



Informatica® Data Quality
10.1.1

Guía sobre aceleradores

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.

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.

Informatica, el logotipo de Informatica y PowerCenter son marcas comerciales o marcas comerciales registradas de Informatica LLC en los Estados Unidos y en muchas otras 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.

Hay fragmentos de este software y/o documentación que están sujetas a copyright perteneciente a terceros, incluido, entre otros: Copyright DataDirect Technologies. Todos los derechos reservados. Copyright © Sun Microsystems. Todos los derechos reservados. Copyright © RSA Security Inc. Todos los derechos reservados. Copyright © Ordinal Technology Corp. Todos los derechos reservados. Copyright © Aandacht c.v. Todos los derechos reservados. Copyright © Genivia, Inc. Todos los derechos reservados. Copyright © Isomorphic Software. Todos los derechos reservados. Copyright © Meta Integration Technology, Inc. Todos los derechos reservados. Copyright © Intalio. Todos los derechos reservados. Copyright © Oracle. Todos los derechos reservados. Copyright © Adobe Systems Incorporated. Todos los derechos reservados. Copyright © DataArt, Inc. Todos los derechos reservados. Copyright © ComponentSource. Todos los derechos reservados. Copyright © Microsoft Corporation. Todos los derechos reservados. Copyright © Rogue Wave Software, Inc. Todos los derechos reservados. Copyright © Teradata Corporation. Todos los derechos reservados. Copyright © Yahoo! Inc. Todos los derechos reservados. Copyright © Glyph & Cog, LLC. Todos los derechos reservados. Copyright © Thinkmap, Inc. Todos los derechos reservados. Copyright © Clearpace Software Limited. Todos los derechos reservados. Copyright © Information Builders, Inc. Todos los derechos reservados. Copyright © OSS Nokalva, Inc. Todos los derechos reservados. Copyright © Edifecs, Inc. Todos los derechos reservados. Copyright © Cleo Communications, Inc. Todos los derechos reservados. Copyright © International Organization for Standardization 1986. Todos los derechos reservados. Copyright © ej-technologies GmbH. Todos los derechos reservados. Copyright © Jaspersoft Corporation. Todos los derechos reservados. Copyright © International Business Machines Corporation. Todos los derechos reservados. Copyright © yWorks GmbH. Todos los derechos reservados. Copyright © Lucent Technologies. Todos los derechos reservados. Copyright © University of Toronto. Todos los derechos reservados. Copyright © Daniel Veillard. Todos los derechos reservados. Copyright © Unicode, Inc. Copyright IBM Corp. Todos los derechos reservados. Copyright © MicroQuill Software Publishing, Inc. Todos los derechos reservados. Copyright © PassMark Software Pty Ltd. Todos los derechos reservados. Copyright © LogiXML, Inc. Todos los derechos reservados. Copyright © 2003-2010 Lorenzi Davide. Todos los derechos reservados. Copyright © Red Hat, Inc. Todos los derechos reservados. Copyright © The Board of Trustees of the Leland Stanford Junior University. Todos los derechos reservados. Copyright © EMC Corporation. Todos los derechos reservados. Copyright © Flexera Software. Todos los derechos reservados. Copyright © Jinfonet Software. Todos los derechos reservados. Copyright © Apple Inc. Todos los derechos reservados. Copyright © Telerik Inc. Todos los derechos reservados. Copyright © BEA Systems. Todos los derechos reservados. Copyright © PDFlib GmbH. Todos los derechos reservados. Copyright © Orientation in Objects GmbH. Todos los derechos reservados. Copyright © Tanuki Software, Ltd. Todos los derechos reservados. Copyright © Ricebridge. Todos los derechos reservados. Copyright © Sencha, Inc. Todos los derechos reservados. Copyright © Scalable Systems, Inc. Todos los derechos reservados. Copyright © jQWidgets. Todos los derechos reservados. Copyright © Tableau Software, Inc. Todos los derechos reservados. Copyright © MaxMind, Inc. Todos los derechos reservados. Copyright © TMate Software s.r.o. Todos los derechos reservados. Copyright © MapR Technologies Inc. Todos los derechos reservados. Copyright © Amazon Corporate LLC. Todos los derechos reservados. Copyright © Highsoft. Todos los derechos reservados. Copyright © Python Software Foundation. Todos los derechos reservados. Copyright © BeOpen.com. Todos los derechos reservados. Copyright © CNRI. Todos los derechos reservados.

Este producto incluye software desarrollado por la Apache Software Foundation (<http://www.apache.org/>) y/u otro software protegido por varias versiones de la licencia Apache License ("Licencia"). Puede obtener una copia de estas licencias en <http://www.apache.org/licenses/>. A menos que las leyes aplicables lo requieran o se haya acordado por escrito, el software distribuido bajo estas licencias se distribuye "TAL CUAL", SIN GARANTÍAS NI CONDICIONES DE NINGÚN TIPO, ya sea expresas o implícitas. Consulte las licencias del idioma específico para conocer los permisos y las limitaciones que rigen según las licencias.

Este producto incluye software desarrollado por Mozilla (<http://www.mozilla.org/>), copyright del software de The JBoss Group, LLC, todos los derechos reservados; copyright del software © 1999-2006 de Bruno Lowagie y Paulo Soares y otro software protegido por licencia por el acuerdo GNU Lesser General Public License Agreement, que se puede encontrar en la dirección <http://www.gnu.org/licenses/lgpl.html>. Los materiales se facilitan gratuitamente por parte de Informatica, "tal cual", sin garantía de ningún tipo, ya sea expresa o implícita, incluidas, entre otras, las garantías implícitas de adecuación para un propósito determinado y de validez para el comercio.

El producto incluye software ACE(TM) y TAO(TM) con copyright de Douglas C. Schmidt y su grupo de investigación de la Washington University, University of California, Irvine y Vanderbilt University, Copyright (©) 1993-2006, todos los derechos reservados.

Este producto incluye software desarrollado por el OpenSSL Project para uso en el OpenSSL Toolkit (copyright The OpenSSL Project. Todos los derechos reservados) y la redistribución de este software está sujeta a los términos especificados en <http://www.openssl.org> y <http://www.openssl.org/source/license.html>.

Este producto incluye software Curl con Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://curl.haxx.se/docs/copyright.html>. La autorización para utilizar, copiar, modificar y distribuir este software para cualquier propósito con o sin tasas se concede por el presente, siempre que el aviso de copyright anterior y este aviso de permiso aparezcan en todas las copias.

El producto incluye copyright de software 2001-2005 (©) MetaStuff, Ltd. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://www.dom4j.org/license.html>.

El producto incluye copyright de software © 2004-2007, The Dojo Foundation. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://dojotoolkit.org/license>.

Este producto incluye software ICU con copyright de International Business Machines Corporation y otros. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

Este producto incluye copyright de software © 1996-2006 Per Bothner. Todos los derechos reservados. Su derecho a utilizar estos materiales está establecido en la licencia que puede encontrarse en la dirección <http://www.gnu.org/software/kawa/Software-License.html>.

Este producto incluye software OSSP UUID con Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Los permisos y las limitaciones relativas a este software están sujetos a los términos disponibles en la dirección <http://www.opensource.org/licenses/mit-license.php>.

Este producto incluye software desarrollado por Boost (<http://www.boost.org/>) o protegido por la licencia de software de Boost. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección http://www.boost.org/LICENSE_1_0.txt.

Este producto incluye copyright de software © 1997-2007 University of Cambridge. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://www.pcre.org/license.txt>.

Este producto incluye copyright de software © 2007 The Eclipse Foundation. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos especificados en <http://www.eclipse.org/org/documents/epl-v10.php> y <http://www.eclipse.org/org/documents/edl-v10.php>.

Este producto incluye software protegido por licencia según los términos que aparecen en <http://www.tcl.tk/software/tcltk/license.html>, <http://www.bosrup.com/web/overlib/?License>, <http://www.stlport.org/doc/license.html>, <http://asm.ow2.org/license.html>, <http://www.cryptix.org/LICENSE.TXT>, <http://hsqldb.org/web/hsqllLicense.html>, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, http://www.gzip.org/zlib/zlib_license.html, <http://www.openldap.org/software/release/license.html>, <http://www.libssh2.org>, <http://slf4j.org/license.html>, <http://www.sente.ch/software/OpenSourceLicense.html>, <http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3-license-agreement>, <http://antlr.org/license.html>, <http://aopalliance.sourceforge.net/>, <http://www.bouncycastle.org/licence.html>, <http://www.jgraph.com/jgraphdownload.html>, <http://www.jcraft.com/jsch/LICENSE.txt>, http://jotm.objectweb.org/bsd_license.html, <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231>, <http://www.slf4j.org/license.html>, <http://nanoxml.sourceforge.net/orig/copyright.html>, <http://www.json.org/license.html>, <http://forge.ow2.org/projects/javaxservice/>, <http://www.postgresql.org/about/licence.html>, <http://www.sqlite.org/copyright.html>, <http://www.tcl.tk/software/tcltk/license.html>, <http://www.jaxen.org/faq.html>, <http://www.jdom.org/docs/faq.html>, <http://www.slf4j.org/license.html>, <http://www.iodbc.org/dataspace/iodbc/wiki/IODBC/License>, <http://www.keplerproject.org/md5/license.html>, <http://www.toedter.com/en/jcalendar/license.html>, <http://www.edankert.com/bounce/index.html>, <http://www.net-snmp.org/about/license.html>, <http://www.openmdx.org/#FAQ>, http://www.php.net/license/3_01.txt, <http://srp.stanford.edu/license.txt>, <http://www.schneier.com/blowfish.html>, <http://www.jmock.org/license.html>, <http://xsom.java.net>, <http://benalman.com/about/license/>, <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>, <http://www.h2database.com/html/license.html#summary>, <http://jsoncpp.sourceforge.net/LICENSE>, <http://jdbc.postgresql.org/license.html>, <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>, <https://github.com/rantav/hector/blob/master/LICENSE>, <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>, <http://jibx.sourceforge.net/jibx-license.html>, <https://github.com/lyokato/libgeohash/blob/master/LICENSE>, <https://github.com/hjiang/jsonxx/blob/master/LICENSE>, <https://code.google.com/p/lz4/>, <https://github.com/jedisct1/libsodium/blob/master/LICENSE>, <http://one-jar.sourceforge.net/index.php?page=documents&file=license>, <https://github.com/EsotericSoftware/kryo/blob/master/license.txt>, <http://www.scala-lang.org/license.html>, <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>, <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>, <https://aws.amazon.com/asl/>, <https://github.com/twbs/bootstrap/blob/master/LICENSE>, <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>, <https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE> y <https://github.com/apache/hbase/blob/master/LICENSE.txt>.

Este producto incluye software desarrollado por la Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), la Common Development and Distribution License (<http://www.opensource.org/licenses/cddl1.php>), la Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), la Sun Binary Code License Agreement Supplemental License Terms, la BSD License (<http://www.opensource.org/licenses/bsd-license.php>), la nueva BSD License (<http://opensource.org/licenses/BSD-3-Clause>), la MIT License (<http://www.opensource.org/licenses/mit-license.php>), la Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) y la Initial Developer's Public License Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

Este producto incluye copyright de software © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://xstream.codehaus.org/license.html>. Este producto incluye software desarrollado por Indiana University Extreme! Lab. Para obtener más información, visite <http://www.extreme.indiana.edu/>.

Este producto incluye software Copyright © 2013 Frank Balluffi y Markus Moeller. Todos los derechos reservados. Los permisos y las limitaciones relativas a este software están sujetos a los términos de la licencia MIT.

Consulte las patentes en <https://www.informatica.com/legal/patents.html>.

EXENCIÓN DE RESPONSABILIDAD: Informatica LLC proporciona esta documentación "tal cual" sin garantía de ningún tipo, ya sea expresa o implícita, incluidas, entre otras, las garantías implícitas de no incumplimiento, de adecuación para un propósito determinado y de validez para el comercio. Informatica LLC no garantiza que este software o esta documentación estén libres de errores. La información proporcionada en este software o en esta documentación puede contener imprecisiones técnicas o errores tipográficos. La información de este software y esta documentación está sujeta a cambios en cualquier momento sin previo aviso.

AVISOS

Este producto de Informatica (el "Software") incluye ciertos controladores (los "Controladores DataDirect") de DataDirect Technologies, una empresa operativa de Progress Software Corporation ("DataDirect") que están sujetos a los términos y condiciones siguientes:

1. LOS CONTROLADORES DATADIRECT SE PROPORCIONAN "TAL CUAL" SIN GARANTÍA DE NINGÚN TIPO, YA SEA EXPRESA O IMPLÍCITA, INCLUIDAS, ENTRE OTRAS, LAS GARANTÍAS IMPLÍCITAS DE NO INCUMPLIMIENTO, DE ADECUACIÓN PARA UN PROPÓSITO DETERMINADO Y DE VALIDEZ PARA EL COMERCIO.
2. EN NINGÚN CASO DATADIRECT NI SUS PROVEEDORES DE TERCEROS SERÁN RESPONSABLES ANTE EL USUARIO FINAL POR NINGÚN DAÑO DIRECTO, INDIRECTO, FORTUITO, ESPECIAL, CONSECUENTE, NI DE NINGÚN OTRO TIPO, RESULTANTE DEL USO DE LOS CONTROLADORES ODBC, INDEPENDIENTEMENTE DE SI SE HA AVISADO O NO DE LOS POSIBLES DAÑOS POR ADELANTADO. ESTAS LIMITACIONES SE APLICAN A TODAS LAS DEMANDAS JUDICIALES, INCLUIDAS, ENTRE OTRAS, AQUELLAS POR INCUMPLIMIENTO DE CONTRATO, INCUMPLIMIENTO DE LA GARANTÍA, NEGLIGENCIA, RESPONSABILIDAD Estricta, TERGIVERSACIÓN Y OTROS AGRAVIOS.

La información contenida en esta documentación está sujeta a cambios sin previo aviso. Si encuentra algún problema en esta documentación, infórmenos por escrito a Informatica LLC 2100 Seaport Blvd. Redwood City, CA 94063.

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: 2018-06-09

Tabla de contenido

Prefacio	8
Documentación de Informatica	8
Informatica Network.	8
Base de conocimiento de Informatica.	8
Documentación de Informatica	8
Matrices de disponibilidad de productos de Informatica.	9
Informatica Velocity.	9
Catálogo de soluciones de Informatica.	9
Servicio internacional de atención al cliente de Informatica.	9
 Capítulo 1: Introducción a los aceleradores.....	10
Resumen de los aceleradores.	10
Estructura de los aceleradores.	10
Estructura del acelerador general.	11
Estructura del acelerador de dominios de datos.	11
Instalación de los aceleradores.	12
Reglas y directrices para la instalación de los aceleradores.	13
Importar reglas y asignaciones.	14
Importar dominios de datos y grupos de dominios de datos.	14
Componentes de los aceleradores.	15
Reglas.	17
Asignaciones de demostración.	18
Dominios de datos.	18
Tablas de referencia.	18
Conjuntos de contenido.	18
Etiquetas y reglas.	19
Uso de los aceleradores en PowerCenter.	19
 Capítulo 2: Acelerador del núcleo.....	20
Resumen del acelerador del núcleo.	20
Reglas de limpieza de datos de direcciones del núcleo.	21
Reglas de limpieza de datos de contacto del núcleo.	22
Reglas de limpieza de datos corporativos del núcleo.	23
Reglas de limpieza de datos generales del núcleo.	23
Reglas de coincidencia y eliminación de duplicados del núcleo.	30
Reglas de limpieza de datos de producto del núcleo.	30
Asignaciones de demostración del núcleo.	31
 Capítulo 3: Acelerador de dominios de datos.....	32
Introducción al acelerador de dominios de datos.	32

Los dominios de datos en el acelerador de dominios de datos.	33
Reglas de nombre de columna en el acelerador de dominios de datos.	41
Reglas de datos en el acelerador de dominios de datos.	45
Capítulo 4: Acelerador de Australia/Nueva Zelanda.	52
Resumen del acelerador de Australia/Nueva Zelanda.	52
Reglas de limpieza de datos de direcciones de Australia/Nueva Zelanda.	53
Reglas compuestas de Australia/Nueva Zelanda.	54
Reglas de limpieza de datos de contacto de Australia/Nueva Zelanda.	56
Reglas de limpieza de datos corporativos de Australia/Nueva Zelanda.	59
Reglas de limpieza de datos generales de Australia/Nueva Zelanda.	60
Reglas de eliminación de duplicados y coincidencia de Australia/Nueva Zelanda.	61
Asignaciones de demostración de Australia/Nueva Zelanda.	64
Capítulo 5: Acelerador de Brasil.	65
Resumen del acelerador de Brasil.	65
Reglas de limpieza de datos de direcciones de Brasil.	66
Reglas compuestas de Brasil.	67
Reglas de limpieza de datos de contacto de Brasil.	68
Reglas de limpieza de datos corporativos de Brasil.	69
Reglas de limpieza de datos generales de Brasil.	69
Reglas de coincidencia y eliminación de duplicados de Brasil.	70
Asignaciones de demostración de Brasil.	72
Capítulo 6: Acelerador de Servicios Financieros.	74
Resumen del acelerador de Servicios Financieros.	74
Reglas de limpieza de datos de contacto de Servicios Financieros.	74
Reglas de limpieza de datos financieros de Servicios Financieros.	75
Reglas de limpieza de datos generales de Servicios Financieros.	78
Reglas de coincidencia y eliminación de duplicados de Servicios Financieros.	79
Capítulo 7: Acelerador de Francia.	81
Resumen del acelerador de Francia.	81
Reglas de limpieza de datos de direcciones de Francia.	81
Reglas compuestas de Francia.	82
Reglas de limpieza de datos de contacto de Francia.	83
Reglas de limpieza de datos corporativos de Francia.	85
Reglas de limpieza de datos generales de Francia.	86
Reglas de coincidencia y eliminación de duplicados de Francia.	87
Asignaciones de demostración de Francia.	89
Capítulo 8: Acelerador de Alemania.	90
Resumen del acelerador de Alemania.	90

Reglas de limpieza de datos de direcciones de Alemania.	91
Reglas compuestas de Alemania.	92
Reglas de limpieza de datos de contacto de Alemania.	93
Reglas de limpieza de datos corporativos de Alemania.	94
Reglas de limpieza de datos generales de Alemania.	95
Reglas de coincidencia y eliminación de duplicados de Alemania.	96
Asignaciones de demostración de Alemania.	98
Capítulo 9: Acelerador de Portugal.	99
Resumen del acelerador de Portugal.	99
Reglas de limpieza de datos de direcciones de Portugal.	99
Reglas compuestas de Portugal.	100
Reglas de limpieza de datos de contacto de Portugal.	101
Reglas de limpieza de datos corporativos de Portugal.	103
Reglas de limpieza de datos generales de Portugal.	103
Reglas de coincidencia y eliminación de duplicados de Portugal.	104
Asignaciones de demostración de Portugal.	106
Capítulo 10: Acelerador de España.	107
Resumen del acelerador de España.	107
Reglas de limpieza de datos de direcciones de España.	107
Reglas de limpieza de datos de contacto de España.	109
Reglas de limpieza de datos corporativos de España.	110
Reglas de limpieza de datos generales de España.	110
Reglas de coincidencia y eliminación de duplicados de España.	111
Asignaciones de demostración de España.	114
Capítulo 11: Acelerador del Reino Unido.	115
Resumen del acelerador del Reino Unido.	115
Reglas de limpieza de datos de direcciones del Reino Unido.	116
Reglas compuestas del Reino Unido.	117
Reglas de limpieza de datos de contacto del Reino Unido.	118
Reglas de limpieza de datos corporativos del Reino Unido.	121
Reglas de limpieza de datos financieros del Reino Unido.	121
Reglas de limpieza de datos generales del Reino Unido.	122
Reglas de coincidencia y eliminación de duplicados del Reino Unido.	122
Asignaciones de demostración del Reino Unido.	125
Capítulo 12: Acelerador de EE. UU./Canadá.	126
Resumen del acelerador de EE. UU./Canadá.	126
Reglas de limpieza de datos de direcciones de EE. UU./Canadá.	126
Reglas compuestas de EE. UU./Canadá.	129
Reglas de limpieza de datos de contacto de EE. UU./Canadá.	131

Dependencias de limpieza de datos corporativos de EE. UU./Canadá.	136
Reglas de limpieza de datos generales de EE. UU./Canadá.	137
Reglas de coincidencia y eliminación de duplicados de EE. UU./Canadá.	138
Asignaciones de demostración de EE. UU./Canadá.	140

Prefacio

La *Guía sobre aceleradores* de Informatica Data Quality está destinada a desarrolladores de calidad de datos. Esta guía supone que tiene cierto conocimiento de los conceptos relacionados con la calidad de los datos, tales como estandarización, análisis, etiquetado y validación.

Documentación de Informatica

Informatica Network

Informatica Network incluye el servicio internacional de atención al cliente de Informatica, la base de conocimiento de Informatica y otros recursos de producto. Para acceder a Informatica Network, visite <https://network.informatica.com>.

Un miembro puede:

- Acceder a todos sus recursos de Informatica en un solo lugar.
- Busque recursos de producto, como documentación, preguntas frecuentes y mejores prácticas en la base de conocimiento.
- Vea la información de disponibilidad del producto.
- Revisar los casos de asistencia.
- Buscar su red de grupos de usuarios de Informatica locales y colaborar con sus iguales.

Base de conocimiento de Informatica

Utilice la base de conocimiento de Informatica para buscar recursos de producto como documentación, artículos de procedimientos, mejores prácticas y PAM en la red de Informatica.

Para acceder a la base de conocimiento, visite <https://kb.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

Para obtener la documentación más reciente del producto, consulte la base de conocimiento de Informatica en https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx.

Si tiene preguntas, comentarios o ideas relacionadas con esta documentación, póngase en contacto con el equipo de documentación de Informatica enviando un correo electrónico a infa_documentation@informatica.com.

Matrices de disponibilidad de productos 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 de datos y destinos admitidos por una versión de un producto. Si es miembro de la red de Informatica, puede acceder a las PAM en <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity es un conjunto de sugerencias y mejores prácticas desarrollado por los servicios profesionales de Informatica. Desarrollado a partir de la experiencia real de cientos de proyectos de administración de datos, Informatica Velocity representa el conocimiento conjunto de nuestros asesores, los cuales han trabajado con organizaciones de todo el mundo para planificar, desarrollar, implementar y mantener con éxito soluciones de administración de datos.

Si es miembro de la red de Informatica, puede acceder a los 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. Al aprovechar cualquiera de los cientos de soluciones de los desarrolladores y los socios de Informatica, puede mejorar la productividad y acelerar el tiempo de implementación en los proyectos. Puede acceder al 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 soporte en línea en la red de Informatica.

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:
<http://www.informatica.com/us/services-and-training/support-services/global-support-centers>.

Si es miembro de la red de Informatica, puede utilizar el soporte en línea en <http://network.informatica.com>.

CAPÍTULO 1

Introducción a los aceleradores

Este capítulo incluye los siguientes temas:

- [Resumen de los aceleradores, 10](#)
- [Estructura de los aceleradores, 10](#)
- [Instalación de los aceleradores, 12](#)
- [Componentes de los aceleradores, 15](#)
- [Etiquetas y reglas, 19](#)
- [Uso de los aceleradores en PowerCenter, 19](#)

Resumen de los aceleradores

Los aceleradores son paquetes compuestos de contenido que abordan problemas comunes de calidad de datos en un país, una región o una industria. Los aceleradores pueden contener mapplets que puede usar para analizar y mejorar los datos de una organización. Los aceleradores también podrían contener dominios de datos que puede usar para detectar los tipos de información que contienen los datos.

Puede añadir los mapplets y los dominios de datos al repositorio de modelos. Informatica configura los mapplets y los dominios de datos para responder a las reglas empresariales que puede definir para los datos de organización. Los aceleradores utilizan los términos *mapplet* y *regla* para identificar los mapplets. Al importar los mapplets en el repositorio de modelos, la herramienta del desarrollador crea los objetos de mapplet en una carpeta llamada *Rules*.

Informatica Data Quality incluye un acelerador del núcleo y un acelerador del dominio de datos del núcleo. Puede comprar y descargar aceleradores adicionales de Informatica.

Estructura de los aceleradores

Un acelerador es un archivo comprimido que contiene archivos de metadatos del repositorio y otros archivos en una estructura de directorios. La estructura de directorios depende del tipo de acelerador. Los aceleradores generales contienen reglas, objetos de datos de referencia, asignaciones de demostración y orígenes de datos de demostración. Los aceleradores de dominios de datos contienen reglas, objetos de datos de referencia, dominios de datos y grupos de dominios de datos.

Estructura del acelerador general

Los aceleradores generales incluyen las reglas que analizan y mejoran los datos de organización y las asignaciones de ejemplo que demuestran las operaciones de la regla. Los aceleradores generales también contienen los archivos de datos de referencia y los archivos de datos de origen que utilizan las reglas y las asignaciones.

Un acelerador general contiene los siguientes directorios:

- Accelerator_Content
- Accelerator_Sources

Directorio Accelerator_Content

El directorio Accelerator_Content contiene los siguientes componentes:

Archivo XML del acelerador

Contiene metadatos de las reglas, asignaciones de demostración, tablas de referencia y objetos de datos.

Archivo de datos de referencia

Contiene los datos de referencia que utilizan las reglas y las asignaciones para identificar formas diferentes de los valores de datos. El archivo de datos de referencia es un archivo comprimido que contiene archivos de diccionario en varios directorios. Especifique el archivo comprimido cuando importe el archivo XML correspondiente. El proceso de importación copia los datos de referencia en tablas de la base de datos de referencia.

Nota: Si exporta una asignación que contiene una regla a PowerCenter®, copie los archivos de diccionario en un directorio que pueda leer el servicio de integración de PowerCenter.

Directorio Accelerator_Sources

El directorio Accelerator_Sources contiene el archivo de datos de demostración. El archivo de datos de demostración es un archivo comprimido que contiene los datos de origen de las asignaciones de demostración. Copie el archivo de datos de origen en el sistema de archivos.

Estructura del acelerador de dominios de datos

Los aceleradores de dominios de datos incluyen los dominios de datos que determinan los tipos de información en un conjunto de datos y las reglas que definen la lógica del dominio de datos. Los aceleradores también contienen los archivos de datos de referencia que utilizan los dominios de datos y las reglas.

Un acelerador de dominios de datos contiene los siguientes archivos:

Archivo de metadatos de dominios de datos

Contiene los metadatos para los dominios de datos y grupos de dominios de datos que añade al glosario de dominio de datos.

Archivo de metadatos de reglas

Contiene los metadatos para las reglas que definen la lógica del dominio de datos y para los objetos de datos de referencia que utilizan los dominios de datos.

Archivo de datos de referencia para los dominios de datos

Contiene los datos de referencia que utiliza un dominio de datos cuando se ejecuta un perfil que contiene el dominio de datos. El archivo de datos de referencia es un archivo comprimido que contiene archivos de diccionario en varios directorios. Especifique el archivo comprimido cuando importe el

archivo XML correspondiente. El proceso de importación copia los datos de referencia en tablas de la base de datos de referencia.

Archivo de datos de referencia para las reglas de dominios de datos

Contiene los datos de referencia que utiliza una regla cuando se ejecuta un dominio de datos que contiene la regla. El archivo de datos de referencia es un archivo comprimido que contiene archivos de diccionario en varios directorios. Especifique el archivo comprimido cuando importe el archivo XML correspondiente. El proceso de importación copia los datos de referencia en tablas de la base de datos de referencia.

Instalación de los aceleradores

Para instalar un acelerador, importe los metadatos del objeto de repositorio en un proyecto del repositorio de modelos y copie los archivos de datos de demostración en el sistema de archivos. Utilice Developer tool para importar los objetos del repositorio.

Cuando importe reglas y asignaciones de demostración, seleccione el proyecto del repositorio en el Explorador de objetos. Cuando importe dominios de datos, seleccione el proyecto del repositorio en el cuadro de diálogo **Preferencias**. En cada caso, la operación de importación le solicitará que seleccione el archivo comprimido que contiene los datos de referencia especificados por el archivo XML.

Ejemplo de acelerador general

Puede importar el siguiente archivo de metadatos para el acelerador del núcleo:

`Informatica_Core_Accelerator_1011.xml`

Cuando importe el archivo de metadatos, seleccione el siguiente archivo de datos de referencia:

`Informatica_Core_Accelerator_1011.zip`

Ejemplo de acelerador de dominio de datos

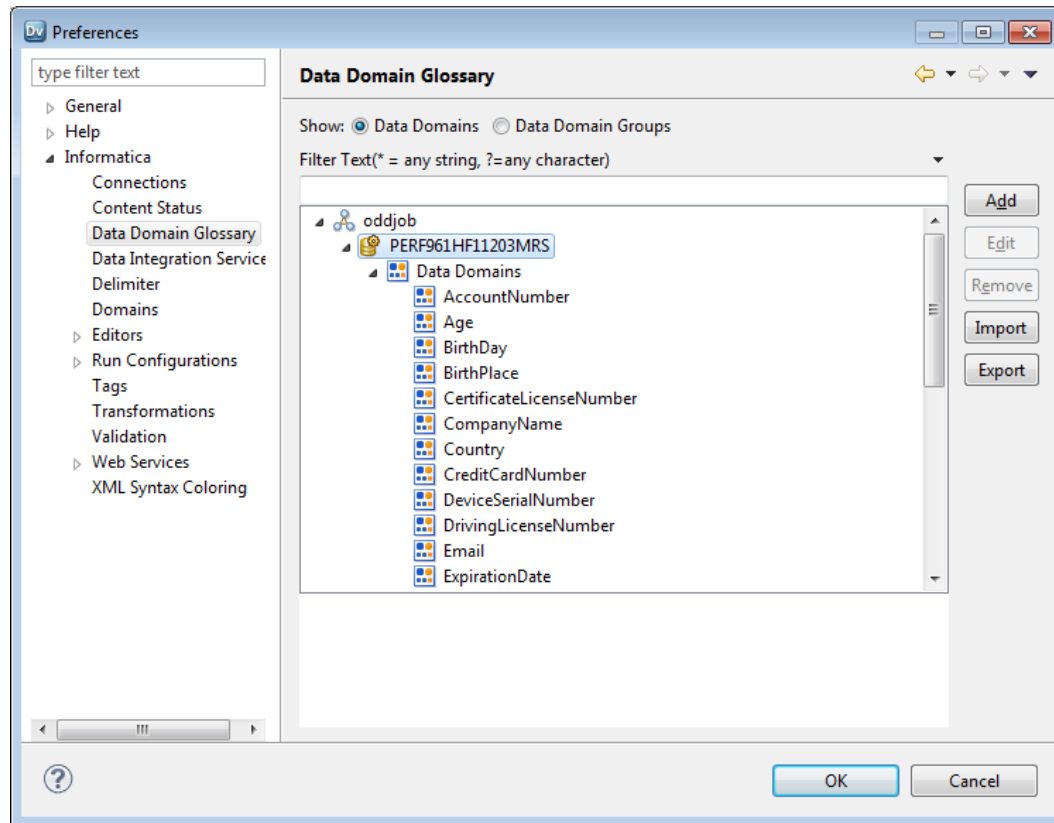
Puede importar el siguiente archivo de metadatos para el acelerador de dominio de datos del núcleo:

`Informatica_IDE_DataDomain_1011.xml`

Cuando importe el archivo de metadatos, seleccione el siguiente archivo de datos de referencia:

`Informatica_IDE_DataDomain_1011.zip`

La siguiente imagen muestra los dominios de datos en el cuadro de diálogo **Preferencias**:



Datos de origen para asignaciones de muestra

Cuando importe un acelerador general, copie los archivos de datos de demostración en el siguiente directorio del equipo del servicio de integración de datos:

<Directorio de instalación de Informatica>\services\DQContent\INFA_Content\demos\source_data

Reglas y directrices para la instalación de los aceleradores

Los objetos del repositorio y los archivos de datos de un acelerador operan en el mismo modo que otros objetos y archivos del sistema de Informatica. Algunas reglas y directrices se aplican al contenido del acelerador.

Tenga en cuenta las siguientes reglas y directrices al instalar un acelerador:

- Antes de importar o copiar archivos, asegúrese de que cuenta con todos los privilegios en el servicio de integración de datos, el servicio de administración de contenido y el servicio del analista.
- Importe los aceleradores en un proyecto único del repositorio de modelos. Cree el proyecto antes de importar los aceleradores.
- Instale el acelerador del núcleo antes de instalar otro acelerador.
- Instale el acelerador de dominios de datos de núcleo antes de instalar el acelerador de dominios de datos.
- Si importa un archivo de metadatos que contiene un objeto que es común a un acelerador que ya importó antes, reemplace el objeto en el repositorio.

- Para utilizar las reglas del acelerador que realizan la validación de direcciones, descargue e instale los archivos de datos de referencia de direcciones del país que especifique el acelerador. Para utilizar las reglas del acelerador que realizan el análisis de coincidencia de identidades, descargue e instale los archivos de poblaciones de identidad del país que especifique el acelerador. Los archivos de datos de referencia de direcciones y los archivos de poblaciones de identidad se pueden adquirir en Informatica.

Importar reglas y asignaciones

Utilice Developer tool para importar los metadatos para reglas, asignaciones de demostración y orígenes de datos de asignaciones. Durante la operación de importación, seleccione el archivo de datos de referencia que utilicen las reglas y asignaciones.

1. En Developer tool, conéctese al repositorio de modelos que contiene el proyecto de destino para los metadatos.
2. En el Explorador de objetos, seleccione el proyecto de destino.
Por ejemplo, seleccione el proyecto *Informatica_DQ_Content*. Si es necesario, cree un proyecto en el repositorio de modelos.
3. Seleccione **Archivo > Importar**.
4. En el cuadro de diálogo **Importar**, seleccione **Informatica > Importar archivo de metadatos de objeto (avanzado)**.
5. Haga clic en **Siguiente**.
6. Busque el archivo de metadatos XML en la estructura del directorio del acelerador y selecciónelo.
7. Haga clic en **Abrir** y, a continuación, haga clic en **Siguiente**.
8. En el panel **Origen**, seleccione los elementos que aparecen en el nodo del proyecto.
9. En el panel **Destino**, seleccione el proyecto de destino.
10. Haga clic en **Añadir a destino**.
 - Si el proyecto del repositorio contiene un objeto que desea añadir, Developer tool le pedirá que fusione el objeto con el objeto actual. Haga clic en **Sí** para fusionar los objetos.
 - Si Developer tool le solicita que cambie el nombre de los objetos, haga clic en **No**.
 - Si un objeto permanece en el panel **Origen**, utilice el puntero para mover el objeto al proyecto de destino.
11. Haga clic en **Siguiente**.
12. Busque el archivo de datos de referencia comprimido en la estructura del directorio del acelerador y selecciónelo.
13. Haga clic en **Abrir**.
14. Compruebe que la página de códigos es UTF-8, y haga clic en **Siguiente**.
15. En el campo **Conexión de destino**, seleccione la base de datos de datos de referencia.
16. Haga clic en **Finalizar**.

Importar dominios de datos y grupos de dominios de datos

Utilice el cuadro de diálogo Preferencias para importar metadatos para los dominios de datos y los grupos de dominios de datos. Durante la operación de importación, seleccione el archivo de datos de referencia que utilizan los dominios de datos.

1. En la herramienta del desarrollador, conéctese al repositorio de modelos que contiene el proyecto de destino para los metadatos.

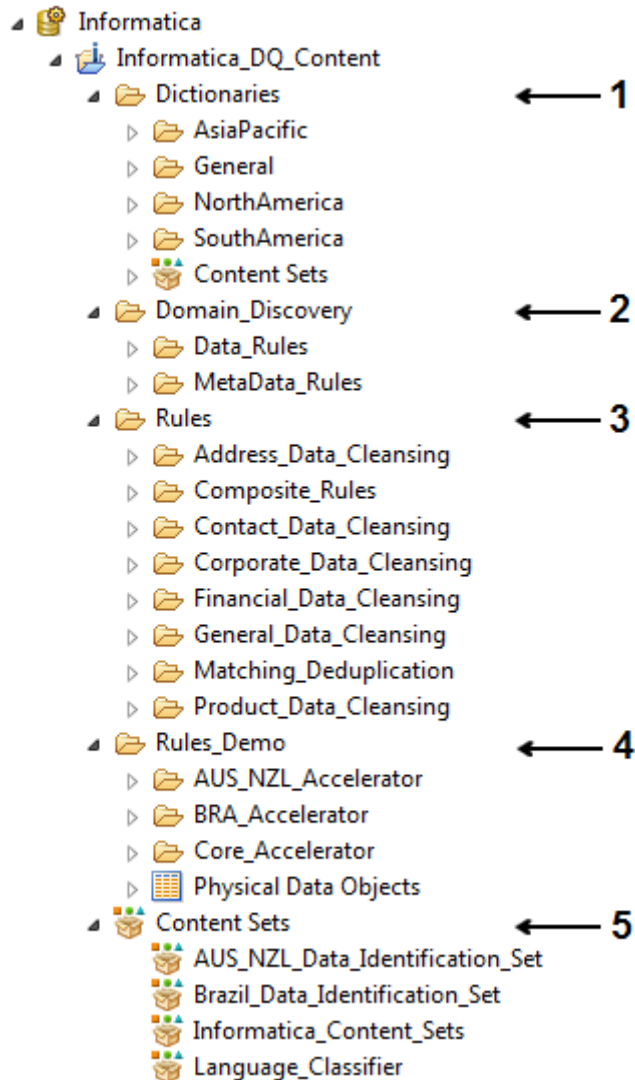
2. Seleccione **Ventana > Preferencias**.
3. En el cuadro de diálogo **Preferencias**, expanda el nodo Informatica y seleccione **Glosario de dominio de datos**.
4. En el panel del repositorio, seleccione el nodo de nivel superior de los dominios de datos o los grupos de dominios de datos.
5. Haga clic en **Importar**.
6. Busque el archivo de metadatos XML en la estructura del directorio del acelerador y selecciónelo.
7. Haga clic en **Abrir** y, a continuación, haga clic en **Siguiente**.
8. En el panel **Origen**, seleccione el proyecto de glosario de dominio de datos.
9. En el panel **Destino**, seleccione el proyecto de destino.
10. Seleccione la siguiente opción en el campo Resolución:
Reemplazar opción en destino
11. Haga clic en **Añadir contenido a destino**.
 - Si la Herramienta del desarrollador le solicita añadir los objetos, haga clic en **Sí**.
 - Si la Herramienta del desarrollador le solicita que cambie el nombre de los objetos, haga clic en **No**.
12. Haga clic en **Siguiente**.
13. Si la operación de importación identifica las dependencias, copie los objetos dependientes del proyecto de origen en el proyecto de destino.
14. Haga clic en **Siguiente**.
15. Busque el archivo de datos de referencia comprimido en la estructura del directorio del acelerador y selecciónelo.
16. Haga clic en **Abrir**.
17. Compruebe que la página de códigos es UTF-8 y haga clic en **Siguiente**.
18. En el campo **Conexión de destino**, seleccione la base de datos de datos de referencia.
19. Haga clic en **Finalizar**.

Componentes de los aceleradores

Al importar un acelerador, Developer tool crea carpetas para las reglas, los dominios de datos y otros objetos que especifique el acelerador. Cada carpeta contiene subcarpetas que organizan los objetos por país y por tipo de operaciones de calidad de datos que realizan.

Utilice el acelerador del núcleo para crear las carpetas en un proyecto de repositorio. Al importar aceleradores adicionales, se añaden objetos y carpetas al proyecto.

La siguiente imagen muestra la estructura de carpetas del proyecto Informatica_DQ_Content al importar varios aceleradores al proyecto:



1. Carpeta Dictionaries
2. Carpeta Domain_Discovery
3. Carpeta Rules
4. Carpeta Rules_Demo
5. Carpeta Content Sets

El proyecto contiene las siguientes carpetas de nivel superior:

Dictionaries

La carpeta Dictionaries contiene objetos de tabla de referencia. Cada objeto hace referencia a una tabla de la base de datos de referencia.

Domain_Discovery

La carpeta Domain_Discovery contiene las reglas que definen los dominios de datos en los aceleradores que instale. La carpeta contiene una carpeta Data_Rules y una carpeta Metadata_Rules. Las reglas de la carpeta Data_Rules corresponden a los dominios de datos que analizan valores de datos de columna.

Las reglas de la carpeta Metadata_Rules corresponden a los dominios de datos que analizan nombres de columna.

Rules

La carpeta Rules contiene las reglas que se utilizan para analizar y mejorar los datos.

Rules_Demo

La carpeta Rules_Demo contiene las asignaciones de demostración y los orígenes de datos de demostración.

Content Sets

En la carpeta Content Sets, se pueden encontrar objetos de datos de referencia que no especifican datos en la base de datos de datos de referencia.

Reglas

Las reglas del acelerador definen un intervalo de operaciones de análisis y transformación de datos. Puede añadir tanto una regla como una serie de reglas a una asignación.

Utilice las reglas del acelerador para realizar las siguientes tareas de calidad de datos:

Validación de direcciones

Validar y mejorar los datos de los registros de direcciones postales. Las reglas requieren archivos de datos de referencia de direcciones.

Análisis de datos

Analizar la información de los registros. Las reglas de análisis pueden extraer varios tipos de información, incluidos nombres de personas, nombres de organizaciones, números de teléfono, fechas y números de identificación.

Estandarización de datos

Estandarizar la ortografía y el formato de los valores de datos. Las reglas de estandarización pueden identificar y corregir varios tipos de información, incluidos nombres de personas, nombres de organizaciones, números de teléfono, fechas y números de identificación.

Análisis duplicado

Detectar registros duplicados en un conjunto de datos. Las reglas de análisis duplicado comparan los registros en un conjunto de datos y generan una puntuación numérica que representa el grado de similitud entre los registros.

Las reglas de análisis duplicado pueden leer registros que contienen datos corporativos generales y registros que contienen datos de identidad. Las reglas de datos de identidad requieren archivos de datos de poblaciones de identidad.

La operación de importación añade las reglas a la siguiente carpeta del repositorio:

```
[Informatica_DQ_Content]\Rules
```

Puede encontrar las reglas que realizan operaciones de validación de direcciones, análisis de datos y estandarización de datos en las subcarpetas *Limpieza de datos* del proyecto del acelerador. Igualmente, puede encontrar las reglas que realizan análisis duplicado en la subcarpeta *Matching_Deduplication* del proyecto del acelerador.

Al importar reglas de un país o una región, se añade una subcarpeta de reglas compuestas. Una regla compuesta combina varias reglas en una única regla de forma anidada.

Asignaciones de demostración

Las asignaciones de demostración son objetos de tiempo de ejecución que aplican una o más reglas a un origen de datos y escriben los resultados en otro origen de datos. Puede utilizar las asignaciones de demostración como plantillas para otras asignaciones.

La operación de importación añade las asignaciones y los objetos de origen de datos a la siguiente carpeta del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo
```

Al importar un acelerador, la operación de importación añade el origen de datos de las asignaciones de demostración a la carpeta `Rules_Demo`. Copie los archivos de origen de datos del directorio `Accelerator_Sources` en el sistema de archivos.

Dominios de datos

Un dominio de datos describe los valores de datos que puede representar un tipo único de información empresarial en una columna. Utilice dominios de datos para determinar el tipo de información de una columna y para encontrar información de un tipo especificado en una columna. Los aceleradores incluyen dominios de datos para una serie de tipos de información, incluidos los números de seguridad social, los números de tarjetas de crédito, las direcciones de correo electrónico y los cargos.

Por ejemplo, una tabla de base de datos podría contener números de seguridad social en una columna Comentarios que puede leer cualquier usuario. Deberá identificar los registros que contienen los números de seguridad social y eliminar o mover estos números. Añada el dominio de datos `SSN` a un perfil y ejecute el perfil en la columna Comentarios.

Puede asignar un dominio de datos a uno o más grupos de dominios de datos. Utilice los grupos de dominios de datos para organizar los dominios de datos según el tipo de análisis empresarial que realizan los dominios de datos. El glosario del dominio de datos muestra una lista de los dominios de datos y los grupos de dominios de datos que se añaden al repositorio de modelos. Utilice el menú **Preferencias** de la herramienta del desarrollador para añadir dominios de datos al glosario del dominio de datos. Para actualizar las definiciones de datos en un dominio de datos, utilice las reglas del acelerador de dominio de datos.

Nota: No podrá ver los objetos del dominio de datos en el Explorador de objetos.

Tablas de referencia

Una tabla de referencia contiene la versión estándar y la alternativa de un conjunto de valores de datos. Las reglas utilizan tablas de referencia para comprobar que los valores de datos son precisos y tienen el formato correcto.

La operación de importación añade las tablas de referencia a la siguiente carpeta del repositorio:

```
[Informatica_DQ_Content]\Dictionaries
```

Conjuntos de contenido

Un conjunto de contenido es un objeto de datos de referencia que no almacena datos en tablas de base de datos. Los conjuntos de contenido incluyen conjuntos de caracteres, conjuntos de patrones, expresiones regulares, conjuntos de tokens, modelos probabilísticos y modelos de clasificador.

La operación de importación añade las reglas a la siguiente carpeta del repositorio:

```
[Informatica_DQ_Content]\Content Sets
```

Nota: Para ver una lista de los elementos en un conjunto de contenido, abra el conjunto de contenido en la herramienta del desarrollador y seleccione la ficha **Etiquetas**.

Etiquetas y reglas

Las reglas del acelerador incluyen etiquetas que indican el tipo de datos que la regla puede leer, así como el tipo de operación que puede realizar.

Para ver las etiquetas que se aplican a una regla, abra la regla en la herramienta del desarrollador y haga clic en la ficha Etiquetas. Puede utilizar las opciones de búsqueda de la herramienta del desarrollador para buscar aceleradores que contengan una etiqueta determinada.

Uso de los aceleradores en PowerCenter

Puede exportar reglas y asignaciones del repositorio de modelos al sistema de archivos y al repositorio de PowerCenter. Cuando exporte objetos, seleccione las tablas de referencia, los objetos de datos y otras dependencias que haya en los objetos que va a exportar.

La operación de exportación copia los datos de la tabla de referencia en el sistema de archivos. Copie los archivos en el equipo host del servicio de integración de PowerCenter. Las ubicaciones de los archivos de datos de referencia en la estructura de directorios de PowerCenter deben corresponder con las ubicaciones de las tablas de referencia en la estructura de carpetas del repositorio de modelos.

La ruta siguiente describe una estructura de directorios de muestra de los objetos de datos de referencia en una instalación de PowerCenter:

```
<Informatica_installation_directory>\services\<Model_repository_project_name>\<Model_repository_project_folder_name>
```

Nota: Si la versión del producto PowerCenter no coincide con la versión de la herramienta del desarrollador, compruebe que el entorno de PowerCenter incluye el complemento de integración de Data Quality.

Para obtener más información acerca de la integración de Data Quality con PowerCenter, consulte la guía del usuario de la integración de Informatica Data Quality para PowerCenter.

CAPÍTULO 2

Acelerador del núcleo

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador del núcleo, 20](#)
- [Reglas de limpieza de datos de direcciones del núcleo, 21](#)
- [Reglas de limpieza de datos de contacto del núcleo, 22](#)
- [Reglas de limpieza de datos corporativos del núcleo, 23](#)
- [Reglas de limpieza de datos generales del núcleo, 23](#)
- [Reglas de coincidencia y eliminación de duplicados del núcleo, 30](#)
- [Reglas de limpieza de datos de producto del núcleo, 30](#)
- [Asignaciones de demostración del núcleo, 31](#)

Resumen del acelerador del núcleo

Utilice las reglas del acelerador del núcleo para comprobar y mejorar los datos empresariales en cualquier país o región.

El acelerador del núcleo incluye reglas que realizan los siguientes procesos de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos corporativos
- Limpieza de datos generales
- Limpieza de datos de coincidencia y eliminación de duplicados
- Limpieza de datos de producto

El acelerador del núcleo contiene mapplets y objetos de datos de referencia que otros aceleradores pueden volver a utilizar. Instale el acelerador del núcleo antes de instalar cualquier otro acelerador.

Reglas de limpieza de datos de direcciones del núcleo

Utilice las reglas de limpieza de datos de direcciones para analizar, estandarizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de direcciones del acelerador del núcleo:

Nombre	Descripción
mplt_Global_AddressValidation5_v2_Discrete_Webservice	Valida direcciones postales de varios países. Utilice el mapplet cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones. El mapplet llama a un servicio web de validación de direcciones. Utilice el mapplet como ejemplo al configurar otros mapplets de servicio web.
mplt_Global_AddressValidation5_v2_Hybrid_Webservice	Valida direcciones postales de varios países. Utilice el mapplet cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones. El mapplet llama a un servicio web de validación de direcciones. Utilice el mapplet como ejemplo al configurar otros mapplets de servicio web.
mplt_Global_AddressValidation5_v2_Multiline_Webservice	Valida direcciones postales de varios países. Utilice el mapplet cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones. El mapplet llama a un servicio web de validación de direcciones. Utilice el mapplet como ejemplo al configurar otros mapplets de servicio web.
rule_Calc_Distance_Between_Geocoordinates	Calcula la distancia entre dos conjuntos de coordenadas geográficas.
rule_Country_Identification	Identifica un país.
rule_Country_Name_Standardization	Estandariza nombres de países. La regla devuelve un nombre de país, así como códigos de país ISO de dos y de tres caracteres.
rule_Geoocordinate_In_Polygon	Comprueba la presencia de puntos de coordenadas geográficas en un área que dos o más puntos de coordenadas geográficas han definido.
rule_Global_Address_Parse_Hybrid	Divide direcciones sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_Global_Address_Parse_Multiline	Divide direcciones sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Nombre	Descripción
rule_Global_Address_Validation_Discrete	Valida la factibilidad de entrega de los registros de direcciones de varios países. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_Global_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de registros de direcciones de varios países y añade coordenadas de latitud y longitud a cada una de las direcciones de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_Global_Address_Validation_Hybrid	Valida la factibilidad de entrega de los registros de direcciones de varios países. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_Global_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de registros de direcciones de varios países y añade coordenadas de latitud y longitud a cada una de las direcciones de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_Global_Address_Validation_Multiline	Valida la factibilidad de entrega de los registros de direcciones de varios países. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_Global_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de registros de direcciones de varios países y añade coordenadas de latitud y longitud a cada una de las direcciones de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Reglas de limpieza de datos de contacto del núcleo

Utilice las reglas de limpieza de datos de contacto para analizar y validar datos sobre contactos empresariales e individuos.

Puede encontrar las reglas de limpieza de datos de direcciones de contacto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de contacto del acelerador del núcleo:

Nombre	Descripción
rule_Email_Parse	Analiza direcciones de correo electrónico en campos de datos.
rule_Email_Parse_and_Validate	Analiza direcciones de correo electrónico en campos de datos y valida el formato de cada dirección de correo electrónico.
rule_Email_Parse_Into_Mailbox_Domain	Divide direcciones de correo electrónico en puertos de buzón, dominio y subdominio. Por ejemplo, la regla divide <i>info@informatica.com</i> de la siguiente manera: <ul style="list-style-type: none">- Buzón: info- Subdominio: informatica- Dominio: com
rule_Email_Validation	Valida el formato de direcciones de correo electrónico. Esta regla no comprueba que las direcciones de correo electrónico sean precisas o estén activas. La regla devuelve Válido o No válido.
rule_Identify_Suspect_Names	Identifica nombres que podrían no ser nombres de persona genuinos. La regla compara los valores de entrada con una tabla de referencia de nombres que es poco probable que sean genuinos. Por ejemplo, la tabla de referencia incluye los nombres de personajes ficticios.

Reglas de limpieza de datos corporativos del núcleo

Utilice las reglas de limpieza de datos corporativos en el acelerador del núcleo para estandarizar datos corporativos.

Puede encontrar las reglas de limpieza de datos corporativos en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos corporativos del acelerador del núcleo:

Nombre	Descripción
rule_Company_Name_Standardization	Utiliza tablas de referencia para estandarizar nombres de empresas.

Reglas de limpieza de datos generales del núcleo

Utilice las reglas de limpieza de datos generales para analizar, estandarizar y validar datos.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador del núcleo:

Nombre	Descripción
mplt_Parse_Tokens_Into_Single_Field	Divide cada palabra en una cadena delimitada por espacios para un puerto independiente.
rule_Add_Leading_Zero	Añade el número "0" al inicio de una cadena.
rule_Add_Parentheses_At_Start_End_ofLine	Añade paréntesis al inicio y al final de una cadena.
rule_Add_Plus_To_Start_of_Line	Añade el signo más al comienzo de una cadena.
rule_Add_Space_Around_Ampersand	Añade un espacio antes y uno después de todos los símbolos et (&) de una cadena.
rule_Add_Space_Around_Hyphen	Añade un espacio antes y uno después de todas las rayas y los guiones de una cadena.
rule_Add_Space_Between_Number_Letter	Añade un espacio entre un par de caracteres que se compone de caracteres numéricos y alfabéticos. El mapplet lee de izquierda a derecha y añade un espacio al primer par de caracteres que se compone de un carácter numérico y un carácter alfabético en los datos.
rule_Add_Spaces_Around_Period	Añade un espacio antes y uno después de todos los puntos de una cadena.
rule_AllTrim	Quita todos los espacios iniciales y finales de los campos de datos de entrada.
rule_Assign_DQ_AddressResolutionCode_Description	Asigna una descripción a la salida del código de resolución de direcciones de la transformación del validador de direcciones.
rule_Assign_DQ_ElementInputStatus_Description	Asigna una descripción a la salida del estado de entrada de elementos de la transformación del validador de direcciones. La descripción corresponde a la salida de las transformaciones de Data Quality en versiones anteriores a Data Quality 9.0.
rule_Assign_DQ_ElementRelevance_Description	Asigna una descripción a la salida de la relevancia de elementos de la transformación del validador de direcciones. La descripción corresponde a la salida de las transformaciones de Data Quality en versiones anteriores a Data Quality 9.0.
rule_Assign_DQ_ElementResultStatus_Description	Asigna una descripción a la salida del estado de resultado de elementos de la transformación del validador de direcciones. La descripción corresponde a la salida de las transformaciones de Data Quality en versiones anteriores a Data Quality 9.0.
rule_Assign_DQ_ExtendedElementStatus_Description	Asigna una descripción a la salida del estado del resultado de elementos extendido de la transformación del validador de direcciones.
rule_Assign_DQ_GeocodingStatus_Description	Asigna una descripción a la salida del estado de geocodificación de la transformación del validador de direcciones. La descripción corresponde a la salida de las transformaciones de Data Quality en versiones anteriores a Data Quality 9.0.

Nombre	Descripción
rule_Assign_DQ_Mailability_Score_Description	Asigna una descripción a la salida de puntuación de viabilidad de envío de correo de la transformación del validador de direcciones. La descripción corresponde a la salida de las transformaciones de Data Quality en versiones anteriores a Data Quality 9.0.
rule_Assign_DQ_Match_Code_Description	Asigna una descripción a la salida del código de coincidencia de la transformación del validador de direcciones. La descripción corresponde a la salida de las transformaciones de Data Quality en versiones anteriores a Data Quality 9.0.
rule_Classify_Language	<p>Clasifica una cadena como uno de los siguientes idiomas: alemán, árabe, español, francés, inglés, italiano, neerlandés, portugués, ruso o turco. La regla utiliza el conjunto de contenido Language_Classifier para identificar los idiomas.</p> <p>Nota: La regla devuelve un idioma por cada cadena que analice. Si una cadena pertenece a un idioma que la regla no reconoce, la regla devuelve el idioma que más se asemeje al texto de la cadena.</p>
rule_Compare_Dates	<p>Calcula la diferencia entre dos fechas. El mapplet utiliza las siguientes unidades de medida:</p> <ul style="list-style-type: none"> - Horas - Días - Meses - Años <p>Cada valor de salida excluye los demás valores. No se pueden sumar las salidas para representar la diferencia entre los valores de datos.</p>
rule_Completeness	Comprueba si hay valores NULL en un solo puerto. Devuelve "Completo" si el puerto contiene datos. Devuelve "Incompleto" si el puerto está vacío o contiene un valor NULL.
rule_Completeness_Multi_Port	Comprueba si hay valores NULL en varios puertos. Devuelve "Completo" si todos los puertos contienen datos. Devuelve "Incompleto" si cualquiera de los puertos está vacío o contiene un valor NULL.
rule_Concatenate_Words	Concatena dos campos. Utiliza un espacio como separador.
rule_Convert_Match_Codes_to_Legacy_Values	Convierte la salida del puerto de código de coincidencia de una transformación del validador de direcciones en el código de coincidencia de validación de direcciones equivalente en Data Quality 8.6.

Nombre	Descripción
rule_CreditCard_Number_Validation	<p>Valida números de tarjeta de crédito para tarjetas de crédito que utilizan el algoritmo de Luhn. La validación incluye, entre otras, las siguientes tarjetas de crédito:</p> <ul style="list-style-type: none"> - American Express - Diners Club Carte Blanche - Diners Club International - Diners Club US & Canada - Discover Card - JCB - Maestro - Master Card - Solo - Switch - Visa - Visa Electron <p>La regla devuelve "Válido" o "No válido".</p>
rule_Date_Complete	<p>Comprueba que la cadena de entrada se ajuste a un formato de fecha que la regla reconozca. La regla lee el siguiente objeto de datos de referencia:</p> <ul style="list-style-type: none"> - user_defined_dates_infa
rule_Date_of_Birth_Validation	<p>Comprueba la cantidad de años que existe entre una fecha de nacimiento y la fecha actual. Devuelve "Adulto" o "Menor de edad", así como "Válido" si el número de años es igual o menor a 120. Devuelve "No válido" si el número de años es mayor a 120.</p>
rule_Date_Parse	<p>Analiza los datos de fecha de una cadena para un puerto que especifica la regla. La regla reconoce fechas con los siguientes formatos:</p> <ul style="list-style-type: none"> - dd/mm/aaaa - mm/dd/aaaa - aaaa/dd/mm <p>La regla devuelve una fecha, así como una cadena que contiene el texto de entrada sin la fecha.</p>
rule_Date_Standardization	<p>Estandariza cadenas de fechas según el formato de salida que especifique. Para definir el formato de salida, abra la transformación de expresión dq_FormatDate de la regla y actualice las variables de expresión Output_Date_Format y Delimiter. Si los datos de entrada no describen una fecha válida, la regla devuelve 0 para cada carácter de entrada.</p>
rule_Date_Validation	<p>Valida cadenas de fechas que aparecen en un formato único en una columna de datos. Para configurar el formato de fecha que la regla utiliza para realizar la validación, abra la transformación de expresión dq_ValidateDate de la regla y actualice la variable de expresión In_Date_Format. El formato es "MM/DD/AAAA". La regla devuelve "Válido" o "No válido".</p>

Nombre	Descripción
rule_Date_Validation_Variable_Format	<p>Valida cadenas de fechas que aparecen en varios formatos en una columna de datos. Utilice la regla cuando un origen de datos incluya las siguientes columnas:</p> <ul style="list-style-type: none"> - Una columna que contiene valores de fecha en varios formatos. - Una columna que identifica el formato del valor de fecha en cada fila. Si la columna no identifica un formato de fecha para una fila, la regla aplica el formato "MM/DD/AAAA" al valor de fecha. <p>La regla lee todos los valores de datos que reconoce la función <i>is_date()</i>. La regla devuelve "Válido" o "No válido".</p>
rule_Days_Between_Dates	Calcula la cantidad de días que existe entre dos fechas.
rule_Days_From_Current_Date	Calcula la cantidad de días que existe entre una fecha determinada y la fecha actual.
rule_EAN13_Algorithm	Valida un número de artículo internacional. La regla devuelve "Válido" si el dígito de comprobación es correcto para el número y "No válido" si el dígito de comprobación es incorrecto.
rule_GTIN_Validation	Valida un GTIN (Global Trade Item Number, número mundial de un artículo comercial). La regla valida números compuestos de ocho, doce, trece o catorce dígitos. La regla devuelve "Válido" si el dígito de comprobación es correcto para el número y "No válido" si el dígito de comprobación es incorrecto.
rule_IsNumeric	Confirma que los datos de entrada son numéricos. La regla devuelve "True" o "False".
rule_LowerCase	Devuelve todos los caracteres alfabéticos en minúsculas.
rule_Luhn_Algorithm	Aplica el algoritmo de Luhn a una cadena numérica. La regla puede validar cadenas numéricas, tales como números de tarjetas de crédito.
rule_Mask_Profanity	Busca lenguaje obsceno en los datos de entrada. Enmascara los casos de lenguaje obsceno con la palabra "CENSORED" en los datos de salida.
rule_Negative_Number_Validation	Comprueba que los datos de entrada sean números negativos.
rule_Numeric_Completeness	Busca valores NULL en las entradas numéricas.
rule_Parse_Alpha_Chars_from_Non_Alpha_Chars	<p>Identifica los caracteres alfabéticos y no alfabéticos en una cadena de entrada y escribe cada conjunto de caracteres en distintos puertos de salida. Por ejemplo, la regla analiza los siguientes valores desde la cadena de entrada <code>teststring_123</code>:</p> <pre>teststring _123</pre>
rule_Parse_First_Word	Analiza la primera palabra en una cadena de entrada para un puerto que especifica la regla.
rule_Parse_Number_At_End_Of_Line	Analiza cualquier número que aparezca al final de una cadena de entrada para un puerto que especifica la regla. La regla lee las cadenas de izquierda a derecha.

Nombre	Descripción
rule_Parse_Number_At_Start_Of_Line	Analiza cualquier número que se aparezca al inicio de una cadena de entrada para un puerto que especifica la regla. La regla lee las cadenas de izquierda a derecha.
rule_Parse_Profanity	Compara cadenas a una tabla de referencia de términos profanos y analiza cualquier término que coincida con un valor de la tabla de referencia para un puerto que especifica la regla.
rule_Parse_Text_Between_Parentheses	Analiza cadenas que están delimitadas por paréntesis para un puerto que especifica la regla. La regla contiene un puerto de salida para las cadena analizadas y un puerto de salida para el texto de entrada sin las cadenas analizadas.
rule_Parse_Text_in_Single_Quotes	Analiza cadenas que están delimitadas por comillas para un puerto que especifica la regla. Cuando los datos de entrada contienen varios elementos entre comillas, la regla analiza el último elemento. La regla lee las cadenas de entrada de izquierda a derecha. La regla contiene un puerto de salida para las cadena analizadas y un puerto de salida para el texto de entrada sin las cadenas analizadas.
rule_Past_Date_Label	Determina si una fecha de entrada es anterior o posterior a la fecha del sistema.
rule_Personal_Company_Identification	Analiza nombres de persona y nombres de empresa en los diferentes puertos que especifica la regla. La regla tiene las siguientes salidas: <ul style="list-style-type: none"> - Nombre de persona - Nombre de empresa - Categoría de datos, como nombre de persona o nombre de empresa - Datos que la regla no puede analizar
rule_Postive_Number_Validation	Confirma que los datos de entrada sean números positivos.
rule_Prepending_Zero_to_Single_Digit	Añade el número "0" a caracteres numéricos simples.
rule_Remove_All_Leading_Zeros	Elimina todas las instancias de caracteres numéricos "0" que se encuentren al inicio de una cadena.
rule_Remove_Apostrophe	Quita los apóstrofes. La regla fusiona las cadenas de texto a cada lado del apóstrofo.
rule_Remove_Control_Characters	Elimina los caracteres de control de las cadenas de texto. La regla devuelve una cadena que contiene los caracteres de control, así como una cadena que contiene el texto de entrada sin los caracteres de control.
rule_Remove_Extra_Spaces	Reemplaza todos los espacios consecutivos con un espacio simple y elimina los espacios iniciales y finales.
rule_Remove_Hyphen	Elimina los guiones.
rule_Remove_Leading_Zero	Elimina una sola instancia del carácter numérico "0" del inicio de una cadena.

Nombre	Descripción
rule_Remove_Limited_Punctuation	Elimina los caracteres superfluos. Los caracteres superfluos incluyen barras diagonales, barras diagonales inversas, puntos, signos de exclamación, guiones bajos y espacios consecutivos múltiples.
rule_Remove_Non_Numbers	Quita todos los caracteres que no son numéricos.
rule_Remove_Parentheses	Quita los paréntesis izquierdos y derechos.
rule_Remove_Period	Quita los puntos.
rule_Remove_Period_Parentheses	Quita los siguientes caracteres: - Paréntesis izquierdo y derecho - Puntos
rule_Remove_Punctuation	Quita los signos de puntuación.
rule_Remove_Punctuation_and_Space	Quita todos los signos de puntuación y los espacios.
rule_Remove_Quotation	Quita las comillas.
rule_Remove_Slashes	Quita las barras diagonales y las barras diagonales inversas.
rule_Remove_Space	Quita todos los espacios.
rule_Replace_Ampersand_With_Space	Reemplaza los símbolos et (&) por espacios.
rule_Replace_Hyphen_Underscore_with_Space	Reemplaza los guiones y los guiones bajos por espacios.
rule_Replace_Hyphen_with_Space	Reemplaza los guiones por espacios.
rule_Replace_Limited_Punct_with_Space	Reemplaza los siguientes caracteres de puntuación por un espacio simple: raya, barra diagonal inversa, punto, signo de exclamación y guión bajo. La regla también reemplaza entre dos y cuatro espacios consecutivos por un espacio simple.
rule_Replace_Non_Alphabetic_with_Space	Reemplaza los números y los signos de puntuación por un espacio simple.
rule_Replace_Period_With_Space	Reemplaza los puntos por espacios simples.
rule_Replace_Punctuation_with_Space	Reemplaza todos los signos de puntuación por espacios.
rule_Replace_Slashes_With_Space	Reemplaza las barras diagonales y las barras diagonales inversas por espacios.
rule_Reverse_String_Input	Invierte el orden de los caracteres en las cadenas de entrada.
rule_String_Completeness	Comprueba la integridad de una cadena. La regla también busca valores de la tabla de referencia string_default_values_inf en las cadenas de entrada. La tabla de referencia contiene valores como NA, DEFAULT y XX. Si una cadena de entrada contiene un valor de la tabla de referencia, la regla determina que la cadena está incompleta.

Nombre	Descripción
rule_TitleCase	Convierte las cadenas en una cadena de tipo título. En las cadenas de tipo título, la primera letra de cada palabra se escribe en mayúscula.
rule_Translate_Diacritic_Characters	Reemplaza los caracteres diacríticos por equivalentes ASCII. Por ejemplo, esta regla convierte "ã" en "a".
rule_UpperCase	Devuelve todos los caracteres alfabéticos en mayúsculas.
rule_URL_Validation	Valida el formato y la estructura de una dirección URL.
rule_Years_Since_Date_of_Birth	Calcula los años que han transcurrido desde la fecha de entrada.

Reglas de coincidencia y eliminación de duplicados del núcleo

Utilice las reglas de coincidencia y eliminación de duplicados para identificar registros duplicados.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador del núcleo:

Nombre	Descripción
mplt_Consolidate_and_Remove_Duplicate_Rows	Reúne clústeres de registros duplicados en un único registro y quita registros duplicados redundantes.

Reglas de limpieza de datos de producto del núcleo

Utilice las reglas de limpieza de datos de producto para analizar, estandarizar y validar datos de producto.

Puede encontrar las reglas de limpieza de datos de producto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Product_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de producto del acelerador del núcleo:

Nombre	Descripción
rule_Color_Parse	Analiza los valores de color para un puerto que especifica la regla.
rule_Parse_Quantity_And_UOM	Analiza la primera instancia de una cantidad y una unidad de medida de una cadena para un puerto que especifica la regla. La regla lee la cadena de izquierda a derecha y devuelve los siguientes datos: <ul style="list-style-type: none">- Cantidad.- Unidad de medida.- La cadena de entrada sin los valores de cantidad ni de unidad de medida.
rule_UOM_Standardization	Estandariza una unidad de medida. La regla devuelve valores estandarizados y no estandarizados de la cantidad y la unidad de medida. También devuelve una cadena que contiene el texto de entrada con una unidad de medida estandarizada.
rule_UPC_Validation	Valida un código de producto universal y devuelve un código de producto universal estandarizado.

Asignaciones de demostración del núcleo

Las asignaciones de demostración del acelerador del núcleo utilizan varias reglas para comprobar los procesos de calidad de datos.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\Core_Accelerator
```

El acelerador contiene las siguientes asignaciones de demostración:

m_customer_data_demo

Analiza, estandariza y valida datos de Estados Unidos y Canadá.

m_product_demo

Analiza descripciones de productos y valida la calidad de las descripciones.

CAPÍTULO 3

Acelerador de dominios de datos

Este capítulo incluye los siguientes temas:

- [Introducción al acelerador de dominios de datos, 32](#)
- [Los dominios de datos en el acelerador de dominios de datos, 33](#)
- [Reglas de nombre de columna en el acelerador de dominios de datos, 41](#)
- [Reglas de datos en el acelerador de dominios de datos, 45](#)

Introducción al acelerador de dominios de datos

Un dominio de datos es un objeto del repositorio de modelos predefinido o definido por el usuario que utiliza reglas para detectar el significado funcional de los datos de columna o los nombres de columna. Las reglas de dominios de datos definen patrones de datos y patrones de nombre de columna que coinciden con los datos y los metadatos de origen. Puede utilizar las reglas de dominios de datos para actualizar la lógica de los dominios de datos.

Utilice los dominios de datos del acelerador de dominios de datos para detectar el significado funcional de los datos de origen según los valores de nombre de columna o dato de columna.

El acelerador de dominios de datos incluye los siguientes tipos de regla:

- Regla de datos. Busca columnas con los datos que coinciden con la lógica que la regla define.
- Regla de nombre de columna. Busca columnas con los nombres de columna que coinciden con la lógica column-name que la regla define.

Las reglas de dominios de datos devuelven valores booleanos que indican si los datos de columna o el nombre de columna cumplen con los criterios de la regla. Las reglas de dominios de datos utilizan expresiones regulares o tablas de referencia para buscar valores o patrones específicos. Por ejemplo, puede utilizar una expresión de regla de nueve dígitos para buscar valores de datos en el formato de número de seguridad social.

Cuando utilice expresiones en las reglas de dominios de datos, algunos de los valores de datos no relacionados también podrían cumplir con los criterios de la expresión de regla. Por ejemplo, los códigos postales de los Estados Unidos en los datos de origen podrían cumplir con el formato de número de seguridad social. Para que la deducción de dominios de datos sea eficaz, deberá revisar si existen discrepancias en los resultados de detección del dominio de datos. Después de revisar y comprobar los resultados de detección del dominio de datos, puede decidir asociar un dominio de datos con una columna de datos.

Los dominios de datos en el acelerador de dominios de datos

Utilice los dominios de datos predefinidos en perfiles para realizar la detección del dominio de datos e identificar características de datos críticos dentro de una empresa.

Nota: En la tabla, el símbolo de asterisco (*) es un carácter comodín.

La siguiente tabla describe los dominios de datos disponibles en el acelerador de dominios de datos:

Nombre	Descripción	Tipo de regla dependiente	Grupo de dominio de datos
Account_Status	Detecta datos de columna que coinciden con los valores del estado de cuenta en una tabla de referencia.	Regla de datos	Account_Bank
AccountNumber	Detecta nombres de columna que contienen las cadenas "a*c*num" o "acc".	Regla de nombre de columna	Account_Bank, PCI, PHI
Age	Detecta nombres de columna que contienen las cadenas "age" o "dob", y detecta los datos de columna con valores de 1 a 120.	Regla de nombre de columna Regla de datos	PII
AlphaNumeric_SpecialCharacter	Detecta datos de columna que contienen datos alfanuméricos sin formato y datos con caracteres especiales.	Regla de datos	General
Austria_NationalID	Detecta datos de columna que coinciden con el ID nacional de Austria.	Regla de datos	NationalID
BIC_SwiftCode	Detecta datos de columna que coinciden con los códigos BIC (Bank Identifier Code, código de identificación del banco) o SWIFT (Sociedad de Telecomunicaciones Financieras Interbancarias Mundiales) mediante reconocimiento de patrones y códigos de país.	Regla de datos	Account_Bank
BinayValue	Detecta datos de columna que contienen valores binarios.	Regla de datos	General
BirthDay	Detecta nombres de columna que contienen las cadenas "dob", "date*of*bir*" o "birth*da*", e identifica los datos de columna que coinciden con fechas de nacimiento válidas.	Regla de nombre de columna Regla de datos	PII

Nombre	Descripción	Tipo de regla dependiente	Grupo de dominio de datos
BirthPlace	Detecta nombres de columna que contienen las cadenas "birth*place" o "location*birth".	Regla de nombre de columna	PII
Brazil_IDDoc	Detecta datos de columna que coinciden con el formato de número de la tarjeta de ID de Brasil que se llama Registro Geral.	Regla de datos	NationalID
Brazil_NationalID	Detecta datos de columna que coinciden con el ID nacional de Brasil.	Regla de datos	NationalID
Bulgaria_NationalID	Detecta datos de columna que coinciden con el ID nacional de Bulgaria.	Regla de datos	NationalID
Canada_SIN	Detecta datos de columna que coinciden con el formato de número de seguridad social de Canadá.	Regla de datos	NationalID
CertificateLicenseNumber	Detecta nombres de columna que contienen las cadenas "cert*lic*number", "cert*lic*no*", "lic*nu*" o "lic*no*".	Regla de nombre de columna	PHI
China_NationalID	Detecta datos de columna que coinciden con el ID nacional de China.	Regla de datos	NationalID
CompanyName	Detecta nombres de columna que contienen la cadena "Company" e identifica los datos de columna que coinciden con los valores de nombre de organización en una tabla de referencia.	Regla de nombre de columna Regla de datos	Contact, PII
ComputerAddress	Detecta datos de columna que coinciden con formatos de dirección de equipos.	Regla de datos	General
Country	Detecta nombres de columna que contienen las cadenas "iso*countr*code", "iso*country" o "countr*", e identifica los datos de columna que coinciden con nombres de país.	Regla de nombre de columna Regla de datos	Address, PII
CountryCode_Phone	Detecta números de teléfono según códigos de marcación internacional.	Regla de datos	Contacto

Nombre	Descripción	Tipo de regla dependiente	Grupo de dominio de datos
CreditCardNumber	Detecta nombres de columna que contienen las cadenas "ccn", "cr*ca*nu" o "credit*no*", e identifica los datos de columna que coinciden con el formato de número de tarjeta de crédito de varias organizaciones de tarjetas de crédito.	Regla de nombre de columna Regla de datos	Account_Bank, PCI, PII
CreditCard_AMEX	Detecta datos de columna que coinciden con el formato de número de tarjeta de crédito American Express.	Regla de datos	Account_Bank
CreditCard_DinersCard	Detecta datos de columna que coinciden con el formato de número de tarjeta de crédito Diners Club International.	Regla de datos	Account_Bank
CreditCard_Discover Card	Detecta datos de columna que coinciden con el formato de número de tarjeta de crédito Discover.	Regla de datos	Account_Bank
CreditCard_JCB	Detecta datos de columna que coinciden con el formato de número de tarjeta de crédito JCB International.	Regla de datos	Account_Bank
CreditCard_Master Card	Detecta datos de columna que coinciden con el formato de número de tarjeta de crédito MasterCard.	Regla de datos	Account_Bank
CreditCard_Visa	Detecta datos de columna que coinciden con el formato de número de tarjeta de crédito Visa.	Regla de datos	Account_Bank
Croatia_NationalID	Detecta datos de columna que coinciden con el ID nacional de Croacia.	Regla de datos	NationalID
Date_AllFormats	Detecta valores de fecha en los datos de columna.	Regla de datos	General
Date_MM_DD_AAAA	Detecta cadenas de fechas en los datos de origen que aparecen en un formato único en una columna de fecha. El formato es "MM/DD/AAAA".	Regla de datos	General
Denmark_NationalID	Detecta datos de columna que coinciden con el ID nacional de Dinamarca.	Regla de datos	NationalID

Nombre	Descripción	Tipo de regla dependiente	Grupo de dominio de datos
DeviceSerialNumber	Detecta nombres de columna que contienen las cadenas "device*number", "device*no*", "serial*number", "serial*no*" o "device*identi*".	Regla de nombre de columna	PHI
DriverLicense_Canada	Detecta datos de columna que coinciden con los números del permiso de conducción de Canadá, excepto para las provincias de Columbia Británica, Quebec, Manitoba e Isla del Príncipe Eduardo.	Regla de datos	NationalID
DriverLicense_GBR	Detecta datos de columna que coinciden con los números del permiso de conducción del Reino Unido.	Regla de datos	NationalID
DriverLicense_USA	Detecta datos de columna que coinciden con los números del permiso de conducción de la mayoría de los estados de Estados Unidos.	Regla de datos	NationalID
DrivingLicenseNumber	Detecta nombres de columna que contienen las cadenas "license" o "driver*license". Identifica los datos de columna que coinciden con los números de la licencia de conducir del Reino Unido, los Estados Unidos y algunas provincias canadienses según los requisitos de longitud y patrón.	Regla de nombre de columna Regla de datos	PII
Email	Detecta nombres de columna que contienen la cadena "email" e identifica los datos de columna que coinciden con un formato de ID de correo electrónico predefinido.	Regla de nombre de columna Regla de datos	Contact, PHI
ExpirationDate	Detecta nombres de columna que contienen las cadenas "exp*da*" o "cr*exp*", e identifica los datos de columna que coinciden con fechas de tarjeta de crédito caducadas.	Regla de nombre de columna Regla de datos	PCI
Finland_NationalID	Detecta datos de columna que coinciden con el ID nacional de Finlandia.	Regla de datos	NationalID

Nombre	Descripción	Tipo de regla dependiente	Grupo de dominio de datos
FirstName	Detecta nombres de columna que contienen la cadena "f*nam*" e identifica los datos de columna que coinciden con los valores en una tabla de referencia de nombres de pila.	Regla de nombre de columna Regla de datos	Contact, PCI, PII
France_INSEE	Detecta datos de columna que coinciden con el formato de número del Institut National de la Statistique et des Études Économiques (INSEE) de Francia.	Regla de datos	NationalID
Gender	Detecta nombres de columna que contienen la cadena "gender" o cadenas como "female" y "male", e identifica los datos de columna que coinciden con los valores de sexo en una tabla de referencia.	Regla de nombre de columna Regla de datos	Contact, PII
Geocode_Latitude	Detecta nombres de columna que contienen la cadena "latitude" e identifica los datos de columna que coinciden con las coordenadas de latitud válidas.	Regla de nombre de columna Regla de datos	Address, General
Geocode_LatitudeLongitude	Detecta los nombres de columna que contienen cadenas como "latitude", "longitude" y "geocode", e identifica los datos de columna que coinciden con coordenadas de latitud o longitud válidas.	Regla de nombre de columna Regla de datos	Address, General
Geocode_Longitude	Detecta nombres de columna que contienen la cadena "longitude" e identifica los datos de columna que coinciden con las coordenadas de longitud válidas.	Regla de nombre de columna Regla de datos	Address, General
Grade	Detecta nombres de columna que contienen la cadena "grade".	Regla de nombre de columna	PII
GreatBritian_NINO	Detecta datos de columna que coinciden con el formato de número de seguridad social de Gran Bretaña.	Regla de datos	NationalID
Nombre de host	Detecta datos de columna que coinciden con nombres de host válidos.	Regla de datos	General

Nombre	Descripción	Tipo de regla dependiente	Grupo de dominio de datos
IBAN	Detecta datos de columna que coinciden con el formato del Código Internacional de Cuenta Bancaria de varios países europeos.	Regla de datos	Account_Bank
India_NationalID	Detecta datos de columna que coinciden con el formato de número de cuenta permanente de la India.	Regla de datos	NationalID
IPAddress	Detecta nombres de columna que contienen las cadenas "ip" o "inter*port*add", e identifica los datos de columna que coinciden con un formato de dirección de IP predefinido.	Regla de nombre de columna Regla de datos	PII
ISBN	Detecta datos de columna que coinciden con el formato del Número estándar internacional de libros.	Regla de datos	General
Italy_FiscalCode	Detecta datos de columna que coinciden con el ID nacional de Italia.	Regla de datos	NationalID
JobPosition	Detecta nombres de columna que contienen las cadenas "title", "position" o "designation".	Regla de nombre de columna	PII
Korea_NationalID	Detecta datos de columna que coinciden con el ID nacional de Corea.	Regla de datos	NationalID
LastName	Detecta nombres de columna que contienen las cadenas "lname", "su*name" o "last*name", e identifica los datos de columna que coinciden con los valores en una tabla de referencia de apellidos.	Regla de nombre de columna Regla de datos	Contact, PCI, PII
Norway_NationalID	Detecta datos de columna que coinciden con el ID nacional de Noruega.	Regla de datos	NationalID
Passport_DEU_MR	Detecta datos de columna que coinciden con el formato de número de pasaporte de lectura automática de Alemania.	Regla de datos	NationalID
Passport_GBR	Detecta datos de columna que coinciden con el formato de número de pasaporte del Reino Unido.	Regla de datos	NationalID

Nombre	Descripción	Tipo de regla dependiente	Grupo de dominio de datos
Passport_India	Detecta datos de columna que coinciden con el formato de número de pasaporte de la India.	Regla de datos	NationalID
Passport_MachineReadable	Detecta datos de columna que coinciden con el formato de número de pasaporte de lectura automática de todos los países.	Regla de datos	NationalID
Passport_USA_MR	Detecta datos de columna que coinciden con el formato de número de pasaporte de lectura automática de Estados Unidos.	Regla de datos	NationalID
PhoneNumber	Detecta nombres de columna que contienen las cadenas "phone" o "fax", e identifica los datos de columna que coinciden con el formato de número de teléfono de Estados Unidos.	Regla de nombre de columna Regla de datos	Contact, PHI
Postcode	Detecta datos de columna que coinciden con los códigos postales de varios países.	Regla de datos	Address
Romania_NationalID	Detecta datos de columna que coinciden con el ID nacional de Rumanía.	Regla de datos	NationalID
Salary	Detecta nombres de columna que contienen las cadenas "compensation", "salary" o "wages".	Regla de nombre de columna	PII
SouthAfrica_NationalID	Detecta datos de columna que coinciden con el ID nacional de Sudáfrica.	Regla de datos	NationalID
SSN	Detecta nombres de columna que contienen las cadenas "SSN", "social*sec*no" o "social*sec*num*", e identifica los datos de columna que coinciden con el formato de número de seguridad social.	Regla de nombre de columna Regla de datos	NationalID, PHI
SSN_General	Detecta datos de columna que coinciden con el formato de número de seguridad social.	Regla de datos	NationalID, PII
State	Detecta nombres de columna que contienen las cadenas "add*sta", "state" o "us*sta*", e identifica los datos de columna que coinciden con los nombres de estado de Estados Unidos.	Regla de nombre de columna Regla de datos	PII

Nombre	Descripción	Tipo de regla dependiente	Grupo de dominio de datos
Street	Detecta nombres de columna que contienen una de las siguientes cadenas: <ul style="list-style-type: none"> - street - road - lane - court - avenue - way - blvd - boule*ard 	Regla de nombre de columna	PII
Sweden_NationalID	Detecta datos de columna que coinciden con el ID nacional de Suecia.	Regla de datos	NationalID
Taiwan_NationalID	Detecta datos de columna que coinciden con el ID nacional de Taiwán.	Regla de datos	NationalID
UniquelidentifyingNumber	Detecta nombres de columna que contienen las cadenas "unique*iden*number" o "iden*num".	Regla de nombre de columna	PHI
UPC	Detecta datos de columna que coinciden con un código de producto universal.	Regla de datos	General
URL	Detecta nombres de columna que contienen las cadenas "uni*res*loc", "URL" o "web", e identifica los datos de columna que coinciden con los formatos de URL predefinidos.	Regla de nombre de columna Regla de datos	PHI
USZip_5digit	Detecta datos de columna que coinciden con códigos postales de Estados Unidos.	Regla de datos	Address
VehicleRegPlateNumber	Detecta nombres de columna que contienen las cadenas "registration", "number*plate", "license*plate" o "vehicle*registration".	Regla de nombre de columna	PII
ZipCode	Detecta nombres de columna que contienen las cadenas "zip" o "pin", e identifica los datos de columna que coinciden con códigos postales de Estados Unidos.	Regla de nombre de columna Regla de datos	PII

Reglas de nombre de columna en el acelerador de dominios de datos

Utilice las reglas de nombres de columnas de dominios de datos para identificar las columnas de datos con nombres que coinciden con la lógica de nombre de columna que definen las reglas. Cada regla utiliza una o más expresiones regulares para buscar cadenas comunes que el nombre de columna pueda incluir.

Por ejemplo, la regla *dataDomain_MetaDataRule_BIC_SWIFTCode* contiene una transformación del etiquetador que busca con las siguientes expresiones regulares:

```
^*[iI][sS][oO].*.[9][3][6][2].*$  
^*[sS][wW][iI][fF][tT]*[bB][iI][cC]$  
^*[bB][iI][cC].*[cC][oO][dD][eE].*$
```

Las reglas de nombre de columna analizan los caracteres de los nombres de columna. Las reglas de nombre de columna no analizan los valores de datos de las columnas.

Puede encontrar las reglas de nombres de columna en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Domain_Discovery\MetaData_Rules
```

La siguiente tabla describe las reglas de nombre de columna en el acelerador de dominios de datos:

Nombre	Descripción
dataDomain_MetaDataRule_ABARoutingNumber	Busca nombres que puedan describir una columna de números de enrutamiento de la Asociación Bancaria de Estados Unidos.
dataDomain_MetaDataRule_AccountNumber	Busca nombres que puedan describir una columna de números de cuenta.
dataDomain_MetaDataRule_AccountStatus	Busca nombres que puedan describir una columna de información de estado de cuentas.
dataDomain_MetaDataRule_Address	Busca nombres que puedan describir una columna de información de direcciones.
dataDomain_MetaDataRule_AdmissionDate	Busca nombres que puedan describir una columna de información de fechas de admisión.
dataDomain_MetaDataRule_Age	Busca nombres que puedan describir una columna de información de edad o fechas de nacimiento.
dataDomain_MetaDataRule_BankAccount	Busca nombres que puedan describir una columna de información de cuentas bancarias.
dataDomain_MetaDataRule_BIC_SwiftCode	Busca nombres que puedan describir una columna de códigos de identificación bancaria. Los códigos de identificación bancaria también se denominan códigos SWIFT y códigos ISO 9362.
dataDomain_MetaDataRule_BirthDay	Busca nombres que puedan describir una columna de información de fecha de nacimiento o cumpleaños.
dataDomain_MetaDataRule_BirthPlace	Busca nombres que puedan describir una columna de información de lugar de nacimiento.
dataDomain_MetaDataRule_CertificateLicenseNumber	Busca nombres que puedan describir una columna de información de número de licencia de certificado.

Nombre	Descripción
dataDomain_MetaDataRule_City	Busca nombres que puedan describir una columna de información de ciudad.
dataDomain_MetaDataRule_CompanyName	Busca nombres que puedan describir una columna de información de nombre de empresa.
dataDomain_MetaDataRule_Computer_Address	Busca nombres que puedan describir una columna de datos de dirección de máquina o MAC.
dataDomain_MetaDataRule_Country	Busca nombres que puedan describir una columna de información de país, incluida la información de códigos de país ISO.
dataDomain_MetaDataRule_CountryCode_Phone	Busca nombres que puedan describir una columna de códigos telefónicos de país.
dataDomain_MetaDataRule_County	Busca nombres que puedan describir una columna de información de condados.
dataDomain_MetaDataRule_CreditCardNumber	Busca nombres que puedan describir una columna de números de tarjeta de crédito.
dataDomain_MetaDataRule_CreditCardTrack1FormatB	Busca nombres que puedan describir una columna de información de pista 1 y formato B de una tarjeta de crédito.
dataDomain_MetaDataRule_Currency	Busca nombres que puedan describir una columna de información de divisa.
dataDomain_MetaDataRule_DeviceSerialNumber	Busca nombres que puedan describir una columna de número de dispositivo o información de número de serie.
dataDomain_MetaDataRule_DischargeDate	Busca nombres que puedan describir una columna de información de fecha de liquidación.
dataDomain_MetaDataRule_DrivingLicenseNumber	Busca nombres que puedan describir una columna de información de permiso de conducir.
dataDomain_MetaDataRule_Email	Busca nombres que puedan describir una columna de información de correo electrónico.
dataDomain_MetaDataRule_ExpirationDate	Busca nombres que puedan describir una columna de información de fecha de caducidad, por ejemplo, información de fecha de caducidad para tarjetas de crédito.
dataDomain_MetaDataRule_FirstName	Busca nombres que puedan describir una columna de información de nombre.
dataDomain_MetaDataRule_FullName	Busca nombres que puedan describir una columna de información de nombre completo.
dataDomain_MetaDataRule_Gender	Busca nombres que puedan describir una columna de información de sexo.
dataDomain_MetaDataRule_Grade	Busca nombres que puedan describir una columna de información de grado.

Nombre	Descripción
dataDomain_MetaDataRule_HealthCareBeneficiaryNumber	Busca nombres que puedan describir una columna de números de beneficiario de atención médica.
dataDomain_MetaDataRule_Height	Busca nombres que puedan describir una columna de información de altura.
dataDomain_MetaDataRule_Hostname	Busca nombres que puedan describir una columna de información del nombre del host de equipos.
dataDomain_MetaDataRule_IBAN	Busca nombres que puedan describir una columna de números de cuentas bancarias internacionales.
dataDomain_MetaDataRule_ICD10	Busca nombres que puedan describir una columna de valores de la décima revisión de la Clasificación Internacional y Estadística de Enfermedades y Problemas Relacionados con la Salud.
dataDomain_MetaDataRule_ICD9	Busca nombres que puedan describir una columna de valores de la novena revisión de la Clasificación Internacional y Estadística de Enfermedades y Problemas Relacionados con la Salud.
dataDomain_MetaDataRule_IPAddress	Busca nombres que puedan describir una columna de información de direcciones IP de equipos.
dataDomain_MetaDataRule_ISBN	Busca nombres que puedan describir una columna de número estándar internacional de libros.
dataDomain_MetaDataRule_ITIN_USA	Busca nombres que puedan describir una columna de número de identificación de contribuyente individual.
dataDomain_MetaDataRule_JobPosition	Busca nombres que puedan describir una columna de información de cargo, puesto o designación.
dataDomain_MetaDataRule_LastName	Busca nombres que puedan describir una columna de información de apellidos.
dataDomain_MetaDataRule_Latitude	Busca nombres que puedan describir una columna de información de latitud.
dataDomain_MetaDataRule_LatitudeLongitude	Busca nombres que puedan describir una columna de información de latitud, longitud o geocoordenadas.
dataDomain_MetaDataRule_Longitude	Busca nombres que puedan describir una columna de información de longitud.
dataDomain_MetaDataRule_MaidenName	Busca nombres que puedan describir una columna de información de nombre de soltero.
dataDomain_MetaDataRule_MiddleName	Busca nombres que puedan describir una columna de información de segundo nombre.
dataDomain_MetaDataRule_NationalId	Busca nombres que puedan describir una columna de números de identidad nacional.
dataDomain_MetaDataRule_NDC_USA	Busca nombres que puedan describir una columna de información del código nacional de medicamentos.

Nombre	Descripción
dataDomain_MetaDataRule_NPI_USA	Busca nombres que puedan describir una columna de números de identificador de proveedor nacional.
dataDomain_MetaDataRule_Passport	Busca nombres que puedan describir una columna de información de pasaporte.
dataDomain_MetaDataRule_PhoneNumber	Busca nombres que puedan describir una columna de números de teléfono o números de fax.
dataDomain_MetaDataRule_Quantity	Busca nombres que puedan describir una columna de información de cantidad.
dataDomain_MetaDataRule_Race	Busca nombres que puedan describir una columna de información de raza o color.
dataDomain_MetaDataRule_Religion	Busca nombres que puedan describir una columna de información de religión, fe o creencias.
dataDomain_MetaDataRule_Salary	Busca nombres que puedan describir una columna de información de sueldo, salario o compensación.
dataDomain_MetaDataRule_SSN	Busca nombres que puedan describir una columna de números de seguridad social.
dataDomain_MetaDataRule_State	Busca nombres que puedan describir una columna de información de estado de Estados Unidos.
dataDomain_MetaDataRule_Street	Busca nombres que puedan describir una columna de información de direcciones físicas.
dataDomain_MetaDataRule_UniqueIdentifying Number	Busca nombres que puedan describir una columna de números de identificación únicos.
dataDomain_MetaDataRule_UPC_EAN	Busca nombres que puedan describir una columna de códigos de producto universales o números de artículo europeos.
dataDomain_MetaDataRule_URL	Busca nombres que puedan describir una columna de información de localizador de recursos uniforme o direcciones web.
dataDomain_MetaDataRule_VehicleRegPlate Number	Busca nombres que puedan describir una columna de registro del vehículo o matrícula del vehículo.
dataDomain_MetaDataRule_Weight	Busca nombres que puedan describir una columna de información de peso.
dataDomain_MetaDataRule_ZipCode	Busca nombres que puedan describir una columna de códigos postales.

Reglas de datos en el acelerador de dominios de datos

Utilice las reglas de datos del dominio de datos para identificar las columnas que contienen datos que coinciden con los criterios de la regla.

Busque las reglas de datos en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Domain_Discovery\Data_Rules
```

La siguiente tabla describe las reglas de datos en el acelerador de dominios de datos:

Nombre	Descripción
dataDomain_DataRule_ABARoutingNumber	Identifica los datos de columna que coinciden con el formato de un número de enrutamiento de la Asociación Bancaria de Estados Unidos. El número de enrutamiento identifica a una institución financiera en una transacción financiera.
dataDomain_DataRule_Account_Status	Identifica los datos de columna que coinciden con los valores de estado de la cuenta en los datos de referencia.
dataDomain_DataRule_Address_Data	Identifica los datos de columna que representan la información de dirección. La regla reconoce los datos de direcciones de varios países a nivel global.
dataDomain_DataRule_Age	Identifica los datos de columna con valores de 1 a 120.
dataDomain_DataRule_Alphanumeric_Special Character	Identifica los datos de columna que contienen datos alfanuméricos sin formato y datos con caracteres especiales.
dataDomain_DataRule_Amount	Identifica los datos de columna que representan una cantidad física.
dataDomain_DataRule_AUT_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Austria.
dataDomain_DataRule_BankAccount_USA	Identifica los datos de columna que coinciden con un formato de número de cuenta bancaria en Estados Unidos.
dataDomain_DataRule_BGR_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Bulgaria.
dataDomain_DataRule_BIC_SWIFTCode	Identifica los datos de columna que coinciden con los códigos BIC (Bank Identifier Code, código de identificación del banco) o SWIFT (Sociedad de Telecomunicaciones Financieras Interbancarias Mundiales) mediante reconocimiento de patrones y códigos de país.
dataDomain_DataRule_BinaryValues	Identifica los datos de columna que contienen valores binarios.
dataDomain_DataRule_BirthDay	Identifica los datos de columna que coinciden con fechas de nacimiento válidas. La regla verifica el número de años entre la fecha de entrada y la fecha actual. La regla devuelve "Adulto" o "Menor de edad" o "Válido" según los valores de 1 a 120. La regla devuelve "No válido" para todos los demás valores.
dataDomain_DataRule_BRA_IDDoc	Identifica los datos de columna que coinciden con el formato de número de la tarjeta de identificación de <i>Registro Geral</i> en Brasil.

Nombre	Descripción
dataDomain_DataRule_BRA_Personal_ID	Identifica los datos de columna que coinciden con el formato de ID personal de Brasil.
dataDomain_DataRule_CAN_SIN	Identifica los datos de columna que coinciden con el formato de número de seguridad social en Canadá.
dataDomain_DataRule_CHN_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de China.
dataDomain_DataRule_City	Identifica los datos de columna que contienen un nombre de ciudad válido. La regla lee los datos de referencia que contienen nombres de ciudades internacionales.
dataDomain_DataRule_CompanyName	Identifica los datos de columna que coinciden con los valores de nombre de organización en los datos de referencia.
dataDomain_DataRule_Computer_Address	Identifica los datos de columna que coinciden con el formato de direcciones IP y direcciones MAC.
dataDomain_DataRule_Country	Identifica los datos de columna que coinciden con un nombre de país ISO.
dataDomain_DataRule_CountryCode_Phone	Identifica los datos de columna que coinciden con números de teléfono basados en códigos de marcación internacionales.
dataDomain_DataRule_County	Identifica los datos de columna que coinciden con un nombre de condado de Estados Unidos.
dataDomain_DataRule_CreditCard_AMEX	Identifica los datos de columna que coinciden con el formato de número de tarjeta de crédito American Express.
dataDomain_DataRule_CreditCard_DinersCard	Identifica los datos de columna que coinciden con el formato de número de tarjeta de crédito Diners Club International.
dataDomain_DataRule_CreditCard_DiscoverCard	Identifica los datos de columna que coinciden con el formato de número de tarjeta de crédito Discover.
dataDomain_DataRule_CreditCard_JCB	Identifica los datos de columna que coinciden con el formato de número de tarjeta de crédito JCB International.
dataDomain_DataRule_CreditCard_MasterCard	Identifica los datos de columna que coinciden con el formato de número de tarjeta de crédito MasterCard.
dataDomain_DataRule_CreditCard_Visa	Identifica los datos de columna que coinciden con el formato de número de tarjeta de crédito Visa.
dataDomain_DataRule_CreditCardNumber	Identifica los datos de columna que coinciden con el formato de número de tarjeta de crédito de las principales organizaciones de tarjetas de crédito, como American Express, Diners Club International y Maestro.
dataDomain_DataRule_CreditCardTrack1FormatB	Identifica datos de columna que coinciden con la información de tarjeta de crédito de pista 1 y formato B.
dataDomain_DataRule_Currency	Identifica los datos de columna que coinciden con un término de moneda en los datos de referencia.

Nombre	Descripción
dataDomain_DataRule_Date_Validation	Identifica las cadenas de fecha en los datos de origen que aparecen en un formato único en una columna de fecha. Para configurar el formato de fecha que la regla utiliza para realizar la validación, abra la transformación de expresión dq_ValidateDate de la regla y actualice la variable de expresión In_Date_Format. El formato es "MM/DD/AAAA". La regla devuelve "Válido" o "No válido".
dataDomain_DataRule_Date_Validation_All_Formats	Identifica los valores de fecha en los datos de columna y estandariza los datos de columna en un único formato de fecha.
dataDomain_DataRule_DEU_Machine_Readable_Passport	Identifica los datos de columna que coinciden con el formato de número de pasaporte de lectura automática de Alemania.
dataDomain_DataRule_DNK_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Dinamarca.
dataDomain_DataRule_DriversLicense	Identifica los datos de columna que coinciden con los números de permiso de conducción de Canadá, Reino Unido y Estados Unidos basándose en la longitud y el patrón de los valores de datos.
dataDomain_DataRule_DriversLicense_Canada	Identifica los datos de columna que coinciden con los números de permiso de conducir de Canadá, excepto los números de las provincias de Columbia Británica, Quebec, Manitoba e Isla del Príncipe Eduardo.
dataDomain_DataRule_DriversLicense_Canada_narrow	Identifica los datos de columna que coinciden con los números de permiso de conducir de Canadá, excepto los números de las provincias de Columbia Británica, Quebec, Manitoba e Isla del Príncipe Eduardo. La regla es similar a la regla dataDomain_DataRule_DriversLicense_Canada. Sin embargo, dataDomain_DataRule_DriversLicense_Canada_narrow realiza un análisis más restrictivo para reducir la probabilidad de falsos positivos.
dataDomain_DataRule_DriversLicense_GBR	Identifica los datos de columna que coinciden con los números del permiso de conducción del Reino Unido.
dataDomain_DataRule_DriversLicense_narrow	Identifica los datos de columna que coinciden con los números del permiso de conducción del Reino Unido y de muchos estados y provincias de Canadá y Estados Unidos. La regla no valida números de las provincias de Columbia Británica, Quebec, Manitoba e Isla del Príncipe Eduardo. Para reducir la probabilidad de falsos positivos, la regla no valida los números que contienen entre cuatro y ocho dígitos.
dataDomain_DataRule_DriversLicense_USA	Identifica los datos de columna que coinciden con los números del permiso de conducción de la mayoría de estados de Estados Unidos.
dataDomain_DataRule_DriversLicense_USA_narrow	Identifica los datos de columna que coinciden con los números del permiso de conducción de la mayoría de estados de Estados Unidos. Para reducir la probabilidad de falsos positivos, la regla excluye los valores de datos que contienen entre seis y ocho dígitos. Por ejemplo, la regla excluye un valor como 01012017.

Nombre	Descripción
dataDomain_DataRule_Email	Identifica los datos de columna que coinciden con un formato de ID de correo electrónico predefinido.
dataDomain_DataRule_ExpirationDate	Identifica los datos de columna que coinciden con las fechas de tarjetas de crédito caducadas. La regla compara la fecha de entrada con la fecha del sistema para la validación.
dataDomain_DataRule_FIN_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Finlandia.
dataDomain_DataRule_FirstName	Identifica los datos de columna que coinciden con los valores de un conjunto de datos de referencia de nombres.
dataDomain_DataRule_FRA_INSEE	Identifica los datos de columna que coinciden con el formato de número del Institut National de la Statistique et des Études Économiques (INSEE) de Francia.
dataDomain_DataRule_FullName	Identifica las cadenas en una columna de datos que contienen nombres, segundos nombres y apellidos. La regla compara las palabras de cada cadena con los datos de referencia.
dataDomain_DataRule_GBR_NINO	Identifica los datos de columna que coinciden con el formato de número de seguridad social del Reino Unido.
dataDomain_DataRule_GBR_Passport_Number	Identifica los datos de columna que coinciden con el formato de número de pasaporte del Reino Unido.
dataDomain_DataRule_Gender	Identifica los datos de columna que coinciden con los valores de sexo en los datos de referencia.
dataDomain_DataRule_Height	Identifica los datos de columna con valores del 1 al 8, donde 8 representa los pies en altura.
dataDomain_DataRule_HostName	Identifica los datos de columna que coinciden con nombres de host válidos.
dataDomain_DataRule_HRV_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Croacia.
dataDomain_DataRule_IBAN	Identifica los datos de columna que coinciden con el formato del Código Internacional de Cuenta Bancaria de varios países europeos.
dataDomain_DataRule_ICD_10	Identifica los datos de las columnas que coinciden con los nombres de las condiciones en la décima revisión de la Clasificación Internacional y Estadística de Enfermedades y Problemas Relacionados con la Salud (ICD). La Organización Mundial de la Salud (OMS) mantiene la clasificación.
dataDomain_DataRule_ICD_9	Identifica los datos de las columnas que coinciden con los nombres de las condiciones en la novena revisión de la Clasificación Internacional y Estadística de Enfermedades y Problemas Relacionados con la Salud (ICD). La Organización Mundial de la Salud (OMS) mantiene la clasificación.
dataDomain_DataRule_IND_NATID	Identifica los datos de columna que coinciden con el formato de número de cuenta permanente de la India.

Nombre	Descripción
dataDomain_DataRule_IND_Passport	Identifica los datos de columna que coinciden con el formato de número de pasaporte de la India.
dataDomain_DataRule_IPAddress	Identifica los datos de columna que coinciden con un formato de dirección IP predefinido.
dataDomain_DataRule_ISBN	Identifica los datos de columna que coinciden con el formato del Número estándar internacional de libros.
dataDomain_DataRule_ISIN	Identifica los datos de columna que coinciden con el formato de número de identificación internacional de valores (ISIN). Un ISIN identifica de forma exclusiva un valor como una acción o un bono.
dataDomain_DataRule_ItalyFiscalCode	Identifica los datos de columna que coinciden con el formato de ID nacional de Italia.
dataDomain_DataRule_ITIN_USA	Identifica los datos de columna que coinciden con el formato de un número de identificación de contribuyente individual (ITIN) en Estados Unidos. El Servicio de Impuestos Internos emite los números de identificación.
dataDomain_DataRule_JobPosition	Identifica los datos de columna que coinciden con los nombres de puesto laboral en los datos de referencia.
dataDomain_DataRule_KOR_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Corea.
dataDomain_DataRule_LastName	Identifica los datos de columna que coinciden con los valores de un conjunto de datos de referencia de apellidos.
dataDomain_DataRule_Latitude	Identifica los datos de columna que coinciden con coordenadas de latitud válidas.
dataDomain_DataRule_LatitudeLongitude	Identifica los datos de columna que coinciden con pares válidos de coordenadas de latitud y longitud, donde cada par está separado por un punto y coma.
dataDomain_DataRule_Longitude	Identifica los datos de columna que coinciden con coordenadas de longitud válidas.
dataDomain_DataRule_Machine_Readable_Passport	Identifica los datos de columna que coinciden con los números de pasaporte de lectura automática de todos los países.
dataDomain_DataRule_NDC_USA	Identifica los datos de columna que coinciden con un valor en el directorio del código nacional de medicamentos (NDC) de Estados Unidos. Cada código identifica de forma exclusiva un fármaco que un fabricante desarrolló para uso humano.
dataDomain_DataRule_NOR_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Noruega.
dataDomain_DataRule_NPI_USA	Identifica los datos de columna que coinciden con un número de identificador de proveedor nacional (NPI) en Estados Unidos. Los centros de servicios de Medicare y Medicaid emiten los números a los proveedores de atención médica.

Nombre	Descripción
dataDomain_DataRule_PhoneNumber	Identifica los datos de columna que coinciden con un formato de número de teléfono de Estados Unidos.
dataDomain_DataRule_PostCode	Identifica los datos de columna que coinciden con los códigos postales de varios países.
dataDomain_DataRule_Quantity	Identifica los datos de columna que describen una cantidad física e incluyen unidades de medida.
dataDomain_DataRule_Race	Identifica los datos de columna que coinciden con el nombre de una raza de personas en los datos de referencia.
dataDomain_DataRule_Religion	Identifica los datos de columna que coinciden con el nombre de una religión en los datos de referencia.
dataDomain_DataRule_ROU_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Rumanía.
dataDomain_DataRule_SouthAfrica_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Sudáfrica.
dataDomain_DataRule_Spanish_NIF	Identifica los datos de columna que coinciden con el formato del número de identificación fiscal (NIF) en España.
dataDomain_DataRule_SSN	Identifica los datos de columna que coinciden con el formato de número de seguridad social en Estados Unidos.
dataDomain_DataRule_State	Identifica los datos de columna que coinciden con los nombres de estado en Estados Unidos.
dataDomain_DataRule_Street	Identifica las cadenas en los datos de columna que describen la información de dirección física; por ejemplo, calle, vía, avenida. La regla utiliza una expresión regular para encontrar descriptores de calle en los datos de columna.
dataDomain_DataRule_SWE_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Suecia.
dataDomain_DataRule_TWN_NATID	Identifica los datos de columna que coinciden con el formato de ID nacional de Taiwán.
dataDomain_DataRule_UPC	Identifica los datos de columna que coinciden con un código de producto universal válido. Un código de producto universal es un tipo de código de barras.
dataDomain_DataRule_UPC_EAN	Identifica los datos de columna que coinciden con un código de producto universal válido o con un número de artículo europeo. Los códigos de producto universales y los números de artículo europeos son tipos de códigos de barras.
dataDomain_DataRule_URL	Identifica los datos de columna que coinciden con formatos de URL predefinidos.
dataDomain_DataRule_US_Zip5	Identifica los datos de columna que coinciden con códigos postales de Estados Unidos.

Nombre	Descripción
dataDomain_DataRule_USA_Machine_Readable_Passport	Identifica los datos de columna que coinciden con un formato de número de pasaporte de lectura automática de Estados Unidos.
dataDomain_DataRule_USA_SSN_post_2011June	Identifica los datos de columna que coinciden con el formato de número de seguridad social en longitud, valores numéricos y valores mínimos y máximos de las secciones de área, grupo y número de serie. Según la iniciativa de aleatorización de SSN que entró en vigor el 25 de junio de 2011, la regla no verifica la emisión de un número de seguridad social ni la combinación de número de grupo y área.
dataDomain_DataRule_Weight	Identifica los datos de columna que describen un valor de peso. La regla busca un número entre 0 y 500.
dataDomain_DataRule_ZipCode	Identifica los datos de columna que coinciden con códigos postales de Estados Unidos.

CAPÍTULO 4

Acelerador de Australia/Nueva Zelanda

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador de Australia/Nueva Zelanda, 52](#)
- [Reglas de limpieza de datos de direcciones de Australia/Nueva Zelanda, 53](#)
- [Reglas compuestas de Australia/Nueva Zelanda, 54](#)
- [Reglas de limpieza de datos de contacto de Australia/Nueva Zelanda, 56](#)
- [Reglas de limpieza de datos corporativos de Australia/Nueva Zelanda, 59](#)
- [Reglas de limpieza de datos generales de Australia/Nueva Zelanda, 60](#)
- [Reglas de eliminación de duplicados y coincidencia de Australia/Nueva Zelanda, 61](#)
- [Asignaciones de demostración de Australia/Nueva Zelanda, 64](#)

Resumen del acelerador de Australia/Nueva Zelanda

Utilice las reglas del acelerador de Australia/Nueva Zelanda para comprobar y mejorar datos de organizaciones de Australia y Nueva Zelanda.

El acelerador de Australia/Nueva Zelanda incluye reglas que realizan las siguientes operaciones de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos corporativos
- Limpieza de datos generales
- Coincidencia y eliminación de duplicados

El acelerador de Australia/Nueva Zelanda también incluye reglas compuestas. Una regla compuesta combina varias reglas en un único objeto.

El acelerador depende de las reglas que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de direcciones de Australia/Nueva Zelanda

Utilice las reglas de limpieza de datos de direcciones para analizar, estandarizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La siguiente tabla describe las reglas de limpieza de datos de direcciones del acelerador de Australia/Nueva Zelanda:

Nombre	Descripción
rule_AUS_Address_Parse_Hybrid	Divide direcciones de Australia sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_AUS_Address_Parse_Multiline	Divide direcciones de Australia sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_AUS_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones australianas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_AUS_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones australianas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_AUS_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones australianas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_AUS_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones australianas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_AUS_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones australianas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_AUS_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones australianas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Nombre	Descripción
rule_NZL_Address_Parse_Hybrid	Divide direcciones de Nueva Zelanda sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_NZL_Address_Parse_Multiline	Divide direcciones de Nueva Zelanda sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_NZL_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones neozelandesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_NZL_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones neozelandesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_NZL_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones neozelandesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_NZL_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones neozelandesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_NZL_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones neozelandesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_NZL_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones neozelandesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Reglas compuestas de Australia/Nueva Zelanda

Utilice las reglas compuestas en el acelerador de Australia/Nueva Zelanda para agregar un conjunto de reglas a una asignación como un único objeto. Una regla compuesta es una regla que utiliza la lógica de otras reglas del acelerador.

Puede encontrar las reglas compuestas en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

En la siguiente tabla se describen las reglas compuestas del acelerador de Australia/Nueva Zelanda:

Nombre	Descripción
rule_AUS_Contact_Data	Analiza, estandariza y valida datos de contacto de Australia, tales como direcciones, números de teléfono y números de contribuyente de Australia.
rule_NZL_Contact_Data	Analiza, estandariza y valida datos de contacto de Nueva Zelanda, tales como direcciones, números de teléfono y números del IRD.

Regla compuesta de datos de contacto de Australia

La regla `rule_AUS_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en `rule_AUS_Contact_Data`:

Regla	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_AUS_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_AUS_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_AUS_Gender_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_AUS_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_AUS_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_AUS_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_AUS_Tax_File_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_AUS_Tax_File_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Prename_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Regla compuesta de datos de contacto de Nueva Zelanda

La regla `rule_NZL_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en rule_NZL_Contact_Data:

Regla	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_AUS_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_AUS_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_NZL_Gender_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_IRD_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_IRD_Number_Validate	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_Phone_Number_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing

Reglas de limpieza de datos de contacto de Australia/Nueva Zelanda

Utilice las reglas de limpieza de datos de contacto para analizar, estandarizar y validar datos sobre los contactos empresariales y los individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
```


La siguiente tabla describe las reglas de limpieza de datos de contacto del acelerador de Australia/Nueva Zelanda:

Nombre	Descripción
rule_AUS_Driver_Licence_Number_Validation	Valida los números de permisos de conducción australianos a partir de los requisitos de longitud y patrón.
rule_AUS_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, esta regla asigna el sexo "M" al nombre "John Smith", ya que es masculino.
rule_AUS_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres australianos.
rule_AUS_Multi_Person_Name_Parse	<p>Analiza valores de nombre de persona en puertos independientes. La regla crea puertos para valores como el tratamiento, el nombre, el segundo nombre y los apellidos.</p> <p>La salida de la regla incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Cuando los datos de nombre identifican más de una persona, la regla crea un puerto de salida para cada nombre completo. Por ejemplo, la regla puede leer el nombre "John y Jane Smith" y crear puertos de salida para "John Smith" y "Jane Smith".</p>
rule_AUS_Personal_Name_Parse_FML	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Nombre, segundo nombre, apellidos <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_AUS_Personal_Name_Parse_LFM	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_AUS_Phone_Number_Parse	<p>Reconoce un número de teléfono australiano en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda.</p> <p>La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla. La regla procesa los siguientes signos de puntuación: el signo más, los paréntesis y una almohadilla. Antes de ejecutar la regla, quite todos los demás signos de puntuación, incluidos los dobles espacios.</p> <p>La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.</p>

Nombre	Descripción
rule_AUS_Phone_Number_Standardization	Estandariza los números de teléfono australianos según formatos de marcación internacional y local. La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla.
rule_AUS_Phone_Number_Validation	Valida el código de área y la longitud de los números de teléfono australianos. La regla devuelve la región del número de teléfono, así como los códigos que determinan la validez del código de área y la longitud de un número de teléfono.
rule_AUS_Tax_File_Number_Parse	Analiza el TFN (Tax File Number, número de contribuyente de Australia).
rule_AUS_Tax_File_Number_Standardization	Estandariza el TFN. Para configurar el formato estandarizado, edite la variable de expresión TFN_Format en la transformación de expresión dq_Format_TFN. El valor predeterminado es "No_punctuation".
rule_AUS_Tax_File_Number_Validation	Valida el TFN a partir del dígito de comprobación de cada número.
rule_NZL_Gender_Assignment	Asigna un sexo a partir de nombres de pila neozelandeses. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, esta regla asigna el sexo "M" al nombre "John Smith", ya que es masculino.
rule_NZL_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres neozelandeses.
rule_NZL_IRD_Number_Parse	Reconoce cadenas numéricas de nueve dígitos como números del IRD (Inland Revenue Department, departamento de Hacienda Pública de Nueva Zelanda).
rule_NZL_IRD_Number_Standardization	Estandariza números del IRD. Para configurar el formato estandarizado, edite la variable de expresión IRD_Format en la transformación de expresión dq_Format_IRD. El valor predeterminado es "No_punctuation". La regla requiere que la entrada sea una cadena de nueve dígitos.
rule_NZL_IRD_Number_Validate	Valida números del IRD a partir del dígito de comprobación de cada número.
rule_NZL_Phone_Number_Parse	Reconoce un número de teléfono neozelandés en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda. La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla. La regla procesa los siguientes signos de puntuación: el signo más, los paréntesis y una almohadilla. Antes de ejecutar la regla, quite todos los demás signos de puntuación, incluidos los dobles espacios. La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.
rule_NZL_Phone_Number_Standardization	Estandariza los números de teléfono neozelandeses según formatos de marcación internacional y local. La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla.
rule_NZL_Phone_Number_Validation	Valida el código de área y la longitud de los números de teléfono neozelandeses. Esta regla devuelve la región del número de teléfono, así como los códigos que determinan la validez del código de área y la longitud de un número de teléfono.

Nombre	Descripción
rule_Prenome_Assignment	Genera un tratamiento según el sexo. Puede cambiar la variable de expresión female_prenome de "Ms." a "Mrs.".
rule_Salutation_Assignment	Genera saludos formales e informales a partir de nombres de pila y tokens de nombre. Por ejemplo, si los datos de entrada contienen "Mr. John Smith," la regla genera el saludo formal "Dear Mr. Smith," y el saludo informal "Dear John,". Puede cambiar el prefijo y la puntuación si edita las variables en la transformación de expresión dq_Generate_Salutation.

Dependencias de reglas de limpieza de datos de contacto del núcleo

El acelerador de Australia/Nueva Zelanda depende de las siguientes reglas de limpieza de datos de contacto del acelerador del núcleo:

- rule_Email_Validation

Para obtener más información sobre estas reglas, consulte ["Reglas de limpieza de datos de contacto del núcleo" en la página 22.](#)

Reglas de limpieza de datos corporativos de Australia/Nueva Zelanda

Utilice las reglas de limpieza de datos corporativos para analizar, estandarizar y validar datos corporativos.

Puede encontrar las reglas de limpieza de datos corporativos en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

La siguiente tabla describe las reglas de limpieza de datos corporativos del acelerador de Australia/Nueva Zelanda:

Nombre	Descripción
rule_AUS_Business_Number_Parse	Reconoce cadenas numéricas de once dígitos como un ABN (Australian Business Number, número de identificación de empresa de Australia).
rule_AUS_Business_Number_Standardize	Estandariza el ABN según el formato NN NNN NNN NNN. Esta regla requiere que la entrada sea una cadena de once dígitos.
rule_AUS_Business_Number_Validation	Valida el ABN a partir del dígito de comprobación de cada número.
rule_AUS_Company_Name_Standardization	Estandariza nombres de empresas según valores de tabla de referencia de Australia.

Reglas de limpieza de datos generales de Australia/Nueva Zelanda

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

La siguiente tabla describe las reglas de limpieza de datos generales del acelerador de Australia/Nueva Zelanda:

Nombre	Descripción
rule_AUS_NZL_NER_Field_Identification	Identifica el tipo de información contenida en un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de Australia y Nueva Zelanda. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea técnicas de coincidencia probabilística para identificar los tipos de información.

Dependencias de reglas de limpieza de datos generales del núcleo

El acelerador de Australia/Nueva Zelanda depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Assign_DQ_GeocodingStatus_Description
- rule_Assign_DQ_Mailability_Score_Description
- rule_Assign_DQ_Match_Code_Description
- rule_Remove_Extra_Spaces
- rule_Remove_Hyphen
- rule_Remove_Leading_Zero
- rule_Remove_Period_Parentheses
- rule_Remove_Punctuation
- rule_Remove_Punctuation_and_Space
- rule_Remove_Space
- rule_Replace_Limited_Punct_with_Space
- rule_UpperCase

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos generales del núcleo” en la página 23.](#)

Reglas de eliminación de duplicados y coincidencia de Australia/Nueva Zelanda

Utilice las reglas de coincidencia y eliminación de duplicados en el acelerador de Australia/Nueva Zelanda para medir los niveles de similitud entre los registros de conjuntos de datos.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador de Australia/Nueva Zelanda:

Nombre	Descripción
mplt_AUS_Firstname_and_TFN_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Australia a partir de los nombres de pila y el TFN (Tax File Number, número de contribuyente de Australia). El mapplet genera claves de grupo a partir de los datos del TFN.
mplt_AUS_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Australia a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_AUS_IMO_Familyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Australia a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_AUS_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Australia a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_AUS_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Australia a partir de nombres de personas y datos personales. Los campos de la columna de datos personales deben contener un único tipo de datos, tales como número de teléfono, dirección de correo electrónico o número de contribuyente de Australia. El mapplet genera claves de grupo a partir de los datos personales.
mplt_AUS_Individual_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y datos de direcciones de Australia. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_AUS_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y fechas de Australia. El mapplet genera claves de grupo a partir de los datos de fechas.
mplt_AUS_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de direcciones de correo electrónico y nombres de personas de Australia. El mapplet genera claves de grupo a partir de los datos de direcciones de correo electrónico.
mplt_AUS_Individual_Name_and_Phone_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y números de teléfono de Australia. El mapplet genera claves de grupo a partir de los datos de números de teléfono.

Nombre	Descripción
mplt_AUS_Individual_Name_and_TFN_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Australia a partir de nombres de personas y números de contribuyente de Australia. El mapplet genera claves de grupo a partir de los datos del TFN.
mplt_AUS_Individual_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas de Australia. El mapplet genera códigos NYSIIS a partir de los valores de apellido y los emplea como claves de grupo.
mplt_AUS_NZL_Company_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas y datos de direcciones de Australia y Nueva Zelanda. El mapplet utiliza una combinación de caracteres de los valores de nombre de empresa y código postal para generar claves de grupo.
mplt_AUS_NZL_Familyname_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de apellidos y datos de direcciones de Australia y Nueva Zelanda. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_Company_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas. El mapplet genera códigos Soundex a partir de los valores de nombre de empresa y los emplea como claves de grupo.
mplt_NZL_Firstname_and_IRD_Match	Utiliza estrategias de la coincidencia de campos para identificar las filas duplicadas en los datos de Nueva Zelanda a partir de nombres de pila y números del IRD (Inland Revenue Department, departamento de Hacienda Pública de Nueva Zelanda). El mapplet genera claves de grupo a partir del número del IRD.
mplt_NZL_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Nueva Zelanda a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_NZL_IMO_Familyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Nueva Zelanda a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_NZL_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Nueva Zelanda a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_NZL_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Nueva Zelanda a partir de nombres de personas y datos personales. Los campos de la columna de datos personales deben contener un único tipo de datos, como número de teléfono, dirección de correo electrónico o número del IRD. El mapplet genera claves de grupo a partir de los datos personales.
mplt_NZL_Individual_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y datos de direcciones de Nueva Zelanda. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_NZL_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y fechas de Nueva Zelanda. El mapplet genera claves de grupo a partir de los datos de fechas.

Nombre	Descripción
mplt_NZL_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de direcciones de correo electrónico y nombres de personas de Nueva Zelanda. El mapplet genera claves de grupo a partir de los datos de direcciones de correo electrónico.
mplt_NZL_Individual_Name_and_IRD_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas de Nueva Zelanda y números del IRD. El mapplet genera claves de grupo a partir del número del IRD.
mplt_NZL_Individual_Name_and_Phone_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y números de teléfono de Nueva Zelanda. El mapplet genera claves de grupo a partir de los datos de números de teléfono.
mplt_NZL_Individual_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas de Nueva Zelanda. El mapplet genera códigos NYSIIS a partir de los valores de apellido y los emplea como claves de grupo.
rule_AUS_NZL_Company_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y direcciones de Australia y Nueva Zelanda.
rule_AUS_NZL_Familyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de apellidos y direcciones de Australia y Nueva Zelanda.
rule_AUS_NZL_Firstname_and_PID_MatchScore	Genera una puntuación de coincidencia a partir de nombres de pila y números de identificación personal.
rule_AUS_NZL_Individual_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de Australia y Nueva Zelanda.
rule_AUS_NZL_Individual_Name_and_PID_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números de identificación personal.
rule_Company_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas.
rule_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.
rule_Individual_Name_and_Phone_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números de teléfono.
rule_Individual_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas.

Asignaciones de demostración de Australia/Nueva Zelanda

Las asignaciones de demostración en el acelerador de Australia/Nueva Zelanda utilizan varias reglas para comprobar los procesos de calidad de datos.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\AUS_NZL_Accelerator
```

El acelerador contiene las siguientes asignaciones de demostración:

m_AUS_customer_data_demo

Analiza, estandariza y valida datos de Australia y Nueva Zelanda.

m_AUS_customer_matching_demo

Analiza y estandariza datos de identidad de Australia y Nueva Zelanda, y realiza un análisis de coincidencia de identidades de los datos.

La asignación analiza las siguientes combinaciones de datos y genera clústeres de coincidencia para cada combinación:

- Nombre de persona y datos de direcciones
- Nombre de persona y número de teléfono

CAPÍTULO 5

Acelerador de Brasil

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador de Brasil, 65](#)
- [Reglas de limpieza de datos de direcciones de Brasil, 66](#)
- [Reglas compuestas de Brasil, 67](#)
- [Reglas de limpieza de datos de contacto de Brasil, 68](#)
- [Reglas de limpieza de datos corporativos de Brasil, 69](#)
- [Reglas de limpieza de datos generales de Brasil, 69](#)
- [Reglas de coincidencia y eliminación de duplicados de Brasil, 70](#)
- [Asignaciones de demostración de Brasil, 72](#)

Resumen del acelerador de Brasil

Utilice las reglas del acelerador de Brasil para comprobar y mejorar datos de organizaciones en Brasil.

El acelerador de Brasil incluye reglas que realizan las siguientes operaciones de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos corporativos
- Limpieza de datos generales
- Coincidencia y eliminación de duplicados

El acelerador de Brasil también incluye una regla compuesta. Una regla compuesta combina varias reglas en un único objeto.

El acelerador depende de las reglas que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de direcciones de Brasil

Utilice las reglas de limpieza de datos de direcciones para analizar, estandarizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de direcciones del acelerador de Brasil:

Nombre	Descripción
rule_BRA_Address_Parse_Hybrid	Divide direcciones de Brasil sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_BRA_Address_Parse_Multiline	Divide direcciones de Brasil sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_BRA_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones brasileñas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_BRA_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones brasileñas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_BRA_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones brasileñas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_BRA_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones brasileñas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_BRA_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones brasileñas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_BRA_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones brasileñas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Reglas compuestas de Brasil

Utilice las reglas compuestas en el acelerador de Brasil para agregar un conjunto de reglas a una asignación como un único objeto. Una regla compuesta es una regla que utiliza la lógica de otras reglas del acelerador.

Puede encontrar las reglas compuestas en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

La tabla siguiente describe la regla compuesta en el acelerador de Brasil:

Nombre	Descripción
rule_BRA_Contact_Data	Analiza, estandariza y valida datos de contacto de Brasil, tales como direcciones, números de teléfono y números del CPF (Cadastro de Pessoas Físicas, registro de personas físicas de Brasil).

La regla `rule_BRA_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en `rule_BRA_Contact_Data`:

Regla	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_BRA_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_BRA_Company_Suffix_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_BRA_Personal_CPF_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Personal_Name_Parse_Validate	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Phone_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Prenome_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Salutation_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing

Reglas de limpieza de datos de contacto de Brasil

Utilice las reglas de limpieza de datos de contacto para analizar, estandarizar y validar datos sobre los contactos empresariales y los individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de contacto del acelerador de Brasil:

Nombre	Descripción
rule_BRA_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, esta regla asigna el sexo "M" al nombre "Joao Coelho", ya que es masculino.
rule_BRA_Given_Name_Standard	Genera nombres a partir de sobrenombres brasileños.
rule_BRA_Personal_CPF_Validation	Valida los dígitos de comprobación de los números del CPF (Cadastro de Pessoas Físicas, registro de personas físicas de Brasil).
rule_BRA_Personal_Name_Parse_Validate	<p>Analiza valores de nombre de persona en puertos independientes. La regla crea puertos para valores como el tratamiento, el nombre, el segundo nombre y los apellidos. La regla también indica si el nombre puede ser un nombre de empresa y valida la ortografía del nombre.</p> <p>La salida de la regla incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_BRA_Personal_PIS_PASEP_Validation	Valida números de la seguridad social brasileña.
rule_BRA_Personal_Voter_Registration_Validation	Valida los dígitos de comprobación de los números de registro electoral de Brasil.
rule_BRA_Phone_Number_Parse	Reconoce un número de teléfono brasileño en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de izquierda a derecha. La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.
rule_BRA_Phone_Number_Standardization	<p>Estandariza números de teléfono brasileños. Esta regla devuelve el número de teléfono con los formatos siguientes:</p> <ul style="list-style-type: none">- Estándar: nn nnnn nnnn- Con guiones: nn-nnnn-nnnn- Sin espacios: nnnnnnnnnn
rule_BRA_Phone_Validation	Valida el código de área y la longitud de los números de teléfono de Brasil. La regla devuelve códigos que determinan la validez del código de área y la longitud de un número de teléfono.

Nombre	Descripción
rule_BRA_Prenome_Assignment	Genera un tratamiento según el sexo. Puede cambiar la variable de expresión female_prenome de "Sra" a "Sta".
rule_BRA_Salutation_Assignment	Genera saludos formales e informales a partir de nombres de pila y tokens de nombre. Por ejemplo, si los datos de entrada contienen "Sr. Joao Coelho,", la regla genera el saludo formal "Prezado Sr. Coelho," así como el saludo informal "Prezado Joao,". Puede cambiar el prefijo y la puntuación si edita las variables en la transformación de expresión dq_Generate_Salutation.

Dependencias de reglas de limpieza de datos de contacto del núcleo

El acelerador de Brasil depende de las siguientes reglas de limpieza de datos de contacto del acelerador del núcleo:

- rule_Email_Parse_Into_Mailbox_Domain
- rule_Email_Validation

Para obtener más información sobre estas reglas, consulte ["Reglas de limpieza de datos de contacto del núcleo" en la página 22.](#)

Reglas de limpieza de datos corporativos de Brasil

Utilice las reglas de limpieza de datos corporativos del acelerador de Brasil para estandarizar y validar datos corporativos.

Puede encontrar las reglas de limpieza de datos corporativos en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos corporativos del acelerador de Brasil:

Nombre	Descripción
rule_BRA_Company_CNPJ_Validation	Valida los números del CNPJ (Cadastro Nacional da Pessoa Jurídica, registro nacional de personas jurídicas de Brasil). Los números del CNPJ identifican a las empresas brasileñas.
rule_BRA_Company_Suffix_Standardization	Estandariza los sufijos de las empresas brasileñas.

Reglas de limpieza de datos generales de Brasil

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en los campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador de Brasil:

Nombre	Descripción
rule_BRA_NER_Field_Identification	Identifica el tipo de información contenida en un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de Brasil. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia y técnicas de coincidencia probabilística para identificar los tipos de información.

Dependencias de reglas de limpieza de datos generales del núcleo

El acelerador de Brasil depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Assign_DQ_GeocodingStatus_Description
- rule_Assign_DQ_Mailability_Score_Description
- rule_Assign_DQ_Match_Code_Descriptions
- rule_Remove_Extra_Spaces
- rule_Remove_Non_Numbers
- rule_Remove_Punctuation
- rule_Remove_Punctuation_and_Space
- rule_Replace_Limited_Punct_with_Space
- rule_TitleCase
- rule_UpperCase

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos generales del núcleo” en la página 23.](#)

Reglas de coincidencia y eliminación de duplicados de Brasil

Utilice las reglas de coincidencia y eliminación de duplicados para medir los niveles de similitud entre los registros de conjuntos de datos.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador de Brasil:

Nombre	Descripción
mplt_BRA_Company_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas y datos de direcciones de Brasil. El mapplet utiliza una combinación de caracteres de los valores de nombre de empresa y código postal para generar claves de grupo.
mplt_BRA_Famillyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Brasil a partir de apellidos y direcciones. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_BRA_Firstname_and_CPF_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir del nombre de pila y el número del CPF (Cadastro de Pessoas Físicas, registro de personas físicas de Brasil). El mapplet genera claves de grupo a partir del número del CPF.
mplt_BRA_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Brasil a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_BRA_IMO_Famillyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Brasil a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_BRA_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Brasil a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_BRA_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Brasil a partir de nombres de personas y datos personales. Los campos de la columna de datos personales deben contener un único tipo de datos, tales como número de teléfono, dirección de correo electrónico o número del CPF. El mapplet genera claves de grupo a partir de los datos personales.
mplt_BRA_Individual_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y datos de direcciones de Brasil. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_BRA_Individual_Name_and_CPF_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas de Brasil y números del CPF. El mapplet genera claves de grupo a partir del número del CPF.
mplt_BRA_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de datos de nombres de personas y fechas de Brasil. El mapplet genera claves de grupo a partir de los datos de fechas.
mplt_BRA_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y direcciones de correo electrónico de Brasil. El mapplet genera claves de grupo a partir de datos de direcciones de correo electrónico.
mplt_BRA_Individual_Name_and_Phone_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y números de teléfono de Brasil. El mapplet genera claves de grupo a partir de números de teléfono.

Nombre	Descripción
mplt_Company_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas. El mapplet genera códigos Soundex a partir de los valores de nombre de empresa y los emplea como claves de grupo.
rule_BRA_Company_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y datos de direcciones de Brasil.
rule_BRA_Famillyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de apellidos y datos de direcciones de Brasil.
rule_BRA_Firstname_and_CPF_MatchScore	Genera una puntuación de coincidencia a partir del nombre de pila y el número del CPF.
rule_BRA_Individual_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y datos de direcciones de Brasil.
rule_BRA_Individual_Name_and_CPF_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y datos de direcciones de Brasil.
rule_BRA_Individual_Name_and_Phone_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números de teléfono.
rule_Company_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas.
rule_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.

Asignaciones de demostración de Brasil

Las asignaciones de demostración del acelerador de Brasil utilizan varias reglas para comprobar los procesos de calidad de datos.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\BRA_Accelerator
```

El acelerador contiene las siguientes asignaciones de demostración:

m_BRA_customer_data_demo

Analiza, estandariza y valida datos de Brasil.

m_BRA_customer_matching_demo

Analiza y estandariza datos de identidad de Brasil, y realiza un análisis de coincidencia de identidades de los datos.

La asignación analiza las siguientes combinaciones de datos y genera clústeres de coincidencia para cada combinación:

- Nombre de persona y datos de direcciones
- Nombre de persona y número de teléfono

CAPÍTULO 6

Acelerador de Servicios Financieros

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador de Servicios Financieros, 74](#)
- [Reglas de limpieza de datos de contacto de Servicios Financieros, 74](#)
- [Reglas de limpieza de datos financieros de Servicios Financieros, 75](#)
- [Reglas de limpieza de datos generales de Servicios Financieros, 78](#)
- [Reglas de coincidencia y eliminación de duplicados de Servicios Financieros, 79](#)

Resumen del acelerador de Servicios Financieros

Utilice el acelerador de Servicios Financieros para comprobar y mejorar los datos de organizaciones del sector de servicios financieros.

El acelerador de Servicios Financieros incluye reglas que realizan los siguientes procesos de calidad de datos:

- Limpieza de datos de contacto
- Limpieza de datos financieros
- Limpieza de datos generales
- Coincidencia y eliminación de duplicados

El acelerador depende de las reglas de limpieza de datos que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de contacto de Servicios Financieros

Utilice las reglas de limpieza de datos de contacto para estandarizar datos sobre contactos empresariales e individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

La tabla siguiente describe la regla de limpieza de datos de contacto del acelerador de Servicios Financieros:

Nombre	Descripción
rule_USA_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres estadounidenses. Por ejemplo, la regla estandariza el sobrenombre "Bob" según el nombre de pila "Robert".

Reglas de limpieza de datos financieros de Servicios Financieros

Utilice las reglas de limpieza de datos financieros para analizar, estandarizar y validar datos financieros.

Puede encontrar las reglas de limpieza de datos financieros en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Financial_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos financieros del acelerador de Servicios Financieros:

Nombre	Descripción
rule_Account_Status_Validation	Valida el estado de la cuenta. La regla requiere datos de referencia sobre el estado de cuenta.
rule_Accrual_Period_Validation	Confirma que la fecha de inicio sea anterior a la fecha de finalización.
rule_Age_For_Account_Validation	Valida la edad del cliente del tipo de cuenta. La regla emplea la tabla de referencia age_per_account_inf. Debe actualizar la tabla de referencia con sus propios datos.
rule_Beta_Coefficient_Validation	Confirma que la cadena de coeficiente beta sea un número. La regla determina si la cadena es un número positivo, un número negativo o un número igual a cero, o bien si no es un número.
rule_BIC_SWIFT_Code_Validation	Valida los códigos BIC (Bank Identifier Code, código de identificación del banco) o SWIFT (Sociedad de Telecomunicaciones Financieras Interbancarias Mundiales) mediante reconocimiento de patrones y validación de código de país.
rule_CAN_Transit_Number_Validation	Utiliza transacciones de fondos en efectivo y electrónicas para validar el formato de un número de tránsito de Canadá.
rule_Credit_Card_Expiry_Check	Valida la fecha de vencimiento de una tarjeta de crédito. La regla compara la fecha de vencimiento de la tarjeta de crédito con la fecha del sistema e identifica las fechas de vencimiento. La regla acepta una cadena de siete caracteres con el formato MM/AAAA.
rule_Credit_Card_Security_Code_Validation	Confirma que el código de seguridad de la tarjeta de crédito sea un número entero de tres o cuatro dígitos.
rule_Currency_Code_Country_Validation	Confirma que el código de divisa sea válido para el código de país ISO compuesto por tres caracteres.

Nombre	Descripción
rule_Currency_Code_Validation	Valida el código de divisa. La regla devuelve "Válido" o "No válido".
rule_CUSIP_Validation	Valida el formato y la longitud del valor de comprobación de dígitos. La regla devuelve un estado que describe la validez del valor del dígito de comprobación y un mensaje que explica el estado.
rule_Delta_Validation	Confirma que el valor delta sea positivo, negativo o igual a cero.
rule_Dividend_Yield_Validation	Confirma que la cadena de rentabilidad de dividendos sea un número igual o superior a cero. La regla determina si la cadena es un número positivo, un número negativo o un número igual a cero, o bien si no es un número.
rule_EAD_Drawn_Balance_Validation	Confirma que la cantidad que figura en la exposición en el momento del impago (EAD) no sea inferior al saldo restante. La regla cumple las directrices para el cálculo de la EAD que ha establecido la Financial Services Authority (autoridad de servicios financieros del Reino Unido).
rule_EAD_Validation	Confirma que la cadena de EAD sea un número. La regla determina si la cadena es un número positivo, un número negativo o un número igual a cero, o bien si no es un número.
rule_EPS_Validation	Confirma que la entrada sea un número igual o superior a cero.
rule_Ex_Dividend_Date_Validation	Confirma que la fecha exdividendo y la fecha de registro sean fechas válidas, y que la fecha exdividendo sea anterior a la fecha de registro. La regla determina que las fechas con una diferencia superior a 15 días no son válidas. La regla devuelve los días diferencia que existen entre la fecha de registro y la fecha exdividendo.
rule_Gamma_Validation	Confirma que la cadena gamma sea un número. La regla determina si la cadena es un número positivo, un número negativo o un número igual a cero, o bien si no es un número.
rule_GBR_Bank_Account_Parse	Reconoce cadenas numéricas de ocho dígitos como números de cuentas bancarias del Reino Unido.
rule_GBR_Bank_Account_Validation	Valida números de cuentas bancarias del Reino Unido. La regla devuelve códigos que determinan si la entrada es numérica y si cuenta con la cantidad de dígitos correcta.
rule_GBR_Bank_Sort_Code_Parse	Analiza cadenas numéricas de seis dígitos como códigos de sucursal bancaria del Reino Unido. La regla analiza cadenas de números con los formatos siguientes: <ul style="list-style-type: none"> - Números consecutivos (999999) - Números delimitados por guión (99-99-99)
rule_GBR_Bank_Sort_Code_Standardize	Estandariza un código de sucursal bancaria del Reino Unido según el formato "NN-NN-NN".
rule_GBR_Bank_Sort_Code_Validation	Valida el formato y la longitud de códigos de sucursales bancaria del Reino Unido que se han estandarizado según el formato delimitado por guiones (99-99-99). La regla devuelve un puerto de estado que describe la validez del código de sucursal y un puerto de nota de validación que explica el estado. Si el prefijo del código de sucursal coincide con una asignación conocida de un banco del Reino Unido, el puerto de nota de validación incluirá el nombre del banco.

Nombre	Descripción
rule_Interest_Rate_Within_Range	Confirma que el valor del tipo de interés decimal se encuentra en el intervalo especificado. Los dos puertos de variable de la transformación de expresión definen este intervalo. La regla devuelve "Verdadero" o "Falso".
rule_ISIN_Code_Validation	Confirma que el valor de entrada sea un número de identificación internacional de valores (ISIN). La regla comprueba la estructura del valor y verifica el dígito de comprobación.
rule_Loan_to_Value_Ratio	Calcula la relación préstamo-valor, la cual se obtiene a partir de la división del importe del préstamo por el valor de la propiedad.
rule_Loss_Given_Default_Validation	Confirma que la cadena sea numérica y que tenga un valor positivo, negativo o igual a cero.
rule_Market_Cap_Validation	Confirma que la entrada sea un número igual o superior a cero.
rule_Maturity_Date_Validation	Confirma que la fecha de vencimiento sea posterior a la fecha del sistema.
rule_Positive_Close_Price_Value_Validation	Confirma que la entrada sea un número mayor que cero.
rule_Positive_Coupon_Percent_Validation	Confirma que la entrada sea un número mayor que cero.
rule_Positive_Last_Price_Value_Validation	Confirma que la entrada sea un número mayor que cero.
rule_Positive_Open_Price_Validation	Confirma que la entrada sea un número mayor que cero.
rule_Positive_Volume_Validation	Confirma que la entrada sea un número mayor que cero.
rule_Price_Earnings_Ratio_Validation	Confirma que la relación precio-ganancia sea un número positivo entre 0 y 100.
rule_Probability_of_Default_Validation	Confirma que el valor de la probabilidad de impago sea numérico e indica si tiene un valor positivo, negativo o igual a cero. Si es positivo, la regla devuelve mensajes de estado para los valores en los siguientes intervalos: - <= 0,1 - > 0,1 y <= 0,5 - > 0,5 y <= 1 - > 1
rule_Rating_Code_Validation	Confirma que la clasificación corresponda a la escala de Standard & Poor's, con la de Moody's o con una lista definida por el usuario.
rule_Rating_Date_Validation	Confirma que la fecha de la clasificación sea un año posterior a la que muestra la fecha del sistema.
rule_Risk_Weighted_Asset_Validation	Confirma que un activo ponderado en función del riesgo sea un número positivo.
rule_SEDOL_Validation	Valida un código SEDOL (Stock Exchange Daily Official List, lista oficial diaria de la bolsa de valores del Reino Unido) mediante la verificación del formato y el dígito de comprobación.

Nombre	Descripción
rule_Stock_Exchange_Validation	Valida la mayoría de las bolsas de valores del mundo mediante el nombre y el símbolo.
rule_USA_Routing_Number_Validation	Valida un número de ruta con formato de una línea de lectura magnética de caracteres (MICR) estándar. Valida el banco de la Reserva Federal asociado, la estructura de la entrada y el cálculo de la suma de comprobación.
rule_Volatility_Validation	Confirma que el valor de volatilidad sea un número igual o superior a cero.

Reglas de limpieza de datos generales de Servicios Financieros

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en los campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\General _Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador de Servicios Financieros:

Nombre	Descripción
rule_Positive_Number_Validation	Valida que el número sea mayor que cero.

Dependencias de reglas de limpieza de datos generales del núcleo

El acelerador de Servicios Financieros depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Remove_Punctuation
- rule_Remove_Punctuation_and_Space
- rule_Remove_Space
- rule_UpperCase

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos generales del núcleo” en la página 23](#).

Reglas de coincidencia y eliminación de duplicados de Servicios Financieros

Utilice las reglas de coincidencia y eliminación de duplicados para generar puntuaciones de coincidencia y detectar los registros duplicados.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador de Servicios Financieros:

Nombre	Descripción
mplt_Company_Name_and_Address_Match	Identifica filas duplicadas a partir de nombres de empresas y datos de direcciones de Estados Unidos. El mapplet utiliza una combinación de caracteres de los valores de nombre de empresa y código postal para generar claves de grupo.
mplt_Company_Name_Match	Detecta filas duplicadas basándose en el nombre de la empresa. El mapplet genera códigos Soundex a partir de los valores de nombre de empresa y los emplea como claves de grupo.
mplt_Familyname_and_Address_Match	Detecta filas duplicadas a partir de apellidos y datos de direcciones de Estados Unidos. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_Individual_Name_and_Address_Match	Detecta filas duplicadas a partir de nombres de personas y datos de direcciones de Estados Unidos. El mapplet genera códigos NYSIIS a partir de los valores de apellido y los emplea como claves de grupo.
mplt_Individual_Name_and_Date_Match	Detecta filas duplicadas basándose en nombres de personas y datos de fechas. El mapplet genera claves de grupo a partir de los datos de fechas.
mplt_Individual_Name_and_Email_Match	Detecta filas duplicadas basándose en direcciones de correo electrónico y nombres de personas. El mapplet genera claves a partir de los datos de direcciones de correo electrónico.
mplt_Individual_Name_and_Phone_Match	Detecta filas duplicadas a partir de números de teléfono y nombres de personas. El mapplet genera claves de grupo a partir de números de teléfono.
mplt_Individual_Name_Match	Detecta filas duplicadas basándose en nombres de personas. El mapplet genera códigos NYSIIS a partir de los valores de apellido y los emplea como claves de grupo.
rule_Company_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y direcciones de Estados Unidos.
rule_Company_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas.
rule_Familyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de apellidos y direcciones de Estados Unidos.
rule_Individual_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de Estados Unidos.

Nombre	Descripción
rule_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.
rule_Individual_Name_and_Phone_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números de teléfono.
rule_Individual_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas.

CAPÍTULO 7

Acelerador de Francia

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador de Francia, 81](#)
- [Reglas de limpieza de datos de direcciones de Francia, 81](#)
- [Reglas compuestas de Francia, 82](#)
- [Reglas de limpieza de datos de contacto de Francia, 83](#)
- [Reglas de limpieza de datos corporativos de Francia, 85](#)
- [Reglas de limpieza de datos generales de Francia, 86](#)
- [Reglas de coincidencia y eliminación de duplicados de Francia, 87](#)
- [Asignaciones de demostración de Francia, 89](#)

Resumen del acelerador de Francia

Utilice las reglas del acelerador de Francia para comprobar y mejorar datos de organizaciones en Francia.

El acelerador de Francia incluye reglas que realizan los siguientes procesos de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos corporativos
- Limpieza de datos generales
- Coincidencia y eliminación de duplicados

El acelerador depende de las reglas de limpieza de datos que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de direcciones de Francia

Utilice las reglas de limpieza de datos de direcciones para analizar, estandarizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de direcciones del acelerador de Francia:

Nombre	Descripción
rule_FRA_Address_Parse_Hybrid	Divide direcciones de Francia sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_FRA_Address_Parse_Multiline	Divide direcciones de Francia sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_FRA_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones francesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_FRA_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones francesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_FRA_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones francesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_FRA_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones francesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_FRA_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones francesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_FRA_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones francesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Reglas compuestas de Francia

Utilice las reglas compuestas en el acelerador de Francia para agregar un conjunto de reglas a una asignación como un único objeto. Una regla compuesta es una regla que utiliza la lógica de otras reglas del acelerador.

Las reglas compuestas del acelerador de Francia se instalan en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

Regla compuesta de datos de contacto de Francia

La siguiente tabla describe la regla compuesta de datos de contacto del acelerador de Francia:

Nombre	Descripción
rule_FRA_Contact_Data	Analiza, estandariza y valida datos de contacto de Francia, tales como direcciones y números de teléfono.

La regla `rule_FRA_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en `rule_FRA_Contact_Data`:

Regla	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_FRA_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_FRA_Gender_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Phone_Number_Standardize	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Reglas de limpieza de datos de contacto de Francia

Utilice las reglas de limpieza de datos de contacto para analizar, estandarizar y validar datos sobre los contactos empresariales y los individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de contacto del acelerador de Francia:

Nombre	Descripción
rule_FRA_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, esta regla asigna el sexo "M" al nombre "Jean Leclerc", ya que es masculino.
rule_FRA_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres franceses.
rule_FRA_INSEE_Parse	Reconoce el número del INSEE (Institut National de la Statistique et des Études Économiques, instituto nacional de estadística y estudios económicos de Francia) en una cadena.
rule_FRA_INSEE_Standardization	Estandariza el formato del número del INSEE como una cadena de 13 dígitos, o bien como una cadena de 13 dígitos seguidos de un espacio y la clave de control de dos dígitos.
rule_FRA_INSEE_Validation	Valida el número del INSEE a partir de los valores del sexo, la fecha y el COG (Code Officiel Géographique, código oficial geográfico).
rule_FRA_Multi_Person_Name_Parse	<p>Analiza valores de nombre de persona en puertos independientes. La regla crea puertos para valores como el tratamiento, el nombre, el segundo nombre y los apellidos.</p> <p>La salida de la regla incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Cuando los datos de nombre identifican más de una persona, la regla crea un puerto de salida para cada nombre completo. Por ejemplo, la regla puede leer el nombre "Jean et Marianne Leclerc" y crear puertos de salida para "Jean Leclerc" y "Marianne Leclerc".</p>
rule_FRA_Personal_Name_Parse FML	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Nombre, segundo nombre, apellidos <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_FRA_Personal_Name_Parse LFM	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>

Nombre	Descripción
rule_FRA_Phone_Number_Parse	<p>Reconoce un número de teléfono francés en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda.</p> <p>La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla. La regla procesa los siguientes signos de puntuación: el signo más, los paréntesis y una almohadilla. Antes de ejecutar la regla, quite todos los demás signos de puntuación, incluidos los dobles espacios.</p> <p>La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.</p>
rule_FRA_Phone_Number_Standardize	<p>Estandariza los números de teléfono franceses según formatos de marcación internacional y local. La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla.</p>
rule_FRA_Phone_Number_Validation	<p>Valida el código de área y la longitud de los números de teléfono de Francia. La regla devuelve la región del número de teléfono, así como los códigos que determinan la validez del código de área y la longitud de un número de teléfono.</p>
rule_FRA_Prenameline_Assignment	<p>Genera un tratamiento según el sexo.</p>
rule_FRA_Salutation_Assignment	<p>Genera saludos formales e informales a partir de nombres de pila y tokens de nombre. Por ejemplo, si los datos de entrada contienen "M. Jean Leclerc," la regla genera el saludo formal "Monsieur Leclerc," y el saludo informal "Cher Jean,". Puede cambiar el prefijo y la puntuación si edita las variables en la transformación de expresión dq.Generate_Salutation.</p>

Dependencias de reglas de limpieza de datos de contacto del núcleo

El acelerador de Francia depende de las siguientes reglas de limpieza de datos de contacto del acelerador del núcleo:

- rule_Email_Validation

Para obtener más información sobre estas reglas, consulte ["Reglas de limpieza de datos de contacto del núcleo" en la página 22.](#)

Reglas de limpieza de datos corporativos de Francia

Utilice las reglas de limpieza de datos corporativos para estandarizar datos corporativos.

Puede encontrar las reglas de limpieza de datos corporativos en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos corporativos del acelerador de Francia:

Nombre	Descripción
rule_FRA_Company_Name_Standardization	Estandariza los nombres de empresas según valores de tabla de referencia.
rule_FRA_SIRET_Number_Parse	Reconoce el número del SIRET (système d'identification du répertoire des établissements, sistema de identificación del directorio de empresas de Francia) en una cadena.
rule_FRA_SIRET_Number_Standardize	Estandariza un número de 14 dígitos según el formato NNN NNN NNN NNNNN, independientemente del espaciado o la puntuación de la cadena. No es posible estandarizar una cadena de menos de 14 dígitos.
rule_FRA_SIRET_Number_Validation	Valida el número del SIRET. Tras aplicar la regla rule_FRA_SIRET_Number_Standardization, la regla asume que el número tiene el formato estándar.

Reglas de limpieza de datos generales de Francia

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en los campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador de Francia:

Nombre	Descripción
rule_FRA_NER_Field_Identification	Identifica el tipo de información contenida en un campo de entrada. La regla puede identificar nombres, identificaciones personales, nombres de empresas, fechas y datos de direcciones de Francia. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia para identificar los tipos de información. La regla emplea técnicas de coincidencia probabilística para identificar los tipos de información.

El acelerador de Francia depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Assign_DQ_GeocodingStatus_Description
- rule_Assign_DQ_Mailability_Score_Description
- rule_Assign_DQ_Match_Code_Description
- rule_Luhn_Algorithm
- rule_Remove_Extra_Spaces
- rule_Remove_Parentheses
- rule_Remove_Punctuation
- rule_Remove_Punctuation_and_Space

- rule_Replace_Limited_Punct_with_Space
- rule_UpperCase

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos generales del núcleo” en la página 23.](#)

Reglas de coincidencia y eliminación de duplicados de Francia

Utilice las reglas de coincidencia y eliminación de duplicados para generar puntuaciones de coincidencia y detectar los registros duplicados.

Las reglas de coincidencia y eliminación de duplicados del acelerador de Francia se instalan en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador de Francia:

Nombre	Descripción
mplt_Company_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas. El mapplet genera códigos Soundex a partir de los valores de nombre de empresa y los emplea como claves de grupo.
mplt_FRA_Company_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas y direcciones. El mapplet utiliza una combinación de caracteres de los valores de nombre de empresa y código postal para generar claves de grupo.
mplt_FRA_Famillyname_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de apellidos y direcciones. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_FRA_Firstname_and_INSEE_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir del número del INSEE (Institut National de la Statistique et des Études Économiques, instituto nacional de estadística y estudios económicos de Francia). El mapplet genera claves de grupo a partir de los datos de números del INSEE.
mplt_FRA_Firstname_SurnameDOB_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas de nombres de personas, fechas de nacimiento y códigos postales. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_FRA_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Francia a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_FRA_IMO_Famillyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Francia a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.

Nombre	Descripción
mplt_FRA_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Francia a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_FRA_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Francia a partir de nombres de personas y datos personales. Los campos en la columna de datos personales deben contener un único tipo de datos, tales como número de teléfono, dirección de correo electrónico o número del INSEE. El mapplet genera claves de grupo a partir de los datos personales.
mplt_FRA_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y datos de fechas de Francia. El mapplet genera claves de grupo a partir de fechas.
mplt_FRA_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y direcciones de correo electrónico de Francia. El mapplet genera claves de grupo a partir de los datos de direcciones de correo electrónico.
mplt_FRA_Individual_Name_and_INSEE_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas de Francia y números del INSEE. El mapplet genera claves de grupo a partir de los datos de números del INSEE.
mplt_FRA_Individual_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas de Francia. El mapplet genera códigos NYSIS a partir de los valores de apellido y los emplea como claves de grupo.
rule_Company_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas.
rule_FRA_Company_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y direcciones de Francia.
rule_FRA_Familyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de apellidos y direcciones de Francia.
rule_FRA_Firstname_and_INSEE_MatchScore	Genera una puntuación de coincidencia a partir de nombres de pila y datos de la columna de datos personales, tales como número de teléfono, correo electrónico o número del INSEE.
rule_FRA_Firstname_Surname_DOB_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de apellidos, fechas de nacimiento y códigos postales.
rule_FRA_Individual_Name_and_INSEE_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números del INSEE.
rule_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.
rule_Individual_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas.

Asignaciones de demostración de Francia

Las asignaciones de demostración del acelerador de Francia utilizan varias reglas para comprobar los procesos de calidad de datos.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\FRA_Accelerator
```

El acelerador contiene las siguientes asignaciones de demostración:

m_FRA_customer_data_demo

Analiza, estandariza y valida datos franceses.

m_FRA_customer_matching_demo

Analiza y estandariza datos de identidad de Australia y Nueva Zelanda, y realiza un análisis de coincidencia de identidades de los datos.

La asignación analiza las siguientes combinaciones de datos y genera clústeres de coincidencia para cada combinación:

- Nombre de persona y datos de direcciones
- Nombre de persona y número de teléfono

CAPÍTULO 8

Acelerador de Alemania

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador de Alemania, 90](#)
- [Reglas de limpieza de datos de direcciones de Alemania, 91](#)
- [Reglas compuestas de Alemania, 92](#)
- [Reglas de limpieza de datos de contacto de Alemania, 93](#)
- [Reglas de limpieza de datos corporativos de Alemania, 94](#)
- [Reglas de limpieza de datos generales de Alemania, 95](#)
- [Reglas de coincidencia y eliminación de duplicados de Alemania, 96](#)
- [Asignaciones de demostración de Alemania, 98](#)

Resumen del acelerador de Alemania

Utilice las reglas del acelerador de Alemania para comprobar y mejorar datos de organizaciones en Alemania.

El acelerador de Alemania incluye reglas que realizan los siguientes procesos de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos corporativos
- Limpieza de datos generales
- Coincidencia y eliminación de duplicados

El acelerador depende de las reglas de limpieza de datos que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de direcciones de Alemania

Utilice las reglas de limpieza de datos de direcciones para analizar, estandarizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de direcciones del acelerador de Alemania:

Nombre	Descripción
rule_DEU_Address_Parse_Hybrid	Divide direcciones de Alemania sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_DEU_Address_Parse_Multiline	Divide direcciones de Alemania sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_DEU_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones alemanas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_DEU_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones alemanas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_DEU_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones alemanas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_DEU_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones alemanas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_DEU_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones alemanas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_DEU_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones alemanas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Reglas compuestas de Alemania

Utilice las reglas compuestas en el acelerador de Alemania para agregar un conjunto de reglas a una asignación como un único objeto. Una regla compuesta es una regla que utiliza la lógica de otras reglas del acelerador.

Puede encontrar las reglas compuestas en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

Regla compuesta de datos de contacto de Alemania

La siguiente tabla describe la regla compuesta de datos de contacto del acelerador de Alemania:

Nombre	Descripción
rule_DEU_Contact_Data	Analiza, estandariza y valida datos de contacto de Alemania, tales como direcciones y números de teléfono.

La regla `rule_DEU_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en `rule_DEU_Contact_Data`:

Regla	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_DEU_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_DEU_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_DEU_Gender_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_DEU_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_DEU_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_DEU_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_DEU_Prename_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_DEU_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Reglas de limpieza de datos de contacto de Alemania

Utilice las reglas de limpieza de datos de contacto para analizar, estandarizar y validar datos sobre los contactos empresariales y los individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de contacto del acelerador de Alemania:

Nombre	Descripción
rule_DEU_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, esta regla asigna el sexo "M" al nombre "Hans Adler", ya que es masculino.
rule_DEU_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres alemanes.
rule_DEU_Multi_Person_Name_Parse	<p>Analiza valores de nombre de persona en puertos independientes. La regla crea puertos para valores como el tratamiento, el nombre, el segundo nombre y los apellidos.</p> <p>La salida de la regla incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Cuando los datos de nombre identifican más de una persona, la regla crea un puerto de salida para cada nombre completo. Por ejemplo, la regla puede leer el nombre "Hans und Maria Adler" y crear puertos de salida para "Hans Adler" y "Maria Adler."</p>
rule_DEU_Personal_Name_Parse_FML	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none">- Nombre, segundo nombre, apellidos <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_DEU_Personal_Name_Parse_LFM	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none">- Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>

Nombre	Descripción
rule_DEU_Phone_Number_Parse	<p>Reconoce un número de teléfono alemán en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda.</p> <p>La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla. La regla procesa los siguientes signos de puntuación: el signo más, los paréntesis y una almohadilla. Antes de ejecutar la regla, quite todos los demás signos de puntuación, incluidos los dobles espacios.</p> <p>La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.</p>
rule_DEU_Phone_Number_Standardization	<p>Estandariza los números de teléfono alemanes según formatos de marcación internacional y local. La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla.</p>
rule_DEU_Phone_Number_Validation	<p>Valida el código de área y la longitud de los números de teléfono de Alemania. La regla devuelve la región del número de teléfono, así como los códigos que determinan la validez del código de área y la longitud de un número de teléfono.</p>
rule_DEU_Prefix_Assignment	<p>Genera un tratamiento según el sexo.</p>
rule_DEU_Salutation_Assignment	<p>Genera saludos formales e informales a partir de nombres de pila y tokens de nombre. Por ejemplo, si los datos de entrada contienen "Herr Hans Adler," la regla genera el saludo formal "Sehr geehrter Herr Adler," y el saludo informal "Lieber Hans,". Puede cambiar el prefijo y la puntuación si edita las variables en la transformación de expresión dq.Generate_Salutation.</p>

Dependencias de reglas de limpieza de datos de contacto del núcleo

El acelerador de Alemania depende de las siguientes reglas de limpieza de datos de contacto del acelerador del núcleo:

- rule_Email_Validation

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos de contacto del núcleo” en la página 22.](#)

Reglas de limpieza de datos corporativos de Alemania

Utilice las reglas de limpieza de datos corporativos para estandarizar datos corporativos.

Puede encontrar las reglas de limpieza de datos corporativos en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos corporativos del acelerador de Alemania:

Nombre	Descripción
rule_DEU_Company_Name_Standardization	Utiliza tablas de referencia para estandarizar nombres de empresas.

Reglas de limpieza de datos generales de Alemania

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en los campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador de Alemania:

Nombre	Descripción
rule_DEU_NER_Field_Identification	Identifica el tipo de información contenida en un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de Alemania. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia para identificar los tipos de información. La regla emplea técnicas de coincidencia probabilística para identificar los tipos de información.

Dependencias de reglas de limpieza de datos generales del núcleo

El acelerador de Alemania depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Assign_DQ_GeocodingStatus_Description
- rule_Assign_DQ_Mailability_Score_Description
- rule_Assign_DQ_Match_Code_Descriptions
- rule_Remove_Extra_Spaces
- rule_Remove_Hyphen
- rule_Remove_Leading_Zero
- rule_Remove_Parentheses
- rule_Remove_Period_Parentheses
- rule_Remove_Punctuation
- rule_Remove_Punctuation_and_Space
- rule_Remove_Space
- rule_Replace_Limited_Punct_with_Space
- rule_UpperCase

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos generales del núcleo” en la página 23.](#)

Reglas de coincidencia y eliminación de duplicados de Alemania

Utilice las reglas de coincidencia y eliminación de duplicados para generar puntuaciones de coincidencia y detectar los registros duplicados.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador de Alemania:

Nombre	Descripción
mplt_Company_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas. El mapplet genera códigos Soundex a partir de los valores de nombre de empresa y los emplea como claves de grupo.
mplt_DEU_Company_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de nombres de empresas y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de nombre de empresa y código postal para generar claves de grupo.
mplt_DEU_Familyname_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de apellidos y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_DEU_Firstname_3CharsSurname_DOB_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de los nombres de personas, los tres primeros caracteres de los apellidos, las fecha de nacimiento y los códigos postales. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_DEU_Firstname_and_PID_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de nombres de personas e id. personales agrupadas. El mapplet genera claves de grupo a partir de los datos de ID personales.
mplt_DEU_Firstname_Surname_2ElementsDOB_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de nombres de personas, dos elementos de la fecha de nacimiento y códigos postales. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_DEU_Firstname_Surname_DOB_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de nombres de personas, fecha de nacimiento y códigos postales. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_DEU_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Alemania a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_DEU_IMO_Familyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Alemania a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.

Nombre	Descripción
mplt_DEU_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Alemania a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_DEU_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Alemania a partir de nombres de personas y datos personales. Los campos de la columna de los datos personales deben contener un único tipo de datos, tales como número de teléfono, dirección de correo electrónico o ID personal. El mapplet genera claves de grupo a partir de los datos personales.
mplt_DEU_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y datos de fechas agrupados por fecha. El mapplet genera claves de grupo a partir de los datos de fechas.
mplt_DEU_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de nombres de personas y direcciones de correo electrónico. El mapplet genera claves de grupo a partir de los datos de direcciones de correo electrónico.
mplt_DEU_Individual_Name_and_Phone_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de nombres de personas y números de teléfono. El mapplet genera claves de grupo a partir de los datos de números de teléfono.
mplt_DEU_Individual_Name_and_PID_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de nombres de personas e id. personales. El mapplet genera claves de grupo a partir de los datos de ID personales.
mplt_DEU_Individual_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Alemania a partir de nombres de personas. El mapplet genera códigos NYSIIS a partir de valores de apellido y los emplea como claves de grupo.
rule_Company_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas.
rule_DEU_Company_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y direcciones.
rule_DEU_Familyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de apellidos y direcciones.
rule_DEU_Firstname_3CharsSurname_DOB_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de los nombres de pila, los tres primeros caracteres de los apellidos, la fecha de nacimiento y los códigos postales.
rule_DEU_Firstname_and_PID_MatchScore	Genera una puntuación de coincidencia a partir de nombres de pila y datos de la columna de datos personales, tales como número de teléfono, correo electrónico o ID personal.
rule_DEU_Firstname_Surname_2ElementsDOB_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas, fechas de nacimiento y códigos postales. Nota: El formato de entrada de la fecha de nacimiento se supone que es DD/MM/AAAA.
rule_DEU_Firstname_Surname_DOB_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de los apellidos, las fechas de nacimiento y los códigos postales.

Nombre	Descripción
rule_DEU_Individual_Name_and_Phone_MatchScore	Genera una puntuación de coincidencia a partir de los nombres de personas y los números de teléfono.
rule_Familyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de los apellidos y las direcciones.
rule_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.
rule_Individual_Name_and_SSN_MatchScore	Genera una puntuación de coincidencia a partir de nombres de pila y datos de la columna de datos personales, tales como número de teléfono, correo electrónico o número de la seguridad social.
rule_Individual_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas.

Asignaciones de demostración de Alemania

Las asignaciones de demostración del acelerador de Alemania utilizan varias reglas para comprobar los procesos de calidad.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\DEU_Accelerator
```

El acelerador contiene las siguientes asignaciones de demostración:

m_DEU_customer_data_demo

Analiza, estandariza y valida datos alemanes.

m_DEU_customer_matching_demo

Analiza y estandariza datos de identidad de Alemania, y realiza un análisis de coincidencia de identidades de los datos.

La asignación analiza las siguientes combinaciones de datos y genera clústeres de coincidencia para cada combinación:

- Nombre de persona y datos de direcciones
- Nombre de persona y número de teléfono

CAPÍTULO 9

Acelerador de Portugal

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador de Portugal, 99](#)
- [Reglas de limpieza de datos de direcciones de Portugal, 99](#)
- [Reglas compuestas de Portugal, 100](#)
- [Reglas de limpieza de datos de contacto de Portugal, 101](#)
- [Reglas de limpieza de datos corporativos de Portugal, 103](#)
- [Reglas de limpieza de datos generales de Portugal, 103](#)
- [Reglas de coincidencia y eliminación de duplicados de Portugal, 104](#)
- [Asignaciones de demostración de Portugal, 106](#)

Resumen del acelerador de Portugal

Utilice las reglas del acelerador de Portugal para comprobar y mejorar datos de organizaciones en Portugal.

El acelerador de Portugal incluye reglas que realizan los siguientes procesos de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos corporativos
- Limpieza de datos generales
- Coincidencia y eliminación de duplicados

El acelerador depende de las reglas de limpieza de datos que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de direcciones de Portugal

Utilice las reglas de limpieza de datos de direcciones para analizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de direcciones del acelerador de Portugal:

Nombre	Descripción
rule_PRT_Address_Parse_Hybrid	Divide direcciones de Portugal sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_PRT_Address_Parse_Multiline	Divide direcciones de Portugal sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_PRT_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones portuguesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_PRT_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones portuguesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_PRT_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones portuguesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_PRT_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones portuguesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_PRT_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones portuguesas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_PRT_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones portuguesas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Reglas compuestas de Portugal

Utilice las reglas compuestas en el acelerador de Portugal para agregar un conjunto de reglas a una asignación como un único objeto. Una regla compuesta es una regla que utiliza la lógica de otras reglas del acelerador.

Puede encontrar las reglas compuestas en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

Regla compuesta de datos de contacto de Portugal

La siguiente tabla describe la regla compuesta de datos de contacto de Portugal del acelerador de Portugal:

Nombre	Descripción
rule_PRT_Contact_Data	Analiza, estandariza y valida datos de contacto de Portugal, tales como direcciones, números de teléfono y NIF (Número de Identificação Fiscal, número de identificación fiscal de Portugal).

La regla `rule_PRT_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en `rule_PRT_Contact_Data`:

Regla	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_PRT_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_PRT_NIF_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_NIF_Validate	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Personal_Name_Parse_Validate	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Reglas de limpieza de datos de contacto de Portugal

Utilice las reglas de limpieza de datos de contacto para analizar, estandarizar y validar datos sobre los contactos empresariales y los individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de contacto del acelerador de Portugal:

Nombre	Descripción
rule_PRT_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, la regla asigna el sexo "M" al nombre "Artur Cruz", ya que es masculino.
rule_PRT_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres portugueses.
rule_PRT_NIF_Parse	Reconoce el NIF (Número de Identificação Fiscal, número de identificación fiscal de Portugal) en cadenas. La regla devuelve los números de ID, así como una cadena que contiene el texto de entrada sin los números de ID.
rule_PRT_NIF_Standardization	Estandariza el formato del NIF como una cadena de nueve dígitos. La regla quita los caracteres alfabéticos, los símbolos y los espacios.
rule_PRT_NIF_Validate	Valida el NIF a partir del dígito de comprobación de cada número. La regla requiere que la entrada sea una cadena numérica de nueve dígitos sin espacios.
rule_PRT_Personal_Name_Parse_Validate	<p>Analiza valores de nombre de persona en puertos independientes. La regla crea puertos para valores como el tratamiento, el nombre, el segundo nombre y los apellidos. La regla también indica si el nombre puede ser un nombre de empresa y valida la ortografía del nombre.</p> <p>La salida de la regla incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_PRT_Phone_Number_Parse	Reconoce un número de teléfono portugués en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda. La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.
rule_PRT_Phone_Number_Standardization	Estandariza los números de teléfono portugueses según formatos de marcación internacional y local.
rule_PRT_Phone_Number_Validation	Valida el código de área y la longitud de los números de teléfono de Portugal. La regla devuelve la región del número de teléfono, así como los códigos que determinan la validez del código de área y la longitud de un número de teléfono.
rule_PRT_Prenome_Assignment	Genera un tratamiento según el sexo. Puede cambiar la variable de expresión female_prenome de "Sra" a "Sta".
rule_PRT_Salutation_Assignment	Genera saludos formales e informales a partir de nombres de pila y tokens de nombre. Por ejemplo, si los datos de entrada contienen "Sr. Artur Cruz," la regla genera el saludo formal "Prezado Sr. Cruz," así como el saludo informal "Prezado Artur,". Puede cambiar el prefijo y la puntuación si edita las variables en la transformación de expresión dq_Generate_Salutation.

Dependencias de reglas de limpieza de datos de contacto del núcleo

El acelerador de Portugal depende de las siguientes reglas de limpieza de datos de contacto del acelerador del núcleo:

- rule_Email_Validation

Para obtener más información sobre estas reglas, consulte ["Reglas de limpieza de datos de contacto del núcleo" en la página 22.](#)

Reglas de limpieza de datos corporativos de Portugal

Utilice las reglas de limpieza de datos corporativos para analizar, estandarizar y validar datos corporativos.

Puede encontrar las reglas de limpieza de datos corporativos en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos corporativos del acelerador de Portugal:

Nombre	Descripción
rule_PRT_Company_Name_Standardization	Estandariza nombres de empresas de Portugal según valores de tabla de referencia.
rule_PRT_NIPC_Parse	Analiza un NIPC (Número de Identificação Pessoa Colectiva, número de identificación de persona jurídica de Portugal). La regla devuelve el NIPC, así como una cadena que contiene el texto de entrada sin el NIPC.
rule_PRT_NIPC_Standardize	Estandariza el formato del NIPC como una cadena de nueve dígitos. La regla quita los caracteres alfabéticos, los símbolos y los espacios.
rule_PRT_NIPC_Validate	Valida un número NIPC a partir del dígito de comprobación de cada número. La regla requiere que la entrada sea una cadena de nueve dígitos.

Reglas de limpieza de datos generales de Portugal

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en los campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador de Portugal:

Nombre	Descripción
rule_PRT_NER_Field_Identification	Identifica el tipo de información contenida en un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de Portugal. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia para identificar los tipos de información. La regla emplea técnicas de coincidencia probabilística para identificar los tipos de información.

Dependencias de reglas de limpieza de datos generales del núcleo

El acelerador de Portugal depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Assign_DQ_GeocodingStatus_Description
- rule_Assign_DQ_ElementResultStatus_Description
- rule_Assign_DQ_Mailability_Score_Description
- rule_Assign_DQ_Match_Code_Descriptions

- rule_Parse_First_Word
- rule_Remove_Extra_Spaces
- rule_Remove_Non_Numbers
- rule_Remove_Punctuation
- rule_Remove_Punctuation_and_Space
- rule_Replace_Limited_Punct_with_Space
- rule_UpperCase

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos generales del núcleo” en la página 23.](#)

Reglas de coincidencia y eliminación de duplicados de Portugal

Utilice las reglas de coincidencia y eliminación de duplicados para generar puntuaciones de coincidencia y detectar los registros duplicados.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador de Portugal:

Nombre	Descripción
mplt_Company_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas. El mapplet genera códigos Soundex a partir de los valores de nombre de empresa y los emplea como claves de grupo.
mplt_PRT_Company_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Portugal a partir de nombres de empresas y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de nombre de empresa y código postal para generar claves de grupo.
mplt_PRT_Famillyname_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Portugal a partir de apellidos y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_PRT_Firstname_and_NIF_BI_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Portugal a partir de nombres de pila y números de identificación personal, tales como el NIF (Número de Indentificação Fiscal, número de identificación fiscal de Portugal) y el BI (Bilhete de Identidade, tarjeta de identidad de Portugal). El mapplet genera claves de grupo a partir de los datos de números de identificación personal.
mplt_PRT_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Portugal a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.

Nombre	Descripción
mplt_PRT_IMO_Famillyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Portugal a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_PRT_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Portugal a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_PRT_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Portugal a partir de nombres de personas y datos personales. Los campos de la columna de datos personales deben contener un único tipo de datos, tales como número de teléfono, dirección de correo electrónico o NIF. El mapplet genera claves de grupo a partir de los datos personales.
mplt_PRT_Individual_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Portugal a partir de nombres de personas y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_PRT_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Portugal a partir de nombres de personas y datos de fechas. El mapplet genera claves de grupo a partir de los datos de fechas.
mplt_PRT_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Portugal a partir de nombres de personas y direcciones de correo electrónico. El mapplet genera claves de grupo a partir de los datos de direcciones de correo electrónico.
mplt_PRT_Individual_Name_and_Phone_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Portugal a partir de nombres de personas y números de teléfono. El mapplet genera claves de grupo a partir de los datos de números de teléfono.
mplt_PRT_Individual_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Portugal a partir de nombres de personas. El mapplet genera códigos NYSIS a partir de los valores de apellido y los emplea como claves de grupo.
rule_Company_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas.
rule_PRT_Company_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y datos de direcciones de Portugal.
rule_PRT_Famillyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de apellidos y datos de direcciones de Portugal.
rule_PRT_Firstname_and_NIF_BI_MatchScore	Genera una puntuación de coincidencia a partir de los datos de nombres de pila, el NIF y el número de BI.
rule_PRT_Individual_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y datos de direcciones de Portugal.
rule_PRT_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_PRT_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.

Nombre	Descripción
rule_PRT_Individual_Name_and_Phone_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números de teléfono.
rule_PRT_Individual_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas.

Asignaciones de demostración de Portugal

Las asignaciones de demostración del acelerador de Portugal utilizan varias reglas para comprobar los procesos de calidad de datos.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\PRT_Accelerator
```

El acelerador contiene las siguientes asignaciones de demostración:

m_PRT_customer_data_demo

Analiza, estandariza y valida datos portugueses.

m_PRT_customer_matching_demo

Analiza y estandariza datos de identidad de Portugal, y realiza un análisis de coincidencia de identidades de los datos.

La asignación analiza las siguientes combinaciones de datos y genera clústeres de coincidencia para cada combinación:

- Nombre de persona y datos de direcciones
- Nombre de persona y número de teléfono

CAPÍTULO 10

Acelerador de España

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador de España, 107](#)
- [Reglas de limpieza de datos de direcciones de España, 107](#)
- [Reglas de limpieza de datos de contacto de España, 109](#)
- [Reglas de limpieza de datos corporativos de España, 110](#)
- [Reglas de limpieza de datos generales de España, 110](#)
- [Reglas de coincidencia y eliminación de duplicados de España, 111](#)
- [Asignaciones de demostración de España, 114](#)

Resumen del acelerador de España

Utilice las reglas del acelerador de España para comprobar y mejorar datos de organizaciones en España.

El acelerador de España incluye reglas que realizan los siguientes procesos de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos corporativos
- Limpieza de datos generales
- Coincidencia y eliminación de duplicados

El acelerador depende de las reglas de limpieza de datos que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de direcciones de España

Utilice las reglas de limpieza de datos de direcciones para analizar, estandarizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de direcciones del acelerador de España:

Nombre	Descripción
rule_ESP_Address_Parse_Hybrid	Divide direcciones de España sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_ESP_Address_Parse_Multiline	Divide direcciones de España sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_ESP_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones españolas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_ESP_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones españolas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_ESP_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones españolas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_ESP_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones españolas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_ESP_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones españolas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_ESP_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones españolas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Reglas de limpieza de datos de contacto de España

Utilice las reglas de limpieza de datos de contacto para analizar, estandarizar y validar datos sobre los contactos empresariales y los individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de contacto del acelerador de España:

Nombre	Descripción
rule_ESP_CIF_Parse	Analiza el CIF (Certificado de Identificación Fiscal de España).
rule_ESP_CIF_Standardization	Estandariza el CIF.
rule_ESP_CIF_Validation	Valida el CIF.
rule_ESP_DNI_Parse	Analiza el DNI (Documento Nacional de Identidad de España).
rule_ESP_DNI_Standardization	Estandariza el DNI.
rule_ESP_DNI_Validation	Valida el DNI.
rule_ESP_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, la regla asigna el sexo "M" al nombre "Juan García", ya que es masculino.
rule_ESP_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres españoles.
rule_ESP_NIE_Parse	Analiza el NIE (Número de Identidad de Extranjero de España).
rule_ESP_NIE_Standardization	Estandariza el NIE.
rule_ESP_NIE_Validation	Valida el NIE.
rule_ESP_NIF_Parse	Reconoce el NIF (Número de identificación Fiscal de España) en una cadena.
rule_ESP_NIF_Standardization	Estandariza el NIF.
rule_ESP_NIF_Validation	Valida el NIF.
rule_ESP_Personal_Name_Parse	Analiza nombres de personas de España.
rule_ESP_Phone_Number_Parse	Reconoce un número de teléfono español en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda. La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.
rule_ESP_Phone_Number_Standardization	Estandariza los números de teléfono españoles según formatos de marcación internacional y local. La regla reconoce números de teléfono que contengan ceros iniciales, códigos de marcación internacional o extensiones que comienzan con una almohadilla.

Nombre	Descripción
rule_ESP_Phone_Number_Validation	Valida el código de área y la longitud de los números de teléfono de España. La regla devuelve la región del número de teléfono, así como los códigos que determinan la validez del código de área y la longitud de un número de teléfono.
rule_ESP_Phone_Parse_Standardize_Validate	Analiza números de teléfono españoles y estandariza sus formatos. Valida el código de área y la longitud de los números de teléfono de España.
rule_ESP_Prename_Assignment	Genera un tratamiento según el sexo.
rule_ESP_Salutation_Assignment	Genera saludos formales e informales a partir de nombres de pila y tokens de nombre. Por ejemplo, si los datos de entrada contienen "Sr. Juan García,", la regla genera el saludo formal "Estimado Sr. García," y el saludo informal "Querido Juan,". Puede cambiar el prefijo y la puntuación si edita las variables en la transformación de expresión dq_Generate_Salutation.

Dependencias de reglas de limpieza de datos de contacto del núcleo

El acelerador de España depende de las siguientes reglas de limpieza de datos de contacto del acelerador del núcleo:

- rule_Email_Validation

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos de contacto del núcleo” en la página 22.](#)

Reglas de limpieza de datos corporativos de España

Utilice las reglas de limpieza de datos corporativos para estandarizar datos corporativos.

El acelerador de España depende de la siguiente regla de limpieza de datos corporativos del acelerador del núcleo:

- rule_Company_Name_Standardization

Reglas de limpieza de datos generales de España

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en los campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]Rules\General_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador de España:

Nombre	Descripción
rule_ESP_NER_Field_Identification	Identifica el tipo de información contenida en un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de España. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea técnicas de coincidencia probabilística para identificar los tipos de información.

El acelerador de España depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Assign_DQ_ElementResultStatus_Description
- rule_Assign_DQ_GeocodingStatus_Description
- rule_Assign_DQ_Match_Code_Descriptions
- rule_Remove_Extra_Spaces
- rule_Remove_Leading_Zero
- rule_Remove_Limited_Punctuation
- rule_Remove_Non_Numbers
- rule_Remove_Punctuation_and_Space
- rule_Remove_Punctuation
- rule_Replace_limited_Punct_with_Space
- rule_Translate_Diacritic_Characters
- rule_UpperCase

Reglas de coincidencia y eliminación de duplicados de España

Utilice las reglas de coincidencia y eliminación de duplicados para generar puntuaciones de coincidencia y detectar los registros duplicados.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador de España:

Nombre	Descripción
mplt_Company_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas.
mplt_ESP_Company_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de nombres de empresas y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de nombre de empresa y código postal para generar claves de grupo.
mplt_ESP_Familyname_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de apellidos y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_ESP_Firstname_and_NIF_BI_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de nombres de pila y números de identificación personal, tales como el NIF (Número de Identificación Fiscal de España). El mapplet genera claves de grupo a partir de los datos de números de identificación personal.
mplt_ESP_IMO_Company_Name_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de España a partir de nombres de empresas. El mapplet genera códigos Soundex a partir de los valores de nombre de empresa y los emplea como claves de grupo.
mplt_ESP_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de España a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_ESP_IMO_Familyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de España a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_ESP_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_ESP_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de nombres de personas y datos personales. Los campos de la columna de datos personales deben contener un único tipo de datos, tales como número de teléfono o dirección de correo electrónico. El mapplet genera claves de grupo a partir de los datos personales.

Nombre	Descripción
mplt_ESP_Individual_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de nombres de personas. El mapplet genera códigos NYSIIS a partir de los valores de apellido y los emplea como claves de grupo.
mplt_ESP_Individual_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de nombres de personas y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_ESP_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de nombres de personas y fechas. El mapplet genera claves de grupo a partir de los datos de fechas.
mplt_ESP_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de direcciones de correo electrónico y nombres de personas. El mapplet genera claves de grupo a partir de los datos de direcciones de correo electrónico.
mplt_ESP_Individual_Name_and_Phone_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de España a partir de nombres de personas y números de teléfono. El mapplet genera claves de grupo a partir de los datos de números de teléfono.
rule_Company_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas.
rule_ESP_Company_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y datos de direcciones de España.
rule_ESP_Familyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de apellidos y datos de direcciones de España.
rule_ESP_Firstname_and_NIF_BI_Matchscore	Genera una puntuación de coincidencia a partir del nombre de pila y el NIF.
rule_ESP_Individual_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas.
rule_ESP_Individual_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y datos de direcciones de España.
rule_ESP_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_ESP_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.
rule_ESP_Individual_Name_and_Phone_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números de teléfono.

Asignaciones de demostración de España

Las asignaciones de demostración del acelerador de España utilizan varias reglas para comprobar los procesos de calidad de datos.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\ESP_Accelerator
```

El acelerador contiene las siguientes asignaciones de demostración:

m_ESP_customer_data_demo

Analiza, estandariza y valida datos españoles.

m_ESP_customer_matching_demo

Analiza y estandariza datos de identidad de España y realiza un análisis de coincidencia de identidades de los datos.

La asignación analiza las siguientes combinaciones de datos y genera clústeres de coincidencia para cada combinación:

- Nombre de persona y datos de direcciones
- Nombre de persona y número de teléfono

CAPÍTULO 11

Acelerador del Reino Unido

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador del Reino Unido, 115](#)
- [Reglas de limpieza de datos de direcciones del Reino Unido, 116](#)
- [Reglas compuestas del Reino Unido, 117](#)
- [Reglas de limpieza de datos de contacto del Reino Unido, 118](#)
- [Reglas de limpieza de datos corporativos del Reino Unido, 121](#)
- [Reglas de limpieza de datos financieros del Reino Unido, 121](#)
- [Reglas de limpieza de datos generales del Reino Unido, 122](#)
- [Reglas de coincidencia y eliminación de duplicados del Reino Unido, 122](#)
- [Asignaciones de demostración del Reino Unido, 125](#)

Resumen del acelerador del Reino Unido

Utilice las reglas del acelerador del Reino Unido para comprobar y mejorar datos de organizaciones en el Reino Unido.

El acelerador del Reino Unido incluye reglas que realizan los siguientes procesos de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos financieros
- Coincidencia y eliminación de duplicados

El acelerador depende de las reglas de limpieza de datos que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de direcciones del Reino Unido

Utilice las reglas de limpieza de datos de direcciones para analizar, estandarizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de direcciones del acelerador del Reino Unido:

Nombre	Descripción
rule_GBR_Address_Parse_Hybrid	Divide direcciones del Reino Unido sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_GBR_Address_Parse_Multiline	Divide direcciones del Reino Unido sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_GBR_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones británicas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_GBR_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones británicas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_GBR_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones británicas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_GBR_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones británicas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_GBR_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones británicas y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_GBR_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones británicas. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_GBR_Postcode_Parse	Analiza códigos postales del Reino Unido.

Nombre	Descripción
rule_GBR_Postcode_Standardize	<p>Estandariza códigos postales del Reino Unido. La regla requiere que la entrada tenga formatos predefinidos.</p> <p>La regla estandariza entradas que coinciden con los siguientes patrones:</p> <ul style="list-style-type: none"> - A9 9AA - A99 9AA - AA9 9AA - AA99 9AA - A9A 9AA - AA9A 9AA - GIR 0AA <p>La letra A representa un carácter alfabético, mientras que el número 9 representa un dígito.</p>
rule_GBR_Postcode_Validate	<p>Valida códigos postales del Reino Unido. La regla hace coincidir los códigos postales estandarizados con los códigos postales válidos del Reino Unido. Si la regla no encuentra ningún código postal que coincida, determina si el código postal sigue el patrón de los códigos postales del Reino Unido.</p>

Reglas compuestas del Reino Unido

Utilice las reglas compuestas en el acelerador del Reino Unido para agregar un conjunto de reglas a una asignación como un único objeto. Una regla compuesta es una regla que utiliza la lógica de otras reglas del acelerador.

Puede encontrar las reglas compuestas en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

Regla compuesta de datos de contacto del Reino Unido

La siguiente tabla describe la regla compuesta de datos de contacto del acelerador del Reino Unido:

Nombre	Descripción
rule_GBR_Contact_Data	Analiza, estandariza y valida datos de contacto del Reino Unido, tales como las direcciones, los números de teléfono y el NINO (National Insurance Number, número de la seguridad social nacional del Reino Unido).

La regla `rule_GBR_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en `rule_GBR_Contact_Data`:

Nombre	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing

Nombre	Ubicación
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_GBR_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_GBR_Gender_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_NINO_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_NINO_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Reglas de limpieza de datos de contacto del Reino Unido

Utilice las reglas de limpieza de datos de contacto para analizar, estandarizar y validar datos sobre los contactos empresariales y los individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

La tabla siguiente describe las reglas de limpieza de datos de contacto del acelerador del Reino Unido:

Nombre	Descripción
rule_GBR_Driver_Number_Parse	Analiza cadenas que coincidan con el formato de permisos de conducción del Reino Unido.
rule_GBR_Driver_Number_Validation	Valida los números de permiso de conducción del Reino Unido a partir de los requisitos del catálogo de estándares de datos del Gobierno del Reino Unido (Government Data Standards Catalog).
rule_GBR_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, esta regla asigna el sexo "M" al nombre "John Smith", ya que es masculino.
rule_GBR_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres británicos.

Nombre	Descripción
rule_GBR_Multi_Person_Name_Parse	<p>Analiza valores de nombre de persona en puertos independientes. La regla crea puertos para valores como el tratamiento, el nombre, el segundo nombre y los apellidos.</p> <p>La salida de la regla incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Cuando los datos de nombre identifican más de una persona, la regla crea un puerto de salida para cada nombre completo. Por ejemplo, la regla puede leer el nombre "John y Jane Smith" y crear puertos de salida para "John Smith" y "Jane Smith".</p>
rule_GBR_NHS_Number_Parse	Reconoce números del NHS (National Health Service, servicio nacional de sanidad del Reino Unido) en una cadena. La regla devuelve un número del NHS, así como una cadena que contiene el texto de entrada sin el número del NHS.
rule_GBR_NHS_Number_Standardize	Estandariza números del NHS según el formato estándar (999 999 9999). La regla requiere que la entrada sea una cadena de diez dígitos.
rule_GBR_NHS_Number_Validate	Valida números del NHS a partir del dígito de comprobación de cada número. La regla requiere que la entrada sea una cadena de diez dígitos.
rule_GBR_NINO_Conformity_Check	Valida el patrón estándar del NINO (National Insurance Number, número de la seguridad social nacional del Reino Unido). La regla no comprueba que el NINO sea preciso o esté activo.
rule_GBR_NINO_Parse	Reconoce el NINO en cadenas. La regla devuelve el NINO, así como una cadena que contiene el texto de entrada sin el NINO.
rule_GBR_NINO_Standardization	<p>Estandariza el NINO con los dos formatos más habituales. La regla devuelve los siguientes formatos, en los que C representa un carácter alfabético y N un número:</p> <ul style="list-style-type: none"> - CC NN NN NN C - CCNNNNNNC <p>La regla cambia todos los caracteres alfabéticos a mayúsculas. La regla requiere que la entrada coincida con el patrón del NINO.</p>
rule_GBR_NINO_Validation	Valida el NINO. La regla no comprueba que el NINO esté activo.
rule_GBR_Passport_Number_MR_Parse	Analiza números de pasaporte del Reino Unido con formato extendido. El formato extendido es el formato de lectura automática de números de pasaporte.
rule_GBR_Passport_Number_Parse	Reconoce números de pasaporte del Reino Unido que utilizan el formato del catálogo de estándares de datos del Gobierno del Reino Unido (Government Data Standards Catalog). La regla analiza todas las cadenas de nueve dígitos.
rule_GBR_Passport_Number_Validation	Valida los números de pasaporte del Reino Unido que utilizan el formato del catálogo de estándares de datos del Gobierno del Reino Unido (Government Data Standards Catalog).

Nombre	Descripción
rule_GBR_Personal_Name_Parsing_FML	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Nombre, segundo nombre, apellidos <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_GBR_Personal_Name_Parsing_LFM	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_GBR_Phone_Number_Parse	<p>Reconoce un número de teléfono británico en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda.</p> <p>Esta regla reconoce los números de teléfono que empiezan con ceros, el código de marcación internacional "+44" y las extensiones que empiezan con una almohadilla. La regla procesa los siguientes signos de puntuación: el signo más, los paréntesis y una almohadilla. Antes de ejecutar la regla, quite todos los demás signos de puntuación, incluidos los dobles espacios.</p> <p>La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.</p>
rule_GBR_Phone_Number_Standardization	<p>Estandariza los números de teléfono británicos según formatos de marcación internacional y local. Esta regla reconoce los números de teléfono que empiezan con ceros, el código de marcación internacional "+44" y las extensiones que empiezan con una almohadilla.</p>
rule_GBR_Phone_Number_Validation	<p>Valida el código de área y la longitud de los números de teléfono del Reino Unido. La regla devuelve la región del número de teléfono, así como los códigos que determinan la validez del código de área y la longitud de un número de teléfono.</p>
rule_Prefix_Assignment	<p>Genera un tratamiento según el sexo. Puede cambiar la variable de expresión <code>female_prefix</code> de "Ms." a "Mrs.".</p>
rule_Salutation_Assignment	<p>Genera saludos formales e informales a partir de nombres de pila y tokens de nombre. Por ejemplo, si los datos de entrada contienen "Mr. John Smith," la regla genera el saludo formal "Dear Mr. Smith," y el saludo informal "Dear John,". Puede cambiar el prefijo y la puntuación si edita las variables en la transformación de expresión <code>dq.Generate_Salutation</code>.</p>

Reglas de limpieza de datos corporativos del Reino Unido

Utilice las reglas de limpieza de datos corporativos para estandarizar datos corporativos.

Puede encontrar las reglas de limpieza de datos corporativos en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos corporativos del acelerador del Reino Unido:

Nombre	Descripción
rule_GBR_Company_Name_Standardization	Estandariza un nombre de compañía y proporciona las siglas para el nombre si es posible hacerlo.

Reglas de limpieza de datos financieros del Reino Unido

Utilice las reglas de limpieza de datos financieros para analizar, estandarizar y validar datos financieros.

Puede encontrar las reglas de limpieza de datos financieros en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Financial_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos financieros del acelerador del Reino Unido:

Nombre	Descripción
rule_GBR_Bank_Account_Parse	Reconoce cadenas numéricas de ocho dígitos como números de cuentas bancarias del Reino Unido.
rule_GBR_Bank_Account_Validation	Valida números de cuentas bancarias del Reino Unido. La regla devuelve códigos que determinan si la entrada es numérica y si cuenta con la cantidad de dígitos correcta.
rule_GBR_Bank_Sort_Code_Parse	Analiza cadenas numéricas de seis dígitos como códigos de sucursal bancaria del Reino Unido. La regla analiza cadenas de números con los formatos siguientes: <ul style="list-style-type: none">- Números consecutivos (999999)- Números delimitados por guión (99-99-99)
rule_GBR_Bank_Sort_Code_Validation	Valida el formato y la longitud de códigos de sucursales bancaria del Reino Unido que se han estandarizado según el formato delimitado por guiones (99-99-99). La regla devuelve un puerto de estado que describe la validez del código de sucursal y un puerto de nota de validación que explica el estado. Si el prefijo del código de sucursal coincide con una asignación conocida de un banco del Reino Unido, el puerto de nota de validación incluirá el nombre del banco.
rule_GBR_Bank_Sort_Code_Standardize	Estandariza un código de sucursal bancaria del Reino Unido según el formato "NN-NN-NN".

Reglas de limpieza de datos generales del Reino Unido

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en los campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador del Reino Unido:

Nombre	Descripción
rule_GBR_NER_Field_Identification	Identifica el tipo de información contenida en un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones del Reino Unido. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia para identificar los tipos de información. La regla emplea técnicas de coincidencia probabilística para identificar los tipos de información.

El acelerador del Reino Unido depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Assign_DQ_GeocodingStatus_Description
- rule_Assign_DQ_Mailability_Score_Description
- rule_Assign_DQ_Match_Code_Descriptions
- rule_Remove_Extra_Spaces
- rule_Remove_Leading_Zero
- rule_Remove_Period_Parentheses
- rule_Remove_Punctuation
- rule_Remove_Punctuation_and_Space
- rule_Remove_Space
- rule_Replace_Limited_Punct_with_Space
- rule_UpperCase

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos generales del núcleo” en la página 23](#).

Reglas de coincidencia y eliminación de duplicados del Reino Unido

Utilice las reglas de coincidencia y eliminación de duplicados para medir los niveles de similitud entre los registros de conjuntos de datos.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador del Reino Unido:

Nombre	Descripción
mplt_GBR_Company_Name_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de nombres de empresas y códigos postales. El mapplet genera claves de grupo a partir del código postal.
mplt_GBR_Famillyname_and_NINO_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de los apellidos y el NINO (National Insurance Number, número de la seguridad social nacional del Reino Unido). El mapplet genera claves de grupo a partir de los datos del NINO.
mplt_GBR_Famillyname_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de apellidos y códigos postales del Reino Unido. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_GBR_Firstname_3CharsSurname_DOB_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de los datos siguientes: <ul style="list-style-type: none"> - Nombre de pila - Los tres primeros caracteres del apellido - Fecha de nacimiento - código postal El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_GBR_Firstname_Surname_2ElementsDOB_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de los datos siguientes: <ul style="list-style-type: none"> - Nombres de personas - Cualquier par de elementos de la fecha de nacimiento, como mes y año - Código postal del Reino Unido El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_GBR_Firstname_Surname_DOB_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas a partir de los datos siguientes: <ul style="list-style-type: none"> - Nombres de personas - Fecha de nacimiento - código postal El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_GBR_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos del Reino Unido a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_GBR_IMO_Famillyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos del Reino Unido a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_GBR_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos del Reino Unido a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_GBR_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos del Reino Unido a partir de nombres de personas y datos personales. Los campos de la columna de datos personales deben contener un único tipo de datos, como número de teléfono, dirección de correo electrónico o número de la seguridad social nacional del Reino Unido. El mapplet genera claves de grupo a partir de los datos personales.

Nombre	Descripción
mplt_GBR_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de nombres de personas y datos de fechas. El mapplet genera claves de grupo a partir de los datos de fechas.
mplt_GBR_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de nombres de personas y datos de direcciones de correo electrónico. El mapplet genera claves de grupo a partir de los datos de direcciones de correo electrónico.
mplt_GBR_Individual_Name_and_NINO_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir los nombres de personas y el NINO. El mapplet genera claves de grupo a partir de los datos del NINO.
mplt_GBR_Individual_Name_and_Phone_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de nombres de personas y números de teléfono. El mapplet genera claves de grupo a partir de los datos de números de teléfono.
mplt_GBR_Individual_Name_and_Postcode_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de nombres de personas y datos de código postal. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_GBR_Individual_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos del Reino Unido a partir de nombres de personas. El mapplet genera códigos NYSIIS a partir de los valores de apellido y los emplea como claves de grupo.
rule_GBR_Familyname_and_NINO_MatchScore	Genera una puntuación de coincidencia a partir de los apellidos y el NINO.
rule_GBR_Familyname_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de los apellidos y los códigos postales del Reino Unido.
rule_GBR_Firstname_3CharsSurname_DOB_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de la siguiente información: <ul style="list-style-type: none"> - Nombre de pila - Los tres primeros caracteres del apellido - Fecha de nacimiento - Código postal
rule_GBR_Firstname_Surname_2ElementsDOB_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de la siguiente información: <ul style="list-style-type: none"> - Nombres de personas - Cualquier par de elementos de la fecha de nacimiento, como mes y año - Código postal del Reino Unido
rule_GBR_Firstname_Surname_DOB_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas, fechas de nacimiento y códigos postales.
rule_GBR_Individual_Name_and_NINO_MatchScore	Genera una puntuación de coincidencia a partir de los nombres de personas y el NINO.
rule_GBR_Individual_Name_and_Phone_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números de teléfono.
rule_GBR_Individual_Name_and_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y códigos postales del Reino Unido.

Nombre	Descripción
rule_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_Individual_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas.
rule_GBR_Company_Name_Postcode_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y códigos postales del Reino Unido.
rule_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.

Asignaciones de demostración del Reino Unido

Las asignaciones de demostración del acelerador del Reino Unido utilizan varias reglas para comprobar los procesos de calidad de datos.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\GBR_Accelerator
```

El acelerador del Reino Unido incluye las siguientes asignaciones de demostración:

m_GBR_customer_data_demo

Analiza, estandariza y valida datos de clientes del Reino Unido.

m_GBR_customer_matching_demo

Analiza y estandariza datos de identidad del Reino Unido, y realiza un análisis de coincidencia de identidades de los datos.

La asignación analiza las siguientes combinaciones de datos y genera clústeres de coincidencia para cada combinación:

- Nombre de persona y datos de direcciones
- Nombre de persona y número de teléfono

CAPÍTULO 12

Acelerador de EE. UU./Canadá

Este capítulo incluye los siguientes temas:

- [Resumen del acelerador de EE. UU./Canadá, 126](#)
- [Reglas de limpieza de datos de direcciones de EE. UU./Canadá, 126](#)
- [Reglas compuestas de EE. UU./Canadá, 129](#)
- [Reglas de limpieza de datos de contacto de EE. UU./Canadá, 131](#)
- [Reglas de limpieza de datos corporativos de EE. UU./Canadá, 136](#)
- [Reglas de limpieza de datos generales de EE. UU./Canadá, 137](#)
- [Reglas de coincidencia y eliminación de duplicados de EE. UU./Canadá, 138](#)
- [Asignaciones de demostración de EE. UU./Canadá, 140](#)

Resumen del acelerador de EE. UU./Canadá

Utilice las reglas del acelerador de EE. UU./Canadá para comprobar y mejorar los datos de organizaciones en Estados Unidos y Canadá.

El acelerador de EE. UU./Canadá incluye reglas que realizan los siguientes procesos de calidad de datos:

- Limpieza de datos de direcciones
- Limpieza de datos de contacto
- Limpieza de datos generales
- Coincidencia y eliminación de duplicados

El acelerador depende de las reglas de limpieza de datos que el acelerador del núcleo ha instalado.

Reglas de limpieza de datos de direcciones de EE. UU./Canadá

Utilice reglas de limpieza de datos de direcciones para analizar, estandarizar y validar datos de direcciones.

Puede encontrar las reglas de limpieza de datos de direcciones en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

La tabla siguiente describe las reglas de limpieza de datos de direcciones del acelerador de EE. UU./Canadá:

Nombre	Descripción
rule_CAN_Address_Certification_Hybrid	<p>Valida la factibilidad de entrega de las direcciones canadienses a los estándares del programa de evaluación y reconocimiento de software (SERP) que Canada Post mantiene para direcciones canadienses. La regla corrige errores en las direcciones de entrada siempre que sea posible e incluye un puerto de estado que identifica las direcciones que cumplen con el estándar SERP. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.</p> <p>Nota: Cuando certifique un conjunto de registros de direcciones con el estándar SERP, debe enviar un informe de certificación a Canada Post. La transformación del validador de direcciones incluye campos de propiedades que se pueden rellenar con información para el informe. Guarde o imprima el informe, e inclúyalo con los detalles de direcciones que envíe a Canada Post.</p>
rule_CAN_Address_Parse_Hybrid	Divide direcciones de Canadá sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_CAN_Address_Parse_Multiline	Divide direcciones de Canadá sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_CAN_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones canadienses y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_CAN_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones canadienses. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_CAN_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de las direcciones canadienses y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_CAN_Address_Validation_Hybrid	Valida la factibilidad de entrega de las direcciones canadienses. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_CAN_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones canadienses y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_CAN_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones canadienses. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Nombre	Descripción
rule_CAN_Postcode_Validation	Valida los códigos postales canadienses. La regla devuelve "Válido" o "No válido".
rule_CAN_Province_Validation	Valida los nombres de las provincias canadienses. La regla devuelve "Válido" o "No válido".
rule_USA_Address_Certification_Hybrid	<p>Valida la factibilidad de entrega de las direcciones estadounidenses a los estándares del sistema de asistencia de precisión de la codificación (CASS) que el servicio postal de Estados Unidos mantiene para las direcciones. La regla corrige errores en las direcciones de entrada siempre que sea posible e incluye un puerto de estado que identifica las direcciones que cumplen con el estándar CASS. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.</p> <p>Nota: Cuando certifique un conjunto de registros de direcciones con el estándar CASS, debe enviar un informe de certificación a USPS. La transformación del validador de direcciones incluye campos de propiedades que se pueden rellenar con información para el informe. Guarde o imprima el informe, e inclúyalo con los detalles de direcciones que envíe a USPS.</p>
rule_USA_Address_Parse_Hybrid	Divide direcciones de Estados Unidos sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_USA_Address_Parse_Multiline	Divide direcciones de Estados Unidos sin estructura en elementos de dirección. La regla no valida las direcciones. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_USA_Address_Validation_Discrete_w_Geocoding	Valida la factibilidad de entrega de las direcciones estadounidenses y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_USA_Address_Validation_Discrete	Valida la factibilidad de entrega de las direcciones estadounidenses. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada discretos en la transformación del validador de direcciones.
rule_USA_Address_Validation_Hybrid_w_Geocoding	Valida la factibilidad de entrega de los registros de direcciones estadounidenses y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_USA_Address_Validation_Hybrid	Valida la factibilidad de entrega de los registros de direcciones estadounidenses. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada híbridos en la transformación del validador de direcciones.
rule_USA_Address_Validation_Multiline_w_Geocoding	Valida la factibilidad de entrega de las direcciones estadounidenses y añade coordenadas de latitud y longitud a cada dirección de salida. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.

Nombre	Descripción
rule_USA_Address_Validation_Multiline	Valida la factibilidad de entrega de las direcciones estadounidenses. La regla corrige errores en las direcciones de entrada siempre que sea posible. Utilice la regla cuando pueda conectar los campos de direcciones de entrada a los puertos de entrada de varias líneas en la transformación del validador de direcciones.
rule_USA_County_Validation	Valida los nombres de condados de Estados Unidos. La regla compara datos de entrada con los nombres de condado de todos los estados. La regla devuelve "Válido" o "No válido".
rule_USA_State_Validation	Valida nombres de estados de Estados Unidos. La regla devuelve "Válido" o "No válido".
rule_USA_ZIPCode_Validation	Valida códigos ZIP (Zone Improvement Plan, código postal de Estados Unidos) de cinco dígitos. La regla devuelve "Válido" o "No válido".

Reglas compuestas de EE. UU./Canadá

Utilice las reglas compuestas en el acelerador de EE. UU./Canadá para agregar un conjunto de reglas a una asignación como un único objeto. Una regla compuesta es una regla que utiliza la lógica de otras reglas del acelerador.

Puede encontrar las reglas compuestas en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

En la siguiente tabla se describen las reglas compuestas del acelerador de EE. UU./Canadá:

Nombre	Descripción
rule_CAN_Contact_Data	Analiza, estandariza y valida datos de contacto de Canadá, tales como las direcciones, los números de teléfono y el SIN (Social Insurance Number, número de la seguridad social de Canadá).
rule_USA_Contact_Data	Analiza, estandariza y valida datos de contacto de Estados Unidos, tales como las direcciones, los números de teléfono y el SSN (Social Security Number, número de la seguridad social de Estados Unidos).

Regla compuesta de datos de contacto de Canadá

La regla `rule_CAN_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en `rule_CAN_Contact_Data`:

Regla	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	Informatica_DQ_Content\Rules\General_Data_Cleansing

Regla	Ubicación
rule_CAN_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_CAN_Gender_Assignment	Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_Multi_Person_Name_Parse	Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_SIN_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_SIN_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Regla compuesta de datos de contacto de Estados Unidos

La regla `rule_USA_Contact_Data` lee mapplets de varias carpetas en el repositorio. La regla también incluye una transformación no reutilizable.

La tabla siguiente enumera los nombres y las ubicaciones en el repositorio de las reglas y la transformación en `rule_USA_Contact_Data`:

Regla	Ubicación
Case_Converter	Transformación no reutilizable
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_USA_Gender_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Regla	Ubicación
rule_USA_SSN_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_SSN_Validation_post_June2011	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Reglas de limpieza de datos de contacto de EE. UU./ Canadá

Utilice las reglas de limpieza de datos de contacto para analizar, estandarizar y validar datos sobre contactos empresariales e individuos.

Puede encontrar las reglas de limpieza de datos de contacto en la siguiente ubicación del repositorio:

[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

La tabla siguiente describe las reglas de limpieza de datos de contacto del acelerador de EE. UU./Canadá:

Nombre	Descripción
rule_CAN_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, esta regla asigna el sexo "M" al nombre "John Smith", ya que es masculino.
rule_CAN_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres canadienses. Por ejemplo, la regla estandariza el sobrenombre "Bob" según el nombre de pila "Robert".
rule_CAN_Multi_Person_Name_Parse	<p>Analiza valores de nombre de persona en puertos independientes. La regla crea puertos para valores como el tratamiento, el nombre, el segundo nombre y los apellidos.</p> <p>La salida de la regla incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Cuando los datos de nombre identifican más de una persona, la regla crea un puerto de salida para cada nombre completo. Por ejemplo, la regla puede leer el nombre "John y Jane Smith" y crear puertos de salida para "John Smith" y "Jane Smith".</p>
rule_CAN_Personal_Name_Parse_and_Standardize_FML	<p>Analiza los valores en un nombre de persona en puertos independientes. La regla también estandariza los valores de nombre.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Nombre, segundo nombre, apellidos <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>

Nombre	Descripción
rule_CAN_Personal_Name_Parse_and_Standardize_LFM	<p>Analiza los valores en un nombre de persona en puertos independientes. La regla también estandariza los valores de nombre.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_CAN_Personal_Name_Parse_and_Standardize_FML	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Nota: La regla no estandariza los valores de nombre. Para estandarizar y analizar los valores de nombre de Canadá en la secuencia que define la regla, seleccione rule_CAN_Personal_Name_Parse_and_Standardize_FML.</p>
rule_CAN_Personal_Name_Parse_and_Standardize_LFM	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Nota: La regla no estandariza los valores de nombre. Para estandarizar y analizar los valores de nombre de Canadá en la secuencia que define la regla, seleccione rule_CAN_Personal_Name_Parse_and_Standardize_LFM.</p>
rule_CAN_Phone_Number_Parse	<p>Reconoce un número de teléfono canadiense en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda. La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.</p>
rule_CAN_Phone_Number_Standardization	<p>Estandariza números de teléfono canadienses. Esta regla devuelve el número de teléfono con los formatos siguientes:</p> <ul style="list-style-type: none"> - Estándar: (nnn) nnn-nnnn - Con guiones: nnn-nnn-nnnn - Sin espacios: nnnnnnnnnn
rule_CAN_Phone_Number_Validation	<p>Valida el código de área y la longitud de los números de teléfono de Canadá. La regla devuelve códigos que indican el tipo y la validez del número de teléfono. Los tipos describen categorías como "número gratuito".</p>
rule_CAN_Phone_Parse_and_Standardize_Validate	<p>Analiza un número de teléfono de una cadena de texto y comprueba que el código de área es válido para Canadá.</p> <p>Si el código de área es válido, la regla devuelve el número de teléfono en tres formatos estándar. La regla también devuelve un valor de estado para indicar si los datos se ajustan al formato estándar para un número de teléfono canadiense.</p>

Nombre	Descripción
rule_CAN_Phone_w_Extension_Parse	Analiza un número de una cadena de texto si el número se ajusta al formato estándar para un número de teléfono canadiense. La regla incluye cualquier dato de extensión de teléfono cuando analiza el número de teléfono.
rule_CAN_SIN_Parse	Reconoce el SIN (Social Insurance Number, número de seguro social de Canadá) en una cadena. La regla devuelve el SIN, así como una cadena que contiene el texto de entrada sin el SIN.
rule_CAN_SIN_Standardization	<p>Estandariza el SIN. La regla puede generar una salida con los siguientes formatos:</p> <ul style="list-style-type: none"> - Sin puntuación: nnnnnnnnn - Con espacios: nnn nnn nnn - Con guiones: nnn-xxx-xxx <p>Para cambiar el formato, edite la variable de la expresión SIN_Format en la transformación de expresión dq_Format_SIN. El valor predeterminado es "No_punctuation".</p>
rule_CAN_SIN_Validation	Valida el SIN. La regla utiliza el algoritmo de Luhn para determinar la validez del SIN. La regla devuelve "Válido" o "No válido".
rule_Prenome_Assignment	Genera un tratamiento según el sexo. Puede cambiar la variable de expresión female_prenome de "Ms." a "Mrs.".
rule_Salutation_Assignment	Genera saludos formales e informales a partir de nombres de pila y tokens de nombre. Por ejemplo, si los datos de entrada contienen "Mr. John Smith," la regla genera el saludo formal "Dear Mr. Smith," y el saludo informal "Dear John,". Puede cambiar el prefijo y la puntuación si edita las variables en la transformación de expresión dq_Generate_Salutation.
rule_USA_Gender_Assignment	Asigna un sexo a partir de nombres de pila. La regla devuelve "M" para nombres masculinos, "F" para nombres femeninos y "U" si el sexo es desconocido. Por ejemplo, esta regla asigna el sexo "M" al nombre "John Smith", ya que es masculino.
rule_USA_Given_Name_Standard	Genera nombres de pila a partir de sobrenombres estadounidenses. Por ejemplo, la regla estandariza el sobrenombre "Bob" según el nombre de pila "Robert".
rule_USA_Multi_Person_Name_Parse	<p>Analiza valores de nombre de persona en puertos independientes. La regla crea puertos para valores como el tratamiento, el nombre, el segundo nombre y los apellidos.</p> <p>La salida de la regla incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Cuando los datos de nombre identifican más de una persona, la regla crea un puerto de salida para cada nombre completo. Por ejemplo, la regla puede leer el nombre "John y Jane Smith" y crear puertos de salida para "John Smith" y "Jane Smith".</p>

Nombre	Descripción
rule_USA_Personal_Name_Parse_and_Standardize_FML	<p>Analiza los valores en un nombre de persona en puertos independientes. La regla también estandariza los valores de nombre.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Nombre, segundo nombre, apellidos <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_USA_Personal_Name_Parse_and_Standardize_LFM	<p>Analiza los valores en un nombre de persona en puertos independientes. La regla también estandariza los valores de nombre.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p>
rule_USA_Personal_Name_Parse_Validation	<p>Valida la asignación de sexo para un nombre. La regla calcula las probabilidades de que un valor de datos sea un nombre masculino o un nombre femenino. Si se desconoce el sexo, la regla utiliza los cálculos de probabilidad para asignar un sexo al nombre.</p>
rule_USA_Personal_Name_Parsing_FML	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Nombre, segundo nombre, apellidos <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Nota: La regla no estandariza los valores de nombre. Para estandarizar y analizar los valores de nombre de Estados Unidos en la secuencia que define la regla, seleccione rule_USA_Personal_Name_Parse_and_Standardize_FML.</p>
rule_USA_Personal_Name_Parsing_LFM	<p>Analiza los valores en un nombre de persona en puertos independientes.</p> <p>La regla crea los puertos en la siguiente secuencia:</p> <ul style="list-style-type: none"> - Apellidos, nombre, segundo nombre <p>La salida de la regla también incluye un puerto que contiene el nombre completo de la persona en el registro. Puede utilizar el puerto del nombre completo como entrada para una transformación de coincidencia en una asignación de análisis de coincidencia de identidades.</p> <p>Nota: La regla no estandariza los valores de nombre. Para estandarizar y analizar los valores de nombre de Estados Unidos en la secuencia que define la regla, seleccione rule_USA_Personal_Name_Parse_and_Standardize_LFM.</p>
rule_USA_Phone_Number_Parse	<p>Reconoce un número de teléfono estadounidense en una cadena. La regla analiza el primer número de teléfono que encuentra en los datos al leer de derecha a izquierda. La regla devuelve un número de teléfono, así como una cadena que contiene el texto de entrada sin el número de teléfono.</p>

Nombre	Descripción
rule_USA_Phone_Number_Standardization	<p>Estandariza números de teléfono de Estados Unidos. Esta regla devuelve el número de teléfono con los formatos siguientes:</p> <ul style="list-style-type: none"> - Estándar: (nnn) nnn-nnnn - Con guiones: nnn-nnn-nnnn - Sin espacios: nnnnnnnnnn
rule_USA_Phone_Number_Validation	<p>Valida el código de área y la longitud de los números de teléfono de Estados Unidos. La regla devuelve códigos que determinan la validez del código de área y la longitud de un número de teléfono.</p>
rule_USA_Phone_Parse_Standardize_Validate	<p>Analiza un número de teléfono de una cadena de texto y comprueba que el código de área es válido para Estados Unidos.</p> <p>Si el código de área es válido, la regla devuelve el número de teléfono en tres formatos estándar. La regla también devuelve un valor de estado para indicar si los datos se ajustan al formato estándar para un número de teléfono de Estados Unidos.</p>
rule_USA_Phone_w_Extension_Parse	<p>Analiza un número de una cadena de texto si el número se ajusta al formato estándar para un número de teléfono de Estados Unidos. La regla incluye cualquier dato de extensión de teléfono cuando analiza el número de teléfono.</p>
rule_USA_SSN_Parse	<p>Analiza el SSN (Social Security Number, número de la seguridad social de Estados Unidos).</p>
rule_USA_SSN_Parse_Standardize_and_Validate	<p>Analiza, estandariza y valida los números de seguridad social de Estados Unidos procedentes de una cadena de texto más grande. La regla puede analizar números que incluyan u omitan guiones.</p> <p>De forma predeterminada, la regla escribe los números de seguridad social sin ninguna puntuación. Para cambiar el formato de estandarización, abra la transformación dq_SSN_Format en la regla y actualice la expresión en el puerto SSN_Format.</p>
rule_USA_SSN_Standardization	<p>Estandariza el SSN. La regla puede generar una salida con los siguientes formatos:</p> <ul style="list-style-type: none"> - Sin puntuación: nnnnnnnnnn - Con espacios: nnn nnn nnn - Con guiones: nnn-nnn-nnn <p>Para cambiar el formato, edite la variable de expresión SSN_format en la transformación de expresión dq_SSN_Format. El valor predeterminado es "No_punctuation".</p>

Nombre	Descripción
rule_USA_SSN_Validation	<p>Valida el SSN. La regla comprueba cada SSN en función de la longitud, los valores numéricos y los valores máximos y mínimos conocidos de las secciones Área, Grupo y Número de serie.</p> <p>La sección Área comprende los tres primeros dígitos del SSN, mientras que la sección Grupo contiene el cuarto y el quinto dígito. La sección Número de serie comprende los últimos cuatro dígitos.</p> <p>Si el SSN se emitió antes de junio de 2011, la regla también comprueba la validez de la combinación de los valores de Área y Grupo. La regla no comprueba que el SSN sea un número expedido. La regla devuelve "Válido" o "No válido".</p>
rule_USA_SSN_Validation_post_June2011	<p>Valida el SSN. La regla comprueba cada SSN en función de la longitud, los valores numéricos y los valores máximos y mínimos conocidos de las secciones Área, Grupo y Número de serie.</p> <p>La sección Área comprende los tres primeros dígitos del SSN, mientras que la sección Grupo contiene el cuarto y el quinto dígito. La sección Número de serie comprende los últimos cuatro dígitos.</p> <p>La regla no comprueba la validez de la combinación de los valores de Área y Grupo. La regla no comprueba que el SSN sea un número expedido. La regla devuelve "Válido" o "No válido".</p>

Dependencias de reglas de limpieza de datos de contacto del núcleo

El acelerador de EE. UU./Canadá depende de las siguientes reglas de limpieza de datos de contacto del acelerador del núcleo:

- rule_Email_Validation

Para obtener más información sobre estas reglas, consulte ["Reglas de limpieza de datos de contacto del núcleo" en la página 22.](#)

Reglas de limpieza de datos corporativos de EE. UU./Canadá

Utilice las reglas de limpieza de datos corporativos del acelerador de EE. UU./Canadá para analizar, estandarizar y validar datos corporativos.

La tabla siguiente describe las reglas de limpieza de datos corporativos del acelerador de EE. UU./Canadá:

Nombre	Descripción
rule_NAICS_Code_Validation	Valida los códigos del sistema norteamericano de clasificación industrial (NAICS).
rule_USA_SIC_Code_Validation	Valida los códigos de clasificación industrial estándar (SIC).

El acelerador de EE. UU./Canadá depende de las siguientes reglas de limpieza de datos corporativos del acelerador del núcleo:

- rule_Company_Name_Standardization

Reglas de limpieza de datos generales de EE. UU./ Canadá

Utilice las reglas de limpieza de datos generales para identificar el tipo de información contenida en los campos de entrada.

Puede encontrar las reglas de limpieza de datos generales en la siguiente ubicación del repositorio:

[Informatica_DQ_Content]\Rules\General_Data_Cleansing

La tabla siguiente describe las reglas de limpieza de datos generales del acelerador de EE. UU./Canadá:

Nombre	Descripción
rule_CAN_Field_Identification	Identifica el tipo de información que contiene un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de Canadá. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia para identificar los tipos de información.
rule_CAN_NER_Field_Identification	Identifica el tipo de información que contiene un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de Canadá. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia para identificar los tipos de información. La regla emplea técnicas de coincidencia probabilística para identificar los tipos de información.
rule_USA_Field_Identification	Identifica el tipo de información que contiene un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de EE. UU. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia para identificar los tipos de información.
rule_Field_North_American_Data	Identifica los siguientes tipos de campos: nombre, cargo, empresa, dirección, ciudad, estado o provincia, código postal, país, identificación personal, correo electrónico, teléfono, tarjeta de crédito y fecha. La regla genera una puntuación que indica el grado de confianza en la identificación del campo. Las puntuaciones más altas indican mayores niveles de confianza. Si la regla no puede asignar un tipo de campo, escribe los datos en el puerto Out_Undetermined.
rule_USA_NER_Field_Identification	Identifica el tipo de información que contiene un campo de entrada. La regla puede identificar nombres, ID personales, nombres de empresas, fechas y datos de direcciones de EE. UU. La regla devuelve una etiqueta que describe el tipo de datos de entrada. La regla emplea datos de referencia para identificar los tipos de información. La regla emplea técnicas de coincidencia probabilística para identificar los tipos de información.

Dependencias de reglas de limpieza de datos generales del núcleo

El acelerador de EE. UU./Canadá depende de las siguientes reglas de limpieza de datos generales del acelerador del núcleo:

- rule_Assign_DQ_GeocodinStatus_Description
- rule_Assign_DQ_Mailability_Score_Description
- rule_Assign_DQ_Match_Code_Descriptions

- rule_Date_Validation
- rule_Remove_Extra_Spaces
- rule_Remove_Punctuation
- rule_Replace_Limited_Punct_with_Space
- rule_UpperCase

Para obtener más información sobre estas reglas, consulte [“Reglas de limpieza de datos generales del núcleo” en la página 23.](#)

Reglas de coincidencia y eliminación de duplicados de EE. UU./Canadá

Utilice las reglas de coincidencia y eliminación de duplicados para medir los niveles de similitud entre los registros de conjuntos de datos.

Puede encontrar las reglas de coincidencia y eliminación de duplicados en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

La siguiente tabla describe las reglas de coincidencia y eliminación de duplicados del acelerador de EE. UU./Canadá:

Nombre	Descripción
mplt_CAN_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Canadá a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_CAN_IMO_Familyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Canadá a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_CAN_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Canadá a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_CAN_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Canadá a partir de nombres de personas y datos personales. Los campos de la columna de datos personales deben contener un único tipo de datos, tales como números de teléfono, correos electrónico o números de la seguridad social de Canadá. El mapplet genera claves de grupo a partir de los datos personales.
mplt_Company_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de nombre de empresa y código postal para generar claves de grupo.
mplt_Company_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de empresas. El mapplet genera códigos Soundex a partir de los valores de nombre de empresa y los emplea como claves de grupo.

Nombre	Descripción
mplt_Famillyname_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de apellidos y datos de direcciones. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_Firstname_and_SSN_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de los nombres de pila y los números de la seguridad social de Estados Unidos. El mapplet genera claves de grupo a partir de datos del número de la seguridad social de Estados Unidos.
mplt_Individual_Name_and_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y datos de direcciones de Estados Unidos. El mapplet utiliza una combinación de caracteres de los valores de apellido y código postal para generar claves de grupo.
mplt_Individual_Name_and_Date_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y datos de fechas. El mapplet genera claves de grupo a partir de los datos de fechas.
mplt_Individual_Name_and_Email_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y direcciones de correo electrónico. El mapplet genera claves de grupo a partir de los datos de direcciones de correo electrónico.
mplt_Individual_Name_and_Phone_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y números de teléfono. El mapplet genera claves de grupo a partir de los datos de números de teléfono.
mplt_Individual_Name_and_SSN_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas y números de la seguridad social de Estados Unidos. El mapplet genera claves a partir de datos del número de la seguridad social de Estados Unidos.
mplt_Individual_Name_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas a partir de nombres de personas. El mapplet genera códigos NYSIIS a partir de los valores de apellido y los emplea como claves de grupo.
mplt_USA_Address_Match	Utiliza estrategias de la coincidencia de campos para identificar filas duplicadas en datos de Estados Unidos a partir de apellidos y datos de direcciones Estados Unidos. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_USA_IMO_Company_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Estados Unidos a partir de nombres de empresas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_USA_IMO_Famillyname_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Estados Unidos a partir de apellidos y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.
mplt_USA_IMO_Individual_Name_and_Address_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Estados Unidos a partir de nombres de personas y direcciones. El mapplet genera claves de grupo a partir de los datos de código postal.

Nombre	Descripción
mplt_USA_IMO_Personal_Name_and_Data_Match	Utiliza estrategias de coincidencia de identidades para identificar filas duplicadas en datos de Estados Unidos a partir de nombres de personas y datos personales. Los campos de la columna de datos personales deben contener un único tipo de datos, tales como número de teléfono, dirección de correo electrónico o número de la seguridad social de Estados Unidos. El mapplet genera claves de grupo a partir de los datos personales.
rule_Company_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas y datos de direcciones de Estados Unidos.
rule_Company_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de empresas.
rule_Familyname_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de apellidos y datos de direcciones de Estados Unidos.
rule_Firstname_and_SSN_MatchScore	Genera una puntuación de coincidencia a partir de nombres de pila y datos de direcciones de Estados Unidos.
rule_Individual_Name_and_Address_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y datos de direcciones de Estados Unidos.
rule_Individual_Name_and_Date_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y fechas.
rule_Individual_Name_and_Email_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y direcciones de correo electrónico.
rule_Individual_Name_and_Phone_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas y números de teléfono.
rule_Individual_Name_and_SSN_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas, números de la seguridad social de Estados Unidos y datos de identificación.
rule_Individual_Name_MatchScore	Genera una puntuación de coincidencia a partir de nombres de personas.
rule_USA_Address_MatchScore	Genera una puntuación de coincidencia a partir de datos de direcciones de Estados Unidos.

Asignaciones de demostración de EE. UU./Canadá

Las asignaciones de demostración del acelerador de EE. UU./Canadá utilizan varias reglas para comprobar los procesos de calidad de datos.

Puede encontrar las asignaciones de demostración en la siguiente ubicación del repositorio:

```
[Informatica_DQ_Content]\Rules_Demo\US_Canada_Accelerator
```

El acelerador de EE. UU./Canadá incluye las siguientes asignaciones de demostración:

m_customer_data_US_demo

Analiza, estandariza y valida datos de Estados Unidos y Canadá.

m_customer_matching_US_demo

Analiza y estandariza datos de identidad de Estados Unidos y realiza un análisis de coincidencia de identidades de los datos.

La asignación analiza las siguientes combinaciones de datos y genera clústeres de coincidencia para cada combinación:

- Nombre de persona y datos de direcciones
- Nombre de persona y número de teléfono