



Informatica®

10.4.0

Guía de introducción a la creación de perfiles

© Copyright Informatica LLC 2010, 2020

Este software y la documentación se proporcionan exclusivamente en virtud de un acuerdo de licencia independiente que contiene restricciones de uso y divulgación. Ninguna parte de este documento puede ser reproducida o transmitida de cualquier forma o manera (electrónica, fotocopia, grabación o mediante otros métodos) sin el consentimiento previo de Informatica LLC.

Informatica y el logotipo de Informatica son marcas comerciales o marcas comerciales registradas de Informatica LLC en Estados Unidos y en las diversas jurisdicciones de todo el mundo. La lista actual de marcas comerciales de Informatica está disponible en Internet en <https://www.informatica.com/trademarks.html>. Otros nombres de productos y empresas pueden ser nombres o marcas comerciales de sus respectivos titulares.

Las bases de datos, el software y los programas de DERECHOS DEL GOBIERNO DE LOS ESTADOS UNIDOS, y la documentación e información técnica relacionadas entregadas a los clientes del Gobierno de los Estados Unidos constituyen "software informático comercial" o "datos técnicos comerciales" de acuerdo con el Reglamento de Adquisición Federal y las regulaciones complementarias específicas del organismo que correspondan. Como tales, el uso, la duplicación, la divulgación, la modificación y la adaptación están sujetos a las restricciones y los términos de licencia establecidos en el contrato gubernamental aplicable, y hasta donde sea aplicable en función de los términos del contrato gubernamental, a los derechos adicionales establecidos en FAR 52.227-19, Licencia de Software Informático Comercial.

Algunas partes de este software o la documentación están sujetas a derechos de autor de terceros. Se incluyen con el producto los avisos obligatorios de terceros.

La información contenida en esta documentación está sujeta a cambios sin previo aviso. Si encuentra algún problema en esta documentación, escríbanos a infa_documentation@informatica.com para notificarnoslo.

Los productos de Informatica gozan de garantía en función de los términos y condiciones de los acuerdos conforme a los cuales se proporcionen. INFORMATICA PROPORCIONA LA INFORMACIÓN DE ESTE DOCUMENTO "TAL CUAL" SIN GARANTÍA DE NINGÚN TIPO, EXPRESA O IMPLÍCITA, INCLUIDAS LAS GARANTÍAS DE COMERCIALIZACIÓN, ADAPTACIÓN A UN FIN PARTICULAR Y CUALQUIER GARANTÍA O CONDICIÓN DE NO INCUMPLIMIENTO.

Fecha de publicación: 2020-02-04

Tabla de contenido

Prefacio	6
Recursos de Informatica	6
Informatica Network.	6
Base de conocimiento de Informatica.	6
Documentación de Informatica.	6
Matrices de disponibilidad de producto de Informatica.	7
Informatica Velocity.	7
Catálogo de soluciones de Informatica.	7
Servicio internacional de atención al cliente de Informatica.	7
 Capítulo 1: Introducción.....	 8
Resumen de creación de perfiles.	8
Presentación de Informatica Analyst.	10
Interfaz de usuario de Informatica Analyst.	10
Interfaz de usuario de Informatica Developer.	12
Página de bienvenida de Informatica Developer.	13
Referencias rápidas.	14
Tareas de Informatica Developer.	14
Historial de la guía de autoaprendizaje.	14
Estructura de la guía de autoaprendizaje.	15
Requisitos previos de la guía de autoaprendizaje.	15
Guía de autoaprendizaje de Informatica Analyst.	16
Guía de autoaprendizaje de Informatica Developer.	16
 Parte I: Primeros pasos con Informatica Analyst.....	 18
 Capítulo 2: Lección 1. Configuración de la herramienta Analyst de Informatica	 19
Introducción a la configuración de Informatica Analyst.	19
Tarea 1. Inicio de sesión en Informatica Analyst.	20
Tarea 2. Creación de un proyecto.	20
Tarea 3. Creación de una carpeta.	21
Resumen de la configuración de Informatica Analyst.	21
 Capítulo 3: Lección 2. Cómo crear objetos de datos	 22
Introducción a la creación de objetos de datos.	22
Tarea 1. Creación del objeto de datos de archivo sin formato.	23
Tarea 2. Visualización de las propiedades del objeto de datos.	24
Resumen de la creación de objetos de datos.	24

Capítulo 4: Lección 3. Creación de perfiles predeterminados.	25
Introducción a la creación de perfiles predeterminados.	25
Tarea 1. Crear y ejecutar un perfil predeterminado.	26
Tarea 2. Ver los resultados del perfil en una vista de resumen.	27
Tarea 3. Ver los resultados del perfil en una vista detallada.	27
Resumen de la creación de perfiles predeterminados.	28
 Capítulo 5: Lección 4. Creación de perfiles personalizados.	 29
Introducción a la creación de perfiles personalizados.	29
Tarea 1. Crear un perfil personalizado.	30
Tarea 2. Ejecutar el perfil.	31
Tarea 3. Obtener detalles de los resultados del perfil.	31
Resumen de la creación de perfiles personalizados.	32
 Capítulo 6: Lección 5. Creación de reglas de expresión.	 33
Introducción a la creación de reglas de expresión.	33
Tarea 1. Crear reglas de expresión y ejecutar el perfil.	34
Tarea 2. Ver la salida de la regla de expresión.	34
Tarea 3. Editar las reglas de expresión.	35
Resumen de la creación de reglas de expresión.	35
 Capítulo 7: Lección 6. Creación y ejecución de cuadros de mando.	 36
Introducción a la creación y ejecución de cuadros de mando.	36
Tarea 1. Crear un cuadro de mando a partir de los resultados de un perfil.	37
Tarea 2. Ejecutar el cuadro de mandos.	38
Tarea 3. Ver el cuadro de mandos.	39
Tarea 4. Editar el cuadro de mandos.	39
Tarea 5. Configurar umbrales.	40
Tarea 6. Ver gráficos de tendencias de puntuación.	40
Creación y ejecución de cuadros de mando.	41
 Capítulo 8: Lección 7. Creación de tablas de referencia a partir de columnas de perfil.	 42
Introducción a la creación de tablas de referencia a partir de columnas de perfil.	42
Tarea 1. Crear una tabla de referencia a partir de columnas de perfil.	43
Tarea 2. Editar la tabla de referencia.	44
Resumen de la creación de tablas de referencia a partir de columnas de perfil.	44
 Capítulo 9: Lección 8. Creación de tablas de referencia.	 45
Introducción a la creación de tablas de referencia.	45
Tarea 1. Crear una tabla de referencia.	46
Resumen de la creación de tablas de referencia.	46

Parte II: Primeros pasos con Informatica Developer.....	47
Capítulo 10: Lección 1. Configuración de Informatica Developer.	48
Introducción a la configuración de Informatica Developer.	48
Tarea 1. Iniciar Informatica Developer.	49
Tarea 2. Añadir un dominio.	49
Tarea 3. Añadir un repositorio de modelos.	50
Tarea 4. Crear un proyecto.	50
Tarea 5. Crear una carpeta.	50
Tarea 6. Selección de un servicio de integración de datos predeterminado.	51
Resumen de la configuración de Informatica Developer.	51
Capítulo 11: Lección 2: Importar objetos de datos físicos.	52
Introducción a la importación de objetos de datos físicos.	52
Tarea 1. Importar el objeto de datos del archivo sin formato Boston_Customers.	53
Tarea 2. Importar el objeto de datos de archivo sin formato LA_Customers.	59
Tarea 3. Importar el objeto de datos del archivo sin formato All_Customers.	60
Resumen de la importación de objetos de datos físicos.	60
Capítulo 12: Lección 3. Ejecutar un perfil con datos de origen.	61
Introducción a la creación de perfiles de datos.	61
Tarea 1. Realizar un análisis de combinación en dos orígenes de datos.	62
Tarea 2. Ver los resultados de un análisis de combinación.	63
Tarea 3. Ejecutar un perfil en un origen de datos.	64
Tarea 4. Ver los resultados de perfil de columna.	64
Resumen de datos de creación de perfiles.	65
Índice.....	66

Prefacio

Siga las instrucciones de la *Guía de introducción a la creación de perfiles* para obtener información sobre cómo crear y ejecutar perfiles en Developer tool y en la Herramienta del analista. La guía sirve de tutorial para ayudar a usuarios noveles a aprender a usar *Informatica Developer* e *Informatica Analyst*.

La guía se ha redactado para analistas de datos y desarrolladores.

Recursos de Informatica

Informatica proporciona una variedad de recursos de productos a través de Informatica Network y otros portales en línea. Use los recursos para sacar el mayor provecho de los productos y las soluciones de Informatica y aprender de otros expertos en la materia y usuarios de Informatica.

Informatica Network

Informatica Network es la puerta de entrada a muchos recursos, entre ellos, la base de conocimientos de Informatica y el servicio internacional de atención al cliente de Informatica. Para entrar en Informatica Network, visite <https://network.informatica.com>.

Como miembro de Informatica Network, tiene las siguientes opciones:

- Buscar recursos de productos en la base de conocimientos
- Ver la información de disponibilidad del producto
- Crear y revisar casos de soporte
- Buscar su red de grupos de usuarios de Informatica locales y colaborar con sus pares

Base de conocimiento de Informatica

Use la base de conocimientos de Informatica para encontrar recursos de productos como artículos prácticos, procedimientos recomendados, tutoriales de video y respuestas a preguntas frecuentes.

Para buscar en la base de conocimiento, visite <https://search.informatica.com>. Si tiene preguntas, comentarios o ideas relacionadas con la base de conocimiento de Informatica, póngase en contacto con el equipo de la base de conocimiento de Informatica en KB_Feedback@informatica.com.

Documentación de Informatica

Use el portal de documentación de Informatica para recorrer una extensa biblioteca de documentación para las versiones de productos actuales y recientes. Para recorrer el portal de documentación, visite <https://docs.informatica.com>.

Si tiene preguntas, comentarios o ideas acerca de la documentación de los productos, póngase en contacto con el equipo de la documentación de Informatica en infa_documentation@informatica.com.

Matrices de disponibilidad de producto de Informatica

Las matrices de disponibilidad de producto (PAM, Product Availability Matrixes) indican las versiones de sistemas operativos, bases de datos y otros tipos de orígenes y destinos de datos admitidos por la versión de un producto. Puede recorrer las PAM de Informatica en <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity es una colección de consejos y procedimientos recomendados desarrollados por los servicios profesionales de Informatica que se basan en experiencias reales de cientos de proyectos de administración de datos. Informatica Velocity representa el conocimiento colectivo de los consultores de Informatica que trabajan con organizaciones de todo el mundo para planificar, desarrollar, implementar y dar mantenimiento a soluciones de administración de datos exitosas.

Puede encontrar recursos de Informatica Velocity en <http://velocity.informatica.com>. Si tiene alguna pregunta, comentario o idea acerca de Informatica Velocity, póngase en contacto con los servicios profesionales de Informatica en ips@informatica.com.

Catálogo de soluciones de Informatica

El catálogo de soluciones de Informatica es un foro donde puede buscar soluciones que aumenten, amplíen o mejoren sus implementaciones de Informatica. Aproveche cualquiera de los cientos de soluciones de socios y desarrolladores de Informatica que se encuentran en el catálogo para mejorar su productividad y acelerar la implementación de los proyectos. Puede encontrar el catálogo de soluciones de Informatica en <https://marketplace.informatica.com>.

Servicio internacional de atención al cliente de Informatica

Puede ponerse en contacto con un centro de atención global por teléfono o a través del Informatica Network.

Para encontrar el número de teléfono local del servicio internacional de atención al cliente de Informatica, visite el sitio web de Informatica en el siguiente vínculo:

<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Para encontrar recursos de soporte en línea en Informatica Network, visite <https://network.informatica.com> y seleccione la opción eSupport.

CAPÍTULO 1

Introducción

Este capítulo incluye los siguientes temas:

- [Resumen de creación de perfiles, 8](#)
- [Presentación de Informatica Analyst, 10](#)
- [Interfaz de usuario de Informatica Developer, 12](#)
- [Historial de la guía de autoaprendizaje, 14](#)
- [Estructura de la guía de autoaprendizaje, 15](#)

Resumen de creación de perfiles

Utilice la creación de perfiles para buscar el contenido, la calidad y la estructura de orígenes de datos de una aplicación, un esquema o una empresa. El contenido de un origen de datos incluye las frecuencias de valor y los tipos de datos. La estructura de un origen de datos incluye las claves y las dependencias funcionales.

Como parte del proceso de detección, puede crear y ejecutar perfiles. Un perfil es un objeto de repositorio que busca y analiza todas las irregularidades de los datos de diferentes orígenes de datos de la empresa, así como los problemas ocultos que pudiese haber con los datos y que ponen en riesgo los proyectos de datos. Ejecutar un perfil en cualquier origen de datos de la empresa le permitirá entender las ventajas y las desventajas de sus datos y metadatos.

Puede utilizar Informatica Analyst e Informatica Developer para analizar los datos de origen y los metadatos. Los analistas y los desarrolladores pueden usar estas herramientas para trabajar conjuntamente, detectar problemas de calidad en los datos y analizar las relaciones entre los datos. Según cual sea su función, puede usar las capacidades que ofrece la Herramienta del analista o Developer tool. El tipo de creación de perfiles que puede hacer depende de la herramienta que use.

Tanto en la Herramienta del analista como en Developer tool puede realizar las siguientes tareas:

- Crear perfiles de columna. El proceso implica detectar la cantidad de valores únicos, valores nulos y patrones de datos de una columna.
- Realizar la detección del dominio de datos. Puede detectar características de datos importantes dentro de una empresa.
- Conservar los resultados de perfiles, incluidos los tipos de datos, los dominios de datos, las claves principales y las claves externas.
- Crear cuadros de mando para supervisar la calidad de los datos.

- Elija un perfil del sistema operativo para crear y ejecutar perfiles de columna, perfiles de detección empresarial y cuadros de mando según los permisos del usuario del sistema operativo que se hayan definido en el perfil del sistema operativo.
- Utilizar bloqueos de activos del repositorio para evitar que otros usuarios sobrescriban el trabajo.
- Utilizar el sistema de control de versiones para guardar varias versiones de un perfil.
- Crear y asignar etiquetas a los objetos de datos.
- Consultar el significado de un nombre de objeto como un término empresarial en Business Glossary Desktop. Por ejemplo, puede buscar el significado de un nombre de columna o nombre de perfil para entender sus requisitos empresariales y actual implementación.

Puede realizar las siguientes tareas en Developer tool:

- Detectar el grado de combinaciones potenciales que hay entre dos columnas de datos en un origen de datos.
- Determinar el porcentaje de superposición de datos en pares de columnas dentro de un origen de datos o de diferentes orígenes de datos.
- Comparar los resultados de perfil de columna.
- Generar un objeto de asignación a partir de un perfil.
- Detectar claves principales en un origen de datos.
- Detectar claves externas en un conjunto compuesto por uno o más orígenes de datos.
- Detectar las dependencias funcionales entre las columnas de un origen de datos.
- Ejecutar tareas de detección de datos en un gran número de orígenes de datos en varias conexiones. Las tareas de detección de datos incluyen perfil de columna, deducción de clave principal y externa, detección del dominio de datos y la generación de un resumen gráfico consolidado de las relaciones de datos.

Puede realizar las siguientes tareas en la Herramienta del analista:

- Realizar la detección empresarial en un gran número de orígenes de datos en varias conexiones. Puede ver un resumen consolidado de resultados de la detección de metadatos de columna y dominios de datos.
- Realizar la búsqueda de detección para ver dónde se encuentran los datos y los metadatos en la empresa. Puede buscar activos específicos, tales como perfiles, reglas y objetos de datos. La búsqueda de detección encuentra activos e identifica las relaciones con otros recursos en las bases de datos y los esquemas de la empresa.
- Ver los resultados de perfil para una ejecución de perfil histórica.
- Compare los resultados de perfil de dos ejecuciones de perfil en un perfil de columna.
- Ver el linaje de cuadro de mando de cada métrica o grupo de métricas de cuadro de mando.
- Vea el panel de cuadros de mando.
- Añadir comentarios a un perfil o a columnas de un perfil.
- Asignar etiquetas a un perfil o a columnas de un perfil.
- Asignar términos empresariales a las columnas de un perfil.

Presentación de Informatica Analyst

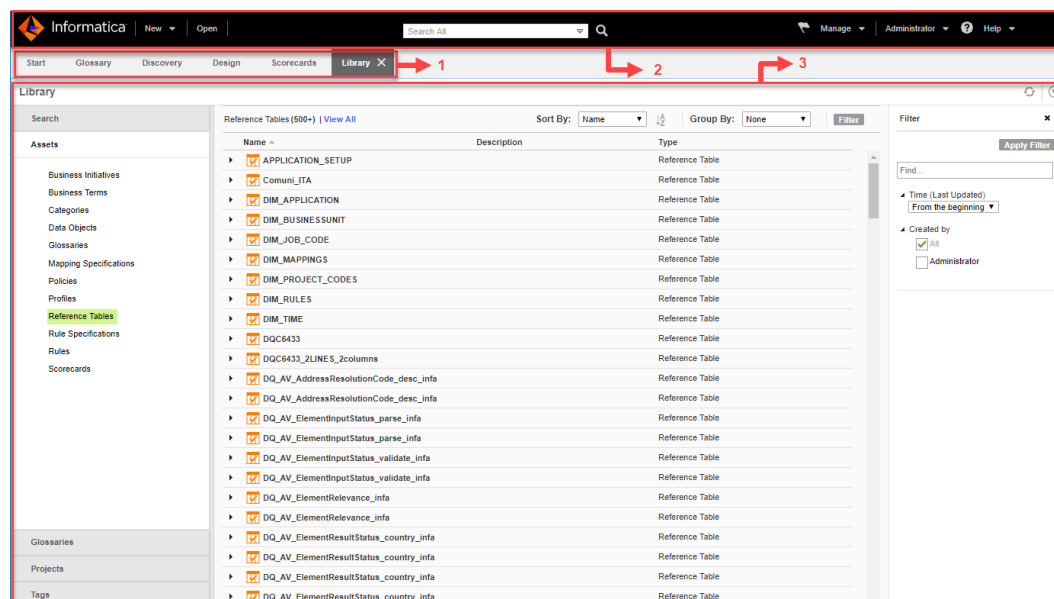
Informatica Analyst es un cliente de aplicación basado en web que los analistas pueden utilizar para analizar, limpiar, estandarizar, perfilar y puntuar los datos de una organización.

Según su licencia, los desarrolladores y analistas empresariales utilizan Analyst Tool para la colaboración controlada por datos. Puede realizar el perfilado de columnas y reglas, crear cuadros de mando y gestionar registros incorrectos y duplicados. También puede gestionar datos de referencia y ofrecer los datos a desarrolladores mediante una solución de calidad de datos.

Interfaz de usuario de Informatica Analyst

La interfaz de usuario de la Herramienta del analista se compone de varios espacios de trabajo que puede utilizar para crear soluciones de integración de datos y calidad de datos federación de datos. Un espacio de trabajo es una página web a la que puede acceder en función de la licencia y los privilegios que posea.

La siguiente imagen muestra la interfaz de usuario de la Herramienta del analista:



1. Panel de acceso al espacio de trabajo
2. Área de encabezado
3. Espacio de trabajo

La Herramienta del analista se abre en el espacio de trabajo **Inicio**. El espacio de trabajo **Inicio** enumera los espacios de trabajo de los que posee licencia y privilegios de uso mediante los paneles de acceso del espacio de trabajo.

La Herramienta del analista muestra los siguientes menús de encabezado:

Nuevo

Cree activos en los espacios de trabajo **Glosario**, **Detección** y **Diseño**.

Abrir

Abra el espacio de trabajo **Biblioteca**.

Alerta de notificaciones

Consulte las notificaciones para los activos de glosario.

Administrar

Abra espacios de trabajo temporales y **Notificaciones**. Puede abrir los espacios de trabajo **Conexiones**, **Dominios de datos**, **Estado de tarea**, **Proyectos** y **Seguridad de Business Glossary**.

Nombre de usuario

Establezca las preferencias del usuario para cambiar la contraseña y para cerrar sesión en la Herramienta del analista.

Ayuda

Acceda a la ayuda en el espacio de trabajo actual.

La Herramienta del analista contiene los siguientes espacios de trabajo permanentes:

Inicio

Acceda a otros espacios de trabajo de los que posea licencia para acceder mediante los paneles de acceso del espacio de trabajo. Si posee la licencia para realizar la administración de excepciones, sus tareas aparecerán en el espacio de trabajo **Mis tareas**.

Glosario

Defina y describa los conceptos empresariales que sean importantes para su organización. Puede crear y administrar términos, categorías, glosarios y directivas empresariales.

Detección

Analice la calidad de los datos y metadatos en los sistemas de origen. Puede crear y administrar perfiles, objetos de datos de archivos sin formato y objetos de datos de tablas.

Diseño

Diseñe la lógica empresarial que ofrezca ayuda para la colaboración de analistas y desarrolladores. Puede crear y administrar especificaciones de asignación, tablas de referencia y definiciones de regla.

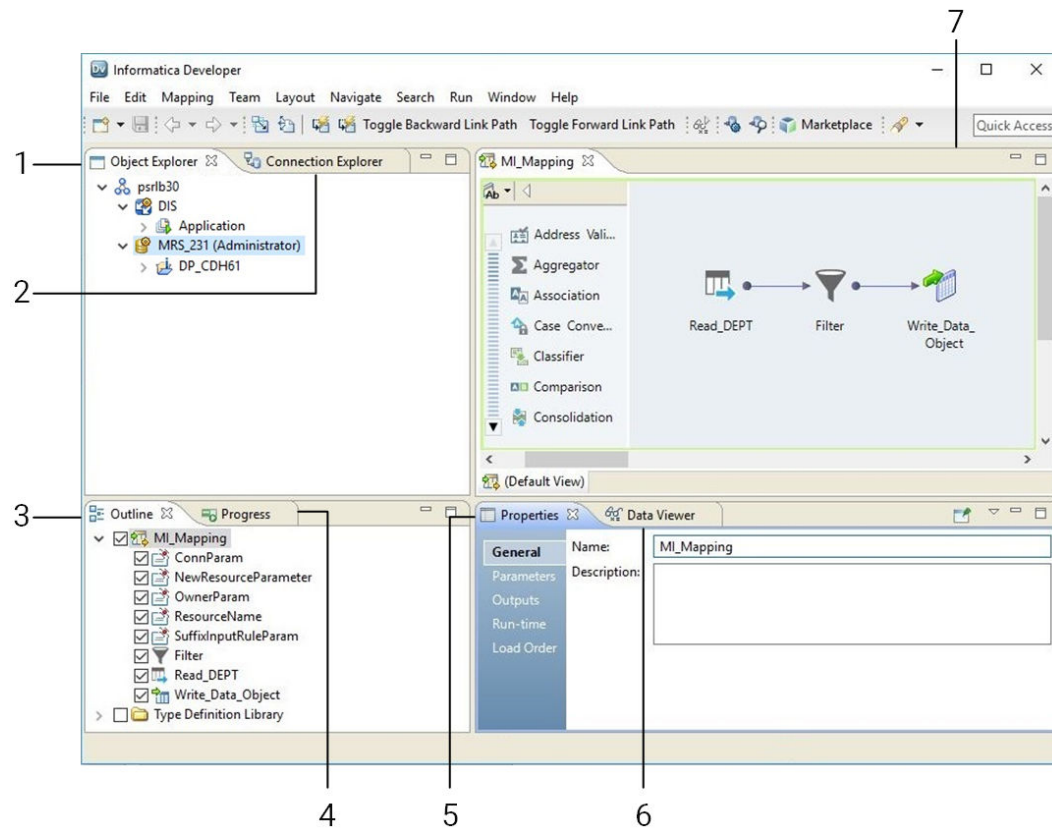
Cuadros de mando

Abra, edite y ejecute cuadros de mando creados a partir de los resultados del perfil. Puede agregar mediciones, obtener más detalles sobre columnas, agregar filtros de cuadros de mando y ver gráficos de tendencias de un cuadro de mando.

Interfaz de usuario de Informatica Developer

La interfaz de usuario de la herramienta Developer consta de un entorno de trabajo con varias vistas que puede utilizar para crear soluciones de integración de datos y calidad de los datos federación de datos .

La siguiente imagen muestra la interfaz de usuario de Developer tool:



1. Vista Explorador de objetos
2. Vista Esquema
3. Vista Propiedades
4. Vista Visor de datos
5. Vista Etiquetas
6. Vista Dependencias de objetos
7. Vista Alertas
8. Vista Explorador de conexiones
9. Editor

La Herramienta del desarrollador puede mostrar las siguientes vistas:

Vista Explorador de objetos

Muestra el dominio y los objetos en tiempo de diseño y en tiempo de ejecución del dominio. Los objetos en tiempo de diseño se almacenan en proyectos y carpetas en el Repositorio de modelos, mientras que los objetos en tiempo de ejecución se almacenan como parte de una aplicación en tiempo de ejecución en un servicio de integración de datos.

Vista Explorador de conexiones

Muestra las conexiones con las bases de datos relacionales.

Vista Esquema

Muestra los objetos que dependen de un objeto seleccionado en la vista **Explorador de objetos**.

Vista Referencias rápidas

Muestra la referencia rápida que haya abierto. Para abrir una referencia rápida, haga clic en **Ayuda > Referencias rápidas** y seleccione una referencia rápida.

Vista Visor de datos

Muestra los datos de origen, los resultados de perfil y las vistas previas del resultado de una transformación. También puede obtener una vista previa de mensajes de servicio web o ejecutar una consulta SQL desde la vista Visor de datos.

Vista Dependencias de objetos

Muestra las dependencias de objetos al visualizar, modificar o eliminar un objeto.

Vista Alertas

Muestra alertas del estado de conexión.

Vista Registro de validación

Muestra los errores de validación de objetos.

La herramienta Developer puede mostrar también otras vistas. Puede ocultar vistas y mover vistas a otra ubicación en el entorno de trabajo de la Herramienta del desarrollador. Haga clic en **Ventana > Mostrar vista** para seleccionar las vistas que desee mostrar.

Página de bienvenida de Informatica Developer

La primera vez que abre Developer tool, aparece la página de bienvenida. Use esta página para aprender cómo configurar y comenzar a trabajar en Developer tool.

La página de bienvenida muestra las siguientes opciones:

- **Resumen.** Haga clic en el botón Resumen para ver un resumen de las soluciones de calidad de datos y servicios de datos.
- **Primeros pasos.** Haga clic en el botón Primeros pasos para saber más sobre cómo configurar Developer tool y acceder a las lecciones sobre Informatica Data Quality e Informatica Data Services.
- **Guías de autoaprendizaje.** Haga clic en el botón Guías de autoaprendizaje para ver las lecciones de la guía de autoaprendizaje sobre soluciones de calidad de datos y servicios de datos.
- **Recursos web.** Haga clic en el botón Recursos web para obtener un vínculo a la base de conocimiento de Informatica. Puede acceder a la biblioteca de instrucciones de Informatica. La biblioteca de instrucciones de Informatica contiene artículos sobre Informatica Data Quality, Informatica Data Services y otros productos de Informatica.
- **Entorno de trabajo.** Haga clic en el botón Entorno de trabajo para comenzar a trabajar en la herramienta Developer.

Haga clic en **Ayuda > Bienvenida** para acceder a la página de bienvenida tras cerrarla.

Referencias rápidas

La herramienta Developer incluye referencias rápidas como parte de la ayuda en línea. Una referencia rápida es una guía paso a paso que ayuda a completar una o más tareas en la herramienta Developer.

Cuando se sigue una referencia rápida, se realizan las tareas y se ven los resultados. Por ejemplo, después de completar una referencia rápida para importar y previsualizar un objeto de datos relacional, ha importado una tabla de base de datos relacional y previsualizado los datos en la herramienta Developer.

Para acceder a las referencias rápidas, haga clic en **Ayuda > Referencias rápidas**.

Tareas de Informatica Developer

Utilice las funciones de calidad de datos de Developer tool para analizar el contenido y la estructura de los datos. Puede mejorar los datos para satisfacer sus necesidades empresariales.

Utilice Developer tool para diseñar y ejecutar procesos para lograr los siguientes objetivos:

- Ejecutar un perfil con datos de origen. La creación de perfiles revela el contenido y la estructura de los datos. La creación de perfiles es un paso clave en cualquier proyecto de datos, ya que permite identificar las ventajas y desventajas de los datos y, además, le ayuda a definir el plan de proyecto.
- Crear cuadros de mandos para revisar la calidad de los datos. Un cuadro de mandos es una representación gráfica de las mediciones de calidad de un perfil.
- Crear y ejecutar reglas de calidad de datos. Informatica proporciona reglas creadas previamente que se pueden ejecutar o editar para adaptarse a los objetivos de cada proyecto. Se pueden crear reglas en Developer tool.
- Colaborar con usuarios de Informatica. Las reglas que se añaden al repositorio de modelos están disponibles para los usuarios de las Developer tool y herramientas Analyst. Los usuarios pueden colaborar en proyectos; asimismo, varios usuarios pueden ser propietarios de objetos en distintas etapas de un proyecto.

Historial de la guía de autoaprendizaje

HypoStores Corporation es una organización minorista nacional con sede central en Boston y almacenes en varios estados. Integra datos operativos de todos los almacenes del país con los datos guardados en la sede de forma periódica. Recientemente, ha abierto un almacén en Los Ángeles.

La sede incluye un equipo ICC central de administradores, desarrolladores y arquitectos responsables de ofrecer una capa de servicios de datos común para todas las aplicaciones integradas y de inteligencia empresarial. Las aplicaciones de inteligencia empresarial incluyen un sistema CRM que contiene los archivos de datos maestros de clientes utilizados para la facturación y el marketing.

HypoStores Corporation quiere hacer un perfil de los datos de Boston y Los Ángeles antes de integrar los grupos de datos. Las operaciones de creación de perfiles identifican los problemas de calidad que HypoStores puede resolver antes de realizarse la integración.

Estructura de la guía de autoaprendizaje

La *Guía Primeros pasos* contiene guías de autoaprendizaje que incluyen lecciones y tareas.

Lecciones

Cada lección presenta conceptos que le ayudarán a comprender las tareas que deben realizarse en la lección. La lección proporciona los requisitos empresariales a partir del historial general. Los objetivos de la lección destacan las tareas que se realizarán para cumplir los requisitos empresariales. Para cada lección, se indica un tiempo estimado de realización. Después de terminar las tareas de la lección, puede revisar el resumen de la lección.

En el caso de que no esté configurado el entorno de la herramienta, la primera lección de cada guía de autoaprendizaje le ayudará a configurarlo.

Tareas

Las tareas ofrecen instrucciones paso a paso. Lleve a cabo todas las tareas en el orden indicado para realizar la lección.

Requisitos previos de la guía de autoaprendizaje

Antes de poder iniciar las lecciones de la guía de autoaprendizaje, el dominio de Informatica debe ejecutarse al menos con un nodo configurado. Debe comprobar que el servicio del analista, el servicio de repositorio de modelos y el servicio de integración de datos están en ejecución.

El instalador incluye los archivos de la guía de autoaprendizaje que se utilizarán para llevar a cabo las lecciones. Puede encontrar todos los archivos tanto en la instalación del cliente como del servidor:

- Puede encontrar los archivos de la guía de autoaprendizaje en la siguiente ubicación de la ruta de instalación de la herramienta Developer:

`<Directorio de instalación de Informatica>\clients\DeveloperClient\Tutorials`

- Puede encontrar los archivos de la guía de autoaprendizaje en la siguiente ubicación de la ruta de instalación de los servicios:

`<Directorio de instalación de Informatica>\server\Tutorials`

Necesita los siguientes archivos para las lecciones de la guía de autoaprendizaje:

- All_Customers.csv
- Boston_Customers.csv
- Customer_Order.xsd
- LA_customers.csv
- orders.csv

Guía de autoaprendizaje de Informatica Analyst

En esta guía de autoaprendizaje, un analista inicia una sesión en la Herramienta del analista, crea proyectos y carpetas, crea perfiles y reglas, crea cuadros de mando en datos y crea tablas de referencia.

La siguiente tabla describe las lecciones que puede realizar:

Lección	Descripción
Lección 1. Configuración de Informatica Analyst	Inicie sesión en la herramienta del analista y cree un proyecto y una carpeta para las lecciones de la guía de autoaprendizaje.
Lección 2. Creación de objetos de datos	Importe un archivo sin formato como un objeto de datos y prevismualice los datos.
Lección 3. Creación de perfiles predeterminados	Cree un perfil predeterminado para hacerse una idea rápida de la calidad de los datos.
Lección 4. Creación de perfiles personalizados	Cree un perfil personalizado para configurar las columnas y las opciones de muestreo y obtención de detalles.
Lección 5. Creación de reglas de expresión	Cree reglas de expresión para modificar y perfilar los valores de columna.
Lección 6. Creación y ejecución de cuadros de mando	Cree y ejecute un cuadro de mando para medir el progreso de la calidad de datos a lo largo del tiempo.
Lección 7. Creación de tablas de referencia a partir de columnas de perfil	Cree una tabla de referencia que pueda utilizar para estandarizar los datos de origen.
Lección 8. Creación de tablas de referencia	Cree una tabla de referencia para establecer relaciones entre los datos de origen y los valores válidos y estándar.

Guía de autoaprendizaje de Informatica Developer

En esta guía de autoaprendizaje, utilizará la herramienta Developer tool para llevar a cabo varias operaciones de calidad de datos. Cree una conexión con un repositorio de modelos, cree proyectos y carpetas, importe un

archivo sin formato como un objeto de datos físico, cree un perfil de detección empresarial, realice un análisis de combinación de dos orígenes de datos y visualice los resultados del perfil.

La siguiente tabla describe las lecciones que puede realizar:

Lección	Descripción
Lección 1. Configuración de Informatica Developer	Añada un dominio y cree una conexión con un repositorio de modelos gestionado por un servicio de repositorio de modelos en el dominio. Cree un proyecto y una carpeta para guardar el trabajo de las lecciones de la guía de autoaprendizaje. Seleccione un servicio de integración de datos predeterminado para ejecutar asignaciones y previsualizar los datos.
Lección 2. Importación de objetos de datos físicos	Importe un archivo sin formato para crear un objeto de datos físico.
Lección 3. Ejecutar un perfil con datos de origen	Cree un perfil de detección empresarial, realice un análisis de combinación de dos orígenes de datos y visualice los resultados del perfil. Es posible realizar un análisis de combinación en un perfil para determinar si se puede producir una combinación válida entre dos columnas de datos.

Parte I: Primeros pasos con Informatica Analyst

Esta parte incluye los siguientes capítulos:

- [Lección 1. Configuración de la herramienta Analyst de Informatica , 19](#)
- [Lección 2. Cómo crear objetos de datos , 22](#)
- [Lección 3. Creación de perfiles predeterminados, 25](#)
- [Lección 4. Creación de perfiles personalizados, 29](#)
- [Lección 5. Creación de reglas de expresión, 33](#)
- [Lección 6. Creación y ejecución de cuadros de mando, 36](#)
- [Lección 7. Creación de tablas de referencia a partir de columnas de perfil, 42](#)
- [Lección 8. Creación de tablas de referencia, 45](#)

CAPÍTULO 2

Lección 1. Configuración de la herramienta Analyst de Informatica

Este capítulo incluye los siguientes temas:

- [Introducción a la configuración de Informatica Analyst, 19](#)
- [Tarea 1. Inicio de sesión en Informatica Analyst, 20](#)
- [Tarea 2. Creación de un proyecto, 20](#)
- [Tarea 3. Creación de una carpeta, 21](#)
- [Resumen de la configuración de Informatica Analyst, 21](#)

Introducción a la configuración de Informatica Analyst

Antes de empezar con las lecciones de esta guía de autoaprendizaje, debe configurar la Herramienta del analista. Para configurar la Herramienta del analista, inicie sesión en la Herramienta del analista y cree un proyecto y una carpeta donde guardar el trabajo.

El dominio de Informatica es un conjunto de nodos y servicios que definen el entorno de Informatica. Los servicios del dominio incluyen el servicio del analista y el servicio de repositorio de modelos. El servicio del analista ejecuta la Herramienta del analista y el servicio de repositorio de modelos gestiona el repositorio de modelos. Cuando se trabaja con la Herramienta del analista, esta herramienta almacena los activos que se han creado en el repositorio de modelos.

Debe crear un proyecto antes de que pueda crear activos en la Herramienta del analista. Un proyecto contiene activos en la Herramienta del analista. Asimismo, un proyecto puede contener carpetas que almacenen activos relacionados, como objetos de datos que formen parte del mismo requisito de negocio.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

- Inicio de sesión en la Herramienta del analista.
- Creación de un proyecto para almacenar los activos que cree en la Herramienta del analista.
- Creación de una carpeta en el proyecto en la que pueda guardar los activos relacionados.

Requisitos previos

Antes de empezar con esta lección, compruebe los siguientes requisitos previos:

- Un administrador ha configurado un servicio de repositorio de modelos y un servicio del analista en la herramienta Administrator.
- Tiene el nombre de host y el número de puerto de la Herramienta del analista.
- Tiene un nombre de usuario y una contraseña para acceder al servicio del analista. Un administrador puede facilitarle esta información.

Duración

Destine de 5 a 10 minutos para realizar esta lección.

Tarea 1. Inicio de sesión en Informatica Analyst

Inicie sesión en la Herramienta del analista para iniciar la guía de autoaprendizaje.

1. Inicie un navegador Microsoft Internet Explorer o Google Chrome.
2. En el campo Dirección, escriba la URL de Informatica Analyst:
`http[s]://<fully qualified host name>:<port number>/analyst`
3. Si el dominio utiliza la autenticación nativa o de LDAP, introduzca su nombre de usuario y contraseña en la página de inicio de sesión.
4. Seleccione **Nativo** o el nombre de un dominio de seguridad específico.
El campo Dominio de seguridad aparece cuando el dominio de Informatica utiliza la autenticación de LDAP o Kerberos. Si no conoce el dominio de seguridad al que pertenece su cuenta de usuario, póngase en contacto con el administrador de dominios de Informatica.
5. Haga clic en **Iniciar sesión**.
La Herramienta del analista se abre en el espacio de trabajo **Inicio**.

Tarea 2. Creación de un proyecto

En esta tarea, se crea un proyecto que contenga los activos creados en Analyst Tool. Cree un proyecto de la guía de autoaprendizaje que contenga la carpeta del proyecto.

1. En el encabezado **Administrar**, haga clic en **Proyectos**.
Aparece el espacio de trabajo **Proyectos**.
2. En el menú **Acciones**, haga clic en **Nuevo > Proyecto**.
Aparece la ventana **Nuevo proyecto**.
3. Escriba su nombre precedido de "Tutorial_" como nombre del proyecto.
4. Haga clic en **Aceptar**.

Tarea 3. Creación de una carpeta

En esta tarea, tiene que crear una carpeta para almacenar los activos relacionados. Puede crear una carpeta en un proyecto o en otra carpeta. Cree una carpeta con el nombre Clientes para almacenar los activos relacionados en el proyecto de calidad de datos.

1. En el panel **Proyectos**, seleccione el proyecto de la guía de autoaprendizaje.
2. En el menú **Acciones**, haga clic en **Nuevo > Carpeta**.
Se abrirá la ventana **Nueva carpeta**.
3. Escriba Clientes como nombre de carpeta.
4. Haga clic en **Aceptar**.
Aparece la carpeta incluida en el proyecto de esta guía de autoaprendizaje.

Resumen de la configuración de Informatica Analyst

En esta lección ha aprendido que Analyst Tool almacena activos en proyectos y carpetas. Un repositorio de modelos contiene proyectos y carpetas. El servicio Analyst ejecuta la herramienta Analyst. El servicio de repositorio de modelos gestiona el repositorio de modelos. El servicio Analyst y el servicio de repositorio de modelos son servicios de aplicación del dominio Informatica.

Inició sesión en la herramienta Analyst y creó un proyecto y una carpeta.

Ya puede utilizar la herramienta Analyst para llevar a cabo otras lecciones de esta guía de autoaprendizaje.

CAPÍTULO 3

Lección 2. Cómo crear objetos de datos

Este capítulo incluye los siguientes temas:

- [Introducción a la creación de objetos de datos, 22](#)
- [Tarea 1. Creación del objeto de datos de archivo sin formato, 23](#)
- [Tarea 2. Visualización de las propiedades del objeto de datos, 24](#)
- [Resumen de la creación de objetos de datos , 24](#)

Introducción a la creación de objetos de datos

En la Herramienta del analista, un objeto de datos es una representación de datos basados en un archivo plano o tabla de base de datos relacional. Se crea un objeto de archivo sin formato o de tabla y luego se ejecuta un perfil sobre los datos del archivo sin formato o de la base de datos relacional. Cuando se crea un objeto de datos a partir de un archivo sin formato en la Herramienta del analista, se puede cargar el archivo en la memoria caché del archivo sin formato en el equipo en el que se ejecuta la Herramienta del analista o se puede especificar la ubicación de red donde está almacenado el archivo sin formato.

Historial

HypoStores guarda los datos de los clientes de Los Ángeles en archivos sin formato. HypoStores necesita crear un perfil y analizar los datos y realizar tareas de calidad de datos.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

1. Cargar el archivo sin formato en la ubicación de memoria caché del archivo sin formato y crear un objeto de datos.
2. Previsualizar los datos para el objeto de datos de archivo sin formato.

Requisitos previos

Antes de empezar con esta lección, compruebe los siguientes requisitos previos:

- Ha terminado la lección 1 de esta guía de autoaprendizaje.
- Tiene el archivo sin formato LA_Customers.csv. Puede encontrar este archivo en la carpeta <Directorio raíz de la instalación>\<Versión de lanzamiento>\clients\DeveloperClient\Tutorials.

Duración

Destine de 5 a 10 minutos para realizar esta tarea.

Tarea 1. Creación del objeto de datos de archivo sin formato

En esta tarea, se utiliza el asistente **Añadir archivo sin formato** para crear objetos de datos de archivo sin formato desde el archivo de datos LA_Customers.

1. En el **Navegador de proyectos**, seleccione la carpeta Clientes de su proyecto de la guía de autoaprendizaje.

Nota: Debe seleccionar el proyecto o la carpeta donde desee crear el objeto de datos de archivo sin formato para poder crearlo.

2. Haga clic en **Nuevo > Objeto de datos de archivo sin formato**.

Aparecerá el asistente **Añadir archivo sin formato** en el espacio de trabajo **Detección**.

3. Seleccione **Explorar y cargar** y haga clic en **Explorar**.

4. Explore la ubicación de LA_Customers.csv y haga clic en **Abrir**.

5. Haga clic en **Siguiente**.

En la página **Elegir tipo de importación** se muestran las opciones **Delimitado** y **De ancho fijo**. La opción predeterminada es **Delimitado**.

6. Haga clic en **Siguiente**.

7. En **Especificar los delimitadores y calificadores de texto utilizados en los datos**, seleccione **Comillas dobles** como calificador de texto.

8. En **Especificar líneas para importar**, seleccione **Importar desde la primera línea** para importar nombres de columnas desde la primera línea que no esté vacía.

9. Haga clic en **Siguiente**.

La sección **Atributos de columna** muestra el nombre, el tipo de datos, la precisión, la escala y el formato de cada columna.

10. Haga clic en **Siguiente**.

El campo **Nombre** muestra **LA_Customers**.

11. Opcionalmente, cambie el nombre del archivo y añada una descripción.

12. En la sección **Carpetas**, seleccione la carpeta **Clientes**.

13. Haga clic en **Finalizar**.

El objeto de datos LA_Customers aparece en el espacio de trabajo **Detección** con la vista previa de datos de los datos de origen.

Tarea 2. Visualización de las propiedades del objeto de datos

En esta tarea, puede ver las propiedades del objeto de datos LA_Customers.

1. Haga clic en **Abrir** para abrir el espacio de trabajo **Biblioteca**.
2. En el espacio de trabajo **Biblioteca**, haga clic en **Objetos de datos** en el panel **Activos**.
Se muestra una lista de objetos de datos en el panel **Objetos de datos**.
3. Haga clic en el archivo sin formato **LA_Customers**.
Se muestra el panel **Vista previa de datos** con los datos obtenidos del objeto de datos LA_Customers.
4. En el panel **Vista previa de datos**, revise la estructura y el contenido del objeto de datos LA_Customers.
La Herramienta del analista muestra las primeras 100 filas del objeto de datos de archivo sin formato.
5. Haga clic en **Propiedades**.
El panel **Propiedades** muestra el nombre, el tipo, la descripción y la ubicación del objeto de datos.
También puede ver los nombres y las propiedades de la columna para el objeto de datos.

Resumen de la creación de objetos de datos

En esta lección, ha aprendido que los objetos de datos son representaciones de datos basadas en un origen de archivo sin formato o de base de datos relacional. Ha aprendido que puede crear un objeto de datos de archivo sin formato y previsualizar los datos que contiene.

Ha cargado un archivo sin formato y creado un objeto de datos de un archivo sin formato, previsualizado los datos del objeto de datos y visualizado las propiedades del objeto de datos.

Después de crear un objeto de datos, cree un perfil predeterminado para el objeto de datos en la lección 3 y cree un perfil personalizado para el objeto de datos en la lección 4.

CAPÍTULO 4

Lección 3. Creación de perfiles predeterminados

Este capítulo incluye los siguientes temas:

- [Introducción a la creación de perfiles predeterminados, 25](#)
- [Tarea 1. Crear y ejecutar un perfil predeterminado, 26](#)
- [Tarea 2. Ver los resultados del perfil en una vista de resumen, 27](#)
- [Tarea 3. Ver los resultados del perfil en una vista detallada, 27](#)
- [Resumen de la creación de perfiles predeterminados, 28](#)

Introducción a la creación de perfiles predeterminados

Un perfil es el análisis de la calidad de datos basado en el contenido y la estructura de los datos. Un perfil predeterminado es un perfil que se crea con opciones predeterminadas. Utilice un perfil predeterminado para obtener resultados de perfil sin tener que configurar todas las columnas y opciones de un perfil.

Cree y ejecute un perfil predeterminado para analizar la calidad de los datos cuando inicie un proyecto de calidad de datos. Cuando cree un objeto de perfil predeterminado, seleccione el objeto de datos y las columnas del objeto de datos que desee analizar. Un perfil predeterminado omite la columna de perfil y la configuración de opciones. La Herramienta del analista crea los perfiles en el archivo sin formato activo del objeto de datos del archivo sin formato.

Historial

HypoStores desea incorporar a su almacén de datos los datos de la recién adquirida oficina de Los Ángeles. Antes de incorporar los datos al almacén de datos, éste se debe limpiar. Usted es el analista responsable de evaluar la calidad de los datos y de retransmitir la información al desarrollador responsable de la limpieza de los datos. Usted desea ver rápidamente los resultados del perfil y hacerse una idea básica sobre la calidad de los datos.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

1. Crear y ejecutar un perfil predeterminado para el objeto de datos de archivo sin formato LA_Customers.
2. Ver los resultados del perfil.

Requisitos previos

Antes de empezar con esta lección, verifique los siguientes requisitos previos:

- Ha terminado las Lecciones 1 y 2 de esta guía de autoaprendizaje.

Duración

Destine de 5 a 10 minutos para realizar esta lección.

Tarea 1. Crear y ejecutar un perfil predeterminado

En esta tarea, creará un perfil predeterminado para todas las columnas del objeto de datos y utilizará opciones predeterminadas de muestreo y de obtención de detalles.

1. En el espacio de trabajo **Biblioteca**, seleccione un objeto de datos en el panel **Activos**.
2. Haga clic con el botón derecho en el objeto de datos y seleccione **Crear perfil**.
Aparecerá el asistente **Nuevo perfil**.
3. En la pantalla **Especificar propiedades generales**, el nombre, la descripción y la ubicación se rellenan de forma predeterminada.
4. Haga clic en **Siguiente**.
5. En la pantalla **Seleccionar origen**, el objeto de datos aparece de forma predeterminada. Puede ver las columnas en el panel **Seleccionar columnas**.
6. Haga clic en **Siguiente**.
7. En la pantalla **Especificar la configuración**, las siguientes opciones están seleccionadas de forma predeterminada:
 - **Ejecutar perfil de columna**
 - **Todas las filas** en el panel **Ejecutar perfil en**
 - **Activo** en el panel **Obtención de detalles**
 - **Excluir de la inferencia los tipos y dominios de datos aprobados en las siguientes ejecuciones del perfil**.
 - Opción **Nativo** en el panel **Entorno en tiempo de ejecución**.
8. Haga clic en **Siguiente**.
9. En la pantalla **Especificar reglas y filtros**, haga clic en **Guardar y ejecutar** para crear y ejecutar el perfil.
La Herramienta del analista crea y ejecuta el perfil. Los resultados de perfil aparecen en la vista de resumen.

Tarea 2. Ver los resultados del perfil en una vista de resumen

En esta tarea, utilizará la vista de resumen del perfil LA_Customers para obtener un resumen rápido de los resultados del perfil.

1. En el panel **Biblioteca > Activos > Perfiles**, haga clic en el perfil **LA_Customers**.
Los resultados de perfil aparecen en la vista de resumen.
2. En la vista de resumen, haga clic en **Columnas** en el panel **Filtrar por** para ver los resultados del perfil por columnas.
3. Mueva el cursor sobre los gráficos de barras horizontales para ver los valores en porcentajes.
4. En la sección **Porcentaje de valores nulos, únicos y no únicos**, podrá ver los valores nulos, únicos y no únicos en porcentajes para una columna.
5. En la sección **Patrón**, podrá ver varios patrones de la columna como gráficos de barras horizontales. Puede ver los caracteres de los patrones y el número de patrones similares de una columna como un porcentaje si mueve el puntero sobre el gráfico de barras.
6. En la sección **Longitud**, podrá ver la longitud mínima y máxima de los valores de la columna.
7. En la sección **Valor**, podrá ver los valores mínimo y máximo de una columna.
8. En la sección **Tipo de datos**, podrá ver todos los tipos de datos deducidos y todos los tipos de datos documentados de una columna si mueve el cursor sobre los valores.
9. En la sección **Dominio de datos**, podrá ver todos los tipos de dominio deducidos de una columna si mueve el cursor sobre los valores.
10. Para ver los datos atípicos, haga clic en **Acciones > Detectar valor atípico** para detectar valores atípicos en los resultados de perfil.
11. Haga clic en los filtros **Valor atípico del patrón** o **Valor atípico de frecuencia del valor** para ver los valores atípicos en los resultados del perfil.
12. Haga clic en un nombre de columna para ver los resultados de perfil de dicha columna en la vista detallada.

Tarea 3. Ver los resultados del perfil en una vista detallada

En esta tarea, utilizará la vista detallada del perfil LA_Customers para obtener un resumen rápido de una columna en los resultados del perfil.

1. Haga clic en un nombre de columna en la vista de resumen para ver los resultados de perfil de dicha columna en la vista detallada.
Los resultados del perfil para la columna se muestran en la vista detallada.
2. En el panel **General**, podrá ver los valores nulos, únicos y no únicos como un gráfico de barras horizontales y expresados en porcentajes.
3. En el panel **Valores**, podrá ver todos los valores del panel superior. Cuando mueva el control deslizante sobre los valores del panel superior, todos los valores comprendidos en el control se mostrarán en el panel inferior.

4. En el panel **Vista previa de datos**, haga clic en **Acciones > Mostrar vista previa** para ver los datos de origen.
5. En el panel **Patrones**, podrá ver los patrones deducidos en la columna como un gráfico de barras horizontales y expresados en porcentajes.
6. En el panel **Tipo de datos**, podrá ver los tipos de datos deducidos en la columna como un gráfico de barras horizontales y expresados en porcentajes.
7. En el panel **Dominio de datos**, podrá ver los dominios de datos deducidos.

Resumen de la creación de perfiles predeterminados

En esta lección, ha aprendido que un perfil predeterminado muestra resultados de un perfil sin tener que configurar todas las opciones de muestreo de columnas y filas para un perfil. Ha aprendido que puede crear y ejecutar un perfil predeterminado para analizar la calidad de los datos cuando inicie un proyecto de calidad de datos. También ha aprendido que la Herramienta del analista crea perfiles en el archivo sin formato activo del objeto de datos de archivo sin formato.

Ha creado un perfil predeterminado y ha analizado los resultados del perfil. Ha obtenido más información sobre las columnas del perfil, incluidos los valores nulos y los tipos de dato. También ha utilizado los valores y patrones de la columna para identificar problemas de calidad de datos.

Después de analizar los resultados de un perfil rápido, puede realizar las siguientes tareas:

- Crear un perfil personalizado para excluir columnas del perfil e incluir solamente las columnas en las que esté interesado.
- Crear una regla de expresión para crear columnas virtuales y crear un perfil para las mismas.
- Crear una tabla de referencias para incluir los valores de una columna.

CAPÍTULO 5

Lección 4. Creación de perfiles personalizados

Este capítulo incluye los siguientes temas:

- [Introducción a la creación de perfiles personalizados, 29](#)
- [Tarea 1. Crear un perfil personalizado, 30](#)
- [Tarea 2. Ejecutar el perfil, 31](#)
- [Tarea 3. Obtener detalles de los resultados del perfil, 31](#)
- [Resumen de la creación de perfiles personalizados, 32](#)

Introducción a la creación de perfiles personalizados

Un perfil es el análisis de la calidad de datos basado en el contenido y la estructura de los datos. Un perfil personalizado es un perfil que crea cuando desea configurar las columnas, las opciones de muestreo y de obtención de detalles para una creación de perfiles más rápida. Configure las opciones de muestreo para seleccionar las filas de muestra en el origen. Configure las opciones de obtención de detalles para obtener detalles de las filas de datos en los datos de origen o provisionales. Puede optar por ejecutar el perfil en un entorno nativo o de Hadoop.

Un perfil se crea y ejecuta para analizar la calidad de los datos después de iniciar un proyecto de calidad de datos. Al crear un objeto de perfil, puede comenzar seleccionando el objeto de datos y las columnas del objeto de datos en los que desea ejecutar un perfil.

Ejemplo

HypoStores necesita incorporar a su almacén de datos los datos de la recién adquirida oficina de Los Ángeles. HypoStores desea acceder a la calidad de los datos de niveles de clientes del archivo de datos de clientes de Los Ángeles. Usted es el analista responsable de evaluar la calidad de los datos y de transmitir la información al desarrollador responsable de la limpieza de los datos.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

1. Crear un perfil personalizado para el objeto de datos de archivo sin formato y excluir las columnas con valores nulos.
2. Ejecutar el perfil para analizar el contenido y estructura de la columna CustomerTier.
3. Obtener detalles de las filas para los resultados del perfil.

Requisitos previos

Antes de empezar con esta lección, verifique los siguientes requisitos previos:

- Ha terminado las Lecciones 1, 2 y 3 de esta guía de autoaprendizaje.

Duración

Destine de 5 a 10 minutos para realizar esta lección.

Tarea 1. Crear un perfil personalizado

En esta tarea, creará un perfil personalizado. Para crear un perfil personalizado, seleccione el objeto de datos y las columnas en los que desee ejecutar un perfil. También configura las opciones de muestreo y de obtención de detalles.

1. Haga clic en **Nuevo > Perfil**.
Aparecerá el asistente **Nuevo perfil**.
2. La opción **Origen único** está seleccionada de forma predeterminada. Haga clic en **Siguiente**.
3. En la pantalla **Especificar propiedades generales**, defina las siguientes opciones:
 - En el campo Nombre, introduzca **Profile_LA_Customers**.
 - En el campo Ubicación, seleccione la carpeta **Customers**.
4. Haga clic en **Siguiente**.
5. En la pantalla **Seleccionar origen**, haga clic en **Seleccionar**.
Se muestra el cuadro de diálogo **Elegir objeto de datos**.
6. En el cuadro de diálogo **Elegir objeto de datos**, seleccione **LA_Customers**. Haga clic en **Aceptar**.
7. En la pantalla **Seleccionar origen**, desactive las columnas Address2, Address3 y City2.
8. Haga clic en **Siguiente**.
9. En la pantalla **Especificar la configuración**, defina las opciones siguientes:
 - **Ejecutar perfil de columna**.
 - Seleccione la opción **Muestra aleatoria** en el panel **Ejecutar perfil en**.
 - Seleccione la opción **Excluir de la inferencia los tipos y dominios de datos aprobados en las siguientes ejecuciones del perfil**.
 - Seleccione la opción **Provisional** en el panel **Obtención de detalles**.
10. Haga clic en **Siguiente**.
11. En la pantalla **Especificar reglas y filtros**, haga clic en **Guardar y finalizar** para crear y ejecutar el perfil.
La Herramienta del analista crea el perfil y lo muestra en el espacio de trabajo **Detección**. Tiene que ejecutar el perfil para ver los resultados.

Tarea 2. Ejecutar el perfil

En esta tarea, se ejecuta un perfil para crear perfiles en el objeto de datos y mostrar los resultados de los perfiles. La Herramienta del analista crea los perfiles en el archivo sin formato temporal para el objeto de datos del archivo sin formato.

1. Asegúrese de que está en el Espacio de trabajo **Detección**.
Puede ver el perfil **Profile_LA_Customers_Custom** en el espacio de trabajo.
2. Haga clic en **Profile_LA_Customers_Custom** en el espacio de trabajo.
3. Aparece la pantalla del perfil, donde puede optar por modificar el perfil o ejecutarlo. Haga clic en **Ejecutar**.
4. Los resultados del perfil aparecen en la vista de resumen.

Tarea 3. Obtener detalles de los resultados del perfil

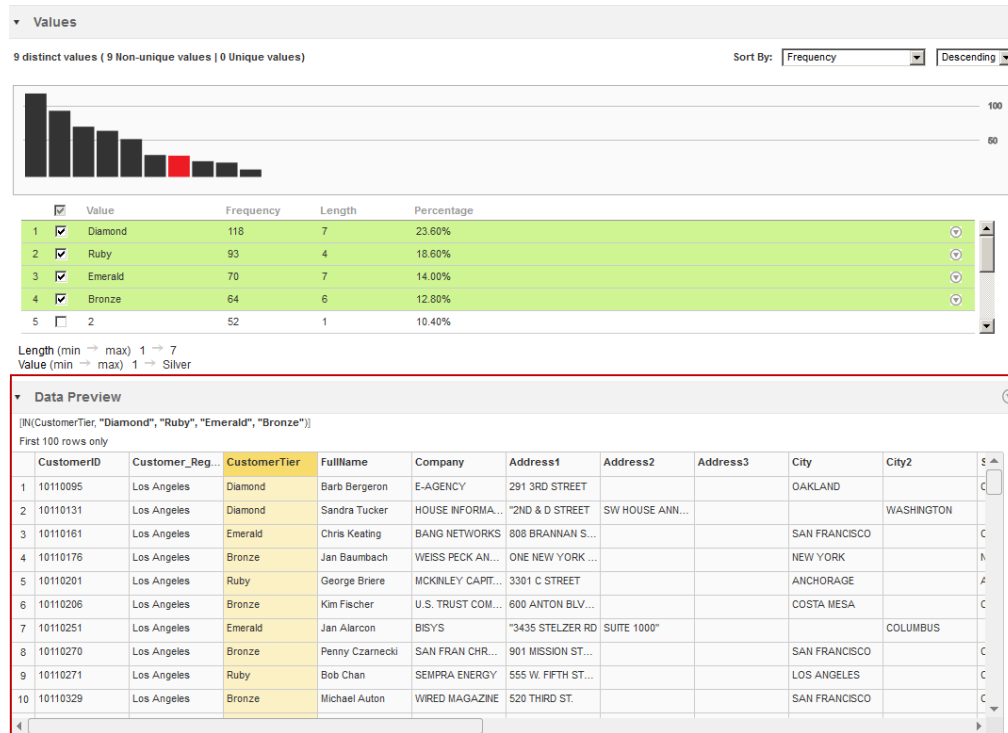
En esta tarea, se obtienen detalles de los valores de la columna CustomerTier para ver las filas de origen en el objeto de datos para el perfil.

1. Compruebe que está en la vista de resumen de los resultados del perfil **Profile_LA_Customers**.
2. Haga clic en la columna **CustomerTier**.
Los resultados de perfil de la columna se mostrarán en la vista detallada.

3. En la vista detallada, seleccione los valores Diamante, Rubí, Esmeralda y Bronce. Haga clic con el botón derecho en los valores del panel **Valores** y seleccione **Obtener detalles**.

Las filas de la columna que tienen el valor Diamante, Rubí, Esmeralda o Bronce aparecen en el panel **Vista previa de datos**.

La siguiente imagen muestra los resultados de obtención de detalles en el panel **Vista previa de datos** de los valores Diamante, Rubí, Esmeralda o Bronce:



El panel **Vista previa de datos** muestra las 100 primeras filas de la columna seleccionada. El título del panel **Vista previa de datos** muestra la lógica utilizada para la columna de origen.

Resumen de la creación de perfiles personalizados

En esta lección, ha aprendido que puede configurar las columnas para las que se crea un perfil y que se pueden configurar las opciones de muestreo y de obtención de detalles. Ha aprendido que puede obtener detalles para ver las filas subyacentes de valores de columna y que puede configurar las columnas que se incluyen en el momento de visualizar los valores de columna.

Ha creado un perfil personalizado que incluía la columna CustomerTier columna, ha ejecutado el perfil y obtenido detalles de las filas subyacentes para la columna CustomerTier en los resultados.

Utilice el objeto de perfil personalizado para crear una regla de expresión en la lección 5.

CAPÍTULO 6

Lección 5. Creación de reglas de expresión

Este capítulo incluye los siguientes temas:

- [Introducción a la creación de reglas de expresión, 33](#)
- [Tarea 1. Crear reglas de expresión y ejecutar el perfil, 34](#)
- [Tarea 2. Ver la salida de la regla de expresión, 34](#)
- [Tarea 3. Editar las reglas de expresión, 35](#)
- [Resumen de la creación de reglas de expresión, 35](#)

Introducción a la creación de reglas de expresión

Las reglas de expresión emplean funciones de expresión y columnas de origen para definir la lógica de las reglas. Puede crear reglas de expresión y añadirlas a un perfil en la Herramienta del analista. Puede asociar una regla de expresión a uno o más perfiles.

La salida de una regla de expresión es una columna virtual del perfil. La Herramienta del analista perfila la columna virtual cuando se ejecuta el perfil.

Puede usar reglas de expresión para validar columnas origen o crear columnas origen adicionales basadas en el valor de las columnas origen.

Ejemplo

HypoStores desea incorporar a su almacén de datos los datos de la recién adquirida oficina de Los Ángeles. HypoStores desea analizar los nombres de los clientes y dividirlos en nombre y apellido. HypoStores desea utilizar reglas de expresión para analizar una columna que contenga nombres y apellidos en columnas virtuales diferenciadas y luego perfilar las columnas. HypoStores también quiere poner las reglas a disposición de otros analistas que necesiten analizar la salida de estas reglas.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

1. Crear reglas de expresión para separar la columna FullName en una columna para el nombre y otra para el apellido. Crear una regla para separar el nombre del nombre completo. Crear otra regla para separar el apellido del nombre completo. Estas reglas se crean para el perfil Profile_LA_Customers.
2. Ejecutar el perfil y visualizar la salida de las reglas en el perfil.

3. Editar las reglas para que las puedan usar otros usuarios de la Herramienta del analista.

Requisitos previos

Antes de empezar con esta lección, verifique los siguientes requisitos previos:

- Ha finalizado las lecciones 1 a 4.

Duración

Destine de 10 a 15 minutos para realizar esta lección.

Tarea 1. Crear reglas de expresión y ejecutar el perfil

En esta tarea, se crean dos reglas de expresión para analizar la columna FullName en dos columnas virtuales llamadas FirstName y LastName. Los nombres de la regla son FirstName y LastName.

1. En el espacio de trabajo **Biblioteca**, haga clic en el perfil **Profile_LA_Customers**.
Los resultados de perfil aparecen en la vista de resumen.
2. Haga clic en **Editar**.
Aparecerá el asistente **Perfil**.
3. Haga clic en la pantalla **Especificar reglas y filtros**.
4. En el panel **Reglas**, haga clic en **Acciones > Crear regla**.
5. En el campo **Nombre**, introduzca **FirstName**.
6. En la sección **Expresión**, introduzca la siguiente expresión para separar el nombre de la columna **Nombre**:

```
SUBSTR(FullName,1, INSTR(FullName, ' ', -1, 1) - 1)
```
7. Haga clic en **Validar**.
8. Haga clic en **Aceptar**.
9. Repita los pasos 4 a 8 para crear una regla denominada **LastName**. Introduzca la siguiente expresión para separar el apellido de la columna **Nombre**:

```
SUBSTR(FullName, INSTR(FullName, ' ', -1, 1), LENGTH(FullName))
```
10. Haga clic en **Guardar y ejecutar** para guardar y ejecutar el perfil.

Tarea 2. Ver la salida de la regla de expresión

En esta tarea, se visualiza la salida de reglas de expresión que separan nombres y apellidos después de ejecutar un perfil.

1. En la vista de resumen, haga clic en **Editar**.
Aparecerá el asistente Perfil.
2. En el asistente Perfil, haga clic en **Seleccionar origen**.
3. En la pantalla **Seleccionar origen**, marque la casilla junto a **Nombre** en la barra de herramientas para borrar todas las columnas.

Una de las columnas está seleccionada de forma predeterminada porque es necesario seleccionar al menos una columna en la sección **Columnas**.

4. Seleccione la columna **FullName**, y las reglas **FirstName** y **LastName**.

5. Borre cualquier otra columna que se haya seleccionado.

6. Haga clic en **Guardar y ejecutar**.

Los resultados de perfil aparecen en la vista de resumen.

7. Haga clic en la regla **FirstName**. Los resultados de perfil de la regla aparecerán en la vista detallada.

8. Seleccione un valor en el panel **Valores**. Haga clic con el botón derecho en el valor y seleccione **Obtención de detalles**.

Los valores para la columna **FullName**, así como las reglas **FirstName** y **LastName** aparecerán en el panel **Vista previa de datos** junto con otros valores de columna. Tenga en cuenta que la Herramienta del analista divide la columna **FullName** en nombre y apellido.

Tarea 3. Editar las reglas de expresión

En esta tarea, define que las reglas de expresión se puedan reutilizar y que estén disponibles para todos los usuarios de la Herramienta del analista.

1. En la vista de resumen del perfil **Profile_LA_Customers**, haga clic en **Editar**.

Aparecerá el asistente Perfil.

2. Haga clic en la pantalla **Especificar reglas y filtros**.

3. En la pantalla **Especificar reglas y filtros**, seleccione la regla **FirstName** y haga clic en **Acciones > Editar regla**.

A continuación, aparecerá **Editar regla**.

4. Seleccione la opción **¿Desea guardar esta regla como una regla reutilizable?** y elija una ubicación para guardar la regla.

5. Haga clic en **Aceptar**.

6. Seleccione la regla **LastName** y repita los pasos 3 a 5.

7. Haga clic **Guardar y finalizar** para guardar el perfil.

Cualquier usuario de la Herramienta del analista puede utilizar las reglas **FirstName** y **LastName** para dividir una columna con nombres y apellidos en columnas diferenciadas.

Resumen de la creación de reglas de expresión

En esta lección, ha aprendido que las reglas de expresión utilizan funciones de expresión y columnas de origen para definir una lógica de regla. Ha aprendido que la salida de una regla de expresión es una columna virtual del perfil. La herramienta Analyst incluye la columna virtual cuando se ejecuta el perfil.

Ha creado dos reglas de expresión, las ha añadido a un perfil y ha ejecutado el perfil. Ha visualizado la salida de las reglas y las ha puesto a disposición de todos los usuarios de la herramienta Analyst.

CAPÍTULO 7

Lección 6. Creación y ejecución de cuadros de mando

Este capítulo incluye los siguientes temas:

- [Introducción a la creación y ejecución de cuadros de mando, 36](#)
- [Tarea 1. Crear un cuadro de mando a partir de los resultados de un perfil, 37](#)
- [Tarea 2. Ejecutar el cuadro de mandos, 38](#)
- [Tarea 3. Ver el cuadro de mandos, 39](#)
- [Tarea 4. Editar el cuadro de mandos, 39](#)
- [Tarea 5. Configurar umbrales, 40](#)
- [Tarea 6. Ver gráficos de tendencias de puntuación, 40](#)
- [Creación y ejecución de cuadros de mando, 41](#)

Introducción a la creación y ejecución de cuadros de mando

Un cuadro de mandos es una representación gráfica de valores válidos para una columna o la salida de una regla en los resultados del perfil. Utilice cuadros de mandos para medir y controlar el progreso de la calidad de los datos a lo largo del tiempo.

Para crear un cuadro de mandos, añada columnas desde el perfil a un cuadro de mandos como mediciones, asigne ponderaciones a las mediciones y configure los umbrales de puntuación. Puede agregar filtros a los cuadros de mandos que se basen en los datos de origen. Para ejecutar un cuadro de mandos, seleccione los valores válidos de la medición y ejecute el cuadro de mandos para ver las puntuaciones de las mediciones.

Los cuadros de mandos muestran la frecuencia de valores de las columnas de un perfil en forma de puntuaciones. Las puntuaciones reflejan el porcentaje de valores válidos de una medición.

Historial

HypoStores desea incorporar a su almacén de datos los datos de la recién adquirida oficina de Los Ángeles. Antes de que la organización fusione los datos, desea comprobar que se analiza la calidad de los datos en los distintos niveles de clientes y estados. Usted es el analista responsable de supervisar el progreso de los análisis de calidad de los datos. Desea crear un cuadro de mandos a partir de las columnas de perfil de nivel de cliente y estado y visualizar gráficos de tendencia de puntuación para determinar hasta qué punto mejoran las puntuaciones a lo largo del tiempo.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

1. Crear un cuadro de mandos a partir de los resultados del perfil Profile_LA_Customers_Custom para visualizar las puntuaciones correspondientes a las columnas CustomerTier y State.
2. Ejecutar el cuadro de mandos para generar las puntuaciones de las columnas CustomerTier y State.
3. Visualizar el cuadro de mandos para ver las puntuaciones de cada columna.
4. Editar el cuadro de mandos para especificar distintos valores válidos para las puntuaciones.
5. Configurar umbrales de puntuación y ejecutar el cuadro de mandos.
6. Visualizar gráficos de tendencia de puntuación para determinar la mejora de las puntuaciones en el transcurso del tiempo.

Requisitos previos

Antes de empezar con esta lección, verifique los siguientes requisitos previos:

- Ha terminado las Lecciones 1 y 5 de esta guía de autoaprendizaje.

Duración

Dedique 15 minutos a completar las tareas de esta lección.

Tarea 1. Crear un cuadro de mando a partir de los resultados de un perfil

En esta tarea, creará un cuadro de mando a partir del perfil Profile_LA_Customers_Custom para puntuar los valores de las columnas CustomerTier y State.

1. En el espacio de trabajo **Biblioteca**, haga clic en el perfil **Profile_LA_Customers**.
Aparecerá la vista de resumen de los resultados de perfil.
2. En la vista de resumen, seleccione la columna **CustomerTier** y haga clic con el botón derecho en la columna; a continuación, seleccione **Añadir a > Cuadro de mando**.
Aparece el asistente **Añadir a cuadro de mando**.
3. En el asistente **Añadir a cuadro de mando**, la opción **Nuevo cuadro de mando** está seleccionada de forma predeterminada. Haga clic en **Siguiente**.
4. En la pantalla **Paso 2 de 8**, introduzca **sc_LA_Customer** como nombre del cuadro de mando y navegue hasta la carpeta **Clientes** para definirla como ubicación del cuadro de mando.
5. Haga clic en **Siguiente**.
6. En la pantalla **Paso 3 de 8**, seleccione las columnas **CustomerTier** y **Estado** para añadirlas al cuadro de mando.
7. Haga clic en **Siguiente**.
8. En la pantalla **Paso 4 de 8** puede crear, editar o eliminar filtros para las mediciones. En esta guía de autoaprendizaje, no vamos a crear un filtro de cuadros de mando. Haga clic en **Siguiente**.
9. En la pantalla **Paso 5 de 8**, seleccione la medición **CustomerTier** en el panel **Mediciones**.
10. En el panel **Puntuación con: valores**, seleccione todos los valores y haga clic en el botón **Añadir todo** para mover los valores a la sección **Valores válidos**.

Utilice la tecla **Mayús** para seleccionar varios valores.

11. En el panel **Mediciones**, seleccione la medición **Estado** y aquellos valores que tengan un código de estado de dos letras en la sección **Puntuación con: valores**.
12. Haga clic en el botón **Añadir** para mover los valores a la sección **Valores válidos**.
Puede ver el número total de los valores válidos y el porcentaje de valores válidos en la parte superior de la sección.
13. En la sección **Mediciones** de cada medición, acepte la configuración predeterminada de los umbrales de puntuación en la sección **Umbrales de medición**.
14. Haga clic en **Siguiente**.
15. En la pantalla **Paso 6 de 8**, puede seleccionar un grupo de mediciones para añadirlo a las mediciones de forma opcional. De forma predeterminada, la Herramienta del analista añade las mediciones al grupo de mediciones **Predeterminado**.
16. Haga clic en **Siguiente**.
17. En la pantalla **Paso 7 de 8**, haga doble clic en la columna **Ponderación** de la medición **CustomerTier** del panel **Predeterminado - Mediciones**.
Al ejecutar un cuadro de mando, la Herramienta del analista calcula la media ponderada de cada grupo de mediciones según la puntuación y ponderación asignadas a cada medición.
18. Especifique la ponderación de las mediciones **CustomerTier** y **State**.
19. Haga clic en **Siguiente**.
20. En la pantalla **Paso 8 de 8**, la opción **Nativo** está seleccionada de forma predeterminada. Haga clic en **Guardar** para crear el cuadro de mando.
El cuadro de mando aparece en el espacio de trabajo **Cuadros de mando**.

Tarea 2. Ejecutar el cuadro de mandos

En esta tarea, ejecutará el cuadro de mandos `sc_LA_Customer` para generar las puntuaciones de las columnas **CustomerTier** y **State**.

1. En el espacio de trabajo **Biblioteca**, haga clic en **Activos > Cuadros de mandos**.
2. Seleccione un cuadro de mandos del panel **Cuadro de mandos**.
3. Haga clic en **Acciones > Abrir**.
Aparecerá el cuadro de mandos.
4. Haga clic en **Acciones > Ejecutar cuadro de mandos**.
Aparece el cuadro de diálogo **Ejecutar cuadro de mandos**.
5. Compruebe los valores del cuadro de diálogo y haga clic en **Ejecutar**.
El espacio de trabajo **Cuadros de mando** muestra las puntuaciones de las columnas **CustomerTier** y **State**.

Tarea 3. Ver el cuadro de mandos

En esta tarea, visualizará el cuadro de mandos **sc_LA_Customer** para ver las puntuaciones de las columnas **CustomerTier** y **State**.

1. Seleccione la fila **Estado** que contenga la puntuación Estado que desee visualizar.

En la sección **sc_LA_Customer - mediciones**, podrá ver las siguientes propiedades del cuadro de mandos:

- Nombre del cuadro de mandos.
- Número total de filas en el cuadro de mandos.
- Número de filas no válidas.
- La puntuación junto con un gráfico de barras horizontales.
- Tendencia de la puntuación. Puede hacer clic en la tendencia de la puntuación para ver una representación gráfica en la pantalla **Detalle del gráfico de tendencias**.
- Ponderación de la medición.
- Coste de los datos no válidos.
- Tendencia de coste.
- Objeto de datos. Haga clic en el objeto de datos para acceder a la vista previa de datos del objeto de datos en el espacio de trabajo **Detección**.
- Nombre de la regla o la columna.
- Tipo de origen.
- Icono de obtención de detalles.

2. Haga clic en el icono Obtención de detalles de la fila **State**.

Las puntuaciones no válidas de la columna **State** aparecerán en la sección **Filas no válidas** del panel **Obtención de detalles**.

3. Seleccione **Filas válidas** para ver las puntuaciones que sean válidas para la columna **Estado**.

4. Haga clic en el icono Obtención de detalles de la fila **CustomerTier**.

Todas las puntuaciones de la columna **CustomerTier** son válidas.

Tarea 4. Editar el cuadro de mandos

En esta tarea, editará el cuadro de mandos **sc_LA_Customer** para especificar el valor **Rubí** como no válido para la puntuación **CustomerTier**.

1. Compruebe que se encuentra en el espacio de trabajo **Cuadro de mandos** y que está abierto el cuadro de mandos **sc_LA_Customer**.
2. Seleccione **Acciones > Editar > Mediciones**.
Aparece el cuadro de diálogo **Editar cuadro de mandos**.
3. En la sección **Mediciones**, seleccione **CustomerTier**.
4. En la sección **Puntuación con: valores**, mueva **Rubí** desde la sección **Valores válidos** hasta la sección **Valores disponibles**.

Acepte la configuración predeterminada que aparece en la sección **Umbrales de medición**.

5. Haga clic en **Guardar y ejecutar** para guardar los cambios realizados en el cuadro de mandos y ejecutarlo.
6. Vuelva a visualizar la puntuación CustomerTier.
La puntuación CustomerTier cambia al porcentaje 81,4.

Tarea 5. Configurar umbrales

En esta tarea, configurará los umbrales de la puntuación State del cuadro de mandos sc_LA_Customer para determinar los rangos válidos de los datos de la columna State. Los valores con un código de dos letras, como CA, son aceptables, y los códigos con más de dos letras, como Calif, no son aceptables.

1. Compruebe que se encuentra en el espacio de trabajo **Cuadro de mandos** y que está abierto el cuadro de mandos sc_LA_Customer.
2. Seleccione **Acciones > Editar > Mediciones**.
Aparece el cuadro de diálogo **Editar cuadro de mandos**.
3. En la sección **Mediciones**, seleccione **Estado**.
4. En la sección **Umbrales de medición**, introduzca los siguientes rangos para las puntuaciones Buena e Inaceptable: del 90 al 100%, buena; del 0 al 50%, inaceptable y entre el 51 y el 89%, aceptable.
Los umbrales representan los límites inferiores de los rangos aceptable y bueno.
5. Haga clic en **Guardar y ejecutar** para guardar los cambios realizados en el cuadro de mandos y ejecutarlo.
En el panel **Cuadro de mandos**, visualice los cambios en el porcentaje de puntuación y la puntuación representada en forma de barra para la puntuación State.

Tarea 6. Ver gráficos de tendencias de puntuación

En esta tarea, se visualiza el gráfico de tendencias para la puntuación Estado. Puede visualizar gráficos de tendencias para evaluar puntuaciones en el transcurso del tiempo.

1. Compruebe que se encuentra en el espacio de trabajo **Cuadro de mandos** y que está abierto el cuadro de mandos sc_LA_Customer.
2. Seleccione la fila **State**.
3. Haga clic en **Acciones > Mostrar gráfico de tendencias** o haga clic en la flecha situada bajo la columna **Tendencia de puntuación**.
Aparecerá el cuadro de diálogo **Detalle del gráfico de tendencias**. Puede visualizar los umbrales **Bueno**, **Aceptable** e **Inaceptable** para la puntuación. Los umbrales cambian cada vez que se ejecuta el cuadro de mandos después de editar los valores de las puntuaciones en el cuadro de mandos.
4. Señale a cualquier círculo del gráfico para ver los valores válidos en la sección **Valores válidos** en la parte inferior del gráfico.
5. Haga clic en **Cerrar** para volver al cuadro de mandos.

Creación y ejecución de cuadros de mando

En esta lección, ha aprendido que puede crear un cuadro de mando a partir de los resultados de un perfil. Un cuadro de mando contiene las columnas obtenidas a partir de un perfil. Ha aprendido que puede ejecutar un cuadro de mando para generar puntuaciones para columnas. Ha editado un cuadro de mando para configurar valores válidos y establecer umbrales para puntuaciones. También ha aprendido a visualizar el gráfico de tendencia de puntuación.

Ha creado un cuadro de mando a partir de las columnas CustomerTier y State de un perfil para analizar la calidad de los datos de estas columnas. Ha ejecutado el cuadro de mando para generar puntuaciones para cada columna. Ha editado el cuadro de mando para especificar distintos valores válidos para las puntuaciones. Ha configurado umbrales para una puntuación y ha visualizado el gráfico de tendencia de puntuación.

CAPÍTULO 8

Lección 7. Creación de tablas de referencia a partir de columnas de perfil

Este capítulo incluye los siguientes temas:

- [Introducción a la creación de tablas de referencia a partir de columnas de perfil, 42](#)
- [Tarea 1. Crear una tabla de referencia a partir de columnas de perfil, 43](#)
- [Tarea 2. Editar la tabla de referencia, 44](#)
- [Resumen de la creación de tablas de referencia a partir de columnas de perfil, 44](#)

Introducción a la creación de tablas de referencia a partir de columnas de perfil

Una tabla de referencia contiene datos de referencia que puede utilizar para estandarizar datos de origen. Los datos de referencia pueden incluir valores válidos y estándar. Cree tablas de referencia para establecer relaciones entre los valores de datos de origen y los valores válidos y estándar.

Puede crear una tabla de referencia a partir de los resultados de un perfil. Tras crear una tabla de referencia, puede editar la tabla de referencia para añadir columnas o filas y para añadir o editar valores estándar o válidos. Puede ver los cambios realizados en una tabla de referencia en una traza de auditoría.

Historial

HypoStores desea crear un perfil de los datos para detectar anomalías y estandarizarlos con valores válidos. Usted es el analista responsable de estandarizar los valores válidos en los datos. Desea crear una tabla de referencia basada en valores válidos a partir de las columnas de un perfil.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

1. Crear una tabla de referencias a partir de la columna CustomerTier del perfil Profile_LA_Customers_Custom seleccionando valores válidos para las columnas.
2. Editar la tabla de referencia para configurar distintos valores válidos para las columnas.

Requisitos previos

Antes de empezar con esta lección, verifique los siguientes requisitos previos:

- Ha terminado las Lecciones 1 y 6 de esta guía de autoaprendizaje.

Duración

Dedique 15 minutos a completar las tareas de esta lección.

Tarea 1. Crear una tabla de referencia a partir de columnas de perfil

En esta tarea, se crea una tabla de referencia y se añade la columna CustomerTier desde el perfil Profile_LA_Customers a la tabla de referencia.

1. En el espacio de trabajo **Biblioteca**, haga clic en **Activos > Perfiles**.
2. Haga clic en el perfil **Profile_LA_Customers** para abrir los resultados de perfil en la vista de resumen.
3. En la vista de resumen, seleccione la columna CustomerTier que desee añadir a la tabla de referencia. Haga clic con el botón derecho y seleccione **Añadir a tabla de referencia**.

Se muestra el cuadro de diálogo **Agregar a la tabla de referencia**.

4. Seleccione **Crear una tabla de referencia**.
5. Haga clic en **Siguiente**.
6. En el campo **Nombre**, introduzca **Reftab_CustTier_HypoStores**.
7. Introduzca una descripción y establezca 0 como valor predeterminado.

La Herramienta del analista utiliza el valor predeterminado para cualquier registro de tabla que no contenga ningún valor.

8. Haga clic en **Siguiente**.
9. En la sección **Atributos de columna**, configure las siguientes propiedades de la columna CustomerTier:

Propiedad	Descripción
Nombre	CustomerTier
Tipo de datos	Cadena
Precisión	10
Descripción	Valores de niveles de cliente de referencia

10. Si lo desea, puede crear una columna de descripción para las filas de la tabla de referencia. Introduzca el nombre y la precisión para la columna.
11. Compruebe los valores de la columna CustomerTier de la sección **Vista previa**.
12. Haga clic en **Siguiente**.

Aparece el nombre de tabla de referencia Reftab_CustomerTier_HypoStores. Puede introducir una descripción opcional.

13. En la sección **Guardar en**, seleccione el proyecto de guía de autoaprendizaje donde desee crear la tabla de referencia.
En el panel **Tablas de referencia**: aparece una lista de las tablas de referencia existentes en la ubicación seleccionada.
14. Introduzca una nota de auditoría opcional.
15. Haga clic en **Finalizar**.

Tarea 2. Editar la tabla de referencia

En esta tarea, se edita la tabla Reftab_CustomerTier_HypoStores para añadir valores alternativos para los niveles de cliente.

1. En el espacio de trabajo **Biblioteca**, haga clic en **Activos > Tablas de referencia**.
2. Haga clic en la tabla de referencia Reftab_CustomerTier_HypoStores.
La tabla de referencia se abre en el espacio de trabajo **Diseño**.
3. Para editar una fila, selecciónela y haga clic en **Acciones > Editar** o en el icono **Editar**.
A continuación, aparecerá **Editar fila**. Opcionalmente, seleccione múltiples filas para agregar el mismo valor alternativo a cada fila.
4. Introduzca los siguientes valores alternativos para las filas Diamante, Esmeralda, Oro, Plata y Bronce: 1, 2, 3, 4, 5.
Introduzca una nota de auditoría opcional.
5. Haga clic en **Aplicar** para aplicar los cambios.
6. Haga clic en **Cerrar**
Los valores modificados de la tabla de referencia aparecen en el espacio de trabajo **Diseño**.

Resumen de la creación de tablas de referencia a partir de columnas de perfil

En esta lección, ha aprendido a crear tablas de referencia a partir de los resultados de un perfil para configurar valores válidos para datos de origen.

Ha creado una tabla de referencia a partir de una columna de perfil seleccionando valores válidos para columnas. Ha editado la tabla de referencia para configurar distintos valores válidos para las columnas.

CAPÍTULO 9

Lección 8. Creación de tablas de referencia

Este capítulo incluye los siguientes temas:

- [Introducción a la creación de tablas de referencia, 45](#)
- [Tarea 1. Crear una tabla de referencia, 46](#)
- [Resumen de la creación de tablas de referencia, 46](#)

Introducción a la creación de tablas de referencia

Una tabla de referencia contiene datos de referencia que puede utilizar para estandarizar datos de origen. Los datos de referencia pueden incluir valores válidos y estándar. Cree tablas de referencia para establecer relaciones entre los valores de datos de origen y los valores válidos y estándar.

Se puede crear una tabla de referencia manualmente mediante el editor de tablas de referencia. Utilice la tabla de referencia para definir y estandarizar los datos de origen. Puede compartir la tabla de referencia con un desarrollador para utilizarla en transformaciones de búsqueda o de estandarizador en Developer tool.

Historial

HypoStores desea estandarizar datos para que tengan valores válidos. Usted es el analista responsable de estandarizar los valores válidos en los datos. Desea crear una tabla de referencia para definir códigos de nivel de cliente estándar que hagan referencia a los datos de clientes de Los Ángeles. Más adelante, podrá compartir la tabla de referencia con un desarrollador.

Objetivos

En esta lección, llevará a cabo la tarea siguiente:

- Crear una tabla de referencia con el editor de tablas de referencia para definir códigos de nivel de cliente estándar que hagan referencia a los datos de clientes de Los Ángeles.

Requisitos previos

Antes de empezar con esta lección, verifique los siguientes requisitos previos:

- Ha terminado las Lecciones 1 y 2 de esta guía de autoaprendizaje.

Duración

Dedique 10 minutos a completar las tareas de esta lección.

Tarea 1. Crear una tabla de referencia

En esta tarea, creará la tabla de referencia Reftab_CustomerTier_Codes para estandarizar los valores válidos del los datos del nivel de clientes.

1. Haga clic en **Nueva > Tabla de referencia**.
Aparece el asistente **Nueva tabla de referencia**.
2. Seleccione **Usar el editor de tablas de referencia**.
3. Haga clic en **Siguiente**.
4. Para cada columna que desee incluir en la tabla de referencia, haga clic en el icono **Añadir columna nueva** y configure la propiedades para cada columna.
Añada los siguientes nombres de columna: CustomerID, CustomerTier y Status. Puede reordenar las columnas o borrarlas.
5. Introduzca una descripción opcional y establezca el valor predeterminado como **0**.
La herramienta del analista utiliza el valor predeterminado para cualquier registro de tabla que no contenga ningún valor.
6. Haga clic en **Siguiente**.
7. En el campo **Nombre**, introduzca **Reftab_CustomerTier_Codes**.
8. En la sección **Carpetas**, seleccione la carpeta **Clientes** en el proyecto de la guía de autoaprendizaje.
9. Haga clic en **Finalizar**.
La tabla de referencia aparece en el espacio de trabajo **Diseño**.
10. En el menú **Acciones**, seleccione **Añadir fila** para llenar cada columna de tabla de referencia con los siguientes cuatro valores:
CustomerID = LA1, LA2, LA3, LA4
CustomerTier = 1, 2, 3, 4.
Estado= Activo, Inactivo

Resumen de la creación de tablas de referencia

En esta lección, ha aprendido a crear tablas de referencia mediante el editor de tablas de referencia para crear valores estándar válidos y usarlos con datos de origen.

Ha creado una tabla de referencia mediante el editor de tablas de referencia para estandarizar los valores de niveles de cliente de los datos de clientes de Los Ángeles.

Parte II: Primeros pasos con Informatica Developer

Esta parte incluye los siguientes capítulos:

- [Lección 1. Configuración de Informatica Developer, 48](#)
- [Lección 2: Importar objetos de datos físicos, 52](#)
- [Lección 3. Ejecutar un perfil con datos de origen, 61](#)

CAPÍTULO 10

Lección 1. Configuración de Informatica Developer

Este capítulo incluye los siguientes temas:

- [Introducción a la configuración de Informatica Developer, 48](#)
- [Tarea 1. Iniciar Informatica Developer, 49](#)
- [Tarea 2. Añadir un dominio, 49](#)
- [Tarea 3. Añadir un repositorio de modelos, 50](#)
- [Tarea 4. Crear un proyecto, 50](#)
- [Tarea 5. Crear una carpeta, 50](#)
- [Tarea 6. Selección de un servicio de integración de datos predeterminado, 51](#)
- [Resumen de la configuración de Informatica Developer, 51](#)

Introducción a la configuración de Informatica Developer

Antes de empezar con las lecciones de esta guía de autoaprendizaje, debe iniciar y configurar Developer tool. Para ello, tiene que añadir un dominio. Añada un repositorio de modelos que esté en el dominio y cree un proyecto y una carpeta para guardar el trabajo desarrollado. También debe seleccionar un servicio de integración de datos predeterminado si el dominio incluye más de un servicio.

El dominio de Informatica es un conjunto de nodos y servicios que definen el entorno de Informatica. Los servicios del dominio incluyen el servicio de repositorio de modelos y el servicio de integración de datos.

El servicio de repositorio de modelos gestiona el repositorio de modelos. El repositorio de modelos es una base de datos relacional que guarda los metadatos para los proyectos que crea en Developer tool. Un proyecto guarda los objetos creados en Developer tool. Asimismo, un proyecto puede contener carpetas que almacenen objetos relacionados, como objetos que formen parte del mismo requisito de negocio.

El servicio de integración de datos realiza las tareas de integración de datos en Developer tool.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

- Iniciar Developer tool e ir al entorno de trabajo de dicha herramienta.

- Añadir un dominio en Developer tool.
- Añadir un repositorio de modelos para poder crear un proyecto.
- Crear un proyecto para almacenar los objetos creados en Developer tool.
- Crear una carpeta dentro del proyecto donde almacenar objetos relacionados.
- Seleccionar un servicio de integración de datos predeterminado para ejecutar las tareas de integración de datos.

Requisitos previos

Antes de empezar con esta lección, compruebe los siguientes requisitos previos:

- Ha instalado Developer tool.
- Dispone de un nombre de dominio, un nombre de host y un número de puerto para conectarse a un dominio. Un administrador del dominio puede facilitarle esta información.
- Un administrador del dominio ha configurado un servicio de repositorio de modelos en la Herramienta del administrador.
- Tiene un nombre de usuario y una contraseña para acceder al servicio de repositorio de modelos. Un administrador del dominio puede facilitarle esta información.
- Un administrador del dominio ha configurado un servicio de integración de datos.
- El servicio de integración de datos se está ejecutando.

Duración

Dedique de 5 a 10 minutos a completar las tareas de esta lección.

Tarea 1. Iniciar Informatica Developer

Inicie la herramienta Developer para comenzar la guía de autoaprendizaje.

1. Inicie la herramienta Developer.
Aparece la página **Bienvenida** de la herramienta Developer.
2. Haga clic en el botón **Entorno de trabajo**.
Aparece el entorno de trabajo de la herramienta Developer.

Tarea 2. Añadir un dominio

En esta tarea, se añade un dominio en la herramienta Developer para acceder a un repositorio de modelos.

1. Haga clic en **Ventana > Preferencias**.
Aparece el cuadro de diálogo **Preferencias**.
2. Seleccione **Informatica > Dominios**.
3. Haga clic en **Añadir**.
Aparece el cuadro de diálogo **Nuevo dominio**.
4. Escriba el nombre de dominio, el nombre del host y el número de puerto.

5. Haga clic en **Finalizar**.
6. Haga clic en **Aceptar**.

Tarea 3. Añadir un repositorio de modelos

En esta tarea, se añade el repositorio de modelos que se desea utilizar para almacenar proyectos y carpetas.

1. Haga clic en **Archivo > Conectar a repositorio**.
Aparece el cuadro de diálogo **Conectar a repositorio**.
2. Haga clic en **Explorar** para seleccionar un servicio de repositorio de modelos.
3. Haga clic en **Aceptar**.
4. Haga clic en **Siguiente**.
5. Escriba su nombre de usuario y la contraseña.
6. Seleccione un espacio de nombres.
7. Haga clic en **Finalizar**.
El repositorio de modelos aparece en la vista **Explorador de objetos**.

Tarea 4. Crear un proyecto

En esta tarea, se crea un proyecto para almacenar los objetos creados en la herramienta Developer. Puede crear un proyecto para todas las guías de autoaprendizaje de esta guía.

1. En la vista **Explorador de objetos**, seleccione un servicio de repositorio de modelos.
2. Haga clic en **Archivo > Nuevo > Proyecto**.
Aparece el cuadro de diálogo **Nuevo proyecto**.
3. Escriba su nombre precedido de "Tutorial_" como nombre del proyecto.
4. Haga clic en **Finalizar**.
El proyecto aparece en el servicio de repositorio de modelos, en la vista **Explorador de objetos**.

Tarea 5. Crear una carpeta

En esta tarea, se crea una carpeta para almacenar los objetos relacionados. Puede crear una carpeta para todas las guías de autoaprendizaje de esta guía.

1. En la vista **Explorador de objetos**, seleccione el proyecto al que desee añadir la carpeta.
2. Haga clic en **Archivo > Nueva > Carpeta**.
3. Escriba un nombre para la carpeta.

4. Haga clic en **Finalizar**.

La herramienta Developer añade la carpeta en el proyecto, en la vista **Explorador de objetos**. Expanda el proyecto para ver la carpeta.

Tarea 6. Selección de un servicio de integración de datos predeterminado

En esta tarea, se selecciona un servicio de integración de datos predeterminado para poder ejecutar asignaciones y previsualizar datos. Este paso es obligatorio si hay más de un servicio de integración de datos en el dominio. Si el dominio contiene un servicio de integración de datos, este servicio se establece como predeterminado.

1. Haga clic en **Ventana > Preferencias**.
Aparece el cuadro de diálogo **Preferencias**.
2. Seleccione **Informatica > Servicios de integración de datos**.
3. Expanda el dominio.
4. Seleccione un servicio de integración de datos.
5. Haga clic en **Establecer como predeterminado**.
6. Haga clic en **Aceptar**.

Resumen de la configuración de Informatica Developer

En esta lección, ha aprendido que el dominio Informatica incluye el servicio de repositorio de modelos y el servicio de integración de datos. El servicio de repositorio de modelos gestiona el repositorio de modelos. Un repositorio de modelos contiene proyectos y carpetas. El servicio de integración de datos realiza tareas de integración de datos.

Ha iniciado y configurado la herramienta Developer. Ha añadido un dominio a la herramienta Developer, ha añadido un repositorio de modelos y ha creado un proyecto y una carpeta. También ha seleccionado un servicio de integración de datos predeterminado.

Ya puede utilizar la herramienta Analyst para llevar a cabo otras lecciones de esta guía de autoaprendizaje.

CAPÍTULO 11

Lección 2: Importar objetos de datos físicos

Este capítulo incluye los siguientes temas:

- [Introducción a la importación de objetos de datos físicos, 52](#)
- [Tarea 1. Importar el objeto de datos del archivo sin formato Boston_Customers, 53](#)
- [Tarea 2. Importar el objeto de datos de archivo sin formato LA_Customers, 59](#)
- [Tarea 3. Importar el objeto de datos del archivo sin formato All_Customers, 60](#)
- [Resumen de la importación de objetos de datos físicos, 60](#)

Introducción a la importación de objetos de datos físicos

Un objeto de datos es una representación de datos basados en un archivo sin formato o tabla de base de datos relacional. Se puede importar un archivo sin formato o una tabla de base de datos relacional como objeto de datos físicos para usarlos como origen o destino en una asignación.

Historial

HypoStores Corporation almacena datos de sus oficinas de Los Ángeles y Boston en archivos sin formato. Usted quiere trabajar con estos datos de clientes en Developer tool. Para ello, tiene que importar cada archivo sin formato como un objeto de datos físicos.

Objetivos

En esta lección, se importan archivos sin formato como objetos de datos físicos. También se especifica el directorio del archivo de origen para que el servicio de integración de datos pueda leer los datos de origen desde el directorio correcto.

Requisitos previos

Antes de empezar con esta lección, verifique los siguientes requisitos previos:

- Ha terminado la lección 1 de esta guía de autoaprendizaje.

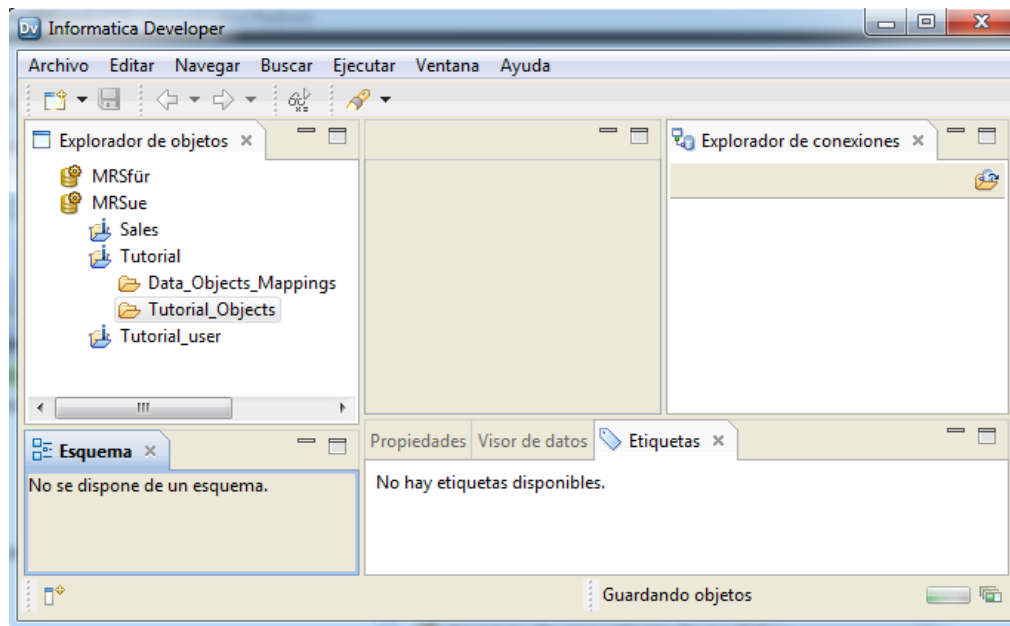
Duración

Destine de 10 a 15 minutos para realizar las tareas de esta lección.

Tarea 1. Importar el objeto de datos del archivo sin formato Boston_Customers

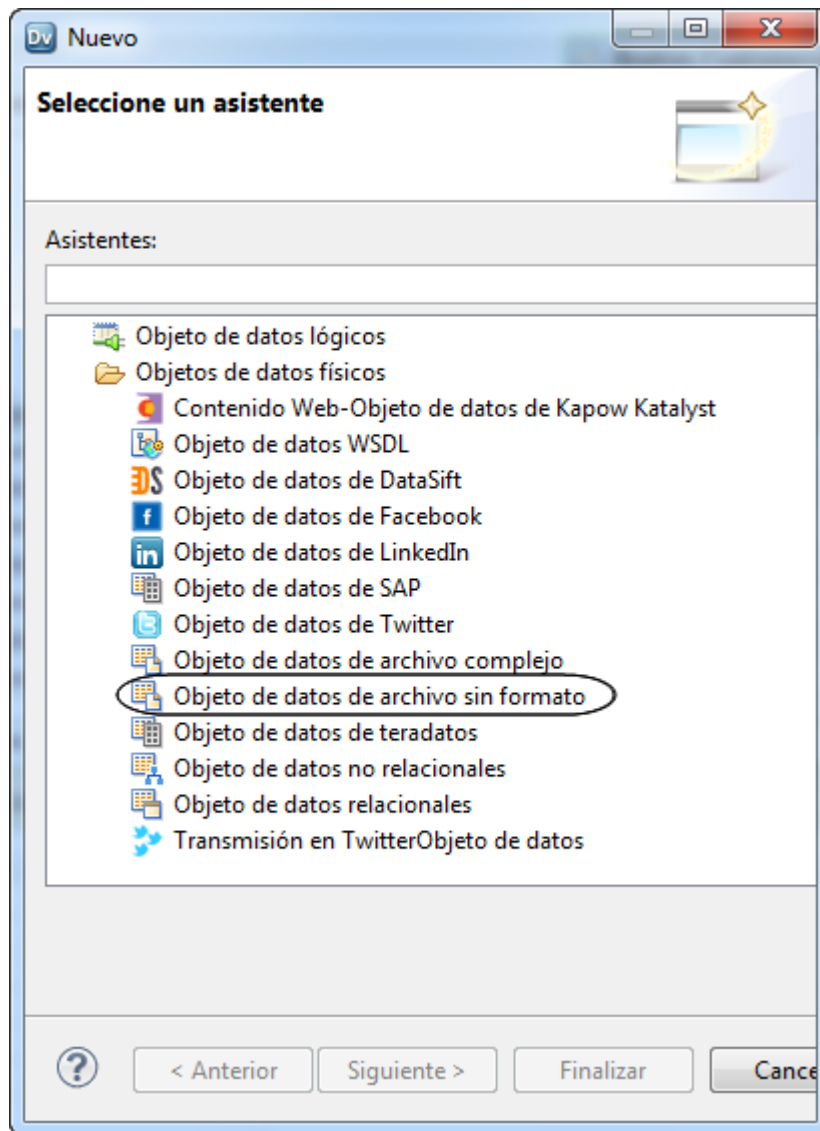
En esta tarea, tiene que importar un objeto de datos físicos desde un archivo que contiene datos de clientes de la oficina de Boston.

1. En la vista **Explorador de objetos**, seleccione la carpeta Tutorial_Objects.



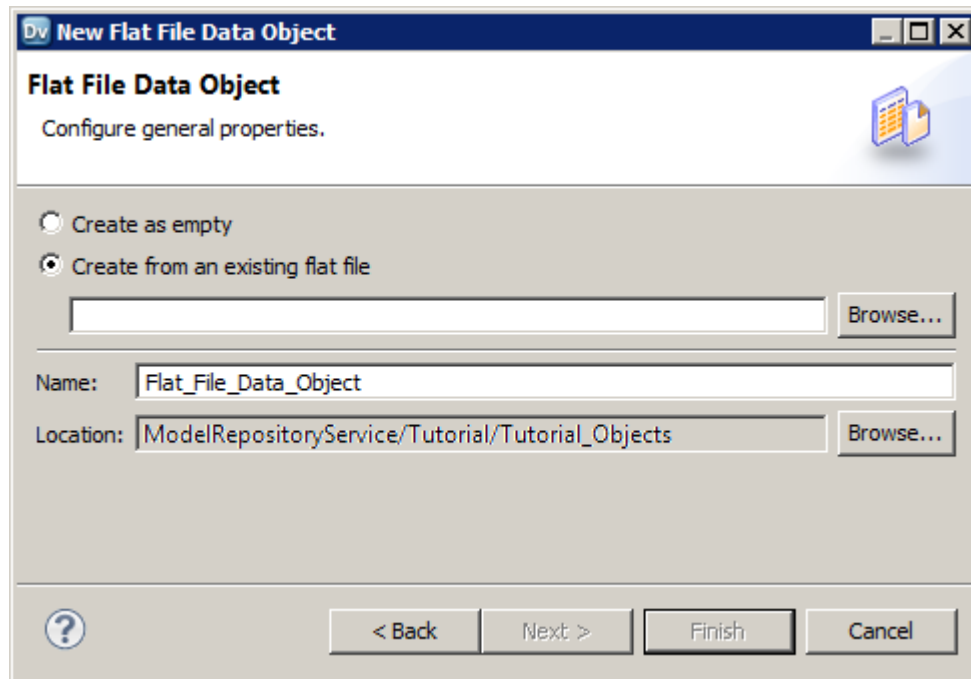
2. Haga clic con el botón derecho en la carpeta Tutorial_Objects y seleccione **Nuevo > Objeto de datos**.

A continuación, aparecerá el cuadro de diálogo **Nuevo**.



3. Seleccione **Objetos de datos físicos** > **Objeto de datos de archivo sin formato** y haga clic en **Siguiete**.

A continuación, aparecerá el cuadro de diálogo **Nuevo objeto de datos de archivo sin formato**.



4. Seleccione **Crear desde un archivo sin formato existente**.
5. Haga clic en **Examinar** y desplácese hasta Boston_Customers.csv en el directorio siguiente en el equipo de Developer Tool: <directorio de instalación de Informatica>\clientes\DeveloperClient\Guías de aprendizaje
6. Haga clic en **Abrir**.
El asistente nombra al objeto de datos como "Boston_Customers".
7. Haga clic en **Siguiente**.
8. Compruebe que la página de códigos esté establecida en **MS Windows Latín 1 (ANSI)**, **superconjunto de Latín 1** y que el formato esté establecido en **Delimitado**.

El cuadro de diálogo **Nuevo objeto de datos de archivo sin formato** muestra la página de códigos predeterminada, el formato y una vista previa de los datos del archivo sin formato.

Nuevo objeto de datos de archivo sin formato

Configure la página de códigos y el formato.

Página de códigos: MS Windows Latin 1 (ANSI), superset of Latin1

Formato

☒ Delimitado (campos separados por delimitadores)

☐ Ancho fijo (campos alineados en columnas)

Máximo de filas para vista previa: 500

	Campo1	Campo2	Campo3	
1	CustomerID	Customer Region	CustomerTier	Lastname
2	10110102	Boston	Bronze	Brosseau
3	10110105	Boston		Anderson
4	10110106	Boston	Emerald	Boons
5	10110107	Boston	Ruby	Chan
6	10110108	Boston	Silver	Chen

? < Anterior Siguiente > Finalizar Cancelar

9. Haga clic en **Siguiente**.
10. Seleccione **Importar nombres de columna desde la primera línea**.

El cuadro de diálogo **Nuevo objeto de datos de archivo sin formato** muestra los nombres de columna en la vista previa de los datos de archivo sin formato.

Objeto de datos de archivo sin formato
Configure las propiedades del formato delimitado.

Delimitadores

☐ Tabulador ☐ Punto y coma ☒ Coma

☐ Espacio ☐ Otro: ...

Calificador de texto

☒ Sin comillas ☐ Comillas simples ☐ Comillas dobles

Opciones de vista previa

☒ Importar nombres de columna desde la primera línea

Iniciar importación en la línea:

Delimitador de fila:

☐ Tratar delimitadores consecutivos como uno

☐ Retener carácter de escape en datos

Carácter de escape:

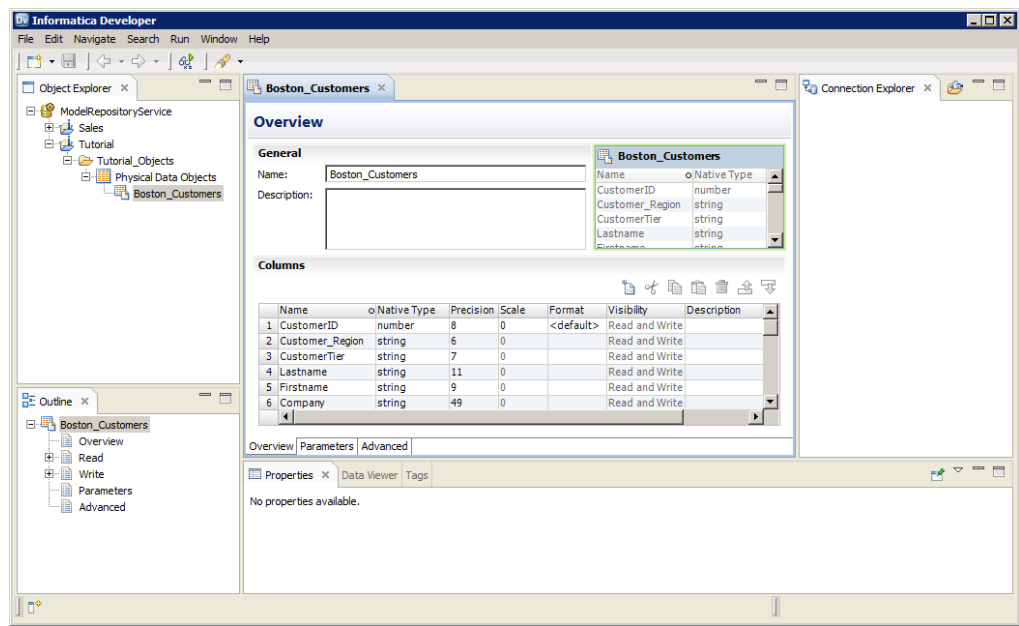
Máximo de filas para vista previa:

	CustomerID	Customer_Region	CustomerTier	Lastname
1	10110102	Boston	Bronze	Brosseau
2	10110105	Boston		Anderson
3	10110106	Boston	Emerald	Boonstra
4	10110107	Boston	Ruby	Chan

1. La opción **Importar nombres de columna desde la primera línea**
2. Nombres de columna

11. Haga clic en **Finalizar**.

El objeto de datos físicos Boston_Customers aparece en la carpeta Objetos de datos físicos dentro de la carpeta Tutorial_Objects. La vista **Resumen** muestra el contenido del archivo y se abre en el editor.



12. Haga clic en la vista **Opciones avanzadas**.
La vista **Avanzadas** muestra las propiedades del objeto de datos físicos.
13. En la vista **Avanzadas**, desplácese hasta la sección **Tiempo de ejecución: Lectura**.
14. En la sección **Tiempo de ejecución: Lectura**, establezca **Directorio de archivo de origen** en el siguiente directorio del equipo del servicio de integración de datos: <directorio de instalación de Informatica>\server\Tutorials.

El servicio de integración de datos buscará el archivo de origen en el directorio del servidor del equipo donde se ejecuta dicho servicio. La instalación del servidor contiene una copia de los archivos de la guía de autoaprendizaje. El servicio de integración de datos no puede leer los archivos desde el directorio de instalación del cliente a menos que se cambien los permisos de acceso en el archivo de origen y el directorio.

La siguiente figura muestra un ejemplo del directorio de archivo de origen:

Advanced	
Name	Value
Runtime : Read	
Input type	File
Source type	Direct
Source file name	Boston_Customers.csv
Source file directory	\\MyMachine\Informatica\10.0\server\Tutorials

Nota: El equipo de Developer Tool debe tener acceso al directorio del archivo de origen en el equipo que ejecuta el servicio de integración de datos. Si Developer Tool no puede acceder al directorio del archivo de origen, dicha herramienta no podrá obtener una vista previa de los datos del archivo de origen ni ejecutar asignaciones que accedan a los datos contenidos en este. Si ejecuta varios servicios de integración de datos, hay un directorio de archivo de origen independiente para cada uno de ellos.

15. Haga clic en la vista **Visor de datos**.

16. En la vista **Visor de datos**, haga clic en **Ejecutar**.
El servicio de integración de datos lee los datos del archivo Boston_Customers y muestra los resultados en la ventana Salida.
17. Haga clic en **Archivo > Guardar** para guardar el objeto de datos físicos Boston_Customers.

Tarea 2. Importar el objeto de datos de archivo sin formato LA_Customers

En esta tarea, tiene que importar un objeto de datos físicos desde un archivo sin formato que contiene datos de clientes de la oficina de Los Ángeles.

1. En la vista **Explorador de objetos**, seleccione el proyecto de la guía de autoaprendizaje.
2. Haga clic en **Archivo > Nuevo > Objeto de datos**.
A continuación, aparece el cuadro de diálogo **Nuevo**.
3. Seleccione **Objetos de datos físicos > Objeto de datos de archivo sin formato** y haga clic en **Siguiente**.
A continuación, aparece el cuadro de diálogo **Nuevo objeto de datos de archivo sin formato**.
4. Seleccione **Crear desde un archivo sin formato existente**.
5. Haga clic en **Examinar** y desplácese hasta LA_Customers.csv en el siguiente directorio: <Directorio de instalación de Informatica>\clients\DeveloperClient\Tutorials
6. Haga clic en **Abrir**.
El asistente nombra el objeto de datos como LA_Customers.
7. Haga clic en **Siguiente**.
8. Compruebe que la página de códigos sea MS Windows Latín 1 (ANSI), supraconjunto de Latín 1.
9. Compruebe que el formato esté delimitado.
10. Haga clic en **Siguiente**.
11. Compruebe que el delimitador establecido se corresponda con la coma.
12. Seleccione **Importar nombres de columna desde la primera línea**.
13. Haga clic en **Finalizar**.
El objeto de datos físicos LA_Customers aparece en **Objetos de datos físicos**, en el proyecto de la guía de autoaprendizaje.
14. Haga clic en **Lectura** y seleccione la transformación de salida.
15. Haga clic en la ficha **Tiempo de espera**, en la vista **Propiedades**.
16. Defina el directorio del archivo de origen como el directorio siguiente en el equipo del servicio de integración de datos: <Directorio de instalación de Informatica>\server\Tutorials
17. Haga clic en **Archivo > Guardar**.

Tarea 3. Importar el objeto de datos del archivo sin formato All_Customers

En esta tarea, tiene que importar un objeto de datos físicos desde un archivo sin formato que contiene datos de clientes de las oficinas de Los Ángeles y Boston.

1. En la vista **Explorador de objetos**, seleccione el proyecto de la guía de autoaprendizaje.
2. Haga clic en **Archivo > Nuevo > Objeto de datos**.
A continuación, aparece el cuadro de diálogo **Nuevo**.
3. Seleccione **Objetos de datos físicos > Objeto de datos de archivo sin formato** y haga clic en **Siguiente**.
Aparece el cuadro de diálogo **Nuevo origen de datos de archivo sin formato**.
4. Seleccione **Crear desde un archivo sin formato existente**.
5. Haga clic en **Examinar** y desplácese hasta All_Customers.csv en el siguiente directorio: <Directorio de instalación de Informatica>\clients\DeveloperClient\Tutorials.
6. Haga clic en **Abrir**.
El asistente nombra el objeto de datos como All_Customers.
7. Haga clic en **Siguiente**.
8. Compruebe que la página de códigos sea MS Windows Latín 1 (ANSI), supraconjunto de Latín 1.
9. Compruebe que el formato esté delimitado.
10. Haga clic en **Siguiente**.
11. Compruebe que el delimitador establecido se corresponda con la coma.
12. Seleccione **Importar nombres de columna desde la primera línea**.
13. Haga clic en **Finalizar**.
El objeto de datos físicos All_Customers aparece en **Objetos de datos físicos**, en el proyecto de la guía de autoaprendizaje.
14. Haga clic en **Lectura** y seleccione la transformación de salida.
15. Haga clic en la ficha **Tiempo de espera**, en la vista **Propiedades**.
16. Defina el directorio del archivo de origen como el directorio siguiente en el equipo del servicio de integración de datos: <Directorio de instalación de Informatica>\server\Tutorials
17. Haga clic en **Archivo > Guardar**.

Resumen de la importación de objetos de datos físicos

En esta lección, ha aprendido que los objetos de datos físicos son representaciones de datos basadas en un archivo sin formato o en una base de datos relacional.

Ha creado objetos de datos físicos a partir de archivos sin formato. También se especifica el directorio del archivo de origen para que el servicio de integración de datos pueda leer los datos de origen desde el directorio correcto.

Los objetos de datos se utilizan como orígenes de asignación en las lecciones sobre calidad de datos.

CAPÍTULO 12

Lección 3. Ejecutar un perfil con datos de origen

Este capítulo incluye los siguientes temas:

- [Introducción a la creación de perfiles de datos, 61](#)
- [Tarea 1. Realizar un análisis de combinación en dos orígenes de datos, 62](#)
- [Tarea 2. Ver los resultados de un análisis de combinación, 63](#)
- [Tarea 3. Ejecutar un perfil en un origen de datos, 64](#)
- [Tarea 4. Ver los resultados de perfil de columna, 64](#)
- [Resumen de datos de creación de perfiles, 65](#)

Introducción a la creación de perfiles de datos

Un perfil es un conjunto de metadatos que describe el contenido y la estructura de un conjunto de datos.

La creación de perfiles y la obtención de datos suelen ser los primeros pasos de un proyecto. Puede ejecutar un perfil para evaluar la estructura de datos y para verificar que las columnas de datos contengan los tipos de información que usted espera. En el caso de que un perfil revele problemas en los datos, puede definir pasos en el proyecto para corregir dichos problemas. Por ejemplo, si un perfil revela que una columna contiene valores de una longitud mayor que la esperada, puede diseñar procesos de calidad de datos para eliminar o corregir los valores problemáticos.

Un perfil que analiza la calidad de datos de las columnas seleccionadas se llama perfil de columna.

Nota: También puede usar Developer tool para descubrir la clave primaria, la clave externa y relaciones de dependencia funcional, así como analizar condiciones de combinación en columnas de datos.

Un perfil de columna proporciona los siguientes hechos sobre los datos:

- La cantidad de valores únicos y nulos de cada columna, expresada en forma de número y de porcentaje.
- Los patrones de datos en cada columna y las frecuencias con que ocurren estos valores.
- Estadísticas sobre los valores de columna, tales como las longitudes máxima y mínima de los valores, y el primer y último valor de cada columna.
- Para perfiles de análisis de combinación, el grado de solapamiento entre dos columnas de datos, mostrado como diagrama de Venn y como valor de porcentaje. Utilice perfiles de análisis de combinación para identificar posibles problemas con condiciones de unión de columnas.

Puede ejecutar un perfil de columna en cualquier etapa de un proyecto para medir la calidad de los datos y comprobar que los cambios efectuados en los datos cumplan los objetivos del proyecto. Puede ejecutar un perfil de columna en una transformación de una asignación para indicar el efecto que ejercerá la transformación en los datos.

Ejemplo

HypoStores desea comprobar que los datos de los clientes no contengan errores, inconsistencias ni información duplicada. Antes de que HypoStores diseñe los procesos para lograr los objetivos de calidad de datos, necesita cuantificar la calidad de sus archivos de datos de origen y confirmar que los datos están listos para procesarse.

Objetivos

En esta lección, llevará a cabo las tareas siguientes:

- Realizar un análisis de combinación en el origen de datos `Boston_Customers` y en el origen de datos `LA_Customers`.
- Visualizar los resultados del análisis de combinación para determinar si se puede o no fusionar correctamente los datos de las dos oficinas.
- Ejecutar un perfil en el origen de datos `All_Customers`.
- Visualizar los resultados de la creación de perfil de columna para observar los valores y patrones de los datos.

Requisitos previos

Antes de empezar con esta lección, verifique los siguientes requisitos previos:

- Ha terminado las Lecciones 1 y 2 de esta guía de autoaprendizaje.

Tiempo necesario

- Dedique 20 minutos a completar esta lección.

Tarea 1. Realizar un análisis de combinación en dos orígenes de datos

En esta tarea, se realiza un análisis de combinación de los orígenes de datos `Boston_Customers` y `LA_Customers` para ver las condiciones de combinación.

1. Seleccione la carpeta del autoaprendizaje y haga clic en **Archivo > Nuevo > Perfil**.
2. Seleccione **Perfil de detección empresarial**.
3. Haga clic en **Siguiente**.
4. En el campo **Nombre**, introduzca **Tutorial_Profile**.
5. Haga clic en **Finalizar**.
El perfil `Tutorial_Profile` aparece en el explorador de objetos.
6. Arrastre los orígenes de datos **`Boston_Customers`** y **`LA_Customers`** al editor en el lado derecho.
Sugerencia: Mantenga pulsada la tecla Mayús para seleccionar varios objetos de datos.
7. Haga clic con el botón derecho sobre el nombre de un objeto de datos y seleccione **Perfil de combinación**.
Aparece el asistente **Nuevo perfil de combinación**.

8. En el campo **Nombre**, introduzca **Análisis de combinación**.
9. Compruebe que Boston_Customers y LA_Customers aparecen como objetos de datos y haga clic en **Siguiente**.
10. Compruebe que la columna **CustomerID** esté seleccionada en los dos orígenes de datos. Desplácese por el panel del asistente para ver las columnas de ambos conjuntos de datos. Haga clic en **Siguiente**.
11. Haga clic en **Añadir** para añadir condiciones de combinación. Aparece el cuadro de diálogo **Condición de combinación**.
12. En la sección **Columnas**, haga clic en **Añadir fila**.
13. Haga doble clic en la primera fila de la columna de la izquierda y seleccione **CustomerID**.
14. Haga doble clic en la primera fila de la columna de la derecha y seleccione **CustomerID**.
15. Haga clic en **Aceptar** y, a continuación, haga clic en **Finalizar**.
16. Si la Herramienta del desarrollador le solicita guardar los cambios, haga clic en **Sí**. La Herramienta del desarrollador ejecuta el perfil.

Nota: No cierre el perfil. Verá los resultados del perfil en la siguiente tarea.

Tarea 2. Ver los resultados de un análisis de combinación

En esta tarea, se visualizan los resultados de un análisis de combinación en la vista Resultado de combinación del perfil JoinAnalysis.

1. Haga clic en la ficha **JoinAnalysis** en el editor.
2. En la sección **Resultado de combinación**, haga clic en la primera fila.

La sección **Detalles** muestra un diagrama de Venn y la clave de colores que detalla los resultados del análisis de combinación.
3. Compruebe que la columna **Filas de combinación** dé como resultado cero para el número de filas que contienen una combinación.

Este valor indica que los campos CustomerID no tienen valores duplicados. Se pueden fusionar correctamente los dos orígenes de datos.
4. Para ver los valores de CustomerID del objeto de datos LA_Customers, haga doble clic en el círculo con el nombre **LA_Customers** en el diagrama de Venn.

Sugerencia: Haga doble clic en los círculos del diagrama de Venn para ver las filas de datos. Si los círculos forman intersecciones en el diagrama de Venn, haga doble clic en la intersección para ver los valores de datos comunes a ambos conjuntos de datos.

El visor de datos muestra los valores de CustomerID del objeto de datos LA_Customers.

Tarea 3. Ejecutar un perfil en un origen de datos

En esta tarea, se ejecuta un perfil en el origen de datos All_Customers para visualizar el contenido y la estructura de los datos.

1. En la vista **Explorador de objetos**, localice los objetos de datos del proyecto de autoaprendizaje.
2. Seleccione el origen de datos **All_Customers**.
3. Haga clic en **Archivo > Nuevo > Perfil**.
A continuación, aparece el cuadro de diálogo **Nuevo**.
4. Seleccione **Perfil**.
5. Haga clic en **Siguiente**.
6. En el campo **Nombre**, introduzca **All_Customers**.
7. Haga clic en **Finalizar**.

El perfil All_Customers se abre en el editor y se ejecuta el perfil.

Tarea 4. Ver los resultados de perfil de columna

En esta tarea, se visualizan los resultados de creación de perfiles de columna del objeto de datos All_Customers y se examinan los valores y patrones contenidos en los datos.

1. Haga clic en **Ventana > Mostrar vista > Progreso** para ver el progreso del perfil All_Customers.
Se abre la vista Progreso.
2. Cuando la vista Progreso comunique que el perfil All_Customers ha terminado de ejecutarse, haga clic en la vista **Resultados** en el editor.
3. En la sección **Creación de perfiles de columna**, haga clic en la columna **CustomerTier**.
La sección **Detalles** muestra todos los valores contenidos en la columna CustomerTier y muestra información sobre la frecuencia de aparición de los valores en el conjunto de datos.
4. En la sección **Detalles**, haga doble clic **Rubí**.
El visor de datos se ejecuta y muestra los registros en los que la columna CustomerTier contiene el valor **Rubí**.
5. En la sección **Creación de perfiles de columna**, haga clic en la columna **OrderAmount**.
6. En la sección **Detalles**, haga clic en la lista **Mostrar** y seleccione **Patrones**.
La sección **Detalles** muestra los patrones encontrados en la columna OrderAmount. La cadena **9(5)** de la columna Patrón se refiere a los registros que contienen importes de pedidos de cinco cifras. La cadena **9(4)** se refiere a registros que contienen importes de pedidos de cuatro cifras.
7. En la columna **Patrón**, haga doble clic en la cadena **9(4)**.
El visor de datos se ejecuta y muestra los registros en los que la columna OrderAmount contiene un importe de pedido de cuatro cifras.
8. En la sección **Detalles**, haga clic en la lista **Mostrar** y seleccione **Estadísticas**.
La sección **Detalles** muestra estadísticas de la columna OrderAmount, incluidos el valor promedio, la desviación estándar, las longitudes máxima y mínima, los cinco valores más comunes y los cinco valores menos comunes.

Resumen de datos de creación de perfiles

En esta lección, ha aprendido que un perfil proporciona información sobre el contenido y la estructura de los datos.

Ha aprendido que puede realizar un análisis de unión de dos objetos de datos y visualizar el grado de solapamiento entre los mismos. También ha aprendido que puede ejecutar un perfil de columna en un objeto de datos y visualizar los valores, patrones y estadísticas relacionados con cada columna del objeto de datos.

Ha creado el perfil `JoinAnalysis` para determinar si los datos del objeto de datos `Boston_Customers` pueden fusionarse con los datos del objeto de datos `LA_Customers`. Ha visualizado los resultados de este perfil y ha determinado que todos los valores de la columna `CustomerID` son únicos y que puede fusionar los objetos de datos correctamente.

Ha creado el perfil `All_Customers` y ha ejecutado un perfil de columna en el objeto de datos `All_Customers`. Ha visualizado los resultados de este perfil para descubrir valores, patrones y estadísticas para las columnas en el objeto de datos `All_Customers`. Por último, ha ejecutado el visor de datos para visualizar filas que contienen los valores y patrones seleccionados, permitiéndole verificar la calidad de los datos.

INDICE

C

configuración de Developer tool
resumen [48](#)
configuración de la Herramienta del analista
resumen [19](#)
creación de cuadros de mandos
resumen [36](#)
creación de objetos de datos
resumen [22](#)
creación de perfiles
resumen [8](#)
creación de perfiles de datos
resumen [61](#)
creación de perfiles personalizados
resumen [29](#)
creación de perfiles predeterminados
resumen [25](#)

creación de reglas de expresión
resumen [33](#)
creación de tablas de referencia a partir de columnas
resumen [42](#)

I

importar objeto de datos físicos
resumen [52](#)

T

tablas de referencia
resumen [45](#)