**Optimización de columnas y metadatos**

**Introducción**

Microsoft Power BI contiene varias funciones que le permiten dar forma, transformar y modelar datos para crear perspectivas inmersivas e interactivas. Sin embargo, si el modelo de datos es grande y complejo, las manipulaciones de datos pueden parecer más lentas de lo habitual. ¿La solución? Necesita optimizar sus columnas y metadatos.

Esta lectura se centra en la optimización de columnas y metadatos en Power BI y explora las funciones que ofrece Microsoft Power BI para este fin. Explica cómo inspeccionar tablas y columnas, renombrar columnas, categorizar datos, cambiar los tipos de datos de las columnas y añadir descripciones a los campos en Power BI.

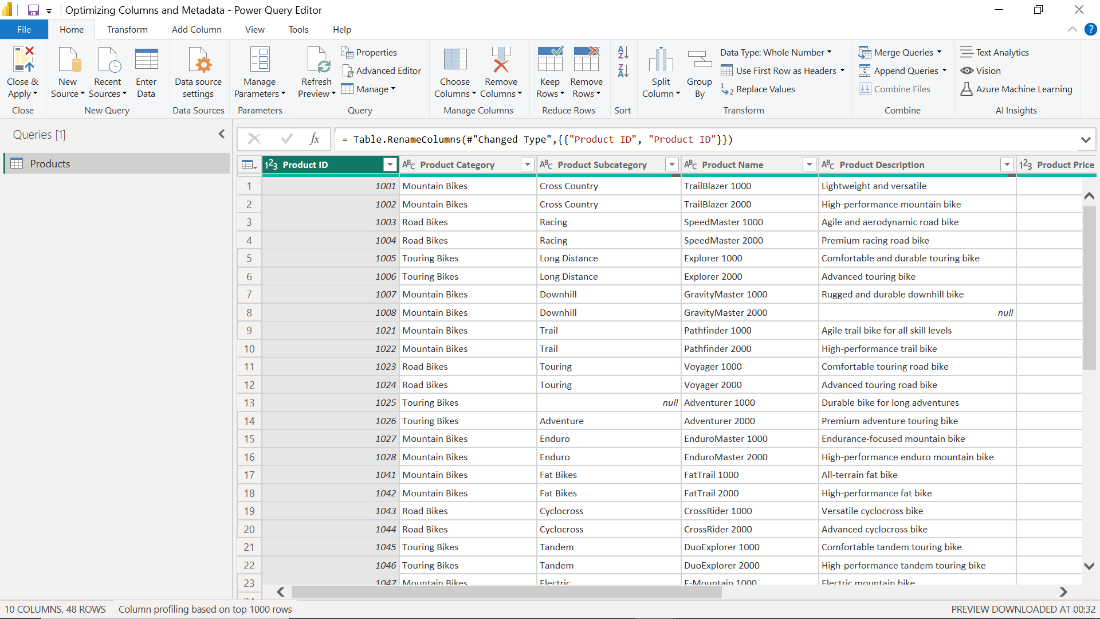
**Optimización de columnas y metadatos**

Los metadatos, en los términos más sencillos, son datos sobre datos. Proporciona información sobre otros datos, facilitando su recuperación, manipulación y gestión. Algunos ejemplos de metadatos en Microsoft Power BI, serían los nombres de las tablas, los nombres de las columnas, las relaciones o los tipos de datos. Los metadatos actúan como un mapa que le guía a través del paisaje de datos, facilitando la navegación por vastos volúmenes de información. Las columnas son componentes fundamentales de las tablas en Power BI, que contienen datos de diversos atributos o variables. Cada columna tiene un tipo de datos específico que determina qué tipo de valores puede almacenar. El tipo de datos puede ser texto, número entero, número decimal o fecha/hora.

Optimizar las columnas y los metadatos es importante. En Power BI, cada dato almacenado consume memoria, y la eficiencia con la que se almacenan y recuperan los datos puede afectar significativamente al rendimiento de su modelo de datos. Un modelo de datos optimizado garantiza una experiencia de Power BI más rápida y fluida y utiliza la memoria de forma más eficiente, lo que contribuye a la rentabilidad. Power BI ofrece múltiples formas de optimizar columnas y metadatos.

**Vista previa de datos**

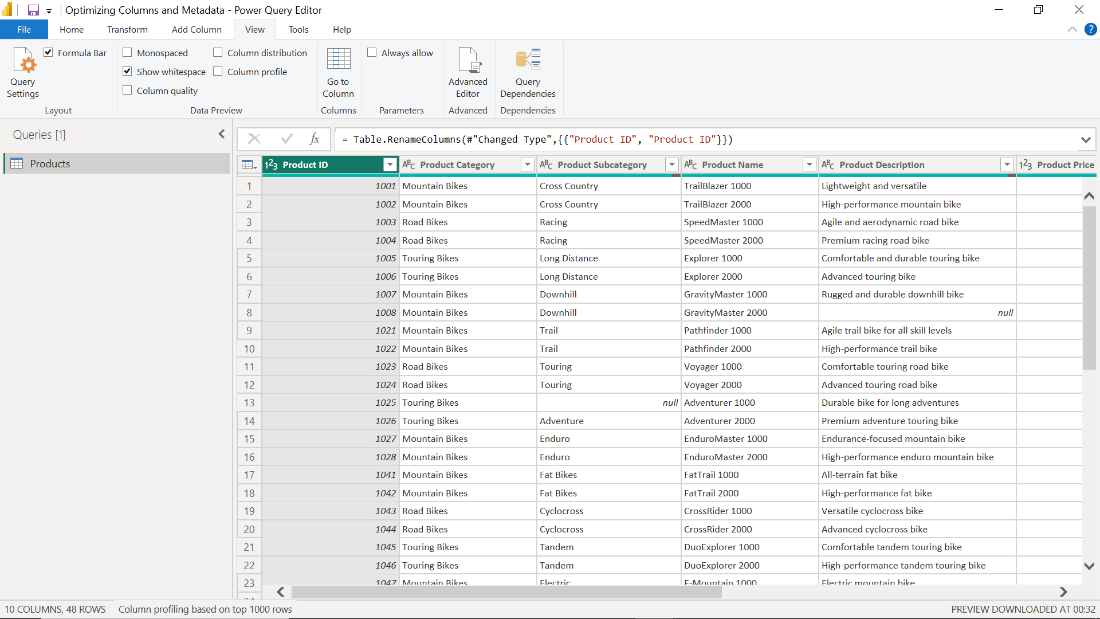
Comprender los metadatos de sus columnas puede mejorar significativamente la calidad de su análisis, ya que ayuda a identificar errores, valores atípicos, patrones y distribuciones que podrían influir en sus interpretaciones de los datos. El **Editor de Power Query** ofrece sólidas opciones para previsualizar sus datos y analizar los metadatos, entre las que se incluyen Calidad de **columna**, **Distribución de columna** y **Perfil de columna.**



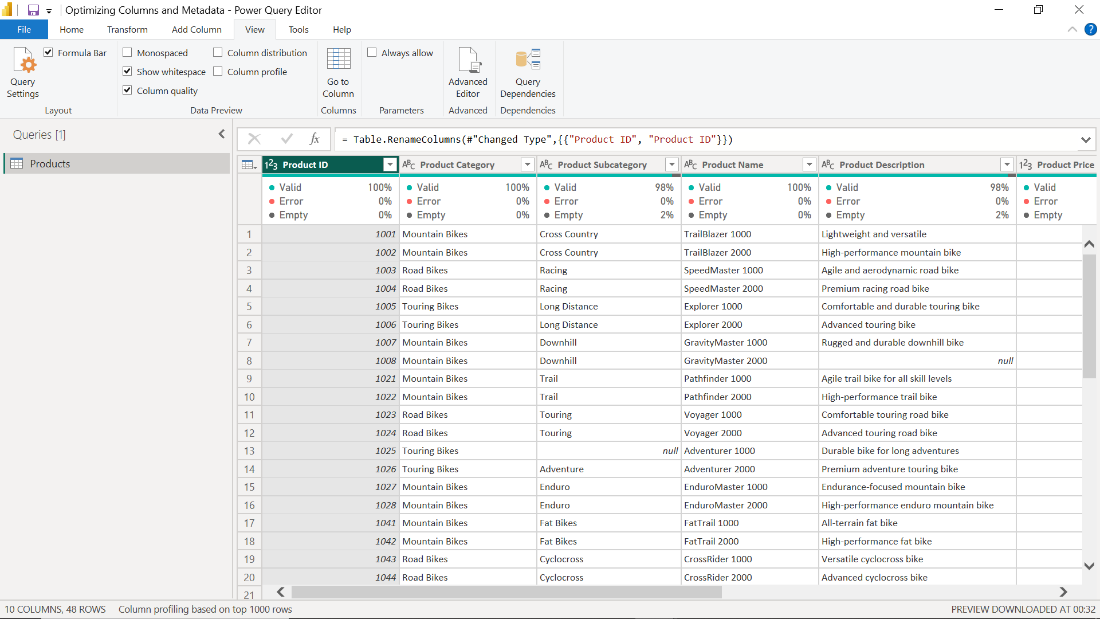
**Calidad de columna**

Esta función le permite evaluar rápidamente la calidad de los datos de sus columnas. Cuantifica el porcentaje de entradas válidas, erróneas o vacías en su columna, ofreciéndole una rápida visión general de los posibles problemas que podrían tener que abordarse. Los pasos para comprobar la calidad **de** las columnas son los siguientes:

1. Seleccione la pestaña **Ver** en las cintas de la parte superior de la ventana **del Editor de Microsoft Power Query**.



1. En el grupo **Vista previa de datos**, seleccione la casilla Calidad de **columna** para activar la función.

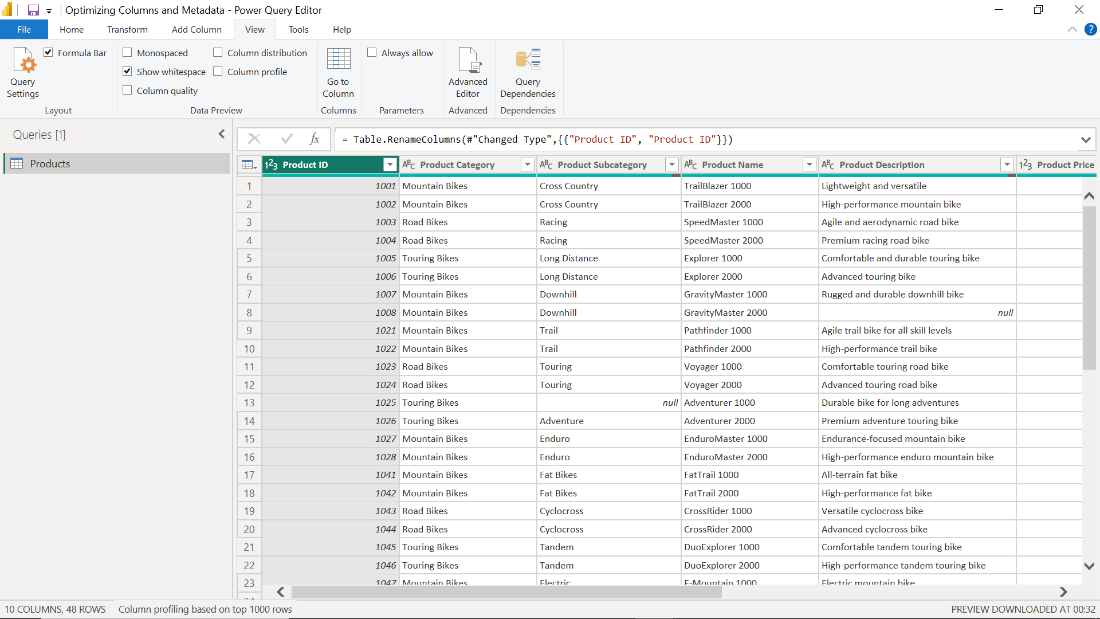


1. Se abre un panel de Calidad de **columna** debajo de cada cabecera de columna en la vista previa de datos. Muestra el porcentaje de entradas válidas, de error y vacías de una columna. Un porcentaje de **Válido** inferior al 100% indica la presencia de errores o celdas vacías que es necesario abordar.

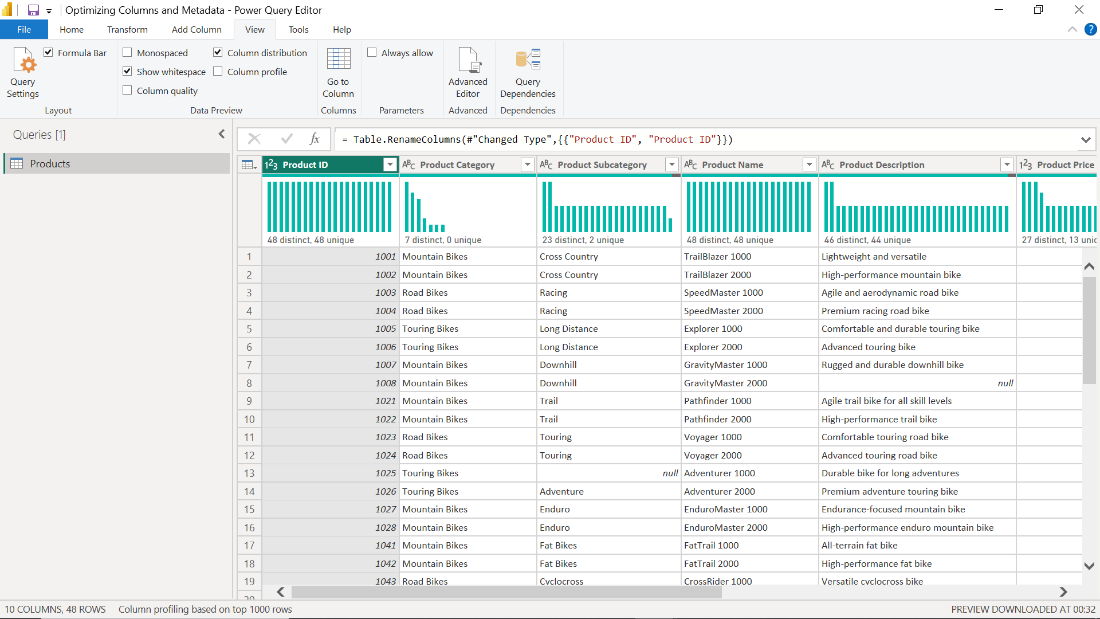
**Distribución de columnas**

Esta función muestra la frecuencia y distribución de los valores dentro de las columnas, visualizados como un histograma. Le ayuda a comprender el rango y la frecuencia de los datos, lo que resulta crucial para el análisis de los mismos. Los pasos para activar esta característica son los siguientes:

1. Seleccione la pestaña **Ver** en las cintas de la parte superior de la ventana **del Editor de Power Query**. En el grupo **Vista previa de datos**, seleccione la casilla Distribución **de** columnas.



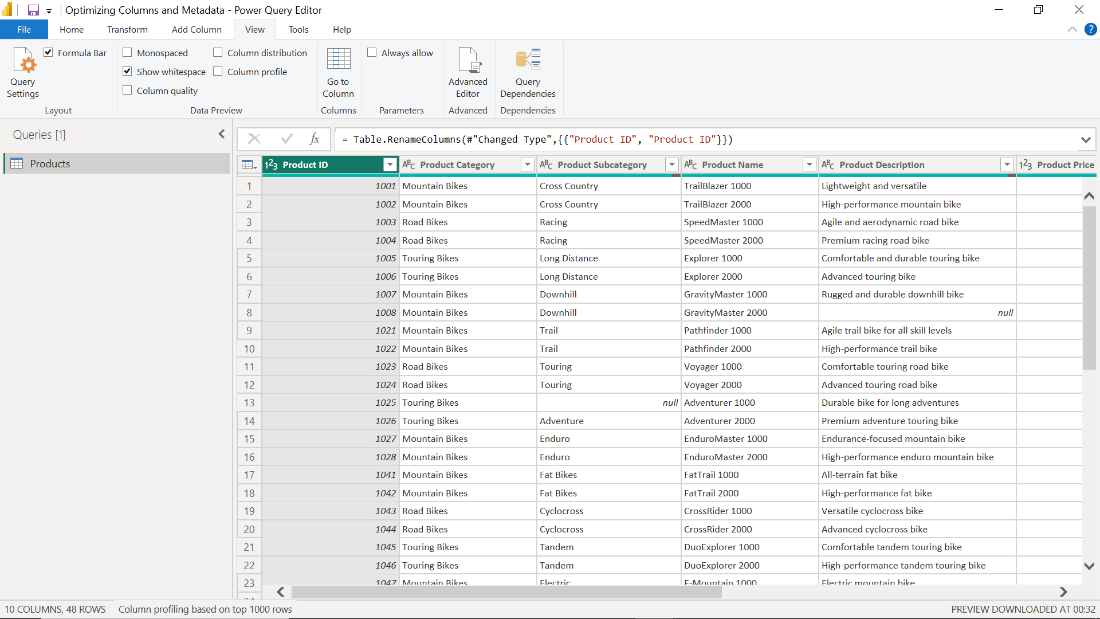
1. El panel Distribución de **columnas** se abre bajo los encabezados de columna en la vista previa de datos. Muestra un histograma que representa la frecuencia de los valores dentro de su columna.



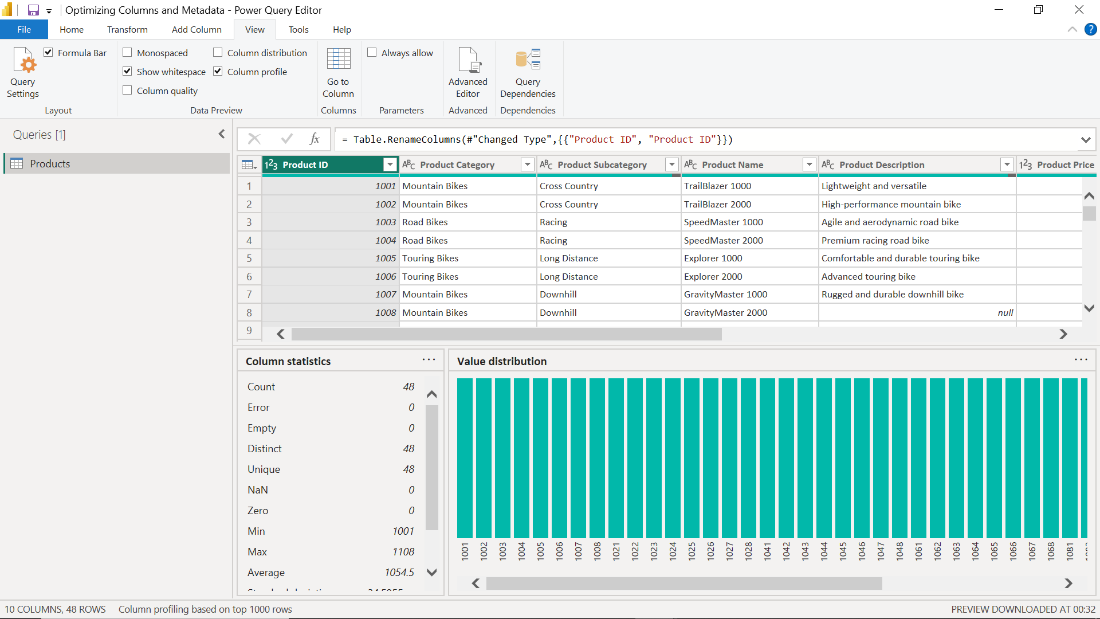
**Perfil de columna**

La función Perfil **de** columna proporciona una visión detallada de sus columnas, mostrando medidas estadísticas como recuento, recuento único, mín., máx. y media. También muestra un gráfico de distribución. Esta característica es muy beneficiosa a la hora de comprender sus datos y ayuda a identificar problemas como los valores atípicos. Los pasos para activar esta característica son los siguientes:

1. Seleccione la pestaña **Ver** en las cintas de la parte superior de la ventana **del Editor de Power Query**. En el grupo **Vista** previa de datos, seleccione la casilla **Perfil de columna** .

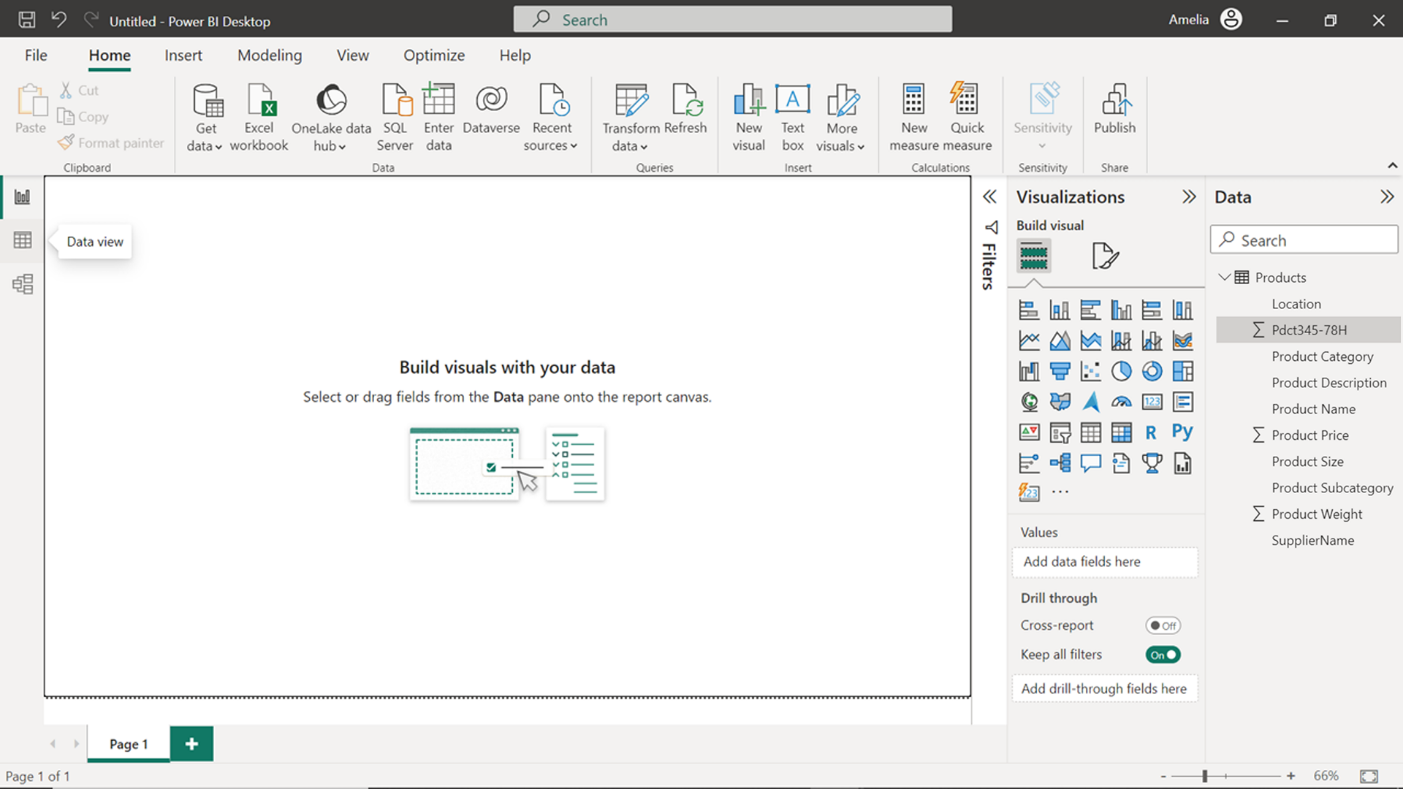


1. En la ventana Vista previa **de** datos, seleccione la columna que desea inspeccionar. Se abre un panel estadístico adicional debajo de la vista previa de datos. El panel estadístico se divide en dos subsecciones: **Estadísticas de columna** y **Distribución de valores**. Muestra una serie de medidas estadísticas y un gráfico de distribución, proporcionando un perfil rico y detallado de su columna seleccionada.



**Inspeccionar las tablas y columnas**

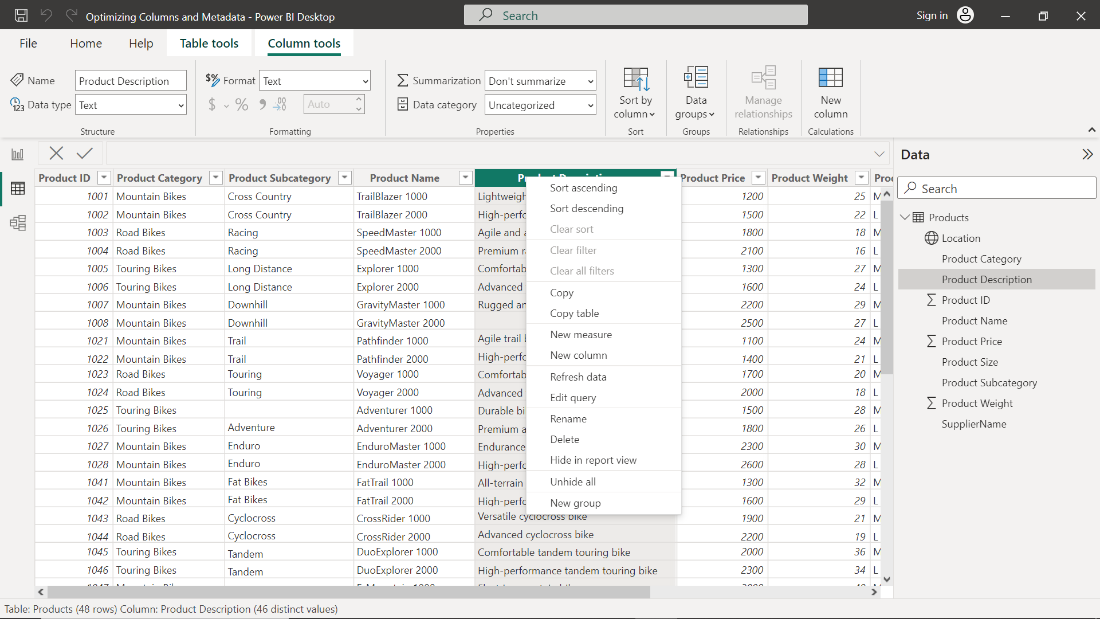
Para examinar la estructura actual de su informe, navegue hasta la **vista Datos**. Aquí podrá ver todas sus tablas en el panel **Datos** de la derecha. Al seleccionar una tabla, ésta se expande para revelar sus columnas. Verá los nombres de las columnas, los tipos de datos y los registros, lo que le dará una idea de los datos que contiene cada una.



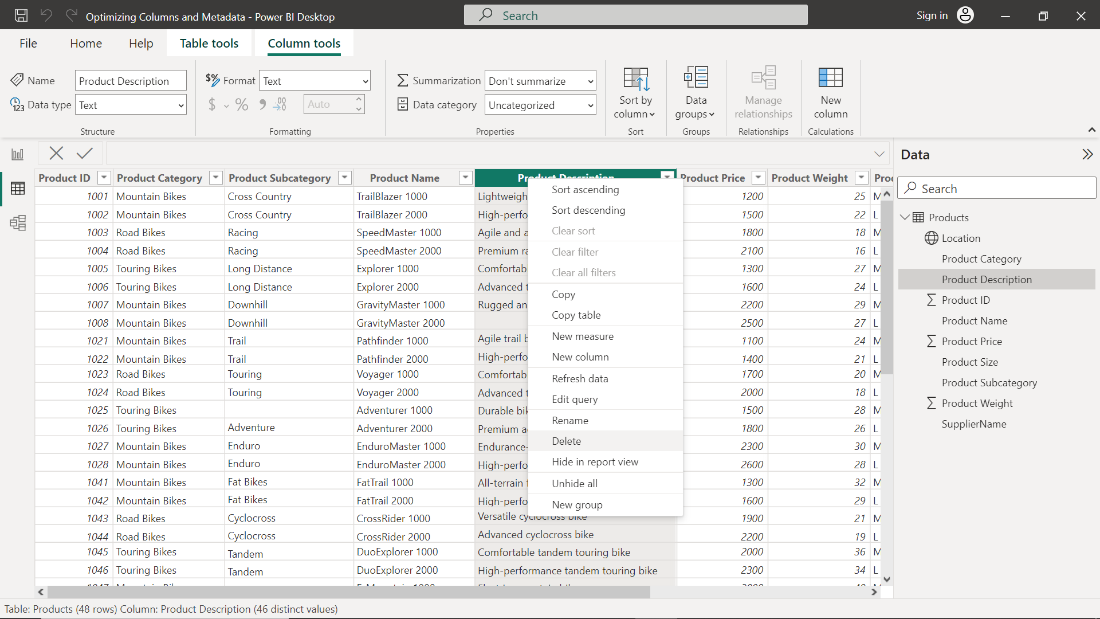
**Eliminación de columnas**

Cuando adquiera un conocimiento más profundo de sus datos, puede darse cuenta de que algunas columnas son innecesarias para su análisis. Las columnas innecesarias ocupan memoria y pueden ralentizar las consultas. Recuerde que menos es más cuando se trata de modelar datos, ya que querrá que su modelo de datos incluya sólo lo necesario para el análisis. Los pasos para eliminar columnas innecesarias son los siguientes:

1. En primer lugar, seleccione el nombre de la tabla en el panel **Datos** para mostrar una lista de sus columnas. A continuación, haga clic con el botón derecho del ratón en la columna que desee eliminar.



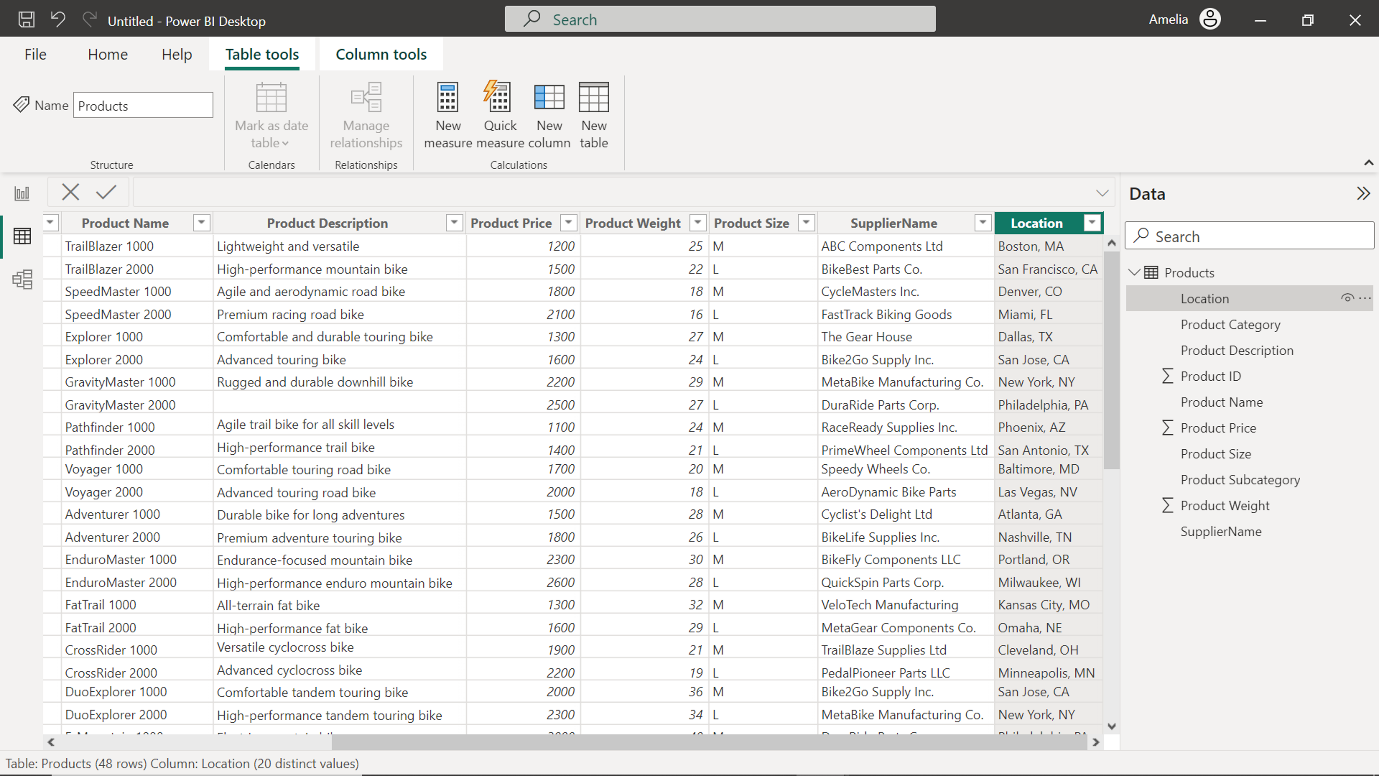
1. En el menú que aparece, seleccione **Eliminar**. Esta acción eliminará la columna de su conjunto de datos, ayudándole a mantener su modelo de datos ágil, mejorando su rendimiento y simplificando su estructura.



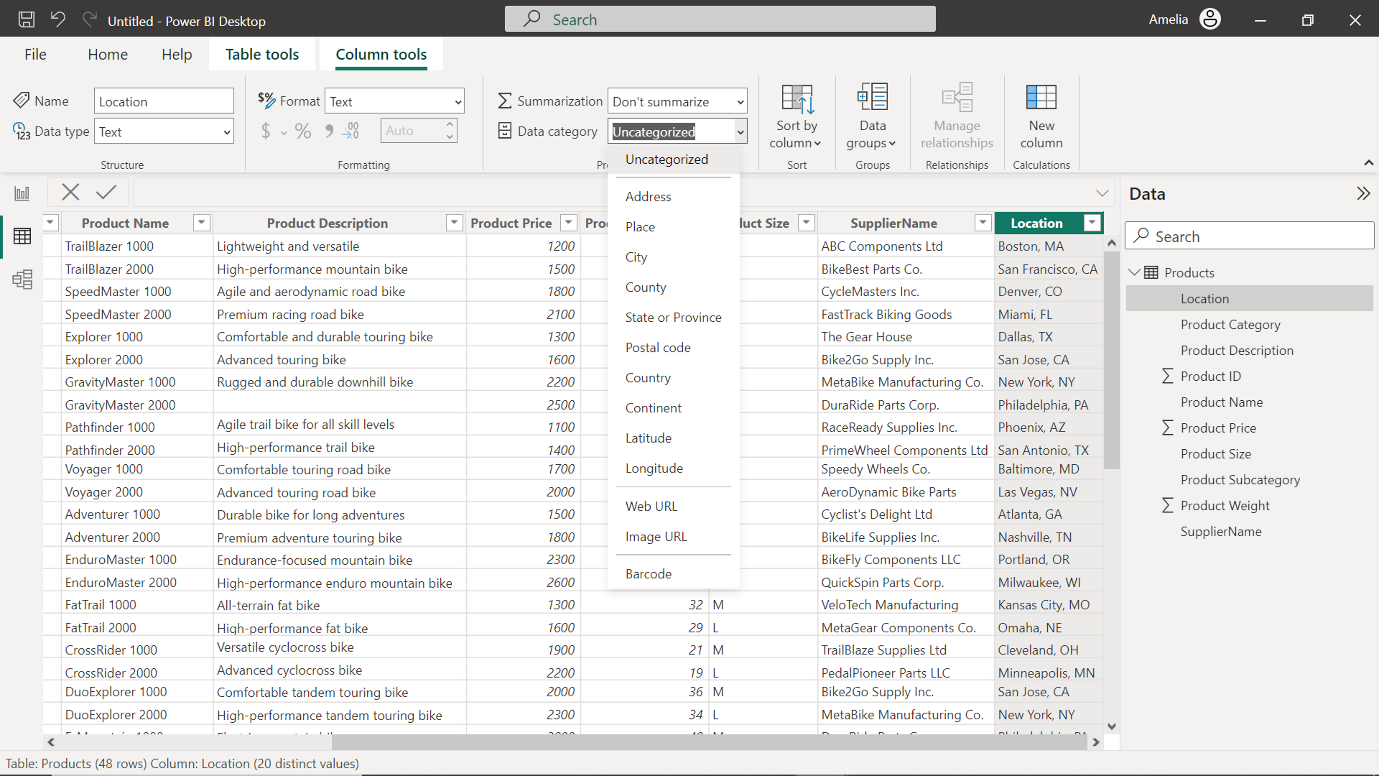
**Categorización de datos**

Una vez que haya eliminado las columnas innecesarias, deberá categorizar sus datos. Categorizar sus columnas en Power BI puede ayudar a que el sistema comprenda mejor sus datos y ofrezca características y perspectivas más relevantes. Por ejemplo, si una columna se categoriza como **Ciudad**, Power BI sabrá que esta columna puede utilizarse en análisis geográficos. Los pasos para categorizar columnas son los siguientes:

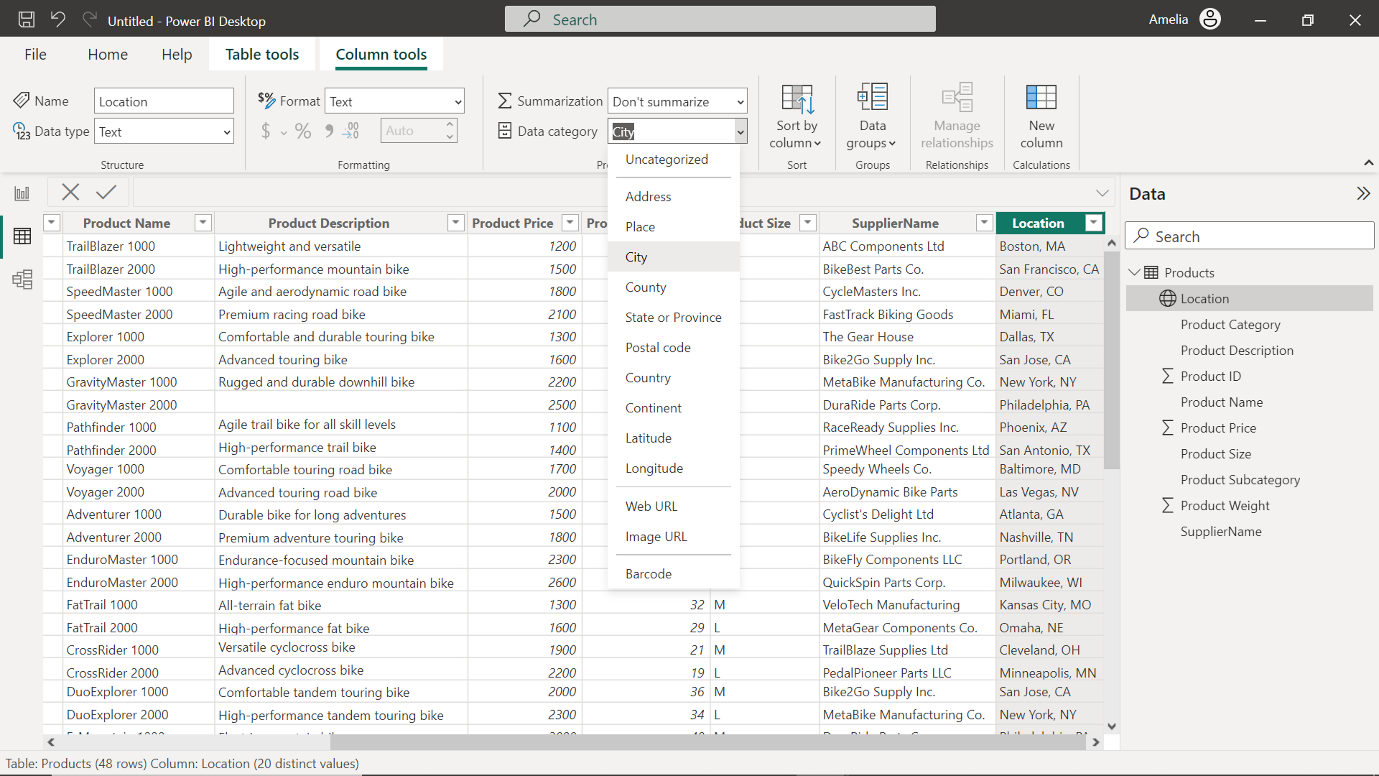
1. En la **vista Datos**, seleccione la columna que desea categorizar.



1. En el área de la cinta, seleccione la pestaña **Herramientas de columna**. En el grupo **Propiedades** , seleccione el desplegable **Categoría de datos**.



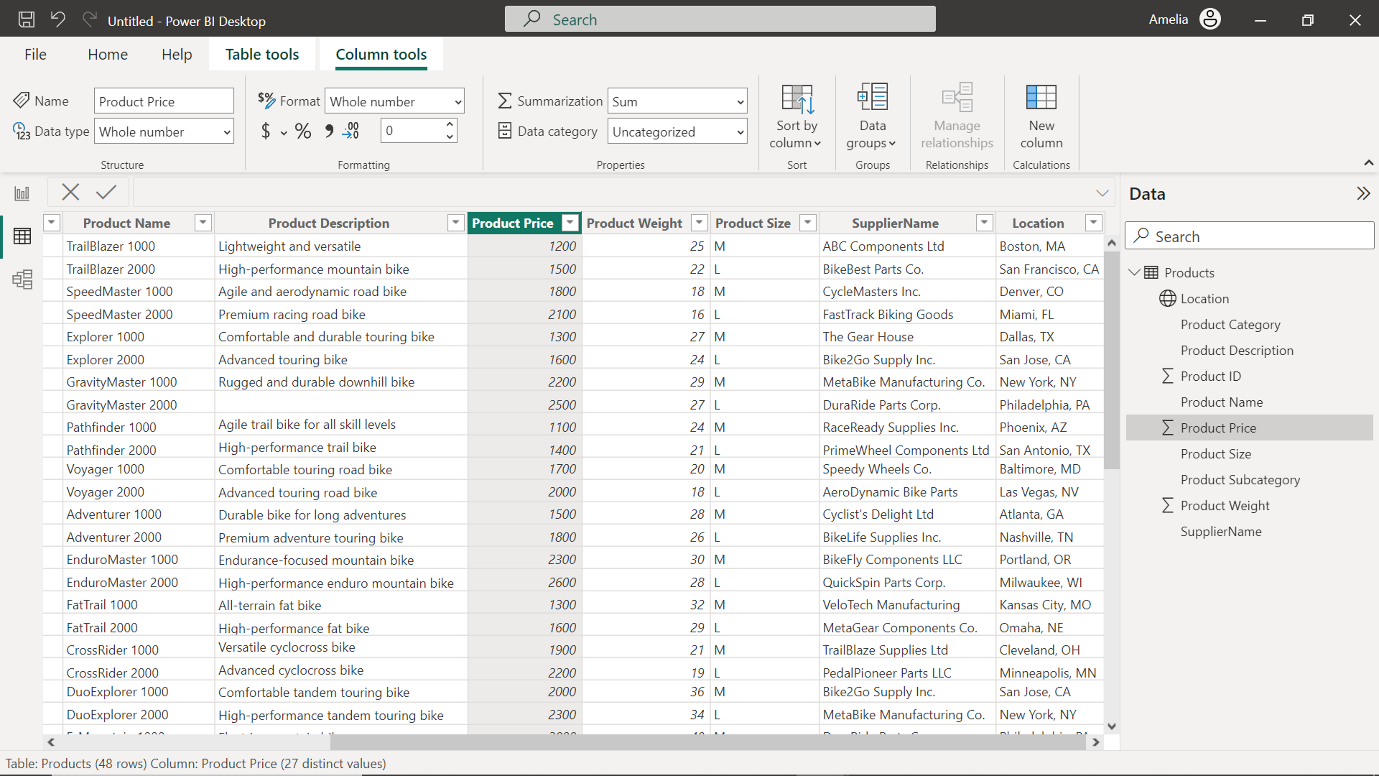
1. Se abre un menú desplegable que contiene categorías como **Ciudad**, **País**, **URL de imagen** y otras. Seleccione la categoría que mejor se ajuste a los datos de su columna.



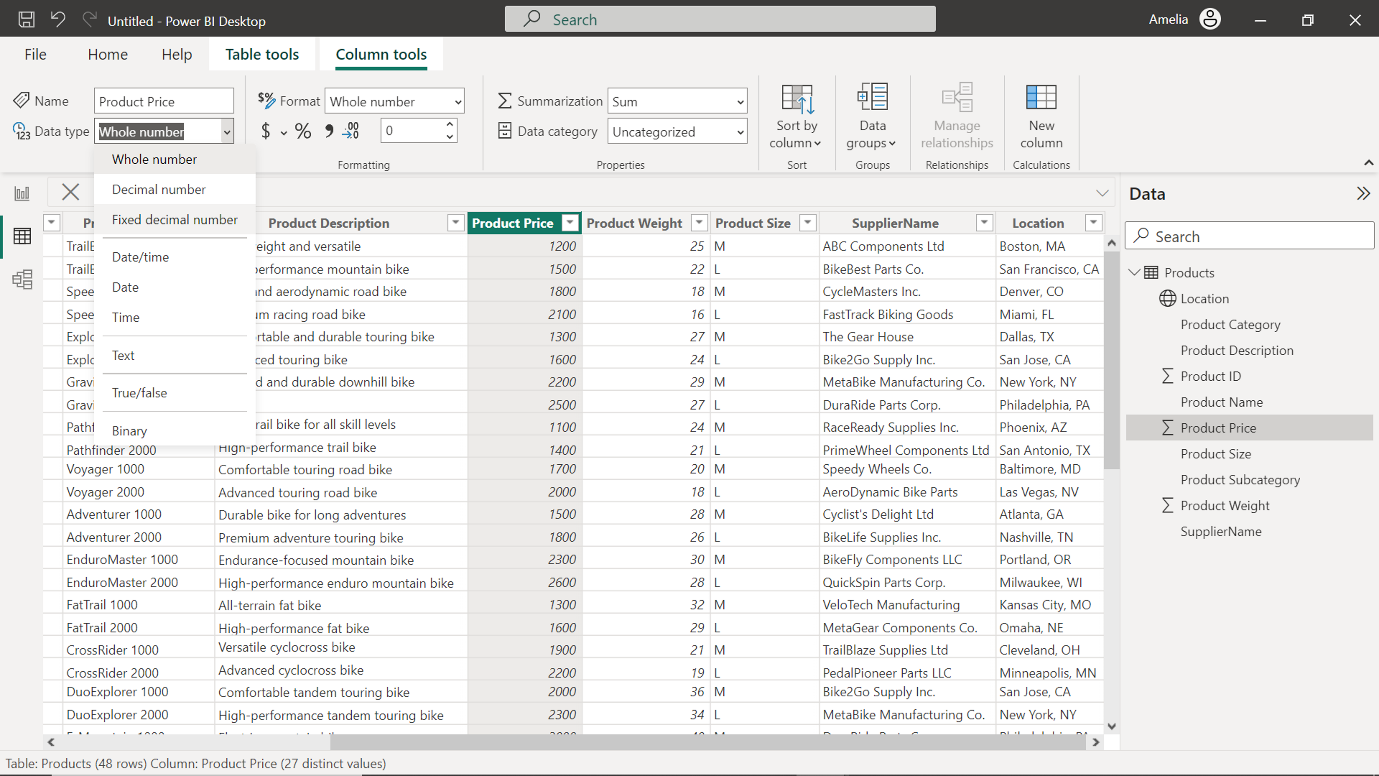
**Cambiar los tipos de datos de la columna**

Microsoft Power BI intenta adivinar el tipo de datos de cada columna cuando importa datos por primera vez. A veces, puede asignar un tipo de datos que no sea el más eficiente. Por ejemplo, si a una columna de números se le asigna un tipo de datos **Texto**, puede ocupar más memoria de la necesaria. Para cambiar el tipo de datos, haga clic con el botón derecho del ratón en el nombre de la columna y seleccione **Cambiar tipo**. Aparecerá una lista de tipos de datos. Seleccione el más apropiado. Por ejemplo, una columna que contenga números enteros puede cambiarse de **Texto** a **Número entero**, que es un tipo de datos más eficiente en cuanto a memoria. Los pasos para definir tipos de datos para columnas son los siguientes:

1. En la **vista Datos**, seleccione la columna para la que desea definir el tipo de datos.



1. En la cinta de opciones, seleccione la pestaña **Herramientas de columna**. En el grupo **Propiedades** seleccione el desplegable **Tipo** de datos.



1. Se abrirá un menú desplegable con varios tipos de datos, como **Texto**, **Número entero**, **Número decimal**, **Fecha/Hora** y otros. Elija el tipo de datos que se corresponda con los datos de su columna.

**Conclusión**

La optimización de columnas y metadatos en Microsoft Power BI puede hacer que un modelo de datos sea más eficiente e inteligible. La optimización del modelo de datos mejora el rendimiento técnico de Power BI y allana el camino para un análisis de datos más preciso, significativo e impactante. Los cambios que ha implementado hoy pueden conducir a mayores conocimientos mañana, y estos conocimientos pueden influir en las decisiones estratégicas tomadas por Adventure Works.