

## Cree un modelo de aprendizaje automático en Power BI

### Crear un flujo de datos con los datos de entrada

Cree un flujo de datos con datos de entrada siguiendo estos pasos.

#### Obtener datos

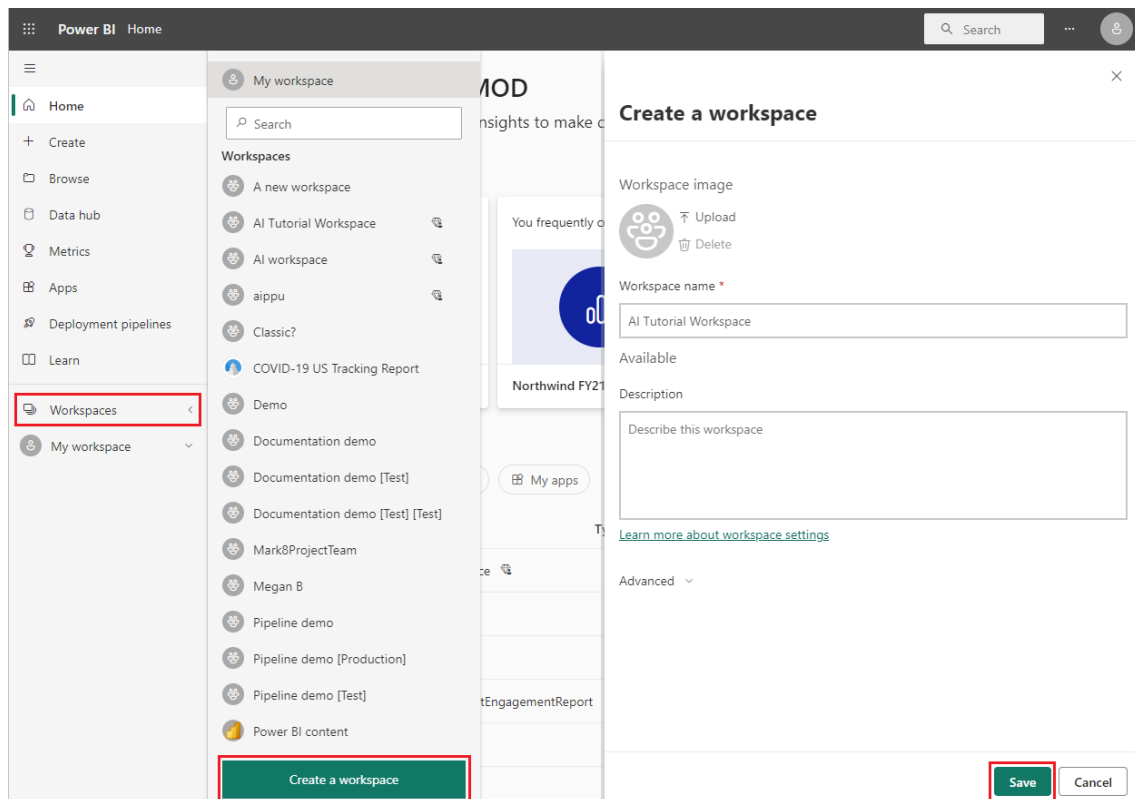
El primer paso para crear un flujo de datos es tener listas las fuentes de datos. En este caso, se utiliza un modelo semántico de aprendizaje automático a partir de un conjunto de sesiones en línea, algunas de las cuales culminaron en una compra. El modelo semántico contiene un conjunto de atributos sobre estas sesiones, que se utilizan para entrenar el modelo.

Puede descargar el modelo semántico desde el sitio web de UC Irvine o descargando [online shoppers intention.csv](#) . Más adelante en este tutorial, se conectará al modelo semántico especificando su URL.

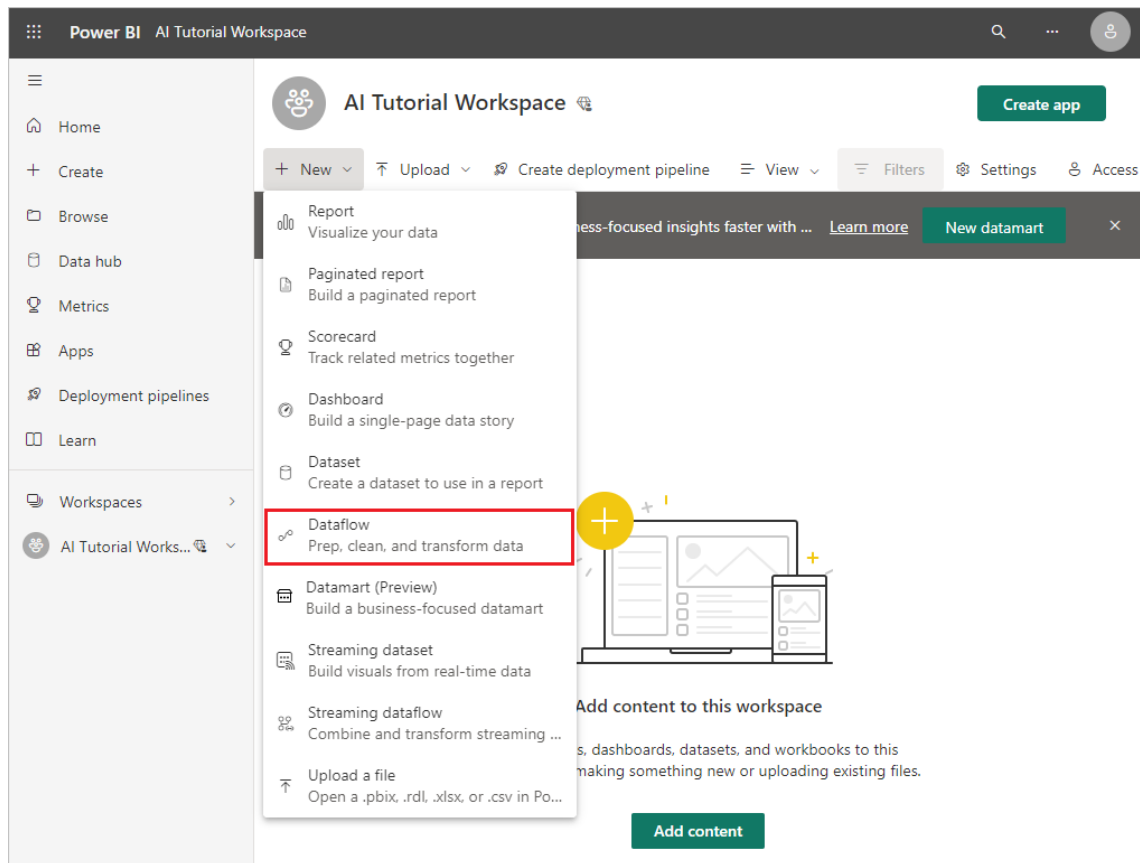
#### Crear las tablas

Para crear las entidades en su flujo de datos, inicie sesión en el servicio Power BI y navegue a un espacio de trabajo.

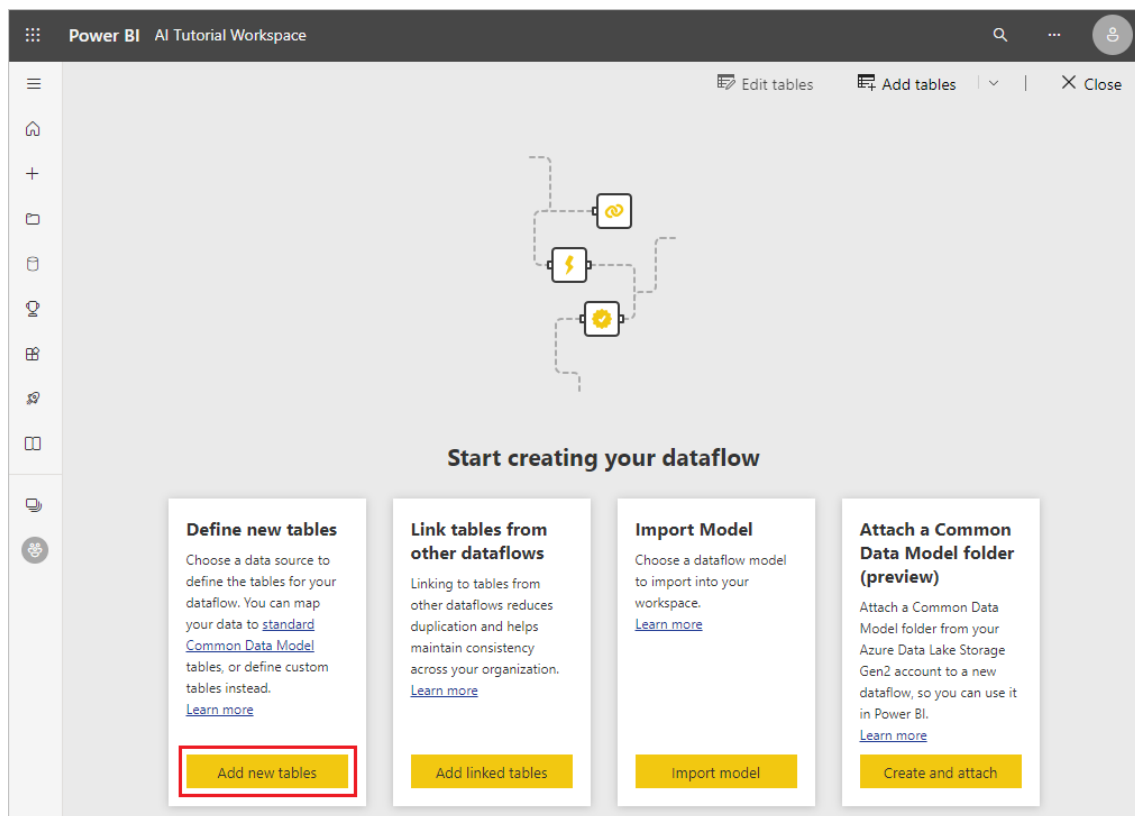
1. Si no tiene un espacio de trabajo, cree uno seleccionando **Espacios de trabajo** en el panel de navegación izquierdo de Power BI y seleccionando **Crear un espacio de trabajo** . En el panel **Crear un espacio de trabajo** , ingrese un nombre para el espacio de trabajo y seleccione **Guardar** .



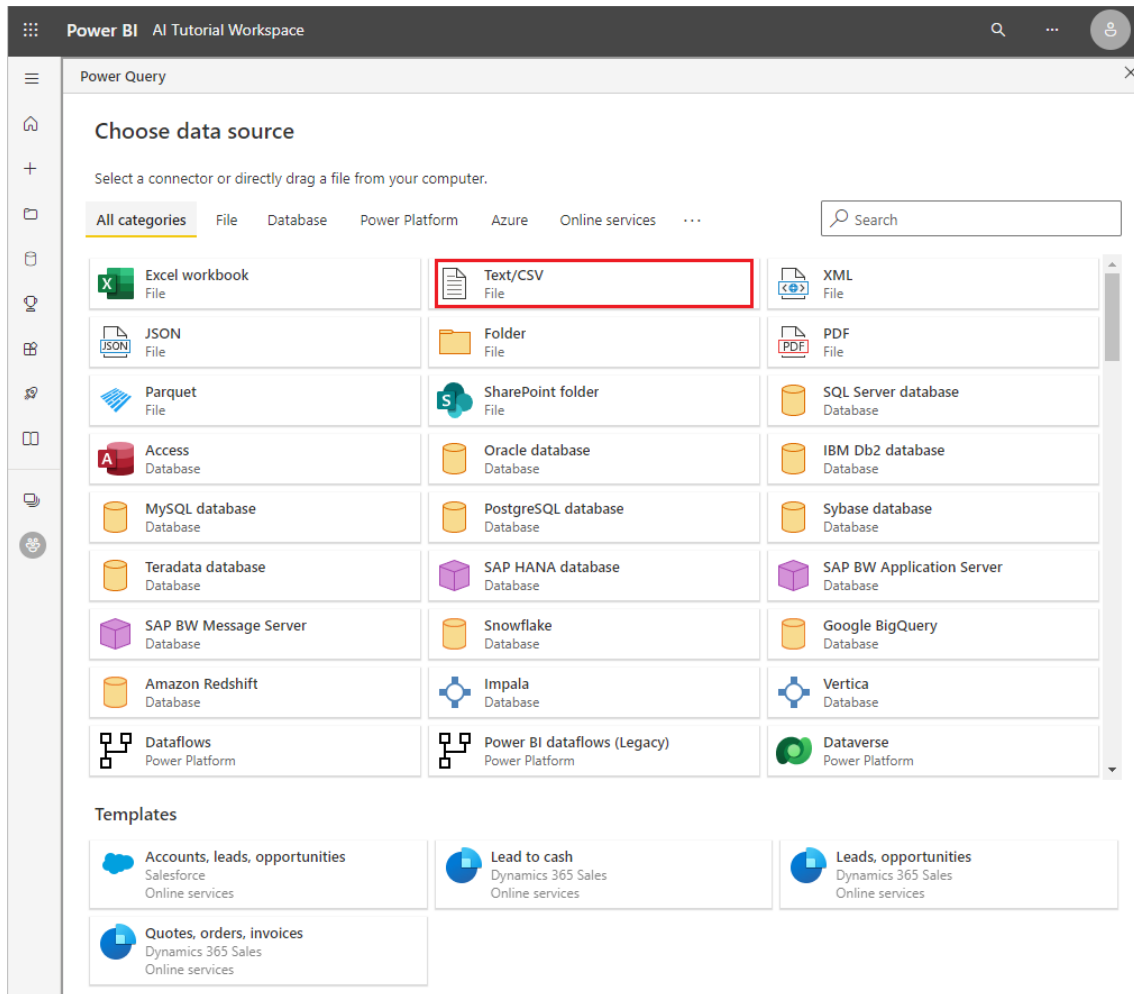
2. Seleccione **Nuevo** en la parte superior del nuevo espacio de trabajo y luego seleccione **Flujo de datos** .



3. Seleccione **Agregar nuevas tablas** para iniciar un editor de **Power Query** en el navegador.



4. En la pantalla **Elegir fuente de datos** , seleccione **Texto/CSV** como fuente de datos.



5. En la página **Conectarse a una fuente de datos** , pegue el siguiente enlace al archivo *online\_shoppers\_intention.csv* en el cuadro **Ruta de archivo o URL** y luego seleccione **Siguiente** .

[https://raw.githubusercontent.com/santoshc1/PowerBI-AI-samples/master/Tutorial\\_AutomatedML/online\\_shoppers\\_intention.csv](https://raw.githubusercontent.com/santoshc1/PowerBI-AI-samples/master/Tutorial_AutomatedML/online_shoppers_intention.csv)

Power BI AI Tutorial Workspace

### Power Query

#### Connect to data source

Text/CSV  
File  
[Learn more](#)

**Connection settings**

☒ Link to file ☐ Upload file (Preview) ⓘ

**File path or URL \***  
 [Browse OneDrive...](#)

**Connection credentials**

Connection  
 ↻

Connection name

Data gateway  
 ↻

Authentication kind

Back Cancel Next

- El editor de Power Query muestra una vista previa de los datos del archivo CSV. Para realizar cambios en los datos antes de cargarlos, seleccione **Transformar datos**.

Power BI AI Tutorial Workspace

Search

### Power Query

#### Preview file data

URL: [https://raw.githubusercontent.com/santoshc1/PowerBI-AI-samples/master/Tutorial\\_AutomatedML/online\\_shoppers\\_intention.csv](https://raw.githubusercontent.com/santoshc1/PowerBI-AI-samples/master/Tutorial_AutomatedML/online_shoppers_intention.csv)

File origin:  Delimiter:  Data type detection:  ↻

1 <sup>2</sup> Administrative	1,2 Administrative_Duration	1 <sup>2</sup> Informational	1,2 Informational_Duration	1 <sup>2</sup> ProductRelated	1,2 ProductRelated_Duration	1,2 BounceRates	1,2 ExitRates	1,2 P
1	0	2	211.25	144	4627.489571	0.001360544	0.020664031	
7	150.3571429	1	9	221	11431.00124	0.011148992	0.021904109	
3	16	3	86	15	2773.5	0	0.03	
0	0	0	0	7	705.8333333	0.028571429	0.085714286	
0	0	0	0	7	212.5	0	0.033333333	
0	0	0	0	44	615	0.013636364	0.036363636	
3	100.5	0	0	95	1453.640882	0.002105263	0.012761333	
0	0	0	0	6	279	0	0.033333333	
0	0	0	0	21	1128.583333	0	0.013043478	
0	0	0	0	8	143.5833333	0.014285714	0.05	
7	191.2	4	654.3333333	31	1204.600433	0.000854701	0.013342491	
2	47	0	0	39	1743.366667	0	0.016410256	
0	0	0	0	10	236	0	0.026666667	
4	724	0	0	2	159	0	0.008	
0	0	0	0	2	24	0	0.1	
8	237.5	4	142	30	1569.625	0	0.008108108	
3	66.33333333	0	0	15	424.8333333	0	0.0125	
3	227	0	0	2	62	0	0.025	
1	23	0	0	15	656.25	0.028571429	0.028571429	
0	0	0	0	29	617.5	0.006896552	0.01954023	
0	0	2	39	15	245	0	0.0125	
0	0	0	0	3	20	0	0.066666667	
0	0	0	0	38	1372	0.005405405	0.018918919	
2	41	0	0	7	139	0.022222222	0.055555556	
4	74	2	629	149	5042.458059	0.003870968	0.007285123	

Back Cancel Add table using examples Transform data

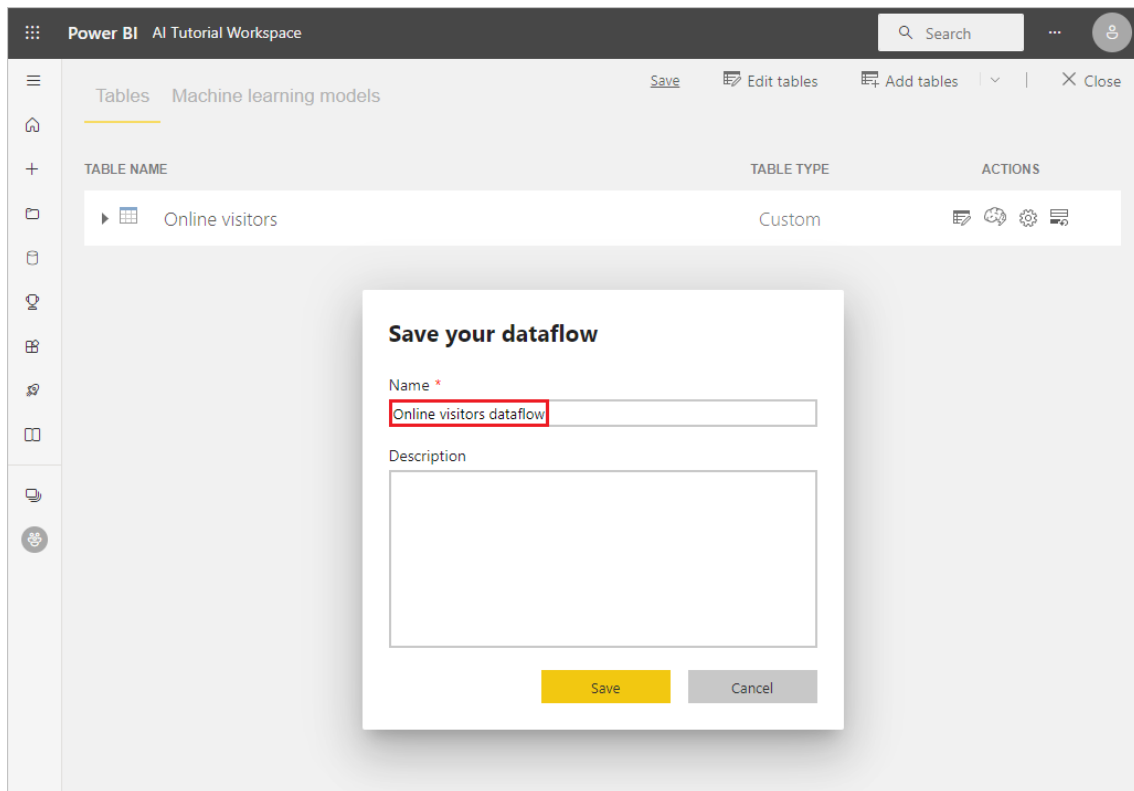
7. Power Query infiere automáticamente los tipos de datos de las columnas. Puede cambiar los tipos de datos seleccionando el icono de tipo de atributo en la parte superior de los encabezados de columna. Cambie el tipo de la columna **Ingresos** a **Verdadero/Falso** .

Puede cambiar el nombre de la consulta a un nombre más fácil de usar modificando el valor en el cuadro **Nombre** en el panel derecho. Cambie el nombre de la consulta a *Visitantes en línea* .

Power Query Editor interface showing the 'Revenue' column header with a data type icon highlighted. The dropdown menu for the icon is open, showing various data types. The 'True/False' option is highlighted. The 'Query settings' pane on the right shows the 'Name' field set to 'Online visitors'.

Visitor	Region	TrafficType	VisitorType	Weekend	Revenue
1	2	1	2 Return	1.2	Decimal number
2	5	1	2 Return	\$	Currency
3	2	1	2 Return	123	Whole number
4	11	4	1 Return	%	Percentage
5	2	1	11 Return		
6	1	3	1 Return		
7	2	3	2 New		
8	2	3	10 Return		
9	2	1	2 Return		
10	2	3	1 Return		
11	2	7	2 Return		
12	3	2	1 Return		
13	2	1	1 Return		
14	2	1	7 New		
15	6	1	2 Return		
16	2	1	9 Returning_Visitor	TRUE	NULL
17	2	6	6 Returning_Visitor	FALSE	NULL
18	2	1	10 Returning_Visitor	TRUE	NULL
19	1	8	2 Returning_Visitor	TRUE	NULL
20	2	3	10 Returning_Visitor	FALSE	NULL
21	2	1	3 Returning_Visitor	TRUE	NULL
22	2	2	1 Returning_Visitor	TRUE	NULL
23	2	1	2 Returning_Visitor	FALSE	NULL
24	5	6	2 Returning_Visitor	TRUE	NULL
25	-	-	-	-	-

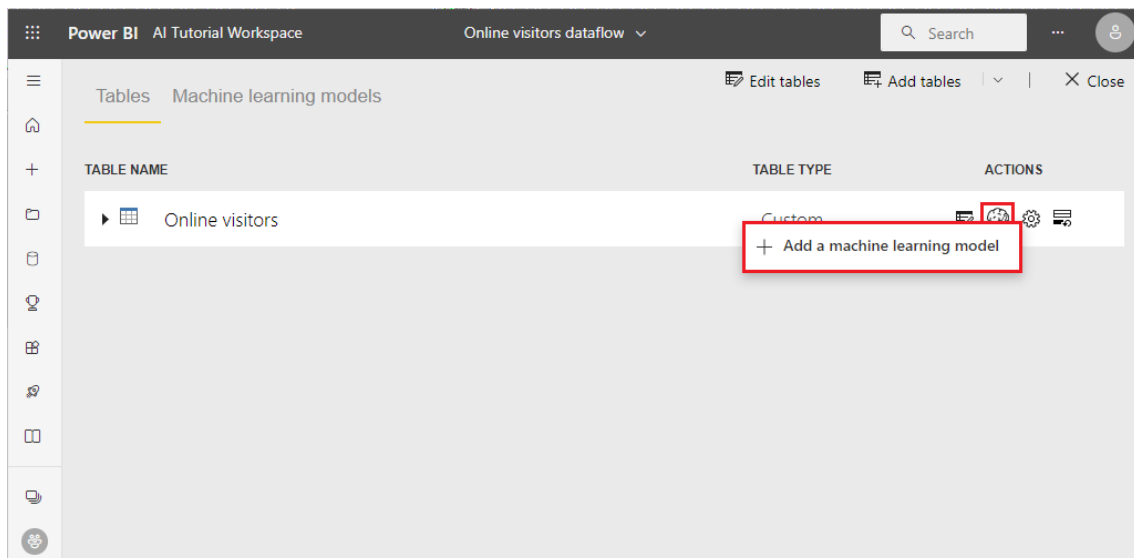
8. Seleccione **Guardar y cerrar** y, en el cuadro de diálogo, proporcione un nombre para el flujo de datos y luego seleccione **Guardar** .



## Crear y entrenar un modelo de aprendizaje automático

Para agregar un modelo de aprendizaje automático:

1. Seleccione el ícono **Aplicar modelo de aprendizaje automático en la lista Acciones** de la tabla que contiene sus datos de entrenamiento y la información de la etiqueta y, luego, seleccione **Agregar un modelo de aprendizaje automático**.



2. El primer paso para crear el modelo de aprendizaje automático es identificar los datos históricos, incluido el campo de resultado que desea predecir. El modelo se crea aprendiendo de estos datos. En este caso, desea predecir si los visitantes realizarán una compra o no. El resultado que desea predecir se encuentra en

el campo **Ingresos** . **Seleccione Ingresos** como el valor de la **columna Resultado** y, luego, seleccione **Siguiente** .

The screenshot shows the Power BI AI Tutorial Workspace interface. At the top, there's a header with 'Power BI AI Tutorial Workspace', 'Online visitors dataflow', a search bar, and a user profile icon. Below the header is a progress bar with four steps: 'Select a column to predict' (active, marked with a checkmark), 'Choose a model', 'Select data to study', and 'Name and train'. The main content area is titled 'What do you want to predict?' and includes the instruction 'Select the table and the outcome column you'd like to make predictions about so we can recommend the best model.' There are two dropdown menus: 'Table' with 'Online visitors' selected, and 'Outcome column' with 'Revenue' selected. The 'Revenue' option in the 'Outcome column' dropdown is highlighted with a red rectangle. At the bottom right, there are 'Next' and 'Cancel' buttons.

3. A continuación, seleccione el tipo de modelo de aprendizaje automático que desea crear. Power BI analiza los valores del campo de resultados que identificó y sugiere los tipos de modelos de aprendizaje automático que puede crear para predecir ese campo.

En este caso, dado que desea predecir un resultado binario de si un visitante va a realizar una compra o no, Power BI recomienda **la predicción binaria** . Dado que está interesado en predecir los visitantes que van a realizar una compra, seleccione **verdadero** en **Elegir un resultado objetivo** . También puede proporcionar diferentes etiquetas para usar en los resultados del informe generado automáticamente que resume los resultados de la validación del modelo. Luego, seleccione **Siguiente** .

Power BI AI Tutorial Workspace Online visitors dataflow

Select a column to predict Choose a model Select data to study Name and train

### Choose a model

Based on the column you selected, we recommend a **Prediction** model. This model learns from your data to predict whether or not an outcome will be achieved. Not what you're looking for? [Select a different model](#)

**Binary Prediction**

Predict whether or not an outcome will be achieved.

#### Choose a target outcome

Enter or select the Revenue outcome that you're most interested in.

#### How should we label predictions in the model training report?

**Match label**

Enter the text you want to display when our prediction matches your target value.

**Mismatch label**

Enter the text you want to display when our prediction doesn't match your target value.

Back Next Cancel

- Power BI realiza un análisis preliminar de una muestra de sus datos y sugiere entradas que podrían generar predicciones más precisas. Si Power BI no recomienda una columna, explica por qué no lo hace junto a ella. Puede cambiar las selecciones para incluir solo los campos que desea que el modelo estudie seleccionando o desmarcando las casillas de verificación junto a los nombres de las columnas. Seleccione **Siguiente** para aceptar las entradas.



Power BI AI Tutorial Workspace Online visitors dataflow

Select a column to predict Choose a model **Select data to study** Name and train

### Select the data your model should study

Based on a sample of your data, we've selected columns that may produce more accurate outcomes. If we don't recommend a column, we've explained why next to it. Change your selections to include only the columns you want the model to study.

Search

[Reset](#) [Clear](#)  
12 columns selected

- ☐ Administrative (low correlation with Revenue)
- ☒ Administrative\_Duration
- ☒ Informational
- ☐ Informational\_Duration (low correlation with Revenue)
- ☐ ProductRelated (low correlation with Revenue)
- ☒ ProductRelated\_Duration
- ☒ BounceRates
- ☒ ExitRates
- ☒ PageValues
- ☒ SpecialDay
- ☒ Month
- ☒ OperatingSystems
- ☐ Browser (low correlation with Revenue)
- ☐ Region (low correlation with Revenue)
- ☒ TrafficType
- ☐ VisitorType (low correlation with Revenue)
- ☒ Weekend
- ☒ Revenue (Outcome column)

Back Next Cancel

- En el paso final, nómbralo como *predicción de intención de compra* y elige el tiempo que deseas dedicar al entrenamiento. Puedes reducir el tiempo de entrenamiento para ver resultados rápidos o aumentarlo para obtener el mejor modelo. Luego, selecciona **Guardar y entrenar** para comenzar a entrenar el modelo.

**Power BI AI Tutorial Workspace** Online visitors dataflow

Select a column to predict Choose a model Select data to study **Name and train**

**Name and train your model**

Model name  
Purchase intent prediction

Description  
(Optional)

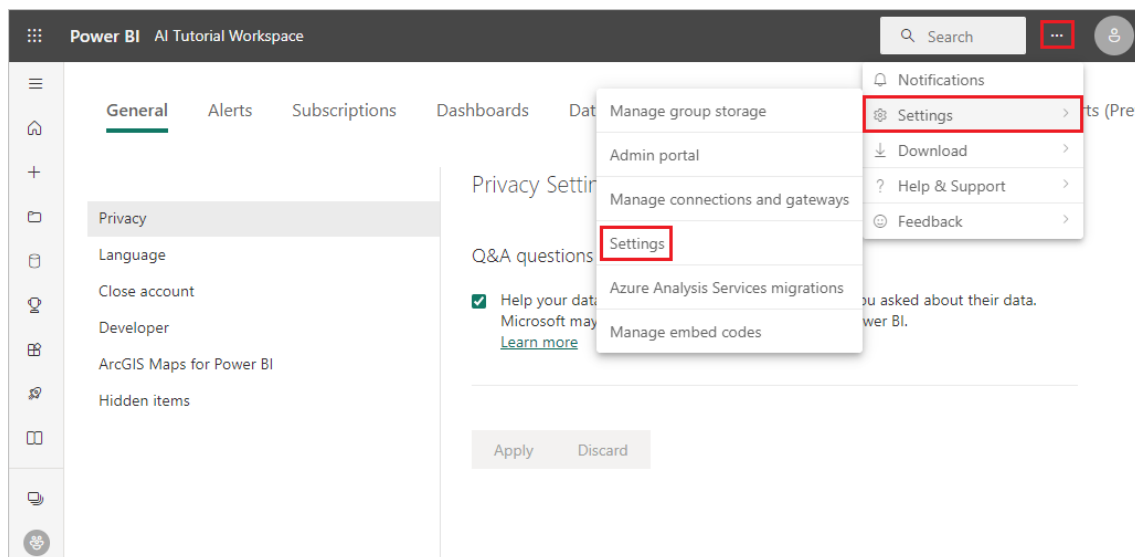
**Training time**  
The longer you train your model, the more accurate the results. Train for a short time if you just want to make sure you've selected the right data. Keep in mind, this won't result in the best model.

5 minutes 360 minutes | 61 minutes

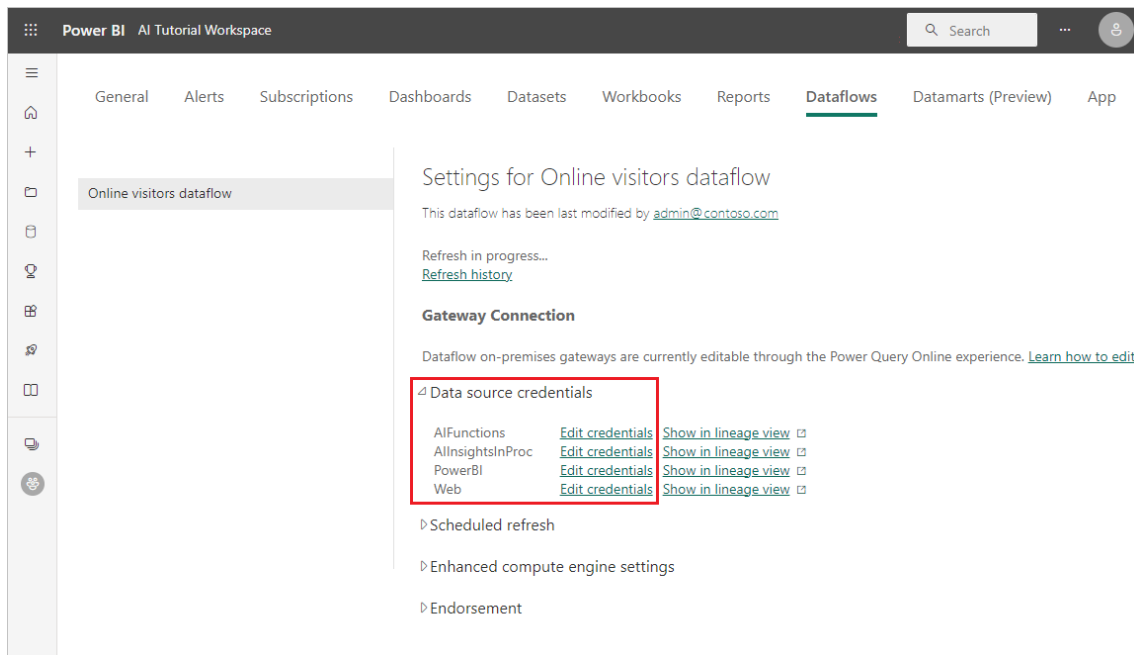
**What happens next?**  
We'll take a statistically significant sample of your data and train the model using 80% of it. We'll then test the model on the remaining 20% and go over the Prediction accuracy in a report. You can find the training and test data we used in your workspace.

Back Save **Save and train** Cancel

Si recibe un error similar a **No se encontraron credenciales para la fuente de datos** , debe actualizar sus credenciales para que Power BI pueda puntuar los datos. Para actualizar sus credenciales, seleccione **Más opciones...** en la barra de encabezado y luego seleccione **Configuración > Configuración** .



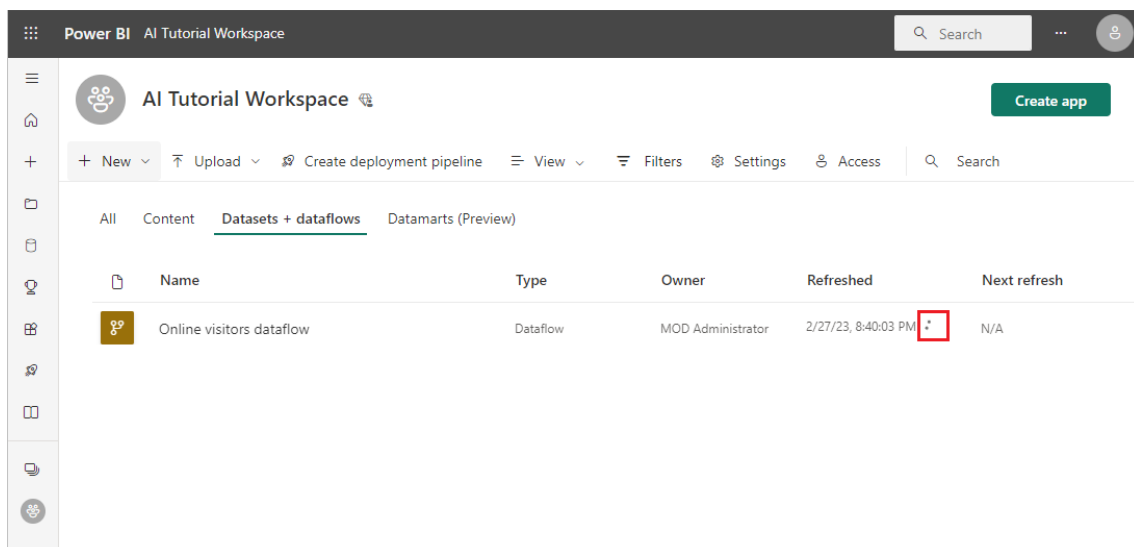
Seleccione su flujo de datos en **Flujos de datos** , expanda **Credenciales de fuente de datos** y luego seleccione **Editar credenciales** .



## Seguimiento del estado del entrenamiento

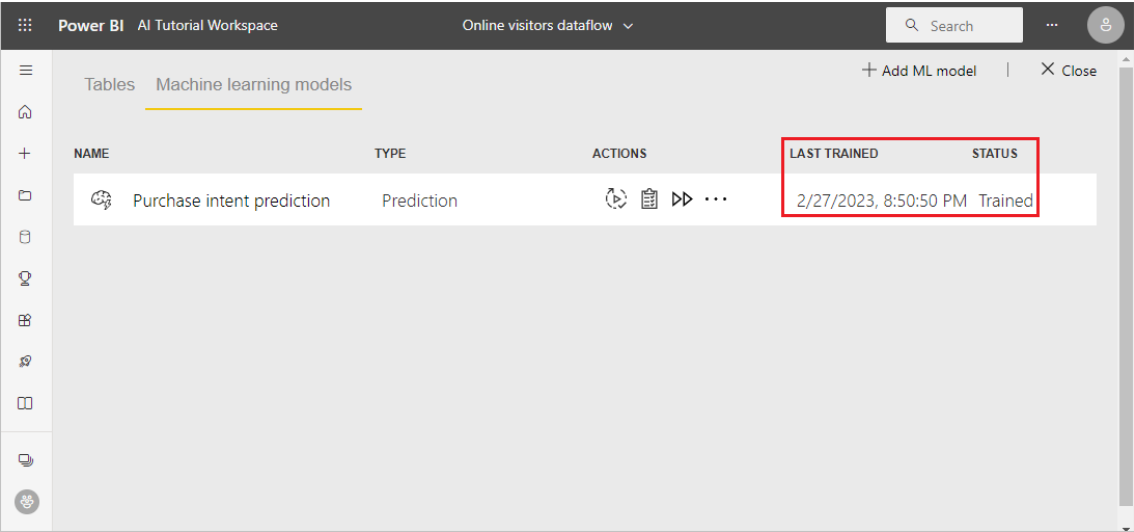
El proceso de entrenamiento comienza con el muestreo y la normalización de sus datos históricos y la división de su modelo semántico en dos nuevas entidades: **Datos de entrenamiento de predicción de intención de compra** y **Datos de prueba de predicción de intención de compra**.

Según el tamaño del modelo semántico, el proceso de entrenamiento puede tardar desde unos pocos minutos hasta el tiempo de entrenamiento que haya seleccionado. Puede confirmar que el modelo se está entrenando y validando a través del estado del flujo de datos. El estado aparece como una actualización de datos en curso en la pestaña **Modelos semánticos + flujos de datos** del espacio de trabajo.



Puede ver el modelo en la pestaña **Modelos de aprendizaje automático** del flujo de datos. **El estado** indica si el modelo se ha puesto en cola para entrenamiento, se encuentra en proceso de entrenamiento o ya está entrenado. Una vez que se completa el

entrenamiento del modelo, el flujo de datos muestra una hora **de último entrenamiento** actualizada y un estado de **Entrenado** .

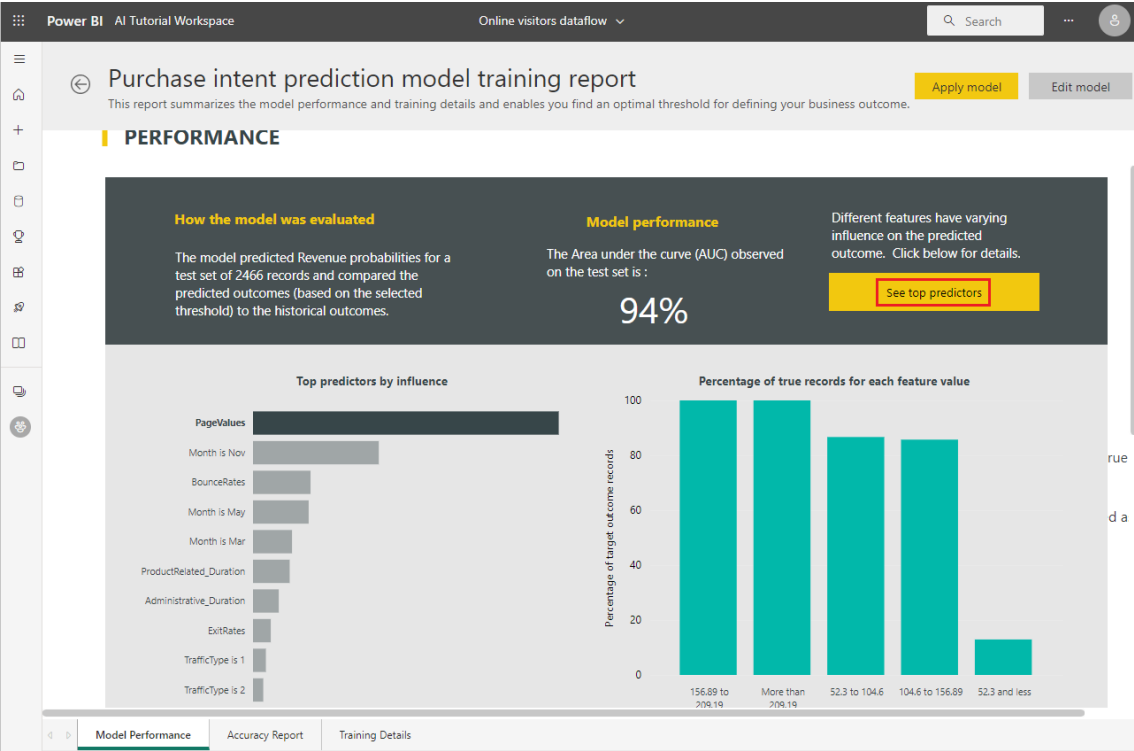


Power BI AI Tutorial Workspace Online visitors dataflow				
Machine learning models				
NAME	TYPE	ACTIONS	LAST TRAINED	STATUS
Purchase intent prediction	Prediction		2/27/2023, 8:50:50 PM	Trained

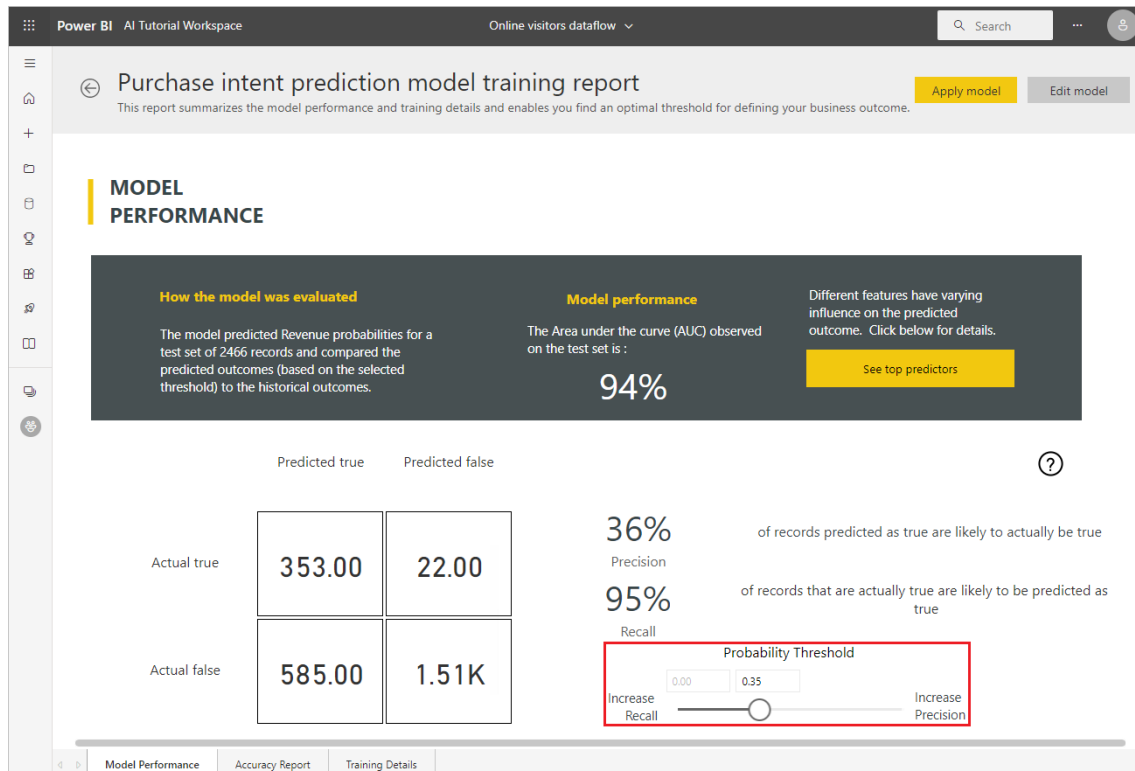
**Revisar el informe de validación del modelo**

Para revisar el informe de validación del modelo, en la pestaña **Modelos de aprendizaje automático** , seleccione el ícono **Ver informe de entrenamiento en Acciones** . Este informe describe cómo es probable que funcione su modelo de aprendizaje automático.

En la página **Rendimiento del modelo** del informe, seleccione **Ver predictores principales** para ver los predictores principales de su modelo. Puede seleccionar uno de los predictores para ver cómo se asocia la distribución de resultados con ese predictor.



Puede utilizar la segmentación de **datos Umbral de probabilidad** en la página **Rendimiento del modelo** para examinar la influencia de la **Precisión** y la **Recuperación** del modelo en el modelo.

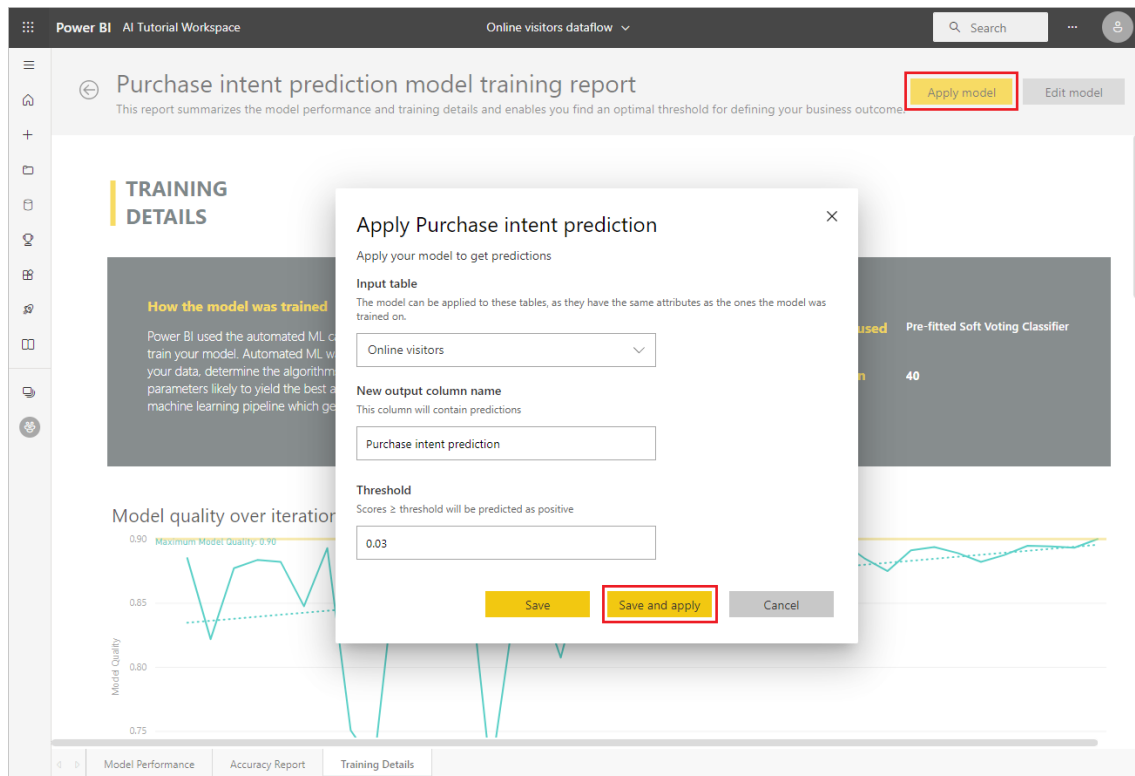


Las otras páginas del informe describen las métricas de rendimiento estadístico del modelo.

El informe también incluye una página **de Detalles de entrenamiento** que describe las **iteraciones ejecutadas**, cómo se extrajeron las características de las entradas y los hiperparámetros para el **modelo final utilizado**.

### Aplicar el modelo a una entidad de flujo de datos

Seleccione el botón **Aplicar modelo** en la parte superior del informe para invocar este modelo. En el cuadro de diálogo **Aplicar**, puede especificar la entidad de destino que tiene los datos de origen a los que se aplicará el modelo. Luego, seleccione **Guardar y aplicar**.



Al aplicar el modelo se crean dos tablas nuevas, con los sufijos **enriched** **<model\_name>** y **enriched <model\_name> explains** . En este caso, al aplicar el modelo a la tabla **Visitantes en línea** se crea lo siguiente:

- **Los visitantes en línea enriquecieron la predicción de la intención de compra** , que incluye el resultado previsto del modelo.
- **Los visitantes en línea enriquecieron las explicaciones de predicción de intención de compra** , que contienen los principales influyentes específicos del registro para la predicción.

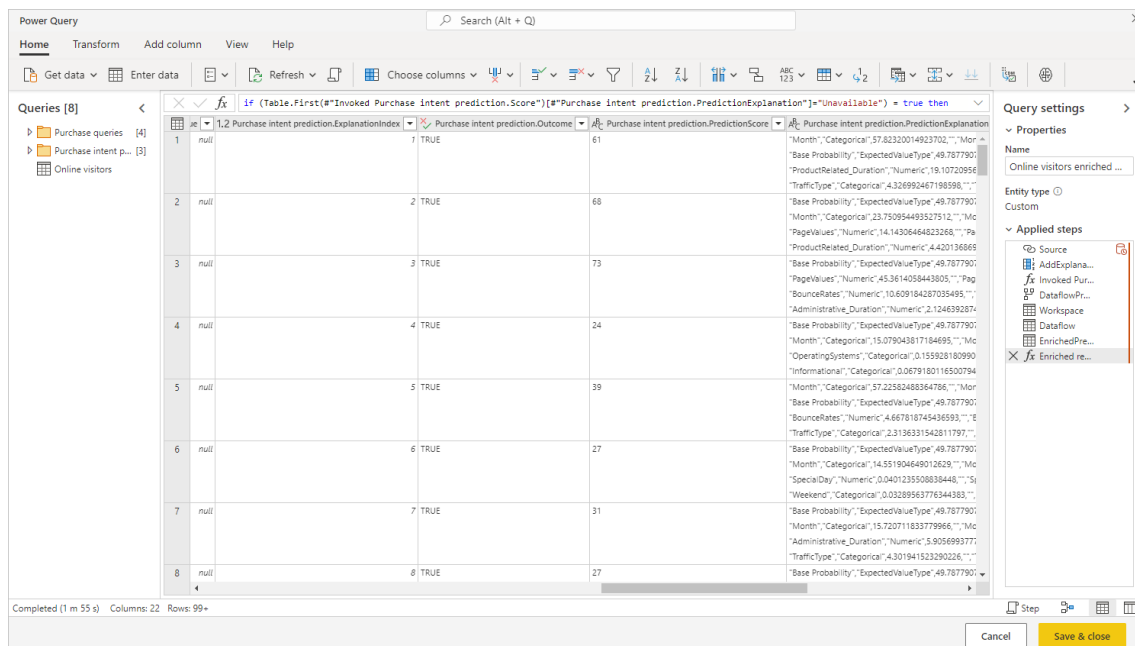
La aplicación del modelo de predicción binaria agrega cuatro columnas: **Outcome** , **PredictionScore** , **PredictionExplanation** y **ExplanationIndex** , cada una con un prefijo **de predicción de intención de compra** .

Tables Machine learning models			Edit tables	Add tables	Close
TABLE NAME	TABLE TYPE	ACTIONS			
ExitRates	Double				
PageValues	Double				
SpecialDay	Double				
Month	String				
OperatingSystems	Int64				
Browser	Int64				
Region	Int64				
TrafficType	Int64				
VisitorType	String				
Weekend	Boolean				
Revenue	Boolean				
Purchase intent prediction.Outcome	Boolean				
Purchase intent prediction.PredictionScore	Decimal				
Purchase intent prediction.PredictionExplanation	String				
Purchase intent prediction.ExplanationIndex	Int64				

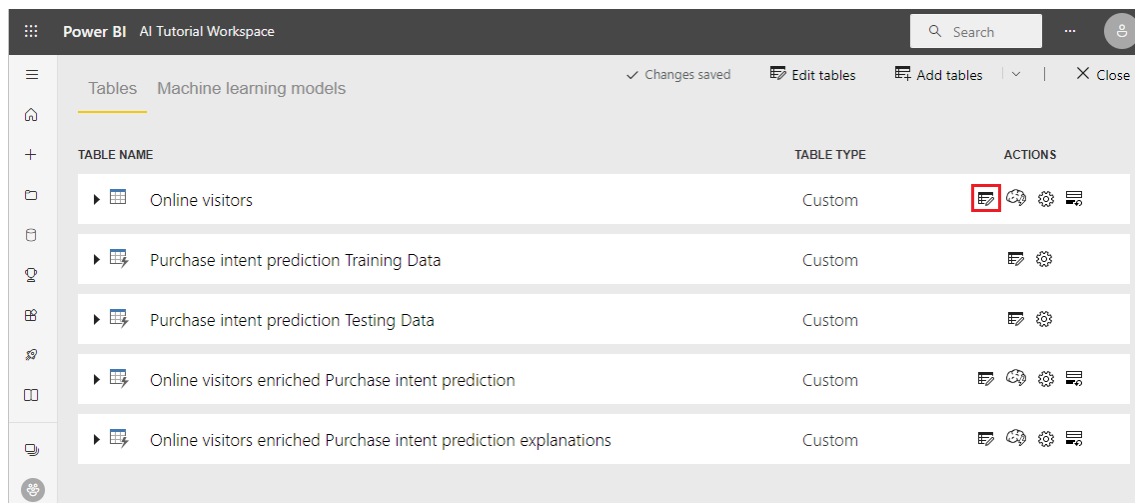
Online visitors enriched Purchase

Custom

Una vez que se complete la actualización del flujo de datos, puede seleccionar la tabla **de predicción de intención de compra enriquecida de visitantes en línea** para ver los resultados.



También puede invocar cualquier modelo de aprendizaje automático automatizado en el espacio de trabajo directamente desde el Editor de Power Query en su flujo de datos. Para acceder a los modelos de aprendizaje automático automatizado, seleccione **Editar** para la tabla que desea enriquecer con información de su modelo de aprendizaje automático automatizado.



En el Editor de Power Query, seleccione **AI Insights** en la cinta.





Power BI AI Tutorial Workspace Online visitors dataflow PPU Trial 17 days left

Home Transform Add column View Help Search (Alt + Q)

Get data Enter data Refresh Choose columns Remove columns Keep rows Remove rows Filter rows Split column

Queries [8]

- Purchase queries
  - Purchase intent prediction Training Data
  - Purchase intent prediction Testing Data
  - Online visitors enriched Purchase intent prediction
  - Purchase intent prediction queries
- Online visitors

	1	2	3	4	5
1	1	1	1	1	1
2	1	2	3	4	5
3	1	2	3	4	5
4	1	2	3	4	5
5	1	2	3	4	5

Query settings

Properties

Name Online visitors enriched Purchase intent prediction

Entity type Custom

Applied steps

- Source
- AddExplorationIndex
- InvokePurchaseIntentPredictionScore
- DataflowPrecalculatedSource
- Workspace
- Dataflow
- EnrichedPreview
- Enriched results

Después de guardar el flujo de datos, el modelo se invoca automáticamente cuando el flujo de datos se actualiza, para cualquier fila nueva o actualizada en la tabla de entidades.