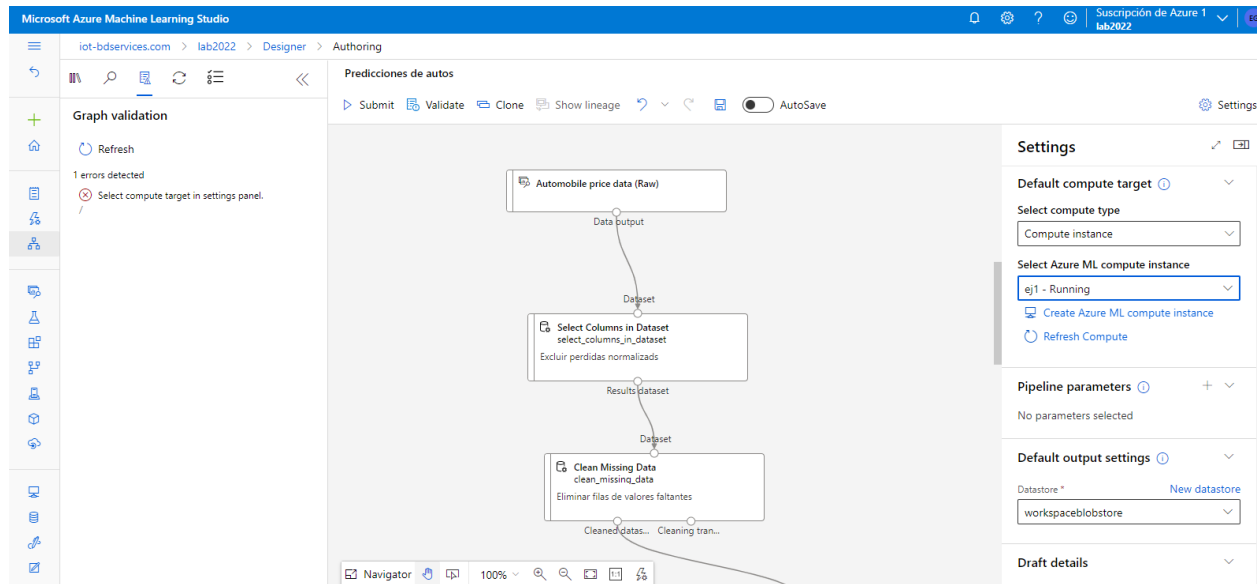
La tubería final debería verse así:



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** .

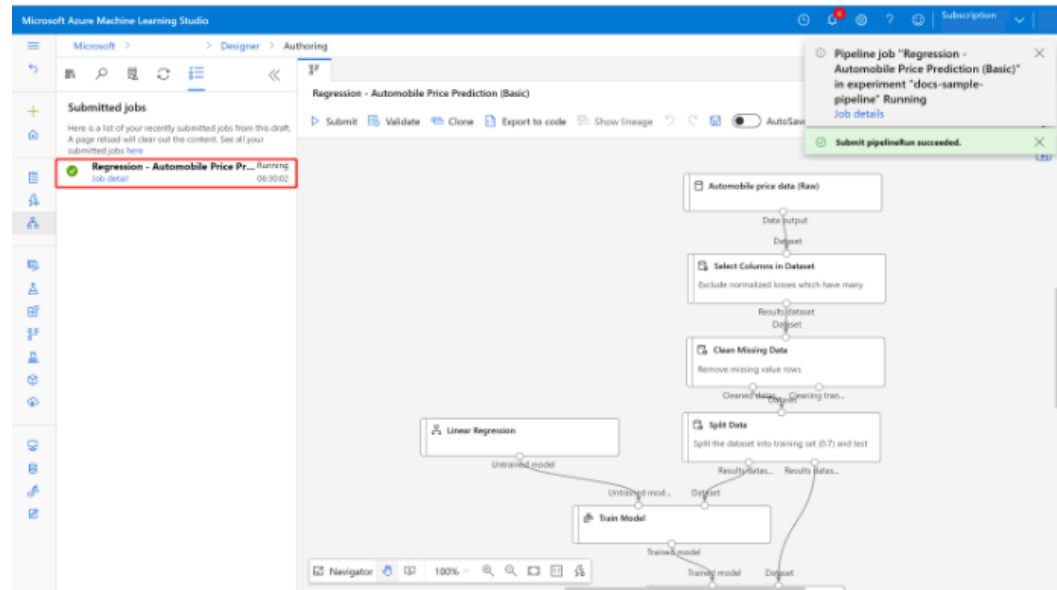
Nota

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** .

The screenshot shows the "Set up pipeline job" dialog box. It has a title bar with a close button. The "Experiment" section has two radio buttons: "Select existing" and "Create new", with "Create new" selected. Below this is a text field for "New experiment name" with the value "tutorial2". There are also fields for "Job display name" (containing "Predicciones de autos") and "Job description" (containing "Pipeline created on 20220707"). The "Job tags" section has a text field followed by a colon and another text field, with a plus icon and a trash icon to the right. At the bottom, there is a checkbox labeled "Continue on failure step" which is checked. The dialog has "Submit" and "Cancel" buttons at the bottom right.

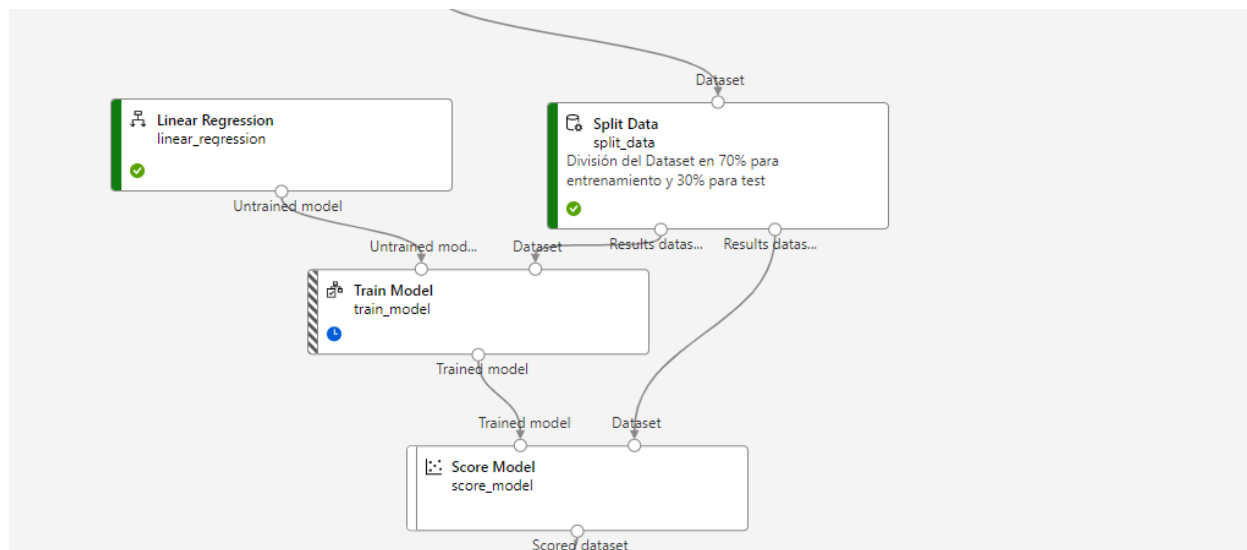
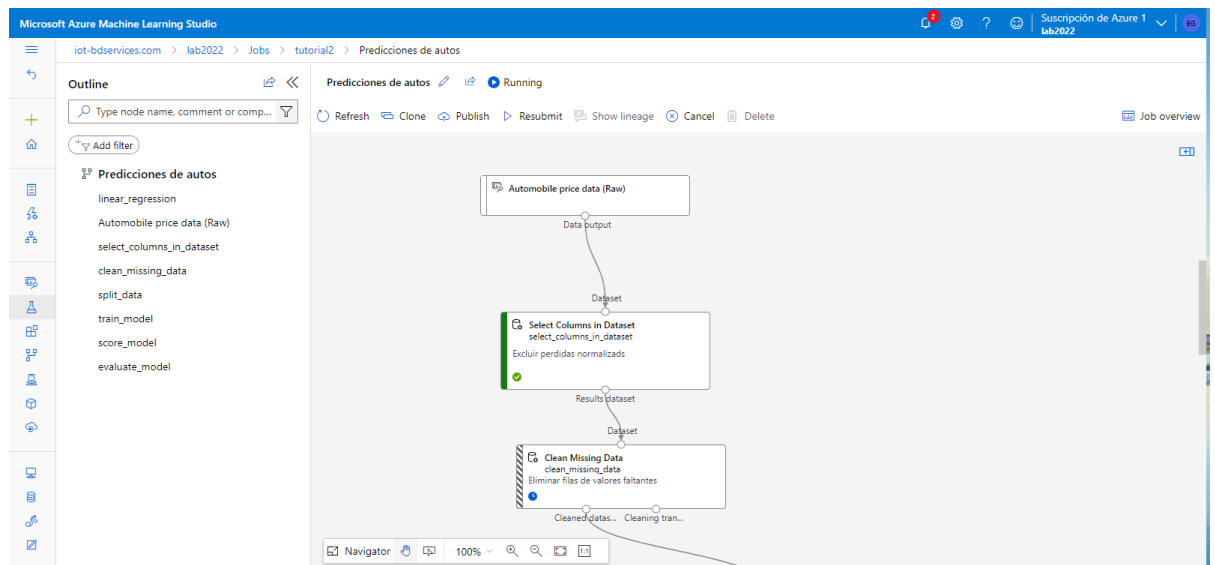
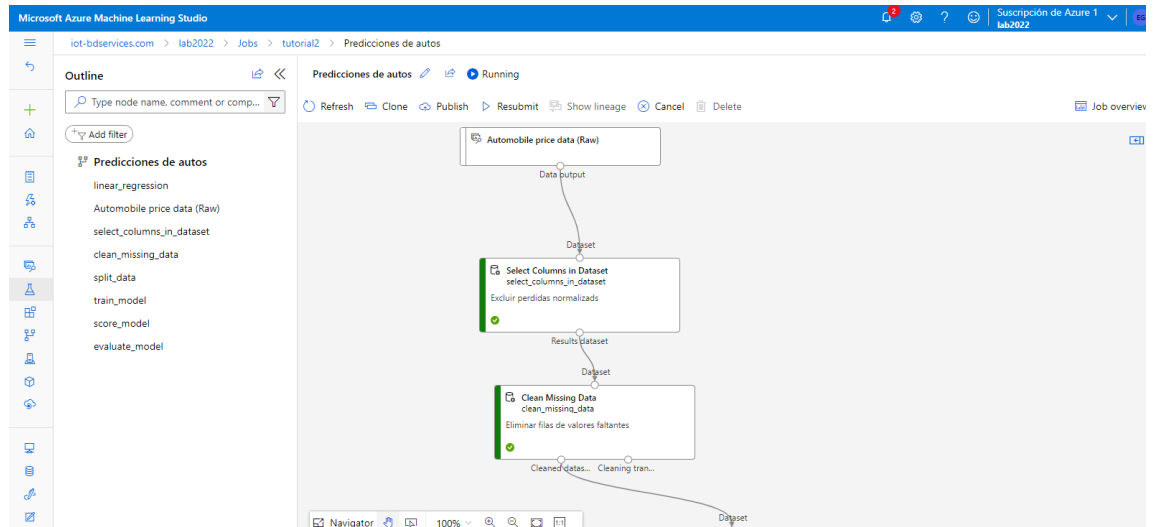
1. Seleccione **Enviar** .
2. Verá una lista de envíos en el panel izquierdo del lienzo y aparecerá una notificación en la esquina superior derecha de la página. Puede seleccionar el enlace Detalles del **trabajo** para ir a la página de detalles del trabajo para la depuración.



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)

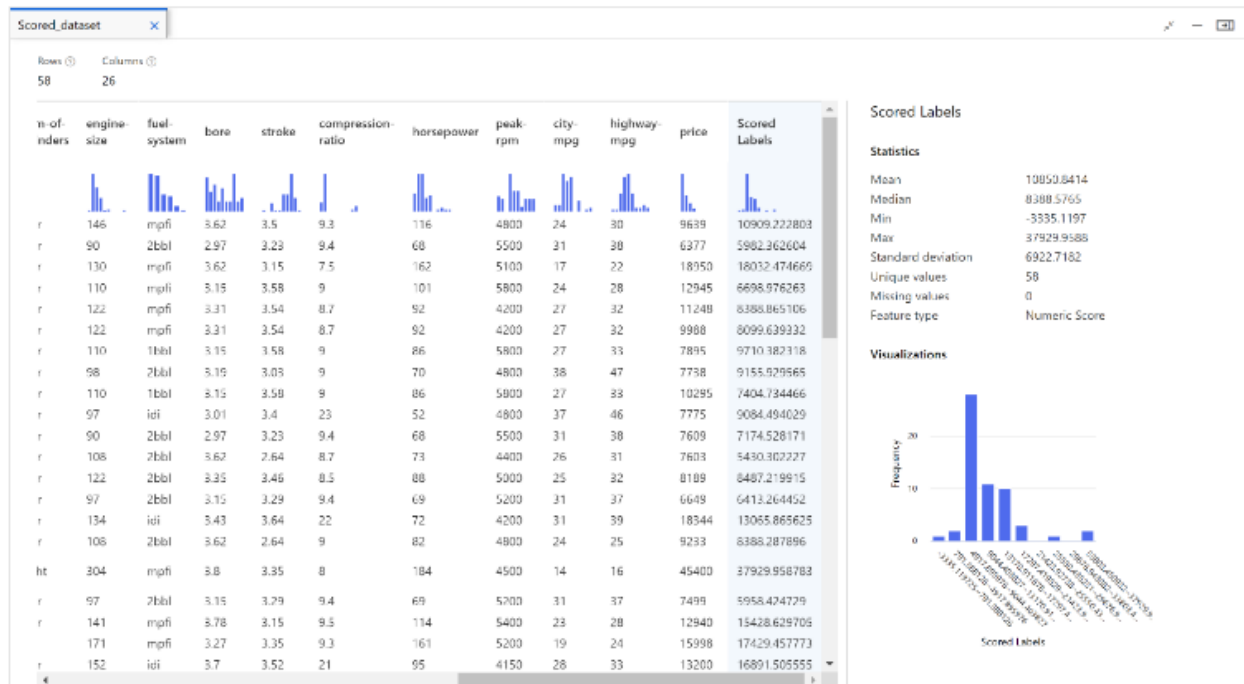
Materia: Minería de Datos

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



Ver etiquetas puntuadas

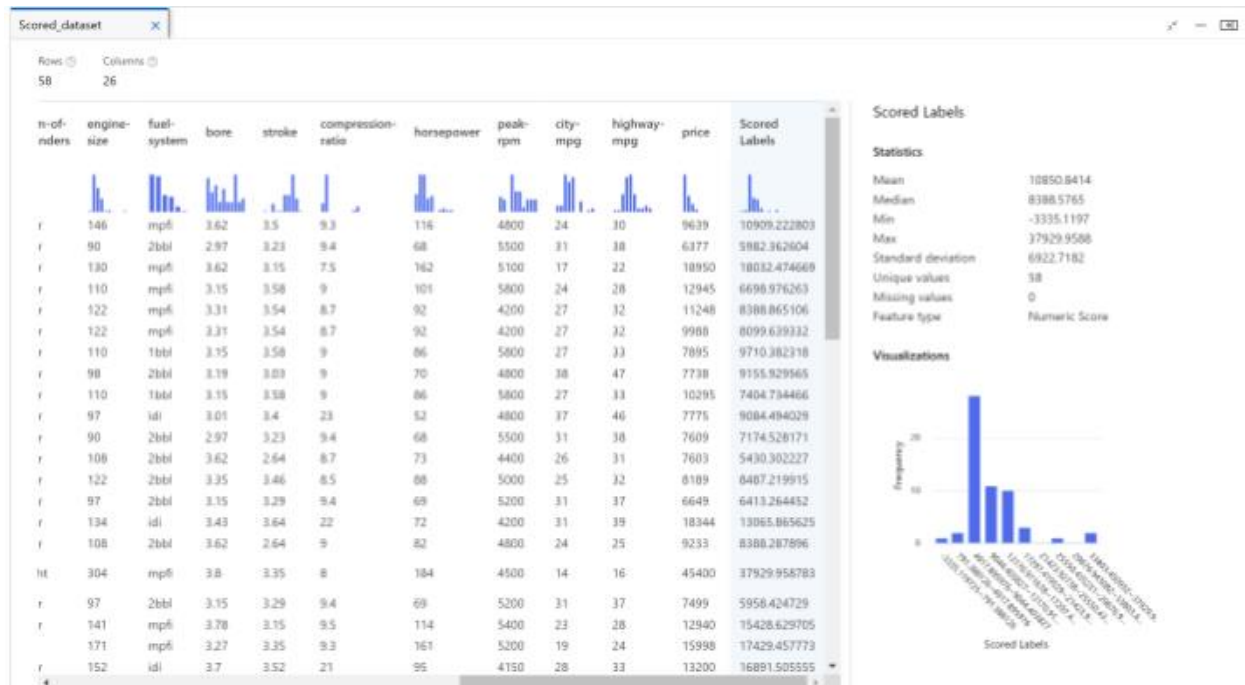
En la página de detalles del trabajo, puede verificar el estado, los resultados y los registros del trabajo de canalización.



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.

Materia: Minería de Datos

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

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](#).

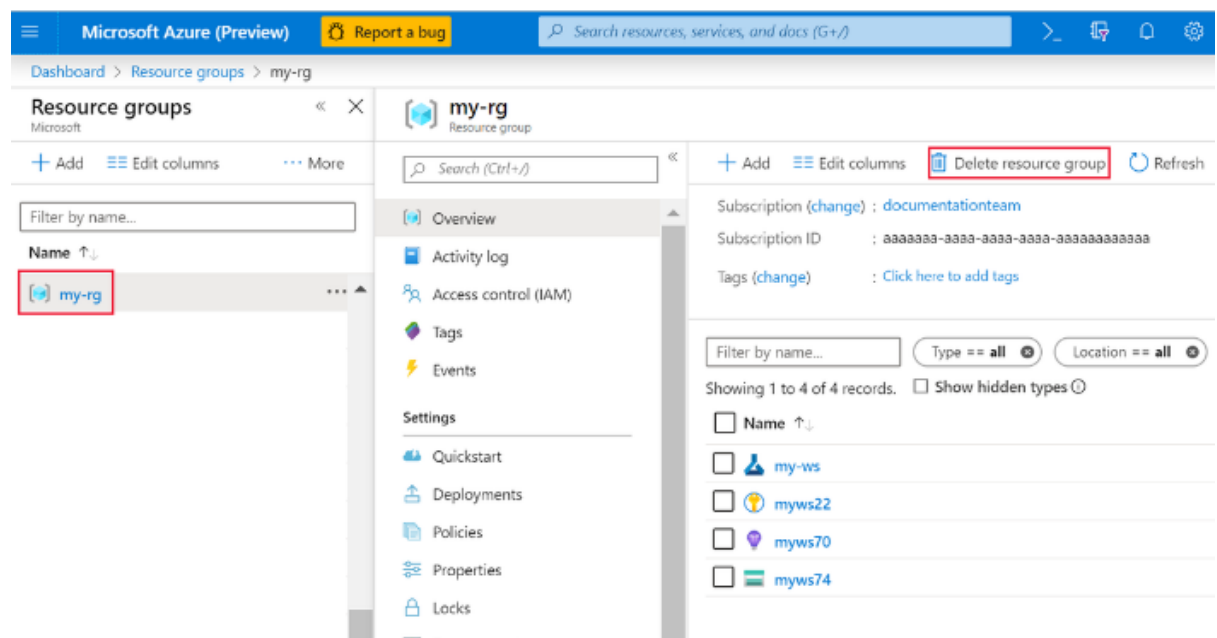
Importante

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.



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

Materia: Minería de Datos

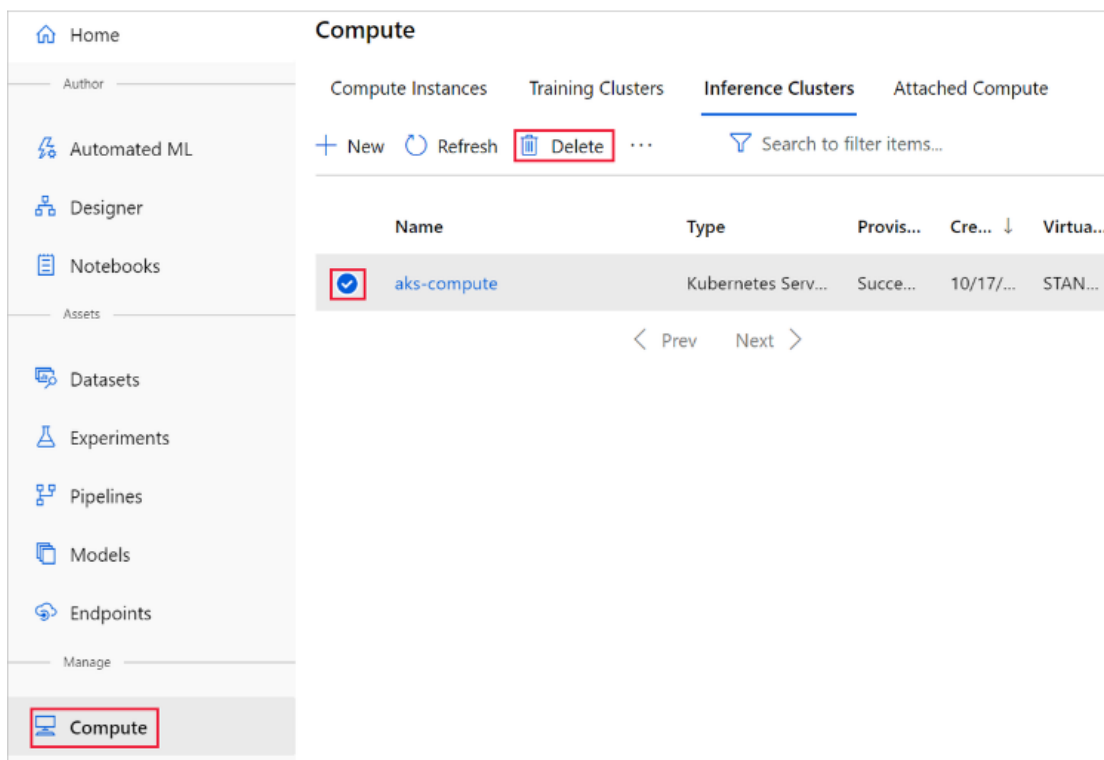
Taller: Entrenar un modelo de regresión sin código usando Azure Machine Learning

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:



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 portal interface. On the left, the 'Datasets' tab is selected in the sidebar. The main panel displays details for a dataset named 'TD-Sample_1: Regression_-_Automobile_Price_Prediction_(Basic)-Clean_Missing_Data-Cleaning_transformation-f6dc0eb1'. The 'Unregister' button is highlighted in the top navigation bar. The dataset details include:

- Attributes:**
 - Properties:** File
 - Description:** This is a dataset promoted by inference graph generation automatically on 11/12/...
 - Datasource:** workspaceblobstore
 - Relative path:** azureml/4393076b-19ff-4e41-81d9-a1146d905696/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 = 'ec85ed72-2b26-48f6-abe8-4b5bc498fbd9'
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>